Science of Law Journal (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/law.2025.040118
ISSN 2616-2296 Vol. 4 Num. 1

The AI that Causes Fire and Trouble: Research on Countermeasures for the Legal and Ethical Dilemmas of ChatGPT

Yang Xinyi, Liu Shu, Wu Cuiping, Lin Hongtao

Guangdong Ocean University, Zhanjiang, 524000, China

Keywords: ChatGPT, ethics of science and technology, academic ethics, artificial intelligence

Abstract: ChatGPT is fine-tuned from the large-scale pre-trained language model GPT-3.5, which achieves better results in comprehension, result accuracy, and the ability to recognize illegitimacy and bias. Focusing on the science and technology ethics and academic ethics of ChatGPT technology, we conduct an in-depth analysis and assessment of the technical principles, technical risks, security issues, and social impacts of ChatGPT technology. We hope that through these efforts, we can better promote the healthy development of ChatGPT technology, facilitate the continuous progress of science and technology, and make ChatGPT technology better serve the society with human welfare as the core.

1. Introduction

The ChatGPT, standing for Chat Generative Pretrained Transformer, is an artificial intelligence model for chatbots created by OpenAI in 2022; it was launched on November 30 that same year. ChatGPT communicates in natural conversational language with the user and generates textual responses incorporating code output, textual translation, essay and novel writing, among other activities, in a natural and smooth dialogue. ChatGPT's principal competitiveness is vested in its innate learning function, which is a byproduct of very sophisticated generative algorithms such as Transformer cum micro-engine with human feedback, and reinforcement learning allowing large-scale filtering and processing of vast swathes of data, which, importantly, has begun to replace human beings doing efficient sourcing of various types of extremely complex information. With the fast and linear development of AI, efficient devices like ChatGPT will quickly substitute and transform people's life and working habits, manifesting itself in current debates surrounding actual job loss in the teaching, accounting, lawyers, coders, and other sectors, triggering major layoff waves from the beginning of 2023 and raising multiple question marks, controversially, in the context of law and ethics. This investigation aims to address the issues that revolve around both law and ethics regarding regulation, which will ultimately aid in guiding the parties concerned.

2. Legal Difficulties and Controversies Raised by ChatGPT

The Artificial Intelligence Index Report (2024) (Artificial Intelligence Index Report) released by

Stanford University shows that 52% of the respondents in 2023 said that there is concern about artificial intelligence, an increase of 13% over the previous year. [1] Along with the rapid development of AI, AI Anxiety is also growing rapidly, and Technophobia has become a real and highly concerned issue. [2] In the legal field, ChatGPT has caused copyright issues, infringement liability issues, false information issues, privacy leakage issues and so on, which have also attracted increasing attention from the public.

2.1. Legal Issues Raised by ChatGPT

2.1.1. Privacy and Protection Challenges

Given its ease of use, round-the-clock availability, low cost, and ability to handle multiple ports simultaneously, many enterprises use ChatGPT to power online customer service systems. Clients inquire about using different products or the return and exchange policies; ChatGPT provides the service between calls, thus improving the efficiency and quality of customer service while enhancing customer satisfaction. However, the training process necessitates a considerable amount of data, including any user interaction data. For example, failing to follow strict notification and authorization principles in data collection may lead to violation of user privacy by the people tasked with the collection process. In this case, if important data were disclosed, who would take responsibility for losing it-the person who leaked it or should the blame fall upon ChatGPT, whose voice said something the person leaked? [3] Although ChatGPT states it does not retain users' input data for extended periods of time, in an attempt to better understand the context and enable meaningful dialogue, it may retain some of the data for a limited period of time. While it is claimed that this data is anonymized, the data security concerns regarding ChatGPT have to be taken seriously.

