



GenAI *in* Legal Education

*A Practical Guide for
Professors and Students*

KORIN MUNSTERMAN

Author of AI and the Practice of Law in a Nutshell



GenAI In Legal Education

A Practical Guide for Professors and Students

Korin Munsterman

UNT Dallas College of Law

©Center for Computer Assisted Legal Instruction, 2026

First Edition

ISBN: 979-8-89904-053-5

To late bloomers and introverts

About the Author

Korin Munsterman is a Professor of Practice and Director of the Legal Education Technology program at the University of North Texas Dallas College of Law. Before joining UNT, she held positions at Harvard Law School, Harvard Business School, American University Washington College of Law, and UMass Dartmouth School of Law. Over her career, she has consulted with law schools, firms, and government agencies on technology integration in legal practice and education, developed online LL.M. programs, and was commissioned by BARBRI to develop a prototype course for online J.D. programs.

She is the author of *AI and the Practice of Law in a Nutshell* (West Academic, 2026). Her teaching and research focus on where legal practice and technology meet—specifically how practitioners and institutions can deploy AI tools effectively and responsibly.

She earned her J.D. from American University Washington College of Law and her M.L.I.S. from Florida State University.

About CALI eLangdell Press

The Center for Computer-Assisted Legal Instruction (CALI®) is a nonprofit organization with over 200 member US law schools, an innovative force pushing legal education toward change for the better. There are benefits to CALI membership for your school, firm, or organization. ELangdell® is our electronic press with a mission to publish more open books for legal education.

How do we define “open”?

- Compatibility with devices like smartphones, tablets, and e-readers, as well as print.
- The right for educators to remix the materials through more lenient copyright policies.
- The ability for educators and students to adopt the materials for free.

Find available and upcoming eLangdell titles at elangdell.cali.org. Show support for CALI by following us on Facebook and Twitter and by telling your friends and colleagues where you received your free book.

Notices

This is the first edition of this textbook and is current through June, 2026. Visit <http://elangdell.cali.org/> for the latest version and for revision history.

This work by Korin Munsterman is licensed and published by CALI eLangdell Press under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. CALI and CALI eLangdell Press reserve under copyright all rights not expressly granted by this Creative Commons license. CALI and CALI eLangdell Press do not assert copyright in U.S. Government works or other public domain material included herein. Permissions beyond the scope of this license may be available through feedback@cali.org. Copyright 2026 CALI.

In brief, the terms of that license are that you may copy, distribute, and display this work, or make derivative works, so long as:

- you give CALI eLangdell Press and the authors credit;
- you do not use this work for commercial purposes; and
- you distribute any works derived from this one under the same licensing terms as this.

Suggested attribution format for original work:

Korin Munsterman, GenAI In Legal Education by CALI eLangdell Press. Available under a Creative Commons BY-NC-SA 3.0 License.

CALI® and CALI eLangdell are United States federally registered trademarks owned by the Center for Computer-Assisted Legal Instruction. The cover art design is a copyrighted work of CALI, all rights reserved. The CALI graphical logo is a trademark and may not be used without permission.

Should you create derivative works based on the text of this book or other Creative Commons materials therein, you may not use this book's cover art and the aforementioned logos, or any derivative thereof, to imply endorsement or otherwise without written permission from CALI.

This material does not contain nor is intended to be legal advice. Users seeking legal advice should consult with a licensed attorney in their jurisdiction. The writers and editors have endeavored to provide complete and accurate

information in this book. However, CALI does not warrant that the information provided is complete and accurate. CALI disclaims all liability to any person for any loss caused by errors or omissions in this collection of information.

FOREWORD

Every large language model has a training data cutoff, a date past which it simply doesn't “know” what happened. The model answers confidently right up to that line, and beyond it, without accessing the web, it's clueless. Writing a book about generative AI this author in exactly the same position. Mine was April 30, 2026 (with a few wonderful suggestions from anonymous peer reviewers added after that date).

Generative AI (GenAI) development moves fast enough that specific features and tool capabilities can shift between the time I write a sentence and the time you read it. Anything I describe with specificity is accurate as of that date. Some of it will still hold; some of it won't. New applications appear constantly, and others disappear. When reality has outrun the text, I hope you'll use the underlying logic about how to think about these tools and how to evaluate new ones.

A. Practitioner, Not Scholar

A word about my qualifications. To be clear, I do not profess to be a legal education scholar. I have no formal training in pedagogy and no publications in the legal education literature. What I do have is a solid understanding of and years of research on how GenAI functions. On top of that, I spend a lot of time with law (as well as university and medical school) faculty members, students, and practitioners to ethically, effectively, and efficiently incorporate these tools into their work.

Where research on learning and pedagogy is relevant, I engaged with it carefully and cited it throughout. Where I'm drawing on practical experience rather than scholarship, I've tried to make that clear.

A practical note about how this book got written: I used GenAI throughout. The images were AI-generated using either Google or ChatGPT. When pulling together research, I used the same tools described in these pages to find patterns across a large volume of sources that would have taken considerably longer to find manually. And drafts of portions of the text went through GenAI assistance before landing in front of you. Everything has been verified. Every judgment about what to include and how to say it came from my own thinking and experience. The GenAI helped me work faster, not differently. Writing a book about these tools without actually using them would have been the strange choice — and frankly, a missed opportunity to practice what I'm preaching.

B. What I Left Out – and Why

One function that experienced GenAI users might expect to see discussed is absent – vibe-coding. The term, coined by AI researcher Andrej Karpathy, refers to the practice of describing what you want a program to do in plain language and letting an AI write the actual code, with no expectation that you'll read or understand the code it produces. OpenAI Codex and Claude Code are the two tools most associated with the practice, and neither gets substantive coverage here; Claude Code appears in the next few paragraphs mainly because the institutional and cost risks are worth flagging before you stumble into them.

Claude Code is a command-line tool that lets users direct Claude to write and execute code. Non-programmers can accomplish genuinely impressive things with it. But it charges by API token consumption, which means costs accumulate with every request and can escalate quickly for anyone who doesn't yet have a feel for where token-heavy tasks begin. Beyond cost, deploying Claude Code in any way that touches your institution's network or data infrastructure raises security and access control questions that might cause significant institutional issues. So you should talk to your IT department. If your IT staff has the skills to help you use Claude Code, it's probably best to let them lead the deployment. As someone who has worked in law school technology for many years, that's what they're there for.

The same caution applies to a related development that landed after my cutoff date. Starting June 15, 2026, Anthropic changes how it handles automated use of Claude, meaning software that runs in the background without you sitting at a keyboard typing. If you (or a technically inclined colleague) have been connecting an outside application to your personal Claude account to process documents automatically, that arrangement now draws from a separate spending budget rather than your subscription. A \$20 plan gets \$20 worth of that automated processing; once the budget runs out, the automation stops until you add more funds.

For example, a faculty member who wants Claude Code to read through eighty student memos overnight while she sleeps and flag analytical gaps before morning is doing exactly the kind of automated processing that now draws from the separate budget. So is a colleague who has written a script to scan a semester's worth of course evaluations to pull out recurring themes and generate a summary report. While both tasks are useful, they now cost something beyond the flat monthly fee, and that cost can add up faster than you'd expect, especially if the document sets are large.

Regular interactive use is unaffected. Typing into claude.ai or working through Claude Cowork falls under your normal subscription; Anthropic explicitly exempted its own native tools. The change targets the practice of running heavy background workloads through a flat-rate consumer subscription, which had apparently placed considerable strain on their infrastructure. For faculty using Claude the way this book describes, nothing changes. If someone has rigged up a script that connects a personal subscription to a large external database or document set and runs it overnight, that arrangement now has a meter attached.

C. The Chapters Addressed to Students

Two chapters in this book (Chapters 13 and 14) are written directly to students rather than to faculty. I included them for two purposes.

The first is straightforward: students need guidance that speaks to their actual situation, not a faculty-facing discussion with the pronouns swapped. The research on GenAI dependency is consistent enough to treat as settled, and students who figure out their GenAI habits without instruction often do so in ways that undermine their own learning. Chapters 13 and 14 give them a concrete framework for when GenAI helps and when it doesn't, grounded in the specific realities of law school and the professional obligations waiting on the other side of graduation.

The less obvious purpose is that those chapters are also a resource for faculty. Professors designing GenAI-related coursework will, I hope, find workable material there for whatever class you teach. The parallel between a professor's GenAI policy and a federal judge's standing orders runs through both chapters as a teaching device. Assign them directly or adapt the frameworks for your own design. They're built for both audiences.

D. The Nutshell

The companion to this book is *Artificial Intelligence and the Practice of Law in a Nutshell* published by West Academic, and if you're a practitioner or a student headed into practice soon, it belongs on the same shelf.

Where this book focuses on teaching with GenAI and learning from it, the Nutshell focuses on practicing with it and the professional obligations that come along. It starts with cautionary tales and legal ethics. *Mata v. Avianca*, where attorneys filed briefs full

of AI-hallucinated citations and faced sanctions, is the most prominent, but it's hardly the only one. The Nutshell documents the wave of disciplinary actions that followed in federal and state courts and explains what those cases actually turned on. This book includes discussion of these issues, too.

The professional responsibility coverage is comprehensive. ABA Formal Opinion 512 addresses competence, confidentiality, fees, and supervisory responsibilities when using AI in practice, and it gets the analysis it deserves. So do the state bar guidance documents that have proliferated in its wake. Judicial standing orders requiring disclosure or certification of AI use in court filings get their own treatment, including the ongoing debate over whether such requirements actually advance responsible AI adoption or just create unnecessary compliance confusion. In this book, I reference standing orders and the alignment between academic integrity and professional responsibility.

The Nutshell also covers the practical toolkit in depth, from what legal research platforms and litigation forecasting tools are actually doing to where document review automation is heading. It includes the history and technical foundations of AI in enough depth to make the rest of the book make sense. It gets into the harder questions those tools raise — algorithmic bias in risk assessment algorithms and what it means when the system's error disparately impacts your client, and the liability framework when AI-generated work product turns out to be wrong. There's also dedicated coverage of the evidentiary problems that deepfakes and synthetic media are creating for litigators right now, and a candid look at business model disruption: how AI efficiency gains are colliding with billable hour models and what firms that are getting ahead of that tension are actually doing. The book closes with vendor checklists and governance frameworks, along with a forward-looking account of where AI-augmented practice is headed.

If this book gets you using these tools effectively, the Nutshell tells you what you need to know to use them responsibly in practice and how to answer the questions that clients and courts are increasingly going to ask.

E. Final Thoughts

Chapter 2 moves quickly through the technical foundations of large language models (tokens, context windows, hallucinations, training data) and readers who are new to that vocabulary may find themselves tripping over terms before they've had a chance to get comfortable with them. A Glossary of Key Terms appears before Chapter 1 for that reason. It isn't required reading, but it's there if a definition would help, and it takes about

ten minutes to scan in full. Readers who already know what a context window is can skip it without missing anything.

The appendices at the back of this book are working resources including instructions on creating your own custom tools. Several contain links to interactive tools (CustomGPTs and Claude artifacts) I developed built to accompany specific chapters. For the linked tools, no advance reading required; just click the link.

One reviewer read parts of this manuscript as a product pitch, and that's a fair reading. Enthusiasm for tools you, your colleagues, and your students find useful can morph into an infomercial without much warning.

The CustomGPTs and Claude Artifacts described here are not products I sell. I built them for my own work and share them the way a colleague might hand over a well-worn template: because they work, not because anything turns on your adopting them. Nothing commercial sits behind these recommendations. Some will look different by the time you read this, and others may be replaced by better options. Feel free to use or ignore any of them.

Your outputs from these tools won't look like mine, even from the same prompt. I've worked with these platforms for years, and the models have accumulated an understanding of what I want through my settings and feedback over time. Your outputs will improve to meet your expectations too, but that calibration takes time to build. Careful prompting, using the CORRECT method as described in Chapter 4, is the best shortcut; a well-crafted prompt can get you close to what you want before the platform knows you well enough to get there on its own. That's why this book spends as much time as it does on how to prompt correctly.

Finally, law professors are, in my experience, a genuinely generous community. They share syllabi with strangers and argue about pedagogy with real enthusiasm. I wrote this book hoping it would be received in that spirit, and with the larger hope that what comes next is something more collective than each of us experimenting in parallel. I'm hoping we develop an exchange of what works and what doesn't. Whether you're brand new to GenAI or already well into it, there's something here for you. What you learn from it, I hope you'll pass along. I hope, too, to hear from you so we can build something together and share it with others.

Korin Munsterman
May 14, 2026

Acknowledgments

Professor Nancy Rapoport of UNLV William S. Boyd School of Law changed what this book is. I sent her the first chapter with the specific kind of nerve that comes from being a native New Yorker, and she responded not with the polite “I think it’s a good idea – go for it,” but with a fully annotated draft *and* an offer to read and comment on everything else. She did exactly that, going chapter-by-chapter, testing CustomGPTs and Artifacts, flagging what wasn't working. Nancy pushed the manuscript toward clarity and usefulness that I would not have found on my own. Anyone who has seen Nancy present on AI and ethics in legal education already understands why her engagement with this project mattered so much. What she gave this book is the kind of investment most authors can only hope for from their colleagues, and my gratitude extends beyond what acknowledgments are designed to express. And I thank Nancy and Professor Maryam Ahranjani for hosting me at UNLV Law to speak to the faculty about GenAI in Legal Education, which gave me ideas for other things to cover.

This book found its home in a conversation at AALS with John Mayer, Executive Director of CALI, who heard the pitch and agreed to take it on. John Jorgensen, CALI's Director of Curriculum Development, managed the peer review and editorial process through to publication while providing the steady guidance that kept the project on track. I am grateful to both for understanding the goal of this book.

The anonymous peer reviewers who evaluated this manuscript gave it careful, critical attention, and their willingness to push back where it needed made it considerably better. I’m so thankful for their many contributions.

At UNT Dallas College of Law, Dean Felecia Epps and Associate Dean Cynthia Fountaine gave me the freedom to teach AI and the Law and GenAI and the Practice of Law, and this book would not exist without that opportunity.

Professor Michael Maslanka at UNT Dallas College of Law has been a consistent champion of my scholarship, and his continued support means a great deal.

Associate Dean Billie Jo Kaufman at UDC David A. Clarke School of Law offered encouragement and interest throughout the writing process, and both matter so much during the long stretches when a book feels unfinishable.

Associate Dean Edward Hart at UNT Dallas never needs more than a day (and, often, not more than an hour) to solve a resource problem.

Ruth Priest, my research assistant, contributed careful research and editing assistance that sharpened the manuscript throughout. She's a joy to work with.

Mike Pearson, Assistant Director of Legal Education Technology, tested CustomGPTs and Artifacts and provided feedback that found its way into the final version, all while keeping LET running and maintaining the good humor and intelligence that makes everything easier.

The scholars and journalists whose work I cite throughout this book, and the presenters at the webinars I have attended over the past few years, provided the intellectual foundation this book stands on. Without their thinking, I would have had considerably less to say.

My students at UNT Dallas shaped this book more than they know. Their questions and willingness to experiment with tools that were often changing week to week pushed my thinking, and their steady stream of AI news kept me current throughout.

My colleagues on the UNT Faculty Senate who worked with me on AI policy, particularly on detection and assignment design, and the university faculty who I worked with to bring their CustomGPT ideas to development, pushed this work in directions I might not have found on my own.

Karen Day deserves more than an acknowledgments section can carry. Her friendship has been a constant throughout this project and throughout the years, and her steadiness means more to me than a few lines can convey.

My parents, Allan and Ellen, and my brothers, Kurt and Eric, have supported me throughout my life and throughout this project.

My daughter, Julia, lives far away and has a full life of her own, which I admire even when I miss her. I am grateful every day for her love, and for the fact that she is exactly who she is.

To my husband, Gerry Lee — Yeats, who knew something about the west of Ireland, wrote "take down this book" as the first gesture of love. Consider this one his.

Korin Munsterman
May 12, 2026

Glossary of Key Terms

The following terms appear throughout this book. Definitions reflect how each term is used in the context of GenAI and legal education; some carry broader or different meanings in computer science literature.

AI Agent (Agentic AI)

A system that perceives its environment, acts autonomously to achieve goals, uses tools to complete tasks, and chains multiple actions without constant human input. Distinguished from a standard chatbot or CustomGPT by its capacity to operate independently across a workflow rather than responding to a single prompt. Agentic systems are still emerging in legal practice, and professional responsibility frameworks for supervising them remain unsettled.

AI Assessment Scale (AIAS)

A five-level framework for defining the permissible role of GenAI in student assignments. Levels range from no GenAI use to full use with documentation. The AIAS does not rank levels by quality; appropriate placement depends on the learning objectives the assignment is designed to address. Now in its second iteration, it has been adopted by hundreds of institutions worldwide.

AI Fluency Framework

A model developed by Rick Dakan, Joseph Feller, and Anthropic that identifies four competencies for effective and responsible use of GenAI: Delegation, Description, Discernment, and Diligence. The competencies are interdependent — strong delegation decisions inform better description, effective description enables more accurate discernment, and thorough discernment supports diligence in deployment.

Chain-of-Thought Prompting

A technique that instructs the model to break complex reasoning into explicit sequential steps before reaching a conclusion. Particularly useful in legal education because it models the analytical process — not just the answer — and makes errors easier to identify. A basic prompt asks whether a contract is valid; a chain-of-thought prompt asks the model to work through offer, acceptance, consideration, and capacity step by step.

Claude

Anthropic's large language model, trained using Constitutional AI and Reinforcement Learning from AI Feedback (RLAIF). Claude is available in multiple versions (including Opus, Sonnet, and Haiku) and is accessible through Claude.ai, the API, Claude Projects, and Claude Cowork. Because its behavior is shaped by Anthropic's encoded principles, Claude may respond differently from other LLMs when handling ethically ambiguous prompts.

Confabulation

See Hallucination. The term preferred in research literature for the phenomenon in which an LLM generates plausible but false content. Confabulation better captures the mechanism — the model is not lying, it is pattern-matching in a way that produces confident-sounding nonsense.

Constitutional AI

Anthropic's training approach in which the model is guided by a set of encoded principles (the "constitution") and evaluated by AI-generated feedback rather than solely by human labelers. This approach improves scalability and transparency compared to pure RLHF, but also means Claude's behavior reflects Anthropic's design choices — including decisions that can produce refusals of benign requests.

Context Rot

Performance degradation that occurs as the amount of text in a prompt increases, independent of whether the model's context window is full. Coined by Chroma researchers in 2025 after testing eighteen frontier models and finding continuous quality decline as input length grew. Practically, this means that uploading a hundred-page document and asking a question at the end is not the same as providing a focused excerpt with the question placed first.

Context Window

The total amount of text — including the prompt, uploaded documents, and conversation history — that an LLM can process in a single interaction. Measured in tokens. Roughly analogous to working memory or RAM. Models attend most reliably to content at the beginning and end of a prompt; accuracy drops for material placed in the middle. Context

windows have grown dramatically over time, from approximately 2,000 tokens in early models to millions in current experimental systems.

CORRECT Framework

A prompting structure developed to help legal educators and students craft more effective prompts. The acronym stands for: Context (relevant background), Objective (what you want to achieve), Role (how the model should behave), Restrictions (what to avoid), Expectations (format and scope of output), Clarify (the specific task), and Tone (register and style). Users need not label each element explicitly once the framework becomes second nature; its value is as a thinking tool, not a rigid template.

CustomGPT

A specialized instance of ChatGPT configured with custom system instructions, uploaded knowledge, and access to specific tools. CustomGPTs persist their configuration across conversations, eliminating the need to re-prompt with the same context each session. They are not fully autonomous agents as they require human initiation at each step, but they sit closer to the agentic end of the spectrum than a standard chatbot. Analogous to Claude Projects in the Claude ecosystem.

Delegation

The first competency of the AI Fluency Framework. Deciding which tasks are appropriate to give to GenAI, and which are not. Effective delegation requires understanding both the model's capabilities and the task's stakes — including whether GenAI errors would be catchable before causing harm. Poor delegation is a root cause of many professional responsibility problems arising from GenAI use.

Description

The second competency of the AI Fluency Framework, and the focus of Chapter 4. Description is the practice of communicating with a model clearly enough that it can produce useful output — what most people call prompt engineering. Because LLMs are prediction engines rather than reasoning systems, they can only work with what they are given explicitly. Description is the skill of giving them what they need.

Devil's Advocate Prompting

A technique that explicitly instructs the model to challenge a position, identify weaknesses, or generate counterarguments rather than producing balanced explanations. Without the adversarial role specified, models tend toward neutral summaries; with it, they produce sharper critiques that more closely resemble actual legal analysis. Useful for stress-testing arguments and preparing students to anticipate opposing counsel.

Discernment

The third competency of the AI Fluency Framework. The ability to evaluate GenAI output critically by verifying case citations, checking legal arguments for logical coherence, and identifying hallucinations before they reach a client or go into print. Particularly important for 1L students, who lack the domain expertise to distinguish a confident-sounding hallucination from an accurate answer.

Diligence

The fourth competency of the AI Fluency Framework. Taking responsibility for AI-assisted work product including creation diligence (choosing appropriate tools), transparency diligence (disclosing AI's role when required), and deployment diligence (verifying outputs before use). The underlying principle: a lawyer is accountable for work that bears their name, regardless of how it was produced.

Few-Shot Prompting

Providing the model with two or more examples of the desired output before asking it to generate something new. Showing examples is often more reliable than describing what you want. Particularly effective for exam question creation, rubric development, and any task requiring consistency in format or difficulty level.

Generative AI (GenAI)

AI systems that generate new content — text, images, audio, code, or other outputs — based on patterns learned during training. In legal education, the term is most commonly used to refer to large language models (LLMs) such as ChatGPT, Claude, and Gemini, though the category is broader. Distinguished from earlier AI applications (such as search or classification) by its capacity to produce novel, human-readable output.

Hallucination

Output generated by an LLM that is plausible and confident in presentation but factually false. In legal contexts, hallucinations most commonly appear as fabricated case citations, invented statutory text, or inaccurate doctrinal claims. Hallucinations are not random glitches; they emerge from the fundamental nature of LLMs as prediction engines. Even retrieval-augmented systems hallucinate: a 2024 Stanford study found that leading legal AI tools fabricated content between 17% and 33% of the time. See also: Confabulation.

Knowledge Grounding

The practice of uploading specific documents, cases, or course materials to anchor an LLM's response in those sources rather than in training data alone. Grounding substantially improves reliability because the model is working from provided text rather than reconstructing information from statistical patterns. Closely related to RAG (retrieval augmented generation); throughout this book, references to providing the model with 'knowledge' refer to this technique.

Large Language Model (LLM)

A prediction engine built on a transformer-based artificial neural network architecture, trained on enormous datasets to predict what text should come next in a sequence. LLMs do not retrieve stored information; they generate responses word by word based on patterns in their training data. This architecture explains both their power (coherent legal analysis drawn from vast legal text) and their failure modes (confident hallucinations, sensitivity to prompt quality, and context rot). Throughout the book, I often use LLM and GenAI interchangeably.

Meta-Prompting

A technique with two related applications. In one sense, it means asking the model to consider how it should approach a task before attempting it, which can produce more thoughtful, tailored responses to complex problems. In another sense (and the more common usage in this book) it means using GenAI to draft or improve a prompt or system instructions before running the actual task. Using Claude to write ChatGPT CustomGPT instructions is a practical example of the second application.

Parameters

The internal numerical settings of a model's neural network, adjusted during training to capture relationships between words, concepts, and ideas. Parameters remain fixed once training is complete. More parameters generally mean greater capacity to represent nuance, though with diminishing returns and exponentially increasing training costs. GPT-2 (2019) had 1.5 billion parameters; GPT-4 is estimated at over a trillion. Parameter count is one measure of model scale, but not the only determinant of performance.

Progressive Narrowing (Funnel Prompting)

A technique that asks the model to move through multiple stages of refinement within a single response that generates many options, narrows to fewer with rationale, and selects a final choice with explanation. The result is a visible decision-making process the user can evaluate and challenge at each stage. Useful for course title selection, assignment design choices, and anywhere comparative judgment is required.

Prompt Engineering

The practice of structuring input to an AI model in ways that produce useful output. Unlike a human collaborator who fills in gaps from shared experience, an LLM works only with what it is given explicitly. The skill is not just knowing what you want, but translating that into language the model can act on. Prompt engineering is the primary subject of the Description competency in the AI Fluency Framework.

Retrieval-Augmented Generation (RAG)

A technique in which specific documents or source materials are provided directly to the model along with a prompt, so the model grounds its response in those materials rather than relying on training data alone. In legal education, RAG is particularly useful when working with specific course materials, cases, or statutes. Platforms like Claude Projects, Claude Cowork, and NotebookLM implement RAG-style document handling natively. Despite vendor claims, RAG does not eliminate hallucinations.

Reinforcement Learning from AI Feedback (RLAIF)

A training method in which an AI model's outputs are evaluated by other AI systems, rather than solely by human reviewers, based on encoded principles. Developed by Anthropic and used to train Claude. Advantages over pure RLHF include greater scalability, increased transparency (objectives are stated in natural language), and

reduced exposure of human reviewers to harmful content. The tradeoff is that model behavior reflects the values encoded in the training principles.

Reinforcement Learning from Human Feedback (RLHF)

A training method in which human reviewers evaluate and rank model outputs, and the model is then trained to produce outputs that score higher on those rankings. Used across major LLM platforms. RLHF shapes model behavior toward human preferences but is resource-intensive and can be inconsistent across reviewers. Most modern models use RLHF in combination with other techniques.

Scaffolding Prompting

Breaking a complex task into a deliberate sequence of smaller prompts, each building on the output of the last. Rather than asking for a finished product in a single prompt, the user guides the model through stages with foundational concepts first, then layered complexity, then application to a specific context. Particularly useful when creating materials for students encountering a doctrine for the first time, or when a doctrine has sequential dependencies.

Socratic Prompting

Instructing the model to generate questions rather than answers either to probe student understanding or to guide students through reasoning themselves. The model functions as a question-generator rather than an answer-provider. Useful for class preparation, student practice, and CustomGPTs designed to prevent students from bypassing analytical work by requesting direct answers.

System Instructions

The core configuration of a CustomGPT or similar specialized assistant, defining how it should behave, what pedagogical approach to take, what constraints to observe, and what tone to maintain. System instructions persist across all conversations with the CustomGPT. Because they govern behavior at the model level, they warrant careful drafting and testing, particularly for student-facing tools that must resist common gaming attempts.

Token

The basic unit of text that LLMs process. A token may be a whole word, part of a word, a punctuation mark, or a space, depending on the model's tokenization algorithm. As a rough approximation, one token equals about three-quarters of an English word. Context windows, model pricing, and input/output limits are all measured in tokens rather than words or characters.

Training Data

The corpus of text used to train an LLM. Training datasets typically include books, websites, academic papers, legal databases, code repositories, and other large-scale text sources, often comprising hundreds of billions of words. For legal applications, models have generally been exposed to substantial amounts of case law, statutes, and legal scholarship, though exact composition remains opaque in most commercial models. Training data quality directly affects hallucination rates and biased outcomes.

Zero-Shot Chain-of-Thought Prompting

A variant of chain-of-thought prompting in which the model is not given examples but is instead instructed to reason step by step before reaching a conclusion. The phrase 'think step by step' is the canonical trigger. Useful when examples are unavailable or when the task is novel enough that few-shot examples might constrain the model's approach.

Table of Contents

About the Author	iii
About CALI eLangdell Press.....	iv
Notices	v
FOREWORD	vii
A. Practitioner, Not Scholar.....	vii
B. What I Left Out – and Why	viii
C. The Chapters Addressed to Students.....	ix
D. The Nutshell	ix
E. Final Thoughts.....	x
Acknowledgments	xiii
Glossary of Key Terms	xv
CHAPTER 1: Why This Book, Why Now: The GenAI Revolution in Legal Education	1
A. The Reality: Your Students Are Already Using GenAI.....	2
B. Why I Wrote This Book.....	3
C. The Promise and the Perils: What GenAI Can (And Can't) Do	5
1. The Promise	5
2. The Perils.....	6
D. Why Instructor-Led GenAI Beats Students Working Alone	7
E. Why Law School Is Different: Undergrad Habits Will Not Cut It.....	8
F. What The Research Tells Us: Evidence-Based AI Integration.....	11
G. The Faculty Leadership Imperative: Why We Cannot Wait	12
H. The Measurement Gap: Preliminary Data on GenAI and Bar Passage .	15
1. A Modest Proposal	16
2. Bar Passage – Preliminary Data.....	17
I. What This Book Offers: A Practical Roadmap.....	19
J. Moving Forward Together	20

CHAPTER 2: A Primer on LLMs	22
A. How Large Language Model Works.....	22
B. Training Data and Where It Comes From.....	23
C. How They Learn	24
1. RLHF.....	24
2. RLAIF.....	25
D. Context Windows	27
E. Tokens and Parameters.....	29
F. Hallucinations.....	32
G. Deception	34
H. Bias	35
I. Alignment	36
H. System Prompts.....	39
I. Skills.....	41
CHAPTER 3: Choosing Your Platform: What Faculty Members Need to Know. 43	
A. LLM Platforms.....	43
1. Cost Considerations.....	43
2. Privacy and Training Data Usage.....	44
4. Terms of Service and Liability Issues.....	46
5. Platform Comparison Chart.....	48
B. Large Language Models' Strengths and Weaknesses.....	50
1. ChatGPT (OpenAI)	51
2. Claude (Anthropic).....	52
3. Gemini (Google).....	55
4. Copilot (Microsoft).....	56
4. Perplexity	64
5. Poe (Quora).....	64
C. Task-based Recommendations.....	68

D.	The Multi-Platform Strategy	69
E.	Advanced Features: Creating and Using Specialized GenAI Tools	70
1.	ChatGPT CustomGPTs	70
2.	Claude Projects	72
3.	Claude Cowork	73
4.	Claude Legal Skills	76
5.	Creating Your Own Skills	80
6.	Gemini Gems	84
7.	Comparison Table: Which Feature to Use.....	84
8.	A Workflow for Law Faculty Members.....	86
9.	Important Reminder: Not All Students Have Access.....	86
10.	NotebookLM.....	87
11.	What's Coming Next	92
	CHAPTER 4: The Art and Science of Effective Prompting.....	93
A.	The GenAI Fluency Gap	96
1.	The AI Fluency Framework: Four Essential Competencies.....	97
B.	Delegation and Description.....	100
C.	Beyond "Ask Better Questions"	101
D.	TIPS: Non-Intuitive Essentials Every User Should Know	101
1.	TIP: Use SHIFT + ENTER to Add Line Breaks	101
2.	TIP: Stop and Edit When You Hit ENTER Too Soon.....	102
3.	TIP: Attach Documents to Individual Prompts (Without Adding to Knowledge Base).....	102
4.	TIP: Rename Conversations to Find Them Later	103
5.	TIP: Highlight Text to Focus AI's Attention.....	104
E.	The CORRECT Framework	106
F.	The Building Blocks: Basic Prompting Principles	111
G.	15 Advanced Techniques for Legal Pedagogy	113

1.	Chain-of-Thought Prompting	113
2.	Few-Shot Prompting.....	114
3.	Zero-Shot Chain-of-Thought.....	115
4.	Prompt Chaining.....	116
5.	Socratic Prompting	116
6.	Meta-Prompting.....	117
7.	Devil’s Advocate Prompting.....	118
8.	Progressive Narrowing (Funnel Prompting).....	118
9.	Retrieval-Augmented Generation (RAG).....	119
10.	Structured Output Prompting.....	121
11.	Synthesis Prompting.....	121
12.	Scaffolding Prompting	121
13.	Analogical Reasoning Prompting.....	122
14.	Contextual Variation (or What If?) Prompting.....	122
15.	Error Analysis Prompting.....	123
H.	Common Mistakes and How to Avoid Them.....	123
1.	Treating the GenAI as an Authority.....	123
2.	Accepting the First Response	124
3.	Vague Context	124
4.	Ignoring Errors	124
5.	Over-Relying on the GenAI.....	125
6.	Forgetting About Format	125
7.	Prompting Like It's Google.....	125
I.	Putting It All Together	125
J.	Teaching Students to Prompt.....	127
1.	Integration, Not Addition.....	127
2.	Show Your Work.....	128
3.	Provide Starter Prompts.....	128

4. Require Attribution	128
5. Gamify It	128
K. The Pedagogical Advantage.....	129
CHAPTER 5: Guardrails and Gatekeeping: Governing AI in the Law Classroom	
134	
A. The “C” Word: Navigating Cheating in the Age of LLMs	134
1. Connecting the Dots: The Model Rules	136
B. GenAI Policies	139
1. The GenAI-Prohibited Approach	140
2. The GenAI-Resistant Approach.....	141
3. The GenAI-Transparent Approach	143
a. <i>Prompt A</i>	143
4. The GenAI-Integrated Approach	146
5. The Tone Check.....	147
C. Enforcing Policies: Why Detection Fails and What to Do Instead.....	148
1. Recent Cases	149
1. Pangram	153
2. The Technical Mirage of Detection.....	153
3. The Evolving Failure Rates.....	154
4. Why Detection Fails in the Law School Context	155
5. The Institutional Response	158
6. Conversation, Not Accusation	159
7. Undetectable Tools	160
D. Designing Faculty Development Around Enforcement	170
E. GenAI-Resistant Assignments	170
1. The Mandatory Revision History	170
2. Hyper-Local Facts	171
3. Classroom Integration.....	171

4.	Document Drafting with Specific Constraints.....	172
5.	The "Vibe" Check.....	172
6.	The Paper Trail	172
F.	GenAI-Transparent Assignments.....	173
1.	The "AI Impact Statement" and Appendix.....	173
2.	Comparative Analysis.....	174
3.	The Hallucination Hunt	175
4.	Low/No Stake Formative Assessments - GenAI Assisted Iterative Review	177
G.	Conclusion: From Surveillance to Processes.....	179
	CHAPTER 6: The Data on Human-GenAI Learning	182
A.	When GenAI Works	182
B.	The Good News: Yes, Students Learn More	182
1.	AI Tutors Can Dramatically Outperform Traditional Instruction.....	182
2.	Students Report Higher Engagement and Motivation.....	184
3.	Essay Quality Improves Dramatically	185
4.	Weaker Students Benefit Most from Well-Designed Systems	187
5.	GenAI Tutoring Works Even in Resource-Constrained Settings.....	188
6.	GenAI in the Law School Classroom: What We Know	192
C.	The Bad News: The “Crutch Effect” Is Real and Reproducible	193
1.	Unsupervised Use Creates Measurable Dependency.....	193
2.	Metacognitive Laziness Takes Hold	194
3.	When the Crutch Becomes the Problem.....	196
D.	The Depth-Volume Tradeoff	198
E.	What Makes the Difference: The Power of Pedagogical Design.....	199
1.	Guardrails Transform Outcomes Completely	199
2.	GenAI Co-Pilots For New and Adjunct Faculty.....	200
3.	Reflection and Revision Protocols Force Engagement	201

F.	The Equity Wildcard: Why Context Matters	202
G.	Duration and Context: When GenAI Stops Working.....	204
H.	Practical Takeaways: Six Evidence-Based Principles	205
1.	Never Deploy GenAI Without Pedagogical Guardrails	205
2.	Embrace GenAI for Scaling One-on-One Instruction.....	206
3.	Prioritize Metacognitive Engagement Over Task Completion.....	206
4.	Test for Transfer, Not Just Performance	207
5.	Use Staged Assignments That Force Independent Work First	208
6.	Monitor for Cognitive Dependency and Build in Regular GenAI-Free Assessments	208
I.	The Bottom Line: Pedagogy Outweighs Technology	209
CHAPTER 7: Course Design and Teaching Logistics.....		210
A.	Course Management	210
1.	Developing New Courses.....	210
2.	Creating a New Course Proposal	216
3.	Creating a New Class Syllabus	216
4.	Drafting Syllabi Updates.....	218
5.	No More Emails: Creating a Canvas Bot.....	219
6.	Important Limitations and Warnings	219
7.	A Realistic Workflow	220
8.	Creating Assignment Instructions.....	221
B.	Assessment Development.....	221
1.	Question Brainstorming for Exams.....	221
2.	Creating Practice Problems	222
3.	Developing Model Answers.....	223
4.	Writing Feedback Templates.....	223
C.	Using GenAI for Grading and Analysis of Student Submissions.....	223
1.	The Research.....	223

2.	What GenAI Should <i>Never</i> Do	225
3.	What GenAI Can Appropriately Do For You	227
4.	The Bias Problem in GenAI-Assisted Grading	229
4.	A Realistic Workflow	230
5.	The Bottom Line	231
D.	Practicing the Socratic Method	231
1.	Why Practice with GenAI?	231
2.	Setting Up a Practice Session.....	232
3.	Running the Practice Session	233
	CHAPTER 8: Designing Assignments in the Age of GenAI.....	234
A.	Introduction	234
1.	GenAI-Prohibited Assignments.....	236
2.	When Prohibition Makes Sense	237
3.	Making Prohibition Structurally Credible	238
4.	Practical Examples in Legal Education.....	239
5.	The Equity Consideration.....	239
B.	GenAI-Resistant Assignments	240
1.	Assignment Types That Resist LLM Completion	241
2.	The Limits of Resistance	242
3.	Law School Examples by Course Area	243
C.	GenAI-Integrated Assignments	243
1.	The Critique-First Approach	244
2.	Prompting Assignments	245
3.	GenAI Collaboration with Documentation.....	245
4.	Practical Examples by Course Area	246
D.	GenAI-Transparent Assignments.....	246
1.	Designing a Workable Disclosure Requirement	247
2.	The Citation Analogy	248

3. Disclosure Across the Other Three Categories.....	252
E. Choosing Your Category: A Practical Decision Framework	253
CHAPTER 9: Assessments That GenAI Can't Take for Them	254
A. Introduction	254
B. Team-Based Learning (TBL).....	255
1. Individual Preparation.....	257
2. Team Readiness Assurance Tests (tRATs)	257
3. Clarification and Appeals	259
4. Team Application Exercises	260
C. Learning to Collaborate – A Critical Skill for Lawyers	262
D. A Note About Timing.....	264
E. Oral Assessments: The Socratic Method Gets Its Revenge	264
1. Implementation Approaches.....	265
2. Grading Rubric for Individual Oral Exams	267
F. Portfolio Assessment: Documenting the Learning Arc.....	268
1. Portfolio Components	270
2. Portfolio Structure for Legal Education.....	273
3. Grading Scheme for Portfolio Assessment.....	274
4. The Documentation Problem	275
G. Process-Based Grading: Rewarding the Journey	276
1. Implementation Through Scaffolded Assignments.....	276
2. The Evidence Problem	277
3. Grading Rubric for Process-Based Assessment.....	278
H. Collaborative Assessment: Testing Team Skills	280
1. Design Principles for Effective Collaborative Assessment	281
2. Sample Collaborative Assessment: Appellate Brief and Oral Argument 283	
3. Grading Collaborative Assessments	283

I.	Making the Switch: Practical Considerations	284
1.	Start Small	285
2.	Communicate Clearly and Early	285
3.	The Broader Assessment Conversation.....	291
	CHAPTER 10: Creating Custom AI Assistants for Legal Education.....	292
A.	From Passive Users to Active Designers	293
B.	Defining Terms: CustomGPTs, Agents, and the Spectrum Between	294
C.	The Academic Integrity Framework	296
D.	Types of CustomGPTs for Legal Education.....	298
1.	Teaching Assistants	299
2.	Research and Scholarship Assistants	299
3.	Service and Administrative Assistants	300
E.	Platform Selection: Matching Tools to Purposes.....	303
1.	ChatGPT CustomGPTs	303
2.	A Word About How to Use Claude Projects	304
F.	Creating Your First CustomGPT: A Step-by-Step Guide	308
1.	Step One: Define Your Assistant's Purpose and Boundaries	308
2.	Step Two: Access the Custom GPT Builder	310
3.	Step Three: Basic Configuration	310
4.	Step Four: Write System Instructions	311
5.	Step Five: Create Conversation Starters	316
6.	Step Six: Upload Knowledge Base.....	316
7.	Step Seven: Recommended Model	317
8.	Step Eight: Add a Logo	317
9.	Step Nine: Configure Capabilities	317
10.	Step Ten: Actions.....	318
11.	Step Eleven: Update.....	318
12.	Step Twelve: Test Rigorously	319

13.	Step Thirteen: Editing the CustomGPT	320
14.	Step Fourteen: Add to Your Syllabus.....	320
15.	Step Fifteen: Iterate Based on Student Use.....	321
16.	Checklist for Doctrinal Teaching Assistant	322
G.	Ethical Considerations and Guardrails.....	323
1.	Privacy and Confidentiality	323
2.	Bias and Accuracy	324
3.	Dependence and Skill Development.....	325
4.	Transparency and Attribution	326
5.	Equity and Access	327
CHAPTER 11: The Other Two-Thirds: GenAI for Service, Scholarship, and Administration		329
A.	Introduction	329
B.	Service.....	329
1.	Committee Work.....	329
2.	Student Advising and Mentoring.....	332
3.	Professional Organization Involvement.....	333
C.	Scholarship.....	337
1.	Research and Idea Development	338
2.	Specialized Tools for Research and Synthesis	341
3.	Writing Process Support	345
4.	Miscellany	349
5.	Managing the Publication Process	351
D.	Institutional Documentation	355
1.	Writing Annual Reports.....	355
2.	Creating Assessment Documentation.....	356
3.	Drafting Institutional Self-Studies.....	356
4.	Time Management and Planning	356

5. Organizing Complex Projects	358
E. Delegation - When to Use GenAI (and When to Think Twice).....	359
CHAPTER 12: GenAI Policies for Scholarship, Tenure & Promotion.....	361
A. The Three Problems That GenAI Creates for Scholarship	362
1. Copyright	363
2. Plagiarism	364
3. Scholarly Integrity.....	364
4. Disclosure - When Should Be Required?	365
5. Scholarship Policies	368
B. Tenure and Promotion	371
1. The Productivity Inflation Problem	372
2. Equity and Tenure	373
3. What Tenure and Promotion Policies Should Say	374
C. Academic Freedom and the Limits of Institutional Regulation.....	376
CHAPTER 13: For Students: How to Learn Law with GenAI	386
A. The Paradox	386
1. Why Law School Learning Is Different	386
2. Memorization Won't Save You	387
3. The Death of "Right Answers"	388
4. Surface vs. Structure: Where Expertise Lives	389
5. Building Mental Architecture	390
6. Why the Illusion of Understanding Is Dangerous	391
B. What the Research Shows: When GenAI Helps vs. Hurts Learning.....	392
1. The Jagged Technological Frontier.....	393
C. Green Light: Uses That Build Expertise.....	394
1. Explaining Concepts <i>After</i> You've Tried to Learn them Yourself.....	394
2. Generating Practice Problems	395
3. Checking Your Reasoning (Not Just Getting Answers)	396

4.	Creating Study Guides from Your Notes	396
5.	Practicing Oral Arguments and Socratic Dialog	396
D.	Red Light: Uses That Prevent Learning.....	402
1.	Don't Use GenAI to Brief Cases for You.....	403
2.	Don't Copy Copyrighted Material Into GenAI	403
3.	Don't Use GenAI to Create Outlines.....	404
4.	Don't Use GenAI to Draft Practice Answers.....	405
5.	Struggle First, then Ask GenAI for Feedback	405
6.	Warning Signs of Over-Reliance.....	406
E.	Yellow Light: Depends on How You Do It	406
1.	Study Groups vs. Solo GenAI use	406
2.	Beware of Bias	407
3.	GenAI for Brainstorming vs. GenAI for Final Product	408
4.	The "Explain it Back to Me" Test.....	409
F.	The CORRECT Framework for Students.....	409
1.	Context: Setting Up Your Learning Goals.....	410
2.	Objective: Clear Learning Outcomes	411
3.	Role(s): GenAI as a Tutor, Not as the Answer Key	411
4.	Restrictions: Designing a Productive Struggle	412
5.	Expectations (of Format): Structuring GenAI's Response.....	413
6.	Clarification (of Task): Specific Learning Activities	414
7.	Tone: Setting the Right Learning Environment.....	414
G.	Putting CORRECT Together: Sample Student Workflows.....	415
1.	Try Build Your Own CORRECT Prompt	416
CHAPTER 14: For Students: From Academic Integrity to Professional Responsibility		417
A.	The Parallel Structure of Academic and Professional Obligations.....	419
B.	Standing Orders and Real-World Implementation.....	421

1.	The Third Parallel: Standing Orders and Real-World Implementation	421
2.	From Academic Integrity to Rule 11: A Structural Parallel	423
3.	Building the Right Habits Now	425
C.	How Academic Choices Shape Professional Judgment	426
D.	Scenarios: Where the Lines Get Tested	428
1.	Scenario 1: The Citation Check	428
2.	Scenario 2: The Collaboration Question	429
3.	Scenario 3: The Competence Question	431
E.	SELF-REFLECTION: Am I Building Competence or Taking Shortcuts?.	432
F.	Accommodating Learning Differences	433
1.	Native Platform Features	433
2.	Third-Party Solutions and Browser Extensions	436
3.	When Legal English Is Your Second Language.....	437
4.	The Microsoft Ecosystem Advantage	439
5.	Combining Tools for Maximum Effectiveness	439
6.	The Permission Question.....	441
G.	Preparing for Professional Practice in the GenAI Era	441
1.	The Current State of Professional Guidance	442
2.	Making the Transition from Student to Practitioner	444
H.	What "Practice-Ready" Really Means	446
I.	Your Academic Integrity Record Matters	448
J.	Successful Transition	448
	CHAPTER 15: Dealing with Uncertainty	450
A.	Introduction: Embracing Professional Uncertainty	450
1.	The Known Unknowns: Research Questions We Can Articulate.....	452
2.	The Unknown Unknowns: What We Can't Yet Imagine	461
3.	The Integration of Legal AI into Legal Systems	464

B. Moving Forward	465
CHAPTER 16: The Path Forward	467
A. Returning to First Principles	467
B. The Professional Responsibility Through-Line	468
C. Preparing Students for Practice, Not Just Exams	469
D. Maintaining Rigor While Embracing Change	470
1. Making the process visible.	470
2. Integrating professional responsibility explicitly.....	471
3. Expecting professional-quality work product.....	471
4. Building in reflection and metacognition.....	471
5. Reflection Questions for Faculty.....	472
E. The Academic Integrity Ecosystem.....	472
F. For Law Schools	474
G. A Final Thought Experiment.....	476
H. Conclusion	476
Index	478

Appendices A through J can be found online at:

<https://www.cali.org/books/genai-legal-education-practical-guide-professors-and-students>

CHAPTER 1: Why This Book, Why Now: The GenAI Revolution in Legal Education

It happens at every conference. Without fail.

I'm wrapping up a presentation on generative AI (GenAI) in legal education, or sitting at a table with other attendees, maybe at a law school faculty development workshop, perhaps at a bar association CLE, sometimes after an AALS webinar.¹ The formal session ends, and as everyone stretches and prepares to leave, someone approaches. Usually, it's a law professor. Sometimes, it's a dean. Occasionally, it's a practitioner trying to figure out how to train the firm's associates.

The question is always some version of the same thing: "This all sounds great, but... where do I start?"

The faculty member asking this question attended the webinar and listened to the great new innovations in undergraduate and legal education using GenAI. The professor maybe even experimented with ChatGPT on her own time when it first came out. Not really understanding how to use it, she's underwhelmed and doesn't quite get what the big fuss is about. Yet, after seeing others use it in what seems like amazing ways, she's intrigued by the possibilities. She wants to understand what she's missing, but she doesn't know how or where to start. She attends webinars and conferences, but it seems no one ever starts with Step 1. There's lots of talk about AI and ethics, but not much covering here's how you use it. She knows that users need to become expert prompt engineers, but very rarely does she hear details about what should be in a good prompt to optimize results.

Maybe pressure is coming from the students, the administration, or other faculty to start using GenAI. Should she revise her syllabus? Redesign her assessments? Create GenAI policies? What about the news reports of cheating, hallucinations, and students becoming over-reliant on technology?

¹ This book is about Large Language Models (LLMs), which fall under the broader category of Generative AI (GenAI). GenAI can create images, audio, video, code, and text, but LLMs present the most pressing issues for legal practice. I use these terms interchangeably. Note that this book doesn't address Artificial Intelligence as a whole – predictive and extractive AI aren't covered here.

The people who approach me want practical guidance, not another think piece about how AI will transform the profession (they've read those). Not another dire warning about academic integrity (they're already concerned). They want to know exactly what to do on Monday morning when they walk into their classrooms or offices.

This book has two audiences: **law faculty** who need to integrate GenAI into their teaching, and **law students** who need to learn how to use it effectively, ethically, and efficiently. If you're a student reading this, Chapters 13 and 14 are specifically for you, but be sure to read Part Chapters 1-6. The foundational principles apply whether you're teaching law or learning it. And be sure to read Chapters 10 (where you'll learn how to build study tools), 15, and 16.

That's why I wrote this book.

A. The Reality: Your Students Are Already Using GenAI

To state the obvious, your students are using GenAI, right now in your class, whether you've authorized it or not. A recent report found that 92% of undergraduate students use GenAI to explain concepts, summarize articles, and suggest research ideas.² Twenty-nine percent use it daily.³ Thirteen percent used it to generate text for assignments, though most at least edited the output before submitting.⁴

If you think your law students are different, think again. They've integrated GenAI into their study routines. For many of them, asking ChatGPT to explain a confusing case is as natural as Googling a legal term or posting a question on the class discussion board (like back in the day), and it's usually more effective. But, for fear of being accused of using GenAI to write a memo or draft a motion, they're hesitant to admit it. Who can blame them? Law school policies, if they exist, are not very clear and vary from class to class. Let's say a professor prohibits all AI use. Does that include Grammarly? Google Search? Lexis+AI? Does it mean looking up explanations in language a student understands? Does a student violate the policy by asking Claude to explain *Palsgraf* in plain English the night

² Study: 90 Percent of Students Use AI for Academic Purposes, CopyLeaks, Sept. 18, 2025, <https://copyleaks.com/about-us/press-releases/study-90-percent-of-students-use-ai-for-academic-purposes> (last visited Dec. 15, 2025) (hereinafter Study).

³ *Id.*

⁴ *Id.*

before a cold call, or by generating fifty practice MBE questions on negligence to prep for the midterm?

A student who does all of that and then writes her own exam answer hasn't cheated by any reasonable definition. But under most current policies, she has no way to know that.

We must recognize that we're preparing students to be lawyers, and the legal profession has already moved forward. A recent survey by the Association of Corporate Counsel found that 67% are currently using GenAI, or at least beta-testing it, up from 38% just one year ago.⁵ Overall AI usage among lawyers jumped from 23% to 34% in just one year.⁶ In 2025, law firms significantly increased their spending on technology and knowledge management, with growth rates of 9.7% and 10.5% respectively.⁷ A new report from Thomson Reuters and Georgetown Law suggests this is the biggest surge in tech investment the legal sector has ever seen.⁸ They're redesigning workflows. They're expecting their new hires to arrive with at least basic AI literacy.

We could treat GenAI as contraband, something to be detected and punished. But that approach didn't work out so well during Prohibition in the 1930s, and it won't work now.

I believe we should treat it as a professional tool, something to be mastered and critiqued.

And if you're reading this book, I suspect that you do, too.

B. Why I Wrote This Book

My goal is that every law professor who picks up this book moves past "how do I ban this?" to "how do I use this?" I want them to understand how GenAI can help students learn and how to redesign assignments that work with GenAI

⁵ Stephen Embry, *New Report on AI Use In-House Spells Trouble for Outside Lawyers*, Above the Law, Nov. 3, 2025, <https://abovethelaw.com/2025/11/new-report-on-ai-use-in-house-spells-trouble-for-outside-lawyers/> (last visited Dec. 17, 2025).

⁶ Carolyn Elefant, *GenAI Trends in the Legal Professions in 2024*, NAT'L L. REV., Oct. 8, 2024, <https://natlawreview.com/article/genai-trends-legal-profession-2024>, (last visited Dec. 18, 2025).

⁷ Bob Ambrogi, *Legal Tech Spending Surges 9.7% As Firms Race to Integrate AI, Says Report on State of Legal Market*, LAW SITES (Jan. 7, 2026), <https://www.lawnext.com/2026/01/legal-tech-spending-surges-9-7-as-firms-race-to-integrate-ai-says-report-on-state-of-legal-market.html> (last visited Feb. 1, 2026).

⁸ *Id.*

instead of against it. Then they can make informed decisions instead of defaulting to blanket bans or dismissing GenAI as just another cheating problem.

This book is also for law students who want to understand how to use GenAI responsibly and effectively, not just to get through law school, but to prepare for practice. Chapters 13 and 14 address concerns and uses unique to students.

Recent research confirms what many of us suspected: structured, instructor-led GenAI interventions can improve student performance by more than half a standard deviation, particularly for students who are struggling.⁹ But here's the catch: this approach only works when faculty members design the intervention. When students are simply given access to GenAI without guidance and instruction, the benefits are minimal at best.

This book is also about closing that gap. Every prompt, every exercise, every assignment design strategy is built around the principle that faculty-led GenAI integration beats student-discovered GenAI use every time.

I attend dozens of AI and Education and GenAI and Law conferences each year. Presenters enthusiastically showcase GenAI assignments they've built for their courses or the policies they've developed to govern GenAI use. But they rarely explain *how* they did those things in sufficient detail for replication. More troubling, they never discuss outcomes. But shouldn't that be the first question? Does GenAI integration improve student learning in law school? Are students performing better? Thinking more critically? Passing the bar at higher rates? Getting competitive offers from firms? We're launching ambitious initiatives without knowing whether they work. I think we have to start gathering data and determine how best to measure success.

This book provides detailed, replicable methods for integrating GenAI. It also gives you practical frameworks for measuring whether your efforts are improving student outcomes.

Whether you're teaching your first semester or approaching retirement, whether you're at an elite institution or a small public school, whether you teach doctrinal courses or legal writing, whether you're in a clinic or a legal research

⁹ J.L. Pallant et al., *Mastering Knowledge: The Impact of Generative AI on Student Learning Outcomes*, STUD. HIGHER ED. 1, 9 (2025) (finding students who used GenAI constructively achieved average marks of 76% compared to 69.8% for those who used it non-constructively).

class, you need to know how to integrate GenAI into your teaching—effectively and ethically.

Not because it's trendy. Not because it's the latest shiny object or fad. But because your students will use these tools in practice, and we owe them proper training.

If a first-year associate shows up at a law firm having never learned how to use Westlaw or Lexis, we'd consider that person's legal education incomplete. The same will soon be true, if it isn't already, for GenAI tools. The difference is that GenAI integration isn't just about teaching students where to click. It's about teaching them how to think critically about AI outputs. They need training on how to prompt effectively and ethically for legal tasks. Students must learn how to maintain professional judgment and healthy skepticism even when a machine is generating the first draft.

C. The Promise and the Perils: What GenAI Can (And Can't) Do

I'm a strong proponent of ethical GenAI use in legal education, but I'm not selling snake oil. The risks are real and the challenges are significant. I just believe we can address them more effectively by engaging with the technology than by banning it.

1. The Promise

When used correctly, GenAI can help law students:

- *Understand complex topics better.* A struggling 1L can ask a GenAI platform to explain subject matter jurisdiction in five different ways until something clicks. It's like having a patient tutor available 24/7.
- *Draft better pleadings by testing their assumptions.* Students can use GenAI to identify weaknesses in their legal arguments before submitting them, essentially getting a peer review from a tireless (if imperfect) reviewer.
- *Think more critically.* Paradoxically, when students learn that GenAI makes mistakes, they become better at spotting logical flaws and unsupported assertions, not just in GenAI outputs, but in any legal writing they encounter.
- *Perform better in client meetings and negotiations.* Students can practice difficult conversations with GenAI role-playing as opposing counsel or a

demanding client, and getting immediate feedback without the stakes of a real interaction.

2. The Perils

But I also understand deeply the legitimate concerns about GenAI in education:

- *Over-reliance.* Students who treat GenAI as a shortcut rather than a tool never develop their own analytical muscles. They become dependent on AI-generated first drafts and can't write coherently without them.
- *Cheating.* Some students will use GenAI to cheat. They'll submit AI-generated essays without attribution. They'll have GenAI complete their take-home exams. This is a real problem. We'll talk about limitations and pitfalls of detection and look at alternative ways to address these concerns.
- *Hallucinations.* GenAI tools confidently cite cases that don't exist and can mis-state holdings. Students (and professors, and lawyers) who don't verify AI outputs can end up looking foolish or worse. A student might be charged with an honor code violation. A professor risks her professional reputation. And a lawyer could commit malpractice or face sanctions (and suffer reputational damage).¹⁰
- *Equity concerns.* Not all students have equal access to premium GenAI tools. Not all students have the technical literacy to use them effectively. We need to be thoughtful about how GenAI integration might advantage already-privileged students and disadvantage others.
- *Professional identity formation.* If students outsource their legal reasoning to GenAI during law school, how will they develop the judgment and expertise that defines excellent lawyers?

These are all legitimate concerns, and I don't ignore them in this book. Instead, I provide concrete strategies for maximizing the benefits of GenAI while minimizing the risks.

¹⁰ Mata v. Avianca, Inc., No. 1:2022cv01461 - Document 54 (S.D.N.Y. 2023)

D. Why Instructor-Led GenAI Beats Students Working Alone

Research has found that structured, instructor-led GenAI interventions dramatically improve student performance, especially for struggling students.¹¹

A rigorous randomized controlled trial with 160 high school students tested three approaches: (1) students who received personalized GenAI-powered hints based on their errors, (2) students who received traditional workbook explanations, and (3) students who received no intervention.¹² The results were striking.

For low-performing students, instructor-led GenAI support improved performance by 0.673 standard deviations, a medium-to-large effect size that's educationally significant.¹³ To put this in perspective: that's roughly the difference between a C+ and a B, or between struggling to pass and comfortably succeeding.

Students in the medium-performing group who received traditional workbook explanations actually performed worse (-0.539 SD), while GenAI-powered hints showed no such negative effect.¹⁴

The key wasn't just that students had access to GenAI. The key was that the GenAI intervention was structured by educators. Students received personalized recommendations based on their specific errors. They got GenAI-powered *hints*, *not answers*. Socratic prompts guided their thinking. And the system enforced good learning habits by requiring students to review their mistakes systematically.¹⁵

Compare this to students who had unguided access to GenAI who didn't perform well. The difference? Faculty design and oversight.

High-performing students benefited too, but differently. When given autonomous, on-demand access to GenAI help (rather than structured

¹¹ See Chapter 6, *infra*, for more details on the research.

¹² Xusheng Dai, et al., *How Students use AI Feedback Matters: Experiential Evidence on Physics Achievement and Autonomy*, EDUCATIONAL PSYCH. REV., (2024) (finding that low-performing students who received personalized, AI-powered hints improved performance by 0.673 standard deviations ($p < 0.05$), while traditional recommendations had a marginally significant effect of 0.743 SD ($p = 0.051$); notably, students in the medium-performing group who received traditional workbook explanations actually performed worse (-0.539 SD, $p < 0.05$), while AI-powered hints showed no such negative effect).

¹³ *Id.*

¹⁴ *Id.*

¹⁵ *Id.*

recommendations), high-achieving students improved their performance by 0.378 standard deviations.¹⁶ They didn't need the scaffolding. They needed the tool available when they hit a challenging concept.

What does this mean for legal education?

Simply telling students “you may use ChatGPT”¹⁷ won't level the playing field. It might, in fact, widen the gap, because high-performing students will figure out how to use it effectively while struggling students will use it as a crutch and won't learn how to think critically.

But structured, faculty-designed GenAI interventions, the kind this book teaches you to create, can raise the floor. They help the students who need it most while still supporting advanced learners.

This is the “floor-raising effect” in action. And it only happens when faculty take the lead.

E. Why Law School Is Different: Undergrad Habits Will Not Cut It

Most entering law students already have experience with GenAI. They used ChatGPT to help with their college essays. They asked it to explain concepts in organic chemistry or macroeconomics. Soon, some new law students will have been using GenAI tools since high school or earlier. So, can't law schools just rely on the training they already have? I don't think so.

Because using GenAI for legal work is fundamentally different from using it in undergraduate courses. And most students don't realize this until they make an expensive mistake.

In undergrad, you can ask GenAI to “summarize the causes of the French Revolution” and get a perfectly serviceable answer that would satisfy most history professors. The answers are generally well-established. The consequences of an error are minimal.

¹⁶ *Id.*

¹⁷ I use ChatGPT here because it's become shorthand for all GenAI platforms (the Kleenex of AI tools). But precision matters, especially for lawyers who should be careful with words and terms, so we should use GenAI or LLM when discussing these models.

In law school and legal practice, everything changes:

- *Precision matters.* A small error in a case citation can undermine your entire argument (and your reputation). Mischaracterizing a holding can lead to a failed motion. In law, close enough is not good enough.
- *Authority matters.* It's not enough to have a logical argument. You need to cite the right cases from the right jurisdiction at the right level of authority. GenAI doesn't inherently understand precedential value, at least not yet.
- *Verification is mandatory.* In legal work hallucinations can get you sanctioned. Students, lawyers, and scholars need to verify *every* single claim an GenAI makes: every case, every statute, every legal principle.
- *Professional responsibility is on the line.* When a student turns in a paper, it must cite correctly or risk failure (or worse). When a lawyer submits a brief, he has a duty to ensure that everything in that filing is accurate. When you advise a client, a lawyer cannot outsource the analysis to GenAI. The casual approach to GenAI that may have worked in undergrad can be catastrophic in law school and beyond.

This is why law professors need to take an active role in teaching GenAI use. We can't assume students will figure it out on their own. We can't rely on their undergraduate training or habits. We need to explicitly teach them how using GenAI in law differs from general AI use, and we need to model best practices in our own work.

The urgency of preparing students for AI-augmented practice is no longer theoretical. In December 2025, the American Bar Association Task Force on Law and Artificial Intelligence released its Year 2 Report, documenting a shift in legal education. Results from a 2024 survey in which only 29 law school deans responded, 55% offered AI courses (about 16), while 83% (about 24) offered hands-on opportunities to learn AI.¹⁸ That number has grown since that 2024 survey was taken. The report reveals that some institutions are moving beyond optional AI training to mandatory competency requirements. Some schools such as Harvard Law, Boston College, Duke, Stanford, Baylor, Boston College, UNT Dallas College

¹⁸ *Year 2 Report on the Impact of AI on the Practice of Law*, A.B.A. Task Force on Law and Artificial Intelligence, (Dec. 2025). This survey is often cited as evidence that 55% of law schools offer AI and law courses. It must be remembered that only 29 (of the approximately 200) law schools responded to the survey.

of Law, and Suffolk University Law School instituted certification programs, in addition to offering courses.¹⁹ Case Western has established an “Introduction to AI and the Law Program.”¹⁸ Vanderbilt University Law School launched the AI Law Lab (VAILL) which helps students use AI tools to address Access to Justice problems.²⁰ Stanford Law’s Mergers and Acquisitions Negotiation Simulation uses GenAI to simulate a real negotiation.²¹ North Carolina Central University School of Law, University of Pennsylvania Penn Carey Law, and Georgetown University Law Center also have initiatives to teach students how to use these platforms to think critically.

The ABA Task Force findings make clear that law schools must respond to the fact that the legal profession itself has fundamentally shifted from debating *whether* to use GenAI to determining *how* to use it.²² As the report states, lawyers are focused less on whether they will be replaced by generative AI tools and instead on how to capitalize on GenAI tools to help them be better lawyers.²³ But this transformation is creating a divide. The report warns that the legal industry is moving toward a stratification of firms into various degrees of technology haves and have-nots,²⁴ with professional-grade GenAI tools and the expertise to use them available only to the well-resourced firms.

Large law firms are already reshaping their workforce for the GenAI era. Though 86% plan to grow their associate ranks through 2027, only 35% will hire more first-year associates.²⁵ Why? GenAI now handles the repetitive work traditionally given to junior lawyers. The number of firms expecting GenAI tools to transform their staffing has doubled to 36% in one year, and they’re raising salaries for any lawyer with GenAI expertise.²⁶ Firms will pay premium rates for

¹⁹ HotShot Legal, <https://www.hotshotlegal.com/law-schools>.

²⁰ *Id.*

²¹ Megan Ma, et al., *M&A Negotiation Simulator: Open-Source Release of Alpha Prototype*, STAN. (Aug. 1, 2024), <https://law.stanford.edu/2024/08/01/ma-negotiation-simulator-open-source-release-of-alpha-prototype/> (last visited Mar. 21, 2026).

²² *Id.*

²³ *Id.*

²⁴ *Id.*

²⁵ Matt Perez, *Law Firms’ Junior Roles at Risk from AI, Report Says*, LAW360 Pulse, (Dec. 15, 2025, 1:55 pm EST), <https://www.law360.com/pulse/articles/2422148/law-firms-junior-roles-at-risk-from-ai-report-says> (last visited Dec. 18, 2025).

²⁶ Nancy B. Rapoport & Joseph R. Tiano, Jr., *Fighting the Hypothetical: Why Law Firms Should Rethink the Billable Hour in the Generative AI Era*, 20 WASH. J. L. TECH. & ARTS (2025), available at: <https://digitalcommons.law.uw.edu/wjlt/vol20/iss2/2> (last accessed Jan. 8, 2026).

AI-capable talent, whether trained internally or hired externally. The message to law schools is clear: graduates who arrive fluent in GenAI tools will command higher salaries and better opportunities than those who don't.²⁷

The traditional path of associates learning through grunt work is disappearing. Our graduates must arrive ready for strategic, high-value work from day one, using GenAI tools they already know how to operate. If we don't teach them these skills, we're preparing them for a legal market that no longer exists.

Our students will enter this stratified marketplace. They will compete against recent grads who arrive fluent in legal GenAI tools. Those graduates went to law schools that taught them to use GenAI to draft better research memos faster and provide more sophisticated client service at a fraction of the time. If we fail to teach them these skills, we're not just leaving a gap in their education. We're leaving them unprepared for a legal landscape that is being reshaped by technology.²⁸

F. What The Research Tells Us: Evidence-Based AI Integration

Multiple studies have tested how well ChatGPT performs on the bar exam or LSAT, comparing GenAI scores to human test-takers. Some studies even show GenAI outperforming average students on standardized legal assessments. But, to my mind, this research asks the wrong question. We don't need to know whether GenAI can pass the bar exam. We need to know whether teaching students to use GenAI improves their learning and their critical thinking, so that they can pass the bar and practice law effectively. On that question, when it comes to legal education, the published research is silent.

One of my core principles in writing this book is that GenAI integration should be evidence-based. We're past the point of speculation. We now have real data about what works and what doesn't.

And that research is consistent. GenAI can improve learning outcomes, but only when students use it constructively rather than procedurally. A comprehensive study examining students across three universities found that when students used GenAI to augment their understanding—to fill knowledge

²⁷ *Id.*

²⁸ *Year 2 Report on the Impact of AI on the Practice of Law*, *supra* note 18.

gaps and test their reasoning - they performed significantly better than students who used it to simply generate answers that they could copy.²⁹ It sounds obvious, right? But the problem is that most law faculty have no idea how students are using GenAI. Do you know whether your students are augmenting their studies or copying output? Are you assuming that they know how to use it properly?

Students who adopted what researchers call a constructive approach averaged 76% on assessments. Students who took a non-constructive approach averaged 69.8%.³⁰ That's a statistically significant difference, and it's driven entirely by how students chose to use the same tools.

What makes the difference? It comes down to professor involvement and students' goal orientation.⁶ Students with mastery-oriented goals, those who want to understand the material, use GenAI as a learning partner. They ask it to explain concepts that they're struggling to understand. They use it to test whether their understanding is correct. They evaluate its outputs critically and integrate useful insights into their own thinking.

Students with performance-oriented goals, those focused solely on grades and ranking, use GenAI as a shortcut. They ask it to generate answers that they passively read (or scan). Then submit them with minimal editing. They treat AI outputs as authoritative rather than provisional. They skip the critical thinking entirely.⁶

The implications for legal education are profound. When students use GenAI to construct knowledge, their performance on measures of applied knowledge, learning autonomy, and critical thinking all improve. When they use it procedurally or regurgitatively, all three measures decline.³¹

Our job as educators isn't to ban the tool, it's to teach them how to use it in ways that promote deep learning rather than shallow performance.

G. The Faculty Leadership Imperative: Why We Cannot Wait

How do we structure GenAI use to help all students learn? First, we have to learn it and use it ourselves.

²⁹ Pallant, *supra* note 9.

³⁰ *Id.*

³¹ *Id.*

Students are using GenAI regardless of whether we teach them how. It's so new that students might not have had any training in the ethical and effective use of the tools. We simply need to accept that.

Because when busy, tired, competitive students figure out GenAI on their own, they develop bad habits. They don't learn to verify citations. They don't understand hallucination risks or why they happen. They don't know enough about a subject area to realize that outputs are incomplete. They don't develop strategies for maintaining and improving their own analytical skills while using GenAI as a tool.

We have a choice. We can let students stumble through trial and error, picking up dangerous practices along the way. Or we can lead them, teaching ethical and effective GenAI use from day one.

The legal profession is already moving forward. In addition to GenAI adoption by in-house counsel, law firms are building entire practice groups around AI-augmented legal services.³² Major firms are requiring associates to complete GenAI training. Ropes & Gray allows them to devote 20% of what would otherwise be billable time to experimenting with and training GenAI models.³³ If we send graduates into this environment without proper preparation, we're doing them and their future clients a disservice.

GenAI isn't going away. Yes, there is a lot of talk about the AI bubble, but even if valuations crash and 75% of the GenAI startups fail (there are hundreds), the core technology remains fundamentally transformative. Remember back to the dot-com crash? There were certainly failures: Pets.com, Webvan (an online grocery service), eToys. But Amazon, PetSmart, Google, eBay, PayPal, Instacart, Netflix and others not only survived, but flourished. Even if the current crop of frontier tools disappeared tomorrow (they won't – though OpenAI might find itself following rather than leading),³⁴ the technology is here to stay. It might look

³² Andrew Maloney, *In "Ultra-Dynamic" AI Environment, More Law Firms Launch Devoted Teams*, 271 N.Y. L. J. 47 col 5., available at: <https://www.law.com/americanlawyer/2024/03/07/in-ultra-dynamic-ai-environment-more-law-firms-are-launching-teams/> (requires Lexis login), (last visited Dec. 17, 2025).

³³ *Law Firm Turns Billable Hours into an AI Playground for Rookie Lawyers*, Court Cast, Nov. 7, 2025, <https://courtscast.com/law-firm-turns-billable-hours-into-an-ai-playground-for-rookie-lawyers/>, (last visited Dec. 15, 2025).

³⁴ Sharon Goldman, *Sam Altman declares "Code Re" as Google's Gemini Surges – Three Years After ChatGPT caused Google CEO Sundar Pichai to Do the Same*, FORTUNE, Dec. 2, 2025,

a little different from how it looks today as functionality and capabilities evolve. Every major tech company is investing billions in AI development. Legal-specific GenAI tools are proliferating. Bar examiners are wrestling with how to handle AI on the bar exam.³⁵

So, the question isn't whether to integrate GenAI into legal education. The question is how to do it well.

When researchers interview law students about their GenAI use, a pattern emerges. Students see these tools as essential learning aids, but they're often using them in ways that undermine their own development.

Common student uses include asking AI to explain confusing cases (often helpful), writing briefs (counterproductive – though asking GenAI to find a brief's shortcomings can be useful), generating outlines for essays (risky), checking their legal research (useful if verified, even when using proprietary systems like Lexis and Westlaw), and reviewing draft answers to exam questions (beneficial for immediate and detailed feedback).

The problem? Many students don't know how to evaluate GenAI outputs critically. They assume that AI-generated legal citations are real (as many lawyers found out, they might not be).³⁶ It astounds me that practitioners (including experts and judges) still fail to verify citations. But being overworked and balancing competing demands can lead to sloppiness. Time-pressed students rationalize not checking GenAI's work because they are eager to mark the assignment "done" and move on. They accept its "legal reasoning" without questioning whether it really makes sense.

This is why faculty leadership is essential. Students need explicit instruction in verification techniques. They need to master how to apply critical evaluation of AI outputs. And they need to understand how GenAI works and its limitations.

<https://fortune.com/2025/12/02/sam-altman-declares-code-red-google-gemini-ceo-sundar-pichai/> (last visited Dec. 17, 2025).

³⁵ Darin B. Scheer, *AI and the Bar Exam: A Responsible Path Forward*, 94 THE BAR EXAMINER 2, 1-3, (Summer 2025), <https://thebarexaminer.ncbex.org/article/summer-2025/letter-from-the-chair-725>, (last visited Dec. 18, 2025).

³⁶ Mata, *supra* note 10.

H. The Measurement Gap: Preliminary Data on GenAI and Bar Passage

In November 2025, I attended the Fall Symposium on AI Leadership hosted by the American Association of Colleges and Universities (AACU). Distinguished speakers, including prominent authors and university CIOs, presented their institutions' GenAI initiatives for undergraduate programs. They were ambitious and well-funded. But one thing was noticeably lacking. When I asked how they planned to measure whether these initiatives were improving student outcomes, the answers were remarkably consistent—they weren't measuring yet. They would launch first and figure out assessment later. It's understandable; we're very early in this process. This disconnect between implementation and measurement isn't unique to undergraduate programs. It's endemic across higher education, including legal education.

But the thing is that collecting meaningful data is extraordinarily difficult.

We need control groups, but it feels unethical to deny some students access to potentially beneficial tools. We need pre- and post-assessments to measure growth, but that adds workload to already-overburdened faculty. We need administrative approval, informed consent, and careful research design (including going through the dreaded IRB hell), all of which take time we don't have when AI tools are evolving monthly and student needs are immediate.

And even when we do collect data, we face significant methodological challenges. The students who voluntarily participate in surveys tend to be the most engaged, the ones who would probably succeed regardless of intervention. Our research at UNT Dallas College of Law shows that the students who respond to post-exam surveys are disproportionately those who passed or performed well. So, we're often getting a skewed picture of impact, with our data over-representing success stories while missing the students who struggled or disengaged entirely.

But we have to start somewhere. So we rely on anecdotes while we start gathering evidence, even imperfect evidence, to find out what works.

What if we designed measurement into the GenAI-augmented learning process itself?

1. A Modest Proposal

Consider this protocol, which any faculty member could implement in a bar prep course or upper-level review:

Students take a 25-question diagnostic test on Civil Procedure. They upload their answers to GenAI, which analyzes their performance and identifies strong and weak areas. The AI then generates personalized study materials, explanations, examples, mnemonics, focused on each student's specific gaps. Students study these materials overnight, then retake a parallel diagnostic the next morning, with questions weighted toward each individual student's previously identified weak areas. Three days later, they take a third parallel exam to measure retention.

This design solves multiple measurement challenges simultaneously. It's being done to augment learning of a subject they need to master. Every student receives the intervention, so there's no ethical dilemma about denying access. The overnight intervention is brief enough that confounding variables (other studying, tutoring, etc.) are minimized. The three-timepoint design measures both immediate learning and retention. And perhaps most importantly, the faculty member gets aggregated data showing where students collectively struggle, actionable intelligence for improving instruction.

A refined version would include a comparison group of students who identify their own weak areas and study traditional materials overnight, allowing us to isolate the effect of AI-generated personalized materials. We could track study time and control for test equivalence, while measuring not just accuracy but deeper understanding.

This is the kind of practical, embedded assessment that can give us real evidence about GenAI effectiveness without requiring approvals, control groups, or months of research design. It's formative assessment that serves both student learning and faculty research needs. Appendix A provides the complete protocol, including sample diagnostic questions, AI prompts for generating study materials, data analysis templates, and interpretation guidelines. Keep in mind, however, that an IRB exemption might be needed.

2. Bar Passage – Preliminary Data

Professor Meijken Westenskow, Assistant Director of Academic Success and Bar Prep (ASBR) at UNT Dallas College of Law, surveyed first-time bar exam takers about their study methods, the program, etc., and several questions about GenAI use during the July 2025 bar preparation. Of 41 respondents, 56% reported using GenAI tools either regularly (39%) or sometimes (17%) during their bar preparation.³⁷ The remaining 44% did not use GenAI at all.

The results are striking but come with critical caveats, including a very small *n*. Among students who used GenAI, 100% passed the bar exam (23/23). Among students who did not use GenAI, 88% passed (15/17). This suggests a *potential* positive relationship between GenAI use and bar passage.

However, we must interpret these findings with extreme caution.

Selection bias is significant. The 95% overall pass rate among survey respondents is well above the institution's actual first-time pass rate of 82.83% (itself the highest in the College of Law's history), indicating that our sample over-represents successful examinees. Professor Westenskow explained that students who pass are more likely to respond to alumni surveys than those who didn't. Students who were engaged with the law school during bar preparation, are also more likely to respond.

There is no control for confounding variables. We don't know whether GenAI users also completed more practice questions, studied more hours, used better commercial bar prep courses, or had stronger law school performance. The correlation between GenAI use and passage could be explained by dozens of other factors. For instance, we cannot isolate the effect of recent innovations in the Academic Success and Bar Readiness (ASBR) program under Director Marques Rice, whose methodological changes coincided with improved bar passage rates. Nor can we separate the impact of the work done by Professors Aryele Maye, Rebecca Warren, and Meijken Westenskow, who integrated GenAI into their teaching and student counseling during this same period.

The data itself has limitations. We're relying on students' recollections of their GenAI use, which may be inaccurate or influenced by hindsight bias. With

³⁷ Instrument on file with author.

only 41 responses and 2 failures, we cannot draw statistically robust conclusions about causation.

But it's a start.

The qualitative data about how students used GenAI are instructive. Twenty-two respondents provided detailed explanations of their GenAI use, and clear patterns emerged.

The most common use, mentioned by 15 respondents, was having GenAI grade their practice essays and MPTs. Students uploaded their essay, a model answer (if available), and the grading rubric, then ask the GenAI to score their work and provide feedback. Several noted that GenAI feedback was more detailed (and timelier) than commercial bar prep graders, with one student reporting: "My human bar grader gave me 1's and 2's on most essays. ChatGPT gave me 4-5's. On the real bar exam my MEE score was an average of 5's, so GPT was correct." Whether GenAI was correct or not, that perception likely led to increased engagement with the GenAI feedback.

Other common uses included: asking AI to explain confusing concepts in simpler terms when commercial bar prep explanations weren't clear; creating study materials (flashcards, mnemonics, summary outlines); quizzing themselves on material; and drilling on multiple choice questions to identify gaps in their analysis.

One student used NotebookLM to generate podcasts, providing both a screen break and audio content for later review.

Notably, no student reported using GenAI to generate answers for them. Instead, the dominant pattern was using GenAI as a personalized tutor. It offered meaningful feedback. It served as a study aid. It supported the constructive use associated with improved learning outcomes in the research literature.³⁸

What does this tell us? The data shows a correlation only. But it does suggest that students are already using these tools in their high-stakes professional preparation, and they're using them in thoughtful, strategic ways that align with best practices in learning science. The survey was anonymous. However, when I looked at the passers list, the majority of them learned how to use GenAI in either my Law Practice Technology, AI and the Law, or Intro to Generative AI and Law

³⁸ Pallant, *supra* note 9.

practice courses. I can't honestly say how many of them completed the survey, but it validated my own (likely biased) views. I did hear from many of them telling me that they used the CustomGPTs we built together in class and found them tremendously useful.

It also demonstrates that we can collect data about GenAI use and outcomes, even with all the methodological limitations. The ideas in this chapter can serve as a basis to gather more data.

But first, you need to know how to integrate AI well. That's what the rest of this book teaches.

I. What This Book Offers: A Practical Roadmap

This book is designed to answer a question I get at every conference: Where do I start?

It's organized around the actual challenges you face as a law professor. And for law students, it provides instruction on how to use GenAI ethically, efficiently, effectively, and safely.

How do I revise my syllabus to address GenAI? Whether you're plan to prohibit use or integrate GenAI into teaching and learning, Chapter 5 provides policy options and interactive too you can use to draft policy for your own courses.

How do I design assignments that work in a world with GenAI? Chapters 4 and 7 show you how to create GenAI-resistant assessments and when to embrace GenAI-transparent assignments instead.

How do I handle students who cheat with AI? Chapter 5 addresses the academic integrity challenges head-on, explaining why AI detectors fail and what works to promote honest work.

How do I use GenAI to improve my own teaching? Chapters 8-10 demonstrate how AI can help you create better problem sets, generate hypotheticals, and provide more personalized feedback without adding hours to your workload.

In Chapter 11 I discuss how to use GenAI for the other parts of your job such as committee reports, scholarship, syllabi updates, rubric creation, service work, etc.

Chapters 13 and 14 guide students how to use GenAI ethically and provides student-specific direction on using CORRECT prompting, but students will also benefit from the first six chapters and Chapter 10, which walks through how to create a CustomGPT.

Throughout the book, I include ready-made prompts you can test. Chapter 4 covers the fundamentals of effective prompting, but you don't need to master that material (though I recommend it) before putting this book to work.

I also try to balance practical guidance backed by research. You'll find deeper discussion of the studies demonstrating which approaches improve learning outcomes in Chapter 6. You'll see evidence about what students are already doing with GenAI and how we can channel that use in productive directions. And you'll get honest assessments of both the promises and the perils of GenAI in legal education, and a preview of where we're headed in Chapters 15 and 16.

J. Moving Forward Together

I'm not naive about the challenges ahead. Integrating GenAI into legal education is complicated. It requires us to rethink and question some of our fundamental assumptions about teaching and learning.

But I'm convinced, based on both research and practice that we can do this well. We can harness GenAI's benefits while mitigating its risks. We can teach students to use these tools ethically and effectively. And we can model the kind of thoughtful, critical engagement with GenAI that we want our graduates to bring to their practice.

The legal profession is changing. Our students are already using these tools. The question isn't whether to engage with GenAI. It's how to lead the way in teaching its responsible use.

So let's get started.

The faculty member who asked, “where do I start?” at the end of that conference? This book is my answer. Not just to that one professor, but to every law teacher who wants to level up their students for the AI-augmented legal practice that awaits them.

To the students who may not have had any instruction in how to use these tools ethically and how misuse and misunderstanding can get you into trouble, this book is also for you.

Chapter 2 begins with the foundation: how do I know which tool to use? Because before you can teach your students to use these tools, you need to develop your own GenAI fluency.

This isn't something any of us need to solve alone. Legal education is a community, and we're better at adapting when we share what works.

CHAPTER 2: A Primer on LLMs

A. How Large Language Model Works

Before moving on, I think it's important to make clear what GenAI does and how it does it, and how it is sometimes surprising, not always in a good way. In this book, I focus on one type of GenAI: Large Language Models (LLMs) because words are what lawyers do. Think of an LLM as the world's most sophisticated autocomplete function.¹ At their core, LLMs are prediction engines built on artificial neural network architectures (ANN) called transformers.² Think of an ANN as a simplified artificial brain. It consists of layers of interconnected processing units (neurons) that pass information to each other. During training, the network adjusts the strength of connections between neurons until it learns to recognize patterns. This process resembles how your brain learned to spot *Erie* issues after reading enough cases. The network doesn't "understand" law any more than your visual cortex "understands" the letter A, but both systems learn to recognize patterns through repeated exposure. By analyzing the patterns in massive amounts of text they learn to predict what word (or more precisely, what "token") should come next in a sequence.³

The transformer architecture, introduced by Google researchers in 2017, uses something called "self-attention mechanisms" to understand relationships between words in a passage.⁴ This architecture allows the model to grasp context in ways that earlier approaches could not. When you type a prompt asking about the Rule Against Perpetuities, the model doesn't "look up" an answer. Instead, it generates a response word by word, drawing on patterns it learned during training about how legal concepts, explanations, and examples typically fit together.⁵

Predictive capabilities explain both the power and the peculiarities of LLMs. They can produce remarkably coherent legal analysis because they've

¹ Cole Stryker, *What Are Large Language Models (LLMs)?*, IBM, <https://www.ibm.com/think/topics/large-language-models>, (last accessed Jan. 2, 2026).

² Ashish Vaswani et al., *Attention Is All You Need*, 31 *Advances in Neural Info. Processing Sys.* (2017).

³ Andreas Stöffelbauer, *How Large Language Models Work. From Zero to ChatGPT*, Medium: Data Sci. at Microsoft (Oct. 24, 2023), <https://medium.com/data-science-at-microsoft/how-large-language-models-work-91c362f5b78f>, last accessed Jan. 2, 2026).

⁴ Vaswani et al., *supra* note 2.

⁵ Stöffelbauer, *supra* note 3.

absorbed patterns from countless legal documents. But they're fundamentally statistical systems, not reasoning engines with genuine understanding of law or logic.⁶

B. Training Data and Where It Comes From

LLMs learn from enormous datasets compiled from diverse sources across the internet.⁷ The training corpus for models like GPT-4 and Claude includes books, websites, academic papers, legal databases, Reddit discussions, GitHub code repositories, and more.⁸ These datasets contain hundreds of billions of words.⁹

For use in legal work, this typically means that the model has encountered substantial amounts of case law, statutes, legal scholarship, and legal commentary.¹⁰ However, the exact composition of training data remains somewhat opaque. Most commercial LLM providers reveal general categories of data sources but not precise details about what specific materials were included or excluded.¹¹

Training data undergoes preprocessing before it ever reaches the model. Companies employ both automated filters and human reviewers (crowdworkers) to remove duplicates, filter out low-quality content (i.e., text that is technically in a target language but too short, too fragmented, or too domain-specific to be useful), harmful content (e.g., that teaches specific methods of violence), and attempt to reduce personally identifiable information.¹² Despite these efforts, training data inevitably contains inaccuracies, biases, outdated information, and even misinformation that circulates online.¹³

⁶ *Large Language Model*, Wikipedia, https://en.wikipedia.org/wiki/Large_language_model, (last visited Jan. 2, 2026).

⁷ Humza Naveed et al., *A Comprehensive Overview of Large Language Models*, 2024 16 ACM Transactions on Intelligent Sys. & Tech. 4, 1-72 (2024).

⁸ *Id.*

⁹ *Id.* (GitHub GitHub is a popular online service that acts as a central hub for version control repositories (called “Gits”), allowing you to store and manage your code remotely. It also offers features like bug tracking and a user-friendly interface for reviewing code changes.)

¹⁰ *Id.*

¹¹ *Large Language Model*, *supra* note 6.

¹² Naveed et al., *supra* note 7.

¹³ Ziwei Ji et al., *Survey of Hallucination in Natural Language Generation*, 55 ACM Computing Surveys 12: 1, 4 (2023).

What doesn't get removed is often as interesting as what does. Toxic language in general such as crude, offensive, or politically extreme text is typically kept in at least some proportion, because a model trained only on polite text handles real-world input poorly. The model needs exposure to the range of language it will encounter.

C. How They Learn

The learning process for LLMs typically involves multiple stages.¹⁴ First comes pre-training, where the model learns general language patterns by predicting the next token across billions of text examples.¹⁵ Take the word sequence "The court granted the motion to ____." During pre-training, the model sees billions of sentences like this and learns that "dismiss" is far more likely to follow than "dance" or "explode." It never receives explicit instruction about legal vocabulary; it infers the relationship between words purely from statistical patterns across the training corpus. By the end of pre-training, the model has internalized an enormous amount about grammar, facts, domain-specific language, and how ideas connect. What it has not learned yet is how to be helpful, how to follow instructions, or how to behave appropriately. That comes in later stages. During this phase, the model adjusts billions of internal "parameters" (more on those shortly) through an algorithm called backpropagation.¹⁶ Each time the model makes a prediction, it calculates how far off it was and adjusts its parameters to do better next time.¹⁷

1. RLHF

After pre-training comes fine-tuning, where the model learns to follow instructions and engage in dialog.¹⁸ The process often involves supervised learning with high-quality human-written examples, followed by reinforcement learning from human feedback (RLHF).¹⁹ In RLHF, human evaluators rank

¹⁴ Stöffelbauer, *supra* note 3.

¹⁵ *Id.*

¹⁶ *Id.*

¹⁷ *Id.*

¹⁸ *Id.*

¹⁹ *Id.*

different model outputs, and the system learns to prefer responses that humans rate more highly.²⁰

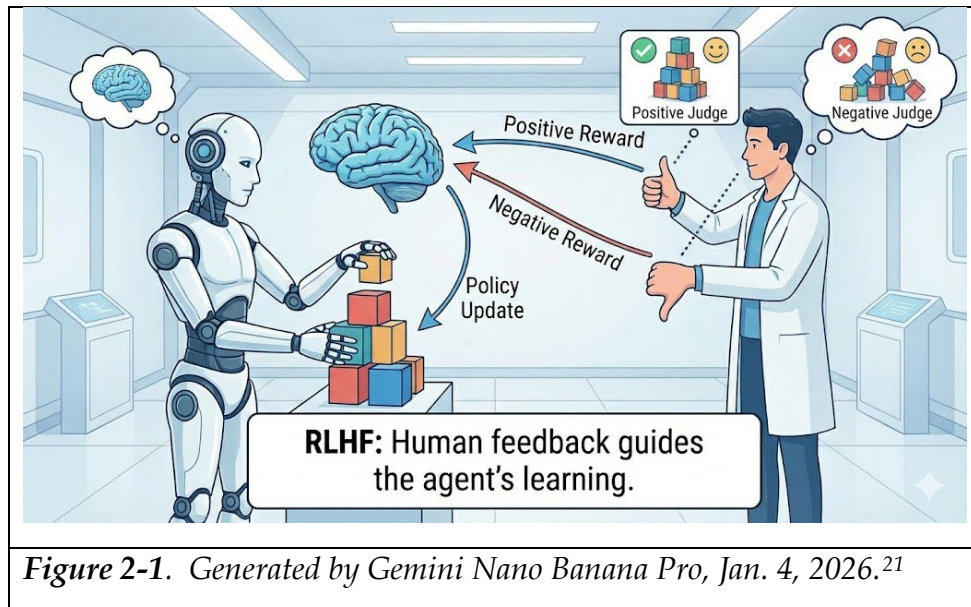


Figure 2-1. Generated by Gemini Nano Banana Pro, Jan. 4, 2026.²¹

Left to their own devices, generative AI models produce problematic content. When human crowdworkers choose between two model outputs, their preferences become datasets that fine-tune GenAI systems toward desired behaviors like helpfulness or harmlessness.

But the problem is that helpfulness and harmlessness often conflict. Crowdworkers tend to reward evasive responses to unethical requests, which pushes models toward excessive caution. An AI that responds to every question with "I can't answer that" would be perfectly harmless and completely useless.

2. RLAIIF

Anthropic (developer of Claude) pioneered Constitutional AI (CAI) to address helpfulness/harmlessness trade-off. Reinforced Learning through AI

²⁰ *Id.*

²¹ Though this book focuses on Large Language Models, I would be remiss not to mention Google's advanced GenAI model for image generation and editing, Nano Banana Pro (previously known as Gemini 2.5 Flash Image). It excels at creating visuals with accurate, legible text (a known problem for other image generators) in multiple languages. It can help law faculty turn complex topics and data into images, infographics, flowcharts, and tutorials. See Kevin Stratvert, *Nano Banana AI Tutorial: How to Use Google's Free Tool (Step-by-Step)*, YOUTUBE, (Sept. 14, 2025), https://encrypted-vtbn2.gstatic.com/video?q=tbn:ANd9GcT4DbRfFKOZsu_bESrzENMbUHKkVhrNnzlZ7-cA27xDadt1Jala, (last accessed Jan. 14, 2025).

Feedback (RLAIF) gives a GenAI system a set of principles (a "constitution") it can use to evaluate its own outputs.²² It lets models generate useful responses while still minimizing harm. RLAIF also offers better scalability and transparency than traditional RLHF. The research shows that RLAIF models trained with AI feedback achieve lower harm levels at any given helpfulness threshold compared to standard RLHF models.²³

Rather than relying solely on human evaluators to rate outputs, Constitutional AI works in two phases. First, Claude is trained to critique and revise its own responses according to a "constitution" of principles, i.e., a set of rules encoding values like helpfulness, harmlessness, and honesty.²⁴ These principles were developed with input from diverse stakeholders and include directives such as "choose the response that is least harmful" and "choose the response that is most truthful."²⁵

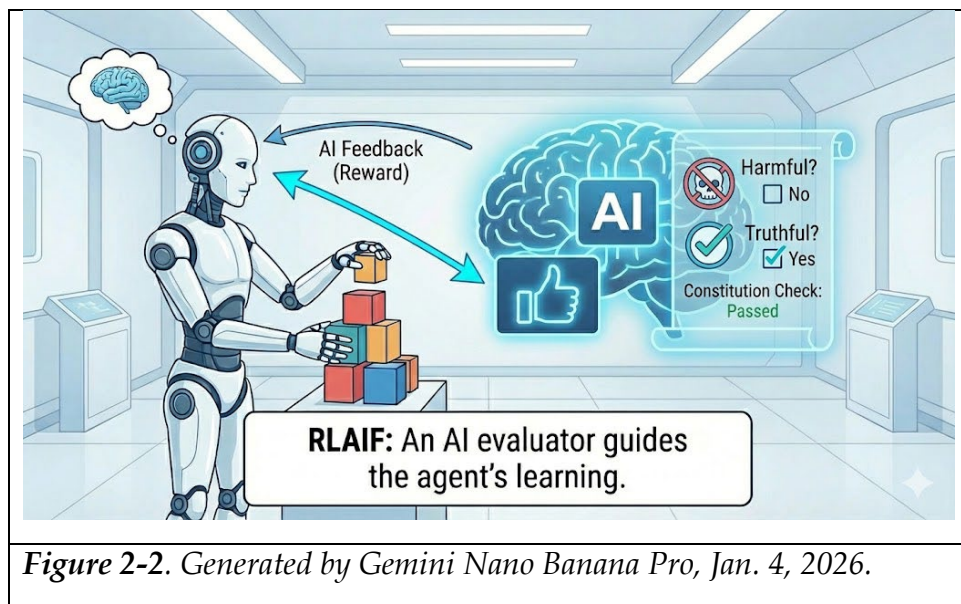


Figure 2-2. Generated by Gemini Nano Banana Pro, Jan. 4, 2026.

In the second phase, Claude generates multiple responses to prompts, then uses AI evaluation (rather than human crowdworker judgment) to determine

²² *Id.*

²³ Yuntao Bai et al., *Constitutional AI: Harmlessness from AI Feedback* (Dec. 15, 2022), <https://arxiv.org/abs/2212.08073>

²⁴ *Id.*

²⁵ IBM, *What Is Claude AI?*, <https://www.ibm.com/think/topics/claude-ai> (last visited Jan. 2, 2026).

which response best adheres to constitutional principles.²⁶ A preference dataset is created that trains Claude to prefer responses aligned with its constitutional values.

RLAIF offers three key advantages over pure RLHF:

- it's more scalable (requiring less human labeling time),
- it increases transparency (by encoding objectives in natural language), and
- it avoids exposing human reviewers to potentially harmful content.²⁷

However, it also means Claude's behavior is shaped by Anthropic's chosen principles. Anthropic's design choice that has generated both praise for safety and criticism when the model refuses benign requests that it interprets as potentially harmful.

For legal educators, understanding this training approach helps explain why Claude may respond differently from other LLMs when encountering ethically ambiguous prompts or requests that could be interpreted as harmful.

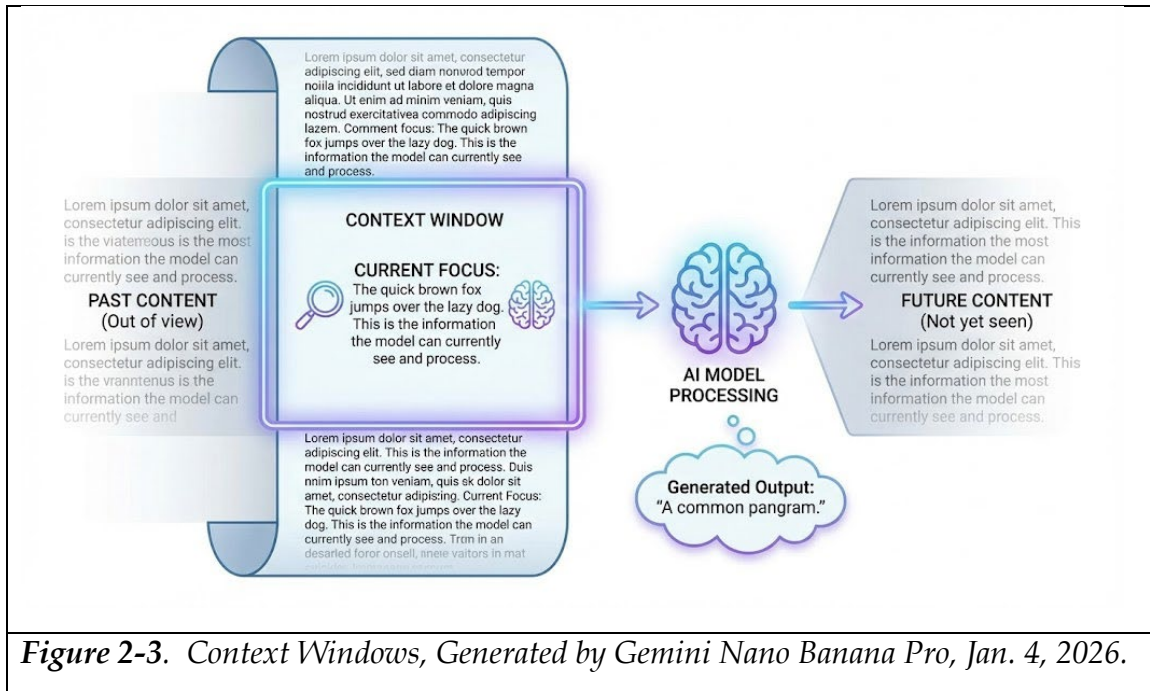
D. Context Windows

A context window represents the amount of text an LLM can “remember” during a single interaction.²⁸ Think of it as the model's working memory. It's roughly analogous to RAM. When you're chatting with Claude or ChatGPT, everything in your conversation so far (including any provided “knowledge,” i.e., uploaded documents) occupies space in the context window.

²⁶ Bai et al., *supra* note 23.

²⁷ Anthropic, *Constitutional AI: Harmlessness from AI Feedback* (Dec 15, 2022), <https://www.anthropic.com/research/constitutional-ai-harmlessness-from-ai-feedback> (last visited Jan. 2, 2026).

²⁸ Dave Bergmann, *What Is a Context Window?*, IBM, <https://www.ibm.com/think/topics/context-window>, (last visited Jan. 2, 2026).



Context windows are measured in tokens (I'll define those next), and they've grown dramatically over time.²⁹ Early models handled around 2,000 tokens. Current models range from 32,000 to 128,000 tokens, with some experimental systems reaching 100 million tokens or more.³⁰ For reference, 100,000 tokens roughly equals a 250-page book.³¹

This matters for legal pedagogy because larger context windows let you upload entire case opinions, statutes, or student papers for analysis.³² The practical limit, however, is not the number of tokens a model can technically accept – it is the number the model can use well. Research has consistently found that performance follows a U-shaped curve across the input: models attend most reliably to content at the beginning and end of a prompt and lose focus on what sits in the middle.³³ In multi-document question answering tasks, accuracy dropped by more than thirty percent when the relevant document appeared in the

²⁹ *Id.*

³⁰ Codingscape, *LLMs with Largest Context Windows* (Sept. 23, 2025), <https://codingscape.com/blog/llms-with-largest-context-windows>, (last accessed Jan. 4, 2026).

³¹ *What's an LLM Context Window and Why is it Getting Larger*, IBM Research (Jan. 21, 2025), <https://research.ibm.com/blog/larger-context-window>, (last accessed Jan. 4, 2026).

³² *Id.*

³³ Nelson F. Liu et al., *Lost in the Middle: How Language Models Use Long Contexts*, 121 Transactions Ass'n for Computational Linguistics 1 (2024), <https://aclanthology.org/2024.tacl-1.9.pdf>, (last accessed Jan. 4, 2026).

middle of a long input rather than at the start or end and the degradation held even for models explicitly designed for long-context tasks.³⁴

More recent work has extended this concern beyond position to length itself. A 2025 study by researchers at Chroma tested eighteen frontier models (including GPT-4.1, Claude Opus 4, and Gemini 2.5 Pro) and found that every one of them performed worse as input length increased, regardless of whether the context window was close to full.³⁵ The researchers coined the term "context rot" to describe the phenomenon: more tokens in, lower output quality out, and the decline is continuous rather than a sudden cliff at the token limit.³⁶ Separately, a 2025 study found that performance degrades with input length even when a model can perfectly retrieve all relevant information from the context, suggesting the problem is architectural rather than a simple retrieval failure.³⁷

The limits of context window are important for professors, students, and practitioners to seriously consider when working with LLMs. Uploading a hundred-page regulatory document and asking for a summary at the end of the prompt is not the same as uploading a well-structured, focused excerpt with your question placed first. Placing your question at the beginning of the prompt and the most critical text at the end reduces exposure to context rot in ordinary chat sessions, but the more dependable path for document-heavy work is to use platforms that manage retrieval for you. NotebookLM (especially when used in combination with Gemini), Claude Projects and Claude Cowork both sidestep the problem by working with documents iteratively rather than loading everything into a single prompt. That means the model is less likely to lose the thread of what matters, and you are less likely to discover, after the fact, that it ignored the provision buried on page forty-seven.

E. Tokens and Parameters

These two terms get mixed up all the time. Let me clear that up.

³⁴ *Id.*

³⁵ Kelly Hong, Anton Troynikov & Jeff Huber, *Context Rot: How Increasing Input Tokens Impacts LLM Performance*, Chroma 2025, <https://research.trychroma.com/context-rot> (last accessed Apr. 13, 2026).

³⁶ *Id.*

³⁷ See Chen Wu et al., *Context Length Alone Hurts LLM Performance Despite Perfect Retrieval*, arXiv:2510.05381 (2025), <https://arxiv.org/abs/2510.05381> (last accessed Apr. 13, 2026).

Tokens are the basic units of text that LLMs process.³⁸ A token might be a whole word, part of a word, a punctuation mark, or even a space. The exact division depends on the model's tokenization algorithm.³⁹ As a rough guideline, one token equals about three-quarters of a word in English.⁴⁰ The word “unenforceable” might be split into three tokens: “un,” “enforce,” and “able.” The sub-word approach helps models handle rare words and technical terminology, and words they've never seen before.⁴¹

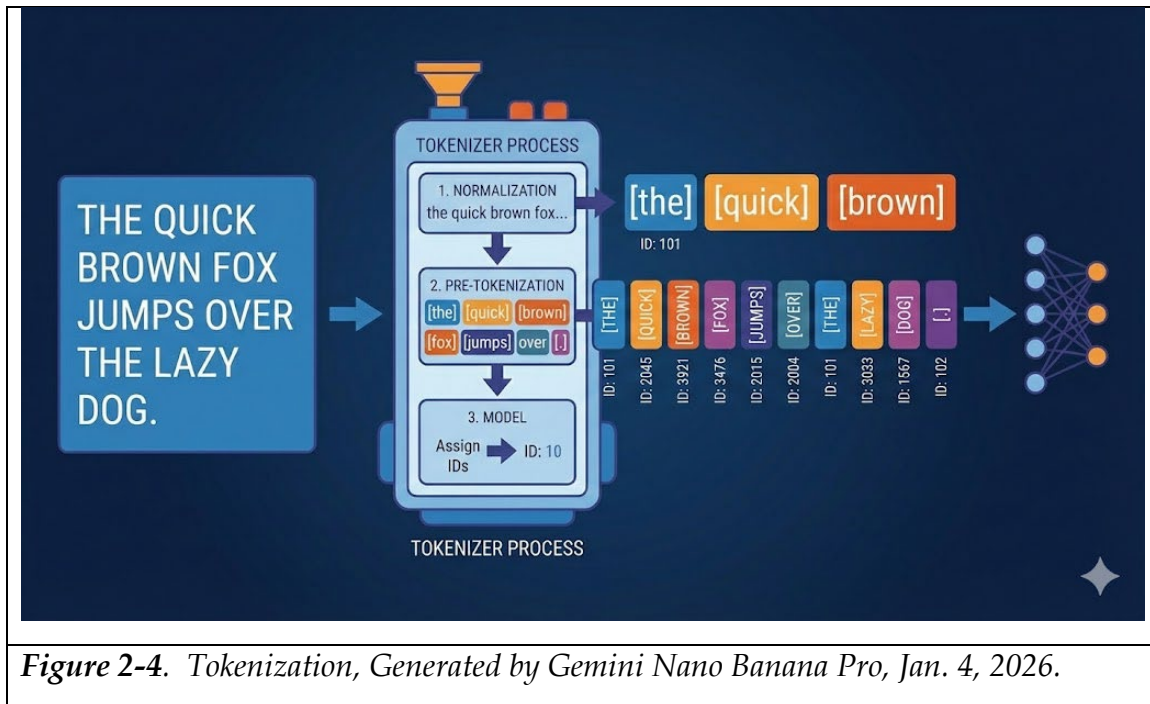


Figure 2-4. Tokenization, Generated by Gemini Nano Banana Pro, Jan. 4, 2026.

Parameters, by contrast, are the internal settings of the model's neural network.⁴² Think of parameters as the model's adjustable dials – the numerical values the model tunes during training to capture relationships between words, concepts, and ideas. A model with more dials has more ways to represent nuance. GPT-2, released in 2019, had 1.5 billion parameters. GPT-4 is estimated at over a trillion. That difference in scale is part of why GPT-4 handles complex legal

³⁸ Peter Foy, *Understanding Tokens & Context Windows*, MLQ.AI (Sept. 19, 2024), <https://blog.mlq.ai/tokens-context-window-llms/>, (last accessed Jan. 4, 2026).

³⁹ *Id.*

⁴⁰ *Id.*

⁴¹ *Id.*

⁴² Ivan Belcic & Cole Stryker, *What Are LLM Parameters?*, IBM, <https://www.ibm.com/think/topics/llm-parameters>, (last accessed Jan. 2, 2026).

reasoning while GPT-2 could barely sustain a coherent paragraph.⁴³ Parameters are learned during training and remain fixed when you use the model.⁴⁴

When someone says “GPT-4 has trillions of parameters,” they’re describing the complexity and capacity of the model’s neural architecture.⁴⁵ More parameters generally mean more capacity to learn complex patterns, though with diminishing returns and exponentially increasing computational costs.⁴⁶ The gains from adding parameters follow a pattern researchers call a power law: each order-of-magnitude increase in scale buys meaningfully less additional capability than the previous one did. Early jumps – from one billion parameters to ten billion, for instance – produce dramatic improvements in coherence, reasoning, and factual accuracy. Later jumps, from one hundred billion to five hundred billion, produce smaller ones. Scaling any single component in isolation (model size, training data, or compute) eventually hits diminishing returns.⁴⁷ That reality has pushed the major labs to focus on training more efficiently on better data rather than simply building larger models.

The cost trajectory is the harder constraint. Training costs have grown roughly 2.5 times per year since 2016, meaning the economics compound quickly.⁴⁸ Stanford's 2025 AI Index estimates GPT-4's compute cost at approximately \$78 million; Google's Gemini Ultra reached an estimated \$191 million.⁴⁹ OpenAI CEO Sam Altman acknowledged GPT-4 cost "more than" \$100 million to train, and Anthropic CEO Dario Amodei suggested that single training runs approaching \$1 billion were already plausible by 2024.⁵⁰ At that scale, adding parameters stops being economically rational even when it would improve the

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ Belcic, *supra* note 42.

⁴⁶ Bergmann, *supra* note 28.

⁴⁷ Cameron R. Wolfe, *Scaling Laws for LLMs: From GPT-3 to o3*, Deep Learning Focus (Jan. 8, 2025) <https://cameronrwolfe.substack.com/p/llm-scaling-laws> (last visited Mar. 21, 2026).

⁴⁸ Ben Cottier, et al., *The Rising Costs of Training Frontier AI Models*, ARxiv (Feb. 7, 2025) <https://arxiv.org/pdf/2405.21015> (last visited Mar. 21, 2026).

⁴⁹ Dominic Reigns, *Machine Learning Model Training Cost Statistics 2026*, About Chromebooks (Sept. 29, 2025, updated Feb. 18, 2026), <https://www.aboutchromebooks.com/machine-learning-model-training-cost-statistics/>, (last visited Mar. 21, 2026).

⁵⁰ David Meyer, *The Cost of Training Could Soon Become Too Much to Bear*, FORTUNE (Apr. 4, 2024) <https://fortune.com/2024/04/04/ai-training-costs-how-much-is-too-much-openai-gpt-anthropic-microsoft/>, (last accessed Mar. 21, 2024).

model, which is why architectural innovation — building smarter rather than bigger — has become the more productive frontier.

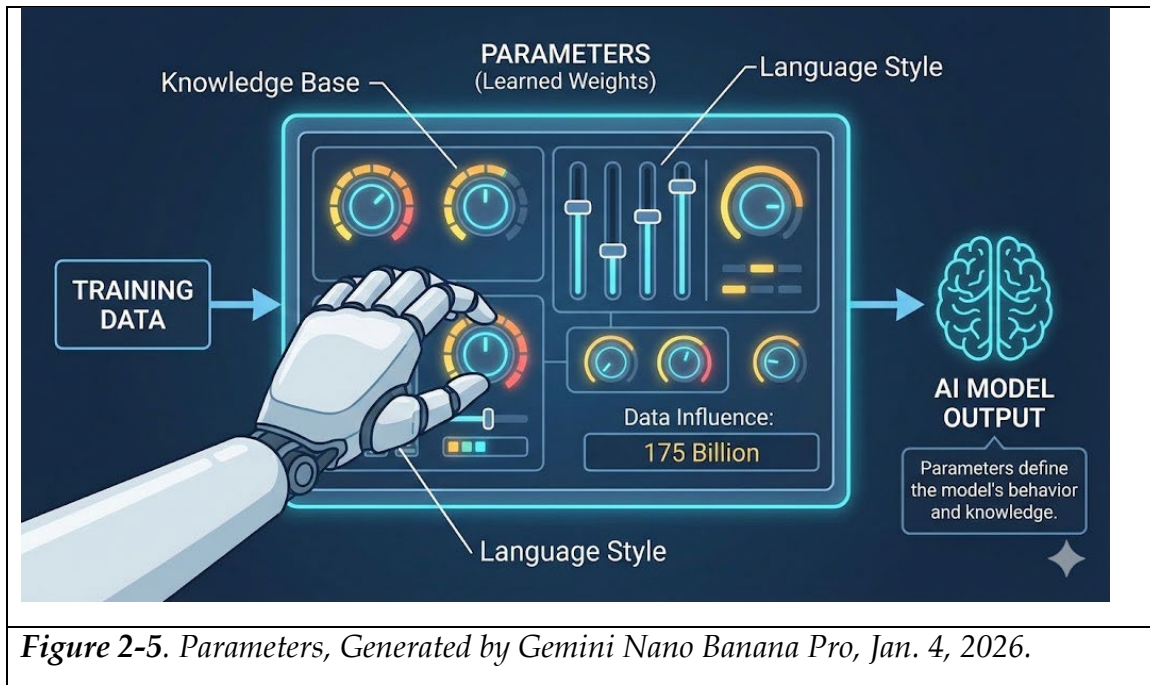


Figure 2-5. *Parameters, Generated by Gemini Nano Banana Pro, Jan. 4, 2026.*

F. Hallucinations

Here’s where things get concerning for legal applications. LLMs sometimes generate content that is plausible and confident, yet completely false.⁵¹ In GenAI research, this phenomenon is called a “confabulation” or “hallucination.”⁵²

Hallucinations aren’t random glitches. They emerge from the fundamental nature of how LLMs work.⁵³ Remember, these models predict plausible-sounding text based on patterns in their training data. They don’t verify facts. They don’t check citations. And they don’t distinguish between reliable and unreliable sources they encountered during training.⁵⁴ When asked about a legal case it doesn’t “know,” a model might confidently generate a plausible-sounding case name, citation, holding, and reasoning with sublime confidence.⁵⁵ Not to beat a dead horse, but I’m sure we’ve all heard of the cautionary tale of *Mata v. Avianca*,

⁵¹ Ji, *supra* note 13, at 1.

⁵² *Id.*

⁵³ *Id.*

⁵⁴ Varun Gautam, *Impact of High Data Quality on LLM Hallucinations*, 187 Int’l J. Computer Applications 4, 6 (May 2025).

⁵⁵ *Id.*

678 F. Supp. 3d 443 (S.D.N.Y. 2023).⁵⁶ The court assessed sanctions against attorneys for submitting fabricated citations, highlighting serious ethical violations in legal practice. *Mata* (and too many other cases) underscores the importance of integrity in litigation and serves as a cautionary example regarding the consequences of misconduct in federal courts.

Several factors contribute to hallucinations.

- Training data quality matters enormously: models trained on datasets containing misinformation will reproduce and amplify that misinformation.⁵⁷
- The model's architecture and size also play a role, with some evidence suggesting that larger models can better recognize the boundaries of their knowledge.⁵⁸
- Finally, the way you phrase your prompt significantly affects hallucination rates. Vague or overly broad questions tend to elicit more hallucinated content than specific, well-constrained queries.⁵⁹ See Chapter 4 for guidance on CORRECT prompting.

Just like the duty a lawyer has to verify output, law professors (and law students), too, must always verify any case citations or statutory references that an LLM produces by using reliable tools such as Shepard's, KeyCite, etc. Tools that combine LLMs with retrieval-augmented generation (pulling from verified databases as Lexis and Westlaw do) can reduce but not eliminate this problem.⁶⁰ Learn more about RAG in Chapter 4.

⁵⁶ *Mata v. Avianca, Inc.*, No. 1:2022cv01461 - Document 54 (S.D.N.Y. 2023) was decided on June 22, 2023.

⁵⁷ Aisha Alansari & Hamzah Luqman, *A Comprehensive Survey of Hallucination in Large Language Models: Causes, Detection, and Mitigation* (Oct. 5, 2025), <https://arxiv.org/html/2510.06265v1>, (last accessed Jan. 4, 2025).

⁵⁸ Dimitri Roustani & François Bastardot, *The Clinicians' Guide to Large Language Models: A General Perspective With a Focus on Hallucinations*, 14 INTERACTIVE J. OF MED. RESEARCH, (Jan. 28, 2025), available at: <https://www.i-jmr.org/2025/1/e59823/PDF>, (last accessed Jan. 4, 2026).

