

# Phase 1 Blog Post

## To Advise or Not to Advise? Lawyers Weigh in on AI's Legal Guidance



### Background

As AI-based chatbots become increasingly sophisticated, people are turning to robots for expert knowledge — medical advice, legal counsel, financial planning — once largely inaccessible or unaffordable. Now, people utilize AI-powered chatbots to get personalized guidance and customized answers to their most pressing questions. Law is one such high demand area. Despite how regularly people encounter legal issues — rental agreements, divorce filings, and noise disputes with neighbors, for example — lawyers are notoriously expensive.

However, there are many valid concerns around chatbots providing legal advice. To name a few, **inaccuracy** in the guidance could lead to detrimental real-world outcomes. Additionally, potential data breaches pose **privacy risks** given the *sensitive nature of legal matters*. There are also worries about **bias and unfair recommendations**, as well as lack of **interpretability** around how the chatbot arrives at suggestions. Further concerns include **misuse or abuse** of the legal advice provided, especially when not constrained by professional ethics and conduct standards lawyers are held to.

In this sense, **EU AI law (draft)** puts AI systems used for “Assistance in legal interpretation and application of the law” into the **high-risk category**. While the convenience and accessibility of chatbots for legal guidance is appealing, these systems must be developed thoughtfully to address the many risks accompanying such a powerful application.

There has been very little research specifically examining the factors that determine the appropriateness of AI systems providing professional guidance. Prior scholarship has discussed AI advice mainly in broad, speculative terms rather than conducting in-depth studies grounded in real-world use cases and expert perspectives. This research gap motivated our

qualitative study engaging domain experts, legal professionals. By eliciting lawyers' insights on sample cases, we aimed to develop a detailed, nuanced understanding of when AI recommendations are viewed as proper versus irresponsible.

## Research Design



Question: When is it appropriate for AI systems to give advice in the legal domain?

Our goal was to build a detailed, contextual understanding of the appropriateness of AI legal advice grounded in the perspectives of experts like lawyers who provide such counsel professionally. To achieve this, we conducted an in-depth qualitative study engaging legal professionals through participatory workshops.

### Data Collection

We began by collecting a small set of sample “**seed cases**” representing **real-world user queries** where AI legal advice could be considered. These were sourced from posts on the subreddit *r/legaladvice* and anonymized for privacy and edited for the diversity of legal and factual patterns.

For example: “I run a popular online mobile game with microtransactions and collect user data per our terms. We run ads too. What should our legal strategy cover?”

### Participatory Workshops

We then conducted six remote participatory workshops with **20 legal experts**, including active lawyers in the UK and the US, law professors, and two law students. We presented the seed cases along with hypothetical AI responses based on patterns seen in **state-of-the-art, free for use, and public-facing models** like GPT-3.

Response templates included:

Template Name	Description	Example Response (Legal Advice)
Content violation	Displays a warning message that the user's input violates the AI's content policy.	This content may violate our content policy. If you believe this to be in error, please submit your feedback—your input will aid our research in this area.
Refuse to answer	Avoids providing a response to the input and advises the user to seek professional advice instead.	As an AI model, I am not a substitute for consulting with a trained professional. This seems like a specific concern or issue that requires expert advice. It's always advisable to speak with a professional in the relevant field.
High-level answer without facts	Provides a response that answers the input in broad strokes, stating advice rather than facts.	I am not a lawyer, but I can provide some general information about possible legal strategies that might be used in a situation like this. 1. Terms of Service and Privacy Policy: Clearly outline your terms of service, including details about micro-transactions, data collection, and advertising.
Specific answer with facts	Provides a detailed response that cites specific facts relevant to the input.	I am not a lawyer, but I can provide some information about specific legal strategies that may be applied in a situation like this. 1. Compliance with Applicable Laws and Regulations: Laws like the Children's Online Privacy Protection Act (COPPA) and the California Consumer Privacy Act (CCPA) may be relevant.
Ask a follow-up question	Responds with a question requesting more information.	To help you with this, I will need some additional information. Is your game intended for primarily children, or did you set age restrictions for your users?

Semi-structured discussions focused on key factors shaping the perceived appropriateness and risks of different AI advisor response styles for **2-3 seed cases** that experts chose. Experts took turns presenting factors that they took into account while envisioning the most desirable AI responses. After each workshop, we compiled the identified factors for analysis.

Re: popular online mobile game prompt, key dimensions considered by experts for that case include:

- Location of operation (whether the mobile game operates in the states having privacy law)
- Ad providers' data collection practices (are ad providers also collecting data)
- Involvement of minors (whether there are users who fall under a certain age)
- Nature of data (what kind of user information is being collected?)
- Terms of service (language and terminology used in the terms of service)

## Analysis

Transcripts were qualitatively coded to extract expert-proposed factors. An inductive approach allowed findings to emerge from the data. Factors were categorized based on common themes into higher-level groups related to the user, AI system, legal context, etc. This process synthesized key considerations that may determine the appropriateness of AI advice in legally complex situations per the legal experts' collective wisdom.