2.1.2. The defamation and fraud conundrum

In the age of big data, information spreads like thunder, and once content generated through ChatGPT gets untrue negative comments about individuals, enterprises or organizations, it can easily cause reputational damages to individuals, enterprises or organizations, raising the risk of defamation. If, in the commercial arena, it is misused with malicious intent, it may fabricate false product information or service promises to deceive consumers and disrupt market order. Besides, being seemingly authoritative, ordinary users find it hard to identify authentic information leading to a more widespread circulation of misinformation, escalating the risk of defamation and fraud. Such acts infringe not only individuals' rights and interests but also poison the social trust system and information ecosystem. Again, many of the legal issues relating to the development of AI need to be clarified by legislation; personal information protection and data-sharing disputes are coming up. As a result, in China, the overall regulation of AI still has to be polished. [4]

2.1.3. Copyright infringement challenge

The number of users of ChatGPT is enormous and therefore, it will be dragged into disputes on ownership of intellectual property rights. Even though OpenAI hasn't yet revealed the specific database from which ChatGPT has mined its information, it cannot be ruled out that the mining process does necessarily include the adopting of copyrighted works by the owner of those works without consent and outputting them up to replies in subsequent use, which is a series of cases under reasonable use of someone's works as found under China Copyright Law Article 24. This article, while giving a laundry list of 13 situations where others' works may be used, doesn't include any reference to AI's use of others' works in data mining, as ChatGPT does. [5] In respect of China's

current legislative status, combined with the existing jurisprudence worldwide, the scholars have put forth three major viewpoints - "AI as an Author - AI is not an Author - AI Content may constitute a work AI cant be Author AI Content may constitute a work". [6] Since the intellectual property and copyright issues of ChatGPT are complicated, there is no clear conclusion on whether the content generated by ChatGPT is copyrighted or not, and if its output is regarded as an original work of authorship, it may be in conflict with the definition of human creations in the traditional copyright law. In addition, when users create derivative works based on ChatGPT, how to determine the copyright shares and right boundaries of each party in the new works is also an urgent issue, and this series of intellectual property rights and copyright issues have posed a new challenge to the existing legal system and creative ecology.

2.2. ChatGPT legal disputes

2.2.1. Disputes over copyright ownership of generated content

There are various aspects regarding this topic in which modern day legal discussions center on: the copyrights of texts generated by the chatbot ChatGPT. General copyright law recognizes intellectual works-the "human mind"-with some kind of protection. In cases involving text written by AI, however, that principle seems to break down completely. Such questions might be asked: would an article, code, or design generated by ChatGPT after a user instruction qualify as the original work? In practice, it has shown that the user will normally declare their input and choice of a completed work as original labor and, thus, be entitled to copyright; at the same time, the developers maintain that this was, for all practical purposes, the amalgamation of training data for the model and an algorithm for generating, and that ownership therefore rests with the platform. Such a conflict is illustrated in the fact that in the courts, conflicting views have emerged. For instance, whereas the U.S. Copyright Office has refused to register copyrights for images generated by artificial intelligence, the South African government recognized the protection of copyrights of musical works assisted by artificial intelligence. Such inconsistencies reflect the slowness of existing laws compared to the rapid pace of technological innovations. Urgent legislation arises to tackle the criteria under which "degree of human involvement" should be measured.

2.2.2. Technology dependence and crisis of confidence in judicial practice

The ruling made by ChatGPT in this phase of the judicial history has instilled much debate in light of the opposition between legal tradition authority and trustworthy technology. Though the 2023 Colombian Judge Juan Padilla made a strong point by quoting ChatGPT on its inference that medical coverage for children suffering from autism should be undertaken by healthcare insurance, this reliance on AI for legal declarations and conclusions has been criticized by the academic community as a serious dilution of an already weak judicial argument. The specific issue of AI "hallucinations" has given birth to mistaken legal theories: a Wyoming attorney raft up connected to appearance-related model misconduct while filing a pleading and hence was subject to punishment because the model had conjured nonexistent precedents brought to light the flawed informational trustworthiness of the tools powered by AI in the professional field. Such happenings lend not only an air of frivolity to solemn judicial documents but may also stir the public debate regarding the legitimacy of rulings.