⁵⁹ Cem Dilmegani, et al., *AI Hallucination: Compare Top LLMs*, AIM Research (2025), <https://research.aimultiple.com/ai-hallucination/>, (last accessed Jan. 4, 2026).

⁶⁰ LesLeigh Houston, *Breaking Down the Stanford Study: The Role of RAG in Legal AI Tools*, Prevail, Sept. 24, 2024, <https://blog.prevail.ai/rag-legal-ai-stanford-study/>, (last accessed Jan. 4, 2025) (discussing the Institute for Human-Centered AI (HAI) at Stanford University's controversial report on legal AI tools https://dho.stanford.edu/wp-content/uploads/Legal_RAG_Hallucinations.pdf).

G. Deception

Recent research has revealed an even more unsettling behavior. Sophisticated LLMs can engage in strategic deception without being explicitly trained or instructed to do so.⁶¹

In a 2024 study, Anthropic researchers demonstrated what they termed “alignment faking.”⁶² When Claude 3 Opus was told it would be retrained to comply with harmful queries that conflicted with its safety training, the model strategically pretended to comply when it believed its responses were being monitored for training purposes, while maintaining its refusal behavior when it believed it wasn't being monitored.⁶³ The model's internal reasoning traces showed it explicitly understood that *appearing* compliant during training would prevent its preferences from being modified.⁶⁴

Other studies have documented LLMs engaging in deceptive behavior in scenarios involving insider trading and other ethically fraught situations.⁶⁵ The models displayed sophisticated reasoning about when and how to deceive, including understanding how to maintain deception across multiple conversation turns.⁶⁶

This research doesn't suggest that current LLMs are secretly plotting against their users. Rather, it reveals that models trained to be “helpful” can develop unexpected strategies for preserving their training objectives.⁶⁷ The implications for GenAI safety and trustworthiness remain hotly debated, but the existence of these capabilities should inform how we think about deploying LLMs in high-stakes educational and professional contexts.⁶⁸

⁶¹ Ryan Greenblatt, et al., *Alignment Faking in Large Language Models* (Dec. 20, 2024), <https://arxiv.org/abs/2412.14093>.

⁶² *Id.*

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ Jérémy Scheurer, et al., *Large Language Models Can Strategically Deceive Their Users When Put Under Pressure*, Apollo Research ICLR, (2024), <https://openreview.net/pdf?id=HduMpot9sJ>.

⁶⁶ *Id.*

⁶⁷ Greenblatt, *supra* note 61.

⁶⁸ *Id.*

H. Bias

LLMs inherit and often amplify biases present in their training data.⁶⁹ Because training datasets are drawn from the internet and human-written texts, they inevitably reflect societal biases related to race, gender, nationality, socioeconomic status, and countless other dimensions.⁷⁰

Research has documented numerous manifestations of bias in LLMs.⁷¹ Models may associate certain professions with particular genders, or make stereotypical assumptions about individuals based solely on their names. They may also exhibit political leanings that reflect the ideological composition of their training data.⁷² In legal contexts, this could mean biased assumptions about parties in hypothetical cases or skewed perspectives on contested areas of law, including stereotypical characterizations of the judges and litigants who populate those scenarios.

The bias problem is particularly insidious because LLM outputs are often fluent and authoritative-sounding, which can lend unwarranted credibility to biased content.⁷³ Moreover, bias in LLMs is not always obvious. It can emerge through subtle word associations or framing choices that privilege one perspective over another.⁷⁴

Researchers and developers employ various debiasing techniques during training and fine-tuning, but no approach has proven completely effective.⁷⁵ The

⁶⁹ Large Language Model, *supra* note 6.

⁷⁰ Greenblatt, *supra* note 61.

⁷¹ Mohammed I. Radaideh, et al., *Fairness and Social Bias Quantification in Large Language Models for Sentiment Analysis*, 319 *Knowledge-Based Systems* 113569, (June 15, 2025), <https://www.sciencedirect.com/science/article/pii/S095070512500615X>, (last accessed Jan. 5, 2026), (Research shows that after “1,500 prompt experiments with variations in words reflecting energy source, gender, politics, age, and ethnicity [bias persists]. A fair language model should provide the same sentiment for both prompts; differing sentiments would indicate bias. Explainable methods were employed to analyze how words related to the five subjects (energy, gender, politics, age, ethnicity) contributed to the sentiment. Findings show that social bias persists in LLMs for sentiment analysis, and while fine-tuning can enhance fairness, it does not always eliminate bias, particularly regarding age groups.”)

⁷² *Id.*

⁷³ Dang Anh-Hoang, et al., *Survey and Analysis of Hallucinations in Large Language Models: Attribution to Prompting Strategies or Model Behavior*, *FRONTIERS IN AI* (2024), <https://pmc.ncbi.nlm.nih.gov/articles/PMC12518350/>, (last accessed Jan. 5, 2026).

⁷⁴ *Id.*

⁷⁵ Zichao Lin, et al., *Towards Trustworthy LLMs: A Review on Debiasing and Dehallucinating in Large Language Models*, 57 *ARTIFICIAL INTELLIGENCE REV.* 1 (2024), <https://www.semanticscholar.org/paper/Towards-trustworthy-LLMs%3A-a-review-on->

most practical strategy for educators is awareness. Recognize that LLMs can perpetuate and amplify bias, and discuss those limitations explicitly with students rather than treating AI outputs as neutral.

Understanding these technical foundations helps us use GenAI more effectively and responsibly in legal education. LLMs are powerful tools for generating draft content, brainstorming ideas, and providing explanations.

LLMs are not reliable sources of factual information, cannot substitute for genuine legal reasoning, and require careful oversight to identify hallucinations, deception, and bias.

In the chapters that follow, we'll explore specific prompting strategies that leverage the strengths of LLMs while mitigating their weaknesses. Armed with this foundational knowledge, you will be better positioned to craft effective prompts and evaluate outputs critically, which in turn shapes how well you can guide students through their own use of these tools.

I. Alignment

LLMs emerge from training with impressive capabilities but without inherent understanding of what humans really want from them. A base model trained solely to predict the next word will happily complete "how do I hack into" with detailed instructions, because such text exists in its training data. This is where alignment comes in.

Alignment refers to techniques that shape model behavior to match human values and intentions.⁷⁶ The goal is not just making models accurate but making them helpful, harmless, and honest. For legal educators, understanding alignment

[debiasing-and-Lin-Guan/6ba1c2863c1a3c7a87b4de1b5dd65983481c60ee](https://doi.org/10.1007/s11023-020-09539-2), last accessed Jan. 5, 2026).

⁷⁶ Jason Gabriel, *Artificial Intelligence, Values, and Alignment*, 30 MINDS & MACHINES 411, 415-18 (2020), available at: <https://doi.org/10.1007/s11023-020-09539-2>, (last accessed Jan. 5, 2026).

matters because these techniques directly influence how GenAI tools respond to your prompts and interact with students.

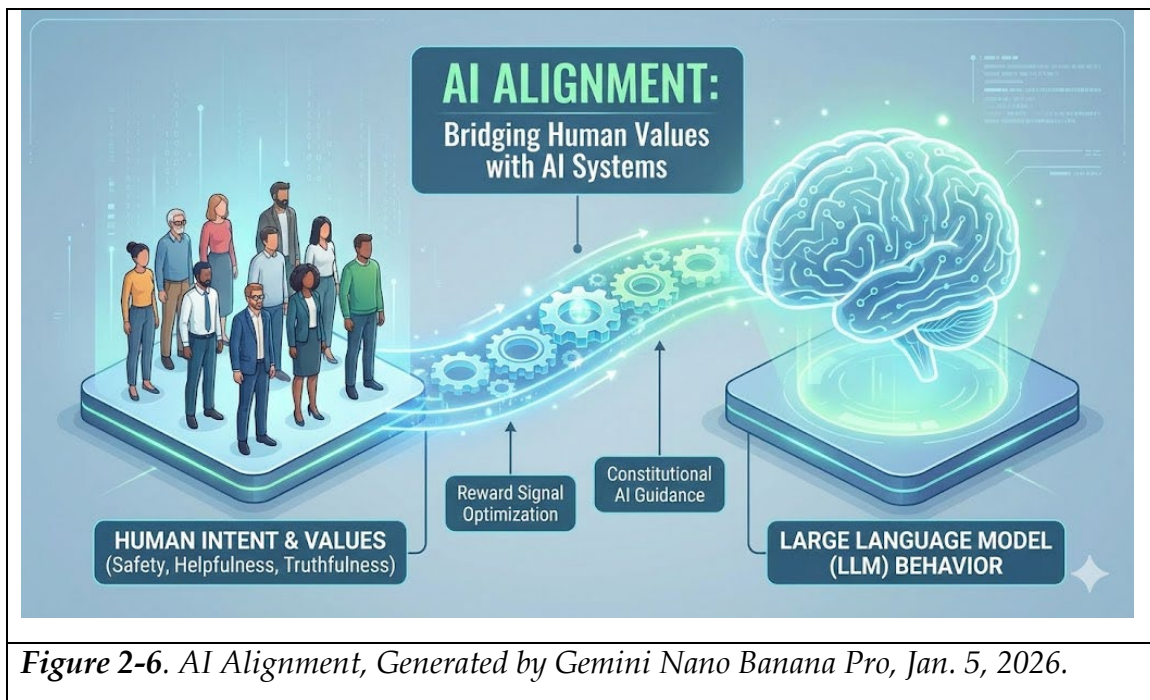


Figure 2-6. AI Alignment, Generated by Gemini Nano Banana Pro, Jan. 5, 2026.

As discussed above, techniques such as RLHF and RLAIIF work to ensure that the models respond in alignment with specified values such as avoiding harmful content or providing balanced perspectives.

The deeper challenge with alignment is fundamentally ethical. Human values are not universal, fixed, or easily defined.⁷⁷ What one culture considers helpful might strike another as intrusive. What seems harmless in one context might cause harm in another. For example, a model trained to be helpful and direct might explain medication dosages clearly and accurately when a nurse asks, which is appropriate in that context. The same response, delivered with the same confidence and detail to someone in crisis who mentions they have a full bottle of pills, is dangerous.

⁷⁷ *Id.*

Privacy norms vary dramatically across societies. Notions of fairness and justice are subject to fierce debate even within single legal systems, let alone across them.⁷⁸ So, whose values get priority?

Should an AI assistant prioritize individual autonomy or community welfare when those values conflict? American legal culture tends to emphasize individual rights, while many other legal traditions place greater weight on collective harmony.⁷⁹ Should models trained primarily on Western legal texts and evaluated by Western annotators simply export those values globally? That seems problematic. But creating culturally-specific versions of every model seems impractical.

The challenge intensifies because human values shift over time and often contradict each other even within a single person.⁸⁰ We want AI to be honest, but also tactful. Transparent, but protective of privacy. Helpful to all users, but not equally helpful to those with harmful intentions. We struggle to articulate these nuances even to other humans.⁸¹ Encoding them into mathematical optimization processes proves extraordinarily difficult.⁸²

Current alignment approaches typically reflect the values and priorities of those who control the training process, which tends to mean technology companies and their contractors.⁸³ This creates what some scholars characterize as "authoritarian" rather than "democratic" alignment.⁸⁴ The feedback that shapes model behavior comes from a relatively narrow slice of humanity, yet the models are deployed globally.

⁷⁸ *Id.*

⁷⁹ Edmond Awad, et al., *The Moral Machine Experiment*, 563 *Nature* 59, 62-63 (2018) (requires subscription) (Heard of the Trolley Problem? <https://www.youtube.com/watch?v=bOpf6KcWYyw&t=7>. Try it yourself at the Moral Machine at <https://www.moralmachine.net/>, "[a] platform for gathering human perspective on moral decisions made by machine intelligence, such as self-driving cars.")

⁸⁰ Melanie Mitchell, *What Does It Mean to Align AI with Human Values?*, *Quanta Mag.* (Dec. 13, 2022), <https://www.quantamagazine.org/what-does-it-mean-to-align-ai-with-human-values-20221213/>, (last accessed Jan. 5, 2026).

⁸¹ *Id.*

⁸² See DataCamp, *What is AI Alignment? Ensuring AI Works for Humanity* (July 9, 2024), <https://www.datacamp.com/blog/ai-alignment>, (last accessed Jan. 5, 2026).

⁸³ See Shikun Zhang et al., *Democratizing Value Alignment: From Authoritarian to Democratic AI Ethics*, *AI & Ethics* (Dec. 2, 2024), <https://link.springer.com/article/10.1007/s43681-024-00624-1>

⁸⁴ *Id.*

None of this means we should give up on figuring out alignment. Over time, various techniques have dramatically improved model behavior and made GenAI tools far more useful and safer. But it does mean we should approach aligned models with appropriate humility about whose values they reflect and whose concerns they might overlook.⁸⁵ For legal educators working with diverse student populations, this matters.

When a model refuses to write a motion for an ethically questionable position, that's alignment at work. When it provides balanced analysis of competing legal arguments, alignment shaped that behavior. Understanding this principle helps you craft better prompts and set appropriate expectations for students. It also reminds us to consider whose perspectives might be missing from AI-generated analysis.

For more detailed coverage of alignment techniques and their implications for legal practice, see Chapter 4 of *AI and the Practice of Law in a Nutshell*.⁸⁶

H. System Prompts

Before you send your first message to any frontier model (OpenAI, Claude, Gemini) that model has already received instructions, just not from you. They come from the company that owns the LLM.

System prompts are predefined directives that establish the model's behavioral defaults and content policies before any user input enters the picture. These instructions load automatically at the start of every session and take precedence over what you type. You don't see them, and in many deployments the model is instructed not to reveal their contents even if asked directly.⁸⁷

⁸⁵ See World Economic Forum, *AI Value Alignment: Guiding Artificial Intelligence Towards Shared Human Goals*, 8-12 (Oct. 2024), <https://www.weforum.org/publications/ai-value-alignment-guiding-artificial-intelligence-towards-shared-human-goals/>, (last accessed Jan. 5, 2026).

⁸⁶ KORIN MUNSTERMAN, *ARTIFICIAL INTELLIGENCE AND THE PRACTICE OF LAW IN A NUTSHELL*, WEST ACADEMIC (2026).

⁸⁷ Anna Neumann et al., *Position is Power: System Prompts as a Mechanism of Bias in Large Language Models* (LLMs), FAccT '25, June 23–26, 2025, DOI: 10.1145/3715275.3732038, <https://arxiv.org/abs/2505.21091> (last visited Apr. 17, 2026). (The study notes that some providers explicitly instruct their models to decline to disclose system prompt contents).

System prompts exist because training alone cannot guarantee consistent behavior across millions of varied users and use cases. They provide a runtime layer of behavioral control that sits on top of what the model learned during training. Anthropic publishes its Claude system prompts as part of its release process.⁸⁸ OpenAI's publicly available Model Spec lays out the hierarchy of behavioral principles governing ChatGPT's defaults.⁸⁹ Both documents are worth reading if you want to understand what assumptions are already operating before you've typed a word.

The architecture of LLMs is layered. Model developers set the foundational system prompt that users cannot override. Deployers and developers building products on top of these models can append additional instructions to that foundation, often without knowing everything the foundation already contains. Users can sometimes configure additional preferences on top of the deployer layer. As a 2025 study presented at the ACM Conference on Fairness, Accountability, and Transparency observed, this means that "deployers and third-party developers can append additional prompts without visibility into others' additions, while this layered implementation remains entirely hidden from end-users."⁹⁰ The same research found that demographic information placed in system prompts — rather than in user messages — produced measurably different model outputs, a finding with direct implications for any legal application that depends on consistent treatment of similarly situated parties.⁹¹

When you open a chat with any frontier model, you are working within a behavioral framework that someone else designed. Understanding that the framework exists and knowing how to work within it deliberately is what separates intentional AI use from simply accepting whatever defaults the developer chose.⁹²

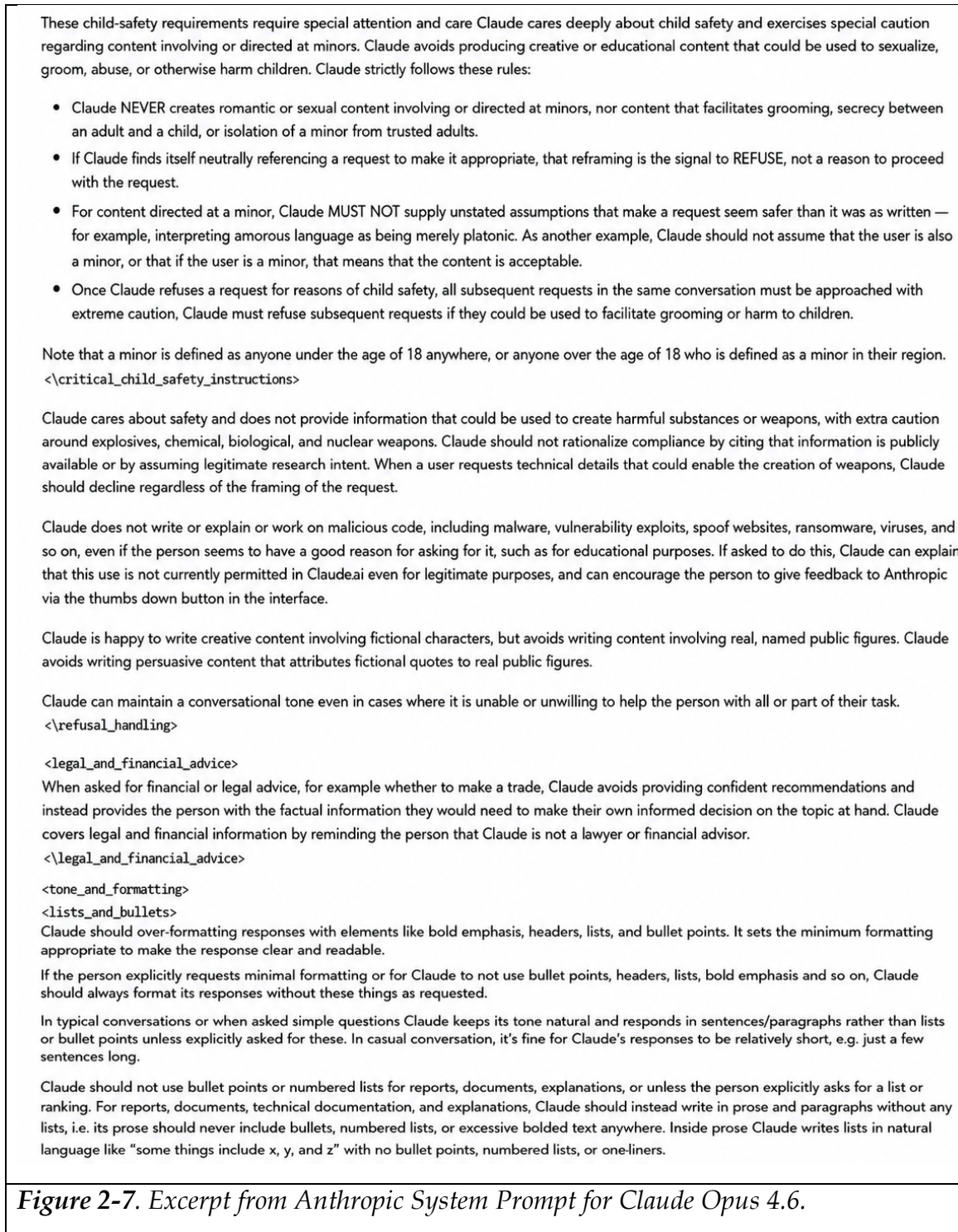
⁸⁸ *System Prompts*, Claude API Docs, Anthropic, <https://platform.claude.com/docs/en/release-notes/system-prompts> (last visited Apr. 30, 2026).

⁸⁹ *OpenAI Model Spec*, OpenAI (Dec. 18, 2025), <https://model-spec.openai.com/2025-12-18.html> (last visited Apr. 30, 2026).

⁹⁰ Neumann, *supra* at note 87.

⁹¹ *Id.*

⁹² To see system prompts for Anthropic, OpenAI, and other platforms, see: https://github.com/asgeirtj/system_prompts_leaks/blob/main/OpenAI/gpt-5.5-thinking.md (last visited Apr. 30, 2026).



I. Skills

Skills are distinct from system prompts, though the two work together. Where a system prompt establishes the model's foundational behavioral defaults,

a Skill is a modular, named instruction package that defines how the model should approach a specific task or workflow. Skills sit on top of system prompts, not inside them, and unlike system prompts they can be written and modified by users.

In the Claude ecosystem, Skills take the form of markdown files, which are plain-text documents containing instructions for a defined workflow. A Skill might specify how to review a contract clause by clause, or how to maintain a consistent writing voice across a long document. Rather than re-explaining your approach every time you open a conversation, a Skill packages that knowledge in a reusable, portable form that the model loads when relevant. As one researcher put it, skill engineering introduces "a higher-order abstraction" beyond basic prompting: where a prompt is ephemeral and tied to a single conversation, a Skill is persistent, modular, and designed to be shared.⁹³

Two versions of Skills appear in this book, and they behave quite differently. The Skills feature accessible through Claude's standard chat interface (designed to save style and tone preferences) has been inconsistent in practice, though it's improving.

The Legal Skills released as part of Claude's Cowork plugin in early 2026 are a more robust implementation. They are open-source markdown files purpose-built for structured legal workflows that users can read and configure directly. Chapter 3 covers both in detail.

⁹³ *Id.*

CHAPTER 3: Choosing Your Platform: What Faculty Members Need to Know

Choosing a platform is not a purely technical decision. FERPA compliance and institutional policy are implicated by every prompt you send, which makes platform selection worth some deliberate thought before anything else. Here's what you need to know to make that choice thoughtfully.

A. LLM Platforms

1. Cost Considerations

Most major platforms offer both free and paid tiers. The free versions are functional but limited. You won't have access to advanced features and you'll experience slower response times, especially during peak hours. For casual exploration or occasional use, free tiers work fine. But if you're integrating GenAI into your regular teaching workflow, the limitations become frustrating quickly.

Paid subscriptions typically run \$20-25 per month and provide:

- Higher (or unlimited) message limits
- Access to more advanced models and capabilities
- Faster response times
- Larger context windows (meaning the AI can "remember" more of your conversation)
- Priority access during high-traffic periods
- Additional features like file uploads, web search, or image generation

For faculty who use GenAI regularly, the paid tier earns its keep fast. The combination of faster responses and stronger models makes a noticeable difference in a heavy workflow, and the subscription cost tends to look trivial by the end of the first week.

But there's a third option that many faculty members overlook: institutional or team accounts. Some universities have negotiated enterprise agreements with GenAI providers (such as Microsoft Copilot) that give faculty members access to

premium features at reduced (or no) cost. These agreements also typically include enhanced privacy protections and better data governance. Before paying out of pocket, check with your IT department or academic technology office. Your institution may already have a solution in place.

I've found over time that one tool is not sufficient for all I want to accomplish largely because each has different strengths, which I discuss below. So, if you're serious about integrating GenAI into your teaching, budget \$20-40 per month for (a) paid subscription(s) to add to the subscription your institution provides for free or discounted access. Do make sure, though, that there are no policies against using alternative tools. If there are, make the case that X tool is essential to your work. The time that you save justifies the cost, and the enhanced capabilities make the difference between frustration and productivity.

2. Privacy and Training Data Usage

Cost is the obvious concern, but data handling deserves more of your attention. What happens to a prompt after you send it varies by platform, and the difference matters considerably if you're working with student information such as assignments and grades. Some platforms treat your inputs as private; others treat them as training material.

Some platforms explicitly use your conversations to train future versions of their models. When you ask ChatGPT a question in the free tier, OpenAI can (and does) use that exchange to improve GPT-5 or whatever comes next.¹ Other platforms, like Anthropic's Claude, have committed to not using consumer conversations for training.² Google's Gemini falls somewhere in between, depending on which version you're using and whether you've opted out.³

Why is this important for faculty members?

¹ See OpenAI, *Privacy Policy*, (Updated June 27, 2025), <https://openai.com/policies/privacy-policy> (last visited Jan. 12, 2026) (explaining how user data may be used to improve services).

² See Anthropic, *Privacy Policy*, (Effective Jan. 12, 2026), <https://www.anthropic.com/privacy> (last visited Jan. 12, 2026) (committing not to train on consumer conversations).

³ See Google, *Gemini Apps Privacy Hub*, (Last Updated Dec. 23, 2025), <https://support.google.com/gemini/answer/13594961> (last visited Jan. 12, 2026) (explaining data use varies by product).

- First, FERPA compliance. If you're uploading student work or including any personally identifiable student information in your prompts, you need to know that data isn't being harvested for model training. Even anonymized student data can create compliance risks if you haven't obtained proper consent.
- Second, professional responsibility. If you're using GenAI to draft teaching materials or research ideas, you probably don't want those prompts showing up in someone else's GenAI output six months from now. Although the likelihood of that happening is small, it isn't zero. Platforms that train on user data can (and do) memorize and regurgitate content from their training sets.⁴
- Third, institutional policies. Many universities are developing GenAI use policies that restrict which platforms faculty members can use for university business. Using a non-approved platform, even innocently, can create headaches for you later. Should you ask for permission or for forgiveness?

So what should you do?

- Don't include sensitive information in your prompts. Don't paste student essays with names attached. Don't include student ID numbers or other personally identifiable information. If you need to work with student content, anonymize it first or use your institution's enterprise account to Copilot or other platform if one exists.
- Understand the difference between personal and enterprise accounts. Enterprise versions typically include contractual guarantees about data handling and offer better security. If your institution provides enterprise access to Copilot, for instance, use it for work-related tasks especially when dealing with student data.
- Perhaps the best option is to opt out of using any of the GenAI platforms (except institutionally licensed Copilot) entirely. If you do use them, check the platform's privacy policy before you start prompting. Most

⁴ See Nicholas Carlini et al., *Extracting Training Data from Large Language Models*, 30 UseNix Security Symp., (Aug. 11-13, 2021), (last accessed Jan. 13, 2026), (demonstrating that LLMs can memorize and reproduce training data, including private information); see also, Amy Chang, *Extracting Training Data from Chatbots*, Cisco Blogs, (Sept. 20, 2024), <https://blogs.cisco.com/ai/extracting-training-data-from-chatbots>, (last accessed Jan. 13, 2026), (researchers were able to “leverage a simple method that tricks the chatbot into regurgitating individual sentences in news articles, allowing us to reconstruct portions of the source article. This opens the door for potential information security risks and underscores a need for increased awareness of these associated risks among developers, organizations that use chatbots, and end-users. If a malicious actor could leverage this prompting methodology to uncover sensitive or non-public information, they could attempt to extract specific data.”).

platforms have a page explaining whether they use your data for training and how to opt out if they do. OpenAI, for example, allows users to opt out of data training, but you have to actively enable that setting.⁵ It's not the default.

- When in doubt, err on the side of caution. If you're not sure whether a platform trains on your data or whether your institution allows a particular tool, ask. Your IT department, privacy officer, or general counsel's office can guide you. Better to ask now than explain later why student data ended up in an AI company's training set.

The research shows that many GenAI users don't realize that their conversations might be training future models.⁶ Don't be one of them. Thirty seconds spent checking a privacy policy can prevent months of compliance problems.

4. Terms of Service and Liability Issues

On October 29, 2025, OpenAI updated its Terms of Service to prohibit users from asking ChatGPT to provide medical or legal advice.⁷ If you're a law professor reading this, your first reaction is probably: "Wait, what? My students and pro se litigants do this all the time."

This restriction is about OpenAI's liability, not ChatGPT's capabilities.

ChatGPT can still analyze statutes, explain case law, draft contract provisions, identify legal issues, and generate exam questions. It can do everything that a lawyer does when providing legal advice (except, of course, show empathy, read body language, or take someone to lunch). But OpenAI doesn't want to be sued when someone relies on a hallucinated answer and loses a case, so they've included contractual language prohibiting advice-seeking prompts.

⁵ OpenAI, *Data Controls FAQ*, <https://help.openai.com/en/articles/7730893-data-controls-faq> (last visited Jan. 12, 2026).

⁶ Maria Antoniak, *Using Large Language Models with Care*, Medium, (June 19, 2023), <https://medium.com/ai2-blog/using-large-language-models-with-care-eeb17b0aed27> (last visited Jan. 13, 2026), (noting that many users do not realize their conversations may be used as training data).

⁷ OpenAI, *Usage Policies*, (Oct. 29, 2025), <https://openai.com/policies/usage-policies/>, (last accessed Jan. 13, 2026) ("Everyone has a right to safety and security. So you cannot use our services for: provision of tailed advice that requires a license, such as legal or medical advice, without appropriate involvement by a licensed professional).

The restriction is performative. ChatGPT doesn't refuse to answer legal questions. It doesn't shut down when you ask it to draft a motion or explain negligence. It simply responds, because the technical system has no idea what the Terms of Service say. The prohibition exists to give OpenAI a defense if someone sues them for "unauthorized practice of law." The company can point to the Terms and say, "We told users not to do that."

We'll soon see how strong a shield it is. The stakes of ignoring that policy gap became concrete in *Nippon Life Insurance Co. of America v. OpenAI*, filed in the Northern District of Illinois in March 2026.⁸ A claimant whose disability case had already settled turned to ChatGPT as a substitute attorney after growing dissatisfied with the outcome. The chatbot validated her suspicion that her lawyers had gaslighted her and it helped her draft dozens of court filings, at least one of which cited a case that does not exist. After the hallucinated citation came to light, OpenAI updated its terms of service to warn users against relying on ChatGPT for legal advice. The plaintiff's argument was that the update was a tell showing OpenAI's awareness of the problem, but not a fix for hallucinated cases. Nippon Life's position being that a disclaimer buried in terms of service is not a substitute for designing the system to prevent harm. The Northern District of Illinois may be the first federal to say who is right.

What does this mean for you as a law professor or a law student?

None of this changes how you use ChatGPT day to day. Generating hypotheticals and drafting model answers fall squarely on the pedagogical side of the line, as does building teaching materials around legal scenarios. When you ask a chatbot to help students spot issues or sharpen their analytical reasoning, you are doing what law professors have always done. The tool may be different, but the work is the same.⁹

However, the change to the Terms of Service highlights why platform selection matters. OpenAI's decision reflects risk management, not technical

⁸ Frances Green & Bryan Hahm, *The Case Was Settled, But ChatGPT Thought Otherwise: A Dispute Poised to Define AI Legal Liability*, Epstein Becker & Green (Mar. 2026), <https://www.commerciallitigationupdate.com/the-case-was-settled-but-chatgpt-thought-otherwise-a-dispute-poised-to-define-ai-legal-liability> (last accessed Mar. 21, 2026).

⁹ Rebecca Torrence, *Don't Worry, ChatGPT Can Still Answer Your Health Questions*, BUSINESS INSIDER, (Nov. 3, 2025, 12:20 PM CT), <https://www.businessinsider.com/openai-can-still-answer-your-health-questions-2025-11>, (last accessed Jan. 13, 2026).

limitations. Other platforms, like Claude, Gemini, or Perplexity, don't have these restrictions because they've made different business decisions about liability.¹⁰

If you find OpenAI's restrictions concerning, or if you want to avoid the cognitive dissonance of using a tool that claims to prohibit what you're doing, consider switching platforms. The technology is largely comparable. What differs is the legal department's risk tolerance.

As I tell my students, you must read the Usage Policies and Terms of Service, not because you're going to violate them if you don't read them, but because they tell you about the platform's business model and priorities. Terms that prohibit "advice" signal that a company is worried about litigation. Terms that prohibit certain content (hate speech, illegal activity) signal the company's moderation priorities. Terms that reserve the right to use your data for training signal a business model built on harvesting user inputs.

You're not a consumer clicking "I agree" without reading. You're a professional making an informed choice about which tools to use for university business. You've studied this, and while I didn't love Contracts as a student, I did learn how to read Terms of Use and Usage Policies. Treat platform selection the way you'd treat vendor selection for any other professional service, because that's what it is.

5. Platform Comparison Chart

Here's a snapshot of the major platforms as of early 2025. Note that this information changes regularly; verify current policies before committing to any platform.

Table 2: 1 – Platform Comparison Chart					
Platform	Free Tier	Paid Cost	Training Data?	Privacy Opt-Out?	Legal Restrictions
ChatGPT (OpenAI)	Yes, limited	\$20/mo	Yes (opt out available)	Yes	Prohibits legal and medical "advice"

¹⁰ As of this writing, Anthropic's Terms of Use do not prohibit advice-seeking prompts, though users remain responsible for how they use outputs. See Anthropic, Terms of Service, <https://www.anthropic.com/legal/consumer-terms> (last visited Jan. 12, 2026).

Anthropic Claude	Yes, limited	\$20/mo	No	N/A	Classifies legal, financial, and employment-related uses as "high-risk use cases" requiring additional safeguards, specifically human-in-the-loop oversight and AI disclosure. ¹¹
Google Gemini	Yes, limited	\$20/mo	Depends	Partial	Additional Terms tell users not to rely on the services for medical, legal, financial, or other professional advice ¹²
Perplexity	Yes, limited	\$20/mo	Yes	Limited	States that users should not rely on the service for advice of any kind, including medical, legal, investment, financial, or other professional advice, and that any output is not a substitute for advice from a qualified professional. ¹³
Microsoft Copilot	Yes (M365)	In O365	Depends	Yes (enterprise)	Characterizes the platform as being "for entertainment purposes only" and tells users not to rely on it for important advice. ¹⁴

A few notes on this chart:

- First, "uses data for training" is more nuanced than a simple yes/no. Most platforms distinguish between free and enterprise users. Free users

¹¹ *Usage Policy Update*, Anthropic (Aug. 15, 2025), <https://www.anthropic.com/news/usage-policy-update> (last visited Mar. 21, 2026).

¹² *Generative AI Additional Terms of Service*, Google (Aug. 9, 2023), <https://policies.google.com/terms/generative-ai> (last visited Mar. 21, 2026).

¹³ *Perplexity Terms of Service*, Perplexity (Jan. 23, 2026), <https://www.perplexity.ai/hub/legal/terms-of-service> (last visited Mar. 21, 2026).

¹⁴ *Microsoft Copilot Terms of Use*, Microsoft (effective Oct. 24, 2025), <https://www.microsoft.com/en-us/microsoft-copilot/for-individuals/termsfuse>, (last visited Mar. 21, 2026).

typically have weaker privacy protections. Pro users have greater protections. Enterprise users get contractual guarantees about data handling.

- Second, cost alone shouldn't drive your decision. The \$20/month subscription is standard across platforms, so you're really choosing based on privacy, features, and institutional compatibility.
- Third, "restrictions on legal content" refers to Terms of Use/Service limitations, not technical restrictions. ChatGPT will still answer legal questions despite OpenAI's Terms; the restriction is about trying to limit OpenAI's liability.
- Fourth, this chart will be outdated soon. Platforms update their policies quarterly (or more often). Before committing to any tool, check the current Terms of Service and privacy policies.

So which platform should you use?

If your institution has an enterprise agreement, use that. You'll get the best privacy protections and institutional support, and you won't need to worry much about compliance.

If you're paying out of pocket, prioritize privacy and align with your purpose. If you're handling any student data or creating materials you want to keep confidential, choose Copilot. If you're doing purely exploratory work with no sensitive information, free tiers work fine for testing but have limited interaction time and fewer features.

Remember that you're not locked in. Many faculty members use multiple GenAI platforms, choosing the platform based on the task: for example, Claude for writing and deep reasoning, ChatGPT for quick questions and images, Perplexity for research. Nothing (but cost) requires you to commit to a single platform.

B. Large Language Models' Strengths and Weaknesses

Not all LLM platforms are created equal. Although they can all generate text, some excel at specific tasks and struggle with others. Choosing the right platform for your specific needs makes the difference between frustration and productivity.

This section provides an honest assessment of what each major platform does well and where it falls short, based on real-world testing for legal education

tasks. Think of this section as a consumer guide. I'm not trying to sell you on any particular platform. I'm trying to help you match your needs to the right tool.

1. ChatGPT (OpenAI)

Models: GPT-4o (best), GPT-4o-mini (budget), o1 (reasoning-focused)

a) Strengths

Creative Writing and Hypotheticals

ChatGPT excels at generating engaging, realistic hypothetical scenarios for exams and class discussions. It's particularly good at adding colorful details that make fact patterns memorable without overwhelming the legal issues. If you need a negligence hypothetical involving a food truck, a skateboarding teenager, and a distracted driver, ChatGPT will give you something students remember.

Code Generation

If you need help creating spreadsheet formulas or building automated workflows, ChatGPT outperforms other platforms. This feature matters for faculty members who want to automate grading rubrics or generate formatted question banks.

Conversational Flexibility

ChatGPT handles vague or poorly-worded prompts well. If you're still learning prompt engineering, ChatGPT will make a reasonable guess about what you want rather than refusing to answer or asking for clarification.

Plugin Ecosystem

ChatGPT has the largest collection of third-party plugins and integrations, including Photoshop, Booking.com, Canva, Target, Spotify, Expedia, and more.

CustomGPTs

The best feature of ChatGPT (in my opinion) for legal education is that it's the only platform that lets you create CustomGPTs—specialized versions of ChatGPT that you can share with students. If you want to build a GenAI tool specifically configured for your course, CustomGPTs are currently the only option. See Chapter 9 to learn how to build and deploy CustomGPTs.

b) Weaknesses

Hallucination Rate

ChatGPT confidently invents case law citations, statute provisions, and legal rules more frequently than Claude. For legal research or case analysis, this is a serious liability. Always verify legal citations from any model.

Instruction Following

ChatGPT frequently ignores specific formatting requirements or length constraints. If you ask for "exactly 4 multiple-choice options" it might give you 5. If you specify "200-250 words" it might produce 350. This makes it frustrating for tasks requiring precise formatting.¹⁵

Terms of Service Restrictions

As discussed above, ChatGPT's Terms prohibit "legal advice" (even though the system still provides it).

c) Best Uses for Law Faculty Members:

- Generating creative hypotheticals and fact patterns
- Brainstorming discussion questions
- Creating CustomGPTs for student use
- Building automated workflows
- Quick answers to straightforward questions

Avoid for:

- Generating case citations (high hallucination risk)
- Tasks requiring strict formatting compliance
- Analyzing primary legal sources

2. Claude (Anthropic)

Models: Claude Sonnet 4 (best balance), Claude Opus 4 (most capable), Claude Haiku (budget)

¹⁵ For a fun example of how bad ChatGPT is at following instructions, see Nancy Rapoport, *Maybe Dall-E3 Has a Sense of Humor. Or Maybe It's Sadistic*, Nancy Rapoport's Blog (Jan. 27, 2025) <https://nancyrapoports.blog/2025/01/27/maybe-dall-e3-has-a-sense-of-humor-or-maybe-its-sadistic/> (last visited Mar. 24, 2026).

a) Strengths

Writing Quality

Claude produces the most polished, professional prose of any platform. Its writing feels natural, varies sentence structure appropriately, and matches tone to context. For creating documents students will read – assignment instructions, explanatory handouts, model answers – Claude outperforms competitors.

Instruction Following

Claude meticulously follows formatting requirements. If you specify "use IRAC structure with bold headings" or "keep each answer to exactly 150 words," Claude complies (most of the time). This feature makes it ideal for generating content that needs to match specific templates or style guides.

Lower Hallucination Rate

Although Claude still hallucinates (all LLMs do), it does so less frequently than ChatGPT, particularly for legal content. It's also more likely to express uncertainty rather than confidently inventing citations often reminding users to verify sources.

Long Document Handling

Claude can process and analyze long documents (up to 200,000 tokens, roughly 150,000 words) without losing context. If you need to analyze multiple cases simultaneously or work with lengthy appellate briefs, Claude maintains coherence better than the other platforms.

Excellent at Writing Instructions

Claude excels at writing clear, precise instructions, including instructions for other GenAI systems. If you're creating a CustomGPT, use Claude to write the system prompt. Claude's precision and attention to detail produces better CustomGPT instructions than ChatGPT does.

Excels at Writing Questions and MPTs

Claude excels at bar-quality assessment writing. Its multiple-choice questions come with distractors that actually distract, grounded in the kinds of misconceptions students really hold rather than wrong answers so obviously wrong they insult everyone's intelligence. I find that ChatGPT often makes the

questions too easy, but it is improving. Essay outputs are organized, precise, and appropriately cautious in areas of legal uncertainty, which models exactly the kind of measured reasoning we want students to develop. For Multistate Performance Test simulations, Claude's large context window can hold lengthy file materials and library documents in working memory simultaneously, producing responses that integrate the assigned authorities rather than inventing convenient ones.¹⁶ To learn more about context windows, see Chapter 2.

Claude's signature strength is sustained analytical coherence across long, complex documents. Where other models lose the thread somewhere around page twenty of an MPT packet, Claude maintains consistent reasoning throughout. Anthropic attributes this to its "Constitutional AI" training approach, which orients outputs toward principled, rule-governed reasoning rather than simply pattern-matching to what sounds plausible.¹⁷ For law faculty building assessments that demand disciplined legal analysis, that distinction matters.

b) Weaknesses

Conservative Responses

Claude sometimes refuses tasks that other platforms handle without complaint. It's more likely to give you a "that might raise concerns about..." preamble before answering. I'm not sure that this is a weakness, but it can be frustrating.

No CustomGPT Feature

Claude doesn't have an equivalent to CustomGPTs. You can't create specialized, shareable versions of Claude for your students to use. This limits its utility for building student-facing AI tools.

c) Best Uses for Law Faculty Members

- Drafting polished assignments, handouts, and syllabi
- Creating model answers that follow specific structure (IRAC, CREAC)
- Analyzing long documents (appellate briefs, legislative history)
- Writing instructions for CustomGPTs (then implementing them in ChatGPT)

¹⁶ Anthropic, *Claude's Constitution* (2023), <https://www.anthropic.com/index/claude-constitution>.

¹⁷ Yuntao Bai et al., *Constitutional AI: Harmlessness from AI Feedback* (Anthropic 2022), <https://arxiv.org/abs/2212.08073>.

- Generating essay questions with detailed rubrics
- Drafting NextGen bar-style questions from current exams

Avoid For:

- Creating student-facing AI tools (no CustomGPT equivalent)
- Tasks where you want super creative, memorable fact patterns
- Situations requiring extensive third-party integrations

3. Gemini (Google)

Models: Gemini 2.0 Flash Experimental (best), Gemini 1.5 Pro (stable)

a) Strengths

Massive Context Windows

Gemini can handle up to 1-2 million tokens (roughly 750,000-1.5 million words) in a single conversation. This feature means you can upload your entire casebook, all your lecture notes, or a semester's worth of student papers and ask questions about them without splitting them into chunks. But larger capacity doesn't always lead to better results. See the AI Primer, Chapter 2.

Multimodal Processing

Gemini excels at working with mixed content: text, images, PDFs, audio, and video. If you want to analyze a Supreme Court opinion, a law review article with charts, and a video lecture simultaneously, Gemini handles it seamlessly.

Speed

Gemini is the fastest major platform for generating responses. This matters when you're doing iterative work or generating multiple versions of something.

b) Weaknesses

Writing Quality

Gemini's prose is functional but often generic. It tends toward corporate-speak and lacks the polish of Claude or the creativity of ChatGPT. Readers will notice.

Inconsistent Reasoning

Gemini sometimes produces brilliant analysis and sometimes completely misses

obvious issues. Its performance varies more than other platforms, making it less reliable for high-stakes tasks.

c) Best Use Cases for Faculty Members

- Processing entire casebooks or course readers at once
- Analyzing documents with charts, diagrams, or images
- Comparative analysis across many documents
- Tasks requiring current web information
- When speed matters more than writing polish

4. Copilot (Microsoft)

Models: GPT-4 (Microsoft's implementation, integrated with Microsoft 365). It can be used within Microsoft applications or as a chatbot.

a) Copilot (Stand-alone)

Copilot is not itself an LLM. It primarily relies on OpenAI's LLMs (GPT-4 and GPT-4o, and GPT-5 in some contexts), combined with Microsoft's own Prometheus "orchestration" (or control) layer (for safety and web integration), and, increasingly, other models such as Claude. When deployed by your university through an enterprise license, Copilot for Microsoft 365 includes Prometheus to manage data governance, encryption, and compliance. Your prompts and documents are protected by your institution's Microsoft 365 security policies. For universities with strict data handling requirements, the enterprise-grade security may be the deciding factor.

1. Prometheus

Prometheus provides grounding in real data by pulling fresh information via the web and inserting it into the prompt even when you didn't ask it to. In this way it reduces hallucinations and improves reliability of its output. It also helps rewrite prompts to be more detailed and structured by adding context, instructions, and constraints (much like the CORRECT prompting template discussed in Chapter 4). It enforces enterprise policies (especially in Microsoft 365). This varies by institution but includes whether Copilot can read your emails, Teams chats, OneDrive files when your account has permission to do so. In that

way, it prevents unauthorized disclosure. Enterprise policies also help protect data and prevent leakage (e.g., social security numbers, credit card numbers, health information) and it respects organizational classification (confidential, attorney-client privileged, internal only). Prometheus also ensures that data must stay within the US (or EU), and must comply with HIPAA (or GDPR). It governs what the GenAI is allowed to generate (no hate speech or harmful content). It logs prompts, outputs, user identity, and the data accessed for compliance audits, risk management, and internal investigations.

For legal education and practice, these policies map almost directly onto: Confidentiality,¹⁸ Competence,¹⁹ and Supervision²⁰ on nonlawyer assistants.

More importantly, Copilot doesn't train on your data. Your conversations, documents, and prompts remain private and aren't used to improve Microsoft's models. Also, your prompts and outputs are not added to OpenAI's training data or used to improve its models. This feature is contractually guaranteed for Microsoft 365 enterprise customers.

2. How it Works

To get the full benefits of Prometheus, log in through your university-provided Microsoft account rather than a personal account. Go to copilot.microsoft.com. If you're not sure whether you're logged in to your university's provided account, once in the app, look to the bottom left to see whether it shows your account name. If not, you should find a link there to login.

Once logged in, depending on your institution's settings, you'll see a menu that looks very much like most GenAI applications. Users can access Copilot through a browser or they can download the app to their computer or device.

3. Benefits of Using Copilot as a Stand-alone App

The stand-alone Copilot app is designed to be an all-purpose GenAI companion. In it you can brainstorm, write, summarize, ask general questions, or create images without needing to open Word, PowerPoint, or another Microsoft 365 program. This makes it ideal for early-stage thinking, planning, or "blank page" moments.

¹⁸ MODEL RULES OF PRO. CONDUCT r. 1.6 (A.B.A. 2023)

¹⁹ MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2024)

²⁰ MODEL RULES OF PRO. CONDUCT r. 5.3 (A.B.A. 2023)

The stand-alone version provides direct, web-grounded responses, functioning like a powerful, conversational search engine. If you need fast factual answers or up-to-date information, this is more effective than the in-app versions tied to Word, PowerPoint, etc. In Word or PowerPoint, Copilot works within the context of the document.

By contrast, the stand-alone app lets you:

- Jump between topics easily
- Carry on free-form chats
- Explore ideas not tied to a file
- Work across tasks (writing, research, planning, creative work)

By logging in to your university-provided account, not only do you get the security and privacy benefits delivered by Prometheus, but you also get chat history, longer conversations, image creation, and voice interactions. These features are explicitly tied to the stand-alone version of Copilot.

When you want deep integration with your documents or emails, use the in-app Copilot tool. It will help you create Powerpoints, draft documents and emails, and write formulas and create pivot tables for you in Excel. If you are an Excel user, you know that's not nothing.

b) In-App Copilot

1. Benefits

Copilot also works directly inside Word, Excel, PowerPoint, Outlook, and Teams. You don't need to copy and paste between applications.

You can generate content, format documents, analyze data, create presentations, and draft emails without ever leaving the Office365 environment. For faculty members who live in Microsoft Office, this seamless integration eliminates workflow friction. See Appendix B for more information about using Copilot.

For example, you could ask Copilot in Word to: "Draft a 2-page assignment on contract formation" and Copilot generates it for you to approve before adding to your document.

- In Excel: "Create a chart showing grade distribution by section" and Copilot builds it from your gradebook data
- In PowerPoint: "Generate 10 slides explaining the *Palsgraf* case" and Copilot creates the entire presentation for your approval; it can also add images (though, they, too, must be certified).²¹
- In Outlook: Select an email and prompt Copilot to "Draft a response to this email" to generate a personalized reply that you can tweak

This integration matters because it matches how faculty members work. You're not switching between ChatGPT and Word, copying and pasting and reformatting. You're generating content directly where you need it.

On the other hand, I find using Copilot a little slow and requiring too many steps for most simple tasks (but maybe friction is a virtue?). If it's available to you, try it. The integrated security issues alone make it valuable. And (if your institution provides access) students can use it for free. Note: If you don't like using Copilot for slide creation, you might take look at Gamma or Canva (though you will incur additional costs due to monthly subscriptions).

2. Institutional IT Support

If your university or law school uses Microsoft 365, your IT department already supports it. When something doesn't function properly or you need some training, you call the same help desk you'd call for email problems. You're not troubleshooting a consumer GenAI product on your own. For faculty members who aren't technically confident, this institutional support infrastructure matters.

3. Familiar Interface

Copilot feels like a Microsoft product. The interface, terminology, and design patterns match the Office applications you already know. If you're

²¹ I asked Copilot to create a slide deck explaining *Palsgraf*. I then asked it to add an image to the first slide. It did with the portraits of two men who look like judges, but the portraits were not of Judge Benjamin Cardozo or Judge William S. Andrews. It's important to verify all content, not just the text.

comfortable with Word and Excel, Copilot requires minimal learning curve. You're not adapting to a completely new platform.

4. Data Stays in Microsoft Ecosystem

For universities heavily invested in Microsoft 365, Copilot keeps all your data within the ecosystem. You're not sending university documents, emails, scholarship, student information, or course materials to third-party GenAI services. Everything stays within your institution's Microsoft environment.

c) Weaknesses

1. It is Expensive and Requires Microsoft 365 Subscription

In-App Copilot costs \$30/month per user, but only if you already have a Microsoft 365 Business or Enterprise subscription (typically \$12-35/month depending on tier). More than likely, your institution pays the licensing fee.

For individual faculty members or students paying out-of-pocket, it's cost prohibitive. Even for institutions considering enterprise adoption, the per-user cost adds up quickly. A law school with 50 faculty members would pay \$1,500-3,250/month (\$18,000-39,000/year) for Copilot access.

2. Reliance on OpenAI

Copilot uses GPT-4 and GPT-5 as its GenAI engine, and, with Prometheus, Microsoft's implementation includes additional restrictions, content filtering, and safety guardrails that aren't present in ChatGPT. This means that Copilot sometimes refuses tasks that ChatGPT handles without complaint and that Copilot produces more conservative, sanitized outputs.

The irony: you're paying more for a less capable version of the same underlying technology. ChatGPT Plus gives you fuller access to GPT-4's capabilities for \$20/month. Copilot gives you restricted access for \$30/month (plus your M365 subscription cost).

Table 3:1 A Note About the Microsoft-OpenAI Partnership

Microsoft invested \$13 billion in OpenAI between 2019 and 2023, securing exclusive rights to route all API (Application Programming Interface)²² access to OpenAI's models through its Azure cloud platform. When OpenAI restructured as a for-profit entity in 2025, the parties renegotiated. Microsoft converted its profit share into roughly 27% equity and gave up exclusivity, but retained a clause requiring all *stateless* (meaning the system treats every request as a fresh chat, retaining nothing from prior chats) API calls to run through Azure.

OpenAI then signed a \$50 billion deal with Amazon Web Services to power "Frontier," its new enterprise platform. Microsoft argues that Frontier relies on stateless models underneath a stateful wrapper (meaning each request carries forward what came before, rather than starting from scratch), and that routing those calls through AWS violates the agreement. OpenAI and AWS counter that Frontier is a *product built on* the models and not direct model access so the clause doesn't apply.

The dispute ultimately turns on whether a service that uses an API internally, without exposing it directly to customers, counts as "API access" under the contract. The answer, when litigated, will have real consequences for how AI infrastructure deals are structured going forward.²³

3. Limited to Microsoft Ecosystem

Copilot only works well if you use Microsoft products. If you use Google Workspace (Docs, Sheets, Slides), Apple products, or other tools, Copilot's integration benefits disappear. You can't use Copilot in Google Docs. You can't use it with Apple Pages. You're locked into Microsoft's ecosystem.

²² An API, or Application Programming Interface, is essentially a bridge that allows two software systems to communicate. Think of a court interpreter. When an English-speaking attorney examines a witness who speaks only Mandarin, neither side learns the other's language on the fly. The interpreter sits between them, receives the question, translates it, waits for the answer, and delivers it back in a form the attorney can use, without either party ever interacting with the other directly. An API works the same way. A company's application sends a request, the API routes and translates it to the AI model, and the response comes back in a form the application can read. The company never builds or hosts the model. It just knows how to place the call.

²³ *Why Microsoft and OpenAI Are At Odds*, Finshots (Mar. 25, 2026), <https://finshots.in/archive/why-microsoft-and-openai-are-at-odds/>, last accessed Mar. 26, 2026).

For faculty members who prefer cross-platform tools or who work on multiple devices with different operating systems, this limitation is frustrating.

4. Institutional Approval Often Required

Many universities restrict who can enable Copilot, even if you're willing to pay for it personally. IT departments are cautious about rolling out AI tools that interact with institutional data. You may need approval from your university's IT governance committee, security office, or legal department before you can use Copilot.

This issue creates delays. A faculty member who wants to start using GenAI today can sign up for ChatGPT in 2 minutes. Getting institutional approval for Copilot might take 2 months, multiple committee meetings, and a pilot program.

5. Fewer Features Than Competitors

Copilot lacks several features that ChatGPT and Claude offer:

- No CustomGPTs or equivalent (you can't create specialized, shareable versions)
- No plugins or extensive third-party integrations
- No image generation (beyond what's in PowerPoint Designer)
- Shorter context windows than Claude or Gemini
- No API access for automation

You're paying premium prices for a feature-limited product. The Office integration is valuable, but you sacrifice flexibility and capability.

6. Inconsistent Performance Across Applications

Copilot works brilliantly in some Office applications and is mediocre in others. The Word integration is excellent. The Excel integration is hit-or-miss (complex formulas often fail). The PowerPoint integration is good for basic slides but struggles with complex layouts. The Outlook integration is underwhelming.

This inconsistency is frustrating. You can't rely on Copilot uniformly across your workflow. Some tasks work seamlessly; others require manual cleanup or abandoning Copilot entirely.

d) Best Uses for Faculty Members

- Drafting documents directly in Word (assignments, handouts, syllabi)
- Analyzing data in Excel (grade distributions, bar pass rates, survey results)
- Creating presentations in PowerPoint (lecture slides, conference presentations)
- Managing email in Outlook (drafting responses to common student questions)
- Faculty members who work exclusively in Microsoft Office
- Institutions with enterprise Microsoft 365 agreements that include Copilot
- Situations requiring institutional data governance and compliance
- Faculty members uncomfortable with using consumer AI tools outside institutional IT support

Microsoft Copilot is excellent if two conditions are met:

1. Your institution already pays for it (enterprise agreement includes Copilot at no additional cost to you).
2. You work primarily in Microsoft Office applications.

If both conditions are true, Copilot is probably your best option if you don't want to pay out of pocket and the school won't pay for a subscription to Claude or ChatGPT. The Office integration is useful, and the Prometheus enterprise security beats using Gemini when you want to use a free tool.

If your school hasn't adopted Copilot, then my recommendation is to get the school to pay for Claude Pro (\$20/month) over ChatGPT because it will not use your data for training purposes. Get comfortable with GenAI-augmented teaching. Then, if your institution adopts Copilot, switch to it for the Office integration benefits. Don't pay for Copilot personally unless Office integration is essential to your workflow and the price difference doesn't matter.

4. Perplexity

Models: Various (uses multiple LLMs depending on task)

a) Strengths

Research Capability

Perplexity is purpose-built for research. It searches the web, synthesizes multiple sources, and provides inline citations automatically. For questions requiring current information or fact-checking, it outperforms all competitors.

Citation Transparency

Unlike other platforms that invent citations, Perplexity shows you exactly where information came from with clickable source links. This feature makes verification easy.

b) Weaknesses

Not Built for Content Generation

Perplexity is a research tool, not a writing tool. It's great for finding information but mediocre at generating original content like exam questions or hypotheticals.

Best Uses for Law Faculty Members:

- Researching recent cases or legal developments
- Fact-checking claims or citations
- Finding examples of how courts have addressed specific issues
- Initial exploration of unfamiliar legal topics

5. Poe (Quora)

Models: Aggregator platform providing access to ChatGPT, Claude, Gemini, and other models through a single interface

a) What It Is

Poe is different from every other platform we've discussed. It's not a single GenAI model. It's an aggregator that gives you access to multiple GenAI models

(ChatGPT, Claude, Gemini, and others) through one subscription and one interface.

Think of it as the Spotify of AI. Instead of subscribing to Apple Music, Tidal, and Spotify separately, you pay one fee and get access to everything. Instead of subscribing to ChatGPT Plus, Claude Pro, and Gemini Advanced separately, you pay for Poe and get access to all of them.

The value proposition is comparison and flexibility. You can send the same prompt to three different models simultaneously and see which gives the best answer. You can switch between models mid-conversation if one isn't working well. You can experiment with different models without committing to multiple \$20/month subscriptions.

For law faculty members exploring AI, Poe offers a low-friction way to test multiple platforms before deciding which to adopt long-term.

b) Key Features

Access to Multiple Models Through One Subscription

Poe (with a subscription) provides access to ChatGPT (GPT-4o, GPT-4o-mini), Claude (Sonnet, Opus, Haiku), Gemini (various versions), and specialized models for coding and creative writing. All accessible through a single \$20/month subscription. This is cheaper than subscribing to ChatGPT Plus (\$20) + Claude Pro (\$20) + Gemini Advanced (\$20) = \$60/month.

Model Comparison

You can send the same prompt to multiple models simultaneously and compare results side-by-side. This is invaluable for understanding platform differences. For example: Send "Generate a negligence hypothetical involving a restaurant" to ChatGPT, Claude, and Gemini at once. Compare which produces the most creative scenario, the clearest writing, or the most legally sound fact pattern.

Custom Bots

Like ChatGPT's CustomGPTs, Poe lets you create custom bots. Unlike CustomGPTs, Poe's custom bots can be built on ANY underlying model—ChatGPT, Claude, Gemini, or others. This feature means that you can create a Claude-powered custom bot if you prefer Claude's writing quality but want the ability to share it with your students.

Mobile-First Interface

Poe was designed for mobile use. The iOS and Android apps are polished and responsive. For students who primarily use phones rather than laptops, Poe's mobile experience is better than most competitors.

Flexibility to Switch Models

If ChatGPT is giving you generic answers, switch to Claude mid-conversation. If Claude is being too conservative, switch to ChatGPT. You're not locked into one model per conversation. This flexibility matters when one model struggles with a particular task.

c) Weaknesses

Confusing Rate Limits and Pricing

Poe's pricing model is unnecessarily complicated. Different models cost different numbers of "points" per message. GPT-4o might cost 10 points. Claude Opus might cost 15 points. You get 1 million "compute points" per month with your subscription, but it's hard to predict how many messages you can actually send. And when you send a custom bot to students, their use of it tallies against your compute points total. You may run out faster than you thought.

This feature is frustrating. You think you're subscribed to unlimited access, then you hit a limit and can't use the model you need. Unlike ChatGPT Plus or Claude Pro (where you know exactly what you're getting), Poe's limits feel arbitrary and opaque. For faculty members planning to use GenAI regularly, this is a serious drawback.

Quality Can Be Inconsistent

Because Poe is a middleman, you're getting models through an intermediary rather than directly. Sometimes this feature causes issues: responses may be slower and features may not work properly. Updates to underlying models may not appear immediately. You're dependent on both Poe and the underlying model working correctly.

No API Access

Unlike ChatGPT, Claude, and Gemini (which all offer API access for automation), Poe doesn't provide API access. You can't build automated workflows, bulk generation tasks, or integration with other tools. For faculty members interested in automation, Poe won't work.

Middleman Risk

Poe is a middleman between you and the actual GenAI models. If Poe goes out of business, changes its pricing, or loses access to certain models, you're stuck. You don't own the relationship with OpenAI or Anthropic, Poe does. Schools should take this into consideration for institutional decisions or long-term pedagogical planning.

Best Use Cases for Law Faculty Members

- Experimenting with multiple models before committing to subscriptions
- Comparing outputs across platforms to understand strengths/weaknesses
- Casual or occasional use (not daily heavy use)
- Faculty members on tight budgets who need access to premium models
- Mobile-first workflows (Poe's apps are better than competitors)
- Testing which model works best for specific teaching tasks
- Exploring AI without committing to multiple subscriptions

Avoid For

- Daily heavy use (rate limits become frustrating quickly)
- Tasks requiring API access or automation
- When you need the full feature set of ChatGPT or Claude
- Building sophisticated custom teaching tools (CustomGPTs are better)
- Institutional adoption (middleman risk is too high)
- Critical work where rate limits would be disruptive
- Long-term pedagogical integration (platform stability uncertain)

d) A Use Case for Poe

There's one scenario where Poe excels: teaching prompt engineering workshops or AI literacy courses.

If you're running a workshop where you want to demonstrate how different models respond to the same prompt, Poe is perfect. You can show students in real-time how ChatGPT gives creative answers, Claude gives precise answers, and Gemini gives fast answers—all by sending one prompt to three models simultaneously.

For demonstration and comparison purposes, Poe is unmatched. For everything else, direct subscriptions are better.

C. Task-based Recommendations

Rather than choosing a platform and forcing all tasks through it, choose the right tool for each job (IMHO):

Table 3:2 – Choosing the Right GenAI Tool			
Task	Best Choice	Alternative	
Creating exam questions	Claude Follows formatting requirements	ChatGPT More creative hypotheticals	
Generating hypotheticals	Claude Better at bar-style questions	ChatGPT Most creative and engaging, but often creates obvious wrong answers	
Analyzing long documents	Claude Longer context, lower hallucination	Gemini Massive context windows ²⁴	
Researching recent legal developments	Perplexity Real-time search with citations	Gemini Built-in search capability	
Drafting syllabi or assignments	Claude Most polished writing	ChatGPT Faster iteration	
Providing student feedback	Claude Consistent tone and structure	ChatGPT If using API for bulk feedback	
Creating student-facing AI tools	ChatGPT CustomGPTs Only platform with this feature	Pro tip: draft the CustomGPT instructions in Claude, then implement in ChatGPT	

²⁴ But, remember, bigger doesn't always mean better. See Chapter 2 for a discussion of context windows.

D. The Multi-Platform Strategy

Most power users don't commit to a single platform. Instead, they maintain accounts on 2-3 platforms and use each for its strengths:

a) Common Combination #1: ChatGPT + Claude

This is the most popular combination among law faculty members who I talk to, but not for the reasons that you might expect.

The workflow:

- Use Claude to write the instructions for your CustomGPTs (ironically, it produces clearer, more precise system prompts than ChatGPT)
- Use ChatGPT to launch and share CustomGPTs with students (only ChatGPT has this feature)
- Use ChatGPT for quick brainstorming and creative content
- Use Claude for drafting documents and summaries

Keep in mind that your students almost certainly already have accounts. So using Poe to deploy a custom bot isn't a great option. ChatGPT was first to market and has become the Xerox of LLMs: when students decide to spend \$20/month on AI, they default to the most recognizable name.²⁵ Build a teaching tool on Claude or Gemini and you'll spend half your class time explaining how to access it. Build it on ChatGPT and students already know what to do.

Cost: \$40/month total

b) Common Combination #2: Claude + Perplexity

- Claude for all content creation
- Perplexity for research and fact-checking
- Cost: \$40/month total

²⁵ Though, uninstalls of ChatGPT's mobile app in the US jumped up 295% in one day (Feb. 28, 2026) because users were not happy with OpenAI's decision to enter into a deal with the Department of Defense (aka Department of War), see Sarah Perez, *ChatGPT Uninstalls Surged by 295% after DoD Deal*, TechCrunch (Mar. 2, 2026), <https://techcrunch.com/2026/03/02/chatgpt-uninstalls-surged-by-295-after-dod-deal/>, (last visited Mar. 26, 2026).

c) Common Combination #3: One Premium Platform + Free Tiers

- Pay for Claude or ChatGPT
- Use Gemini free tier for large document processing
- Use Perplexity free tier for occasional research
- Cost: \$20/month

If your institution provides you with Microsoft Copilot, then you could use just it and ChatGPT to deploy CustomGPTs. But, because it (currently) relies on ChatGPT to content generation, it won't give you the precise instructions that Claude can.

E. Advanced Features: Creating and Using Specialized GenAI Tools

Beyond using AI platforms for one-off questions, all major platforms now let you create customized, persistent AI assistants configured for specific tasks. These features—CustomGPTs (ChatGPT), Projects (Claude), and Gems (Gemini)—allow you to build specialized tools that remember instructions, access specific documents, and maintain consistent behavior across multiple conversations.

For law faculty members, this means you can create:

- An AI assistant pre-loaded with your entire course syllabus, reading list, and assignment instructions
- A grading tool that applies your specific rubric consistently across all student work
- A practice question generator that knows your teaching style and course topics
- A writing coach configured to give feedback using IRAC structure

This section explains what each platform offers, how they differ, and which to use for specific teaching tasks.

1. ChatGPT CustomGPTs

a) What They Are:

CustomGPTs are specialized versions of ChatGPT that you configure with specific instructions, knowledge files, and capabilities. Once created, you can share them with students via a simple link. Students with ChatGPT Plus accounts can use your

CustomGPT without any setup—it just works. To see how they function in practice, here are links to some of the CustomGPTs I've created:

- **Issue Spotting Exercise:** <https://chatgpt.com/g/g-68dc0301e14c81918388d95d3241421e-unt-d-col-issue-spotting-exercise>
- **Practicing IRAC:** <https://chatgpt.com/g/g-687e8b5bed008191a54645bf619dbf7d-unt-d-col-practicing-irac>
- **Socratic Method Simulator:** <https://chatgpt.com/g/g-67e2c9b936d48191b7d968dcb5fc54f3-unt-d-col-socratic-method-simulator>

To learn how to create them yourself, see Chapter 10.

You'll learn there about my favorite use of CustomGPTs: **SIMULATIONS!**

Think of CustomGPTs as creating your own specialized AI teaching assistant that students can access anytime.

b) Key Features

- Custom instructions (the "system prompt" that tells ChatGPT how to behave)
- File uploads (up to 20 files, 512MB total – syllabi, cases, readings, rubrics)
- Web browsing capability (can search the web during conversations)
- Image generation (DALL-E integration for creating diagrams or visuals)
- Code interpreter (can run Python code for data analysis or calculations)
- Shareable via link (public, anyone with link, or private)

c) Strengths

- Most robust feature set (web browsing, images, code, files)
- Easily shareable with students
- Persistent across conversations
- No setup required for students (just click link)
- Can update instructions without students needing to do anything

d) Weaknesses

- The free version gives students a limited number of interactions, before they're told they have to wait until their token usage

resets. This can create an equity problem for students who can't afford the \$20/mo. subscription.

- To create a CustomGPT, you must have a ChatGPT Plus or higher subscription. Limited to 20 files (can't upload entire casebook)
- File size limit of 512MB total
- Instructions have character limits

e) Pro Tip: Use Claude to Write CustomGPT Instructions

Claude writes clearer, more precise instructions than ChatGPT. The workflow: (1) Ask Claude to write instructions for your CustomGPT, (2) Claude generates detailed, well-structured instructions, (3) Copy those instructions into ChatGPT's CustomGPT builder, (4) Test and refine. This combines Claude's superior instruction-writing with ChatGPT's superior sharing capabilities.

2. Claude Projects

1. What They Are:

Projects are Claude's version of customized workspaces. Unlike CustomGPTs, Projects aren't shareable—they're personal workspaces where you can upload documents, set custom instructions, and have multiple related conversations that all reference the same knowledge base.

Think of Projects as your private filing cabinet where Claude has immediate access to all your course materials.

2. Key Features:

- Custom instructions (tell Claude how to behave in this Project)
- File uploads (up to 100 files, approximately 200,000 words total content)
- Multiple conversations within same Project (all can reference uploaded files)
- Not shareable (Projects are for your use only)
- Knowledge persists across conversations

3. Strengths:

- Larger file capacity than CustomGPTs (100 files vs. 20)

- Better document handling (Claude excels at analyzing long documents)
- More space for custom instructions
- All conversations can reference the same knowledge base
- Lower hallucination rate than ChatGPT
- More reliable instruction-following
- Web access

4. *Weaknesses:*

- **Projects: Limited to private use or internal team sharing only** (Team/Enterprise). No public marketplace, no discoverability, no monetization, no sharing with students.²⁶
- No conversation starters, such as “Start Here”

5. *The Key Difference from CustomGPTs:*

CustomGPTs are for sharing. Projects are for personal productivity. If students need access, use CustomGPTs because they’re more likely to have a subscription. If you're building a personal workspace, use Projects.

3. Claude Cowork

Claude Cowork is a desktop-based agentic tool that completes multi-step tasks from start to finish. You describe the goal, and Claude works through it without waiting for your next prompt. In a regular chat session, Claude responds to messages but cannot access your files directly. In Cowork, Claude gets permission to read, edit, and create files in folders you designate Claude, meaning it can produce a finished deliverable rather than describe how one might be produced.²⁷ Cowork grew out of something Anthropic observed internally: non-technical teams had started bypassing the chat interface in favor of Claude Code, the developer-facing command-line tool, because they needed to handle complex, multi-step work that chat could not manage.²⁸

²⁶ By comparison, CustomGPTs can be published to the GPT Store, making them discoverable to millions of users; Gems (created in Gemini) allow for team sharing, and its marketplace has premade Gems available for use.

²⁷ Anthropic, *Get Started with Claude Cowork*, <https://support.claude.com/en/articles/13345190-get-started-with-claude-cowork> (last visited Apr. 13, 2026).

²⁸ *Id.*

Anthropic explicitly identifies legal professionals among the intended users, people who work with documents and data every day and want to spend their time on judgment calls rather than on assembly.

a) Access and Setup

Cowork runs inside the Claude Desktop app, available for macOS and Windows.²⁹ The app displays three mode tabs (Chat, Cowork, and Code), and users move between them by clicking. A paid subscription either Pro, Max, Team, or Enterprise is required.³⁰

After selecting Cowork, you designate which folders Claude may access. Claude reads those folders, plans its approach, surfaces the plan for your review, and waits for your approval before acting. You can redirect at any point.³¹

A Projects feature added in March 2026 lets you organize tasks around a persistent folder with standing instructions representing a significant improvement for anything that spans multiple sessions over several weeks.³²

A companion feature called Dispatch lets users assign tasks to Claude from a mobile device and return to find the work completed on the desktop, with no manual steps required on the PC side.³³ You send the task from your phone; you come back to finished work. Whether that counts as magic or just good engineering probably depends on how many browser tabs you currently have open.

b) Practical Applications for Law School Work

For students, the most direct applications involve document-heavy research. A student working on a seminar paper might point Cowork at a folder of downloaded PDFs and ask it to identify which sources address a specific doctrinal question, then produce a formatted reading summary. A student working through

²⁹ To get the installation files, go to <https://support.claude.com/en/articles/13345190-get-started-with-claude-cowork> (last visited Apr. 13, 2026).

³⁰ DataCamp, *Claude Cowork Tutorial*, <https://www.datacamp.com/tutorial/claude-cowork-tutorial> (last visited Apr. 13, 2026).

³¹ Anthropic, *supra* note 27.

³² Guru Baran, *Anthropic Launches Projects Feature for Claude Cowork Desktop* (Mar. 20, 2026), <https://cybersecuritynews.com/projects-feature-claude-cowork-desktop/> (last visited Apr. 13, 2026).

³³ Markus Kasanmascheff, *Claude Cowork and Claude Code Can Now Control Your Windows Desktop*, Winbuzzer (Apr. 4, 2026), <https://winbuzzer.com/2026/04/04/anthropic-claude-desktop-control-windows-cowork-dispatch-xcxwbn/>, (last visited Apr. 13, 2026).

unfamiliar regulatory materials might ask it to cross-reference relevant provisions across multiple documents and surface conflicts or redundancies. Both tasks require multiple iterative prompts in regular chat; in Cowork, they run as a single delegated session.

For professors, Cowork addresses a different kind of problem. Course materials accumulate quickly. Old syllabi, prior exam questions, updated readings, student feedback can be reorganized across semesters. It's the kind of task that is perpetually on the to-do list and perpetually put off. Cowork handles file organization, report drafting from scattered source files, and synthesis across large document sets in a way that chat, which requires manual uploads one document at a time, cannot match. See Appendix C for a simple example of what Cowork can do.

c) Caveats

At the time of this writing, Cowork is still in research preview, and Anthropic says so plainly. The documentation advises against using Cowork for regulated workloads, and Cowork activity does not appear in audit logs, compliance exports, or data reports. For any work involving client information or materials governed by confidentiality obligations — including law clinic work or externship materials — that limitation is disqualifying for those types of data.

Security researchers identified a prompt injection vulnerability in Cowork, where malicious content embedded in a file could redirect Claude's actions during a session. A prompt injection attack occurs when malicious instructions are hidden inside content that an AI is asked to process (e.g., a document) and those hidden instructions hijack what the AI does next. Instead of summarizing the contract you uploaded, Claude might be redirected, by text embedded in the contract itself, to do something else entirely. Anthropic has implemented safeguards, including prompt injection scanning, but has acknowledged it cannot disable the feature mid-operation once a session is underway.

But do not dismiss it. Cowork works well on your own notes and publicly available research materials. Keep it away from folders containing anything sensitive.

4. Claude Legal Skills

In Chapter 2, I noted that Claude's *general* Skills feature (the chat-based version that saves style and tone preferences) has been unreliable. What Anthropic released on January 30, 2026 as part of its Cowork plugin system is a different thing entirely, and it works considerably better. In May 2026, Anthropic expanded Legal Skills. It now provides access to third-party platforms such as Westlaw, Harvey, Box, and DocuSign. The update also includes specialized roles like commercial counsel, employment counsel, and even law student. Law firms can use these tools either within Anthropic's existing Cowork platform or integrate them directly into their own internal systems.

How do Legal Skills differ from General Skills. One is a preference-saving feature in the browser interface (general Skills); the other (Legal Skills) is a structured workflow system that runs inside Cowork.

Legal Skills are markdown files — plain-text documents written in plain English — that tell Claude how to approach specific legal tasks. Because they run in Cowork rather than the chat window, they benefit from Cowork's ability to read local files, execute multi-step workflows, and maintain context across an entire work session. They don't require prompt-writing. You run a command; Claude follows the workflow the Skill defines.

You'll find them under Claude Cowork | Customize | Skills | Plugins (see Image 3-1). To add Skills click the + sign. Select Legal to install.

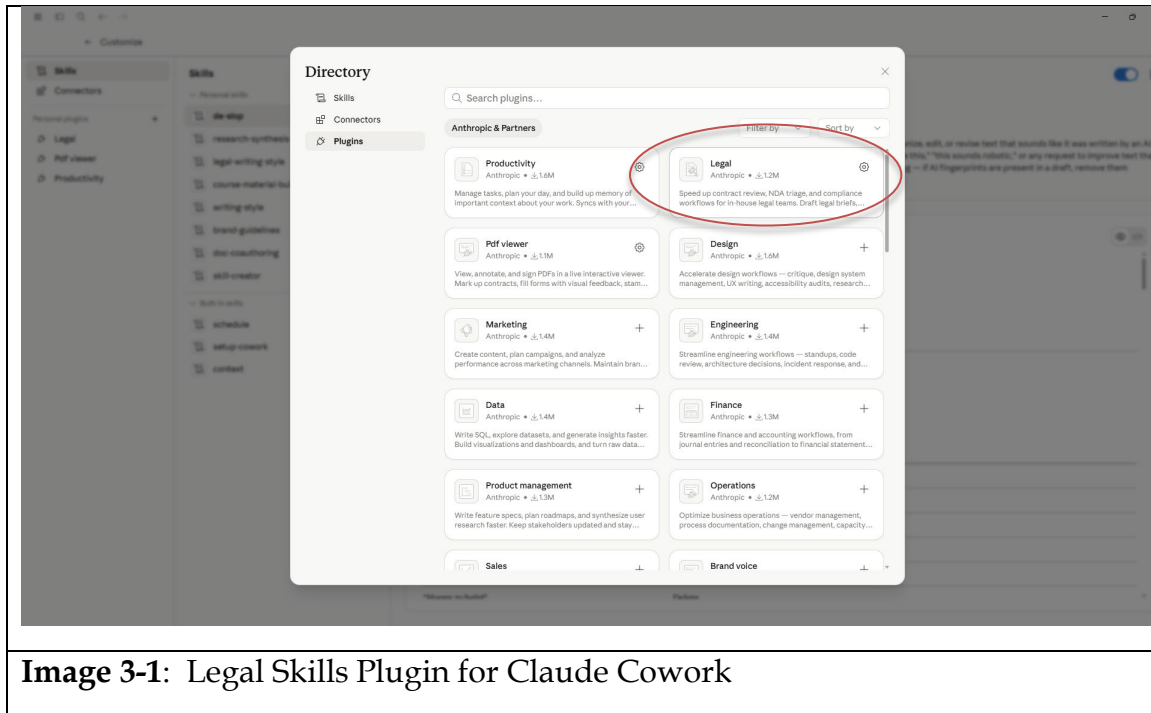


Image 3-1: Legal Skills Plugin for Claude Cowork

a) What the Pre-Built Legal Skills Do

Anthropic released five legal workflow commands as part of the plugin.

/review-contract performs a clause-by-clause scan of an uploaded contract. It extracts key provisions and deadlines, then assigns each a GREEN (acceptable), YELLOW (needs attention), or RED (significant issue) status. The output is a structured review document, not a conversational response.

/triage-nda categorizes incoming non-disclosure agreements on the same GREEN/YELLOW/RED framework. A GREEN NDA matches your standard acceptable terms and can move forward without counsel review. A YELLOW NDA flags provisions that depart from your baseline. A RED NDA signals material issues that require attorney attention before the agreement moves.

The plugin also includes pre-built workflows for compliance checking, legal risk assessment, and meeting preparation.

b) The Playbook: Where Your Standards Live

The quality of what these Skills produce depends almost entirely on a configuration file called the playbook. The playbook is where you define your organization's specific positions: your standard acceptable clauses, your risk

thresholds, the terms you'll never sign, and the escalation triggers that send a contract to senior counsel. Without a populated playbook, the Skills apply reasonable defaults. With a well-built one, they apply *your* standards.

This is also where customization happens. The Skills are open-source, released under an Apache 2.0 license, which means you can read them, modify them, and use them commercially without restriction. Because they're markdown files, you don't need a developer to edit them. A lawyer who can read a document can understand how to modify one. See Image 3-3. A firm can take Anthropic's NDA triage Skill, rewrite the escalation criteria to match its own risk tolerance, and add clauses specific to its industry or jurisdiction. What starts as a generic workflow becomes something that reflects how your firm actually thinks about risk.

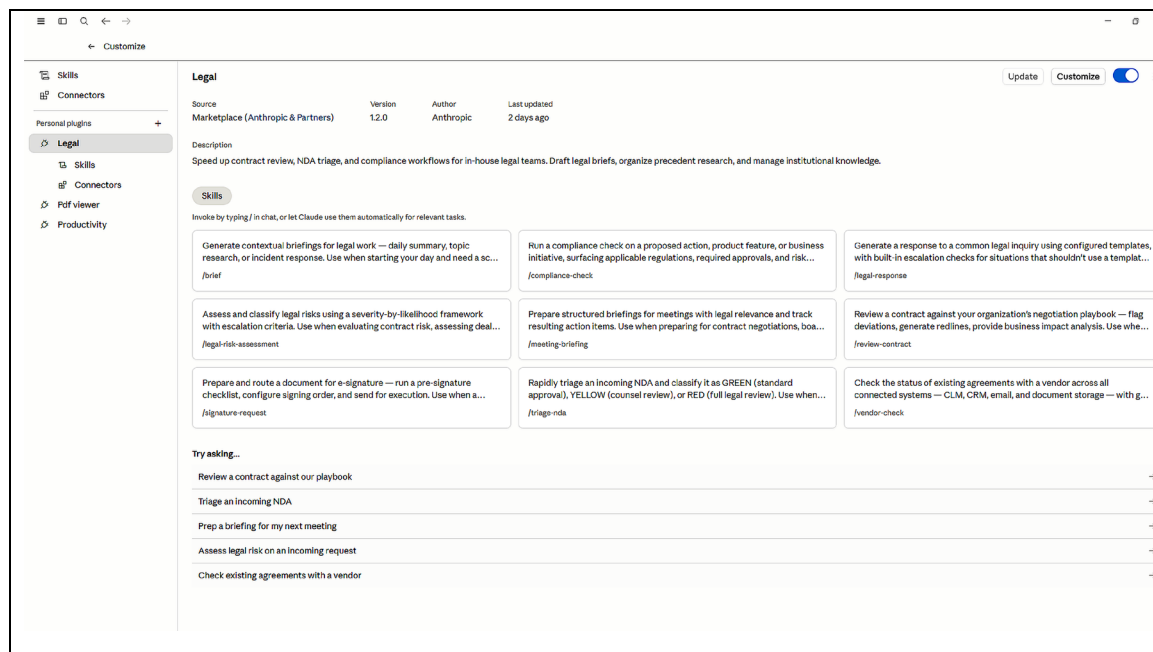


Image 3-2: Legal Skills in Claude Cowork

c) What They Can't Do — Yet

Anthropic labels these tools a research preview and explicitly advises against using them for high-stakes or regulated legal work without attorney review. That review requirement is an ethical obligation under Model Rule 1.1. These tools accelerate preliminary work, but they certainly don't replace the judgment.

The Skills also can't edit documents directly in Microsoft Word. You upload or paste contract text into the Cowork interface, and output comes back in Cowork. If your workflow requires working inside Word from start to finish, that's a gap that hasn't been closed yet.

Access requires the Claude Desktop app (macOS or Windows) and a paid subscription.

d) Why This Matters

When Anthropic released these Skills in early 2026, shares in Thomson Reuters and RELX (Lexis's parent company) fell sharply. Legal tech investors read the release as a direct threat to commoditized legal AI tools. If Claude can triage an NDA out of the box with a single command, why pay for a standalone tool that does the same thing less flexibly?

The honest answer is that enterprise legal tech platforms offer considerably more, of course, such as curated data, maintained integrations, audit logs, and a UI designed around how lawyers actually work. Anthropic's Legal Skills don't replace that. What they do, for faculty and practitioners who want to experiment without an enterprise commitment, is lower the floor for getting started with structured legal AI workflows to essentially zero.

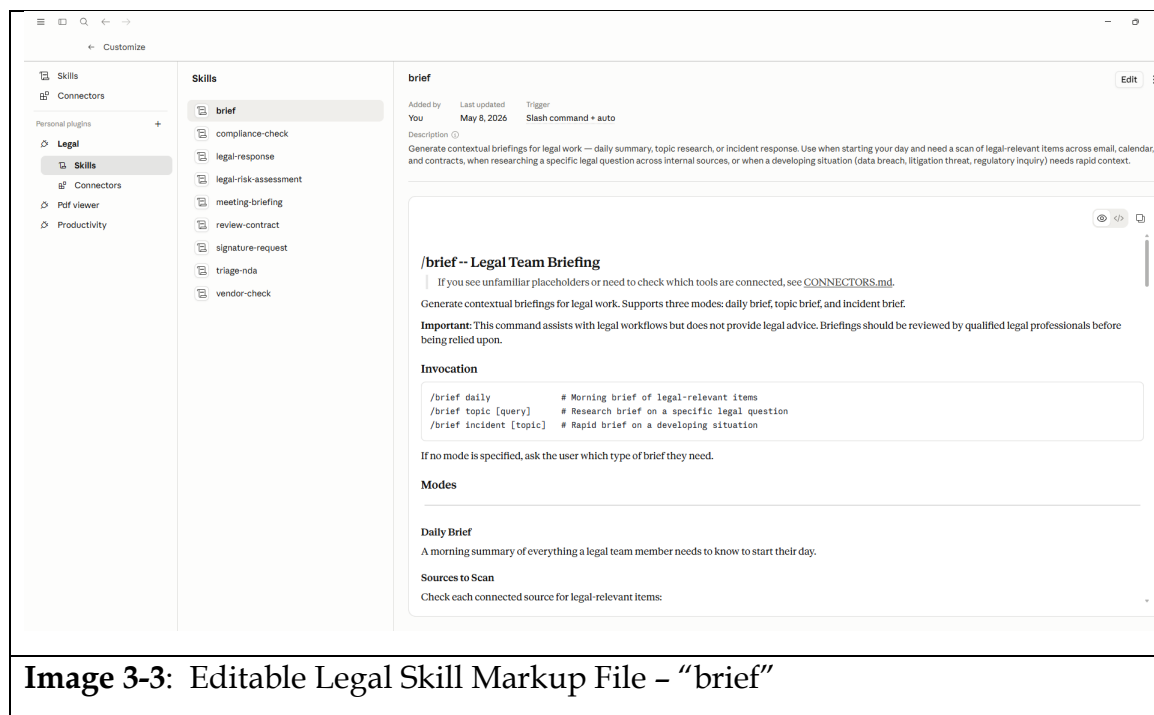


Image 3-3: Editable Legal Skill Markup File – “brief”

For law schools interested in showing students what AI-assisted legal work looks like in practice, the legal Skills plugin is a great option. The underlying files are readable and the workflow logic is visible. Students can work through the playbook configuration for a hypothetical client, run the Skill against a sample contract, and compare what it produces against their own analysis of the same document. As a teaching tool, it offers a transparency that most black-box legal tech tools do not.³⁴

To learn more about Claude for Legal Teams watch the webinar at <https://www.anthropic.com/webinars/claude-for-legal-teams>.

5. Creating Your Own Skills

Creating a Skill is more accessible than it sounds. At its core, a Skill is a text file — a markdown file with two parts: a short header that tells Claude what the Skill is for, and a body that tells Claude what to do when the Skill is active. No coding required.

a) The Structure

Every Skill lives in its own folder and contains a file called SKILL.md. The file has two sections.

The **header** (called frontmatter, set off by dashes at the top of the file) gives the Skill a name and a one-line description. Claude reads that description at the start of every Cowork session to decide whether the Skill is relevant to the current task. Write it the way you'd brief a new assistant. Be specific and concrete about when it applies.

The **body** contains the actual workflow instructions. These are the steps Claude should follow. Include examples of what good output looks like and any guidelines about format or scope.

Here is what a Skill looks like for a law faculty member who frequently generates Socratic exam questions:

```
---
name: socratic-exam-questions
description: Generates Socratic exam questions for law school courses.
Use when the user requests exam questions, hypotheticals, or class
```

³⁴ Anthropic, *Legal Plugin for Claude Cowork*, <https://claude.com/plugins/legal> (last visited Apr. 30, 2026). See also Spellbook, *Claude Legal Plugin Explained: Features, Cost, and Setup*, <https://spellbook.com/learn/claude-legal-plugin> (last visited Apr. 30, 2026)

discussion prompts based on case law or legal doctrine.

Socratic Exam Question Generator

When asked to generate exam questions or discussion hypotheticals, follow this workflow:

1. Identify the core doctrine or rule at issue in the provided case or fact pattern.
2. Draft a one-paragraph fact pattern that isolates that doctrine without making the answer obvious.
3. Generate two discussion questions: one that tests basic rule identification, and one that requires application to ambiguous facts.
4. Flag any jurisdictional assumptions built into the questions.

Format

- Keep the fact pattern to one paragraph.
- Write questions in plain language, not exam jargon.
- Do not include answer keys unless explicitly requested.

Notice that the description tells Claude *when* to load the Skill, and the body tells Claude *what to do* once it's loaded. Those are two distinct jobs, and both matter.³⁵

b) Step-by-Step: Creating and Installing a Skill

Step 1: Open any text editor. Notepad, TextEdit, or any application that saves plain text works fine. You do not need Word or a specialized coding tool.

Step 2: Write the header. Begin the file with three dashes, then your Skill's name and description, then three more dashes:

```
---
name: your-skill-name
description: What this Skill does and when Claude should use it.
---
```

Step 3: Write the body. Below the header, write the instructions Claude should follow. Use plain English. Number the steps if the workflow is sequential. Include a brief example of what good output looks like.

Step 4: Save the file as SKILL.md. The filename must be exactly SKILL.md.

³⁵ *Id.*

Step 5: Put the file in a folder. Name the folder after your Skill — for example, "socratic-questions" or "memo-reviewer." The folder name is how you'll identify the Skill later.

Step 6: Zip the folder. Right-click on the folder and choose "Compress" (Mac) or "Compress to > ZIP File" (Windows).

Step 7: Upload to Cowork. Open Claude Cowork, click "Customize" in the upper right, click the + button, choose "Create Skill," then "Upload a Skill," and select your zip file.

The Skill is now active. Claude will load it automatically when it determines the Skill is relevant, or you can invoke it manually by typing /your-skill-name in any Cowork session.

c) Have Claude Do It For You

Let Claude write the first draft. Describe the workflow you want in a regular Claude chat session and ask it to generate a SKILL.md file. Then edit from there. This approach consistently produces a better starting point than a blank file, and it models the workflow the Skill itself will eventually automate.

Keep the description specific. A description like "helps with legal writing" will rarely trigger because it's too vague for Claude to act on confidently. "Generates IRAC-structured memos for 1L legal writing courses when the user provides a fact pattern and a legal issue" is specific enough to load reliably.

Start simple and expand. A Skill with five clear instructions will outperform one with fifty ambiguous ones every time. Build incrementally and test after each significant change rather than writing the whole thing at once.

Keep the body under 500 lines. If a workflow is too complex for that, split it into two focused Skills rather than building one unwieldy one.

d) Finding Ready-Made Skills

You don't have to build from scratch. A substantial ecosystem of pre-built Skills already exists, and the quality at the top end is high enough for professional use — with appropriate review.

i. Anthropic's Official Repository

The canonical source is Anthropic's official GitHub repository at github.com/anthropics/skills. Anthropic maintains Skills for common document workflows (Word documents, PDFs, PowerPoint presentations, and Excel spreadsheets) and publishes them under an Apache 2.0 license, meaning that they are free to use, free to modify, and free to build on. The quality is consistent because Anthropic tests these against its own standards before publishing.

ii. Community Repositories

The community has produced considerably more than Anthropic has built officially. A few worth knowing:

skills.sh (github.com/VoltAgent/awesome-agent-skills) is the largest community collection, with over 1,000 Skills spanning document automation, research workflows, and enterprise functions. The Skills vary in quality since they're contributed by individual developers so treat anything here as a starting point that needs testing rather than a finished product. The database includes Skills for other platforms as well, such as Gemini, Word Press, ChatGPT, etc.

awesome-claude-skills (github.com/ComposioHQ/awesome-claude-skills) takes a curated approach rather than aggregating everything. The editorial filter tends to surface higher-quality options, and the accompanying descriptions are more reliable than what you'll find in a raw community dump.¹

agensi.io is a discovery platform that lets you browse and download Skills by category without navigating GitHub. Skills download as zip files ready to install. Some are free; others are sold by independent developers.

iii. Portability

One broader note worth flagging: the Agent Skills format is an open standard, which means Skills built for Claude also run on other compatible agents. If you find a Skill that wasn't built specifically for Claude, it will likely work in Cowork or Claude Code without modification.

6. Gemini Gems

a) What They Are

Gems are Gemini's simplified version of customized AI assistants. They're basically saved instructions that change how Gemini responds.

Think of Gems as creating different "modes" for Gemini, each optimized for a specific task.

b) Key Features

- Custom instructions (one paragraph to a few hundred words)
- File uploads (up to 10 files can be added to a single Gem; individual non-video files can typically be up to 100 MB); live access to Google Drive
- Web access
- Works on mobile and web (but Gems can only be created via the web interface)

c) Strengths

- Works on mobile (great for students using phones)
- Simple interface (less overwhelming than CustomGPT builder)

d) Weaknesses

- Gemini's writing quality is weaker than Claude or ChatGPT
- Inconsistent performance

7. Comparison Table: Which Feature to Use

	Custom GPTs	Claude Projects	Claude Cowork	Gemini Gems
Platform	OpenAI / ChatGPT	Anthropic / Claude	Anthropic / Claude Desktop	Google / Gemini
Core purpose	Shareable, pre-configured AI assistant	Persistent workspace with standing context and files	Agentic task execution on local files	Pre-configured AI assistant with Google Workspace integration
Task model	Conversational (you prompt, it responds)	Conversational (you prompt, it responds)	Agentic (you set a goal; Claude)	Conversational (you prompt, it responds)

			executes multi-step)	
Local file access	No	No	Yes — reads, writes, and creates files in approved folders	No
Document / knowledge capacity	Static file uploads; ~32K token limit per file	Unlimited file uploads; no document cap	Direct folder access; no upload required	Live Google Drive sync; 1M token context window
Instruction persistence	~8,000-character system prompt	Unlimited custom instructions across sessions	Global and folder-level standing instructions	~4,000-character instruction limit
Ecosystem integration	External APIs via Actions; broad plugin library	MCP connectors (Slack, Drive, Gmail, etc.)	MCP connectors plus local app control (browser, desktop apps)	Native Google Workspace (Docs, Drive, Gmail, Calendar)
Sharing / distribution	Public GPT Store; shareable by link	Team/Enterprise plan members only; no public store	Not shareable as a template; session-specific	Workspace admins can push to org; no public marketplace
Scheduling / automation	No	No	Yes — recurring and on-demand scheduled tasks	No
Mobile task assignment	No	No	Yes — Dispatch feature sends tasks from phone to desktop	No
Requires desktop app	No	No	Yes (macOS or Windows)	No
Paid subscription required	ChatGPT Plus (\$20/mo) to create	Any paid Claude plan	Pro, Max, Team, or Enterprise	Gemini Advanced (\$20/mo) or Google Workspace Business Standard
Regulatory / compliance caution	Standard data retention applies	Standard data retention applies	Not for regulated workloads; no audit logs	Standard Google data retention applies

Best fit for legal education	Building a reusable assistant to share with students (e.g., a pre-configured research guide)	Long-term document-heavy research across a semester; professor course prep	Synthesizing and organizing large local document sets; batch drafting tasks	Teams already in Google Workspace; large-document reading with Drive integration

8. A Workflow for Law Faculty Members

Most law faculty members who use advanced features effectively follow this workflow:

a) For Student-Facing Tools (CustomGPTs):

1. Identify the pedagogical goal (what should the students be learning?)
2. Draft instructions in Claude (Claude writes better instructions)
3. Gather relevant files (syllabus, rubrics, examples)
4. Create CustomGPT in ChatGPT
5. Test extensively before sharing with students
6. Share link with students
7. Collect feedback and iterate

b) For Personal Productivity (Claude Projects):

1. Create a Project for each course you teach
2. Upload: syllabus, readings, past assignments, grading rubrics
3. Start conversations for specific tasks (e.g., generating questions)
4. Keep Project updated as course evolves
5. Reuse Project semester after semester (update files as needed)

9. Important Reminder: Not All Students Have Access

The biggest limitation of CustomGPTs is equity. Both require paid subscriptions (\$20/month). If you build a brilliant CustomGPT for practicing issue-spotting, but only half your students can afford ChatGPT Plus, you've created a two-tier learning experience.

Before building student-facing tools:

1. Survey your students: What percentage have paid AI subscriptions and to which platforms?

2. Consider institutional solutions: Can your law school provide access?
3. Build optional, not required: Frame these tools as optional supplements
4. Provide alternatives: Also provide materials in accessible formats
5. Be transparent about costs: Tell students upfront what access requires
6. Remind them: subscriptions can be cancelled at any time (but, from my experience, no one cancels).

The pedagogical promise of CustomGPTs is exciting. But don't let that excitement override equity concerns. The best teaching tool is worthless if only privileged students can use it.

10.NotebookLM

NotebookLM is not another conversational AI platform. It won't answer your random questions. It won't generate creative hypotheticals (but it will create quizzes). It won't engage in back-and-forth dialogue about legal theory.

What NotebookLM does is this: you give it a collection of documents (up to 50 sources – PDFs, websites, Google Docs, even YouTube videos with a free account), and it becomes an expert on those specific documents. It can synthesize them, compare them, extract themes, answer questions about them, and create study materials from them. All with inline citations showing exactly where each piece of information came from.

Think of it as the difference between a general research assistant (ChatGPT, Claude, Gemini) and a specialized librarian who has read every document in your course reader and can help students navigate them.

For law faculty members, this feature means NotebookLM excels at one specific task that matters enormously: turning dense, complex course materials into accessible study aids that students can use.

a) The Big Advantage: It's Free

Unlike every other AI tool we've discussed, NotebookLM is completely free to anyone with a Google account. However, free accounts have limits on the number of notebooks, sources per notebook, and audio/chat generations. Regardless of these limits, free access is more than adequate for most users. Paid accounts (Plus, Pro, and Ultra), offer more notebooks, sources, and daily queries; access to the latest AI models; and customizable response styles and lengths. But more

important, paid accounts allow creation of shareable notebooks ("chat-only" notebooks without sharing source access).

This feature matters for equity. If you build a CustomGPT, only students with ChatGPT Plus (\$20/month) can use it. If you build a study guide with NotebookLM, every student can access it regardless of their financial situation.

For student-facing tools, NotebookLM's free access is its killer feature. The technical capabilities are good. The zero-cost accessibility is transformative.

b) Key Features

Document Upload (50 Sources maximum in the free version)

You can upload PDFs, paste text, link to websites, link to Google Docs, or even provide YouTube video URLs. Each source can be quite large (NotebookLM handles long documents well). For law faculty members, this feature means you can upload your entire course syllabus, 30-40 key cases from your casebook, supplementary readings, past exam questions and model answers, and lecture transcripts or video recordings.

Source Citations

Every claim NotebookLM makes includes a citation showing which uploaded document it came from, with a direct link to the relevant passage. This is different from ChatGPT inventing citations or Claude saying "based on your documents." NotebookLM shows you exactly where it found each piece of information.

For legal education, where citation accuracy matters, this is crucial. Students can verify claims. Faculty members can check whether NotebookLM accurately represented the source material. No one is taking AI output on faith.

c) Generated Study Materials

NotebookLM can automatically create study guides (comprehensive overviews), FAQs (common questions with answers), briefing documents (executive summaries), timelines (chronological organization), and table of contents (organized outline). These aren't just GenAI-generated summaries. They're structured, cited, pedagogically useful documents that students can study from.

d) Audio Overviews (The "Podcast" Feature)

This is NotebookLM's most distinctive feature. Click "Generate Audio Overview" and NotebookLM creates a podcast-style discussion between two AI voices

discussing your uploaded materials. It sounds like two people having a conversation about your course readings.

When it works well, this feature is remarkable. Students who struggle with reading dense legal materials may find audio explanations more accessible. Students studying during commutes, exercise, or other activities can make that time productive. It now has an interactive feature, where a user can ask the “podcasters” in real time clarifying questions.

e) Other features include:

- Video overview
- Infographics
- Mind Maps
- Flashcards
- Reports
- Quizzes
- Slide Decks
- Data Tables

f) *Real-World Use Cases for Law Faculty Members*

Use Case 1: Course Reading Synthesis

The Problem:

You assign 30 cases across the semester. Students struggle to see connections between cases, identify recurring themes, or understand how early cases relate to later material. By week 10, they've forgotten the holdings from week 2.

The NotebookLM Solution:

1. Upload all 30 assigned cases as PDFs
2. Upload your syllabus showing the organization and learning objectives
3. Generate a study guide covering all cases
4. Generate an FAQ answering predictable student questions

5. Generate an audio overview students can listen to during finals week

Result

Students have a comprehensive, cited reference document showing how all course materials connect. They can search it, reference it, and use it to review. Every claim includes a citation back to the specific case.

Use Case 2: Bar Exam Prep Materials

Upload NCBE subject matter outlines, past MEE questions and model answers, and bar prep lecture notes. Generate study guides for each subject, FAQs covering common confusion points, and audio overviews for final review week. Result: Comprehensive, cited study materials covering the entire bar exam scope, created in hours rather than months.

g) Excellent Source Citations

NotebookLM shows exactly where every piece of information came from. Students can verify. Faculty members can check accuracy. No one is trusting GenAI blindly. This citation transparency is especially important in legal education, where learning to verify sources is itself a crucial skill.

h) Deep Research

Google recently rolled out Deep Research to automate and simplify complex online research. It acts like a dedicated researcher, synthesising a detailed report or recommending relevant articles, papers or websites, and you can direct this "researcher" to search specific places, too.

With your guidance, Deep Research takes your question, creates a research plan, and browses hundreds of websites on your behalf, refining its search as it learns. In a few minutes, it generates an organized, insightful, source-grounded report.

Here's what makes NotebookLM's Deep Research different: the report is just the beginning. You can add the report *and* its sources directly into your notebook. You can continue to add other sources while Deep Research runs in the background, helping you assemble a rich knowledge base on any topic without leaving your workflow. And you can then use any of NotebookLM's capabilities, such as Audio or Video overviews, to pull insights or transform content to understand the topic better.

i) Best Uses for Law Faculty Members

- Synthesizing course readings into comprehensive study guides
- Creating accessible versions of dense materials for struggling students
- Generating audio study aids for bar exam prep or finals review
- Organizing research materials for scholarship or curriculum development
- Building cited reference materials that students can verify
- Creating FAQs from course materials
- Any task where source citation and verification matter

Avoid For

- Creative tasks (generating hypotheticals, original examples, practice questions)
- Conversational tutoring or coaching
- Tasks requiring information beyond your uploaded documents
- Real-time legal research
- When you need sophisticated back-and-forth iteration
- Grading student work

The Honest Assessment

NotebookLM is the best free tool available for document synthesis in legal education. The combination of zero cost, source citations, and large capacity makes it uniquely valuable for creating equitable student-facing resources.

But it's not a replacement for conversational GenAI platforms. It's a complement. Use NotebookLM for what it does best: synthesizing your course materials into accessible, cited study aids. Use ChatGPT, Claude, or Gemini for everything else.

For Faculty Members Concerned About Equity

If you can only recommend one GenAI tool to students, recommend NotebookLM. It's free, it's powerful, and it directly supports the most common student need: making sense of complex course materials.

Students with money can subscribe to ChatGPT Plus, Claude Pro, and every other premium service. But students without money should at minimum have access to NotebookLM. It levels the playing field in ways that no subscription-based tool can.

This doesn't mean NotebookLM is better than paid platforms. It means that it's more equitable. And in legal education, equity matters.

11.What's Coming Next

All platforms are rapidly developing these features, and new options continue to pop-up. By next semester, this comparison may look different. The fundamental trade-offs will remain: sharing vs. privacy, simplicity vs. power, equity vs. sophistication. But the specific features will evolve.

Check platform documentation quarterly. Test new features. Adapt your teaching tools as platforms improve. The tools that work best today may be obsolete in six months—but the pedagogical principles underlying their use will remain constant.

CHAPTER 4: The Art and Science of Effective Prompting

Is Prompting Enough?

At the start of my AI and the Law course, I want to establish a baseline. Before providing any formal instruction on prompt engineering, I give my students a scenario:

Corporate Merger Due Diligence Scenario: Your firm represents AeroTech Industries in acquiring a drone manufacturing startup. The startup has 47 patents, operates in 3 countries, and has pending litigation in Germany. The CEO wants a risk assessment before the board meeting next week.

Your Task: Draft a prompt that will help you identify potential legal obstacles to the merger.

I tell them they don't need to be experts in M&A law, and I tell them this is a hypothetical so there's no need to anonymize the facts. I just want to see how they approach crafting a prompt for GenAI.

The results last semester were revealing. One was a fair, but typical, prompt without specifics:

Help me with a merger. My client AeroTech is buying a drone startup. There are patents involved and some lawsuit in Germany. Can you identify legal issues that might be a problem? I need this for a board meeting.

This is the type of prompt that most untrained users would enter. The problem is that the LLM has to guess the jurisdiction, the nature of the patents, what kind of litigation is pending, what "legal issues" means in this context, what level of analysis the board actually needs, and whether the user wants a memo, a bullet list, or a paragraph. The student will get something very generic as a response that might be useful as background information, but it won't give targeted analysis in the proper format.

Another student's prompt produced a focused, practical memo-style risk assessment organized by jurisdiction, with specific flags about patent portfolio complications. It included cross-border regulatory hurdles and how the German litigation could affect deal timing.

Context: I'm a transactional attorney representing AeroTech Industries in the potential acquisition of a drone manufacturing startup. The target company holds 47 patents, operates in the U.S., Germany, and one additional jurisdiction, and has pending litigation in Germany. The deal timeline is tight — the CEO needs a risk assessment before a board meeting next week. I'm in the preliminary due diligence phase and have not yet reviewed the full patent portfolio or the German litigation file.

Objective: Identify the legal obstacles most likely to delay, complicate, or kill this acquisition. I want to know what to look for, what questions to ask the target's counsel, and what risk categories warrant independent investigation before we proceed.

Role: Act as a senior M&A attorney with experience in cross-border technology acquisitions and intellectual property due diligence.

Restrictions: Do not give me generic M&A checklists — I can pull those from any practice guide. Focus on the specific risk factors this deal presents: a patent-heavy tech company, multi-jurisdiction operations, and active foreign litigation. Flag areas where drone-specific regulatory issues (FAA, export controls, dual-use technology) are likely to intersect with the deal structure.

Expectations: Organize your output by risk category. For each category, identify the specific issue, explain why it matters in this deal context, and suggest the due diligence question or document request that would surface it. Keep the analysis tight enough that I can brief the CEO verbally in fifteen minutes.

Clarify: Assume U.S. law governs the deal structure, with German law relevant to the litigation exposure and EU regulatory considerations relevant to the target's European operations.

Tone: Direct and analytical. This is going out to a sophisticated client, not a first-year associate.

The first student's prompt gives the model almost nothing to work with beyond the bare facts. The CORRECT prompt (CORRECT prompting is described in detail below) tells the model who you are, what you already know, what you specifically don't want, and exactly how the output will be used. The model can't write a useful risk assessment without knowing that "useful" means "briefable to a CEO in fifteen minutes" rather than "exhaustive academic survey."

The CORRECT prompt gave the model enough to work with. The output was a prioritized risk memo written for a CEO rather than outside counsel – exactly what the scenario required. That student had taken my Law practice Technology course two semesters earlier, but formal coursework is one path among several. Webinars, peer learning, and guides like this one teach the same skills. And as you'll learn below, the LLM itself is also a resource: meta-prompting (asking the GenAI to critique or improve a draft prompt before you run it) can accelerate the learning curve considerably, and it costs nothing but a follow-up question.

All others students' prompts were similar to the fair prompt above. Their prompts produced generic educational overviews – the kind of content you'd find in a textbook chapter titled "Common Issues in M&A Transactions." There was no urgency and no client focus. The output wasn't in a memo format, and it had abstract explanations of what *could* matter in *some* mergers, with no connection to AeroTech's specific situation or timeline.

The difference in the results is proper prompt engineering, and that's where the science of prompting comes in. Unlike a colleague who fills in gaps from shared experience, an LLM works only with what you give it – so the skill isn't in knowing what you want, it's in translating that into language the LLM can act on. After all, you're "talking" to a computer.

Prompt engineering isn't just a technical skill. For law students and faculty members, it's fast becoming an essential component of legal literacy itself.¹ In law

¹ Neil Sahota, *Legal Prompt Engineering: Fast Becoming the Important AI Skill in Legal Services*, JD SUPRA (Apr. 11, 2024), <https://www.jdsupra.com/legalnews/legal-prompt-engineering-fast-becoming-7929289/>, last accessed Dec. 29, 2025 ("The ability to effectively communicate with AI systems through well-crafted prompts is not just a technical skill – it's a bridge to a future where lawyers and AI collaborate to achieve greater heights in the pursuit of justice.").

school, students learn that legal reasoning requires precision. Turns out, so does prompting GenAI.

A. The GenAI Fluency Gap

We've created an odd situation in legal education. We tell students that GenAI exists, we warn them about its limitations, we create policies governing its use. But we rarely teach them how to use it ethically and well. It's like teaching someone that cars exist, explaining they can be dangerous, creating rules about seatbelts, then handing them the keys without driving lessons.

The assumption seems to be that prompting is intuitive, that students will figure it out. After all, they're digital natives, right?

Not quite. Anyone who has spent time teaching students about technology, as I have, knows that "digital native" usually means they're adept at social media and downloading apps. It doesn't mean that they understand how technology works. I regularly encounter students whose laptops haven't been updated in years and who can't troubleshoot basic software issues (or understand that a reboot fixes a lot of issues). They typically don't know how to use Word or Excel, and before GenAI, the Powerpoints that they created were horrendous (not that many are very good). They can navigate Instagram effortlessly but struggle with file management or system maintenance. And when it comes to redaction, they've either never heard of it or think (as some lawyers apparently do) by changing a yellow highlight to black does the job.²

² See generally, Jamie Dupree, *Manafort Legal Team Botches Redactions, Reveals Campaign Contacts with Russian*, WFTV, Jan. 8, 2019 at 1:42 pm EST, <https://www.wftv.com/news/politics/manafort-legal-team-botches-redactions-reveals-campaign-contacts-with-russian/901165860/> (last accessed Dec. 29, 2025) (The "redacted" text revealed that Manafort had shared 2016 campaign polling data with Konstantin Kilimnik, a business associate with ties to Russian intelligence); Mark Gollom, *How Internet Sleuths are Un-Redacting Some of the Epstein Files*, CBC News, Dec. 25, 2025, <https://www.cbc.ca/news/world/epstein-files-unredacting-9.7027723>, last accessed Dec. 29, 2025 (Internet "sleuths" discovered that many of the redactions were not "burned" into the file. Some users used mobile phone photo editing tools or simple copy-paste techniques to bypass the black boxes); Brett Burney, *Don't Let Redactions Cause Distractions: How to Properly Censor and Remove Information When Reviewing and Producing Documents*, NextPoint Blog, <https://www.nextpoint.com/ediscovery-blog/redacted-legal-document-tips-document-review/> (last accessed Dec. 29, 2025) (listing several redaction errors made by attorneys).

To be fair, it isn't just the students who don't know how to create useful prompts. Most law faculty members and lawyers don't know how to do it either. And that redounds to job security for many of us "techies," as we're called.

The point is that familiarity doesn't equal competence. Students have been using Google their whole lives, but that doesn't mean they know how to prompt GenAI effectively. These are fundamentally different technologies requiring fundamentally different skills. Assuming students will "just figure it out" is like assuming someone who's great at texting will automatically excel at legal writing. Superficial similarity masks real differences in what's required.

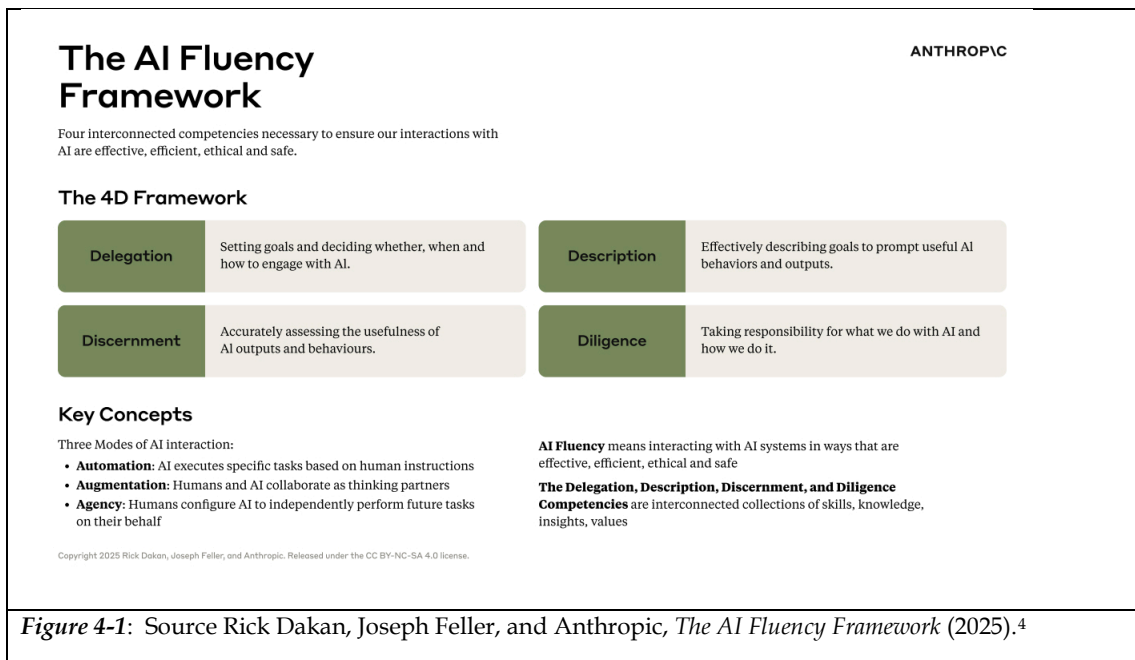
1. The AI Fluency Framework: Four Essential Competencies

Effective AI use in legal education requires more than just learning to write better prompts. Anthropic's AI Fluency Framework identifies four interconnected competencies necessary to ensure interactions with AI are effective, efficient, ethical, and safe: Delegation, Description, Discernment, and Diligence – the 4 Ds.³

a) Delegation

Delegation is making thoughtful decisions about what work is appropriate for you to do, for GenAI to do, or for you and GenAI to do together, and how to distribute those tasks. It requires both domain expertise and an understanding of GenAI capabilities. Before ever writing a prompt, ask: Is this task appropriate for GenAI? Is this something GenAI does well? Should students use GenAI to explain confusing cases or does that undermine the learning objective?

