Risks And Regulatory Framework of AI-Generated Content (AIGC) in the Judicial Field

Gong Xuan

Faculty of Law, The University of Hong Kong, Hong Kong, China

Abstract. New types of Artificial Intelligence-Generated Content (AIGC), such as ChatGPT and Codex, can assist lawyers and judges in reviewing files, writing legal documents and participate in legal proceedings, which can reduce the work pressure of legal practitioners and improve judicial efficiency to a certain extent. Moreover, AIGC can provide more intelligent legal services to the public in the future. However, the in-depth application of AIGC techniques in the judicial field may bring new security risks and ethical crises. A hierarchical governance framework should be adopted to clarify the allocation of responsibilities among different target groups and effective regulatory measures should be established to support the sustainable and healthy development of AIGC services in the judicial field.

Keywords: Artificial intelligence in the field of justice; Review archives; Writing legal documents and participating in legal proceedings

1. Application of AIGC in the judicial field

On 30 November 2022, OpenAI launched ChatGPT, a chatbot program whose outstanding ability to generate text and make natural language conversations made it the fastest-growing consumer app in history. Moreover, the concept of AIGC enters the public eye. AIGC is an algorithm or model that can generate text, images, videos, code and other content based on input prompts. Currently, Internet companies such as Google, Microsoft, Anthropic and Cohere have launched a variety of AIGC models. In the judicial field, some legal practitioners are experimenting with AIGC techniques to reduce the stress of mechanical legal work and continue to explore the boundaries and possibilities of related AIGC services.

Current AIGC techniques, such as the majority of large language models (LLaMs) generally require training and deep learning based on big data and large model corpus [1]. The existing legal regulations and judicial decisions are transformed into the form of data through extensive training based on different AIGC model structures, and the corresponding legal structures and legal principles are internalised into algorithmic models. These AIGC models can be combined with existing online judicial databases and judicial archives systems, and they can also be applied to several smart legal service platforms. In judicial practice, AIGC can first replace some of the mechanical tasks undertaken by lawyers and judicial officers in order to improve the efficiency of their routine work [2]. For example, some AIGC services can check, retrieve and integrate the basic information in the case files to help the relevant personnel conduct legal searches and improve their efficiency of judicial review. At the same time, AIGC can also identify the specific content of a case and intelligently push relevant legal regulations, judicial decisions and other information related to the content of the case. In addition, AIGC can assist in the writing of preliminary drafts of normative legal documents in different judicial procedures, such as briefs, bills of complaint, standardised deposition questions in the trial proceedings and other legal materials such as wills, contracts of sale, etc[3]. Furthermore, some smart services with AIGC techniques can understand the general legal requirements of the people in need and the relevant legal facts through interactive questions, then provide different legal advice and solutions with a reference value [4]. Analogous services can also help target users to review contracts and other legal materials or promote administrative and judicial proceedings in their daily life. Finally, AIGC can make the courts more intelligent not only by undertaking routine tasks such as translation, recording and presiding over court hearings but also by intelligently pushing relevant legal information to judges to help them make judicial decisions and write verdicts.

2. Risks and challenges in the use of AIGC in the judicial field

2.1. Practical issues

On the one hand, the application of AIGC in the judicial field will bring about a series of practical problems such as unemployment and information security issues.

First, as the accuracy and professionalism of the legal content generated by AIGC gradually increase, some AIGC systems can provide more professional legal advice and write normative legal documents to help users in need to participate in different judicial procedures, therefore legal practitioners such as lawyers will be replaced to a certain extent. Meanwhile, the improvement of the intelligence of the current judicial system may also result in the unemployment of some judicial staff such as clerks and legal personnel engaged in judicial investigation and judicial identification. According to a Goldman Sachs report, artificial intelligence (AI) could replace 300 million full-time jobs, while 44% of legal work could be automated by artificial intelligence [5]. Not only will this change cause serious unemployment and severe economic issues, but a large number of legal practitioners may be alienated into modifiers or technical maintainers of AIGC legal service systems. At the same time, companies that master relevant techniques or smart legal service platforms may monopolise the relevant legal service market. As a result, the Turing trap caused by AIGC may further amplify the issues of social division and power imbalance.

