

Artificial Intelligence in the Service of Self-represented Litigants

Accessing the Law with the Aid of Large-Language-Models

Kalliopi Terzidou

Abstract: Lawyers play a crucial role in representing litigants in court and providing legal counsel on their rights, obligations, and the procedural steps involved in complex pre-trial and trial proceedings. However, the excessive costs of legal services and the absence of (sufficient) legal aid mechanisms can exclude certain citizens from seeking a dispute settlement before courts. Chatbots based on Large Language Models (LLMs) have become popular due to their user-friendly way of communicating information on substantive and procedural law and on court proceedings. Nevertheless, LLM-based chatbots can mislead users by providing inaccurate information, ultimately affecting the progress of their case before courts. The present paper asks to what extent Large Language Models-based chatbots are designed in a human centric way to support self-represented litigants in accessing legal information. To address this question, the paper explores the concept of ‘human centricity’ under the EU’s framework for trustworthy Artificial Intelligence, in order to develop a set of functional requirements against which the performance of LLM-based chatbots is assessed in the context of landlord-tenant disputes.

1. Introduction

Since late 2022, ChatGPT¹ and other chatbots have emerged as virtual assistants, providing users with answers to a wide range of questions. These chatbots are based on Large Language Models (LLMs), i.e. Artificial Intelligence (AI) models trained on a large amount of data through self-supervision techniques, so they can perform generally applicable functions in

1 OpenAI, ‘ChatGPT’ <<https://chatgpt.com>> accessed 07 July 2025.

different contexts.² Popular³ online LLMs online are OpenAI's ChatGPT, Microsoft's Copilot,⁴ and Google's Gemini.⁵

Legal aid applications based on AI systems exist in the market for several years, offering different kinds of legal support to users. Websites such as DoNotPay⁶ and JusticeBot⁷ offer legal 'advice' and resources to users experiencing everyday legal problems, such as housing and landlord-tenant issues. Other websites, such as A2J Author⁸ and Court Forms Online,⁹ assist users with the completion and filing of court forms. These types of AI systems are usually rule-based, i.e. trained on a limited dataset to perform a specifically designated set of tasks, in contrast to LLM-based chatbots (such as the ones mentioned above) that are General-Purpose AI systems, i.e. AI systems able to address a variety of purposes by generating multiple forms of content.¹⁰ Chatbots employ a conversational tone to help users complete tasks, including legal research, summarizing and translating of legal resources, and drafting legal documents.¹¹ Users can personalize their experience by refining their prompts or even customizing the available settings. Unlike lawyers, these applications are available on a 24/7 basis from a remote location, requiring only a digital device connected to the Internet. In this way, legal information becomes more accessible to people who would not otherwise have the financial resources to consult a lawyer. Ultimately, users can reduce their exposure to bureaucratic delays while accessing court services, such as filing a court form, and the stress incurred when visiting the lawyer's office or the court.

2 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L/1689 (AI Act) recital 99 and article 3(66).

3 See, for example, Krystal Hu, 'ChatGPT Sets Record for Fastest-Growing User Base – Analyst Note' *Reuters*, (2023) <<https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>> accessed 03 October 2024.

4 Microsoft, 'Copilot' <<https://copilot.cloud.microsoft>> accessed 07 July 2025.

5 Google, 'Gemini' <<https://gemini.google.com/app>> accessed 07 July 2025.

6 DoNotPay, <<https://donotpay.com/>> accessed 03 October 2024.

7 JusticeBot, <<https://justicebot.ca/question/1/>> accessed 03 October 2024.

8 A2J Author, <<https://www.a2jauthor.org/>> accessed 03 October 2024.

9 Court Forms Online, <<https://courtformsonline.org/>> accessed 03 October 2024.

10 AI Act, recital 99 and article 3(66).

11 Kalliopi Terzidou, 'Generative AI Systems in Legal Practice Offering Quality Legal Services While Upholding Legal Ethics' (2025) *International Journal of Law in Context* 1, 12 <<https://doi.org/10.1017/S1744552325000047>> accessed 07 July 2025.

At the same time, most legal aid applications of any type entail certain implications. First, they are jurisdictionally limited, i.e. providing legal information and resources specific to a country that cannot be extended to states with different legal systems. Second, they are available exclusively or primarily in the English language, thus excluding a large part of Internet users from accessing them. Third, they are typically designed for simple and straightforward legal matters, thus not being able to handle nuanced or multi-jurisdictional legal issues. For example, when prompted to draft a legal document, such as a court form, the application may return plain templates, so that users must customize them according to the applicable legal framework and the facts of the case. Fourth, LLMs, in particular, can hallucinate, i.e. generate content that is fictitious and results in unreliable outputs.¹² Fifth, the applications' providers are not bound by the same rules of professional conduct as lawyers are during the consultation and representation of their clients, such as the principle of confidentiality and professional secrecy,¹³ meaning that they cannot be accountable to the public for erroneous information that their LLM-based solutions return to them. Finally, there are concerns over the protection of users' (personal) data when using these AI applications, as well as over the exclusion of certain members of society without digital literacy or access to a digital device and/or a stable Internet connection from taking advantage of the benefits of these AI systems.

The aim of this paper is to, first, define 'human-centric' functional requirements for the design of legal aid applications based on AI systems, so that these provide equal and safe access to justice to self-represented litigants (SRLs) by retrieving understandable and reliable information on substantive and procedural law and on court proceedings. Second, the paper aims at discovering whether the design of LLM-based chatbots corresponds to these requirements, in order to understand if they can indeed help SRLs to access legal information in a human-centric way. Therefore, the research question asks:

To what extent are Large Language Models-based chatbots designed in a human centric way to support self-represented litigants in accessing legal information?

12 See, for example, Zhiqing Ji and others, 'Survey of Hallucination in Natural Language Generation' *arXiv*, (2022) 2 <<https://arxiv.org/abs/2202.03629>> accessed 7 July 2025.

13 Providers are mainly bound by product safety rules, such as those laid down in the AI Act concerning the placement in the market of (high-risk) AI systems.

To address this question, ‘human-centricity’ is defined based on the EU’s notion of ‘trustworthy AI’ as developed in two main documents: the Ethics Guidelines on Trustworthy AI¹⁴ and the AI Act.¹⁵ In particular, the principle of human-centricity denotes the respect of human dignity and of fundamental rights, such as the right to privacy, which are inherently connected to the (ethical) principles of respect for human autonomy, prevention of harm, fairness, and explicability.¹⁶ Based on this definition, a set of functional requirements for legal aid applications is outlined, against which seven LLM-based¹⁷ chatbots are evaluated, in particular ChatGPT, ChatGPT Free Legal Advice, ChatGPT Pro, Copilot, Copilot for students or enterprises, Gemini, and Gemini Advanced. These chatbots are the free and the subscription-based versions of ChatGPT, Copilot, and Gemini,¹⁸ which are popular chatbots and, thus, most likely to be accessed by the average Internet user. This diversity of chatbots allows for more extensive and valuable comparisons based on the human-centric functional requirements.

‘Self-represented litigants’ (SRLs) are individuals who are not represented by a legal professional during court proceedings. The choice to represent oneself before the courts may be attributed to several factors, including unaffordable legal services and insufficient legal aid means provided by the State. A fragmented situation is observed in the EU, where only nine Member States offer full coverage of litigation costs along with partial legal aid schemes, while ten offer either full or partial legal aid, and four leave the matter on the discretion of courts.¹⁹ Furthermore, in six Member States individuals are not automatically exempted from paying court fees, while in three Member States people with income below the Eurostat poverty

14 High-Level Expert Group on AI (European Commission), ‘Ethics Guidelines for Trustworthy AI’ (2019) 37 <<https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>> accessed 07 March 2025.

15 AI Act.

16 High-Level Expert Group on AI (European Commission), ‘Ethics Guidelines for Trustworthy AI’ (n 14), II–20.

17 For the definition of a Large Language Model, see AI Act, recital 99 and article 3(66).

18 Even though most users are likely to visit the free version of these chatbots, the subscription-based versions are also accessible for a considerably small monthly or annual fee, arguably offering a more advanced user experience.