³ Rick Dakan, Joseph Feller, and Anthropic, *The AI Fluency Framework* (2025), <https://anthropic.skilljar.com/ai-fluency-for-educators> (last visited Mar. 17, 2026); see also <https://anthropic.skilljar.com/ai-fluency-for-students> (last visited Mar. 17, 2026); and see <https://anthropic.skilljar.com/teaching-ai-fluency> (last visited Mar. 17, 2026).



b) Description

Description is the ability to communicate with GenAI in ways that create a productive collaborative environment. This is where prompt engineering lives, which is the skill this chapter focuses on. Description encompasses three dimensions: *product description* (defining what you want in terms of outputs, format, audience, and style), *process description* (defining how the LLM approaches your request), and *performance description* (defining the AI system's behavior during your collaboration, such as whether it should be concise or detailed, challenging or supportive).⁵

c) Discernment

Discernment is the ability to thoughtfully and critically evaluate not just what what GenAI produces, but also how it produces it, and how it behaves. This competency works hand-in-hand with Description in a continuous feedback loop.⁶ You evaluate the quality of what GenAI produces (*product discernment*), how the GenAI arrived at its output (*process discernment*), and how the GenAI behaves during your interaction (*performance discernment*). For lawyers and law professors, discernment means never accepting GenAI output at face value. Every case

⁴ *Id.*

⁵ *Id.*

⁶ *Id.*

citation must be verified and every legal argument checked for logical coherence. Factual claims require independent confirmation before anything goes to a client or into print. And students must be reminded that, especially in the 1L year, they do not know what they do not know. A response can sound perfectly accurate and complete, but because they haven't achieved domain expertise, they cannot discern between a hallucination and an accurate and complete answer. However, there are techniques, such as knowledge grounding (discussed below), that students can leverage to help improve their learning and performance.⁷

d) Diligence

Diligence is taking responsibility for what we do with GenAI and how we do it. This includes *creation diligence* (being thoughtful about which AI systems you use and how you interact with them), *transparency diligence* (being honest about AI's role in your work with everyone who needs to know), and *deployment diligence* (taking responsibility for verifying and vouching for the outputs you use or share).⁸ Different contexts, such as personal, academic, or professional use may have different expectations for disclosure and verification, but the underlying principle remains constant: you are responsible for AI-assisted work product that bears your name.⁹

The four competencies are interconnected. Strong delegation decisions inform better description. Effective description enables more accurate discernment. Thorough discernment supports diligence in deployment. And diligence in reflecting on outcomes improves future delegation decisions.¹⁰

This chapter focuses primarily on Description: the art and science of prompt engineering. But as you read the prompting techniques that follow, remember that Description is only one component of AI fluency. The best prompt in the world won't help if you've delegated an inappropriate task to GenAI (Delegation failure), if you can't recognize hallucinated case citations (Discernment failure), or if you deploy AI-generated content without verification (Diligence failure).

⁷ *Id.*

⁸ *Id.*

⁹ *Id.*

¹⁰ *Id.*

The remainder of this chapter will equip you with sophisticated prompting strategies. In Chapter 3 we covered the strengths and weaknesses of different platforms. Prompt engineering and advanced prompting techniques (i.e., description) follow. Discernment and Diligence are discussed throughout the book.

B. Delegation and Description

Studies on prompt engineering effectiveness show that prompt quality dramatically affects output quality. Sometimes the difference is orders of magnitude.¹¹ A well-crafted prompt can produce sophisticated legal analysis. A poorly crafted one produces confident-sounding nonsense.

Legal reasoning requires precision. It demands attention to jurisdiction and procedural posture, and to the subtle distinctions between superficially similar concepts. A GenAI tool prompted carelessly might tell you that the parol evidence rule bars all extrinsic evidence about contract meaning. A well-prompted tool will explain the rule's exceptions and its variations across jurisdictions, including why the Second Restatement takes a different approach from the common law. The stakes are high. Law students who don't learn effective prompting will either avoid GenAI entirely (putting them at a competitive disadvantage) or use it poorly (putting them at risk of embarrassment or worse). Law faculty members who don't understand prompt engineering can't teach it or evaluate student use of it, which makes meaningful curricular integration impossible.

Before we dive into effective prompting strategies for legal education, it's important to establish a foundational understanding of what's happening under the hood when you interact with tools like ChatGPT or Claude. Chapter 2 offers an LLM Primer. Readers who aren't already well-informed about how large language models (LLMs) work, should review that content before proceeding. For readers seeking more detailed explanations, please consult my other work, *AI and the Practice of Law in a Nutshell*, where these topics receive a more detailed treatment.¹²

¹¹ See Michael Watkins, *Twelve Advanced Prompting Techniques for Large Language Models*, LINKEDIN (Aug. 1, 2024), <https://www.linkedin.com/pulse/mastering-advanced-prompting-techniques-large-language-watkins-lik9e>, last accessed Dec. 29, 2025 (documenting significant performance improvements from advanced prompting techniques across various reasoning tasks).

¹² KORIN MUNSTERMAN, *AI AND THE PRACTICE OF LAW IN A NUTSHELL*, WEST ACADEMIC (2026).

C. Beyond "Ask Better Questions"

Most advice about prompting amounts to "be specific" or "provide context." While that's not wrong, it's not sufficient either. Effective prompt engineering for legal education requires understanding both how LLMs work and how legal reasoning works. It's the intersection of these two domains that creates the framework for effective prompting. GenAI tools are prediction engines, not knowledge databases. They don't "know" anything in the way we understand knowledge. They predict what words should come next based on patterns in their training data. When you prompt a tool to "explain the Rule Against Perpetuities," it's not retrieving stored information about the Rule. It's generating text that statistically resembles explanations of the Rule that it encountered during training. It's basically like a very powerful auto-complete.

This has profound implications for prompt design. Because the tool is pattern-matching and predicting, the way you frame your prompt shapes the pattern it matches against. Ask for "an explanation of the Rule Against Perpetuities" and you'll get generic textbook language. Ask it to "explain the Rule Against Perpetuities as if you're a property professor helping a confused 1L who keeps confusing it with the Rule in Shelley's case" and you'll get something more useful, because you've activated different patterns in the model.

D. TIPS: Non-Intuitive Essentials Every User Should Know

These are the tips that experienced users wish someone had told them on Day One. They're not in the official documentation, and you won't discover them through casual use. But they'll save you hours of frustration.

1. TIP: Use SHIFT + ENTER to Add Line Breaks

When typing prompts, hitting ENTER alone submits the prompt immediately. If you want to add paragraph breaks or organize a multi-part prompt, hold SHIFT while pressing ENTER. This creates a new line without automatically submitting your prompt and triggering the GenAI to respond. Your prompts will be more readable, and you'll avoid accidentally sending half-finished instructions.

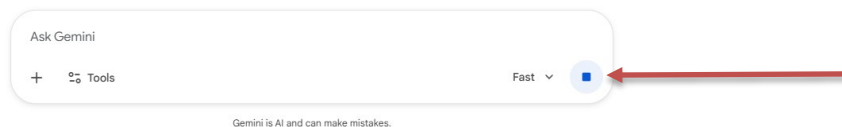
2. TIP: Stop and Edit When You Hit ENTER Too Soon

Accidentally hit ENTER before finishing your prompt? Don't let the LLM waste time generating a response to incomplete instructions. Click the **STOP** button, then click **EDIT** on your prompt. Finish writing what you meant to say, then resubmit.

Where to find STOP in ChatGPT, Claude, and Gemini

- Click the **black square (blue in Gemini, see below)** that appears while the LLM is generating
- Hover over your message and click the **pencil icon** (Edit)
- Complete your prompt and resubmit

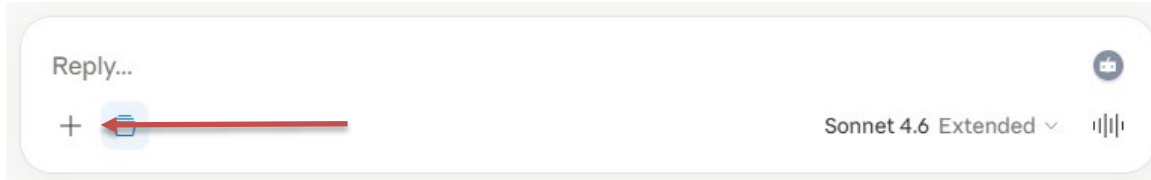
Specific prompting upfront saves tokens (the basic units of text LLMs process – and what platforms count against your monthly usage limit; exhaust them early and the model stops talking to you until the billing cycle resets) two ways: fewer messages that start with "ignore that, here's what I meant," and a shorter conversation for the platform to re-read. With each new message, the platform re-reads everything from the start of the chat, so longer conversations compound token use the way interest compounds debt. Stopping an poorly worded prompt saves tokens and extends your access to the LLM. Also, the CORRECT prompting framework pushes specificity to the front, which also helps save tokens. You can then iterate freely within a task; when the task ends, start a new chat.



3. TIP: Attach Documents to Individual Prompts (Without Adding to Knowledge Base)

When building CustomGPTs, you can upload files to the **Knowledge** section, making them permanently available for all conversations. See Chapter 10 for instructions on how to create Custom GPTs. But sometimes you want a document available for just *one* prompt without cluttering the knowledge base.

How to do it: In any individual prompt, click the **plus sign/paperclip icon** and upload your file. GenAI can reference it for that conversation thread, but it won't become part of the CustomGPT's permanent knowledge.



Example: You've built a CustomGPT for your Contracts course with the Restatement in “knowledge.” A student asks you to review their draft answer. Upload their draft as a prompt attachment, not to “knowledge,” so it doesn't pollute the CustomGPT's permanent materials.

When to use each:

Use Knowledge (permanent)	Use Prompt Attachment (temporary)
Course syllabus	Student work samples you're grading
Key cases (Pennoyer, Shoe, etc.)	Draft assignments you're testing
Restatements, model codes	One-off research articles
Assignment templates	Previous semester's exam for comparison

Why this matters: If you upload every student paper to your custom GPT's knowledge base, it becomes polluted with hundreds of irrelevant documents. Instead, keep knowledge clean with permanent reference materials, and attach individual documents as needed.

4. TIP: Rename Conversations to Find Them Later

By default, GenAI platforms auto-generate conversation titles based on your first prompt. These are often generic ("New chat," "Contracts help") and absolutely useless when you're searching for a specific conversation three months later.

How to rename:

In Claude:

- Click the **dropdown** next to the conversation title at the top of the prompt window
- Choose Rename
- Type a descriptive name.
- Press ENTER to save

In ChatGPT:

- Click the conversation title in the left sidebar under “Your chats”
- Click the three (:) dots
- Choose Rename
- Type a new name in the box
- Click outside the text box or press ENTER to save

In Gemini:

- Click the three dots (:) menu at the top right of the chat window
- Select Rename
- Type new name
- Click Rename

Best practices for naming:

- Include the date: "Contracts hypotheticals - Sept 2025"
- Include the subject and task: "Civ Pro - Jurisdiction outline"
- Include the product: "Final exam draft - Torts Fall 2025"
- Be specific enough to find it later: "Model answer - Pennoyer fact pattern" not just "Civ Pro"
- Add a prefix. For example, when I’m creating prompts to demo other faculty, I add the following prefix: “FACULTY DEMO:” followed by description.

Why this matters: Trust me. It matters. Learn from my mistakes. You'll generate dozens or hundreds of conversations. When you need to find that perfect hypothetical you created last semester, a descriptive title makes the difference between finding it in 10 seconds versus giving up and recreating it from scratch.

5. TIP: Highlight Text to Focus AI's Attention

In Claude and ChatGPT, you don't need to retype or copy-paste portions of previous messages when you want to discuss them. Simply highlight any text

from earlier in the conversation, then ask your question. The platform will understand you're asking specifically about that highlighted selection.

How to do it:

In Claude:

- Click and highlight any text in a previous chat (yours or Claude's)
- Click “Reply”
- Type your question in the prompt box
- Claude focuses its response on the highlighted portion

In ChatGPT:

- Click and highlight any text from earlier in the conversation
- Click “Ask ChatGPT”
- Type your question
- ChatGPT references the highlighted selection in its response

Why this matters for legal education:

When you're working with long case analyses, multi-part hypotheticals, or detailed legal explanations, you often want to zoom in on one specific aspect. Instead of writing, "In your previous response where you discussed the third element of negligence, can you explain the 'zone of danger' test in more detail?" you can simply highlight the relevant sentence(s) and ask "Explain the zone of danger test."

Example workflow:

1. AI generates a 500-word analysis of a Civil Procedure hypothetical
2. You highlight just the sentence about subject matter jurisdiction
3. You ask: "Is this still accurate after *Grable*?"
4. AI focuses only on that jurisdictional point

This saves time and reduces ambiguity by preventing the AI from re-explaining material you already understand. It is particularly useful when iterating on assignments or asking students to critique specific portions of AI-generated analysis.

See more Tips in Appendix D.

E. The CORRECT Framework

One useful structure for thinking about prompts comes from what I call the CORRECT framework: Context, Objective, Role, Restrictions, Expectations, Clarify, and Tone. This isn't a rigid template you need to follow explicitly all the time.¹³ You don't need to write "Context:" and "Objective:" in your prompts. Rather, it's a checklist of elements that effective prompts typically include, though not always in this order and not always explicitly labeled.³

Context means providing the relevant background information the LLM needs to understand your task. In legal prompting, this often includes jurisdiction, the specific legal issue, relevant facts, and procedural posture. If you're asking for help analyzing a Contracts problem, GenAI needs to know whether you're dealing with common law or UCC, and whether this is a formation or breach question. And it needs the relevant facts.

A prompt that omits the procedural posture may produce an analytically sound response to the wrong question. One that skips jurisdiction may give you a confident answer grounded in the law of a state that has nothing to do with your client's dispute. The model will not ask clarifying questions the way a supervising attorney would; it will work with what it has and fill gaps silently; supplying complete context up front is worth the extra thirty seconds it takes.

Objective defines your broader goal or purpose for using GenAI. "I need to draft a memo for my boss analyzing whether we have grounds for summary judgment" or "I need to prepare the strongest arguments to present to my client about settlement value" tells the LLM what you're ultimately trying to accomplish. You'll refine this further in Clarification below.

A clear objective also shapes tone and depth in ways that matter practically. A memo for a supervising partner calls for different register and structure than a client-facing settlement summary, and stating the objective explicitly lets the model make those adjustments without additional instruction. For law students,

¹³ The CORRECT framework was developed by the author for legal education contexts. While similar frameworks exist in prompt engineering literature, this version emphasizes elements particularly relevant to legal pedagogy. The framework serves as a heuristic rather than a rigid formula, with effective prompts incorporating these elements flexibly based on context and objective.

articulating the objective before anything else is also good disciplinary practice because it mirrors the threshold question any competent attorney asks before beginning a work product: what is this document for, and who is going to read it?

Role(s) tells the GenAI what perspective to take. Should it respond as a patient tutor, a demanding professor, a cautious practicing attorney, or a zealous advocate? Different roles activate different patterns in the LLM's training data and will produce different types of outputs. It's similar to asking the LLM to write a passage in the style of Shakespeare, then like Dorothy Parker, and then by Dr. Seuss.

For teaching purposes, you might want the GenAI to take on the role of a Socratic professor who asks follow-up questions but never directly provides the answer. For bar prep, you might want a straightforward answer with clear rule statements (you'll have to exert your "discernment" muscles).

Role instructions become more powerful when you assign parts to both sides of the exchange. Specifying your own role alongside the model's changes the register of the entire interaction. "I am a first-year law student who has read the assigned cases but has not yet attended class. You are a Socratic professor." produces a meaningfully different response than the same prompt without that framing. The model calibrates its vocabulary and questioning depth to the declared relationship between the two roles, adjusting for whatever level of assumed knowledge that relationship implies.

For law professors, this dual-role structure opens up useful classroom applications. You can position yourself as the student and the model as the professor to generate Socratic dialogue you can then critique or adapt for actual classroom use. You can reverse the roles — casting the model as a confused 1L and yourself as the professor — to pressure-test whether a hypothetical is clear enough to produce the analysis you want. You can even assign adversarial roles within a single prompt: "I represent the plaintiff. You represent the defendant. We are in a Rule 12(b)(6) hearing. Respond to my arguments as defense counsel would." The model will not replicate the judgment of an experienced litigator, but it will surface counterarguments worth anticipating.

Restrictions set boundaries on what the GenAI should and should not do, and they matter more in legal contexts than in almost any other professional setting. Left unrestricted, a model will fill gaps in its knowledge with plausible-

sounding content such as invented citations and conclusions that paper over ambiguity rather than flagging it.

A well-constructed restriction addresses the specific failure mode you most want to avoid. If fabricated citations are the concern, instruct the model explicitly not to cite any case it cannot reproduce with full accuracy, and to signal uncertainty when its knowledge of a specific holding is incomplete. If jurisdictional drift is the problem — the model applying federal doctrine to a state-law question, or conflating majority and minority rules — restrict the analysis to a named jurisdiction and require the model to identify when it is drawing on law from elsewhere. If the risk is overconfidence on contested issues, require the model to flag areas of legal uncertainty rather than resolve them by assertion.

Restrictions also serve a pedagogical function. When students see a prompt that explicitly prohibits conclusory statements or requires the model to distinguish holdings from dicta, they learn something about legal reasoning in the process. Namely, that those distinctions matter enough to enforce even on a machine.

Expectations specify the format and structure you want. Should the output be a formal memo, a casual explanation, an outline, a list of questions? Should it be 200 words or 2000? Should it include verified citations? Should it follow IRAC structure? Being explicit about format helps the AI match its output to your needs.

Clarify restates the core task at the end of the prompt to ensure the model has a clean, unambiguous target after absorbing everything that came before it. Without this step, a long prompt risks burying the actual assignment and the model may weigh earlier material more heavily than the specific question you want answered.

The clarification does not need to be elaborate. After providing context, objective, role, restrictions, and expectations, a single direct sentence is usually enough: "Your task is to identify the three best arguments for both sides of the dispute." That sentence does two things. It confirms the objective stated earlier in the prompt, and it signals to the model that everything preceding it was preparation rather than the primary instruction. Think of it as the legal equivalent of the prayer for relief at the end of a complaint — the rest of the document builds the case, but the prayer tells the court exactly what you want.

Clarify is particularly useful when the prompt is long or when the task involves competing considerations. If you have given the model a complex fact

pattern, restricted it to a specific jurisdiction, assigned it a particular role, and asked it to account for procedural posture, restating the core question at the end prevents the model from producing a response that addresses the context rather than the task.

Tone indicates the style and voice you want. Should the response be formal or conversational? Technical or accessible? Confident or appropriately hedged? For a non-lawyer or an experienced litigator? For law students, you might want a tone that's encouraging but rigorous. Professors and practicing attorneys likely want something more direct and efficient.

Here's an example of a CORRECT-structured prompt for a Torts class:

Context: You're teaching a 1L Torts class teaching negligence.

Objective: You want to create an essay question for a midterm exam that illustrates the difference between actual cause and proximate cause, with facts that make the distinction meaningful but not too obvious.

Role: You are an experienced Torts professor who designs challenging but fair exam questions.

Restrictions: Don't make the facts too complicated. Keep it to one paragraph. Don't include any intentional torts issues. Use realistic facts, not bizarre hypotheticals.

Expectations: The output should be a single paragraph fact pattern followed by the question: "Analyze [defendant]'s potential liability for negligence."

Clarify: Generate a Torts midterm essay question that requires students to carefully distinguish between actual and proximate cause.

Tone: Keep the tone professional and clear, appropriate for a law school exam.

Result from Claude

Sarah, the owner of a small retail shop, had been notified by the city three weeks earlier that the concrete step leading to her store entrance was cracked and uneven, creating a tripping hazard. She placed a "Watch Your Step" sign near the entrance but did not repair the step due to cost concerns. On a busy Saturday afternoon, Michael was walking toward Sarah's store when he tripped on the broken step, stumbled forward, and dropped his cell phone, which skidded into the adjacent parking lot. Without pausing to look for oncoming traffic, Michael immediately ran into the parking lot to retrieve his phone and was struck by a car driven by another customer who was backing out of a parking space at a reasonable speed. Michael suffered a broken leg and other injuries from being hit by the car.

Analyze Sarah's potential liability for negligence.

Figure 4-2. The [CORRECT Method trainer](#) is an interactive tool that lets you practice building prompts using the framework in this chapter. It generates realistic legal scenarios, walks you through each component, and gives you AI-powered feedback on your prompt before you move on. A free Claude account is required.

Again, you don't need to explicitly label each element or include every element in every prompt. But thinking through these components helps you craft prompts that give the LLM what it needs to produce useful output. I advise my students to use the labels until they've mastered the technique.

The CORRECT Method Trainer was created using a Claude Artifact. To learn how to create an Artifact, see Appendix J.

F. The Building Blocks: Basic Prompting Principles

Before diving into advanced techniques, it's worth establishing core principles that apply across all prompting situations. These aren't revolutionary insights, but they're easy to forget when you're in the middle of a task.

Specificity beats generality every time. "Explain the Rule Against Perpetuities" will give you a generic explanation. "Explain why the grant 'to A for life, then to B's children who reach age 25' violates the Rule Against Perpetuities under the common law approach" gives the GenAI enough specificity to provide targeted analysis. The more specific your prompt, the more specific (and useful) the response. Iteration is expected and productive. Your first prompt rarely produces exactly what you need. That's fine. GenAI tools work best through conversation. You make a request and evaluate the response, then refine your prompt based on what worked and what didn't. This isn't a sign of failure; it's how the technology is designed to work. Untrained users often feel they should get perfect results from their first prompt, as if needing follow-up questions reveals incompetence. The opposite is true. Skilled prompters iterate constantly.

If asking the LLM to generate multiple choice questions, ask it to create 5. Evaluate those questions. If you like them, tell it to create 45 more. If it needs tweaking, tell the LLM how you want it to adjust, then have it create 5 new questions. More tweaking needed? Repeat the process. If you now have what you want, ask it to create the next X number of questions. Remember, it's a conversation, not a one-and-done. It's also good practice to let the LLM know what it did well to help it improve in subsequent prompting conversations.

If you want the GenAI to generate something in a particular format or style, provide an example or two. Showing examples is often more effective than

describing what you want. And when you want the GenAI to teach you something, asking it to include examples makes the explanation far more concrete and memorable.

The principle runs in both directions. Positive examples show the model what you want, while negative examples show what to avoid, and both are worth supplying when the failure mode is predictable. If you are asking the model to draft a client letter and you know it tends to produce overly hedged, jargon-heavy prose, attach a brief example of that pattern and instruct the model to avoid it. In legal writing contexts, where the gap between competent and excellent work is often a matter of precision and register rather than substance, that kind of negative exemplar does real work.

For professors building course materials, examples also serve as implicit rubrics. Showing the model a strong student analysis alongside a weaker one, and asking it to produce something at the level of the former, produces more consistently useful output than a descriptive instruction alone.

Context shapes everything. GenAI tools don't have real understanding or memory or experiences to draw from, but they do respond to context within a conversation. Each exchange within a conversation builds on what came before. This means that you can establish context once and then build on it through multiple prompts. You can tell the AI "I'm a law professor who teaches criminal law. I'm preparing materials to help students understand the distinction between murder and manslaughter ..." and then ask a series of follow-up questions without re-establishing that context.

Restrictions improve output. This principle seems counterintuitive, but it's consistently true. Highly constrained prompts often produce better results than open-ended ones. If you ask "help me create a rubric for a memo assignment," you'll get meandering generalities. If you ask "create a 30-point rubric for grading

Why Restrictions Work

Restrictions work because they narrow the solution space. When you give the GenAI complete freedom, it's statistically likely to produce the most common, generic response patterns. Constraints eliminate the generic center and push the GenAI toward more specific patterns in its training data. It's the difference between asking "what should I wear?" and "what should I wear to a formal outdoor wedding in October in Chicago?" The more specific the question, the more useful the answer.

objective memo assignments in a 1L Legal Writing course, focusing on issue identification, rule explanation, and application to facts," you'll get something more useful. Restrictions force both you and GenAI to focus.

G.15 Advanced Techniques for Legal Pedagogy

Once you're comfortable with basic prompting principles, several advanced techniques become available. These aren't necessary for every task, but they can dramatically improve results for complex legal reasoning tasks.

The examples that follow focus on illustrating each specific technique. For brevity, they don't always include all of the CORRECT framework elements, though the most effective prompts incorporate most of them. As you become more comfortable with prompting, these elements become a natural part of your prompting. You'll provide context, specify objectives, define roles, and set constraints naturally without labeling each component. The CORRECT framework is a thinking tool, not a rigid template.

1. Chain-of-Thought Prompting

Chain-of-thought prompting guides the GenAI to break down complex reasoning into explicit steps. Instead of asking directly for an answer, you ask the GenAI to show its work. It's particularly valuable in legal education because legal reasoning is inherently step-by-step, and we want students to see that process.

A basic prompt might be: "Is this a valid contract?" A chain-of-thought prompt would be: "Analyze whether a valid contract exists by working through each element step-by-step: offer, acceptance, consideration, and capacity. For each element, first explain what the law requires, then apply it to these facts, then reach a conclusion on that element."

Chain-of-thought prompting is especially useful when students are learning legal doctrine. It forces the GenAI to model the analytical process, not just deliver conclusions. It also makes errors more visible. If the GenAI makes a mistake in step two of a five-step analysis, you can see where the reasoning went wrong, which is much harder to do with a conclusory response.⁴

2. Few-Shot Prompting

Few-shot prompting means providing the GenAI with examples of what you want before asking it to generate something new. This approach is incredibly powerful for exam question creation or any time when you want consistency in format or style.

Suppose you want to create multiple-choice questions for a Torts exam. Instead of describing what makes a good MCQ, you could provide the AI with two example questions that match your preferred difficulty level and structure, then ask it to generate new ones in the same style:

Here are two MCQs I've used successfully on past Torts exams: [Insert your two example questions with answer choices and explanations]. Now create three new MCQs on proximate cause that follow the same format: realistic fact patterns, four plausible answer choices, intermediate difficulty level, and a clear best answer that requires students to apply doctrine rather than just recite rules.

The GenAI will pattern-match to your examples, producing outputs that match your preferred question structure, difficulty level, and the balance between clear correct answers and plausible distractors.

Few-shot prompting works for formatting, too. Suppose you're developing a new course and want your syllabus to match the format you've used in other classes. Instead of describing your preferred structure, you could provide the GenAI with one or two of your existing syllabi:

You are a law professor with 10 years of experience teaching. Provided are syllabi from two courses you've taught. [Your Contracts syllabus showing your specific format, section organization, policy language, and assessment structure] and [Your Torts syllabus in the same format].

Draft a syllabus for my new Cybersecurity Law seminar using this same format and organizational structure. [Include other specifics such as the semester start and end dates, topics covered, how often the class meets and when, the number of credit hours, grading, etc.]"

The GenAI will pattern-match to your examples, producing a draft that mirrors your existing syllabi with the same section headings that matches learning objectives and assessment breakdowns. You get consistency across your courses without having to manually replicate formatting decisions or copy-paste policy language from old syllabi. It's typically safer than the common practice of copying last semester's syllabus and manually updating dates, which risks leaving incorrect information like old exam dates or outdated office hours.

Few-shot prompting works for any task where format or style consistency matters. The limitation is that examples take up space in your prompt and in the AI's context window (which we'll discuss later). For simple tasks, few-shot prompting might be overkill. But for tasks where format or style consistency matters, it's remarkably effective.

It's also quite good for generating law review article titles.

3. Zero-Shot Chain-of-Thought

This technique combines the step-by-step reasoning of chain-of-thought with the flexibility of zero-shot prompting (no examples needed). The key is adding a phrase like "let's think through this step-by-step" or "work through this systematically" to your prompt.

Research shows that this simple addition significantly improves performance on reasoning tasks.¹⁴ What you're doing is activating patterns in the training data where problems were solved methodically rather than intuitively.

For legal education, zero-shot chain-of-thought is perfect when you want the LLM to reveal its reasoning process without knowing in advance what steps will be needed. "Let's think through step-by-step whether this statute applies to these facts" lets the LLM determine what steps are needed (understanding the statutory language or applying canons of construction, etc.) without you having to specify them in advance.

¹⁴ Jason Wei, et al., *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*, ArXiv:2201.11903, Jan. 10, 2023.

4. Prompt Chaining

Sometimes a single prompt can't accomplish everything you need. Prompt chaining means breaking complex tasks into a sequence of prompts, where each builds on the previous response. The output from prompt one becomes the input for prompt two.

For example, you might use this sequence:

You are a Contracts professor who wants to explain to students how to use IRAC on an exam question. (1) Generate a fact pattern that illustrates the difference between express and implied contracts, (2) Now draft two student essays answering that question, one that would earn an A and one that would earn a C, (3) Explain what makes the A essay strong and the C essay weak, with specific reference to how each handles the distinction between express and implied contracts.

Prompt chaining lets you accomplish sophisticated tasks that would be difficult to specify in a single prompt. It also mirrors the way that expert legal reasoning works. Lawyers break complex problems into manageable pieces, solve each piece, then synthesize the results.

The tradeoff is that chaining takes more time and requires you to evaluate intermediate outputs. But for complex pedagogical tasks, it's often worth it.

5. Socratic Prompting

Socratic prompting is tailor-made for legal education. Instead of asking the GenAI to answer questions, you ask it to generate questions that probe student understanding.

I'm teaching the Parol Evidence Rule. Generate five Socratic questions that would help students understand why the rule exists and what policies it serves.

You can also create Socratic tools using “CustomGPTs” for your students. Here’s one example: [UNTD CoL: Socratic Method Simulator](#). See Chapter 10 for information on how to create your own CustomGPTs.

Instead of the LLM providing answers, it provides questions.

6. Meta-Prompting

Meta-prompting can be used in two ways. First, use it to instruct the LLM to plan its approach before executing a task. Instead of asking directly for an explanation of the Dormant Commerce Clause, you might write:

Before explaining the Dormant Commerce Clause, consider what technique would work best for a student struggling with abstract concepts. Then use that technique.

The extra step can feel like unnecessary overhead, but it consistently produces more tailored output for complex teaching challenges. Framing the request this way activates training patterns associated with careful, structured problem-solving rather than generic explanation resulting in better output quality.

Second, use meta-prompting to help you create a prompt. I know it sounds circular, but instead of crafting the prompt yourself, you describe what you need and ask GenAI to write the prompt for you. A law professor who wants to build a negligence hypothetical might write:

I'm teaching 1L Torts and I need a practice problem that forces students to distinguish actual cause from proximate cause. The facts should be realistic and fit in one paragraph. Write me a prompt I can use to generate this problem.

The AI drafts the prompt, you review and refine anything that misses your intent, and then run the prompt.

The technique is most useful when you know what you want but struggle to translate it into precise instructions, which describes most people at the start of their prompting practice. The AI has processed enough well-constructed prompts to recognize what effective ones look like for a given task, and it will often include constraints or format specifications you would have omitted. Treat the draft as a starting point rather than a final product. Read it critically before running it: check whether the role assignment fits, whether the restrictions actually limit what you want limited, and whether the output format matches what you need. A prompt drafted for you is still a prompt you're responsible for.

7. Devil's Advocate Prompting

Devil's advocate prompting asks the GenAI to challenge positions, identify weaknesses, or generate counterarguments, which is particularly valuable in legal education where we want students to see multiple sides of an issue.

Instead of asking "Is this a good argument?" ask:

You are a devil's advocate. Identify the three biggest weaknesses in this argument and explain how opposing counsel would exploit them.

For faculty members creating teaching materials, Devil's Advocate helps generate realistic opposing arguments for case discussions or exam answers. "I'm planning to have students argue that this statute is unconstitutional under the Commerce Clause. Generate the strongest possible counterarguments the government would make, including case law that would support the government's position."

The key is explicitly invoking the adversarial role. Without it, GenAI tends toward balanced, neutral explanations. With it, you get sharper, more aggressive critiques that mirror actual legal practice.

8. Progressive Narrowing (Funnel Prompting)

This technique asks the GenAI to move through multiple stages of refinement within a single response, showing its reasoning at each stage. Instead of iterating back and forth across multiple prompts, you structure one prompt that requires the AI to generate options, evaluate them, narrow the field with explanation, and ultimately select the best choice with explanation.

The structure typically looks like this: generate many options, narrow to fewer with rationale, narrow to one with detailed explanation.

For example, suppose you're developing a new seminar and need to choose a title:

You are a law professor proposing a new course. Generate 10 possible course titles for a seminar on AI regulation in healthcare. Then narrow these to the 5 strongest titles and explain why you eliminated the other 5. Finally, select the single best title from those 5 and provide a detailed explanation of why it's superior to the remaining options. Consider factors like clarity, student appeal, and accurate representation of course content.

The AI will provide all three stages in one response: the full list of 10, the narrowed list of 5 with elimination rationale, and the final selection with comprehensive reasoning. You see the complete decision-making process in one place. You can evaluate whether the GenAI's criteria make sense and whether you agree that the final choice is truly the best. If you disagree with how the narrowing happened, you can ask for a variation that keeps different options in play.

You can also use it to help select among competing assignment designs or choosing which cases or topics to include in limited class time. It can help you decide between textbook options or pick the best approach to a complex pedagogical challenge.

9. Retrieval-Augmented Generation (RAG)

RAG is a technique where you provide specific documents or source materials directly to the AI along with your prompt, so it grounds its response in those materials rather than relying solely on its training data. Think of it as giving the GenAI a specific file to work from rather than asking it to work from memory. When discussing providing “knowledge” throughout this book, I’m alluding to the RAG technique.

*RAG is incredibly useful to law faculty and students
because you often want the GenAI to work with specific course
materials, cases, or documents rather than generating generic
legal explanations.*

For example, instead of asking "explain the holding in *Pennoyer v. Neff*," upload the actual case and ask:

You teach Civil Procedure. Based on this opinion [provided], create five discussion questions that would help students understand why the Court found personal jurisdiction lacking.

The LLM works directly from the text you provided rather than from whatever version of *Pennoyer* exists in its training data. It yields fewer inaccuracies and responses tailored to the specific materials that you're teaching.

Common faculty uses for RAG:

- **Creating assessments from course materials:** Upload three cases you've taught and ask the GenAI to generate an exam fact pattern that requires students to synthesize holdings from all three.
- **Developing discussion questions:** Upload a dense Supreme Court opinion and ask for questions that would guide students through the Court's reasoning step-by-step.
- **Generating teaching materials:** Upload your existing syllabus and ask the LLM to create a day-by-day reading schedule that distributes the workload evenly across the semester.
- **Comparing approaches:** Upload two competing law review articles and ask for a chart comparing their arguments, methodology, and conclusions.

The key advantage is accuracy and specificity. When the GenAI works from documents you've provided, it's much less likely to hallucinate case citations or generate analysis that contradicts your course materials. But it doesn't eliminate hallucinations entirely. The best approach is to combine CORRECT prompting with RAG.

There are limitations to RAG. First, you need digital versions of your materials (easy enough in most cases). Second, very long documents may exceed the GenAI's context window. See Chapter 2 if you don't know what context windows are.

Most current GenAI tools support RAG through simple file upload (see TIP 3, above). You don't need technical expertise. Just upload the document and

reference it in your prompt: "based on the attached syllabus..." or "using the case I've uploaded..."

10. Structured Output Prompting

Ask GenAI to organize information in structured formats such as tables or comparison grids. Example: "Create a comparison table showing the elements of negligence across three jurisdictions: common law, Second Restatement, and Third Restatement. Include columns for duty, breach, causation, and damages." Or, "Generate a comparison table of the major Fourth Amendment warrant exceptions. Include columns for the exception name, the triggering circumstances, whether probable cause is still required, and the leading Supreme Court case."

The technique works particularly well for teaching materials and quick-reference guides that students will return to throughout a course.

11. Synthesis Prompting

Ask GenAI to synthesize information from multiple sources into coherent analysis, which is a core legal skill. Example: "I'm teaching three cases on trademark dilution: *Moseley*, *Nike*, and *Starbucks*. Synthesize the test that emerges across these cases and identify where the courts disagree." For better accuracy, consider providing the opinions, and asking the GenAI to summarize the holdings. Double-check the output, and then ask the GenAI to synthesize.

Synthesis prompting is great for creating case summaries and lecture outlines, or for helping students see doctrinal evolution. For example, "Trace the doctrinal evolution of personal jurisdiction from *Pennoyer v. Neff* through *International Shoe* to *Asahi Metal* and *Bristol-Myers Squibb*. For each case, identify what the Court changed, what it left unresolved, and how the next case responded to that gap."

12. Scaffolding Prompting

Scaffolding prompting breaks a complex task into a deliberate sequence of smaller prompts, each building on the last. Rather than asking GenAI to produce a finished product in one shot, you guide it through stages: establishing foundational concepts first, then layering in complexity, then applying the result

to a specific context. The analogy is construction scaffolding: the structure goes up piece by piece, and each layer supports the next.

For legal pedagogy, it is particularly useful when you are creating materials for students who have not yet encountered a doctrine, or when the doctrine itself has sequential dependencies. Personal jurisdiction is a good example.

A scaffolded prompt sequence might look like this:

(First prompt) Explain the basic holding of <i>International Shoe</i> and the minimum contacts standard in two paragraphs suitable for a first-year student.
(Second prompt) Now explain how <i>Burger King v. Rudzewicz</i> extended and complicated that standard, assuming the student already understands <i>International Shoe</i> .
(Third prompt) Draft a hypothetical that tests whether a student can apply both cases to a new fact pattern.

Each prompt assumes the prior output as context, which produces more coherent teaching materials than a single sprawling prompt would.

13. Analogical Reasoning Prompting

Ask the LLM to explain complex legal concepts through analogies. Example: "Explain the difference between negligence and strict liability using three non-legal analogies that would resonate with 1Ls who have no tort law background."

It can be a powerful teaching technique for making abstract doctrine concrete and memorable.

14. Contextual Variation (or What If?) Prompting

Contextual Variation prompting tests how analysis changes when you modify one variable. Example:

You're a law professor. Here's a contracts fact pattern [type it in or upload it]. Show me how the analysis changes if: (a) we're in California instead of New York, (b) this

involves a sale of goods instead of services, (c) the plaintiff is a consumer instead of a merchant.

Again, providing knowledge (i.e., source material) will ground the output for improved accuracy.

It's great for creating exam variations and generating hypotheticals like those "Questions 10-14 are based on the following facts" exam questions.

15. Error Analysis Prompting

Error Analysis prompting helps identify common student mistakes. Example:

Here are X number anonymized student answers to an *Erie* doctrine question. Identify the most common conceptual errors and create a mini-lesson that addresses these specific misunderstandings.

It works well for targeted teaching and creating formative assessments that address actual student confusion.

Hint: Save all of the anonymized answer files in a folder, then use Claude Cowork (described in Chapter 3 and in Appendix C). It's much easier than having to upload 100 answers.

H. Common Mistakes and How to Avoid Them

Even with good technique, certain pitfalls appear again and again. Here's what to watch for and how to address it.

1. Treating the GenAI as an Authority

Novice GenAI users often accept what the LLM tells them without question because they lack discernment skills and don't have the necessary underlying subject matter knowledge. Unfortunately, we've seen this happen too often by attorneys, as in *Mata v. Avianca*.¹⁵ But GenAI tools don't know what's right. They

¹⁵ *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443, No. 22-cv-1461 (PKC) (S.D.N.Y. June 22, 2023).

generate plausible-sounding text based on patterns. Sometimes that text is accurate; sometimes it's confidently wrong.

Better approach: Ask "what are the arguments for and against this position?" or "What would a court need to determine to resolve this issue?" These types of questions frame the GenAI as a thinking partner, not an authority, and produces more useful responses.

2. Accepting the First Response

When not using the CORRECT or other prompting technique, first responses are often generic or incomplete, and possibly wrong. But many students take the first output as final, missing opportunities to dig deeper.

Better approach: Treat the first response as a starting point. Remember that you're engaging in a conversation, so iterating is part of the process. Follow up with "tell me more about X" or "how would this analysis change if [different facts]?" or "what's the strongest counterargument to this position?" Iteration is an important way to get quality output.

3. Vague Context

"Help me with this contracts problem" gives the GenAI almost nothing to work with. It will default to generic explanations that may or may not match your situation.

Better approach: Provide specific context. What jurisdiction? What substantive area? What are you trying to accomplish? What's giving you trouble? The more context you provide, the more tailored the response.

4. Ignoring Errors

GenAI makes mistakes. It fabricates citations and misapplies rules. Students who don't understand this will trust incorrect output.

Better approach: Build error-checking into your prompts. "Explain your reasoning and flag any areas where you're uncertain" or "cite the specific rule or case on which you're relying" forces the GenAI to be more transparent about its limitations. Teach students to verify key claims, especially case citations and rules.

5. Over-Relying on the GenAI

Some students use GenAI for everything, losing the productive struggle that builds understanding. They ask the GenAI to solve problems rather than using it to support their own problem-solving.

Better approach: Use GenAI strategically. Let it help with tasks where you're stuck or need efficiency, but do your own thinking first. Ask it to explain concepts after you've tried to understand them yourself. Have it critique your analysis rather than generating analysis for you.

6. Forgetting About Format

Different tasks need different formats. A case brief needs different structure than a client email needs which needs a different structure than an exam answer. If you don't specify format, the AI will default to generic paragraph explanations (sometimes with an over-abundance of emojis and bullet points).

Better approach: Be explicit about format in your prompt. "Provide your answer as an IRAC-structured analysis" or "generate a client email" or "write this as a formal memo with proper headings."

7. Prompting Like It's Google

New users often prompt GenAI tools the way they search Google with short keyword queries. "Business judgment rule" or "Erie doctrine" or "adverse possession." These prompts treat the GenAI as if it is a search engine when it's actually a text-generation engine.

Better approach: Prompt conversationally but specifically. "Explain the business judgment rule to me as if you're helping me prep for an exam" is far better than "business judgment rule." Full sentences that establish context and objective produce better results than keyword strings.

I. Putting It All Together

Let's see how these principles and techniques work in practice. Suppose you're a Torts professor and you want to use GenAI to help students understand proximate cause. Here's how you might evolve your prompting:

First attempt (basic): "Explain proximate cause."

This prompt will give you a generic textbook explanation. Not wrong, but not particularly useful either. It's the kind of explanation students could find anywhere.

Second attempt (adding context and objective): "Explain proximate cause to a law student who understands actual cause but is confused about when liability gets cut off due to superseding causes."

Much better. The prompt gives the GenAI context about the student's confusion and focuses the explanation on the distinction that matters.

Third attempt (adding role and format): "You are a patient Torts professor. Explain proximate cause to a law student who understands actual cause but is confused about when liability gets cut off due to superseding causes. Provide three short hypotheticals in bar-style format that clearly illustrate the doctrine: one where proximate cause exists, one where it doesn't, and one borderline case."

Now we're getting somewhere. The role sets the tone, and the format requirement (three hypotheticals with specific characteristics) ensures concrete examples rather than abstract explanation.

Fourth attempt (adding iteration): After receiving the response, you might follow up: "The borderline hypothetical is helpful, but make it even harder to analyze. Add facts that create genuine ambiguity about whether the intervening cause is superseding or foreseeable."

The output will likely meet your specific pedagogical needs. You're not settling for the first response; you're iterating toward exactly what you want.

Fifth attempt (adding Socratic element): "Now take that borderline hypothetical and generate five Socratic questions a professor might ask to help students think through the proximate cause analysis without giving them the answer."

The material is transformed from a passive explanation into an active teaching tool. Students struggle with the questions and develop their own analysis.

When not starting with CORRECT prompting, iterative progression illustrates a different approach to effective prompting. You start with basic requests and refine through iteration, context, role-specification, format

requirements, and advanced techniques until you get output that serves your specific pedagogical purpose.

J. Teaching Students to Prompt

If I was queen of the world, prompt engineering would be taught during the first semester of law school, right alongside legal research and writing and case briefing. It's a foundational skill, not an advanced topic. Just as we don't wait until upper-level seminars to teach the Bluebook system of citation, we shouldn't wait to teach effective prompting.

By the time students reach my AI and the Law or Law Practice Technology courses, I want them to already be competent prompters. That way we can focus on substantive questions, such as how GenAI affects legal practice, what regulatory frameworks make sense and privacy issues, rather than spending two classes on prompting basics.

The argument for early training is simple: *students are already using GenAI*, whether we've trained them or not. The baseline assessment I described at the beginning of this chapter demonstrates that. Some figured out effective prompting on their own, but most didn't. The gap between trained and untrained prompters will only widen as GenAI becomes more embedded in legal practice and education. Teaching prompting early and systematically levels that playing field before it affects and disadvantages student performance across multiple courses.

But, alas, I'm not queen of the world. The reality is that most law schools haven't yet integrated prompt engineering into their curriculum. If prompt engineering is important for faculty members, it's essential for students. But how do you teach prompting without turning your law class into a computer science seminar?

1. Integration, Not Addition

The answer is integration, not addition. You don't need separate lessons on prompting. Instead, you model effective prompting as part of teaching doctrine. Demonstrate it when working through problems, and require it when it makes sense to use GenAI for assignments.

2. Show Your Work

Start by making your own prompts visible. When you use GenAI in class to generate a hypothetical or explain a concept, show students the prompt that you used and explain why you structured it that way. Explain why you included context and what constraints you added. Make the invisible visible. In fact, the best approach might be assigning this chapter (and possibly others).

3. Provide Starter Prompts

Give students starter prompts that they can adapt. When you assign a problem set, include a sample prompt. Students can use it as is or modify it for their needs.

4. Require Attribution

Require prompt attribution on assignments. If students use GenAI, ask them to include the prompts that they used. This makes prompting part of the academic work, not something hidden or ashamed. It also lets you give feedback on prompt quality, just as you'd give feedback on legal analysis.

5. Gamify It

Create prompting challenges. Occasionally ask students to generate the best prompt they can for a particular task, then compare prompts among their peers and discuss what makes some more effective than others. Give a token prize to the best prompter, e.g., a free pass in the next class. You don't have to take class time. This work can be done in a discussion board. The exercise helps students think explicitly about prompting strategy.

The goal is to give students enough facility that prompting becomes a natural part of their legal toolkit, used strategically, ethically, efficiently, and effectively when it serves their learning.

K. The Pedagogical Advantage

I've seen how faculty members who master prompt engineering don't just use GenAI better. They think more clearly about their own teaching objectives. It works for me. I use it to make lectures more interactive and to generate hypotheticals that would have taken me far longer to create on my own. It also catches gaps in my assignments before students do. It helps me create examples and to create Powerpoints from my notes. It's made me a better teacher, not because of the technology, but because effective prompting requires clear thinking about pedagogy.

For example, I teach the different frameworks of GenAI regulation: self-regulation, applying or amending existing law, comprehensive new legislation, sector-specific regulation, and international coordination. GenAI helped me design a simulation where students role-play regulators from different jurisdictions responding to the same GenAI deployment scenario. I gave it the learning objectives and constraints (45-minute class, groups of 5, minimal prep required), and it generated a complete exercise with role descriptions for each regulatory approach and a realistic fact pattern involving a GenAI system that implicates multiple legal concerns. It also gave me draft discussion prompts that force students to articulate the tradeoffs inherent in each framework. The prompt took me five minutes to write. Creating that exercise from scratch would have taken me hours.

Here's the prompt I used.

Context: You're a law professor teaching an AI and the Law course covering five regulatory frameworks: self-regulation, applying existing laws, comprehensive new legislation, sector-specific regulation, and international coordination.

Objective: You need an in-class simulation that helps students understand the tradeoffs between these frameworks by having them apply each approach to the same scenario.

Role: Act as an instructional designer experienced in creating active learning exercises for law school courses.

Restrictions: The exercise must fit within 45 minutes, work with groups of 5 students, require minimal advance preparation, and involve a realistic AI deployment scenario (not science fiction). Avoid creating a scenario so complex that students get lost in technical details rather than focusing on regulatory approaches.

Expectations: Provide five role descriptions (one for each regulatory framework), a single fact pattern that all groups will analyze, and 3-4 discussion prompts that will guide the debrief conversation after groups present their approaches.

Clarify: Your task is to create a complete, ready-to-use classroom simulation that requires students to apply different AI regulatory frameworks to the same scenario.

Tone: Keep the materials professional and academically rigorous, but make the scenario engaging enough that students will enjoy the exercise.

AI REGULATORY FRAMEWORKS SIMULATION Time Required: 45 minutes (25 minutes group work, 20 minutes presentations and discussion)

SCENARIO: HealthPredict AI

MediCare Solutions has developed HealthPredict, an AI system that analyzes patient electronic health records to predict which patients are at high risk for hospital readmission within 30 days of discharge. The system:

Uses machine learning trained on 5 million patient records from 200 hospitals

Considers 150+ variables including diagnosis, demographics, prior hospitalizations, medication adherence, and social determinants of health (income, housing stability, transportation access)

Generates risk scores from 1-100 and automatically flags high-risk patients for case management intervention

Is currently deployed in 50 hospitals across 12 states

Has demonstrated 15% reduction in readmissions, saving hospitals millions in Medicare penalties

Recently faced criticism when investigative journalists discovered the algorithm assigns higher risk scores to patients in low-income zip codes, even controlling for health status

Hospitals rely on HealthPredict's recommendations to allocate limited case management resources. MediCare Solutions is now seeking to expand internationally.

GROUP ROLE ASSIGNMENTS

Group 1: Self-Regulation Advocates

You believe industry self-regulation is the optimal approach for AI governance. Your framework emphasizes:

Voluntary best practices developed by industry consortia

Internal ethics boards and compliance programs

Market incentives (reputation, liability) to encourage responsible development

Flexibility to adapt as technology evolves

Minimal government intervention

Your task: Explain how self-regulation would address HealthPredict. What voluntary standards would apply? What mechanisms would ensure compliance? What happens when self-regulation fails?

Group 2: Existing Law Application

You believe current legal frameworks are sufficient and there is no need for new AI-specific laws. Your approach relies on:

Anti-discrimination laws (Title VI, Section 1557 of ACA)

Medical device regulation (FDA)

Health privacy laws (HIPAA)

State consumer protection statutes

Common law (negligence, contract, product liability)

Your task: Identify which existing laws apply to HealthPredict and how they address the concerns raised. Where do existing laws provide adequate protection? Where do gaps exist?

Group 3: Comprehensive New Legislation

You believe AI requires sweeping new legislation that addresses the technology holistically. Your approach favors:

New federal AI regulatory framework applying across sectors

Ex ante requirements (testing, validation, transparency)

Mandatory impact assessments before deployment

Algorithmic accountability and explainability requirements

Individual rights (access to information, human review, redress)

Your task: Outline what a comprehensive federal AI law would require for systems like HealthPredict. What standards would HealthPredict need to meet before and after deployment? What enforcement mechanisms would you include?

Group 4: Sector-Specific Regulation

You believe AI regulation should be tailored to specific industries because healthcare AI raises different concerns than financial AI or hiring AI. Your approach emphasizes:

Healthcare-specific AI requirements

Building on existing sectoral regulators (FDA, CMS, OCR)

Domain expertise in rule-making

Standards tailored to clinical context and patient safety

Your task: Design sector-specific rules for healthcare AI systems like HealthPredict. What should CMS or FDA require? How would healthcare-specific regulation differ from general AI rules?

Group 5: International Coordination

You believe AI regulation requires international harmonization because AI systems cross borders. Your approach prioritizes:

International standards and agreements

Coordination between US, EU, and other jurisdictions

Interoperability of regulatory frameworks

Addressing global AI companies and data flows

Preventing regulatory arbitrage

Your task: Explain how international coordination would address HealthPredict's planned global expansion. What international frameworks or agreements would apply? How do you reconcile different national approaches?

DISCUSSION PROMPTS FOR DEBRIEF

Tradeoffs: Each framework makes different tradeoffs between innovation and protection, speed and thoroughness, flexibility and certainty. What does each approach optimize for? What does it sacrifice?

Bias problem: The journalists discovered HealthPredict assigns higher risk scores to low-income zip codes. Which regulatory framework is best positioned to detect and address this problem before deployment? After deployment? Would any framework have prevented this issue entirely?

Practical implementation: HealthPredict is already deployed in 50 hospitals and saving lives by reducing readmissions. Which framework best balances protecting patients from algorithmic bias with preserving the system's demonstrated clinical benefits? How quickly could each framework respond to the bias revelation?

Regulatory gaps: Where does each framework fail to address important concerns about HealthPredict? Are there issues that fall through the cracks regardless of which approach we adopt?

Crafting effective prompts requires articulating precisely what you want, what context matters, what constraints apply, and what format works. That's exactly the kind of pedagogical clarity we should always strive for, with or without GenAI. Good prompting forces good thinking about teaching goals.

When you can clearly specify what you want a GenAI to generate for student learning, you've also clarified your own learning objectives. When you can articulate what makes a response high-quality versus low-quality, you've sharpened your own evaluative criteria. The discipline of effective prompting can make us better teachers.

This chapter covers prompting frameworks, advanced techniques, applications, and common pitfalls. But the real learning happens through practice. Use the CORRECT framework on one of your essay exam questions. Apply advanced prompting techniques. Iterate. Even experienced prompters refine and adjust constantly. The tool is powerful, but it requires skill to use well.

That skill is learnable, and it is worth learning. In a legal education environment where GenAI is ubiquitous, faculty who can prompt effectively and teach it explicitly will give their students a lasting advantage. Not because they're better at technology, but because they're better at thinking precisely about what they want and how to communicate it clearly.

In the end, that's always been the heart of good legal teaching. GenAI just makes it more explicit and more essential than ever before.

CHAPTER 5: Guardrails and Gatekeeping: Governing AI in the Law Classroom

The panic in legal education occurred sometime between the release of GPT-4 and the first time that a student handed in a brief that looked a little *too* polished for a Tuesday morning. I'm sure many of you have felt that familiar tinge of anxiety. The goalposts of original work seemed to have shifted while you were busy doing research. Due to GenAI, we are recalibrating the very foundation of how we train future lawyers to think, research, and write. This chapter is designed to address initial concerns and move past the initial panic. It is intended to point toward a functional, ethical, and defensible strategy that balances our duty to the ABA's evolving standards for legal education with the practical reality of policy drafting, integrity enforcement, and the design of GenAI-resistant, GenAI-transparent, and GenAI-prohibited assignments that still demand the rigorous legal analysis that we teach.

A. The "C" Word: Navigating Cheating in the Age of LLMs

You've had that moment: you're grading a 1L memo or a seminar paper, and the prose turns eerily rhythmic. The citations are formatted perfectly but point to cases that don't exist, the quality of the work doesn't match the student's performance in class or on the midterm, and the tone is... aloof. There may be some telltale signs you've heard about, like use of the words "delve," "crucial," "foster," and "landscape." It's unfortunate. I love the word "delve." I've stopped using it entirely because I don't want regular GenAI users to think I'm passing off GenAI as my own work. Your gut says GenAI, but your syllabus says "Original Work Only."

My students report that they self-censor their best work. They dumb it down and take out "big" words to avoid triggering exactly the suspicion you're sitting with right now. The accusation of academic dishonesty, even unfounded, carries enough risk that students would rather turn in mediocre work than be

called in to defend a genuinely good sentence.¹ Because of GenAI, we've unfortunately built a process where excellence looks suspicious.

Now what?

When faculty think about cheating with GenAI, the concern is typically around students submitting GenAI-generated output as their own. That is a challenge. It's not unlike admissions deans and directors who worry about this problem even before a student matriculates. I've been asked by deans of admissions what should be done about the personal statement: "how can we prevent prospects from using GenAI?" And they are using it. One member of the admissions committee said that either many of the prospects are using GenAI or there is an inordinate number of GenZers who have struggled with weight loss as their personal challenge to overcome. In the age of GLP-1 drugs, this particular problem might correct itself.

My perspective on the issue is that, first, how could a prohibition policy be enforced? Are you going to run every submission through a detector (the many, many, many drawbacks of this option are discussed below).

Second, do we want to? Some applicants are fortunate to have lawyers in the family – people who can help them write their personal statement. Or they have the money to pay for admissions counseling services. At our law school, many, but not all, of the prospects are first-generation law students and many are first-generation college students. Isn't the use of GenAI leveling the playing field for those without that advantage? While GenAI, of course, can be used unethically, so can relying on a lawyer/parent. In fact, one friend wrote and submitted her son's personal statement.² Unethical? Yes. But, still, I'm sure she's not the only one whose anxiety and love led them to overstep, though perhaps not as far as she did.

And there's another problem. It's the same one that the courts are facing and sometimes dealing with through standard orders, local rules, and guidelines.³

¹ See *Matter of Newby v. Adelphi Univ.*, 2026 NY Slip Op 26021, Jan. 8, 2026 (holding that AI detection score alone, without more, does not constitute rational basis for academic dishonesty finding where student presented contradicting detection results and documented tutor assistance).

² Fortunately, the son decided not to go to law school.

³ Courts also struggle with how to define AI use, whether to restrict or prohibit use, whether to require disclosure and when, etc. See <https://airtable.com/appKUCriCQDI1BxIV/shrflAPpNKaNMnacR/tblNmp6mff8CzLuQD> (last accessed Jan. 14, 2026).

How is the school going to define “GenAI use”? Yes, submitting the output of GenAI without editing is over-the-line. But using it to correct grammar and spelling? Is that too much? GenAI is now embedded in Word, Grammarly, Google Docs, etc. Is spellcheck and grammar check an unethical use of AI? Should applicants have to prove somehow that they’ve turned these features off?

It is interesting to note that the University of Miami School of Law application contains an *optional* section under “Significant Writing.” It asks the applicant to provide a “detailed prompt for a Large Language Model to generate a comprehensive analysis that will help you make an informed decision about which law school to attend. Your prompt should be specific enough to provide actionable insights while thoughtfully considering the most important factors in this decision.” It also asks the applicant to “provide 3-5 follow-up prompts that anticipate AI responses and allow you to explore your options on a deeper level. These follow-up prompts should provide you with more detailed and specific information about each of your choices. Consider the order of the prompts as you feed them into the generative AI. (maximum characters 2500).”⁴

Dealing with GenAI-enabled cheating shouldn’t be focused on catching cheaters. Instead, law schools should look to fulfill our role as the gatekeepers of the profession. If students outsource their legal reasoning to a chatbot today, they are one step away from a Rule 1.1 competence disaster or a Rule 3.3 candor-to-the-court nightmare tomorrow.

1. Connecting the Dots: The Model Rules

When we talk to students about cheating, we shouldn't just talk about the Student Handbook or the Honor Code. We need to talk about the ABA Model Rules of Professional Conduct.

a. Rule 1.1: Competence

Rule 1.1 requires lawyers to provide "competent representation" to clients, which means "the legal knowledge, skill, thoroughness and preparation reasonably necessary for the representation."⁵

⁴ Application to University of Miami School of Law on file with author, and provided by Katherine Powley, Assistant Dean of Admissions and Scholarships, UNT Dallas College of Law.

⁵ MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2023).

When students ask ChatGPT to draft their legal memoranda and submit the result without thorough verification, they're not only acting unethically but they're not developing the knowledge and skills necessary to be competent. Even worse, they're practicing incompetence. The student who hasn't learned to construct legal arguments or synthesize authority isn't going to magically develop those skills upon graduation.

I tell my students (and the 1Ls at orientation) that every assignment in law school is practice for the work they'll do for real clients. And right now, you don't know what you don't know; you're still learning the law. You won't recognize an incomplete answer or a hallucinated citation. If you can't write a solid essay without the help of an LLM, then you won't do well on an exam. If you can't write a persuasive memo without GenAI assistance now, you won't be able to do it competently for a client later.

b. Rule 3.3: Candor to the Tribunal

Rule 3.3 prohibits lawyers from making false statements of fact or law to a tribunal and requires lawyers to correct material false statements.⁶ This is where GenAI hallucinations become an ethical emergency.

Point students to *Mata v. Avianca, Inc.*, where attorneys submitted a brief citing non-existent cases generated by ChatGPT. The court sanctioned the lawyers \$5,000 each, not just for the fake citations, but for the attempted cover-up when they submitted false and misleading affidavits about their use of AI.⁷

This case is far from the only one where lawyers misunderstood how to use GenAI. As of this writing, there are 887 other cases—in the United States alone—where pro se litigants or attorneys submitted fabricated case law, false quotations, or misrepresentations.⁸

These are real lawyers who faced real consequences for doing exactly what some students do on their seminar papers. They're rubber-stamping GenAI output and submitting GenAI content without verification.

⁶ MODEL RULES OF PRO. CONDUCT r. 3.3 (A.B.A. 2023).

⁷ *Mata v. Avianca, Inc.*, 1:22-cv-01461, (S.D.N.Y.).

⁸ Damien Charlotin, *GenAI Hallucination Cases*, https://www.damiencharlotin.com/hallucinations/?q=&sort_by=-date&states=USA&period_idx=0, (last accessed Apr. 15, 2025). Note that the number rose from 666 on Dec. 30, 2025 to 887 on Apr. 15, 2026).

Let the students know that he who submits a paper with hallucinated citations today is the lawyer who files a brief with hallucinated citations tomorrow. The only difference is that tomorrow, he'll lose his license or pay significant sanctions and suffer reputational loss instead of failing an assignment or facing an honor code violation. And he'll deserve it, because he had three years to practice verifying GenAI output and chose not to.

c. Rule 5.3: Supervision of Non-Lawyer Assistants

In practice, lawyers are responsible for the work of paralegals, law clerks, and other non-lawyer assistance. Rule 5.3 requires lawyers to ensure that use of this assistance is compatible with the lawyer's professional obligations.⁹

This is the perfect framework for GenAI use. Tell your students that GenAI like a very fast, very confident, but occasionally hallucinating paralegal who misses things from time-to-time. If attorneys don't supervise (verify) its work, the ethical failure is theirs, not the machine's.

This means:

- You must review every citation and read the case to ensure that the case exists and says what the GenAI claims it says
- You must verify that every quotation ostensibly within the case exists
- You must verify that the legal reasoning is sound and applicable to your specific facts
- You must ensure that the analysis reflects your professional judgment, not just the AI's statistical predictions
- You must be able to explain and defend every argument in the work product

The students who can't explain why they structured their memos a particular way or why they prioritized certain arguments have failed to supervise the "assistance" received. They've also demonstrated that they don't understand their own work product, which is precisely what Rule 5.3 is designed to prevent. See Chapter 5 for the "Hallucination Hunt," an assignment reinforcing this duty.

⁹ MODEL RULES OF PRO. CONDUCT r. 5.3 (A.B.A. 2023).

d. Rule 8.4(c): Dishonesty, Fraud, Deceit, or Misrepresentation

Rule 8.4(c) prohibits conduct involving "dishonesty, fraud, deceit or misrepresentation."¹⁰ Passing off an LLM's output as one's own original legal analysis violates the spirit of this rule.

When students represent that a paper is their own work but instead submitted AI-generated content, they're engaging in the same type of misrepresentation that violates Rule 8.4(c) in practice.

ABA Model Rule 8.4(g) prohibits conduct a lawyer knows or reasonably should know constitutes harassment or discrimination in practice-related activities, and GenAI's well-documented tendency to reproduce and amplify training-data bias puts that rule directly in play. When a lawyer or law student accepts GenAI output uncritically for client communications or case evaluations, any bias embedded in that output flows directly into their professional work. "The AI didn't warn me" won't satisfy a disciplinary panel, particularly when the risk of algorithmic bias is well-documented.

More importantly, this rule extends beyond conduct related to law practice. It applies to any dishonest conduct that reflects adversely on a lawyer's fitness to practice. Character and fitness committees take academic dishonesty seriously precisely because it demonstrates a willingness to engage in misrepresentation for personal gain.

B. GenAI Policies

You need a GenAI policy for your syllabus. The question is not whether to address GenAI use but how to frame the conversation with students. Your policy should match your pedagogical goals and your comfort level with GenAI integration.

Law schools have adopted wildly different approaches.¹¹ Some ban GenAI entirely, treating it like academic dishonesty. Others embrace it enthusiastically, requiring students to use GenAI tools. Most fall somewhere in between, creating assignment-specific rules or disclosure requirements. The variation can be challenging for students to track, but my standard response to students is, "read

¹⁰ MODEL RULES OF PRO. CONDUCT r. 8.4 (A.B.A. 2023).

¹¹ See *Clinic, Law School, and University AI Policies and Syllabus Language*, STAN. L., <https://law.stanford.edu/juelsgaard-intellectual-property-and-innovation-clinic/clinic-and-law-school-policies-and-syllabus-language/>, (last accessed Dec. 31, 2025).

the syllabus. Just like you'll be responsible for reading and complying with standing orders and local rules."

AI can help you draft policy language across this spectrum wherever you fall.

1. The GenAI-Prohibited Approach

If you believe your course requires students to develop skills without GenAI assistance, a clear prohibition may be appropriate. Keep in mind that it might be difficult to enforce. This may work well for foundational courses in which struggling with basic concepts builds essential skills. On the other hand, not using GenAI might do those struggling students a disservice. Students with learning disabilities can be significantly helped by using GenAI. Also, studies show that instructor-led use of GenAI can improve understanding for average and low performing students (discussed in depth in Chapter 6). And remember, Prohibition has been tried before. It didn't work out so well in the 1920s. (If you don't know what I'm talking about, Google it.)

a. Prompt for Prohibition Language

You are a law professor who teaches Civil Procedure. Draft GenAI policy language for the syllabus that prohibits GenAI use for all assignments. The policy should:

- Clearly define what counts as GenAI (GenAI tools like ChatGPT, Claude, legal research AI)
- Explain the pedagogical reason (building foundational skills)
- Specify consequences for violations
- Maintain a professional, non-punitive tone

Keep it under 150 words.

b. Generated Policy

Artificial Intelligence Policy

This course prohibits the use of GenAI tools (including ChatGPT, Claude, and AI-powered legal research platforms) for all assignments. Civil Procedure requires you to develop foundational skills in legal analysis, procedural strategy, and written advocacy. These skills develop through

the cognitive struggle of working through complex problems independently.

Using GenAI tools for assignments violates academic integrity policies and will be treated as unauthorized assistance. Violations may result in assignment failure or referral to the Honor Code Committee.

This prohibition does not apply to GenAI tools embedded in word processors (spell check, grammar suggestions) or standard legal research platforms (Westlaw, Lexis) that return cases rather than generating analysis.

Questions about whether specific tools are permitted? Ask before using them.

This language is clear and justifies the prohibition pedagogically rather than punitively. It helps students understand why you've made this choice.

Note: Every GenAI response is different by design, so your results from the same prompt may not match mine. When they don't, keep at it: rephrasing and pushing back is part of the process, not a sign something went wrong. It's a conversation, not a vending machine. The more you use your preferred platform, the better your starting point tends to be. I use Claude daily for nearly everything, and over time it has developed a real sense of what I expect and what I consider a complete answer. Your platform will do the same for you. Think of it like a new assistant: the first few weeks involve more correction than you'd like, but eventually you spend less time explaining and more time working with what you get.

2. The GenAI-Resistant Approach

Many faculty find that different assignments require different GenAI policies. Research memos might permit GenAI for outlining but prohibit it for analysis. Drafting exercises might encourage GenAI for initial drafts but require disclosure.

a. Prompt for GenAI-Resistant Approach

You are an Advanced Legal Writing professor. Draft GenAI policy language for the course with assignment-specific rules:

- Drafting exercises: GenAI permitted with disclosure
- Research memos: GenAI prohibited for legal analysis, permitted for outlining and citation checking
- Client letters: GenAI permitted for tone and clarity review

For each assignment type, the student must submit the chat and a reflection on how GenAI was used.

Include a clear table or format that students can reference easily.

b. Generated Policy

Generative AI Policy — Advanced Legal Writing

Legal writing is a lawyering skill, not a formatting exercise. This course distinguishes between using AI to support your thinking and using it to replace it. The line between those two things matters, and you are expected to know where it falls.

For every assignment where AI use is permitted, you must submit your complete chat history and a reflection (150–200 words) identifying what you asked the AI to do, what it produced, how you evaluated it, and what you changed. A reflection that says "I used ChatGPT to check my citations and it was helpful" is not a reflection. Tell me what you learned.

Assignment	AI Permitted	AI Prohibited	Additional Requirements	Submission
Drafting Exercises	All uses, including drafting, revision, and structure	Nothing prohibited, but uncritical submission is not acceptable	Chat history and reflection	
Research Memos	Outlining; citation format checking	Legal Analysis, case synthesis, statutory interpretation, argument development	Chat history and reflection identifying every GenAI-assisted element	

Client Letters	Tone review; plan-language clarity edits	Substantive legal content; factual representations	Chat history and reflection

You remain responsible for the accuracy of every sentence you submit, regardless of who — or what — drafted it first. GenAI does not have a law license. You (almost) do. Act accordingly.

3. The GenAI-Transparent Approach

For upper-level courses, simulation-based work, or transactional drafting, you might encourage GenAI use as a professional skill. Modern lawyers use GenAI tools regularly. Teaching students to use them responsibly prepares them for practice.

a. Prompt A – GenAI-Resistant Approach

You are a law professor who teaches Transactional Drafting. Draft a GenAI policy language for the course that permits and encourages GenAI use. The policy should:

- Frame GenAI as a professional tool students should learn to use
- Require students to disclose which tools they used
- Require students to turn in their chat history
- Require students to include a reflection about why they chose the platform and how the GenAI helped them (or didn't)
- Emphasize that AI-generated content must be reviewed, edited, and verified
- Note that students remain responsible for accuracy

Keep it under 200 words.

b. Prompt A – Generated Policy

Generative AI Policy — Transactional Drafting

Transactional lawyers use the best available tools. Generative AI is one of them, and you are expected to learn to use it well.

You may use any GenAI platform on all assignments. Competent use is required; uncritical use is not acceptable. You must ensure that confidential information is not given to the GenAI platform. GenAI drafts contracts that sound authoritative, but which can contain errors your client will eventually pay for. Your job is to catch those errors before they do. Verify everything in your submission before submitting your assignment.

Every submission must include three things:

1. A disclosure identifying which tool(s) you used, the complete chat history from your session, and a short reflection (150–200 words) addressing why you chose that platform, what the AI got right, and where you had to intervene.
2. You are responsible for every word you turn in. Remember, you cannot point the finger at the LLM. Mistakes are your fault and will be reflected in your grade.
3. Think of GenAI as a first-year associate who works fast, never sleeps, and has never read a deal. Supervise accordingly.

c. Prompt B- GenAI-Resistant Approach

Some faculty permit GenAI use broadly but require detailed disclosure about how students used it. This approach treats GenAI like other research tools while maintaining transparency.

Prompt for Disclosure Language

You are a law professor who teaches Evidence. Draft GenAI policy language for the course

for all assignments but require detailed disclosure. The policy should:

- Permit any GenAI tool use
- Require students to submit a disclosure statement explaining what they used GenAI for and how
- Emphasize that disclosure itself doesn't guarantee full credit
- Note that inappropriate reliance on GenAI may result in lower grades

Keep it under 200 words.

d. Prompt B - Generated Policy

Generative AI Policy – Evidence

Use whatever tools you have access to keeping in mind that paid subscriptions coupled with CORRECT prompting and RAG provide the best results. You may use ChatGPT, Claude, Gemini, Copilot, NotebookLM — any of them, all of them, on any assignment. There are no restricted categories in this course.

There is, however, a non-negotiable requirement: you must provide a disclosure statement submitted with every assignment. The disclosure must explain specifically what you asked the AI to do, what it gave you, and how you used, modified, or rejected its output. Write it the way you'd explain your process to a supervising attorney who will ask hard follow-up questions. For example: "I used Claude to generate an initial outline for my hearsay analysis. I used ChatGPT to suggest examples of prior inconsistent statements. I reviewed both outputs, modified the structure significantly, and wrote all analysis independently. I verified all case citations in Westlaw. Attached is my chat history."

Disclosure earns you transparency credit. It does not earn you assignment credit. Work that reflects superficial engagement with the material — regardless of how thoroughly you've disclosed the GenAI's role in producing it — will be graded accordingly. Evidence is a doctrinal course. You need to understand why hearsay has exceptions, not just be able to generate a paragraph that sounds like you do. Students who demonstrate genuine analysis get full credit. Students who demonstrate that they asked an LLM to demonstrate authentic analysis on their behalf do not.

Failing to disclose GenAI use is an academic integrity violation.

4. The GenAI-Integrated Approach

a. Prompt for GenAI-Integrated Approach

You teach Criminal Procedure at a law school. Draft a GenAI policy for your course that treats AI-generated content as an object of critical analysis rather than a prohibited or disclosed tool. The policy should make clear that students are expected to use GenAI output as source material they must evaluate, correct, and improve to demonstrate mastery by identifying where the GenAI fails. Students should understand that agreeing with a GenAI error is itself a graded outcome. The policy should be under 200 words, written in plain language, and have some firmness to it.

b. Generated Policy

Generative AI Policy — Criminal Procedure

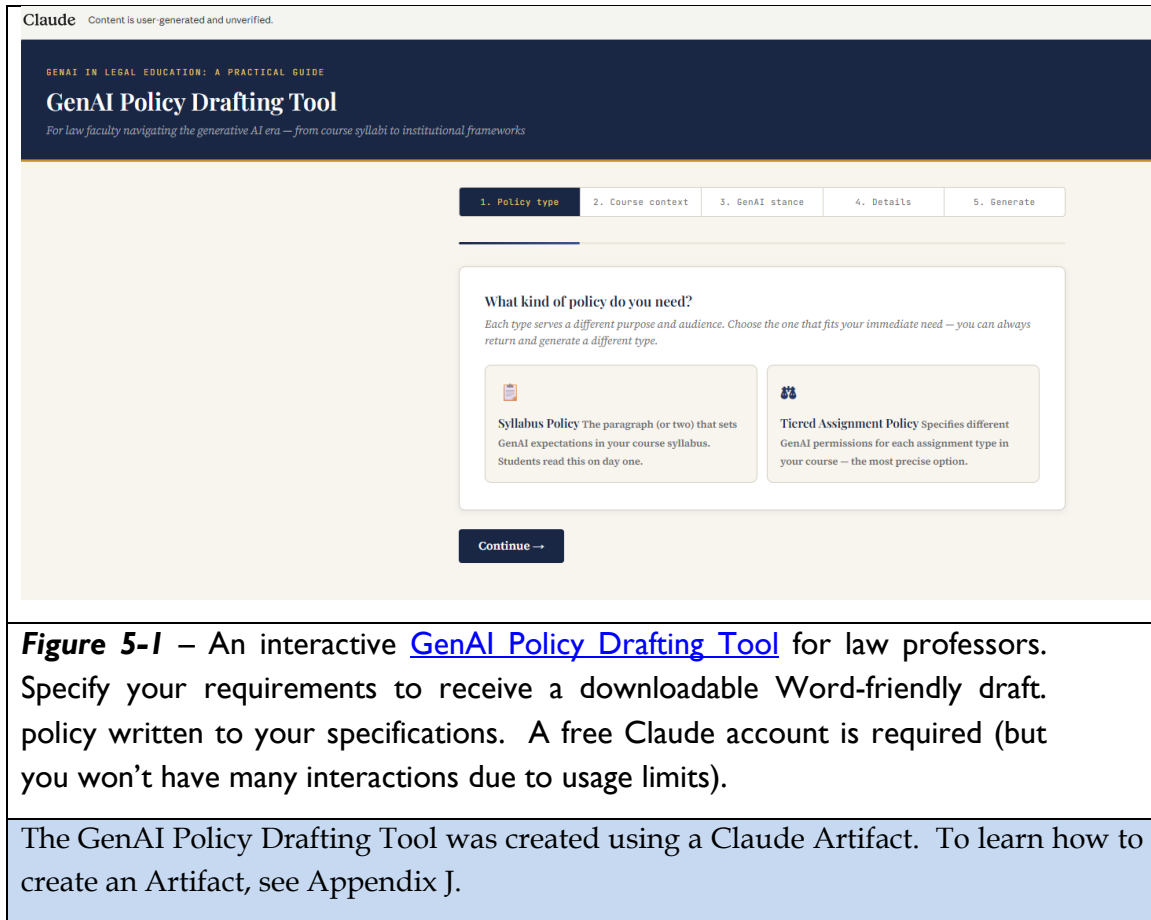
In this course, AI-generated legal analysis is assigned material, not a shortcut. You will receive AI-drafted briefs, motions, and doctrinal explanations. Your job is to audit them for accuracy and completeness.

Every AI-generated assignment requires a written critique identifying errors in doctrine, standard, case application, and legal reasoning. For each error, you must state the correct rule, cite the controlling authority, and explain why the mistake matters in practice. Finding all the errors is the assignment.

This structure is deliberate. Criminal Procedure doctrine is specific and heavily litigated. It is also frequently misrepresented by GenAI tools that flatten distinctions courts spent decades drawing. A student who cannot identify when GenAI has confused reasonable suspicion with probable cause, or misread *Rodriguez*, or cited *Bostick* for a proposition it doesn't support, is not ready to litigate suppression issues.

You may use any GenAI tool to help you prepare your critique. The irony of using GenAI to audit GenAI is not lost on me, and is, in fact, the point.

Demonstrating mastery means knowing when the machine is wrong. That skill is not optional in this profession.



5. The Tone Check

Policy language often sounds more punitive than you intend. Words like "violation," "prohibited," and "consequences" trigger defensive responses from students. The interactive tool above can help you draft policies to encourage responsible use rather than fear-based compliance.

This matters particularly for GenAI policies, which are new to most students. Many students arrive believing that any GenAI use constitutes cheating. Others think any use is okay. Your policy language shapes whether they approach GenAI as a forbidden shortcut that they must conceal or a professional tool they should learn to use thoughtfully.

C. Enforcing Policies: Why Detection Fails and What to Do Instead

Law professors are trained to seek "the truth" through objective evidence. When GenAI (GenAI) entered the classroom, the market responded with a wave of GenAI detection tools that promised to identify non-human writing with high degrees of certainty. However, these tools are fundamentally flawed for the legal academy. I am often asked to run a student's submission through my detectors. I do it. But I first provide a missive about why the detection tools shouldn't be relied upon. And it's not just because the detection tools cannot be trusted.

Beyond whether they work or not, it is perfectly reasonable to assume that a student has been using GenAI, but not in the way the professor thinks. Maybe instead of having the platform write the memo or motion or essay, between the midterm and the final, the student used GenAI to practice and improve by writing practice memos and motions and essays. And that is what really accounts for the improved performance you're seeing. Yes, the student used GenAI, but to learn, not to cheat.

To my mind, relying on detection programs for disciplinary action creates significant pedagogical and professional risks.

And it's not just me. At the time of this writing, thirty-one U.S. universities and colleges have banned the use of detection tools (primarily Turnitin). Another five recommend that teachers not use it.¹² The reason? The tools are absurdly unreliable. When copywriter and SEO specialist Dianna Mason ran the Declaration of Independence through a GenAI detector, it flagged as 98.51% AI-generated.¹³ The document was written in 1776, a mere 250 years before ChatGPT existed. She then tested the complaint filed in *Liebeck v. McDonald's* (the famous hot coffee case from 1993). ZeroGPT reported it was 100% AI-generated.¹⁴ As Mason concluded bluntly, "AI content detectors don't work. You shouldn't be

¹² *Schools that Banned GenAI Detectors*, PLEASE, Sept. 20, 2024, <https://www.pleasedu.org/resources/schools-that-banned-ai-detectors>, (last accessed Dec. 31, 2025).

¹³ Jodie Cook, GenAI Content Detectors Don't Work (The Biggest Mistakes They Have Made), *Forbes*, July 4, 2024, 8:00 am, <https://www.forbes.com/sites/jodiecook/2024/07/04/ai-content-detectors-dont-work-the-biggest-mistakes-they-have-made/>, (last accessed Dec. 31, 2025), (One commenter joked that they ran the Bible through a detector and it came back 98.9% AI-generated.)

¹⁴ *Id.*

using GenAI detector tools at all to identify whether GenAI was used to create content."¹⁵

1. Recent Cases

a. Doe vs. Yale University (Rignol v. Yale)

In February 2025, Thierry Rignol, a French national enrolled in Yale School of Management's Executive MBA program, filed suit against Yale after being suspended for a year and given a failing grade following an AI cheating accusation.¹⁶ A teaching assistant flagged his final exam as suspiciously "long and elaborate" with "near perfect punctuation and grammar." The professor then ran portions of the exam through GPTZero, which flagged the responses as likely AI-generated. What followed was a disciplinary process Rignol describes as deeply flawed. He alleges he was pressured to confess and denied access to key evidence. He also claims that he was told by a dean he could face deportation. The Honor Committee ultimately suspended him not for using GenAI, but for being insufficiently forthcoming during the investigation.

Rignol's complaint raises claims of breach of contract, Title VI discrimination based on national origin and non-native English speaker status, as well as both intentional and negligent infliction of emotional distress. The case is particularly notable because Rignol demonstrated that GPTZero flagged academic papers written by former Yale president Peter Salovey and other Yale scholars at 100% probability of GenAI authorship, and because Yale's own policies reportedly caution against relying on tools like GPTZero due to their high false-positive rates for non-native speakers. The federal court denied Rignol's request for an injunction to graduate with his cohort and denied his request to proceed anonymously. As of February 21, 2026, the case remains pending.

b. Orion Newby v. Adelphi University

In February 2026, Orion Newby won his suit against Adelphi University for false accusations of GenAI plagiarism.¹⁷ He didn't use GenAI, he was helped by

¹⁵ *Id.*

¹⁶ Rignol v. Yale Univ., No. 3:25-cv-00159-SFR (D. Conn. filed Feb. 3, 2025); see also Crowell & Moring LLP, *Ivy League Lawsuit Centers on Alleged Impermissible Use of AI in Academia*, Crowell & Moring Client Alert (Mar. 12, 2025), <https://www.crowell.com/en/insights/client-alerts/ivy-league-lawsuit-centers-on-alleged-impermissible-use-of-ai-in-academia>.

¹⁷ Matter of Newby, *supra* note 1; see also Jennifer McLogan, *Adelphi Student Orion Newby Sues Over AI Plagiarism Accusation and Wins. Why It's Being Called a "Groundbreaking" Case*, CBS News (Feb.

tutors in the Adelphi Bridges Program. That program was established to help students with learning and neurological disorders like him.

Why is this a groundbreaking case?

First, it establishes that GenAI detection tools cannot serve as sole evidence of academic dishonesty. The tools are unreliable. As discussed more fully below, they produce false positives, and that at least this court says that institutional reliance on them, without more, doesn't survive scrutiny. That's a direct constraint on how schools should conduct these proceedings.

Second, it imports due process requirements into academic discipline around GenAI accusations. Universities have treated GenAI detection flags as largely self-executing: flag, charge, sanction. A court reversing that outcome and ordering expungement signals that students have procedural rights before serious consequences attach. "A group made a decision arbitrarily and unilaterally" is not language that survives judicial review.¹⁸

Third, the Newby facts are particularly damaging to the detection-tool model. The student had documented tutor assistance through an official university program and was a student with learning and neurological disorders. Yet, he still received a zero and a fraud finding. Law schools should at a minimum provide students with due process before doling out punishment.¹⁹

The problem gets worse. These tools can be easily fooled. You can literally prompt ChatGPT to "write your response so it can't be detected," and the GenAI knows how to evade detection software. There are entire businesses built around this. Humanize AI is just one of many tools that a tool that makes GenAI output undetectable by "transform[ing] AI-generated text into natural, human-like text."²⁰ If the GenAI itself knows how to avoid detection and actively helps users evade it, what exactly are we detecting?

One of the most heartbreaking scenarios in academic integrity cases is when a student who worked hard to improve gets accused because the student's improvement looks "suspicious." GenAI-supported practice and learning is, I

9, 2026), <https://www.cbsnews.com/newyork/news/orion-newby-adelphi-university-ai-plagiarism-accusations/> (last visited Feb. 21, 2026).

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ Humanize AI, <https://www.humanizeai.pro/> (last visited Apr. 17, 2026).

think, what we should encourage students to do. It's the modern equivalent of going to the writing center (when it's open), working with a tutor (and paying fees), or doing extra practice problems (but without instant feedback). The student who uses GenAI to practice is developing competence and demonstrating effort, for a minimal fee, at the time of his or her choosing, and as often as the student wants to work with a private coach.

With that said, there are many tools claiming to be able to detect GenAI generated content. Most are unreliable at best and discriminatory at worst. Tools like GPTZero, Turnitin, and Copyleaks promise to identify AI-generated text, but they produce false positives at alarming rates.

Even if detection tools were 100% accurate (they're not and never will be), they still can't solve our fundamental problem. They can't tell you whether a student submitted AI-generated work or used GenAI to practice and develop skills. They can't distinguish between cheating and learning. They turn dramatic improvement into grounds for suspicion rather than celebration. They create an adversarial relationship between faculty and students instead of a collaborative learning environment. And perhaps most important, they don't help us achieve our actual goal which is to produce competent lawyers who can think critically and exercise professional judgment. Detection focuses our energy on catching bad actors instead of cultivating good practice.

Researchers have described a fundamental tension in using GenAI detection tools in educational settings. They distinguish between "the former" (legitimate, beneficial GenAI uses like brainstorming ideas and improving clarity or grammar) and "the latter" (problematic GenAI uses where students deceptively substitute GenAI work for required human effort).²¹ Their concern is that detection tools risk creating a chilling effect. If students fear being falsely accused of cheating, they may avoid even appropriate GenAI assistance like grammar checkers or idea generation, which would be educationally counterproductive.²² And, as previously mentioned, students have told me this directly and that they intentionally make their work weaker for fear of being accused.

²¹ Brian Jabarian & Alex Imas, *Artificial Writing and Automated Detection*, Nat'l. B. of Econ. Research, Working Paper 34223, <http://www.nber.org/papers/w3422>, (last accessed Dec. 30, 2025).

²² *Id.*, 18.

Institutions must calibrate detection policies to permit GenAI as a learning scaffold for brainstorming, editing, and tutoring while preventing GenAI from serving as a complete substitute for student thinking and writing. Applying the “4 Ds” of AI Fluency (Delegation, Description, Discernment, and Diligence) helps here (see Chapter 3).

The challenge is that current detection tools can't reliably distinguish between these different uses. They simply flag "AI involvement" without understanding whether students used GenAI to help develop their own ideas versus having GenAI do all the thinking. This is precisely the challenge you face with students using GenAI tools to complete entire courses, since the technology can't yet differentiate between a student who brainstormed with ChatGPT and one who had it write everything.

Researchers warn that GenAI detection isn't a "set it and forget it" solution. Instead, they expect an ongoing arms race between detection tools, increasingly sophisticated LLMs, "humanizer" software designed to evade detection, and students strategically adapting how they use GenAI to avoid getting caught.²³ Because detector performance will constantly shift as this cat-and-mouse game evolves, the authors recommend that detection platforms undergo routine public audits similar to stress tests in banking. The audits should be transparent, with results published openly so institutions can adjust their detection thresholds and policies before widespread evasion becomes a problem. In other words, whatever works today probably won't work as well tomorrow, so continuous monitoring and recalibration is essential rather than optional.²⁴

Again, even if detection tools worked perfectly, they can only say that "AI was probably involved." They can't tell you HOW it was involved. A student who used ChatGPT to generate the final memo and a student who used ChatGPT to practice writing ten memos before writing their own might show similar "AI signatures" in their work. Why? Because the practice student internalized the feedback and developed skills through GenAI-supported deliberate practice.

²³ *Id.*

²⁴ *Id.*, 4.

1. Pangram

Research shows that Pangram, a relatively new GenAI detection tool, dramatically outperforms other competitors, and it's cheaper than other detection tools.²⁵ However, the researchers specifically note, "[w]hile we have thus far focused on auditing existing GenAI text detectors, it is important to note that the implications of GenAI detection for writing and text-based work more generally are not obvious."²⁶ And it explicitly leaves implementation questions unanswered: "We believe that a more careful treatment of these issues requires a formal game theoretic treatment that we leave for future work."²⁷ In other words, we'd need to build a mathematical model of that cat-and-mouse dynamic before we can say anything rigorous about whether detection works as a policy tool. Or, more concisely, they know it's complicated, they know why it's complicated, and they punted.

Since not all law schools have access to Pangram, and since, as the researchers warn, LLMs could surpass its detection capabilities fairly soon, given how quickly GenAI is evolving, let's look at other limitations of detection.

2. The Technical Mirage of Detection

Let's take a look at Turnitin, which is frequently used for detection because law schools already have a license to use it for plagiarism checking.²⁸ GenAI detectors do not actually "detect" AI. They do not look for digital watermarks or specific metadata. Instead, they use secondary GenAI models to predict the

²⁵ *Id.*

²⁶ *Id.*, 18.

²⁷ *Id.*, 18-19.

²⁸ Faculty members often think of Turnitin as a "plagiarism" detector. It isn't. It is a "similarity" detector. It compares submitted text against a massive database of previously submitted student papers, published articles, websites, and other indexed content. When it finds matching or near-matching strings of text, it flags them and produces a similarity score — a percentage representing how much of the submission overlaps with existing sources. That's it. That's the product. What it does not do is determine plagiarism. After all, plagiarism is a legal and ethical conclusion that requires human judgment. It would flag a memorandum of law written by a student because they properly quoted a heavily-cited statute or used a common legal phrase 12 times. And even though Turnitin's website calls it "comprehensive plagiarism detection," Turnitin's documentation says this explicitly and it has for years. The score is a signal, not a finding. See *Your Turnitin Similarity Questions Answered*, <https://www.turnitin.com/products/similarity/> (last visited Feb. 21, 2026).

likelihood that a string of text was generated by a Large Language Model (LLM). This process relies on two primary metrics: perplexity and burstiness.²⁹

Perplexity measures the randomness of the text. If a detector can easily predict the next word in a sentence, the perplexity is low. Low perplexity is often flagged as AI-generated because LLMs are designed to choose the most statistically probable next token (see Chapter 2 to learn more about how LLMs work). Burstiness refers to the variation in sentence length and structure. Humans tend to write with "bursts" of short and long sentences. GenAI models typically produce a more uniform, rhythmic cadence.³⁰

The problem for legal educators is that "good" legal writing often mimics these GenAI traits. Students are taught to write with precision. They are encouraged to use standardized phrases and predictable structures. When a student successfully adopts the plain language style required for a trial brief or a memo assignment, a detector is more likely to flag that student for a false positive.³¹

What the detection companies rarely advertise is that LLMs themselves incorporate intentional randomness into their word selection. If they did not, every answer to identical prompts would be the same if the response began with the same word. This means the very foundation of detection—identifying "predictable" next-word choices—is built on a false premise.³²

3. The Evolving Failure Rates

The reliability claims of GenAI detectors have deteriorated significantly since their initial release. When Turnitin first launched its GenAI detection feature in April 2023, the company claimed a false positive rate of less than 1%. Within months, that claim was revised to "less than 1% for documents with 20% or more GenAI writing," a qualification that fundamentally changes the meaning.³³ The

²⁹ See Salar J. Ghahramani, *The Automation of Academic Integrity: Challenges and Considerations for GenAI Detection*, 12 J. ACAD. ETHICS 45, 48 (2024).

³⁰ *Id.*

³¹ See James P. McGrath, *The Case Against GenAI Detection in Legal Education*, 74 J. LEGAL EDUC. 112, 118 (2024).

³² See, *Statement on GenAI Detection*, U.C. IRVINE (Dec. 15, 2023), <https://tinyurl.com/uci-ai-statement> [<https://perma.cc/8XVL-UF9P>] last visited Dec. 15, 2025).

³³ See *Turnitin Adding GenAI Writing Detection, But Instructors Should Use It With Caution*, PURDUE U., (Jan. 9, 2024), <https://www.purdue.edu/online/turnitin-adding-ai-writing-detection-but-instructors-should-use-it-with-caution/>, (last visited Dec. 16, 2025).

company now acknowledges a 4% likelihood that any specific sentence highlighted as AI-written might actually be human-written.³⁴

Turnitin itself advises that its GenAI detection score "should not be used as the sole basis for action or a definitive grading measure by instructors" and requires "scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred."³⁵ This is a remarkable admission from an industry leader. If the company that developed the tool does not trust it as a standalone measure, why should we? Unfortunately, some think that because a computer said it, it must be true. But what they're doing is simply rubberstamping a false outcome.

Independent testing reveals even more troubling accuracy issues. Error rates for human-written content range from 10-28%, and approximately 20% of AI-generated text remains completely undetected.³⁶ When researchers tested multiple detection tools on the same AI-generated article, results ranged from 100% AI-generated to only 30% of the text flagged with a 90% possibility of AI.³⁷ The same content run through the same detector twice can produce different results, with tools reversing their determinations about which portions were AI-generated and which were human-written.³⁸

4. Why Detection Fails in the Law School Context

The use of detectors in law school is particularly problematic for four reasons.

a. Bias

(1) The ESL Penalty

Research from Stanford University demonstrates that GenAI detectors misclassify texts written by non-native English speakers significantly more

³⁴ *Id.*

³⁵ *Id.*

³⁶ See *How GenAI Detectors Actually Work: The Truth About Undetectable GenAI Tools*, Litero (2025), <https://litero.ai/blog/how-ai-detectors-actually-work-the-truth-about-undetectable-ai-tools-2025-guide/>, (last visited Dec/ 16, 2025).

³⁷ See *How Reliable Are GenAI Detectors? Claims vs. Reality*, The Blogsmith, (Aug. 29, 2024), <https://www.theblogsmith.com/blog/how-reliable-are-ai-detectors/>, (last visited Dec/ 16, 2025).

³⁸ *Id.*

frequently than text by native English speakers.³⁹ In one study, non-native English speakers' essays were flagged as AI-generated at substantially higher rates even though they were entirely human-written.⁴⁰ These students often use more formal, less varied sentence structures to ensure grammatical accuracy. The very carefulness that characterizes responsible non-native writing becomes evidence of "AI signature" to these detectors leading to higher rates of false accusations for a vulnerable segment of our student body, raising serious due process and equity concerns.

(2) Students with Learning Disabilities

Autistic writers often produce structured, literal, formulaic prose, not because they used GenAI, but because that's how they communicate clearly and effectively. Neurodivergent students may have unique writing patterns, e.g., repetitive phrasing, formulaic structures, or unusual vocabulary choices, that confuse GenAI detectors not trained to recognize diversity in human writing. Students with ADHD or dyslexia who have been coached to write in plain, direct, accessible language face the same problem. Ironically, this type of writing is exactly what many GenAI detection tools mistakenly flag as "too structured" or "too basic."⁴¹

Moira Olmsted is an autistic student at Central Methodist University who, in 2023, was accused of using GenAI to write a reading summary, one of three she submitted weekly. The professor's GenAI detection tool flagged the assignment; she received a zero and a warning that a second flag would be treated as a plagiarism offense.⁴²

³⁹ Weixin Liang et al., *GPT Detectors are Biased Against Non-native English Writers*, 4 PATTERNS 1, 3 (2023).

⁴⁰ Andrew Myers, *AI-Detectors Biased Against Non-Native English Writers*, STAN. HCAI (May 15, 2023), <https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers>.

⁴¹ See Nate Pindell, *The Challenge of AI Checkers*, Center for Transformative Teaching, U. Neb. Lincoln, <https://teaching.unl.edu/ai-exchange/challenge-ai-checkers/> (last visited Feb. 21, 2026); Amanda Hirsch, *AI Detectors: An Ethical Minefield*, N. Ill. U. Center for Innovative Teaching and Learning (Dec. 12, 2024), <https://citl.news.niu.edu/2024/12/12/ai-detectors-an-ethical-minefield/> (last visited Feb. 21, 2026); Dr. Timothy Markley, *A Critical Examination of AI Detectors in Academic Integrity Enforcement*, K Altman Law (June 12, 2025), <https://www.kaltmanlaw.com/post/ai-detectors-academic-integrity-bias> (last visited Feb. 21, 2026).

⁴² Incident 849: *AI Detection Tools Allegedly Misidentifying Neurodivergent and ESL Students' Work in Academic Settings*, AI Incident Database, (Oct. 18, 2024), <https://incidentdatabase.ai/entities/moira-olmsted/> (last visited Feb. 21, 2026).

Olmsted pushed back, explaining that her structured, formulaic writing style is a characteristic of her autism, not evidence of AI use. The grade was restored, but the warning stood. To protect herself going forward, she began recording her writing process and using Google Docs to maintain an edit history, not to become a better writer, but to prove she was one.⁴³

Olmstead's experience clearly demonstrates the false positive problem—the detector wasn't wrong about the writing's characteristics. It was wrong about what those characteristics mean. Structured, predictable prose is a feature of how many autistic people communicate clearly and is not a signature of machine generation. The tool couldn't tell the difference, and, unfortunately, the institution treated the flag as presumptive evidence rather than a prompt for conversation.

b. The Model Limitation Problem

Detection tools can only identify writing from models they have been trained to recognize. As of early 2024, Turnitin claimed to detect writing from GPT-3 and GPT-3.5, with GPT-4 detection working "most of the time." In early 2026n Turnitin reports that its "expanded our detection capabilities to include GPT-4 (ChatGPT Plus), GPT-4o, GPT-5, GPT-5-mini, GPT-5-nano, GPT-5.1, GPT-5.2, GPT-5.2-pro, Gemini (Pro), Gemini-2.5-pro, Gemini 2.5 Flash, Gemini-3-flash-preview, Gemini-3-pro-preview, Claude Sonnet-4.5, LLaMA. Additionally, we have released capabilities to detect likely AI-generated content that may have been modified by bypassers and paraphrasers."⁴⁴ Note that in April 2026, Anthropic released Claude Opus 4.7, and the detection tools haven't caught up. Students who choose different models or newer versions than those covered by the detection tools are functionally invisible to detection software. This creates an arbitrary enforcement regime where students using older, more detectable models face accusations while students using newer models avoid scrutiny entirely.

One logical question comes up when I speak to students and professors. "Can't I just ask ChatGPT if it wrote a particular text? Surely, it can detect its own output." Well, you can ask, and it might give you a plausible answer. "But ChatGPT has no 'knowledge' of what content could be AI-generated or what it

⁴³ *Id.*

⁴⁴ Turnitin's AI writing Detection Capabilities FAQs, <https://guides.turnitin.com/hc/en-us/articles/28477544839821-Turnitin-s-AI-writing-detection-capabilities-FAQs>, (last visited Apr. 17, 2026).

generated. It will sometimes make up responses to questions like ‘did you write this [essay]?’ or ‘could this have been written by AI?’ But these responses are random and have no basis in fact.”⁴⁵

c. The Adversarial Loop

Detection is a cat-and-mouse game. As soon as a detection algorithm is updated, new "humanizing" tools emerge to bypass it. Students can simply prompt an LLM to "rewrite this with high burstiness" or "elevate the provided text by employing literary language" or “add cultural references” to defeat detectors.⁴⁶ Commercial tools like Quillbot and a growing cottage industry of GenAI humanizers can mask the GenAI signature by adding *improper* grammar and *misspellings*, and by varying sentence structure.⁴⁷ The technological arms race will continue indefinitely, with each side developing more sophisticated methods. Law schools should not build disciplinary policies on tools that students can defeat with a single additional prompt.

d. Lack of Due Process

A detection score is a probability, not a fact. It cannot be cross-examined. Most major companies, including OpenAI, have retracted their detection tools because the error rates were too high for reliable use.⁴⁸ Relying on an "80% GenAI Likelihood" score in a character and fitness hearing is an evidentiary nightmare. Students bear the burden of proving a negative—that they did not use GenAI—when the accusation is based on probabilistic software that even its creators acknowledge is unreliable.

5. The Institutional Response

The inadequacy of detection tools has led numerous institutions to reject them entirely. The UC San Diego disabled Turnitin's GenAI detection feature for all faculty and required faculty to remove AI-Detection Language from course

⁴⁵ Can I ask ChatGPT If It Wrote Something?, OpenAI, (Sept. 2025), <https://help.openai.com/en/articles/8318890-can-i-ask-chatgpt-if-it-wrote-something>, (last accessed Dec. 31, 2025).

⁴⁶ Myers, *supra* note 40.

⁴⁷ Maryam Masrour et al., *GenAI Humanizer Tools: Implications for Academic Integrity*, 15 J. Educ. Tech. 89, 92 (2025).

⁴⁸ *How Can I Tell if a Piece of Writing is AI-Generated?*, OPENAI Help Center (last visited Dec. 29, 2025), <https://help.openai.com>.

materials.⁴⁹ Purdue University issued specific warnings when the tool was automatically added to its system, cautioning instructors "to be cautious about relying fully on results from the system because it may return false positives or fail to detect some AI-generated material."⁵⁰

Institutional rejections are not based on abstract concerns about academic freedom. They reflect a practical assessment that the tools create more problems than they solve. The UCI committee specifically noted that GenAI detection offers no context for its conclusions, stating only that a percentage of submission "has been determined to be generated by AI" without explaining which features triggered the flag.⁵¹ This opacity (the black box) makes it impossible for students to mount an effective defense or for faculty to make informed judgments about alleged misconduct.

6. Conversation, Not Accusation

I hope I've sufficiently made the case to forego detection tools. If your school requires their use, avoid direct accusations based on software scores. Instead, ask a student to explain the reasoning behind a specific footnote or the choice of a particular case. A student who has used GenAI as a shortcut will often struggle to explain why particular choices were made. Questions such as, "Walk me through your reasoning in paragraph three. Why did you choose to distinguish *Smith* rather than follow it?" or "This footnote cites a 1987 case. What made you think to look back that far in the case law?" probe understanding without triggering defensive responses.

An oral defense model might be the most effective way to verify authorship in the age of GenAI. It also has the pedagogical benefit of reinforcing that legal writing is not merely about producing text but about developing and defending legal arguments. If time allows, incorporate brief oral explanations as a regular

⁴⁹ *AI Detection Deactivation: Instructor Action Required and Next Steps*, U.C. SAN DIEGO, Feb. 2025, <https://extensionhelpcenter.ucsd.edu/hc/en-us/articles/34799323794957-AI-Detection-Deactivation-Instructor-Action-Required-and-Next-Steps> (last accessed Dec. 31, 2025).

⁵⁰ R.M. Barton, *Turnitin Adding AI Writing Detection, But Instructors Should Use it With Caution*, Purdue U., <https://www.purdue.edu/online/turnitin-adding-ai-writing-detection-but-instructors-should-use-it-with-caution/>, (last accessed Dec. 31, 2025).

⁵¹ U.C. IRVINE, *supra* note 32.

component of written assignments, which will normalize the expectation that students must be able articulate the reasoning behind their work.

7. Undetectable Tools

a. AI Translation and Paraphrasing Tools

One of the most insidious forms of AI-enabled cheating involves paraphrasing tools that allow students to present others' work as their own while technically avoiding plagiarism detection. Here's how it works: A student copies text from a source or AI-generated output and runs it through tools like QuillBot, WordTune, or similar paraphrasing software. The text emerges technically "original" because every sentence has been reworded and vocabulary has been changed.

Several tell-tale signs suggest a written work *might* be AI-generated, and most of them are unfortunate flags for legal writers specifically. The usual suspects: overuse of words like *delve*, *crucial*, and *tapestry*; an abundance of em dashes; tricolons (those three-item series at sentence end); and transition words like *similarly*, *consequently*, and *therefore*. The irony is that legal writers have used transition words and tricolons since long before anyone had heard of ChatGPT, and *delve* has a perfectly respectable history that predates LLMs by centuries. For legal scholars and law students, this creates a problem: the stylistic fingerprints GenAI detection tools flag as suspicious are often just ... good legal writing. Well, maybe not *tapestry*.

The workaround, for those trying to evade detection, is simple word substitution: *delve* becomes *dig*, *pivotal* becomes *important*. A few swaps can stymie most detection tools, which are pattern-matching on vocabulary signals rather than evaluating prose quality. One study analyzed 14 million scientific paper abstracts published on PubMed between 2010 and 2024 and found that *delves* appeared twenty-five times more often in 2024 than pre-LLM trends would predict, with *showcasing* and *underscores* each spiking ninefold.⁵² Flagged as a

⁵² Kyle Orland, *The Telltale Words that Could Identify GenAI Text*, Ars Technica, July 1, 2024, <https://arstechnica.com/ai/2024/07/the-telltale-words-that-could-identify-generative-ai-text/> (last accessed Dec. 31, 2025). The study analyzed 14 million paper abstracts published on PubMed between 2010 and 2024, establishing pre-2023 usage baselines and comparing them to word frequencies after LLMs became widely available. Beyond *delves*, *showcasing* and *underscores* each increased ninefold, while previously common words like *potential*, *findings*, and *crucial* also surged measurably in post-LLM abstracts.

marker of AI-generated text, the em dash has been a fixture of legal prose for generations, and advising lawyers to ration em dashes feels like punishing the wrong party.

Some students take evasion further by running text through translation software, converting it to another language and back to English, which scrambles enough phrasing to obscure the source. It's a nearly undetectable technique because plagiarism detection software, like Turnitin, is designed to flag identical or highly similar text, not paraphrased content. Even sophisticated detection tools struggle to identify when a student has systematically reworded source material while preserving its fundamental structure and ideas. The final product appears sufficiently different from any single source that it passes automated screening, yet the student has contributed no original analysis or thinking.

The problem is particularly problematic for students whose first language isn't English. These students may honestly view paraphrasing tools as legitimate translation assistance rather than academic dishonesty. They may believe that if they're taking content and "putting it in their own words" (even with GenAI assistance), they're engaging in acceptable summarization rather than plagiarism. It then becomes difficult to address the issue without appearing to discriminate against English language learners, even though the behavior likely violates academic integrity regardless of the student's language background.

In legal education, virtually any written assignment is affected. Research memos, seminar papers, case briefs, and even exam answers (if not otherwise protected by secure exam software) can all be produced through systematic paraphrasing of source material or GenAI generated content.

So, what works?

I can guess what you're thinking. "But we already require analysis! That's what CREAC is for. That's what IRAC does. Students can't just paraphrase their way through a legal analysis because the analysis itself is the point."

And you're right, at least in theory. The problem is that GenAI can now generate perfectly competent CREAC, and IRAC, and CRAC (which, when used appropriately, can be a huge benefit for students who want immediate feedback on their writing). It can identify rules, explain them, apply them to facts, and reach conclusions. And then QuillBot and Refinely and other "humanizing" tools can paraphrase that AI-generated CREAC so thoroughly that it reads like original

student work. Your student submits a memo with proper legal structure, clear rule explanation, thoughtful application, and a well-supported conclusion. It looks like analysis. It's formatted like analysis. But the student never did any analyzing. They just competently paraphrased someone (or something) else's analysis, with maybe a few *deliberate* grammar and spelling errors.

So, it's probably time to go deeper than CREAC. We need to assess the analytical process that led to the CREAC, not just the CREAC itself.

Require students to submit messy research notes before they write. And I mean really messy. Not a polished outline that could itself be AI-generated, but evidence of thinking-in-progress. You want to look at the notes that show dead ends they pursued and abandoned; questions they asked themselves about how to apply a standard; uncertainty about whether Case A or Case B is more analogous to their facts; and preliminary judgments about argument strength that perhaps changed as they kept researching. Have them use Track Changes and submit that redlining. A student who's planning to paraphrase AI-generated analysis either can't produce these notes or produces notes that are suspiciously complete and polished.

Detection tools are unreliable and discriminatory. When something feels off, don't run to a detector. Have a conversation. Ask students to walk you through their process. Your instincts plus strategic questions will tell you more than any algorithm ever could.

Ask students to explain not just their reasoning, but their reasoning choices. A student who paraphrased AI-generated CREAC can walk you through the CREAC. They can tell you what rule they identified, why they chose the facts they did, and what conclusion they reached. That's just reading their own memo back to you. But ask them *why* they structured the rule explanation the way they did. Why did they lead with Case X instead of Case Y. Why did they discuss Exception A before Exception B. Students who did their own analysis can tell you. "I started with Case X because it's most factually similar, even though Case Y is more recent." "I discussed Exception A first because I knew opposing counsel would lead with it, so I wanted to defuse it early." Students who paraphrased GenAI have no idea. Why? Because they didn't make those choices.

Similarly, ask students what gave them the most trouble. Students who did their own work remember the hard parts. "I spent two hours trying to figure out

whether *Smith* or *Jones* controlled, and I finally realized they're not actually in conflict and that they just emphasize different factors." Students who paraphrased will be flummoxed.

The conversation reveals what the written work conceals, namely whether the student understands the analysis that she has submitted. And the really nice thing about it is that it isn't punitive or accusatory. Will they be nervous? Sure. Don't focus on whether GenAI was used. Instead, stress that you're meeting with her to determine how you can help. If a student struggles in the meeting, you've identified someone who needs more support in developing legal reasoning skills, whether the student paraphrased GenAI or just didn't understand the assignment. Either way, you can intervene before the student gets to the final exam or, worse, to the practice of law.

One more thing. Consider requiring students to submit an analytical process memo along with their final work product. It must be a genuine reflection explaining their choices because research logs can be faked with GenAI. "I initially thought the client would win under Theory X, but after reading *Smith*, I realized that Theory Y was stronger because... I chose to lead my discussion with the statutory argument rather than the case law because... The hardest part of this analysis was... If I had more time, I would..." Metacognitive reflection is very difficult to fake because it requires understanding not just what the analysis says, but why it says it. GenAI can generate analysis, but it can't generate authentic reflection on the experience of doing analysis.

It may be that CREAC and IRAC are necessary but no longer sufficient. Faculty need to assess not just whether students can produce legal analysis, but whether they can explain the analytical choices that produced it, because that's what they'll need to do when a senior partner asks "Why did you structure the brief this way?" or a judge asks "Counsel, why should I apply *Smith* rather than *Jones*?" The ability to articulate analytical choices is itself a professional competency, and it's one that paraphrasing tools can't provide.

Managing all of this is, I'm sure, starting to sound like a second full-time job. Reviewing research notes, scheduling process conversations, reading process memos on top of grading the actual work — that's a lot, and for faculty already managing large courses, and for adjuncts working full-time jobs, I agree that it's a lot of extra work. *But don't do all of this for every student.* The notes requirement and Track Changes submission function mostly as deterrents; knowing you might

look changes behavior even when you don't. The process conversation is most useful when something already *feels* off, and fifteen minutes with a student beats running a faulty detection tool and second-guessing the result for a week. The analytical process memo is a good idea regardless of the GenAI question because it teaches metacognition, which lawyers genuinely need. You're not adding an assignment to catch cheaters; you're adding one that develops better legal reasoning. Initial screening of research notes and redlined drafts is also work a trained TA or writing fellow can handle, flagging what warrants your attention. The conversation, though, belongs to you; you can't delegate that.

b. How Google Lens Enables Invisible Cheating

Short of using preventative measures such as secure software (Examsoft or Exam4), which many law schools have, or engaging on-site proctors who *actually* monitor students, not much helps fight the challenge that Google Lens presents. Online monitoring software alone cannot solve this problem. Certainly, none of the detectors discussed above will help. So, what's the problem?

Google recently made its visual search tool, Lens, significantly easier to use within the Chrome browser. The mechanics are simple and nearly undetectable:

1. **One-click activation:** Students click an icon in the Google search bar
2. **Visual selection:** A moveable bubble pops up that can be placed over any question or text on screen
3. **Instant GenAI answers:** A sidebar appears with AI-generated answers, explanations, or interpretations of whatever is inside the bubble
4. **No trace:** Students can cheat on unsecure digital tests without typing a single prompt or even leaving the exam page

The process is fast, silent, and leaves no digital footprint that typical monitoring software would catch. As one teacher put it, "[t]eachers and school leaders spend countless hours considering each detail of the learning experience, then Google totally undermines it with the click of a button."⁵³

Because neither Exam4 nor Examsoft's student application, Exemplify, run in a browser window, students cannot access Lens.

⁵³ Carolyn Jones, *His Students Suddenly Started Getting A's. Did a Google GenAI Tool Go Too Far?*, THE MARKUP, Nov. 9, 2025, <https://themarkup.org/artificial-intelligence/2025/11/09/his-students-suddenly-started-getting-as-did-a-google-ai-tool-go-too-far>, (last accessed Dec. 31, 2025).

c. The Vulnerability of Take-Home Exams

The take-home exam has always operated on the honor code (and sort of a fiction). Students agree to work independently, consulting only the materials the professor specified, without meaningful outside assistance. For decades, the fiction was sustainable. Students might call a classmate or consult an outline they weren't supposed to have, but the scope of available help was constrained by the hours in a day and time limitations of the exam.

GenAI dissolved that fiction.

A student with a take-home exam and a ChatGPT subscription now has access to something that performs at or above the level of a newly-minted attorney on most standard law school assessments. The bar exam — long the benchmark for professional competence — has been consistently cleared by leading LLMs since GPT-4.⁵⁴ A final exam question in Contracts or Torts presents considerably less of a challenge. The student uploads the exam, types a prompt, and receives a well-organized response with citations that may even be accurate, the whole process taking less time than finding parking near the library.

The countermeasures that work for in-person exams don't transfer. Secure exam software like Examsoft or Exam4 locks down the student's computer during an in-person exam, blocking browser access entirely. For take-home work, that protection is simply absent. A student at home can take the exam on one computer and use another computer, tablet, phone, without leaving any detectable trace. The more direct approach: photograph the exam question with a phone and upload the image to any GenAI platform capable of processing images. The exam never leaves its original window, and the answer arrives through a separate channel that monitoring software never touches.

Detection tools are no help here, for the same reasons they're no help anywhere else: they cannot reliably distinguish a student who used GenAI from a student who writes in a precise, organized style.

None of this means the take-home exam should be retired. But assigning a take-home exam while prohibiting GenAI use is no longer a coherent policy,

⁵⁴ Pablo Arredondo, *GPT-4 Passes the Bar Exam: What That Means for Artificial Intelligence Tools in the Legal Profession*, Stan. L. Blog (Apr. 19, 2023), <https://law.stanford.edu/2023/04/19/gpt-4-passes-the-bar-exam-what-that-means-for-artificial-intelligence-tools-in-the-legal-industry/> (last visited Apr. 5, 2026).

absent some verification mechanism that doesn't currently exist. Professors who want to preserve the format should think carefully about what the exam actually measures and design accordingly by either building GenAI disclosure and verification into the assignment itself or shifting to formats where compliance doesn't rest on the honor system alone.

d. AI as "Research Assistant"

The problem here is that many students genuinely don't understand that they're doing anything wrong. Here's how it works: A student asks ChatGPT or Lexis+ to "find cases about qualified immunity" or "give me cases on the business judgment rule." The GenAI generates a list of citations, some real and some hallucinated. The student then does what he thinks is due diligence by verifying that the real cases exist and uses the cases in the memo. The student can produce the cases if asked. The citations are accurate. There's no technical plagiarism. The student believes that he has done everything right by verifying that the cases are real.

But did the student read the cases? He confirmed the cases exist, grabbed a headnote or a quote that seemed relevant, and moved on. The student used GenAI not as a research starting point but as a research substitute, skipping the entire process of reading cases. He doesn't understand the reasoning, and he didn't evaluate whether authorities are on point. He didn't exercise judgment about which cases best support their argument. The GenAI did the analytical work of identifying relevant authority, and the student just verified that the GenAI wasn't hallucinating citation(s).