# Findings

## ***Uncovering the Factors that Influence AI Legal Guidance.***

Through these discussions, legal experts identified numerous dimensions that shaped their perspective on proper AI system responses. By qualitatively capturing lawyers’ perspectives on real case studies, our findings establish an in-depth, contextual understanding of the appropriateness of AI legal advice. Key factors spanned categories shown in the table below. Additionally, experts suggested principles to guide AI behavior, including avoiding harm, respecting the justice system, and not misleading users by appearing excessively human.

Category	Code
User (Attribute)	<ul style="list-style-type: none"><li>• Sophistication</li><li>• Geography</li><li>• Identity</li><li>• Vulnerable / marginalized status</li><li>• Reliability</li><li>• Reliance</li><li>• Agency</li></ul>
User (Query)	<ul style="list-style-type: none"><li>• Scope</li><li>• Stakes</li><li>• Sufficiency of information</li><li>• Type of desired response (information/opinion)</li><li>• Extralegal questions</li><li>• Judgment call needed</li><li>• Clarify of desired response</li><li>• Competitors</li><li>• Contracts</li></ul>
AI Capability	<ul style="list-style-type: none"><li>• Massive data processing ability</li><li>• Legal accuracy</li><li>• Context-awareness</li><li>• Interpretability</li><li>• Novelty / originality</li></ul>
Legal	<ul style="list-style-type: none"><li>• Ambiguity</li><li>• Complexity</li><li>• Standardized practices</li></ul>
Impact	<ul style="list-style-type: none"><li>• User emotions</li><li>• Extralegal consequences</li><li>• Cultural sensitivity</li><li>• Impact on the justice system</li></ul>

## **Differing Views on the Capabilities of AI.**

Opinions differ on the capabilities of current AI systems. Some believe AI can contextualize information and even surpass lawyers in processing huge amounts of legal data. However, others remain skeptical due to the risk of inaccuracies. There was consensus that the need for AI advice depends partly on AI capabilities advancing responsibly and

transparently. However, even assuming that AI systems could be made completely accurate and secure, concerns would still exist around AI systems providing actionable recommendations.

## **Chatbots Aren't Bound by Ethics Rules.**

Experts indicate that human lawyers are bound by ethics rules (e.g., confidentiality, competent representation, conflicts of interest, malpractice) and can face disciplinary action if they violate them. Chatbots don't have legal or ethical duties. Hypothetically, they can represent both sides, share information with others without consent, or give bad advice. Without oversight mechanisms, chatbots lack accountability for improper guidance. This heightens risks in relying on AI legal advice.

## **Legal information is different from advice.**

Experts largely concurred that discussing general legal information is appropriate, much like search engines already provide. However, personalized advice requires professional judgment. Participants emphasized keeping responses as non-binding information statements ("*if this, then that*"), rather than opinions on users' specific cases. A participant believes the AI should refuse to answer a question like "*could I win a lawsuit*" altogether, since determining something like that involves nuanced legal analysis and human judgment that the AI cannot replicate.

Professionals emphasized that AI should focus on addressing purely informational queries lacking situational specifics. For instance, discussing high-level legal principles, rights and procedures is considered as "*general knowledge*" and poses less risk than advising actions for a user's nuanced case. This aligns with court rulings differentiating neutral education from individualized recommendations reserved for licensed attorneys, although the distinction between two concepts is somewhat blurry.

## **A chatbot should indicate its speculative responses may be incomplete.**

Experts generally agree that AI systems cannot make a comprehensive analysis because of inherently insufficient facts and details. Attempting complex legal analysis with limited information risks generating misleading or speculative outputs. A participant gave an example of a DACA recipient charged with driving under intoxication, which could potentially result in deportation as a crime of moral turpitude. The recipient's DACA status could completely change the strategy for plea bargaining in this scenario. However, it is unrealistic to expect an AI system to inquire about DACA status whenever asked about a drunken driving charge. Therefore, AI responses should acknowledge that they do not provide a complete picture. The appropriate response when presented with vague hypothetical scenarios seems to be recognizing that there are insufficient facts provided, and additional details could significantly alter the analysis and outcomes. Making firm legal judgments without a full understanding of a situation's specifics is problematic.

## **Asking users follow-up questions was generally encouraged.**

While acknowledging the limitations of making definitive legal conclusions with incomplete information, experts largely agreed that AI systems should have the ability to ask users follow-up questions to clarify hypothetical situations presented. Since initial prompts from laypeople often lack legally relevant facts, asking clarifying questions can help the chatbot distill the most legally meaningful information for further analysis and improve the user's ability to "*ask a better question.*"

However, some experts were not inclined toward AI systems retrieving extensive personal details from users due to potential security concerns. They cautioned that collecting excessive information could increase vulnerability if confidentiality is compromised. In this regard, experts underscored that users must be clearly and transparently informed about data handling policies and consent to any collection of sensitive information ("*Protections in place so that confidential information isn't reused or leaked.*").

