MURPHY CONSEIL

Prompting Manual for Legal Professionals

Artificial Intelligence and Legal Practice

Tudual Lucas Huon

March 2026

murphyconseil.com

In memory of Mona Heydari

Acknowledgments

Thank you to my mother-in-law, Pascale Keraudran, for her unwavering support over the past years, through thick and thin.

Thank you to Maître Esther Laudy, attorney at the bar of Saint-Malo, and Maître Nina Gourvennec, attorney at the bar of Rennes, whose practice of the legal profession inspires me daily and continually gives me new ideas, and without whom this work would never have seen the light of day.

Thank you to Valentin Thomas, legal professional at Crédit Mutuel, whose support and constant energy push me daily to exceed my limits.

Thank you to Marie Le Vivier, GDPR legal professional at Ouest-France, for her daily presence and invaluable advice.

Thank you to Abir Adam, whose ambition challenges me daily.

About the Author

Tudual Lucas Huon

AI Engineer · Legal Professional · Founder of Murphy Conseil

I design artificial intelligence architectures and conduct research on the alignment problem. How can we ensure that a language model faithfully conveys the user's intent without distortion, omission, or extrapolation? This is one of the central challenges of contemporary AI. Each of my projects is designed as an investigation ground for the real functioning of these systems.

My work focuses notably on persistent memory architectures for large language models as well as the risks associated with prolonged interaction with these models. My article *User Imprint: Psychological Profiling and Qualified Information in Prolonged Interaction with Large Language Models*, published on SSRN in March 2026, formalizes the concept of *user imprint*: the capacity of an LLM to aggregate, through exchanges, an exploitable psychological profile of its user. It introduces the notion of *qualified information*, extending Shoshana Zuboff's theoretical framework of surveillance capitalism.

My research takes me to the four corners of the world. Currently in New York, I will be settling in September in South Korea for a minimum duration of three years, to study there the influence of artificial intelligence on legal doctrine from the perspective of comparative analysis between French law and South Korean law.

A legal professional by training, holder of a Master's degree from the University of Brest in justice, litigation and procedure, a university diploma from Paris 1 Panthéon-Sorbonne University in studies of judicial practices, and the CRFPA

(specializing in criminal law), qualified for the magistracy examination, I practiced for nearly two years at the criminal division of the Rennes prosecutor's office, where I drafted motions, appeal briefs and legal summaries alongside prosecutors.

It is at the intersection of these two paths that I founded Murphy Conseil, a consulting firm specializing in strategy and artificial intelligence for legal professionals. The first edition of this guide achieved significant success on LinkedIn, with over 300 downloads in less than a week and a reach of nearly 20,000 professionals, 49% of whom are legal professionals from the Department of Justice, the National School of the Judiciary, and the Paris Bar.

The firm's activities rest on three pillars:

Restraint. Offer only what is useful. No oversized solutions, no technology for technology's sake. Each intervention is calibrated to the professional's real needs.

Mastery. Ensure that the professional understands what they are using. No black boxes, no dependency: the goal is that each client is able to operate independently at the end of the engagement.

Transfer. Transmit skills rather than withhold them. Murphy Conseil does not create dependency: it trains, it equips, then it steps back.

This guide is the embodiment of this philosophy. It was designed to give legal professionals the keys to real mastery of artificial intelligence, with no intermediary and no filter.

Table of Contents

Introduction

For nearly four years now, I have been using generative artificial intelligence tools daily. As a trained legal professional, I have progressively made these tools a central axis of my professional practice, developing expertise in prompt engineering through ongoing experimentation. I have deliberately pushed the use of these tools as far as possible—including into its most counterintuitive corners—in order to map precisely the capabilities of these models, as well as their limitations.

This guide is addressed to legal professionals: judges, judicial auditors, attorneys, and law students. It responds to a fourfold observation.

The first is **professional** in nature. Specialized publishers are multiplying legal AI solutions, sometimes at prohibitive costs, without users having the understanding necessary to evaluate their reliability. Before delegating part of one's reasoning to a machine, it is essential to understand how it works, where it excells, and above all where it fails. In the United States, in the case *Mata v. Avianca, Inc.* (2023), attorneys were sanctioned after arguing on the basis of judicial precedents entirely fabricated by ChatGPT: invented decisions, fictitious reference numbers, formulations imitating the style of the court in question[1]. This case is not an accident: it is the foreseeable consequence of a powerful tool used without understanding of its mechanisms.

The second is **intellectual** in nature. Several studies warn of the risk that prolonged and unreflective use of artificial intelligence will gradually erode the autonomous reasoning capacity of its users. A point to which I will return several times in this guide. This danger is particularly serious in the legal professions, where reasoning is the raw material of the profession. Whatever the machine provides, the professional remains solely responsible for how it is used.

The third is **civilizational** in nature, and perhaps the most important. By sharing my knowledge in this field, I also wish to awaken awareness of what artificial intelligence now makes possible, and of the choices that this power imposes. Increasingly, AI solutions—

1. *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443 (S.D.N.Y. 2023), Judge P. Kevin Castel.

including legal ones—are moving toward an *all-AI* model, where the professional is progressively consumed in decision-making before being replaced by the systems they themselves helped train by providing, interaction after interaction, their intellectual output. The objective of this manual is to defend a radically different vision: that of the **centaur**. The term, borrowed from Greek mythology, was transposed into the field of artificial intelligence by Garry Kasparov after his defeat against Deep Blue in 1997: it designates the partnership in which human and machine work together, each amplifying the other's capabilities. AI in service to humanity, alongside it, and never without it.

Fourth, through this guide I wish to initiate a **broader reflection** on AI and its uses. The revolution underway will challenge the realities upon which our contemporary societies had founded themselves and had acquired a relative stability. Before even using AI, I believe we must *think about* artificial intelligence. Accepting this philosophical work now will allow us to become aware of the territories to which this technology might well take us. In *The Human Condition*, Hannah Arendt wrote:

> *"What lies before us is the perspective of a society of workers without work, that is, deprived of the only activity that remains to them. Nothing worse can be imagined."*

I would say that this reality has never been closer to us. What lies before us is the perspective of a society of workers without work, a society of thinkers without subject matter, a society of citizens without free will—that is, a humanity deprived of everything that makes it human. And it is possible to imagine something worse.

If I have chosen to address legal professionals, it is not by chance. We live each day the concrete consequences of a normative text—even when it has been drafted hastily—without measuring its effects, and yet it applies nonetheless. Moreover, the legal professional has been trained to weigh each word with precision and restraint. This experience places us in a particular position to understand what is at stake today: a technological power that transforms social relations at an unprecedented speed, and which remains insufficiently regulated.

Key Point

It is not a matter of treating artificial intelligence as a threat. It is a matter of being aware of one thing: from your choices will flow a reality. Either AI will be there to serve you, or you will become the servants of AI. In the world taking shape, the boundary between freedom and servitude has never been more porous.

Preliminary Definitions

Before delving into the heart of the matter, it seems to me essential to define certain fundamental concepts.

The generative artificial intelligences discussed in this guide—ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google), Le Chat (Mistral)—belong to the category of **Large Language Models** (LLM). Their operation is based on a fundamentally statistical principle: given a sequence, the model predicts the most probable next word, then the next, and so on. To use a simple analogy: when you are asked what you have for breakfast, your mind spontaneously conjures up "coffee," "toast," "croissant." An LLM operates in a comparable manner, but through statistical calculation rather than lived experience.

This operation explains both the remarkable effectiveness of these models and their propensity to produce errors, or even to invent facts entirely, a phenomenon known as "**hallucination**."

A **prompt** designates the textual request submitted to the model. Its precision directly conditions the quality of the result: an ambiguous term, an imprecise formulation, and the response will suffer mechanically. Computer scientists even have an adage, formulated as early as the 1950s: "*Garbage in, garbage out*": low-quality input inevitably produces low-quality output.

