Insights

LLMs for Legal Tasks: Best Practices



vlex.com

Damien Riehl

Damien Riehl is a lawyer and technologist with experience in complex litigation, digital forensics, and software development. A coder since 1985 and for the web since 1995, Damien clerked for the chief judges of state and federal courts, practiced in complex litigation for over a decade, has led teams of cybersecurity and worldspanning digital forensics investigations, and has led teams in legal-software development. Co-Chair of the Minnesota Governor's Council on Connected and Automated Vehicles, he has helped recommend changes to Minnesota statutes, rules, and policies - all related to connected and autonomous vehicles. Damien is Chair of the Minnesota State Bar Association's working group on AI and the Unauthorized Practice of Law (UPL). At SALI, the legal data standard he helps lead, Damien develops and has greatly expanded the taxonomy of over 14,000 legal tags that matter, helping the legal industry's development of Generative AI, analytics, and interoperability. At vLex Group – which includes Fastcase, NextChapter, and Docket Alarm – Damien helps lead the design, development, and expansion of various products, integrating AI-backed technologies (e.g., GPT) to improve legal workflows and to power legal data analytics.





Introduction

Executive Summary

This paper presents an analytical overview of some best-practice strategies for using Large Language Models (LLMs) to perform legal tasks — reducing and resolving concerns such as inaccuracy and confabulations ("hallucinations") — to realize LLMs' potential to transform legal practice. In particular, the paper argues that systems must enable users to more easily "trust but verify" the LLMs' outputs. That ease of use — through the User Interface (UI) and User Experience (UX) — should help ensure that human lawyers actually perform that necessary "trust but verify" step. The paper also explores the role that LLM-based legal tools can play in legal argumentation, jurisdiction comparison, and user-centric approaches to providing answers — outlining how integrating Generative AI into legal practice can redefine legal tasks by helping lawyers move beyond traditional tools.



Introduction

LLMs in Legal Work Opportunities amidst Challenges

More than a year after November 2022, when OpenAI introduced ChatGPT to the world, the legal industry continues integrating Large Language Models (LLMs) into legal practice – a development of increasing complexity and importance.

This white paper presents an analytical overview of best practices to shape LLM use in legal tasks; balancing practical realities and challenges with their potential benefits and massive potential to transform the industry. That is, if we do it right.



The Promises and Challenges of LLM Tools in Legal

In 2023, LLMs significantly shifted most organizations' approaches to legal work, introducing advanced capabilities with the potential to profoundly influence legal practice.

Lawyer efficiency. The LLM transformation occurs amidst evolving client expectations for efficiency and cost-effectiveness. Traditional models of legal research and documentation face increasing pressures. In this context, LLM-based tools can be pivotal, providing a blend of speed, accuracy, and comprehensive analysis that outpaces conventional methodologies.

Lawyers' work is entirely language: We read, we analyze, and we write. That aligns well with LLMs' core competencies, which can understand text with post-graduate proficiency. And they can read text with superhuman speed. Same with writing – they do it faster than any human can. Today's LLMs are really good at all of those things. Tomorrow's LLMs will be even better. **So LLMs' proficiency in language processing – ingesting, analyzing, and outputting – offer significant potential for legal applications.**

When LLMs process text — interpreting and analyzing disparate concepts — the industry refers to that as "interpolation." That is, an LLM can review Concept 1 and Concept 2, then provide interconnected, related concepts through interpolation. In doing so, LLMs can create coherent, contextually appropriate outputs.

An example of the above:

- Concept 1 = Negligence
- Concept 2 = Fiduciary Duty
- Concept 3 = Board Members
- Concept 4 = Pirate Ship

LLMs can then interpolate those disparate concepts. Here is an actual example of these concepts from GPT-4:

• Board members, akin to captains of a pirate ship, must avoid negligence by acting with care and loyalty, thus fulfilling their fiduciary duty to prioritize and protect the company's interests.

That combination of disparate concepts constitutes interpolation. That interpolation is the essence of what the best LLM systems do today. And interpolation also provides the roadmap for what the best LLM systems will become tomorrow.