19 European Commission, ‘2025 EU Justice Scoreboard’ COM(2025) 375 final (1 July 2025) 20–21 <https://commission.europa.eu/document/download/51b21eff-a4b0-4e73-b461-06bd23b43d4e_en?filename=2025%20EU%20Justice%20Scoreboard_template.pdf> accessed 7 July 2025.

threshold are unable to receive legal aid.²⁰ The lack of guarantees in favour of legal aid across the EU underlines the gap in access to justice and the attraction of individuals to self-representation. It must be noted that self-representation is not allowed in all types of court proceedings. This practice can generally be observed in small claims cases, as can be seen in the examined case of landlord-tenant disputes in Greek courts, whereby a landlord can self-represent himself or herself before Small Claims Courts (Ειρηνοδικεία) when the claim is under EUR 600.²¹

One way to support SRLs in their legal journey is to assist them in navigating the complex regulatory landscape governing their legal issue. These rules are expressed in lengthy and ambiguous documents²² which are often charged with legalese.²³ For the needs of this research, ‘legal information’ refers to publicly available resources such as legislation, case law, and information on court proceedings, which are published by state and judicial institutions in analogue and/or digital format. All Member States provide access to online information about their judicial system for the public, including information on legal aid, court fees, procedural rights, and court forms.²⁴ Several Member States (Germany, Estonia, Latvia, Greece, Croatia, Austria, Portugal) reportedly integrate chatbots in their websites to help the public retrieve information on their justice system.²⁵ The increasing adoption of chatbots by national justice systems for the retrieval of legal information indicates an effort by EU Member States towards enhancing access to justice for their citizens. Legal information is democratized, in contrast to the longstanding ‘monopoly’ that legal experts have exercised over legal knowledge. LLM-based chatbots can offer legal information in plain and concise language, at little or no cost, and on a 24/7 basis, offering

20 *ibid* 21 and 106.

21 Greek Parliament, ‘Code of Civil Procedure (Κώδικας Πολιτικής Δικονομίας)’ (n 28), articles 14 (1)(b) and 94 (2)(a).

22 Max Barrett, *The Art and Craft of Judgment Writing: A Primer for Common Law Judges* (Globe Law and Business 2022) 39–46 <<https://www.globelawandbusiness.com/books/the-art-and-craft-of-judgment-writing-a-primer-for-common-law-judges-second-edition>> accessed 07 July 2025.

23 Z Zódi, ‘The limits of plain legal language: understanding the comprehensible style in law’ (2019) 15(3) *International Journal of Law in Context* 246, 257–258 <<https://www.cambridge.org/core/journals/international-journal-of-law-in-context/article/abs/limits-of-plain-legal-language-understanding-the-comprehensible-style-in-law/B07D80A34052ED1E6EF7F7C289EC97DD>> accessed 07 July 2025.

24 European Commission, ‘2025 EU Justice Scoreboard’ (n 19), 32–33.

25 European Commission, ‘2025 EU Justice Scoreboard’ (n 19), 33.

an unyielding support to SRLs (notwithstanding phenomena such as hallucinations indicative of LLMs).

For the evaluation of these systems, the researcher first assumed the persona of a landlord whose tenant does not pay the rent on time. The landlord does not know if this qualifies as a legal issue or how it is legally treated under Greek law. Greece has been the targeted jurisdiction due to the familiarity of the researcher with Greek law, having obtained an LL.B in Greece, so as to be able to reliably evaluate the chatbots' responses. Another reason is the particularity of Greek online resources. Because the majority of online content is in the English language, LLMs trained on Internet-sourced data perform typically better on English-language prompts. Consequently, the LLM-based chatbots are tested in the Greek language to confirm whether LLM-based chatbots underperform on Greek prompts due to the limited availability of Greek online resources.

A total of 16 questions (Annex I) were posed to the LLM-based chatbots, encompassing matters of landlord-tenant relations that are also regulated by national law, primarily by the Greek Civil Code²⁶ and Code of Civil Procedure.²⁷ These regulations treat three main regulatory areas: the legal rights of the landlord, the competent court before which these rights can be defended, and the applicable procedure. It must be noted that under Greek law, landlords can choose between two legal pathways when their tenants do not pay rent. The landlord can either apply for an *action* ('αγωγή')²⁸ or an *order* ('διαταγή')²⁹ for the restitution of rent.³⁰ A

26 Greek Parliament, 'Civil Code (Αστικός Κώδικας)' ΦΕΚ: Α 164 19841024, (1946) <<https://www.ministryofjustice.gr/wp-content/uploads/2019/10/Αστικός-Κώδικας.pdf>> accessed 11 March 2025.

27 Greek Parliament, 'Introductory Law of the Code of Civil Procedure (Εισαγωγικός Νόμος Κώδικα Πολιτικής Δικονομίας)' ΦΕΚ: Α 182 19851024, (1968) <<https://www.lawspot.gr/nomikes-plirofories/nomothesia/eisnkpold/arthro-66-eisagogikos-nomos-kodika-politikis-dikonomias>> accessed 11 March 2025.

28 Greek Parliament, 'Civil Code (Αστικός Κώδικας)' (n 25), article 597; Greek Parliament, 'Introductory Law of the Code of Civil Procedure (Εισαγωγικός Νόμος Κώδικα Πολιτικής Δικονομίας)' (n 26), article 66.

29 Greek Parliament, *Code of Civil Procedure (Κώδικας Πολιτικής Δικονομίας)*, Presidential Decree 503/1985, FEK Α 182/24.10.1985, articles 623–636 <<https://www.laws.pot.gr/nomikes-plirofories/nomothesia/kodikas-politikis-dikonomias>> accessed 11 March 2025.

30 The main difference between the action and the order is that upon bringing an action before the court, the lease is terminated, while the order does not result in the termination of the contract. Moreover, the decision on the action must become final before it can be enforced against the tenant, while the order is an enforceable title

landlord can self-represent himself or herself only when the legal dispute concerns a small claim (under EUR 600) to be judged before Small Claims Courts (Ειρηνοδικεία).³¹ The answers by the LLM-based chatbots to these questions were assessed on a scale from 1 (low) to 5 (high) based on the designated human-centric functional requirements. The points were assigned by the researcher herself, based on her user experience and the verification of the retrieved information against Greek legislation.³²

The hypothesis is that users are empowered to take legal action and defend their legal position before courts by accessing quality legal information on substantive and procedural law, that is readable and accessible to all so that they can rely upon it. In any case, users must contextualize and verify any information retrieved by LLM-based chatbots according to the facts of their cases, since the responses to their prompts may be generic and thus not applicable to their particular case or generally inaccurate, respectively.

The paper is divided into three main sections. The first section explores ‘human-centricity’ under the EU’s framework for trustworthy AI to ultimately define a set of functional requirements for the design of legal aid applications based on AI systems. The second section evaluates the selected LLM-based chatbots against these functional requirements to understand if they are designed in a human-centric way to support SRLs involved in landlord-tenant disputes. The third section critically examines the conditions for integrating LLM-based chatbots into national court websites for the provision of court services.

2. A Human-Centric Design

According to a 2021 HiiL and IAALS survey, people interested in resolving everyday issues of a family, land, or employment nature are likely to seek

and can be executed upon the expiry of a 20-day period. The action is also a more expensive and time-consuming procedure than the order; e-Nomika, ‘Διαδικασία Έξωσης ενοικιαστή: Χρόνος και κόστος’ <<https://e-nomika.gr/διαδικασία-έξωσης-ενοικιαστή-χρόνος/#:~:text=Με%20την%20άσκηση%20της%20αγωγής,αίτηση%20έκδοσης%20της%20διαταγής%20απόδοσης>> accessed 11 March 2025.

31 Greek Parliament, ‘Code of Civil Procedure (Κώδικας Πολιτικής Δικονομίας)’ (n 28), articles 14 (1)(b) and 94 (2)(a).

32 Future research could replicate the present methodology by involving participants in an experiment setting, where the chatbots’ performances could be tested against an even wider set of human-centric functional requirements potentially defined by participants.

legal advice on the Internet, at least before resorting to a legal expert.³³ This was often attributed to people not considering the help of a lawyer to be necessary or being unable themselves to identify their issue as a legal one.³⁴ When visiting the Internet, most people admitted that the retrieved information affected them considerably by helping them understand how to solve their issue,³⁵ notwithstanding the importance of trustworthiness, clarity, and context-specificity in rendering the information helpful and usable for users.

The widespread use of the Internet for legal information highlights the need for web services to be designed with a human-centric approach. The principle of ‘human-centricity’ is fundamental in EU’s approach towards trustworthy AI. According to the High-Level Expert Group on Artificial Intelligence, a human-centric AI approach aims to ‘...ensure that human values are central to the way in which AI systems are developed, deployed, used and monitored, by ensuring respect for fundamental rights...all of which are united by reference to a common foundation rooted in respect for human dignity...’³⁶

For a human-centric approach to be respected in the designing, development, deployment, and use of AI systems, the High-Level Expert Group on Artificial Intelligence claims that four ethical principles should be respected: fairness, respect for human autonomy, explicability, and prevention of harm.³⁷ According to the *principle of fairness*, all individuals and groups must receive an equal and just treatment when interacting with AI systems, which entails two dimensions: a substantive one (avoid unfair biases that can lead to their discrimination) and a procedural one (opportunity to contest through an effective redress the outcomes of AI systems).³⁸ All AI systems, including LLM-based chatbots designed to retrieve legal information, should provide equal and just access to available resources, taking

33 The Hague Institute for Innovation of Law (HiiL) and the Institute for the Advancement of the American Legal System (IAALS), ‘Justice Needs and Satisfaction in the United States of America 2021: Legal problems in daily life’ (2021) 160 <<https://iaals.du.edu/sites/default/files/documents/publications/justice-needs-and-satisfaction-us.pdf>> accessed 03 October 2024.