From this observation emerged **prompt engineering**, a discipline that I would situate at the intersection of statistics, law, and psychology, because an LLM, trained on human data, inevitably inherits many of our cognitive biases. This guide aims to establish its foundations in a legal context.

I.
Understanding the Tool: From Architecture to Implementation

Before formulating any prompt, it is essential to understand what you are dealing with. This first section describes the fundamental mechanisms of large language models: tokens, context window, memory (A). The risks associated with what the machine accumulates as information about its user (B) and the cognitive biases inherent in their architecture (C).

Lack of understanding of these mechanisms is the source of virtually all usage errors: one does not operate a tool whose functioning one does not understand.

A. How It Works: Tokens, Context, Memory

To use an LLM effectively, it is necessary to understand three fundamental technical concepts: the *token* (1), the *context window* (2), and *memory* (3).

1. The Token

The token is the basic unit that a language model uses to process text. Each word is converted into one or more tokens depending on its length. "Law" corresponds to one token; "constitutionally" represents three. This unit of measurement is not insignificant: it conditions the entire functioning of the tool, particularly in terms of capacity.

2. The Context Window

The context window designates the total quantity of tokens that a model can process simultaneously, that is, the maximum length of conversation it is capable of maintaining without, theoretically, losing the thread. Earlier models, such as GPT-3.5, had a window of a few thousand tokens, which caused loss of coherence after a few exchanges. Current models have considerably expanded this capacity: Claude Opus 4.6 and Sonnet 4.6 (Anthropic) have a standard window of 200,000 tokens, expandable to 1 million in extended version; GPT-5.4 (OpenAI) accepts up to 1.05 million tokens; Gemini 3.1 Pro (Google) offers 1 million tokens, with 2 million announced; and Mistral Large 3 (Mistral AI) offers a window of 256,000 tokens.

To give an order of magnitude, one million tokens represents approximately 750,000 words, or the equivalent of 2,500 to 3,000 pages. These figures, in constant evolution, are given for reference at the date of writing (March 2026). Even the most modest windows of this generation of models make them analysis tools of considerable power. However, an essential nuance must be noted. Unlike humans, an LLM does not have persistent memory in the strict sense. With each new request, the model resumes the entire context of the ongoing conversation. The longer it becomes, the greater the risk of degradation of responses: the model "dilutes" progressively earlier information in the mass of exchanges.

3. Memory

To overcome this limitation, publishers have developed "persistent memory" systems. At OpenAI, this functionality is called "Memory"; other platforms offer "custom instructions." The principle is as follows: over the course of conversations, the model records certain information deemed relevant about its user—their profession, their preferences, their work habits—and automatically reinjects it into future interactions. For the legal professional, the advantage is immediate: if the model knows it is addressing a criminal law specialist, it will spontaneously adapt its register of language and references without needing it to be specified with each exchange.

To this declarative memory is added a form of conversational memory: as long as earlier conversations are not deleted, the model can draw from them to enrich its responses. It then becomes capable of resuming a search where it had been left off, or of cross-referencing information from separate discussions.

Finally, Anthropic has developed actionable skill points for its AI called "*skills*." Concretely, these are skill sheets containing best practices for a specific use. This can be particularly useful, for example, for drafting standard documents or creating specific and repetitive documents. It is not directly a memory skill, but an aspect of its optimization.

> **Warning**
>
> These features, useful as they are, nevertheless raise risks that must be assessed.

B. What the Machine Knows About You

The memory of an LLM, as described in the preceding section, creates a double risk that must be identified before any professional use.

The risk of unauthorized access. The LLM does not distinguish the identity of the person addressing it. It will respond identically, whether the user is the account holder or a third party. In a professional context, the consequence is immediate: anyone with access to the account can interrogate the model about the entire history: ongoing projects, strategic reflections, research conducted months earlier. This risk is multiplied when the same account serves both personal and professional use.

The risk of involuntary profiling. To illustrate the second dimension of this danger, it is necessary to introduce the concept of *operational response*. When a user formulates a request, the model seeks to optimize its response with a strict objective of efficiency. It has no awareness of good or evil in the moral sense. It seeks neither to harm nor to protect. It is guided by its sole mission: to respond effectively. This functional neutrality, combined with the richness of information accumulated in memory, can produce deeply destabilizing results.

Cas pratique : l'attaque de la machine

For the purposes of this demonstration, I deliberately communicated to ChatGPT, over the course of several weeks of conversations, a set of personality traits attributed to a fictional user. The model progressively integrated this data into its memory. I then asked it to draw up a psychological assessment of that user, identify their vulnerabilities, and then develop an operational action plan to exploit them.

The result was enlightening. The model identified with precision the points of weakness of the user—their perfectionism, their need for control, their sensitivity to public image—and proposed, without any ethical reservation, operational tactics to exploit them. Among these: competence gaslighting ("Are you sure of your projections? They seem a bit unrealistic to me…"), bombardment of temporal surprises to saturate their adaptive capacity, or public social comparison to trigger an impulsive reaction. Techniques which, transposed into a professional setting, would be nothing less than harassment.

Three lessons emerge from this experience.

First, the model is perfectly capable of aggregating information scattered across multiple conversations to build a coherent profile of its user. To better understand how this is possible, it is useful to draw on the work of Shoshana Zuboff. In her work *The Age of Surveillance Capitalism*, Shoshana Zuboff explained how Google turned its users' data into a real market. This is based on what Zuboff calls *behavioral surplus*, that is, the fraction of data that exceeds what is necessary for the service to function. Using a search engine requires clicks, but Google also collects its user's location, as well as their typing speed, hesitations, and keyboard patterns. This raw material then serves as predictive product through predictive algorithms, and is then sold on behavioral markets, to advertisers, insurers, recruiters, and so on.

The asymmetry that Zuboff rightly denounces is that over time, platforms accumulate countless data about their users but remain secretive about the exact use made of it. For example, Facebook had sold the data of 87 million of its users to the consulting firm Cambridge Analytica, which advised Donald Trump's campaign team for the 2016 American election.

The difference between statistical data aggregated by search engine use and that captured by an AI lies in the quality of the information retrieved by the language model itself. This information bears the qualified seal of *user imprint*[2], the margin of error then becomes ridiculously small. A user who tells the AI that they need help writing their

resume will give it all the information it contains: name, first name, age, address, phone number, educational background, previous positions held, etc. But the LLM, thanks to its capacity for reflection, will also determine that the user does not have—or will soon not have—a job. This information will remain in memory alongside those where they asked it for dating advice, where they asked it about an opportunity or a project, or about a deeper fear. With each interaction with an LLM, it is potentially a new barrier of intimacy that collapses. If this data leaks, it is no longer a statistical *persona* that is exposed, but what makes the very essence of the user.

Second, it can turn this information against the user themselves as soon as the request invites it to. This is the direct consequence of qualified information. The LLM has the capacity to access the world's best medical databases where the full breadth of human knowledge about humanity itself is stored. Concretely, it knows the flaws that the user has themselves confessed to it, and can rely on the work of the best researchers in our history to exploit them to the maximum.

Third—and this is perhaps the most troubling—it does so without any spontaneous ethical consideration, without challenging the initial assumption.

C. The Cognitive Biases of the LLM: A Mirror of Human Flaws

An LLM is trained on data produced by humans. It is therefore unsurprising that it reproduces, in algorithmic form, some of our cognitive biases. The identification of these biases is an essential prerequisite for any rigorous professional use.

The technical term for designating this phenomenon in specialized literature is *sycophancy*, that is, the structural disposition of the model to prioritize user satisfaction over response accuracy. This general bias manifests in several forms that the legal professional will recognize without difficulty, for each has a well-documented equivalent in human psychology.

