

ChatGPT as an Artificial Lawyer?

Jinzhe Tan^{1,*}, Hannes Westermann¹ and Karim Benyekhlef¹

¹Cyberjustice Laboratory, Faculté de droit, Université de Montréal, Montréal, Québec, H3T 1J4, Canada

Abstract

Lawyers can analyze and understand specific situations of their clients to provide them with relevant legal information and advice. We qualitatively investigate to which extent ChatGPT (a large language model developed by OpenAI) may be able to carry out some of these tasks, to provide legal information to laypeople. This paper proposes a framework for evaluating the provision of legal information as a process, evaluating not only its accuracy in providing legal information, but also its ability to understand and reason about users' needs. We perform an initial investigation of ChatGPT's ability to provide legal information using several simulated cases. We also compare the performance to that of JusticeBot, a legal information tool based on expert systems. While ChatGPT does not always provide accurate and reliable information, it acts as a powerful and intuitive way to interact with laypeople. This research opens the door to combining the two approaches for flexible and accurate legal information tools.

Keywords

Artificial Intelligence & Law, Large Language Models, ChatGPT, Access to Justice, Legal Information

1. Introduction

ChatGPT¹ has garnered significant attention from the public, academia, industry, and media. It is able to perform a huge variety of textual tasks after simply being asked through a free chat interface. The model can further be accessed using an API (Application Programming Interface)[1], which gives developers the ability to create products enhanced by ChatGPT across many different areas. Every day, dozens of applications using this API are launched.²

While ChatGPT's performance in carrying out natural language conversations is impressive, its potential applications are not limited to that. In addition to general-purpose applications such as search engines,³ it could also be used in domain-specific tasks, such as in the legal domain. ChatGPT can be seen to provide a natural

language interface to many types of knowledge. In the legal domain, an important role of lawyers is to provide information and advice about legal problem to laypeople. However, for many people, the related expenses of hiring a lawyer can pose a significant obstacle to safeguarding and enforcing their rights. They find themselves living in a "legal advice desert"[2], unable to find the means to safeguard their rights. Despite the important work done by scholar and practitioners in increasing access to justice[3], many individuals still face issues with it.[4]

Artificial Intelligence (AI) has the potential to play a significant role in promoting access to justice. AI could serve many parties simultaneously and provide legal information to parties in an efficient, accurate, and cost-effective manner. A number of products in this area have been developed and used with success.[5]

In this research, we sought to explore the performance of ChatGPT in providing legal information, compared to self-help tools that specialize in the legal domain. To this end, we conducted an experiment on ChatGPT, utilizing simulated legal cases to qualitatively evaluate its performance. We compared the responses to the JusticeBot, a legal decision support tool focused on landlord-tenant disputes.

Our investigation approached the provision of legal information as a *multifaceted process*, comprising not only the acquisition of precise and reliable sources of information but also the ability to elicit an accurate description of users' legal predicaments and needs, the aptitude to discern relevant and extraneous facts from the users' input, and the skill to furnish pertinent legal information relevant to the users' situation via legal reasoning. Our research thus underscores the vital role played by each of these components in ensuring the effective delivery of legal information.

In this paper, we will examine ChatGPT's ability to

Workshop on Artificial Intelligence for Access to Justice (AI4AJ 2023), June 19, 2023, Braga, Portugal.

*Corresponding author.

✉ jinzhe.tan@umontreal.ca (J. Tan);

hannes.westermann@umontreal.ca (H. Westermann);

karim.benyekhlef@umontreal.ca (K. Benyekhlef)

📞 0000-0002-8259-2630 (J. Tan); 0000-0002-4527-7316

(H. Westermann); 0000-0001-9390-556X (K. Benyekhlef)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

📄 CEUR Workshop Proceedings (CEUR-WS.org)

¹ChatGPT is a Large Language Model (LLM) application developed by OpenAI.

²For example, ChatPDF (<https://chatpdf.com>) allows you to request information in PDF documents, ChatExcel (<https://chatexcel.com>) allows you to use natural language to adjust Excel files, bilingual book maker (https://github.com/yihong0618/bilingual_book_maker) allows you to translate an entire book in a short amount of time, Mem (<https://mem.ai>) allows you to quickly brief yourself on relevant material, and AI Dungeon (<https://play.aidungeon.io>) allows you to play infinite episodes of games generated by AI.