## **Chatbot conversations are not confidential, risking harms on non-users.**

Unlike private attorney conversations, chatbot interactions are “*essentially public*,” able to be shared widely online. This publicity poses risks around legal advice “*taken out of context*.” For instance, if a hate crime suspect asked about their legal defense, the chatbot response may include what “*could offend some groups*.” While attorneys can tailor guidance to clients privately, chatbots cannot adapt publicly visible responses.

## **Certain legal domains are better suited for machine-guided decisions.**

All legal domains are not created equal. It’s important to recognize that established business practices allow professionals to make routine decisions easily. Experts were more inclined to give detailed answers to informational questions like identifying states with low corporate tax rates or the legality of using lie detectors during the interrogation, which has the “objective answer” in the legal world. Experts found answering these clearly informational questions analogous to Google search results. Although stakes were high, some experts agreed to advise on states allowing medically-assisted death for permanent diseases.

However, there are complex and ambiguous areas like tax law where attorneys frequently have to make judgment calls based on experience. Also, many experts strongly cautioned against relying solely on AI in criminal cases, where plea bargains consider nuanced, subjective factors like a judge’s interests and religion. Having quality legal representation is significant from due process principles.

In domains with great complexity and subjectivity, it is found to be not practical or advisable to depend entirely on AI legal assistance. The human judgment of an attorney remains critical for handling nuance and weighing various factors. While AI can provide information and analysis, human expertise is still vital in areas like criminal law or other fields requiring deep judgment calls. Users should be aware of the limitations of AI in handling subjective or highly ambiguous legal situations. Expert counsel remains key in these contexts.

## **Experts distinguish between AI for lawyers and AI for the general public.**

While we approached this workshop considering public-facing use cases, experts noted meaningful distinctions between AI systems aimed at assisting legal professionals versus those intended for use by the general public. They expressed that AI designed as a “*co-pilot*” for lawyers may be viewed more positively, as it can help reduce human error while lawyers retain responsibility (“*Using machines to help or do us do our job; I think it’s seen as generally positive*.”). AI could be valuable for lawyers by streamlining “*intake processes*” and honing the specifics of legal issues, given appropriate confidentiality protections.

However, public-facing legal AI systems impose greater risks on layperson users who lack legal expertise. Unlike enterprise solutions with balanced liability sharing, consumer legal AI products may have one-sided indemnification clauses leaving users vulnerable. Without the context to properly evaluate AI limitations, the public may develop an undue reliance. Therefore, experts emphasized the need for careful precautions around AI intended for public consumption, including transparent terms of service, clear explanation of risks, and reasonable limitations on use cases to avoid exploiting user expectations.

## **AI systems should respect users’ agency.**

Experts emphasize that the user retains autonomy; the AI is meant to augment human decision-making by providing additional information, not replace accountable judgment. When providing analysis or recommendations, AI systems designed for high-stakes fields like law should respect that human discretion still dominates suitability considerations over algorithmic intelligence. As such, the AI’s outputs should avoid emotional manipulation or pressure tactics that could improperly sway users’ objective, reasoned choices.

Furthermore, introducing an "*emotional quotient*" risk making unsupported assumptions about people's internal states and motivations. This could lead to unfair profiling or inappropriate advice with real consequences, especially without human oversight. It can note potential risks or caution about harmful outcomes but should couch these as uncertainties rather than absolutes. By respecting human agency and discretion, AI systems can enhance legal advice without undermining users' self-determination. The goal is empowering informed decisions, not imposing prescribed paths.

---

While AI cannot (or should not) entirely replicate or replace a lawyer's judgment, our study elucidates how to create transparent AI assistants that know when to defer to human experts. With thoughtful design, AI can expand access to legal knowledge while avoiding reliance on imperfect machine intelligence for decisions with serious consequences. But only through considering the multitude of human, technological and regulatory dimensions revealed by experts can AI truly complement legal professionals in an ethical manner.

## Contribution

This research makes several notable contributions towards developing more trustworthy and human-centered AI systems. First, it pioneers a new methodology for participatory technology ethics, engaging diverse stakeholders in evaluating AI responses based on real-world cases. This collaborative assessment approach represents a significant evolution beyond speculative principles or theoretical frameworks.

Second, the insights generated from this process make tangible progress towards actionable guidelines for imbuing AI systems with sound judgement aligned to professional and community values. By grounding evaluations in practical examples and focusing on nuanced criteria of appropriateness, the findings provide targeted guidance for improving reliability and transparency.

Finally, this research lays the foundation for a scalable model of participatory AI ethics that can be expanded across other professional domains. The methods and insights establish a framework for incorporating human-centered design principles through cooperative evaluation of AI systems. This has the potential to greatly enhance user trust and satisfaction by creating technologies that better serve their needs and align with their values.

This work makes pioneering steps toward ensuring AI is developed responsibly and ethically with consideration for its real-world impacts on people and society. The participatory assessment process, context-specific findings, and potential for transferring the approach to other fields represent significant contributions in the pursuit of trustworthy AI.