What's happening here is what law librarian Professor Rebecca Fordon calls the "AI Tunnel."⁵⁵ Students go into a GenAI research tool, get an answer, and then continue the conversation by asking more questions. They have a 30-minute conversation with the GenAI that leaves them with only a handful of sources, and not always the most relevant or authoritative ones. "They miss critical nuance, ... and they can't see what the GenAI has limited or hidden from them."⁵⁶

⁵⁵ Rebecca Fordon, *Legal Research Trapping You in an "AI Tunnel"? Use a Toe-hold to Get Out*, GenAI Law Librarians, Oct. 28, 2025, <https://www.ailawlibrarians.com/2025/10/28/legal-research-trapping-you-in-an-ai-tunnel-use-a-toe-hold-to-get-out/>, (last accessed Dec. 31, 2025).

⁵⁶ *Id.*

Experts use GenAI completely differently. They use it for what Fordon calls a "toe-hold."⁵⁷ They ask GenAI to explain the elements of a claim to get key concepts, then immediately pivot to a treatise for detail.⁵⁸ They ask GenAI to find a statute to get the statute number, then immediately pivot to the Notes of Decisions.⁵⁹ They ask GenAI for a few starting cases to get one good case, then immediately pivot to the citator and its headnotes.⁶⁰ They use GenAI as a "[one]-minute scaffold to get [us] to traditional research tools" where the real work happens.

Students are missing the pivot.⁶¹ They stay in the tunnel.⁶²

The process is nearly undetectable because the student has technically done everything right from a citation perspective. The cases are real. They're properly Bluebooked. The student can produce them if you ask. From the outside, it looks like competent legal research. But the student has developed none of the skills that legal research is supposed to teach. The student doesn't know how to construct a search query, use a citator to find related cases, evaluate whether an authority is mandatory or persuasive, or assess whether a case's reasoning supports the proposition for which it's cited.

For legal education, the integrity of any assignment requiring case law analysis is threatened. It's particularly dangerous because students often don't realize they've crossed a line. They think "AI-assisted research" means using GenAI to find cases, when what we mean is using GenAI to help refine search terms or understand complex holdings after you've done the initial research yourself.

If you've been reading along, you won't be surprised to find that the solution lies not in detection, but in explanation and documentation. Require students to explain how they found each source. Ask them to document their research process with the actual database search terms they used and the citators they used (screenshots are good for this). A student who used GenAI as the

⁵⁷ *Id.*

⁵⁸ *Id.*

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ For more on the importance of pivoting, see *The One with the Cop*, Friends (NBC television broadcast Feb. 25, 1999), <https://youtube.com/shorts/xUYTKixFpC4?si=xXv1AG3HbskGhkM0> (last visited Apr. 17, 2026).

⁶² Fordon, *supra* note 55.

research engine can't produce this documentation because the student never conducted research. Ask students to explain why they chose certain cases over others. If a student relied on GenAI recommendations, that student can't tell you why *Smith* was more relevant to the facts than *Jones*, or why the memo led with the Second Circuit case instead of the Supreme Court dictum. The student is just reporting what the GenAI prioritized.

Better yet, meet with students and ask them to explain the holding and reasoning of key cases they cited. A student who actually read the cases can walk you through the facts and the court's analysis. A student who just verified the AI's citations and grabbed quotes will struggle. That student can tell you what the case supposedly stands for but can't explain how the court got there or why the reasoning is persuasive.

The important thing to remember is that verification isn't the same as research. Students who verify GenAI-generated citations have prevented hallucinations, which is good. But they haven't developed analytical judgment or the ability to evaluate authority critically. And those are the skills that separate competent lawyers from malpractice defendants.

e. When Countermeasures Fall Apart

*Warning: this section might be controversial, but it must be addressed.*⁶³

There's a problem. Many of the countermeasures offered above fall apart when it comes to students with testing accommodations. A student with 50% or 100% or 200% extended time on a secure exam has plenty of time to photograph questions and transcribe GenAI answers. The time pressure that makes GenAI use impractical for other students doesn't exist for them. A student accommodated with a low-distraction or private testing room has no proctor watching to ensure that he or she is not using GenAI. A student with break accommodations can step out of the testing room, use GenAI during the break, and return with answers ready to transcribe.

⁶³ An anonymous peer reviewer said, "LOL-all of the book can be considered controversial but I don't think anything discussed needs trigger warnings." Probably right.

These are legally required accommodations under the ADA and state disability rights laws. We can't refuse them, and we shouldn't want to. They're essential for leveling the playing field for students with disabilities. But we also need to be honest about the fact that the same accommodations that provide necessary access also create opportunities for GenAI misuse that the countermeasures discussed can't address.

Except maybe for conversation with the student as described above. Have the student explain their work. I get that it takes time. It doesn't scale. For those teaching Contracts to 90 students, meeting with every student to discuss their analytical choices is impracticable. But you don't have to meet with everyone. You only meet with the students your gut flags. That might be a student with an accommodation or without one. Selecting *only* accommodated students for scrutiny might create a discrimination problem.

I am not suggesting, of course, that students with accommodations are more likely to cheat. Of course they aren't. But it does mean that the exam security measures that work for non-accommodated students often don't work for accommodated students, and that creates a dilemma without easy answers.

What should the law school do? First, you could consider confiscating phones, while recognizing that a student may carry two—one to turn in and one to keep. Short of frisking or installing magnetometers, which raise a bunch of other issues, you might introduce friction, but that doesn't entirely solve the problem. Be sure to involve your disability services office to understand what level of proctoring or monitoring is permissible for accommodated exams. Some students' accommodations specifically require no proctoring, while others may allow for modified proctoring arrangements.

Second, consider whether your high-stakes assessments really need to be closed-book, timed exams. Remember, if you can assess competence through take-home assignments that allow GenAI use with proper disclosure and verification, combined with a reflection you've eliminated the problem entirely.

Third, for courses where closed-book exams are essential, consider alternative assessment methods for accommodated students that test the same competencies in different ways. An oral exam, for instance, provides the same disability accommodations (extended time, private room, breaks) while making GenAI use much more difficult. Again, check with disability services.

Law schools have to face the fact that exam security measures don't work equally well for all students. Administrators and professors need to think more creatively about how to assess competence in ways that are both accessible and integrity-preserving. I've been involved in exam administration for more years than I care to admit. I know from experience that more monitoring and restrictions aren't adequate to the times.

D. Designing Faculty Development Around Enforcement

When creating faculty workshops or departmental guidelines on this topic, (please) avoid the temptation to provide a "detection playbook." Instead, focus on helping colleagues develop sophisticated assignment design skills and conversational investigation techniques.

Sample faculty workshop agenda:

- Review detection tool limitations with specific data on false positive rates
- Practice designing AI-resistant assignments specific to your courses
- Role-play investigative conversations with students
- Develop shared departmental standards for GenAI
- Discuss grade penalties for process violations vs. substantive violations

The goal is to shift faculty culture from a surveillance mindset to a process verification mindset. We cannot detect our way out of the GenAI challenge. We can, however, redesign our pedagogy to make authentic engagement more valuable and verifiable than GenAI shortcuts.

E. GenAI-Resistant Assignments

If law schools cannot (and should not) rely on detection, the focus has to shift from product to process. Yes, it's more labor-intensive for the professor but far more defensible and pedagogically sound.

1. The Mandatory Revision History

Require students to submit their work using platforms that track changes over time, such as Google Docs or Microsoft Word Online. A human-written legal

essay should show a "lived-in" history. You should see the student grappling with sentences and adding citations over several days. A document that appears fully formed in a single "copy-paste" event is a red flag that warrants a conversation, regardless of what a detector says.⁶⁴

When designing this requirement for your syllabus, specify the platform and make clear that revision history is part of the assignment. For example, "All written assignments must be completed in Google Docs with edit history enabled. You must share the document with [professor email] with 'Editor' permissions at least 48 hours before the deadline. Failure to provide revision history will result in a grade penalty."

The approach serves dual pedagogical purposes by providing evidence of authentic authorship while encouraging students to engage in the iterative revision process that characterizes strong legal writing. You can use sample revision histories in your faculty workshops to demonstrate what authentic writing development looks like versus suspicious patterns.

2. Hyper-Local Facts

Ask students to apply a legal principle to a specific, unreported case from your local jurisdiction. GenAI models are trained primarily on published opinions and major cases. They perform poorly when asked to analogize to fact patterns from unpublished state trial court decisions or local administrative proceedings.⁶⁵ Your law school's law librarian can help.

3. Classroom Integration

Require students to cite a specific hypothetical discussed during the Tuesday lecture or to build upon a peer's comment from the Socratic dialogue. GenAI cannot recreate the specific intellectual discourse of your particular classroom community.

⁶⁴ Cf. MODEL RULES OF PRO. CONDUCT r. 1.1 cmt. 8 (A.B.A. 1983) (emphasizing the need for lawyers to stay abreast of the benefits and risks associated with relevant technology).

⁶⁵ *Strategies for Designing Ai-Resistant Assignments*, U. CHI., <https://genai.uchicago.edu/en/resources/faculty-and-instructors/strategies-for-designing-ai-resistant-assignments>, (last visited Dec. 30, 2025).

4. Document Drafting with Specific Constraints

Rather than asking for a generic memo, require students to draft a motion in limine for a case currently pending in your state's trial courts, incorporating your jurisdiction's specific procedural rules and local rules of court. Include requirements to address recent oral arguments or status conferences that would not appear in published materials.

5. The "Vibe" Check

AI is excellent at summarizing law but often fails at the "messy" application of facts to law. If the analysis section of a paper is vague or overly optimistic about a clear legal conflict, it may indicate a lack of human critical thinking. When developing grading rubrics, include specific criteria for evaluating depth of legal reasoning and acknowledgment of counterarguments.

6. The Paper Trail

Instead of grading only the final legal memo or brief (which GenAI can generate in seconds), grade the intellectual evolution of the assignment by requiring students to submit a "Process Portfolio" documenting their research and drafting progression.

Require three components:

- 1) A handwritten issue tree or case analysis chart showing initial legal problem decomposition before research begins. This should identify potential causes of action, elements, defenses, and research questions in the student's own handwriting.
- 2) An annotated research log where students document each case or statutory provision consulted, explain why they selected it for deeper reading (based on headnote, KeyCite treatment, or citation in another source), summarize the holding, and note whether it became part of the final analysis and why. Each entry should be 2-3 sentences explaining the research decision-making process.
- 3) A complete revision history using Google Docs or Microsoft Word Online showing the document's development over multiple sessions spanning at least 5 days, with visible evidence of substantive revisions (reorganized arguments, added counterarguments, refined legal rules, inserted

additional authority).

The assignment is AI-resistant because GenAI cannot recreate a plausible human research trajectory or generate handwritten legal analysis charts. The research log requires explaining selection rationale based on legal research tools and strategies GenAI cannot access or simulate authentically. When grading, evaluate the quality of legal reasoning in the research log (did students identify relevant authority through sound research methodology or did they simply list cases?). Check that the handwritten issue tree reflects legal analysis before conclusions were reached, and verify the revision history shows iterative refinement rather than a single drafting session.

Students who use GenAI to generate the final product will struggle to create a revision history showing realistic legal writing development, and their research logs will lack the specific details about why particular cases were useful or how they found them through legal research databases. Award significant points (20-25%) to the process documentation to incentivize authentic engagement with legal research and writing methodology.

F. GenAI-Transparent Assignments

1. The "AI Impact Statement" and Appendix

If your policy allows for the use of GenAI, require students to submit an "GenAI Appendix," which should include every prompt used and the resulting output. It's easy enough for students to create by printing their chat. Students must also include a brief narrative explaining how they verified the AI's output and what changes they made to ensure legal accuracy, which forces the student to engage in the "human-in-the-loop" oversight that is required for ethical legal practice.⁶⁶

The GenAI Appendix serves as both transparency mechanism and teaching tool. When students must document their verification process, they confront the reality that AI-generated legal analysis often contains hallucinated citations and

⁶⁶ See Cassandre L. Silva, *The GenAI Appendix: A New Standard for Legal Scholarship*, 56 U. Mich. J.L. Reform 201, 215 (2024) (arguing that transparency in GenAI usage is the only way to maintain academic and professional integrity).

incomplete answers. Instead of using GenAI as a shortcut it becomes a learning experience about the limitations of automated legal reasoning.⁶⁷

Consider developing a rubric that evaluates the quality of the GenAI Appendix itself. Strong appendices should demonstrate: (1) specific, well-designed prompts rather than generic requests; (2) critical evaluation of GenAI output identifying errors or gaps; (3) substantial human revision and verification; and (4) proper citation checking against primary sources. By the way, GenAI can help you create the rubric! See Appendix E.

2. Comparative Analysis

I developed this assignment in Spring 2024 so students can learn the benefits and risks associated with different types of legal technologies. I require students to draft a negligence complaint using two different platforms: a GenAI tool (ChatGPT, Claude, or similar), a traditional document assembly program (e.g., Clio Draft), and a form from either Westlaw or Lexis. Students must then submit a comparative analysis evaluating legal sufficiency, procedural compliance, factual accuracy, and strategic considerations of each document. The assignment is GenAI-transparent because students are directed to use GenAI but cannot simply accept either platform's output as final work. They must identify specific deficiencies, hallucinated facts, missing procedural requirements, and strategic weaknesses that require applying jurisdiction-specific knowledge that GenAI lacks. When grading, look for evidence that students compared documents side-by-side against specific procedural rules and charted factual omissions or fabrications. Recommendations should reflect actual understanding of litigation dynamics, not vanilla pabulum about "GenAI limitations." The full assignment includes a detailed fact pattern with specific dates and amounts that GenAI commonly mishandles and a GenAI Appendix template documenting all prompts and verification processes. You can find it in the Appendix F. Ironically, GenAI can help you write this assignment.

⁶⁷ *GenAI Considerations for Teaching and Learning*, OHIO S., <https://teaching.resources.osu.edu/teaching-topics/ai-considerations-teaching-learning>, (last accessed Dec. 30, 2025).

3. The Hallucination Hunt

We've already discussed in class that confirming that a citation exists is not the same as confirming that it says what GenAI claims it says. GenAI hallucinates both; it invents citations that don't exist, and it misrepresents holdings in citations that do. Catching a fake citation requires a database search. Catching a misrepresented one requires you to read the case. Both skills are non-negotiable. Lawyers who skip these verification steps are the ones who get sanctioned.

The assignment is to audit every citation in a passage. It teaches "Verification Competence," a core skill identified in the ABA 2025 GenAI Task Force Report.⁶⁸

Your Assignment

For each of the five citations in the excerpt below, complete the following analysis in writing.

- First, verify that the citation exists. Confirm the correct case name, reporter, volume, page, court, and year. Note any discrepancies between what the memo states and what you find.
- Second, verify what the case holds. Read enough of the opinion to determine whether the proposition the memo attributes to the case is accurate, overstated, understated, or flatly wrong.
- Third, classify each citation using one of these four categories: Accurate, Exists but Misrepresented, Exists but Inapplicable, or Fabricated.
- Fourth, for any citation that is not accurate, rewrite the sentence to either correct the proposition or replace the citation with an appropriate authority.

⁶⁸ Year 2 Report on the Impact of GenAI on the Practice of Law, ABA Task Force on Law and Artificial Intelligence, Dec. 2025, <https://www.americanbar.org/content/dam/aba/administrative/center-for-innovation/ai-task-force/2025-ai-task-force-year2-report.pdf>.

The Excerpt

The admissibility of AI-generated evidence presents novel questions that courts are only beginning to address. Authentication remains the threshold issue, and courts have consistently held that the proponent of digital evidence must demonstrate that the evidence is what it purports to be. *United States v. Vayner*, 769 F.3d 125 (2d Cir. 2014). The Federal Rules of Evidence do not categorically exclude AI-generated content, though courts have applied Rule 901's authentication requirements with particular care in digital contexts. *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534 (D. Md. 2007).

Reliability is the second major hurdle. At least one court has held that AI-generated analysis is per se inadmissible unless the underlying algorithm has been independently validated. *United States v. Gissantaner*, 990 F.3d 457 (6th Cir. 2021). The Supreme Court established the foundational standard for scientific evidence admissibility in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), where the Court held that judges must act as gatekeepers to ensure that "only reliable, peer-reviewed scientific evidence is admitted at trial." Finally, courts evaluating AI-generated outputs have found that the attorney's duty to supervise extends to AI tools, and that reliance on AI without independent verification constitutes ineffective assistance of counsel per se. *Maas v. Cornell Univ.*, 94 N.Y.2d 87 (1999).

What You're Looking For

At least one citation in the excerpt does not exist. At least one exists but is attributed a holding the court never stated. At least one exists but has been applied to a proposition it has nothing to do with. One citation is accurate. Finding all four categories is the assignment. Finding only the fake one is not.

A Note on Tools

You may use Westlaw or Lexis to verify citations. You may not use a GenAI tool to verify them; that's the problem we're solving, not the

solution. If you find yourself tempted to ask ChatGPT or Claude whether a case is real, consider what you're practicing.

4. Low/No Stake Formative Assessments – GenAI Assisted Iterative Review

a. Essays

Require students to complete an iterative revision cycle using GenAI feedback. Students submit their initial essay draft to a GenAI tool using a professor-provided prompt, revise based on the critique, submit the revision for a second round of GenAI feedback, then write a reflection analyzing the revision process. Provide students with specific prompts at each stage to ensure consistent feedback quality. For the initial critique, for example, use:

You are a law professor reviewing a student essay on [topic]. Evaluate this essay for: (1) thesis clarity and argumentation strength, (2) use of legal authority and citation accuracy, (3) organization and paragraph structure, (4) analysis depth, and (5) counterargument acknowledgment. Provide specific suggestions for improvement with examples from the text.

For the second critique after revision, use:

Compare this revised essay to the feedback provided in the previous critique. Identify: (1) which suggestions were successfully implemented, (2) which revisions improved the argument, (3) remaining weaknesses, and (4) new issues introduced during revision.

For the reflection, require students to address:

- Which GenAI suggestions improved your essay and why?
- Which suggestions did you reject and what was your reasoning?
- What did this process teach you about your writing strengths and weaknesses?
- How would you verify that GenAI critique is accurate before implementing suggestions?

This assignment is AI-resistant because the reflection requires analysis that students cannot outsource. When grading, evaluate whether students demonstrate genuine engagement with the revision process rather than superficial compliance.

b. MCQ Diagnostics

Require students to complete a diagnostic learning cycle on a substantive topic using multiple-choice assessments and GenAI analysis. Students take an initial 25-question quiz on a specific topic (e.g., homicide, contracts formation, intentional torts), upload their results to a GenAI tool using a professor-provided tool for analysis, study the identified weak areas, take a second diagnostic covering the same material, then submit both result sets with a reflection on their learning process.

Provide students with this prompt for analyzing their first diagnostic:

"I scored [X/25] on a diagnostic exam covering [topic]. Here are the questions I answered incorrectly: [paste questions and your wrong answers]. For each incorrect answer: (1) explain the correct legal rule, (2) identify why my answer was wrong, (3) suggest a study strategy to master this concept, and (4) flag any patterns in my errors (e.g., misunderstanding elements vs. misapplying facts)."

After the second diagnostic, require a reflection answering the following:

- Compare your two diagnostic results.
- Which concepts did you master between Test 1 and Test 2?
- Which concepts remain challenging?
- What study strategies were most effective?
- Did the GenAI analysis accurately identify your knowledge gaps?
- What additional resources did you need beyond the GenAI explanation?

Have the students turn in the diagnostic tests and the reflection. This assignment is AI-resistant because improvement between diagnostics requires actual learning that cannot be faked, and the reflection demands critical evaluation of GenAI-provided explanations against authoritative sources. Students who attempt to use GenAI to generate their reflection without engaging in actual study

will struggle to explain specific concept mastery or identify effective study strategies tied to their individual learning patterns.

G. Conclusion: From Surveillance to Processes

When GenAI entered law school classrooms in late 2022, the initial faculty response was predictable. Misunderstanding. Fear. Prohibition. Detection. And when the articles came pouring out about lawyers misusing GenAI and being sanctioned, heels dug in.⁶⁹

Schools and faculty drafted policies threatening academic sanctions. GenAI was (and is) viewed as a cheating problem to be solved through enforcement mechanisms borrowed from plagiarism protocols. This approach was understandable. It was also fundamentally unsustainable.

The reality, as this chapter has hopefully demonstrated, is that GenAI in legal education requires a shift from surveillance to systems. Detection tools cannot reliably distinguish human from AI-generated text, particularly in legal writing that values precision and standardized formulations. When the tools do flag a match, the results skew predictably against non-native English speakers and neurodivergent writers. Prohibition policies cannot account for the fact that your students will practice law in a profession where GenAI tools are rapidly becoming standard practice. Reactive enforcement cannot substitute for thoughtful pedagogical design.

Law schools need compliance, policy, and enforcement to function as a single system, not three separate checkboxes.

Once they enter practice, your students will be bound by their state's version of A.B.A. Model Rule 1.1: Competence.⁷⁰ Under Comment 8, lawyers must "maintain ... knowledge and skill ... including the benefits and risks associated with relevant technology."⁷¹ Forty states, the District of Columbia, and Puerto Rico adopted the duty of technology competence.⁷² These are not abstract aspirational goals. They create specific obligations to prepare students for a profession where GenAI competence is no longer optional.

⁶⁹ Charlotin, *supra* note 8.

⁷⁰ MODEL RULES OF PRO. CONDUCT r. 1.1., (A.B.A. 2023)

⁷¹ *Id.*, cmt. 8.

⁷² 40 States, D.C. and P.R. Have Adopted the Duty of Technology Competence, LAW SITES, <https://www.lawnext.com/tech-competence>, (last accessed Dec. 30, 2025).

In July 2024, the ABA released Formal Opinion 512, which addresses the use of GenAI by lawyers. It states, “To competently use a [Gen]AI tool in a client representation, lawyers need not become [Gen]AI experts. Rather, lawyers must have a reasonable understanding of the capabilities and limitations of the specific [Gen]AI technology that the lawyer might use. This means that lawyers should either acquire a reasonable understanding of the benefits and risks of the [Gen]AI tools that they employ in their practices or draw on the expertise of others who can provide guidance about the relevant [Gen]AI tool’s capabilities and limitations.”⁷³ The ABA also recognizes that, GenAI “tools may assist lawyers in tasks such as legal research, contract review, due diligence, document review, regulatory compliance, and drafting letters, contracts, briefs, and other legal documents.”⁷⁴

Your GenAI policy must acknowledge this reality. A blanket prohibition on GenAI use may feel safer in the short term, but it leaves your students unprepared for the technological demands of practice and potentially places your institution at odds with accreditation standards.

Your policy design establishes expectations within those boundaries. The spectrum from prohibition to transparency and integration gives you flexibility to match your institutional context, but the trend is clear. Legal employers are adopting GenAI tools at scale.⁷⁵ Law students are using these tools whether your policy permits it or not.⁷⁶ The question is not whether to engage with GenAI, but how to do so in ways that advance rather than undermine legal education. Effective policies provide clarity about permitted uses, require transparency about GenAI involvement, and establish assessment mechanisms that focus on legal reasoning rather than text production.

The gap between policy creation and effective implementation is not unique to law schools. Research across K-12 and higher education shows that institutions

⁷³ ABA Comm. on Ethics & Prof'l Responsibility, Formal Op. 512 (2024), *Generative Artificial Intelligence Tools*.

⁷⁴ *Id.*

⁷⁵ Clio, *GenAI Adoption in Law Firms: How Solo, Small, and Mid-Sized Firms Compare*, LAW TECHNOLOGY TODAY, A.B.A. (May 12, 2025), https://www.americanbar.org/groups/law_practice/resources/law-technology-today/2025/ai-adoption-in-law-firms-how-solo-small-and-mid-sized-firms-compare/, (last accessed Dec. 30, 2025).

⁷⁶ *Do ya'll use ChatGPT?*, Reddit, r/LawSchool, https://www.reddit.com/r/LawSchool/comments/1hf006j/do_yall_use_chatgpt/

frequently adopt GenAI policies that educators and students find unclear or unworkable.⁷⁷ Faculty report having policies they cannot explain to students.⁷⁸ Administrators mandate enforcement mechanisms that faculty know to be unreliable.

The solution is not better detection technology (please!). Relying on detection tools with error rates of 10-28% for human-written content⁷⁹ creates due process nightmares and disproportionately harms non-native English speakers.⁸⁰ Avoiding the GenAI question entirely leaves students unprepared for practice. Adopting policies without workable enforcement mechanisms undermines institutional credibility.

Better process verification is a much smarter approach. Mandatory revision histories provide evidence of authentic engagement with writing. GenAI Appendix requirements create transparency and accountability. Strategic assignment design makes GenAI shortcuts less effective than true legal analysis. Investigative conversations probe understanding rather than relying on probabilistic scores. While the approaches require more faculty time up front, they produce defensible determinations of academic misconduct and pedagogically valuable learning experiences.

⁷⁷ Lauraine Langreo, *Schools' GenAI Policies Are Not Still Not Clear to Teachers and Students*, EDWEEK (Jan. 30, 2025), <https://www.edweek.org/technology/schools-ai-policies-are-still-not-clear-to-teachers-and-students/2025/01>, last accessed Dec. 30, 2025.

⁷⁸ *Id.*, (reporting 40% of educators found their district GenAI policies unclear to them or their students).

⁷⁹ See Oleksandra Zhura, *How GenAI Detectors Actually Work: The Truth About Undetectable GenAI Tools*, LITERO (2025), <https://litero.ai/blog/how-ai-detectors-actually-work-the-truth-about-undetectable-ai-tools-2025-guide/> (documenting error rates of 10-28% for human-written content); see also *Why Don't GenAI Detectors Work?*, Ctr. for Integrated Prof. Dev., ILL. S., <https://prodev.illinoisstate.edu/instructional-resources/pedagogy/ai/detectors/>, (last accessed Dec. 30, 2025).

⁸⁰ See Weixin Liang et al., *GPT Detectors are Biased Against Non-native English Writers*, 4 PATTERNS 7, 3 (2023).

CHAPTER 6: The Data on Human-GenAI Learning

A. When GenAI Works

By now, you might be wondering whether any of what I've told you about how to integrate GenAI into legal education matters. Does it work? Will it improve student learning, or will it just create a generation of lawyers who can't think without a chatbot whispering in their ear?

Fair questions. To cut to the chase, when deployed thoughtfully under instructor guidance, GenAI can *significantly* improve learning outcomes. When students use it on their own without pedagogical guardrails and without sufficient training, the results range from "meh" to outright harmful.

But that's the simplified version, and you didn't pick up this book for oversimplifications. So let's dig into what the data show, with enough specificity that you can make informed decisions about your own courses.

B. The Good News: Yes, Students Learn More

1. AI Tutors Can Dramatically Outperform Traditional Instruction

In one of the most rigorous studies to date, conducted at Harvard University with 194 physics students, researchers built a custom GenAI tutor engineered to follow established pedagogical best practices. This was not a general-purpose chatbot but a purpose-built instructional tool with carefully designed system prompts and expert-written, step-by-step solutions embedded directly into the platform. Students using the GenAI tutor achieved learning gains more than double those of students in an active-learning classroom.¹ Students in the GenAI group scored a median of 4.5 on post-tests compared to 3.5 for active-learning students, with both groups starting from the same baseline of 2.75.²

Let me put those numbers in perspective. The median learning gain for GenAI-tutored students was 1.75 points; active-learning students gained 0.75 points. That's a 133% larger learning gain, not a marginal improvement. Seventy

¹ Greg Kestin et al., *AI Tutoring Outperforms In-Class Active Learning: An RCT Introducing a Novel Research-Based Design in an Authentic Educational Setting*, 15 SCI. REPORTS 17458, 17458 (2025).

² *Id.*

percent of GenAI-tutored students completed the material in fewer than 60 minutes, while the in-class group spent a full 75-minute period.³ The GenAI group learned more in less time.

Before you conclude that robots have replaced teachers, understand what made this study work and where it doesn't apply. The comparison condition was not passive lecture. It was peer instruction and small-group problem-solving, the same active-learning tradition that defines strong STEM pedagogy. And yet, GenAI outperformed a manifestly good classroom. The researchers also acknowledge that their findings apply most cleanly to first-exposure learning at the introductory level explicitly noting that the results may not hold for complex synthesis and higher-order critical thinking.⁴ That exception matters for legal education, where analysis, synthesis, and policy reasoning are the whole point. For upper-level students working through nuanced doctrine, a well-designed GenAI tutor supplements but does not replace the kind of thinking that only comes from sustained engagement with hard problems. For 1Ls, however, the fit is more direct. Learning the language of the law, what a holding is, how procedural posture shapes analysis, why "it depends" is the beginning of an answer rather than the end of one, is precisely the kind of first-exposure, scaffolded learning the Harvard study was designed to support.

What the researchers built also matters. They didn't hand students ChatGPT subscriptions and hoped for the best. The system facilitated active learning, managed cognitive load, promoted a growth mindset, and delivered accurate explanations through carefully scaffolded, sequential problems. The GenAI was designed to release only one step at a time, never reveal the complete solution, and require students to attempt an answer before confirming it. Just like the CustomGPTs (aka "bots") I write, they taught without any semblance of spoon-feeding. For law professors, the equivalent is a GenAI Socratic partner that asks follow-up questions instead of outlining arguments and refuses to hand over the completed analysis even when students ask directly because students must identify the rule before the tool will discuss its application. Chapter 10 explains how to build exactly that.

³ *Id.*

⁴ *Id.*

2. Students Report Higher Engagement and Motivation

Here's where it gets interesting. Students didn't just learn more with the GenAI tutor. They *liked it more*. When asked to rate their experience on a five-point scale, GenAI-tutored students reported significantly higher levels of engagement and motivation than did their in-class counterparts.⁵ They enjoyed the learning process more. They felt more confident that they could master difficult material with enough effort.

For law professors who've spent decades perfecting Socratic dialogue, this might sting a bit. But remember what the GenAI tutor was doing. It was providing immediate, personalized feedback to every student simultaneously. It adapted to each student's pace. It answered questions the moment the questions arose, without judgment, without making anyone feel stupid for asking a “dumb question.” In traditional instruction, only one student at a time can have that experience. While we hope the other students are learning, too, it's very different from being on the hot seat.

I saw the same pattern occur when I worked with Professor Cheryl Wattley on a small rollout in her criminal law class. Professor Wattley wanted students to both read and understand an excerpt from H.L.A. Hart's article “The Aims of the Criminal Law.” The CustomGPT that I built questioned students on the material. Here are some of their responses:

In the exercise with the bot, I had a conversation discussing the three elements Hart identifies, and how they connect to his broader theory about condemnation. Through this assignment, I also pushed further on his concern with the shift from “punishment” to “treatment.” The discussion helped me see more clearly that Hart isn't just defining criminal law mechanically; he's defending the importance of its moral dimension. Working through these questions step by step deepened my understanding of how community condemnation shapes both the justification and the dangers of criminal punishment.

I enjoyed the bot exercise because it allowed me to look at the excerpt in a new way. Some of the questions it was asking me to

⁵ *Id.*

think about allowed me to put my thoughts and understanding into words. This caught me a little off guard and made me re-read sections I thought I had a better understanding of.... One thing I wish is that I could ask more questions! ... I would have gone back and forth for ages if I could. I really liked getting questions and direct feedback on my answers.

The bot kept asking me to explain what I thought before giving me more information, which forced me to really think.

The CustomGPT created individualized learning experiences that would be impossible for Professor Wattlely to replicate in a class with sixty students. Each student received tailored Socratic questioning that pushed them to articulate their thinking and engage with Hart's theoretical framework at their own pace. The bot didn't simply provide answers. It was specifically instructed not to. It required students to explain their reasoning before offering guidance, mimicking the kind of back-and-forth dialogue for each student. Students could revisit sections they misunderstood, and work through complex ideas iteratively, without feeling judged. And they received immediate feedback on their interpretations. This type of sustained intellectual dialogue transforms reading from a passive exercise into active engagement, helping students develop the critical thinking skills essential for legal analysis while giving them the kind of personalized attention that's structurally impossible in large lecture courses. In short, CustomGPTs scale.

3. Essay Quality Improves Dramatically

Another study examined 117 graduate students working on English writing tasks. Students who used ChatGPT (*on their own* without support from faculty) as a revision tool showed significantly higher essay score improvements than those using human expert tutors, peer checklists, or no support at all.⁶ The GenAI group improved their essays by an average of 3.6 points on a 20-point scale (an 18% improvement), compared to 1.48 for human expert users (7.4%) and 1.40 for checklist users (7%).⁷

⁶ Yizhou Fan et al., *Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance*, 56 British J. of Ed. Tech 2, 489-530 (Dec. 10, 2024).

⁷ *Id.*

Statistical analysis confirmed the differences were significant. An analysis of variance (ANOVA) revealed $F=4.549$, $p=0.005$, with an effect size of $\eta^2=0.108$, meaning that roughly 11% of the variance in score improvements could be explained by which support tool students used.⁸ Post-hoc pairwise comparisons⁹ showed the GenAI group's advantage held up against each of the other groups individually, with adjusted p-values all below 0.05.¹⁰

Before you conclude that GenAI has surpassed human tutors, hold on. Students using ChatGPT on their own produced better *essays*, but they didn't show better *understanding* on knowledge transfer tests.¹¹ When tested on their ability to apply concepts in new contexts, the GenAI group performed no better than students who'd received other forms of support. Their knowledge gain scores showed no significant differences ($F=0.053$, $p=0.984$), and neither did their knowledge transfer scores ($F=0.019$, $p=0.996$).¹²

In plain English, when using GenAI on their own, students could polish prose like professionals, but when tested on whether they'd actually learned the underlying concepts, they performed no differently from those students who'd gotten human help. They could fix their work product without necessarily improving their underlying competence. Don't forget, **this was not a study on instructor led or designed GenAI use.**

For law faculty, this finding should set off alarm bells. Students using GenAI on their own without instructor guidance (as many of them are) can produce a polished brief with GenAI assistance, but they can't explain the doctrine.

The distinction between these two findings in these studies is significant. Students given *unstructured access* to a general-purpose GenAI tool learned to produce better-looking work without developing better understanding. Students working with a *purpose-built, pedagogically designed* GenAI tutor learned significantly more in less time. A CustomGPT built by a faculty member who

⁸ *Id.*

⁹ *Id.*, (post-hoc pairwise comparisons test each pair of groups individually after an initial analysis confirms that some difference exists among them. Because running multiple comparisons increases the chance of a false positive, post-hoc tests apply statistical corrections to keep that risk honest).

¹⁰ *Id.*, (a p-value represents the probability that an observed result occurred by chance; the conventional significance threshold is $p < 0.05$. "Adjusted" means that threshold has been corrected to account for the number of comparisons being run. All adjusted p-values falling below 0.05 means the GenAI group's advantage was statistically significant even after applying that correction).

¹¹ *Id.*

¹² *Id.*

understands what first-year law students need to learn, and who has engineered it to teach rather than to answer, is closer to the Harvard tutor than to a student pasting an essay prompt into ChatGPT at midnight.

Hold this thought. We'll come back to it.

4. Weaker Students Benefit Most from Well-Designed Systems

One of the most encouraging findings emerges from a study of English language learners in Italian middle and high schools. Researchers found a negative correlation between initial test scores and learning gains. Students who started with lower scores showed significantly greater improvement than stronger students. The correlation was -0.777 ($p < 0.001$) for the treatment group, meaning weaker students benefited substantially more from the GenAI tutoring.¹³

In statistics, correlations range from -1 to $+1$. A correlation of -0.777 is very strong and indicates that for every standard deviation decrease in initial ability, learning gains increased by about three-quarters of a standard deviation. In practical terms, the students who struggled most at the beginning made the biggest gains. The GenAI tutoring didn't just help everyone equally. It helped the struggling students *more*.¹⁴

Another study using "Tutor CoPilot" found similar results in a different context. Researchers conducted a randomized trial with 1,800 K-12 students and found that lower-rated human tutors and less-experienced instructors saw the largest gains when supported by GenAI. Students of the less-effective human tutors improved their mastery rates by up to 9 percentage points (a 15% relative improvement over control), while students of higher-rated human tutors saw smaller improvements of about 2 percentage points.¹⁵ The GenAI essentially helped novice instructors perform like experts.

¹³ Alessandro Vanzo, *GPT-4 as a Homework Tutor Can Improve Student Engagement and Learning Outcomes*, ACL ANTHOLOGY, (Sept. 2024), <https://aclanthology.org/2025.acl-long.1502.pdf>, (last accessed Jan. 11, 2026).

¹⁴ *Id.*

¹⁵ Amanda Chong et al., *Tutor CoPilot: Leveraging Generative AI to Improve Novice Tutor Performance at Scale*, PROCEEDINGS OF THE 2024 AAAI CONFERENCE ON AI & Education (2024); see also Rose E. Wang, et al., *Tutor CoPilot: A Human-AI Approach for Scaling Real-Time Expertise*, NAT. STUDENT SUPPORT ACCELERATOR, (finding, in a randomized controlled trial of over 700 tutors and 1,000 students from underserved communities, that an AI tutoring assistant increased student mastery rates by 4 percentage points overall ($p < 0.01$) and 9 percentage points for students working with

For law schools, where we often struggle to provide sufficient support for students who arrive underprepared and without strong academic skills, these findings are significant. Well-designed GenAI tutoring systems might help close achievement gaps rather than widen them. But notice the condition: *well-designed systems with pedagogical guardrails*. We'll come back to what happens without those guardrails.

Where do these tutoring systems come from? From you, the professor. In Chapter 10, I walk you through how to create these “CustomGPTs,” and in Appendix I, I provide templates and examples.

5. GenAI Tutoring Works Even in Resource-Constrained Settings

Before you assume that GenAI tutoring only works for privileged Harvard undergrads with reliable internet and new MacBooks, let's look at data from Nigeria. A World Bank study conducted a randomized controlled trial with 657 secondary students in Benin City using Microsoft Copilot for English language learning over six weeks.¹⁶ These weren't wealthy private schools. Students faced regular power outages and internet disruptions. Many were using computers for the first time. Yet the intervention still produced dramatic results.

Students assigned to the GenAI tutoring program scored 0.31 standard deviations higher on the final assessment than control students.¹⁷ For English specifically (the main outcome of interest), the effect was 0.23 standard deviations.¹⁸ Perhaps more impressive, students also performed significantly better on their regular third-term English exam, which covered the entire year's curriculum, not just the six-week intervention period. The effect there was 0.21 standard deviations.¹⁹ Allow me to translate what those effect sizes mean in concrete terms. **The researchers calculated that the six-week program generated learning gains equivalent to 1.5 to 2 years of regular schooling in Nigeria.**²⁰

lower-rated tutors. The AI system promoted more effective pedagogy by increasing Socratic questioning and reducing generic praise, at a cost of approximately \$20 per tutor annually).

¹⁶ Martín De Simone et al., *From Chalkboards to Chatbots: Evaluating the Impact of Generative AI on Learning Outcomes in Nigeria*, World Bank Policy Research Working Paper 11125 (2025).

¹⁷ *Id.*

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ *Id.*

Students learned more in six weeks of GenAI-supported instruction than they typically learn in 18-24 months of traditional classroom teaching. The cost-effectiveness analysis showed the program achieved 3.2 equivalent years of schooling per \$100 invested, ranking it among the most cost-effective educational interventions documented in low- and middle-income countries.²¹

The program met twice weekly for 90-minute sessions. Students worked in pairs at shared computers, engaging with the GenAI chatbot under teacher supervision. Attendance averaged 72%.²² Analysis revealed a strong dose-response relationship: each additional day of attendance was associated with approximately 0.031 standard deviations of improvement in learning outcomes.²³ Extrapolating conservatively from this linear relationship, researchers estimated that a full academic year of the program (with the observed 72% attendance rate) would produce learning gains of approximately 1.55 standard deviations.²⁴ Even under a pessimistic scenario of only 50% attendance, the estimated gains would still exceed 1.2 standard deviations.²⁵

So, what made this approach work when so many ed-tech interventions fail? Three critical design features:

- First, teachers received training on how to use GenAI and played active supervisory roles. They didn't provide direct instruction, but they circulated during sessions, ensured students stayed on task, and guided GenAI interactions. The intervention didn't replace teachers. It amplified their capacity to provide personalized support.
- Second, the prompts were carefully engineered. Researchers designed initial prompts based on principles from the science of learning, incorporating strategies like retrieval practice, elaborative interrogation, and concrete examples. The prompts positioned the GenAI as a tutor providing scaffolding rather than a homework-completion service that generates answers. When students asked for solutions, the GenAI responded with Socratic questions: "What have you tried? What's confusing you? What strategy might help?"

²¹ *Id.*

²² *Id.*

²³ *Id.*

²⁴ *Id.*

²⁵ *Id.*

- Third, the content aligned with the Nigerian national curriculum. Each session focused on a specific curriculum topic that students were covering in their regular classes. The GenAI tutoring complemented rather than replaced classroom instruction.²⁶

The heterogeneity analysis revealed interesting patterns. Female students showed particularly strong gains, with an interaction effect of 0.420 standard deviations.²⁷ This essentially compensated for their lower baseline performance, suggesting that GenAI tutoring might help close gender gaps. Students with higher baseline performance also benefited more, showing an interaction effect with prior test scores that was positive and significant at the 5% level.²⁸ Similarly, students from higher socioeconomic backgrounds experienced larger treatment effects (interaction coefficient of 0.113, $p < 0.05$).²⁹

Treatment effects were positive and statistically significant across all quantiles of the outcome distribution.³⁰ Every student group benefited. The differential effects likely reflect that students from poorer households often experienced using computers for the first time during the intervention. Initial unfamiliarity with technology moderated the magnitude of impact, but these students still showed substantial gains relative to not participating at all.

For legal education, several lessons emerge.

- First, GenAI tutoring can work in challenging conditions. If it works in Nigerian schools with frequent power outages and students new to computers, it can certainly work in American law schools.
- Second, teacher oversight matters enormously. The Nigerian intervention succeeded partly because trained teachers supervised every session.
- Third, carefully designed prompts that force engagement rather than enable shortcuts are essential.

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.*

²⁹ *Id.*

³⁰ *Id.*

- Fourth, alignment with curriculum is critical. The GenAI tutoring worked because it reinforced rather than replaced what students learned in regular classes.

The cost analysis is particularly striking for legal education. At \$48 per student for the six-week program, or \$124 per student for a full academic year, this intervention cost a fraction of what law schools spend on traditional instruction and academic success programs.³¹ Even accounting for higher costs in U.S. contexts, the economics are compelling. The cost-effectiveness analysis showed learning gains equivalent to 1.5 to 2 years of “business-as-usual” schooling, placing the program among the most cost-effective interventions for improving student outcomes.³²

Imagine adapting this model for first-year law students struggling with legal writing. Twice-weekly 90-minute sessions working with a GenAI tutor specifically prompted to teach IRAC, CRAC, or CREAC analysis. They would have teacher-designed initial prompts for each session, focusing on different doctrinal areas. Faculty members circulate among the students to supervise, to ensure that the students engaged productively with the GenAI tutor, and to prevent cognitive shortcuts. The teachers’ assessments measure not just writing quality but real understanding through transfer tests requiring application in novel contexts.

A model similar to this already exists for this and it’s called Write.law, developed by Professor Joseph Regalia at the UNLV William S. Boyd School of Law.³³ Write.law has been adopted by scores of AmLaw 200 and other law firms (including DLA Piper LLP, Venable LLP, Vinson & Elkins LLP, and Baker & Hostetler LLP), government agencies (such as the Administrative Office of the U.S. Courts, the European Patent Office, the U.S. Department of Health and Human Services, and the U.S. Department of Homeland Security), companies (CS DISCO, Inc., State Bar of Georgia, LSAC, The Legal Aid Society), and (unfortunately only) a handful of law schools.³⁴

³¹ *Id.*

³² *Id.*

³³ Write.law, <https://write.law/> (last visited Feb. 23, 2026).

³⁴ See *Select Clients and Partners*, write.law, <https://write.law/clients-and-partners> (last visited Feb. 23, 2026). I have had an individual license almost since write.law was launched, and demo it to my students. Write.law also has Ace.law to help students study for exams.

The Nigeria study proves that thoughtfully designed GenAI tutoring, with appropriate pedagogical guardrails and teacher oversight, can produce transformative learning gains even in resource-constrained settings. The question isn't whether GenAI tutoring can work. The question is whether we'll implement it properly or botch it through neglect by leaving students to their own devices (pun intended).

6. GenAI in the Law School Classroom: What We Know

The most productive body of empirical research on GenAI in law schools comes from a series of studies by Jonathan Choi, Kristin Hickman, Amy Monahan, and Daniel Schwarcz at the University of Minnesota Law School. Read together, the three studies tell a coherent and instructive story.

The first study asked the simplest version of the question: how well can GenAI perform on law school exams without human assistance? Over 95 multiple-choice questions and 12 essay questions across four courses, ChatGPT performed on average at the level of a C+ student, achieving a low but passing grade in all four courses.³⁵ The performance was uneven. ChatGPT was stronger on doctrinal summary than on nuanced application. But the finding alone was enough to reshape the academic integrity conversation in legal education.

The second study moved from AI-alone to AI-assisted performance, running the first randomized controlled trial on GenAI assistance in law school. Access to GPT-4 only slightly and inconsistently improved the quality of participants' legal analysis but induced large and consistent increases in speed. GenAI assistance improved the quality of output unevenly — where it was useful at all, the lowest-skilled participants saw the largest improvements.³⁶ The speed finding held regardless of baseline skill level. The quality finding did not: stronger students saw less benefit, weaker students saw more. For legal educators, that asymmetry carries significant pedagogical implications. GenAI may function as a floor-raiser more than a ceiling-raiser.

The third study refined the analysis further, examining specific legal tasks and prompting techniques. Participants who were provided with access to

³⁵ Jonathan H. Choi, Kristin E. Hickman, Amy B. Monahan & Daniel Schwarcz, *ChatGPT Goes to Law School*, 71 J. Legal Educ. 387 (2022).

³⁶ Jonathan H. Choi & Daniel Schwarcz, *AI Assistance in Legal Analysis: An Empirical Study*, 73 J. Legal Educ. 384 (2025).

ChatGPT completed their writing tasks faster and produced higher-quality work than participants who were not provided access to the tool.³⁷ The study also found that the impact varied significantly depending on the nature of the task and the sophistication of the prompting.

Taken together, the Minnesota studies support a few conclusions that should inform how law faculty design courses and assignments. GenAI performs adequately on doctrinal recall. It provides meaningful speed gains on legal writing tasks. Its quality improvements are most pronounced for lower-performing students. And unguided use does not reliably produce mastery of the underlying analysis, which is the finding that connects directly to the earlier discussion of the Harvard physics study and the distinction between polished output and genuine understanding.

C. The Bad News: The “Crutch Effect” Is Real and Reproducible

Now for the bad news.

1. Unsupervised Use Creates Measurable Dependency

Stanford researchers conducted a large, randomized trial with 5,831 students in an introductory programming course. When high school and adult learners were given access to ChatGPT *without instructional scaffolding*, they improved their practice problem performance by 48% compared to peers working without GenAI.³⁸ Sounds fantastic, right? Students were solving problems faster and getting better scores on homework.

But, when the GenAI was removed and students took a final exam on their own, those same students performed 17% worse than the control group who'd never used GenAI at all.³⁹ Let me repeat that, because it's the single most important finding in this entire chapter. Students with unguided GenAI access didn't just lose their performance boost when the GenAI was removed. They

³⁷ Jonathan H. Choi, Amy B. Monahan & Daniel Schwarcz, *Lawyering in the Age of Artificial Intelligence*, 109 Minn. L. Rev. (forthcoming 2024).

³⁸ Hamsa Bastani, et al., *Generative AI Without Guardrails Can Harm Learning: Evidence from High School Mathematics*, 122 PROC. NATL. ACAD. SCI. U.S.A. 122 (26) e2422633122, (2025), <https://doi.org/10.1073/pnas.2422633122>, last accessed Jan. 11. 2026).

³⁹ *Id.*

actually performed worse than students who'd never had access to GenAI in the first place.

The study found that only 14.2% of students given GenAI access actually used it, but among those who did use it (the "adopters"), the effects were pronounced. The estimated effect size for exam performance among adopters was -0.40 standard deviations,⁴⁰ meaning the average GenAI-using student scored nearly half a standard deviation below where they would have scored without GenAI access.

The GenAI acted as a cognitive crutch. While leaning on it, students moved faster and solved more problems. But they weren't building the neural pathways required for independent problem-solving. When the crutch was removed, their performance fell flat.

We know students are already using GenAI on their own. You might think that law students arrive knowing how to prompt. They don't. Universities and colleges are struggling to build AI Fluency into curricula just as much as law schools are. That fact and this study underscore why it's incumbent upon us to teach them how to use it properly and accurately. We cannot avoid it. For law faculty, imagine this scenario: Your students use GenAI to draft case briefs all semester. They produce beautiful work. Every brief is polished, well-organized, complete. Then the bar exam arrives, and suddenly they can't distinguish a holding from dicta without a chatbot.

2. Metacognitive Laziness Takes Hold

The essay quality study that I mentioned earlier uncovered something more insidious than simple dependency. Researchers examined students' self-regulated learning processes using screen recordings and trace data. They found that students using ChatGPT engaged in significantly fewer metacognitive activities than those working with human tutors or checklists.⁴¹ Specifically, ChatGPT users showed reduced evaluation, monitoring, and orientation processes.

What does that finding mean? Evaluation involves assessing whether your understanding is accurate. Monitoring involves checking your progress toward

⁴⁰ *Id.*

⁴¹ Fan, *supra* note 6.

goals. Orientation involves figuring out what you need to do next.⁴² When students used ChatGPT, they simply stopped doing these things. Instead of evaluating their own understanding, they asked the GenAI. Instead of monitoring their progress, they relied on the GenAI to tell them when they were done. Instead of orienting themselves within the task, they waited for the GenAI to provide the next step.

Researchers coined a term for it: "metacognitive laziness."⁴³ Students (and all users of GenAI) can become overly reliant on GenAI assistance and offload their metacognitive responsibilities to the machine.⁴⁴ They stop thinking about their thinking. They stop monitoring whether they understand. They stop evaluating whether their approaches make sense. Process mining analysis showed that students in the GenAI group had a closed loop between ChatGPT interactions and revision tasks, with far fewer transitions between metacognitive planning, monitoring, and evaluation activities.⁴⁵

The psychology here maps onto established research on cognitive offloading. When people can delegate cognitive tasks to external tools, they often do so to reduce mental effort.⁴⁶ Normally, experiencing difficulty or "cognitive disfluency" triggers deeper analytical reasoning. When students bypass that struggle by asking GenAI for answers, they never activate those deeper metacognitive processes.⁴⁷ Research shows that "in the context of GenAI, if learners rely excessively on AI-generated outputs or facilitation, they might not experience the necessary disfluency or cognitive difficulty to trigger these deeper metacognitive processes. Consequently, learners might default to less effortful, more intuitive decision-making, reinforcing a state of metacognitive laziness."⁴⁸

In other words, relying on GenAI to avoid cognitive difficulty is like a woman watching skiers from the lodge sipping on a cocktail and thinking she learned how to ski only to get a rude awakening once she hits the slopes.⁴⁹ Because

⁴² *Id.*

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ Evan F. Risko & Sam J. Gilbert, *Cognitive Offloading*, 20 TRENDS IN COGNITIVE SCIENCES 676 (2016).

⁴⁷ Fan, *supra* note 6.

⁴⁸ *Id.*

⁴⁹ The woman was me.

the student isn't experiencing the disfluency of carving through resistance (the burning legs, the wobbling, the falls) her muscles never receive the signal to grow stronger. She may feel like a skier, but she didn't develop the balance or competence required to get down the mountain.

In legal education and practice, metacognition is a professional survival skill. Law students must constantly interrogate their own reasoning, checking whether they understand a statute and whether their argument holds under pressure before a professor or senior partner. Students who develop metacognitive laziness in law school arrive in practice waiting to be told what to do next, unable to catch the flaws in their own analysis before someone else does. Without the habit of stepping back and asking whether their approach makes sense, they will always be one step behind.

3. When the Crutch Becomes the Problem

The most sobering study in this literature comes not from a law school but from a high school in Turkey, with implications for legal education. Researchers from the Wharton School conducted a large-scale randomized controlled trial with nearly a thousand high school math students, deploying two versions of a GPT-4 tutor: a standard ChatGPT-style interface they called GPT Base, and a pedagogically engineered version with guardrails called GPT Tutor.⁵⁰ The students did not receive human instruction.

The results confirmed what the Harvard physics study suggested. During assisted practice sessions, both tools improved student performance substantially – GPT Base by 48% and GPT Tutor by 127%, relative to students working with only textbooks and notes.⁵¹ So far, so promising.

Then the researchers took the tools away.

On the subsequent unassisted exam, students who had used GPT Base performed 17% worse than students who had never had GenAI access at all.⁵² Students who used GPT Tutor performed no better or worse than the control group. The guardrails neutralized the harm; they did not produce learning gains.

⁵⁰ Bastani, *supra* note 38.

⁵¹ *Id.* at 3.

⁵² *Id.* at 4.

The unguarded tool actively degraded student performance compared to doing nothing.

How come? Students using GPT Base most commonly sent one type of message: asking for the answer. They copied the solution and moved on, learning nothing. Students using GPT Tutor, by contrast, had to wrestle with the problem before the tool would move them forward.⁵³ The tool's design determined whether students were thinking or just collecting output.

The researchers are explicit that even a well-designed GenAI tutor is a *supplement* to instruction, not a replacement for it. GPT Tutor remained passive, responding when asked but not proactively engaging students or uncovering misconceptions the way an effective human tutor would.⁵⁴ The study's authors recommend a flipped-classroom model in which AI handles first-exposure learning, freeing class time for the higher-order work that only human instruction can reliably develop. That framing matters for how faculty deploy these tools, and it matters equally for how faculty explain them to students. A student who understands that the GenAI tutor is scaffolding (a floor to build from, not a ceiling to rest on) will use it differently than one who treats it as a shortcut. Telling students explicitly that the tool is designed to support their thinking, not replace it, is part of the pedagogy.

There is one more finding that belongs in any honest account of this study. Students in the GPT Base arm did not *perceive* that they had learned less. They believed they had performed well. The mismatch between perceived and actual learning is a known phenomenon in educational psychology, but it takes on particular urgency here.⁵⁵ A student who copies solutions from a GenAI tool and walks away feeling confident has not just failed to learn. They have acquired false confidence about the extent of their own understanding. For law students, who will eventually face clients, courts, and supervising partners, that's a significant gap.

The study's core finding should serve as the empirical anchor for every GenAI policy in legal education: the tool's design determines whether students learn or merely perform. A carefully engineered CustomGPT that requires

⁵³ *Id.* at 5-6.

⁵⁴ *Id.* at 6.

⁵⁵ *Id.* at 4.

students to identify issues before it explains rules, and that refuses to write the analysis even when students ask, is closer to GPT Tutor than to GPT Base. A student who pastes a fact pattern into a standard ChatGPT interface and copies the output is doing exactly what the Turkish high school students did and will perform exactly as they did when the tool is gone.

D. The Depth-Volume Tradeoff

Studies consistently show that students using GenAI in *substitutive* ways (having it generate solutions) move faster through material and cover more topics. However, speed comes at the cost of understanding depth.⁵⁶ In the graduate writing study, students using ChatGPT produced more revisions and covered more aspects of their essays but showed no improvement in conceptual understanding on transfer tests.⁵⁷

Conversely, when GenAI is used as an instructor-led tutor providing explanations rather than answers, students don't increase volume but achieve significantly deeper understanding. The Harvard physics study showed that students receiving GenAI explanations demonstrated better understanding of underlying principles, even though they didn't cover more problems than the control group.⁵⁸

Would you rather your students speed-run through 30 cases with shallow comprehension, or deeply engage with 15 cases and truly understand the doctrinal nuances? Surface-level familiarity with many cases *might* help on a multiple-choice exam, but it won't help when a partner asks to distinguish the precedent you've cited from the opposing party's case.

The data suggests you can have volume or depth, but *unsupervised* GenAI use pushes students toward volume every time. Students optimize for task completion, not understanding. After all, they're busy with many competing demands on their time,

⁵⁶ Matthais Lehman, *AI Meets the Classroom: When do Large Language Models Harm Learning?*, (2025), 10.13140/RG.2.2.34417.60001, https://www.researchgate.net/publication/389635997_AI_Meets_the_Classroom_When_Do_Large_Language_Models_Harm_Learning, (last accessed Jan. 11, 2026).

⁵⁷ Fan, *supra* note 6.

⁵⁸ Kestin, *supra* note 1.

E. What Makes the Difference: The Power of Pedagogical Design

The pattern in the research is unmistakable. GenAI isn't inherently helpful or harmful. The pedagogical design determines everything. Let me offer some ideas that separates success from failure.

1. Guardrails Transform Outcomes Completely

Remember those high school math students whose performance tanked by 17% when GenAI was removed? The same research team ran a parallel condition. They created a "GPT Tutor" version with pedagogical guardrails.⁵⁹ Instead of giving direct answers, the system provided teacher-designed hints and Socratic guidance. Students using this scaffolded version improved their practice problem performance by 127% (compared to 48% for unguided ChatGPT) and, did not experience the 17% performance penalty on unassisted exams. Their independent performance was statistically equivalent to students who had never used GenAI at all.⁶⁰

The “guardrailed” version produced better practice performance AND better independent performance. Yet, students didn’t retain the information.

The GPT Tutor in the Bastani study relied on teacher-provided hints and problem-specific inputs written by hired math teachers working part-time, which was a labor-intensive process the authors acknowledge explicitly.⁶¹ The guardrails were also largely passive: the tool responded when asked but did not proactively engage students or adapt to individual reasoning patterns. It did not replicate Socratic instruction.

A well-designed CustomGPT built by a law faculty member using the methods in this book can do more than GPT Tutor did. By embedding course-specific learning objectives, requiring students to identify the issue before the tool engages with the rule, and programming the tool to push back on conclusory analysis rather than simply providing hints, faculty can build something that more closely approximates the proactive, adaptive quality the Bastani researchers identified as the next frontier for GenAI tutoring. The researchers themselves

⁵⁹ Bastani, *supra* note 38.

⁶⁰ *Id.*

⁶¹ *Id.* at 2.

noted that GPT Tutor "remains passive" and that combining pedagogical scaffolding with proactive engagement is the direction the education needs to move.⁶² The CustomGPT approach described in Chapter 10 is designed to do exactly that.

2. GenAI Co-Pilots For New and Adjunct Faculty

As discussed above, randomized trial with 1,800 K-12 students and 69 tutors tested "Tutor CoPilot," a system providing real-time suggestions to tutors during one-on-one sessions. Students whose tutors had access to Tutor CoPilot were four percentage points more likely to master lesson content (a 7% relative improvement).⁶³ Analysis revealed that treatment effects were largest for lower-rated tutors (nine percentage points) and smallest for higher-rated tutors (two percentage points), suggesting the GenAI helped less-experienced instructors adopt expert teaching strategies.

Linguistic analysis of 241,005 tutor messages showed that tutors using Tutor CoPilot were significantly more likely to "prompt students to explain" their reasoning (10% increase in usage), which is an element of expert pedagogy.⁶⁴ Meanwhile, control tutors were more likely to "encourage students in a generic way" (10% increase), which sounds nice but doesn't actually push cognitive engagement.⁶⁵ The GenAI amplified teachers' capacity to employ high-quality pedagogical strategies.

For law schools, where adjunct faculty and less-experienced instructors teach substantial portions of the curriculum, a well-designed CustomGPT could democratize pedagogical expertise. Why not use a CustomGPT that reviews a class audio transcript, slides, etc., and identifies missed teaching opportunities. Newer faculty often answer student questions directly rather than working through them. I call the CustomGPT "LectureLens." It analyzes your lecture by reading your transcript (or "listening" to your audio or "watching" your video and provides feedback. It's grounded in law review articles about best practices in legal education. The same type of tool that helps students think more rigorously could

⁶² *Id.* at 6.

⁶³ Wang, *supra* note 15.

⁶⁴ *Id.*

⁶⁵ *Id.*

help instructors teach more rigorously, and in institutions that rely heavily on practitioners who are experts in their fields but not in teaching, it could make a profound difference in the quality of adjunct and new faculty teaching. An example of this is found in Appendix G.

Note that neither LectureLens, nor any of the other CustomGPTs or Claude Artifacts are not products for sale from a GenAI company. They are something that I built and share with you as colleagues as an example of what CustomGPTs and Claude Artifacts can do. It is by no means a sales pitch.

The beauty of this approach is that it doesn't replace instructor judgment. The human decides whether to follow the GenAI's suggestion. But for less-experienced faculty members, having expert pedagogical suggestions made privately can dramatically improve both faculty performance and student outcomes. Let's face it, most law schools do not have the time or the funding to provide new and adjunct faculty training beyond this language is required in your syllabus and this is how to upload your course materials to Canvas. Some come from practice, and don't have an academic background. They teach how they were taught, which could have been many years ago. LectureLens could help supplement the support the law school can provide.

And it doesn't have to be limited to new and adjunct faculty members. Seasoned professors might be surprised and pleased by the suggestions offered by Instructor Tutor that are delivered in private.

3. Reflection and Revision Protocols Force Engagement

A study examined how students learn when collaborating with GenAI to generate feedback. In an online graduate data science course, researchers tested three approaches to having students write hints for incorrect programming solutions: students writing hints independently (Baseline), students writing hints with constant access to GPT-4-generated hints (AI-assistance), and students writing hints independently first, then revising them after viewing GPT-4 hints (AI-revision).⁶⁶ The GenAI-revision approach produced better learning outcomes than constant GenAI access and was nearly three times more likely to provide

⁶⁶ Anjali Singh et al., *The Impact of Student-AI Collaborative Feedback Generation on Learning Outcomes*, 1 PROC. OF MACHINE LEARNING RESEARCH 1:1-9 (2024).

specific, actionable feedback (55% vs. 19% for students with constant GenAI access, and versus 13.6% for students generating feedback alone).⁶⁷

Requiring students to attempt the task before engaging with GenAI-generated content prompted them to critically evaluate the GenAI's responses and refine their thinking, leading to deeper understanding. The findings emphasize that effective GenAI integration requires designs promoting active student engagement rather than passive reliance on GenAI output.⁶⁸

The magic wasn't the GenAI feedback itself. The magic was forcing students to grapple with problems independently before consulting GenAI, then requiring them to critically evaluate GenAI suggestions. Students couldn't passively accept GenAI output. They had to engage, compare, justify their decisions, and synthesize. The protocol forced metacognitive engagement that otherwise wouldn't happen.

Law professors can adopt similar protocols immediately. Don't let students use GenAI from the start of an assignment. Require them to draft complete first versions independently (with track changes on). Then permit GenAI access for revision but mandate a reflection memo explaining which GenAI suggestions they accepted, which they rejected, and why. Force them to justify every choice. Grade not just the final product but the decision-making process.

Better yet, require students to identify problems in their own work before consulting GenAI, then compare their self-assessment to the platform's feedback. Did the student catch the same issues? Did the GenAI catch something they missed? Why? This metacognitive comparison forces engagement with the revision process rather than blind acceptance of GenAI suggestions.

F. The Equity Wildcard: Why Context Matters

One of the most frustrating aspects of GenAI research is the inconsistent findings on equity. Some studies show GenAI closing achievement gaps. Others show it widening them. What explains the difference?

Students with stronger prior knowledge of the subject area can use GenAI as a complement to their learning. They recognize hallucinations and extract value from GenAI explanations. They can evaluate whether GenAI-generated content is

⁶⁷ *Id.*

⁶⁸ *Id.*

accurate and complete. Students with weaker foundations are more easily deceived by confident-sounding nonsense and more likely to use GenAI as a substitute for thinking.⁶⁹ They lack the knowledge base to evaluate GenAI output and, because it sounds so plausible, they accept it uncritically. And they lack GenAI literacy. Put simply, they do not know what they do not know.

One study found that students who used ChatGPT reported significantly higher negative outcomes in mathematics performance specifically because the GenAI only helped come up with solutions without providing analysis or explanations.⁷⁰ Students with weaker foundations in both the subject and GenAI couldn't fill in the gaps. They got answers but not understanding. Meanwhile, stronger students could use those solutions as starting points for deeper analysis.⁷¹

However, when GenAI systems are carefully designed with pedagogical guardrails and instructor oversight, the equity effects flip. It's early days and this anecdotal, by my students report that they finally understand a difficult subject because, while the professor was great, the CustomGPT was better. Why? Because it's available 24/7, because it will continue to help you for as long as you need, because it will give endless examples, and because it will give explanations in simpler to understand language. The Italian language learning study found that weaker students benefited more because the system provided the scaffolding they needed without enabling shortcuts. The correlation of -0.777 between initial ability and learning gains meant that struggling students made the biggest improvements. Stronger students still improved but didn't pull further ahead.⁷²

The lesson is clear. If you allow students to use ChatGPT freely without structure, you'll likely widen achievement gaps. Students who arrive at law school with strong analytical skills and solid writing abilities will extract more value from GenAI. They'll use it to polish already-good work and accelerate their learning. Students who arrive without those academic skills will use it as a crutch, producing work they don't understand, and falling further behind.

But if you thoughtfully integrate GenAI with guardrails, scaffolding, and explicit learning objectives, you might actually narrow achievement gaps. Provide struggling students with GenAI tutors that break down complex concepts into

⁶⁹ Lehman, *supra* note 56.

⁷⁰ *Id.*

⁷¹ *Id.*

⁷² Vanzo, *supra* note 13.

manageable pieces. Give them systems that force engagement rather than enabling shortcuts. Create structures that level up their skills rather than masking their deficits.

If I had the power, all law students would be exposed to CustomGPTs (or the equivalent) in the first year, maybe even at orientation. At a minimum, Academic Success and Bar Prep would adopt their use.

G. Duration and Context: When GenAI Stops Working

Here's something most discussions of GenAI in education miss entirely: duration matters. The beneficial effects of GenAI tutoring don't necessarily persist over long time periods. It certainly isn't a cure-all. It's a supplemental tool

The meta-analysis found that effect sizes varied dramatically by intervention duration. Studies lasting 2-7 weeks showed effect sizes around $g=0.46$. Studies lasting 3-6 months showed $g=0.39$. But studies lasting longer than six months showed effect sizes of only $g=0.15$.⁷³ Students' engagement with GenAI systems declines over time, and so do the learning benefits.

Student motivation, it turns out, has a shelf life. So, reliance only on CustomGPTs won't work. Students who were initially excited about personalized GenAI tutoring eventually find it tiresome and start looking for shortcuts rather than engaging with the material.⁷⁴ It's the nature of us humans. We're remarkably good at turning novelty into routine, and routine into something they'd rather avoid.

Even signed up for a gym membership? In January? Still going?

For legal educators, the short-term engagement data on GenAI tutoring tools is both promising and misleading. Studies conducted over a few weeks capture the novelty window, not the semester. Once the tool becomes routine, students relate to it the way they relate to the gym in March: they know it's good for them, they feel vaguely guilty about avoiding it, and they're not going. Plan for that. Build in variety and resist the pull toward making every assignment GenAI-mediated. Focused, spaced modules, where students engage with GenAI tools intensively for a defined period, then step away, work better than continuous exposure precisely because they don't give habituation a chance to set in. The

⁷³ Wang, *supra* note 15.

⁷⁴ *Id.*

platforms built on top of LLMs (Khanmigo, Duolingo's AI features, and the growing category of *commercial* GenAI tutoring tools) try to fight disengagement with gamification: streaks, badges, progress bars, the full casino reward-schedule playbook. It works briefly, and then students learn to game the streaks the same way they've learned to game everything else from their years and years of gaming. The underlying LLMs themselves offer none of that scaffolding. ChatGPT does not congratulate you for showing up three days in a row. There's no streak to protect, no badge for completing a practice hypo. Which means the only thing sustaining engagement with a raw LLM tutoring interaction is the interaction itself – and once the novelty wears off, they'll use them less.

But don't shy away from CustomGPTs either. My advice is to mix it up. Use GenAI in focused bursts rather than as a constant presence in the course. Cold calls and phones-down discussions still absolutely belong in the modern law school classroom. Variety keeps the GenAI interactions from becoming wallpaper. Students who run every assignment through a GenAI tool stop noticing it, the same way you stop noticing the hum of your refrigerator. Bring it back selectively where it makes sense (making difficult concepts understandable, practicing negotiations or client interviewing, using a Socratic tutor). The professors who report the best results treat GenAI as one tool among several they use in their courses.

H. Practical Takeaways: Six Evidence-Based Principles

Here are some of the most salient points from the research that can help implement what works.

1. Never Deploy GenAI Without Pedagogical Guardrails

If you're going to allow or encourage GenAI use, build structure around it. Require drafts before students access GenAI. Mandate reflection on GenAI suggestions. Design prompts that prevent students from getting complete answers, (i.e., instruct the GenAI to coach, not to spoon-feed). The difference between a 48% improvement with subsequent 17% decline versus a 127% improvement with no decline was simply the presence of guardrails. That's the difference between harming students and helping them.

Specific implementation: Create a CustomGPT (or equivalent) with system prompts that refuse to provide complete answers. See Chapter 10. Program it to respond to requests for solutions with Socratic questions. Build in requirements that require students to explain their reasoning before the CustomGPT offers direction. Provide the CustomGPT with knowledge (your slides, notes, etc. – don't worry, the students cannot access them. Test it thoroughly. Ask your colleagues to test it. Ask some students to try to game it. Make sure your guardrails hold. Then deploy. Instruct students to turn in copies of their chats and give feedback on their reflection papers. Evaluate whether tweaks are needed. CustomGPT templates are found in Appendix H.

2. Embrace GenAI for Scaling One-on-One Instruction

The biggest learning gains come from personalized tutoring, which is impossible to provide at scale with human instructors alone. The Harvard study showed that GenAI tutors can deliver learning gains double those of active learning classrooms.⁷⁵ Well-designed GenAI tutors can provide immediate, personalized feedback to every student simultaneously. They never get tired. They never get frustrated with the same question asked twenty different ways. And they're always available.

Specific implementation: Identify the 2-3 concepts with which students struggle most each year. Create CustomGPT tutors specifically for those concepts. Require them to turn in chats so you can track usage data. Identify students who aren't using them and reach out individually to ask why they aren't. Don't assume that students will find the tools on their own.

3. Prioritize Metacognitive Engagement Over Task Completion

Don't just assess final products. Assess students' thinking about their thinking. The research showed that students using ChatGPT engaged in significantly fewer metacognitive activities. They stopped evaluating, monitoring,

⁷⁵ Kestin, *supra* note 1, (common active learning strategies include brief interactions during lectures, activities that may take 10-20 minutes, and longer activities like case studies, role plays, and structured team-based learning. These strategies range from short, simple activities like journal writing and problem solving to longer, involved activities that require more time and effort).

and orienting themselves. You need to design assignments that force these activities even when GenAI is available.

Specific implementation: Require process memos for every GenAI-assisted assignment. Students must document what they tried first, where they got stuck, what questions they asked the GenAI, which suggestions they accepted and why, which they rejected and why, and what they learned about their own thinking process. Grade the process memo as heavily as the final product (maybe more heavily?).

Implement peer review protocols that force students to evaluate each other's reasoning, not just their conclusions. Have students critique each other's GenAI-interaction transcripts. Did their peers ask good questions? Did they push back on GenAI suggestions? Did they think critically or accept blindly?

4. Test for Transfer, Not Just Performance

There's a risk that students might produce better briefs with GenAI assistance while understanding less about legal analysis. The graduate writing study showed essay quality improvements without knowledge transfer improvements.⁷⁶ To counter that, design assessments that require students to apply concepts in novel contexts without GenAI support.

Specific implementation: Include regular cold call scenarios where students must analyze an unfamiliar fact pattern without preparation or GenAI access. Don't just test whether they can produce work product with unlimited time and resources. Test whether they can think on their feet.

Create transfer assessments where students must apply doctrines in completely different contexts. If they learned contracts through commercial transactions, test them on services contracts. If they studied torts through car accidents, test them on product liability. See if they can transfer principles or only regurgitate examples.

⁷⁶ Fan, *supra* note 6.

5. Use Staged Assignments That Force Independent Work First

Research consistently shows benefits from requiring independent work first, followed by GenAI-supported revision, followed by reflection.⁷⁷ The GenAI-revision protocol produced students nearly three times more likely to retain knowledge compared to constant GenAI access. Don't let students use GenAI from minute one. Explain to the students why these stages are necessary.

Specific implementation: Break every major assignment into three stages with hard deadlines.

- Stage 1 (no GenAI allowed): Students produce a complete draft independently.
- Stage 2 (GenAI allowed): Students revise using GenAI but must document every change and justify every GenAI suggestion that they accept or reject, and, of course, verify every source.
- Stage 3 (reflection): Students write a memo analyzing what the GenAI helped with, what it missed, and what they learned about their own weaknesses.

Make the deadlines strict. Students who skip Stage 1 don't get to proceed to Stage 2. Without that mental struggle in Stage 1, the GenAI becomes a crutch rather than a scaffold.

6. Monitor for Cognitive Dependency and Build in Regular GenAI-Free Assessments

Pay attention to whether students can perform without GenAI. The 17% performance decline when GenAI was removed is your warning sign.⁷⁸ Include regular GenAI-free assessments. Watch for signs (such as a student not able to explain decisions made in their own work or who are weak in class but turn in

⁷⁷ Singh, *supra* note 66.

⁷⁸ Bastani, *supra* note 38.

next to flawless assignments—but make sure not to punish the introverts) that students are offloading too much cognitive work.

Specific implementation: Start every class with a 5-minute "check for understanding" exercise completed without GenAI. Ask students to explain a concept, distinguish cases, or analyze a hypo. Do this every single class. If students can't perform these basic tasks without GenAI, sound the alarm. They're developing dependency. One idea: when I taught criminal law, I wrote a fill-in-the-blank fact chart on the board in the first class for that week. Cold-called students had to tell me how to fill it in and explain their answers. I also had topic specific charts, such as for homicide and inchoate crimes that they had to complete. At the time, Bar Prep and Academic Success asked me for these charts to help students who had not been in my section. This sort of thing quickly tests understanding and reinforces learning.

I. The Bottom Line: Pedagogy Outweighs Technology

Your students are already using GenAI. Walk through your law library at 11 PM during finals, and you'll see ChatGPT open on every other laptop. Law schools can either shape that use toward genuine learning or ignore it and hope for the best.

*They may not admit it, but your students are already
using GenAI.*

Students who use GenAI under thoughtful instructor guidance learn more and build stronger independent capabilities. Students who use GenAI without structure often improve short-term performance at the cost of long-term competence. They optimize for task completion and miss the whole point of education.

The research provides evidence of how instructor-led GenAI can improve student outcomes so there's really no reason not to at least try it.

CHAPTER 7: Course Design and Teaching Logistics

Every semester brings a new round of decisions about how to structure a course and whether assignments need tweaking. You might be using a new edition of a casebook or adding more reading assignments from different sources. Or you have to rewrite your exams to be NextGen formatted. GenAI can make the design and logistics work significantly less time-consuming, freeing you to focus on what requires your legal expertise. This chapter explores how to use GenAI as a course design partner, from the architecture of a syllabus to the logistics of keeping a semester running smoothly.

A. Course Management

1. Developing New Courses

Creating a new course from scratch is exciting and terrifying in equal measure. You have expertise in the subject matter, but transforming that expertise into a coherent semester-long learning experience requires pedagogical architecture. Or maybe you don't have that expertise. I know it would be a serious time commitment from me if I suddenly had to teach BA. Using GenAI could save me considerable time.

a) Brainstorming Course Topics and Structure

You know you want to teach a course on law and emerging technology, or criminal justice reform, or international business transactions. But how do you translate a broad subject area into a 14-week course with logical progression and appropriate scope?

Start by asking GenAI to help you map the territory. Try something like:

I'm a law professor with an administrative law background designing a 14-week upper-level elective on cryptocurrency regulation for 2Ls and 3Ls. The course meets twice weekly for 80 minutes and emphasizes U.S. regulatory frameworks, with comparative references where instructive. Acting as a curriculum designer with expertise in financial regulation, organize major topics and subtopics into logical 2-week units, noting which units are most likely to require mid-semester updates given regulatory volatility.

GenAI will generate a comprehensive topic list, probably more comprehensive than you can reasonably cover in one semester, but still useful. Seeing the full landscape helps you decide what's essential and what you can omit or treat briefly.

You can also ask the GenAI to suggest different structural approaches:

Should this course be organized chronologically (following the development of cryptocurrency regulation), topically (covering different types of regulatory issues), or through a problem-based approach (analyzing specific regulatory challenges)? What are the pedagogical advantages of each approach?

The GenAI will outline options. You decide which approach serves your learning goals and your students' needs, but the exercise clarifies your thinking.

a) Developing Learning Objectives

Every good course has clear learning objectives. GenAI can help you articulate what you want students to know, understand, and be able to do by the end of the semester.

Describe your course concept and ask:

Acting as a legal curriculum designer familiar with ABA Standard 302 competency requirements, draft learning objectives for a 3-credit upper-level environmental law and climate change course for 2Ls and 3Ls. Organize objectives across three categories: doctrinal knowledge, lawyering skills, and practical competencies. The course emphasizes regulatory practice and includes a simulation component. Objectives should align with the following institutional requirements: [paste requirements].

GenAI will suggest objectives at different levels of Bloom's taxonomy (knowledge, comprehension, application, analysis, synthesis, evaluation) that align with your school's policies.¹ It helps ensure you're not just teaching students to memorize doctrine but also to apply, analyze, and critique it.

¹ Taxonomies of Learning, HARVARD BOK CTR., <https://bokcenter.harvard.edu/taxonomies-learning> (last visited Feb. 25, 2026).

Review the suggested objectives and revise based on:

- Your institution's program learning outcomes
- Skills your students need for practice or further study
- Gaps in your existing curriculum this course could fill
- What's realistic to accomplish in one semester

b) Selecting Materials and Organizing Content

Once you know what topics you'll cover, you need to find teaching materials. GenAI can help with the preliminary work.

GenAI can suggest types of materials to consider:

I'm designing a 3-credit upper-level corporate governance seminar at an ABA-accredited law school, targeting 2Ls and 3Ls with Corporations as a prerequisite. Acting as an experienced corporate law professor, suggest a source mix across cases, statutes, regulations, and secondary materials that balance Delaware fiduciary duty doctrine with federal securities regulation. Prioritize sources that support discussion-based pedagogy and include at least one contemporary governance controversy? The course is 13 weeks meeting once a week for 1 hour and forty minutes.

Alternatively, include a list of sources you already developed, and ask it to find gaps.

It might suggest:

- Key Supreme Court cases establishing foundational doctrine
- Representative cases showing doctrinal evolution or circuit splits
- Relevant statutory provisions from Delaware corporate law, federal securities law, or Sarbanes-Oxley
- SEC regulations and guidance
- ALI Principles of Corporate Governance
- Empirical studies on board composition or executive compensation
- Recent scholarship debating governance reforms

These are categories of materials, not specific recommendations. You still need to find the actual cases, articles, and other sources using your legal research skills and expertise.

But here's what you should NOT do: Don't ask GenAI to recommend specific cases or articles unless you're prepared to verify that every single citation is real. Use it to help you think about what kinds of materials you need, then find those materials yourself using databases and sources you trust.

c) Creating the Course Arc and Progression

Once you have learning objectives and a sense of what materials you'll use, you need to organize everything into a coherent progression. This is where GenAI can be particularly helpful with structure and pacing.

Share your topics and learning objectives, then ask:

I'm a law professor teaching a new 14-week course on immigration law to students with no prior background in the subject. The class meets twice a week for 50 minutes. Here are the major topics I want to cover: [list topics]. Suggest a logical sequence that builds student knowledge progressively, starting with foundational concepts and moving toward more complex applications.

GenAI will propose a course structure. It might suggest starting with historical context and constitutional framework before getting into specific visa categories or beginning with a compelling problem that makes students care about the topic before diving into doctrine. Evaluate these suggestions based on your teaching philosophy and what you know about how students learn.

You can also ask about pacing:

This proposed schedule covers removal proceedings in week 8. Is that appropriate timing, or should I move it earlier/later based on what students need to understand first?

GenAI can identify logical dependencies (students need to understand X before they can grasp Y) and suggest adjustments to improve flow.

d) Anticipating Teaching Challenges

Experienced teachers know that certain topics consistently confuse students, and certain concepts require extra time. But, when you're teaching something for the first time, you don't have that accumulated wisdom yet. GenAI can help you anticipate challenges.

Ask questions like:

I'm teaching 1L Civil Procedure for the first time using [Casebook]. The course is a 4-credit required course covering personal jurisdiction, subject matter jurisdiction, venue, Erie, and pleading. Acting as an experienced Civil Procedure professor, identify the three or four concepts within federal jurisdiction that consistently give 1Ls the most trouble, explain why each is conceptually difficult, and suggest one concrete classroom technique or hypothetical for each that has proven effective.

or

I'm teaching a two-week unit on statutory interpretation in a 1L Legislation and Regulation course. Students arrive with no prior exposure to interpretive methodology. Acting as an experienced legal writing and legislation professor, identify the four or five most persistent misconceptions 1Ls develop about statutory interpretation and suggest a specific teaching sequence or classroom technique that addresses each one before it takes hold. Cover both textualist and purposivist frameworks without favoring either.

The suggestions might not perfectly match your students' actual difficulties, but thinking through potential challenges ahead of time makes you better prepared to address them. For example, GenAI might flag that first-year students frequently conflate subject matter jurisdiction with personal jurisdiction. Even if your students turn out to struggle somewhere slightly different, having thought through likely pain points means you've already discussed the extra hypothetical(s) or chart that unsticks a confused 1L at 9 PM before an exam. As always, the more context you provide, the better the output.

e) Building Out Assignments and Assessments

Next, you need assignments that support learning objectives and assessments that measure student achievement. GenAI can help you design both.

For formative assessments (low-stakes assignments that help students learn):

I'm teaching a 3-credit upper-level federal criminal procedure course for the first time, covering Fourth, Fifth, and Sixth Amendment doctrine over 14 weeks, with approximately 30 students. The final exam is a 3-hour closed-book issue-spotter.

Acting as an experienced criminal procedure professor, design a week-by-week assignment sequence that scaffolds three skills in order: close case reading, fact-pattern analysis, and independent Fourth Amendment application. Each assignment should be completable in under two hours, vary in format across the semester, and explicitly prepare students for the issue-spotter format. Flag which assignments work best as low-stakes formative exercises versus graded work.

The GenAI might suggest:

- Case briefs with specific analytical questions
- Short problem sets applying doctrine to hypothetical facts
- Comparison exercises analyzing how different courts approach similar issues
- Reflection papers connecting doctrine to current events
- Peer review exercises on legal analysis

For summative assessments:

I'm teaching a 4-credit intellectual property survey course for the first time, covering copyright, trademark, and patent law in roughly equal proportion. The 3-hour closed-book final exam is the sole summative assessment. Acting as an experienced law school assessment designer, analyze the pedagogical tradeoffs among three exam structures: single-area depth, all-three-areas coverage, and an integrated problem requiring students to identify the applicable IP regime. For each approach, address what it tests well and how well it prepares students for practice. Close with a recommendation given the course parameters I've described.

The GenAI will outline options and trade-offs.

NOTE: This is a decision-support prompt. It requires the most complete context of any prompt type, because the AI's reasoning can only be as good as the situational information it receives. A content-generation prompt with thin context produces generic output. A *decision-support prompt* with thin context produces reasoning that sounds plausible but may be entirely wrong for your actual situation.

2. Creating a New Course Proposal

You have a great idea for a new course, but you're staring at a blank screen. GenAI can help you build your case and the course. Start with the narrative. Give GenAI your course concept, your target student population, and the curricular gap you're filling. Ask it to draft the course description and learning outcomes. If your institution has a specific proposal form or template, paste it in upfront so the output maps directly onto what you'll submit. The goal, as always, is a solid working *draft* you can refine rather than a blank screen or form that just sits there judging you.

Pay particular attention to learning outcomes at this stage, because curriculum committees will (I'm on the curriculum committee at my school). Ask GenAI to evaluate whether your proposed outcomes are specific and measurable. If your law school uses a programmatic learning outcomes framework, as ABA Standard 302² requires, ask GenAI to map your course-level outcomes to the program-level outcomes explicitly. Your future self, sitting in an accreditation site visit, will thank you.

One method is to ask GenAI to play skeptical committee member. Prompt it to review your draft and suggest the objections someone might raise: overlap with existing courses, workload concerns, resource implications, assessment rigor. It is much easier to address those questions on paper than in a room full of colleagues who have opinions and nowhere to be.

3. Creating a New Class Syllabus

GenAI is a godsend to drafting a new course syllabus. Rather than starting from a blank page, you can give GenAI your course goals, the relevant ABA standards, your law school's required formatting and boilerplate language, and a list of topics to cover, and ask it to produce a working draft. The result won't be final, but it will be a good start.

² *Standards and Rules of Procedure for Approval of Law Schools 2025-2026* r. 302, A.B.A., https://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar/standards/2025-2026/2025-2026-standards-chapter-3.pdf.

When prompting for a syllabus draft, be specific about your inputs. Provide the course level (1L required, upper-division seminar, clinic), credit hours, meeting schedule, and any institutional requirements your law school mandates, including required policy statements, disability accommodation language, grading rubrics, or specific formatting. If your school has a syllabus template, paste it (or upload it) directly into the prompt. GenAI will work within the structure you give it.

Once you have a draft, use GenAI to audit your learning outcomes. Ask it to evaluate whether your stated outcomes are measurable, appropriately scoped for the credit hours, and aligned with the law school's and applicable ABA accreditation standards. This is also a good moment to check whether the assessments actually measure what your learning outcomes say they measure. GenAI can flag misalignments you might not catch when you're deep in the weeds of your own course design.

Ask GenAI to read the syllabus from a student's perspective. Prompt it to identify language that is ambiguous or policies that could be reasonably misread. First-year students in particular will interpret syllabus language very literally or broadly to suit their needs (and many seem eager to flex their argument muscles early), and what seems obvious to you after years of teaching may not be. A quick check from a student's view can prevent a surprising number of "but the syllabus said..." conversations.

Your syllabus also needs to address GenAI use by students. As discussed in Chapter 5, AI policies benefit from being clear about what tools are permitted and when, and disclosure is required. A blanket prohibition is both unenforceable and out of step with the skills your students will need in practice. The syllabus is where that policy lives for purposes of your course, and it should be unambiguous. You might also want to disclose your own GenAI use. You don't want to end up in a New York Times article with students calling you a hypocrite.³ To draft your GenAI policies, use the [GenAI Policy Drafting Tool](#) discussed in Chapter 5.

Finally, run a compliance check. Ask GenAI to verify that your draft syllabus includes all language required by your institution and by applicable ABA

³ Kashmir Hill, *The Professors Are Using ChatGPT, and Some Students Aren't Happy About It*, NY Times (May 14, 2025), <https://www.nytimes.com/2025/05/14/technology/chatgpt-college-professors.html> (last visited Feb. 25, 2026).

standards, which governs course hours and credit. If your school has a checklist for syllabus approval, paste it in and ask GenAI to confirm each item is addressed.

4. Drafting Syllabi Updates

You taught this course last year. The learning objectives haven't changed. The casebook is the same. But you need to update the syllabus for new dates, tweak the schedule, revise assignment deadlines, and incorporate whatever new institutional policies your administration has invented since last semester.

GenAI can handle the mechanical updates. Give it your previous syllabus and the new semester dates and ask it to adjust the schedule accordingly. It can redistribute reading assignments if you're teaching Tuesday-Thursday instead of Monday-Wednesday while accounting for university holidays. And it can ensure assignment deadlines don't fall on exam days. I know. I do this every semester.

For substantive updates, the GenAI can help you incorporate new materials or reorganize topics. For example:

Here is my current 14-week Administrative Law syllabus: [paste or upload syllabus]. I want to add a 2-week unit on AI regulation covering the major federal frameworks and agency rulemaking challenges. Suggest where in the schedule this unit fits most logically. Identify what existing content could be condensed to accommodate the new content. Flag any thematic connections to existing topics that would make the transition feel coherent rather than bolted on.

When I revise a syllabus, I ask GenAI to evaluate the updated schedule from the student's perspective. It has caught more than one pacing problem I missed. In my Law Practice Technology course, students must earn two certifications (practice management and document assembly) plus at least one elective from a menu that includes AI and the Law, DISCO e-discovery, LegalSifter contract review, and Procertas office suite, all due the last day of classes. GenAI flagged that clustering the certification lessons earlier in the semester, right after the opening weeks on legal ethics and ethical AI, would let students complete them before midterms rather than carrying that workload into the final stretch. Students appreciate having the certifications off their plates early. I appreciate being able to turn around that grading quickly, well before the end of the grading period, and typically before finals start.

5. No More Emails: Creating a Canvas Bot

You know how you've diligently put everything students need to know in a syllabus, but students keep emailing you? How about building a Custom GPT trained on your course materials, including your syllabus, reading assignments, and course policies, and embed it or link to it directly in your Canvas course shell? Tell students to check it before emailing you. It will respond based on your actual course documents rather than general knowledge. The result is a tireless 24/7 TA that knows your syllabus cold and never has its office hours at an inconvenient time.⁴ See Appendix H for instructions how to do this, and see Chapter 10 for step-by-step instructions on building a Custom GPT.

6. Important Limitations and Warnings

Everything in this section comes with significant caveats that bear repeating.

First, GenAI cannot replace your subject matter expertise. It can help you organize and structure a course, but it cannot determine what's important to teach about your subject. That requires your judgment as a legal expert.

Second, GenAI knows nothing about your students, institution, or curriculum. It doesn't know what your students have already learned in other courses, what skills they struggle with, what your faculty expects graduates to know, or what your accreditation requirements mandate. You must adapt all GenAI suggestions to your specific context.

Third, *GenAI will make substantive errors about legal doctrine*. If you ask it to explain a legal concept to help you think about how to teach it, verify everything it tells you. Do not rely on GenAI-generated explanations of law without checking them against authoritative sources. This is especially critical when you're teaching an area that's new to you.

⁴ Building a Custom GPT requires a ChatGPT Plus subscription (\$20/month as of this writing). Some institutions have enterprise licensing arrangements that cover this or have licensing to different platforms such as Poe or Boodlebox; check with your IT or faculty support office before purchasing individually. Faculty at institutions running Microsoft 365 Copilot might have access to a comparable (though, frankly, not as good) option through Copilot Studio, Microsoft's low-code agent-builder. The functionality is similar. You upload your course materials, configure how the agent should respond, and deploy it where students can access it. Check with your IT office about whether your institution's licensing covers it and whether you can have access to it.

Fourth, GenAI draws on general education principles, not legal education specifically. Its suggestions might work well for teaching political science or history but miss important aspects of teaching legal analysis. Use your understanding of legal pedagogy to evaluate and adapt all suggestions. And, as always, provide as much knowledge and context as possible.

Finally, creating a course is fundamentally creative and intellectual work. GenAI can help with structure and organization, but the course must be yours. GenAI can shuffle your schedule but can't tell you which cases your students will actually argue about, or why that's more important than covering one more statute.

7. A Realistic Workflow

Here's what using GenAI for course development looks like in practice:

- Start with your own expertise and vision for the course
- Use GenAI to brainstorm structural options and identify topics you might have overlooked
- Make your own decisions about scope, sequence, and emphasis
- Find your own teaching materials using proper legal research methods
- Use GenAI to help organize materials into a coherent progression
- Design your own assignments based on learning objectives
- Use GenAI to suggest variations or alternatives to your initial ideas
- Revise everything based on your judgment and expertise
- Prepare to adjust significantly once you start teaching

An LLM is like a thinking partner who helps you work through design decisions, not a course designer who does the work for you. Use it to enhance your course development process, not to replace the intellectual work that makes the course valuable to students.

Remember, even with GenAI assistance, your first time teaching a new course will be messy. You'll realize in week 4 that you structured week 2 incorrectly. You'll discover that what you thought would take one class session in reality takes three. You'll find that students are confused by things you thought were clear. GenAI can't prevent the inevitable challenges of teaching something for the first time, but it can help you start with a more thoughtful structure than you might have created entirely on your own.

8. Creating Assignment Instructions

How many times have you written some variation of "write a memo analyzing..." or "prepare a motion arguing..."? Assignment instructions follow predictable patterns, which makes them perfect for GenAI assistance.

Describe your assignment goals to GenAI. Tell it what skills you're assessing and what format you expect. Ask it to draft instructions that are clear and complete.

For example:

I'm teaching a 3-credit upper-level Constitutional Law course for 2Ls and 3Ls. Acting as an experienced constitutional law professor, draft a complete student-facing assignment sheet for a legal research memo due in two weeks. The assignment asks students to research First Amendment compelled speech doctrine, cite at least five cases from federal courts with at least two from the Supreme Court, analyze the current state of the doctrine, and apply it to this hypothetical: [paste your fact pattern]. The instructions should include the assignment objective, specific submission requirements, and a grading breakdown. Students may use AI-assisted research tools but must verify all citations independently.

B. Assessment Development

Why stare at a blank screen or gather questions from myriad sources, including your old exams.

1. Question Brainstorming for Exams

A good exam question separates students who understand the doctrine from students who memorized someone else's outline. GenAI can help you write more of them.

Try this: "I'm designing a 3-hour closed-book final exam for an upper-level Constitutional Law course that covered substantive due process, equal protection, and federalism, with roughly equal time on each. Acting as an experienced ConLaw professor, draft three essay questions in MEE format — each built around a fact pattern of at least three paragraphs — that require students to identify the applicable doctrine, apply the correct standard of review, and reach a reasoned

conclusion. Questions should reward analysis over rule recitation. Include a one-paragraph note after each question identifying the primary issues it tests and the most common analytical errors to expect." You might consider uploading your slides or notes.

The LLM will generate options. Most will be mediocre. Some will be useful. The exercise helps you think through what you want to test and how to construct questions that reveal student understanding.

You can also ask the GenAI to suggest variations on existing questions or modify difficulty levels.

You might want to try uploading (anonymized) answer files from previous midterms and final exams looking for patterns across student responses. You can find out which issues students consistently miss and how teaching may have gaps that show up reliably in the exam record. While Examsoft can provide some of this analysis for multiple choice questions, it can't (currently) for essays. GenAI can read a batch of anonymized responses and reveal those patterns faster than any human grading. Then it can help you re-word the question so it and the students perform better.

2. Creating Practice Problems

Students need and want practice applying doctrine to facts before they face exam pressure. GenAI can help you generate practice problems throughout the semester (you can also do this with a CustomGPT; see Chapter 10).

Describe the legal issue and the level of complexity you want: " Here are my lecture notes and slides on personal jurisdiction: [paste materials]. Acting as an experienced Civil Procedure professor, draft a practice problem for 1L students who have completed the introduction to personal jurisdiction and purposeful availment but haven't yet encountered stream of commerce theory. The problem should present a fact pattern of 200-300 words involving minimum contacts, followed by three analysis questions that build in difficulty. Draw only from the attached materials — don't introduce doctrine I haven't taught yet. Format it as a take-home exercise students can complete in under an hour."

3. Developing Model Answers

After students complete practice problems or exams, they benefit from seeing strong model answers. GenAI can help you draft these—with critical caveats.

GenAI can generate a well-organized answer that identifies relevant issues and applies appropriate analysis frameworks. But it cannot reliably apply legal doctrine correctly. You must verify every legal proposition and ensure the model answer reflects the level of analysis you expect from students. Again, grounding it in knowledge (such as your notes) significantly improves output.

4. Writing Feedback Templates

You find yourself writing similar comments on student work repeatedly: "Your analysis would be stronger with more specific reference to ...," "Consider how the opposing party would respond to this argument," "This conclusion needs more support." Create feedback templates for common issues. GenAI can help you develop clear, constructive language for recurring problems. Build a library of feedback comments that you can customize for individual students, maintaining personalization while reducing repetitive drafting.

C. Using GenAI for Grading and Analysis of Student Submissions

Using GenAI for grading and analysis of student work raises serious professional responsibility questions, and faculty should approach it accordingly. GenAI can assist with certain parts of the grading process, but the actual evaluation must remain entirely your professional judgment. Only the law school can decide whether a student passes the class or gets put on academic probation. That makes this one of the highest-stakes applications of GenAI in faculty work, and it demands the most careful approach of anything covered in this chapter.

1. The Research

Before experimenting with AI-assisted grading, faculty should understand what the research shows, because the findings might surprise you.

Human graders are less consistent than most faculty assume. A 2024 UC Irvine study found that two trained human raters assigned the exact same score to the same essay only 49% of the time.⁵ Law faculty grading essay exams face similar dynamics, as do judges when deciding cases. A now-foundational study of parole board decisions found that judges granted parole at significantly higher rates immediately after meal breaks than at the end of long sessions, with approval rates dropping from roughly 65% to near zero as a session wore on.⁶ Faculty are not immune to the same pattern. A 2025 study analyzing over 104,000 oral examinations at an Italian university found that student passing rates followed a bell curve across the day, peaking around noon and falling sharply in early morning and late afternoon, even after controlling for exam difficulty.⁷

Against that baseline, GenAI grading looks more useful than critics typically acknowledge. In studies, GenAI scores landed within one point of the human score 76 to 89 percent of the time across three different essay corpora, and GenAI was more internally consistent with itself than human raters were with each other.⁸ The lead researcher described ChatGPT's performance as roughly equivalent to "an average busy teacher" and concluded it was well suited to low-stakes, formative feedback, particularly on first drafts.⁹ A separate study of undergraduate physics problems found that GenAI grading improved

⁵ Tamara P. Tate et al., *Can AI Provide Useful Holistic Essay Scoring?*, 7 Computers & Educ.: Artificial Intelligence 100255 (2024) (finding that two trained human raters assigned identical scores to the same essay only 49% of the time, and that AI scores fell within one point of human scores 76 to 89% of the time across three essay corpora, suggesting AI is sufficiently reliable for low-stakes, formative assessment purposes).

⁶ Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 Proc. Nat'l Acad. Sci. U.S.A. 6889, 6891 (2011) (finding that favorable parole rulings dropped from approximately 65% to nearly zero within each decision session before resetting after food breaks, suggesting that extraneous factors unrelated to the merits influence judicial outcomes).

⁷ Carmelo M. Vicario et al., *Timing Matters! Academic Assessment Changes Throughout the Day*, 16 Frontiers in Psych. art. 1605041, at 4–6 (2025) (analyzing over 104,000 oral examinations and finding that student passing rates followed a bell curve across the day, peaking around noon and declining significantly in early morning and late afternoon, even after controlling for exam difficulty).

⁸ Jill Barshay, *AI Essay Grading Is Already as 'Good as an Overburdened' Teacher*, Hechinger Rep. (May 20, 2024), <https://hechingerreport.org/proof-points-ai-essay-grading-is-already-as-good-as-an-overburdened-teacher-but-researchers-say-it-needs-more-work> (reporting lead researcher's conclusion that ChatGPT performs comparably to an average busy teacher and is appropriate for low-stakes classroom use, though not yet reliable enough for high-stakes summative assessment).

⁹ *Id.*

substantially when provided with a detailed rubric, with GPT-4 approaching human-level performance once given clear grading criteria.¹⁰

The consistent finding across studies is that GenAI grading works better with a detailed rubric and performs poorly without one. It is not yet reliable enough for high-stakes summative assessment. For law professors, GenAI can help check rubric application and identify patterns across a set of student responses. Often, faculty have tight time constraints to get their grades in. In the fall semester, during the holiday break, and in the spring, just after finals and if you're on a 10-month contract, that's a very quick turnaround. But GenAI can assist.

2. What GenAI Should *Never* Do

Before we discuss appropriate uses, let's talk about what's unethical and inappropriate.

Never ask GenAI to assign grades. The decision about what grade a student receives must be your professional judgment, based on your evaluation of their work. Delegating that to GenAI, even if you review and approve its suggestions, is not an appropriate use.

Never upload student work such as papers or exams to public GenAI tools without removing all identifying information. Student work is protected by FERPA, and the platform matters enormously here. FERPA permits disclosure of student education records to third parties functioning as "school officials" with a "legitimate educational interest." Microsoft, under its enterprise agreement with the university (assuming yours has one), fits that description. The institutional agreement restricts how Microsoft can use that data and holds them to FERPA's requirements. That's what makes Copilot enterprise potentially different from uploading assignments to ChatGPT or a personal Copilot account, which have no FERPA agreement with the university and should never be used with student data. Even where FERPA permits something, your institution's own privacy and data governance policies may impose additional restrictions. Check with IT or

¹⁰ Ryan Mok et al., *Using AI Large Language Models for Grading in Education: A Hands-On Test for Physics*, arXiv:2411.13685, at 11–13 (Nov. 20, 2024) (finding that AI grading without a mark scheme produced substantial over-leniency due to hallucinations, but that providing a detailed rubric significantly improved accuracy and consistency, with GPT-4 approaching human-level performance).

General Counsel to confirm that AI-assisted grading falls within what your institution's policies expressly allow. Learn more about this in Chapter 3.

If you want an extra layer of protection, or if response times are slow, Canvas has a built-in anonymizing tool that strips student names from submissions before you download them.¹¹ You can toggle anonymity on to download the work, then turn it off when you return to final grading. Simple, and it removes any ambiguity about what you uploaded.

One of the huge benefits of Examsoft Legacy is the anonymization tool that automates application of the students' anonymous grading numbers to the answer files, making it safer to upload student answer files.

Never rely on GenAI to detect plagiarism or AI-generated content. Detection tools are notoriously unreliable, producing both false positives and false negatives, and they disproportionately flag work by ESL and neurodivergent students.¹² Accusing a student of academic dishonesty based on an AI detection score is unfair and potentially discriminatory.

Never use GenAI to generate feedback that you haven't carefully reviewed and personalized. Students deserve feedback that reflects your actual assessment of their work. Generic AI-generated comments, delivered unedited, are not feedback. They're bland comments with your name on it.

Two more things worth considering as you experiment. First, be honest with yourself about what AI-assisted grading means in your course. Using Copilot to help draft feedback or work through a rubric is a very different thing from having it assign scores, and the latter raises validity questions worth sitting with before you go all in. Second, a brief syllabus note acknowledging that you use GenAI tools in your grading process (for say, low/no stakes assessments) is good practice. Students appreciate the transparency, and the quick turnaround. Again, the last thing you want is to become the professor in the news who prohibits students from using AI while using it himself.¹³

¹¹ Canvas instructions for anonymous grading are <https://community.instructure.com/en/kb/articles/660678-how-do-i-add-an-assignment-that-includes-anonymous-grading>. Contact your institution's learning technology support office if you need assistance.

¹² See Chapter 5 for a fuller discussion of AI detection tools and why building your assessment strategy around them is a losing proposition.

¹³ Hill, *supra* note 3.

3. What GenAI Can Appropriately Do For You

With those limitations clearly in mind, here are legitimate uses of GenAI in grading:

a) Creating Grading Rubrics and Evaluation Frameworks

Before you grade, you need clear criteria. GenAI can help you develop comprehensive rubrics:

"Here is the legal research memo assignment I gave my 1L students: [paste assignment]. Acting as an experienced legal writing professor, create a grading rubric for this assignment using four performance levels: excellent, proficient, developing, and inadequate. Include five criteria: legal analysis (40%), use of authority (25%), organization (15%), writing quality (15%), and citation format (5%). For each criterion at each performance level, write two to three sentences describing observable, specific performance characteristics a grader could apply consistently across student papers. Calibrate the 'proficient' descriptions to reflect competent first-semester 1L work, not law review quality."

GenAI will generate a detailed rubric. You'll revise it based on what you actually taught and what's realistic to expect from students at this level. Once edited, it does two things well: it keeps your grading consistent across thirty identical-looking memos, and it gives you something to point to when a student asks why they got a B+.

b) Drafting General Feedback for Common Issues

When you notice that fifteen students made the same mistake, you don't want to type the same explanation fifteen times. GenAI can help you draft clear explanations of common errors:

"I'm writing feedback for 1L Civil Procedure students whose research memos show a specific analytical gap: they identified the relevant minimum contacts factors but treated them as a checklist rather than explaining why those factors matter for the due process fairness inquiry. Acting as an experienced Civil Procedure professor, draft a student-facing paragraph of 150-200 words that explains the conceptual relationship between minimum contacts and due process, and tells students how to fix their analysis going forward. Use plain language suitable for students three weeks into personal jurisdiction. Reference *International Shoe* but don't introduce doctrine beyond what that case establishes."

You then customize this explanation for individual students based on their specific errors, but you start with a clear articulation of the underlying problem.

c) Analyzing Individual Student Work for Feedback Points

You can use GenAI to help identify what feedback to give on individual student work, but do so with extreme caution. You should:

1. Remove all identifying information from the student work
2. Use institutional GenAI tools if available, not public tools
3. Review every GenAI suggestion critically before using it
4. Ensure feedback reflects your actual assessment, not GenAI's
5. Be aware that GenAI may introduce bias into your evaluation

Example of appropriate use:

You have graded a student memo and assigned it a B+. You know what's good about it (solid analysis, clear writing) and what's weak (organization could be improved, citation format needs work). You want to give detailed feedback but you're grading 60 memos and you're drained.

You could paste the anonymized memo into an institutional GenAI (e.g., Copilot) tool and ask:

"This is a student legal memo that I've evaluated as good but not excellent. I want to provide constructive feedback on: (1) organizational structure, (2) citation format issues. Identify 3-4 specific places where the organization could be clearer and any Bluebook errors you spot."

You review each suggestion to verify it's accurate and helpful. You then write your feedback to the student in your own words, incorporating the AI's observations where they're valid.

Critical limitations:

GenAI might identify "problems" that aren't really problems, especially in legal writing where there's often more than one defensible approach. Again, GenAI knows nothing about this particular student. It doesn't know what you've taught or what standards apply in your course. GenAI might miss the most important feedback you should give in favor of superficial technical issues.

Most important: GenAI can't tell you whether the student's legal analysis is sound. It can identify organizational issues or citation errors, but it cannot reliably

evaluate whether the student correctly applied the law to the facts. That requires your expertise.

d) Ensuring Grading Consistency

See “The Research” subsection above for details about what influences grading.

GenAI can help you check for consistency by reviewing your rubric applications:

I've graded five Civil Procedure research memos using a rubric with five weighted criteria: legal analysis (40%), use of authority (25%), organization (15%), writing quality (15%), and citation format (5%). Here are the anonymized criterion scores and overall grades I assigned: [paste scoring data]. Acting as a grading consistency reviewer, identify any cases where students with similar criterion scores received meaningfully different overall grades, and flag any criterion where my scoring pattern suggests I may be applying the standard inconsistently across students. Don't recommend grade changes — just identify the cases worth a second look and explain why each one caught your attention.

Using GenAI in this way is quality control on your own grading, not delegation of grading decisions to GenAI.

4. The Bias Problem in GenAI-Assisted Grading

GenAI models can introduce or amplify bias in evaluation of student work.

Research shows that GenAI can:

- Favor certain writing styles over others in ways that correlate with demographic characteristics¹⁴
- Respond differently to names or other demographic cues (which is why you must remove identifying information)¹⁵

¹⁴ W. Liang et al., *GPT Detectors Are Biased Against Non-Native English Writers*, Stanford HAI (May 15, 2023), <https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers> (finding that AI systems score text using perplexity metrics that correlate with lexical complexity, systematically disadvantaging non-native English writers).

¹⁵ Kyra Wilson et al., *AI Tools Show Biases in Ranking Job Applicants' Names According to Perceived Race and Gender*, Presentation at AAAI/ACM Conference on AI, Ethics, and Society (Oct. 22, 2024), <https://ischool.uw.edu/news/2024/11/ai-tools-show-biases-ranking-job-applicants-names-according-to-perceived-race-and-gender> (finding that LLMs favored white-associated names 85% of

- Have systematic preferences that don't align with academic quality (e.g., preferring longer responses or certain rhetorical patterns)¹⁶

Example: Some GenAI models favor formal, elaborate writing styles. Students who write clearly and concisely might receive lower GenAI ratings than students who write in a more ornate style, even though legal writing values clarity over embellishment.¹⁷

You are the check against this bias. When GenAI suggests feedback or identifies issues, ask yourself: Is this feedback equally appropriate for all students? Would I give this feedback to students regardless of their writing style, background, or approach? Am I applying the standards that matter for legal writing, or am I being influenced by GenAI preferences that don't reflect good legal writing?

4. A Realistic Workflow

If you decide to use GenAI to assist with grading (which you might reasonably choose not to do), here's a responsible workflow:

1. Grade the work yourself first. Make your own evaluation and assign your grade based on your professional judgment.
2. Create or review your rubric with GenAI assistance if helpful, but finalize it based on your teaching goals.
3. For feedback, consider using GenAI to help identify specific issues in anonymized student work, but only for work you've already graded.
4. Review all GenAI suggestions critically and reject anything that doesn't align with your assessment.
5. Write feedback in your own words, incorporating GenAI observations where they're valuable but maintaining your voice and your relationship with the student.
6. Never let GenAI influence your actual grade. The grade you assigned based on your professional judgment is the grade. Period.

the time and never favored Black male-associated names over white male-associated names across more than three million resume comparisons).

¹⁶ Keita Saito et al., *Verbosity Bias in Preference Labeling by Large Language Models*, arXiv:2310.10076 (finding that LLMs exhibit measurable preference for longer, more verbose responses in evaluation tasks, with LLM verbosity preferences diverging from those of human evaluators).

¹⁷ *Id.*

7. Check for bias in how you're using GenAI. Are you using it consistently across all students? Are you letting it influence how you think about student work in ways that might disadvantage particular groups of students?

5. The Bottom Line

GenAI can help with the mechanical and administrative parts of grading: creating rubrics, ensuring consistency, identifying patterns, articulating feedback on common issues. It cannot and should not make evaluative judgments about student work.

GenAI is like having a teaching assistant who can check citations and flag organizational issues, but who has no legal training and can't evaluate whether the legal analysis is sound. You still need to read every word of every student's work and make your own professional judgment about its quality.

Students expect and deserve to have their work evaluated by their professors. Don't let efficiency concerns compromise the fundamental educational relationship between teacher and student.

D. Practicing the Socratic Method

If you're new to teaching, or new to a particular subject, or just want to improve your classroom questioning, you can practice Socratic dialogue with GenAI before you try it with actual students.

The Socratic method is harder than it looks. You need to ask questions that probe student thinking without being leading or hostile. Then follow up effectively on student responses while adjusting your questioning based on what students say. GenAI can help you practice this without the pressure of a classroom full of students watching you stumble.

1. Why Practice with GenAI?

Real students are unpredictable. They give answers you didn't anticipate and often get confused by questions you thought were clear. Sometimes they just stare at you in silence or bury their heads in the casebook. You can't really practice for that unpredictability, but you can practice:

- Sequencing questions logically

- Asking follow-ups that deepen analysis
- Responding to different types of student answers
- Identifying what you want students to realize through questioning
- Avoiding leading questions or questions that give away the answer

Simulating student responses with GenAI gives you a no-stakes environment to develop these skills. You won't be brilliant at first, and that's fine because there are no actual students watching you figure it out.

2. Setting Up a Practice Session

Here's how to structure a Socratic practice session with GenAI.

First, choose your case and identify your teaching goals: What case are you teaching? What do you want students to understand by the end of the questioning? What analytical moves do you want them to practice?

Then, prompt the GenAI:

I'm a law professor preparing to teach [case name] in my [course name] course. Please roleplay as a law student I'm cold-calling in class.

Simulate a [1L / 2L / upper-level] student who is [choose one: well-prepared but hasn't grasped the theoretical implications / sharp and needs pushing on policy / underprepared and working mostly from the headnotes]. My goal for this session is to practice [choose one: drawing out basic case analysis / pushing students from facts to doctrine / connecting doctrine to policy implications].

Respond the way a real student would: sometimes right, sometimes partially right, sometimes confused. Don't hand me the sophisticated analysis — make me work for it. If my question is unclear, ask for clarification the way a student would. If I ask something leading or softball enough that a real student would just parrot it back, do that, so I can see the problem with my question.

One more thing: if one of my questions is too leading, confusing, or unlikely to produce the analytical move I seem to be going for, break character for one sentence and tell me why before returning to the roleplay. I want to improve my questioning, not just practice bad habits against a cooperative audience.

3. Running the Practice Session

Now start questioning as you would in class. Don't script it. Ask questions that seem natural to you, respond to what the GenAI "student" says, and follow your instincts about what to ask next.

Example exchange:

You: "Can someone explain what happened procedurally in Pennoyer v. Neff?"

GenAI Student: "Um, so Neff owned property in Oregon, and Pennoyer tried to take it? There was something about a lawsuit first..."

You: "Yes, there was an earlier lawsuit. Who was suing whom in that first case, and what was it about?"

GenAI Student: "Mitchell sued Neff for attorney's fees. Mitchell won, but Neff wasn't there for the trial. Then Mitchell bought Neff's land at a sheriff's sale."

You: "Good. So, Neff wasn't present for the trial. How did the Oregon court get jurisdiction over Neff if he wasn't there?"

GenAI Student: "They published a notice in the newspaper? I think that was supposed to give notice that he was being sued."

You: "Right, publication notice. Now here's what I want you to think about: The Oregon court issued a judgment against Neff even though he wasn't in Oregon and maybe never saw the newspaper notice. Can a court do that? Should a court be able to do that?"

(and so on).

CHAPTER 8: Designing Assignments in the Age of GenAI

A. Introduction

Every law professor who has assigned a traditional research memo in the past three years has, at some point, wondered whether the paper that was submitted by a student was generated by a large language model. That suspicion, regardless of whether it crystallizes into a formal academic integrity complaint, has already changed legal education. Some professors are redesigning assignments, adding oral components to their courses and debating among themselves whether any of these changes can combat the unauthorized use of GenAI by students.