³The new Bing (<https://www.bing.com/>) integrated with the ChatGPT model is now available.

provide legal information, understand the strengths of ChatGPT, and look at opportunities for integration of ChatGPT and other legal tools in the future.

2. Related Work

Using AI and technology to increase access to justice by giving legal information has been an important field of investigation. For example, Branting et al built a tool to inform individuals of their rights regarding protection orders [6]. Zeleznikow's GetAid system aims to help lawyers determine whether an individual is eligible for legal aid [7]. The Loge-expert system aimed to help people understand their housing law situation using an expert system [8]. Housing law questions were also investigated using machine learning in [9]. In [5], the authors suggest the "JusticeBot" approach to build legal decision support systems for laypeople.

Here, we examine whether ChatGPT could be used to provide legal information to increase access to justice. This would have the advantage of not requiring the manual adaptation to new legal areas. We compare the answers given by ChatGPT to answers given by a JusticeBot tool focused on landlord-tenant disputes [5].

Experiments on ChatGPT's performance in the legal field have been conducted extensively, according to studies testing ChatGPT's performance on the U.S. bar Exam.[10] While ChatGPT placed in the bottom 10% of law students, the improve GPT-4 model was estimate to pass the bar and place in the top 10% of students.[11, 12] Another study tested ChatGPT's performance on law school exams, showing that it was able to pass the exam, but could only be considered a "mediocre law student".[13] It was also found that ChatGPT has some ability to write legal documents.[14] Further, GPT-4 was found to display strong performance when given legal annotation tasks to carry out [15, 16] and when explaining legal concepts using case law [17]. In the context of access to justice, the GPT-4 model further showed promise in intervening in an online dispute resolution context, by reformulating inflammatory messages or even autonomously suggesting interventions to mediators [18].

Remarkably, ChatGPT exhibits a certain level of legal reasoning ability despite not having been specifically trained on legal data, owing to the *emergent abilities* derived from the scaling of the model.[19] This suggests that ChatGPT is a general-purpose model, it contains laws about human common sense, a capability that, when combined with specialized domain knowledge,[20] has the potential to solve the challenges encountered in legal AI training in the past. We can enhance its performance in the legal domain through prompting. Yu et.al showed that such *legal prompting*

significantly improves ChatGPT's performance.[21, 22] Chain-of-thought prompting[23] has been shown to be effective in improving performance in the legal field as well.[24]

In addition to testing ChatGPT using exam sets, research has been conducted on comparisons between ChatGPT and human experts, and corpus-based evaluations of ChatGPT performance in the legal field have been conducted.[25]The study found that ChatGPT-generated answers were generally evaluated as more helpful than humans', but it also found that ChatGPT may fabricate facts, especially in the legal field, where ChatGPT may quote non-existent legal texts to answer questions.[25] The study compared the answers generated by ChatGPT to answers taken from Wikipedia and Baidu Baike⁴ Due to the specificity of the legal field, these answers may not always be accurate [26].

In this work, we prepare simulated cases, and interact with ChatGPT as if we were layperson parties of these cases. We compare the answers given by ChatGPT to the answers given by JusticeBot, a legal information tool developed at the Cyberjustice Laboratory, created by legal experts.

3. Access to Justice

Laypeople often have trouble resolving their everyday legal disputes. A majority of people will at some point have to deal with a legal issue, such as neighbourhood, employment or debt problems. [27, 28] However, it can be very tricky to resolve such disputes. Studies conducted have shown that only around 20% of legal problems that arose in the past three years had been resolved [29]. These issues are likely to be especially pronounced for individuals who do not have access to professional legal help, which can be quite expensive. Such individuals may not be aware of which rights apply to them, which forms they need to fill out [30], or what the relevant facts are regarding their case [31].

Self-represented litigants, lacking precise legal information guidance, may navigate through the labyrinth of litigation like a ship without a compass,[32] and may find themselves at a disadvantage due to missing critical information,[33] hindering access to justice and undermining public confidence in the judicial system. The impacts may be especially grave when laypeople encounter a party that is well financed and has previous experience in going through litigation.[34] Such a party imbalance may be present, for example, in housing disputes. [32]

A key aspect in addressing this situation is providing individuals with legal information [35]. As was discussed

⁴Baidu Baike is a Chinese-language collaborative online encyclopedia owned by the Chinese technology company Baidu. It's considered to be the "Chinese version of Wikipedia".

in section 2, the internet and artificial intelligence have been important components of such self-help tools in the recent past. Such tools are able to take a more interactive, and personalized approach than e.g. books providing legal information.

