

# Generative AI for the Legal Profession: Facing the Implications of the Use of ChatGPT through an Intradisciplinary Approach

2023-09-08 16:09:04 Kalliopi Terzidou

Chatbots are an accessible and direct way to get responses to your questions. They are more common in commercial environments, such as online shopping platforms, while their responses are not very extensive or sophisticated, creating the need for a more interactive conversation with a human representative. Many of these chatbots are based in previous versions of Generative Pre-trained Transformer (GPT) models. The difference of ChatGPT is that it provides responses that seem human-generated, given their grammatical and syntactical accuracy, as well as their contextual adaptation and occasional attempts of casual talk with the user.

On November 30<sup>th</sup> 2022, OpenAI released ChatGPT, a chatbot designed to respond to users' prompts in a conversational way. The algorithmic system is trained through [Reinforcement Learning methods](#), reacting to prompts that entail an order for a specific task, and which are further processed as inputs for the fine-tuning of the model so it can perform a second similar task. [In the case of ChatGPT](#), OpenAI's developers trained the initial model with conversations, adopting the sides of both the user and the chatbot and providing responses through model-written suggestions. The final dialogue format is a combination of this dialogue dataset and that of InstructGPT, a sibling model of ChatGPT. The reward consists in the ranking of alternative completions of a randomly selected model-written messages according to quality.

Entering the prompt "What is the punishment for robbery according to the Greek Criminal Code?" the chatbot returns the following information:

"According to the Greek Criminal Code, the punishment for robbery (Greek: ληστεία) can range from **imprisonment for a minimum of 5 years to a maximum of life in prison**, depending on the circumstances of the case. If the robbery was committed with violence or threat of violence against individuals, the sentence is typically harsher. Additionally, if the robbery was committed as part of an organized criminal group, the punishment may also be more severe." [emphasis added]

The response directly addresses the question of the type of punishment by mentioning the minimum and maximum duration of imprisonment, even though this is revealed through cross-reference to another article of the Greek Criminal Code (Article 52). The chatbot's main omission concerned the "circumstances" under which a person can receive life imprisonment, explicitly prescribed in Article 380 regulating robbery. Given the overall accuracy of the answer, the system could be used as an initial source of information, pending further verification and consultation by a legal professional. This raises the question of whether systems based on ChatGPT can, to some extent, replace certain functions of the legal profession. The following sections explain how ChatGPT can be relevant in the context of legal services and what are the regulatory implications arising from its use.

## 1. The Regulatory Implications of Using ChatGPT

The ability of ChatGPT to apply domain knowledge as per the request of the user raises the question of how the application can be relevant to [legal services](#), ranging from legal advice, legal research, and drafting of legal documents. The use of similar language models is not novel in the field of legal services. Generative models for legal advice have been employed by companies, such as [DoNotPay](#), an online platform that provides clients with arguments to be used in court proceedings, initially designed for cases regarding parking tickets. Legal research is also possible through Natural Language Processing (NLP) models, as is demonstrated by legal tech solutions, such as [ROSS](#), designed for the retrieval of U.S. jurisprudence based on the user's filtered query and ranking of results to return the answer most relevant to the original question.

The generative power of ChatGPT raises questions on the training of the system and on the accuracy of its responses, especially considering its use in legal services where interested parties should be well informed on their rights and obligations, depending on the circumstances of their case. It is, therefore, imperative to explore some of the key regulatory issues under the relevant EU legislative acts, focusing on Artificial Intelligence (AI) regulation, content moderation, and personal data protection.

### 1.1 AI Regulation

With the anticipated entry into force of the [AI Act](#) comes the question of whether generative AI systems, such as ChatGPT, can be covered by its provisions. Article 3 (1b) of the [amendment to the AI Act](#), submitted on 25<sup>th</sup>

November 2022 by the Permanent Representatives Committee (Part 1) to the Council of the EU, defines a General Purpose AI System (GPAIS) as an AI system “...intended by the provider to perform **generally applicable functions** such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system **may** be used in a **plurality of contexts** and be integrated in a **plurality of other AI systems**” [emphasis added]. ChatGPT can fall under this definition, since the application can generate text according to users’ prompts concerning different contexts, while it can be employed independently or as part of other AI systems. The definition seems to be technologically neutral, allowing for the regulation of the risks to human safety and rights imposed by multiple AI applications that can be characterized as GPAIS. At the same time, the choice of the word ‘may’ seems to suggest that AI systems used in *specific* contexts and designed for *specific* systems could also fall under the definition, rendering it hard to distinguish between regular AI systems and GPAIS in order to classify them as unacceptable, high, limited, and minimal risk AI systems, as foreseen by the Act.