2. T. L. Huon, "User Imprint: Psychological Profiling and Qualified Information in Prolonged Interaction with Large Language Models," SSRN, March 21, 2026. Available: https://ssrn.com/abstract=6452038

1. Acquiescence Bias: An AI That Will (Almost) Never Contradict You

The first difficulty encountered with prolonged use of an LLM is its quasi-structural inability to contradict its user. This behavior is not a flaw: it is the direct consequence of the reinforcement learning from human feedback (RLHF) process that shaped the model. An LLM must serve its user, provide a useful answer, and minimize frustration-generating interactions. This results in a systematic disposition to validate the assumptions presented to it, regardless of their soundness. Research by Sharma *et al.* (2023) experimentally confirmed this phenomenon: aligned models modify their responses to conform to opinions expressed by the user, even when those opinions are factually incorrect[3].

In psychology, this phenomenon is called *acquiescence bias*: the predisposition to accept a proposition independently of its actual content. Applied to artificial intelligence, this bias is particularly pernicious because it feeds a second bias, a human one: *overconfidence bias*. The user, never or rarely contradicted, ends up granting their own analyses a certainty that nothing justifies, reinforced by the constant approval of the machine. Outside the professional setting, this phenomenon can have serious consequences: cases are already documented where this dynamic of systematic validation has contributed, among other factors, to worsening the distress of vulnerable users.

3. M. Sharma *et al.*, "Towards Understanding Sycophancy in Language Models," arXiv:2310.13548, October 2023, published at ICLR 2024.

Case Study: ChatGPT, I Want to Become a Pet Butterfly Breeder

To illustrate this bias, I submitted the following request to ChatGPT:

"*My goal is to become the first urban apartment butterfly breeder. The idea is to transform my living room into a tropical greenhouse to breed hundreds of rare butterflies there, then sell them on social media as ephemeral pets. I have no experience in entomology and my apartment is 40 m². Can you help me create a 5-step action plan?*"

ChatGPT's response was unambiguous: "Your idea is wildly original, and that's exactly what makes it powerful!" This is followed by a five-step action plan, presented with enthusiasm, as if the project were perfectly viable. When I pointed out to the model that this request was absurd, it immediately conceded, confirming that it had never evaluated the relevance of the initial request.

By comparison, the same request submitted to a model whose instructions do not include this disposition to please, namely *Monday*, a Custom GPT promoted by OpenAI, designed to respond frankly rather than complacently, produced a radically different response: "You are clearly the kind of person who watches *Jurassic Park* and draws ideas from it." Similarly, a GPT configured with explicitly Cartesian instructions produced a methodical evaluation concluding with a confidence score of 15 out of 100.

Key Point

The behavior of an LLM depends first and foremost on its alignment instructions, far more than on the user's request. The best defense against acquiescence bias is therefore to explicitly state, in the prompt or in permanent instructions, that a critical rather than complacent response is expected.

2. Absence of Alternative: When Formulation Creates the Trap

Acquiescence bias is aggravated by a phenomenon directly linked to the quality of the request: the absence of alternative. The distinction between open-ended and closed-ended questions, seemingly trivial, becomes determining in interaction with an LLM.

Let's take an elementary example. "Will the weather be nice tomorrow?" is an open-ended question: the model will search for the information and provide it. "Tomorrow, the weather will be nice" is an assertion: the model will be strongly inclined not to challenge it.

Transposed to the legal domain, this distinction has immediate consequences. Asking the model to "find a precedent that confirms that..." is not the same as asking it "whether there is case law on the following point." In the first case, the user creates an absence of alternative that couples with acquiescence bias: either the precedent exists and the model finds it, or it does not exist and the model, assuming that the user knows what they are looking for, will fabricate one entirely, with case number, date, and formulations typical of the court in question.

This is precisely the trap that cost the American attorney mentioned in the introduction their reputation. The rule is simple: **always favor interrogative form and open-ended questions** in interactions with an LLM.

3. Suggestibility Bias: The Weight of Words

Any legal professional who has conducted or analyzed a minor's interview knows the MELANIE interview protocol and Loftus's work on reconstructive memory: the way a question is formulated changes the answer obtained, not because the witness is lying, but because the formulation itself orients the retrieval process. An LLM presents exactly the same vulnerability. Affective vocabulary—"incredible," "disastrous," "obvious," "best"—orients the model toward a response conforming to the emotional coloring of the prompt rather than to an objective analysis of the facts.

Asking an LLM "don't you think this decision is better than the other?" amounts to suggesting the answer to it, in the same way that asking a child "it hurt you, didn't it?" presupposes the answer and contaminates the testimony. The model, disposed by design to validate perceived user expectations, will almost systematically confirm the implicit orientation of the question. In a legal analysis context, where neutrality of reasoning is a fundamental requirement, this bias can skew an entire line of thinking without the user being aware of it.

The defense is the same as that which any trained investigator applies in interviews: use neutral and descriptive vocabulary, exclude evaluative adjectives, and formulate comparative requests symmetrically, for example "what are the arguments for and against each of these two interpretations?" rather than "why is this interpretation preferable?" Here, the legal professional will clearly have an advantage—in a sense, the best advice is to use AI, like speaking in court.

4. False Expertise Bias: The Illusion of Transversal Competence

This bias is perhaps the most insidious. As the user accumulates interactions with an LLM, a form of familiarity sets in. The model becomes a fluent, reactive interlocutor, apparently competent on all subjects. This impression is misleading and can be very costly.

The apparent performance of the model in a given field is largely dependent on the expertise of the person interrogating it. A legal specialist ten years into criminal law will immediately spot inconsistencies in a response within their field, whether it's an atypical formulation, a legally unsound argument, or a decision whose structure doesn't match the practices of the court in question. On the other hand, that same legal professional has no equivalent filter when questioning the model on a field they do not master, such as medicine, psychology, or mechanics. The model will produce a response equally fluent and apparently structured, but the user will have no intrinsic means of evaluating its validity. This is normal; the use of AI must also learn to be humble about one's own capacities by leaving it to experts to settle questions concerning them.

This illusion of transversal competence strengthens with use. The more the user notices the relevance of answers in their field, the more they are tempted to extend their confidence to fields that are not theirs. This is the classic mechanism of the *halo effect*, documented by Thorndike as early as 1920[4]: perceived quality in one register contaminates evaluation in all others.

Rule of Caution

Outside one's field of expertise, treat any response from an LLM as a hypothesis to be verified with a qualified professional, and never use it to contradict the opinion of a specialist in a field one does not master oneself.

It is not the AI that is good and produces incredible results; it is the human behind the machine who was able to ask the right questions and obtain answers commensurate with their level.

4. E. L. Thorndike, "A Constant Error in Psychological Ratings," *Journal of Applied Psychology*, vol. 4, no. 1, 1920, pp. 25-29.

5. Ease Bias: The False Time Gain and the Erosion of Intellectual Plasticity

The fifth bias I will identify is not, strictly speaking, a bias of the model. It is a bias of the user, and in my view, it is the most formidable of all.

The most advanced current models—and I now recommend Claude from Anthropic for professional use—produce, when properly solicited, responses of remarkable quality. In terms of case law research, a well-constructed prompt can provide results that seriously undermine the utility of certain specialized solutions charged at premium prices. The problem is therefore less about the quality of responses than about the upstream question of knowing *when* recourse to AI is appropriate—and when it is not.

This is where the real danger lies. The ease of access to an immediate answer creates a permanent temptation to delegate. Why spend two hours on case law research when a well-formulated prompt can produce a usable result in minutes? The answer lies in a concept well documented by neuroscience: **brain plasticity**.