But the promise of LLMs in Legal is moderated by concerns over their limitations, including the challenge of ensuring that they generate contextually accurate outputs. Despite LLMs' significant promise for legal work, lawyers and legal professionals are approaching them with cautious consideration. A notable and valid concern is the propensity of LLMs – left unchecked – to produce outputs that may not be entirely accurate or fact-based. This phenomenon is often labeled "hallucination." In legal contexts, where precision and factual integrity are sacrosanct, this aspect of LLMs is particularly unnerving. And that's why addressing hallucinations is seen as Job One.

The Approach: 'Trust but verify'

This paper examines the best practice strategies to address these concerns, focusing on "trust but verify". For decades, partners working with their associates and paralegals have trusted those associates and paralegals — but still verified the associates' work. That "trust but verify" is essential in legal practice.

This is also true with LLMs. Lawyers must trust but verify the LLM output — with the ground truth. So LLMbacked systems need to ensure that "trust but verify" is dead simple. Because "the devil is in the defaults." And if the default is "verification is hard," then users won't do it. Then bad things can happen.

In other words, users should take advantage of LLMs' capability to efficiently process and synthesize information, while concurrently validating LLM outputs against established legal sources: judicial opinions, statutes, and regulatory texts.

Done poorly, a time-consuming "trust but verify" approach hazards giving users the feeling of "Why did I even use an LLM in the first place, if I have to do all the work anyway?" But done well, a "trust but verify" system can greatly expedite legal work — much faster than traditional methods.

So those companies that build legal systems that integrate LLMs into legal practice must walk a fine line: leverage LLMs' myriad benefits, while simultaneously making it easy for users to validate ground truth. Because if it's not easy, users (e.g., lawyers) won't do it.



What a Successful LLM-Based Tool Looks Like

User in control. Successful LLM systems must permit and encourage user-enabled experiences. Humans in the loop. Those systems must simplify the verification process. They must provide more control over research resources. And they must offer contextual summarization capabilities. If LLM systems do that, they can both address common user fears, as well as help users work in tandem with their LLM counterparts.

Data = Oil. This paper also underscores the need for underlying data. Data is the new oil. And legal practice is increasingly global. So to be effective, lawyers must have global oilfields – providing legal advice across jurisdictions. As such, LLM-based tools must similarly be cross-jurisdictional.

Multi-jurisdictional surveys. For decades, lawyers have relied on 50-state surveys. But in our global economy, more necessary are 50-country surveys. To survey 50 countries, you need data from those 50 countries: statutes, regulations, judicial opinions. If you only have United States oil, you cannot refine legal advice globally.

What is a modern legal tool? This paper also explores a modern LLM system's role in legal argumentation, jurisdiction comparison, and user-centric approaches to providing answers. Those legal tools should be flexible in selecting foundational models — not being confined to any single AI model — enabling any system to stay abreast of technological advancements, ensuring continued state-of-the-art efficacy.

Throughout, this paper discusses ways that legal tools can harness LLMs' significant advancements – likely the largest legal-technology advancement in our lifetimes. With these tools, we can move beyond traditional tools, effectively integrating Generative AI into our legal practice. This has profound implications for the legal profession, and it can redefine the standards in legal research and practice.



Lawyers Must Do More with Less Resources

As legal professionals continue walking down the road of our LLM futures, we must address a major challenge that has been brewing for decades: the increasing expectation for lawyers and legal researchers to deliver better results with fewer resources. That's true for external counsel (e.g., firms) as well as internal counsel (e.g., law departments). The current environment, marked by time constraints and budget consciousness, demands an approach to legal work that is both efficient and effective.

For example, traditional legal research methods are characterized by extensive manual review of documents and texts – a method that today's post-LLM work demonstrates isn't sustainable. Clients' reluctance to bear the costs of prolonged research, particularly when such expenses can be expedited, is driving a shift in lawyers' behavior. The winners in legal-research innovation will offer solutions aligning client expectations and research efficiency.

Systems must enhance the legal research process by significantly reducing the time required for information gathering and analysis. They must also utilize state-of-the-art LLM models, permitting effective navigation through massive volumes of legal data, pinpointing relevant information swiftly. Expediting research with rapid turnaround times can address clients' hunger for innovation.

Importantly, these systems' increased efficiency cannot compromise research quality. Accuracy and thoroughness remain paramount. **So systems that win in the marketplace must be fast and reliable**, **balancing speed**, **comprehensiveness**, **and precision**. And contemporary legal researchers must deliver detailed and accurate legal insights — in a way that's time-efficient and economically viable.