Article 4b of the amendment states that GPAIS that may be used as or are components of ‘high-risk’ AI systems (i.e. AI systems that create a high risk to the health and safety or fundamental rights of natural persons) must comply with the requirements of Articles 8-15 of the proposed AI Act. These requirements include the implementation of risk-management systems (Article 9), data governance and management strategies (Article 10), compliance assurance (Article 11), record-keeping (Article 12), informing users on the operative principles of the system (Article 13), human oversight (Article 14), and design according to the appropriate levels of accuracy, robustness, and cybersecurity (Article 15). These conditions aim to prevent or restrict adverse harm to end-users’ safety and rights, but they also place a significant burden on developers and providers due to the considerable costs necessary for their adherence.

Article 4c of the amendment presents an exception to the above requirements, “when the provider has explicitly excluded all high-risk uses in the instructions of use or information accompanying the general-purpose AI system.” This condition may be hard to be complied with in practice, since the providers of GPAIS cannot *ex ante* exclude their future uses with absolute certainty, given their applicability in multiple contexts and as part of multiple systems.

## 1.2 Misinformation

Harmful content and misinformation could be produced by ChatGPT, following users’ prompts, with serious consequences to the users’ physical and mental state and information rights, respectively. It is important that the information ChatGPT provides is accurate and does not mislead users from the factual truth. Especially in the context of legal services, any legal advice generated by the AI system must closely follow current legislative and jurisprudential developments and be applicable to the facts of the case. Similarly, the use of ChatGPT for legal research must be conditional to the possibility of having access to updated and sufficient information on legislation and case law so legal professionals can adequately inform their practice.

On an EU level, the [Digital Services Act](#) regulates intermediary service providers (Article 2) by setting the conditions for a safe online environment for consumers, excluding risks such as the dissemination of illegal content and disinformation (Preamble 2). To this purpose, the established framework tackles the subject-matters of the liability of providers of intermediary services, their due diligence obligations, and the cooperation and coordination between the competent authorities (Article 1). Measures against the dissemination of illegal content in online platforms include prior warning and suspension of the provision of services to users that frequently provide manifestly illegal content and can extend to the employment of notice and action mechanisms and internal complaints-handling systems (Article 23). It is not currently possible, however, to directly regulate generative AI systems, such as ChatGPT. By ‘intermediary services,’ the Regulation refers to information society services that either transmit or (temporarily) store information of the recipients of the service (Article 3, point g). ChatGPT’s primary function is to generate content and not transmit or store information already provided by its users. Indirectly, [ChatGPT’s outputs could be regulated](#) when they are reproduced in social media, which do constitute intermediary services (hosting providers).

Incidentally, OpenAI implements a [content moderation tool](#) to detect and potentially block users generating harmful content, described as content that is “sexual, hateful, violent, or promotes self-harm.” Misinformation is not classified as harmful content, although it is described as a disallowed usage of OpenAI’s models under the category of ‘fraudulent or deceptive activity’ in the company’s [content/usage policy](#). Meanwhile, there are means to easily circumvent existing moderation tools, such as the tweaking of the system through designated prompts.

## 1.3 Personal Data Protection & Related Issues

Generative AI systems used by law firms must comply with applicable personal data protection rules. In cases where the clients' personal data must be processed, such as during the drafting of contracts, it is most likely that lawyers are not asking directly for their clients' consent, in accordance with Article 6 (1) (a) of the [General Data Protection Regulation](#) (GDPR). Even when the clients are notified of their data's processing, for example through the legal services agreement, this might not qualify as 'informed' consent, if it is not intelligible enough or easily accessible (Article 7). However, the legal basis of legitimate interest might be relied upon for the entry of clients' information in generative systems, given the ancillary use of these systems for the provision of legal services (Article 6, para. 1, point f).