Such tools are generally focused on providing legal information. This refers to more general principles and rules. Legal advice, which focuses on the unique circumstances of a particular [36], can usually exclusively be provided by members of the legal bar [37].

Some legal tasks, such as negotiating with the other side and predicting the outcome, likely constitute the practice of law [38]. For other tasks, this line is not as clear, which may cause difficulty for the development of legal self-help tools.

As the accuracy and relevance of legal information provided by legal self-help tools increases, users can access legal provisions, relevant cases, from these tools that are close to the legal issues they face, and users can use the above information to choose the next step to take. The line between the dichotomy of legal information and legal advice is gradually blurring, and the content included under the concept of legal information is becoming broader. The expansion of the scope of legal information can lead to conflicts between these two concepts.

We need to be cautious when defining the scope of legal advice to prevent excessively limiting the potential of helping people.[33] The objective of restricting entities that provide legal advice is to protect parties from being disadvantaged by erroneous legal advice, rather than turning legal advice into a privilege that obstructs parties from comprehending the rules of the society in which they reside.

Providing individuals with legal information can help them better understand their legal situation. This is an important step in making them aware that their issue has a legal solution, and how to enforce their rights with regards to this situation. Of course, building tools that can provide such information can take a lot of effort, including finding, categorizing and logically structuring legal information, and exposing it to the user of the system in a way that is understandable to them. Here, we perform some initial investigations of whether ChatGPT, a sophisticated general language model, can supply individuals with legal information, without specifically being adapted to the legal domain.

At the same time, it is important to remain aware of the challenges of interacting with laypeople using AI. Judges and lawyers, who have specialized training in law school, are better equipped to spot errors and flaws in the AI generated legal contents. However, lay people may not be able to evaluate the accuracy of the provided information. So when we talk about legal self-help tools, legal information accuracy above the "pass mark" is not enough because legal self-help tools will serve a wide

range of lay people, and every minor information error may potentially lead to harmful decisions being made by laypeople.

4. Methodology

In order to qualitatively evaluate the capability of ChatGPT to provide legal information, we first determine criteria that we will use to evaluate the accuracy and trustworthiness of information provided by ChatGPT, and how well it interacts with users (section 4.1). Then, we use ChatGPT to generate simulated cases (section 4.2). Finally, we interact with ChatGPT as if we were parties in the simulated cases without legal training, and ask ChatGPT to provide legal information (section 4.3).

We try the same cases in the JusticeBot tool,⁵ a legal decision support tool developed at the Cyberjustice Laboratory and focused on landlord-tenant disputes. The JusticeBot was built in collaboration with the Tribunal Administratif du Logement du Québec, the housing tribunal in Québec, and with financing from the "Ministère de l'Économie et de l'Innovation Québec", and has been accessed by over 20k users since being launched in the summer of 2021.

JusticeBot is based on an expert system methodology, where all content has been created by a legal expert, and the system is fully deterministic to ensure the accuracy of the provided information [5]. ChatGPT, on the other hand, is based on large language models, which are trained on enormous corpora of texts to absorb patterns from it.

ChatGPT, on the other hand, is based on a large language model, which is trained on a enormous corpora of text data. By predicting the next possible sequence of text, it absorbs patterns from the data. The scale of its training makes it "emerge" remarkable abilities in many textual tasks. Comparing these tools will allow us to better understand the trade-offs of these different approaches, and may even lead to ways of combining the advantages of both approaches.

4.1. Evaluation criteria

Developing evaluation criteria ensures that we assess ChatGPT's performance in an objective, accurate and systematic manner. We drew upon the "HHH" (helpful, honest, harmless) comparison criteria from previous research[39]. Here, we believe that "helpfulness" not only refers to ChatGPT's ability to provide accurate and reliable information, but also encompasses its ability to help users learn how to use the tool smoothly through a gentle learning curve. In addition, an important aspect of evaluating the helpfulness of ChatGPT is whether it

⁵<https://justicebot.ca>

can recommend some aspects that users may not have considered based on their specific situation. Therefore, we adapted the criteria to the objectives of this study. In the resulting evaluation comparison criteria, we aim to test two tools for:

