

# Generative AI & Legal Research: A Mismatch?



Practical  
applications of  
generative AI in  
legal research.

BY LELAND SAMPSON

// How do you use generative artificial intelligence (GenAI) for legal research?" The question usually surfaces in the context of continuing legal education presentations. The answer for most law librarians is, "I do not really use GenAI for research, but I use it for other job-related tasks." Articles describing how GenAI will disrupt the legal field routinely cite legal research as a category ripe for disruption, but how does that disruption manifest in a law librarian's daily research? This article explores the technology behind GenAI, examines strategies for reducing hallucinations, and highlights its applications in legal research.

## How Does GenAI Work?

The key to making the best use of a new tool is to understand the basics of how the tool operates. A large language model (LLM)—the underlying technology of chatbots like ChatGPT—is a fancy word prediction machine. The LLM has been “trained” by analyzing trillions of word associations found in things like books, internet articles, videos, etc. The words in the source material are turned into numbers, called “tokens.” The associations between tokens are assigned specific weights during training. The associations between words are further adjusted by humans through a process called “fine-tuning” to arrive at the final weights—the strength of connections between words. The total number of connections, called “parameters,” determines the size of the final model. Many variables can be adjusted to create an LLM optimized for generating computer code, translating old English, or writing legal documents.

A general purpose LLM, such as Google’s Gemini, which has billions of parameters can respond to inquiries in any domain of human knowledge. They are useful for general tasks such as drafting a letter, summarizing a large document, brainstorming creative ideas, or generating structured outputs. A general purpose LLM should not be used for specific knowledge domains where nuance and ambiguity matter. Querying Gemini for a general summary of the legal doctrine *fruit of the poisonous tree* is more likely to provide a reliable response than asking the model to summarize the dicta from a Supreme Court opinion. A general-purpose LLM has broad knowledge across many topics but limited depth in each area.

## Improving the LLM

Because LLMs are word prediction machines, the greatest danger is a plausible sounding untrue statement. There are two primary approaches to improving reliability in the legal research environment: fine tuning and retrieval augmented generation (RAG).

The fine-tuning approach revises the model weights, similar to how a general LLM is trained. Here, the LLM views a curated set of legal material. The model adjusts the word associations

based on the legal texts. Humans review sample outputs and reinforcement of good responses further refines the model. The result is a LLM that is more likely to respond with accurate outputs relevant to legal questions.

RAG is the other common method of improving the reliability of responses. Legal research tools that incorporate RAG first compile a knowledgebase of reliable information such as statutes, cases, etc. When a user enters a prompt, the research tool first selects relevant snippets from the knowledgebase. It then sends the user’s prompt and the relevant snippets to the LLM. The goal is to improve the reliability of the LLM’s response by supplying it with information to help synthesize a response.

## Type 1 and Type 2 Thinking

The conversational capabilities of LLMs captured the imagination years ago. Engaging with a machine that appears to communicate naturally has led users to infer a capacity for genuine thought. Examining human cognition offers a valuable framework for evaluating the reasoning capabilities of LLMs.

Psychology has, for years, developed a distinction between two types of human thinking. When a person is making quick, instinctual decisions, they are engaging in Type 1 thinking. When asked to sum two times two, most people will instinctually blurt out “four!” In contrast, Type 2 thought engages the logical and analytical thinking that must step through an analysis to arrive at a conclusion. When asked to sum 17 times 26 (without a calculator), a person would use pencil and paper to write out the multiplication function and arrive at an answer.

An experienced legal researcher could likely provide an overview of the exclusionary rule without much thought, demonstrating Type 1 thinking. Distinguishing between dicta and holding in a Supreme Court opinion requires Type 2 thinking. The difference between Type 1 and Type 2 thinking is important when evaluating GenAI legal research tools that incorporate LLMs. The current set of legal research tools are only capable of Type 1 style responses. This limitation is the result of the underlying technology described above. Type 2 thinking requires

an ability to connect concepts and abstractions—something a LLM cannot do because its associations are based on numbers.

Recently a new type of LLM has emerged that is often described as a “reasoning” model. These models mimic Type 2 thinking by including special instructions with the prompt. The result is a step-by-step analysis before a conclusion. This strategy has shown improved accuracy in certain types of tasks such as math and computer coding, but knowledge-based tasks remain prone to hallucinations.

### Thinking Required for Legal Research

There are many facets of legal research. The types of questions researchers work on day to day vary depending on the setting. At a government law library, many questions arise from experts such as judges and lawyers. Other research needs come from members of the public who have a legal problem they are trying to solve. Experienced practitioners are comfortable engaging with primary sources and dense legal treatises. Members of the public likely find more value in secondary sources such as legal encyclopedias or self-help articles written for a general audience. At the beginning of the reference interview, the librarian considers the types of resources that may be responsive while engaging in Type 2 thought.

The inability of LLMs to engage in Type 2 thought is the primary reason they are not of much use to the knowledgeable researcher. A GenAI tool can easily find a resource that directly answers an inquiry, just as traditional search methods can. However, identifying a less obvious but potentially persuasive authority requires Type 2 thought—something beyond the capabilities of an LLM.

### LLMs in Legal Research

As the initial novelty of talking to a human-sounding chatbot wore off, the focus of many turned to the question of “What tasks are these LLMs useful for?” Rather than serving as an all-encompassing solution, time and experience have demonstrated that current LLM technology is best suited to tasks that align with Type 1 thinking. Here are a few legal research tasks

where incorporating LLMs into workflows can enhance efficiency:

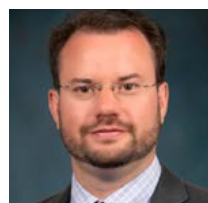
- **Overviews of topics where there is existing authoritative material.** Often a researcher’s existing knowledge of a topic may only be cursory. When faced with a narrow facet of a niche topic, beginning with a GenAI summary is often a helpful first step in understanding the relevant legal landscape.
- **Suggest steps for how to research a particular resource type.** For example, Gemini provides an accessible and user-friendly summary of how to search Maryland legislative intent materials. Though an expert in this topic would dispute a few citations and offer some critiques, the outline is helpful for a novice getting started.
- **Confirming a lack of sources.** Some research questions are truly novel and there are no primary or secondary sources that apply the law to the facts of the present case. At this stage, it is disconcerting to a researcher to come up empty. GenAI can help confirm that no obvious alternative research sources have been overlooked.

### Final Thoughts

The distinction between Type 1 and Type 2 thinking is crucial in evaluating the reliability and usability of GenAI legal research tools. While LLMs are limited to Type 1 responses and are less useful for knowledgeable researchers, they can still provide value in specific areas. The rapid evolution of GenAI research tools requires researchers to remain informed about advancements of the underlying technology. Newer reasoning models may improve at emulating Type 2 thinking. Experimentation with these GenAI tools is the only way to find the most beneficial uses. Although LLMs may never engage in true Type 2 thinking, their current capabilities can enhance legal research efficiency when used appropriately. ■

---

 Research + Analytics



**LELAND SAMPSON**  
**HEAD OF WEB CONTENT AND SERVICES**  
Thurgood Marshall State Law Library  
Annapolis, MD  
leland.sampson@mdcourts.gov

© 2025 BY LELAND SAMPSON