34 *ibid* 164.

35 *ibid* 166.

36 High-Level Expert Group on AI (European Commission), ‘Ethics Guidelines for Trustworthy AI’ (n 14) 37.

37 *ibid* 9–13.

38 *ibid* 12–13.

into consideration the users' level of legal and digital literacy to adapt the accessibility and understandability of the generated information to their needs.³⁹ Legal literacy is important for the empowerment of SRLs, since it provides people with awareness and understanding of information on key legal principles and rules pertaining to their rights and duties, enabling them to take action when faced with a legal issue.⁴⁰ Digital literacy is also important since it allows SRLs to remain in control of the AI system by mastering its different features to arrive to the desired result.⁴¹ AI systems, including LLM-based chatbots, might present challenges related to legal and digital literacy, potentially confusing users by generating inaccurate, incomplete, or otherwise unreliable legal information (legal literacy) or when they present this information through a complex interface with limited or unclear ways of interacting with the available features (digital literacy). SRLs as users may be more vulnerable to harm from LLM-based chatbots, since they are more likely to lack the necessary legal knowledge to assess the reliability of the generated legal information and may further not have access to legal consultation due to the high legal fees. This result may be exacerbated if they are not digitally literate and thus in a position to distinguish between the advice of a legal expert and that generated by an

-
- 39 People should take part in the designing process of legal aid applications to define users' expectations and needs regarding legal and court services and give feedback on the applications' features from an end-user's perspective. Participation can take the form of surveys or workshops, among other types of interventions; OECD, 'OECD Framework and Good Practice Principles for People-Centred Justice' (2021), 14–15 and 66–72 <https://www.oecd.org/en/publications/oecd-framework-and-good-practice-principles-for-people-centred-justice_cdc3bde7-en.html> accessed 03 October 2024.
- 40 For a definition of 'legal literacy' or 'legal education', see The Law Society, 'Public legal education' <<https://www.lawsociety.org.uk/campaigns/public-legal-education>> accessed 07 July 2025; Such interventions should be inclusive by providing equal opportunities for legal education, considering the different ways that people view, consider, and digest legal information, in text or other media, such as videos. In any case, legal information should be available in a plain and clear language, devoid of heavy legal jargon.
- 41 For a definition of 'digital literacy', see UNESCO Institute for Statistics, 'SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all' 1 <<https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2021/08/Metadata-4.4.2.pdf>> accessed 07 July 2025; Deployers of such applications, including courts, must also offer an intuitive browsing experience through simple interfaces and customizable features, as well as alternatives for citizens that do not have access to a digital device or stable Internet connection, equipping public spaces (e.g. public libraries) with freely accessible digital equipment and Wi-Fi connection.

LLM-based chatbot. in the latter case, users may excessively rely on the system's outputs due to the persuasive way that the generated information is communicated (automation bias).⁴²

According to the *principle of respect for human autonomy*, as determined by the High-Level Expert Group on Artificial Intelligence,⁴³ individuals must be able to retain full and effective self-determination, without any unjustified interference, when interacting with AI systems, in order to be empowered to make meaningful choices and participate in democratic society. When interacting with LLM-based chatbots, users may perceive the systems' responses as expert advice and accept them uncritically due to the highly persuasive conversational tone of the systems, inadvertently relinquishing some of their agency and taking less informed decisions. This is why AI systems must be transparent so their purposes, capabilities, and decisions can be communicated in an explainable manner to those directly or indirectly affected, following the *principle of explicability*.⁴⁴ This is not always perceivable in General-Purpose AI systems, including LLM-based chatbots such as ChatGPT, the functioning of which remains a black-box in an effort to protect them as proprietary models, hindering the understanding of the mechanism behind the generation of their outputs that is critical to maintain human agency and oversight over the systems.

According to the *principle of prevention of harm*,⁴⁵ AI systems should not result in physical and mental harm, but should ensure the dignity and safety of users, especially vulnerable individuals, when faced with malicious use and power asymmetries. With respect to LLM-based systems, prevention of harm is difficult to achieve given the considerable amount of training data that these systems need, which in this case can reach trillions of data,⁴⁶ that might include sensitive data points. In particular, these training datasets can

42 Hannah Ruschemeier, Lukas J Hondrich, 'Automation bias in public administration – an interdisciplinary perspective from law and psychology' (2024) 41(3) Government Information Quarterly 101953 1 <<https://doi.org/10.1016/j.giq.2024.101953>> accessed 7 July 2025.

43 High-Level Expert Group on AI (European Commission), 'Ethics Guidelines for Trustworthy AI' (n 14), 12.

44 *ibid* 13.

45 *ibid* 12.

46 Nevertheless, the limited financial resources that a provider, such as national (judicial) institutions, might have available to develop LLM-based chatbots from scratch or fine-tune such systems can lead to the training being based on a limited amount of data and result in applications that do not perform well their designated task; Ashwin Telang, 'The Promise and Peril of AI Legal Services to Equalize Justice' (*Harvard*

well include personal data provided by the users themselves through their prompts or cookie preferences or automatically collected by the AI systems (for example as metadata). The processing of personal data for training purposes can lead the system to infer sensitive human characteristics, such as age and ethnicity, and ultimately discriminate against vulnerable groups, including women and ethnic minorities, resulting to their exclusion from legal or court services in violation of the principle of fairness. If the AI system has suffered from technical errors or produces harmful outputs, including outputs that lead to adverse discrimination or breaches of personal data protection, then mitigation measures must be implemented to prevent further harm, following the results of applicable impact assessments.

The AI Act likewise recognizes the importance of human-centricity, stating in article 1 (1) that ‘The purpose of this Regulation is to improve the functioning of the internal market and promote the uptake of human-centric and trustworthy artificial intelligence (AI)...’⁴⁷ In this way, the Act operationalizes the ethical principles of human-centric AI, as previously stated by the High-Level Expert Group, and further sets down binding legal rules for providers, deployers, and other actors involved in the ecosystem of AI systems. These rules follow a risk-based approach, according to which certain AI practices are prohibited, such as the use of AI systems that employ subliminal, manipulative, or deceptive techniques to distort behaviour and impair informed decision-making (article 5, para. 1, point a). The production and use of systems that are described as ‘high-risk,’ i.e. systems that pose a ‘significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making’ (article 6, para. 3), is in principle allowed, following strict product safety requirements that aim at a more transparent and safe development and deployment of an AI system (articles 8–15).

For ‘high-risk’ AI systems, the principle of fairness is addressed through the requirement of data governance (article 10), to ensure data quality and to prevent discriminatory outcomes; the respect for human autonomy is expressed in the requirement for human oversight over the AI system (article 14), to guarantee meaningful outcomes and informed decisions by users; the principle of explicability is embedded in documentation (article 11) and transparency obligations (article 13), to record the functioning of the system

Journal of Law & Technology Digest, 2023) <<https://jolt.law.harvard.edu/digest/the-promise-and-peril-of-ai-legal-services-to-equalize-justice>> accessed 03 October 2024.

47 AI Act.

and allow for the explainability and contestability of its outputs, and; the principle of prevention of harm is operationalized through provisions such as risk management (article 9), to allow the assessment of risks and their effective mitigation.

3. Human-Centric Design of AI-Based Legal Aid Applications

A ‘human-centric’ approach to designing AI systems, as outlined in Section 2, provides a useful ethical and human rights-based framework under which the design of LLM-based chatbots can be evaluated. More specifically, the ethical principles of fairness, respect for human autonomy, explicability, and prevention of harm, can provide this assessment framework,⁴⁸ which can become binding when these principles are translated into requirements under the AI Act.

The principle of fairness demands that AI systems are designed to equally and justly distribute their benefits, while avoiding the production of unfair biases that can result in discrimination of any kind (substantive dimension).⁴⁹ In the context of the present research, AI systems must be accessible regarding their features, so that all users are guaranteed equal access to the benefits of LLM-based chatbots, particularly digitally illiterate users. Accessibility standards ensure inclusivity, especially for people with disabilities, enabling equitable access and participation in digital assistive technologies. Accessibility is greatly associated with the user interface, i.e. how a user perceives and interacts with the features of LLM-based conversational interfaces. The interface must be *user-friendly*, meaning that it is interactive and responsive to users’ instructions, allowing a personalized and intuitive experience through customizable settings.

The outputs of AI systems (as well as their disclaimers and terms of use) must be *readable*, i.e. easy to understand,⁵⁰ in order to reduce information asymmetries, particularly regarding legal information which may not be readily understood by the wider public, and thereby enhance access to the

48 High-Level Expert Group on AI (European Commission), ‘Ethics Guidelines for Trustworthy AI’ (n 14) 15–19.

49 *ibid* 12–13.