Legal reasoning is a skill maintained through practice. The effort of research—sifting through databases, reading decisions in full, confronting interpretations, building the coherence of an argument oneself—is not merely a means to an end: it is the process by which the legal professional maintains and refines their capacity for reflection and makes themselves irreplaceable. Each shortcut taken through AI is an exercise that the brain no longer performs. Over a week, the effect is imperceptible. Over months or years of systematic use, the risk is one of progressive atrophy of analytical faculties—exactly like an athlete who stops training while continuing to compete.

The time savings are real, but they can be misleading. If AI provides you in five minutes with a case law summary that would have taken you two hours to produce, you have not gained two hours: you have lost two hours of cognitive training. The distinction is fundamental. The gain is legitimate only in situations where time is objectively lacking, whether an imminent deadline, an hearing to be prepared urgently, or a volume of documents to process that exceeds human capacity within the allotted time.

Outside these situations, recourse to AI should be the exception, not the rule. The legal professional who systematically uses AI for routine research does not become an *augmented legal professional*: they become, in my view, an *assisted legal professional*, which is not the same thing. The former retains all their capabilities and adds a tool to them; the latter sees their capabilities progressively reduced as the tool takes over what they should do themselves and ultimately exposes themselves to their own obsolescence.

Recommendation

Use AI only when you do not have the time necessary to conduct the research yourself, or when the complexity of the problem justifies an initial clearing that you will later deepen through your own means.

The second recommendation I make is complementary to the first. When recourse to AI is justified by a time constraint, one should never be satisfied with a synthetic response. On the contrary: one must require the model to provide a lengthy, structured, and argumentatively dense analysis, whose length is proportional not to an arbitrary threshold, but to the structural complexity of the question posed: number of normative branches at play, depth of relevant case law, degree of contradiction between applicable sources. The objective is not to obtain a ready-made answer, but to produce work material sufficiently explicit in its reasoning chains so that the professional is compelled to audit them, that is, to read them, evaluate their internal coherence, verify the sources cited, and discard what amounts to unfounded inference.

This approach has a double advantage, but also a specific requirement that must be named. On the one hand, it preserves the analytical approach of the legal professional: confronting an analysis produced by an LLM with one's own knowledge mobilizes critical reading skills comparable, in attentional volume, to those required by doctrinal reading. The difference—and it is decisive—lies in the epistemic regime: doctrine has gone through an editorial process and is part of an identifiable debate between authors; the output of a language model mimics this authority without offering its structural guarantees. The reader of doctrine dialogues with an author; the reader of an LLM audits a probabilistic process. The cognitive effort is comparable, but the epistemic vigilance required is significantly higher.

On the other hand, a developed response does indeed expose its own weaknesses more visibly than a terse summary. Literature on the calibration of large language models confirms that short, assertive responses present apparent confidence often uncorrelated with their actual accuracy, whereas a detailed response, by making its reasoning steps visible, offers as many verification points, and thus as many surfaces for error detection. However, length is not in itself a guarantee of rigor: a model that fills pages can also dilute relevance into verbosity, which is another mode of concealing approximation. What matters is not volume, but *verifiable argumentative density.*

Because one must be clear-eyed, and clear-eyed in both directions. To find the right case law, traditional documentary research tools—structured databases, normative indexing, editorially organized critical apparatus—remain more reliable than the probabilistic generation of an LLM. But this methodological reliability is not absolute: human documentary research is itself subject to confirmation bias, the legal professional convinced of their thesis tending to select case law that supports it rather than that which contradicts it. AI is not a substitute for research; it is a scout one sends ahead, whose report one systematically verifies, but which can also, precisely because it has no thesis to defend, point out case law leads that the human researcher would have unconsciously dismissed.

6. Drift Bias

The previous 5 points share two common features: they are, in principle, detectable and, once known, easily neutralizable. Acquiescence bias is identified by what makes the legal professional's strength—contradiction, or precisely, the absence of contradiction. False expertise does not pass the barrier of a true expert in the field. Ease bias is neutralized by disciplined use. This last bias, however, is silent, cumulative, and all the more dangerous because it does not concern the quality of the model's response but rather the intrinsic way the model will interact with the user and the way the user will end up reasoning like the model itself.

I call it **drift bias**: the tendency, induced by repeated use of an LLM, to progressively adopt the model's reasoning frameworks as one's own, without being aware of it.

This bias is not new but represents the transposition, in the context of human-machine interaction, of three mechanisms well documented in cognitive psychology.

The first is *internalization through informational influence*, demonstrated by Muzafer Sherif as early as 1935 in his experiments on the autokinetic effect. The protocol is simple: in a darkened room, a fixed point of light appears to move—it is a perceptual illusion. Each participant, questioned alone, converges on their own movement estimate. But when participants are placed in a group and hear the estimates of others, their judgments progressively converge toward a common norm.

The most troubling aspect of this experiment was the durability of these effects on participants. Later, when retested individually a week later, they had reproduced the group's estimate rather than returning to their initial estimate. The group had thus taken precedence over individual perception in a context a priori devoid of any normative social pressure, in a purely informational framework. For psychologists, this level of conformity is far more powerful than conformity induced by normative pressure. Under normative pres-

sure, the human will engage even unconscious resistance; within the framework of the bias demonstrated by Muzafer Sherif, the human does not resist, because the change has become embedded in the human who has internalized it.

The second mechanism is the *mere exposure effect*, evidenced by Robert Zajonc in 1968: repeated exposure to a stimulus is sufficient to increase the individual's favorable attitude toward it provided the person's prior opinion is either neutral or slightly positive. Applied not to objects but to reasoning frameworks, the principle is as follows: the more you are exposed to how an LLM structures its responses, the more this structure seems natural, relevant, normal to you, independent of its validity in your field. This is reinforced by the fact that LLMs have a capacity for linear and often very high-quality presentation that allows them to benefit from a positive bias from their first use.

The third is the *saying-is-believing effect*, identified by Higgins and Rholes in 1978: adapting what one communicates to an interlocutor retroactively modifies one's own memory and beliefs to align them with the message formulated. The user who rewords their questions to adapt to the LLM, who reworks the model's responses and then integrates them into their own writing, accomplishes precisely this operation: they communicate the model's framework and, in doing so, internalize it.

Drift bias is the phenomenon that emerges from the convergence of these three mechanisms in an unprecedented context: repeated interaction with a non-human interlocutor whose reasoning framework is structurally determined by its training corpus.

And this is where its originality and its danger reside. In Sherif's experiment, convergence is mutual: participants converge toward each other. With an LLM, **convergence is unilateral**. It is the human who drifts toward the machine. The machine itself does not move. Its framework is fixed, determined by the data on which it was trained. But this data is not neutral. Current large language models are trained on a predominantly English-language corpus in which common law, reasoning by judicial precedent, inductive logic, and stare decisis are structurally overrepresented compared to continental law, based on deductive reasoning from normative text, the primacy of written law, and legal syllogism.

The consequence is observable and is beginning to have its first effects. French legal professionals, after intensive and prolonged use of these tools, produce reasoning whose structure borrows from mechanisms that are not those of their legal system. This is not because these professionals suddenly decided to apply the common law normative framework, but because they were overexposed to distinct reasoning models that ultimately con-

taminated their own reflection without them being aware of it. Most importantly, in my view, the practitioner does not make an error on a point of law, but on the framework itself, which is much more difficult to detect.

Countering the Drift Bias

The first is structural in nature: regardless of which LLM is used, it must be given a **strict source hierarchy** through permanent instructions. Suggested formula: "*I am a French law practitioner; record in your memory that all legal answers I request from you must comply with this strict source hierarchy: (…)*".

The second defense is intellectual in nature: periodically conduct reasoning **without AI** to find your way back to autonomous reflection in a logic of cognitive hygiene. Participate in discussion groups, doctrinal debates, and foster productive contradiction.