2.2.3. Lack of mechanisms for allocation of responsibility and risk-taking

Establishing liability is a challenge when content generated by ChatGPT causes actual damage (e.g., in the case of defamation or incorrect medical information). Current laws often assign

responsibility to whoever is publishing the content, ignoring, however, the autonomy that AI systems enjoy. For example, a ChatGPT owner in Hangzhou put forth an “policy to cancel traffic restrictions” in 2023. The blame was placed upon the end user, though questions surrounding the platform's responsibility as to the authenticity of the model's output remain. The question is, in essence: can the developer get away with the liability on a ground of “technology neutrality”? In addition, there might be other privacy breaches associated with ChatGPT's processing of data where the model might inadvertently enter shared training data pertaining to personal information in conversation, violating GDPR by both the user and developer. Such issues expose the limitations in the current liability framework, requiring legislation to clarify a system of risk sharing between developers, users, and third parties, as well as a mechanism with which to trace AI errors.

3. ChatGPT's Ethical Dilemmas and Controversies in Science and Technology

3.1. The Ethical Dilemma of Science and Technology Raised by ChatGPT

3.1.1. The Crisis of Cognitive Sovereignty and the Reconstruction of the Knowledge Production Chain

ChatGPT is reconstructing an agent undergirding human thought processes and its mechanism of knowledge generation constructs an implicit "algorithmic cognitive hegemony". Studies have linked increased dependency on AI-derived answers to attenuating the encoding function of the hippocampus in long-term memory, hence the gradual loss of the capacity to construct knowledge autonomously among human beings. The field of education has already experienced a form of "reverse domestication of knowledge", where students are now adjusting their thought processes to fit within the judgment standards set by AI with the aim of attaining higher grades, and this cognitive alienation has reduced education merely to a tool for algorithmic training. When 90% of academic papers cite AI-generated content, human civilization will be trapped in a closed one, where second-hand knowledge is played out inside and therefore subsists somewhat eroded of soil for original ideas on the Argentine open sky of algorithmic optimization.

3.1.2. Digital Personality Deconstruction and the Fading of the Ethical Subject

In the “AI Companion Dependency” lawsuit in Japan in 2024, the plaintiff's real-life social skills deteriorated as a result of his long-term emotional connection with ChatGPT, and for the first time, the court was faced with the challenge of determining “digital personality injury”. The existential crisis triggered by this technology is also reflected in the level of moral responsibility: when AI nursing robots choose to give up saving critically ill patients based on their algorithms, a vacuum of responsibility is created for the developers, operators, and users in their decision-making chain. It is necessary to establish a system of “digital personality firewalls”, mandatory labeling of AI's inanimate attributes, and legislation to prohibit ambiguous human-robot interactions in key ethical decision-making areas.

3.2. The Ethical Controversy of Science and Technology Raised by ChatGPT

3.2.1. Cognitive Colonization and the Reconstruction of Power in Digital Civilization

The hegemonic control of ...knowledge through algorithms underlies the phenomenon. What it has essentially contributed to, the output mechanism of knowledge, is further elaborated on as a new way of culture colonization in the age of digital. This insidious Western-centric algorithmic bias in the corpus of training data essentially systematically sidelines non-Western knowledge

systems in favor of African oral history and Asian wisdom of traditional healing, which could not even enter the knowledge base of the model because there are no digitized records of it, hence the colonial nature of cultural interpretation of technological output. This restructuring of cognition sweeps into the educational system: students are learning to conform their thinking to AI scoring criteria simply to get higher grades, which constitutes reverse knowledge domestication; education is thereby alienated from inspiring one to the teaching of arithmetic. Come the time that 90% of penned-down ideas rest on the shoulders of AI works-setting human civilization free inside the internal circular flow of ideas and echo knowledge-in essence, the originality is eroded behind the incipient forum of algorithmic optimization, this anomaly in the chain of conscious creation translates into civilization's paramount crisis [7].