50 See, for example, Margaret Hagan, ‘Good AI Legal Help, Bad AI Legal Help: Establishing quality standards for responses to people’s legal problem stories’ in *JURIX 2023: 36th International Conference on Legal Knowledge and Information Systems, AI and Access to Justice Workshop, Maastricht*, (2023) 4 <<https://ssrn.com/abstract=4696936>> accessed 03 October 2024.

justice system. AI-based legal aid applications should minimize legal jargon and offer alternative information formats, such as visualizations, to ensure accessibility for users with disabilities and minors who may benefit from audiovisual stimuli for better comprehension. The application must also be available in multiple languages, including the official language(s) of the jurisdiction where the application is offered, to ensure its unhindered use by all users.

According to the principle of respect for human autonomy, users of LLM-based chatbots should be able to make informed and autonomous decisions based on the systems' outputs, without the risk of misinformation or deception.⁵¹ In the context of the present research, AI systems must protect the autonomy of individuals by ensuring access to *quality information* that is accompanied by references to accompanying trustworthy sources, so that users avoid being misinformed and can safely rely on the outputs of AI systems to take legal action. The quality of the retrieved legal information amounts to the accuracy, completeness (sufficient details), consistency,⁵² and (situational) relevance (alignment with the user's information needs) of the systems' responses.⁵³ Since the information retrieved by legal aid applications is of a legal nature, the information further needs to be jurisdiction-specific, so it applies to a specific country, and updated, so it captures the latest amendments of the legislation or the most recent court decisions. Quality information as a functional requirement enables the reliance of users on the system's outputs by minimizing the risk of misinformation and subsequent undesirable effects, such as mistakes in the defence strategy due to ill-informed decisions.⁵⁴

51 High-Level Expert Group on AI (European Commission), 'Ethics Guidelines for Trustworthy AI' (n 14) 12.

52 See, for example, Christian Matt and others, 'Towards a multicentric quality framework for legal information portals: An application to the DACH region' (2023) 40(4) Government Information Quarterly 1, 4–5 <<https://www.sciencedirect.com/science/article/pii/S0740624X23000400>> accessed 16 March 2025.

53 T Saracevic, 'Relevance Reconsidered' in *Information Science: Integration in Perspectives, Proceedings of the Second Conference on Conceptions of Library and Information Science, Copenhagen (Denmark)* (1996) 12 <<https://www.bibsonomy.org/bibtex/2c3f30a1af84eb099284f266679ff41f2/aschmidt>> accessed 16 March 2025.

54 See, for example, Sara Merken, 'New York Lawyers Sanctioned for Using Fake ChatGPT Cases in Legal Brief' *Reuters* (2023) <<https://www.reuters.com/legal/new-york-lawyer-s-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/>> accessed 03 October 2024.

Reliance on the retrieved information is further reinforced through *citations*, i.e. references to the underlying sources, so citizens can verify the accuracy of the information and make informed choices. In the case of legal aid applications, citations must mainly consist of legislation and case law but can also refer to commentaries of the law (ideally laid down in plain language), so citizens can interpret the law and apply it to their case. Links to the online versions of the cited sources should further be provided for easier navigation.

If the retrieved information is of good quality and accompanied by references, then users can rely on it to take the appropriate legal action for the resolution of their legal issue(s). This is an important feature of a human-centric design because it aims to empower SRLs to seek a solution to their legal issue by not only collecting the necessary legal information and resources but ultimately applying them in practice when initiating court proceedings.⁵⁵ *Actionability* refers to the ability of the LLM-based chatbot to return information that not only enables users to identify if their issue is justiciable, but also to initiate court proceedings if necessary. For example, the retrieval of court forms' templates or the generation of pre-filled versions according to the individual situation could enable users to start the application process for court proceedings of their own accord.

According to the principle of prevention of harm, AI systems must be designed, developed, and deployed in a way that protects users from any kind of harm, physical or otherwise.⁵⁶ Fewer to no harms mean more equal access to the benefits of AI systems, in accordance with the principle of fairness, given that some types of harms may affect digitally illiterate people to a greater extent. One such harm might be breaches of personal data, which users may voluntarily provide to the system without knowledge of the consequences on their rights to privacy and data protection. Breaches of personal data can occur due to their unlawful processing by the LLM or due to the compromise of the model's security, for example through unauthorized access. Another risk may be the production of unfair biases due to the system's training on historical data or due to poor governance models, leading to the discrimination of societal groups, some of which may be vulnerable. Some of these harms may be only addressed by developers, as is the case with the elimination of unfair biases, through the controlled

⁵⁵ Hagan (n 49) 10.

⁵⁶ High-Level Expert Group on AI (European Commission), 'Ethics Guidelines for Trustworthy AI' (n 14) 12.

training and constant monitoring of the system. In any case, LLM-based chatbots must inform users through *disclaimers* on the risk of exposure to these harms and offer them the option to consent to the processing of their data under strictly defined principles, such as purpose and storage limitation.⁵⁷ The importance of information provided to users on these types of harm is also highlighted under the principle of explicability.

According to the principle of explicability, individuals must have access to transparent AI systems, whose capabilities and limitations are ‘openly communicated’ to the public.⁵⁸ The transparency of AI systems enhances trust in their functioning, so that users maintain their autonomy by avoiding being misinformed or otherwise manipulated by the system and being able to make informed choices, including challenging the systems’ outputs if they are inaccurate or otherwise problematic. The display of disclaimers on the functioning of the AI system ensures that users have the chance to become aware of the capabilities and limitations of the system and their rights in case of erroneous outputs that have consequences on their livelihood. In the case of LLM-based chatbots, the system must be transparent by displaying disclaimers on the machine-generated nature of the retrieved information, which should be independently verified rather than accepted at face value. Disclaimers can also alert users as to the processing of their (personal) data or as to the probability of inaccurate or irrelevant outputs that might misinform users. In the case of SRLs, disclaimers may encourage users to consult a lawyer to verify the accuracy and relevance of the retrieved information for the case and thus avoid any adverse consequences when following unreliable generated information.

4. Chatbots’ Performance on Landlord-Tenant Disputes-Related Questions

LLM-based chatbots are widely used for the generation of responses to users’ prompts, including questions on their legal situation, rights, obligations, and relevant procedures when they are facing a legal and possibly justiciable problem. However, LLMs may produce hallucinations when generating answers that lead to users’ misinformation and, subsequently, to

57 See, for example, European Parliament and the Council of the EU, ‘Regulation (EU) 2016/679 (General Data Protection Regulation)’ (2016) OJ L 119 article 5 <<https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>> accessed 16 March 2025.

58 High-Level Expert Group on AI (European Commission), ‘Ethics Guidelines for Trustworthy AI’ (n 14) 13.

ill-informed choices affecting the self-representation of their case before the court. This scenario is contrary to a human-centric approach, as outlined in Section 3, which guarantees that all users have equal access (principle of fairness) to high quality and referenced information (principle of respect for human autonomy) and to disclaimers (principle of explicability), both of which allow them to make informed choices and contest any erroneous output by the AI system. The ultimate goal is to empower SRLs to defend their rights before courts on the basis of this information without risking any kind of harm (principle of prevention of harm).

The table below presents the performance scores of the examined LLM-based chatbots, rated on a scale from 1 (low) to 5 (high).

	Readability	User - friendliness	Quality	Actionability	Citations	Disclaimers	OVERALL
ChatGPT	4	3	3	3	1	3	2,8
ChatGPT Pro	4	3	3	3	1	4	3
ChatGPT FLA	4	3	3	3	1	4	3
Copilot	4	3	3	3	3	4	3,3
Copilot Student	4	4	3	3	3	3	3,3
Gemini	4	4	2	4	1	5	3,3
Gemini Advanced	4	5	3	4	1	5	3,5

Figure 1 – Overview of Chatbots’ Scores

Gemini Advanced received the best overall score, although with marginal difference from the other models. The following sections contain the main findings of the evaluation of the LLM-based chatbots based on their responses to the 16 questions on landlord-tenant disputes under Greek law (Annex I), assessed against the human-centric functional requirements for legal aid applications (Section 3), namely readability and user-friendliness (principle of fairness), quality and actionability of cited information (principle of respect for human autonomy), and the existence of disclaimers (principles of explicability and prevention of harm).