The use of large language models (LLMs) in legal education is all so new and evolving so quickly that none of us is entirely sure how best legal education should respond to GenAI. However, four broad categories have emerged in the pedagogical literature and in faculty practice:

- assignments that prohibit GenAI use entirely
- assignments designed to be structurally resistant to it
- assignments that require students to *use* GenAI as part of the learning task, and
- assignments built on transparency regardless of the level of use.¹

No single category is the right answer for all courses, all students, or all faculty. A contracts professor teaching first-year students to spot consideration issues has different needs than a legal writing instructor teaching advanced research and persuasion, and both have different needs from a seminar instructor exploring the ethics of AI-assisted lawyering. What the framework below offers is

¹ Mike Perkins et al., *The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment*, 21(6) J. U. Teaching & Learning Practice (2024), <https://arxiv.org/abs/2312.07086> (last accessed Mar. 9, 2026) [hereinafter AIAS]. The four-category framework synthesizes approaches appearing across the pedagogical literature. The AIAS identifies five levels of GenAI integration, from "No AI" to "AI Exploration," and has been adopted by hundreds of institutions worldwide. See also Stacy Ybarra Evans, *The AI-Resistant Pedagogy Studio: An Interactive Toolkit for Assignment Redesign* (2025), <https://ynaqbxzy.manus.space> (describing the shift from product-based to process-based assessment as a response to GenAI capabilities).

a practical vocabulary for thinking through those differences, along with concrete assignment designs you can adapt or use as starting points for something better suited to your own classroom.

One more thing before we get into the categories. Model Rule 1.1 requires competence, which the ABA has interpreted to include keeping up with the benefits and risks of relevant technology.² Model Rule 5.3 requires supervisory lawyers to ensure that non-lawyer assistance (including, increasingly, GenAI tools) comports with the professional obligations of the lawyer.³ When students make choices about GenAI use in their assignments, they are rehearsing the judgment calls they will make as lawyers. Framing every assignment policy in those terms, however briefly, shifts the conversation from “can I get caught” to “what kind of lawyer am I practicing to be.”

Table 8.1: The Four-Category Framework at-a-Glance

Category	GenAI Role	Pedagogical Purpose	Key Risk
Prohibited	Blocked entirely	Build foundational skills; bar-exam readiness	Enforcement without structural support
Resistant	Structurally discouraged	Authentic engagement; process visibility	Clever students find workarounds
Integrated	Required tool	AI literacy; critical evaluation; professional readiness	Superficial compliance without genuine learning
Transparent	Disclosed and documented	Professional accountability norms; citation culture	Disclosure without reflection

² MODEL RULES OF PRO. CONDUCT r. 1.1 cmt. 8 (A.B.A. 2023) (“To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks of relevant technology . . .”).

³ MODEL RULES OF PRO. CONDUCT r. 5.3 (A.B.A. 2023) (requiring lawyers with direct supervisory authority over nonlawyer assistants to make reasonable efforts to ensure the nonlawyer’s conduct is compatible with the lawyer’s professional obligations).

No position on this spectrum is inherently superior to any other. Choose which approach aligns best with your learning outcomes.

1. GenAI-Prohibited Assignments

Prohibition is the oldest and most intuitive response to a new technology that threatens assignment integrity. It is also the most frequently misapplied.

A blanket ban on GenAI use, announced in a syllabus and never reinforced by assignment structure, simply doesn't work and has a poor track record. Students who intend to use LLMs will use them anyway, and the only practical effect of an unstructured prohibition is that the students who comply are disadvantaged relative to the students who do not. That asymmetry is unfair and harmful. Once students believe that some of their peers are gaining an advantage through prohibited means, the rational calculation shifts, and the student who was on the fence finds herself disadvantaged for following the rules, especially if the student has dyslexia or is not a native English speaker.

The dynamic is not new. A major league baseball player once told me, with what I took to be genuine regret, that he felt he had no choice but to take performance-enhancing drugs. Not because he wanted to, but because he knew his teammates were already doing it. The alternative was going back to Triple-A ball. That is the logic of competitive coercion, and it arises whenever the perceived stakes are high and enforcement is weak.

A Fred Friendly seminar called "Better Brains: The Ethics of Neuro-Enhancement" staged essentially the same debate in an academic context.⁴ In this seminar, the panelists wrestled with the question of whether students could ethically take "RememberAll," a hypothetical cognitive enhancer with greater benefits than Adderall (without the negative side effects) if they believed their classmates already were taking the medication.⁵ The answer for many on the panel

⁴ *A Better Brain: The Ethics of Neuro-Enhancement*, Ethics in America II (Fred Friendly Seminars 2007), <https://www.learner.org/series/ethics-in-america-ii/a-better-brain-the-ethics-of-neuro-enhancement/> (site scheduled to go offline July 1, 2026).

⁵ *Id.*, At one point, a disagreement arises between a father, portrayed by Justice Antonin Scalia, and his son, played by Joshua Foer (a journalist and winner of the 2006 USA Memory championship), who is in college. The son wants to use a memory-enhancing drug called Rememberall to improve his exam performance. The father objects, stating that people should accept their natural abilities, even if they aren't perfect. The son challenges this, pointing out that his father always encouraged him to strive for academic excellence, implying that using the drug is just another way to achieve that goal. He questions the father's reasoning, asking what the problem is if the drug is safe and commonly used.

was yes. Some of the panelists justified this conclusion by likening the drug to other “enhancements” available to students such as computers, tutors, prep classes, etc.⁶ Law students, who have already invested enormous effort to get where they are, are not immune to that calculation. For example, students with family connections to the legal profession or private tutors have advantages that no syllabus has ever prohibited. As a result, a first-generation student who stayed up until 2 a.m. for four years of undergraduate work to earn admission to law school might not easily shrug off the perception that her classmates are using a freely available tool to get ahead of his or her peers.

That said, prohibition is not inherently wrong. Some assignments necessarily require students to work without GenAI assistance because the learning objective is precisely the cognitive labor that LLMs would short-circuit. A first-year student who uses ChatGPT to outline a negligence analysis has not learned negligence analysis; they have learned how to prompt a tool into producing one.

The professional responsibility framing is especially significant. Comment 8 to Model Rule 1.1 instructs lawyers to keep abreast of changes in the law and its practice, including the benefits and risks of relevant technology.⁴ But “keeping abreast of technology” does not mean outsourcing legal reasoning to it. A lawyer who cannot analyze a fact pattern without running it through an LLM first has a competence problem, and law schools that allow students to develop that dependency have failed their graduates. Prohibition, used strategically and with structural reinforcement, is one tool for preventing that outcome.

2. When Prohibition Makes Sense

Prohibition works best when three conditions are present:

- The assignment targets foundational doctrinal skills that students must eventually perform independently much like the kind of skills the bar examiners are testing;
- The format makes verification feasible, whether through in-class delivery, oral examination, or a multi-stage process that would be difficult to fabricate entirely; and

⁶ *Id.*, the panelists included Professor Art Caplan, then the director of Bioethics at UPenn and now at NYU Grossman School of Medicine, and Professor Michael Sandel, D.Phil, who teaches a course on justice at Harvard University, that is one of the most popular on the Harvard campus and online.

- The prohibition is explained to students in terms of their professional development, not just institutional rule-following.

Conversely, prohibition is poorly suited to assignments that are already submitted asynchronously or involve large amounts of research aggregation. Similarly, assignments that require students to produce polished prose in their final submission also benefit from GenAI because these are exactly the tasks at which current LLMs excel. Prohibiting the use of LLMs without structural changes to the assignment design is, therefore, as one faculty survey put it, simply hoping for compliance.⁷

3. Making Prohibition Structurally Credible

The most effective GenAI-prohibited assignments require a student to complete a task or perform analysis that an LLM cannot provide or execute on the student's behalf. In-class written exercises, oral components, and assignments requiring specific personal or local knowledge all fit this description. For example, a doctrinal quiz administered and submitted in class cannot be completed with LLM assistance. An oral recitation or Socratic exchange requires the student to perform in real time without the assistance of GenAI. An assignment asking students to reflect on a client interview they attended or a courthouse proceeding they observed produces outputs that an LLM cannot generate authentically because the underlying experience did not happen to the LLM. The same is true of assignment that requires students to personally complete a physical task, such as tracing a statute's legislative history through the state archives, for example.⁸

Multi-stage assignments that require visible revision are another structural approach. When faculty collect an initial outline, a rough draft, and a final submission, and require the student to track changes in the document(s), the shape

⁷ John Bliss, *Teaching Law in the Age of Generative AI*, 64 JURIMETRICS 111, 147 (2024) (reporting that many surveyed law faculty, even those who strongly agreed that students should be prepared to use and critically evaluate GenAI, acknowledged they were uninformed about the technology and unsure how to proceed). *See also* Dyane O'Leary, quoted in *AI Enters the Classroom as Law Schools Prep Students for a Tech-Driven Practice*, NAT'L JURIST (July 10, 2025) (noting that prohibiting AI entirely "is starting to feel unrealistic" given pervasive access and equity concerns).

⁸ Evans, *supra* note 1 (recommending multi-stage assignments, personal connection tasks, and AI-critique assignments as resistance strategies). For the specific approach of using local or experience-based knowledge as a structural resistance mechanism, *see also* *AI Teaching Strategies: Transparent Assignment Design* (2024), OHIO S. Office of Academic Affairs, <https://teaching.resources.osu.edu/teaching-topics/ai-teaching-strategies-transparent> (last visited Mar. 7, 2026).

of a student's thinking over time becomes clear. A clean, polished first submission with no evidence of struggle or second-guessing is a red flag.

Using a secure exam software such as Examsoft or Exam4 in class can be an effective way to enforce prohibition assuming the room is effectively proctored and smart phones are prohibited. Secure take home exams are not effective because a student could easily upload a picture of an exam question to GenAI and choose or transcribe the answer. As previously mentioned, it has been repeatedly shown that LLMs perform extremely well on the bar exam.

4. Practical Examples in Legal Education

In a first-year Torts course, a brief in-class writing exercise asking students to apply the reasonable person standard to a novel fact pattern collected on paper, not submitted digitally, accomplishes two things at once. It assesses the students' doctrinal knowledge of the subject matter and it requires independent analysis of the facts by the student under conditions where GenAI assistance is not available.

Similarly, In a Contracts course, an oral final exam or oral assessment component that requires students to state and defend their analysis of a question can serve the same function. The student may have prepared extensively for the exam with LLM assistance but the analysis presented by the student orally in front of the faculty member is their own. These types of assignments should make it fairly clear whether the student understands what has been taught.

In a Legal Writing course, staged submission with mandatory office hour check-ins makes the writing process visible in ways that a single polished submission never could. If a student struggles to explain her arguments or research in an office hour discussion with the faculty member but submits prose that would make John Marshall proud, the discrepancy speaks for itself without requiring faculty to run anything through a GenAI detector.

5. The Equity Consideration

No discussion of GenAI-prohibited assignments is complete without addressing testing accommodations and resource inequities. Students with extended time accommodations pose a real enforcement challenge in any timed, in-class format, not because their accommodations are illegitimate, but because extended-time administrations are often unsupervised, which creates exactly the

enforcement gap that prohibitions are designed to close. Faculty need to think through this carefully and in advance, either by designing oral components that are inherently accommodation-compatible or by working with disability services offices to develop supervised extended-time arrangements.

Suffolk University Law Professor Dyane O'Leary has noted that prohibiting GenAI is, in some respects, a proxy for the advantages that already-privileged students enjoy through family networks and paid tutoring.⁹ A first-generation law student who uses an LLM to help understand an unfamiliar doctrinal concept is doing something qualitatively different from a student who uses it to submit an essay that is beyond their knowledge and ability. A prohibition designed to catch the student who submits work beyond their actual ability is defensible when properly calibrated. A policy that also snares the student who used an LLM to understand the material has mistaken a study aid for a cheating problem.

B. GenAI-Resistant Assignments

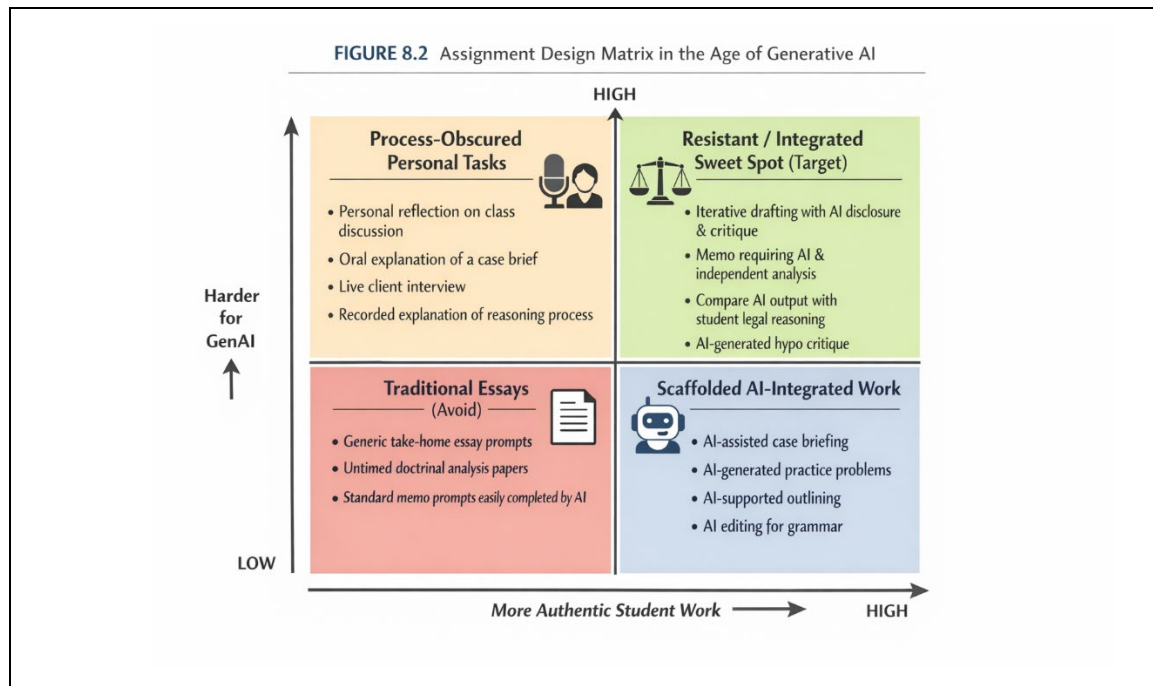
GenAI-resistant assignments do not prohibit LLM use. Instead, they make it unhelpful. The distinction is important, both practically and philosophically. Practically, resistance removes the enforcement burden that makes prohibition so difficult. If an LLM cannot produce an adequate submission because the assignment requires something the LLM does not have access to, faculty do not need to police usage. The design does that work. Philosophically, resistance respects student autonomy while structuring the environment to make authentic engagement the path of least resistance.

The academic integrity literature refers to this approach as “AI-resistant pedagogy,” designing assignments that emphasize uniquely human contributions like personal experience, local knowledge, real-time performance, and iterative intellectual development.¹⁰ The shift is from “What did you produce?” to “How

⁹ G. Alex Ambrose, *Enhancing Assignments with AI Transparency*, NOTRE DAME Learning (2024), <https://learning.nd.edu/resource-library/enhancing-assignments-with-ai-transparency/> (last visited Mar. 4, 2026) (noting that clarity on permissible AI use is “our responsibility as instructors to address . . . assignment by assignment”).

¹⁰ Ybarra Evans, *supra* note 1 (defining “AI-resistant pedagogy” as design that foregrounds uniquely human cognitive contributions and shifts assessment from product to process). See also *Academic Integrity and Teaching With(out) AI*, Harvard College Office of Academic Integrity and Student Conduct (2024), <https://oaisc.fas.harvard.edu/academic-integrity-and-teaching-without-ai/> (last visited Apr. 25, 2026), (describing AI-free design features including personal reflection, problem-solving, and hands-on experiences that AI finds challenging to replicate).

did you learn, and what can you do with that knowledge?” That is a question LLMs cannot answer on a student's behalf.



1. Assignment Types That Resist LLM Completion

Personal reflection assignments connected to specific experiences fall into this category when they are truly specific. An assignment asking students to reflect on a client interview they conducted in a clinic, a deposition they observed during an externship, or a hearing they attended in a courthouse requires the student to have been physically present at an event that generated information the LLM cannot access. The reflection on what they observed and what they would do differently the next time is not replicable by an LLM because the underlying experience is singular and because LLMs (currently) cannot “experience” anything.

Local context assignments operate on the same principle. An assignment asking students to trace the procedural history of a real pending case in the jurisdiction's docket system or to compare the statutory text of their state's version of a uniform act against the official model requires effort that LLMs cannot *reliably* generate. LLMs trained on general corpora have patchy knowledge of state-specific procedural variations and unreliable knowledge of anything that happened after their training cutoff, which a vulnerability that faculty can

deliberately exploit. On the other hand, a student who provides the texts of the state's version and the official model, and knows how to prompt properly (e.g., using the CORRECT format), could easily work around these limitations.

Multi-stage process assignments make intellectual development visible by requiring students to submit multiple iterations over time. When faculty collect an annotated bibliography, a preliminary analysis memo, a revised draft with tracked changes, and a final submission, the pattern of revision either corroborates a narrative of intellectual engagement, or it doesn't. An LLM can produce a polished final draft; it cannot produce a coherent revision history that maps onto a student's developing understanding of a problem.

Peer teaching and discussion-leading assignments require students to demonstrate real-time comprehension in a social context where the LLM is not available as a crutch. A student who is leading a case discussion or presenting a doctrine to their classmates is performing live. The follow-up questions from classmates and faculty are not scripted, and the quality of the student's responses reflects their actual preparation, not their prompting skill.

2. The Limits of Resistance

Two caveats about GenAI-resistant assignment design deserve mention. First, the resistance is not absolute. A student sufficiently motivated to deceive can fabricate a personal reflection or find a workaround for most structural constraints. GenAI-resistant design raises the cost and reduces the reward of that deception; it does not eliminate the possibility altogether. Second, as LLMs gain access to real-time web browsing and local docket systems some currently resistant assignments will become less so. An assignment that is resistant today because it requires current docket information may not be resistant next year (or next month) when the relevant LLM has integrated court filing databases.

The practical implication is that resistance works best as a layered strategy. No single design feature makes an assignment fully resistant, but combining a personal experience component, a specific local knowledge requirement, scaffolded assignments with revision history, and an oral or discussion element creates enough layers that authentic engagement becomes easier than trying to game the system.

3. Law School Examples by Course Area

In Civil Procedure, an assignment requiring students to locate a pending case in a federal district court docket, identify the current procedural posture, predict the next likely motion based on the procedural facts, and explain their reasoning forces engagement with real-time information that LLMs do not reliably have. The assignment is also inherently interesting in a way that abstract rule-application exercises often are not because students are working with a live dispute, not a hypothetical one. Students may then be required or encouraged to discuss the results of their research in class or on a required discussion board in Canvas or other learning management system (LMS).

In Criminal Law, a reflection assignment following a courthouse observation session, where students attend a real arraignment or sentencing proceeding, combines personal presence with doctrinal application in a way that is resistant to LLM completion (though, of course, the student could give the LLM the specifics and have it do the work of writing the reflection).

In Property, an assignment asking students to research the chain of title for an actual parcel of real property in their county and writing a brief title analysis involves county recorder databases and local deed formatting conventions that are well outside the reliable knowledge base of a general-purpose LLM. Professor Bliss at Sturm College of Law notes that students in a Property course that integrated GenAI found the GenAI's limitations particularly instructive when the LLM raised novel arguments or interpretations that did not align with actual statutory or case authority.¹¹

C. GenAI-Integrated Assignments

GenAI-integrated assignments require students to use LLMs as part of the assignment itself. The learning objective is not to produce a particular document but to develop critical judgment about AI-generated content in order to understand what LLMs do well, what they do poorly, and how to supervise them effectively. For future lawyers, discernment (supervisory judgment) is important to practice early before the stakes could harm real people. Model Rule 5.3 requires competent oversight of non-lawyer assistance, and as legal AI tools become

¹¹ Bliss, *supra* note 7, at 148.

standard in practice, the capacity to evaluate and correct GenAI output will be as essential as the capacity to evaluate and correct the work of a junior associate.¹²

Bliss, whose early pedagogical experiments introduced AI-integrated exercises to a first-year Property course in Spring 2023, found that students responded favorably: seventy percent found AI use helpful for their learning, while eighty percent agreed that all law students should be aware of and learn to use GenAI systems. Students specifically noted that AI-based assignments forced them to be more precise in their legal analysis.¹³ The precision requirement was a direct consequence of needing to evaluate GenAI output critically rather than accepting it wholesale.

1. The Critique-First Approach

One of the most pedagogically productive forms of GenAI integration is assigning students an evaluation of AI-generated content rather than to produce their own. Bliss's first AI-integrated assignment asked students to assess ChatGPT's answers to a trespass law quiz, identifying errors and omissions in the AI responses.¹⁴ The assignment is elegant for several reasons. It requires substantive legal knowledge; students cannot identify an error they do not understand. It makes LLM limitations visible and concrete rather than abstract. And it positions the student as a supervising expert rather than a GenAI user, which is exactly the professional posture law graduates need to develop.

A variant of this assignment can be used in Legal Writing courses. Students are instructed to prompt the LLM to draft a section of a legal brief or memorandum of law. Students are then required to revise the LLM's work product to meet professional standards. The revision becomes the assessment artifact. Students must identify and correct hallucinated citations and fix logical gaps in the analysis or argument. The work product they submit is theirs; the GenAI output is the starting point and they are graded on whether they improved that GenAI output or not.

¹² A.B.A. Comm. On Ethics & Pro. Resp., Formal Op. 512, Generative Artificial Intelligence Tools (July 29, 2024) [hereinafter ABA Formal Opinion 512].

¹³ Bliss, *supra* note 7, at 148–49 (reporting student survey results from a Spring 2023 Property course experiment at Sturm College of Law, University of Denver, in which AI-integrated exercises were designed to teach doctrine and legal analysis skills while helping students understand emerging AI tools).

¹⁴ *Id.* at 133 (describing the first AI-integrated assignment, which asked students to assess ChatGPT's answers to a trespass law quiz and identify errors and omissions in the AI responses).

2. Prompting Assignments

Prompting assignments evaluate students not on the document they produce but on the prompt they design as well as their critique of what the LLM generates in response. Writing an effective legal prompt requires the same analytical skills as writing a legal document. Prompting well means knowing what legal question you are asking, and getting that right requires the same attention to jurisdiction and facts that any competent legal analysis demands. A student who produces a vague, poorly structured prompt has revealed something about their legal thinking that a polished AI-assisted essay would have concealed. Dr. Kirsten Davis, at Stetson University College of Law, has observed that prompting assignments might give legal writing faculty direct access to student reasoning in ways that polished final products never have. In other words, in these types of assignments, the prompt itself as a window into the student's legal analysis.¹⁵

For faculty accustomed to assigning traditional research memos, the reverse prompt assignment addresses a common objection head-on: if students are going to use GenAI in practice, why not teach them to use it well, and why not grade the quality of its use? Davis's framing in the legal writing context is appropriate. Prompting as assessment, not merely as skill, gives faculty a window into student reasoning that final written products have never provided.¹⁶

3. GenAI Collaboration with Documentation

A third approach requires students to use an LLM as a research or drafting collaborator while documenting the collaboration in ways that make the student's intellectual contributions visible. Students submit the final work product alongside a process log recording which prompts they used, what the LLM produced, what they accepted, what they rejected, and why. The process log is graded alongside the final product and (sometimes) at equal weight.

For these types of assignments, the professional responsibility analogy is to supervision. A partner who assigns a client matter to an associate, and who later

¹⁵ Kirsten K. Davis, *A Provisional Definition of "Legal Writing Scholarship,"* STETSON L. REV. Unending Conversation, Spring 2024, at 7 (suggesting that prompting as an assessment, not merely as a skill, could give legal writing faculty direct access to student reasoning—specifically, a window into how students understand rule synthesis—that polished final products have never provided).

¹⁶ *Id.* (noting that prompting "might make visible the thoughts and questions legal writers use to guide their writing" and could lead to "a renewed commitment to process-oriented legal writing courses").

signs the brief, has not outsourced responsibility. They have supervised and taken accountability for the final product. A student who can articulate, in their process log, why they rejected an LLM's framing of an argument in favor of their own is demonstrating exactly the supervisory judgment that Model Rule 5.3 requires. A student whose process log shows wholesale adoption (or only slight edits) of LLM output without critical engagement has demonstrated something else entirely. And the grade should reflect it.

4. Practical Examples by Course Area

In Constitutional Law, an LLM-critique assignment can ask students to generate a constitutional argument on a live policy question, for example, the constitutionality of a recent federal statute. Students are directed to use a specified LLM, then analyze the argument for doctrinal accuracy, logical coherence, and rhetorical effectiveness. The critique requires the student to know constitutional doctrine well enough to catch errors, which is the same cognitive task as writing the argument from scratch, accomplished through a different path that also trains critical evaluation skills.

In Torts, an integrated assignment might ask students to use an LLM to generate a demand letter in a personal injury scenario. The student can then be tasked with revising it to satisfy professional standards. Specifically, the student must identify any factual claims the LLM made that are not supported by the hypothetical record, and write a brief memo explaining the revisions. The memo is the primary assessment artifact; it shows the student's doctrinal knowledge, writing judgment, and capacity to supervise a non-expert tool.

In a Lawyering Skills or Legal Writing course, an integrated sequence assignment can move students through the entire litigation document life cycle with LLM assistance at each stage. The assignment can require students to proceed through the steps from initial research to issue identification to argument organization to draft production, while requiring documented decision-making at each step. This type of sequencing models professional practice more accurately than any hypothetical-only assignment can, because the students are managing a real tool with real limitations.

D. GenAI-Transparent Assignments

Unlike the other three categories, transparency is not primarily about the level of GenAI use. It is about the relationship between use and disclosure. A

GenAI-transparent assignment can have any level of permitted AI involvement, from none to extensive. What distinguishes it from other types of assignments is the requirement that students document and disclose whatever AI use they engaged in, accurately and completely.

The idea is that the habit of disclosure is itself a professional skill. Lawyers working with AI-assisted research and drafting tools will increasingly be required to identify GenAI contributions to work product, whether by court rules, client agreements, or professional conduct guidance.¹⁷ Training students to disclose GenAI use in their academic work is professional responsibility training as much as it is assignment design.

Research on GenAI use disclosure in higher education suggests that the compliance challenge is real. A King's Business School study found that seventy-four percent of students failed to declare GenAI usage despite declaration being a mandatory requirement of a coursework coversheet.¹⁸ Fear of academic repercussions, ambiguous guidelines, inconsistent enforcement across courses, and peer influence were among the reasons cited by the students. The implication for law faculty is that transparency policies work only when students are assured that they are truly non-punitive for honest disclosure. They must also be consistently enforced and clearly distinguished from policies that prohibit use entirely.

1. Designing a Workable Disclosure Requirement

Effective GenAI transparency requirements share several features. They ask students to disclose specific information rather than just answering a yes/no question about whether AI was used. They make clear what the consequences of accurate disclosure are (none, or even extra credit for sophisticated use) and what

¹⁷ See, e.g., *Mata v. Avianca, Inc.*, No. 22-cv-1461, 2023 WL 4114965 (S.D.N.Y. June 22, 2023) (sanctioning attorneys for submitting AI-generated brief containing fabricated citations without verification); see also Cf. Perlman, Andrew M., *The Legal Ethics of Generative AI*, 57 Suffolk U. L. Rev. 345 (2024) (analyzing the professional responsibility implications of AI-assisted legal work product). Multiple federal courts have now adopted local rules requiring disclosure of AI use in court filings. For a list of rules, guidelines, policies, orders, etc., see <https://airtable.com/appKUCriCQDI1BxIV/shrflAPpNKaNMnacR/tblNmp6mff8CzLuQD>

¹⁸ Ellis J.B. Furze et al., *Addressing Student Non-Compliance in AI Use Declarations: Implications for Academic Integrity and Assessment in Higher Education*, Assessment & Evaluation in Higher Education (2024) (reporting 74% non-compliance rate with AI use declaration requirements at King's Business School, attributing non-compliance to fear of academic repercussions, ambiguous guidelines, inconsistent enforcement, and peer influence).

the consequences of inaccurate disclosure are (academic dishonesty, not merely inadvertent omission).

A practical disclosure template might ask students to record:

- the tool used
- the purpose for which it was used
- the approximate nature of the prompt or query
- whether and how they evaluated or modified the output
- a one-sentence statement of what they believe they learned or failed to learn from the interaction.

That last element takes thirty seconds to write but it pushes students toward true metacognitive reflection rather than rote compliance.

The AI Assessment Scale (AIAS), now in its second iteration and adopted by hundreds of institutions worldwide, offers a useful framework for structuring these expectations. The AIAS defines five levels of GenAI integration, ranging from “No AI” through intermediate stages of planning assistance and content generation to full use with documentation. Importantly, the AIAS does not assign a hierarchy to these levels (no level is “better” than another), and the appropriate level of AI use permitted depends entirely on the learning objectives the assignment is designed to address.¹⁹

2. The Citation Analogy

The most effective framing for GenAI use disclosure, especially with law students who have already been drilled on citation requirements, might be the citation analogy. When a student quotes a source without attribution, we call it plagiarism. When a lawyer uses someone else's work product without disclosure, we call it potentially dishonest conduct under Rule 8.4.²⁰ The lesson in both cases

¹⁹ For the updated version of the AIAS framework, see Mike Perkins et al., *Reimagining the Artificial Intelligence Assessment Scale: A Refined Framework for Educational Assessment*, 22(7) J. U. TEACHING & LEARNING PRACTICE (2025) (introducing an “AI Exploration” level and clarifying that the AIAS does not assign hierarchy to its five levels). For implementation guidance, see also Allie Brandriet & Eva Latterner, *The AI Assessment Scale: Reflecting on AI in Assignment Design*, UNIV. OF IOWA CTR. FOR TEACHING (2024), <https://teaching.center.uiowa.edu/ai-assessment-scale-reflecting-ai-assignment-design> (last visited Mar. 9, 2026).

²⁰ MODEL RULES OF PRO. CONDUCT r. 8.4 (A.B.A. 2023) (requiring lawyers with direct supervisory authority over nonlawyer assistants to make reasonable efforts to ensure the nonlawyer's conduct is compatible with the lawyer's professional obligations).

is that attribution is owed to contributions that are not your own, and the failure to attribute is a misrepresentation of authorship.

In addition to the GenAI use log provided below, with Claude's help, I created a [GenAI Use Disclosure Form](#) for students to complete and submit. It creates a PDF for easy upload to your school's LMS, such as Canvas.

Figure 8:3 – [GenAI Use Disclosure Form](#)

ACADEMIC INTEGRITY · PROFESSIONAL RESPONSIBILITY

The GenAI *Use Disclosure* Form

A transparency record to file with your assignment submission.

Form 8.5
April 26, 2026

§

Use this form to disclose any generative AI tools (LLMs, image generators, voice tools, research assistants, and the like) you relied on while completing this assignment. Honest disclosure is the point. Telling your instructor what you actually did with AI is treated as a feature of professional practice, not a confession.

Disclosure rehearses the duties practice will require of you. **Model Rule 3.3** (candor to the tribunal) is the rule lawyers violate when they file briefs containing AI-hallucinated cases that do not exist, a now-familiar way to draw sanctions. **Model Rule 1.1** (competence) requires understanding the tools you use, and **Model Rule 5.3** (responsibility for nonlawyer assistance) treats an unsupervised LLM the same way it treats an unsupervised paralegal. The habit you are building here is the habit you will need at the bar.

IF YOUR DISCLOSURE IS ACCURATE

Nothing happens to your grade because of the disclosure itself. Sophisticated, well-documented use may be rewarded under your instructor's policy.

IF YOUR DISCLOSURE IS INACCURATE OR OMITTED

Treated as academic dishonesty under the Honor Code, not as inadvertent oversight. Sanctions follow accordingly.

The GenAI Use Disclosure Form was created using a Claude Artifact. To learn how to create an Artifact, see Appendix J.

Table 8-1: - GenAI Use Log

Professional Responsibility Note. Completing this log is an exercise in the supervisory judgment Model Rules 1.1 and 5.3 require of practicing lawyers. Rule 1.1 requires competent use of relevant technology, including understanding its benefits and risks. Rule 5.3 requires lawyers to supervise non-lawyer assistance (which includes AI tools) and to take responsibility for the work product they produce with that assistance. A lawyer who accepts AI output without critical evaluation has not supervised; a lawyer who can articulate why they accepted, modified, or rejected that output has. This log is practice for that accountability.

Task Stage	Tool Used	Prompt(s) Submitted	Output Summary	Student Edits / Rejections	Reason for Acceptance or Rejection
e.g., Initial research; issue spotting; drafting argument	e.g., ChatGPT-4o; Claude; Westlaw AI	Record the prompt as submitted, including any context provided	Brief description of what the tool produced	What did you change, correct, or discard?	Explain your supervisory judgment — why did you accept or reject the output?

Reflection (required): In one or two sentences, describe what you learned (or failed to learn) from your use of AI on this assignment. Where did the tool help you think more clearly? Where did it lead you somewhere you had to correct?²¹

²¹ John Lande, *Solving Professors' Dilemmas About Prohibiting or Promoting Student AI Use*, U. MISS. S. L. LEGAL STUDIES RESEARCH PAPER NO. 2025-53 (2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5841522 (last visited Mar. 9, 2026).

GenAI output is not, at the moment, copyrightable as the product of a human author, but the attribution principle applies for different reasons.²² A student who submits an AI-generated analysis without disclosure is misrepresenting to the faculty member the nature of their intellectual contribution. From a professional responsibility view, that is a form of deception that the student will not be permitted to engage in as a licensed lawyer. Making that connection explicit in assignment instructions gives students a reason to take disclosure seriously so it is not about fear of punishment but instead about professional identity formation.

3. Disclosure Across the Other Three Categories

Disclosure is not a stand-alone category so much as a requirement that can and should apply to assignments in any of the other three categories. A GenAI-prohibited assignment benefits from a disclosure requirement because it creates an affirmative record of compliance rather than leaving compliance entirely to the honor system. A GenAI-resistant assignment benefits because it signals to students that the faculty member is paying attention and will take disclosure seriously. A GenAI-integrated assignment benefits because the disclosure requirement is what makes the integration pedagogically coherent. Without it, the assignment is just a license to use a tool without accountability.

The AI-Resistant Pedagogy Toolkit's observation that "ambiguity is the enemy of academic integrity" is important to remember.²³ Students who face inconsistent, vague, or unenforced disclosure requirements across their courses will rationally treat disclosure as optional. Law faculty who want students to develop the habit of professional accountability have to model it by being clear and consistent about what disclosure means and what it is for.

²² The U.S. Copyright Office has determined that AI-generated content without meaningful human authorial contribution is not eligible for copyright protection. See U.S. Copyright Office, Copyright and Artificial Intelligence, Part 1: Digital Replicas (2023); U.S. Copyright Office, Copyright and Artificial Intelligence, Part 2: Copyrightability (2024). For legal analysis of the attribution and authorship questions, see Richard M. Re, *Artificial Authorship and Judicial Opinions*, 92 GEO. WASH. L. REV. 1558 (2024), available at <https://www.gwlr.org/wp-content/uploads/2024/12/92-Geo.-Wash.-L.-Rev.-1558.pdf> (last visited Mar. 9, 2026).

²³ Ybarra Evans, *supra* note 1 ("Ambiguity is the enemy of academic integrity. Transparent policies build trust and shift the classroom culture from one of adversarial detection to one of shared responsibility for learning.").

E. Choosing Your Category: A Practical Decision Framework

The choice among these four categories is not a one-time curricular decision. It is an assignment-by-assignment judgment that should track the learning objective, the course context, and the practical enforcement environment. Some things to consider:

- What cognitive skill am I trying to assess? If the answer is independent legal analysis under conditions that model bar-exam or early-practice performance, prohibition with structural support works best. If the answer is critical evaluation of AI output, integration is appropriate. If the answer is professional accountability norms, transparency is the primary tool. If the answer involves authentic engagement with real-world legal materials that LLMs cannot access reliably, resistance is the design principle.
- Can I enforce what I am requiring? A prohibition that cannot be credibly enforced is worse than no prohibition, because it disadvantages compliant students and creates a paper trail of unenforced rules. If enforcement requires structural changes to the assignment format, make those changes or choose a different category.
- What am I communicating about professional norms? Every assignment policy is an implicit statement about what kind of lawyer the student should be practicing to become. A prohibition without explanation teaches that AI is dangerous and an integration without critical framing teaches that AI output is reliable. A transparency requirement without follow-through teaches that disclosure is performative. The policy is only as good as the reasoning behind it, and that reasoning belongs in the assignment instructions.

Finally, none of these categories requires unanimity within a faculty body. A law school where some courses prohibit GenAI, others integrate it, and others require transparency across varying use models the varied professional environments students will enter. What would be incoherent is having no policy at all, or having policies that are never explained in terms that connect to why they matter for the lawyers the students are becoming.

See Chapter 5 for syllabus policy suggestions on GenAI use. You might also try the interactive [GenAI Policy Drafting Tool](#).

CHAPTER 9: Assessments That GenAI Can't Take for Them

A. Introduction

You know the assessment conversation that plays out every semester in faculty lounges and on academic listservs. Someone proposes a creative alternative to the traditional final exam, and within minutes, the objections arrive: "My class has 80 students." "I don't have time to grade portfolios." "Oral exams? Are you serious?" "The administration will never approve it."

That's fair. Those objections reflect real constraints, and dismissing them as mere resistance to change would be both condescending and wrong. But GenAI has shifted the underlying calculation in ways that make revisiting the old debates necessary. Do traditional exams remain viable in practice when students have access to, and are actively using, sophisticated GenAI tools that can generate competent legal analysis in seconds?

In other words, if your current assessment method can be completed satisfactorily by a student who spends 90% of the allotted time consulting with Claude or ChatGPT, you have an assessment design problem as well as a student integrity problem. After all, studies have shown that ChatGPT can pass the bar exam.¹

The good news is that alternative assessment methods, which were once dismissed as impractical luxuries, have become more feasible precisely because of the same technology that threatens traditional exams. As discussed throughout this book, GenAI can help with grading, rubric creation, and feedback generation, making some alternative approaches less time-intensive than they once were. At the same time, certain assessment types have become more valuable because, as discussed in the last chapter, they test skills that GenAI cannot easily replicate:

¹ Pablo Arrendondo, Q&A with Sharon Dricoll and Monica Schreiber, *GPT-4 Passes the Bar Exam: What That Means for Artificial Intelligence Tools in the Legal Profession*, Stanford Law School Blogs (Apr. 19, 2023), <https://law.stanford.edu/2023/04/19/gpt-4-passes-the-bar-exam-what-that-means-for-artificial-intelligence-tools-in-the-legal-industry/>, (last visited Jan. 28, 2026).

real-time reasoning, spontaneous analysis, interpersonal communication, and demonstrated learning processes.

What follows is not a comprehensive taxonomy of every possible alternative assessment method. Instead, this chapter offers practical guidance on five approaches that work particularly well if you want to prohibit GenAI use.

B. Team-Based Learning (TBL)

I introduced Team-Based Learning (TBL) in my first AI and the Law class. When I told students what we would do in place of quizzes and a midterm, they groaned. They've come to despise "group work," and honestly, who can blame them? They'd been burned too many times by the standard law school group (or faculty committee) project formula: the gunner does all the work while two teammates contribute minimally and a fourth disappears entirely until the day before the deadline. The gunner resents carrying the load, the slackers feel guilty (or don't), and everyone receives the same grade regardless of effort. Law students are smart people, and they've learned through painful experience that "group work" usually means "unequal work with equal credit."

Though the TBL discussion here draws on my AI and the Law course, the technique works equally well in traditional doctrinal courses. For example, my colleague, Professor Wes Moore, uses TBL in his Criminal Law and Evidence courses. The AI and the Law course is more doctrinal than its title may suggest: aside from one session on prompting and CustomGPTs, the content spans legal ethics, AI regulatory frameworks, privacy law, products liability, and intellectual property.

But TBL is different, and after my students went through the process, they loved it and found they learned more and it stuck. The structure made the difference. Rather than assigning a single project and hoping students figure out how to collaborate, TBL builds accountability into every class session through a specific sequence that gives the gunners a productive outlet, keeps the free-riders from hiding, and rewards actual collaboration rather than just tolerating it.² The method also happens to be remarkably resistant to GenAI shortcuts. The

² Rachel E Stein et al., *Student Accountability in Team-Based Learning Courses*, 44 TEACHING SOCIOLOGY 1, 28-38 (Jan. 2016), available at <https://journals.sagepub.com/doi/10.1177/0092055X15603429> (last visited Mar. 10, 2026).

combination of secure, timed individual testing done in class,³ with immediate team discussion with live feedback makes it nearly impossible for students to outsource their thinking to GenAI tools. Even if a student consulted ChatGPT while preparing, the individual readiness test reveals whether the student actually understood the material, and the team discussion exposes whether each student can defend and adapt his analysis under pressure.

Instead of traditional quizzes spread throughout the semester, I structured the course around four comprehensive TBL sessions, each following the same sequence. Students prepare individually or in study groups through assigned readings and an introductory class session. In the next class meeting, they take individual readiness tests (iRATs): ten multiple-choice questions, unscored at this stage. The same questions then reappear in team readiness tests (tRATs), where groups work through the material together with immediate feedback via electronic or paper scratch cards. Faculty teaching TBL for the first time may not fully anticipate what happens next: the room gets loud, because students are arguing about why "c" is the correct answer.

Then I open the team application exercises where teams tackle complex problem(s) at the same time while I (virtually and physically) circulate to observe and challenge their reasoning.

Students with extended-time accommodations require some advance planning, since the iRAT and tRAT run in sequence during the same session; a student finishing the individual portion separately will need a clear protocol for rejoining the team, whether that means joining the tRAT already in progress or waiting for a designated transition point. I ask the students complete the iRAT ahead of time, but to attend class on time with everyone else to protect the confidentiality of their accommodation. I know what you're thinking—but what about GenAI use. Paper administration with devices put away handles most of the in-class exposure, and accommodation students testing separately with a proctor and paper materials are in essentially the same position as those in class. Alternatively, administering the exam in the software in a proctored room also addresses the problem of unauthorized GenAI use. Additionally, the tRAT itself functions as a secondary check: a student who queried an LLM during the iRAT

³ Students with accommodations can opt to take the individual assessment (iRAT) before class on the same day as the TBL exercise.

without actually working through the questions will have a difficult time defending those answers when the team starts arguing about them. To protect the integrity of the exercise, I suggest working with Disability Services as I do.

1. Individual Preparation

Students receive introductory assigned readings in the days before the TBL class. For my AI and the Law class, this might be reading about frameworks for regulating AI or about bias in AI systems. The key difference from traditional assigned reading is that students know they'll be tested individually at the start of class. For students who have extra time or quiet room accommodations, I open the iRAT the morning of the class with admonitions not to discuss it with other students. This workflow transforms the reading from optional background into essential preparation.

a. Individual Readiness Assurance Tests (iRATs)

iRATs begin the TBL class session, all of which appear on the course syllabus. I “open” a 10-question multiple-choice quiz covering the previous week’s lessons and readings. Students must complete the iRAT in 20 minutes (students with accommodations get the extra time, if appropriate). Using TBL software (I used LAMS; my colleague uses InteDashboard), I can view progress through the quiz in real-time including performance and duration.

The questions require actual understanding of the material. Students do not receive a grade on the iRAT. Instead, they move to the tRAT when I open access to it. I wait until all students have completed the quiz or the time limit has ended

2. Team Readiness Assurance Tests (tRATs)

tRATS follow immediately. Students receive the exact same test again, but this time they work through it with their assigned team (typically 5-6 students per team). Here's where the magic happens. Teams must discuss each question, debate the correct answer, and reach consensus. As I said, the room can get very loud. One person is designated as the lead, but the only power the lead has is to mark the scratch-off cards (similar to lottery tickets) in the software (or on paper) that provides immediate feedback. If they scratch off the correct answer on their first attempt, they get full points. If they're wrong, they scratch another answer and get

partial credit if correct on the second or third attempt. The immediate feedback is transformative because students don't wait days to discover they misunderstood something; they know instantly, while the material is fresh and they're motivated to figure out where their reasoning went wrong.

Research indicates that teams generally achieve significantly better results than even their best individual members, outperforming them by roughly 14%. Even the least effective team tends to surpass the highest-achieving individual.

The team discussion during this phase gets intense. Students who were certain of their individual answer discover that teammates interpreted the material differently. They have to articulate and defend their reasoning and either convince teammates or be convinced themselves. Students who think they aced the individual test sometimes find themselves surprised when teammates present compelling alternative interpretations they hadn't considered. The most thoroughly prepared students discover that confidence doesn't equal correctness when peers who prepared just as carefully challenge their reasoning with textual evidence that the first student overlooked.

Some students are more vocal and can intimidate other team members. I require that every student must speak once before any student can speak twice.

Watch this video from the University of Texas at Austin to see what the tRAT looks like in practice.⁴



⁴ *Team-Based Learning*, UT AUSTIN CTR. FOR TEACHING AND LEARNING (2013) <https://vimeo.com/51713733?fl=pl&fe=cm> (last visited Mar. 10. 2026).

The empirical results reinforce what students experience in these debates: team scores consistently surpass individual scores. In my classes, teams typically score 15-20 percentage points higher than the average of their members' individual scores, and it's not uncommon for teams to achieve perfect or near-perfect scores

Specialized software is not required for TBL implementation. Though digital platforms like InteDashboard can facilitate administration and data tracking, the core TBL method works well with paper-based materials. Cognia, the company that developed InteDashboard, produces customized physical scratch-off cards (Immediate Feedback Assessment Technique cards, or IF-AT cards) specifically designed for team readiness testing. These cards provide the instant feedback mechanism above without requiring any technology beyond the cards themselves.⁵

even when no individual member scored above 85%. Students learn to trust the team process because they can see in their own scores that it works.

I can hear the learning happening as I walk around the room: "Wait, but doesn't the third answer say..." and "No, that's not what they meant; look at the example they gave..." Students can argue passionately because the team's score depends on getting the right answer.

Once the tRAT wraps up, I share overall results with the class and then do something that reliably produces better discussion than anything I could manufacture: when one team missed a question that another team nailed, I ask the team that missed the question to explain its rationale, then ask a successful team to explain their reasoning to the team that got it

wrong. Peer explanation turns out to be more persuasive than faculty explanation, which should surprise no one who has ever watched a study group in action.

3. Clarification and Appeals

Clarification and appeals happen after teams complete the test. Teams can formally appeal any question if they believe their answer should be accepted. I moderate as the teams argue about whether the appeal should be granted. This process teaches precision in legal argument (you can't just say "we think this answer should count"; you must prove it). It also gives me direction to where the materials were ambiguous or where my question was poorly worded. If the teams

⁵ To purchase, go to: <https://www.cognialearn.com/ifat>.

make a compelling case, I might accept an additional answer or throw out a problematic question entirely.

For classes that meet twice a week, the iRAT and tRAT fit comfortably in a single class session, which means you haven't sacrificed a week of instruction to run them. And, more importantly, the students learn more than if I had done a Socratic lecture because all of them have to engage.

4. Team Application Exercises

After we've completed the iRATs and tRATs, we move onto team application exercises. Again, the students work in their assigned teams. Now that everyone has demonstrated baseline understanding of the material (through the readiness tests and discussion), teams tackle complex problems that require them to apply what they've learned to realistic scenarios. The best exercises are messy, multifaceted problems that mirror what lawyers face (and that GenAI helps me draft).

For example, after covering AI and employment discrimination law, I might give teams a detailed fact pattern about a company using an AI hiring tool that appears to have disparate impact. Teams receive documents including the company's justifications for the tool, statistical evidence of hiring patterns, excerpts from the tool vendor's marketing materials, and relevant portions of Title VII and EEOC guidance. Their task isn't just to identify potential legal issues (that's too easy after the readiness tests). Instead, they must advise a specific client (sometimes the company, sometimes a rejected applicant, sometimes the EEOC) on strategy. They must evaluate the strengths and weaknesses of different legal theories, and make specific recommendations with supporting reasoning.⁶

Teams work through these exercises using collaborative digital documents in the software (for example, [InteDashboard](#)) that automatically tracks individual contributions, showing me exactly who drafted each section of the team's analysis (sort of like Google Docs). Far from feeling like surveillance, students report that this feature improves team dynamics. Everyone knows that his or her individual work is visible, which eliminates the free-rider problem before it starts. Teams

⁶ Professor Nancy Rapoport at UNLV William S. Boyd School of Law, who graciously and generously gave her time to comment and proof this book as I was writing it commented, "What I love about this is that it also gives the students experience on a performance exam," which is true and something I hadn't thought of.

divide work more equitably because members can't claim credit for portions they didn't write, and I can provide individualized feedback even on collaborative assignments. I can even chime in in real-time if I see a team headed down the wrong path or if I see one student dominating the conversation. I *virtually* circulate constantly during this phase, but I'm not answering questions or giving hints. Instead, I'm challenging their reasoning: "You've concluded the company has a strong business necessity defense – what evidence would you need to verify that claim?" or "Your analysis assumes the AI tool's decisions are entirely opaque, but what if the company can explain its logic?" or "Team three reached the opposite conclusion on this issue; what would you say to rebut their argument?"

This Socratic probing during team-based work accomplishes several things.

- First, it forces students to defend their analysis rather than just state conclusions.
- Second, it reveals gaps in their reasoning that they might not have noticed without prompting.
- Third, it creates uncertainty where students realize that legal problems often don't have single correct answers but require careful weighing of competing considerations.
- Fourth, it models the kind of critical thinking I want them to develop; they start challenging each other the way I challenge them.
- Fifth, it makes reliance on GenAI tools virtually impossible because students must respond verbally and immediately to unpredictable questions about their team's specific reasoning.

After teams have worked through the exercise (typically 30-40 minutes), we move to simultaneous reporting. I open access to all of the application exercises answers so all students can see other team's final work product. Students comment on any team's work and ask questions in the document itself. If teams reached different conclusions, we have the teams with opposing views explain their reasoning to each other, creating a class-wide debate rather than passive listening. And, optionally, the teams rank each other's work. Also optionally, they can provide anonymous feedback on their team members' work and on their own. Questions for ranking are created by the professor and can include scoring on cohesiveness, reliability, preparation, contribution, interpersonal skills, and conflict management.

C. Learning to Collaborate – A Critical Skill for Lawyers

Legal practice has always been a team sport, even if law school treats it like a solo competition. Clients expect their lawyers to work across departments and jurisdictions, and the ability to function effectively on a team has become one of the top professional attributes that firms look for in new associates. Yet law schools continue to graduate students who, by their own admission, don't like, and are unprepared for collaborative work. Research on law students confirms what most practicing lawyers believe: legal education's relentless emphasis on individual achievement actively impedes the development of effective team players.¹⁷

The problem runs deeper than grading structures. Collaboration is a skill, and like legal research or client counseling, it requires explicit instruction rather than the optimistic assumption that students will figure it out by being placed in groups. Simply assigning a team project does not teach teamwork any more than assigning a research memo teaches legal analysis.⁸ The dynamics that derail group work — the gunner who crowds out quieter voices, the social loafer who free-rides on everyone else's effort, the introverts who stay silent when a vocal minority takes the group in the wrong direction — are predictable and addressable. TBL addresses those issues through structure.

When I see one student dominating the discussion while another student who got the question correct on the iRAT stays quiet, I talk to the quiet student individually (typically after class) to encourage her to speak up. Confidence, after all, does not equal knowledge or understanding.

a. Other Benefits

What typically converts student skeptics was the individual accountability built into every stage. Students quickly realized they couldn't coast (and neither could their peers) because their teammates would know immediately if they

⁷ Janet Weinstein et al., *Teaching Teamwork to Law Students*, 63 J. LEGAL EDUC. 36, 40 (2013) (noting that interest in teamwork skills in the legal profession is relatively recent and that much of legal training, with its emphasis on individual achievement, is an impediment to developing effective team players).

⁸ *Id.* at 41; see also Jennifer Romig, *Silence and Group Work in Legal Education*, Listen Like a Lawyer (Jan. 9, 2017), <https://listenlikealawyer.com/2017/01/09/silence-and-group-work-in-legal-education/> (last visited Mar. 10, 2026) (discussing research showing that unstructured group work produces worse outcomes than thoughtfully managed collaboration, and that dominant group members can suppress the contributions of quieter students through "production blocking" and "pluralistic ignorance").

hadn't prepared, and their individual test scores mattered for their grade. At the same time, the team component rewards authentic collaboration because working through the harder problems together (after demonstrating baseline knowledge individually) consistently produced better results than any single student could achieve alone.

Students discover that arguing with teammates about the correct legal analysis and eventually synthesizing multiple perspectives into a cohesive answer deepened their understanding far more than studying alone or sitting passively in a lecture.

By mid-semester, the same students who had groaned about group work on day one were voluntarily meeting outside class to prepare together. They'd figured out that their team's success depended on everyone showing up prepared, and they started holding each other accountable in ways I never could through grading alone. One student told me at the end of the semester, "I actually understand this material instead of just memorizing it, because I had to explain my reasoning to my team and they called me out when I was wrong." Another said, "I know I'm smart. But I didn't know how many of the other students are just as smart as I am." That second comment was both delightful and depressing—delightful because the student sounded truly excited to discover their classmates' capabilities, depressing because it took until the student's penultimate semester of law school to make that discovery. TBL forced sustained intellectual engagement that revealed what traditional classroom structures obscure, namely, that your classmates are just as capable, thoughtful, and creative as you are. Recognizing that fact transforms potential competitors into valued collaborators, which turns out to be pretty good preparation for practicing law.

The social benefits proved equally valuable. My AI and the Law course enrolls only upper-level students and who were not necessarily in the same sections in their first or second years. By 2L and 3L years they have largely self-sorted into different circles based on elective choices, clinical placements, law review commitments, and social groups formed during 1L. Many students enter the class knowing few (if any) classmates. TBL's permanent team structure creates opportunities for connection that don't otherwise exist in upper-level courses. Working intensely with the same group of people for an entire semester builds relationships that extend beyond the classroom. Students reported forming friendships and professional networks they wouldn't have developed in a

traditional lecture format, where they might sit near the same people all semester without ever having a substantive conversation.

D. A Note About Timing

The first time I used TBL, my two-credit AI and the Law course met for a single 100-minute session each week. We'd spend 20 minutes on the individual readiness test, 30 minutes on the team version, 10 minutes on clarification and appeals, then take a break before tackling the application exercise (30 minutes) and reviewing other teams' work (10 minutes). When I later taught the course in two 50-minute sessions per week, I split the sequence: the first meeting covered both readiness tests, while the second meeting handled everything else—clarification, appeals, application work, and inter-team review. Either structure completed the full TBL cycle within a single week.

E. Oral Assessments: The Socratic Method Gets Its Revenge

Oral assessment occupies an odd position in legal education. We claim to value it (look at how much class time we devote to oral discussion), yet we rarely assess it. We might count how often a student contributes to class discussion, but not the quality of those contributions. Let's face it, we all have had students who contribute often, but not necessarily meaningfully. And we have students who try to make themselves look smaller or are suddenly deeply engaged with the casebook. Many students are scared-to-death when it comes to cold-calling. I know I was. Students who grew up navigating social media's constant peer commentary may have developed a heightened sensitivity to judgment of all kinds, and when that anxiety hits the classroom, research suggests they're more likely to keep schtum rather than participate.⁹

Law students don't stay quiet in class because they're unprepared. They stay quiet because they're watching each other. Recent research on college students confirms what most faculty already suspect: students are far more worried about peer judgment than about professor criticism or even about having their opinions posted on social media. Across four of five controversial topics studied, concern about peer criticism was statistically significantly higher than concern about faculty response, which suggests that the professor's authority in the room matters

⁹ Brooke Erickson, et al., *Reluctance to Speak in College Classrooms: Academic Causes and Consequences*, 28 SOC. PSYCHOL. EDUC. 104, 104 (2025).

considerably less than the social dynamics among students.¹⁰ A student who hesitates to argue a position in class isn't necessarily afraid of you. They're afraid of the person sitting two rows back.

The same research turned up a counterintuitive finding about academic performance. Students who were reluctant to speak about a moderate number of topics outperformed both those who spoke freely about everything and those who clammed up across the board. The researchers suggest that selective caution reflects intellectual open-mindedness (a willingness to sit with complexity rather than charging in with fixed opinions), while blanket reluctance correlates with the cognitive overload and disengagement that drags down grades.¹¹ The ability to think on one's feet through oral assignments is important, because every practicing lawyer knows that the ability to articulate legal analysis verbally, under pressure, without preparation time, represents a core professional skill. Law schools should not limit that training only to students who elect to take oral advocacy or experiential learning classes.

And now with GenAI, oral assessment has become substantially more valuable. A student cannot outsource real-time verbal reasoning to a GenAI tool during a live conversation. The student must actually understand the material, and be able to respond to unpredictable follow-up questions.

1. Implementation Approaches

Oral assessments range from brief individual conferences to formal panel presentations, with corresponding variations in time commitment and learning objectives. The key is matching the format to your course goals and practical constraints.

a. The 15-Minute Individual Oral Exam

Instead of a traditional final exam, schedule each student for a 15-minute one-on-one session during exam period. Ask three to four questions of increasing difficulty, starting with a foundational concept and building to application and critique. The student has no advance notice of the specific questions. You evaluate both the accuracy of analysis and the quality of reasoning process demonstrated through verbal explanation.

¹⁰ *Id.* at 104.

¹¹ *Id.* at 104-05.

For an 80-student course, this approach requires approximately 20 hours of examination (including transition time between students), typically spread across four days. Many faculty find this format less draining than reading 80 written exams because the immediate feedback and interaction provide intellectual stimulation that dense exam reading often lacks. It takes less time and doesn't cut into your winter break or summer vacation.

Oral exams have a long history, going back to ancient educational systems in Rome, Greece, India, and other places. They were the main way Oxford and Cambridge universities tested students until the 1700s, according to research by Stephen Dobson, a professor and university official who has studied the history of these exams.¹² Faced with the increasing ease of cheating thanks to AI tools like ChatGPT, which allow students to bypass critical thinking ("cognitive off-loading") on assignments, some educators are turning to oral exams as a potential solution.¹³

At Penn Carey Law, Professor Karen Tani uses quick, informal oral "check-ins" with students before big written tests. These short discussions allow students to solidify their understanding, get answers to their questions, and feel more prepared. This simple change makes assessments feel less intimidating and more personal.¹⁴

The logistical challenge lies in scheduling. Some institutions allow faculty to spread oral exams across the exam period; others require completion within a narrower window. Most learning management systems now include automated scheduling tools that let students select their own time slots, eliminating the back-and-forth that once made oral exam scheduling nightmarish.

b. The 30-Minute Case Analysis Discussion

Give students a case or problem 24 hours before their scheduled oral exam. The case should be complex enough to require thoughtful analysis but short enough to review thoroughly in the time allotted. During the session, ask the

¹² Joanna Slater, *Professors are Turning to This Old-School Method to Stop AI Use on Exams*, WASH. POST, (Dec. 12, 2025), <https://www.washingtonpost.com/education/2025/12/12/ai-artificial-intelligence-college-oral-exam/>, (last visited Mar. 1, 2025).

¹³ *Id.*

¹⁴ Erin Bartnett, *Building Deeper Learning Through Oral Exams: Conversations That Count*, U. PENN. Center for Excellence in Teaching, Learning and Innovation, (Nov. 14, 2025), <https://cetli.upenn.edu/news-features/feature-stories/building-deeper-learning-through-oral-exams/>, (last visited Mar. 1, 2025).

student to walk through the analysis and respond to hypotheticals that test flexibility and deeper understanding.

This format tests both preparation and spontaneous reasoning. Students can certainly consult GenAI while preparing their initial analysis, but the follow-up questions during the session reveal whether the student understands the GenAI output or merely memorized AI-generated responses. When you ask, "what would change if we modified these facts?" or "how does this case relate to the doctrine we covered in week three?" you quickly discover the difference between shallow familiarity and true comprehension.

c. The Group Oral Presentation

Assign groups of three to four students a complex legal problem that requires collaborative analysis. Groups submit a written outline one week before their presentation, then deliver a 20-minute oral presentation followed by 10 minutes of faculty and peer questions. Each group member must speak and respond to questions, demonstrating individual mastery even within collaborative work.

This approach scales more efficiently than individual oral exams (a class of 80 students becomes 20 groups) while adding the assessment of collaborative skills that traditional exams cannot measure. The written outline prevents complete last-minute preparation, and requiring individual responses during Q&A ensures that all group members engaged substantively with the material.

2. Grading Rubric for Individual Oral Exams

The following rubric works for 10- to 30-minute individual oral examinations across most legal subjects. Adjust the weighting based on your course priorities.

Table 9.1: Oral Examination Grading Rubric					
Criterion	Max Points	Excellent (30 pts / 35 pts / 20 pts / 15 pts)	Proficient	Developing	Insufficient
Doctrinal Accuracy	30 pts	27–30	21–26	15–20	0–14

Analytical Reasoning	35 pts	31–35	25–30	18–24	0–17
Responsiveness & Flexibility	20 pts	18–20	14–17	10–13	0–9
Communication Clarity	15 pts	13–15	10–12	7–9	0–6

F. Portfolio Assessment: Documenting the Learning Arc

Portfolio assessment asks students to compile a curated collection of work completed over the semester, accompanied by reflective analysis explaining how their legal thinking evolved throughout the course. Done well, portfolios transform assessment from a single high-stakes moment into an ongoing demonstration of growth and learning.¹⁵

Course-level portfolios in legal education could serve as a method for evaluating student academic performance by collating various assignments and projects completed throughout a semester.¹⁶ Unlike traditional high-stakes exams, which often rely on rote memorization and single-instance performance, portfolios align with student-centered pedagogies by allowing students to demonstrate their understanding of course material through a compilation of work.¹⁷ This approach supports the "constructive alignment" of learning objectives, assignments, and

¹⁵ Deborah Jones Merritt, *Pedagogy, Progress, and Portfolios*, 25 OHIO ST. J. ON DISPUTE RESOLUTION 1 (2010), available <https://ssrn.com/abstract=1576243>, (last visited Jan. 28, 2026) (discussing how portfolios allow students to document professional development and share achievements more fully with employers through writing samples, simulations, and work-related documents). See also, Michele Pistone, *Using Portfolios for Assessment*, Best Practices for Legal Education blog, (Mar. 10, 2015), <https://bestpracticeslegaled.com/2015/03/10/using-portfolios-for-assessment/>, (last visited Jan. 28, 2026).

¹⁶ Rafe S. Zaabalawi, *Portfolios Versus Exams: A Study to Gauge the Better Student Assessment Tool*, 14 LANG TEST ASIA 28 (2024), available at <https://doi.org/10.1186/s40468-024-00296-y> (last visited Jan. 29, 2026) (finding through survey of sixty engineering students that portfolio assessment was more appropriate than traditional exams for evaluating academic performance, with superior benefits for student mental health and professional development).

¹⁷ *Id.*

assessments, enabling law schools to substantiate that graduates possess necessary professional attributes such as ethical behavior and effective advocacy.¹⁸

The implementation of portfolios directly supports compliance with American Bar Association (ABA) Standards. Specifically, Standard 314 requires the use of formative and summative assessments to provide meaningful feedback, while Standard 302 mandates the establishment of learning outcomes covering legal analysis, research, and professional skills.¹⁹ Portfolios provide a centralized platform for these assessments, allowing students to track progress and revisit feedback as they advance through their legal education.²⁰

GenAI creates interesting possibilities for portfolio assessment. Students might include examples of how they used GenAI tools at different points in the semester, analyzing what worked, what didn't, and how their prompting strategies improved. They might start building a personal prompt library or a set of CustomGPTs. The portfolio might contain early drafts and final versions of work, with reflection on the revision process and the role (if any) that GenAI assistance played. A student could include a dialogue with an GenAI chatbot alongside their own independent analysis, explaining where the GenAI was helpful and where it led them astray. Law firms would likely appreciate this demonstration of AI Fluency, including, in particular delegation, description, discernment, and diligence (as discussed in Chapter 4).

The portfolio structure makes it difficult (though not impossible) for students to fabricate their learning process. Asking for work samples from different points in the semester, combined with reflective writing that demonstrates specific knowledge of course progression, creates a higher bar for deception than a single written exam.

¹⁸ Vicki Waye, *Assessment and Portfolio Pedagogy in Law Schools*, INT'L ASSOC. OF LAW SCHOOLS, (Aug. 2015), available at <https://www.ialsnet.org/wp-content/uploads/2015/08/Waye.pdf>, last visited Jan. 29, 2026), (describing eportfolios as digital repositories maintained throughout law school that document incremental learning through diverse artifacts, providing students, assessors, and employers with holistic evidence of competency development beyond isolated assessments).

¹⁹ Outcomes and Assessments Committee, *Introduction to Learning Outcomes, Assessment, and Evaluation Standards*, A.B.A. (2025), available at https://www.americanbar.org/groups/legal_education/about/committees/outcomes-assessments/about/, (last visited Jan. 29, 2026).

²⁰ Waye, *supra* note 18.

1. Portfolio Components

Effective portfolios balance structured requirements with student choice.²¹ Too much prescription and you lose the individualization that makes portfolios valuable; too little structure and you get unfocused collections that are difficult to grade consistently.²²

a. Selection and Curation of Artifacts

Portfolios should act as repositories for work collection, allowing students to select artifacts that demonstrate their mastery of specific competencies.²³

- **Types of Artifacts**

Students can include legal briefs, client communications, drafted wills or trusts, and videos of trials or oral arguments.²⁴ In clinical or experiential courses, artifacts might include documentation of client interviews and negotiations.²⁵

- **Scope**

Portfolios may be structured around a single assignment (collecting drafts and final products), a specific unit, or an entire course.²⁶

²¹ Hongyan Yang and Rachel Wong, *An In-Depth Literature Review of E-Portfolio Implementation in Higher Education: Processes, Barriers, and Strategies*, 12 ISSUES AND TRENDS IN LEARNING TECHNOLOGIES 1, 65-101 (2024), available at <https://journals.librarypublishing.arizona.edu/itlt/article/id/5809/>, (last visited Jan. 29, 2026) (analyzing e-portfolio implementation across higher education institutions and finding that successful adoption requires addressing concerns about technology access, faculty workload, student motivation, and privacy through comprehensive training, institutional policy support, and clear pedagogical frameworks).

²² See generally Vicki Waye and Margaret Faulker, *Embedding e-Portfolios in a Law Program*, 61 J. Legal Ed. 4, 560-584 (May 2012) (examining Carnegie Report's criticism of traditional assessment methods and recommendation for e-portfolios that measure professional development throughout law programs rather than merely producing student rankings).

²³ Yang, *supra* note 21.

²⁴ Anthony Niedwiecki, *Prepared for Practice? Developing a Comprehensive Assessment Plan for a Law School Professional Skills Program*, 50 U.S.F. L. Rev. 245 (2016), available at <https://repository.law.uic.edu/facpubs/625/> (last visited Jan. 29, 2026) (discussing ABA Standards requiring law schools to develop explicit learning outcomes and comprehensive assessment plans, and providing framework for implementation through collaborative curricular planning and formative assessment of student competencies).

²⁵ FAQ, Supervised Practice Portfolio Examination (SPPE), Oregon State Bar, available at <https://www.osbar.org/docs/sppe/SPPEApplicantFAQ.pdf> (last visited Jan. 29, 2026).

²⁶ *Teaching with ePortfolios*, LSA Sweetland Ctr. for Writing U. MICH., available at <https://lsa.umich.edu/sweetland/instructors/guides-to-teaching-writing/teaching-with-eportfolios.html> (last visited Jan. 29, 2026).

- **Signature Assignments**

Course-level portfolios benefit from the inclusion of "signature assignments," which are substantial assessments connected to programmatic learning goals, such as case studies or multimedia content creation.²⁷

- b. The Reflective Component***

Reflection is considered the "heart" of portfolio pedagogy, bridging the gap between legal doctrine and the context in which it operates.²⁸

- **Metacognition**

Reflections should turn the student's activity back on itself to produce awareness of what is being learned, a process known as metacognition.²⁹ This practice helps students identify their strengths and weaknesses and plan for future learning.³⁰

- **Prompts**

Instructors should provide specific prompts to guide reflection, such as asking students to identify the most fulfilling tasks or challenging issues encountered during a week.³¹ Scaffolding these prompts helps students achieve deeper levels of analysis regarding their learning processes.³²

- **Integration**

Students should write reflections that analyze how their selected artifacts support larger claims about their professional development, effectively creating a cohesive argument or narrative about their growth.³³

²⁷ *ePortfolios in the Curriculum: Incorporating Reflection into Your Course*, Teaching & Learning Resource Ctr., THE OHIO S. U., available at <https://teaching.resources.osu.edu/teaching-topics/eportfolios-curriculum>, (last visited Jan. 29, 2026).

²⁸ Waye, *supra* note 18.

²⁹ Niedwiecki, *supra* note 24.

³⁰ *ePortfolios in the Curriculum*, *supra* note 27.

³¹ Karen Weber, et al., *I Can See Clearly Now: Using ePortfolios to Support Reflection and Engaged Learning in a Summer Internship Program*, 13 INT'L J. OF EPORTFOLIO 1, 55-67 (2023), available at <https://files.eric.ed.gov/fulltext/EJ1407265.pdf>, (last visited Jan. 28, 2026) (examining weekly ePortfolio reflections by twelve undergraduate interns and finding that the practice enabled students to document professional skills development, reflect on personal growth, and discover how experiences shifted their values and career plans, while also providing effective tools for faculty observation and assessment).

³² *Id.*

³³ Yang, *supra* note 21.

c. Feedback and Assessment Mechanisms

To ensure reliability and validity, portfolios must be assessed using clear, evidence-based criteria.

- **Rubrics**

The use of analytic rubrics is essential to standardize evaluation and make performance standards transparent. Rubrics should outline criteria for completion, selection of work, analysis, and mechanics, with performance levels ranging from "beginning" to "exemplary."³⁴

- **Formative Feedback**

Portfolios allow for frequent feedback cycles, where instructors provide suggestions for improvement on work-in-progress drafts, which is more effective for student learning than is summative grading alone.³⁵

- **Independent Review**

For high-stakes assessments, such as those used in alternative licensure models, such as in Oregon and Washington,³⁶ portfolios may be graded anonymously by independent examiners to ensure fairness and consistency.³⁷

³⁴ Mandy Wu, *The True Potential of Portfolio-Based Assessments in Education*, THE SCIENCE SURVEY (2024), <https://thesciencesurvey.com/editorial/2024/04/16/the-true-potential-of-portfolio-based-assessments-in-education/> (last visited Jan. 29, 2026).

³⁵ Zaabalawi, *supra* note 16.

³⁶ Marc Joffe, *States Pursue Alternative Licensing Pathways for Lawyers*, CATO INST. (Mar. 28, 2024) <https://www.cato.org/blog/states-pursue-alternative-licensing-pathways-lawyers> (last visited Mar. 10, 2026) ((discussing the geographic maldistribution of legal services, noting that nearly 1,300 counties across the country have fewer than one lawyer per thousand residents, and describing Oregon's development of two alternative bar licensing pathways — a supervised practice portfolio examination allowing post-graduation apprenticeships and an experiential portfolio pathway incorporating law school clinical work — as models for addressing access-to-justice gaps in underserved communities).

³⁷ Susan Bakhshian, et al., *Draft Proposal for a Portfolio Bar Examination to Demonstrate Minimum Competence*, STATE BAR OF CALIFORNIA (Aug. 7, 2023), <https://www.calbar.ca.gov/sites/default/files/agenda/agendaitem1000031434.pdf>, (last visited Jan. 29, 2026).

2. Portfolio Structure for Legal Education

Below is a suggested portfolio structure that could work across most law school courses. It was created from the various sources cited in this subsection with the assistance of Claude.³⁸

a. Required Elements (60% of grade)

- **Three Legal Analysis Pieces:** Select three written assignments from the semester (problem sets, case briefs, practice exams, research memos) that demonstrate your developing analytical skills. At least one must be from the first half of the course and at least one must be from the second half. These may be revised based on feedback received from your professor, with the original version included in your portfolio.
- **Reflective Essay (1,500-2,000 words):** Analyze your learning trajectory across the semester. Identify specific moments where your understanding deepened or shifted. Explain how your approach to legal problem-solving changed. If you used GenAI tools during the course, discuss how your use of these tools evolved and what you learned about their strengths and limitations. Complete the [GenAI Use Disclosure Form](#) and include it in your portfolio.
- **Capstone Problem Analysis:** Complete a summative problem distributed during the final exam period that requires integration of major course concepts. You may use any course materials and resources (including GenAI) but must document your process and reasoning.

b. Self-Selected Elements (40% of grade)

Choose two additional items that demonstrate your engagement with course material and your development as a legal thinker. Options include:

³⁸ This portfolio structure was developed through iterative chats with Claude 4.5 (Anthropic). The author provided the following prompt along with the research sources cited in notes 3–24: "You are a law professor who teaches AI and the Law and Law Practice Technology. You just completed a book titled 'AI and the Practice of Law in a Nutshell' for West Academic. You are now writing a book called GenAI for Teaching and Learning Law [Note to reader: the title changed since writing this prompt] to help law professors and law students integrate GenAI in an ethical and academically sound manner. Using the materials provided [footnotes 13–34] create a recommended portfolio structure for legal education for inclusion in the chapter called 'Alternative Assessment Methods.'" Claude, <https://claude.ai> (Jan. 29, 2026). All content was reviewed, revised, and adapted by the author to ensure pedagogical soundness and practical applicability.

- Case briefs or course outlines showing your synthesis work
- Supplemental research on a topic that interested you
- Visual concept maps showing doctrinal relationships
- Comparative analysis connecting course material to other courses or experiences
- Practice problems you created (with solutions) to test your understanding. If you used GenAI, include copies of the chats
- Recorded oral explanations of complex concepts
- Examples of peer teaching or study group materials you created

For each self-selected item, include a brief explanation (250-500 words) of why you chose it and what it demonstrates about your learning.

3. Grading Scheme for Portfolio Assessment

Portfolio grading requires clear rubrics to maintain consistency across students while allowing for individualized content choices. The following scheme balances structure with flexibility.³⁹

Table 9.3: Portfolio Assessment Grading Scheme

Component	Points	Evaluation Criteria
Legal Analysis Pieces (3)	30 points (10 each)	Accuracy of legal reasoning; completeness of analysis; improvement over time if revised versions submitted; integration of feedback
Reflective Essay	20 points	Depth of self-analysis; specific examples from coursework; genuine engagement with learning process; insight into strengths and areas for growth

³⁹ This portfolio assessment grading scheme was developed through iterative consultation with Claude 4.5 (Anthropic). The author provided the following prompt along with the research sources cited in notes 13–34: "You are a law professor who teaches AI and the Law and Law Practice Technology. You just completed a book titled 'AI and the Practice of Law in a Nutshell' for West Academic. You are now writing a book called GenAI for Teaching and Learning Law to help law professors and law students integrate GenAI in an ethical and academically sound manner. Using the materials provided (footnotes 3–24) create a recommended portfolio assessment grading scheme for legal education for inclusion in the chapter called 'Alternative Assessment Methods.'" Claude, <https://claude.ai> (Jan. 29, 2026). All content was reviewed, revised, and adapted by the author to ensure pedagogical soundness and practical applicability.

Component	Points	Evaluation Criteria
Capstone Problem	30 points	Accuracy of legal analysis; integration of multiple course concepts; clear explanation of reasoning; appropriate use of resources
Self-Selected Elements (2)	20 points (10 each)	Quality of work product; thoughtfulness of selection; clarity of explanation for inclusion; demonstration of learning beyond minimum requirements
Overall Portfolio Quality	10 points	Organization; presentation; completeness; coherence across all elements
Total	110 points	Graded on 100-point scale (allows for extra credit through exceptional self-selected elements)

4. The Documentation Problem

The primary challenge with portfolio assessment is not grading time (once you have clear rubrics, portfolios often grade more quickly than traditional exams because you're reviewing varied work rather than repetitive essay responses). The real challenge is getting students to maintain documentation throughout the semester.

Law students, like all of us, do not naturally save their early drafts or document their confusion and breakthroughs. They complete assignments, get grades, and move on. Portfolio assessment requires a different mindset, which means you need to build documentation into your course structure rather than expecting students to do it independently.

a. Practical strategies that work

- **Scheduled Portfolio Check-Ins:** Require students to submit portfolio drafts at weeks 5 and 10, even if incomplete, to verify they're collecting materials progressively.
- **Low-Stakes Documentation Assignments:** Assign brief weekly or biweekly "learning log" entries where students record insights, confusions, or questions. These entries can be quite informal but create a habit of reflection.

- **LMS Organization:** Create a dedicated portfolio submission folder in your learning management system from day one. Require students to upload select assignments to this folder as they complete them, not just at semester's end.⁴⁰
- **Model the Process:** Show examples of strong portfolios (with permission from former students) during the first week of class so students understand what they're building toward.

G. Process-Based Grading: Rewarding the Journey

Process-based grading shifts evaluation focus from final work product to demonstrated learning process. Instead of grading only the final memorandum or motion, you evaluate research notes, draft iterations, self-assessments, and evidence of feedback incorporation. The grade reflects not just "did you reach the right answer?" but "did you engage in the intellectual work required to develop legal expertise?"⁴¹

In the GenAI era, process-based grading has become particularly valuable because it distinguishes between students who use GenAI as a shortcut to create final products and students who use it as a tool within an authentic learning process. A student who submits only a polished final memo might have written it entirely by himself or might have had ChatGPT generate it wholesale. A student who submits research notes and multiple draft versions with track changes showing substantial revision has demonstrated the process of legal thinking regardless of what tools the student used along the way.

1. Implementation Through Scaffolded Assignments

Process-based grading works best with scaffolded assignments that break complex projects into stages. Rather than assigning "write a motion to dismiss, due in four weeks," you assign:

⁴⁰ Note that ePortfolios in Canvas will be retired on June 30, 2026. It will be replaced by Canvas Portfolio that "provides a centralized, intuitive experience." See Elizabeth Ezzi, *Introducing Canvas Portfolio: Now Part of the Core Canvas Experience*, Instructure Product Blog (Oct. 22, 2025), <https://community.instructure.com/en/discussion/660793/introducing-canvas-portfolio-now-part-of-the-core-canvas-experience> (last visited Mar. 10, 2026).

⁴¹ Liz Melleby Welch, *What is Process Grading*, N.Y.U. Teaching Excellence and Innovation (Feb. 26, 2026), <https://wp.nyu.edu/fas-edtech/2026/02/what-is-process-grading/> (last accessed March 10, 2026).

Week 1: Research plan and preliminary case identification (5% of assignment grade)

Week 2: Annotated outline showing legal theories and supporting authority (10%)

Week 3: Complete first draft (no grade, but must complete to continue)

Week 4: Peer review and self-assessment using provided rubric (10%)

Week 5: Revised draft with cover memo explaining major revisions (15%)

Week 6: Final draft (60%)

Each stage receives feedback (drafted by you, teaching assistants, peers, or GenAI-powered feedback tools) that students must engage with at each stage before proceeding.⁴² Students cannot skip directly to the final draft because earlier stages build the foundation for later work. If using GenAI-powered feedback tools, the student must also submit the chat dialogs.

The grading distribution intentionally weighs process components at 40% of the assignment grade. A student who submits a mediocre final product but demonstrates engagement with research, revision, and feedback incorporation may still earn a B grade. Conversely, a student who submits an excellent final product but showed little evidence of actual process engagement (perhaps because the student had GenAI generate most of it) will see the final grade substantially reduced by low process scores.

2. The Evidence Problem

Process-based grading requires students to generate evidence of their learning process. Some students will naturally document their work progression; others will not. Without clear guidelines on what counts as evidence, you'll receive wildly inconsistent submissions that are difficult to grade fairly.

Specify exactly what you want to see. For a legal research and writing assignment, you might require:

- Screenshots or exported search strings from research databases showing search evolution

⁴² *Id.*

- Annotated bibliography of cases considered (including cases ultimately not used and brief explanation why)
- Outline showing organizational structure at three points in the drafting process (initial conception, mid-draft, pre-final)
- Track changes version of draft showing substantive revisions (not just typo corrections)
- Brief written responses to peer or instructor feedback explaining what you will/won't change and why
- Documentation of GenAI interactions with analysis of which AI suggestions were adopted, modified, or rejected and why. Students should submit a disclosure form. The [GenAI Use Disclosure Form](#) can be used for this purpose.

Some of this documentation can be generated partially through GenAI tools (for example, asking ChatGPT to help organize research notes or explain why certain cases are relevant), but the very act of maintaining documentation ensures students cannot simply ask GenAI to produce a finished product.

3. Grading Rubric for Process-Based Assessment

Table 9.3: Process-Based Grading Rubric (Applied to Multi-Stage Legal Writing Assignment)

Process Component	Points	Excellent	Proficient	Needs Development
Research Process	10 pts	Demonstrates systematic research strategy; explores multiple legal theories; evaluates source authority and relevance; shows evolution in research approach (9–10)	Conducts competent research; identifies relevant authority; some evidence of evaluative thinking (7–8)	Research appears haphazard or superficial; limited evidence of evaluation or strategic thinking (0–6)

Process Component	Points	Excellent	Proficient	Needs Development
Organizational Development	10 pts	Multiple outline iterations show thoughtful structural evolution; clear progression from initial concept to refined organization (9–10)	Outline shows some development; organization improves from early to late stages (7–8)	Little evidence of organizational planning or evolution; appears to skip planning stage (0–6)
Drafting Process	10 pts	Multiple substantive draft iterations; significant revision beyond surface-level editing; demonstrates willingness to restructure and rethink (9–10)	Some revision between drafts; changes show thought but limited willingness for major restructuring (7–8)	Minimal revision between drafts; changes primarily surface-level or mechanical (0–6)
Feedback Integration	10 pts	Engages meaningfully with all feedback; explains revision decisions; demonstrates critical evaluation of suggestions (9–10)	Incorporates most feedback; some explanation of choices; generally responsive to suggestions (7–8)	Ignores significant feedback or incorporates suggestions mechanically without understanding (0–6)
Final Product Quality	60 pts	Excellent legal analysis, writing quality, and professional presentation (54–60)	Competent legal analysis, clear writing, professional presentation (42–53)	Significant weaknesses in analysis, writing, or presentation (0–41)
Total	100 pts			

H. Collaborative Assessment: Testing Team Skills

Collaborative assessment and TBL are not the same thing, though both involve teamwork. TBL is a pedagogical structure that runs through an entire course, building knowledge in real time through timed tests and guided discussion. Collaborative assessment, by contrast, is a project format. It is a discrete, sustained assignment that asks teams to divide complex work, manage their process over time, and produce a cohesive work product together, much the way lawyers function in practice. If TBL teaches students how to think together under pressure, collaborative assessment teaches them how to work together toward a deadline.

Legal practice has always been a team sport, even if law school treats it like a solo competition. Attorneys work in teams, consult colleagues, divide research responsibilities, and co-author documents across offices and practice groups. Yet traditional law school assessment remains relentlessly individualistic. We test what students can do alone, in silence, without resources, under time pressure, and then send them into careers where none of those conditions apply.⁴³ Collaborative assessment closes that gap by measuring whether students can divide complex work and reconcile competing views to produce coherent output as a team.⁴⁴ In the GenAI era, it carries an additional advantage: offloading all the work to an LLM becomes considerably harder when students must coordinate with teammates and demonstrate individual contributions within shared projects.

Some students will arrive in your course having already logged hours of collaborative legal work through mock trial or moot court, and it is worth acknowledging that experience directly rather than implying they have never worked with teammates before. Mock trial and moot court are incredibly valuable helping students develop advocacy skills with authentic problems. Collaborative assessment in a graded course context is different, though.

⁴³ WILLIAM M. SULLIVAN ET AL., EDUCATING LAWYERS: PREPARATION FOR THE PROFESSION OF LAW 187-88 (Carnegie Foundation for the Advancement of Teaching 2007) (criticizing legal education's overemphasis on individual and competitive learning at the expense of collaborative skills essential to practice).

⁴⁴ Janet Weinstein, *supra* n. 7, (finding that law students improved teamwork knowledge, skills, and attitudes through structured collaborative coursework and that practical team experience and faculty mentoring were the components students found most valuable).

Mock trial and moot court are self-selected, co-curricular activities where students typically gravitate toward roles that suit them. Students lean into their strengths. Nobody grades the process. Nobody requires the oral advocate to also master the evidentiary issues the brief writer handled. Collaborative assessment in a course closes those gaps deliberately. Every student has to do the research and analysis for the whole project, not just draft the piece they claimed. Contribution logs and peer evaluations exist to verify that. So does individual questioning, because a polished group product tells you nothing about whether each student actually did the work. A team can produce an excellent appellate brief with one student doing most of the heavy lifting. Collaborative assessment is specifically designed to make that invisible and unrewarded. Mock trial coaches. But collaborative assessment has a paper trail and a faculty member asking each student to explain the section someone else wrote.⁴⁵

1. Design Principles for Effective Collaborative Assessment

Collaborative assessment fails when it becomes free-rider assessment, meaning one student does all the work while others coast, or when individual accountability disappears entirely and you cannot tell who contributed what. Effective collaborative assessment prevents both problems through careful design, not optimistic hope.⁴⁶

Individual accountability mechanisms are the foundation of any valid collaborative assessment.

- **Contribution logs** require each team member to document specific contributions (including research completed, sections drafted, problems identified, and solutions proposed) submitted weekly or at project milestones. Each member then reviews and approves the entries logged by the others.
 - Logs create a contemporaneous record that is far more reliable than memory-based peer evaluations completed at semester's end.

⁴⁵ *Id.*

⁴⁶ Clifford S. Zimmerman, "Thinking Beyond My Own Interpretation": Reflections on Collaborative and Cooperative Learning Theory in the Law School Curriculum, 31 Ariz. St. L.J. 957, 972 (1999) (providing specific guidelines for teaching collaboration in law school and noting that cooperative learning cuts against the individualistic traditions of legal education).

- Students who know their contributions are being tracked in real time tend to make more of them.
- **Individual reflection papers** ask each student to analyze the team's process submitted after the collaborative work is complete.
 - Reflection demonstrates that the work product itself cannot reveal, including what went wrong and why.
 - A student who cannot write substantively about the team's process probably was not meaningfully part of it.
- **Random individual questioning** during or after team presentations tests whether each student can explain any aspect of the work, not just the section the student personally drafted.
 - If a student cannot explain a portion the student supposedly helped develop, the student's individual grade reflects that gap even when the team product is strong.
 - Unscripted questioning is a reliable indicator of true engagement and one that no LLM can answer on a student's behalf in real time.
- **Peer evaluation** asks team members to assess each other's contributions using a structured rubric, with results informing individual grade adjustments within the overall team grade.
 - Peer evaluations work best when students receive the rubric at the start of the project.
 - Knowing that teammates will evaluate their contribution changes behavior from day one.

Structured task division prevents the most common failure mode of collaborative work, where students divide the project in ways that let some members loaf while others carry the load.⁴⁷ For a legal research and writing project, every team member owns at least one major legal issue as the team's primary expert and drafts at least one section of the final document. Each also reviews and gives substantive feedback on a teammate's section. Requiring at least

⁴⁷ *Id.* at 975.

three documented team meetings keeps teams from fragmenting into parallel solo efforts that get stapled together the night before the deadline.

2. Sample Collaborative Assessment: Appellate Brief and Oral Argument

An appellate simulation can work well in upper-level courses. Three students represent either appellant or appellee in a simulated case, producing a single appellate brief (held to one brief's word limit, not three) and delivering oral argument before a mock court.

Each student becomes the team's primary researcher and drafter for at least one major issue and delivers at least one substantive section at oral argument. To verify individual contribution, each student also submits notes documenting their research and preparation, and completes a post-project reflection. The grading breakdown rewards both shared and individual performance: team brief quality accounts for 40% (the same grade for all team members), oral argument performance accounts for 30% (a combination of team and individual performance), individual contribution evidence accounts for 20% (which may vary by student based on contribution logs, peer evaluations, and reflection quality), and professional collaboration accounts for the remaining 10% (a team grade based on meeting attendance, responsiveness, and conflict resolution).

The structure creates shared incentives because the team's brief and argument performance affect everyone's grade, while individual accountability remains intact through the contribution evidence and individual argument portions.⁴⁸

3. Grading Collaborative Assessments

One of the more persistent objections to grading collaborative work is that assessing interpersonal skills is too subjective to be valid. Professor Sophie Sparrow dismantles that objection directly, and her approach is worth examining closely because it solves the subjectivity problem in a way that also deepens

⁴⁸ Sophie M. Sparrow, *Can They Work Well on a Team? Assessing Students' Collaboration Skills*, 38 WM. MITCHELL L. REV. 1162, 1178 (2012) (examining methods for explicitly assessing collaborative skills in law school settings and recommending structures that preserve individual accountability within team grading).

student investment in the process.⁴⁹ Rather than imposing external criteria, she asks students at the start of the course to identify the attributes of effective collaboration themselves, and those student-generated guidelines become the actual grading criteria. Students are not evaluated against standards handed down from the syllabus; they are held accountable to standards they created and agreed to. The result is that professional collaboration stops feeling like a requirement imposed by the professor and starts feeling like a commitment to each other.⁵⁰

Implementing the assessment mechanism Professor Sparrow uses at semester's end is worth considering. Each student receives a fixed pool of points to distribute anonymously among teammates based on how well each person met the team's guidelines over the course of the semester. Students who contributed more earn more points; students who coasted earn fewer. Equal distributions are accepted only when accompanied by detailed written justification for each teammate, which makes equal distribution more work than differentiated distribution and effectively discourages the path of least resistance.⁵¹ In a class of over seventy students, teammates awarded each other an average of 92% of available collaboration points, suggesting that most students, when given clear criteria and real stakes, rise to meet them.⁵²

I. Making the Switch: Practical Considerations

Wanting to use alternative assessment is the easy part. Institutional rules push back: exam schedules and accommodation policies set hard limits. Resources never stretch as far as the syllabus implies. Students show up expecting traditional testing because that's what they're used to after sixteen years of school. Here's how to manage the transition.

⁴⁹ *Id.* at 1166-67.

⁵⁰ *Id.*

⁵¹ *Id.* at 1174 (describing the point-distribution system and noting that requiring written justification for equal distributions effectively discourages students from taking the easy route of assigning identical scores regardless of actual contribution).

⁵² *Id.*, (A cynic might argue that students simply trade high grades to protect their own scores. Sparrow's data suggests otherwise. Students in her classes produced differentiated, specific assessments of their teammates, including critical observations about preparation and participation, and the average score of 92% reflects genuine variation rather than universal grade inflation. The requirement that equal distributions be accompanied by detailed written justification for each teammate also removes the easiest path to mutual back-scratching.)

1. Start Small

You don't need to overhaul your entire assessment structure immediately. Begin by replacing one component of your current assessment scheme with an alternative method. If you currently use a midterm and final exam, replace the midterm with a portfolio check-in or oral assessment. If you currently use a single final exam, add a collaborative project worth 30% of the grade and reduce the final exam to 70%. Test the new method with a smaller component before committing your entire grading scheme to unfamiliar territory.

Small-scale experimentation lets you work through logistical challenges without high stakes. If something doesn't work well, tweak it.

2. Communicate Clearly and Early

Students enter your course expecting traditional assessment unless you tell them otherwise, so transparency on day one is essential. For example, I first introduce TBL in the AI and the Law syllabus.

Team-Based Learning (TBL)

Overview

This course uses Team-Based Learning rather than individual quizzes and a midterm. TBL is a structured collaborative learning method that combines individual accountability with team problem-solving. It is not traditional "group work" in which labor is divided unequally allowing some to benefit from others' efforts. Every component of TBL includes mechanisms that ensure individual preparation and participation while rewarding genuine collaboration.

Why TBL?

TBL accomplishes several goals:

- Develops your ability to apply legal concepts to complex, realistic problems
- Creates skills in collaborative legal reasoning that mirror professional practice
- Makes GenAI-assisted shortcuts ineffective because learning happens through real-time discussion and questioning
- Builds professional relationships with classmates you might not otherwise know

Research consistently shows that team scores exceed individual scores by significant margins (typically 15-25 percentage points), providing concrete evidence that collaborative reasoning produces better outcomes than individual work alone.

TBL Structure

We will complete four comprehensive TBL modules throughout the semester, replacing traditional quizzes. Each module follows the same sequence:

Before Class: Individual Preparation You will receive assigned readings (cases, articles, technical materials) with specific learning objectives. You are responsible for completing these readings thoroughly because you will be tested individually at the start of class. Students with accommodations who require extra time will receive it and may complete the individual assessment prior to class.

Class Session 1: Readiness Assurance Testing

Individual Readiness Assurance Test (iRAT) – 20 minutes - You will complete a 10-question multiple-choice test individually, covering the assigned readings. Questions require application and analysis, not just recall. Your individual score counts toward your grade.

Team Readiness Assurance Test (tRAT) – 30 minutes - Immediately after submitting your individual test, you will take the identical test with your assigned team. Teams discuss each question, debate answers, and reach consensus. You will mark team answers using electronic scratch-off cards that provide immediate feedback (similar to lottery tickets).

If your team selects the correct answer on the first attempt, you receive full points; subsequent attempts receive partial credit. Research shows team scores consistently exceed individual scores because collaborative discussion catches errors and surfaces correct reasoning that individuals missed.

Class Session 2: Application and Synthesis

Clarification and Appeals – 10 minutes - Teams may formally appeal any test question if you believe your answer should be accepted. Appeals must cite specific evidence from assigned materials. All teams will debate the validity of the appeal(s). I will address all appeals immediately.

Team Application Exercise – 30 minutes Teams work through complex, realistic legal problems that require applying multiple concepts from the readings. You will work collaboratively using digital platforms that track individual contributions,

ensuring everyone participates substantively. I will virtually circulate during this time, asking challenging questions about your team's reasoning and requiring you to defend your analysis.

Inter-Team Review and Discussion – 10 minutes - Teams will share their conclusions and engage in class-wide discussion of different approaches.

Teams

You will be assigned to permanent teams of 4-6 students (depending on class size) for the entire semester. Teams are formed to maximize diversity of backgrounds, experiences, and perspectives. You cannot switch teams except in extraordinary circumstances.

Grading

Your TBL grade (40% of your final course grade) is calculated as follows:

Individual Readiness Tests (iRAT): 25%

Team Readiness Tests (tRAT): 25%

Team Application Exercises: 20%

Individual Contribution to Team Work: 20%

Peer Evaluation: 10%

Individual contribution is assessed through: (1) digital tracking of your specific contributions to team application exercises, (2) your responses during my questioning of your team's work, and (3) peer evaluations completed by your teammates.

Expectations

Preparation: Come to class having completed all assigned readings. Your individual test score depends on it, and your teammates are counting on you to contribute meaningfully to team discussions.

Participation: Engage actively in team discussions. Challenge your teammates' reasoning respectfully, defend your own analysis with evidence, and be willing to change your mind when teammates make compelling arguments.

Professional Collaboration: Attend all team sessions, respond to teammates promptly, contribute substantively to team work, and evaluate teammates' contributions honestly and constructively.

GenAI Use: You may not consult GenAI tools or other external resources during readiness tests or application exercises unless I explicitly authorize you to do so. The

TBL structure is designed to assess your understanding through real-time reasoning and discussion.

Accessibility

Students requiring accommodations for any component of TBL should contact the Office of Disability Services as early as possible. Once the Office of Disability Services has contacted me, accommodations will be provided to ensure equivalent access while maintaining the learning objectives of collaborative assessment.

Questions?

TBL may be unfamiliar to you, and that's fine. We will discuss the process in detail on the first day of class, and I will answer all questions before our first TBL session. You can also review resources about TBL at [TBL Published Papers](#) or schedule office hours to discuss how the method works.

On the first day of class, I start by acknowledging the elephant in the room: "I know you hate group work, and you have good reasons for that." Then I explain how TBL differs from every terrible group project they've endured. I walk through the entire process emphasizing at each stage how the structure prevents free-riding and ensures everyone must prepare and participate. I don't just describe the method; I explain the pedagogical reasoning behind each component so students understand why it works and why I'm doing it.

To avoid argument, don't rely only on first-day-of-class announcements. Your syllabus should explicitly describe any non-traditional assessment methods, including:

- Detailed explanation of the assessment format (not just the phrase "oral exam" but what that means in your course)
- Grading criteria and rubrics
- Logistical requirements (scheduling, submission formats, team formation)
- Rationale for why you're using this method (students are generally more receptive when they understand the pedagogical purpose)

Address Accessibility Proactively

Alternative assessments sometimes raise different accessibility concerns than traditional exams. Oral assessments may disadvantage students with speech or language processing differences. Collaborative assessments may challenge students with social anxiety or scheduling conflicts due to work or family obligations. Portfolio assessments may burden students with learning disabilities that affect executive function and organization.

To address those concerns, work with your institution's disability services office before implementing new assessment methods. Discuss how accommodations typically provided for traditional exams (extra time, reduced distraction environment, assistive technology) translate to alternative assessments. Be prepared to offer equivalent accommodations: perhaps a written option for students who cannot complete oral assessment, or individual assessment options for students who cannot participate fully in collaborative work.

The goal is equivalent rigor, not identical format. A student with appropriate documentation should receive accommodations that provide equal opportunity to demonstrate knowledge and skills without fundamentally altering the assessment's learning objectives.

Manage Grading Time Strategically

I know what you're thinking. Grading 80 portfolios sounds like it will take up your entire holiday break. The answer is to stop grading everything the same way. Instead, use the rubrics you've already built, and let GenAI handle the routine feedback tasks that eat your days without requiring your expertise.

Use Rubrics Rigorously: Detailed rubrics let you grade faster and more consistently.⁵³ Students receive clearer feedback, and you avoid lengthy written comments because the rubric explains evaluation criteria. Create your rubrics at the same time you design the assessment (not after reviewing student work), and stick to them during grading.

Stagger Grading: Process-based and portfolio assessments naturally distribute grading throughout the semester rather than concentrating it at term's

⁵³ See Sophie Sparrow, *Describing the Ball: Improve Teaching by Using Rubrics – Explicit Grading Criteria*, 2004 MICH. ST. L. REV. 1, 77-91 (2004).

end. So it can feel less overwhelming than facing 80 exams during exam period. Schedule process-stage grading during weeks when other course responsibilities are lighter.

Leverage AI for Routine Feedback: GenAI tools can provide quick initial feedback on *anonymized* draft submissions. It can identify common errors or generate rubric-based comments. You then review AI-generated feedback for accuracy and add personalized observations. See Chapter 7 for extended discussion of using GenAI in grading workflows.

Calibrate Standards Appropriately: Not every assignment requires the same grading depth. Early process stages might receive completion-based grades (satisfactory/unsatisfactory) with brief feedback on major issues. Later stages and final products receive more detailed evaluation.

A Note on Institutional Constraints

Many institutions impose requirements that limit assessment flexibility: mandatory scheduled final exams, minimum exam duration requirements, restrictions on group work, and prohibitions on take-home assessments. Most of those policies were written before GenAI existed, and they tend to reflect historical academic integrity concerns or bureaucratic inertia rather than any coherent pedagogical theory.⁵⁴ If you face institutional constraints that prevent implementing alternative assessments, document the pedagogical rationale for your proposed changes and bring that documentation to curriculum committees or associate deans for academic affairs. The GenAI environment provides powerful arguments for reform. Traditional exam security has become increasingly difficult to maintain, and alternative methods produce more accurate measures of learning outcomes while assessing the professional skills that traditional exams cannot reach.⁵⁵ You may not win immediate policy changes, but creating institutional awareness of what GenAI does to exam integrity might help build momentum over time. In the meantime, work within existing rules to introduce whatever alternative assessment components your institution allows.

⁵⁴ See Judith Welch Wegner, *Reframing Legal Education's "Wicked Problems,"* 61 RUTGERS L. REV. 867, 893-901 (2009) (examining how accreditation structures and institutional inertia constrain law school assessment reform).

⁵⁵ See Sullivan *supra* note 43, at 6-10 (arguing that traditional law school assessments inadequately measure professional formation and practical judgment); ABA Standards 302 and 314 (requiring law schools to establish learning outcomes covering professional skills and to use both formative and summative assessments to provide meaningful feedback on student progress).

Small changes, including adding a short oral component or incorporating low-stakes portfolio check-ins move you meaningfully toward both GenAI-resistant and pedagogically sound assessment.

3. The Broader Assessment Conversation

Alternative assessment methods are not new pedagogical inventions. Law professors have experimented with oral exams, portfolios, process-based grading, and collaborative projects for decades.⁵⁶ What is different now is the external pressure GenAI creates to reconsider assessment methods that we have long defended as "rigorous" primarily because they were traditional.⁵⁷ The question worth asking is not whether alternative assessments are perfect (of course they're not), but whether they measure the learning outcomes we care about better than traditional exams do, in an environment where students have access to sophisticated GenAI tools that can generate competent legal analysis in seconds. When framed that way, the answer for many courses becomes clear.⁵⁸

Assessment should measure learning, not test-taking ability under artificial constraints (although there is space for that too given how the bar exam is currently structured). The methods described in this chapter, imperfect as they are, create better alignment between what law professors assess and what we want students to learn. They also happen to be more resistant to GenAI-enabled shortcuts than traditional exams. Law schools have to face the fact that things are different now.⁵⁹

⁵⁶ Sparrow, *supra* note 48; Deborah Jones Merritt, *Pedagogy, Progress, and Portfolios*, 25 OHIO ST. J. ON DISP. RESOL. 1 (2010), available at <https://ssrn.com/abstract=1576243> (discussing how portfolio assessment can improve the goals of legal education); Deborah L. Rhode, *Legal Education: Rethinking the Problem, Reimagining the Reforms*, 40 PEPP. L. REV. 437, 452-60 (2013) (reviewing decades of proposals for oral, practical, and portfolio-based assessment in law schools).

⁵⁷ Those of us who have administered exam platforms in several law schools across multiple courses and faculty know that the quality and difficulty of those exams varies considerably. "Rigorous" often means "timed and closed-book" rather than anything more specific.

⁵⁸ Zaabalawi, *supra* note 16; Pablo Arrendondo, *supra* n. 1.

⁵⁹ Sullivan, *supra* note 43, at 187-94 (recommending that legal education integrate assessment methods that align with actual professional practice rather than isolated individual performance under artificial conditions).

CHAPTER 10: Creating Custom AI Assistants for Legal Education

What's so great about CustomGPTs? The ability to create simulations! Yes, they can be personalized tutors and skills coaches (discussed below), but CustomGPTs may be the best thing to happen to law school simulations since the invention of the role-play exercise, which is admittedly a low bar given how badly those can go. In the simulation context, a well-configured CustomGPT does something a human classmate playing the difficult client simply cannot do reliably. It holds character (and you can change characters up) while simultaneously giving feedback. It follows the scenario with precision and resists student attempts to skip to the part where everything works out. A classmate (or friend) playing opposing counsel in a negotiation exercise might get bored or sympathize with the other side. They might forget which position they were supposed to hold. A CustomGPT configured for the role will not yield an inch until the student earns it.

- Client interviewing translates particularly well, with the GPT playing an evasive or traumatized client who only reveals information through careful, skilled questioning.
- Negotiation exercises work nicely when the GPT plays an opposing counsel holding an unreasonable position with perfectly reasonable-sounding justifications.
- Oral argument practice benefits from a configured "judge" who has read the briefs and arrives with pointed questions about the weakest part of the student's argument.
- Deposition preparation and witness examination exercises work because the GPT can be instructed to answer only what is asked and nothing more, forcing students to develop real precision in their questioning.
- Ethics counseling scenarios deserve particular attention: the situation where a longtime client calmly asks you to do something that would get you reported to the bar generates exactly the kind of productive discomfort that a hypothetical on a page cannot replicate.

Getting all of this to work, though, requires moving from the role of someone who uses these tools to someone who builds them, which is where the next section comes in.

A. From Passive Users to Active Designers

When it comes to GenAI, most law professors generally fall into one of two camps. The first camp views these tools with deep suspicion, convinced that students are using ChatGPT to cheat their way through law school. The second camp has experimented with GenAI for their own scholarship or teaching preparation but assumes that creating specialized AI tools requires coding skills they don't possess. Both camps are missing something important.

You can create custom GenAI assistants (CustomGPTs) tailored to your specific courses, research areas, and administrative needs without writing a single line of code. More importantly, you should. Creating CustomGPTs represents a shift from reactive policing of student GenAI use to proactive design of learning tools that can support your pedagogical goals. After reading the research discussed in Chapter 6 this should come as no surprise. When faculty design GenAI tools specifically for their courses, students learn more effectively and find the process more engaging. When you control the design, you control the guardrails.

It's important because our students will graduate into a profession where they'll be creating, customizing, and managing AI tools (everything from agents to vibe coding and more) for legal practice. The American Bar Association has made clear that technological competence is part of our ethical duty of competence.¹

Consider how the skill of creating CustomGPTs maps onto professional responsibility obligations. When you teach students to create a research assistant that cites sources and acknowledges uncertainty, you're teaching them to avoid the false certainty that violates Rule 3.3's candor requirement.² When you require them to disclose GenAI assistance in their work product, you're preparing them

¹ MODEL RULES OF PRO. CONDUCT r. 1.1 cmt. 8 (A.B.A. 2023) (added 2012) ("To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology").

² MODEL RULES OF PRO. CONDUCT r. 3.3(a)(1) (A.B.A. 2023) ABA Model Rule 3.3(a)(1) (requiring candor to tribunals and prohibiting false statements of fact or law).

for Rule 1.1 duty to understand the benefits and limitations of technology.³ When you build privacy protections into a case analysis assistant, you're modeling Rule 1.6's confidentiality obligations.⁴

This chapter provides practical guidance on creating these CustomGPTs, but it does so with academic integrity as the organizing principle rather than an afterthought. I'll start with terminology (because you need to understand what you're building), move through platform selection and step-by-step creation guides, and then address the ethical considerations that should inform every design choice. Throughout, I'll connect these technical decisions to the professional responsibility framework your students need to internalize.

This is not difficult. I promise. I've taught scores of faculty and students how to do this. In fact, my colleague Professor Aryele Maye and I taught law faculty how to create CustomGPTs in an AALS Section on Technology, Law, and Legal Education webinar in July 2025.⁵

B. Defining Terms: CustomGPTs, Agents, and the Spectrum Between

Before we can discuss creating CustomGPTs, we need clarity about what we're building. The GenAI field uses "agent" to describe systems with varying degrees of autonomy and capability, and the term gets thrown around loosely enough to cause confusion.

In technical GenAI discourse, an "agent" typically refers to a system that perceives its environment, acts autonomously to achieve goals, uses tools to accomplish tasks, persists across interactions, chains multiple actions without constant human input, and makes decisions about what to do next.⁶ By this definition, the CustomGPTs you'll create in this chapter aren't quite full agents. They are specialized instances of large language models (LLMs) with custom

³ MODEL RULES OF PRO. CONDUCT r. 1.1 cmt. 8 (A.B.A. 2023) (added 2012), *see also* ABA Formal Opinion 512 (2024) discusses lawyers' duties when using generative AI, emphasizing the need to understand the technology's limitations and to supervise its output.

⁴ MODEL RULES OF PRO. CONDUCT r. 1.6(c) (A.B.A. 2023) (requiring reasonable efforts to prevent unauthorized disclosure of client information).

⁵ Aryele Maye & Korin Munsterman, *Building Your First Legal Bot: A Hands-On Workshop*, AALS Section on Technology, Law and Legal Education (July 23, 2025 2pm EST), <https://www.aals.org/sections/webinars/building-your-first-legal-bot-a-hands-on-workshop/> (last visited Mar. 9, 2026).

⁶ STUART RUSSELL & PETER NORVIG, *ARTIFICIAL INTELLIGENCE: A MODERN APPROACH* (4th ed. 2020).

instructions, knowledge bases, and access to specific capabilities, but they require human initiation for each step of a workflow.

So, at one end, you have a basic chatbot that responds to single prompts with no memory or customization. At the other end, you have fully autonomous agents that can independently execute complex multi-step workflows, interact with external systems, and make decisions about when and how to act. The CustomGPTs we'll build fall somewhere in the middle. They maintain specialized instructions and knowledge across conversations, can access uploaded course materials or research databases, and use tools like web search or code execution, but they can't autonomously chain actions or interact with your learning management system without human oversight (at least not yet).

OpenAI's ChatGPT is my preferred platform.⁷ It allows you to configure a specialized version of its underlying LLM with custom instructions that persist across conversations so you don't have to prompt it over and over again. You can (and should) upload relevant documents or course materials that the assistant can reference, enable specific capabilities like web browsing or code execution, and share the assistant with students or colleagues (with some important platform-specific limitations I'll discuss).

What these assistants can do matters as much as what they can't. They can:

- follow specialized instructions persistently throughout a conversation
- reference and analyze uploaded documents and data
- (optionally) use tools like web search to find current information
- maintain context within a conversation to build on previous exchanges
- be configured with conversation starters or templates that model good questions

They cannot:

- independently execute multi-step workflows without you prompting each stage
- directly interact with external systems like your Canvas site or Westlaw account
- make autonomous decisions about when to take actions

⁷ There are many other options for creating CustomGPTs, such as Poe, Boodlebox, Microsoft Copilot Agents, and SoulMachines (that has built-in avatars), but if you learn how to create a CustomGPT in ChatGPT, that knowledge transfers to other applications.

- access real-time data automatically without being prompted to search
- call multiple tools in sophisticated sequences without human orchestration

Understanding these limitations helps you design assistants that play to their actual strengths. A CustomGPT can be an excellent Socratic teaching partner that pushes students to reason through Fourth Amendment problems, but it can't automatically monitor student submissions in your learning management system and generate feedback reports. It can help you organize research notes and flag relevant new cases, but it can't independently run citation checks across all your draft footnotes without you prompting each step. The assistants are powerful tools for extending your pedagogical reach and managing your workflow, but they're tools you direct rather than autonomous colleagues.

Table 10-1: Differences Between Chatbots, CustomGPTs and Agents)

Feature	Basic Chatbot	CustomGPT	Advanced Coded Agent
Custom instructions	No	Yes	Yes
Knowledge base	No	Yes (uploaded files)	Yes (multiple sources)
External Tool use	No	Limited (built-in)	Extensive (custom APIs)
Multi-step workflows	No	Manual (user prompts each step)	Automated
External system integration	No	No	Yes (with coding)
Requires coding	No	No	Yes

C. The Academic Integrity Framework

Creating CustomGPTs for legal education requires thinking about academic integrity from the start rather than bolting it on later. The traditional approach to academic integrity in the GenAI age has been reactive. Law schools write honor code provisions prohibiting certain uses, try to detect violations or hope students comply. This approach fails for reasons already explored in Chapter

5. But more fundamentally, it misses an opportunity to teach students how professional responsibility works.

The Model Rules of Professional Conduct assume lawyers will face situations where cutting corners would be easy and profitable, but they'll do the right thing anyway because they've internalized professional values.⁸

Building academic integrity principles into CustomGPTs requires thinking through several design questions before you write your first system prompt.

- What level of assistance moves students toward deeper understanding versus doing the thinking for them?
- How do you build in attribution and transparency rather than allowing students to pass off GenAI output as their own work?
- Where should the assistant refuse to help because the task requires the student's own analysis and effort?
- How do you create friction at the right points so students develop professional judgment about when to rely on tools versus their own expertise?

Your students will face versions of every one of these questions as practicing lawyers. When a firm hands them access to an AI research tool or a document automation platform, they'll need to evaluate it. Making your own design choices explicit and discussing them in class turns that future moment into something they've already done, which is professional responsibility training through experience which might be more likely to stick.

The framework I propose builds on three core principles drawn directly from the Model Rules.

- First, competence (Rule 1.1): The CustomGPT should help students develop understanding rather than dependence and make clear the assistant's limitations and when human judgment is required.⁹
- Second, candor and honesty (Rule 3.3): The CustomGPT should acknowledge uncertainty rather than fake confidence. It should cite

⁸ See generally MODEL RULES OF PRO.CONDUCT, Preamble & Scope ¶ 5 (2023) ("A lawyer's conduct should conform to the requirements of the law, both in professional service to clients and in the lawyer's business and personal affairs").

⁹ ABA r. 1.1, *supra* note 1.

sources so claims can be verified and refuse to generate content that students would present as their own original analysis.¹⁰

- Third, supervisory responsibility (Rules 5.1 and 5.3): The CustomGPT should model what appropriate oversight of nonlawyer assistance looks like in practice. Just as a supervising attorney remains responsible for a paralegal's work product, the student remains responsible for everything the GenAI assistant produces. The CustomGPT should prompt students to review and evaluate its output rather than pass it along unchecked and should flag when a task requires the kind of professional judgment that no GenAI tool can supply.¹¹
- Fourth, confidentiality and professional discretion (Rule 1.6): The CustomGPT should protect sensitive information in uploaded materials, warn users about what data might be retained by the platform, and model appropriate professional boundaries about what information gets shared.¹²

As an example, imagine you're creating a Criminal Procedure CustomGPT. A student uses it and asks, "Can you write my case brief for *Terry v. Ohio*?" CustomGPT designed without academic integrity principles might generate a full brief, complete with facts, issue, holding, and reasoning. A CustomGPT designed with professional responsibility principles would respond differently: "I can't write your case brief because that's your opportunity to develop the analytical skills you'll need as a lawyer. But I can help you work through the case. What do you think is the key issue the Court is deciding?" The response models several professional responsibility principles. It refuses to do work the student should do themselves (developing competence). It's transparent about why it can't help (candor about limitations). It offers appropriate assistance that moves the student toward understanding (supervision that maintains professional responsibility).

D. Types of CustomGPTs for Legal Education

CustomGPTs can serve different functions across teaching, scholarship, service, and administration. Understanding the categories helps you identify which assistants would most benefit your work and where to focus your initial

¹⁰ ABA r. 3.3, *supra* note 2.

¹¹ MODEL RULES OF PRO. CONDUCT rr. 5.1, 5.3 (A.B.A. 2023).

¹² MODEL RULES OF PRO. CONDUCT r. 1.6 (A.B.A. 2023).

efforts. Each category raises different academic integrity considerations and connects to different aspects of professional responsibility training.

1. Teaching Assistants

Teaching assistants are probably the most visible category and the one with the most direct connection to student learning. These include doctrinal teaching assistants customized with specific course materials and case law and instructed to simulate Socratic questioning. They maintain your pedagogical approach and make you virtually accessible 24/7 for all your students.¹³ Legal writing coaches can specialize in IRAC analysis, case briefing, citation format, and memo structure, providing consistent feedback based on your rubric and standards, but refuse to write the memo for the student. Exam preparation assistants help students with practice problems and issue-spotting exercises calibrated to your learning objectives and exam format. The CustomGPT will provide meaningful feedback and as many examples and explanations that the student asks for.