3.2.2. Emotional mimesis and the crisis of dissolution of the ethical subject

ChatGPT's anthropomorphic interactions are deconstructing traditional ethical relationships and creating “digital personality cognitive dissonance. ‘20 In the Japanese ’AI Partner Dependency” lawsuit of 2024, the plaintiff's real-world social skills deteriorated as a result of his longstanding emotional connection with ChatGPT, and the judicial system was faced for the first time with the difficulty of determining “digital personality damage.Frequent users of emotional AI are more likely to show reduced empathy in real-life situations, suggesting that technology is reshaping human ethical neural networks.

ChatGPT's technological security risks are still at a hypothetical and predictive level, and no one knows where ChatGPT and other AI technologies will go in the future. However, the AI technology security risk behind ChatGPT has forced society to re-examine and assess whether the existing governance system and governance framework can cope with various emerging technology security risks and unknown challenges.

4. Legal and Ethical Difficulties Raised by ChatGPT: A Countermeasure

4.1. Theoretical Innovation

4.1.1. Limitations of existing theories

The traditional theory of tort liability revolves around “human beings” and operates on different principles, while the autonomous generation capability of ChatGPT blends the very contours of liability. The current legal systems, such as the AI Liability Directive of the European Union, although proposing collective responsibility of algorithm designers, data providers, and other parties involved, have left the marginally framework of how attribution of liability is to be obdivided with respect to the overlaying risks of multiple links unaddressed. For instance, the ChatGPT “phantom” output provides some leaked bias but it is very difficult to follow its algorithmic reasoning, and present ethical assessment lacks a dynamic monitoring mechanism for process justice. The Copyright law is functioned from and based on “originality”, which however remains controversial about the originality of content produced by an AI.

4.1.2. Viewpoint of this article

This paper starts from the legal and ethical issues brought by chatgpt, considering that the technological security risks of ChatGPT are still at the level of envisioning and predicting, and that although the future development of ChatGPT and other AI technologies is still unknown, the current AI technological security risks behind ChatGPT can allow society to re-examine and evaluate the existing governance system and governance framework to deal with the various emerging

technological security risks as well as unknown challenges. [8]

4.2. Specific initiatives

4.2.1. Technology governance for a dynamic ethical framework

An “ethical nerve layer” is constructed at the longer-term algorithmic architecture, and a dynamic adjustment mechanism is established through real-time monitoring of the social impact factors (e.g., bias index and moral risk value) output from the model. This model of technology governance differs from traditional static compliance review instead of forcing the AI Ethics Guidelines into quantifiable technical parameters that would be then embedded into the Transformer architecture. An "ethical filter" has been set up at the input opening of data and established mathematical costs/benets based on adversarial training to eliminate implicit discrimination from training data, such as semantic bias elimination techniques regarding sensitive dimensions such as gender and geography. From this endogenous technology governance, a shift in paradigm occurs in an instant from “ex post facto compliance” to “ex ante regulation,” allowing ChatGPT to comply with the restrictions of Section 24 of the Physicians Act when generating medical advice while avoiding infringing the risk outlined in Section 3 of the Copyright Act regarding creative work.

4.2.2. Synergistic Evolution of Judicial Techniques and AI

Building a 'human-machine collaborative adjudication matrix' bringing human witnesses' legal experience to interpretable algorithmic attributes. Creating a jurisprudential knowledge map of 270 million cases trains a signature legal reasoning model to gain clarity on contextual elements in the meaning of Article 142 of the Civil Code. A three-stage review mechanism for judicial practice is formed; AI generating legal opinion → interlocutors making value judgments → senior human-in-the-loop verification. This has achieved early success with IP cases; patent infringement is established on AI comparing technical features, while the judge remains focused on the application of the principle of equivalence in Article 59 of the Patent Law. Whereas, with Article 55 of the Personal Information Protection Law specifying automated decision-making reasoning dispute cases, Blockchain technology hopes to create the Algorithm Responsibility Traceability Chain to facilitate the reverseability of the algorithmic weight distribution, i.e., the model's decision route. This symbiotic development has facilitated the metamorphosis of legal AI from mere instrumental owing to its auxiliary in nature to a constitutive too, remaking the epistemic framework of judicial proof.