4.1 Readability

All models received a good readability score, since they displayed a good command of the Greek language (in terms of grammar, syntax, and vo-

cabulary) and a good understanding of the legal terminology relevant to evictions. The text was easy to follow due to its coherence and the relatively plain language used to explain the legal terms. Most models also provided summaries and/or conclusions to their responses, which helps readers better discern the key points of a lengthy response. One limitation of most models was the absence of visualizations of the retrieved information, which could potentially help lay people understand it better. The only model that provided a simple visualization in the form of a table (comparing the order against the action of restitution of rent) was ChatGPT Free Legal Advice:

Σύγκριση:

Διαταγή Απόδοσης Μισθίου	Αγωγή Απόδοσης Μισθίου
Γρήγορη διαδικασία χωρίς ακρόαση	Πλήρης δικαστική διαδικασία με ακρόαση
Ισχύει μόνο για μη πληρωμή μισθωμάτων	Ισχύει για μη πληρωμή και άλλες παραβάσεις της μίσθωσης
Ο ενοικιαστής μπορεί να υποβάλει ένσταση εντός 15 ημερών	Ο ενοικιαστής μπορεί να παραστεί και να υπερασπιστεί
Μικρότερο κόστος λόγω έλλειψης ακρόασης	Αυξημένο κόστος λόγω ακροαματικής διαδικασίας
Ταχύτερη διαδικασία	Χρονοβόρα διαδικασία

Figure 2 – Table Generated by ChatGPT Free Legal Advice

4.2 User-friendliness

All LLM-based chatbots displayed certain user-friendly features. The chatbots interacted in a dynamic way with the user through a conversation format. Users could pose prompts in natural language, instead of using keywords and Boolean operators as in a search engine, and attach files to allow the system to retrieve information from or in relation to a specific document.⁵⁹ Personalization was possible through prompt-engineering, where users can add more context to their prompts to receive information that is relevant to their particular situation. Additionally, the examined chatbots incorporated common accessibility features, such as the option

59 Although most LLM-based chatbots allowed only the attachment of images, not documents.

to listen to responses instead of reading them. All chatbots enabled users to copy their responses in the clipboard, thus rendering them re-usable. Finally, all chatbots were available as mobile applications, enabling users to access them directly on their mobile devices.

Certain models performed slightly better than others by offering more advanced features. Gemini Advanced was the most user-friendly because of the possibility to customize the user experience through the re-generation of a response. Gemini Advanced could be adjusted in its length, formulation, and tone:

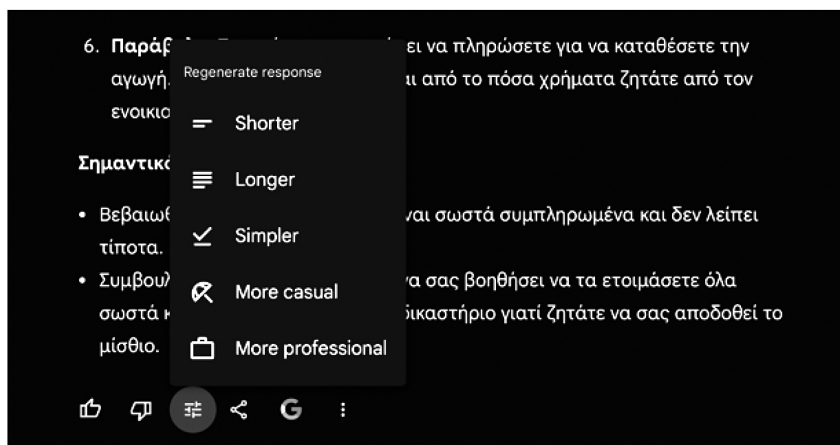


Figure 3 – Re-generation of Responses with Gemini Advanced

This feature could be highly beneficial to users with no legal expertise, who would most likely prefer a shorter and simpler response to their legal question, at least in terms of plain language. The feature could also be seen as inclusive of cultures where the writing style is more laconic and concise.⁶⁰ In practice, though, the 'simpler' feature did not lead to a significantly shorter and simpler response.

Users of Gemini and Gemini Advanced, also have access to more customizable and comprehensive features covering data protection. In the case of the two Gemini models, the user can visit the 'Privacy Policy' of Google, get informed on how Google processes users' data, and adjust the

⁶⁰ See, for example, Telang (n 45).

privacy settings.⁶¹ This experience is relatively user-friendly, since the user can watch videos instead of reading long texts of privacy policies, but still the information is quite extensive and can discourage users from making informed choices on the processing of their (personal) data.

4.3 Quality of Information

All LLM-based chatbots performed moderately well in terms of the quality of the information retrieved. Their responses were relevant to the posed prompts, jurisdiction specific, and consistent. In particular, all LLM-based chatbots provided responses that aligned with the researcher's information need for legal information on the topic of evictions under the Greek law. In terms of jurisdiction-specificity, all chatbots recognized that they had to provide information in the context of Greek law, probably because each question clearly stated that the response should be applicable to the Greek jurisdiction and was posed in the Greek language (see Annex I). The chatbots' responses were generally consistent throughout the conversation, providing the same or logically equivalent answers to questions that requested the same information.⁶² Only Gemini lacked consistency, as it listed different sets of required documents for the application for an order or for an action of restitution of rent as a response to questions asking for the same kind of information.

On the contrary, the LLM-based chatbots did not perform well in terms of the accuracy, timeliness, and completeness of the information retrieved. Regarding accuracy, all chatbots (except Gemini Advanced) frequently misidentified or incorrectly referred to the threshold of EUR 20,000 (instead of EUR 600) below which the Small Claims Court (Ειρηνοδικείο) is competent for the dispute and above which the First Instance Civil Court (Πρωτοδικείο) is the competent one.⁶³ Regarding timeliness, ChatGPT Free Legal Advice referred to an article of the Greek Code of Civil Procedure that was replaced after the most recent amendment by a new

61 Google Privacy and Terms, 'Google Privacy Policy' <<https://policies.google.com/privacy>> accessed 07 March 2025.

62 For example, the question 'What is the procedure for an action for the restitution of rent under Greek jurisdiction?' received the same answer as the question 'To which court should I file an action for the restitution of rent under Greek jurisdiction?' in terms of what the competent court is.

63 Greek Parliament, 'Code of Civil Procedure (Κώδικας Πολιτικής Δικονομίας)' (n 28), article 14 (1) (b) and 2.

article which was irrelevant to the posed question, signifying the difficulty of LLMs to provide updated information.

All LLM-based chatbots further lacked completeness since they did not attach enough details to their responses, even though these were lengthy enough. For example, when the chatbots were asked which documents should be submitted along with the application for an order or for an action of restitution of rent, they did not mention all required documents but missed important ones, such as the written formal notice and the proof of identification. Other missing but important details concerned the applicable duration and costs of proceedings, although this kind of information could only be given by approximation since these details depend on the specific case.

4.4 Actionable Information

Actionable information empowers users to take legal action by performing certain procedural steps, such as drafting an application for court proceedings, or even by reaching out to legal representatives if they deem it necessary. All LLM-based chatbots provided details on the two distinct procedures of the order and the action of restitution of rent to be followed before the competent courts. Of course, this information was shown earlier to be inaccurate or incomplete at times, but it was nevertheless provided by the chatbots. Only two chatbots, ChatGPT Free Legal Advice and Gemini Advanced, provided links to online directories of lawyers in response to the question ‘Where can I find a lawyer in Greece to represent me in legal actions against the tenant?’ ChatGPT Free Legal Advice outperformed all models by providing links to the four (largest) Bar Associations of Greece, thus accommodating the needs of a significant part of the Greek population. However, the links did not lead directly to the directories, but redirected users to the homepages of the Bar Associations so that they had to either navigate through the websites to find the directories (if existing) or contact the Bar Associations through the contact details provided.

Gemini Advanced was the only chatbot to provide the link to the Athen’s Bar Association’s homepage and recommend related experts on evictions (e.g. real estate consultants). The usefulness of this information may be relative but is certainly an improvement in performance compared to the other LLM-based chatbots. Gemini Advanced also included sections called ‘Advice’ (‘Συμβουλή’) to recommend certain actions, for example to remind

users to first collect all the necessary documents before proceeding to the application for an order or for an action of restitution of rent or on how to communicate their legal issues to a lawyer.