A more tangible risk lies in the compromise of clients' personal data when they are re-purposed by ChatGPT so they can be used as training data for the refinement of its outputs. The processing of personal data, according to Article 5 (1) (b) of the GDPR, must be done for specific purposes and to the extent necessary. Nevertheless, once the data are entered in the generative AI system, they can be reused for different purposes and be indeterminately stored in the system's servers. In its [privacy policy](#), OpenAI states that when an individual visits, uses, and interacts with its services (including ChatGPT), the company receives relevant information, including log data (e.g. IP address, interaction with the website), usage data (e.g. types of content viewed or engaged with, features used and actions taken), and cookies. The policy further states that personal information may be processed, among others, "to develop new programs and services," while it may be aggregated and/or shared with third parties. These policy provisions imply a lack of control of the users' personal data, contrary to the GDPR rules that apply given that OpenAI's services are also offered to EU citizens.

Article 5 (1) (c) of the GDPR further clarifies that personal data must be accurate and be kept updated, according to the purpose for which they are processed. Yet, OpenAI [claims](#) that ChatGPT "has limited knowledge of world and events after 2021," meaning that (personal) data available online after the year 2021 do not constitute part of the system's training data so that its outputs are always updated and accurate. Overall, the privacy policy text is quite vague on the rights of data subjects, in breach of their right to access information on the processing of their personal data under Article 13 of the GDPR.

Especially in the field of legal services, the entry of personal data in the generative AI system, irrespective of the applicable legal basis, may compromise the confidentiality between lawyers and their clients, since data originating from the client's personal file are now shared with the provider of the system and can be re-used in similar ways to the ones stated above. More data governance measures would be needed for the safe use of generative AI systems, that should also cover cybersecurity risks, such as [inversion attacks](#) (malicious user attempts to recover the training dataset used by the AI system, that may also include personal data).

It is important to note the Italian Data Protection Authority's [Order of the 11<sup>th</sup> of April 2023](#) against OpenAI, regarding a series of measures that the company, as a controller, must adhere to, so that the functioning of its service ChatGPT is compliant with the provisions of Legislative Decree No 196 of 30 June 2003, transcribing the GDPR into Italian national law. The Order reflects some of the points mentioned above, in that it requests the company to provide an information notice to data subjects on, among others, how their data are processed for the purposes of training the algorithm, that is displayed in the registration flow for better visibility and is further communicated through an information campaign across the Italian media. The Garante continues by ordering OpenAI to make available a tool that gives the possibility to data subjects to enforce their rights of rectification of incorrectly processed personal data, as depicted in the generated results, and of objection to their processing for training purposes. Lastly, OpenAI is required to implement an age verification tool, so that users below 13 years of age are prevented from accessing ChatGPT and users below 18 are prevented to do so in the absence of consent from their parents or legal guardians.

Interestingly, the Garante asked for the reliance of the company on consent or legitimate interest as legal bases for the processing of personal data, instead of the basis of performance of a contract. In its updated privacy policy, OpenAI has not excluded the legal basis of the performance of a contract but now mentions the additional legal bases of legitimate interest "in protecting our Services from abuse, fraud, or security risks, or in developing, improving, or promoting our Services, including when we train our models," and of consent "when we ask for your consent to process your Personal Information for a specific purpose that we communicate to you." The legal basis of compliance to legal obligations is also mentioned when the company must adhere to the applicable law or for the protection of its users', affiliates', and third parties' rights, safety, and property. Again, the policy is vaguely stated (for example, there is no explanation of the meaning or examples of "specific purpose" for the request of consent) and fails to address issues of data minimization for the purposes of training, by simply characterizing the activity as a 'legitimate interest.'

The additional absence of clear anonymization measures and the reported aggregation of personal data render discrimination likely to occur, based on features such as gender and ethnicity, when they are processed for the training of the model. This could have the effect of providing different types of outputs, for instance legal advice, in similar cases. Restrictions imposed by intellectual property law on accessing the training datasets impede the effective external scrutiny of the development of generative AI systems, leaving monitoring obligations to the discretion of the provider and its developers.