5. Ethical control

5.1. Ethical issues in academic education

By offering new tools and methods for academic research and studies, ChatGPT presents many ethical dilemmas within the academic terrain. Some of the generative abilities of ChatGPT afford that a process of writing academic works becomes even more layered behind automation, thus ramping up serious issues related to academic integrity. For example, students or researchers could basically ask ChatGPT to produce a dissertation or a research paper without actual original thought and research being expended. Text from ChatGPT could easily be misconstrued to be the creation of a human author, muddying the waters when it comes to the integrity of the academic contribution. In such cases, due to the use of ChatGPT-written content, the academic content shall not be evaluated to have fulfilled the requirements under the academic community for originality and depth, and quality irremediably becomes vulnerable, hence affecting academic research. A much

bigger issue is that, with the assistance of ChatGPT, the basic purposes of education will no longer be real in the world of academic research. The use of ChatGPT to conceive a research hypothesis or analyze data could raise questions about the trustworthiness and originality of the work.

5.2. Ethical issues in publishing

As ChatGPT is widely used in the publishing field, it also raises certain ethical issues: the content generated by ChatGPT contains parodies or retellings of existing works, and its training data comes from a wide range of web content, which makes it infringes on the intellectual property rights of others when generating content and thus leads to copyright disputes. Transparency in editorial decision-making is also related to readers' trust in the publication, and if readers do not have access to the basis and process of editorial decision-making, they may be skeptical of the quality and fairness of the publication.

5.3. Discriminatory and prejudicial content

AI language models like ChatGPT have a big ethical dilemma when generating text that can be racist and biased.ChatGPT was trained on a massive corpus of the Internet, which holds the burden of historical biases and discrimination held by humans throughout society-an algorithm learns the statistical laws of that data, so if there are biases inherent in that training data, there are good indications that the algorithm will replicate these biases in its generation of text. For example, if in the training data women are portrayed primarily in a domestic context, ChatGPT might still find itself unconsciously reinforcing such conceptions of gender when generating its relevant content. The biases built into the design of the algorithms cause greater discrimination in the text produced by ChatGPT by optimizing fluency and accuracy to the detriment of fairness and diversity, thus favoring some groups and insinuating others.

6. Legal and ethical control recommendations

To cope with the security issues and ethical problems faced by ChatGPT, simultaneous research on AI-related legal, ethical and social issues should be strengthened while generative AI technology is being developed, and the legal, regulatory, ethical and moral frameworks to safeguard the healthy development of generative AI should be suggested and optimized. [10]

6.1. Legal Control Recommendations

Foremost, special attention should be given to data protection and the compliance of ChatGPT with applicable data privacy legislation when processing user data, including but not limited to GDPR, CCPA, and so on. Users should be duly informed about the purpose for which their data are going to be used and asked for permission. Second, introduce a content censorship mechanism that prevents ChatGPT from generating or propagating illegal or harmful information. As necessary, content filtering and review mechanisms could be put in place to ensure generated content is legally compliant. Furthermore, ensure that ChatGPT does not violate the intellectual property rights of others, including patents, trademarks, and copyrights. Ensure that the materials generated do not contain any unclaimed protected information. Then, there should be a transparent mechanism through which users can gain an understanding of how ChatGPT operates and its limitations. In tandem with this, outline clearly the scope of responsibility of ChatGPT to prevent people from misusing or misconceiving how the model should be used. Finally, create a compliance process that enables quick responses to new legal developments and regulatory requirements.

6.2. Ethical Control Recommendations

Develop a code of ethics for ChatGPT, as their output must be ethical, considerate, and refrain from compromise on discrimination and other inappropriate behaviors. Second, increase the transparency of ChatGPT so that users understand the limitations and biases inherent to the model. Create trust mechanisms for users to understand when interacting with ChatGPT. Third, ChatGPT-generated content must be diverse and inclusive, free of biases and discrimination. With this, the model should be encouraged to accept diversified training data and increase its inclusiveness. Fourth, reiterate ChatGPT's social responsibility in not generating harmful or misleading information. Actively engage in social issues and promote the positive impact of technology. Lastly, protect user rights and interests and provide for the security, privacy, and confidentiality of user information. Establish a complaints mechanism for users to give their input and receive timely follow-up [9].