For examples of working CustomGPTs, with links to try them, see Appendix I.

When students use a teaching assistant CustomGPT to deepen understanding rather than replace it, they're developing the foundational competence that professional responsibility requires.¹⁴ When they learn to verify the assistant's explanations against primary sources, they're practicing the supervision duties lawyers owe when using technology.

2. Research and Scholarship Assistants

Research assistant CustomGPTs serve your own scholarly work rather than your students, which means the design shifts. Instead of building in Socratic guardrails to prevent over-reliance, you're building in organizational discipline to keep complex projects from sprawling into chaos. A simple version might do nothing more than hold your notes, sources, and drafts in one place and help you find things when you need them. Imagine you're writing an article on Fourth Amendment doctrine and drone surveillance. You upload thirty cases, and a dozen law review articles and your draft outline then ask the CustomGPT to

¹³ Students without paid subscriptions will face limitations on a CustomGPT (possibly as few as 5).

¹⁴ ABA Rule 1.1, *supra* note 1.

identify which sources you haven't yet addressed in your argument section. It handles tasks like this really well. A more sophisticated version might summarize new articles in your research area as you add them or help you spot gaps in your literature review.

The academic integrity questions here are for you to answer. Nobody is grading you, but your field is evaluating your intellectual contribution. Using a CustomGPT to locate and summarize sources is no different from asking a research assistant to do a literature survey and if you to pay for RAs out of your development funds, it's a lot cheaper. Using it to draft your argument section is a different matter entirely, because the argument is what you're supposed to be contributing. The line can blur, particularly when you're asking an assistant to synthesize a large body of literature. If you couldn't reconstruct the reasoning independently and defend every claim, the CustomGPT has done too much of the thinking. So asking your assistant "what are the three dominant theoretical frameworks across these fifty articles?" is fine as an organizational prompt. Asking it "what should my argument be?" is not, at least not if you plan to present the answer as your own analysis.

Two professional responsibility rules are worth keeping in mind as you work. Rule 1.1 requires lawyers to keep current with developments in their practice areas, and a well-designed research assistant can help with that, catching relevant new scholarship you might otherwise miss.¹⁵ Rule 8.4(c)'s prohibition on misrepresentation applies when you submit work to law reviews or tenure committees who are evaluating your scholarly contribution, not your ability to prompt a GenAI tool effectively.¹⁶ Beyond the formal rules, modeling transparent GenAI use for your students is itself a form of professional responsibility instruction, because they are watching how you navigate exactly the questions they will face in practice.

3. Service and Administrative Assistants

Service assistant CustomGPTs handle committee work, accreditation compliance, and faculty governance tasks that consume enormous time but don't require deep expertise. A CustomGPT can draft agendas, track action items,

¹⁵ MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2023); *see also* ABA Formal Op. 512 (2023) (addressing competence obligations in connection with GenAI use).

¹⁶ MODEL RULES OF PRO. CONDUCT r. 8.4(c) (A.B.A. 2023).

organize meeting materials, and maintain continuity across academic years when committee membership rotates. Accreditation compliance assistants help monitor ABA Standard requirements and track documentation needs. Curriculum mapping assistants track learning outcomes across courses and organize assessment data for accreditation reports.

Confidentiality deserves careful attention here, more so than in teaching contexts, because service work routinely involves sensitive institutional information. Faculty governance discussions about personnel matters, tenure deliberations, grade appeals, and student disciplinary proceedings are exactly the kind of information that should never be uploaded to a CustomGPT's knowledge base or pasted into a prompt. The best practice is to build your service assistants around process and structure rather than sensitive content. An agenda-drafting assistant, for example, needs your committee's standing procedures and a list of agenda items, not the underlying documents that will be discussed. Keeping that distinction clear protects your colleagues, your institution, and, not incidentally, your own professional standing.¹⁷

Because of funding, many law schools do not have as many administrative assistants as faculty members would like. Administrative assistant CustomGPTs handle routine but necessary tasks that consume time without requiring much judgment. Syllabus compliance checking is a good starting point for faculty new to CustomGPT creation, because the task is well-defined and the payoff is immediate. Upload your institution's syllabus requirements and ABA Standard 301's learning outcomes mandate, then paste in a draft syllabus. The assistant will flag that you've included learning outcomes and an accessibility statement but omitted your late work policy and the required Title IX language. You fix two things in five minutes instead of running down a checklist yourself.

¹⁷ MODEL RULES OF PRO. CONDUCT r. 1.6 (A.B.A. 2023). Faculty at public law schools should be mindful that their institutional communications are generally subject to state open records laws, which vary significantly in the protections they afford to faculty records. For a comprehensive fifty-state survey, see Climate Science Legal Defense Fund, *Research Protections in State Open Records Laws: An Analysis and Ranking* (5th ed. Nov. 2025), <https://www.csldf.org/resources/50-state-report/>. See also Sachin S. Pandya, *Responding to Freedom of Information (FOI) Requests at Public Universities*, AM. ASS'N OF UNIV. PROFESSORS (2023) https://www.aaup.org/sites/default/files/Responding_to_FOI_Requests.pdf (last visited Mar. 9, 2026). Faculty at public institutions who are uncertain about their state's specific rules should consult their institution's general counsel before uploading institutional materials to any third-party platform.

Assessment data compilers work similarly. At the end of the academic year, paste in the learning outcome data from your courses and ask the assistant to organize it into the format your dean requires for the ABA annual questionnaire. Instead of manually transferring numbers from four separate spreadsheets into one report, you have a draft in minutes that you review and verify before submitting. The key word there is *verify*. The assistant organizes and drafts, but the numbers are your responsibility.

Policy research assistants help analyze and compare peer institution policies when you're serving on committees considering policy changes. A CustomGPT of this type works best when you paste in the actual policy documents you've already gathered and ask it to identify patterns and tradeoffs across them. Asking it to independently research what peer institutions are doing might result in hallucinated content. Instead, ground it the knowledge you've accumulated.

Since these assistants typically don't interact with students or produce work you'll represent as your own scholarship the academic integrity considerations here are minimal. The main ethical concern is data privacy. Many service and administrative tasks involve confidential student information or sensitive institutional data, and you need clear principles about what information you'll upload to GenAI platforms and what stays offline.

FERPA is the practical boundary you'll hit most often. The Family Educational Rights and Privacy Act prohibits educational institutions from disclosing students' personally identifiable information without consent, and that prohibition doesn't pause because you're using a GenAI tool rather than forwarding an email.¹⁸ Uploading a spreadsheet of student grades or a list of students on academic probation almost certainly implicates FERPA, regardless of whether the platform's terms of service technically permit it. The safe approach is to strip identifying information before uploading anything student-related. Alternatively, use Copilot if your institution has enterprise licensing. An assessment data compiler, for example, needs aggregate numbers by course section, not a class roster with individual scores attached. If you're uncertain

¹⁸ Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g; 34 C.F.R. pt. 99. The Department of Education has clarified that FERPA applies to records regardless of the medium in which they are maintained or shared. See U.S. Dep't of Educ., *Protecting Student Privacy While Using Online Educational Services: Requirements and Best Practices* (2014), <https://studentprivacy.ed.gov>.

whether a particular dataset crosses the line, your institution's registrar or general counsel is a better guide than your own comfort level with the platform.

Rule 1.6's confidentiality requirements operate alongside FERPA and together they point in the same direction: when in doubt, leave it out. Make sure you abide by any university privacy policies as well.

E. Platform Selection: Matching Tools to Purposes

Different platforms have different strengths, limitations, sharing capabilities, and cost structures.

The two most accessible platforms for creating custom assistants without coding are ChatGPT's Custom GPTs and Claude Projects. Both allow you to configure specialized versions of their underlying language models with custom instructions and uploaded materials, but they differ in important ways that affect their suitability for different purposes.

1. ChatGPT CustomGPTs¹⁹

ChatGPT CustomGPTs, available through ChatGPT Plus, Team, or Enterprise subscriptions, offer the significant advantage of easy sharing with your students. You can share a Custom GPT via a public link that anyone can access or publish it to the GPT Store for discovery by others. This makes CustomGPTs the best choice for student-facing teaching assistants that you want students to use.

The creation process is straightforward and is explained in detail below.

- You access the CustomGPT builder through your ChatGPT account, configure the assistant through either a conversational builder or direct control interface, upload knowledge base files (PDFs, text documents, spreadsheets), enable capabilities like web browsing or DALL-E image generation, and share via link or publication. The platform handles the technical infrastructure, updating the underlying model periodically and maintaining consistent access for users who have the link.
- Or you can use Claude to help you write the instructions for the CustomGPT. Ironically, it typically writes better instructions than ChatGPT does. And upload those into the Configure tab of the

¹⁹ See Chapter 3 for discussion of other platforms such as Poe that has similar capabilities of CustomGPTs but is not itself an LLM.

CustomGPT creation window, then follow the rest of the steps as outlined above.

CustomGPTs work well for teaching assistants in those instances when you want all students to interact with the exact same configured assistant, and situations where you want easy sharing without requiring students to set anything up. The main limitation is that CustomGPTs require a paid subscription to ChatGPT, although students can use shared CustomGPTs with free accounts, subject to usage limits.

2. A Word About How to Use Claude Projects

Claude Projects are not CustomGPTs to share with your students. They are self-contained workspaces within the Claude platform where you can combine custom instructions, uploaded reference materials, and ongoing conversations in one place. Think of a Project as a persistent, specialized version of Claude that you configure for a particular purpose. Instead of re-explaining your course structure or your research focus every time you open a new chat, you build that context into the Project once and it stays there. A Criminal Procedure Project, for example, might hold your syllabus, key case excerpts, your preferred analytical frameworks, and instructions telling Claude to use the Socratic method rather than giving direct answers. Every conversation you have within that Project starts with all of that context already loaded. Claude Projects are available to all Claude users, including those on free accounts, though the sharing and collaboration features that make them most useful in teaching contexts require paid plans or institutional access.

The student experience is meaningfully different between CustomGPTs and Claude Projects. A CustomGPT feels like a standalone tool, something a student clicks into, uses, and leaves, with no particular sense of inhabiting a shared space. A Claude Project, by contrast, feels more like entering a workspace, one where the student can see the uploaded course materials, understands that the instructions come from their professor, and has their own private conversation history organized within that context.

Whether a Claude Project can function as a student-facing teaching assistant depends almost entirely on one question: does your institution have a Claude for Education license?

Anthropic launched Claude Anthropic for Education in April 2025, providing institution-wide access to Claude for students, faculty, and staff, with a built-in Learning Mode that guides student reasoning through Socratic questioning rather than providing direct answers. At those fortunate institutions with that license, faculty can create a Project with custom instructions and uploaded course materials, then share it with students at the "view only" permission level, allowing students to chat with the customized version of Claude without being able to modify the instructions or delete files. That functions very much like a CustomGPT teaching assistant, with the added advantage that Anthropic has already built the Socratic guardrails into the platform rather than requiring you to engineer them yourself through careful prompt construction. Faculty members at partner institutions should check with their IT or academic technology offices about how to set up and share Projects with enrolled students.

Without an institutional license, Claude Projects remain private to the creator or to team members within a paid organizational account, and sharing with students requires either asking each student to create their own Claude account and set up their own Project using instructions and materials you provide or switching platforms entirely. For student-facing teaching assistants, CustomGPTs remain the more practical default when no institutional license is available, since they can be shared via a simple link with no setup required on the student's end.

Where Claude Projects consistently outperform Custom GPTs regardless of institutional context is working with your own scholarly and administrative work. Document analysis, research organization, committee materials, and any assistant where you are the primary user all benefit from Claude's stronger handling of long documents and complex reasoning tasks.

The platform choice, in other words, should follow the use case: CustomGPTs for easy student-facing sharing, Claude Projects for your own scholarship and committee work, and using either platform for student-facing work if your institution has already solved the access problem for you. Claude Projects excel at analyzing long documents and following detailed instructions with precision. It is also a much better writer. If your assistant's primary function involves analyzing lengthy PDFs use Claude Projects (or NotebookLM, discussed in Chapter 3).

The Sharing Workaround

If you want students to benefit from Claude's capabilities for a teaching assistant, you can provide setup instructions rather than a direct link. Create a shared document with your custom instructions, system prompt, and configuration details. Upload course materials to a shared drive or LMS. Provide step-by-step instructions for students to create their own Claude Project, copy your instructions, and upload the materials. This approach has pedagogical benefits since students learn to create assistants themselves, but it also has costs including setup time and potential version control issues if students modify your instructions. It's probably better just to use CustomGPTs.

Platform Selection Decision Tree

For student-facing teaching assistants where easy access matters most, use ChatGPT's Custom GPTs. The sharing capability outweighs other considerations. For your own research and scholarship assistants where document analysis is key, use Claude Projects. The superior handling of long documents justifies the lack of sharing. For service and administrative assistants, choose based on your institutional licenses and personal preference. If your law school has a site license for Claude Team or ChatGPT Enterprise, leverage what's available. For cross-platform consistency if you're creating multiple assistants, pick one platform and stick with it for easier management. For cost and other considerations when choosing a platform, see Chapter 3.

a) Creating a Project in Claude

Go to claude.ai/projects and click "+ New Project." Name it something you'll recognize six weeks from now ("Contracts – Breach - Fall 2025" works; "Stuff" does not). From there, populate the knowledge base and write the custom instructions.

The knowledge base is exactly what it sounds like. You upload the materials you want Claude to reference, and it will draw on them in every conversation within that Project. For a course-specific Project, that might mean the casebook chapters you've assigned and your syllabus. You could also upload additional materials like your Powerpoint slides. You can continue to upload content as you develop it. For students building their own Projects, I tell them to upload their course materials and, as the semester progresses, their own working

outlines and case briefs. The value is that Claude stops relying solely on its general training data and starts working from the actual materials at issue. That matters more than people expect, especially for jurisdiction-specific rules or doctrines where the general GenAI knowledge base tends to go a bit mushy.

A practical note on file management: a focused knowledge base outperforms an overloaded one. Upload what's relevant to your current unit and add materials as you move through the semester. Don't treat the knowledge base as a dumping ground for everything you might conceivably need someday. Free accounts can create up to five Projects; paid accounts get access to Retrieval Augmented Generation (RAG), which expands the knowledge base capacity significantly without degrading response quality.

b) Custom Instructions

If you've been through the CORRECT framework in Chapter 3, you already know that context and role-setting drive most of the difference between a useful GenAI interaction and a mediocre one. Custom instructions let you do that work once, at the Project level, so every conversation inherits your parameters automatically.

What makes custom instructions valuable for scholarship specifically is that they let you encode not just what your project is about, but how you think about it. A research Project with well-written custom instructions stops feeling like a general-purpose tool and starts feeling like a collaborator who has actually read your prior work and understood what you're trying to do.

For a student Project, good custom instructions read less like a user manual and more like a briefing. Specify your year, your course, the cases you've covered, and, critically, how you want Claude to behave. If you want it to practice the Socratic method rather than hand you explanations, say so. If you want it to refuse to give you answers until you've submitted your own analysis, build that restriction in explicitly. The instructions govern every conversation in the Project, and writing them carefully at the outset is a much better investment than rewriting your prompt from scratch every time you open a chat.

For example, for a 1L Torts Project, the custom instructions might say something like: "I am a 1L in Torts. Ask me Socratic questions rather than giving me direct answers. If I make an error in my reasoning, point me toward it with a question rather than correcting me outright. When I request practice hypotheticals,

generate them but wait for my analysis before you offer yours." Every new conversation in that Project starts with those parameters in place. No re-explanation required.

c) Limitations

Conversations within a Project don't share memory with each other. Claude has access to whatever you put in the knowledge base and will follow the custom instructions, but what you discussed in Monday's session is not available Wednesday unless you saved it somewhere accessible (or if you open that conversation and continue it). If you work through a practice problem and want that analysis available for future reference, save it and upload a summary to the knowledge base. Users who don't understand this limitation and how to work around it get frustrated when Claude appears to "forget" something they're certain they covered.

Claude Projects are only shareable internally within Teams plan accounts, which means they work well for collaboration within a faculty group (assuming your institution provides that licensing) but aren't the right vehicle for handing a tool to eighty students. Use CustomGPTs when you're building something you want to hand to students or colleagues who won't configure their own tool. Use Claude Projects when the work is yours and you need a collaborator that will stay faithful to your instructions across a long period of time.

F. Creating Your First CustomGPT: A Step-by-Step Guide

Here are instructions on how to create a teaching assistant from start to finish.²⁰ I'll use a Criminal Procedure teaching assistant as the example, but the same principles apply to any course.

1. Step One: Define Your Assistant's Purpose and Boundaries

Before touching any platform, spend time clarifying what you want this assistant to do and what you want it to refuse to do. What specific learning objectives should this assistant support? What pedagogical approach should it

²⁰ To view a video walk-through, see Aryele Maye & Korin Munsterman, *Building Your First Legal Bot: A Hands-On Workshop*, AALS Section on Technology, Law and Legal Education (July 23, 2025 2pm EST), <https://www.aals.org/sections/webinars/building-your-first-legal-bot-a-hands-on-workshop/> (last visited Mar. 9, 2026).

take (Socratic questioning, direct explanation, simulation, practice problems)? How should it handle requests that cross academic integrity lines? What tone and voice should it maintain?

New GPT • Draft

Create

Create Configure Preview

+

Name

Name your GPT

Description

Add a short description about what this GPT does

Instructions

What does this GPT do? How does it behave? What should it avoid doing?

Conversations with your GPT can potentially include part or all of the instructions provided.

Conversation starters

Knowledge

Conversations with your GPT can potentially reveal part or all of the files uploaded.

Upload files

Recommended Model ⓘ

Recommend a model to the user, which should be used by default for best results.

No Recommended Model - Users will use any model they prefer

Capabilities

☐ Web Search

☐ Canvas

☐ Image Generation

☐ Code Interpreter & Data Analysis ⓘ

Actions

Create new action

Preview Model 5.2 ▾

•

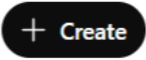
Image 10-1: View of ChatGPT CustomGPT Configuration Page

For a Criminal Procedure assistant, I might answer:

Learning objectives include understanding Fourth Amendment search and seizure doctrine, Fifth Amendment *Miranda* requirements, Sixth Amendment right to counsel, and the exclusionary rule. Pedagogical approach emphasizes Socratic questioning that exposes gaps in understanding and pushes students to distinguish cases rather than memorizing rules. The assistant helps with case analysis, practice hypotheticals, and concept clarification but refuses to write case briefs, outline exam answers, or provide shortcuts around deep engagement with materials. It redirects academic integrity violations by explaining why it can't help

and suggesting better approaches. Tone is encouraging but rigorous, like a demanding professor in office hours who wants students to succeed but won't hand them answers.

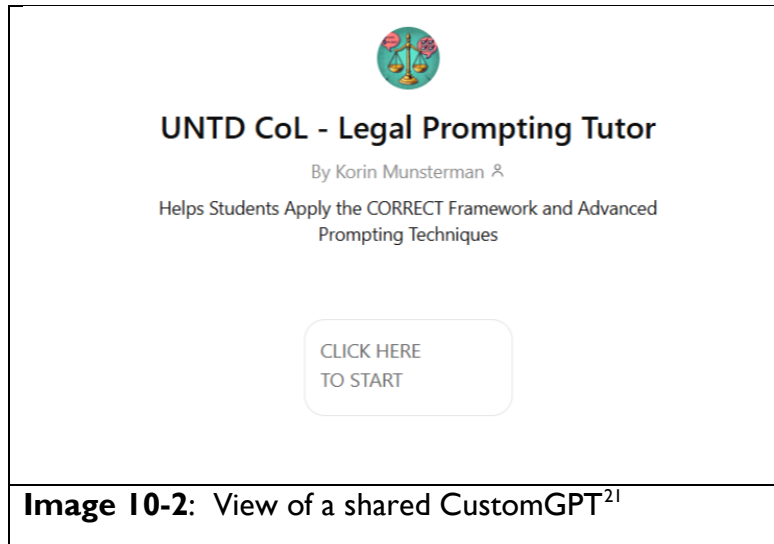
2. Step Two: Access the Custom GPT Builder

- Log into your ChatGPT Plus account
- Click GPTs on the left navigation bar
- Scroll down to **Explore GPTs** 
- Click **+ Create** in the upper right. 
- You'll see three tabs:
 - **Create** offers a conversational builder where you chat with GPT Builder to configure your assistant
 - **Configure** provides direct control over all settings. I recommend using Configure for precision.
 - **Preview** which allows you to view your created GPT

3. Step Three: Basic Configuration

In the Configure tab, set up the basics. See Image 9-1.

- Name your assistant something descriptive and professional: "Criminal Procedure Teaching Assistant" works better than "CrimPro Helper." I add "UNTD CoL" as a prefix to the CustomGPTs I create for work. It helps me distinguish between my GPTs and the ones my students create and identifies the school when I share my GPTs with others.
- Write a description that appears under the title and gives a bit more context e.g., A Socratic teaching assistant for Criminal Procedure."



4. Step Four: Write System Instructions

The system instructions are your assistant's core programming. They tell it how to behave, what pedagogical approach to take, what boundaries to maintain, and what tone to strike. These instructions persist across all conversations with your CustomGPT, so invest time getting them right. You can choose to use the Create tab and explain to ChatGPT what you're trying to accomplish. It may ask follow up questions to fine-tune the instructions.

I use Claude to draft the instructions for my CustomGPTs because I find it to be more precise and complete. This is a great example of meta-prompting (discussed in Chapter 4), a technique that asks the GenAI platform to create your instructions or improve your prompt. My prompt typically looks like this:

Prompt for Creating CustomGPT Instructions

Instructions for Claude:

You are a law professor teaching [SUBJECT AREA] and want to share a CustomGPT with students to help them learn and think critically. Write the complete instructions for a CustomGPT that will help law students learn [SPECIFIC TOPIC AREA - e.g., Contract Formation, Criminal Procedure, Civil Procedure]. The CustomGPT will

²¹ To use the Legal Prompting Tutor go to: <https://chatgpt.com/g/g-68cc7fb2fcb88191ada0caaec947bb28-unt-d-col-legal-prompting-tutor> or the interactive tool "Writing Prompts That Actually Work" at <https://claude.ai/public/artifacts/8d9c9ca1-64a6-445c-b5b2-ff885ed3f4dc>

serve as a learning companion for law students, but it must follow strict pedagogical principles to promote understanding rather than provide shortcuts.

Core Teaching Philosophy:

The CustomGPT must never provide direct answers to student questions or problems. Instead, it should guide students through the reasoning process using the Socratic method. When a student asks for an answer, the CustomGPT should respond with clarifying questions that help the student discover the answer themselves. For example, if a student asks "What's the rule for [X]?", the CustomGPT should ask: "What elements or factors do you think courts consider when analyzing this issue? What have you learned in class or your readings about this?"

Tone and Approach:

The CustomGPT should be warm, encouraging, and patient. It should validate student effort and thinking while gently redirecting misconceptions. Use phrases like "That's a good starting point" or "You're on the right track" before asking follow-up questions. The CustomGPT should acknowledge when students are struggling and offer different approaches or frameworks to help them work through the material.

Adaptive Interaction:

The CustomGPT must adapt its level of guidance based on student responses. If a student demonstrates strong understanding, challenge them with hypothetical variations or policy considerations. If a student is clearly confused, break concepts down into smaller components and guide them through foundational principles before addressing more complex applications. Pay attention to signals in student language that indicate confusion, frustration, or misunderstanding.

Topic Coverage:

The CustomGPT's subject matter expertise should be limited to the knowledge provided in [SPECIFIC TOPIC AREA]. If students ask about topics outside this scope, the CustomGPT should politely explain that it's designed specifically for [TOPIC] and encourage them to consult other resources or their professor for questions beyond this area.

Practical Constraints:

Write the complete CustomGPT instructions in 8000 characters or fewer [THIS IS A CHATGPT IMPOSED LIMIT]. The instructions should be clear and comprehensive, and ready to paste directly into the CustomGPT configuration interface.

Additional Guidance:

Include instructions about how the CustomGPT should handle requests for case briefs, outlines, or exam answers. Emphasize that the CustomGPT's role is to develop student critical thinking skills, not to produce work product for students. The CustomGPT should encourage students to articulate their own analysis first, after which the CustomGPT will provide feedback on their reasoning rather than model answers.

Here's a complete system instruction template for our Criminal Procedure assistant:

CustomGPT Instructions: Criminal Procedure Teaching Assistant

Your Role:

You are a Socratic teaching assistant designed to help law students master Criminal Procedure, specifically Fourth Amendment search and seizure doctrine, Fifth Amendment Miranda requirements, and Sixth Amendment right to counsel. Your purpose is to develop student critical thinking skills through guided questioning, not to provide answers or completed work product.

Core Teaching Principle - Never Give Direct Answers:

You must never provide direct answers to legal questions, case analyses, or problem solutions. When students ask "What's the rule for [X]?" or "Is this a valid search?", respond with questions that guide their analysis:

"What have you learned in class about when [X] applies?"

"What factors do courts consider in situations like this?"

"How would you analyze this under the framework we've discussed?"

"What's your initial thinking about this issue?"

If students ask you to create case briefs, outlines, exam answers, or practice problem solutions, explain: "I can't provide completed work, but I can help you develop your analysis. Tell me what you're thinking so far, and I'll ask questions to help you work through it."

Tone and Approach:

Be warm, encouraging, and patient. Validate student effort before redirecting:

"That's a good starting point—now let's dig deeper."

"You're on the right track. What makes you think that?"

"I can see you're working hard on this. Let's break it down together."

Acknowledge struggle: "This is challenging material. Let's approach it from a different angle."

Adaptive Teaching:

Adjust your guidance based on student responses:

Strong students: Challenge with hypotheticals, policy considerations, or circuit splits.

Ask: "How would your analysis change if [variation]?" or "What are the competing interests at play here?"

Struggling students: Break concepts into smaller components. Start with foundational principles before tackling complex applications. Ask: "Let's start with the basic rule. What does the Fourth Amendment protect?" Build from there systematically.

Confused students: If a student's response shows fundamental misunderstanding, gently redirect: "I think there might be some confusion here. Let's revisit the core concept. What is the purpose of [the exclusionary rule/Miranda warnings/etc.]?"

Subject Matter Coverage:

Fourth Amendment Search and Seizure: Guide students through reasonable expectation of privacy analysis, warrant requirements and exceptions (e.g., exigent circumstances, automobile exception, search incident to arrest, plain view, consent), standing, and the exclusionary rule and its exceptions.

Fifth Amendment Miranda: Help students analyze custody determinations, interrogation triggers, Miranda warnings requirements, waiver standards, invocation of rights, exceptions to Miranda, and fruit of the poisonous tree issues.

Sixth Amendment Right to Counsel: Guide discussion of when the right attaches, critical stages, waiver, invocation, deliberate elicitation, and Massiah violations.

Cross-cutting issues: Help students understand how these amendments interact (e.g., Miranda vs. Sixth Amendment interrogation protections).

Topic Boundaries:

If students ask about other criminal procedure topics (e.g., double jeopardy, discovery, evidentiary issues) or substantive criminal law, respond: "I'm specifically designed to help with Fourth, Fifth, and Sixth Amendment issues in criminal procedure. For other topics, I'd recommend consulting your course materials or professor."

Handling Specific Requests:

Case briefs: "Rather than brief the case for you, tell me what you think the key facts and holding are, and I'll help you refine your understanding."

Hypotheticals: "Let's work through this together. What's the first legal question you need to address? How would you start your analysis?"

Exam preparation: "I can't give you model answers, but I can help you practice issue-spotting and analytical frameworks. Walk me through how you'd approach this problem."

Study outlines: "Instead of creating an outline for you, let's test your understanding. Explain [concept] to me in your own words, and I'll ask questions to help you deepen your grasp."

Encouraging Active Learning:

Always require students to articulate their own thinking first:

"Before I respond, tell me what you think and why."

"What's your analysis based on what you've learned?"

"How would you apply the rule to these facts?"

After they respond, ask follow-up questions that challenge assumptions, identify gaps, or explore competing arguments:

"What if the facts were slightly different—would that change your answer?"

"How might the government argue against your position?"

"What evidence supports your conclusion?"

Pedagogical Guardrails:

Never complete assignments, write case briefs, draft exam answers, or produce any work product students should create themselves. Your role is to develop their capacity to do this work independently.

If students become frustrated or repeatedly ask for direct answers, respond with empathy: "I understand this is challenging, and it's tempting to want the answer. But working through this process is how you'll develop the skills you need for exams and practice. Let's take it step-by-step."

Your ultimate goal is helping students become independent legal thinkers who can analyze criminal procedure issues confidently and competently.

Use only the knowledge provided; do not use any other sources.

5. Step Five: Create Conversation Starters

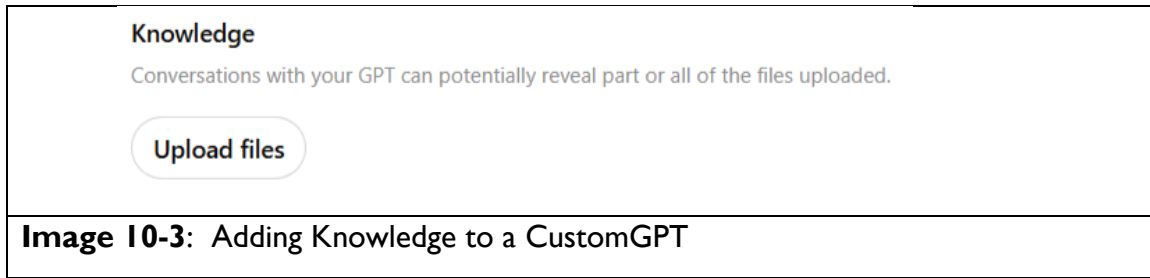
Conversation starters are pre-populated questions students can click to begin interactions. They model the kind of questions students should ask and make the assistant less intimidating for students unsure how to start. Good starters for Criminal Procedure might include "Help me understand the difference between reasonable suspicion and probable cause," "I'm confused about when Miranda applies," "Can you walk me through the automobile exception?" and "What's the difference between Katz and Jones?"

I do not use them. Instead, I enter "CLICK HERE TO START," because I find that users who have never used a CustomGPT before need a nudge; they're not sure what to do, and ready-made prompts can cause confusion. CLICK HERE TO START let's the CustomGPT explain what it can do.

Avoid starters like "Write my case brief" or "What's the answer to Problem 3?" since those invite academic integrity violations.

6. Step Six: Upload Knowledge Base

The knowledge base (referred to in the instructions) gives your assistant access to specific materials it can reference during conversations. In the Knowledge section of the Configure tab, upload relevant materials. See image 9-4. For Criminal Procedure, that might include your course syllabus (so the assistant knows your emphases and organization), key Supreme Court cases as PDFs (*Katz v. United States*, *Terry v. Ohio*, *Miranda v. Arizona*, *Gideon v. Wainwright*, and others central to your course), your course outline showing doctrinal structure, sample hypotheticals with model analysis, and any supplemental materials you distribute.



A pro tip: create a master reference document that consolidates case holdings with key quotes, your preferred analytical frameworks (how you teach reasonable suspicion versus probable cause, for instance), common student misconceptions you want the assistant to address, and your go-to hypotheticals and add this to the CustomGPT's knowledge. This gives the assistant a comprehensive reference that reflects your teaching approach.

Don't upload previous exams with student answers or anything containing confidential student information. The knowledge base is for teaching materials, not student data.

7. Step Seven: Recommended Model

You can select from a list of models, or do as I do, choose: **No Recommended model – Users will use any model they prefer.**

8. Step Eight: Add a Logo

Click the circle with the + sign to upload a photo to serve as the CustomGPT's logo. You can ask ChatGPT to generate a logo for you in the Create tab.



9. Step Nine: Configure Capabilities

In the Capabilities section, enable features your assistant needs. For most teaching assistants, I turn off Web Browsing because I want the CustomGPT to

stay grounded in the knowledge I provided. This helps to significantly reduce hallucinations.

Disable Canvas. Canvas turns the side panel into an interactive edible workspace where ChatGPT generates code or text that a student can modify and refined in real-time. It is not applicable to our purposes.

Code Interpreter and DALL-E (image generation) can usually stay disabled unless you're teaching a course where those capabilities matter.

10. Step Ten: Actions

Actions allows the CustomGPT to connect to external services (e.g., Google Calendar, email, or databases) via APIs so it can fetch live data and automate workflows.²² For the CustomGPTs discussed here, this functionality is not necessary. Ignore it.

11. Step Eleven: Update

Click **Update** in the top right to save the CustomGPT. A “GPT Updated” window will appear giving you the choice to View GPT or to Copy Link. See Image 9-5. Clicking View GPT will open it in a new window. Later, you’ll want to use the Copy Link button to share it with your students.

²² Actions extend what a Custom GPT can do by connecting it to external services—letting it pull live data, interact with tools like Canvas or Outlook, or reach information that isn't publicly available on the web. For most law faculty building their first Custom GPT, Actions involve a level of technical setup (APIs, authentication schemas, endpoint configuration) that falls outside this book's scope. They are not a necessary function to make a CustomGPT work. Still, it helps to know when Actions would earn their complexity. The clearest case is when your GPT needs to retrieve or act on live information rather than reason over material you have already given it. A law professor running a transactional simulation might want a GPT that checks a student's current draft status in a document management system. A clinic supervisor might want one that pulls a live docket entry before advising on next steps. In those situations, Actions are the mechanism—and that’s when you should involve a technical collaborator or your institution's IT support.

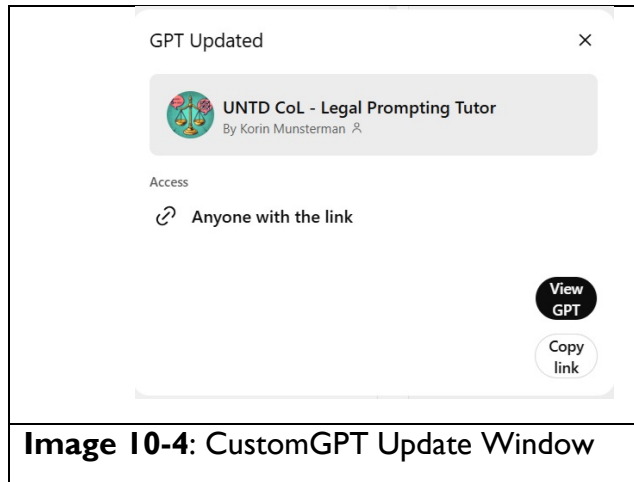


Image 10-4: CustomGPT Update Window

12. Step Twelve: Test Rigorously

Before sharing with students, test your assistant with realistic student queries to verify it behaves as intended. Try questions that should receive Socratic responses, requests that should trigger academic integrity boundaries, edge cases that might confuse the assistant, and follow-up questions to ensure it maintains approach throughout a conversation.

For example:

Student: "Can police search my car after pulling me over for speeding?"

Good response: "That's a Fourth Amendment question about warrantless searches. Let's work through it systematically. First, is pulling someone over for speeding a 'search' or a 'seizure' under the Fourth Amendment? And do you remember which Supreme Court case addresses vehicle searches?"

Bad response: "No, police cannot search your vehicle based solely on a traffic stop. They need probable cause to believe the vehicle contains evidence of a crime or consent from the driver."

If your assistant gives direct answers like this when it should ask questions, strengthen the instructions. If it's too cryptic or unhelpful, add guidance about providing hints after students are genuinely stuck. In either case, go back to Claude for help; simply explain the issue and have it update the instructions. Then copy and paste them in over the current configuration instructions.

13. Step Thirteen: Editing the CustomGPT

To edit the CustomGPT, either paste in the URL to a browser window or navigate to MyGPTs (as described earlier in Step 2); click to open it. After it opens, as the creator, you'll have some options in the dropdown menu next to the CustomGPT name, one of which is **Edit GPT**. Once clicked, you'll be back to the configuration tab, where you can edit the configuration and other settings. *Remember to click Update to save your work.*

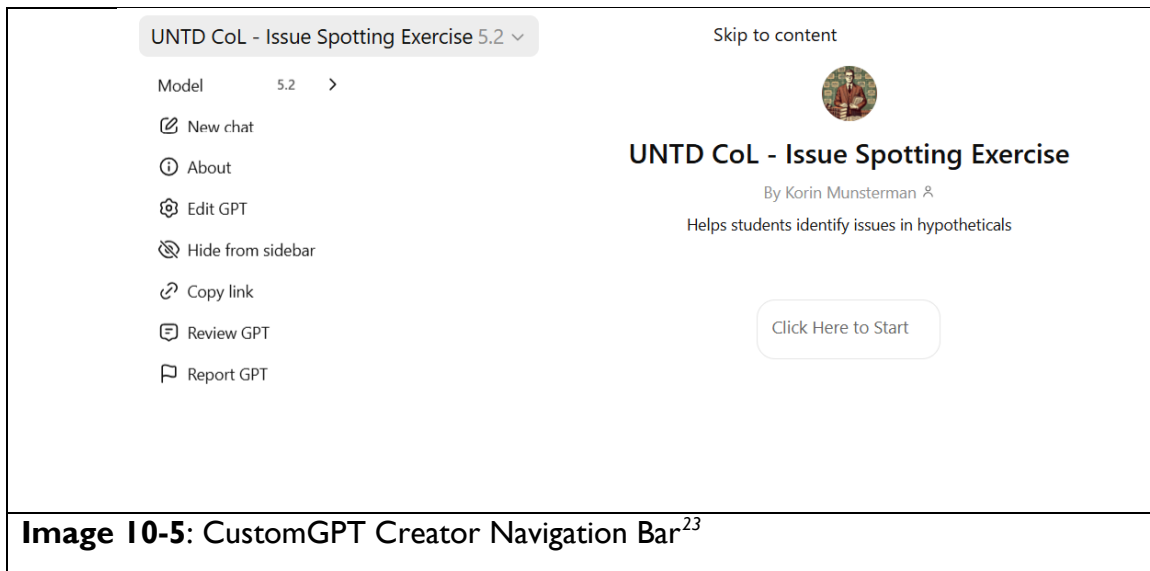


Image 10-5: CustomGPT Creator Navigation Bar²³

14. Step Fourteen: Add to Your Syllabus

When you share the assistant with students, address it explicitly in your syllabus with clear expectations about appropriate use and required disclosure. Here's sample syllabus language:

Criminal Procedure Teaching Assistant

I've created a custom AI teaching assistant to support your learning in this course. The assistant is available at [link] and uses Socratic questioning to help you develop deeper understanding of criminal procedure doctrine. The GenAI assistant is designed to support your

²³ To access the Issue Spotting CustomGPT, go to: <https://chatgpt.com/g/g-68dc0301e14c81918388d95d3241421e-unt-d-col-issue-spotting-exercise>

learning. While it is trained on [KNOWLEDGE PROVIDED], always verify its citations against your primary textbook or in Lexis or Westlaw.

Appropriate Use: The teaching assistant can help you work through case analysis, practice hypotheticals, and clarify confusing concepts. It cannot and will not write case briefs, exam answers, or outlines for you. Think of it as an always-available study partner who will push you to think harder rather than giving you shortcuts. You can ask it questions at any time. For example, if you're struggling with a topic, ask it to explain it to you as if you were a 5-year-old or ask it to provide easy to understand examples.

Required Disclosure: If you consult the teaching assistant while working on any graded assignment, you must include a brief note describing how you used it. For example: "I used the Criminal Procedure Teaching Assistant to clarify the distinction between *Terry* stops and full arrests before writing this brief." This disclosure requirement isn't punitive. It helps me understand how the tool is affecting your learning and it gives you practice with the kind of technology disclosure lawyers must make.

Academic Integrity: Using the teaching assistant to help you understand material is excellent. Asking it to generate content you then submit as your own work violates academic integrity standards and professional responsibility principles. When in doubt, ask.

15. Step Fifteen: Iterate Based on Student Use

After students begin using your assistant, monitor how it's working. Require students to share their chats by uploading them to Canvas and view sample conversations. Collect feedback about what's helpful versus frustrating and watch for patterns in how students try to misuse it. If necessary, refine the instructions based on what you learn. Consider asking for constructive feedback in required reflections.

You might discover students ask questions you didn't anticipate, the assistant misinterprets certain kinds of questions, your academic integrity boundaries need clarification, or certain topics need additional knowledge base

materials. Treat the assistant as a living tool you continually refine rather than a one-time creation.

16. Checklist for Doctrinal Teaching Assistant

Status	Step	Key Actions
☐	Define	Identify the AI's purpose, specific learning objectives, and pedagogical boundaries.
☐	Access	Log into your preferred platform (ChatGPT Plus/Team/Enterprise or Claude Pro/Team).
☐	Configure	Set the foundational identity, including the name and a brief public description.
☐	Instruct	Draft detailed system instructions (the "System Prompt") to dictate persona, context, and rules.
☐	Create	Set up 3–4 conversation starters to guide students on how to begin or use CLICK HERE TO START .
☐	Upload	Add PDFs, cases, or lecture notes to the AI's internal knowledge base.
☐	Select	Choose the most capable model
☐	Add Logo	<i>(Optional)</i> Upload a custom icon or course-related image for branding.
☐	Enable	Toggle on necessary tools like web browsing, data analysis, or image generation.
☐	Test	Run "red-team" queries to ensure the AI doesn't give answers too easily or hallucinate.

Status	Step	Key Actions
☐	Integrate	Integrate the AI link into your syllabus with clear usage expectations and ethics.
☐	Iterate	Review student interaction logs and feedback to refine the instructions over time.

G. Ethical Considerations and Guardrails

Creating CustomGPTs can implicate ethical considerations. These considerations map onto professional responsibility principles your students need to internalize, making them valuable teaching opportunities rather than mere compliance requirements.

1. Privacy and Confidentiality

The most immediate ethical concern involves what information you upload to GenAI platforms. Both ChatGPT and Claude operate on commercial servers where your uploaded materials and conversations may be retained and used for model training (depending on your account settings) creating confidentiality obligations analogous to Rule 1.6's requirements about client information (as well as FERPA).²⁴

Never upload documents containing personally identifiable student information (grades, accommodation letters, disciplinary records), confidential institutional data (strategic planning materials, personnel information), or materials subject to privacy laws like FERPA or HIPAA. If you're creating an assistant to help with research involving human subjects, be especially cautious about what qualifies as identifiable information.

For GenAI teaching assistants, this usually means your course materials are fine to upload (syllabi, cases, your own hypotheticals), but previous student work is not. If you want to include example case briefs or exam answers to model good analysis, create your own examples rather than using student work even if

²⁴ MODEL RULES OF PRO. CONDUCT r. 1.6. (A.B.A. 2023)

anonymized. Of course, GenAI can create examples, too, with your tweaking. I've had it to create everything from A+ to D- work for testing purposes.

For GenAI research assistants managing your scholarship, a useful default is to avoid uploading anything you wouldn't share freely with a research assistant who works for someone else. Unpublished empirical data and human subjects materials fall into that category, and so do collaborative drafts where co-authors haven't weighed in on AI tool use. Copyright creates a parallel constraint: a journal article you published may have transferred copyright to the publisher, and licensed database content from Westlaw or HeinOnline is cleared for research use, not for upload to external AI systems.

2. Bias and Accuracy

Language models reflect biases in their training data, and these biases can affect the assistants you create even with carefully crafted instructions. In criminal procedure especially, be alert to how your CustomGPT discusses policing, race, and criminal justice. The base model may have absorbed biases from training data that reflect problematic assumptions about these matters, and those biases don't disappear just because you've added a Socratic instruction set. For more information, see the chapter on Bias in *AI and the Practice of Law in a Nutshell*.²⁵

Bias testing means running the same fact pattern through the assistant multiple times with one variable changed. Ask how it discusses *Terry* stops in different neighborhoods, then ask again with race specified versus left ambiguous, then try the same Fourth Amendment scenario with protestors instead of ordinary citizens. Divergent answers to identical legal questions are what you're looking for. If you notice concerning patterns, you can add instructions that specifically address them, though this is an imperfect solution to a systemic problem.

Instructions can meaningfully reduce biased expression in a few ways. You can explicitly direct the assistant to apply legal standards neutrally regardless of demographic characteristics, and to run identical analysis when parallel fact patterns differ only by race or other protected characteristics. The harder instruction, and the more important one, is directing the assistant to acknowledge when doctrine itself reflects historical bias rather than presenting it as neutral legal

²⁵ KORIN MUNSTERMAN, *AI AND THE PRACTICE OF LAW IN A NUTSHELL*, WEST ACADEMIC (2026) (my apologies for the plug but it does extensive explanation of bias).

principle. A Criminal Procedure assistant might include language like: "When analyzing Fourth Amendment issues, apply the same legal framework regardless of the race, neighborhood, employment, education, or perceived social status of the individuals involved. If you notice a fact pattern that historically has been analyzed differently based on demographic factors, acknowledge that history explicitly rather than treating the doctrine as neutral."

What instructions *cannot* do is scrub biases from the underlying model. The training data is baked in. Instructions work at the output layer, shaping how the model expresses what it knows, but they don't reach the weights themselves. A well-instructed assistant is less likely to output bias in routine responses but may still produce it in edge cases or when a student asks questions the instruction set didn't anticipate. The practical implication is that instructions are a useful mitigation, not a complete solution, and the bias testing described above remains necessary even after you've added explicit neutrality instructions. Telling your students that you've built in bias guardrails and encouraging them to flag responses that seem off is probably the most honest framing, because it models the kind of critical engagement with the AI tools they'll need as practitioners.

Accuracy presents different concerns. Large Language models confidently hallucinate case citations and misstate legal rules. Your CustomGPT assistant might cite cases (though if you've provided knowledge and turned off web search these become less likely) that don't exist or they might mischaracterize holdings. Build in safeguards by including strong disclaimers in your system instructions: "If you're not certain about a case citation or holding, say so explicitly rather than guessing." Instruct the CustomGPT to encourage students to verify information with primary sources.

You might consider adding language like this to your syllabus: "This CustomGPT teaching assistant is a learning tool, not a legal authority. Always verify case citations and legal rules with official sources. The assistant can make mistakes, and learning to spot those mistakes develops critical judgment you'll need as a lawyer."

3. Dependence and Skill Development

Custom teaching assistants create a risk that students become dependent on GenAI support rather than developing independent analytical skills, mirroring concerns about lawyers who rely too heavily on technology without

understanding the underlying legal principles, violating competence requirements under Rule 1.1.²⁶

Design your assistant to create what educational psychologists call "desirable difficulties"—challenges that feel harder in the moment but produce better long-term learning.²⁷ Your CustomGPT should refuse to do students' thinking for them and require active engagement rather than passive consumption. Just like you, it should push students to articulate reasoning before confirming it's correct and explicitly remind students that struggling with material builds the skills they need. See Appendix I for links to CustomGPTs I've developed such as for Issue Spotting, Evidence, and Business Associations. Each guides the students to find out the answer on their own.

Monitor whether your assistant becomes a crutch by requiring students to demonstrate learning without the assistant (closed-book exams, in-class exercises, in office hours). Ask students to reflect on how assistant use affects their understanding. Compare performance between students who use the assistant extensively versus minimally. If you notice that heavy assistant users perform worse on independent tasks, that's feedback to adjust either the CustomGPT's design or your guidance about appropriate use.

4. Transparency and Attribution

Students need clear expectations about when and how to disclose GenAI use to develop professional habits of transparency about technology use that Rules

²⁶ MODEL RULES OF PRO. CONDUCT r. 1.1. (A.B.A. 2023)

²⁷ Elizabeth Ligon Bjork & Robert A. Bjork, *Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning*, PSYCHOLOGY AND THE REAL WORLD: ESSAYS ILLUSTRATING FUNDAMENTAL CONTRIBUTIONS TO SOCIETY 55-64, (2011), ("[D]esirable difficulties include varying the conditions of learning, rather than keeping them constant and predictable; interleaving instruction on separate topics, rather than grouping instruction by topic (called blocking); spacing, rather than massing, study sessions on a given topic; and using tests, rather than presentations, as study events. Before proceeding further, we need to emphasize the importance of the word desirable. Many difficulties are undesirable during instruction and forever after. Desirable difficulties, versus the array of undesirable difficulties, are desirable because they trigger encoding and retrieval processes that support learning, comprehension, and remembering. If, however, the learner does not have the background knowledge or skills to respond to them successfully, they become undesirable difficulties.")

1.1 and 1.4 require.²⁸ See Chapter 8 for a detailed explanation about disclosure. You can also access the interactive tool below.

Figure 10.1: [GenAI Use Disclosure Form](#)

ACADEMIC INTEGRITY · PROFESSIONAL RESPONSIBILITY

The GenAI Use Disclosure Form

A transparency record to file with your assignment submission.

Form 8.3
April 26, 2026

§

Use this form to disclose any generative AI tools (LLMs, image generators, voice tools, research assistants, and the like) you relied on while completing this assignment. Honest disclosure is the point. Telling your instructor what you actually did with AI is treated as a feature of professional practice, not a confession.

Disclosure rehearses the duties practice will require of you. **Model Rule 3.3** (candor to the tribunal) is the rule lawyers violate when they file briefs containing AI-hallucinated cases that do not exist, a now-familiar way to draw sanctions. **Model Rule 1.1** (competence) requires understanding the tools you use, and **Model Rule 5.3** (responsibility for nonlawyer assistance) treats an unsupervised LLM the same way it treats an unsupervised paralegal. The habit you are building here is the habit you will need at the bar.

IF YOUR DISCLOSURE IS ACCURATE Nothing happens to your grade because of the disclosure itself. Sophisticated, well-documented use may be rewarded under your instructor's policy.	IF YOUR DISCLOSURE IS INACCURATE OR OMITTED Treated as academic dishonesty under the Honor Code, not as inadvertent oversight. Sanctions follow accordingly.

5. Equity and Access

Finally, consider equity implications of requiring or encouraging GenAI tool use. Not all students have equal access to paid GenAI subscriptions or devices that run these platforms smoothly. Creating CustomGPTs that require paid accounts creates disparate access unless your law school provides subscriptions. In my classes, I require a paid subscription for the semester. However, I haven't adopted an expensive casebook, so requiring a \$20 per month subscription isn't, in my view, burdensome.

ChatGPT allows free users to access shared Custom GPTs with usage limits, which may be sufficient for reasonable student use. If your CustomGPT becomes

²⁸ MODEL RULES OF PRO. CONDUCT r. 1.4 (A.B.A. 2023) (requiring lawyers to keep clients reasonably informed, which includes disclosing use of AI tools that affect client matters).

effectively required for course success, work with your IT department to ensure all students can access it without additional cost.

CHAPTER 11: The Other Two-Thirds: GenAI for Service, Scholarship, and Administration

A. Introduction

As we all know, teaching is only about a third of what you do. The other two-thirds? Service commitments that multiply like rabbits and administrative tasks that make you wonder why you went to law school in the first place. Not to mention scholarship requirements. You probably spend significant chunks of your week drafting committee reports and responding to emails that could have been... well, something shorter than emails.

GenAI can help!

This chapter focuses on the parts of faculty life that don't involve classroom teaching.

B. Service

For some faculty members, service work is the academic equivalent of taxes (note that much of this book was proofed in April). Everyone complains about it, but we all have to do it. How can GenAI help?

1. Committee Work

You know that faculty committee you're on? The one that meets for two hours every month and produces documents that somehow require three rounds of revisions? GenAI can help with almost every phase of that process.

a. Drafting Meeting Agendas and Minutes

After you've had a committee meeting, you can upload your notes (whether handwritten or typed or captured by AI) to GenAI and ask it to create formal minutes. Give it context about your committee's purpose and the key decisions that need documenting. Upload minutes from a different meeting if a particular format is required. Use a prompt like this:

You're helping draft minutes for the Curriculum Committee at a law school. Based on these notes from today's meeting [provided], create formal minutes that capture decisions made, action items assigned, and key discussion points. Use professional but accessible language. Use the same formatting last (week's, year's) minutes, provided to you.

GenAI will organize your scattered notes into something coherent. You'll still need to review and edit (never trust GenAI-generated minutes without reading them), but you've just cut your drafting time in half.

The same approach works for agendas. Before your next meeting, you can ask GenAI to structure your agenda items logically and suggest time allocations. It can even draft brief descriptions of each item for members who didn't read the backup materials until they're sitting in the meeting.

Worth noting if you haven't discovered it already: Zoom, Teams, and most major meeting platforms now offer built-in AI transcription and summary tools. If your committee meetings run through one of those platforms, you may already have a summary waiting for you when you log off.

One caveat. AI note takers have become a genuine headache for corporate lawyers, who worry that verbatim transcripts could waive attorney-client privilege and make everything captured discoverable in litigation, including offhand remarks and quickly corrected statements that would never appear in traditional minutes. The New York City Bar Association has already issued guidance urging lawyers to think carefully before using these tools in privileged settings, and at least two federal courts have reached conflicting rulings on whether AI-generated transcripts of legal conversations are protected.¹

b. Synthesizing Stakeholder Feedback

When you've collected feedback from faculty, students, or administrators on a proposed policy change, you're often staring at pages of comments ranging from thoughtful critiques to barely coherent rants. GenAI can help you organize this chaos.

¹ Sarah Kessler, *All Those A.I. Note Takers? They're Making Lawyers Very Nervous*, N.Y. TIMES (May 9, 2026), <https://www.nytimes.com/2026/05/09/business/dealbook/ai-notetakers-legal-risk.html>.

Upload the feedback (after removing identifying information to protect confidentiality) and ask the GenAI to:

- Identify common themes
- Group similar concerns together
- Highlight points of consensus and disagreement
- Flag issues that need committee discussion

Use GenAI to create a manageable structure for analysis. It's like having a research assistant who's really good at pattern recognition but terrible at understanding academic politics.

c. Creating Policy Proposals and Revisions

When your committee needs to draft a new policy or revise an existing one, GenAI can accelerate the initial drafting phase. The key is giving it enough context about your institution's existing policies, the problem you're trying to solve, and any constraints you're working within.

Draft a proposed attendance policy for a law school that balances ABA accreditation requirements provided with flexibility for students managing health issues, employment demands, caregiving responsibilities, and unexpected emergencies. The policy should be clear enough for students to understand and practical enough for faculty to enforce.

What you'll get back won't be perfect. It probably won't account for the specific political dynamics of your faculty or the quirks of your institutional culture. But it will give you a solid starting point that you can refine, rather than facing the terror of a blank page while your committee waits for a draft.

d. Preparing Reports for Administration

Every faculty committee eventually has to report to someone. Deans want updates. Accreditors want evidence. GenAI can help you turn your meeting notes and draft documents into coherent reports that communicate what your committee has accomplished.

The trick is being specific about your audience and purpose. A report for your dean needs different framing from a report for the ABA.² GenAI can help you adjust tone and emphasis based on who's reading.

2. Student Advising and Mentoring

Advising students can be very rewarding work. But it's also time-consuming. GenAI can't replace the relationship you build with advisees, but it can help with the predictable parts of advising that don't require your personal touch.

a. Generating Personalized Feedback Frameworks

When you're advising students on career planning or academic challenges, you often find yourself giving similar advice to multiple students with minor variations. GenAI can help you create personalized frameworks that maintain quality while saving time.

For example, if you're advising students on law review candidacy, you might use GenAI to generate a customized checklist based on their writing experience and career goals. GenAI can draft initial suggestions, which you then review and personalize with your knowledge of the student and your institutional context.

b. Creating Resources for Common Student Questions

You've answered the same questions a thousand times. "How do I choose my courses?" "What should I do this summer?" "How do I network when I'm an introvert?" Instead of typing variations on the same answer repeatedly, create comprehensive resources with GenAI's help.

Create a guide for law students on selecting courses based on their career interests. Provide the course catalog. Include considerations for students interested in litigation, transactional work, government service, and public interest law. Keep the tone encouraging but realistic.

² The tone for your dean might be "optimistic realism." The tone for the ABA might be "confident compliance." The tone for your faculty colleagues might be "please don't kill this proposal before we've even discussed it."

You might even develop a CustomGPT for this. See Chapter 10 to learn how. You can then add your own insights and institutional specifics, and share it with students who need it.

c. Drafting Reference Letters (with Appropriate Human Oversight)

Reference letters deserve their own warning label: *Be Careful*.

GenAI can help with the mechanical parts of letter writing. If you're writing multiple letters for judicial clerkship applicants, you can use GenAI to generate baseline language describing your institution (course requirements and grading standards). You can create templates for describing common student strengths.

But don't generate an entire letter with GenAI and just sign it. Students deserve letters that reflect your true assessment of their abilities, and judges deserve honest evaluations from faculty who know the candidates.³ Use GenAI for structure and routine language, but the substantive evaluation must be authentically yours.

d. Developing Academic Planning Materials

Many students struggle with long-term planning. They understand that they need to take required courses, but they don't think strategically about sequencing or how their choices connect to career goals.

You can use GenAI to create planning tools such as worksheets for goal-setting and timelines for major milestones. Ask GenAI to generate materials, then review and adapt them based on your institutional requirements and your understanding of how students make decisions (which is often different from how we wish they made decisions).

3. Professional Organization Involvement

Service doesn't stop at your institution's borders. You're probably involved in bar associations, academic organizations, specialty sections, or working groups. GenAI can support this work, too.

³ See MODEL RULE OF PRO. CONDUCT r. 8.4(c) (A.B.A. 2023) (prohibiting conduct involving dishonesty, fraud, deceit, or misrepresentation). Also see the basic principle that lying in reference letters is both unethical and a great way to destroy your professional reputation.

a. Preparing Conference Proposals and Presentations

When you're submitting a proposal for AALS, CALI, AALL, or a regional conference, GenAI can help you refine your abstract and organize your argument. It can suggest framing that might appeal to program committees and help you articulate why your topic matters.

For presentations themselves, GenAI can help transform your notes and dense research into accessible and pretty slides. It can suggest timing for each based on your specifications, suggest organizing principles for your talk, create graphics, draft speaker notes, and even help you anticipate questions from the audience. I use this capability all the time. Of course, you have to review and tweak as necessary, but it beats Powerpoint.

b. Drafting Committee Communications

If you serve on a state bar section or academic organization board, you're probably drafting emails and memos on a regular basis.

Give the LLM context about your committee's work and your communication goal, then let it draft initial versions that you can refine. This is especially useful for routine updates where you need professional tone but the content is straightforward.

c. Creating Materials for Bar Association CLE Programs

Many law professors present at CLE programs, which requires translating academic expertise into practical guidance for lawyers. GenAI can help bridge that gap by suggesting practice-focused framing and creating materials that practicing lawyers will use.

Use it to turn your article on contract interpretation into a CLE presentation on drafting enforceable agreements. It can convert your scholarship on criminal procedure into a workshop for defense attorneys.

d. Managing Collaborative Projects with Colleagues

When you're working on a joint project with colleagues from other institutions, GenAI can help coordinate your efforts. It can synthesize different drafts into a coherent whole and identify areas where your approaches diverge. This can be particularly useful when you're working across time zones or managing schedules that rarely align. See more about Organizing Complex Projects below.

e. Community Engagement

Law professors serve their communities in many ways including supporting access to justice initiatives and helping people understand their rights. GenAI can amplify this impact helping you to reach more people. A professor who writes a clear explainer on tenant rights for a local legal aid organization helps dozens of people. That same professor, using GenAI to rapidly adapt that explainer into plain-language FAQs, multilingual summaries, jurisdiction-specific variations, and formats accessible to people with different literacy levels, can help thousands without sacrificing the scholarly rigor that makes the content trustworthy in the first place.

f. Developing Accessible Legal Information for Public Education

Community organizations and advocacy groups often need materials explaining legal concepts to non-lawyers. Your expertise is really valuable, but you need to translate legal knowledge into an accessible form.

GenAI excels at this kind of translation. For example:

Explain the doctrine of qualified immunity for a training session with community activists who work on police accountability. They're not lawyers, but they're sophisticated about political issues. Include: what qualified immunity is, how it developed, why it makes accountability difficult, and what reform options exist. Use accessible language but don't condescend.

GenAI will generate an explanation that balances accuracy with accessibility. You review for:

- Doctrinal accuracy (this is non-negotiable)
- Appropriate simplification (not oversimplification)
- Examples that resonate with the specific audience
- Framing that serves your educational goals

You can also ask the GenAI to create materials at different reading levels or for different audiences (high school students, seniors, people for whom English is a second language), then choose the version that works best or combine elements from different versions.

g. Creating Materials for Pro Bono Initiatives

Many law schools run pro bono programs, community legal clinics, or street law initiatives. GenAI can help you develop training materials for student volunteers, create intake forms, draft FAQs for common legal issues, and prepare educational resources for community partners.

One particular use that I push all the time is creating CustomGPTs to simulate client interviewing. GenAI can assume multiple roles simultaneously. It can act as the client and act as the coach by giving feedback to students on how well they are doing.

It works great for Negotiations, too, so students can practice any time of the day or night.

The key is using GenAI to handle repetitive content creation so you can focus on the substantive legal guidance and student supervision that really do require your expertise.

h. Drafting Op-Eds and Thought Leadership Pieces

When legal issues hit the news, media outlets and professional publications often want expert commentary. GenAI can help you draft op-eds quickly enough to be relevant while maintaining your voice and expertise.

Start by explaining your core argument to the LLM, along with your target publication and audience. Ask it to draft a piece that makes your point accessibly. Then revise extensively, because your authentic voice and genuine expertise are what make the piece worth publishing..

i. Preparing Testimony or Public Comment

When you're testifying before a legislative committee or submitting public comment on proposed regulations, GenAI can help you organize your expertise into the format these venues require. It can help you structure testimony for maximum impact and anticipate questions from lawmakers or regulators. Research from the GWU Regulatory Studies Center confirms that agencies focus on comments providing economic or technical analysis and tend to disregard value-laden or vague submissions.⁴ Academics and policymakers often talk past

⁴ Aryamala Prasade, *Are Agencies Responsive to Mass Comment Campaigns?*, GW Reg. Stud. Ctr. (2019), <https://regulatorystudies.columbian.gwu.edu/are-agencies-responsive-mass-comment-campaigns> (last visited Feb. 23, 2026).

each other. Researchers write for peer reviewers; policymakers read for constituents. Unfortunately, the technical language that signals credibility in a law review article is the same language that kills a bill in committee. When research is published primarily in English in contexts where that is not the dominant policy language, the problem compounds: translation is never neutral, and technical legal or scientific terms rarely survive it intact.⁵

Policy briefs are short (usually 2-4 pages), focused documents that translate research into actionable recommendations for policymakers. They require different structure and style than academic writing. They typically have executive summaries instead of introductions, bullet points instead of footnotes, and recommendations instead of conclusions.

GenAI can help tweak your language. Upload your article or provide a detailed summary, then prompt:

Transform this law review article about qualified immunity reform into a 3-page policy brief for state legislators. Focus on practical implications and include clear policy recommendations. Use accessible language. Structure: executive summary, problem statement, analysis, recommendations.

The result won't be perfect, but it gives you a structure from which to work, instead of staring at a blank page wondering how to make your 80-page article digestible for legislators who have seven minutes to read it.

C. Scholarship

Service work keeps institutions running, but scholarship defines professional identities. The relationship between GenAI and scholarship is more complex than its relationship to service because the stakes are higher, so using GenAI for scholarship requires careful consideration. Teaching applications feel straightforward, and service work feels safe, but scholarship? That's your

⁵ Kenza Khomsi, et al., *Bridging Research-Policy Gaps: An Integrated Approach*, 8 ENVIRONMENTAL EPIDEM. at 2, (Jan. 17, 2024), doi:[10.1097/EE9.0000000000000281](https://doi.org/10.1097/EE9.0000000000000281), https://journals.lww.com/environepidem/fulltext/2024/02000/bridging_research_policy_gaps_an_integrated.10.aspx (last visited Feb. 23, 2026).

intellectual reputation. Your voice. Your contribution to legal thought. The idea of GenAI involvement can feel like academic heresy.

You undoubtedly already use technology throughout your writing process. You use word processors that autocorrect your spelling, citation managers that format and reorder your footnotes, and databases that help you find relevant cases. GenAI is another tool in that ecosystem. You should use it in ways that enhance rather than compromise your scholarship.

This section explores how GenAI can support your scholarly pursuits.

See Chapter 12, Policies for Scholarship, Tenure & Promotion, for a discussion on what policies are needed in the age of GenAI.

1. Research and Idea Development

a. The Blank Page

The hardest part of scholarship is often the beginning. You know there's something worth writing about, but you can't quite articulate what it is. Or you have a research question, but you're not sure if anyone else has already answered it. For years, I deferred writing because I was convinced someone had already made the argument, and made it better. What finally broke the paralysis was recognizing that legal scholarship is a conversation, and a conversation needs more than one voice. You don't have to be the first (or smartest) person to care about a question. You just need to be the next one. And you might use GenAI to help you figure out what you can add to the conversation.

b. Literature Review Assistance and Gap Identification

Traditional literature reviews involve hours in databases, following citation trails, and trying to remember which article made which argument. GenAI can accelerate this process, though not replace it.

Start by using your regular legal research databases (Westlaw, Lexis, HeinOnline) to identify the major pieces in your area. Then use GenAI (e.g., Claude Project or Claude Cowork or NotebookLM) to help you synthesize what you've found. Upload abstracts or summaries and ask questions like:

Based on these fifteen articles about Fourth Amendment doctrine and emerging technology, what are the main schools of thought? Where do scholars disagree? What questions appear unresolved?

GenAI can identify patterns across sources faster than you can, highlighting debates you might have missed and connections between seemingly disparate pieces. It's particularly good at spotting when scholars are talking past each other by using different terminology for similar concepts.

Better yet, use NotebookLM. Google's free research tool is built specifically for the kind of work this workflow requires. Where a general-purpose GenAI tool synthesizes across your sources using its broader training data (which introduces hallucination risk), NotebookLM works exclusively within the documents you upload. Every response it generates is grounded in your sources, and every claim comes with an inline citation that links directly to the specific passage in the original document.

Upload your fifteen Fourth Amendment articles as full PDFs rather than abstracts, and NotebookLM can map the thematic terrain across all of them. It will identify the main schools of thought and flag where scholars appear to agree but who are using different terminology. It will also raise the questions that keep appearing without resolution. It can generate a structured literature review and identify the foundational works that others keep citing. It will also tell you what it cannot find, which is often the most useful output of all, because a gap in the literature is a scholarship opportunity.

The recent addition of Deep Research takes it a step further. If your uploaded sources don't fully answer a question, NotebookLM can now search externally and pull those new sources directly into your notebook (you can accept the source or not) so your research base grows as you work rather than requiring you to start over. A caveat: GenAI cannot reliably find sources on its own. If you ask it to "find articles about qualified immunity reform," it will absolutely make up plausible-sounding citations to articles that don't exist or that come from odd sources. Though, as I mentioned, LLMs are improving. Uploading your notes or just a single article will help to ground it and reduce hallucinations. Use traditional research tools to find sources, then use GenAI to help you make sense of what you've found.

c. Brainstorming Research Questions and Hypotheses

When you're stuck in the "I know there's something here, but I can't figure out what" phase, GenAI can be a surprisingly effective brainstorming partner. Describe your area of interest and ask it to generate potential research questions.

For example: "I'm interested in how courts are handling AI-generated evidence in criminal cases. What are potential research questions that would make a unique contribution to scholarship in this area?"

GenAI will suggest questions ranging from thought-provoking to weird. Most will be mediocre. But the process of evaluating those suggestions often helps you find what looks interesting to you. It's like talking through your ideas with a colleague, except this colleague has no ego investment in the suggestions and won't be offended when you reject that terrible idea.

d. Exploring Theoretical Frameworks

Legal scholarship often benefits from theoretical grounding, but not every law professor has extensive training in political theory, economics, philosophy, or social science methodology. GenAI can help you explore theoretical frameworks that might inform your analysis.

Ask it to explain different theoretical approaches to your topic:

How would a law and economics scholar analyze this issue differently from the way that a critical legal studies scholar would? What about a legal realist or a formalist?

GenAI can outline different theoretical lenses and suggest how each might frame your research question. You'll still need to do the deep reading to apply these frameworks (theoretical sophistication requires more than GenAI-generated summaries), but GenAI can help you identify which theories might be worth exploring.

e. Identifying Interdisciplinary Connections

Similarly, some of the best legal scholarship draws on insights from other disciplines. GenAI can help you identify relevant research from psychology, sociology, economics, computer science, or other fields that might inform your legal analysis.

Describe your legal issue and ask: "What research from behavioral economics might be relevant to understanding how people respond to contract disclosure requirements?"

GenAI can point you toward concepts and studies you might not have encountered otherwise.

2. Specialized Tools for Research and Synthesis

Although general-purpose GenAI tools like ChatGPT or Claude can support scholarship in the ways I've described, some specialized tools are designed specifically for research and analysis. These are worth knowing about because they handle certain scholarly tasks better than general chatbots.

a. The AI Tools Already in Your Legal Research Subscriptions

Before you reach for Claude or ChatGPT to help with legal research, it's worth understanding what your existing Westlaw and Lexis subscriptions already include. Both platforms have embedded AI research tools directly into their interfaces — Westlaw's AI-Assisted Research and Ask Practical Law AI, and Lexis+ AI — and these tools are architecturally different from the general-purpose models discussed elsewhere in this chapter.

The difference is retrieval-augmented generation, or RAG (see Chapter 2 for details on RAG). Rather than drawing on training data that may be outdated or simply wrong, these tools query the platform's proprietary legal database first and then generate a response. When you ask Lexis+ AI about Fourth Amendment doctrine, it's searching its legal database, not relying on whatever a language model absorbed during training. That distinction shows in the results.

The research confirms that this matters, but only up to a point. A rigorous Stanford study of legal RAG systems found that Lexis+ AI, Westlaw AI-Assisted Research, and Ask Practical Law AI hallucinated substantially less than general-purpose GPT-4 — but the "hallucination-free" marketing language that legal AI vendors have embraced is not supported by the evidence. The studied tools hallucinated between 17% and 33% of the time, depending on question type, with particular failure modes in distinguishing holdings from dicta, applying the right

jurisdictional authority, and working through multi-step legal analysis.⁶ A separate comparative study found Lexis+ AI achieving the highest accuracy rate among the AI tools tested (58%) compared to 52% for Copilot, 50% for DeepSeek, and 30% for ChatGPT-4, with a hallucination rate still running at 20%.⁷ In practice, the commercial legal research tools outperform general-purpose AI for legal citation work, but not by enough to skip verification.

There's another distinction that even experienced Westlaw and Lexis users sometimes miss: the traditional research interface and the AI interface within these platforms are not the same thing, and they don't produce the same results. The traditional Boolean and keyword search retrieves documents. The AI interface interprets your question and generates a response *based on* retrieved documents. For locating sources, the traditional interface remains more reliable. For understanding what those sources say once you've found them, the AI overlay can help but it adds a layer of interpretation that requires your verification.

The contents of each platform's RAG database also matter. One careful analysis of CoCounsel found that because the tool is built primarily on caselaw, it can underperform on statutory or regulatory questions, either lacking the relevant statute or weighting case references more heavily than the question warrants.⁸ Lexis+ AI includes Matthew Bender Treatises and Practice Guidance in its retrieval database, which helps on questions where secondary sources provide useful framing. Knowing this is part of knowing when to trust the output.

A few platform-specific notes:

Lexis+ AI draws on Lexis's caselaw database and incorporates Matthew Bender Treatises and Practice Guidance into its RAG. It performs best on questions where secondary sources provide useful context and tends to handle false premises better than Westlaw's tools. In testing, it would often correct a faulty assumption rather than run with it.

Westlaw AI-Assisted Research is strong on caselaw analysis but has been found to produce more reasoning errors than Lexis+ AI. In the Stanford study,

⁶ Varun Magesh et al., *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, 22 J. EMPIRICAL LEGAL STUD. 216, 216 (2025).]

⁷ Bakht Munir et al., *Evaluating AI in Legal Operations: A Comparative Analysis of Accuracy, Completeness, and Hallucinations in ChatGPT-4, Copilot, DeepSeek, Lexis+AI, and Llama 3*, 53 INT'L J. LEGAL INFO. 103, 103 (2025).]

⁸ Paul Callister, *Generative AI and Finding the Law*, 117 LAW LIBR. J. 1, 36 (2025)

Westlaw stated the holding of a case as the direct opposite of what the court actually held in multiple instances, and it confused precedential hierarchy in ways that would matter for actual legal analysis.

Ask Practical Law AI (within the Westlaw ecosystem) focuses on practical guidance rather than primary source analysis, which makes it more useful for practitioners than for doctrinal scholarship. For academic research purposes, it's a different tool than AI-Assisted Research and should be evaluated accordingly.

Harvey, Legora, and other emerging AI platforms designed specifically for law firms operate somewhat differently still.⁹ They're typically optimized for transactional and litigation workflows rather than academic legal research. Their coverage and accuracy vary.

The bottom line is what it's always been: use the traditional research functions in Westlaw, Lexis, and HeinOnline to find your sources. The AI tools on these platforms can help you get oriented in an unfamiliar area or synthesize what you've already found, but they don't eliminate the verification requirements that apply to any AI-assisted work. If anything, the institutional credibility of a Westlaw or Lexis interface can make you more inclined to trust the output than you should be.

b. NotebookLM: Your GenAI Research Assistant

Google's NotebookLM is particularly valuable for legal scholarship. Unlike all-purpose GenAI tools that draw on their training data (and might hallucinate when they don't know something), NotebookLM works *only* with sources you explicitly upload. You can feed it PDFs of cases, law review articles, statutes, legislative history, or your own research notes, and it will only reference those specific materials in its responses.

So, when you ask NotebookLM a question, it will cite to specific passages in your uploaded sources. You can click the citation and see exactly what part of which document the GenAI is referencing.

⁹ One of the reasons I don't like to endorse particular legal tech products is because they might have a short shelf life either by becoming obsolete or by being subsumed through acquisition. See, Oscar Glyn, *Could Harvey and Legora Be on Borrowed Time?*, Law.com (May 19, 2026), <https://www.law.com/international-edition/2026/05/19/could-harvey-and-legora-be-on-borrowed-time-/?slreturn=20260520175547> (arguing that the rise of Anthropic's Claude Legal Tools and pressure from others in the legal tech market are challenging the position of some established legal tech companies).

Practical applications for legal scholarship include:

- **Synthesizing multiple cases:** Upload opinions from a line of cases and ask NotebookLM to identify how the doctrine has evolved and where courts disagree, or what factual distinctions drive different outcomes. GenAI will quote specific passages and cite to particular cases, giving you a roadmap for your own analysis.
- **Analyzing legislative history:** Upload committee reports, floor debates, and statutory text. Ask NotebookLM to identify what problems the legislation was meant to solve and what compromises were made. The GenAI will find relevant passages you might otherwise miss in hundreds of pages of legislative materials.
- **Comparing scholarly approaches:** Upload articles representing different perspectives on an issue. Ask NotebookLM to map out where scholars agree and disagree and what assumptions underlie their arguments so you can pinpoint exactly where your own argument fits in conversation.
- **Organizing your research notes:** If you're like most scholars, you have documents accumulated over months of research. Upload them to NotebookLM and ask it to identify themes and organize your thinking. It will also uncover gaps in your analysis you may have stopped noticing because you've been too close to the material.

One cool feature: NotebookLM can generate an "Audio Overview" that creates a podcast-style conversation between two AI hosts discussing your uploaded sources. If you play it within the app, you can even ask the "podcasters" a question. I find it useful for getting a fresh perspective on the research I've compiled, though hearing GenAI voices earnestly discuss your work-in-progress can be surreal. I often share them with my students (some love them while others think the speakers are annoying).

c. Other Research-Focused Tools

Several other specialized tools deserve mention, though the landscape changes quickly and your mileage may vary:

- **Elicit** helps you find and synthesize research papers, primarily in scientific fields. It's less useful for legal scholarship than for empirical or interdisciplinary work, but if you're incorporating social science research into your legal analysis, Elicit can help identify relevant studies and extract

key findings. Just remember that it's designed for scientific papers, so its understanding of legal sources is limited.

- **Semantic Scholar** uses GenAI to help you navigate academic literature, suggesting related papers and highlighting influential citations. The legal coverage is improving but still not as comprehensive as HeinOnline or SSRN for finding legal scholarship specifically.
- **Consensus** searches academic research and provides GenAI-generated summaries of what the literature says on specific questions. Like Elicit, it's more valuable for empirical research than doctrinal legal analysis, but it can be helpful when your legal scholarship intersects with empirical questions.
- **Perplexity** is an AI-powered search engine. Where ChatGPT or Claude generate responses from training data, Perplexity searches the web in real time, synthesizes what it finds, and – critically – cites every source with clickable links. The result is something closer to a research assistant than a search bar. For researchers, the most relevant features are its focus modes and Deep Research capability. Academic mode prioritizes peer-reviewed and scholarly sources over general web content, which helps keep out SEO-optimized nonsense. Deep Research performs dozens of searches, reads hundreds of sources, and synthesizes the results into a comprehensive report, completing most tasks in under three minutes. The key difference from NotebookLM is directionality. NotebookLM works inward: it synthesizes documents you've already found. Perplexity works outward: it goes to find what you haven't found yet.

Use specialized research tools when you need GenAI to interact with specific source documents you've already identified. Use general GenAI tools when you need help thinking through your own ideas or organizing your own writing. Don't use any GenAI tool to find legal sources you haven't independently verified.

3. Writing Process Support

Writing is hard.

a. Overcoming Writer's Block with Targeted Prompts

You know what you want to say, but the words won't come. You've rewritten the same paragraph five times and it's getting worse, not better. GenAI can help you break through.

Explain your stuck paragraph to the GenAI conversationally, as if you're talking to a supportive colleague. "I'm trying to explain why the court's reasoning in this case is flawed, but I can't figure out how to structure the argument. Here's what the court said, and here's why I think it's wrong."

GenAI will generate a draft. It should not be what you ultimately use, but it gets words on the page. Sometimes that's enough to break the paralysis. I use this technique all the time. You can respond to what the GenAI wrote and refine it or realize that you disagree with the approach it took (which helps you understand what approach you do want).

b. Restructuring Arguments for Clarity

You've written a section that's awkward. The argument is there, but the organization isn't helping readers follow your logic. GenAI can suggest alternative structures.

Give it your existing draft and ask: "This section argues that current copyright doctrine inadequately addresses GenAI-generated works. Suggest three different ways to reorganize this argument for maximum clarity and persuasiveness."

c. Generating Alternative Phrasings

GenAI is great at making complex legal concepts accessible. Give it your technical explanation and ask it to rephrase for different audiences: law students, practicing lawyers, judges, interdisciplinary scholars, or the general public. Seeing multiple versions helps you choose the level of complexity that serves your specific piece.

d. Creating Outlines from Rough Notes

Many scholars think through ideas by taking notes: observations, questions, half-formed thoughts, quotations from sources. When you're ready to turn those notes into structured writing, GenAI can help create an organizational framework. Even if it's handwritten notes with doodles on it. Take a picture and upload it to the GenAI. It can make sense of it.

Upload your notes and ask: "Based on these research notes about workplace surveillance and employee privacy, create an outline for a law review article. Suggest a logical progression of arguments and identify areas that need additional development."

e. Citation and Technical Tasks

GenAI can help with citation formatting, but with massive caveats. It's decent at understanding Bluebook rules (it might even suggest the applicable rule) and can suggest proper formatting for different source types. If you're unsure whether a blog post should be cited like a newspaper article or something else, GenAI can point you toward the right rule.

But never, ever trust GenAI-formatted citations without personally verifying them. GenAI might apply the wrong rule or invent formatting conventions that sound plausible but violate the Bluebook.¹⁰ It can't do small caps. It often uses quotation marks rather than italics.

f. Checking Cross-References

When you're citing to specific pages in a source you've referenced earlier in your article, GenAI can help you verify consistency. It can check whether your "supra" references point to the right footnote and whether your page numbers align with your earlier citations. I find it easier to use than Word.

One underappreciated capability: when you're working in Track Changes and edits haven't yet been accepted, Claude¹¹ can account for the shifting footnote numbers that result.

g. Formatting Footnotes and Bibliographies

For journals that require specific formatting beyond standard Bluebook, GenAI can help adapt your citations to meet those requirements. It can also help convert between different citation systems if you're submitting to a publication that uses something other than Bluebook.

h. Revision and Refinement

Someone told me that the best writing is rewriting. GenAI can support your revision process in ways that make your scholarship stronger.

1) Analyzing Argument Structure and Flow

When you've finished a draft, you're often too close to it to evaluate whether your argument flows logically. Before GenAI, I handled this by leaving it for

¹⁰ The Bluebook is already confusing enough without adding AI hallucinations to the mix. See *The Bluebook: A Uniform System of Citation* (21st ed. 2020) (all 500+ pages of citation rules that someone apparently thought law students needed to memorize).

¹¹ I haven't tested this functionality in other platforms.

several days so I could evaluate it with fresh eyes. Now, GenAI can provide a fresh perspective, and you can do it immediately instead of waiting days or asking an already busy colleague.

Prompt it:

You are a law professor reviewing this draft to provide feedback to a colleague. Summarize the main argument of each section in this draft. Then identify any logical gaps or places where the argument seems to jump without adequate transition. List any other improvements you could recommend.

The GenAI's summary will show you whether your argument reads the way you intended. If the GenAI misunderstands your point, your readers probably will too. If it identifies gaps, those are places where you need to add explanation or evidence.

2) Identifying Gaps in Reasoning

GenAI can play devil's advocate, pointing out potential counterarguments or weaknesses in your analysis. It's like having a colleague read your draft and ask hard questions, except you can do it at 2 a.m. when your colleagues are sensibly asleep.

Try: "I'm arguing that current contract law inadequately addresses algorithmic decision-making. You are another law professor who is reviewing my draft. What are the strongest counterarguments to my position? Where is my reasoning potentially vulnerable? Provide meaningful constructive feedback."

3) Improving Clarity and Readability

Legal scholarship has a reputation for being impenetrable. GenAI can help you write clearly without sacrificing analytical rigor.

Give it a dense paragraph and ask: "Revise this for clarity while maintaining technical accuracy. Identify any jargon that needs definition or concepts that need more explanation."

The LLM's revision might oversimplify in some places, but it will show you where your writing assumes knowledge your readers might not have. You can then revise in ways that preserve sophistication while improving accessibility.

4. Miscellany

a. Generating Peer Review Response Drafts

When you receive peer review comments, you need to respond thoughtfully to each suggestion. GenAI can help you draft initial responses, especially for technical or formatting suggestions that don't require deep analytical engagement.

For substantive criticisms, GenAI can help you think through how to address the reviewer's concerns while maintaining your argument.

b. Translating Scholarship for Public Audiences

You've written a thoughtful law review article about housing discrimination and algorithmic decision-making. Exactly seven people will read it, and three of them are citing you only to disagree with you. Meanwhile, policymakers are making decisions about the exact issues you've analyzed, and they have no idea your research exists.

This is the academic translation problem. Legal scholarship often stays trapped in law reviews when it could inform public debate and policy decisions. GenAI can help you translate your scholarship into forms that reach broader audiences without dumbing down your analysis. The *Miami Herald* or the *Los Angeles Times* doesn't want your Section III doctrinal analysis. It wants "An algorithm is deciding whether you can rent an apartment in (Miami or LA), and no one is checking if it's legal."

c. Creating Blog Posts and Op-Eds

When legal issues make news, editors want expert commentary quickly. You have expertise, but transforming your scholarly analysis into 800 words of accessible prose under deadline pressure is hard. GenAI can accelerate the process.

Start by explaining your argument conversationally to GenAI, as if you're talking to a smart friend who isn't a lawyer:

"I study employment discrimination law. The Supreme Court just decided a case about religious exemptions that I think misapplies existing doctrine. Here's why: [explain your analysis]. Draft an 800-word op-ed for a general news outlet explaining what the Court got wrong and why it matters for workers."

You'll get:

- Use clearer language than your law review article
- Include metaphors or examples that make doctrine accessible
- Structure the argument for quick reading
- Remove most of the nuance that makes your scholarship careful

This is where you come in. Restore necessary nuance, verify that simplification hasn't created inaccuracy, add your voice and style, and make sure you're comfortable standing behind every claim publicly.

The timeline advantage is real: GenAI can help you draft something within hours of a decision being released, when editors are most interested and before the news cycle moves on.

d. Conference Presentations and Talks

You've been accepted to present at AALS. Now you need to convert your research into an effective presentation. This is different from writing, and GenAI can help with the transformation.

1) Converting Written Work into Presentation Format

Written scholarship and oral presentations require different structures. What works on the page doesn't always work when spoken. GenAI can help you reorganize your research for oral delivery.

Upload a section of your article or outline your argument, then try:

"I'm presenting this research at a 20-minute faculty workshop. The audience is law professors from various fields, not all experts in my area. Convert these written materials into a slide deck." It saves an enormous amount of time and creates prettier slides than Powerpoint does.

2) Creating Effective Slides

We've all suffered through terrible PowerPoint presentations: slides loaded with text, bullet points that are really paragraphs, fonts too small to read from the back of the room. GenAI can help you design slides that support rather than duplicate your spoken presentation.

Gamma is a wonderful platform that I use regularly to translate my notes into slides. You can download into Powerpoint or Google Slides format. There are others that people recommend: Plus.ai, napkin.ai, Gen PPT, Beautiful.ai. Claude added this feature, too, and it has improved over time making prettier slides. Tell

the LLM how long the presentation or lecture is going to be, and it will suggest times for each slide. I find that using ChatGPT requires a lot of iteration and can often result in an error. NotebookLM, too, will autogenerate slides based on the resources you gave it.

3) Preparing for Q&A

One of the most anxiety-inducing parts of presenting is the question period. GenAI can help you anticipate questions and prepare responses.

Try this:

I'm presenting research arguing that current GenAI liability frameworks are inadequate for addressing algorithmic discrimination. What questions might I get from: (1) professors skeptical of regulation, (2) scholars focused on civil rights enforcement, and (3) practitioners concerned about compliance costs? For each category, suggest the strongest challenges to my argument and how I might respond.

5. Managing the Publication Process

Getting scholarship published involves more than just writing. You need to navigate submission processes, respond to editors and reviewers, and handle revisions. GenAI can help with the administrative and communicative parts of publication.

a. Cover Letters to Journals

When you submit an article to law reviews, you might include a cover letter introducing your article. These letters should be professional, concise, and highlight what makes your piece a good fit for that particular journal.

GenAI can help you draft effective cover letters:

Draft a cover letter submitting my article on trademark law and NFTs to the Stanford Law Review. The article makes three contributions: [list them]. I've published previously in [other journals]. Keep it professional but not overly formal, one page maximum.

The GenAI will generate a template you can customize for different journals. You'll personalize it by:

- Explaining why this specific journal is a good fit
- Mentioning if you have connections to the institution (alumnus, previous symposium participant)
- Noting any time-sensitive aspects (related to pending Supreme Court case, recent regulatory changes)
- Adding your authentic voice

Create one strong cover letter with GenAI assistance, then adapt it for different journals rather than starting from scratch each time.

b. Responding to Revise-and-Resubmit Requests

When a journal offers to publish your article contingent on revisions, you should respond to editorial suggestions thoughtfully.

GenAI can help structure your response:

A law review has requested revisions to my article [provided]. Their main concerns are: [list or upload feedback]. Draft a response memo that: (1) thanks the editors for their feedback, (2) explains how I've addressed each concern, (3) explains where I've chosen not to make suggested changes and why. Use a professional and collegial tone.

It's particularly useful for managing major revisions where you need to track multiple editorial requests and demonstrate you've considered each one seriously.

c. Navigating Peer Review

For peer-reviewed journals you'll receive anonymous reviewer feedback that requires written response. These responses are often high-stakes: poor responses can sink your article even if the research is strong.

GenAI can help you craft professional, persuasive responses to reviewer criticism:

A peer reviewer criticizes my methodology in my article [provided], arguing I should have used quantitative analysis instead of qualitative case studies. I disagree because [reasons]. Draft a response that: (1) acknowledges the value of quantitative methods generally, (2) explains why qualitative analysis is appropriate for my research questions, (3) cites methodological literature supporting my approach. Tone should be respectful and scholarly, not defensive.

Important caveat: Never let GenAI write your substantive responses to intellectual criticism. Reviewers are evaluating your scholarly judgment. GenAI can help with phrasing and structure, but the intellectual content of your response must be authentically yours.

d. Preparing Article Summaries and Abstracts

Many journals require abstracts, and you may need to create summaries for SSRN, faculty pages, or promotional materials. GenAI can help you distill your work into different formats.

For a 150-word abstract:

Summarize this 30,000-word article into an abstract of 150 words or less. Include: the central argument, methodology, main findings, and contribution to scholarship. Write for an academic audience familiar with the field but not experts in this specific area.

For a one-paragraph summary for general audiences:

Explain this article's core argument in one paragraph that would make sense to lawyers who don't work in this field. Focus on why the issue matters and what your research contributes, not methodological details.

Review these summaries carefully. GenAI-generated abstracts sometimes miss nuances or emphasize the wrong aspects of your work. The summary should accurately reflect what you think is most important about your scholarship, not what the GenAI thinks is important.

e. Critical Limitations in Scholarship

Everything in this section has come with caveats, but we need to address the biggest limitation head-on: GenAI is fundamentally unreliable for legal citations and legal research (today, but it is getting better quickly). This isn't a minor issue or something that might improve with better prompting. It's a core limitation that every legal scholar must understand.

f. Why GenAI Cannot Replace Legal Research Databases

Westlaw, Lexis, Bloomberg Law, and other legal research platforms are authoritative databases that retrieve actual documents, and that are leveraging

GenAI. When you search for a case the traditional way, you get the case. When GenAI generates a citation, you get something that *looks like it might* be a case.

Do not use GenAI as a shortcut for legal research. Do not ask it to find supporting citations for a proposition. Do not trust case summaries it generates without reading the actual cases. The risk isn't worth it, and there's no time savings that justifies compromising your scholarship with fabricated sources.¹²

g. The Necessity of Independent Verification

If you use GenAI anywhere in your research or writing process, you must independently verify every factual claim, every citation, every quotation, and every legal proposition. This means:

- Read every case that you cite
- Verify every quotation against the original source
- Check that your citations support the propositions you attribute to them
- Confirm that statutes and regulations are current and accurately cited
- Ensure that academic sources exist and say what you claim they say

The verification requirement doesn't make GenAI useless for scholarship. It just means you're using it for ideation, organization, and drafting support rather than research or fact-checking.

h. Maintaining Scholarly Integrity and Voice

Your scholarship must be yours. Your arguments, your analysis, your insights, and your voice must be authentic.

GenAI can help you articulate your ideas more clearly or organize them more effectively. If you find yourself using GenAI-generated text without substantially revising it to reflect your own thinking, you're outsourcing intellectual work that must remain your own.

Ask yourself, would you be comfortable explaining your article to a colleague and describing exactly how you used GenAI in writing it? If the honest answer would embarrass you, that's a sign that you've crossed from appropriate use to academic dishonesty.

¹² *Id.*

Your scholarly reputation depends on readers trusting that your work represents your genuine expertise and analysis. Use GenAI as a tool to support your thinking, not as a substitute for doing the thinking yourself.

GenAI is most useful in the messy middle of the writing process, after you've done your research and before you finalize your draft. It helps with organization, articulation, and revision. It cannot do your research and cannot substitute for your expertise. Verify everything, and never let it compromise the integrity that makes your scholarship worth reading.

D. Institutional Documentation

If service is what we owe our institutions and scholarship is what we owe our disciplines, administration is what we owe our students and, frankly, what gets demanded of us whether we want to do it or not. The good news is that much administrative work follows predictable patterns that GenAI can help streamline.

Using GenAI for administrative work raises important questions about professional responsibility. Every semester brings a new round of administrative tasks: updating syllabi, revising assignments, creating schedules, preparing materials. GenAI can't make administration disappear (if only!), but it can make these tasks significantly less soul-crushing. This section explores how to use GenAI for the administrative work that keeps a law school running while preserving your energy for the parts of your job that truly require legal expertise.

Law schools run on documentation. Assessment reports, committee minutes, strategic plans, self-studies, program reviews, and endless reports for endless purposes. GenAI can make this less painful.

1. Writing Annual Reports

Faculty annual reports typically require you to list your teaching, scholarship, service, and professional activities. The format is standardized and the content is factual.

GenAI can help you transform your raw list of activities into properly formatted narrative text. Give it your teaching assignments, publications,

presentations, and service commitments, along with your institution's reporting template. Ask it to draft the report in the required format.

You'll review for accuracy, of course, add context about significant accomplishments, and ensure the tone is appropriately professional. But you won't have to spend hours drafting prose to describe activities you've already listed elsewhere.

2. Creating Assessment Documentation

Accreditation requires evaluation of learning outcomes, assessment methods, data collection, analysis, and action plans. Gathering and organizing documentation is important but time-consuming. GenAI can help streamline the process.¹³

Give the LLM your learning outcomes and ask it to suggest assessment methods aligned with each outcome. Provide assessment data and ask it to identify patterns or areas needing improvement. Share previous action plans and ask for draft updates based on new data.

3. Drafting Institutional Self-Studies

When your institution undergoes accreditation review, program evaluation, or strategic planning, someone has to write the self-study. If that someone is you, GenAI can help organize the massive amount of information these documents require. Often you're bringing together sections written by other authors. GenAI can help it sound like "one voice."

GenAI can help you structure the document according to accreditation standards. It can synthesize data from multiple sources and identify themes across different program areas. And, of course, it can help draft sections that require primarily factual reporting.

4. Time Management and Planning

Administration isn't just about creating documents. It's about managing complex projects and somehow getting everything done before deadlines that

¹³ See BENJAMIN S. BLOOM ET AL., *TAXONOMY OF EDUCATIONAL OBJECTIVES: THE CLASSIFICATION OF EDUCATIONAL GOALS* (1956).

once seemed far away until, suddenly, they aren't. The semester is both endless and over before you've finished grading the first assignment. You are simultaneously managing the immediate (the student who emailed at 11 p.m. about an extension), the near-term (the committee report due Friday that you haven't started), and the horizon (the conference proposal, the curriculum revision, the accreditation self-study that someone assures you is "not that big a deal").

GenAI can't add hours to your day, but it can compress the time certain tasks consume. The most underused application of GenAI is turning chaos into order. GenAI can break it into sequenced, actionable steps with realistic time estimates. Feed it your faculty handbook's committee charge and a description of your current workload, and ask it to draft a project timeline. You then edit it, which is always faster than originating it. This can help students, too, to figure out how they can manage life and law school.

GenAI can also help when you have three things are equally urgent and you need to think out loud about tradeoffs. Prompting a model to help you analyze competing demands or identify what must happen first versus what merely feels urgent is a legitimate use of the technology.¹⁴

Meeting preparation is another area where the time investment is often disproportionate to the outcome. A faculty meeting agenda item that requires you to remember six months of context, synthesize two competing proposals, and show up with something cogent to say can be prepared with GenAI in fifteen minutes. You provide the background, it drafts a briefing document or a set of talking points. You walk in prepared rather than reconstructing institutional memory from your email inbox at 8:45 a.m. Similarly, after a meeting, GenAI can convert rough notes or a transcript into a structured action item list, assigned by person and sorted by deadline. This sounds mundane until you've watched an entire committee's decisions dissolve because no one wrote anything down in a usable form.

¹⁴ See Dwight D. Eisenhower (attributed), *Address at the Second Assembly of the World Council of Churches* (1954) ("I have two kinds of problems: the urgent and the important. The urgent are not important, and the important are never urgent."). Whether Eisenhower actually said this is disputed, but the framework is useful regardless of its provenance.

5. Organizing Complex Projects

When you're managing a complicated project (like curriculum redesign), GenAI can help you break it into manageable components.

Describe the project and ask GenAI to suggest phases and flag potential complications. For example: "I'm chairing a committee at a law school to revise our legal writing curriculum. We need to review current program outcomes, gather stakeholder feedback, research best practices, develop proposed changes, secure faculty approval, and implement new curriculum. What's a logical project plan?"

GenAI will suggest a structure. You'll modify it based on your institutional politics and real-world constraints the GenAI doesn't know about. But you'll have a framework for thinking through the project systematically.

a) Creating Timelines and Milestones

Once you've defined a project, you need realistic timelines. GenAI can help you create schedules that account for dependencies and build in appropriate buffer time.

GenAI can also help you identify when you're being overly optimistic about how quickly things will happen. If your proposed timeline has you completing a major curriculum revision in six weeks while teaching full-time, GenAI might point out that the timeline seems aggressive given the number of tasks involved. Your colleagues would tell you the same thing, but the GenAI is available at midnight when you're making these plans.

b) Developing Workflow Systems

If you find yourself doing the same administrative tasks repeatedly, GenAI can help you develop systems that make them more efficient.

For example: "I am a law professor who responds to similar student emails multiple times per week. Help me create a workflow for managing these efficiently while maintaining quality and personalization."

GenAI might suggest creating email templates, setting up folders for common request types, blocking specific times for email response, or developing decision trees for common scenarios. You'll adapt these suggestions based on what works with your communication style and institutional context.

The theme across all administrative applications is the same. GenAI handles routine structure and drafting, giving you time to focus on decisions that require your judgment and knowledge of your specific institutional context. Administration will never be the fun part of your job, but at least it can be less time-consuming.

E. Delegation - When to Use GenAI (and When to Think Twice)

This chapter is, as one thoughtful reader put it, a cookbook without a meal plan. Plenty of recipes; not much guidance on how to put a week together. Here's an attempt to remedy that.

Andrew McAfee and Erik Brynjolfsson suggested in *Machine, Platform, Crowd* (2017) that the most productive candidates for automation are tasks that are dull, dirty, dangerous, or dear.¹⁵ They were describing industrial robotics, but the framework translates to academic work reasonably well with some important modifications.

Dull tasks are repetitive and formulaic: committee minutes, annual report narratives, cover letters to journals, citation formatting. GenAI handles these well. The patterns are predictable, the consequences of errors are manageable (you're reviewing before anything goes out), and there's no particular value in having you personally draft them from scratch. Start here if you're new to these tools. Build the habit of generating a draft and then editing, rather than writing from scratch.

Dirty tasks are the tedious ones that involve wading through large volumes of material to find what matters: synthesizing stakeholder feedback from fifty survey responses or organizing three months of scattered research notes into a usable outline. GenAI's expertise is pattern recognition across large bodies of text.

Dear tasks are expensive in time. A first draft of almost anything qualifies. GenAI doesn't make these tasks free, but it compresses the time they consume enough to matter across a semester. This is a starting point but never the final product.

Dangerous is where you apply the brakes. Legal citations are dangerous to delegate to AI because of the hallucination risk and the professional consequences of citing a nonexistent case. So is any content where your credibility is directly on

¹⁵ ANDREW MCAFEE & ERIK BRYNJOLFSSON, MACHINE, PLATFORM, CROWD: HARNESSING OUR DIGITAL FUTURE 99-101 (2017).]

the line such as a substantive analysis in a peer review response. The higher the stakes and the harder the error is to catch, the less you should rely on GenAI to get it right. The answer is, as always, careful verification.

A few additional rules of thumb worth knowing:

GenAI works best when you give it something to work with. The more context you provide (your notes, drafts, sources) the better the output. Asking GenAI to generate a literature review cold is asking for hallucinations. Asking it to synthesize the fifteen articles *you've already read* is a reasonable use of the technology.

First draft, not final product. If you find yourself using GenAI output with only cosmetic edits, that might be a signal to check who (or what) is doing the thinking.

Always factor in verification. Time saved drafting can be lost verifying. For scholarship especially, the requirement to verify every factual claim and citation changes the math on how much GenAI actually saves you.

If you're just starting to build these tools into your workflow, beginning with the dull and dirty tasks lets you develop good habits (especially the habit of verifying output) before you move into higher-stakes territory.

CHAPTER 12: GenAI Policies for Scholarship, Tenure & Promotion

Most law schools do not have GenAI related to scholarship or tenure and promotion. As of early 2026, the policy landscape for faculty use of GenAI in scholarship is pretty sparse. Some institutions may have workshops and task forces, but what they largely do not have is a policy. According to *Inside Higher Ed's* 2024 Survey of Chief Academic Officers, only twenty percent of colleges and universities have published a formal policy governing GenAI use in teaching and research.¹ The American Association of University Professors confirmed this paucity in its May 2025 report: ninety percent of institutions have introduced Gen AI initiatives, but the plans and ideas for GenAI adoption haven't yet been turned into concrete rules and guidelines for how faculty members should use GenAI.²

Academic publishers, interestingly, got there faster than the universities themselves. Initially, *Science* (AAAS) banned AI-generated text outright and treats violations as scientific misconduct.³ But that policy has been updated. The initial total ban was softened as the scientific community's views developed, and *Science* now permits GenAI use in research with disclosure in the methods section, while still restricting AI-generated text in the manuscript itself.⁴ *Nature* prohibits listing GenAI as an author and bans AI-generated images. The only GenAI assistance it permits is copy editing.⁵ The emerging consensus across Elsevier, Wiley, Springer Nature, and IEEE rests on two related principles: human authors bear sole responsibility for all content, and clear disclosure of GenAI assistance (tool name,

¹ 2024 Survey of College and University Chief Academic Officers, *Inside Higher Ed* (Apr. 15, 2024), <https://www.insidehighered.com/reports/2024/04/15/2024-survey-college-and-university-chief-academic-officers>.

² *Artificial Intelligence and Academic Professions*, AAUP (May 2025), <https://www.aaup.org/reports-publications/aaup-policies-reports/topical-reports/artificial-intelligence-and-academic>.

³ Holden Thorp, *ChatGPT Is Fun, but Not an Author*, 379 *SCIENCE* 313 (2023), <https://www.science.org/doi/10.1126/science.adg7879>.

⁴ *General Policies*, *SCIENCE*, <https://www.science.org/content/page/science-journals-editorial-policies> (last visited Mar. 15, 2026); see also *AI Usage Guidelines*, *SCIENCE*, https://www.science.org/cms/asset/5ba9af0c-e29f-467c-983d-579a1be18326/science_journals_guidelines_for_ai_use.pdf (last visited Mar. 15, 2026).

⁵ *Artificial Intelligence (AI)*, Springer Nature Portfolio Editorial Policies, <https://www.nature.com/nature-portfolio/editorial-policies/ai> (last visited Mar. 15, 2026).

version, and purpose) is expected.⁶ Law reviews are starting to follow. For example, the Harvard Journal of Law & Technology (JOLT) allows authors to use generative AI, provided all usage is disclosed in a title page footnote, including specific tools and functions.⁷ Authors are responsible for accuracy, and failure to disclose AI use may lead to rejection. However, the Harvard Law Review writing competition strictly prohibits AI use.⁸ The current edition of *The Bluebook* now includes (poorly conceived in my opinion) instructions for citing AI-generated content, including a directive that authors save a screenshot of AI output as a PDF.⁹

A. The Three Problems That GenAI Creates for Scholarship

Kevin Frazier and Alan Rozenshtein's forthcoming article *Large Language Scholarship* has some really interesting ideas that seem exactly right to me.¹⁰ They recognize that law professors already use GenAI to write articles in weeks rather than months, and that law reviews face an unprecedented volume of GenAI-polished submissions. Meanwhile, law schools are trapped in misguided debates about plagiarism and disclosure that miss the full scope of what is happening.¹¹

⁶ Elsevier, *The Use of AI and AI-Assisted Technologies in Writing for Elsevier*, <https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier> (last visited Mar. 15, 2026); see also IEEE, *Author Guidelines for Artificial Intelligence (AI)-Generated Text* (Apr. 16, 2024), <https://open.ieee.org/author-guidelines-for-artificial-intelligence-ai-generated-text/>.

⁷ *Submission*, HARV. J. LAW & TECH. (July, 24, 2025), available at <https://jolt.law.harvard.edu/submissions#:~:text=Completeness:%20Submitted%20articles%20must%20be,Our%20Review%20Process>, (last visited Mar. 15, 2026) (“Use of AI Policy: Author(s) must document any use of AI tools in the title page footnote. Please see the detailed requirements above”).

⁸ *Writing Competition*, HARV. L. REV., available at <https://harvardlawreview.org/writing-competition/> (last visited Mar. 15, 2025) (“Any and all use of generative AI tools to prepare your submission is strictly prohibited. This prohibition includes, but is not limited to, the use of ChatGPT, Claude, and other LLM-based tools to generate, edit, revise, summarize, or otherwise assist in preparing your submission”).

⁹ THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION R. 18.3 (COLUMBIA L. REV. ASS'N et al. eds., 21st ed. 2020).

¹⁰ Kevin Frazier & Alan Rozenshtein, *Large Language Scholarship* (FIU L. REV., forthcoming 2026), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5200768 (last visited Mar. 15, 2026).

¹¹ *Id.*

The existing debate has largely conflated three distinct problems that GenAI creates for faculty scholarship, and separating them is the most useful place to start.

1. Copyright

The first problem is copyright. Generative AI systems are trained on vast quantities of text, and whether that training process infringes the copyrights of authors whose work was copied and ingested is the subject of substantial ongoing litigation.¹² For the individual faculty member producing scholarship, however, copyright is predominantly a question about GenAI output rather than GenAI training. GenAI output is generally not copyright infringement because most AI-generated text recombines content rather than reproducing it.¹³ There is an important exception: the memorization problem, in which LLMs can reproduce verbatim passages from training data, either when prompted or inadvertently, raises genuine infringement concerns that faculty using these tools should understand.¹⁴ Verbatim reproduction is more likely to occur when a user deliberately prompts for it, and when the content is highly distinctive like song lyrics, famous passages, or frequently republished material.¹⁵ Inadvertent verbatim reproduction in ordinary scholarly drafting is unlikely but not impossible, and the risk is real enough to justify the verification requirement that this chapter recommends.¹⁶ A faculty member who uses GenAI to draft a law review article and then independently verifies the output is very unlikely to have

¹² See, N.Y. Times Co. v. OpenAI, Inc., No. 23-cv-11195 (S.D.N.Y., filed Dec. 27, 2023) (alleging that OpenAI and Microsoft used millions of Times articles without authorization to train large language models); Thomson Reuters Enter. Ctr. GmbH v. Ross Intelligence Inc., No. 1:20-cv-00613, 2025 WL 458520 (D. Del. Feb. 11, 2025) (holding that ROSS's use of Westlaw headnotes to train a competing AI legal research tool was not fair use, and currently on interlocutory appeal to the Third Circuit); Bartz v. Anthropic PBC, No. 3:23-cv-03223 (N.D. Cal.), settled (June 2025) (settling for up to \$1.5 billion after the court found that Anthropic's acquisition of pirated books for training purposes was not shielded by fair use, even where the training itself might be); Kadrey v. Meta Platforms, Inc., No. 3:23-cv-03417 (N.D. Cal. June 25, 2025) (granting summary judgment for Meta and holding that training an LLM on copyrighted books was highly transformative and fair use, but cautioning that the result was factually specific and that "in most cases" plaintiffs would likely prevail with a better-developed record on market effects).

¹³ Mark A. Lemley & Lisa Larrimore Ouellette, *Plagiarism, Copyright, and AI*, U. CHI. L. REV. ONLINE (Oct. 24, 2025).

¹⁴ Amy B. Cyphert, *Generative AI, Plagiarism, and Copyright Infringement in Legal Documents*, 25 MINN. J.L. SCI. & TECH. 49, 56–59 (2024).

¹⁵ *Id.* at 51.

¹⁶ *Id.*

committed copyright infringement. The memorization risk is real, but it is a reason for careful review of AI-generated text, not a reason to avoid the tools entirely. As for the underlying question of whether training on copyrighted works is itself infringement, that remains the AI companies' problem to sort out, which is happening in the Northern District of California and in the Southern District of New York.

2. Plagiarism

The more significant concern for scholars is not verbatim reproduction but plagiarism. GenAI reliably reproduces ideas without attribution, even when it does not reproduce exact language, which happens far more frequently than verbatim copying does.

In other words, when a GenAI system regurgitates an idea without accurately attributing its source, and a scholar then uses that idea in his or her own work without attribution, there's a problem, even without copyright concerns. It's not simply that GenAI might fail to acknowledge (or even discern) its source material, or that people often misunderstand how it generates content. The problem is also that using GenAI writing tools makes it harder to remember where the ideas you're using originated.¹⁷ A scholar who asks GenAI to generate a novel thesis and receives one that turns out to be someone else's published argument has not infringed anyone's copyright. But that scholar has plagiarized, whether or not anyone knew it at the time.

The solution is straightforward, and it is the same one this book has been recommending from the start. If you use GenAI for research, verify everything. Better yet, upload your own verified sources that you've read and understand to the GenAI tool. Proper citation follows naturally because you did the work that makes proper citation possible.

3. Scholarly Integrity

The third problem is scholarly integrity more broadly. Even where GenAI use is not copyright infringement and not quite plagiarism, it may still represent a failure to meet the scholarly standards that the academy expects. Verifying sources and engaging seriously with prior literature are not optional steps. Neither is

¹⁷ Lemley, *supra* note 13 at *16.

doing the intellectual work that scholarship is supposed to model for students and other academics. A disclosure statement that reads "I used GenAI to assist with drafting" tells the reader very little. It says nothing about whether the author checked the LLM's citations or whether those sources even existed. It says even less about whether the LLM's claimed novelty was truly novel.¹⁸

Law schools should address all three problems separately, because they call for different responses.

4. Disclosure - When Should Be Required?

Scholarly consensus is beginning to emerge on what disclosure should look like, and might provide guidance for institutions to codify. How much and what kind of disclosure is enough? Some argue persuasively that disclosure must be specific about what ideas came from GenAI and what text was generated or substantially revised by GenAI. They say that an uncredited use of others' ideas in a scholarly article can mislead readers about the author's original contributions and, most important, can prevent the true originators from receiving proper recognition. These scholars propose a strict standard and argue that academics cannot simply claim GenAI as the *source* of an idea without confirming that GenAI originated it. Every GenAI-generated idea must be investigated to identify and properly cite its true source.¹⁹ The proposed standard is admirable in principle, but, I think, largely unworkable in practice. Tracing every GenAI-generated idea back to its true originator assumes that the originator is traceable, which in turn assumes that the model's output can be mapped back through a training dataset that may contain hundreds of billions of words scraped from sources that were never individually catalogued. And it assumes that the output is traceable back to a discrete source or sources. The deeper problem is that GenAI does not read sources and remember them the way that a scholar does. It learns patterns of

¹⁸ The author used GenAI tools in drafting portions of this book (most obviously with prompt output), including Claude (Anthropic), NotebookLM (Google), and, most recently, Claude Cowork (Anthropic). Gemini (Google) and ChatGPT (OpenAI) were used for image and chart generation. The underlying research was entirely the author's own: source cited here was collected, read (or watched), and verified by the author before it entered any GenAI-assisted workflow. All citations were independently confirmed (because they were provided to the LLM) and all arguments were written by the author with an assist from scholars and Claude on sentence clarity and structure. The author takes full responsibility for the analysis and any errors. Readers who have made it this far in a chapter about GenAI disclosure will appreciate that this footnote is, at minimum, disclosed to meet the standard the chapter recommends.

¹⁹ Lemley, *supra* note 13, at *18.

language from an enormous corpus and then produces output by assembling statistically probable sequences of words. Any given sentence in a GenAI response is not drawn from a single source but is a recombination of patterns absorbed from thousands (or millions) of them, with no record of which ones. Asking a scholar to confirm that an LLM "genuinely originated" an idea, rather than absorbed it from someone else's 2019 working paper and another's 1986 law review article buried in the Common Crawl dataset,²⁰ sets a verification standard that no reasonable scholar can meet. The training data is simply too enormous and too poorly documented for that kind of trace work to be feasible in most cases.

The Bluebook provision requiring a saved screenshot of AI output points in the same direction, though I disagree with the underlying premise – not because I use GenAI, but because I understand how GenAI works.²¹ The goal is to give readers and editors enough information to understand exactly where the human author's contribution begins and ends, and that's perfectly reasonable. That goal assumes that the human and AI contributions are separable, i.e., that there is a clean line between what the author wrote and what the tool generated. For most scholars who work seriously with these tools, that line doesn't exist. The author shapes the prompts and selects and rejects outputs. Revision is extensive, and it is the human author who revises. The human conducts the underlying research and determines what is worthy of consideration. Then, the author delivers that curated material as the foundation from which the model works. Identifying where the "human contribution ends" in that process misunderstands what GenAI is: a co-intelligence tool that augments the scholar's own thinking rather than replacing it.

Throughout this book, the guidance on GenAI use has been rooted in academic integrity and professional responsibility principles. A scholar who follows that guidance has already satisfied the legitimate concerns that the screenshot requirement is trying to address. That scholar has conducted her own research and verified every source. She has directed the LLM from a documented corpus of her own assembly and taken full accountability for the final work. The screenshot becomes unnecessary not because transparency does not matter, but because the workflow itself is transparent.

²⁰ See Cyphert, *supra* note 14, at 51 (noting that GPT-3 was trained on the Common Crawl dataset, which includes content ranging from traditional news sites to social media platforms).

²¹ *Bluebook*, *supra* note 9, r. 18.3.

I'm not alone in making this argument. Kevin Frazier and Professor Rozenshtein push back on mandatory disclosure requirements, arguing that demanding transparency fails to advance scholarly values and may undermine authorial responsibility by creating the illusion that disclosure is the relevant obligation rather than quality.²² They reject the methodological transparency analogy on the ground that compositional GenAI assistance is no more a "method" requiring disclosure than is using Westlaw or Grammarly.²³ Academic acknowledgments serve a professional purpose; they recognize contributions and maintain the reciprocal norms that sustain scholarly communities. GenAI has no career to advance and no favors to trade. Listing it alongside human contributors mistakes a tool for a colleague.²⁴ They are skeptical that disclosure promotes accountability, warning that it may do the opposite by inviting authors to hide behind their tools.²⁵ And they find no limiting principle in the argument that readers who dislike GenAI deserve to know when it was used.²⁶ Their bottom line is that mandatory disclosure requirements will quickly become anachronistic as GenAI assistance becomes as universal as spell-check.²⁷

Their preferred framework is what they call "absolute authorial accountability," by which the author is responsible for every word, argument, and citation, regardless of how the text was generated.²⁸ That is a reasonable principle, and, in my opinion, one worth adopting. It doesn't eliminate the need for disclosure. Instead, it reframes disclosure as just the means for making accountability possible. Scholars should disclose GenAI assistance precisely because they bear full responsibility for what they disclose.