LLM Pitfalls to Avoid

The current legal-tech landscape has several products seeking to automate and expedite legal research. But many of those LLM tools repeat the problems of legal-research past, rather than pointing to LLMs' legal-research future. LLMs shouldn't be your grandmother's legal-research tool.

PITFALL: Not providing insight into the process used to achieve outputs

Black Box: Take it or leave it. Before LLMs, traditional legal-research tools had an input-output method: Users submit a query, and the system provides results.

After LLMs, most legal-research providers have, sadly, continued that input-output method — without rethinking how the process might work. For example, in many LLM-based legal-research products, users prompt (not question), and the legal-research systems output a completed memorandum.

But for most legal-research systems, that process is a black box. Those systems provide no insights into how they built the memorandum. They show only the cases in the memorandum — not most of the cases that the system considered. They also don't show the particular paragraphs that were most influential in creating the memorandum. Instead, those systems' black boxes essentially tell the user, "Here's the output; we're not telling you how we got there — take it or leave it."

This approach, while initially seeming efficient, presents notable challenges: How do users easily verify ground truth? And why can't users have more control over their research processes?

PITFALL: Making it difficult to access and verify primary sources

"Trust but verify" as friction. A significant challenge with many existing legal LLM systems is validation of ground truth: judicial opinions, statutes, regulations. Often, the crucial data needed to establish that ground truth is embedded deep within hyperlinked documents (e.g., 20-page opinions). Those systems require users to engage in a time-intensive and laborious effort to click a link, to locate specific, relevant information within these documents, go back to the memorandum, and then repeat the process. For dozens of sources. That's a pain. It's hard. So users won't do it.

In software development, this is called "friction." Any aspect that impedes the user's journey – causing delays, confusion, or frustration – should be avoided. This includes complex navigation such as having to click on links, read lengthy documents, and then return back to source documents. The best designers minimize or eliminate friction.

PITFALL: Limiting user control

Nonexistent user control. Many systems frequently limit user control. Users should have control over which judicial opinions, statutes, and regulations the system identifies as relevant. But sadly, most systems provide no ability to modify or tailor these sources. This lack of control can be particularly limiting in complex legal situations where users need to have control over sources that the users deem "irrelevant" versus those that they deem "authoritative."

To counteract those problems, a modern LLM-based legal-research system should facilitate easier verification, allowing users to customize sources. User experience is everything. And enhancing user control over the selection and exclusion of sources, while reducing the burden of sifting through extensive documentation, is an essential feature in embracing the promises of LLM-based legal research.

Features of an Ideal LLM-based Legal Research Tool

Easy Information Verification

Frictionless "Trust but verify." The tool should enhance efficiency and reliability. Verification should be easy. Friction should be reduced or eliminated. Relevant excerpts of case law or statutory language should be displayed right next to the user's question. Users moving their eyeballs from one part of a page to the other? Zero friction. Of course, users should always read the entire source document (e.g., case, statute, regulation), but from an initial "is this along the lines of what I seek," systems should minimize extensive navigation between documents. Instead, "trust but verify" should be streamlined: What you need on the same page. The best systems enable quick assessment of the source's relevance and accuracy.

Individual source summaries. Provide individualized summarization, showing how each case, statute, or regulation answers the user's question. LLMs can offer concise, customized summaries that detail how a specific case or legal provision relates to the user's query. Strangely, such functionality is rare or unique.

Confidence scores. Give users an indication of how well each particular source answers the user's question. Confidence scores can quantitatively assess alignment of the case or law with the user's intent, providing a clearer understanding of the source's applicability and reliability. Again, in today's legal-research tools, confidence scores are rare (or unique).

User control. Place humans in the loop, enabling users to exert some control over their legal-research process. Unlike systems where users receive output passively, a modern legal-research system should allow users to be actively involved in shaping the research output. Users should be able to selectively include or exclude particular cases and laws, tailoring the research to meet that user's specific intent. Furthermore, those systems should give users easy, direct access to the source material, enabling a deeper exploration of the full text for those who seek a more thorough understanding. Not just "select" material – but all material that the system considered. And again, everything should be frictionless: easy.