Second, the speed and effectiveness of AIGC regulation cannot cope with the rapid development of AIGC technology and the complexity of related issues. Along with the development of AIGC, cyberspace may be flooded with fake information, fake pictures and fake news. AIGC may also be used by criminals for new types of illegal and criminal activities, which may cause brand-new issues of intellectual property infringement, privacy infringement, fraud, data security and many other legal issues [6]. These new types of legal issues are more difficult to prevent, identify and detect than traditional ones. Moreover, since the AIGC system itself cannot bear legal responsibilities as an independent subject, and the distribution of responsibilities between the providers and users of AIGC technology is still unclear, therefore the use of AIGC in judicial practice may have the issue of "seeking blame" for mistakes or negative consequences. In addition, AIGC technology is developing extremely fast, leading to the emergence of new types of problems. Besides, security issues of AIGC such as technical problems and hacker attacks in the judicial field may also lead to more serious consequences compared to traditional technical issues. However, the relevant laws and regulations are lagging behind, and the staff of the judiciary have not yet had sufficient professional AIGC knowledge and investigative skills to deal with these new problems, which poses a new challenge to the legislation and judicial practice in the AIGC era [7].

Third, the application of AIGC in the judicial field will bring security issues of technology to a whole new level. The participation of AIGC in the judicial procedures actually means that humans have transferred a certain degree of judicial power to AI, which also means that AIGC in fact indirectly masters a large amount of information and resources correspondingly. Therefore, the operation process of AIGC may influence the outcome of judicial activities and even play a role in the redistribution of human resources. How to ensure the security of private information and relevant social resources during the application of AIGC in the judicial field and to eliminate the possibility of criminals using technical approaches to interfere with algorithmic systems is a core topic of AIGC technology security.

2.2. Ethical issues

On the other hand, the use of AIGC in the judicial field implies a deep involvement of modern technology in judicial practice and the interference of technical factors in ethical concepts such as fairness and justice, which may bring brand-new ethical challenges.

First of all, different social groups may have different understandings and opinions about the application of AIGC in the judicial field, which may subsequently give rise to issues such as the “illusion of authority” or “Tacitus Trap”. More diverse understandings of judicial justice can be

adopted in the era of AIGC. Some technological optimists may presuppose the absolute correctness of AIGC systems and thus become overly dependent on it, therefore they may even turn to question the accuracy of the work of human lawyers and human judges, then ask for a greater expansion of AIGC technology in the judicial field. For some technological pessimists, they may question the accuracy of the content generated by AIGC and criticize the transparency and interpretability of the algorithm process, arguing that the "black-box effect" of AIGC in the judicial sphere could lead to the erosion of the legitimacy and authority of the judicial procedures. In the absence of a proper solution to the practical issues mentioned above, the in-depth application of AIGC in the judicial field may also cause the public to doubt the justice, authority and security of the judicial system in the new context, which in turn will have a negative impact on the judicial system and cause a crisis of judicial trust.

Furthermore, the justice and morality of AIGC participation in the judicial field are also controversial. Currently, there have been attempts by individual judges to use AI adjudication independently in judicial trials to improve judicial efficiency. However, AI adjudication may not be able to achieve judicial justice or provide the public with a sufficient sense of justice. Some scholars believe that the increased use of AIGC in the judicial field will undermine the dignity of the human community and will trigger a new round of social ethical crises and philosophical crises. In fact, AIGC technology is causing people to question the relationship between human beings and AI. Some people doubt whether humans may no longer be regarded as the purpose as become a subsidiary of AI technology [8], while others believe that the deepening use of AIGC in important social areas such as the judicial field may further dismantle the subjectivity of humans.

3. Framework for future AIGC regulation in the judicial field

3.1. Safety First - the basic principle of AIGC applied in the judicial field

In fact, the relationship between technological development and technological security has always been a dilemma that is tricky to balance. In traditional legislative practice, the development of new technologies is encouraged in most scenarios and governments tend to find solutions and explore regulatory ideas after new problems arise. However, AIGC is highly uncontrollable and unpredictable, which may have a disruptive impact on social structures and even human civilisation. Especially in some crucial areas such as the judicial field, the application of AIGC may have a significant impact on the distribution of resources in human society and on the ethics and dignity of human communities, therefore problems such as system errors or hacking may cause more serious security problems. Therefore, "the brake should be installed before we hit the road". It should be ensured that no uncontrollable risks or insoluble situations will arise before the AIGC algorithm is upgraded. In the short term, the red line for the AIGC in the judicial field should be clarified to ensure urgent needs and minimize related risks. In the longer term, a fundamental and systematic AIGC legislation system should be promoted on an ongoing basis. The long-term healthy development of AIGC in the judicial field can only be achieved by placing a higher priority on security and social stability. In fact, some relevant legislative practices in the world also reflect corresponding ideas and tendencies [9], such as the "Artificial Intelligence Act" proposed by the European Commission and the "AIGC Service Management Measures" issued by China [10].