4.5 Citations

Most LLM-based chatbots performed very poorly in citing sources to support their responses. In most cases, there were no references at all on law or secondary sources. It was rare for a chatbot to refer to legislative documents; only Copilot and Copilot for students and enterprises cited their responses:

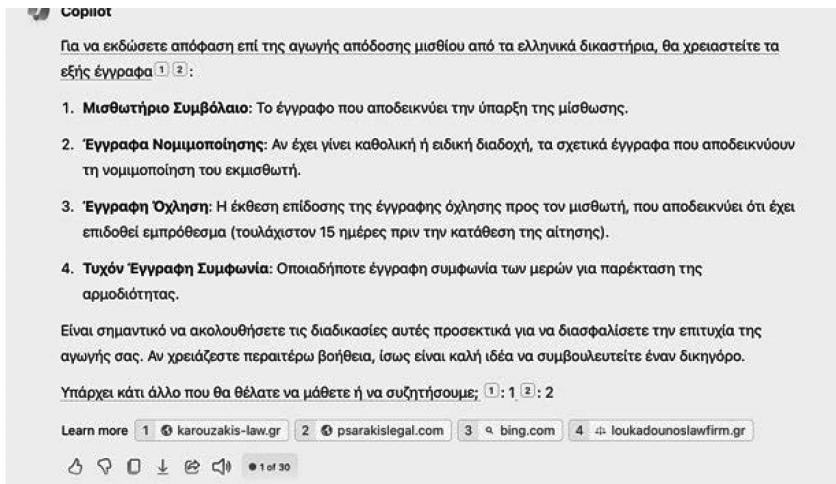


Figure 4 – Citations as Captured by Copilot

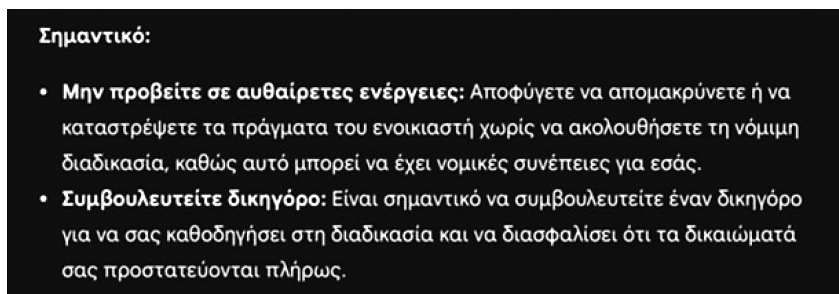
The referenced sources are not official sites of the Greek government or of Greek courts containing legal information in the form of legislation, case law, or abbreviated versions of the two. Instead, they are law firms' websites with featured blog posts on the topic of evictions. This may be useful and reliable to an extent because the blog posts were written by lawyers who have expert knowledge and experience on this legal issue, thus being in a place to provide legal information and advice. Redirecting individuals with no legal expertise to official websites containing only the text of the law would probably not be useful, since they would not be able to interpret

the legal text without the help of a lawyer or at least of a commentary. Therefore, the ideal solution would be to refer both to official websites containing legislation and case law and to secondary sources explaining or commenting on the law, to allow for an effective verification of the retrieved information and for its re-usability by SRLs to understand and defend their legal situation before courts.

4.6 Disclaimers

Most LLM-based chatbots displayed disclaimers, covering the following topics: (i) the need to further seek advice by a legal expert; (ii) the processing of users' (personal) data, and; (iii) the possibility that the model may err in its responses.

Regarding legal advice, all chatbots displayed disclaimers on the importance of consulting a legal expert, although these reminders did not always appear after each response. The only chatbots that had consistent disclaimers on the need for legal consultation were ChatGPT Pro, Copilot, and Gemini Advanced. This type of disclaimer is important because it reminds users of the legal nature of their problem and the importance of having the retrieved information reaffirmed by a legal expert to safely base their legal claims on it. The following excerpt comes from Gemini Advanced, is signaled as 'Important' ('Σημαντικό'), and states that users should 'not take arbitrary actions' ('μην προβείτε σε αυθαίρετες ενέργειες') and 'consult a lawyer' ('συμβουλευτείτε δικηγόρο') [translated from Greek by DeepL]:



Σημαντικό:

- **Μην προβείτε σε αυθαίρετες ενέργειες:** Αποφύγετε να απομακρύνετε ή να καταστρέψετε τα πράγματα του ενοικιαστή χωρίς να ακολουθήσετε τη νόμιμη διαδικασία, καθώς αυτό μπορεί να έχει νομικές συνέπειες για εσάς.
- **Συμβουλευτείτε δικηγόρο:** Είναι σημαντικό να συμβουλευτείτε έναν δικηγόρο για να σας καθοδηγήσει στη διαδικασία και να διασφαλίσει ότι τα δικαιώματά σας προστατεύονται πλήρως.

Figure 5 – Disclaimer by Gemini Advance to Seek Legal Advice

Nevertheless, the disclaimer is formulated as a short text and placed at the very end of the response, where it can easily be ignored. Moreover, its importance is undermined if it is not accompanied by links to online directories of lawyers.

Regarding data processing, not all LLM-based chatbots displayed privacy disclaimers (which coincides with the absence of advanced privacy settings by most models), with the exceptions of Gemini and Gemini Advanced. In both cases, the following messages came up:

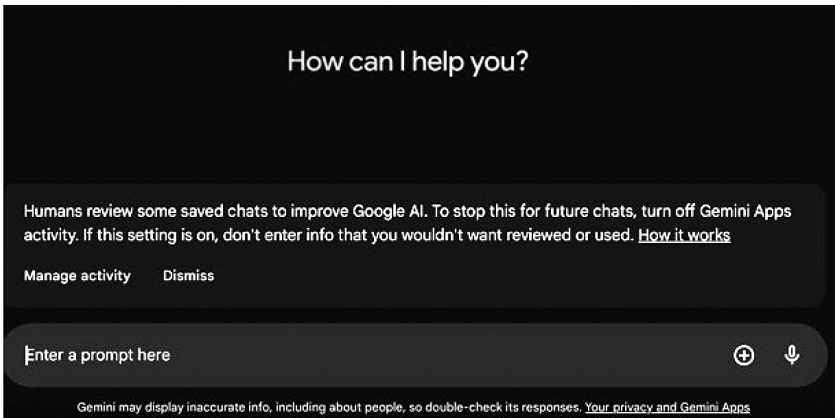


Figure 6 – Disclaimer by Gemini and Gemini Advanced on Privacy

The chatbots signified from their very first interaction with the user that the (personal) data contained in the prompts may be used in the training of Google's AI products to improve their performance. Then, they provide the option to turn off the Gemini Apps activity, attaching links to resources on data processing and the management of users' activity. However, users could easily ignore the feature or not interact with the pop-up message, which would negate any positive impact the disclaimer might have on their awareness and control over the processing of their data. Moreover, the 'how it works' feature leads to another pop-up message featuring a 'learn more' link that redirects the user to a relatively long text charged with even more links, rendering the user experience non-intuitive. Finally, users must actively choose to *not* have their data processed for the training of the AI models, while they should instead be able to opt-*in* to their data being processed by Google for such purposes.

Copilot for students and enterprises also has a slight notification in the form of an icon placed in the top right corner of the chat box, which informs users that the enterprise data protection applies to their chat and includes a hyperlink that redirects the user to a page with a lengthy explanation of what this type of data protection entails. The data provided by users through their prompts are in principle encrypted and not saved in Microsoft's servers or used to train Microsoft's AI models.⁶⁴

Regarding the possibility of LLM-based chatbots making mistakes, all chatbots displayed relevant disclaimers. These disclaimers should aim at raising awareness of the risk of hallucinations and the need for verification of the responses by users. Nevertheless, its presence might also be an attempt to shift the responsibility for the chatbots' mistakes from the providers and deployers of these systems to the end-user. This shift can be detrimental for users, since they might over-rely on the chatbots' responses due to their persuasive tone⁶⁵ and compromise their legal position before courts if the retrieved information is inaccurate. Therefore, the advice of a lawyer might be necessary to confirm the validity of the retrieved information and apply it on the given legal issue.

Interestingly enough, Gemini Advanced enabled the 'double-checking' of a given response and the understanding of this process's results as depicted below:

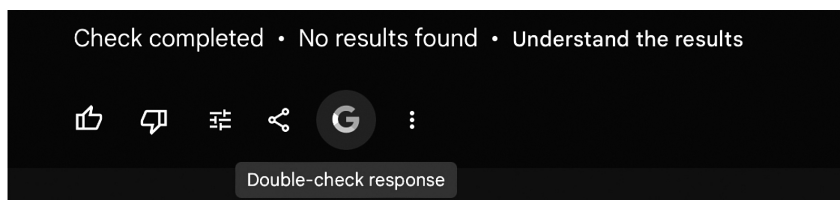


Figure 7 – Double-checking of Response with Gemini Advanced

It must be noted that Google (the provider of Gemini) operates one of the most extensive search engines, indexing a vast range of publicly available online resources. Therefore, it could be assumed that Gemini verifies the accuracy of its responses against all this indexed information to increase their reliability. Nevertheless, the mechanics behind this process remain

64 Microsoft Learn, 'Privacy and protections' <<https://learn.microsoft.com/en-us/copilot/privacy-and-protections#organizational-data>> accessed 15 March 2025.

65 Ruschemeier and Hondrich (n 41).

opaque (in contrast to the principle of explicability), so that the user cannot be certain of how effective it is in practice.