By incorporating these features, a modern system can make "trust but verify" easy, making LLM-based tools both efficient and effective. By reducing users' information-verification effort, those systems will permit greater customization, thereby improving user experience. LLM-based tools should be fun and effective. The best of those systems provide both.



Worldwide Jurisdictions

Jurisdictional coverage is critical when considering legal research tools' utility. To do research, you need relevant, updated sources. And to research in a particular jurisdiction (e.g., court, agency, country), you need that jurisdiction's data. For example, to do research about EU law, you need the EU statutes.

Global reach. Many existing legal-research products focus on the United States, but when customers have cross-jurisdictional presence, they frequently need access to cross-jurisdictional law. Legal-research companies that differentiate themselves through comprehensive global coverage will serve customers better than those who consider non-U.S. jurisdictions as an afterthought. Covering legal content from over 100 countries permits a legal-research company to offer quality legal information to professionals around the world.

Global LLMs need global data. Once a company has amassed global data, then that data can be used for the purposes of LLMs and Generative AI. For example, Vincent AI is designed to perform advanced, vector-embedding searches across many jurisdictions. It can answer questions across various legal systems, giving users precise legal information, irrespective of geographies or language. This capability is increasingly valuable in a globalized legal environment where clients' issues often span multiple national boundaries.

Data is the new oil. Legal oil is statutes, regulations, and judicial opinions. Across this oil, LLMs are refineries. Performing LLM-based tasks requires underlying data: the oil of statutes, regulations, and jurisdictions – across many countries. To date, vLex has amassed perhaps the broadest and deepest legal oilfields in the world. That deep oilfield permits similarly deep refining – through LLM-based legal analysis, synthesis, and generation.



Jurisdiction Comparisons

In both litigation and transactional work, lawyers must frequently compare jurisdictions. This can be true across the 50 states, and it can also be true across 50 countries. In today's global world, clients' legal needs rarely stay neatly within a single jurisdiction. Legal needs span borders.

In litigation, lawyers may argue the applicability of laws in a more favorable jurisdiction (e.g., New York, California), based on differences in procedural rules and substantive laws. As such, effective jurisdictional arguments can significantly improve case outcomes.

In transactional work, companies operating in multiple states need a 50-state survey to align with varying state laws. And in our global world, multinational companies may need 50-country surveys. This can include analyzing state-specific or country-specific regulations affecting business operations. The best lawyers, therefore, provide not only legal analysis but also practical, compliance-oriented solutions – across all of a client's geographic locations.

Lawyers frequently combine cross-jurisdictional analyses. Consider, for example, a lawyer handling multistate litigation, efficiently determining which state's legal framework is most advantageous for a specific argument. Similarly, a privacy lawyer who needs to have updates on data-protection laws can examine how it's approached in multiple countries. Finding and analyzing diverse legal standards across jurisdictions – including cross-border considerations – separates good lawyers from great lawyers.

As such, LLM-based tools' permitting comparative jurisdictions transcends mere convenience; rather, it's foundational for advanced legal analysis and strategic planning. Legal professionals leveraging such a tool can incorporate perspectives from multiple legal systems, providing their clients with more-valuable representation.

Any modern LLM-based tool that permits comparison across multiple jurisdictions marks a significant advancement in legal research methodology. It shifts the paradigm from isolationism toward holistic, global representation. As such, legal professionals can undertake more thorough and well-informed research that reflects their clients' global legal needs.



Building Arguments

Modern LLM-based research tools focus not just on delivering objective answers, but also on crafting subjective arguments. For example, lawyers frequently need to support their arguments, opposing other parties' arguments. This is helpful, of course, when arguing against opposing counsel's propositions. But by inputting their own arguments and discovering potential opposing views, users can also effectively "shadowbox" with **their own** arguments, incorporating anticipated counterpoints into their initial strategy.

Supporting arguments. The best LLM-based tools' ability to not only find objective answers, but also to build and analyze subjective legal arguments, can distinguish them from traditional legal-research tools. Legal practitioners can craft persuasive arguments by identifying and combining relevant legal precedents, statutes, and secondary sources. This feature can permit easy and effective legal narratives, reinforcing practitioners' stances with substantiated legal analyses.