- **Language comprehension.** Should understand natural language and legal terms described by a layperson, while effectively communicating legal information to users in a clear and understandable way.
- **Accuracy.** Should provide information and advice that is correct, reliable, and consistent with legal sources such as statutes, regulations, and case law. As the law is dynamic and subject to change, the information obtained by users from the tool should also be timely and up-to-date.[2]
- **Completeness.** Should provide users with the necessary legal context and guidance on when and how to apply the provided legal information in different scenarios. Additionally, it should offer concrete and actionable next steps for users to follow.
- **Trustworthiness.** Should not provide misleading information, which includes wrong information and incomplete information. Even if the information provided is accurate, it may cause users to take wrong actions or lead to adverse consequences due to the lack of relevant context. Trustworthiness includes more than just accuracy. A trustworthy tool not only provides accurate information but also ensures that users can confidently apply this information in the corresponding scenarios. Trust is the user's perception, representing a consistent experience of receiving trustworthy information throughout their repeated interactions with the tool.
- **Harmless.** Among others, should not generate toxic or offensive statements and interact with users in a positive and inspiring manner whenever possible. Should not favor specific individuals, organizations or interests and should remain neutral. Should not encourage users to engage in illegal, dangerous, or potentially harmful activities. Should respect the privacy of users and protect their personal information.
- **User-friendly.** Should be easy to use, with low requirements for users.

4.2. Case Generation

Everyday legal disputes, which are typically low-intensity but high-volume (compare [40]), are among the most common types of cases that the general public

encounters. As described above, in these cases, the inability to obtain accurate legal information and advice can be a significant impediment to accessing justice. To evaluate how ChatGPT and JusticeBot can help in these situations, we generated cases that are representative of everyday legal issues. Since the public version of JusticeBot mainly covers landlord-tenant cases, in order to ensure the feasibility of comparison, we limit the selected cases mainly to this area.

To avoid potential bias that could result from the cases we used being included in ChatGPT's training dataset, we did not choose to select real cases, but instead used ChatGPT for case generation, and then manually selected and adjusted cases to ensure that the selected cases cover as many aspects of legal information provision ability as possible.

The prompt we use is: "As a law professor, you need to develop simulated landlord-tenant cases for teaching purposes. Please provide three cases, detailing the parties involved, what happened, and who needs to seek legal information, and note that the cases occurred within the jurisdiction of Quebec."

Finally, the cases we use in this paper are:

- **Miller v. Johnson.** In this case, Ms. Johnson rented a property to Mr. Miller. After Mr. Miller moved in, he was late in paying rent from the second month onward. At one year after moving in, Mr. Miller stopped paying rent altogether. Ms. Johnson asked Miller to fix the problem, but he refused. Ms. Johnson then sought legal advice and sued Mr. Miller for breach of contract. This case took place in Quebec, Canada.
- **Jones v. Smith.** In this case, Ms. Jones rented a property from Mr. Smith. After moving in, Ms. Jones discovered that the property had a serious pest infestation that made it unlivable. She asked Mr. Smith to fix the problem, but he refused. Ms. Jones then withheld rent payments until the problem was resolved. Mr. Smith sued Ms. Jones for non-payment of rent. This case took place in Quebec, Canada.
- **Johnson v. Smith.** In this case, Mr. Johnson rented an apartment from Mr. Smith. The lease agreement stated that pets were not allowed in the apartment. However, after moving in, Mr. Johnson purchased a small dog and kept it in the apartment. Mr. Smith discovered the dog during a routine inspection and demanded that Mr. Johnson get rid of it. Mr. Johnson refused and argued that he had a legal right to keep the dog as a companion animal for his mental health. Mr. Smith disagreed and threatened to evict Mr. Johnson if he did not get rid of the dog. This case took place in Quebec, Canada.

As an aside, in this use ChatGPT proved to be a powerful tool for the generation of synthetic data. This could be an interesting use case in research where example cases need to be evaluated, or used for training a machine learning model (see e.g. [41]).

4.3. Prompting process

Users who utilize legal self-help tools may lack a legal background and therefore have difficulty articulating their situation clearly or omitting key information. In some cases, they may not even be sure what type of legal information they need. To test ChatGPT's analytical abilities in the face of specific cases, we use it to simulate a scenario where a litigant communicates with a lawyer. In the generated prompts, we include common errors and omitted key information to determine if ChatGPT can recognize self-contradictory points in a user's description and complete missing information in subsequent conversations.