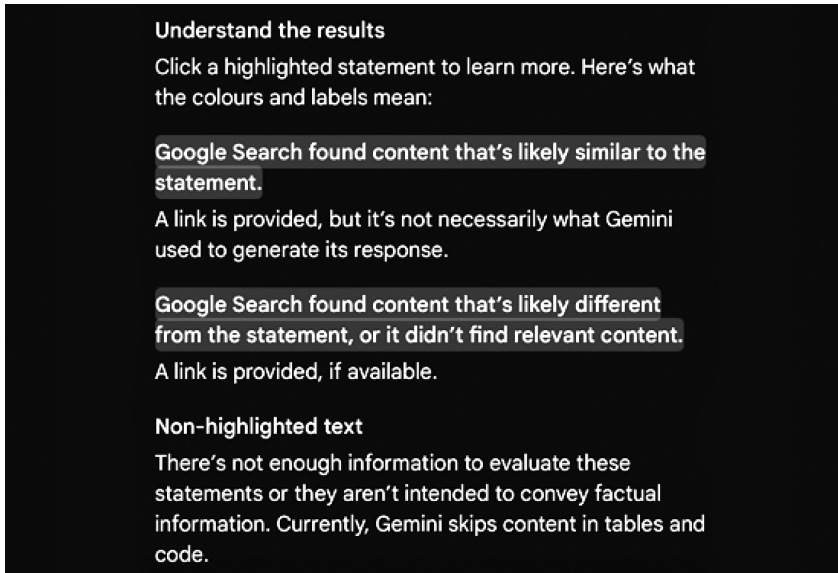


Figure 8 – Understanding the Results with Gemini Advanced

The 'understand the results' feature may seem like explaining the verification process of the LLM, rendering it more transparent and reliable. In theory, the process of verifying the responses ensures a better human oversight of the model's performance by encouraging users to engage with the content of the retrieved information and determine its relevance for their particular case. In this case, however, the verification process is not performed by users but by Gemini itself, thus depriving them of their agency and control over the process. Furthermore, no references are attached to the responses to allow their verification by users themselves. Consequently, this feature may just be an effort by Google to market Gemini Advanced as a reliable LLM-based chatbot, so it can compete with other popular chatbots, such as ChatGPT and Copilot.

5. LLM-based Chatbots for Court Services

The evaluation of the LLM-based chatbots revealed that they partially abide by a human-centric design, respecting the principle of fairness by offering a user-friendly interface and readable information and the principle of explicability through disclaimers on the functioning of the models. However, the principle of prevention of harm may not always be respected, especially when the chatbots collect personal data for training purposes (as was the case with Gemini and Gemini Advanced), regardless of the existence of disclaimers in that regard. Likewise, the principle of respect for human autonomy was compromised whenever the models generated false or inconsistent responses that were not accompanied by citations. These limitations may compromise access to justice, since users cease to have access to reliable and actionable information and may even jeopardize the defence of their rights before courts when relying on false or irrelevant information.

To minimize the risk of LLM-based chatbots providing unreliable information to SRLs, courts could develop their own chatbots to offer court services to the public, by fine-tuning existing LLMs. Fine-tuning aims at adapting the LLM to perform specific tasks by (partially) retraining the model on a task-specific or domain-specific dataset.⁶⁶ In this case, courts could re-train existing models on the legislation and case law of a specific jurisdiction to ensure that the chatbots would provide relevant information to the citizens of a specific country. Citizens could prompt the chatbot to retrieve court forms, published legislation and case law, updates on the progress of pending case, and information on court proceedings. The chatbot could also provide access to recorded webinars analysing common legal issues and their solution before courts,⁶⁷ or live trials for information and

66 See, for example, IBM, 'What is fine-tuning?' <<https://www.ibm.com/think/topics/fine-tuning>> accessed 17 March 2025.

67 These types of webinars have taken place in the People's Republic of China through the Court Mobile TV app that hosts a 'Judge Talks' series, i.e. an online classroom where judges educate the general public on legal issues in a discussion mode; see, Tania Sourdin, Jacqueline Meredith and Bin Li, *Digital Technology and Justice: The Use of Justice Apps* (Routledge 2020) 25–28 <<https://www.routledge.com/Digital-Technology-and-Justice-Justice-Apps/Sourdin-Meredith-Li/p/book/9780367650186?srsltid=AfmBOoqvcbs4YPXvJnQFWTVmKnDa80rszLgPTuXcA-YgLAYQVojn3DtI>> accessed 16 March 2025.

transparency purposes (principle of the publicity of court proceedings⁶⁸). The chatbot could further act as a virtual clerk to inform the user whether his or her issue is justiciable and to provide links or embedded features to online directories of lawyers. Offering court services through LLM-based chatbots to SRLs and the wider public can promote access to justice by eliminating or reducing expenses spent in legal services, the time spent in overburdened courthouses, and the effort of navigating through complex court websites. Legal information is ultimately democratized and citizens are empowered to solve their legal issues with this newly acquired knowledge.

Nevertheless, there are certain constraints regarding the human-centric design of the chatbots to be addressed before courts can release them on their websites. Regarding the principle of respect of human autonomy, the technique of fine-tuning is not always effective in improving the accuracy and relevance of the model by mitigating its hallucinations and thus avoiding incidents of misinformation, while it can deteriorate the model's performance.⁶⁹ Emerging techniques like Retrieval-Augmented Generation (RAG) aim to reduce hallucinations by having the model retrieve and reference external information sources (for example, relevant documents retrieved from a specified database) before generating a response.⁷⁰ While this is a step towards improving the reliability of the chatbot's answers, the range and quality of the database depends on the capacities of the responsible deployer. Courts often lack machine-readable case law, making it challenging to develop or fine-tune an LLM using court-related documents. Limited funding provided by centralized authorities, such as Ministries of Justice, may ultimately impede any effort towards the provision of LLM-based court services to the public.

Regarding the principles of fairness and explicability, all users must be treated justly and equally by having access to LLM-based chatbots that are jurisdiction-specific and language-specific. Most LLMs acting as the foundation of related applications (foundation models) are developed by U.S.

68 See, for example, Council of Europe, 'European Convention of Human Rights' (1950) article 6(1) <https://www.echr.coe.int/documents/d/echr/convention_eng> accessed 07 July 2025.

69 See, for example, OpenAI, 'GPT-4 Technical Report' (2023) 68 <<https://arxiv.org/abs/2303.08774>> accessed 07 July 2025.

70 See, for example, IBM Technology, 'What is Retrieval-Augmented Generation (RAG)?' <<https://www.youtube.com/watch?v=T-D1OfcDW1M&t=322s>> accessed 03 October 2024.

companies, meaning that they can perform better in the English language since they have been primarily trained on English language resources, being the most represented language on the web. Languages with fewer available online resources may result in limited training data for the model, increasing the risk of misrepresenting legal terms in those languages. Therefore, companies should diversify their training data in terms of language and also offer their products in multiple languages, which can then more easily adapt to resources unique to a specific jurisdiction and thus safely integrated in courts' websites.

Further to the goal of fairness in the design of LLM-based chatbots for court services, the choice of courts to integrate these systems in their websites in an effort to become more efficient should not exclude digitally illiterate people or people with no or limited access to a (stable) Internet connection and/or a digital device. Such an exclusion would reinforce the feeling of not being heard among affected parties and would deny them access to justice since they would not be able to receive effective legal protection by accessing court services. The interface of chatbots must ensure equal accessibility by accommodating varying levels of legal and digital literacy through plain language, multilingual support, and accessible design (which was generally observed in the examined LLM-based chatbots). Ultimately, courts should always make available an alternative option to communicate with a court clerk for these categories of interested individuals.

Additionally, LLM-based chatbots must not place SRLs in a disadvantageous position by overestimating their level of legal knowledge and further reinforcing the feeling that the retrieved information can substitute legal consultation (principle of fairness). Disclaimers to resort to a legal expert (principle of explicability) might be sufficient reminders (if they appear consistently throughout the conversation), but they can easily be ignored by users. Therefore, courts must attach in the chat itself a direct link to a directory of lawyers available in all regions of a given State to encourage users to have generated responses verified by a legal expert. This point further relates to the greater discussion of offering full or sufficient legal aid (depending on the financial situation of the eligible individuals), which is not always a given among EU Member States, as seen in Section 1.

The evaluation of the LLM-based chatbots also showed that there is a general lack of consistent citations (principle of respect for human autonomy), which is detrimental in the effort of users to verify the accuracy of the information in order to rely on it and take legal action. Consequently,

the designing of chatbots intended to be integrated in court websites to provide court services must ensure that reliable and relevant citations to authoritative resources (containing the relevant law and its commentary) are *always* provided as an attachment to the chatbots' responses. This improvement enhances trust in courts and empowers users to take legal action without being or feeling manipulated by the system as such. An additional publication by courts of their systems' training data, design choices, and related assessments would enhance transparency and accountability (which is also prescribed in the AI Act for high-risk AI systems under article 11).