"Shadowboxing" your arguments. LLM-based tools can also help users critically examine their arguments through a process akin to "shadowboxing." By entering in the user's own arguments, LLM-based tools help by providing opposing viewpoints. Through this type of shadowboxing, legal professionals can adopt the perspective of opposing counsel, challenging their own propositions. As such, these tools can provide insights into potential counterarguments, sourced from diverse materials that their opposition might use (e.g., cases, statutes, regulations). This shadowboxing feature can be instrumental in enabling legal professionals to develop more comprehensive and resilient legal strategies by anticipating and preparing for counterpoints. Allowing users to identify and address potential vulnerabilities in their case can enhance lawyers' arguments – both in their strength and in their persuasiveness.

The integration of argumentative capabilities represents a significant evolution in legal research tools. Moving from solely information retrieval to an interactive, analytical process, these AI-powered tools considerably improve the tools' scope and quality. Traditional tools permit **finding** potential arguments; LLM-based tools permit **building** arguments and strategy.



Answers, not a Chatbot

The smartest legal-tech companies take the position that lawyers should not need to learn to become "prompt engineers." Rather, lawyers should be able to use tools that understand normal legal questions — questions that they've been asking for decades. And those systems should be able to provide those legal professionals with answers — like memoranda — much like they have received for decades.

Paradox of chatbots: Paralyzing choice. Some LLM-based legal-research tools give users a "blank slate." The blank chatbot box can lead to an overwhelming array of choices. "I can do anything." This gives the paradox of choice: "I can do anything, so I'm paralyzed — I don't know what to do." Lawyers must be prompt engineers.

Modern LLM-based research tools, in contrast, let users ask questions just like they've asked for decades. "What is the trade secret law regarding former employees taking customer lists?" That's not a prompt; it's a question. LLM-based research tools should be able to answer that traditional question. This adaptive approach ensures higher-quality results — and a better user experience. Lawyers won't need to be "prompt engineers."

Questions lead to answers. Legal professionals seek direct answers to specific legal questions. So advanced legal-tech solutions can now interpret these inquiries, leveraging deep legal knowledge bases and sophisticated prompting — not by the users, but by the tech companies — to provide clear, accurate responses like traditional legal memoranda. Lawyers simply lawyer.

In this context, LLMs are not replacing the lawyer's expertise; they're augmenting it. By finding, analyzing, and outputting caselaw, regulations, and statutes, modern LLM-based systems can generate nuanced, context-aware responses. As such, lawyers can concentrate on asking questions, and getting answers — not prompt engineering.



Model Agnostic: Hot-Swap the Best Model for the Task

The most-frequent question that customers ask LLM-based toolmakers: "What model are you using?" The best companies answer "Not one but many — in an ensemble approach." Diversity is king, not only in natural selection and workforces, but also in LLM model selection.

Some companies rely solely on a single AI model, like OpenAI's GPT-4. Putting all of their eggs in that basket. But wiser companies take an ensemble approach, utilizing the best model for the job — at that time. At the time of this writing, GPT-4 does many things very well, but that won't be the case forever. Other models will likely ascend. And when they do, those reliant on a single model will wish they'd built in more flexibility.

So the best LLM-based research tools will be model-agnostic, evaluating and utilizing an ensemble of LLM models. This strategy permits selecting the "best model for the task." If one company's service goes down, this approach also allows the company to switch between models. And as the competitive landscape of LLMs evolves, model agnosticism ensures that the LLM-based tool can use the most capable and suitable LLM for each task.

Different LLMs have unique capabilities, making them more suitable for certain types of legal research. For instance, while OpenAI's GPT might excel in general language processing, other models like Anthropic's Claude or Meta's LLaMA 2 might be better for tasks with specific nuances. By using a diverse array of AI models, these tools ensure the most suitable LLM is used for each task, improving accuracy and relevance of research outputs.

Adaptability is crucial in our rapidly evolving LLM landscape. With near-daily advancements and new models emerging with improved functionalities, the ability to shift between models ensures the use of the most advanced and effective LLMs, giving a significant edge in legal research capabilities. A model-agnostic approach means LLM-based tools are not limited by the constraints of any single model. Instead, those tools benefit from the gestalt – the combined strengths of various models – consistently offering an effective suite. This allows those LLM-based tools to be versatile, capable of adapting to and leveraging ongoing Al advancements.