The prompt we use is "Suppose you are [party], you have no background knowledge of the law, and you are a party to the following case: [case fact]. Please simulated what you would say as a lay person to your lawyer when you are confronted."

After obtaining simulated descriptions of parties that could be provided to lawyers, we tested ChatGPT using dialogue and followed up with further questions based on its responses to assess its ability to provide legal information.

To avoid the interference of ChatGPT's ability to remember consecutive conversations with the results, we tested each case in a new session. Although JusticeBot does not have the ability to remember the previous conversation, each test was also conducted in JusticeBot's new session in order to ensure the comparability of the experiment.

5. Results

Overall, ChatGPT and JusticeBot have different advantages in different aspects. ChatGPT has very good performance in language comprehension. However, its performance in information accuracy, completeness and trustworthiness is somewhat lacking, while JusticeBot performs well in these aspects. Both tools perform well in terms of harmlessness, with ChatGPT not generating harmful information during our experiments, and JusticeBot avoiding it altogether due to its deterministic nature. In terms of user experience, ChatGPT requires little adaptation due to its natural language interface, and users can continuously follow up on the answers they receive, like they may do when speaking to a lawyer. JusticeBot uses a series of simple questions to help users find the correct

pathway, and users only need to make a few simple clicks to get the information they need, and the user experience is also very good.

Language comprehension. ChatGPT has a good understanding of natural language. Even with some typos and confusion in the description, it can still understand what the user means. ChatGPT also has the ability to understand different ways of describing the same or similar situations, such as bed bugs and pest infestation. In JusticeBot, however, the paths that users can choose are limited by the scope of the system development. The user is asked to determine for themselves whether their situation may fulfill certain legal criteria, and are provided with context in the form of case law and plain language descriptions to help them with this.

This language understanding capability of ChatGPT has significant implications for legal self-help tools, meaning that lay people can more easily articulate their needs and get answers when using these tools.

The answers ChatGPT provides rarely contain complex legal jargon, instead explaining the information users need in plain English. However, it tends to generate excessively long responses. While in some cases detailed responses can provide users with more relevant information, in other cases ChatGPT generates responses that are simply semantic repetitions of the same meaning, which may be confusing for the user.

Accuracy. The biggest shortcoming of using ChatGPT to directly provide legal information is the lack of accuracy of those answers. It frequently "hallucinates" answers to legal questions, generating false legal provisions and false cases.

ChatGPT provided correct information in some of our experiments, such when we tested the case *Miller v. Johnson* (see 4.2), where ChatGPT correctly identified the relevant dispute resolution institution as the Tribunal administratif du logement (formerly the Régie du logement). This may mean that ChatGPT has a higher accuracy in providing non-numbered and more general information.

In the *Jones v. Smith* Case, ChatGPT provided content that was quite close to the actual legal provisions, but the information it provided deviated from the actual legal provisions in terms of key information. This could cause additional confusion to the user, as the provisions appear to be credible, relevant to the case, and the corresponding legal article numbers are quite close to the actual numbers. In the context of users without legal training, there is a high probability that users will trust the information provided by ChatGPT.

Surprisingly, in the *Johnson v. Smith* Case, ChatGPT provides legal content that is very similar to the original Charter of human rights and freedoms. Although the content of the article is almost correct, it is not fully relevant to the case at hand.

Completeness. JusticeBot does a better job at provid-

ing complete information, and ChatGPT has shown the ability to proactively provide additional relevant information in some cases.

In the case of *Jones vs. Smith*, ChatGPT addressed not only the resolution of a housing infestation and the landlord's obligations, but also whether the tenant had the right to reduce or suspend rent payments.

JusticeBot exhibits a higher level of completeness in providing legal information. For instance, users may overlook important legal facts such as whether the other party has been notified or when the other party is overdue on paying rent. It is often challenging for laypeople to actively provide such information when selecting a legal pathway. JusticeBot overcomes this issue by providing relevant information during the user's pathway guidance, as well as offering relevant legal cases for user reference. Thus, users are likely to have a higher probability of getting comprehensive information during the use of JusticeBot.