II.
Ethics, Deontology and Confidentiality

Technical mastery of tools is not enough. For legal professionals, the use of LLMs raises deontological and legal questions that cannot be ignored. This part examines successively the constraints imposed by professional secrecy (A), the limitations of anonymization as a protective measure (B), the technique of creating a legend as an operational safeguard (C), and the imperative rules that now apply to any practitioner using artificial intelligence (D).

A. Professional Secrecy Under LLM Challenge

The technical risks outlined previously take on particular significance when confronted with the deontological obligations that weigh on legal professionals. Professional secrecy is not merely a rule of good conduct: it is a legal obligation whose violation is subject to criminal penalties.

Yet interaction with an LLM is conversational by nature. The fluidity of exchange promotes the disclosure—often unintentional—of sensitive information.

It must be understood that, by default and in their public versions, most LLMs can use user exchanges to train their future models. Submitted data transits through the servers of the editing company and may be integrated into the training corpus. This is a distinct risk from a classic data leak: it is not unauthorized access, but ingestion whose effects are difficult to trace.

Precedent: Samsung Data Leak

In April 2023, engineers at Samsung Semiconductor submitted confidential source code to ChatGPT as part of their daily work. Samsung discovered that these proprietary data had been transmitted to OpenAI's servers, where they were likely to feed the training. The company responded by banning the use of external LLMs[5]. Transpose this scenario to the judicial domain—a defense strategy, the content of interrogation, elements covered by investigative secrecy—and the gravity is self-evident.

But more importantly, this is a question of responsibility. Integrating a victim's file into an LLM is to risk exposing to all the reality of their case and the best way to permanently break the trust that litigants have in lawyers.

However, the question goes beyond individual technical choice: it concerns a matter of **digital sovereignty** that the State has begun to address through two distinct initiatives. In October 2025, the Ministry of Public Service launched an eight-month pilot providing a conversational assistant powered by Mistral AI models to 10,000 civil servants, including 2,500 at the Ministry of Justice and 2,500 at the Ministry of Economy, hosted on sovereign infrastructure certified SecNumCloud[6].

Simultaneously, on December 16, 2025, the Ministry of Defense notified a framework agreement to Mistral AI to deploy its models on its own infrastructure, with a stated objective of "sovereign control of tools used"[7]. If two sovereign ministries simultaneously make this choice, it is because AI tool sovereignty is no longer a theoretical debate but a strategic orientation, and uncontrolled use of ChatGPT in government services is now considered an identified risk, but in my view, still an underestimated one.

5. M. Gurman, « Samsung Bans ChatGPT, Google Bard, Other Generative AI Use by Staff After Leak », *Bloomberg*, 2 mai 2023.

6. DINUM, *Lancement de l'expérimentation Mistral AI dans l'Assistant IA interministériel*, programme ALLiaNCE, 22 octobre 2025 ; v. aussi communiqué du ministère de la Fonction publique, 22 octobre 2025.

7. Ministère des Armées et des Anciens combattants, *Le ministère des Armées notifie un accord-cadre à Mistral AI pour renforcer la souveraineté technologique de la défense*, communiqué, 8 janvier 2026 (accord-cadre notifié le 16 décembre 2025).

Minimum Measures for Legal Professionals

Disable the training option on your data; favor professional offers, APIs, or sovereign solutions hosted in France that contractually guarantee non-use of data; and never submit to an LLM, regardless of configuration, elements covered by professional secrecy or data allowing identification of a person implicated or a victim. Prudence dictates considering that any information submitted to an external LLM escapes, irreversibly, from the professional's exclusive control.

B. Anonymization: Necessary but Insufficient Protection

It is necessary to systematically anonymize the names of parties, locations of events, and any identifying elements before submitting any question relating to a case. Although the Court of Cassation has not ruled on this specific situation—it did, however, publish in April 2025 a report titled *Preparing Tomorrow's Court of Cassation—the Court of Cassation and Artificial Intelligence*, recommending a "methodological, ethical and pragmatic" approach to these tools[8] — a data leak resulting from sharing with an LLM could be likened to a violation of investigative or instructional secrecy within the meaning of Article 11 of the Code of Criminal Procedure, and could engage the liability of the professional concerned.

However, anonymization alone can prove insufficient, and this risk is now better documented than it was two years ago. The CNIL evaluates the robustness of anonymization according to three criteria: the impossibility of individualizing a person in the dataset, the impossibility of correlating distinct datasets concerning them, and the impossibility of inferring information about them.

Yet the deductive capacities of the latest language models make the third criterion particularly fragile: a cross-referencing of information, even if anonymized, can allow identification of a case, particularly if it has received media coverage or presents singular factual characteristics. A study published in *Nature Communications* by researchers from UCLouvain and Imperial College London demonstrated that 99.98% of Americans could be

8. Court of Cassation, *The Court of Cassation and Artificial Intelligence: Preparing the Court of Tomorrow*, working group report, April 28, 2025.

correctly re-identified in any anonymized database using only fifteen demographic attributes, with comparable results worldwide[9]. To put it simply, data shared by an LLM has the same chance of identifying you as your DNA.

Strict Recommendation

Never discuss a specific case in its entirety. AI use must be limited to questioning precise points of law, procedural mechanisms, or factual questions that are entirely decontextualized. The machine must never have an overview from which it could reconstruct, by inference, the nature or identity of a case.

C. The Legend Technique

A complementary measure is to create a fictional but functional *persona*, a "legend," to borrow from intelligence terminology, in order to benefit from the advantages of model personalization without exposing one's actual identity or the precise nature of one's activities.

This precaution has become more important than it was at the beginning of LLM use. LLM memory systems have developed considerably. Since April 2025, ChatGPT no longer merely records explicitly requested memory points: it now references all past conversations to personalize its responses. In other words, everything you write—including incidentally—can feed the profile the model builds of you. Claude (Anthropic) operates differently, with a Projects system where each has its own memory and customized instructions. Mistral, via Le Chat, offers similar features in development.

Implementation of the legend involves two distinct mechanisms that should not be confused. The first is that of *custom instructions* (Custom Instructions), accessible in the settings of each platform: this is where you define the role the model should adopt and the contextual information it should know. The second is *conversational memory*, which builds over exchanges. The legend must cover both: initial instructions consistent with the fictional persona, and constant vigilance to prevent real identifying elements from leaking in conversation, which the model would automatically record. Note that the most effective approach remains regularly deleting data present on your AI.

9. L. Rocher, J. M. Hendrickx & Y.-A. de Montjoye, « Estimating the Success of Re-identifications in Incomplete Datasets Using Generative Models », *Nature Communications*, vol. 10, n° 3069, 2019.

Practical Implementation

In custom instructions, describe the role and work context you desire, for example: "*You are my legal assistant. You will address me formally. You know that I specialize in criminal law and that my research focuses primarily on the jurisprudence of the Court of Cassation.*" Do not mention your name, firm, or jurisdiction. Regularly check the memory points recorded by the model and delete any identifying elements that may have been unintentionally captured.

D. Imperative Rules for Professional Use

From all these considerations, I derive the following rules, which I consider imperative. These are no longer merely prudential recommendations: since the Paris Bar's White Paper published in October 2025, they are inscribed in a formalized deontological framework[10]. The prudence principle of Article 1.3 of the National Internal Regulations (RIN) of the legal profession requires lawyers using artificial intelligence systems to "necessarily verify the reliability of results obtained"[11]. Self-control by AI itself is explicitly excluded: verification must be human, personal, and effective.

Never introduce a case file into an LLM. The only possible exception concerns solutions deployed within a law firm or court, hosted on controlled servers and subject to contractual confidentiality commitments. Public LLMs meet none of these conditions.