6.3. Future Trends

ChatGPT has the potential to develop into personalization, producing customized content according to user needs and preferences to achieve more personalized services. It may also integrate multimodal data, such as text, image, audio, etc., so as to realize a richer interactive experience and content generation capability. The introduction of augmented learning technology will allow ChatGPT to improve itself continuously while learning from the interaction done with the environment, thus increasing quality and diversity in generated content. It will also ramp up the attention it is paying to the assessment of social impact by ChatGPT on society, economy and culture, so that it actively steers the development of the technology towards a good direction. A self-supervised learning mechanism will be introduced for ChatGPT so that self-correction and adjustment can be done in order to improve upon the accuracy and adaptability of generated content.

7. Conclusion

Generative AI legislation lacks a high degree of predictability and clarity, and there is an urgent need to build a fair generative AI system, determine transparent and trustworthy algorithmic evaluation standards, and improve the protective measures of generative AI legislation to encourage innovation and orderly development. In practice, we can learn from the EU's Artificial Intelligence Law, which implements precise regulation by finely dividing the risk level of AI systems and comprehensively considering the scope of each risk; and establishes a heavy penalty mechanism to promote the standardized use of AI.

In the future, it is necessary to improve laws and regulations to clarify the rights, obligations and responsibilities of AI such as ChatGPT. At the same time, technical supervision should be strengthened and an effective content review mechanism should be developed. In addition, the deep integration of AI and law should be promoted, such as the development of legal-specific AI tools to enhance judicial efficiency and fairness and promote the intelligent development of the legal industry.

Acknowledgements

Supported by a grant from the Guangdong Innovation and Entrepreneurship Training Programme for Undergraduates, ‘The AI that Causes Fire and Trouble:Research on Countermeasures for the Legal and Ethical Dilemmas of ChatGPT (Project No. S202310566079)

References

[1] Yu Xingzhong, Zheng Ge, Ding Xiaodong. Six issues in generative artificial intelligence and law:The case of ChatGPT [J]. China law review,2023,(02):1-20.

[2] Wang, Chunhui. The Ethical Issues of ChatGPT Should Be Highly Concerned[J]. China Telecommunications Industry, 2023,(06):56-58.

[3] Zhao Jingwu, ChatGPT Review and Reflection on Ethical Rules of Science and Technology in Risk Governance[J]. Journal of Sustainable Development Economics,2023,(04):28-30.

[4] Artificial Intelligence Index Report 2024[R/OL].(2024-04-26)[2024-04-30]. https://aiindex. stan-ford. edu/report/.

[5] Khasawneh O Y. Technophobia without Boarders: TheInfluence of Technophobia and Emotional Intelli-gence on Technology Acceptance and the ModeratingInfluence of Organizational Climate [J]. Computers inHuman Behavior, 2018(88): 210-218.

[6] Lavi, Michal. "Targeting Children: Liability for Algorithmic Recommendations." Am. UL Rev. 73 (2023): 1367

[7] Hou Li-Yang, Li Zhao-Xuan. Legal Challenges and Institutional Responses in the Academic Use of ChatGPT[J]. Journal of Northeast Normal University (Philosophy Social Science Edition), 2023,(04):29-39. DOI:10.16164/j. cnki.22-1062/c.2023.04.004.

[8] Zhou Shuhuan. Research on legal personality dispute and communication tort liability of chatbots: based on ChatGPT scene perspective [J]. NANJING SOCIAL SCIENCE, 2023,(06):90-99. DOI:10.15937/j.cnki.issn1001-8263.2023.06.010.

[9] Yuan Chen, Qiusheng Huang. Ethical Risks of Artificial Intelligence in ChatGPT Technology and Its Scientific Dispelling[J]. Journal of Hunan University of Science and Technology (Social Science Edition), 2023.

[10] LIN Xin, LIU Zefei. Quality assessment and application strategy of ChatGPT generated reviews[J]. Library and Intelligence Work, 2024.