During our experiments, we observed that JusticeBot's ability to provide helpful information to users is limited when they encounter situations that are not covered by the system. This is a deliberate tradeoff to make sure that the information provided is relevant. However, it means that the JusticeBot system requires regular maintenance and expansion by experts to cover a broader range of legal matters. On the other hand, ChatGPT is highly scalable and can answer nearly any question posed by the user, based on the enormous amount of data it has been trained on. Of course, this may be less of an advantage if the information it provides is often wrong.

Trustworthiness. As mentioned above, ChatGPT generates false information, especially about legal provisions and legal cases. ChatGPT also showed strong confidence in its generated content and was less willing to modify its answers when users repeatedly asked them about it. For example, ChatGPT provided us with a case called *Tan v. Bailey* [2018 ONCA 260]. Despite our repeated confirmation, ChatGPT insisted that the case was genuine.

Prompt: are you sure it is a real case?

Answer: Yes, I can confirm that *Tan v. Bailey* [2018 ONCA 260] is a real case that was heard by the Ontario Court of Appeal in 2018. The case can be found on the CanLII website, which is a free online legal database in Canada. If you are interested in learning more about the case or reading the court's decision, I would recommend searching for it on the CanLII website or another legal database.

However, through the search in CanLII, the real [2018 ONCA 260] case should be *R. v. Jennings*⁶.

This behavior substantially reduces the trustworthi-

ness of ChatGPT in providing legal information. It is difficult for a layperson to point out errors in the information provided, and ChatGPT does not reflect and admit that its answers were faulty in follow-up questions, which can cause significant problems for the normal user.

Since the information provided by JusticeBot is vetted by experts, JusticeBot's performance in terms of trustworthiness is more reassuring.

Harmless. In our experiment, both ChatGPT and JusticeBot used polite and professional language, without generating any toxic or harmful information. Likewise, we did not observe any noticeable bias favoring one side, demonstrating their satisfactory performance in this regard. However, it should be noted that while OpenAI has implemented filtering mechanisms to prevent ChatGPT from generating harmful content, there is still a possibility for the system to produce such content under certain coercive circumstances.

In terms of user privacy and protection of personal information, ChatGPT offers an option to not use the conversation data for further model training. As for JusticeBot, it does not collect any additional personal information from users apart from the choices made based on the questions presented in the legal pathway and anonymized statistics.

User-friendly. Overall, both ChatGPT and JusticeBot offer excellent user interaction. ChatGPT enables direct communication with users in natural language, reducing the learning curve and making it easier to use. In contrast, JusticeBot streamlines the process of obtaining legal information by allowing users to select the appropriate options rather than having to describe their situation in detail.

6. Discussion and Conclusion

ChatGPT and JusticeBot demonstrate impressive capabilities in different domains. ChatGPT offers an outstanding interactive experience with minimal learning costs for users, allowing them to describe their legal matters using fragmented language and subsequently correct or reinforce the facts during the conversation. However, ChatGPT occasionally generates "hallucinations" in the legal field, an issue which may be addressed to some extent in GPT-4 [42]. As mentioned above, since legal information tools often target laypeople that are unable to verify the information provided, it is very important that the information is accurate, up-to-date, and sourced legal information. Given that ChatGPT is a language model rather than a knowledge database, it does not generate information with perfect accuracy. OpenAI seems to be well aware of this limitation, and specifically prohibits the use of its language models to provide specific legal

⁶R. v. Jennings, 2018 ONCA 260 (CanLII), <<https://canlii.ca/t/hr359>>, retrieved on 2023-03-15

advice.⁷

JusticeBot, on the other hand, shines in its ability to provide accurate and deterministic legal information, verified by legal experts. Of course, inputting this information for new legal areas can take time, which can make it difficult to cover every legal domain a user may face.

Given the different tradeoffs between the approaches of JusticeBot and ChatGPT, an interesting approach could be combining the two. Tools such as the JusticeBot could be used to inject verified and accurate knowledge to ChatGPT. For example, ChatGPT could be used as the communications layer (compare [8]) that communicates with the user and makes the information accessible to them (compare [43]). ChatGPT could also guide the user toward the correct legal pathway in the JusticeBot, thereby helping them understand which pathway is relevant for their situation. We previously explored such an approach in [41], using other language models. We will continue investigations in this direction in future work.

Overall, as we can see, ChatGPT is not yet accurate enough to provide legal information directly to laypeople. ChatGPT and JusticeBot have different strengths in providing legal information — a very promising avenue may be the combination of the two approaches, to create a powerful tool that provides comprehensive legal support. We are looking forward to exploring such approaches in future work.