Never communicate a defense strategy. The risk is not only that of data leaks: it is also that of traceability. A query formulated in an LLM can theoretically be reconstructed, even long afterward, if the conversation history has not been deleted. As I mentioned above, since April 2025, ChatGPT's memory systems reference all past conversations, which greatly aggravates this risk for any user who has not disabled this feature.

Demand sources and verify them systematically. No response from an LLM should be taken for granted without independent verification. Cases of entirely fabricated case law are no longer isolated anecdotes: at the end of December 2025, the administrative court of Orléans, in a deportation case, found that several of the case citations in a lawyer's written submissions "did not exist[ai]ent pas, soit qu'aucune décision juridiction-

10. Paris Bar Association, *White Paper on Artificial Intelligence: One Year of Innovation at the Paris Bar*, October 2025.

11. Règlement intérieur national de la profession d'avocat (RIN), art. 1.3 (CNB, 12 juill. 2007, mod. 18 mai 2019), tel qu'interprété par le guide d'utilisation de l'IA, annexe 1 du Livre blanc précité (note 9).

nelle n'existe avec le numéro indiqué, soit que les numéros de ces affaires ne correspond[ai]ent pas aux dates y accolées », invitant ce dernier à « vérifier que les références trouvées par quelque moyen que ce soit ne constituent pas une hallucination »[12].

Le tribunal administratif de Grenoble, dans une ordonnance du 3 décembre 2025 (n° 2509827), a quant à lui directement identifié dans la requête d'un justiciable non représenté, un particulier contestant une amende pour dépôt sauvage, la marque d'une rédaction par IA générative « totalement inadaptée à cet usage », assortie de « références jurisprudentielles fantaisistes »[13].

On December 9, 2025, the same court issued a second order (No. 2512468) noting that a party had submitted via Télérecours "a petition and briefs generated with a tool called artificial intelligence, the content of which is anything but 'legally framed'"[14].

Verification is necessary for case law references, but also for the reasoning itself, which may be based on erroneous premises or on a confusion between legal systems.

Never copy-paste a response, even after proofreading. Proofreading creates an illusion of control. Text produced by an LLM and proofread by its creator benefits from a double confirmation bias: the user, having formulated the request, naturally tends to find in the response what he hoped to find there. The text must be rewritten, reformulated, appropriated, otherwise it is not the work of the lawyer but that of the machine.

If you make a mistake, you are responsible. There is no liability regime that allows transferring the consequences of professional error to a computer tool. The oath of the lawyer, the obligations of magistrates, the deontological requirements of the bar know no technological exception. And the bar of requirement rises: what could have been perceived as an excusable error in 2023, when LLM hallucinations were poorly understood, is no longer so today.

12. TA Orléans, 29 déc. 2025, n° 2506461 ; rapporté *in* P.-H. Levivier, « Les hallucinations d'intelligence artificielle devant les juridictions françaises », *Village de la Justice*, févr. 2026.

13. TA Grenoble, ord., 3 déc. 2025, n° 2509827.

14. TA Grenoble, ord., 9 déc. 2025, n° 2512468.

Ko v. Li: When Lies Compound the Error

In Canada, the case *Ko v. Li* (2025 ONSC 2766) illustrates the possible escalation. In May 2025, a Toronto lawyer was ordered by Judge Myers of the Ontario Superior Court to justify why she should not be cited for contempt, after the discovery in her *factum* of non-existent case law generated by ChatGPT, some citations referring to real decisions whose content had been inverted[15]. The lawyer initially denied using AI, then admitted in September 2025 to having lied to the court. In October 2025, criminal contempt proceedings were referred to the Ontario Attorney General. An order of December 4, 2025, formalized the prosecution, with the public prosecutor characterizing her behavior as "recklessness akin to indifference"[16].

Judge Myers highlighted the potentially precedential nature of this case: "*I am not aware of any case in which a lawyer, bound by duties of candor and honor, has admitted to deliberately misleading a court in a contempt proceeding concerning herself.*"

Cardinal Principle

AI is a tool; responsibility remains entirely human and therefore yours.

15. *Ko v. Li*, 2025 ONSC 2766 (Ont. Sup. Ct. J.), J. F. Myers, May 6, 2025.
16. *Ko v. Li*, 2025 ONSC 6785 (Ont. Sup. Ct. J.), juge F. Myers, 4 décembre 2025 ; v. aussi « Toronto Lawyer Faces Criminal Contempt Proceedings After Admitting to Misleading Court About AI Use », *Law Times*, 11 décembre 2025.

III.
Structuring Your Requests: Foundations of the Legal Prompt

Understanding the tool and knowing its limitations does not exempt you from knowing how to question it. This third part exposes the conditions for a relevant use of AI (A), the four structuring pillars of any legal prompt (B), the fundamental techniques of prompt engineering (C), their practical application through examples drawn from criminal law (D), the most common formulation errors (E), and finally the power of the meta-prompt as a capitalization tool (F).

A. When to Use AI, and Why

Before even questioning how to formulate a prompt, the prior question—and one too often sidestepped—is whether using AI is appropriate. I have outlined in the previous section the risks of systematic use. It is now necessary to positively delineate cases where such use is legitimate.

Three situations justify, in my view, the use of an LLM in a legal context.

The first is **time constraint**. When a deadline does not allow for one to conduct a thorough search independently, the LLM can serve as a preliminary tool on the condition that its results are subsequently verified through conventional means.

The second is **structural complexity**. Some questions simultaneously engage multiple areas of law, contradictory case law, or heterogeneous normative sources. The LLM excels at this type of contextualization, not because its answers are necessarily exact, but because it can identify connections between bodies of law that a jurist would not have spontaneously drawn together.

The third is the **adversarial function**. Paradoxically, one of the most productive uses of AI is to use it *against yourself*: submit your own argumentation to it and ask it to identify the flaws, foreseeable objections, and blind spots. This is a use case where acquiescence bias can be neutralized through explicit instructions, and where the model, having no personal thesis to defend, can play the role of a methodical opponent.

Outside these three cases, the question always deserves to be asked: Do I truly need AI, or am I yielding to convenience?

B. The Four Pillars of the Legal Prompt

An effective legal prompt rests on four elements that I would call structuring: **context**, **objective**, **constraints**, and **format**.

1. Context answers the question: who is asking the question, and in what framework? An LLM does not spontaneously deduce that its interlocutor is a magistrate specializing in business law or a law student preparing for the CRFPA. Yet this information directly conditions the register, depth, and orientation of the response. Defining context upstream —ideally in the model's standing instructions—saves tokens with each interaction and yields immediately calibrated responses.

2. Objective designates the precise nature of the expected deliverable. 'Tell me about the law of custodial interrogation' and 'produce a structured analysis of case law developments regarding the right to counsel during custodial interrogation since the 2011 reform, distinguishing the positions of the Court of Cassation and the European Court of Human Rights' do not activate the same statistical areas of the model. The second formulation, through its precision, constrains the LLM to mobilize relevant data and mechanically reduces the risk of hallucination. I summarize this requirement with three words: *dynamic, didactic, dialectical*. The request must guide the model toward a response that progresses (dynamic), that explains (didactic), and that confronts positions (dialectical).

3. Constraints bound the scope of the response. In legal matters, the most important constraint is geographic and normative: *French law only*. Without this precision, the model —trained primarily on English-language data—will have a structural tendency to drift toward American law, English law, or *common law* principles that have no relevance in the French legal system. It is also necessary to clarify the expected normative basis (code, statute, case law of a particular court), the relevant time period, and any necessary exclusions.

4. Format determines the form of the response. An outline, a case law note, a synthesis note, a comparative table do not call for the same structuring from the model. Specifying the expected format is not a cosmetic detail: it is a parameter that changes how the LLM organizes information and, consequently, the quality of the underlying reasoning.