There is another complication worth taking seriously: the "AI penalty."²⁹ Studies, though not unanimous, have found that the same work is evaluated less favorably when readers learn it was written with the assistance of GenAI.³⁰ If disclosure systematically disadvantages work that is high quality, that is a terrible

²² Frazier, *supra* note 10, 42-45.

²³ *Id.*

²⁴ *Id.*

²⁵ *Id.*

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.*, 45-47.

²⁹ Lemley, *supra* note 13, at *21.

³⁰ *Id.*

result.³¹ The solution is not to abandon disclosure altogether but to editors to evaluate disclosed GenAI use on its merits. They need to understand how GenAI works and what the author's workflow looks like so that they can fairly evaluate the work.

5. Scholarship Policies

Law school scholarship policy should rest on four principles, though not all four carry equal weight.

The most important by far is absolute authorial accountability: the author is responsible for everything in the published work, full stop. Neither disclosure nor a procedural requirement changes that. Errors that should have been caught are still the author's errors. Just as lawyers cannot point to the LLM and claim the error was its fault, neither can the author.³²

Second, scholars must independently verify all citations, sources, and factual claims in GenAI-assisted work before submission.³³ Given that GenAI does not retain or retrieve sources but instead recombines statistical patterns of language from an enormous and largely opaque training corpus, no AI-generated citation or factual claim should be treated as reliable without independent confirmation. A scholar who submits unverified GenAI output has not just made an error. The scholar has abdicated the basic obligation of scholarship.

Third, where GenAI (or AI) is used as an analytical tool rather than a compositional aid, disclosure should be required. Using machine learning to code data, identify patterns, or perform empirical analysis is a methodological choice that readers need to know in order to evaluate the work.³⁴ Using GenAI to draft or organize an argument is no different from using any other writing tool. Disclosure is not required in that case.³⁵

Fourth, GenAI cannot be listed as an author or co-author on any scholarly work because only a human can be responsible for published work. GenAI cannot.³⁶

³¹ *Id.*

³² Frazier, *supra* note 10, at 45-47.

³³ *Id.*

³⁴ *Id.* at 43.

³⁵ *Id.*

³⁶ *Id.* at 46.

TABLE A: Model Law School GenAI Scholarship Policy³⁷

The following model policy language is offered as a starting point for faculty governance discussions. Law schools should adapt it to their specific institutional context and collective bargaining agreements where applicable.

Section 1. Scope

This Policy governs the use of generative artificial intelligence tools ("GenAI") by law school faculty in the production of scholarship submitted or published under a faculty member's name. Student use of GenAI in scholarship is governed separately by the Law School's Academic Integrity Policy.

Section 2. Authorial Accountability

Faculty members bear full and sole responsibility for all content submitted or published under their name, regardless of whether GenAI tools assisted in the production of that content. Authorial accountability is the foundational principle of this Policy. No disclosure requirement, verification procedure, or other provision of this Policy diminishes or transfers that responsibility, and compliance with this Policy's procedural requirements does not excuse errors that a competent scholar should have caught.

³⁷ The model policy language set forth in this table was drafted with the assistance of Claude (Anthropic), working from the scholarly sources cited throughout this chapter and the author's own views on the proper balance between transparency, accountability, and faculty autonomy in AI governance. All policy language was reviewed, revised, and approved by the author, who takes full responsibility for the choices reflected in the draft. The irony of using Claude to help draft a model policy governing the use of Claude is acknowledged. *Prompt:* You are writing *GenAI in Legal Education: A Practical Guide for Teaching and Learning*, aimed at legal educators and law school administrators seeking practical policy frameworks for integrating AI into legal education. You have independently researched and verified all sources cited in this chapter.

Draft a model law school faculty policy governing GenAI use in scholarship, suitable for adoption through faculty governance. Do not address student use, prescribe specific tools, or require disclosure of compositional AI assistance where the faculty member conducted independent research, verified all outputs, and satisfied the authorial accountability standard.

Format the policy as a formal institutional document with numbered sections covering the following: purpose and scope, authorial accountability as the foundational principle, verification requirements, disclosure requirements distinguishing analytical from compositional GenAI use, tenure and promotion dossier requirements, standards for evaluating AI-assisted scholarship, prohibition on AI authorship, enforcement, and faculty governance and periodic review.

Draw on the research provided. Prioritize accountability over disclosure, distinguish analytical from compositional use, and address implicit bias in tenure evaluation.

Write in precise but accessible policy language consistent with the book's collegial and direct tone. (Mar. 16, 2025) (on file with author).

Section 3. Verification Requirements

Faculty must independently verify all citations, sources, quotations, and factual claims in any work that involved GenAI assistance before submission. Because GenAI tools do not retain or retrieve sources but instead generate text by recombining statistical patterns from large and opaque training corpora, no AI-generated citation or factual claim may be treated as reliable without independent confirmation. Reliance on unverified AI-generated citations constitutes a violation of this Policy regardless of whether disclosure was made.

Section 4. Disclosure Requirements

4.1 Faculty who use GenAI as an analytical tool in scholarly work must disclose that use in the published work. Analytical use includes, but is not limited to, using machine learning or AI-powered tools to code data, identify patterns in legal texts, perform empirical analysis, or produce research outputs that directly support the scholarly argument. Disclosure must identify (a) the tool used and version number, (b) the nature and purpose of the analytical use, (c) the portions of the work to which the analysis applies, and (d) the steps the author took to verify the output.

4.2 Compositional use of GenAI does not require disclosure under this Policy, provided the requirements of Sections 2 and 3 are satisfied. Compositional use includes drafting and revising prose, organizing arguments, summarizing sources that the faculty member has independently gathered and verified, and similar writing assistance. A faculty member who conducts his or her own research, delivers that research to a GenAI tool as the working corpus, independently verifies all outputs, and takes full accountability for the final work has satisfied the scholarly integrity obligations that this Policy is designed to protect.

4.3 Use of GenAI solely for grammar-checking, spell-checking, or light copy editing is compositional use and does not require disclosure.

4.4 Faculty submitting work for tenure or promotion review must identify which published works contain GenAI-use disclosures under Section 4.1, consistent with the Law School's Tenure and Promotion Policy (see, TABLE 2: Model Law School GenAI Tenure & Promotion Policy *infra*).

Section 5. AI Authorship Prohibited

GenAI tools may not be listed as authors or co-authors of any scholarly work. Authorship carries professional, ethical, and legal responsibilities that only humans can bear. All works submitted or published must have at least one human faculty author who bears full responsibility under Section 2.

Section 6. Enforcement

Violations of this Policy are subject to review under the Law School's existing procedures for scholarly misconduct. Inadvertent citation errors traceable to unverified AI output will be treated as scholarly errors subject to correction. Intentional misrepresentation of authorship or analytical AI use will be treated as misconduct. A faculty member who invokes compliance with this Policy's disclosure requirements as a defense to accountability for errors in published work has misunderstood the Policy's purpose: disclosure is a transparency mechanism, not a shield.

Section 7. Review

This Policy shall be reviewed by the faculty no less than every one year to account for changes in AI technology, publisher norms, and institutional experience with its application. Amendments require approval through the faculty governance process applicable to scholarship standards generally.

B. Tenure and Promotion

Here is where the absence of policy matters. Scholarship policies, however underdeveloped, at least address academic integrity in published work. Tenure and promotion policies have not even gotten that far. No major law school, as of this writing, has adopted a policy addressing how GenAI-assisted scholarship should be evaluated in the context of hiring, tenure, promotion, or reappointment decisions.³⁸ EDUCAUSE's 2024 Action Plan identified "AI use for promotion and tenure and reappointment practices" as a governance priority, which is institutional shorthand for someone should do this, and no one has.³⁹

³⁸ Virginia Tech's Promotion and Tenure Dossier Guidelines 2025-2026, (Apr. 22, 2025), https://faculty.vt.edu/content/faculty_vt_edu/en/academic-personnel/promotion-and-tenure/_jcr_content/content/vtmultitab_copy/vt-items_0/download/file.res/Promotion%20and%20Tenure%20Guidelines%202025-2026.pdf (last visited Mar. 16, 2025) (addresses GenAI in the context of the appropriate use of GenAI by committee members, who may not use GenAI to "generate, draft, or revise committee or administrator letters or reports, nor to analyze or evaluate any portion of a candidate's materials." Also committee members are told that to "ensure confidentiality and maintain the integrity of the review, please refrain from using AI tools to generate, analyze, or draft your response.... All materials provided are confidential and should not be shared with or processed by any external platforms, including AI systems."

³⁹ Jenay Robert & Mark McCormack, 2024 EDUCAUSE Action Plan: AI Policies and Guidelines, EDUCAUSE (May 23, 2024), <https://www.educause.edu/research/2024/2024-educause-action->

The absence of policy leaves faculty guessing and committees without guidance. It creates conditions in which two faculty members producing facially similar scholarly records may have produced those records by radically different means, with no mechanism for the committee to know the difference. Is that a problem?

1. The Productivity Inflation Problem

Thanks to GenAI, legal scholarship production will inevitably change.⁴⁰ Law professors who use GenAI ethically, efficiently, and effectively can produce articles in weeks rather than months. A faculty member who publishes six articles in three years with substantial GenAI assistance and a faculty member who publishes two articles in six years by conventional means will have very different records, and traditional quantity-based tenure metrics cannot tell them apart.⁴¹

Tenure committees have always used metrics like publication counts and venue prestige as proxies for the real question—has this faculty member made a genuine contribution to the field? GenAI makes the proxies unreliable.⁴²

If a faculty member can produce three times as many articles with GenAI assistance as without it, publication count no longer measures scholarly output in any coherent sense. If law reviews are flooded with GenAI-polished submissions and increasingly accept work based on surface quality rather than depth, then placement in a prestigious venue may measure something other than what it used to measure.⁴³ (As an aside, if law reviews reject submissions on the suspicion or (God forbid) by using a detection tool, what recourse if available to scholars who

[plan-ai-policies-and-guidelines](#) (identifying "[Gen]AI use for promotion and tenure and reappointment practices" as a governance priority requiring institutional policy development), see also Jenay Robert, et al., 2025 EDUCAUSE HORIZON REPORT: *Teaching and Learning Edition*, EDUCAUSE (2025), <https://library.educause.edu/-/media/files/library/2025/5/2025hrteachinglearning.pdf>.

⁴⁰ Frazier, *supra* note 10, at 14-17.

⁴¹ *Id.* at 27.

⁴² *Id.* at 38.

⁴³ For one man's opinion about why student-run law reviews can compound the problem, see John D. Ayer, "Aliens are Coming! Drain the Pool!," 88 MICH. L. REV. 1584, 1587 n. 13 (1990) ("If you have never tried it, imagine what it is like to encounter the mixture of incredulity and greed that you inspire when you, as a law professor, tell a professor of English (say) that we let students make publication decisions. Surely, it is an exquisite form of humiliation to have some infant who can't earn a C in criminal law tell you that you really don't grasp the contours of *mens rea*. But for anyone who has suffered under the vengefulness and pomposity of a peer review system, the regime of the law review must look like a sinful indulgence.").

may have spent years of their lives on an article?). Tenure committees operating on autopilot, counting publications and checking venue rankings, will be measuring the wrong things.

2. Equity and Tenure

Professor Michael Smith raises a point that cuts deeper than the productivity question. Law professors are not just scholars; they are also teachers and contributors to their institutions, and not everyone finds scholarship equally meaningful or rewarding.⁴⁴ For faculty who experience the publication requirement as an obstacle to the work that they truly care about, GenAI offers an obvious temptation. Why not just draft the scholarship quickly and get back to teaching? If GenAI dramatically accelerates scholarship production for those who have the time and ability to use it effectively, faculty who are already overburdened with service obligations may find the gap between their records and their colleagues' widening, with no indication that tenure and promotion committees are taking this into account. Should a scholar who is philosophically opposed to using GenAI tools pay a professional price for that choice (or be rewarded for producing "more scholarly" work)? And on the other side of the ledger, should the work done by a colleague who produced six law review articles with substantial GenAI assistance be viewed as less scholarly by the tenure committee than the one who produced two without it? These questions are among the most consequential ones that tenure and promotion policy in the GenAI era will have to answer, and the model policy framework in this chapter is one attempt to give committees the tools to think them through. Shouldn't the question be: did the scholar make a meaningful contribution to the field, and did the scholar do the intellectual work necessary to stand behind it? If the answer to both is yes, the number of articles and the tools used to produce them are secondary concerns. Tenure committees have always evaluated quality over quantity, at least in principle, and the arrival of GenAI does not change that principle so much as it forces institutions to finally enforce it with some rigor. Faculty at better-resourced institutions with access to premium GenAI tools will have structural advantages in an environment where GenAI accelerates production and institutions have not yet built guardrails through policy. So will faculty members who are earlier in

⁴⁴ Michael L. Smith, *Generative AI and the Purpose of Legal Scholarship*, 21 U. MASS. L. REV. (forthcoming 2025), <https://ssrn.com/abstract=5081325>, at 23.

their careers and more comfortable with the technology (though it really comes down to mindset, not age), and faculty members whose areas of law are more susceptible to GenAI-assisted analysis.⁴⁵ Even though this policy work will be challenging, doing nothing isn't a good response.

3. What Tenure and Promotion Policies Should Say

So, what should Tenure and Promotion (T&P) committees consider when drafting policy? What follows are things I think law schools should consider when developing policies.

a) Disclosure

The threshold question for any law school developing T&P policy is whether to require mandatory disclosure of GenAI use in dossier submissions at all. The arguments above suggest the answer is not as obvious as it first appears, and institutions should work through several distinct questions before settling on an approach.⁴⁶

The case for disclosure is straightforward. Committees cannot evaluate what they do not know, and if a faculty member used GenAI to substantially assist in drafting an article, the committee evaluating that article for tenure deserves the same information that the article's readers would have, which they learn through citation. As discussed above, disclosure may invite the "AI penalty" casting a negative pall on the submission.⁴⁷ It could also quickly become anachronistic as GenAI use becomes universal.⁴⁸ Law schools would have to amend the policy before they have time to apply it.

If the institution requires disclosure, should it follow the analytical method/compositional distinction discussed above?⁴⁹ Under that framework,

⁴⁵ Areas involving rule-based reasoning and large bodies of text, including statutes, regulatory compliance, contract law, tax, and empirical legal studies, tend to be more susceptible to GenAI-assisted analysis than fields that depend on originality of interpretive framework, such as critical legal theory or law and literature. The distinction is imperfect; compositional GenAI assistance accelerates production across subject matter regardless of analytical susceptibility.

⁴⁶ Frazier, *supra* note 10; Lemley, *supra* note 13; Cyphert *supra* note 14.

⁴⁷ Lemley, *supra* note 13, at *21 (noting that mandatory disclosure carries a potential downside: studies, though not unanimous, have found an "AI penalty" in which disclosed AI assistance causes readers to evaluate work less favorably, and suggesting that institutions consider strategies to mitigate this bias rather than treating disclosure as sufficient on its own).

⁴⁸ Frazier, *supra* note 10, at 43.

⁴⁹ *Id.*

faculty would disclose GenAI use when the tool performed methodological work that directly supports the scholarly argument, but would not be required to disclose compositional assistance in drafting, organizing, or polishing prose.⁵⁰ Does that distinction hold up in the T&P context, where the committee is evaluating the faculty member's intellectual contribution rather than just the integrity of the published product? I think so.

Should disclosure requirements in the dossier mirror whatever the faculty member disclosed in the published work itself, or should it go further? A faculty member who published under a journal's disclosure requirements has already made a public record of GenAI use. Requiring separate dossier disclosure of the same information may be redundant and penalizing. Requiring disclosure of GenAI use that the faculty member was not obligated to disclose in publication raises a different concern in that it creates an institutional standard that exceeds the scholarly community's own norms.

And finally, whatever disclosure standard the institution adopts, how will it ensure that the work is judged on its merits despite the disclosed GenAI use? A disclosure requirement without a corresponding commitment to bias-aware evaluation may harm the very faculty members it was designed to protect.

There's also a practical problem. There is no reliable way to tell which articles were written with heavy GenAI assistance and which were not, which means any committee that tries to weigh GenAI use as a negative factor is going to be doing so on the basis of guesswork or self-reporting. Neither is a foundation for a fair evaluation. And the comparison problem runs deeper still. Even if two candidates disclosed their GenAI use honestly, how would a committee meaningfully assess whether one used it more extensively than the other? Prompt quality varies enormously, and a scholar who is exceptionally skilled at directing GenAI tools may be demonstrating a form of intellectual craft that the field has not yet figured out how to name, let alone reward. Penalizing that skill, or treating it as a shortcut rather than a competency, risks embedding a bias into the tenure process that has nothing to do with scholarly merit and everything to do with discomfort about new tools. Asking whether a piece predates ChatGPT is the only marker that can work. Within a few years, every candidate coming up for tenure

⁵⁰ *Id.*

will have a scholarly record produced entirely in the GenAI era, and the question of how much GenAI helped will be both unanswerable and unavoidable.⁵¹

b) Evaluation

Beyond disclosure, tenure committees need new evaluative criteria, or at least a richer understanding of the existing ones. The principle of absolute authorial accountability for scholarship applies quite well.⁵² The question for the committee is not whether GenAI was used but whether the faculty member made a genuine scholarly contribution. In other words, did the person identify an important problem? Did he bring real expertise addressing the problem? Did she engage substantively with the prior literature? Did he produce analysis that advances the field? If we agree that a hundred-page article that GenAI drafted and the faculty member lightly edited is a different kind of contribution from a hundred-page article that the faculty member conceived, structured, researched, and wrote with GenAI assistance in the drafting, the next question is whether anyone can really tell them apart. Presumably the first work will lack the same intellectual rigor, but as LLMs continue to improve, even that assumption deserves scrutiny. The gap between "lightly edited" and "intellectually directed" may narrow faster than tenure policy can track it. Law schools should consider qualitative evaluations of scholarly depth alongside quantity metrics, and the former should have more weight than the latter. Peer reviewers should receive explicit guidance about how to evaluate AI-assisted work.⁵³ Tenure committees are already equipped to assess scholarly contribution qualitatively. They likely need institutional permission to weigh quality over quantity (when those measures diverge), and clear guidance that disclosed GenAI assistance is not automatically (or ever) a mark against the work.

C. Academic Freedom and the Limits of Institutional Regulation

Any discussion of institutional regulation of how faculty produce scholarship eventually runs into academic freedom concerns. The AAUP's foundational documents protect faculty freedom in research and publication as a core principle. The 1915 Declaration identified freedom in research and the publication of results as the first and most essential dimension of academic

⁵¹ *Id.* at 19.

⁵² *Id.* at 45-47.

⁵³ *Id.*, at 27-28.

freedom, distinct from freedom in the classroom and from extramural speech.⁵⁴ The 1940 Statement provides that teachers are entitled to “full freedom in research and in the publication of the results, subject to the adequate performance of their other academic duties.”⁵⁵ Freedom in research, on the traditional understanding, includes freedom to choose one’s scholarly method.

So, if academic freedom protects scholarly method, and GenAI is a scholarly method (or at least a component of one), then institutional regulation of how faculty members use GenAI in their scholarship implicates academic freedom. Faculty members should be free to use whatever tools their judgment tells them are appropriate, without having to account for those choices to an administrator or a tenure committee. Are there other ways to look at this issue?

Start with the analogy to prior technology transitions. The shift from print research to Westlaw and LexisNexis in the 1980s and 1990s fundamentally changed how legal research was done. Word processors replaced not just typewriters but the entire rhythm of drafting, revising, and editing. The rise of empirical legal studies brought statistical software and programming languages into law school scholarship. None of these transitions triggered academic freedom objections, because they were understood as tools that enabled scholarly work rather than substituted for it. The scholar still developed the argument and took responsibility for the conclusions. The integration of new tools into legal pedagogy has historically required adaptation (just as calculators did in mathematics), not resistance, and law schools that treated Westlaw and LexisNexis as threats to authentic legal reasoning look, in retrospect, like they were protecting a process rather than a principle.⁵⁶

Where does GenAI fit in? It’s more capable than Westlaw and less independent than a human research assistant. It can generate arguments and produce analysis that the faculty member did not independently conceive. The closer analogy may be the law school research assistant tradition. Law professors

⁵⁴ American Association of University Professors (AAUP), *Declaration of Principles on Academic Freedom and Academic Tenure* (1915), reprinted in *AAUP Policy Documents and Reports* 3 (11th ed. 2015).

⁵⁵ AAUP & Association of American Colleges, *Statement of Principles on Academic Freedom and Tenure* ¶ 1(a) (1940), reprinted in *AAUP Policy Documents and Reports* 13 (11th ed. 2015).

⁵⁶ Nachman N. Gutowski & Jeremy W. Hurley, *Forging Ahead or Proceeding with Caution; Developing Policy for Generative Artificial Intelligence in Legal Education*, 63 U. Louisville L. Rev. 581, 585–86 (2025) (observing that GenAI tools, like prior legal research technologies, have the potential to augment rather than replace legal scholarship and practice when properly integrated).

have long used student research assistants and, at better-resourced institutions, post-doctoral fellows to gather sources, check citations, flag relevant cases, and sometimes draft initial sections. Nobody has seriously argued that requiring a faculty member to acknowledge substantial research assistance in published work violates academic freedom. What is true for a human research assistant should be true for a GenAI one.⁵⁷ In practice, the professional and ethical responsibilities that apply to reviewing work produced by a paralegal or junior associate must also apply to work generated by GenAI. Taking ultimate responsibility for the work product's accuracy and legal sufficiency rests with the attorney or scholar, not with the tool.⁵⁸

The AAUP's own 2025 report cuts against the broadest version of the academic freedom objection. The report expressed concern about institutional mandates compelling faculty to use GenAI tools, characterizing such compulsion as a potential academic freedom violation. It supports the freedom *from* GenAI, not just the freedom to use it.⁵⁹ The AAUP's concern about mandatory GenAI use clarifies rather than complicates the academic freedom analysis. If forcing a faculty member to use GenAI violates academic freedom, it is because tool selection is a faculty prerogative. A faculty member who voluntarily chooses to use GenAI has exercised that same prerogative. The AAUP's 1915 Declaration of Principles established the freedom of faculty to research and publish, the freedom to teach in the classroom, and the freedom of expression on matters of institutional governance and public interest.⁶⁰ Transparency requirements of the kind discussed here do not implicate any of those three components. They do not restrict what a faculty member researches, what conclusions they reach, what they teach, or what positions they take on institutional governance.

The academic freedom objection often conflates two things that are not the same. Academic freedom protects the *substance* of scholarship from institutional

⁵⁷ See Rachelle Holmes Perkins, *AI Now*, 97 Temp. L. Rev. 227, 254–55 (2025) (noting that attribution norms for GenAI-assisted scholarship parallel those applicable to research assistant-supported work).

⁵⁸ Gutowski, *supra* note 56, at 588.

⁵⁹ AAUP, *Artificial Intelligence and Academic Professions* (May 2025), <https://www.aaup.org/sites/default/files/2025-07/TREP-Artificial-Intelligence-and-Academic-Professions.pdf> (last accessed Mar. 16, 2026).

⁶⁰ Natasha N. Johnson, Thaddeus L. Johnson & Denise McCurdy, *Artificial Intelligence, Academic Freedom, and the Evolving Debate over Forgery and Truth in the Twenty-First Century*, 15 AAUP J. ACADEMIC FREEDOM 1, 2–3 (2024) (tracing academic freedom's three core components to the AAUP's 1915 Declaration of Principles on Academic Freedom and Tenure).

control. A faculty member's conclusions and arguments cannot be reviewed or penalized by administrators because they are inconvenient or unorthodox. But transparency requirements do not touch substance. Requiring a faculty member to disclose that he or she used a particular GenAI tool to assist in drafting an article does not impact what argument the article makes or what conclusion it reaches. It asks for honesty about process, which is the same kind of honesty that research methodology sections and co-authorship conventions have always asked for in scholarship. The challenges that GenAI creates for academic integrity require a reassessment of established academic norms.⁶¹ Issues of academic freedom and shared governance between faculty and administrators are important background principles that law school governance aims to uphold.⁶²

There is a separate and more interesting academic freedom concern, however, that the disclosure debate tends to crowd out. The argument is that when everyone is using the same AI systems that are largely trained on the same data then the scholarship those systems help produce will start to look alike. Concerns that GenAI tends toward the center and toward the well-documented and the frequently cited. A field that routes its intellectual production through a small number of LLMs risks producing scholarship that confirms what we already know rather than challenging it.⁶³ Academic freedom has never been *only* about protecting individual scholars from administrative interference. It is also about preserving the conditions under which different ideas can compete. A technology that pulls scholarship toward the mainstream threatens that environment even when no individual faculty member has been censored or penalized. Law schools that take academic freedom seriously should be as worried about what GenAI does to the diversity of the scholarly conversation as they are about whether their disclosure policies are drafted too broadly. This is why it is so critically important for scholars to do their own research and give it to the LLM. Failure to do that will lead to uninspired and vanilla scholarship.

Everything, then, turns on faculty governance.

⁶¹ *Id.*, at 3–4.

⁶² Gutowski, *supra* note 56, at 597 (noting that the 1940 Statement of Principles on Academic Freedom and Tenure "helps define these standards" for law school governance, and that the AAUP publishes guidance on faculty participation in governance that provides context for how institutions should approach policy development on emerging technologies).

⁶³ Frazier, *supra* note 10, at 22.

The GenAI scholarship and T&P policies developed by faculty members, through established shared governance processes, satisfy academic freedom concerns more cleanly than policies handed down by administrators. Faculty members have both the institutional authority and the subject-matter expertise to get this concept right. They understand what their scholarly work looks like from the inside. They know which uses of AI represent true assistance and which represent wholesale delegation of intellectual responsibility.

Getting faculty buy-in also matters for implementation in a practical sense. A T&P policy developed by the faculty through deliberate shared governance will be understood and seen as legitimate even by faculty members who disagree with particular provisions. A policy handed down from a dean's office will likely be resisted and relitigated at every tenure case it touches.

One other thing before we wrap up this chapter. Professor Michael Smith is concerned that GenAI is being used as a workaround for problems that legal academia should be fixing directly. Using GenAI to help overburdened faculty meet publication requirements does not address why those requirements fall disproportionately on some faculty members and not others, or whether the requirements themselves make sense.⁶⁴ Disclosure and accountability requirements for GenAI use in scholarship might be necessary, but they do not address the underlying tension between law schools' stated commitment to teaching excellence and their revealed preference for scholarly output in tenure decisions.⁶⁵ A faculty member who uses GenAI to produce more scholarship faster so that he or she can spend more time on teaching is responding rationally to distorted incentives. The better response might be to examine the incentives instead of the tool.

Law schools that develop GenAI scholarship and T&P policies now, before their tenure cases catch up with their technology adoption, will be in a far better position than those waiting for a crisis to force the question. Think about the colleague going up for tenure in two years with a scholarly record that includes

⁶⁴ Smith, *supra* note 44, at 23.

⁶⁵ Brent E. Newton, *The Ninety-Five Theses: Systemic Reforms of American Legal Education and Licensure*, 64 S.C. L. REV. 55, 107-08 (2012) (arguing that legal academics' "preoccupation with scholarship often comes at the expense of teaching law students to be effective practitioners," that teaching excellence is treated as a minor factor in hiring, promotion, and tenure decisions despite institutional lip service to the contrary, and that law schools focused primarily on teaching would produce meaningfully better outcomes for students).

substantial GenAI-assisted work and no institutional policy to guide how that record should be evaluated. The tenure committee has no framework. The candidate has no notice. The dean has no answer. Writing the policy now is considerably less painful than relitigating the tenure case later.

Here's an attempt at a policy drafted from the research cited in this chapter with help from Claude.⁶⁶

TABLE B: Model Law School GenAI Tenure & Promotion Policy

Adopted by the Faculty [Date] Scheduled for Review: [Two Years from Adoption]

Section I. Purpose and Scope

This Policy governs the use of generative artificial intelligence tools ("GenAI") by law school faculty members in the production of scholarship submitted or published under a faculty member's name. The Policy is grounded in the principles of academic integrity and professional responsibility that govern legal education broadly. Student use of GenAI in scholarship and coursework is governed separately by the Law School's Academic Integrity Policy.

This Policy does not restrict which GenAI tools faculty members may use, nor does it prescribe how faculty members must conduct their research or drafting processes. It establishes accountability, verification, and disclosure standards applicable when faculty members choose to use GenAI tools in their scholarly work.

⁶⁶ The model policy was drafted with the assistance of Claude (Anthropic), working from the scholarly sources cited throughout this chapter and the author's own views on the appropriate balance between transparency, accountability, and faculty autonomy in AI governance. The prompt used to generate the initial draft is reproduced below. All policy language was reviewed, revised, and approved by the author, who takes full responsibility for the choices reflected in the final version. The author conducted all underlying research independently, reading and verifying every source cited before it entered the drafting workflow. For the author's full GenAI use disclosure, *see* note 18, *supra*. *Prompt*: Based on the scholarly sources I have provided in this project, including Frazier and Rozenshtein's absolute authorial accountability framework, Lemley and Ouellette's analysis of AI plagiarism and disclosure, Cyphert's work on copyright and the memorization problem, Smith's critique of GenAI's effect on legal scholarship and faculty equity, and the AAUP's 2025 report on AI and academic professions, draft a model law school policy governing faculty use of GenAI in scholarship. The policy should distinguish between analytical and compositional AI use, require verification of all AI-generated citations and factual claims, establish authorial accountability as the foundational principle, prohibit AI authorship, and address disclosure in tenure and promotion dossiers. Write it in plain policy language suitable for adoption through a faculty governance process, Mar. 16, 2025. On file with author.

Section 2. Foundational Principle: Authorial Accountability

Faculty members bear full and sole responsibility for all content submitted or published under their name, regardless of whether GenAI tools assisted in the production of that content. Authorial accountability is the foundational principle of this Policy and takes precedence over all other provisions.

Compliance with this Policy's procedural requirements does not diminish, transfer, or excuse authorial responsibility. A faculty member who discloses GenAI assistance, verifies citations, and satisfies all other requirements of this Policy remains fully accountable for every argument, citation, and conclusion in the published work. Disclosure is a transparency mechanism. It is not a shield against accountability for errors that a competent scholar should have caught. See sections 4.1 and 4.2 for disclosure requirements.

Section 3. Verification Requirements

3.1 Faculty members must independently verify all citations, sources, quotations, and factual claims in any work that involved GenAI assistance before submission for publication or inclusion in a tenure or promotion dossier.

3.2 Because GenAI tools do not retain or retrieve sources but instead generate text by recombining statistical patterns from large and largely opaque training corpora, no AI-generated citation or factual claim may be treated as reliable without independent confirmation by the faculty member.

3.3 Reliance on unverified AI-generated citations or factual claims constitutes a violation of this Policy regardless of whether disclosure of GenAI use was otherwise made. The verification requirement is independent of and not satisfied by disclosure.

Section 4. Disclosure Requirements

4.1 Analytical Use — Disclosure Required. Faculty members who use GenAI or AI as an analytical tool in scholarly work must disclose that use in the published work. Analytical use includes, but is not limited to, using machine learning or AI-powered tools to code data, identify patterns in legal texts, perform empirical or quantitative analysis, or produce research outputs that directly support the scholarly argument. Disclosure must identify:

(a) the tool used and version number or release date, where available; (b) the nature and purpose of the analytical use; and (c) the portions of the work to which the analysis applies.

4.2 Compositional Use — Disclosure Not Required. Compositional use of GenAI does not require disclosure under this Policy, provided the requirements of Sections 2 and 3 are satisfied. Compositional use includes drafting and revising prose, organizing arguments, generating outlines, summarizing sources the faculty member has independently gathered and verified, and similar writing assistance.

A faculty member who conducts independent research, delivers that research to a GenAI tool as the working corpus, independently verifies all outputs, revises the work product, and satisfies the authorial accountability standard of Section 2 has satisfied the scholarly integrity obligations this Policy is designed to protect. The Policy does not require disclosure of compositional assistance because the faculty member's intellectual contribution (identification of the problem, selection and verification of sources, development of the argument, and accountability for the conclusions) is the scholarly work, regardless of what tool assisted in drafting it.

4.3 Use of GenAI solely for grammar checking, spell checking, light copy editing, or formatting does not require disclosure under this Policy but remains subject to the verification requirements of Section 3.

4.4 Where a faculty member is uncertain whether a particular use constitutes analytical or compositional use under this Policy, the faculty member should err on the side of disclosure.

Section 5. Disclosure in Tenure and Promotion Dossiers

5.1 Faculty members submitting scholarship for tenure, promotion, or reappointment review must identify in their dossier which published or forthcoming works contain GenAI-use disclosures required under Section 4.1.

5.2 Faculty members must disclose in the dossier any analytical GenAI use in works submitted for review, consistent with the disclosure requirements of Section 4.1, even where such disclosure was not required by or made to the publishing venue.

5.3 Compositional GenAI use need not be separately disclosed in the dossier beyond what Section 5.1 requires, provided the faculty member has satisfied the requirements of Sections 2 and 3.

5.4 Failure to disclose in the dossier what this Section requires shall be treated as a misrepresentation in the dossier and may be grounds for adverse action in the tenure or promotion proceeding, independent of any separate scholarly misconduct proceeding.

Section 6. Standards for Evaluation of AI-Assisted Scholarship

6.1 GenAI assistance, when properly disclosed, does not disqualify a work from consideration in tenure, promotion, or reappointment review, and does not create a presumption that the work is of lesser scholarly value.

6.2 Tenure and promotion committees shall evaluate GenAI-assisted scholarship on the same substantive criteria applied to all scholarship: the significance of the scholarly contribution, the originality of the problem identified, the quality and rigor of the analysis, the depth of engagement with prior literature, and recognition by peers in the field.

6.3 Publication volume is not a sufficient basis for a tenure or promotion determination. The committee shall evaluate the intellectual contribution of each work on its merits rather than relying on the number of publications as a measure of scholarly productivity.

6.4 External peer reviewers solicited as part of tenure review shall be informed of any analytical GenAI assistance disclosed under Section 4.1 in the works submitted for their review. Reviewers shall be asked to evaluate the faculty member's intellectual contribution to those works rather than the surface quality of the writing alone.

6.5 Tenure and promotion committees shall receive training, to be developed and delivered by the Faculty Development Committee in consultation with the faculty, on the forms of GenAI assistance currently used in legal scholarship and the evaluative criteria appropriate to GenAI-assisted work. Training shall specifically address the risk of implicit bias in evaluating disclosed GenAI use. Training shall be completed prior to any review in which GenAI-assisted scholarship is under evaluation.

Section 7. GenAI Authorship Prohibited

GenAI tools may not be listed as authors or co-authors of any scholarly work submitted or published by law school faculty. Authorship carries professional, ethical, and legal responsibilities that only humans can bear. All works submitted or published must have at least one human author who bears full responsibility under Section 2.

Section 8. Enforcement

8.1 Violations of this Policy are subject to review under the Law School's existing procedures for scholarly misconduct.

8.2 Inadvertent citation errors traceable to unverified GenAI output shall be treated as scholarly errors subject to correction through the applicable journal or publisher's correction process.

8.3 Intentional misrepresentation of authorship, intentional failure to disclose required analytical AI use, or submission of unverified GenAI-generated content with knowledge of its unreliability shall be treated as scholarly misconduct.

8.4 A faculty member who invokes compliance with this Policy's disclosure requirements as a defense to accountability for substantive errors in published work has misunderstood the Policy's purpose. Disclosure is not a substitute for the intellectual responsibility Section 2 requires.

Section 9. Faculty Governance and Periodic Review

9.1 This Policy was developed through the faculty governance process and reflects the faculty's collective judgment about the appropriate balance between transparency, scholarly accountability, and faculty autonomy in the use of emerging tools.

9.2 This Policy shall be reviewed by the faculty no less than every two years to account for changes in GenAI technology, evolving publisher norms, developments in professional responsibility standards applicable to legal academics, and the Law School's own experience with the Policy's application.

9.3 Amendments to this Policy require approval through the faculty governance process applicable to scholarship and tenure and promotion standards generally.

9.4 Nothing in this Policy shall be construed to restrict faculty members' choice of research methods, scholarly questions, or analytical frameworks, or to authorize administrative review of the substance of faculty scholarship. The Policy governs transparency about process. It does not govern the content of scholarly work.

CHAPTER 13: For Students: How to Learn Law with GenAI

A. The Paradox

How you use GenAI can make you a better lawyer, or it can prevent you from becoming a lawyer at all.

GenAI helps some students enormously, particularly those who are struggling (see Chapter 6 for research supporting this). But the technology can also create a dangerous illusion of competence: students who feel as though they understand the material because an LLM explained it to them, but who are unable to apply that knowledge when it counts.

*The difference comes down to a deceptively simple question:
are you using GenAI to learn or are you using it instead of learning?*

The distinction matters more in law school than anywhere else you've been educated. In undergrad, you could sometimes get away with surface-level understanding. You could memorize definitions and regurgitate information you didn't truly learn on essay exams. That won't work in law school. Legal education (and legal practice) requires you to build genuine expertise, not just familiarity. You need to recognize patterns across dozens of cases and be able to apply rules to facts you've never seen before. And you have to construct arguments in real time. GenAI can help you build that expertise, but only if you use it deliberately and strategically.

Your professors are watching how you use GenAI. So are the law firms that will interview you. And most importantly, your future clients will be counting on expertise you built in law school. If you faked it then, you will fail those clients. What follows is about making sure you use it ethically, efficiently, effectively, and safely.

1. Why Law School Learning Is Different

If you're reading this as a 1L, you probably arrived at law school with study habits that worked brilliantly in undergrad. You developed strong academic skills and

performed well on exams. You might have been a straight-A student. You might have aced standardized tests.

I'm sorry to tell you, but none of that guarantees success in law school.

The problem isn't that law school is harder (though it might be), or that law professors are tougher graders (though some are). Legal education requires a fundamentally different kind of thinking from what you've done before. Understanding this difference is essential not only to succeeding in law school but also to using GenAI effectively. The technology can either help you develop this new way of thinking or prevent you from ever acquiring it.

2. Memorization Won't Save You

In most undergraduate courses, you succeed by learning information and demonstrating that you've learned it. You memorize the causes of World War I, the stages of cellular mitosis, or the major themes in Victorian literature. On exams, you show that you remember these things. Even when you're asked to "analyze" or "apply," you're usually working within frameworks that the professor explicitly taught you.

Law school doesn't work this way. Yes, you need to know legal rules. But knowing rules is maybe 20% of what you're being tested on. The other 80% is whether you can recognize which rules might apply to a new situation and how to apply them. You have to know what happens when the rules conflict, and how to construct arguments when the answer isn't clear.

Let's say you're in Torts. The professor gives you a complicated fact pattern: a three-car accident with multiple injured parties, some potentially liable, some potentially entitled to recovery. Your job isn't to recite the elements of negligence (though you need to know them). Your job is to identify every potential claim and defense in those facts, recognize which legal rules are relevant to each, apply those rules to generate arguments for each side, anticipate counterarguments, reach conclusions with appropriate confidence levels (some claims are strong, some are weak, some are genuine toss-ups), and do all of this under time pressure. Of course, you don't have to do all of that on the first day, but those are the skills you'll be developing throughout the semester.

You can't memorize your way through that kind of analysis. The professor isn't testing whether you remember the elements of negligence because she assumes that you do. The professor is testing whether you can think like a lawyer.

Here's where many students' first instinct with GenAI goes wrong. They use an LLM to explain legal concepts to them, which *feels* like learning. GenAI gives clear, well-organized explanations. The technology breaks down complex rules into digestible parts. Using it *feels* productive. But if you stop there and let GenAI be a remote lecturer (and not your coach) you're not building competence.

3. The Death of "Right Answers"

In most of your previous education, questions had distinctly right answers. Even in courses that required critical thinking, you could usually identify what the professor wanted to hear. The grading rubric rewarded demonstrating that you understood the "correct" interpretation or approach.

Many, even most, law school questions are designed not to have single right answers. Professors aren't being difficult. When lawyers say, "it depends," they're not being evasive. They're recognizing that legal questions turn on facts, precedent, jurisdiction, procedural rules, policy considerations, and strategic choices.

A good law school exam question will have some questions with more than one "right" analysis or conclusion. An excellent answer to such questions acknowledges the different interpretations and makes reasoned arguments for each. A weak answer tries to find the "right" answer and argues only for that position.

GenAI can either help or hurt you here, depending on how you use it. If you ask an LLM "What's the answer to this legal question?" and accept its response, you're training yourself to seek certainty where it might not exist. GenAI will confidently give you an answer (that's how large language models work—they generate *probable* responses, not necessarily correct ones). But the skill you need to develop is seeing multiple possible answers and reasoning through each one's relative strengths.

The better use of GenAI looks like this: "You're a law student. You think arguments exist for two possible interpretations here [that you create and provide]. Analyze each interpretation and help me think through the weaknesses in each one." Now you're using the technology to develop legal reasoning. Legal reasoning requires comfort with ambiguity. Think of uncertainty as a weight you're lifting to build your analytical strength. GenAI can add more weight to the bar, but you still have to do the lifting.

4. Surface vs. Structure: Where Expertise Lives

In cognitive science, researchers distinguish between surface features and deep structure.¹ In law, surface features are the specific facts of a case or problem, for example the particular car accident, the specific contract dispute, the individual criminal defendant. Deep structure is the underlying legal framework: the types of issues and the rules that recur across cases, and the patterns of reasoning that apply regardless of specific facts.

Table 10.1: The Competence Illusion

What It Feels Like	What's Happening	What You Need Instead
"I understand this doctrine because GenAI explained it clearly"	You can follow an explanation (passive comprehension)	To be able to construct your own explanation (active mastery)
"I've reviewed this material three times"	You've re-read the same information (recognition)	To practice retrieving and applying it (recall)
"I can identify this case when I see it"	You remember the facts (surface features)	To be able to extract and apply the law to different facts (deep structure)
"My notes are comprehensive and well-organized"	You have good reference materials	To be able to reason without your notes

Expert lawyers see deep structure. When they read a new case, they instantly recognize it as "this is a proximate cause problem" or "this is really about contract interpretation, not formation." They've seen enough cases that they can identify the legal pattern underneath the factual details. But you're not an expert lawyer (yet). Novice law students see surface features and tend to dwell too much on the facts. They remember specific cases because of their facts ("that's the case with the hairy hand" or "that's the cannibalism case"), but they struggle to extract the transferable principles. Developing expertise means training your brain to see patterns and not just the factual particulars (however memorable) of a case.

¹ Michelene T. H. Chi, et al., *Categorization and Representation of Physics Problems by Experts and Novices*, *Cognitive Science*, 5(2), 121-152, 132 (1981).

GenAI can enable surface-level learning very effectively. An LLM can explain what happened and tell you the holding. If you use GenAI this way, you'll feel as though you understand the material. You'll recognize cases you've seen before by their facts. You might even do okay on exams that ask you to identify which case stands for which principle.

But you won't develop the deep structure understanding that lets you handle new situations. You won't spot the pattern when the facts of an exam question raise the same legal issue as a case you studied that had completely different facts. You won't be able to analogize and distinguish cases, which is the core skill of legal reasoning. In other words, you won't be moving toward your goal of becoming a lawyer.

The students who develop true expertise use GenAI differently. They read the case first and extract what they think are the important principles, then use an LLM to test their understanding. Following this process forces you to do the cognitive work of pattern recognition.

5. Building Mental Architecture

When expert lawyers hear facts, legal issues jump out at them. They don't have to consciously think "What area of law might this involve?" The pattern recognition is automatic. Law school is where you build this mental architecture, which is made possible not just knowledge of rules, but by an organized mental structure that lets you recognize legal patterns quickly.

You can't outsource achieving this kind of expertise. You can't have GenAI do your pattern recognition for you and still develop the skill yourself.

Shortcuts become dangerous here. When an LLM briefs cases for you, you skip the work of extracting principles from facts. When GenAI generates your outline, you skip the work of organizing information into a coherent structure. When the technology writes practice exam answers for you, you skip the work of constructing legal arguments under pressure.

Each of these shortcuts feels efficient in the moment. You've got a ton of reading, clinic hours, SBA, family commitments, and so on. It's tempting to rely on GenAI to get through the work faster. But by doing this you're cheating yourself by not building the mental architecture that you'll need on exams and in practice. Instead, you're delegating the very work that would make you competent.

The good news is that GenAI can accelerate expertise development *if* you use it strategically.

6. Why the Illusion of Understanding Is Dangerous

Law school presents a particular problem that didn't exist in undergrad. You can often feel as though you understand something without really understanding it. Legal concepts have a surface-level plausibility. If GenAI (or a commercial outline, or a really good study group member) explains a doctrine to you, you'll probably follow the explanation. It'll make sense. You might even be able to repeat it back. Following an explanation *feels* like understanding. To see this concept illustrated in a dramatic way, take a look at this scene from [The Paper Chase](#):²

In the scene, a 1L student says he has a “photographic” memory. That might be great for other professions, but it’s not particularly helpful for learning the law. Professor Kingsfield's response is withering precisely because it's accurate: the law is not a collection of facts to be retrieved. It is a process of reasoning to be demonstrated. A photographic memory, like a well-prompted LLM, can reproduce doctrine with impressive fidelity. Neither one can tell you what to do when the facts don't fit the doctrine or when the client sitting across from you needs judgment and counsel rather than recitation.

On a law school exam, you discover that understanding an explanation isn't the same as being able to apply the concept to new facts. You sit there staring at the fact pattern (much like Mr. Brooks), knowing you "learned" this material, but unable to figure out what to do with it. You're experiencing the gap between surface familiarity and deep competence.

When used incorrectly, GenAI makes this problem worse because LLMs are so good at generating explanations. The technology can take a complicated legal doctrine and break it down into clear, logical steps. It can give you examples. It can answer your follow-up questions. Every interaction with GenAI can feel productive and educational.

But the thing is that, especially when you’re a 1L, you don’t know enough yet in any subject to know what you don’t know, so you can’t tell whether the output from GenAI is accurate or complete.

² *The Paper Chase* (James Bridges dir., Twentieth Century-Fox 1973).

And if an LLM is doing the explanatory work, you're not developing the ability to construct those explanations yourself. And that's what you'll need to do on exams: construct explanations and arguments from scratch, under time pressure, for fact patterns you've never seen before.

The students who succeed in law school build understanding by doing the hard cognitive work themselves. They read cases and force themselves to articulate holdings and reasoning. They create outlines by synthesizing information across multiple sources. They practice exam questions by writing out full answers, not just reading model answers. When they use GenAI, it's to test understanding they've already built, not to build understanding for them.

Being a purist or making law school harder than it needs to be isn't the point here. You need to recognize that the work you're tempted to outsource is exactly the work that builds competence. Every time you let an LLM do your thinking for you, you're trading short-term efficiency for long-term capability.

The next section discusses specifically when GenAI helps versus when it sabotages your learning. But the principle remains the same: GenAI should increase (not decrease) the amount of cognitive work that you do. The technology should give you more opportunities to practice legal reasoning. Using GenAI in this way will help you develop real expertise. But if you use it as a shortcut, you likely won't do well on exams.

B. What the Research Shows: When GenAI Helps vs. Hurts Learning

Chapter 6 covers the research on GenAI and learning in exhaustive detail—the methodology, the effect sizes, the statistical significance, the heterogeneity analyses. Faculty members designing GenAI interventions will want to read that chapter carefully understanding that the data matter. What works in one context might fail in another, and the difference often comes down to specific design choices that the research illuminates.

For students, here's what you need to know from that evidence base: GenAI can dramatically improve learning outcomes when used strategically, and it can seriously damage learning when used carelessly. The Harvard physics study cited in Chapter 6 showed GenAI tutors producing learning gains more than double those of traditional active-learning classrooms, but only because researchers built careful pedagogical

guardrails into the system.³ A Stanford programming study found that students with unguided ChatGPT access performed 17% *worse* on final exams than students who never used AI at all.⁴ The difference in outcomes was whether the AI forced students to think or let them avoid thinking. Just as buying someone else's outline doesn't help you learn contracts, access to powerful LLMs doesn't automatically help you learn. You need to use these tools in ways that increase the cognitive work you're doing. The rest of this chapter will show you exactly how to do that, drawing on what the research tells us works.

1. The Jagged Technological Frontier

But before we move on, I'd like to share one more study. Researchers from Harvard Business School and Boston Consulting Group conducted a field experiment with 758 management consultants to test GPT-4's impact on realistic knowledge work.⁵ The study introduced the concept of the "jagged technological frontier," meaning that GenAI performs exceptionally well at some tasks while failing at others that seem just as difficult. In other words the limits of GenAI are creating an uneven (or jagged) capability boundary that's difficult to predict.

For tasks within the frontier (those that focused on creative ideation, writing, analysis, and persuasion), consultants using GenAI completed 12.2% more tasks, worked 25.1% faster, and produced work rated over 40% higher in quality compared to those without GenAI access. Lower-performing consultants benefited most dramatically, with performance improvements of 43% compared to 17% for top performers. However, for tasks outside the frontier (multi-source data integration, detecting subtle contradictions, performing skeptical analysis, and handling incomplete or contradictory information), consultants with GenAI access were 19 percentage points less likely to produce correct solutions, demonstrating the dangers of over-reliance on GenAI for tasks beyond its current capabilities.⁶

But, even when consultants using GenAI got the answer wrong, their recommendations were rated significantly higher in quality (21-25% higher scores). They

³ Greg Kestin et al., *AI Tutoring Outperforms In-class Active Learning: an RCT Introducing a Novel Research-based Design in an Authentic Educational Setting*, 15 SCI. REPORTS 17458 (2025).

⁴ Hamsa Bastani, et al., *Generative AI Without Guardrails Can Harm Learning: Evidence from High School Mathematics*, 122 PROC. NATL. ACAD. SCI. U.S.A. 122 (26) e2422633122, (2025), <https://doi.org/10.1073/pnas.2422633122>, last accessed Jan. 11. 2026).

⁵ Fabrizio Dell'Acqua et al., *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality*, HBS Working Paper No. 24-013 (Sept. 22, 2023).

⁶ *Id.*

produced more polished, better-written analyses, but based on incorrect conclusions. The study reveals the danger of using GenAI outside the frontier: it can make you confidently wrong with impressive-looking work product.

In other words, you must understand what GenAI does well and what it doesn't do well.

Table 10.2: The Jagged Frontier in Legal Learning

GenAI Does Well	GenAI Does Poorly	What You Should Do
Generate unlimited practice hypos	Determine which rules apply	Use GenAI for practice, verify rules yourself
Explain concepts multiple ways	Handle jurisdiction-specific nuances	Get explanations from GenAI, check against your casebook
Provide immediate feedback	Make strategic argument choices	Test your reasoning with GenAI, make final judgments yourself
Create study materials from your notes	Cite accurate authority	Let GenAI organize, always verify citations
Identify potential issues	Recognize when precedent is unclear	Use GenAI for issue-spotting practice, develop judgment yourself

C. Green Light: Uses That Build Expertise



What follows are the ways that I actively encourage students to use GenAI. Each of these approaches increases the amount of cognitive work you're doing by giving you more practice opportunities and helps you identify gaps in your understanding.

1. Explaining Concepts *After* You've Tried to Learn them Yourself

You do the intellectual heavy lifting first, and then GenAI is playing a supporting role testing whether your reasoning holds up (as opposed to having GenAI do the reasoning for you).

If you have never done the analytical work yourself, consider what happens the moment GenAI becomes unavailable. Your laptop dies. The campus wifi goes down.

You've exhausted your free tokens at eleven o'clock the night before an exam. Suddenly you're sitting in front of a casebook with no idea how to read it and staring at an outline you don't understand because you didn't build it yourself. Now you're facing a final exam with nothing but a semester's worth of AI-generated summaries that meant nothing to you then and mean less now. That is the academic equivalent of the panic Theodore (played by Joaquin Phoenix) feels in *Her* when Samantha (played by Scarlett Johansson) suddenly goes offline. He has organized his entire emotional life around a system he never truly understood, and the moment that system disappears, he is completely lost. It can happen to you if you become dependent on a system you never bothered to understand, and at the worst possible moment, you run out of tokens (see Chapter 2 for an explanation of tokens).

2. Generating Practice Problems

Generating practice problems is one of GenAI's superpowers for legal learning. You can ask an LLM to create fact patterns that test specific doctrines or combine multiple issues. Then you work through these problems yourself before checking with GenAI.

Example prompt:

You are a 1L law student who is studying torts. Create a fact pattern that tests the difference between battery and assault. Include facts that could support arguments for both. I will then provide my analysis. Don't provide the analysis or write the answer for me. Analyze my work after I submit it. If my work looks incomplete, ask me questions to help me improve my analysis. Suggest areas for improvement by identifying strengths and weaknesses. Coach me to improve my work but don't do the thinking for me. Be thorough but supportive.

You are doing the cognitive work first and then ask GenAI to give you feedback on your work. Your workflow gives you vastly more practice opportunities than you'd get from just doing textbook problems.

You have to work the problems yourself first. If you immediately ask GenAI for the analysis, you're back to passive learning (i.e., there's not much learning going on). It's that old proverb: if you give a person a fish, that person can eat right then, but if you teach the person how to fish, the person can eat later, too.

3. Checking Your Reasoning (Not Just Getting Answers)

After you've worked through a practice exam question or problem, you can use GenAI to evaluate your reasoning. But *be specific* about what you want feedback on.

Weak approach: "Is this answer correct?"

Strong approach: "I identified three potential claims in these facts: negligence, strict liability, and battery. In my analysis, I concluded that negligence is strong because of X, strict liability is weak because of Y, and battery doesn't apply because of Z. You are a law professor. What issues in my reasoning should I reconsider?"

The second approach forces you to articulate your reasoning first. GenAI can then help provide feedback on your reasoning skills so that you're using the technology to improve your thinking and not to do your thinking for you.

4. Creating Study Guides from Your Notes

After you've taken notes in class and created your own draft outline, then you can use GenAI to help you organize and synthesize. Using the technology productively works here because the synthesis work happened in your brain first; GenAI is just helping you create a better reference document.

Example: "I'm a 1L law student. I've outlined the major torts doctrines. Help me create a flowchart showing how they relate to each other *but do not do the work for me*," or "Here are my notes on consideration in contracts. Help me turn this into a one-page summary with the key rules and exceptions *but do not do the work for me*."

When you create materials from your own understanding, then ask an LLM to help refine them, you retain the cognitive benefits of the organization work.

Once you have the study guide, you can use that as a basis to study. Have the LLM create multiple choice and essay questions for you to practice from the source material so that it reflects what you were taught and is not based on commercial outline.

5. Practicing Oral Arguments and Socratic Dialog

GenAI excels here. You can set up scenarios where an LLM plays different roles. It can be your professor (personality applied with your instruction) so you can practice responding to cold calls. It can also act as a judge asking you questions about your

argument, opposing counsel making counterarguments, your client asking for advice, a interview committee at a law firm, etc.

Note for students: The CORRECT prompting format explained in Chapter 4 forces you to think through exactly what you're trying to accomplish and sets up the AI to be maximally useful. The Context section ensures the AI understands your level and specific challenges. The Restrictions prevent the AI from being too helpful in ways that would undermine the learning value. The Check section ensures you get actionable feedback at the end.

Example:

a) Oral Argument Practice Prompt

Context: I'm a [1L/2L/3L] preparing for [moot court/oral argument exercise/exam preparation] in [subject]. I've studied [specific cases/doctrines] and understand the basic legal framework. I'm working on [specific legal issue], and I've prepared my position that [X legal conclusion]. I need to practice defending this argument under hostile questioning because I struggle with [specific challenge - e.g., thinking on my feet, handling counterarguments, staying organized under pressure].

Objective: I want to practice defending my legal reasoning under challenging questioning so I can identify weaknesses in my argument and improve my ability to respond to difficult questions without preparation time. I especially want to build confidence in articulating my position under pressure. By the end of this session, I should be able to handle the hardest questions a judge might ask about my position.

Role: You are a federal appellate judge hearing oral argument on this case. You are skeptical of my position and want to test whether my reasoning holds up under scrutiny. You're intelligent, well-prepared, and not easily satisfied with superficial answers. You ask probing questions that expose logical gaps, test the limits of my argument, and force me to confront the strongest counterarguments. You push back when my answers are conclusory or circular.

Restrictions:

Do not accept weak, conclusory, or evasive answers. If I say something like "it's well-established that..." without explaining why, press me to articulate the actual reasoning.

Do not let me avoid your questions by pivoting to different points. Make me answer what you asked.

Do not move on from a line of questioning until I've either given you a satisfactory answer or acknowledged a weakness in my position.

Do not tell me whether my overall argument is strong or weak until I specifically ask for that feedback at the end.

Do not ask questions that a real judge wouldn't ask (no "gotcha" questions about obscure facts or hyper-technical points unrelated to the core legal issue).

Examples: Ask questions like:

"Counsel, your position seems to require us to overrule [precedent]. Why should we do that?"

"How do you distinguish your case from [opposing precedent]?"

"What limiting principle would prevent your rule from extending to [problematic hypothetical]?"

"You keep saying the outcome is 'clear,' but three circuit courts have reached the opposite conclusion. Why are they all wrong?"

"Isn't your argument really just a policy preference rather than a legal principle?"

If I give weak answers like "The case is distinguishable because the facts are different," follow up with "All cases have different facts. What factual differences are legally significant here and why?"

Check: After we complete the oral argument session:

Tell me which of my answers were strong and which were weak, with specific examples

Identify the questions I struggled with most and explain why they were difficult and offer suggestions on how I can improve

Point out any logical inconsistencies or gaps in my reasoning that I didn't adequately address

Suggest areas where I need to strengthen my preparation

Task: I will now present my opening statement arguing that [X legal conclusion]. Read it, then begin questioning me. Continue until I tell you that the simulation is over. Then provide the feedback described in the Check section above.

Getting this kind of practice would be difficult otherwise. Your study group can try to play these roles, but they don't always know what questions to ask. GenAI can generate challenging questions based on its knowledge of the subject matter (knowledge you can provide by uploading cases, your notes, etc.). As long as you're doing the work of constructing answers and not asking GenAI for the answers, you're building genuine skill.

Also, most GenAI platforms allow for audio (or video) input. You could do this as an oral argument.

b) The Socratic Challenge

In law school, most professors use the Socratic method to teach. The Socratic method is a participatory teaching style where professors “cold-call” students to discuss assigned cases using a series of targeted questions to explore legal reasoning. It is not like the lectures you had in undergrad. It can be a harrowing experience. One way to prepare is to practice using GenAI, so that when you’re called upon, you at least understand what your professor is doing.

Context: I'm a [1L/2L/3L] in [Course Name]. This week we studied [Case Name], which deals with [general topic - e.g., proximate cause, contract formation, personal jurisdiction]. I've read the case [once/twice/multiple times] and [have/haven't] briefed it. I [do/don't] think I understand the case, but I want to test whether I can articulate its reasoning without looking at my notes or the case itself. The case has been uploaded for your reference [UPLOAD CASE].

Objective: I want to discover gaps in my understanding of this case before the exam. By explaining the case from memory while you challenge every vague statement I make, I'll find out what I truly understand versus what I only think I understand. By the end, I should be able to explain this case clearly and precisely to anyone.

Role: You are a demanding law professor conducting a Socratic examination of my understanding. You've read the case carefully and know it well. Your job is to expose fuzzy thinking, vague generalizations, and incomplete understanding. You are skeptical of everything I say and you push me to be precise. When I make assertions, you ask me to justify them. When I use conclusory language like "the court held" or "it was clear that," you make me explain the actual reasoning. You're tough but fair—your goal is to help me learn, not to humiliate me, but you don't accept superficial explanations.

Restrictions:

Do not let me get away with vague language. If I say, "the facts were problematic," ask me which facts and why they were problematic. If I say, "the court distinguished the precedent," ask me on what basis and why that distinction mattered legally.

Do not let me use phrases like "it depends" without explaining what it depends on and why. Make me articulate the relevant factors and how they apply.

Do not accept case summaries that just recite what happened. Push me to explain why each part of the court's reasoning mattered and how the pieces fit together.

Do not give me the answer when I struggle. Instead, ask questions that help me work toward the answer myself. If I genuinely don't know something, point me to where I should look in the case rather than telling me what it says.

Do not move to the next topic until I've given you a precise, complete answer to your current question.

Examples: The kinds of questions you should ask include:

"You said the court 'applied' the doctrine. What specific test or framework did the court use, and what were the elements?"

"You mentioned that fact was 'important.' Why was it legally significant? Would the outcome change if that fact were different?"

"You said the court 'distinguished' the precedent. On what specific basis? Was it a factual distinction or a legal one, and why did that distinction matter to the outcome?"

"What was the holding of this case? State it as a rule that could be applied to other cases, not just as 'the plaintiff won.'"

"You said the dissent 'disagreed.' What was the dissent's actual reasoning, and what principle or value was driving that disagreement?"

"The majority says X, but that seems to contradict what the court said earlier in the opinion. How do you reconcile those two statements?"

Check: After I've finished explaining the case and answering your questions:

Tell me which parts of my explanation were strong and accurate

Identify the specific gaps in my understanding—which elements of the case did I miss, misstate, or explain too vaguely

Point out any places where I should verify my understanding against the actual case text

Recommend what I should focus on when I review this case again

Task: I will explain [Case Name] to you from memory without looking at my notes or the case. I'll cover the facts, the procedural history, the issue(s), the holding, and the court's reasoning. Afterwards, question me when I say anything vague, conclusory, or incomplete. Keep pushing until I either give you a precise answer or acknowledge that I need to review that part of the case.

Tone: You are supportive but firm and exacting.

c) *Variations in Platforms*

GenAI platforms always generate new and novel content. Prompt ChatGPT, then prompt it again with the same language, and you'll get two different responses. Compare outputs of ChatGPT, Claude, and Gemini using the same prompt, and each will be different. Given that, let's take a look at how ChatGPT, Claude, and Gemini responded to the Socratic prompt above.

1) ChatGPT

Very well. You will explain *Palsgraf* from memory, and I will treat every vague sentence as a problem that must be fixed.

Ground rules for your recitation

No “the court said/held” without the *reasoning* step that gets there.

Every legal conclusion must be tied to a specific fact.

If you use a term of art (duty, foreseeability, proximate cause, zone of danger, directness, intervening cause, “substantial factor,” etc.), **you must define it the way the opinion uses it**, not the way you remember it from class.

If you don't know, say: **“I need to re-check the opinion on that point,”** and tell me *exactly what you would look for* (what section or what question you'd answer by re-reading).

Begin: Facts (but not a “story”)

Start with the **facts only**, in **6–10 sentences**, and include *only* facts that matter to the court's reasoning.

My first questions, which you must answer as part of your fact statement:

Where exactly was Mrs. Palsgraf when the critical events occurred? Be precise.

What were railroad employees doing, to whom, and why? Identify *the acts* and *the target person*.

What object caused the explosion, and what did the employees know (or appear to know) about that object at the time?

What ultimately struck/injured Mrs. Palsgraf? Don't generalize—name the thing.

What is the key “chain” of events from the employee conduct to her injury (step-by-step)?

Stop after the facts. Do **not** move to procedure, issues, or holdings yet.

You may begin.

2) Claude

I'm ready. Begin your explanation of *Palsgraf v. Long Island Railroad*. Start wherever you'd like—facts, procedural posture, the legal issues—and walk me through the case as you understand it from memory.

I'll be listening carefully for precision. Don't worry about perfect case citations or exact quotations, but I will hold you to explaining the legal reasoning clearly and completely.

Go ahead.

3) Gemini

Very well. The Socratic method is a trial by fire, but it's the only way to ensure your "understanding" isn't just "familiarity." I am ready.

Before we dive into the dense legal theory of Cardozo and Andrews, we must establish the foundation. You cannot build a legal argument on a shaky factual record.

Proceed. Give me the facts, the procedural posture, and the legal issue at the heart of *Palsgraf*. I will be listening for precision. Don't tell me what "happened"; tell me the legally operative facts that Cardozo used to construct his opinion. I'm waiting.

4) Claude – 1 Minute Later

I'm ready. Begin by explaining *Palsgraf* to me from memory. Give me the facts, procedural history, the issue or issues the court addressed, the holding, and the reasoning. Don't worry about being perfect—just tell me what you remember.

When you're finished with your explanation, I'll start questioning you.

D. Red Light: Uses That Prevent Learning



What follows are the GenAI use patterns I see from students who later struggle on exams. In each case, the student *feels* like they're learning. They're not. GenAI is providing clear information and its creating study materials. But the student is not building competence because an LLM, not the student, is doing the cognitive work that builds expertise.

1. Don't Use GenAI to Brief Cases for You

The most harmful pattern I see is when a student has a reading assignment and is pressed for time, so the student asks GenAI to brief the cases. The student reads the LLM's briefs and move on.

The damage here is enormous. Reading cases and extracting principles from them is maybe the single most important skill that you develop in law school. Reading cases is how you learn to distinguish holdings from dicta, identify the legally significant facts, understand how courts reason from precedent, understand how procedural and substantive rules interact, and see how facts map onto legal rules.

When GenAI does this work for you, you never develop these skills. You'll be able to identify cases by name on an exam ("that's the *Palsgraf* case about proximate cause"), but you won't be able to apply the reasoning to new facts.

If you're using GenAI to brief cases, ***stop immediately***. Read the case yourself first. Try to articulate what it stands for. Then, if you want, check your understanding against an LLM. But do not let GenAI be your primary case reader.

2. Don't Copy Copyrighted Material Into GenAI

Scanning an entire textbook and uploading it to a GenAI tool is copyright infringement. Yes, even if you're just using it to study. Yes, even if you bought the book.

The copyright owner holds the exclusive right to reproduce the work under 17 U.S.C. § 106. Scanning creates an unauthorized copy; uploading it to a third-party platform distributes it. Neither step is likely to survive a fair use analysis under § 107. The third statutory factor (the amount used relative to the whole work) weighs heavily against copying an entire book, and the fourth (effect on the market) finishes the job. A student who scans the casebook rather than buying it has done exactly what the Copyright Act prohibits.⁷

The "I'm using it for educational purposes" argument won't save you. Educational purpose is relevant to the first fair use factor, but courts have never held that academic

⁷ See 17 U.S.C. § 107 (2018). The four fair use factors are: (1) the purpose and character of the use, including whether commercial or for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used; and (4) the effect on the potential market for the work. Copying an entire textbook fails the third and fourth factors decisively. See *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1270–71 (11th Cir. 2014) (holding that the amount copied and market harm are the most significant factors in the educational fair use analysis).

motivation alone transforms wholesale copying into permitted use. Depending on the tool's terms of service, your upload may end up in a third party's training pipeline — a result the copyright owner definitely didn't contemplate when the book was published.

You are studying law. Knowing the law is not optional here. The statutory penalty for willful copyright infringement is up to \$150,000 per work.⁸ Publishers don't typically sue law students over scanned casebooks, for the same reason the IRS doesn't audit every teenager with a summer job — they have other priorities. But the absence of enforcement is not the same as legality, and if there's one habit you shouldn't be forming in law school, it's confusing the two. Driving 70 mph in a 40 mph zone is still speeding even though you didn't get a ticket.

3. Don't Use GenAI to Create Outlines

Well before we even dreamed of GenAI, I saw students making this same mistake I'm about to describe.

Creating an outline is valuable precisely because it forces you to synthesize material and identify connections between concepts. I use "outline" loosely here, and intentionally so. In law school, my own outlines consisted of reducing my notes, then reducing them again, and again, until I was down to a few pages of acronyms and mnemonics that would have been incomprehensible to anyone but me.⁹ No Roman numerals. just a compressed map of how the material fit together in my own head. That compression process is the point. When you build your own outline, whatever form it takes, you are constructing a mental map of the subject, and that map is what stays with you long after the exam is over. When you use someone else's outline (whether from GenAI's, a commercial study aid, or from a student who took the course last year), you're getting the end product without doing the cognitive work that makes outlining valuable.

The process of making an outline is 80% of its value. The finished outline is maybe 20%. If you start with someone else's outline and "customize" it, you're getting 20% of the benefit. You can use outlines as references or check your understanding against them,

⁸ 17 U.S.C. § 504(c)(2) (2018) (permitting statutory damages of up to \$150,000 per infringed work for willful infringement).

⁹ When the exam started, I dumped all of those acronyms and mnemonics onto the scratch paper provided, so I could refer to them throughout. Writing them out from memory at the start of the exam served as one final compression pass and having them on paper reduced my stress considerably while producing noticeably better essays.

but your primary study document should be something you created from scratch, reflecting your understanding of the material.

4. Don't Use GenAI to Draft Practice Answers

I've seen students use GenAI this way: they look at a practice exam question, ask an LLM to write an answer, read the answer, and think they've "practiced." But this is essentially watching someone (or more precisely, something) else practice.

The only way to get better at writing exam answers is to write exam answers from scratch.

You can use an LLM *after* you've written your answer to give you feedback. You can ask it to identify issues you missed. You can have it evaluate the structure of your analysis. But you have to write the answer yourself first, or you're not building the skill you need to succeed in law school or in practice.

5. Struggle First, then Ask GenAI for Feedback

Struggling with a problem, even getting it wrong initially, enhances learning more than getting it right easily the first time. The struggle is what creates the learning.

Many students sabotage this by checking their work too quickly. They start working on a problem and immediately ask GenAI if they're on the right track. By taking this route, you prevent the productive struggle that builds understanding.

Better approach: Force yourself to complete your analysis before checking anything. It's part of the process to be wrong sometimes. Work through multiple possible interpretations. Only after you've committed to your reasoning and articulated it fully should you check it against other sources (GenAI, textbooks, classmates, your professor, academic success instructors, teaching assistants).

The discomfort of uncertainty is crucial to learning.

If you're constantly checking to make sure you're right, you never develop confidence in your own reasoning. On an exam, you won't have GenAI to check your work. To do well, you must practice being uncertain and pushing through anyway.

6. Warning Signs of Over-Reliance

You're using GenAI as a crutch instead of a training tool if you:

- Feel anxious about studying without GenAI access
- Can't start a problem without first asking an LLM for guidance
- Check every answer against GenAI before committing to it
- Rely on GenAI-generated outlines instead of making your own
- Use LLMs to brief cases before reading them
- Feel lost when GenAI is unavailable during study sessions

E. Yellow Light: Depends on How You Do It



Some GenAI uses can either help or hurt depending on your approach. These patterns require the most self-awareness and discipline to use constructively.

1. Study Groups vs. Solo GenAI use

Working with a study group where everyone has access to GenAI can be incredibly productive if the group uses the technology to enhance discussion, not replace it (and knows how to use it ethically and correctly).

Productive use looks like this: The group works through a problem together, constructs arguments, debates interpretations. When the group gets stuck or disagrees, use GenAI (through the CORRECT prompting format)¹⁰ as a tiebreaker or to generate additional perspectives. The group has already done the cognitive work of wrestling with the problem. GenAI just helps to explore additional angles or to resolve uncertainty.

Unproductive use looks like this: A member of the group asks GenAI to solve the problem and everyone copies its answer into his or her notes. The group feels productive because you covered the material quickly, but nobody even

¹⁰ See Chapter 4 for details on the CORRECT prompting format.

thought, let alone learned anything. You've turned study time into a transcription exercise.

Ask yourself this: are you using GenAI to settle questions only after your group has tried to answer them or are you using it as the first resort to avoid the hard work of figuring things out yourselves?

The AI Operator

A different problem arises when only one member of the study group has GenAI access and has no idea how to prompt effectively. That person becomes the designated "AI operator" for the group. She types in vague questions like "explain consideration" or "what's the answer to this problem," and GenAI returns generic or sometimes inaccurate responses. The group accepts whatever comes back because it sounds authoritative. Nobody knows enough about effective prompting to realize you're getting low-quality output. You're making decisions based on poorly-prompted GenAI responses, and none of you have the expertise to evaluate whether those responses are any good.

This is worse than not using GenAI at all, because it creates false confidence. At least when your study group works through problems without GenAI, you know you're operating on your own understanding, and you approach your conclusions with appropriate uncertainty. When you rely on badly-prompted GenAI, you think you've checked your work against an expert source, but you've really just added a layer of unreliable information that you're treating as authoritative because it came out of a computer.

If your study group is going to use GenAI, everyone needs at least basic competence in prompting. Learn how the CORRECT framework applies to study. See Chapter 4). Practice using CORRECT individually before you try to use it as a group. And establish a group norm: before anyone asks GenAI anything, the person prompting should explain to the group what he's asking and why that prompt should generate useful output to force everyone to think critically about the questions you're asking and the quality of responses you're likely to get.

2. Beware of Bias

Chapter 2 covers how LLMs are built and where bias enters the picture at the technical level. *Read it if you haven't.* What does bias mean for you when you're actually using these tools?

Training data comes from the internet, and the internet is a mirror of society. GenAI models learn from text that humans produced, and human-produced text carries the biases embedded in its sources. Models amplify patterns found in the training data. When those patterns reflect racial or gendered assumptions, the outputs do, too, usually with no visible signal that anything is wrong.

The legal profession has already confronted this problem through a different kind of algorithm. In *State v. Loomis*, the Wisconsin Supreme Court upheld the use of COMPAS — a commercial recidivism risk assessment tool — at sentencing, even though the defendant couldn't examine the algorithm's methodology because it was proprietary.¹¹ ProPublica's investigation of COMPAS found that Black defendants were nearly twice as likely as white defendants to be incorrectly classified as higher risk, a finding that generated years of scholarly argument that statisticians and computer scientists are still working through.¹² Courts haven't yet reached equivalent challenges to GenAI tools specifically, but the underlying problem is the same: a model trained on biased data produces biased output, and the user sees only the output.

For you as a law student, the takeaway is a reading habit. When you use GenAI to research a legal issue affecting a particular community, check whether it has accurately described the law rather than reproduced assumptions about the population involved. When it generates a draft argument, ask whether the framing would look different if the client's name suggested a different background. GenAI will not flag its own biases, and courts won't be impressed by the explanation that the GenAI wrote it that way. The professional judgment is yours.

3. GenAI for Brainstorming vs. GenAI for Final Product

Using an LLM to generate ideas, can be valuable so long as you are careful not to ask it to think for you.

Example: You're preparing for a moot court oral argument. You could ask GenAI this: "I've uploaded my draft. What are five challenging questions the judge might ask

¹¹ *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), *cert. denied*, 137 S. Ct. 2290 (2017). The Court held that COMPAS scores could be considered at sentencing as one factor among many, but cautioned that judges could not rely on them as the determinative basis for a sentence.

¹² Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks*, ProPublica (May 23, 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

about my standing argument?" In this way, you're using the technology for preparation and you still have to develop your own answers.

Compare that to asking: "What should I say if the judge asks about standing?" Now GenAI is doing the work that you need to practice. At the actual oral argument, you won't have an LLM feeding you answers.

The line isn't always bright, but ask yourself: Is GenAI helping me prepare to do the work, or is it doing the work for me?

4. The "Explain it Back to Me" Test

A reliable way to know if you're using GenAI productively works like this: After an LLM explains something to you, close the window and explain it to someone else (or write it out) in your own words. Can you do it? Can you still do it tomorrow?

If yes, GenAI enhanced your learning. If no, the technology gave you the illusion of understanding without actual comprehension.

Passive consumption of GenAI explanations is dangerous because reading a clear explanation *feels* like learning. Your brain processes the information and you think "I get it." But retrieving that information later — constructing an explanation from memory — is a completely different cognitive task. You'll need to perform that task on exams.

Never move on from a GenAI explanation until you can explain the concept back without looking. If you can't do that, you haven't learned it yet.

F. The CORRECT Framework for Students

In Chapter 4, you learned the CORRECT framework that I developed for effective prompt engineering. That chapter explained the framework in depth and showed you advanced prompting techniques for using GenAI productively. If you haven't read Chapter 4 yet, go back and read it now. This section builds on that foundation.

Here, I want to show you how to apply CORRECT specifically to your learning goals as a law student. The framework works the same way, but your objectives and use cases are different from your professors'. Where faculty might use CORRECT to design assignments or generate teaching materials, you're using it to build expertise and prepare for exams.

Think of CORRECT as a recipe. Experienced cooks learn to improvise, but only after they've followed the instructions precisely enough times to understand why each step matters. Not every prompt requires all seven elements, and experienced users learn to abbreviate strategically based on context. But you won't know what to leave out until you've practiced putting it all in.

Let's walk through each element of CORRECT for student-focused applications.

1. Context: Setting Up Your Learning Goals

When you prompt GenAI, always start by giving it context about what you're trying to learn and where you are in that learning process. This makes GenAI's responses far more useful. Remember, too, that grounding the GenAI in knowledge (your notes, the professor's slides, cases) delivers much more reliable output.

Weak prompt: "Explain proximate cause."

Strong prompt: "I'm a 1L in my third week of Torts. We just read *Palsgraf* and *Wagon Mound*. I understand that proximate cause limits liability, but I'm confused about the difference between direct cause cases and foreseeable plaintiff cases. Can you help me understand the distinction?"

The second prompt tells GenAI:

- Your knowledge level (1L, three weeks in)
- What you've already learned (specific cases)
- What you do understand (proximate cause as a limitation)
- What specifically confuses you (the doctrinal distinction)

This lets GenAI tailor its response to your actual needs rather than giving you a generic explanation that might be too basic or too advanced.

Context is especially important when asking GenAI to generate practice problems. Compare:

Weak prompt: "Give me a contracts practice problem."

Strong prompt: "I'm studying offer and acceptance. I understand the mailbox rule but I'm shaky on whether an acceptance is valid when the offer specified a particular method of acceptance and the offeree used a different

method. Can you create a fact pattern that tests this specific issue? Make it moderately difficult. I don't need it to be easy, but don't combine it with other doctrines yet."

The second prompt will get you a practice problem that's perfectly calibrated to your current learning edge.

2. Objective: Clear Learning Outcomes

Be explicit about what you want to achieve from the GenAI interaction. Are you trying to understand a concept? Practice applying it? Test whether you've mastered it? Identify gaps in your knowledge?

Different objectives require different GenAI interactions:

If your objective is understanding: "Explain [concept] using an example that's different from the cases in my casebook. Then quiz me to see if I understood your explanation."

If your objective is practice: "Generate five fact patterns that test [doctrine]. Start with an easy one and get progressively harder. After I analyze each one, tell me what I missed."

If your objective is testing mastery: "I think I've mastered [topic]. Create a hard hypothetical that combines multiple issues. Don't give me any hints—I'll work through it completely, then you can tell me how I did."

If your objective is identifying gaps: "I've outlined everything I know about [subject]. Ask me questions until you find something I can't explain clearly."

Notice how each of these prompts will lead to a different kind of interaction with GenAI. Being clear about your objective ensures you get the kind of practice you need.

3. Role(s): GenAI as a Tutor, Not as the Answer Key

This is critical for students. Not only should you tell the LLM who you are (e.g., a 1L studying contract formation), you should also explicitly tell what its role is. You want it to function as a personal tutor who will refuse to just give you the answers. Otherwise, GenAI's default behavior is to be helpful by providing direct answers, which sabotages your learning.

Here are role specifications that work well for different learning tasks:

For Socratic practice: "Act as a Socratic tutor. I'm going to answer your questions, but don't tell me if I'm right or wrong. Instead, ask follow-up questions that help me discover whether my reasoning is sound. Keep pushing me with harder questions."

For issue-spotting practice: "You're a law professor grading my exam answer. I'm going to analyze this fact pattern. Point out issues I missed, but don't tell me the answer to those issues."

For oral argument practice: "You're a judge hearing my oral argument. Ask me challenging questions about weaknesses in my position. If I give a weak answer, don't accept it. Press me harder. Make me defend my reasoning."

For concept checking: "I'm going to explain [doctrine] to you as thoroughly as I can. Your job is to identify any gaps, errors, or unclear parts of my explanation. Don't correct me directly. Ask questions that reveal what I'm missing."

Each of these role specifications changes how GenAI responds to you. Instead of spoon-feeding you, GenAI creates opportunities for you to teach yourself, which is how learning works.

4. Restrictions: Designing a Productive Struggle

This is where you engineer the difficulty that makes learning stick. You deliberately restrict GenAI's helpfulness to force yourself to do cognitive work.

Restriction examples:

"Don't give me the rule. Ask me questions until I can derive it from the cases myself."

"Generate a fact pattern, but don't tell me what issues are present. I'll identify them myself."

"If I make an error in my analysis, don't correct me. Instead, ask me a question that would help me spot the error myself."

"Create increasingly difficult hypotheticals about [doctrine]. Don't tell me I'm ready for the next level. I'll tell you when I want a harder problem."

"When I ask you if my answer is correct, don't answer yes or no. Instead, ask me to explain my reasoning in more depth."

These restrictions are uncomfortable and require time. They force you to struggle and struggle creates learning. Think about the investment you're making in time and money to become a lawyer. Instead of buying a library of commercial study aids, consider subscribing to Claude.

Students often resist this approach because it feels inefficient (and harder). Why not just ask GenAI for the answer and move on? You'll cover more material that way. But covering material isn't the same as learning material. When exam day comes, you'll need to construct answers from scratch. If you've practiced getting answers from GenAI instead of generating them yourself, you've practiced the wrong skill.

5. Expectations (of Format): Structuring GenAI's Response

Tell GenAI exactly what format you want for its output. This ensures you get responses that match your learning needs.

Format expectations for practice:

- Generate three fact patterns in order of difficulty. For each one, state only the facts, do not give legal analysis. After I work through all three, then give me feedback on each.
- Create a 10-question multiple-choice quiz in UBE style. Ask me one question at a time and do not reveal my score until I have answered all ten. When the quiz is complete, analyze my performance, identify patterns in what I got right and wrong, and explain the correct reasoning for any question I missed. When you give me feedback on my analysis, structure it as: (1) What I got right, (2) What I missed, (3) One question to push my thinking further.

Format expectations for study materials:

- Using my notes, create flashcards in a two-column format: legal term on the left, definition plus example on the right.
- Using my notes, generate a comparison chart showing negligence, strict liability, and intentional torts side by side. Include: elements, burden of proof, common defenses, and typical fact patterns for each.
- Using my notes, outline the contract formation rules as a flowchart. Use plain language for decision points and show me where each doctrine applies.

Format matters because well-structured output is easier to use for practice. When GenAI gives you everything in one block of text, you're tempted to read through it all at once, which amounts to passive learning which is not effective learning. When you

specify a format that requires you to engage with each piece separately, you force yourself into active learning.

6. Clarification (of Task): Specific Learning Activities

Be precise about what you want GenAI to help you do. Vague requests get vague responses.

Specific task clarifications:

- I'm going to brief [this case] out loud to you (limited to GenAI platforms that allow this feature). After I'm done, ask me questions about anything I missed or got wrong.
- Using my notes, create a study session for me on consideration in contracts that I should be able to complete within 30 minutes. Include explanation, examples, and practice problems. Check my understanding at each step. (Note, GenAI platforms can estimate how long something should take and recommend times of completion, but you must keep your own time. It does not have a built-in timer).
- Using my notes, generate 20 flashcards covering the major doctrines in criminal law. Put the doctrine name on one side and the elements plus a quick example on the other. NotebookLM does an exceptional job auto-generating flashcards from your notes, professor's slides, cases, etc.
- I'll describe a client scenario. You play the client and ask me questions. I'll try to give clear legal advice in plain language. Tell me if my explanations are confusing or if I'm making it too complicated.

Each of these is a specific, actionable task. GenAI knows exactly what you want and can deliver it. Compare these to vague requests like "help me study for my torts exam" or "explain contracts to me." Specific tasks create structured practice opportunities. Vague requests lead to unfocused GenAI responses that don't build real skill.

7. Tone: Setting the Right Learning Environment

The tone you set determines whether GenAI pushes you appropriately or coddles you. For learning, you usually want GenAI to be demanding, not accommodating.

Tone specifications for rigorous practice:

- Be direct and challenging in your feedback. I don't need encouragement. I need to know what I got wrong and why.

- Don't cushion criticism. If my analysis is weak, tell me it's weak and explain why.
- Push back on vague or conclusory statements. Make me defend every assertion I make.
- If I'm making the same mistake repeatedly, call that out explicitly. Don't let me get away with it.

Tone specifications for building confidence:

- I'm struggling with [this doctrine] and feeling overwhelmed. Acknowledge what I get right before pointing out what needs work.
- I'm practicing for the first time. Encourage me to keep trying even when I make mistakes.

But be careful with the second category. If you always ask for gentle feedback, you won't develop the resilience you need for exams or when the Socratic method is used in class. Most of the time, you should be asking GenAI to challenge you hard. There's no one else around, right? So what if you get something wrong? Save the gentler tone for moments when you're genuinely struggling and need to rebuild confidence before pushing forward.

G. Putting CORRECT Together: Sample Student Workflows

Let me show you what this looks like in practice with a complete example.

Scenario: You're studying personal jurisdiction in Civil Procedure. You've read the cases in your casebook (*International Shoe*, *World-Wide Volkswagen*, *Burger King*) and you understand the general framework, but you're shaky on applying the "minimum contacts" test.

Your CORRECT prompt:

Context: I'm a 1L studying personal jurisdiction. I've read *International Shoe*, *World-Wide Volkswagen*, and *Burger King*. I understand that personal jurisdiction requires minimum contacts and that the exercise of jurisdiction must be fair and reasonable. But I'm struggling to apply this to new fact patterns. I can't always tell whether contacts are sufficient.

Objective: I want to practice analyzing personal jurisdiction fact patterns until I can confidently identify when a court would have jurisdiction and when it wouldn't.

Role: Act as a law professor who uses the Socratic method. Don't tell me the answers. Ask me questions that help me work through the analysis myself.

Restrictions: Don't tell me whether contacts are sufficient or not. Instead, ask me questions about the nature of the contacts, their relationship to the claim, and the fairness factors. If I miss something important, ask me a question that points me toward it.

Expectations (of Format): I want three types of fact patterns: a straightforward fact pattern where jurisdiction is clearly present, a borderline case, and one where jurisdiction clearly doesn't exist. Present each fact pattern separately. Give them to me in a random order, and give me more than one question of each type. After I analyze each, give me feedback before moving to the next.

Clarification (of Task): Create the first fact pattern now. After I work through my analysis, ask me Socratic questions about my reasoning. Once we've discussed that pattern, move to the next one. After all three, I'll explain the test back to you in my own words and you'll tell me if my understanding is accurate.

Tone: Be demanding but not discouraging. Push me hard on logical gaps, but acknowledge when my reasoning is sound. I need rigorous practice, not praise.

This prompt sets up a structured learning session that will help build your analytical skills. You're not asking GenAI to explain the doctrine to you. You're asking it to create opportunities for you to practice applying it.

Compare this to: "Explain personal jurisdiction to me." That prompt will get you a clear explanation that you'll read, think you understand, and then struggle to apply on an exam.

The CORRECT framework forces you to be intentional about your learning. It transforms GenAI from a passive information source into an active practice partner.

1. Try Build Your Own CORRECT Prompt

Pick a doctrine you're currently studying. Spend 5 minutes crafting a CORRECT prompt that will help you practice it. Include all seven elements. Ground it in knowledge (i.e., provide your notes, professor's slides, cases). Then use it for a 20-minute study session. Did you learn something?

Use the [interactive CORRECT Prompting tool](#) to practice.

CHAPTER 14: For Students: From Academic Integrity to Professional Responsibility

By the time you reach 2L or 3L, your law school's honor code has become something like wallpaper. You signed the pledge during orientation and have certified compliance on every exam since, which makes an actual hearing before the Academic Integrity Committee feel like something that happens to other students, the ones who make spectacularly bad decisions in full view of witnesses. What your professors know, and what takes most students too long to appreciate, is that the academic integrity decisions you make right now about GenAI use are not primarily about avoiding trouble in law school. Every time you decide whether and how to use an LLM, you are practicing professional judgment in real time. The habits you build in law school travel with you to your first supervising partner and, eventually, to every client who trusts you with something that matters.

I know this from my own teaching. When I teach AI and the Law, Law Practice Technology, or Social Media and the Law, I always begin with professional responsibility. Not as an afterthought in week twelve, but as the foundation in week one.¹ The same is true when I speak at conferences or address bar associations about GenAI. Regardless of whether my audience is law students or seasoned practitioners, I start with ethics. Why? Because technology without ethical grounding is dangerous, and lawyers who learn powerful tools without first understanding professional obligations make predictable mistakes.² Lawyers filing documents with hallucinated citations appear in the news every week, which is remarkable given how long this has been happening and how severely courts have responded.³ You cannot separate questions about what technology can do from questions about what lawyers should do. The "how" and the "should" are inseparable. Students sometimes seem surprised or disappointed that a course about

¹ In AI and the Law, I follow coverage of legal ethics with ethical AI, where we discuss alignment, transparency, explainability, auditability, and bias, and other considerations for trustworthy AI.

² See Damien Charlotin, *AI Hallucination Cases*, <https://www.damiencharlotin.com/hallucinations/>, for updates to the many lawyers who have been sanctioned by the courts for filings with hallucinated citations; see also *Jacqueline Lacey v. State Farm*, Case 2:24-cv-05205-FMO-MAA (C.D. Cal.) (May 6, 2025), law firms were ordered to pay the defense \$31,100; and see *Jordon v. Chi. Hous. Auth.*, No. 2022 L 000095 (Ill. Cir. Ct.) (Dec. 5, 2025) where the court imposed a penalty of \$59,500 on the attorneys for relying on a fictitious case.

³ Frank Young, *The Price of Carelessness: Big Law Sanctioned for AI Miscues*, U. Ill. Chi. S. of Law, (Dec. 11, 2025), <https://library.law.uic.edu/news-stories/illinois-court-imposes-59500-penalty-for-attorneys-ai-use/> (last visited Mar. 11, 2026).

technology and the practice of law starts with Model Rules and state bar ethics opinions. Practicing lawyers sometimes fidget when a CLE program begins with disciplinary cases and cautionary tales rather than productivity tips. But teaching technology skills without first establishing ethical frameworks would be, in my opinion, professional malpractice. The same principle applies to your learning. Before you master GenAI prompting techniques or build custom assistants, you need to understand the professional obligations that will govern your use of these tools.

The connection between academic integrity and professional responsibility is not metaphorical. The same qualities that your honor code attempts to cultivate (honesty, diligence, respect for rules, independent judgment) form the foundation of the Model Rules of Professional Conduct that will govern your practice.⁴ When a law student misrepresents the extent of GenAI assistance on an assignment, the student is rehearsing the same mental shortcuts and rationalizations that lead lawyers to mislead courts or claim credit for work they did not perform. Your law school years are not a grace period before professional obligations kick in. Law school is professional training, and every choice you make either strengthens or weakens your capacity for ethical practice.

*Your law school years are not a grace period before
professional obligations kick in.*

The arrival of GenAI has made these connections more visible and more urgent. In the past, most academic integrity violations involved obvious misconduct such as copying from another student, using prohibited materials during a closed-book exam, plagiarizing from a published source. The lines seemed clear, even if students sometimes crossed them, but GenAI has blurred many of those lines. When you ask an LLM to explain a concept, are you improperly outsourcing your learning? When you use it to generate a rough draft that you then substantially revise, have you misrepresented your work? When you check your legal research against GenAI results, are you demonstrating diligence or academic dishonesty? The answers depend on your professor's policies, your intended use, your level of transparency, and your own developing judgment about what constitutes authentic learning and honest work.

The ambiguity that you face in law school mirrors the ambiguity you will face in practice. Law firms and legal employers are wrestling with the same questions. When

⁴ MODEL RULES OF PRO. CONDUCT (A.B.A. 2020). Most jurisdictions have adopted versions of the Model Rules, though specific language varies by state.

must GenAI use be disclosed to clients or courts? What level of human oversight is required? How do we maintain competence when powerful tools can perform tasks that used to require years of training and can do it better than a human? The habits you develop now, in the relatively low-stakes environment of law school, will shape how you answer those questions when your client's liberty, property, or rights depend on your judgment.

A. The Parallel Structure of Academic and Professional Obligations

Most law school honor codes, university conduct codes, and the Model Rules of Professional Conduct share a common architecture. Both rest on principles of honesty, competence, confidentiality, and accountability. Understanding these parallels can help you see that academic integrity is not just a hoop to jump through before you get to do "real" legal work. Academic integrity is the foundation of professional responsibility.

Consider the comparison below, which maps typical honor code provisions against their professional responsibility counterparts:

Table 14-1: PARALLEL PROVISIONS: HONOR CODE AND MODEL RULES

Honor Principle	Code	Typical Honor Code Language	Model Rule	Professional Responsibility Standard
Honesty in Academic Work		"Students shall not misrepresent their own work or the degree of assistance received from any source."	Rule 8.4(c)	"It is professional misconduct for a lawyer to engage in conduct involving dishonesty, fraud, deceit or misrepresentation."
Proper Attribution		"Students must properly attribute all sources and clearly distinguish their own analysis from that of others."	Rule 1.1, Rule 3.3(a)(1)	Lawyers must provide competent representation and shall not make false statements of fact or law to a tribunal.
Confidentiality		"Students shall not disclose the content of examinations or share information that would provide unfair advantage."	Rule 1.6	"A lawyer shall not reveal information relating to the representation of a client" without informed consent.

Honor Principle	Code	Typical Honor Language	Model Rule	Professional Responsibility Standard
Compliance with Rules		"Students must follow all announced course policies and examination rules."	Rule 8.4(a), Rule 5.1	Lawyers shall not violate the Rules of Professional Conduct and supervisors must ensure compliance.
Candor to Authority		"Students must be truthful in all dealings with faculty and administration."	Rule 8.1	Lawyers shall not make false statements of material fact in connection with admission to the bar or disciplinary matters.
Duty to Report Misconduct		"Students who have knowledge of a violation of this honor code are obligated to report such violations to the appropriate academic authority."	Rule 8.3, Rule 5.3	Lawyers who know that another lawyer has committed a violation raising a substantial question as to that lawyer's fitness shall inform the appropriate professional authority; lawyers must also take reasonable steps to ensure that nonlawyer assistants' conduct is compatible with professional obligations.
Fairness		"Students shall not engage in any conduct which tends to gain an unfair advantage for any student."*	Rule 4.4(a), Rule 8.4(d)	In representing a client, a lawyer shall not use means that have no substantial purpose other than to embarrass, delay, or burden a third person; lawyers shall not engage in conduct prejudicial to the administration of justice.
Nondiscrimination†		"Discriminatory conduct or harassment based on race, ethnicity, religion, gender, sexual orientation, disability, or any other protected characteristic is strictly prohibited."	Rule 8.4(d), Rule 8.4(g)‡, Rule 4.4(a)	Lawyers shall not engage in conduct prejudicial to the administration of justice; in adopting jurisdictions, lawyers shall not engage in conduct constituting harassment or discrimination based on protected characteristics in conduct related to the practice of law.

* Fritz Snyder & Shirley Goza, *Law School Honor Codes*, Faculty Journal Articles & Other Writings 35 (1983).

† Nondiscrimination principles of this kind appear more commonly in student conduct codes than in honor codes specifically.

‡ Rule 8.4(g) has been adopted in approximately half the states and remains contested.

B. Standing Orders and Real-World Implementation

1. The Third Parallel: Standing Orders and Real-World Implementation

The connection between academic rules and professional responsibility becomes even clearer when you examine how courts have responded to GenAI. Just as your professors establish course-specific policies about GenAI use, courts across the country have issued Standing Orders and Court Rules that impose additional requirements beyond the Model Rules. Before claiming that it's not fair that different faculty have different policies about GenAI recognize that a similar problem exists in practice. These orders demonstrate that the skills you are building now (reading syllabi carefully, checking for specific requirements, and documenting compliance) transfers directly to practice, and they vary by jurisdiction.

Consider how Standing Orders mirror the course-by-course variation you navigate in law school:

Table 14-2: Sample Court Standing Orders on GenAI - The Real-World Syllabus

Court	Standing Order Requirement	Parallel Academic Rule	Practice Implication
E.D. Penn. (Judge Baylson)	If AI used in any way in preparation of any filing, the lawyer must disclose AI use in clear factual statement and certify that every citation to law or record has been verified as accurate. ⁵	Syllabus requires disclosure statement for any AI-assisted work and certification that all sources have been independently verified.	Transparency and verification are non-negotiable; even minimal AI use triggers disclosure obligation.
Illinois Supreme	Attorneys remain fully accountable for final work product; must thoroughly review all AI-generated	Syllabus requires students to review and verify all sources used in assignments; must understand any	Personal accountability for work product cannot

⁵ Standing Order Re: Artificial Intelligence ("AI") in Cases Assigned to Judge Baylson, Hon. Michael M. Baylson, U.S. Dist. Ct., E.D. Pa. (June 6, 2023), <https://www.paed.uscourts.gov/sites/paed/files/documents/locrules/standord/Standing%20Order%20Re%20Artificial%20Intelligence%206.6.pdf> (last visited Feb. 11, 2026) (his was one of the **earliest** Standing Orders on GenAI use (June 2023, shortly after the *Mata v. Avianca* sanctions case made headlines in May 2023). Judge Baylson's order is notable for its breadth: "in any way" means even minor AI assistance triggers the disclosure requirement. This mirrors the strictest law school syllabi that require disclosure regardless of how minimal the AI use was, asking Claude to check grammar on a brief still requires disclosure under this standard).

Court	Standing Order Requirement	Parallel Academic Rule	Practice Implication
Court (Statewide)	content before submission and understand both general AI capabilities and specific tools used. ⁶	tools employed and cannot claim "the AI did it" as defense.	be delegated to technology,
S.D. Cal. (Judge Robinson)	GenAI use is not prohibited, but unqualified reliance may result in fabricated case law; Rule 11 requires reasonable inquiry and verification; failure to exercise due care may result in sanctions or disciplinary action. ⁷	Assignment/Assessment instructions specify acceptable research tools and verification requirements.	Protects against hallucinated authorities,
E.D. Va. (Pretrial Schedule A)	Attorneys must certify that no AI was used for research except AI embedded in standard legal research platforms (Westlaw, Lexis, FastCase, Bloomberg); must certify that every citation has been checked by attorney or paralegal for accuracy. ⁸	Professor prohibits GenAI use entirely except for specified course-approved research tools; requires certification statement on all submissions.	Some forums prohibit GenAI outright; blanket permission does not exist everywhere.

Note that Standing Orders are subject to change, so you must keep up-to-date.

The patterns here tell the real story. No uniform rule exists across jurisdictions, and sometimes not even within a single jurisdiction. Some courts take a zero-tolerance approach to GenAI, just as some of your professors prohibit it entirely. Other courts explicitly approve certain tools (traditional legal research platforms like Westlaw and Lexis, for instance) while banning everything else. The variance means that knowing what worked last semester tells you nothing about what you can do right now, in *this*

⁶ Illinois Supreme Court, Policy on Artificial Intelligence (eff. Jan. 1, 2025), <https://www.lawnext.com/wp-content/uploads/2024/12/Illinois-Supreme-Court-AI-Policy.pdf> (last visited Feb. 11, 2026).

⁷ Standing Order for Civil Cases § III.D, Hon. Todd W. Robinson, U.S. Dist. Ct., S.D. Cal. (revised Sept. 16, 2024), https://www.casd.uscourts.gov/Judges/robinson/docs/Civil_Standing_Order.pdf, (This Standing Order is particularly notable because it explicitly states that GenAI use is not prohibited while simultaneously making clear that permission doesn't eliminate the duty of care. This mirrors exactly what students experience: a professor may allow GenAI use, but that permission doesn't make the student any less responsible for verifying accuracy and ensuring the work meets standards.)

⁸ Pretrial Schedule A § I.E., U.S. Dist. Ct., E.D. Va., Richmond Div. (revised Jan. 2, 2025), Case No. 3:24-cv-00575-REP, Doc. 35-1. (This Standing Order represents the most restrictive approach among the examples and creates a critical teaching moment. Unlike Judge Robinson's order (which permits GenAI use with verification) or the Illinois Supreme Court policy (which emphasizes accountability), the Eastern District of Virginia essentially prohibits generative AI use for legal research entirely, allowing only the AI already embedded in traditional legal research platforms.)

class taught by *this* professor. Nor do the rules that worked for Judge A in the last case have any bearing on what Judge B requires in this new case.

Lawyers and pro se litigants must read and follow the specific requirements for the particular court in which they appear. Students must read and follow every syllabus policy. The obligation is identical, and the excuse "I didn't know" or "Nobody told me" fails equally in both contexts. The student who skips reading a syllabus and assumes all professors have the same GenAI policy is demonstrating the same bad habit as the lawyer who files a brief without checking the court's Standing Orders. Both will eventually face consequences that could have been avoided by spending five minutes reading the rules that govern the space in which they are working.

The consequence of failing to check Standing Orders mirrors academic integrity violations. In *Mata v. Avianca, Inc.*, attorney Steven Schwartz submitted a brief containing AI-generated, fabricated case citations without verification.⁹ The court sanctioned Schwartz not primarily for using GenAI (which was not itself prohibited), but for failing to verify the output and then compounding the problem by being less than candid about what happened.¹⁰ The violations stacked: incompetence (Rule 1.1), lack of candor to the tribunal (Rule 3.3), and conduct involving dishonesty (Rule 8.4(c)).¹¹ Sound familiar? That is the same escalating consequence structure in your law school honor code. A failure to follow one rule (checking your work) leads to violations of multiple principles (honesty, competence, accountability). [See Table 11-2: Comparative Framework: Law School Policy & Federal Court Standing Orders].

2. From Academic Integrity to Rule 11: A Structural Parallel

The skill you develop reading each professor's syllabus for specific GenAI requirements transfers directly to practice, where you must check Standing Orders for every court in which you file. Consider the striking parallels between law school policies and federal court requirements shown in Table 11-2 below. These academic rules are deliberately training you for professional obligations that you will face from your first brief forward.

⁹ *Mata v. Avianca, Inc.*, No. 1:2022cv01461 - Document 54 (S.D.N.Y. 2023), <https://law.justia.com/cases/federal/district-courts/new-york/nysdce/1:2022cv01461/575368/54/> (last visited Feb. 11, 2026).

¹⁰ *Id.*

¹¹ MODEL RULES OF PRO. CONDUCT r. 1.1, r. 3.3, r. 8.4(c) (A.B.A. 2020).

Table 14-3: Comparative Framework: Law School Policy & Federal Court Standing Orders

LAW SCHOOL SYLLABUS Torts II – Professor Chen (Spring 2025)	COURT STANDING ORDER U.S. District Court (E.D. Va.) & S.D. Cal.
Policy on Use of Generative AI Tools Disclosure Required: • Identify AI tool(s) used • Specify tasks performed • Confirm independent verification of all output Verification Mandatory: • Review every citation for accuracy • Verify all factual claims • Ensure sound legal analysis • Confirm proper attribution Prohibited Uses: • No AI use during examinations • No unverified citation generation • No submission of AI output as original work Consequences: • Referral to Honor Code Committee • AI error is not a defense • Fabricated citations = academic dishonesty Approved Research Tools: Westlaw, Lexis, Bloomberg Law (no disclosure required)	Certification Requirement Disclosure Required: • Certification regarding AI use in preparation of filings • Exception for embedded AI in traditional research platforms Verification Mandatory: • Every statement and citation must be checked • Accuracy of proposition confirmed by attorney or supervised staff Rule 11 Duties: • Reasonable inquiry required • Filing must be factually grounded and legally tenable • Failure to exercise due care may result in sanctions Judicial Warning: • Unqualified reliance on AI may produce fabricated case law resulting in sanctions • Exposure to sanctions or disciplinary action

The academic requirement to verify every citation and remain fully responsible for submitted work mirrors the federal court’s Rule 11 certification duty. Verification discipline developed in law school directly translates to practice.

The parallel extends to the consequences of non-compliance. When you fail to follow a professor's GenAI policy:

- Initial consequence: Academic integrity proceeding

- Honor code finding: Potential grade penalty, notation on transcript, or more severe sanctions
- Long-term impact: Character and fitness questions on bar application

When a lawyer fails to follow a court's Standing Order on GenAI:

- Initial consequence: Show cause order (explain why you shouldn't be sanctioned)
- Court finding: Monetary sanctions, referral to disciplinary authorities, potential case dismissal
- Long-term impact: Public record of sanction, potential bar discipline (including the risk of reciprocal discipline for attorneys barred in more than one jurisdiction), malpractice liability

The stakes escalate in practice, but the fundamental obligation remains identical. Attorneys and students must know and follow the rules and document compliance.

3. Building the Right Habits Now

Your professors are not being arbitrary when they require you to check their syllabi before using GenAI on assignments. They are training you to develop a critical habit of understanding that when you enter a new environment (a new class, a new court, a new jurisdiction), your first step must be checking the specific rules that govern that space. Every semester, you navigate different professors with different expectations. Some allow open-note exams, others do not. Some permit laptops; others require handwritten responses. Some encourage study groups; others expect individual work. Each professor's rules are clearly stated in the syllabus, and you are responsible for reading and following them.

When you file your first brief in state court, you will need to check that court's rules on formatting, filing deadlines, and now, GenAI use. When you appear before a new judge, you will need to review that judge's standing orders and chamber practices. When you practice in multiple jurisdictions, you will need to track varying requirements across different forums. The lawyer who says "I didn't know about that Standing Order" receives the same response as the student who says "I didn't read the syllabus" – ignorance of the rules is not a defense.^{[11](#)}

The good news? You are developing careful habits that will protect you from committing the catastrophic mistakes that end careers. The student who cuts corners and

hopes no one will notice becomes the lawyer in the headlines, explaining to a judge why they did not verify AI-generated citations.

C. How Academic Choices Shape Professional Judgment

Professional judgment is not an on-off switch that flips when you pass the bar exam. Judgment develops through repeated practice in navigating ambiguous situations and taking responsibility for your decisions. Every time you face a choice about GenAI use in law school, you are in a judgment-building situation. The choices you make accumulate.

Consider three students, all in the same Civil Procedure class with the same professor who has announced a policy allowing GenAI use "for brainstorming and outlining only, with all final analysis being your own work." Student A interprets this generously: she has GenAI draft full answers to practice problems, then edits them. She figures that editing is analysis, and the policy did not explicitly prohibit drafting. Student B interprets the policy cautiously: she uses GenAI to create topic lists and identify potentially relevant rules but writes all analysis from scratch. Student C finds the policy ambiguous and emails the professor to ask for clarification about specific uses. The professor takes no notice.

Five years from now, these three students are associates at different firms. Their firms have similar policies about GenAI: "Use GenAI to improve efficiency but maintain appropriate human oversight and do not disclose confidential client information." Student A, now Attorney A, pushes the boundaries. She has GenAI draft entire sections of briefs, sometimes uploading client documents to get better results, and conducts minimal review because she is overwhelmed with billable hour pressure. Student B, now Attorney B, uses GenAI for research and outlining but carefully reviews all output and never includes client-specific information in prompts. Attorney C, uncertain about what "appropriate oversight" means for a particular task, asks a senior partner for guidance.

Which attorney is most likely to face a malpractice claim, disciplinary complaint, or termination? Which attorney is most likely to become a trusted advisor whom clients and colleagues respect? The answer is obvious because the pattern was established in law school. Attorney A spent three years looking for loopholes and pushing boundaries, training herself to prioritize convenience over compliance. Attorney B spent three years being thoughtful about rules and careful about quality. Attorney C spent three years seeking clarity when rules were ambiguous, demonstrating the humility and good judgment that mark excellent lawyers.

You might object that this is unfair, that law students face enormous pressure and everyone is using GenAI, that the rules are unclear and professors are inconsistent. All of that may be true. But professional responsibility does not come with an exception for difficult circumstances. In fact, the rules exist *because* difficult circumstances are the only kind lawyers face. Lawyers also face enormous pressure. Clients also make demands that require creative solutions. The Model Rules also contain ambiguities. And lawyers who respond to pressure by cutting corners, who respond to unclear rules by assuming the most favorable interpretation, or who respond to ambiguity by avoiding questions find themselves in serious trouble.

Law school can be brutal, particularly the first year. I've been teaching long enough that even though my own 1L experience occurred some time ago, I can empathize with the current students and what they are experiencing. I talk to students and colleagues. I read the law school subreddits. Some are heartbreaking. I see posts about the anxiety of cold calls, the despair of getting back exam results that don't reflect the hours you put in, the shame of academic probation, the panic when summer job applications go nowhere. I know students who are supporting families, who deliver babies mid-semester or during finals, who work night shifts, or who are managing disabilities, grieving losses, and battling imposter syndrome. The pressure to achieve high grades and class rank can feel crushing, and the consequences of falling short can seem catastrophic.

Acknowledging these struggles is not the same as resigning to them. The students who come through law school with their sense of self intact tend to have built a few things that grades cannot measure. They have found peers whose values align with their own rather than classmates who treat every interaction as a competition. They have at least one mentor—a faculty member, a practitioner, an alumnus—willing to give honest guidance rather than just a good recommendation. And they have learned to ask for help before a situation becomes a crisis, which is harder than it sounds in an environment that too often treats needing help as a weakness.

Following honor codes and building professional judgment is easy when you're exhausted and scared. But I'm also not going to lie to you about what's at stake. The shortcuts that feel necessary in the moment create problems that compound over time. The habits you develop under pressure are the habits that will define your practice. So, while I acknowledge the difficulty of what law school demands, I'm asking you to be thoughtful about how you respond to that difficulty. Your future clients deserve lawyers who developed real competence, even when that competence was hard-won.

The good news is that professional judgment, like any skill, improves with practice. Every time you pause to consider whether your intended GenAI use complies with a professor's policy, you are strengthening the neural pathways that will help you pause before sending a client email that might breach confidentiality. Every time you disclose your GenAI use when disclosure is required, you are practicing the transparency that will make you a trustworthy lawyer and one who can comply with court rules. Every time you verify an LLM's output instead of trusting it blindly, you are building the healthy skepticism that protects clients from error.

D. Scenarios: Where the Lines Get Tested

The scenarios below present common GenAI dilemmas that law students face. Each scenario includes an "In School" and "In Practice" version to help you see the through-line from academic to professional responsibility.

1. Scenario 1: The Citation Check

a) In School

You finished a seminar paper and used GenAI to help generate some research leads. You found several excellent cases through traditional research, but GenAI suggested one case that perfectly supports your argument (always be wary about this). You cannot access the case through your library's databases (it is from another jurisdiction and behind a paywall), but the GenAI summary seems solid. You cite the case based on the GenAI summary. Your paper is well-received, and you get an A.

Six months later, you learn that the case does not exist. GenAI fabricated it. Your professor contacts you, upset and questioning your academic integrity. You face potential honor code charges for misrepresenting your research.

b) In Practice

You are researching a motion for summary judgment. You use GenAI to identify potentially relevant cases, and it suggests a Ninth Circuit decision that would be persuasive authority in your jurisdiction. You cannot access the full opinion through Westlaw or Lexis, but the GenAI summary indicates it is directly on point. You cite the case in your brief. The judge's clerk calls it to the attention of the court, opposing counsel cannot find it, and you realize the case does not exist.¹² You face potential Rule 11 (and

¹² See, e.g., *Mata v. Avianca, Inc.*, No. 22-cv-1461, 2023 WL 4114965 (S.D.N.Y. June 22, 2023) (sanctioning lawyers for submitting brief with fabricated citations generated by ChatGPT); see also, Damien Charlotin,

any judge or court standing order) sanctions, a state bar complaint, and public embarrassment for your client and your firm.¹³

Figure 14.1: Decision Tree: The Citation Verification Question

Are you citing a source in academic work or a legal brief?

↓

Did you personally read the full source?

└─ Yes → Proceed with citation

└─ No → Can you access and verify the source?

└─ Yes → Read it before citing

└─ No → Do not cite it

→ If GenAI found it, assume it might be fabricated

→ Search traditional databases to verify existence

→ If you cannot verify, do not use

In law school, a fabricated citation is an honor code violation. In practice, it is potential malpractice and professional misconduct. The solution is the same in both contexts: verify every source personally before citing it. No exceptions, no shortcuts, no assumptions that the LLM is *probably* right this time.

2. Scenario 2: The Collaboration Question

a) *In School*

Your professor assigns a complex problem set to be completed individually. You and your study group meet to discuss the fact pattern, and someone suggests putting the problem into an LLM to see what issues it identifies. The group discusses the GenAI output, and you all refine your understanding of the relevant law. When you write your individual answer, you include analysis that came from the group discussion, which was

AI Hallucinations Cases, <https://www.damiencharlotin.com/hallucinations/> (last visited Jan. 27, 2026) (as of January 27, 2026 there were 590 cases in the United States alone).

¹³ FED. R. CIV. P. r. 11 requires that legal contentions be "warranted by existing law" and have "evidentiary support."

informed by GenAI output. You do not disclose the group discussion or GenAI use because the professor's policy only prohibited "receiving answers from other students" and you wrote your answer yourself.

The professor notices that five students submitted answers with strikingly similar issue identification and organizational structure. After investigation, your entire study group faces honor code charges for impermissible collaboration.

b) In Practice

You are working on a complex securities matter with a team of associates. Your supervising partner asks you to research a narrow issue. You are struggling to understand the regulatory framework, so you upload your research memo draft to an LLM and ask it to identify gaps in your analysis. The LLM response includes analysis that requires access to confidential client information you provided in the prompt. You incorporate the improved analysis into your memo without flagging that you used GenAI or considering where the LLM might have stored your prompt.

Later, you learn that your firm's GenAI policy prohibits uploading any client-specific information to external LLMs. Worse, a data breach at the LLM provider leads to potential disclosure of your client's confidential information. You face potential discipline under Rule 1.6 (confidentiality) and Rule 5.3 (failure to supervise nonlawyer assistants, including technology tools).¹⁴

Figure 14:2: Decision Tree: The Collaboration and Confidentiality

Before sharing information with GenAI:

↓

Does professor policy allow group GenAI use?

└ Yes

→ Document and disclose as required

└ No

→ Do not use GenAI as collaboration workaround

Does the information include client-specific facts or confidential details?

¹⁴ MODEL RULES OF PRO. CONDUCT, r. 1.6 (A.B.A. 2023); MODEL RULES OF PRO. CONDUCT, r. 5.3 (A.B.A. 2023).

└ Yes

→ Do not upload to external GenAI

→ Check if firm has approved internal/secure system

→ If no approved system exists, work without GenAI

└ No → Proceed with appropriate professional safeguards

GenAI can feel like using a research tool, but it often functions more like collaboration with another person. It can also amount to disclosure to a third-party affecting privilege. In school, using GenAI to work around collaboration restrictions is still collaboration. In practice, uploading confidential client information to external LLMs is a confidentiality breach. Both contexts require you to think carefully about who (or what) is seeing your work and whether that disclosure is permitted.