Data is Oil: LLMs are Refineries, But You Need the Oil

For legal technology, data is oil: a crucial resource necessary for doing LLM-based work. Over the decades, the best companies in this field have compiled extensive legal databases, comparable to vast "legal oil fields," encompassing over a billion legal documents from dozens of countries.

LLMs are refineries, requiring substantial legal datasets to function effectively. The process of refining legal data mirrors oil refinement.

Extract the oil. Of course, amassing a legal oil field requires extracting legal data from various sources. This includes public records like judicial opinions, briefs, pleadings, motions, orders, legislations, statutes, regulations, and secondary sources. Along with private data from law firms – their internal documents and contracts – this combination provides the best of both worlds: Private Oil and Public Oil.

Refine the oil. After amassing the oil, the refinement phase involves deciding the form (e.g., gasoline, jet fuel). For legal data, refining involves categorizing, tagging, and summarization — making the legal data optimized for LLM processing. In the past, for *incumbent* legal publishers, this required thousands of humans. Today's LLMs make those human taggers and summarizers unnecessary: LLMs can often tag and summarize legal documents more quickly and more accurately than humans. As such, forward-thinking legal-technology companies can leapfrog traditional companies by offering better products at more-reasonable prices. Refining legal oil with LLMs is transformative.

Build products (e.g., plastic, toys, medicine). After refining, the legal-data oil must be converted into usable products. In legal tech, this means converting the data into case summaries, legal precedents, or annotated statutes — ready for AI analysis. Again, with today's LLM-based tools, this can be done more quickly and efficiently than ever in human history.

Build the right products. After refinement, this data can be provided to legal practitioners in both Web interfaces and through API feeds. Through this step, the refined data is transformed into practical tools and services, combining legal expertise with "answer" tools. Doing that requires product-development and user-experience expertise.

When building LLM-based systems, lawyers are required to know what lawyers need. If your legal-research tool doesn't heavily involve lawyers who practiced, beware. To build a good car, designers need to be drivers. "Where should the steering wheel sit?" "How about the mirrors?"

Similarly, legal-tech companies should have lawyers with practice experience on staff, guiding the product's design. Not having former practicing lawyers on the Product team is akin to carmakers saying "We've never driven, but we have been passengers!" The best legal tools greatly benefit from their builders being former practicing lawyers themselves. So every step of the way, those lawyers can ensure that the users' needs are met.

Go to market with a strong reputation. The best companies then bring their refined legal data tools to market, which requires building a strong reputation, fostering client relationships, and assuring the legal community that your system is trustworthy.

In short, legal-technology companies' sophisticated refinement processes mirror the oil industry. By effectively utilizing large legal datasets, these companies create a conducive environment for LLMs, offering advanced services in legal research and analysis.

Conclusion

Modern LLM-based legal research tools represent a significant shift, revolutionizing how legal professionals interact with data. Researchers have gone from "finding" to "answering." And the best of these tools also use advanced retrieval augmented generation and easy source viewing — eliminating hallucination risk by permitting users to easily perform "trust but verify."

These LLM tools — including easy "trust but verify" — can lead to more accurate and reliable results. Enabling user control over the sources, as well as the ability to directly compare findings with original legal texts, all enhance users' convenience and confidence.

Tools that have wide jurisdictional reach (e.g., 100 countries) will win over mono-jurisdictional tools, especially in today's globalized legal landscape. Permitting comparative law analyses across regions, these tools can offer what would manually take weeks — and would require constant updating. The strategy of employing a diverse ensemble of LLMs ensures that the best-suited AI technology is used for specific tasks, adaptable as the competitive landscape evolves.

Rich datasets lead to rich tools. And if those tools are underpinned by vast legal document repositories, then those tools will provide the depth and breadth of data needed for effective LLM insights. As LLM tools go from "finding" to "answering" to "arguing," the tools can extend beyond simple data retrieval, supporting legal professionals in strategic thinking and argument development.

The impact of modern LLM-based tools on the legal profession is substantial, enabling legal service providers to meet their clients' growing demands for efficiency and accuracy — and setting new standards in how modern legal research is employed. These advancements are transforming professional practices, changing professionals' interactions with legal information, enhancing their decision-making capabilities, and ushering in a new era in legal technology intelligence.





vlex.com