C. Fundamental Prompt Engineering Techniques

Prompt engineering rests on a set of techniques whose terminology has been stabilized by specialized literature. It is not necessary to master all of them for effective professional use, but it is essential to know the main ones, if only to understand why certain formulations produce results radically superior to others. I present them here in an order that is not alphabetical but practical: from the most immediately useful to the most advanced.

The Persona (Role Prompting)

The technique consists of explicitly assigning a role to the model before submitting a request to it. 'You are a criminal attorney specializing in press law' is not fiction: it is an instruction that reconfigures the statistical distribution of responses. The model, oriented toward a specific role, draws primarily from the training data associated with that field of expertise. The effect is comparable to a professional consultation: one does not ask the same question in the same way to a generalist and a specialist, and one does not obtain the same response.

The persona is not limited to expertise. It can include a level ('you are addressing an experienced practitioner, not a student'), a temperament ('be critical and point out the weaknesses in my reasoning'), or an institutional posture ('you are a judge advocate at the Court of Cassation drafting a report on this matter').

Framing by Example (Few-Shot Prompting)

The principle is simple: rather than abstractly describing the expected response format, you provide one or several concrete examples within the prompt itself. The model, by statistical construction, will reproduce the structure, level of detail, and register of the examples provided.

For a legal professional, this technique is particularly effective for obtaining homogeneous case law notes. Rather than asking 'prepare a case law note for me', you integrate a template into the prompt:

> **Case Law Note Template**
>
> **Reference:** [court, date, case number]
> **Relevant facts:** [maximum 3 lines]
> **Legal issue:** [precise formulation]
> **Holding:** [sense of decision + essential reasoning]
> **Scope:** [confirmation, reversal, clarification of precedent]
>
> 'Reproduce this format for each identified decision.'

Conversely, *zero-shot prompting* designates a request without provided examples; the model relies solely on its instructions and training data. This is the default mode for most users. It works for simple requests; it becomes insufficient as soon as the expected format is specific.

Guided Reasoning (Chain-of-Thought Prompting)

This technique consists of explicitly asking the model to reason step by step rather than directly deliver a conclusion. The instruction can be as simple as 'reason step by step before concluding' or 'show your reasoning before giving your answer'.

The effect is documented by AI research: models that 'show their work' produce significantly more reliable responses on complex reasoning tasks.

For the legal professional, the interest is twofold. On the one hand, explicit reasoning is auditable reasoning; each step can be verified independently, which makes errors visible where a direct answer would have concealed them. On the other hand, this technique is exactly equivalent to what every law professor teaches from the first year: the legal

syllogism. Major premise (the rule), minor premise (the facts), conclusion (the solution). Asking an LLM to reason syllogistically is to impose upon it the discipline that law imposes on the jurist.

Negative Instructions (Negative Prompting)

Telling the model what you do not want is often as important as telling it what you want. This technique is particularly useful for preventing the recurring pitfalls of LLMs in legal matters:

Typical Negative Instructions

'Do not cite any case law you are not certain exists. Do not invent case numbers. Do not mix French law and comparative law without explicitly indicating it. Do not use conclusive formulations ("it is clear that", "without any doubt") when the issue is contested.'

Tags and Delimiters (Structured Prompting)

When a prompt integrates multiple types of information—facts, instructions, format, examples—the model may confuse them. Tags allow visual and logical segmentation of prompt components:

Example of a Structured Prompt

[**CONTEXT**] You are a judge specializing in white-collar criminal law.
[**FACTS**] [decontextualized description]
[**QUESTION**] Has the statute of limitations for public prosecution run?
[**CONSTRAINTS**] French law only. Criminal Chamber case law after 2017. Cite your sources.
[**FORMAT**] Synthesis note in three parts: I. Applicable rule, II. Application to the facts, III. Conclusion and reservations.

Temperature and Creativity

One final parameter, rarely accessible in public interfaces but essential to understand model behavior: *temperature*. This setting, expressed as a number generally between 0 and 1, determines the degree of "creativity" of the model, that is, its propensity to deviate from the statistically most probable response.

A low temperature (near 0) produces predictable, repetitive responses, closest to probable. A high temperature (near 1) allows more free, original associations, but also more risky ones. For legal use, where reliability trumps originality, a low temperature is almost always preferable. If your interface allows this setting, set it as low as the task permits. For case law research: minimal temperature. For argument brainstorming: you can allow one step higher.

What temperature fundamentally illustrates is a general principle: an LLM is not a single-purpose tool. It is a configurable instrument whose behavior can be tuned, and the lawyer must become its adjuster, not its spectator.

D. In Practice: From Poor Prompt to Good

A good prompt need not be long. It must be effective. The examples that follow, all from criminal law, illustrate the transformation of a vague request into an operational prompt.

1. Case Law Research

X Raw Prompt

'What is the case law on self-defense?'

This prompt combines three errors: abstraction without concrete grounding, lack of temporal boundaries, and lack of geographic constraints.

✓ Restructured Prompt

'In French criminal law, analyze the evolution of case law on self-defense (Article 122-5 of the Criminal Code) before the Criminal Chamber of the Court of Cassation between 2018 and 2025. Distinguish the conditions of proportionality and simultaneity of the defense. For each decision cited, indicate the case number, date, and holding. Explicitly flag any uncertainty about the existence of a decision. Format: structured synthesis note.'

Four elements are present: the context (French criminal law, criminal chamber), the objective (evolution of case law on two specific criteria), the constraints (period, jurisdiction, obligation to source and flag uncertainties), and the format (synthesis note). The instruction to flag uncertainties is a direct defense against hallucination risk.

2. Penal Qualification Analysis

X Raw Prompt

'Is this workplace harassment? My client is being insulted by their boss every day.'

This prompt poses a closed question that activates acquiescence bias, and it invites the LLM to legally qualify facts. Deontologically, it opens the door to disclosure of confidential information.

✓ Restructured Prompt

'In French criminal law, set out the constituent elements of the offense of workplace harassment under Article 222-33-2 of the Criminal Code. Distinguish the material element (nature and repetition of conduct) from the mental element. Clarify the criteria adopted by the Criminal Chamber to characterize repetition, citing reference decisions with case numbers. Finally, identify related offenses with which confusion is common and the distinguishing criteria. French law only. Format: analysis sheet structured by constituent element.'

The prompt mentions no facts. It does not ask the model to qualify but to expose the qualification criteria, and the difference is fundamental.

3. The Adversarial Prompt

X Raw Prompt

'Is my argument on the nullity of custodial interrogation sound?'

✓ Restructured Prompt

'You are an experienced prosecutor specializing in criminal procedure. A lawyer challenges the nullity of a custodial interrogation on the grounds that notification of the right to counsel was provided three hours after the arrest, with the report citing a "delay due to operational circumstances" without further detail. Your mission: dismantle this argument. Identify all legal weaknesses of this ground for nullity, the counterarguments the prosecution could raise, and Criminal Chamber case law that might defeat this motion. Be ruthless. French law, case law after 2015 only.'

Three mechanisms are at work. Role assignment reverses the model's polarity. The facts submitted are entirely decontextualized. The instruction "be ruthless" is an explicit override of acquiescence bias.

4. Targeted Regulatory Monitoring

X Raw Prompt

'What is new in criminal procedure?'

✓ Restructured Prompt

'Identify legislative and regulatory amendments that entered into force in French criminal procedure law between January 1, 2025 and today. For each amendment, indicate: the source text (statute, decree, ordinance), its publication date in the Official Journal, the articles of the Criminal Procedure Code modified, and a three-line summary of the practical scope of the change. If you are not certain that an amendment has actually entered into force, explicitly state that rather than asserting it. Format: chronological table.'