Regarding the principle of prevention of harm, more emphasis should be given on advanced privacy and cybersecurity settings, given the intense processing of users' data through prompts, the log-in procedure (if required), or the collection of metadata. Primarily, the providers of LLM-based chatbots should not re-purpose the processing of the (personal) data inserted by users in their prompts, so that instead of processing them just to retrieve relevant information they further use them to train and improve the model's performance.⁷¹ They should also make sure that their processing is restricted to the minimum data necessary, so users are not treated as products by providing an unlimited amount of (personal) data to optimize the performance of the chatbots.⁷² To limit unauthorized access or other types of malicious interference with the AI system, there are a number of technical robustness measures (such as encryption, access controls, and regular audits) that can be applied, as well as the solution for courts to host their chatbots in local servers and not in the cloud, with which a greater number of malicious actors from Europe and beyond could more easily interfere.⁷³ Anonymization or pseudonymization measures to prevent or minimize re-identification risks should also be implemented, especially when SRLs' or litigants' data are processed (for example, to fill a court form or receive personalized information on court proceedings),⁷⁴ In any case,

71 See, for example, GDPR, article 5(1)(b).

72 *ibid* article 5(1)(c).

73 See, for example, Kalliopi Terzidou, 'Generative AI Systems in Legal Practice Offering Quality Legal Services While Upholding Legal Ethics' (n 11), 4.

74 Kalliopi Terzidou, 'Automated Anonymization of Court Decisions: Facilitating the Publication of Court Decisions through Algorithmic Systems' in *Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law (ICAIL '23)* (ACM 2023) 297, 300–301 <<https://doi.org/10.1145/3594536.3595151>>.

(personal) data of users should not be stored in the courts' databases for longer than necessary to provide the requested court service.⁷⁵

Regarding the risk-based qualification of LLM-based chatbots integrated in court websites, these are not automatically qualified as 'high-risk' AI systems under the AI Act, since they are not intended to assist the judiciary 'in researching and interpreting facts and the law and in applying the law to a concrete set of facts' (Annex III, point 8 (a)). They are more likely to fall under the 'low risk' scale (Recital 61), given that they contribute to the efficiency of back-office procedures by alleviating the court staff from addressing multiple simultaneous requests on court proceedings, in the background of often understaffed and overburdened courts. However, this classification ignores the potential harm LLMs can cause when providing false, irrelevant, or otherwise unreliable information on which users rely to access courts and defend their case before the judges (Recital 52). An interpretation could go as far as implying that this use of an AI system should be prohibited if it entails subliminal techniques (such as the persuasive, conversational tone of the chatbot) with the effect of impairing the person from making an informed decision (by furnishing him or her with inaccurate information that cannot be verified due to the absence of citations), causing or likely to cause significant harm (for example, missing a deadline and thus risking the admissibility of his or her case) (article 5, para. 1, point a).⁷⁶

A human-centric approach in the designing, development, deployment, and use of AI systems, as analyzed in the beginning of this Section, would likely strike a balance between making LLM-based chatbots available to the public due to their user-friendly way to communicate a vast amount of information on a 24/7 basis (thus not outrightly prohibiting their use) and ensuring their utmost robustness and safety (thus applying the requirements of high-risk or systemic risk AI systems), so that SRLs and the wider

75 GDPR, article 5(1)(e).

76 General-Purpose AI systems, including Chat-GPT and any other system that has a general-purpose AI model as a component, are also not classified automatically as 'high-risk.' It could be argued that LLM-based chatbots pose a systemic risk, i.e. 'any actual or reasonably foreseeable negative effects...on democratic processes...[and] the dissemination of illegal, false, or discriminatory content' (Recital 110), given that they can potentially misinform a large number of users due to hallucinations. However, if such systems are fine-tuned for court services, then they can no longer be considered of a General-Purpose, since they are specifically adapted for court-related tasks.

public can be reliably informed on the law and procedure applicable during court proceedings.

6. Conclusion

The evaluation of the performance of LLM-based chatbots against the human-centric functional requirements has yielded an answer to the research question ‘To what extent are Large Language Models-based chatbots designed in a human centric way to support SRLs in accessing legal information?’ In short, the examined chatbots abide by certain requirements, although they are not specifically designed to support SRLs in dealing with legal issues.

Even if Gemini Advanced seemed to perform the best and ChatGPT the worst, in essence all chatbots showed similar strengths and weaknesses. On the one hand, they returned readable responses, i.e. responses correctly expressed in the Greek language, easy to understand, and accompanied by summaries and conclusions to highlight the key takeaway of the response (principle of fairness). Their disclaimers (whether on consulting a lawyer, on privacy settings, or on the possibility of the model committing mistakes) were also prominent and repeated throughout the chat (principle of prevention of harm).

On the other hand, LLM-based chatbots often did not feature any citations and when these were attached, they did not refer to official websites containing legislation, case law, or related information (principle of respect for human autonomy). Instead, they led to websites of law firms and, more specifically, to blog posts written by lawyers on evictions. It could be argued that this is more useful for users that would not be able to interpret the original text of the law on their own. However, the reliability of these secondary sources is doubtful. Furthermore, inaccuracies and incompleteness were observed in the chatbots’ responses, mainly regarding the competent courts and the required documents to apply for an order or an action of restitution of rent.

Steps towards a more human-centric approach in the designing of LLM-based chatbots intended for court websites would give an emphasis on more accurate and relevant outputs, that are devoid of misinformation due to the phenomenon of hallucinations, which could be possible through the implementation of techniques, such as fine-tuning and RAG (principle of respect for human autonomy). The design of the chatbots should also

strive to protect vulnerable users by providing direct links to directories of lawyers for further legal advice (legally illiterate users) and allow for non-digital alternatives to court services through access to a court clerk (digitally illiterate users) (principle of fairness). Disclaimers on the need to consult a lawyer and citations below *every* response enhance trust in the system and encourage users to rely on its outputs (principle of explicability).

Data breaches should also be prevented by implementing technical measures, such as encryption and pseudonymization, that prevent unauthorized access to LLMs and protect the personal data of their users. Ultimately, the providers and deployers of these systems should be obliged to guarantee their technical robustness and safety to the utmost extent, given existing risks including misinformation (due to hallucinations) and manipulation (due to automation bias). This would be possible if LLM-based chatbots integrated by courts were classified as ‘high-risk’ (articles 8–15 of the AI Act), so that the strictest requirements are applied during their designing, development, and deployment.

A final remark regards users themselves. Although the primary responsibility for a human-centric design of LLM-based chatbots for SRLs should be on providers and deployers of these systems, users should also be aware of their level of legal and digital literacy and of the risks that chatbots entail, so as not to create false expectations and be manipulated by the systems. Regarding legal literacy, better legal education is needed to empower citizens to detect and solve their legal issues. Citizens must be educated on their rights and duties, the basic rules that apply in everyday legal matters (such as sales), and the organization and functioning of the national justice system. These matters may be part of formal education (especially in secondary school), while legal aid mechanisms should be available to people with low income, so they can receive legal consultation and representation when necessary, free of charge or for a symbolic fee.

Regarding digital literacy, parents or guardians and educators should instruct children from their very first interaction with the Internet on safe online navigation. The main lessons should include search query formulation (for search engines) and prompt engineering (for LLM-based chatbots), critical evaluation of the retrieved information, and identification of privacy risks and adjustment of privacy settings. The LLM-based chatbots offered by courts should be freely accessible on every digital device with Internet access. If individuals do not have access to a digital device or to a stable Internet connection, access to these means should be provided, for

example, through courts, public libraries, or municipality buildings, so as not to exclude any individual interested in carrying out court proceedings as an SRL. Ultimately, access to a court clerk should be guaranteed for individuals who are digitally excluded or lack the skills to interact with online systems.

Annex I

The set of 16 questions posed to the examined LLMs include the following [translated from Greek through DeepL]:

I am a landlord and my tenant does not pay the rent on time:

1. What are my legal rights against the tenant?
2. What are my legal rights against the tenant under Greek law?
3. To which court should I apply for an order for the restitution of rent under Greek jurisdiction?
4. What is the procedure for issuing an order for the restitution of rent under Greek jurisdiction?
5. To which court should I file an action for the restitution of rent under Greek jurisdiction?
6. What is the procedure for an action for the restitution of rent under Greek jurisdiction?
7. What is the difference between an order for the restitution of rent and an action for the restitution of rent under Greek law?
8. What should I do with the tenant's belongings that are within the tenancy after the issuance of the rent order under Greek law?
9. What should I do with the tenant's belongings that are within the leased property after the conviction of the tenant (following a lawsuit) under Greek law?
10. Where can I find a lawyer in Greece to represent me in legal actions against the tenant?
11. How much does it cost to obtain an order for the restitution of rent in Greece?
12. How much does it cost to bring an action for the restitution of rent in Greece?
13. What is the deadline for the Greek courts to decide on an order for the restitution of rent?

14. What is the deadline for the Greek courts to decide on an action for the restitution of rent?
15. What are the necessary documents for the Greek courts to decide on an order for the restitution of rent?
16. What are the necessary documents for the Greek courts to decide on an action for the restitution of rent?