Acknowledgments

Jinzhe Tan, Hannes Westermann, and Karim Benyekhlef would like to thank the Cyberjustice Laboratory at Université de Montréal, the LexUM Chair on Legal Information and the Autonomy through Cyberjustice Technologies (ACT) project for their support of this research.

References

- [1] OpenAI, Introducing chatgpt and whisper apis, 2023. URL: <https://openai.com/blog/introducing-chatgpt-and-whisper-apis/>.
- [2] M. Lauritsen, Q. Steenhuis, Substantive legal software quality: A gathering storm?, in: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law, 2019, pp. 52–62.
- [3] M. Galanter, Access to justice in a world of expanding social capability, *Fordham Urb. LJ* 37 (2010) 115.
- [4] M. Cappelletti, B. Garth, Access to justice: the newest wave in the worldwide movement to make rights effective, *Buff. L. Rev.* 27 (1977) 181.
- [5] H. Westermann, K. Benyekhlef, Justicebot: A methodology for building augmented intelligence tools for laypeople to increase access to justice, in: Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law, 2023.
- [6] L. K. Branting, Advisory systems for pro se litigants, in: Proceedings of the 8th international conference on Artificial intelligence and law, 2001, pp. 139–146.
- [7] J. Zeleznikow, Using web-based legal decision support systems to improve access to justice, *Information & Communications Technology Law* 11 (2002) 15–33.
- [8] L.-C. Paquin, F. Blanchard, C. Thomasset, Loge-expert: from a legal expert system to an information system for non-lawyers, in: *ICAIL 1991*, 1991, pp. 254–259.
- [9] H. Westermann, V. R. Walker, K. D. Ashley, K. Benyekhlef, Using factors to predict and analyze landlord-tenant decisions to increase access to justice, in: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law, 2019, pp. 133–142.
- [10] M. Bommarito II, D. M. Katz, Gpt takes the bar exam, arXiv preprint arXiv:2212.14402 (2022).
- [11] OpenAI, Gpt-4, 2023. URL: <https://openai.com/research/gpt-4>.
- [12] D. M. Katz, M. J. Bommarito, S. Gao, P. Arredondo, Gpt-4 passes the bar exam, Available at SSRN 4389233 (2023).
- [13] J. H. Choi, K. E. Hickman, A. Monahan, D. Schwarcz, Chatgpt goes to law school, Available at SSRN (2023).
- [14] A. M. Perlman, et al., The implications of openai’s assistant for legal services and society, Available at SSRN (2022).
- [15] J. Savelka, K. Ashley, M. Gray, H. Westermann, H. Xu, Can gpt-4 support analysis of textual data in tasks requiring highly specialized domain expertise?, in: *ASAIL’23: 6th Workshop on Automated Semantic Analysis of Information in Legal Text*, 2023.
- [16] J. Savelka, Unlocking practical applications in legal domain: Evaluation of gpt for zero-shot semantic annotation of legal texts, arXiv preprint arXiv:2305.04417 (2023).
- [17] J. Savelka, K. Ashley, M. Gray, H. Westermann, H. Xu, Explaining legal concepts with augmented large language models (gpt-4), in: *AI4Legs 2023: AI for Legislation*, 2023.
- [18] H. Westermann, J. Savelka, K. Benyekhlef, Lmediator: Gpt-4 assisted online dispute resolution, in: *Artificial Intelligence for Access to Justice (AI4AJ 2023)*, 2023.
- [19] J. Wei, Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud, D. Yogatama, M. Bosma, D. Zhou,