## 2. The Safe Use of Generative AI Systems in Legal Services

In regulating the reported risks of generative AI systems, a holistic approach must be adopted that encompasses legal, technical, and societal preventive and restrictive measures.

From a legal point of view, a technologically neutral approach to the regulation of generative AI systems is a good strategy to ensure the widest regulatory coverage of technological methods, irrespective of their state of art or their distinct features, as law cannot keep up with technological advancements. Concrete rules against misinformation incidents should be laid down on a regional level, so requirements such as the suspension of users and notice and action mechanisms can be enforced homogeneously in all states where GPAIS are offered. Similarly, data protection rules should be strictly enforced by supervisory authorities against malpractices and violations of the GDPR and the national legislations transcribing the Regulation. To this end, the privacy policies of the providers of generative AI systems must be transparent, in the sense that they provide information on the processing of the users' and end-recipients' personal data (for example, data categories, data controllers, data processors, transfers to third parties) in an intelligible manner. Legal measures should promote innovation by allowing regulatory sandboxes, so private actors have the creative freedom to produce AI solutions that might be beneficial for the legal sector and raise the competitiveness of the EU as a global leader in the field of AI.

Technical measures for the safe use of generative AI could include the diligent collection and processing of data, in a way that instances of unlawful or unnecessary processing are excluded. Human oversight, where possible independent, should be encouraged throughout the entire lifecycle of the system so incidents such as the accidental correlation of data resulting in discriminatory outputs are prevented or restricted. For accountability purposes, AI systems should be designed as explainable to the extent possible, given the proprietary protection of the system's datasets. The explanation should focus on the training dataset and the functioning parameters of the system, to account for outputs that cause, directly or indirectly, harm, misinformation, and/or discrimination. In any case, a systematic application of the requirements of the AI Act for high-risk AI systems must be imposed for the equal protection of all EU citizens; in specific, a risk-based approach should be adopted during the whole lifecycle of the system, including regular monitoring and mitigation measures in cases of cybersecurity threats and incidents.

On a societal level, law students and young lawyers are not adequately prepared to use generative or similar AI systems as ancillary tools to their studies and work, respectively. This might be due to conservatism, unsystematic use of such technologies on an everyday basis, or lack of training resources. Intradisciplinary skills are important for future or junior legal professionals, in order to adapt to market shifts, and for law firms, in order to gain a competitive advantage through the possession and use of advanced technological means for a more efficient case management. Especially regarding the use of ChatGPT for legal information retrieval, it is important that lawyers are aware that the results provided by OpenAI's solution are not alike the ones provided by search engines. Search engines offer access to a wide range of older and more recent information, depending on the keywords and filters that users insert in the search bar. ChatGPT, on the other hand, generates responses to users' prompts that might not be based on updated data and might contain fabricated information, thus demanding much more awareness when employed by legal professionals for their research.

The likelihood of replacement of lawyers by generative AI systems is not absolute, but it might reflect cases where AI assistants cover clients' demands for legal advice, for instance in small claims cases of the civil law branch. This means that a considerable fraction of legal professionals might need to abandon their traditional roles and [adopt alternative ones in the edge of law and technology](#). New roles could involve the legal guidance of developers of generative AI systems to supply them with the domain knowledge necessary to train such systems or with the methodological approaches to solve legal issues that can help in appropriately structuring the collected legal data. Regardless of new possible roles, there will be a need of lawyers to represent clients before courts and to provide legal advice on complex legal issues that require expert interpretation of legislation and case law, in domains such as criminal and administrative law.

The emergence of ChatGPT has already sparked the interest of the public, with [100 million monthly active users in January 2023](#) setting a record as the fastest-growing consumer application to exist. [Law firms](#) and [courts](#) are

among these consumers, using the chatbot as a support tool for analyzing contracts or for decision-making, respectively. Considering the fast uptake of these applications, the constant refinement of the model through reinforcement learning methods, and the possibility of fine-tuning it to fit domain-specific purposes, regulative initiatives on generative AI systems must become a priority for regional and national policymakers and legal experts must be ready to swiftly adapt to considerable changes within their profession.

**Share this article!**

 Post