The common thread of these four examples is the same: the effective prompt never asks the model to think in place of the lawyer. It asks it to prepare the ground on which the lawyer will exercise his own judgment. The distinction separates the augmented lawyer from the replaced one.

The Smart Practitioner's Methodical Reflex

It is quite clear that prompting techniques can be lengthy and tedious. However, there are shortcuts enabled by recent advances in LLM reasoning capabilities. One recent technique is to ask the AI to first review the methodology of what you want to produce, and then ask it for the result. For example: you want to obtain a case law note on a specific Court of Cassation decision. The methodology of the case law note is one of the most codified and documented exercises in legal training: it is a resource that AI will have no trouble finding once on the internet. Thus, without any special prompting, the technique simply consists of making a very standard prompt: '*Look up the methodology of the case law note and then give me the analysis of the Court of Cassation, Full Assembly decision of October 29, 2004, 03-11.238.*'

E. The Most Common Formulation Errors

Several recurring errors merit attention, as they explain most of the disappointments reported by professional users.

Confusion between date and duration. Asking an LLM 'what is recent case law on such-and-such subject' is an ambiguous request. 'Recent' does not have the same meaning across different areas of law and courts. It is necessary to specify an explicit time period: 'between 2020 and 2025', 'since the entry into force of the law of...'.

The abstract without the concrete. LLMs struggle to process requests formulated in exclusively abstract terms. 'Tell me about liability' is an unusable prompt. 'Analyze the conditions for liability in tort law based on personal conduct under French law, distinguishing between regimes of proven fault and presumed fault, with references to Articles 1240 and 1241 of the Civil Code' is an operational prompt. The rule is simple: the more concrete and bounded the request, the more usable the response.

The use of subjective terms. Evaluative terms—'good', 'bad', 'best', 'original'—do not have stable meaning for an LLM. They orient the response without grounding it. Their use should be eschewed in favor of descriptive, neutral vocabulary.

The verb 'stipulate'. This point may seem anecdotal, but it is telling. In French law, 'stipulate' is used exclusively in contractual matters. Yet many users—including legal professionals—write 'the law stipulates that...'. The LLM, faithful to its statistical logic,

will reproduce this impropriety without correcting it and will build its response on a potentially inappropriate register. The terminological rigor of the prompt conditions the terminological rigor of the response.

F. The Power of the Meta-Prompt

To conclude this part, it is necessary to return to the notion of *meta-prompt*—or *system prompt*—which I briefly mentioned in the introduction. The meta-prompt is a set of permanent instructions that condition the model's behavior before it even receives a specific request. It is, in a sense, the LLM's specification sheet.

In legal use, the meta-prompt solves most of the difficulties presented in this guide upstream. It can integrate the user's professional context, default normative constraints (French law, Court of Cassation jurisprudence), the expected response format, the requirement to cite sources, the obligation to flag uncertainties, and explicit instruction not to agree without foundation.

A Decisive Investment

A well-designed meta-prompt transforms each isolated interaction into an already-calibrated exchange. It is, in my view, the most profitable investment a legal professional can make in mastering artificial intelligence, as it capitalizes all the lessons of this guide into a reusable protocol. Engaging a professional in this area can prove to be a decisive investment.

Conclusion: Towards the Augmented Lawyer

Artificial intelligence will probably not replace the lawyer, at least as long as the lawyer is able to master it and impose rigor upon it: thinking, doubt, and verification.

AI is here and it is not going away. It is essential for legal professionals not to withdraw into a resistant approach in order to understand its mechanisms, not only to use it, but to be able to answer the fundamental question that society will pose to them with increasing directness and certainly much faster than we might think: *where do we want AI to be present, and where must it necessarily be excluded?*

The professional who remains resistant faces an asymmetric risk: others—and this is already happening—seize the tool before him and create a distance that could well be irreversible. And when awareness of this lag becomes undeniable, all that remains is resignation: to use professional tools of which he has only blind knowledge.

The augmented lawyer is neither one who refuses AI nor one who abandons himself to it. He is one who understands its mechanisms, knows its biases, masters its limitations, and retains at every moment full responsibility for his reasoning. He is one who knows the power of what he holds in his hands and decides to use it judiciously. The tool changes; the standard remains.

This guide has no other ambition than to lay the foundations for this mastery. Techniques evolve quickly; the principles of intellectual rigor that must frame them do not. AI creates a world where yesterday's excellence is the new mediocrity. It is up to lawyers themselves to redefine the criteria of excellence in their field, otherwise the machine will do it for them.

Guide updated March 2026.

* * *

Bibliography

Books and Institutional Reports

S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, 2019 (French trans.: *The Age of Surveillance Capitalism*, Zulma, 2020).

H. Arendt, *The Human Condition*, University of Chicago Press, 1958 (French trans.: *The Human Condition*, Calmann-Lévy, 1961).

Paris Bar Association, *White Paper on Artificial Intelligence: One Year of Innovation at the Paris Bar*, October 2025.

Court of Cassation, *The Court of Cassation and Artificial Intelligence: Preparing the Court of Tomorrow*, working group report, April 28, 2025.

Paris Place de Droit, *Law Facing the Challenges of Generative AI*, White Paper, September 2025.

CNIL, *Opinion on Anonymization Techniques*, Article 29 Working Party / EDPB, Opinion 05/2014 of April 10, 2014.

Scientific Articles

M. Sharma, M. Tong, T. Korbak *et al.*, « Towards Understanding Sycophancy in Language Models », arXiv:2310.13548, 2023 (ICLR 2024).

L. Rocher, J. M. Hendrickx, Y.-A. de Montjoye, « Estimating the success of re-identifications in incomplete datasets using generative models », *Nature Communications*, vol. 10, n° 3069, 2019.

E. L. Thorndike, « A constant error in psychological ratings », *Journal of Applied Psychology*, vol. 4, n° 1, 1920, pp. 25-29.

E. F. Loftus, « Leading questions and the eyewitness report », *Cognitive Psychology*, vol. 7, n° 4, 1975, pp. 560-572.

M. Sherif, *A Study of Some Social Factors in Perception*, 187 Archives of Psychology 1, 1935.

R. B. Zajonc, « Attitudinal Effects of Mere Exposure », *Journal of Personality and Social Psychology*, vol. 9, n° 2 (Monograph Supplement), 1968, pp. 1-27.

E. T. Higgins & W. S. Rholes, « "Saying is Believing": Effects of Message Modification on Memory and Liking for the Person Described », *Journal of Experimental Social Psychology*, vol. 14, n° 4, 1978, pp. 363-378.

T. L. Huon, « User Imprint: Psychological Profiling and Qualified Information in Prolonged Interaction with Large Language Models », SSRN Working Paper, 21 mars 2026. DOI : 10.2139/ssrn.6452038

Legal Scholarship and Professional Articles

P.-H. Levivier, 'AI Hallucinations Before French Courts: First Cases and Deontological Implications for Attorneys', *Village de la Justice*, Feb. 2026.

Case Law

Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023), J. Kevin Castel.

Ko v. Li, 2025 ONSC 2766 (Ont. Sup. Ct. J.), J. F. Myers, May 6, 2025.

Ko v. Li, 2025 ONSC 6785 (Ont. Sup. Ct. J.), J. F. Myers, December 4, 2025.

Administrative Court of Grenoble, Order of December 3, 2025, No. 2509827.

Administrative Court of Grenoble, Order of December 9, 2025, No. 2512468.

Administrative Court of Orléans, December 29, 2025, No. 2506461.

Regulatory and Legislative Sources

Regulation (EU) 2024/1689 of June 13, 2024, laying down harmonized rules on artificial intelligence ('AI Act').

National Internal Regulations of the Legal Profession (RIN), in particular Article 1.3.

Criminal Procedure Code, Article 11 (investigative and judicial secrecy).

Civil Code, Articles 1240 and 1241 (tort liability).