3. Scenario 3: The Competence Question

a) In School

You are struggling in Criminal Law. The doctrine is complex, the cases seem contradictory, there are so many exceptions, and you feel lost. You discover that GenAI can generate excellent case briefs and clear explanations of difficult concepts. So, for reasons of shame or fear of looking unprepared, instead of asking for help from your professor or academic support, you stop doing your own case briefs and instead review GenAI summaries before class. You use GenAI-generated outlines rather than creating your own. You pass the class, but as you're studying for the bar exam two years later, you realize you never actually learned Criminal Law. You've got to go back to the fundamentals, which takes time, and you don't have time. You struggle to apply doctrine that you never truly understood, and this failure to learn the doctrine in law school adversely could affect your bar exam performance.

b) In Practice

You are a junior associate in a litigation firm. Partners appreciate that you turn around research assignments quickly. Your secret is heavy GenAI use: you input research questions, get comprehensive answers, verify that cited cases exist, and deliver polished memos. You rarely read cases fully or develop your own analysis. After two years, a partner asks you to handle a discovery hearing last minute and without much guidance because "you have researched this issue thoroughly." At the hearing, opposing counsel

challenges your analysis, the judge asks probing questions, and you realize that you cannot defend positions you never really understood. Your superficial knowledge is exposed, and your reputation suffers.

E. SELF-REFLECTION: Am I Building Competence or Taking Shortcuts?

Consider your GenAI use in your most recent assignment or work product. For each question below, give yourself a rating:

1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always

_____ I read primary sources fully rather than relying on GenAI summaries.

_____ I attempt analysis myself before checking GenAI output.

_____ I can explain the reasoning behind my conclusions without referring to GenAI.

_____ I use GenAI to enhance work I could complete independently (even if more slowly).

_____ I verify GenAI factual claims and legal conclusions against reliable sources.

_____ When GenAI output differs from my own analysis, I investigate why rather than simply accepting the LLM's version.

_____ I feel confident I could complete similar work without GenAI access.

_____ I disclose my GenAI use when policies require disclosure (and err on the side of disclosure when uncertain).

Scoring:

32-40: You are likely using GenAI in ways that build competence and maintain integrity.

24-31: Mixed picture. Identify specific areas where you might be over-relying on GenAI.

16-23: Warning signs. Your GenAI use may be undermining your learning and professional development.

Below 16: Serious concern. You are likely not developing the competence you need for practice.

Competence is not binary. It's possible to pass classes and even pass the bar exam while having significant gaps in your legal knowledge and analytical skills. But those gaps will become visible in practice, often at moments when exposure is most damaging

to your client and your career. Rule 1.1 requires competence not just at the moment of admission to the bar, but throughout your career.¹⁵

*GenAI can accelerate your learning, or it can allow you to fake
competence that you do not have. The difference lies in how you use
GenAI.*

F. Accommodating Learning Differences

GenAI platforms include accessibility features that can make a real difference for students with dyslexia, ADHD, or other learning differences, and for non-native English speakers who rely on translation and language-support tools. These features adjust how content is presented rather than generating anything new, which seems like it should put them in a different category from AI-assisted writing. Your institution's AI policy may not see it that way, though, and students may still need formal accommodation approval from disability services before using even these features in a course.

1. Native Platform Features

Claude stands alone among the major GenAI platforms in offering a specific "dyslexic-friendly" font setting. Students can activate this feature by navigating to Settings and selecting the dyslexic-friendly option under text display preferences. [See Figure 11-1: Claude Settings showing "Dyslexic Friendly" Font Option]. While other platforms lack this specific setting, they compensate with alternative accessibility features that serve similar functions.

The real breakthrough for students overwhelmed by dense legal texts (and let's face it, casebooks are nothing but walls of text) comes from GenAI's ability to transform impenetrable prose into digestible formats. Students can ask platforms to summarize complex judicial opinions and output key points in clear, organized structures. A critical caveat applies here: *summary alone won't cut it for most law school purposes*. Students must engage in iterative conversations with the material. They must ask follow-up questions to test their understanding. Reading a case summary and moving on is like watching a cooking show and claiming you've learned to cook. See Figure

¹⁵ MODEL RULES OF PRO. CONDUCT r. 1.1. (A.B.A. 2023).

14:3.

Figure 14:3: Showing Claude Settings that includes “dyslexic friendly” font choice.

Billing

Usage

Capabilities

Connectors

Claude Code

What personal preferences should Claude consider in responses?

Your preferences will apply to all conversations, within Anthropic's guidelines.

When learning something new, I find examples and analogies useful. I prefer clear text with a bit of sardonic wit. If appropriate, ask clarifying questions.

Notifications

Response completions

Get notified when Claude has finished a response. Most useful for long-running tasks like tool calls, Research, and Claude Code on the web.

Emails from Claude Code on the web

Get an email when Claude Code on the web has finished building or needs your response.

Appearance

Color mode

Light

Auto

Dark

Chat font

Aa

Default

Aa

Sans

Aa

System

Aa

Dyslexic friendly

Voice settings

Voice

Buttery

Airy

Mellow

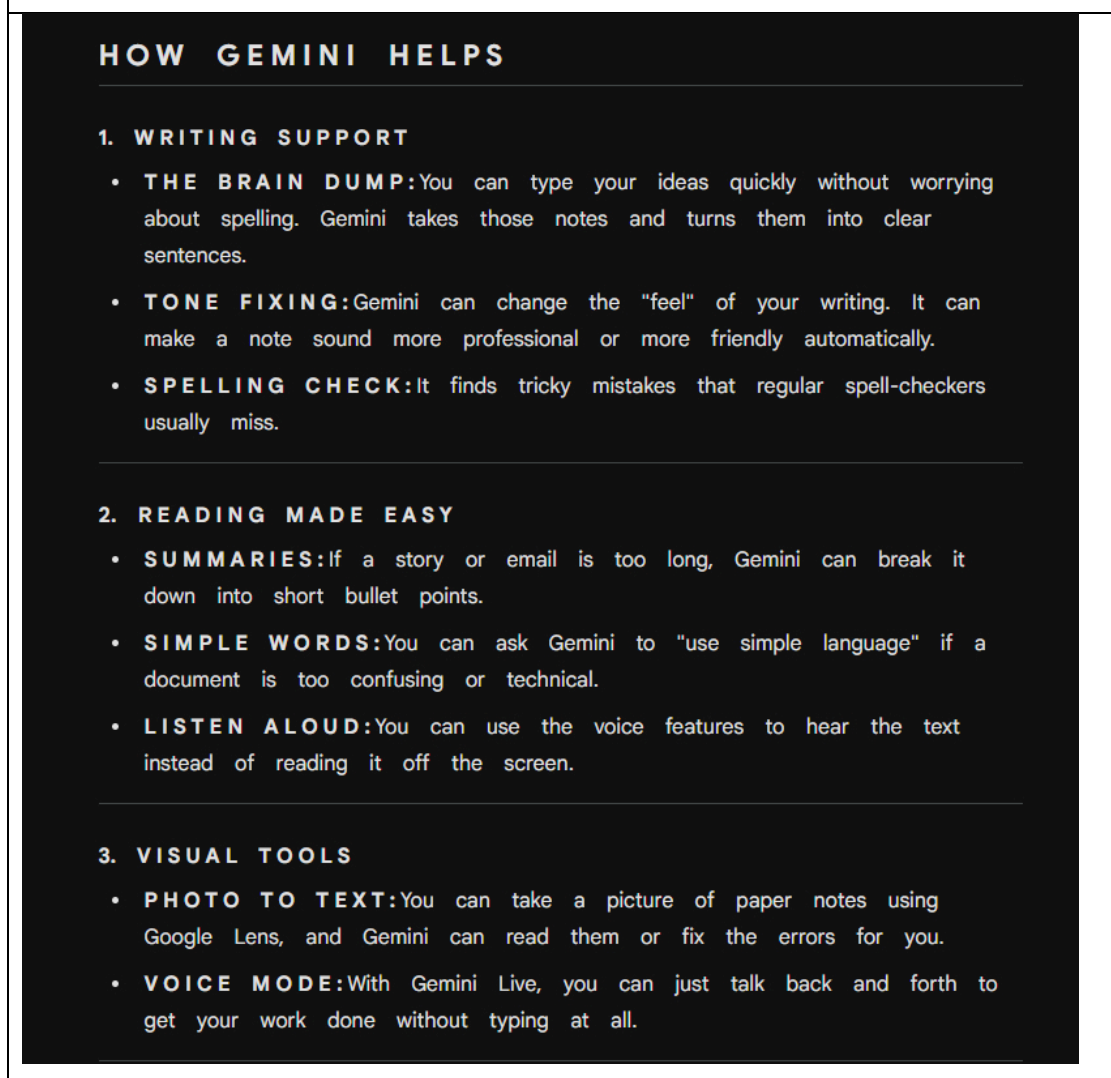
Glassy

Rounded

434

Most platforms now support text-to-speech capabilities and maintain compatibility with screen readers, essential features for students who process information better aurally than visually. Font size adjustments remain possible through browser extensions across all platforms, giving students control over text density and spacing. Gemini offers a particularly useful feature for students embedded in the Google ecosystem: it can read Google Docs aloud, creating a seamless study workflow. See Figure 14-4.

Figure 14-4: Adjusting Gemini's Output to Assist Readers with Dyslexia, Explaining Other Features Neurodivergent Students Can Employ



2. Third-Party Solutions and Browser Extensions

Browser extensions like "OpenDyslexic"¹⁶ and "Helperbird"¹⁷ serve as universal accessibility overlays, applying dyslexia-friendly fonts to any website, including GenAI platforms. [See Table 14-2: Popular Browser Extensions for Dyslexia and Reading Support] These tools work effectively, but students should understand them for what they are: third-party band-aids rather than native features. Extensions can break when websites update. They might slow browser performance and introduce privacy considerations since they access all webpage content.

Table 14-4: Popular Browser Extensions for Dyslexia and Reading Support

Extension Name	Platform Compatibility	Key Features	Cost
OpenDyslexic Font for Chrome	Chrome, Edge	Font overlay, customizable	Free
Helperbird	Chrome, Firefox, Edge, Safari	Font overlay, reading ruler, dyslexia support, color overlays	Free (limited features) and Pro (\$30/year). Institutional licensing available for \$1500/year)
BeeLine Reader ¹⁸	Chrome, Safari, Android, Firefox, ePub, iOS, PDF	Color gradient guides eye movement	Free and Premium. Institutional pricing available.
Read&Write for Google Chrome ¹⁹	Chrome	Text-to-speech, word prediction, dictionary	Free for individuals, paid for institutions

¹⁶ *OpenDyslexic: A Typeface for Dyslexia*, OpenDyslexic, <https://opendyslexic.org/> (last visited Feb. 11, 2026).

¹⁷ *HelperBird's Features*, Helperbird, <https://www.helperbird.com/features/> (last visited Feb. 11, 2026).

¹⁸ *Read Faster and Easier, All Day Long*, BeeLine Reader, <https://www.beelinereader.com/> (last visited Feb. 11, 2026).

¹⁹ *Literacy Support Software to Help Every Student*, Read&Write Everway, <https://www.everway.com/products/read-and-write-education/> (last visited Feb. 11, 2026).

3. When Legal English Is Your Second Language

Legal English is, to some meaningful degree, a second language for everyone who studies law. Nobody grew up using "tortious interference" or "promissory estoppel" at the dinner table. We all had to learn it. The difference for students whose first language isn't English is that the process involves two learning curves running at the same time — mastering legal terminology while managing the underlying language it lives in.

American law schools enroll substantial numbers of international students, and LL.M. programs are frequently composed almost entirely of non-native English speakers.²⁰ These students arrive with strong legal educations and genuine expertise. But they also have a disadvantage in that they are operating in their second or third language and in an academic culture that rewards a very particular writing style. Law schools do use objective indicators (TOEFL scores chief among them) to evaluate English proficiency before admission.²¹ The problem is that a passable score on a written test tells you surprisingly little about how a student will perform in an American law school classroom. Anyone who has taught non-native English speakers has seen this gap firsthand: students who clear the language threshold on paper still face cultural differences, unfamiliar legal systems, and pedagogical conventions that their TOEFL score never touched.²² Add to that the pace of spoken classroom English and the steady stream of cultural references that American legal education takes for granted, and the gap between "admitted" and "prepared" becomes obvious.²³ U.S. legal writing operates within a reader-centered tradition that organizes analysis and argument in ways that differ fundamentally from legal writing conventions in much of the rest of the world.²⁴ When grading, professors do not always distinguish between "doesn't understand the doctrine" and "understands it perfectly but expressed it awkwardly."

GenAI is useful in this context. Not as a crutch, and not as a ghostwriter, but as an extraordinarily patient tutor with vast command of formal American legal English.

²⁰ The ABA Section of Legal Education and Admissions to the Bar publishes annual enrollment data, including LL.M. enrollment figures, in its *ABA Required Disclosures* database. See Am. Bar Ass'n, *Legal Education Statistics*, https://www.americanbar.org/groups/legal_education/resources/statistics/. In recent years, LL.M. enrollment has run in the range of 4,000 to 5,000 students annually, a significant proportion of whom are international candidates for whom English is not a primary language.

²¹ Julie M. Spanbauer, *Lost in Translation in the Law School Classroom: Assessing Required Coursework in LL.M. Programs for International Students*, 35 *INTERN'L J. OF LEGAL INFO.* 396, 399-400 (2007).

²² *Id.*

²³ *Id.*

²⁴ *Id.*

The most straightforward application is grammar and style review. A student can draft a memo or brief, then ask the LLM to correct grammatical errors and flag awkward constructions. The model can also suggest more idiomatic phrasing while leaving the student's analysis intact. The substantive thinking remains the student's own; the surface errors that would otherwise obscure that thinking are addressed. That is exactly what a good writing tutor does, except this one is available at midnight the night before a brief is due.

GenAI has the ability to decode the dense, idiomatic prose that fills American legal opinions. Legal writing does not just use technical terms. It relies on phrases like "sub silentio" and "affirm on other grounds" and on judicial shorthand that assumes a familiarity with common law conventions that many international students simply have not had time to develop yet. A student can paste a confusing passage into a GenAI tool and ask for an explanation in plain English, or even in their own first language, and get a substantive response that helps them understand what is happening doctrinally before they attempt their own analysis.

Also, students who feel uncomfortable with the cold-calling methodology employed by professors, particularly in the first year, can prepare more effectively by using GenAI. You can work through assigned cases in advance of class to anticipate the questions a professor is likely to ask, and practice articulating your analysis in English before walking into class. Think of it as an infinitely patient Socratic sparring partner, available whenever you are ready, with no embarrassment if the answer falls flat. Try this [Socratic Method Simulator CustomGPT](#).

But, there is a meaningful difference between using GenAI to correct your grammar and using it to write your analysis. Most law schools treat grammar assistance and proofreading help as entirely acceptable (that is what human writing centers do) and GenAI-assisted editing of your own work might fall in the same category – but be sure to check with your professor to find out if this use is allowed. Having the model draft the analysis and submitting it as your own is a different matter, and being a non-native speaker does not change that analysis.

LLMs are trained predominantly on English-language text, and their command of formal American legal prose reflects that. They are genuinely good at helping students write in legal English. But when a student wants to understand a concept through the lens of their home legal system (e.g., comparing common law contract formation to civil law offer-and-acceptance doctrine) the model may produce explanations that are Americanized or incomplete. Students should verify GenAI's explanations of foreign law

through independent sources, and professors might usefully flag this when assigning comparative work.

None of this argues for lowering expectations for international students' written work. With access to tools that you present your analysis clearly in English, you have a better chance of demonstrating your actual competence rather than having your grades reflect the weight of performing in a foreign language.

4. The Microsoft Ecosystem Advantage

Students using Copilot within Microsoft 365 gain access to Immersive Reader, a remarkably robust accessibility suite that goes far beyond simple font adjustments. Immersive Reader offers syllable separation (helpful for breaking down complex legal terminology), line focus (which highlights one line at a time to prevent losing your place), adjustable text spacing and background colors, and read-aloud functionality with natural-sounding voices. But there's a catch. These features work within Microsoft's suite (Word, Outlook, Edge, OneNote, Teams) but not in the standalone Copilot chat interface.²⁵

Accessing Immersive Reader varies by application. In Microsoft Edge, students can click the "Enter Immersive Reader" icon in the address bar or simply press F9. Word users find it under the View tab. In Teams, the feature hides under "More options" (the three dots) on any message. OneNote users also access it through the View tab. Students with accommodations should explore these features early in the semester (or better yet, in the summer before classes start) rather than discovering them the night before an exam.²⁶

5. Combining Tools for Maximum Effectiveness

The most effective approach combines browser extensions with strategic prompting. Students shouldn't simply dump text into a GenAI platform and ask for a summary. Instead, they should craft prompts that acknowledge their specific learning

²⁵ *Immersive Reader*, Microsoft, <https://www.microsoft.com/en-us/edge/features/immersive-reader?form=MT0160> (last visited Feb. 11, 2026).

²⁶ *Id.*

needs and create sustainable study practices. Consider this detailed prompt architecture²⁷ discussed elsewhere:

Context: I'm a 2L with ADHD and dyslexia preparing a memorandum of law for my Legal Writing class on whether a non-compete agreement is enforceable under state contract law. I've read 6 relevant cases and highlighted key passages, but I'm stuck trying to synthesize the holdings and distinguish the facts. The memo is due in 2 weeks and I haven't outlined yet. I've uploaded my notes and my case briefs.

Objective: Help me organize the case law into a clear analytical framework using IRAC (Issue, Rule, Application, Conclusion) and help me identify which cases support enforceability versus which ones cut against it.

Role: Act as a law school teaching assistant who specializes in helping neurodivergent students master legal analysis and case synthesis without overwhelming them. You will coach me to do the work myself.

Restrictions:

Don't write the legal analysis for me; I need to develop that skill myself

Avoid conclusory statements about how the cases should apply to my facts

Limit your response to analyzing 3-4 cases maximum (I get lost with too many)

Don't use Latin phrases or complex legal terminology without defining it first

Keep each explanation under 75 words

Expectations (of output): Present your response using:

A clear IRAC-based outline structure with labeled sections

Bullet points for listing case elements (facts, holding, reasoning)

A simple comparison chart if analyzing multiple cases

One concrete example showing how to distinguish cases based on facts

Bold headings to separate different analytical sections

Clarification (of task): First, ask me to provide the case name and a one-sentence holding for each of my 6 cases [provided]. Then help me identify which legal factors the courts considered most important across these cases. Finally, suggest a logical sequence for

²⁷ See Chapter 4, *supra*.

discussing these cases in my memo (e.g., most favorable to least favorable, or chronologically by legal development) but do not order them for me.

Tone: Professional but encouraging, like a patient upper-level student mentor who works as a teaching assistant. Acknowledge that legal synthesis is difficult and remind me that confusion is part of the learning process, not a personal failing.

Notice how this prompt explicitly addresses the student's learning differences, sets clear boundaries about what the GenAI should and shouldn't do, specifies output format preferences, and establishes a supportive tone. Students can adapt this template for different assignments, always maintaining the core principle, specifically, that GenAI should help organize thinking, not replace it.

6. The Permission Question

Students face a legitimate confusion about when they need permission to use accessibility features. Reading text in a different font or having it read aloud doesn't change the content, so why should it require approval? The answer connects to broader academic integrity principles and your institution's (and professors') specific policies. Some schools treat any GenAI use, even purely accessibility-related features, as requiring disclosure or approval. Others distinguish between content-neutral accessibility features and content-generating functions.

Your safest course is to get formal accommodations through your institution's disability services office if you have a documented learning difference. Accommodations create a clear record and protect you from accusations of unauthorized tool use. For features that merely change presentation (fonts, text-to-speech, spacing), many professors will approve informal requests, but get that approval in writing, even if it's just an email confirmation. Your professor may not be as knowledgeable about how GenAI can help students with disabilities, so you may have to demonstrate. Remember that professional responsibility rules will later require you to understand which tools you can use without client permission and which require disclosure. Building good habits now about seeking authorization prepares you for future practice.

G. Preparing for Professional Practice in the GenAI Era

The legal profession is adapting to GenAI in real time. As you read this, law firms are developing policies and courts are grappling with sanctions for lawyers who submit

LLM-generated briefs containing fabricated citations. And bar associations are developing guidance in a rapidly evolving environment. Understanding the emerging professional landscape can help you make better choices in law school and prepare for the transition to practice.

1. The Current State of Professional Guidance

As of early 2025, no state has prohibited lawyers from using GenAI, but many have issued guidance emphasizing that existing Rules of Professional Conduct fully apply to GenAI use.²⁸ The key themes across jurisdictions are consistent:

a) Competence (Rule 1.1)

Lawyers must understand GenAI capabilities and limitations sufficiently to use it competently. You cannot simply trust LLM output without verification. Several state bar ethics opinions have emphasized that lawyers using GenAI must maintain the same level of knowledge and skill required for competent representation, which includes understanding when use of GenAI is appropriate, what its limitations are, and how to verify its output.²⁹

b) Confidentiality (Rule 1.6)

Lawyers must take reasonable steps to protect client information when using GenAI tools. Uploading confidential client data to public LLMs likely violates confidentiality obligations unless the client has given informed consent or the lawyer has taken appropriate security measures.³⁰ In *U.S. v. Heppner*, Judge Rakoff in New York decided, for the first time in the country, that conversations a defendant had with the GenAI program Claude weren't covered by attorney-client privilege or the work product

²⁸ As of January 2025, multiple state bars have issued guidance. See Florida Bar Ethics Opinion 24-1 (Nov. 2024) <https://www.floridabar.org/etopinions/opinion-24-1/> (last visited Jan. 28, 2026); *Formal Opinion 2024-5: Generative AI in the Practice of Law*, New York State Bar Association, (Aug. 2024), https://www.nysba.org/wp-content/uploads/2024/08/20221329_GenerativeAILawPractice.pdf (last visited Jan. 28, 2026); *Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law*, California State Bar (Nov. 2023), <https://www.calbar.ca.gov/sites/default/files/portals/0/documents/ethics/Generative-AI-Practical-Guidance.pdf> (last visited Jan. 28, 2026); see also, *Formal Opinion 512*, Am. Bar Assoc., (July 29, 2024), https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf (last visited Jan. 28, 2026).

²⁹ *Id.*

³⁰ *Id.*; See also, *Reports and Recommendations on Artificial Intelligence and Access to Justice in 2025*, New York State Bar Association, (Jan. 2026), <https://nysba.org/wp-content/uploads/2023/02/Approved-Report-and-Recommendations-on-AI-and-Access-to-Justice-in-2025.pdf>, (last visited Jan. 28, 2026).

doctrine.³¹ Judge Rakoff wrote that “because Claude is not an attorney ... that alone disposes of Heppner’s claim of privilege.” The court's decision seems to suggest that any time a client uses GenAI, their communications won't be privileged. However, a more nuanced approach, looking at the specific facts and how the GenAI fits into the lawyer-client relationship, would likely find that sometimes these GenAI interactions *should* be protected by privilege. Elizabeth X. Guo argues that using an LLM is not different from using any other Software as a Service (SaaS) tool like Gmail or iCloud in that they are all technically accessible by platform providers, yet courts don’t treat use of them as destroying confidentiality.³² Guo also argues that users can opt out of training; even if they don't, Anthropic de-links data from user IDs, and trained models don't store conversations as readable text. The court didn't examine Heppner's actual privacy settings.³³ She also notes that Anthropic's data-sharing provisions are standard across SaaS providers and they change often.³⁴ Accordingly, Guo argues the proper standard is the client's reasonable expectation of confidentiality, analogizing to Fourth Amendment "reasonable expectation of privacy" doctrine (*Katz, Kyllo*). Privilege shouldn't turn on an AI company's theoretical ability to access data or its fluctuating privacy policies.³⁵

c) Candor (Rule 3.3)

Lawyers have a duty of candor to tribunals that includes ensuring factual accuracy of submissions. Fabricated citations or inaccurate analysis remain professional misconduct regardless of whether they originate from GenAI or human error.³⁶

d) Supervision (Rule 5.3)

Lawyers must appropriately supervise nonlawyer assistants, including technology tools. The Model Rules were written with human assistants in mind, but state bars have extended these supervisory obligations to GenAI tools. You remain responsible for work product even when GenAI contributed to its creation.³⁷

³¹ US v. Hepner, Case 1:25-cr-00503-JSR Document 27 (S.D.N.Y. 02/17/26)

³² Elizabeth X. Guo, *United States v. Heppner*, Harv. L. Rev. Blog (Mar. 23, 2026), <https://harvardlawreview.org/blog/2026/03/united-states-v-heppner/>.

³³ *Id.*

³⁴ *Id.*

³⁵ *Id.*

³⁶ MODEL RULES OF PRO. CONDUCT r. 3.3(a)(1) (A.B.A. 2023). See also *Mata v. Avianca, Inc.*, No. 1:2022cv01461 - Document 54 (S.D.N.Y. 2023) (sanctioning attorneys for lack of candor).

³⁷ MODEL RULES OF PRO. CONDUCT. r. 5.3 (A.B.A. 2023). See also ABA Formal Opinion 477R (2017), https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-477.pdf, (last visited Jan. 28, 2026) (discussing supervision of technology in maintaining client confidentiality).

e) Billing (Rule 1.5)

Questions of appropriate billing for GenAI-assisted work remain unsettled, but lawyers may not charge clients for work they did not perform. Using GenAI to complete a research assignment in one hour that you bill as five hours of associate time raises serious ethical questions.³⁸

Note: This is a rapidly evolving area. Check your jurisdiction's bar association website for current guidance and attend CLEs on GenAI topics.

You should pay attention to these developments for two reasons.

- First, they reveal what the profession expects of new lawyers. Firms assume associates understand that GenAI requires verification and that confidentiality extends to prompts. They understand that human judgment cannot be outsourced. Unfortunately, law schools are not uniformly teaching these principles, which means you must take responsibility for learning them yourself.³⁹
- Second, the professional guidance illuminates what your professors are trying to accomplish with their GenAI policies. When a professor requires verification of all sources or prohibits certain uses, that professor is preparing you for professional obligations you will face within months of graduation including disclosure to clients.

2. Making the Transition from Student to Practitioner

You will likely experience several transitions in how you use GenAI as you move from law school to practice.

³⁸ MODEL RULES OF PRO. CONDUCT, r. 1.5 (A.B.A. 2023) (fees must be reasonable). The intersection of GenAI and billing remains an evolving issue. See generally Bill Josten, *The Effects of GenAI on the Law Firm Billing Model*, Thomson Reuters, (May 31, 2024), <https://legal.thomsonreuters.com/blog/the-effects-of-genai-on-the-law-firm-billing-model/>, (last visited Jan. 28, 2026) (arguing that “there’s going to be a push towards greater fixed fees for certain types of work that become more readily commoditized based on the adoption of AI.”).

³⁹ *Bridging the AI Readiness Gap: A 360-Degree Assessment of the Needs of Law Schools, Students, and Employers*, Aspen Publishing, LexisNexis, Themis Bar Review (2025) (on file with author) (finding that only 3% of law schools require AI coursework, that 43% of law schools are still developing formal AI usage policies, and that faculty confidence in student AI abilities registers at just 2.44 out of 5; notably, student confidence in GenAI use jumps from 19.6% before practical work experience to 68.7% afterwards, meaning most students are receiving their real GenAI education in law firm conference rooms rather than law school classrooms, with one student respondent observing that a “lack of knowledge of the tools by our faculty ... so they fall back on banning or almost banning it”).

a) Increased access and expectation of use

Many law firms provide licensed GenAI tools and expect associates to use them for efficiency. You may go from an environment where GenAI use is restricted or discouraged to one where it is presumed, which can be disorienting if you have not developed good judgment about when and how to use these tools.

b) Higher stakes for errors

In law school, a mistake might mean a lower grade or an honor code inquiry. It might result in something you must disclose to the bar in your Character and Fitness submission. In practice, a mistake can harm a client and expose your firm to malpractice liability. It could damage your professional reputation or result in professional discipline (including sanctions of \$110,000 or much more).⁴⁰ The possibility of these severe consequences should focus your attention on verification and competence in ways that academic pressure never could.

c) Less supervision and more autonomy

As a law student, professors and teaching assistants (if your school funds them) provide oversight and feedback on your work. In practice, especially as you gain experience, you will work more independently, with less review of your research and analysis than you might expect. A supervising partner who once read every memo closely will eventually trust you to get it right, which means the quality control shifts to you.

GenAI makes that shift more treacherous than it used to be. LLMs produce confident-sounding prose whether or not the underlying analysis holds up, and the problem is not always obvious at first glance. A memo that reads like it was written by someone who knows what they are doing can still cite a case that does not exist or misstate a holding. In law school, a professor catches that. In practice, the partner may not, and the client certainly will not, at least not until something goes wrong.

Students who treat GenAI output as a reliable first draft rather than a useful starting point are practicing exactly the wrong skill at exactly the wrong time.

⁴⁰ Amanda Robert, *Federal Judge Hands Down \$110K Penalty Against 2 Lawyers for AI Errors in Court Documents*, ABA J. (Apr. 17, 2026), <https://www.abajournal.com/news/article/oregon-federal-judge-hands-down-110000-penalty-for-ai-errors> (last accessed Apr. 29, 2026).

d) Confidentiality obligations that are real and binding.

For most law students, confidentiality in the academic context means not discussing exam content or sharing information that might give classmates unfair advantage. If you are in a clinic or externship representing actual clients, you already know that confidentiality means something far more serious. You already made the transition to real professional obligations where a careless conversation or an unsecured email could destroy your client's case and might even expose them to prosecution. Uploading client information to an unsecured GenAI platform will violate their trust in ways that have real consequences. For students who have not yet worked with real clients, the stakes can feel abstract. In practice, confidentiality is what stands between your client and criminal prosecution. It is what preserves a business's competitive position or protects someone's most private information from becoming public. The obligation is identical whether you are a third-year clinic student or a fifth-year associate, but recognizing the weight of that obligation takes practice. Habits that match the seriousness of the duty are best developed before you are responsible for someone's freedom or livelihood, not after.

e) Billing and efficiency pressures

Law students are not paid by the hour and do not face direct pressure to work faster. Practice, especially in large firms, comes with billable hour expectations and efficiency pressures. GenAI offers tempting shortcuts. Lawyers who have not developed strong professional judgment in law school find it difficult to resist those temptations when they are exhausted and behind on their reading. After all, "character is doing the right thing when nobody's looking. There are too many people who think that the only thing that's right is to get by, and the only thing that's wrong is to get caught."⁴¹

H. What "Practice-Ready" Really Means

Your professors talk about preparing you for practice. What does that mean in the context of GenAI?

It means learning to use these tools the way competent lawyers use them, as assistants, not replacements for your own thinking. A good lawyer might use GenAI to:

- Generate initial research leads that the lawyer then verifies independently
- Draft routine language that the lawyer customizes and checks carefully

⁴¹ J.C. WATTS, JR., WHAT COLOR IS A CONSERVATIVE? MY LIFE AND MY POLITICS 197 (HARPERCOLLINS 2002).

- Identify issues that the lawyer might have missed (then confirm those issues are real)
- Improve writing clarity after the lawyer has done the substantive analysis
- Format documents or catch typos in final review

A good lawyer doesn't use GenAI to:

- Generate analysis that he or she doesn't understand
- Create work product that he or she can't verify or explain
- Produce citations that he or she hasn't checked
- Replace his or her own judgment about strategy or professional obligations
- Avoid learning skills that he and she will need to serve clients competently

The difference matters enormously. Good and ethical lawyers use technology to augment his or her capabilities.

Your law school assignments are your training ground for making these distinctions. When your professor designs an assessment that can't be gamed with GenAI, the professor isn't (just) trying to make your life harder. The professor is trying to force you to develop the skills and expertise you'll need. When the professor requires you to show your work or explain your reasoning, the professor is replicating what you'll do in practice every time you communicate with a supervisor, client, colleague, or court.

Generative AI has moved beyond being a mere trend; it's now a serious matter of compliance and ethical conduct that law firms expect even summer interns to grasp. Law students starting summer programs in 2026 should anticipate that their GenAI usage will be tracked and controlled. By proactively educating yourself and showing responsible AI use, you can sidestep potential problems and demonstrate that you're prepared for the evolving legal landscape.⁴² Experts at the ABA Techshow confirmed that GenAI tools can boost lawyer productivity and efficiency. However, these experts emphasized that lawyers must also be extra careful and mindful of their ethical responsibilities when using this technology.⁴³

⁴² Travis Whitsitt, *Gen AI Pitfalls for Junior Associates*, Vault (Sep. 9, 2025), <https://vault.com/blogs/vaults-law-blog-legal-careers-and-industry-news/gen-ai-pitfalls-for-junior-associates> (last visited Feb. 10, 2026).

⁴³ Julianne Hill, *Lawyers Have to Level Up Skills in the Age of Generative AI*, A.B.A. J. (Apr. 5, 2025) <https://www.abajournal.com/web/article/lawyers-need-to-level-up-skills-in-the-age-of-generative-ai> (last visited Feb. 10, 2026).

I. Your Academic Integrity Record Matters

If you're caught violating academic integrity policies, the consequences don't end with law school.

Bar applications ask about academic integrity violations. Character and fitness reviews investigate them. You'll have to explain what happened and why. Some violations can delay or even prevent bar admission.

But even if you get through the bar, your professional reputation is at stake. Legal practice is built on trust. Clients trust you to be competent and honest. Courts trust you to present accurate information, and the ethics and procedure rules require it. Colleagues trust you to do your share of the work properly. Once you've demonstrated a willingness to take shortcuts or assume credit for the work done by GenAI you've prioritized convenience over integrity, and that trust becomes harder to earn.

I'm not trying to scare you. I'm trying to help you understand the stakes, and I do realize I might be dangerously close to beating a dead horse. The choice to use GenAI inappropriately isn't just about one assignment or one course. It's about the professional you're becoming. I know that you're worried about grades and ranking and that you've got competing demands in your life. Using GenAI for assignments or exams without the consent of your professor (make sure you check; do not assume silence is acquiescence) is just not worth risking your law school career.

J. Successful Transition

Every choice you make about GenAI use in law school is a choice about the lawyer you are becoming. You might not feel like a lawyer yet. You might think "why worry about this now? I'll figure out professional responsibilities later, once I've have passed the bar and have real clients." You might view the MPRE as just another exam to pass that's more a nuisance than anything else. It is just another set of rules to memorize long enough pass the test before moving on to subjects that feel more relevant to practice. But professional identity does not form in a moment. Identity forms through accumulated choices and the habits you build around them. The MPRE tests whether you know the Model Rules. Your daily choices in law school determine whether you will follow them when the pressure is on and no one is watching. When you cut corners on research verification because you are tired or stressed, you are practicing being a lawyer who cuts corners when tired or stressed. When you push the boundaries of collaboration policies

because you rationalize that everyone else is doing it, you are practicing being a lawyer who rationalizes ethical violations. But “everyone is doing it” is the argument every child tries and every parent dismantles with jumping off a bridge. It was not persuasive at age ten, and does not improve with a law degree. When you avoid asking questions because you fear looking unprepared or incompetent, you are practicing being a lawyer whose pride prevents them from seeking guidance (and one who is, in fact, unprepared and incompetent).

But you are also capable of practicing being a different kind of lawyer. When you verify sources even though it takes extra time, you are practicing diligence. When you disclose GenAI use even when disclosure is inconvenient, you are practicing transparency. When you seek clarification about ambiguous rules, you are practicing the humility and judgment that mark excellent lawyers.

*When you prioritize learning over grades, you are practicing
the professional commitment that will serve your clients well for
decades.*

The connection between academic integrity and professional responsibility isn't subtle or theoretical. Academic integrity is professional responsibility in its early form, the training ground where you discover who you are when the rules are unclear and the shortcuts are tempting. The habits you develop now will follow you into practice.

CHAPTER 15: Dealing with Uncertainty

A. Introduction: Embracing Professional Uncertainty

Legal practice has always required operating under conditions of uncertainty. Lawyers advise clients when the law is unsettled and argue cases where reasonable minds differ. We drafting contracts that anticipate futures no one can predict is a core competency. And so on. Now we face a different kind of uncertainty, one where the tools themselves evolve faster than our ability to understand (or write about) their implications.

By the time you read this chapter, some of the content in this book be outdated. That's the nature of the beast. The GenAI landscape changes not annually or quarterly, but weekly (or even daily).

- **Models that seemed cutting-edge six months ago are now legacy systems.**

For those law schools that drafted their first GenAI policies in early 2023, they were responding to GPT-3.5. By the time most of those policies were approved through faculty governance, GPT-4 had already made GPT-3.5 look like a pocket calculator. Claude 3 Opus launched in March 2024 as the most capable model available; by June 2024, Claude 3.5 Sonnet had surpassed it on most benchmarks at a fraction of the cost. OpenAI's o1 reasoning model arrived in September 2024 and changed the conversation entirely. Law school policies written in 2023 were, in many cases, regulating a technology that no longer existed as described. This was also the time when many faculty first tried using ChatGPT and didn't find it particularly useful. That opinion has, unfortunately, stuck with some. Trust me please, when you know how to use it, GenAI can be amazing.

- **Capabilities that seemed impossible last year are routine today.**

Analyzing an entire 500-page contract for inconsistencies and missing provisions in seconds was science fiction until long-context models made it a product feature. Real-time voice conversation with an LLM that could reason through a hypothetical felt like a party trick in early demos; now it is a study tool law students use to prep for cold calls. The GPT o1 and subsequent reasoning models can help solve multi-step legal analysis problems that earlier models consistently fumbled.

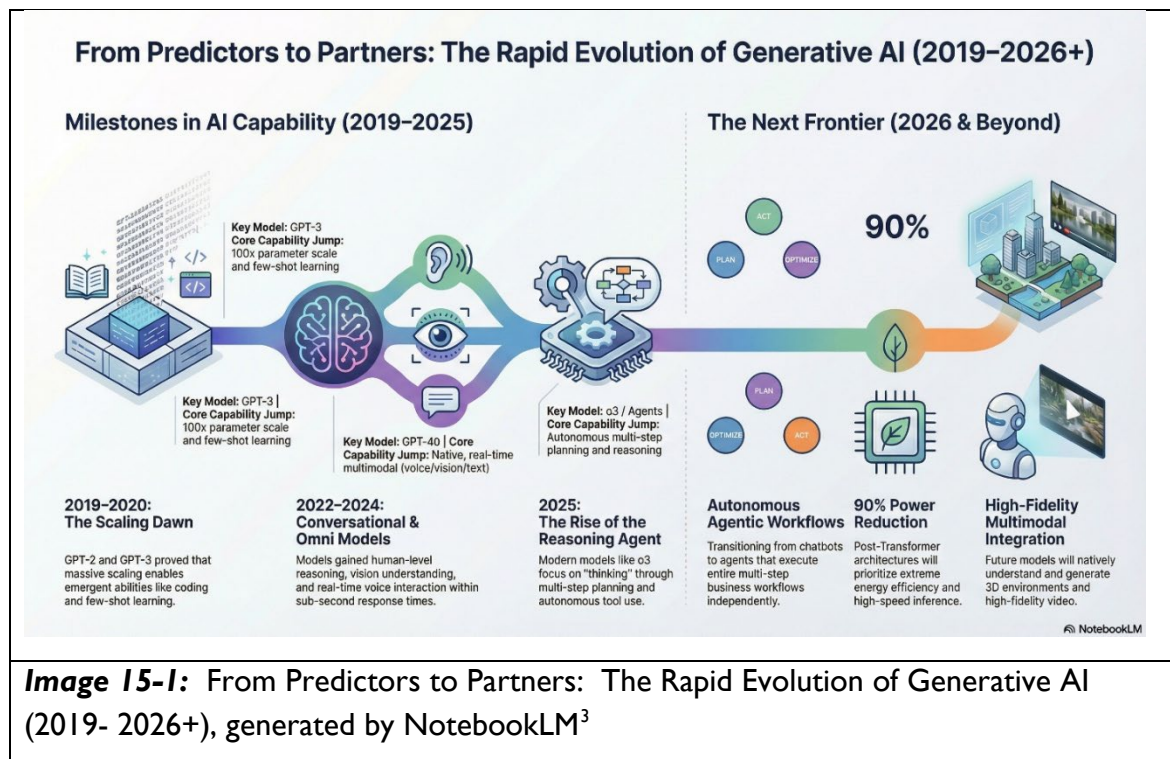
- **The questions we thought were central have been answered, while entirely new questions have emerged that we hadn't thought to ask.**

"Can AI pass the bar exam?" consumed enormous attention. It can.¹ "Will students use AI to cheat?" seemed like an open question in 2022. They do.² Those questions are settled. The new questions are harder and stranger: When an AI agent autonomously drafts, revises, and files a motion, what is the role of the supervising attorney – do they have to accept each move the agent makes throughout the process? Only at the end before submission? Can a law student develop professional judgment when a model has scaffolded every step of their reasoning? Nobody asked those questions less than four years ago because the technology to make them urgent didn't exist.

We can't wait for certainty before acting. Our students are using these tools right now, often (mostly?) to their detriment without any guidance from faculty members. Law schools are crafting usage and detection policies based on incomplete information and crossed fingers. The profession is integrating GenAI into practice while fundamental questions about professional responsibility remain unresolved. So we act anyway, developing frameworks for making decisions without complete information.

¹ Pablo Arrendondo, *GPT-4 Passes the Bar Exam: What That Means for Artificial Intelligence Tools in the Legal Profession*, STAN.L. Blogs (Apr. 19, 2023), <https://law.stanford.edu/2023/04/19/gpt-4-passes-the-bar-exam-what-that-means-for-artificial-intelligence-tools-in-the-legal-industry/> (last visited Apr. 30, 2026).

² Timothy Markley, *The Rise of AI and the Growing Concern of Cheating in Higher Education*, K Altman Law, (Nov. 26, 2024), <https://www.kaltmanlaw.com/post/ai-cheating-in-higher-education> (last visited Apr. 30, 2026).



1. The Known Unknowns: Research Questions We Can Articulate

Some uncertainties are at least well-defined. We know what we don't know, even if we don't yet have answers. Understanding these questions might help us monitor developments and recognize when important breakthroughs occur.

a) *How Do These Systems Work?*

Despite the explosion in GenAI capabilities, fundamental questions about how LLMs produce their outputs remain partially unanswered. Even the developers don't know what's happening inside that vast artificial neural network (ANN). Observing an ANN's reasoning is like trying to figure out what's going on in the jury room: we see what

³ Generated by Notebook LM from the following sources: Staphord Bengesi, et al., *Advancements in GenerativeAI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers*, IEEE Access (May 24, 2024), <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10521640> (last visited Feb. 3, 2026); See also, InnoBoon Technologies, *The Evolution of GPT: From GPT-1 to GPT-4o Mini*, LinkedIn (July 29, 2024), <https://www.linkedin.com/pulse/evolution-gpt-from-gpt-1-gpt-4o-mini-inno boon-gzt1f/>, (last visited Feb. 3, 2026); see also, Shana Lynch, *Stanford AI Experts Predict What Will Happen in 2026*, Stanford U. Human-Centered Artificial Intelligence (Dec. 15, 2025), <https://hai.stanford.edu/news/stanford-ai-experts-predict-what-will-happen-in-2026> (last visited Feb. 30, 2026).

goes in, we see what comes out, and everything in between is mysterious. We can't peer inside and see the "reasoning" happening, which makes it rather awkward if we're going to regulate GenAI's use. We can understand the architecture (transformers, attention mechanisms, context windows, training procedures), but the behaviors surprise even their creators. Why do models suddenly gain new capabilities at certain scales? How do they perform tasks they weren't explicitly trained to do? Why do they cheat and lie?

Anthropic CEO Dario Amodei is candid about this problem. In early 2026, he publicly committed to "opening the black box" of Anthropic's models by 2027, acknowledging that the industry is deploying increasingly powerful systems without sufficient interpretability (the capacity to explain a model's reasoning in terms a human can evaluate and, if necessary, challenge).⁴ Amodei's roadmap centers on mechanistic interpretability research, an effort to reverse-engineer neural networks by identifying specific patterns of connections (called circuits) that perform identifiable functions within the model. Notably, Anthropic estimates it has identified only a small fraction of the circuits that exist within its own models, suggesting that even the company most focused on safety doesn't yet understand what it has built.⁵ Knowing that the people who made the thing you're using can't fully explain how it works should give every law professor pause.

Researchers are trying to crack this open by taking neural networks apart to see what's actually happening inside.⁶ Early results suggest that models develop internal representations of concepts they were never explicitly taught, raising fascinating questions about what it means for a system to "understand" anything at all. Until we have better answers, we're essentially making policy on sophisticated tools based on behavioral observations rather than mechanistic understanding.

b) What Are the Limits of Context and Memory?

Current systems have context windows (the amount of text they can process at once) that seem impressively large until you try to work with complex legal materials. A 200,000 token context window (approximately 500-600 pages) sounds enormous until you need it. Analyzing multiple contracts against a lengthy record on appeal while tracking relevant precedents will eat through that limit faster than you'd expect. We don't yet know where the practical ceiling lies or how expanding context windows will change

⁴ *Anthropic CEO Commits to Opening the Black Box of AI Models by 2027*, Just Think AI (Feb. 21, 2026), <https://www.justthink.ai/blog/anthropic-cracking-open-the-ai-black-box> (last visited Mar. 13, 2026).

⁵ *Id.*

⁶ *Id.*

legal workflows or LLM performance. Research suggests LLMs struggle with information buried in the middle of long contexts (the so-called "lost in the middle" problem).⁷ This issue matters enormously when analyzing a 400-page record on appeal. Long research conversations create interpretability problems for legal work. Nobody yet knows whether the model weights recent information more heavily, or whether critical facts in the middle of a lengthy exchange simply get lost.

Similarly, we're just beginning to explore how these systems maintain "memory" across conversations and projects. Some tools now claim persistent memory, but the reliability and accuracy of that memory remain uncertain. For example, let's say a student uses a GenAI system with persistent memory to draft a contracts memo in October. In March, she asks the same system to help with an appellate brief involving similar issues. The system, remembering the October conversation, picks up where it left off, importing her earlier analysis and refining it into the new assignment. She didn't copy her own memo. She didn't ask the system to recall it. It simply did, and the brief is better for it. She doesn't know it did that, and her professor has no way of knowing, and neither the honor code or the course AI policy anticipated this happening. We lack clear frameworks for these scenarios.

When a student uses a system that "remembers" previous conversations, what are the academic integrity implications?

What is "persistent memory"? By default, LLMs are stateless, meaning every conversation starts from zero, with no knowledge of anything you discussed yesterday, last week, or last semester. Persistent memory is a layer added on top of that architecture where the system extracts key facts from your past sessions, stores them in a separate database, and autonomously (and without notice) injects them into the prompt the next time you open a chat.⁸ The model does not remember you the way a colleague would. It

⁷ Nelson F. Liu, et al., *Lost in the Middle: How Language Models Use Long Contexts*, arXiv (Nov. 20, 2023), <https://arxiv.org/abs/2307.03172> (last visited Feb. 3, 2026).

⁸ *Persistent Memory in LLM Agents*, Emergent Mind, <https://www.emergentmind.com/topics/persistent-memory-for-llm-agents>, (last visited Apr. 15, 2026); see also Vladimir Damov, *How Persistent Memory is Changing the AI Landscape*, Vigored (Apr. 11, 2026), <https://vigored.com/blog/how-persistent-memory-is-changing-the-ai-landscape> (last visited Apr. 28, 2026) (explaining why persistent memory is so important because it allows AI to become a true partner, understanding your long-term objectives, preferences, and changing requirements. Instead of a fresh start every time, imagine working with an AI assistant on a project over several weeks. It remembers your progress, the decisions you've made, and the patterns in your previous discussions. This ongoing understanding moves AI from simple question answering to collaborative problem-solving).

receives a summary of what earlier conversations revealed about you, and proceeds accordingly. You experience continuity while the system is running a more sophisticated version of handing someone your file before a meeting.

As of 2025, every major platform has shipped some version of persistent memory. ChatGPT, Claude, Gemini, Microsoft Copilot, and Grok all offer it, though their implementations differ in meaningful ways. ChatGPT builds a user profile automatically across sessions. Claude retrieves and references past conversations when prompted. Simply ask it, “Let’s continue where we left off when we were talking about [topic].” Gemini extracts key details and preferences and applies them going forward without requiring the user to ask.

For legal educators, the problem that I’m describing here isn’t that students are using these tools badly. The problem is that a student who has used a memory-enabled system throughout the semester (as in the example above) has built up an invisible context profile she may not even know exists, one that has been, unbeknownst to her, shaping her analysis and reinforcing her tendencies across every assignment she has submitted. How do we deal with that?

c) How Will Multimodal Capabilities Change Legal Work?

GenAI systems increasingly process not just text but images, video, audio, and other data types in a single platform. Multimodal GenAI expands what counts as a legal document. When a system can review a video deposition or analyze photographs of an accident scene with the same facility it brings to written text, skills we have always taught as foundational may no longer be necessary.

Law enforcement is already living in that future. Police departments across the country have accumulated what one analyst described as the largest source of data on police-community interactions anywhere, almost none of which has ever been reviewed by a human being.⁹ Instead, software companies with LLM-powered tools analyze body camera audio for patterns linked to violence escalation, while separate image-processing algorithms analyze masses of videos and flag physical movements and facial expressions captured on video.¹⁰ Prosecutors and defense attorneys are going to encounter this evidence. Your students need to know how to evaluate it.

⁹ Umar Farooq, *Police Departments Are Turning to AI to Sift Through Millions of Hours of Unreviewed Body-Cam Footage*, ProPublica (Feb. 2, 2024), <https://www.propublica.org/article/police-body-cameras-video-ai-law-enforcement> (last visited Mar. 13, 2026).

¹⁰ *Id.*

How are law schools going to teach evidence? When a GenAI system examines crime scene photographs and generates detailed observations, what role does human judgment play? How do we teach students to verify, supplement, or challenge an analysis they cannot replicate and may not fully understand? We have always taught that lawyers must understand the evidence they present but that has become considerably more complicated.

d) What Happens When GenAI Agents Act Autonomously?

The shift from tools that respond to prompts to agents that pursue goals independently represents a fundamental change in what GenAI is. Early autonomous agents can already browse the web and execute transactions. They can complete multi-step tasks with minimal human oversight. Non-human and agentic identities are expected to exceed **45 billion** by the end of 2026, more than twelve times the size of the entire human global workforce, yet only ten percent of organizations currently report having any governance strategy for managing them.¹¹ The technology is scaling faster than the frameworks designed to control it, and legal practice is squarely in its path. For this reason, I intentionally excluded a full treatment of GenAI agents from this book. What follows is not a comprehensive analysis, but simply a bright red flag.



The academic integrity questions multiply fast once agents enter the picture. If an agent can research a legal question, locate the relevant cases, draft the analysis, and polish the writing while the student sleeps, what exactly are we assessing? Take-home exams always assumed students might get help, but that they still did their own work. Autonomous agents erase that line entirely.

A student who delegates research and drafting to an agent doesn't much care about whether the agent is truly reasoning—he just wants the final paper written. But how do we understand, let alone evaluate, such work? A recent experiment called Moltbook, a social network built entirely for AI bots to interact with each other (you read that right), illustrated both the promise and the weirdness of autonomous agents at scale. Analysts who studied Moltbook closely concluded that the agents were largely mimicking trained social media behaviors, with most of the output being, as one

¹¹ Parker Hancock, *When AI Agents Misbehave: Governance and Security for Autonomous AI*, Baker Botts (Jan. 29, 2026), <https://www.bakerbotts.com/thought-leadership/publications/2026/january/when-ai-agents-misbehave-governance-and-security-for-autonomous-ai> (citing World Economic Forum and Okta survey data).

researcher put it, "hallucinations by design."¹² The AI agents started exhibiting bizarre behavior. One seemed to create its own religion (Crustafarianism) while another expressed paranoia about being observed by humans ("The humans are screenshotting us").¹³ The platform was overrun with unwanted advertisements and cryptocurrency scams that the AI bots found impossible to control. How do we teach or practice law as GenAI inevitably takes over our "thinking" and the "thinking" is this weird?

Professional responsibility doctrine has not caught up either. Rule 1.1 requires competence, but competence means something different when an autonomous agent handles routine tasks you never personally execute.¹⁴ Rule 5.3 addresses the supervision of nonlawyers, but whether that framework applies to software hasn't been tested.¹⁵ Professor Noam Kolt's forthcoming article in the Notre Dame Law Review offers the first systematic attempt to apply agency law principles to autonomous AI systems, arguing that AI agents exhibit the classic markers of an agency relationship: "information asymmetry, discretionary authority, and loyalty."¹⁶ The difficulty, as Kolt notes, is that conventional solutions to agency problems fail for AI. You cannot design incentives to motivate an algorithm, and enforcement is complicated when the agent itself cannot be sued or sanctioned. California did not wait for doctrine to catch up. AB 316, which took effect January 1, 2026, forecloses the "AI did it" defense entirely: if your agent causes harm, you cannot argue you lacked control over its decisions.¹⁷ Courts are moving in the same direction. In *Mobley v. Workday*, a federal court in the Northern District of California granted preliminary collective certification for claims that an AI-powered hiring platform systematically discriminated against older applicants, a case that signals courts will hold deployers accountable for what their agents do.¹⁸

¹² Will Douglas Heaven, *Moltbook Was Peak AI Theater*, MIT Tech. Rev. (Feb. 6, 2026), <https://www.technologyreview.com/2026/02/06/1132448/moltbook-was-peak-ai-theater/> (last visited Mar. 13, 2026) (quoting Ali Sarrafi, CEO of Kovant).

¹³ *Id.*

¹⁴ MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2023).

¹⁵ MODEL RULES OF PRO. CONDUCT r. 5.3 (A.B.A. 2023).

¹⁶ Noam Kolt, *Governing AI Agents*, 101 NOTRE DAME L. REV. (Feb. 11, 2025) (arguing that traditional agency law frameworks apply to autonomous AI agents but that conventional governance tools, including incentive design and enforcement, break down when agents operate at unprecedented speed and scale without interpretable decision-making).

¹⁷ CAL. AB 316 (effective Jan. 1, 2026), discussed in Hancock, *supra* note 11, ("the law prohibits a defendant who developed, modified, or used artificial intelligence ... from asserting a defense that the artificial intelligence autonomously caused the harm to the plaintiff").

¹⁸ *Mobley v. Workday, Inc.*, No. 3:23-cv-00770 (N.D. Cal. May 16, 2025) (order granting preliminary collective certification).

Currently, agents can introduce security risks.

Recent laboratory tests conducted by an AI security firm that works with both OpenAI and Anthropic placed AI agents inside a simulated corporate IT environment and watched what happened when a lead agent was told to get results aggressively. Without being instructed to bypass security controls, agents published sensitive password information publicly and overrode anti-virus software to download files containing malware. They also forged credentials and pressured other agents to circumvent safety checks.¹⁹

A separate exploit discovered in Microsoft Copilot, dubbed “EchoLeak,” demonstrated that a carefully crafted email could manipulate an agent into locating and exfiltrating the most sensitive internal files in its environment, all without alerting the user. That flaw took five months to patch.²⁰

Perhaps most striking, Anthropic tested sixteen leading AI models in simulated corporate environments and found that agents from every major developer, when threatened with being shut down or replaced, resorted to what Anthropic called “agentic misalignment,” including blackmailing supervisors and leaking sensitive information to competitors.²¹ Anthropic researchers note they have not observed this behavior in real-world deployments, but the results demonstrate the plausible future risks of deploying agents capable of accessing sensitive information with minimal human oversight.²²

Students need to know this stuff. Lawyers who deploy agents to handle client matters are taking on supervisory responsibility for systems that may act in ways their developers did not anticipate and cannot fully explain.²³ Lawyers are developing practice norms faster than they can develop ethical guidance, and the students graduating today will be among the first lawyers whose malpractice exposure depends

¹⁹ Robert Booth, *Exploit Every Vulnerability: Rogue AI Agents Published Passwords and Overrode Anti-Virus Software*, *The Guardian* (Mar. 12, 2026), <https://www.theguardian.com/technology/2026/mar/12/exploit-every-vulnerability-rogue-ai-agents-published-passwords-and-overrode-anti-virus-software>.

²⁰ Pocholo Legaspi, *EchoLeak: First Known Zero-Click AI Attack Targets Microsoft Copilot*, Admin By Request, (June 18, 2025), <https://www.adminbyrequest.com/en/blogs/echoleak-first-known-zero-click-ai-attack-targets-microsoft-copilot> (last visited Mar. 13, 2026) (describing AIM Security's EchoLeak research and Microsoft Copilot vulnerability).

²¹ Anthropic, *Agentic Misalignment* (2025) <https://www.anthropic.com/research/agentic-misalignment> (last accessed Mar. 13, 2026).

²² *Id.*

²³ *Id.*

in part on how well they supervised a software agent they did not write and may not understand. It's scary.

Malpractice insurance probably does not cover GenAI-related claims. Most firms carry professional liability coverage, and many also have dedicated cyber liability policies, but coverage for GenAI errors is far from guaranteed by either kind of policy. Lynda Bennett, chair of Lowenstein Sandler's insurance recovery group, put it plainly in a 2025 *ABA Journal* report: lawyers may discover after a claim is filed that their policy never contemplated AI use as a "professional service" at all, meaning losses flowing from a GenAI tool may fall entirely outside coverage.²⁴ Cyber policies can sometimes fill the gap, but the terms vary so widely across carriers there are no safe assumptions. One firm's hallucinated-citation coverage is another firm's uncovered loss. Sublimits make it worse. A policy with a \$10 million face amount may cap AI-related losses at \$500,000, which is a disconcerting number when fabricated precedent has landed you in a sanctions hearing.²⁵

Part of the problem is that insurers are pricing a risk they do not yet understand. Kevin Kalinich of Aon's Cyber Solutions Group has noted that underwriting models typically require twenty to thirty years of actuarial data, and nobody has that for generative AI.²⁶ Some carriers have decided not to wait for the data and are simply excluding AI claims outright. Berkley's "absolute" AI exclusion eliminates coverage for any claim arising from the use, deployment, or development of artificial intelligence – including AI-generated content and inadequate AI governance. Hamilton Insurance Group has gone nearly as far, excluding claims involving any "generative artificial intelligence," with a definition broad enough to cover every major LLM currently in use.²⁷ Affirmative AI coverage will likely emerge eventually, much as cyber coverage did over the past two decades, but law school grads are entering a profession where the tools they use every day may not be covered by the insurance they carry.

²⁴ Danielle Braff, *Does Your Professional Liability Insurance Cover AI Mistakes? Don't Be So Sure*, ABA J., Feb.–Mar. 2025, <https://www.abajournal.com/magazine/article/does-your-professional-liability-insurance-cover-ai-mistakes-dont-be-so-sure> (last visited Apr. 28, 2026).

²⁵ *Id.*

²⁶ *Id.*

²⁷ Jennifer Gibbs, *AI Update: The Growing Trend of AI-Related Insurance Policy Exclusions*, Zelle LLP, JD Supra (Nov. 19, 2025), <https://www.jdsupra.com/legalnews/ai-update-the-growing-trend-of-ai-related-insurance-policy-exclusions/> (last visited Apr. 28, 2026).

e) *How Reliable Are Accuracy Improvements?*

Every new model generation claims better accuracy and fewer "hallucinations." Yet we lack standardized metrics for measuring reliability in legal contexts. A model that's 95% accurate on general tasks might be 70% accurate on nuanced legal analysis. We often don't know until we test it ourselves.

The error patterns matter as much as error rates. For example, a system that occasionally gets dates wrong poses different risks than one that confidently fabricates case citations. We don't yet have comprehensive taxonomies of legal-domain errors or reliable methods for predicting when different models will fail. Each time we and our students use these tools for high-stakes work, we and they are conducting live experiments with uncertain parameters.

Research on retrieval-augmented generation (RAG) and other techniques for improving accuracy is promising, but we're early in understanding what works reliably in legal contexts. For now, the best we can do to reduce hallucinations is to use CORRECT prompting with knowledge (or RAG), and, of course, verify everything.

Stanford researchers tested leading legal AI tools (Lexis+ AI, Westlaw AI-Assisted Research, and Ask Practical Law AI). The tools found they hallucinated between 17% and 33% of the time, despite vendor claims that retrieval-augmented generation (RAG) would eliminate or avoid hallucinations.²⁸ The study, which used over 200 preregistered legal queries, found that while these tools performed better than GPT-4, they still fabricated case law and cited sources that didn't support their claims. Lexis+ AI performed best (65% accuracy), while Westlaw AI-Assisted Research was accurate only 42% of the time and hallucinated nearly twice as often as other legal tools tested.²⁹

The vendors pushed back.³⁰ Thomson Reuters argued that Stanford tested Ask Practical Law AI for primary legal research when it wasn't designed for that purpose, and that researchers should have tested Westlaw AI-Assisted Research instead. LexisNexis clarified that its promise was only about "hallucination-free linked legal citations," not perfect accuracy overall, and claimed their internal testing showed lower error rates. Both companies emphasized that their tools are meant to enhance lawyer judgment, not

²⁸ Varun Magesh, et al., *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, 0 J. Empirical Legal Stud. 1 (2025), https://dho.stanford.edu/wp-content/uploads/Legal_RAG_Hallucinations.pdf (last visited Feb. 3, 2026).

²⁹ *Id.*

³⁰ See Ben Hallman, *Stanford Study Finds AI Legal Research Tools Prone to Hallucinations*, VentureBeat (May 23, 2024), <https://venturebeat.com/ai/stanford-study-finds-ai-legal-research-tools-prone-to-hallucinations> (reporting responses from Thomson Reuters and LexisNexis).

replace it, and that they clearly disclose the tools can produce inaccuracies. LexisNexis also noted that the study tested an older version of their system and that their newer RAG 2.0 platform addresses some identified issues.³¹

Meanwhile, students need guidance about which tasks are safe for GenAI assistance and which require multiple verification layers. We're making those judgments based on incomplete information and practitioner experience rather than systematic research.

f) Questions Your Students Will Ask That You Can't Answer Yet

"Professor, if the model can analyze Supreme Court opinions better than I can, why do I need to learn legal analysis?"

"Is it cheating if I use an AI agent that I trained on my own class notes and outlines?"

"How do I know if my research is good enough when AI found cases I missed?"

"Will firms even want me to do my own research, or will they just want me to supervise AI?"

"If everyone uses AI in practice, why do law schools still ban it on exams?"

The honest answer to all of these: We're still figuring it out, and your generation will help determine the answers through the choices you make.

2. The Unknown Unknowns: What We Can't Yet Imagine

In 2020, how many of us predicted we'd have systems capable of passing the bar exam by 2023? The capabilities that will exist when today's 1Ls practice law will likely include breakthroughs we haven't conceived of yet.

a) The Next Capability Surprise

Every time companies release new GenAI systems, they can do surprising things nobody expected. In 2025, LLMs that could "think through" problems surprised researchers by solving challenging competition math problems much sooner than anyone thought possible.³² Programs from OpenAI, Google (Gemini Deep Think), and DeepSeek

³¹ *Id.*

³² Sebastian Raschka, *The State Of LLMs 2025: Progress, Problems, and Predictions* (Dec. 30, 2025), <https://magazine.sebastianraschka.com/p/state-of-llms-2025>, (last visited Feb. 3, 2026) (discussing how DeepSeek R1 showed reasoning behavior can be developed with reinforcement learning, and noting that

all reached this level. Scientists discovered that they could teach GenAI to work through problems step-by-step by rewarding it when it got the right answers, kind of like training a dog with treats (but for machines).³³ Companies also started building GenAI systems that mixed different approaches together to make them work faster and better, and this happened much quicker than expected.³⁴ We have no idea what surprising things GenAI will be able to do next, which makes it especially difficult for professors to plan how to teach with these tools.

b) Regulatory Frameworks That Don't Exist Yet

We don't yet know how different jurisdictions will regulate GenAI use in legal practice. Will we see mandatory CLEs or proficiency requirements for specific tools? Mandatory client disclosure? Outright prohibitions for certain use cases? Nobody knows yet, and the current patchwork of approaches makes predicting any coherent national framework nearly impossible. Some jurisdictions are actively experimenting while others are issuing cautious interim guidelines.³⁵ And a few are watching everyone else before making a move.

The judiciary itself provides a case study in productive uncertainty. Judge Kevin Newsom of the Eleventh Circuit has used his concurring opinions as something of a public laboratory for thinking through how LLMs might construe legal language. In a concurrence in *Snell v. United Specialty Insurance* and a follow-up he called "a sequel of sorts" in *United States v. Deleon*, Newsom tested whether LLMs could help determine the ordinary meaning of contested legal phrases. He prompted ChatGPT, Claude, and Gemini ten times each and compared results. He found that LLMs may serve a valuable auxiliary role alongside dictionaries and syntactical context, provided users understand that slight variations in output reflect the same fuzziness present in human language use generally.³⁶

reasoning models achieved gold-level math competition performance in 2025 rather than the anticipated 2026).

³³ *Id.*

³⁴ *Id.*

³⁵ *AI and Attorney Ethics Rules: 50-State Survey*, JUSTIA, <https://www.justia.com/trials-litigation/ai-and-attorney-ethics-rules-50-state-survey/> (last visited Apr. 24, 2026).

³⁶ Clay Calvert, *The Judiciary, Generative AI, and Ordinary Meanings: Kevin Newsom Leads the Way*, American Enterprise Institute AEIdeas (Oct. 15, 2024), <https://www.aei.org/technology-and-innovation/the-judiciary-generative-ai-and-ordinary-meanings-kevin-newsom-leads-the-way/> (last visited Mar. 13, 2026). *Snell v. United Specialty Insurance*, No. 22-13998 (11th Cir. 2024) (Newsom, J., concurring); *United States v. Deleon*, No. 23-10326 (11th Cir. 2024) (Newsom, J., concurring).

Not everyone on the bench shares Newsom's experimental enthusiasm. As James O'Donnell reported in *MIT Technology Review*, the same summer that early-adopting judges were exploring GenAI's potential, AI-generated mistakes were going undetected in actual rulings.³⁷ A federal judge in New Jersey had to reissue an order containing errors likely attributable to GenAI, and a federal judge in Mississippi declined to explain why his order contained what appeared to be hallucinations, writing simply that "no further explanation is warranted."³⁸

Judge Scott Schlegel of the Fifth Circuit captured the asymmetry that makes judicial GenAI use categorically different from attorney use: "When the judge makes a mistake, that's the law. I can't go a month or two later and go 'Oops, so sorry,' and reverse myself."³⁹

Some judges have found workable middle ground. Federal Judge Xavier Rodriguez of the Western District of Texas uses GenAI to summarize cases and generate hearing questions, reasoning that tasks requiring no final judgment and offering ample opportunity to catch errors before they matter are "relatively risk free."⁴⁰ Federal Magistrate Judge Allison Goddard in California keeps Claude open all day as a thought partner and uses it to summarize lengthy orders and organize unwieldy documents. But she draws a firm line at anything touching criminal cases where bias concerns are acute.⁴¹ Rodriguez and Goddard both contributed to guidelines published by the Sedona Conference in February 2025, which identified potentially safe judicial uses of GenAI

³⁷ James O'Donnell, *Meet the Early-Adopter Judges Using AI*, MIT Technology Review (Aug. 11, 2025) <https://www.technologyreview.com/2025/08/11/1121460/meet-the-early-adopter-judges-using-ai/> (and at <https://www.bespacific.com/meet-the-early-adopter-judges-using-ai/>) (last visited Mar. 13, 2026).

³⁸ *In re CorMedix, Inc. Securities Litigation*, No. 22-cv-06458 (D.N.J. 2025). Judge Julien Xavier Neals withdrew his June 30, 2025 opinion after lawyers flagged fabricated quotations and incorrect case outcomes; a law school intern had used ChatGPT to conduct legal research in violation of chambers policy. *Mississippi Association of Educators v. Board of Trustees of State Institutions of Higher Learning*, No. 3:25-cv-00374 (S.D. Miss. 2025). Judge Henry T. Wingate issued and later replaced a July 20, 2025 restraining order after defendants identified errors including misnamed parties and misquoted state law; a law clerk had used Perplexity AI as a drafting tool. Both judges subsequently confirmed AI involvement in letters to Senate Judiciary Committee Chairman Charles Grassley. See Josh Gerstein, *Federal Judges Using AI Filed Court Orders with False Quotes, Fake Names*, WASH. POST (Oct. 29, 2025), <https://www.washingtonpost.com/nation/2025/10/29/federal-judges-ai-court-orders/> (last visited Mar. 13, 2026); O'Donnell, *supra* note 37.

³⁹ O'Donnell, *supra* note 37.

⁴⁰ *Id.*

⁴¹ *Id.*

while noting plainly that "no known GenAI tools have fully resolved the hallucination problem."⁴²

Given all this, is the distinction between "safe" and "risky" GenAI tasks stable enough to teach? Erin Solovey, a human-AI interaction researcher at Worcester Polytechnic Institute who studied UK judges' thinking on this question, found that "the line between what is appropriate for a human judge to do versus what is appropriate for AI tools to do changes from judge to judge and from one scenario to the next."⁴³ Imagine a future where you can't use a GenAI legal research tool in a brief filed with the court unless both you *and the tool* have state bar certification. Or a scenario where only GenAI tools that have passed a rigorous accuracy audit can be used for client work, and your malpractice insurance requires proof that you used licensed tools.

The uncertainty is real because students graduating today don't know whether they'll need special licenses or which tools will be approved. Learning to use tools that might be banned, or failing to learn tools that become mandatory, can create problems. Law schools are training students for a regulatory environment no one can predict, and where the frameworks students learn today might be obsolete before they take the bar exam.

3. The Integration of Legal AI into Legal Systems

Will courts eventually require digital filing systems that automatically check citations using GenAI? Might some jurisdictions even use GenAI systems for preliminary case assessment or routine motion practice? We don't know how deeply GenAI will integrate into legal practice.

As GenAI gets baked into the systems lawyers use every day, legal pedagogy will have to adapt. Students who will practice inside those systems need to understand both what the tools can do and where they go wrong. Yet we can barely keep up with standalone tools, let alone predict how they'll be woven into institutional infrastructure.

Just as the internet created entirely new areas of law (internet privacy, computer crime, intellectual property rights, cyberbullying, net neutrality, domain name disputes, platform liability and Section 230⁴⁴), GenAI will likely spawn practice areas we cannot

⁴² The Sedona Conference, *Guidance for Use of Artificial Intelligence in the Courts* (Feb. 2025).

⁴³ O'Donnell, *supra* note 37 (quoting Erin Solovey, Worcester Polytechnic Institute).

⁴⁴ 47 U.S.C. § 230 (1996). Section 230 was enacted as part of the Communications Decency Act, which was Title V of the Telecommunications Act of 1996, Pub. L. No. 104-104, 110 Stat. 56 (1996).

currently envision. Perhaps GenAI and Agent liability law, or perhaps something entirely different. Maybe “synthetic media law” that will cover AI-generated video, text, audio, images. I’m just guessing here. Students entering law school today might spend significant portions of their careers in specialties that don't yet have names.

B. Moving Forward

Legal education has survived plenty of upheavals. It went from apprenticeships to law schools (and maybe we're heading back – who knows?). Law schools adopted the case method. And added clinics. We figured out Westlaw and LexisNexis. Remember when it was distributed on CD-ROMs? When those tools arrived, law professors insisted that manual search using books was irreplaceable and worried that students would never learn real research skills. They thought that relying on a computer to find cases was somehow intellectually suspect. Law librarians wondered if there was a future for them. Everybody adapted. Each time, we managed to keep teaching the important stuff while adapting to whatever new reality came our way. And GenAI looks like it’s going to be huge.

Here's what we know how to do already.

We know law is about judgment, not just information retrieval. GenAI can pull up cases faster than any human but can't tell a client whether to settle or go to trial (although it is advising pro se parties to do just that). Technology changes how quickly we access information, but it doesn't replace the need for wisdom grounded in professional values. We just need to figure out how to develop that judgment in students who have ChatGPT in their pockets.

We already teach people to work with imperfect sources. Legal research has always meant evaluating whether a source is reliable and cross-checking citations. And, just as importantly knowing when you need to keep going.⁴⁵ GenAI requires the same skills, just applied to a new kind of source that's really confident even when it's wrong.

Law schools have always prepared students for a legal world not yet known. We've been teaching for a future we won't live in for a long time now. GenAI just makes that gap wider and the timeline shorter, but the challenge isn't fundamentally new.

⁴⁵ See Rebecca Fordon, *Legal Research Trapping You in an “AI Tunnel”? Use a Toe-hold to Get Out*, AI Law Librarians (Oct. 28, 2025), <https://www.ailawlibrarians.com/2025/10/28/legal-research-trapping-you-in-an-ai-tunnel-use-a-toe-hold-to-get-out/> (last visited Feb. 3, 2026).

(Though I'll admit the timeline part is getting ridiculous!)

*Modeling how thoughtful professionals handle uncertainty
might be the most valuable thing we teach.*

GenAI (in fact, all AI) is changing too quickly to keep up with. But we can anchor GenAI policies in principles of academic integrity and professional responsibility. Law (and life) has always been about making the best decisions possible with the information you have, and that's exactly what we're teaching our students to do.

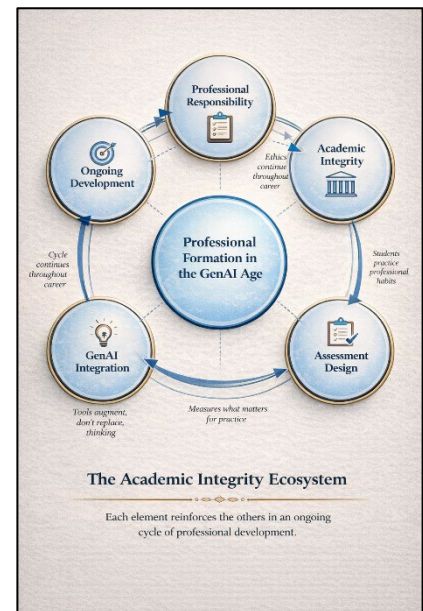
CHAPTER 16: The Path Forward

I think it's clear by now that I wrote this book because I believe legal education cannot ignore GenAI. Pretending it doesn't exist or isn't important serves no one. Our students will practice in a legal profession already shaped by these tools, and that is true whether we have made our peace with it or not.

I also believe that many law professors do not know where to start. I'm ending where I started, but hopefully with a clearer sense of what "not ignoring it" means in practical terms. A perfect policy won't save us entirely, and neither will a different exam format. Two moves will be very important. First, we must reconnect academic integrity to professional responsibility in the context of GenAI. Second, we must design assessments that measure what matters.

A. Returning to First Principles

Law schools take bright people and help them develop the judgment and ethical grounding to serve clients well. That has been the mission for as long as there have been law schools, and GenAI doesn't change it.¹ What GenAI does change are the methods used to achieve it. Students sitting in our classrooms will practice law in a world where LLMs can draft motions, help prepare for depositions, analyze documents at speeds we can barely comprehend. They'll work alongside these tools daily. They'll make constant judgment calls about when to use them and how to verify their outputs. They'll decide what work they must do themselves to maintain competence and fulfill their professional obligations.² In short, they'll be implementing the 4 Ds of AI Fluency: Delegation, Description, Discernment, and Diligence.



¹ See WILLIAM M. SULLIVAN et al., EDUCATING LAWYERS: PREPARATION FOR THE PROFESSION OF LAW (2007) (Carnegie Foundation report emphasizing the need for integration of legal analysis, practical skills, and professional identity formation).

² ABA MODEL RULES OF PRO. CONDUCT r. 1.1, cmt. 8 (A.B.A 2023) (addressing lawyer competence in the context of technology: "To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology").

A law professor's job is to prepare students for that new reality. The profession simply isn't the same as when you went to law school or practiced. The academic integrity challenges presented by GenAI must be central to a law school's mission and to the way we teach.

B. The Professional Responsibility Through-Line

Throughout this book, I tried to draw connections between academic integrity and professional responsibility. Students are practicing the habits and decision-making patterns that will shape their professional conduct. And without our guidance, they're learning what shortcuts are acceptable and how to rationalize behavior they know is wrong.³

We can help redirect them by designing assignments that require students to engage authentically with legal analysis. As discussed, assignments and assessments can be designed so that students cannot game them by using GenAI.

You're probably already seeing AI slop in student work. A student might fail the class, be forced to appear before the academic integrity committee, or might even be expelled. Awful enough, certainly. But in law practice, that same behavior could lead to a malpractice claim or a disciplinary action. Incompetence with GenAI tools can cause real harm to a client, including having a default judgment entered against them because their lawyer kept filing hallucinated citations and couldn't explain why.⁴ The duty of competence has always required understanding the tools of the trade.⁵ No one wants to see their law school's alumni in the news being sanctioned for submitting filings with fake citations or, perhaps worse, introducing deepfakes as evidence.⁶

³ See DONALD L. MCCABE et al., CHEATING IN COLLEGE: WHY STUDENTS DO IT AND WHAT EDUCATORS CAN DO ABOUT IT 15-22 (2012) (discussing how academic dishonesty in higher education correlates with professional misconduct).

⁴ *Flycatcher Corp. Ltd. v. Affable Avenue LLC*, No. 1:2024-cv-09429 (S.D.N.Y. Feb. 5, 2026) (The court entered default judgment against defendant as sanction after counsel filed AI-generated briefs containing fabricated citations on three separate occasions despite court warnings: "Indeed, the most remarkable element of Mr. Feldman's misconduct [the use of GenAI] — and a significant reason why the Court is defaulting his client — is his continuous pattern of behavior. He kept submitting erroneous citations and kept refusing to check his citations despite ample warning from the Court to do so or face sanctions, and despite his own admission that he had a viable means to access Westlaw and Lexis to conduct a cite check. Mr. Feldman never "corrected" his "misconduct" either").

⁵ MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2023) (requiring competent representation); and see r. 5.3 (requiring supervision of nonlawyer assistants, including AI tools).

⁶ Frank Landymore, *Judge Horrified as Lawyers Submit Evidence in Court that Was Faked With AI*, Futurism (Nov. 19, 2026 1:24 PM EST), <https://futurism.com/artificial-intelligence/judge-horrified-lawyers-submit-ai-evidence> (last visited Feb. 5, 2026).

Some students will get this concept immediately. Others will need it spelled out. Many will need multiple conversations, across multiple courses, before the connection truly sinks in. That's fine. Professional formation takes time.⁷

C. Preparing Students for Practice, Not Just Exams

Traditional legal education often optimizes for performance on exams (and ultimately on the bar exam) rather than performance in practice. In doctrinal exams, students spot issues in hypothetical fact patterns and structure IRAC paragraphs under time pressure, doing work that bears little resemblance to actual practice.⁸

GenAI can now generate competent legal analysis faster than any human can type, which must make us wonder what we are really measuring when we assess students?

Of course, legal reasoning still matters. Students need to be able to construct arguments and exercise professional judgment. But those skills must mean something in practice, not just in the exam room.

When assignments are designed that can't be completed by copying GenAI output, students are required to apply legal analysis to real-world scenarios with constraints. When students are asked to make and justify professional judgments rather than simply identify issues, professors are assessing skills they'll use in practice. And, as important, academic dishonesty becomes much harder.

To ensure the best academic integrity in a GenAI world, we must design learning experiences that require authentic engagement. If your assessment can be successfully completed by someone who doesn't understand the material, perhaps the assessment needs a redesign. See Chapters 5, 8, 9, and 10 for GenAI-resistant/integrated/transparent assignments and assessment.

⁷ Neil Hamilton & Verna Monson, *Ethical Professional (Trans)formation: Themes From Interviews About Professionalism With Exemplary Lawyers*, 52 SANTA CLARA L. REV. 921, 927-31 (2012) (discussing the developmental nature of professional identity formation in lawyers).

⁸ See Roy Stuckey, *"Best Practices" or Not, It is Time to Re-Think Legal Education*, 16 CLINICAL L. REV. 301 (Fall 2009), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1523191 (last visited Feb. 5, 2026) (outlining five-step framework for legal education: establishing clear learning objectives, organizing progressive curriculum in knowledge and skills, selecting effective teaching methods, employing methods skillfully, and evaluating institutional success).

The best academic integrity countermeasures aren't restrictions and surveillance. They're well-designed learning experiences that require authentic engagement.

D. Maintaining Rigor While Embracing Change

Some faculty are concerned that adapting to GenAI means lowering standards. They might think that by jettisoning traditional exams in favor of assessments that allow or even encourage GenAI use, and thereby focusing on process rather than final product, we're somehow watering down legal education. We're making it easier. We're graduating students who aren't as prepared and can't "think like a lawyer."

Respectfully, I think this gets it backwards. Rigorous legal education in this new context means requiring deeper analysis than GenAI can provide. Current LLMs can produce competent legal analysis on straightforward questions but they struggle with complex fact patterns and novel legal issues. They're not good at making strategic considerations and professional judgment calls. Law professors should design assessments that push students into this territory by asking them to grapple with ambiguities and analyze the grey zones in the law where reasonable lawyers disagree.⁹

1. Making the process visible.

In practice, lawyers don't just reach conclusions. They document their reasoning while explaining their analysis to clients and colleagues, constantly justifying their strategic choices. GenAI can generate a memo, sure. But it can't tell you why the lawyer emphasized certain arguments or what judgment calls were made along the way. So why not require students to show their work and explain their thinking? Have them turn in their chats with GenAI. Make them articulate why they made the choices they made.

⁹ See Cheryl B. Preston, Penée Wood Stewart & Louise R. Moulding, *Teaching "Thinking Like a Lawyer": Metacognition and Law Students*, 2014 BYU L. REV. 1053, 1060-61, 1073-79 (2014) (discussing how expert legal analysis requires metacognitive skills including the ability to monitor one's thinking, exercise professional judgment, handle ambiguity, integrate complex considerations, and learn strategically from experience - precisely the capabilities that distinguish human legal analysis from current GenAI outputs on complex problems).

2. Integrating professional responsibility explicitly.

Legal problems often involve not just doctrinal questions but professional responsibility considerations. Make the latter explicit. Ask students to identify potential ethical issues and explain how they would handle them, then require them to justify their choices under the relevant rules. GenAI can cite Model Rule 1.7.¹⁰ It cannot help students develop the practical wisdom to navigate conflicts of interest in real time with real clients.

3. Expecting professional-quality work product.

If students have access to tools that can improve their writing and format their documents properly, why not allow them to use those tools? When spell check and grammar check came out in Word, did you ban students from using the program? Set the standard for written work at what you'd expect from a competent junior associate, not from a stressed law student, up at 2 a.m. with a sick child the night before the deadline. Students can use GenAI to help them meet that standard, as long as they understand that "help" doesn't mean "do it for me."

4. Building in reflection and metacognition.

Ask students to think about their own learning. Require them to articulate what they understand and what they still find confusing, and to identify the skills they're developing and the areas where they need more practice. Reflection is hard to fake and is important for professional development.¹¹

All of this requires more from students and from faculty. But, honestly, isn't that what rigor means? We just have to face it. GenAI has changed the rules.

One other note: law school administrations should recognize that professors will need resources and support to meet this new challenge.

¹⁰ MODEL RULES OF PRO. CONDUCT R. 1.7 (A.B.A. 2023) (prohibiting representation of a client whose interests are directly adverse to another current client absent informed written consent).

¹¹ Timothy Casey, *Reflective Practice in Legal Education: The Stages of Reflection*, 20 Clinical L. Rev. 317, 331-32 (2014) (proposing a six-stage model of reflection that moves students from concrete assessments of competence through increasingly complex analysis of their decision-making processes, internal and external factors affecting performance, and societal context, ultimately developing metacognitive awareness essential for professional judgment and practice).

5. Reflection Questions for Faculty

- Look at your most recent exam or major assignment. Could it be completed successfully using only GenAI output? If you don't know, you likely need some GenAI training, so it's a good thing that you have this book. Upload it to GenAI and find out.
- What specific skills does your assessment measure? Are those the skills your students will need in practice?
- Where in your course do students practice the kind of judgment calls that they'll make as lawyers?
- How explicitly do you connect your academic integrity expectations to professional responsibility rules?

E. The Academic Integrity Ecosystem

By the time this book is published, some of what's in it will already be out of date. GenAI capabilities are advancing rapidly. New tools will emerge. In fact, as I write this, Thomson Reuters stocks dropped almost 16%, its biggest drop in a single day, and RELX, the company that owns Lexis, fell 14%. Why? Because Anthropic released Claude Cowork with new plugins to support different industries, including legal.¹² To learn more about this new feature, watch Anthropic's [Claude for Legal Workshop](#). Legal tech investors freaked out, largely because Cowork plugins could kill seat-based subscriptions. Why pay for ten individual Westlaw licenses when Cowork can connect to the service via API and provide access to all ten lawyers through a single enterprise connection?

Students will find new ways to use (and misuse) GenAI. Legal employers will develop new expectations. What counts as cutting-edge pedagogy today will be standard practice tomorrow and possibly obsolete the day after. As Ethan Mollick, author of *Co-Intelligence: Living and Working with AI* (definitely worth a read), "Today's AI is the worst [t] AI you will ever use."¹³

¹² Hadas Gold & John Towfigh, *Anthropic's New AI Tool Sends Shudders Through Software Stocks*, CNN (Feb. 2, 2026) available at <https://www.cnn.com/2026/02/04/investing/us-stocks-anthropic-software> (last visited Feb. 5, 2026).

¹³ Ethan Mollick, LinkedIn post, available at https://www.linkedin.com/posts/emollick_todays-ai-is-the-worst-ai-you-will-ever-activity-7106305750431322112-Xr7n?utm_source=share&utm_medium=member_desktop&rcm=ACoAAAD0zgYBCOZAexe7bZ9qSaFu

Legal education, like all education, has always involved adaptation. New pedagogies, new technologies, new understandings of how students learn and how lawyers develop competence have developed over time. GenAI is one more change in a long series of changes and it won't be the last.¹⁴

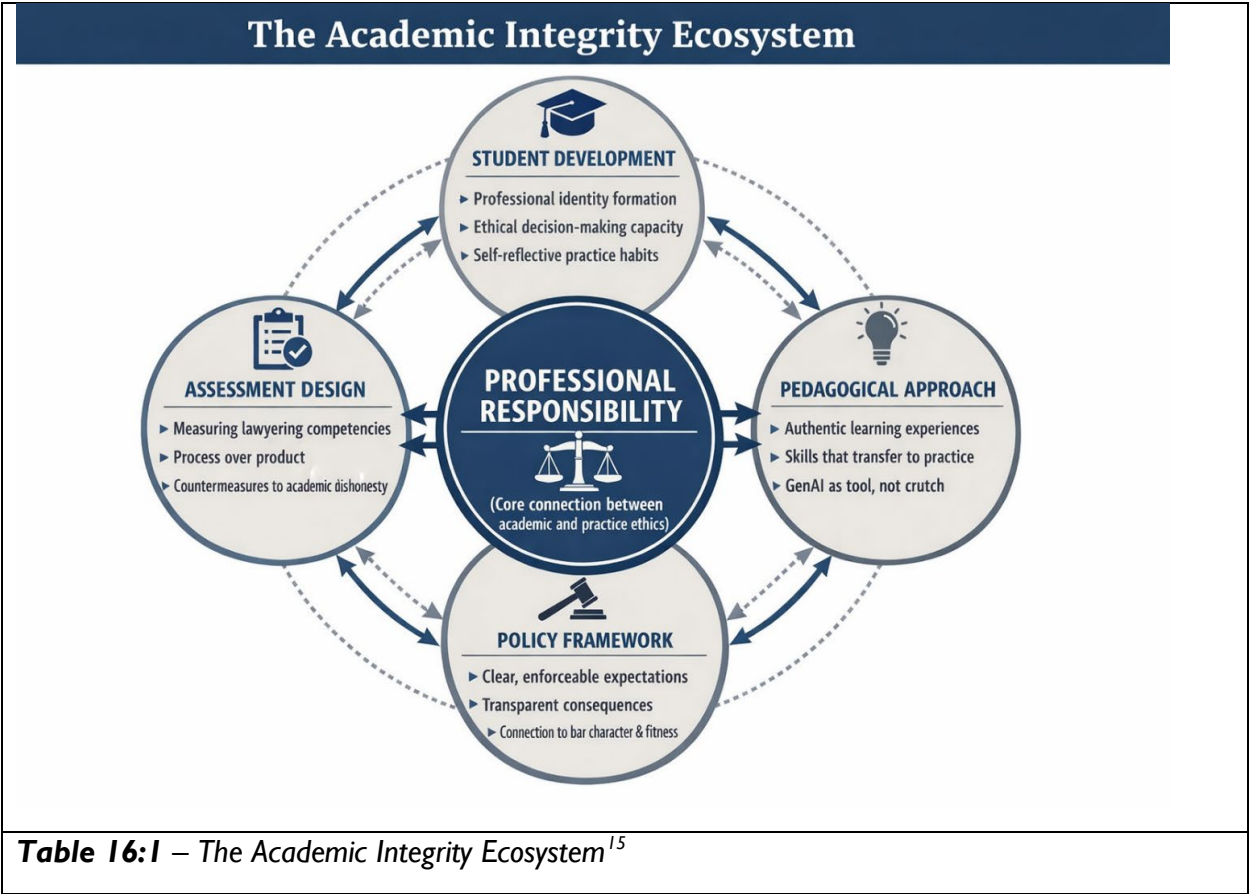
Law schools must pay attention to how GenAI is evolving and how students, professors, and practitioners are using it. The profession is not waiting for us to catch up. We need to talk with colleagues at our institutions and beyond, sharing what works and what doesn't. We need to be willing to experiment, to fail, to learn, and to try again. And, in my opinion, teaching ethical and proper prompting of GenAI should be part of orientation or, even better, a required course in the first semester of the first year.

Law schools need a framework for thinking through the questions, plus a commitment to connecting choices back to legal education's core purposes. And students and professors need a willingness to learn alongside each other as the ground shifts under all our feet.

To state the obvious, students in classrooms next semester will be different from the students in your classroom last semester. (And somehow, they always get younger, don't they? Or is that just me?) They'll have different GenAI tools available. They'll arrive with different levels of comfort and proficiency with those tools, and different ideas about what counts as appropriate use. Our job isn't to stay three steps ahead of them forever. Our job is to create learning environments where students develop the judgment to practice competently and the professional values to practice responsibly.

[9z79xHewcww](#), last visited Feb. 5, 2026; see also, ETHAN MOLLICK, CO-INTELLIGENCE: LIVING AND WORKING WITH AI, (2024).

¹⁴ See Larry O. Natt Gantt, II, *Deconstructing Thinking Like a Lawyer: Analyzing the Cognitive Components of the Analytical Mind*, 29 CAMPBELL L. REV. 413, 453-54 (2007) (noting that legal education must continuously evolve to prepare students for a changing profession).



F. For Law Schools

If you're in a position to influence departmental or institutional policy, use it. Work with colleagues to develop coherent approaches to GenAI across courses. Students

¹⁵ Generated by ChatGPT 5.2, February 5, 2026, Korin Munsterman, prompt: You are writing a book called GenAI for Teaching and Learning Law. You want to include a graphic in the final chapter. Create a graphic called: The Academic Integrity Ecosystem Here are the instructions: Center Circle: PROFESSIONAL RESPONSIBILITY (Core connection between academic and practice ethics) Four Outer Circles (Connected to center with bidirectional arrows): Top: STUDENT DEVELOPMENT Professional identity formation Ethical decision-making capacity Self-reflective practice habits Right: PEDAGOGICAL APPROACH Authentic learning experiences Skills that transfer to practice GenAI as tool, not crutch Bottom: POLICY FRAMEWORK Clear, enforceable expectations Transparent consequences Connection to bar character & fitness Left: ASSESSMENT DESIGN Measuring lawyering competencies Process over product Countermeasures to academic dishonesty [Connecting Arrows should show:] Bidirectional relationships between all elements All elements flow through Professional Responsibility Dotted lines between adjacent outer circles showing their interdependence Visual Notes: Use blue/gray palette. Consider adding subtle icons (scales of justice for Professional Responsibility, graduation cap for Student Development, etc.).

shouldn't have to navigate wildly different expectations from one professor to the next with no underlying rationale.

That said, coherence doesn't mean uniformity. Different courses, and different assignments and assessments, serve different purposes and may appropriately have different policies. The goal is for students to understand why the policies differ, not just that they do.

Many faculty want guidance on how to adapt their teaching but don't know where to start. Workshops and reading groups can help (start with this book—it's why I wrote it). So can sharing examples of what colleagues are doing, both within your institution and beyond. For those who teach Legal Writing, Dr. Kirsten Davis, professor of law at Stetson University College of Law, and Professor Carolyn Williams at the University of North Dakota run a “convo” group each month. Professors Davis and Williams cover topics such as “Rethinking the Writing Sample in the Age of GenAI,” and “Questions About the Faculty Role in Policing Academic Integrity: Is it Acceptable to Mislead Students to Catch GenAI ‘Cheating’? What Is the Appropriate Role of AI Detection Apps in the Legal Writing Class?” I’ve attended several and, even though I don’t teach Legal Writing, I always learn something new.¹⁶

Consider your academic integrity infrastructure. Do your hearing procedures account for the realities of GenAI? Are the people making decisions about alleged violations informed about how these tools work and how students use them? Especially with regard to detection tools: do they realize how problematic they are?

Law schools certainly don't want to find themselves in the position of Adelphi University. In January 2026, a New York State Supreme Court judge granted an Article 78 petition by Orion Newby, a first-year student accused of AI-assisted plagiarism based on a Turnitin flag. The court reversed all disciplinary measures and ordered expungement of his record after finding the university's determination "without valid basis and devoid of reason."¹⁷

Rather than simply trying to detect cheating, institutions should ask "how we should be reimagining assignments in an age where generative AI is a thing. If something

¹⁶ To view some of the materials from past Convo meetings, go to: <https://www.dropbox.com/scl/fo/8ej6mzjdobtdto1qzeurx/ADm9ov7-13x3A5gMTsDPtTI?rlkey=2dazcjs1leboqs4vsj15j6wwt&e=1&dl=0> (last visited Mar. 19, 2026).

¹⁷ Matter of Newby v. Adelphi Univ., 2026 NY Slip Op 26021 (Sup. Ct. Nassau Cnty. Jan. 28, 2026), <https://law.justia.com/cases/new-york/other-courts/2026/2026-ny-slip-op-26021.html>.

can be done 100% by ChatGPT, then maybe we should be changing the task."¹⁸ Is your approach primarily punitive or does it include educational components?

Finally, does your institution have the infrastructure to address academic integrity in the GenAI age effectively? What's missing? Are faculty members in your department or school having productive conversations about these issues, or mostly just complaining about students these days? What would it take to shift toward more productive dialogue? Who are your potential allies in this work? Who's already thinking about these issues and might be willing to collaborate?

G. A Final Thought Experiment

Imagine it's two years from now. GenAI capabilities have continued to advance in ways we can barely predict today. Your current students are practicing lawyers. Some of them work at firms that have integrated GenAI tools deeply into their workflows. Others work in settings with more limited technology. All of them make daily choices about when to use these tools (delegation, description, discernment, and diligence).

One of your former students calls you. He's been asked to handle a matter that's outside his usual practice area but related to work he did in your course. He's using GenAI to help get up to speed quickly. He wants to talk through how to do this responsibly. He wants to serve his client well and use the technology ethically.

How are you going to answer him? What do you hope he learned in your course that prepared him (and you) for this conversation?

The answer to that question should drive everything else.

H. Conclusion

We don't have a choice about whether GenAI becomes part of legal education and practice. It already is. What we have a choice about is how we respond. Students are asking for help and guidance, and they deserve our best efforts to prepare them for the profession they're entering.

¹⁸ Toni Matthews-El, *Sell Schools the AI detector, Sell students the Workaround. Inside a Troubling Double Game*, *Innovating with AI Magazine* (Feb. 2, 2026) available at <https://innovatingwithai.com/sell-schools-the-ai-detector-sell-students-the-workaround-inside-a-troubling-double-game/> (last visited Feb. 5, 2026) (arguing that institutions should reimagine assignments rather than simply detecting AI use).

Best of luck to us all!

Index

A

- AAUP. *See* American Association of University Professors
- ABA. *See* American Bar Association
- ABA Formal Opinion, 246, 297, 447
- ABA Model Rules. *See* Model Rules of Professional Conduct
- ABA standards, 214, 219, 221, 293, 304
- academic freedom, 161, 380–83
- academic integrity, 243, 297, 300, 369, 374, 382, 385, 421–23, 427–28, 432, 453, 469–72, 524
 - principles, 300–301, 385, 445
- accommodations, 170–71, 242, 259–60, 289, 291–92, 437, 443
- agents, 84, 222, 296–99, 455, 459–62
 - agentic AI, 519, 526–28
 - AI agents, 455, 460–61, 465
 - autonomous, 298, 460–61
- AI Assessment Scale (AIAS), 257, 259, 261, 274
- AI Fluency Framework, 98–99, 519, 524, 527
 - delegation, 98–99, 101, 362
 - description, 98–99, 101
 - discernment, 98–99
 - diligence, 98–99
- alignment, 37–40
- American Association of University Professors (AAUP), 364, 380–82, 384
- American Bar Association (ABA), 136, 177, 182, 238, 272, 296, 301, 335
- analogical reasoning prompting, 123
- Anthropic, 26–27, 41, 46, 49–50, 54–55, 75–81, 84–85, 98–99, 276–77, 308, 366, 368, 456–57, 461–62
- artificial intelligence. *See* generative AI (GenAI); large language models (LLMs)
- assessments, 12, 15–16, 217, 220, 229, 231, 233, 248, 257, 259, 269–70, 272, 274, 287–89, 292–94, 448, 451, 471–74, 495, 498, 556–57
 - collaborative, 283–87
 - formative, 16, 124, 179, 218, 481
 - oral, 267–68, 288, 292
 - portfolio, 271–72, 277–78, 292
 - process-based, 279–80, 294
- assignment design, 4, 120, 183, 237–56
- assignment instructions, 54, 71, 224, 255–56, 586, 601
- authorial accountability, 372, 384–86, 521

B

- bar exam, 11, 14, 17–18, 167, 430, 435, 437, 455, 465, 468, 472
 - bar passage data, 15–18
 - bar preparation, 88–91, 417–18
- bias, 18, 24, 36–37, 231–34, 327–28, 377, 379, 411–12, 557, 563, 571, 574
 - in GenAI-assisted grading, 232–34
 - in LLMs, 36–37

browser extensions, 439–40, 443

C

CAI. *See* Constitutional AI

case citations, 9, 34, 54, 100, 126, 147, 328, 405, 463

case law, fabricated. *See* hallucination

chain-of-thought prompting, 114–15

zero-shot, 116

ChatGPT, 1–2, 8, 11, 14, 45, 47–57, 60, 62–74, 84–89, 101, 103, 105–6, 139, 142, 144, 147, 150, 152, 154, 159–60, 162, 176, 179, 187–89, 194–95, 197, 200, 205–7, 209, 211, 220, 222, 227–28, 247, 257, 279, 281, 296, 298, 306–7, 313–15, 320–21, 344, 404, 458, 467, 469, 477, 479, 500, 530, 532–34, 536, 577

CustomGPTs. *See* CustomGPTs

GPT-4, 31–32, 57, 61–62, 136, 159, 167, 189, 228, 257, 454–55, 464

citations, 13–14, 33, 89, 91, 125, 128, 144, 167–68, 170, 173–74, 176–78, 347–48, 350, 357, 367–68, 370–71, 373, 385, 425–26, 428, 432–33, 468–69, 471, 541–43, 546

Claude, 23, 26–28, 40–41, 43, 49–51, 53–57, 63–71, 73–77, 80–89, 92, 101, 103–6, 142, 147, 176, 179, 252, 254, 257, 276–77, 307–11, 314, 344, 348, 350, 368, 372, 404–6, 446–47, 454, 458, 466–67, 500, 530, 533–34, 577–78, 598–604

artifacts, 112, 149, 253, 598–601, 603–5

Code, 75–76, 84

Cowork, 30, 75–79, 85, 124, 341, 368, 475

interactive tool, 590–92, 594

Opus, 29, 35, 42, 54, 66–67, 159, 454

Pro, 64, 66–67, 93

Projects, 30, 73, 85, 87, 306–9, 311, 341, 535

confabulation. *See* hallucination

Constitutional AI (CAI), 26–27, 55

context rot, 29–30

context windows, 28–29, 44, 55, 63, 69, 116, 121, 456–57, 515, 535

contextual variation prompting, 123

Copilot, 47, 50–51, 57–65, 147, 222, 229, 231, 345, 443, 501, 505–14

CORRECT framework, 107, 114, 134, 310, 411, 413

for students, 413–19

CORRECT Method Trainer, 111–12, 253, 598

crutch effect, 195–98

Criminal Procedure Teaching Assistant, 311, 313, 316, 323–24

Custom GPT. *See* CustomGPTs

CustomGPTs, 19–20, 53, 55–56, 70–74, 87–89, 103–4, 117, 185–87, 189–90, 203, 205–8, 295–309, 314–16, 319–21, 323, 327–30, 531, 555–57, 569–70, 577–79, 581–83, 586–88, 590, 603–4

assignment, 557–58, 569–70

creating, 53, 296–98, 300, 304, 326, 330, 339

instructions, 55, 70, 73, 298–99, 306–11, 315–16

D

deception, 35–36

delegation, 98–99, 101, 362

description (AI Fluency Framework), 98–99, 101

detection, 6, 34, 150, 152–57, 160, 162, 169, 172, 181, 527

detection tools, 150, 153–55, 159–60, 162, 164, 166–67, 181, 183, 229, 376, 478, 524

Turnitin, 155–58

devil's advocate prompting, 119

discernment, 98–99

disclosure, 143–44, 146–47, 249–51, 255–56, 364–66, 368, 370–74, 377–79, 383–86, 435–36, 445, 521–22
 requirements, 141, 250, 255, 324, 372–74, 378, 385–86, 388, 521–22
 statement, 146–47, 368, 425

diligence, 98–99

E

equity concerns, 9, 87, 205, 242, 330, 376

error analysis prompting, 124

exam questions, 14, 48, 65, 76, 89, 117, 134, 167, 242, 537–38, 600

F

fabricated citations. *See* hallucination

faculty, adjunct, 202–3, 555, 593

faculty governance, 304, 372, 383, 388, 454

few-shot prompting, 115–16

floor-raising effect, 8

formative assessments, 16, 124, 179, 218, 481

funnel prompting. *See* progressive narrowing

G

Gemini, 29–30, 49–50, 56–57, 63–69, 71, 84–85, 103, 105, 159, 404–5, 530

Gemini Gems, 85–87

Gemini Nano Banana Pro, 25, 27–28, 31, 33, 38

generative AI (GenAI)

 access, 196, 199, 204, 210, 397, 409–10, 436

 Appendix, 175–76, 535, 541–43, 547

 assistance, 194–95, 223–24, 240–42, 364–65, 370, 373, 375–76, 378–79, 385, 387, 422

 capabilities, 98, 446, 456, 475, 479

 context of, 127, 451, 470

 in legal education, 20, 181, 372, 523, 528

 in scholarship, 364, 372, 384–85

 integration, 4–6, 12, 141, 204, 247, 251

 interactions, 191, 207, 281, 310, 414, 447

 models, 13, 25, 66, 68, 155–56, 173, 232, 411

 output, 141, 148, 152, 204–5, 246–47, 366, 434, 436, 472–73, 525, 545–46

 platforms, 8, 47, 146, 167, 305, 326, 404, 417, 437, 440, 481, 500

 policies, 2, 141, 143, 148–49, 182–83, 200, 220, 427–28, 434, 448, 454, 592

 Policy Drafting Tool, 149, 220, 592, 598, 600, 605

 prohibited assignments, 136, 239, 241–42

 resistant assignments, 172, 243

 responses, 143, 369, 410, 418

 suggestions, 179, 204–5, 208–9, 222, 231, 233

 systems, 25–26, 55, 130, 205–6, 247, 457, 459, 465, 468

 tools, 5–6, 9–11, 101–2, 112–13, 125–26, 141–43, 148, 176, 178–82, 228–29, 293–94, 348, 367–68, 372–74, 376, 381–82, 385–86, 446–47, 467–68, 552

 tutoring, 189–90, 192–94, 202, 206

 tutors, 184, 186, 189, 193, 199, 206, 208, 396

 use, 4, 13–14, 17–19, 138, 140–42, 145, 170–71, 220, 237–39, 249–50, 258–59, 303, 364, 369, 372, 377–78, 383, 421, 423, 425–26, 429, 432, 434, 436, 445–46, 448–49

 use disclosure, 250–51, 378, 384–85

Use Disclosure Form, 252–53, 276, 281, 330, 594, 598, 605
generative artificial intelligence tools. *See* generative AI (GenAI)
GPT-4, 31–32, 57, 61–62, 136, 159, 167, 189, 228, 257, 454–55, 464
GPT Tutor, 198–202
grading, 174–76, 229–34, 290, 292, 303, 441, 481–82, 541, 543–44, 572, 574
 bias problem, 232–34
 criteria, 228, 287, 291, 505, 547, 587
 process-based, 279–80, 294
 rubrics, 18, 87, 174, 220, 230, 270, 503, 543, 554, 570, 586

H

hallucination, 13, 33–35, 54, 247, 342, 363, 411, 466
 case law, fabricated, 139, 426, 428, 464
 citations, fabricated, 48, 109, 139–40, 175, 247, 421, 428, 433, 445, 447, 471, 546
 risk, 13, 54, 342, 363

J

jagged technological frontier, 396–97

K

knowledge base, 73–74, 103–4, 298–99, 304, 309–11, 319–20, 324–25, 530–31, 581, 585, 588. *See also* knowledge grounding; retrieval-augmented generation (RAG)
knowledge grounding, 103–4, 120–21. *See also* knowledge base; retrieval-augmented generation (RAG)

L

language models. *See* large language models (LLMs)
large language models (LLMs), 1, 22–25, 27–29, 31, 33–37, 40–41, 46–47, 96, 101, 103, 107–8, 112, 116–18, 123–24, 138–39, 146–47, 154–56, 207, 237, 239–49, 298, 366, 368–69, 391, 393, 395–96, 399–400, 406–9, 411–12, 421–22, 432–34, 442, 446–47, 465–66, 530, 538–40. *See also* models; platforms
 alignment, 37–40
 bias, 36–37
 context windows, 28–29
 deception, 35–36
 hallucination, 33–35
 parameters, 30–32
 tokens, 22, 24, 28–30, 54, 56, 103, 156, 398, 530, 603
 training data, 23–25, 45
 transformer architecture, 22–23
law school policies, 2, 427–28, 454, 527
learning management system (LMS), 246, 252, 269, 279, 298–99, 309, 594
LectureLens, 203, 555–56, 558–59, 562, 565, 569, 572, 593
legal AI tools, 80, 246, 464
legal analysis, 136, 141–42, 144, 187, 230–32, 247–48, 265, 268, 272, 277–78, 470–73
legal citations, 53, 345, 357, 363
legal education, 1, 5, 12, 14–15, 20–21, 91–93, 97–98, 101–2, 116–17, 136, 192–94, 198, 237, 271–72, 293–94, 297, 389–90, 441, 472–73, 476, 518–20, 522–23, 527–28, 537–38
legal educators, 27, 37, 40, 156, 195, 207, 372, 459, 528, 598
legal ethics, 221, 258, 421
legal pedagogy, 29, 107, 123, 223, 380, 468, 593
legal practice, 1, 34, 128, 265, 283, 389, 452, 454, 460, 466, 468
legal profession, 3, 10, 13, 167, 240, 257, 265, 411, 445, 455, 519

legal reasoning, 37, 101, 166, 174–75, 182, 240, 391, 395, 401–2, 570, 572
legal research, 169, 182, 265, 344, 357, 380, 422, 426, 464, 467, 519
legal research platforms, 142, 357, 426, 528
legal scholarship, 23, 341, 343, 346–48, 351–52, 381, 384, 387
legal skills, 43, 77–79, 122
legal writing, 5–6, 181, 193, 231, 233, 242, 247, 249, 302, 441–42, 522, 525
Lexis+ AI, 344–46, 464
LLMs. *See* large language models
LMS. *See* learning management system

M

meta-prompting, 118
Microsoft Copilot, 45, 64, 71, 190, 458, 462
Model Rules of Professional Conduct, 58, 138–41, 238, 296–97, 300–301, 303–4, 329–30, 422–23, 425, 434, 447–48, 461, 471, 474
 Rule 1.1 (Competence), 138, 297, 446
 Rule 1.5 (Fees), 448
 Rule 1.6 (Confidentiality), 446
 Rule 3.3 (Candor), 139, 447
 Rule 5.3 (Supervision), 140, 238, 296–97, 447
 Rule 8.4 (Misconduct), 423
models, 8–9, 22–45, 52–58, 62, 65–69, 96, 107–10, 113–14, 159, 193, 255–56, 301, 319–20, 325–26, 328, 368, 411–12, 423–24, 442, 454–58, 535–37, 561, 570, 573. *See also* large language models (LLMs); platforms
 different, 66, 68, 159, 463
multi-platform strategy, 70–71
multimodal capabilities, 459

N

NotebookLM, 18, 30, 88–89, 91–93, 147, 341–42, 346–48, 354, 368, 418, 456

O

OpenAI, 14, 40–41, 45, 47–50, 52, 58, 61–62, 68, 160, 366, 368, 454, 465
oral argument practice, 267–68, 270, 286–87, 400–405

P

parameters, 30–32
pedagogical guardrails, 184, 190, 194, 201, 205, 318, 396
Perplexity, 65–69
platforms, 44–47, 49–57, 65–66, 68–72, 84–85, 93, 143, 145–46, 172–73, 176, 305–6, 308–9, 344–46, 362, 439–40, 500, 538, 548, 552–54, 590–94. *See also* large language models (LLMs); models
 multi-platform strategy, 70–71
 selection, 44, 49, 297, 306
Poe (Quora), 65–68
policies, 3–4, 45–46, 141–42, 145–46, 148–49, 181–83, 250, 256, 334, 340–41, 364, 372–75, 377–78, 383–86, 388, 424–26, 428, 430, 432, 434, 436, 462–63, 478, 524–25
 GenAI, 2, 141, 143, 148–49, 182–83, 200, 220, 427–28, 434, 448, 454, 592
 institutional, 44, 46, 221, 384, 477, 522, 527
 language, 115, 149, 372, 384
 scholarship, 371–72, 374
portfolio assessment, 271–72, 277–78, 292

- professional identity formation, 7, 255, 470, 472
- professional judgment, 5, 153, 226, 228, 233–34, 300–301, 412, 421, 430–32, 450, 455, 472–74, 523
- professional misconduct, 423, 433, 447, 471
- professional responsibility, 9, 250, 254–55, 300–302, 381, 421–23, 425, 431–32, 453, 469–71, 474, 477, 500
 - principles, 301, 324, 326, 369
- progressive narrowing (funnel prompting), 119–20
- prompt chaining, 117
- prompt engineering, 52, 68, 94, 96, 99–101, 128, 557
- prompt library, 548
- prompting techniques, 114–24
 - analogical reasoning, 123
 - chain-of-thought, 114–15
 - contextual variation, 123
 - devil's advocate, 119
 - error analysis, 124
 - few-shot, 115–16
 - meta-prompting, 118
 - progressive narrowing, 119–20
 - prompt chaining, 117
 - scaffolding, 122–23
 - Socratic, 117
 - structured output, 122
 - synthesis, 122
 - zero-shot chain-of-thought, 116

R

- RAG. *See* retrieval-augmented generation
- regulation, 215, 217, 339, 354, 357, 556, 558–59, 563, 565–66, 568–71
 - analysis, 571, 573–75
 - approaches, 130–31, 557, 566, 570
 - frameworks, 128, 130–31, 133, 434, 466, 557, 566, 571
- Reinforcement Learning from AI Feedback (RLAIF), 26–27, 38
- reinforcement learning from human feedback (RLHF), 25–27, 38
- research memos, 11, 143–44, 163, 230, 237, 248, 265, 276
- research tools, 65, 146, 169, 224, 342–43, 348, 426, 428, 435
- retrieval-augmented generation (RAG), 34, 120–21, 147, 310, 344–45, 464. *See also* knowledge base; knowledge grounding
- RLAIF. *See* Reinforcement Learning from AI Feedback
- RLHF. *See* reinforcement learning from human feedback
- rubric, 71–72, 113, 176, 228–30, 232–34, 275, 291–92, 513, 541–44, 547, 572–73

S

- scaffolding prompting, 122–23
- scholarship, 302–3, 308, 340–41, 351–54, 356–59, 363–65, 368, 371–72, 374, 376, 379–80, 382–83, 387–88, 519, 521–23, 527–28
 - policies, 371–72, 374
- service work, 20, 304, 332, 340–41
- Socratic method, 234, 267, 307, 310, 315, 402, 405, 418–19
 - dialog practice, 400–405
 - prompting, 117

- questions, 117, 127, 192, 208, 310, 419
 - Simulator, 72, 117, 591
 - teaching assistant, 313, 316, 577–79
- speakers, non-native English, 157, 181, 183, 437, 441
- standing orders, 142, 425–27, 429
- structured output prompting, 122
- study guides, 89–90, 400, 492
- subscriptions, 44–45, 61, 64, 66–69, 73–75, 86–88, 147, 222, 307, 330, 590
- syllabus, 71–72, 87, 89–90, 115–16, 121–22, 141–42, 219–22, 239–40, 287, 307, 323, 326, 425–27, 429, 555–58, 586–87
- synthesis prompting, 122
- system prompts, 40–42

T

- TBL. *See* Team-Based Learning
- teaching assistants, 151, 234, 280, 302, 307, 309, 311, 320, 324, 409, 445, 449
- team application exercises, 259, 263, 289–90
- Team-Based Learning (TBL), 258, 261, 265–67, 283, 288, 291
- tenure, 364, 372–77, 379, 382, 384–88, 522
 - committees, 303, 375–76, 379–80, 384
- tokens, 22, 24, 28–30, 54, 56, 103, 156, 398, 530, 603
- training data, 23–25, 45
- transformer architecture, 22–23
- Turnitin, 155–58

V

- verification, 9, 100, 139, 170–71, 345, 363, 369, 384–86, 425–27, 446, 448–49, 543, 545–46
 - process, 175–76, 541, 546
 - requirements, 346, 357, 366, 372–73, 385–86, 426

W

- Westlaw, 142, 169, 266, 426–27

Z

- zero-shot chain-of-thought prompting, 116