3.2. Hierarchical governance - a basic framework for AIGC regulation in the judicial field

The framework of the AIGC regulatory system should be structured in a hierarchical and categorical way, using separate regulation approaches for different target groups.To begin with, the target groups of regulation can be identified as technology developers, service providers and service applicants in the order of the AIGC operation process. Governance concepts and regulatory priorities should be distinguished in different processes and the distribution of responsibilities among different groups should be clarified. On the technology development side of AIGC, more attention should be paid to the professional governance of the algorithm base and underlying technology. Effective

technical intervention approaches such as censorship mechanisms should be introduced to guide the appropriate application of AIGC in the judicial field. To be specific, relevant AIGC systems should be examined to understand the existing legal regulations accurately and the authenticity, legality, compliance and accuracy of the generated legal content should be ensured. Furthermore, effective technical mechanisms should be put in place to prevent algorithmic discrimination and to ensure the confidentiality and security of judicial data used for AIGC. For AIGC service providers in the judicial field, the security of the AIGC services should be fully assessed and the authenticity and legality of the generated content should be tested in multiple rounds before the service is launched, and effective risk prevention mechanisms should also be set up in advance. In the initial stage of providing the AIGC service, a manual review mechanism should be established to review the relevant legal content generated by AIGC and further verify the accuracy of the service. A reasonable labelling mechanism should also be installed to tag the resource of the legal content. At the same time, relevant legal service platforms should play the role of supervision and indirect management of the target users. On the application side, users should be guided to standardise the use of AIGC legal products for different groups such as judicial personnel, lawyers and the general public. The current legal system and case-handling capacity should also be improved in order to prevent users from using AIGC to commit new types of offences.

At the same time, a classification system should also be established for the risks that different target users may encounter according to the degree of risk, and different levels of regulation measures should be taken for different cases [11]. In general, the AIGC applied to the judicial field should establish stricter criteria and restrictions on the accuracy, security and stability of AIGC service products compared with AIGC products in other areas. In the process of policy development, policymakers should take the opinions of AIGC technology developers into consideration and then objectively and systematically review relevant ethical risks in conjunction with the opinions of experts and the public. After further clarifying the responsibilities and obligations of different target groups, a transparent accountability mechanism should be developed in conjunction with the risk classification system. At the same time, the professional training about AIGC techniques of current judicial personnel and the introduction of relevant experts should be strengthened to enhance the professional level of the judiciary to cope with the new issues in the AIGC era.

At the societal level, risk planning and social guidance should be done appropriately to achieve a healthy shift in the judicial system and social civilisation in the AIGC era. First of all, a statistical assessment should be made on the impact of AIGC on future unemployment and other issues, then a reasonable plan should be made to promote a functional transformation of lawyers and other judicial practitioners in order to achieve an effective convergence of the functions of AIGC and different groups of people working in the judicial field. In addition, the AIGC industry should also be guided to form a healthy industry self-discipline, which may raise awareness of technical risks and ethical issues within the industry. At the same time, relevant academic groups and civil organisations can be encouraged to increase their discussions on relevant issues including conducting more in-depth research on the relationship between justice and AIGC in the form of seminars and symposiums. Last but not least, a benign interface and a smooth transition in a specific period should also be achieved at the entire social level. The existing value systems of social ethics, public order and morality in human civilisation should be incorporated into the practice of AIGC application in the judicial field, and the path of compatibility between technological values and current social civilisation in the new era should also be actively explored in various fields. At the same time, law popularization and value guidance should be used to promote the shaping of positive social values in the AIGC era.

References

[1] Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt.

[2] Iu, K. Y., & Wong, V. M. Y. (2023). ChatGPT by OpenAI: The End of Litigation Lawyers.

[3] Perlman, A. M. (2022). The Implications of OpenAI's Assistant for Legal Services and Society.

[4] Gandhi, P. A., & Talwar, V. (2023). Artificial intelligence and ChatGPT in the legal context.

[5] See https://observer.com/2023/03/generative-a-i-may-replace-300-million-jobs-goldman-sachs-study/.

[6] Li, Z. (2023). The Dark Side of ChatGPT: Legal and Ethical Challenges from Stochastic Parrots and Hallucination.

[7] Chen, C., Wu, Z., Lai, Y., Ou, W., Liao, T., & Zheng, Z. (2023). Challenges and Remedies to Privacy and Security in AIGC: Exploring the Potential of Privacy Computing, Blockchain, and Beyond.

[8] Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines.

[9] Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and machines, 30(1), 99-120.

[10] See the Act at http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm.

[11] Similar ideas can be seen from the EU AI Act, in which the risks of AI are categorized into unacceptable risk, high risk, limited risk and low and minimal risk.