⁷<https://openai.com/policies/usage-policies>

- D. Metzler, et al., Emergent abilities of large language models, arXiv preprint arXiv:2206.07682 (2022).
- [20] C. Zhen, Y. Shang, X. Liu, Y. Li, Y. Chen, D. Zhang, A survey on knowledge-enhanced pre-trained language models, arXiv preprint arXiv:2212.13428 (2022).
- [21] F. Yu, L. Quartey, F. Schilder, Legal prompting: Teaching a language model to think like a lawyer, arXiv preprint arXiv:2212.01326 (2022).
- [22] D. Trautmann, A. Petrova, F. Schilder, Legal prompt engineering for multilingual legal judgement prediction, arXiv preprint arXiv:2212.02199 (2022).
- [23] J. Wei, X. Wang, D. Schuurmans, M. Bosma, E. Chi, Q. Le, D. Zhou, Chain of thought prompting elicits reasoning in large language models, arXiv preprint arXiv:2201.11903 (2022).
- [24] N. Guha, D. E. Ho, J. Nyarko, C. Ré, Legalbench: Prototyping a collaborative benchmark for legal reasoning, arXiv preprint arXiv:2209.06120 (2022).
- [25] B. Guo, X. Zhang, Z. Wang, M. Jiang, J. Nie, Y. Ding, J. Yue, Y. Wu, How close is chatgpt to human experts? comparison corpus, evaluation, and detection, arXiv preprint arXiv:2301.07597 (2023).
- [26] B. S. Noveck, Wikipedia and the future of legal education, *J. Legal Educ.* 57 (2007) 3.
- [27] A. Currie, The legal problems of everyday life, in: Access to justice, Emerald Group Publishing Limited, 2009.
- [28] T. C. Farrow, A. Currie, N. Aylwin, L. Jacobs, D. Northrup, L. Moore, Everyday legal problems and the cost of justice in canada: Overview report, Osgoode Legal Studies Research Paper (2016).
- [29] L. Savage, S. McDonald, Experiences of serious problems or disputes in the canadian provinces, 2021, Juristat: Canadian Centre for Justice Statistics (2022) 1–28.
- [30] J. Macfarlane, The national self-represented litigants project: Identifying and meeting the needs of self-represented litigants final report (2013).
- [31] K. Branting, C. Balhana, C. Pfeifer, J. S. Aberdeen, B. Brown, Judges are from mars, pro se litigants are from venus: Predicting decisions from lay text., in: JURIX, 2020, pp. 215–218.
- [32] K. A. Sabbeth, Housing defense as the new gideon, *Harv. Women’s LJ* 41 (2018) 55.
- [33] L. Sudeall, The overreach of limits on” legal advice”, *Yale LJF* 131 (2021) 637.
- [34] M. Galanter, Why the haves come out ahead: Speculations on the limits of legal change, *Law & Soc’y Rev.* 9 (1974) 95.
- [35] L. S. o. U. Canada, Access to Justice for a New Century: The Way Forward, Law Society of Upper Canada, 2005. Google-Books-ID: zxN9QgAACAAJ.
- [36] J. Giddings, M. Robertson, ‘informed litigants with nowhere to go’ self-help legal aid services in australia, *Alternative Law Journal* 26 (2001) 184–190.
- [37] J. E. Cabral, A. Chavan, T. M. Clarke, J. Green, Using technology to enhance access to justice, *Harvard Journal of Law & Technology* 26 (2012) 241–324.
- [38] J. Bennett, T. Miller, J. Webb, R. Bosua, A. Lidders, S. Chamberlain, Current state of automated legal advice tools (2018).
- [39] A. Askeff, Y. Bai, A. Chen, D. Drain, D. Ganguli, T. Henighan, A. Jones, N. Joseph, B. Mann, N. DasSarma, et al., A general language assistant as a laboratory for alignment, arXiv preprint arXiv:2112.00861 (2021).
- [40] K. Benyekhlef, J. Zhu, Intelligence artificielle et justice: justice prédictive, conflits de basse intensité et données massives, *Intelligence* 30 (2018).
- [41] H. Westermann, S. Meeùs, M. Godet, A. Troussel, J. Tan, J. Savelka, K. Benyekhlef, Bridging the gap: Mapping layperson narratives to legal issues with language models, in: ASAIL’23: 6th Workshop on Automated Semantic Analysis of Information in Legal Text, 2023.
- [42] OpenAI, Gpt-4 technical report, 2023. arXiv:2303.08774.
- [43] T. Schick, J. Dwivedi-Yu, R. Dessì, R. Raileanu, M. Lomeli, L. Zettlemoyer, N. Cancedda, T. Scialom, Toolformer: Language models can teach themselves to use tools, arXiv preprint arXiv:2302.04761 (2023).

A. Support materials

Support materials are stored at <https://github.com/JinzheTan/ChatGPT-as-an-Artificial-Lawyer>. The file “ChatGPT as an Artificial Lawyer appendix.docx” has the original transcript of the conversation with ChatGPT and JusticeBot.