Jus Corpus Law Journal

Open Access Law Journal – Copyright © 2023 – ISSN 2582-7820
Editor-in-Chief – Prof. (Dr.) Rhishikesh Dave; Publisher – Ayush Pandey

Algorithm Guardian: Legal Monitoring and ChatGPT

Kavya Sanjay Singh[a]

[a]National Law University, Delhi, India

Received 10 July 2023; *Accepted* 04 August 2023; *Published* 07 August 2023

The present article analyses the evolution and expansion of ChatGPT, an artificial intelligence (AI) technology developed by OpenAI, and its implications for data protection and legal compliance, particularly from an Indian perspective. It discusses the rapid growth of ChatGPT and its increasing usage in various industries, highlighting concerns about data privacy violations and the absence of comprehensive data protection legislation in India. The article explores OpenAI's privacy policy and the challenges it poses to data confidentiality and handling personal information. It also delves into the European Union's General Data Protection Regulation (GDPR) as a model for data protection laws and its potential influence on Indian legislation. The legal limitations and the need for reform in India are evaluated, emphasizing the importance of striking a balance between innovation and data protection. The article briefly discusses the Personal Data Protection (PDP) Bill of 2023 and its provisions related to data processing while noting the need for further amendments. It concludes by emphasizing the significance of establishing a robust data security regime, implementing robust regulations, and fostering international cooperation to protect personal data, ensure user privacy, and combat cyber threats in India's AI landscape.

Keywords: *artificial intelligence, data protection, privacy, gdpr, personal data protection bill, data security.*

INTRODUCTION: EVOLUTION AND EXPANSION OF CHATGPT

ChatGPT is a recent development in the area of artificial intelligence (AI) that was released by OpenAI, a company founded by Elon Musk. The technology has been utilised in order to simplify works relating to writing, which involves drafting emails, essays, research and even

writing codes.[1] The project was a research preview, and subsequently, the use cases of this technology are to be expanded. The project also provides a subscription plan in the form of ChatGPT Plus, which incorporates better features and claims a more user-efficient mechanism.

The arrival of this technology has led to generative AI in workplaces (Amazon, Salesforce, and Oracle) to increase efficiency. However, two months after its launch, the estimates suggest that ChatGPT has more than 30 million users and receives approximately five million visits daily. As a result, it is one of the fastest-growing software products in memory. Even Instagram took almost a year to reach a 10 million user base.[2]

The outcome here has paved the way for the ChatGPT API waitlist. The GPT-4 model is an improved version of previous versions that can not only comprehend natural language or code but also generate it.[3] Such expansion highlights the underlying issues associated with the use of this technology. There have been allegations of violations of the personal data protection of users. The data shared with the interface under these allegations could be utilized by OpenAI without adhering to the proper procedural requirements of notice, consent, and just compensation.[4]

THE LEGAL SAFETY NET: EVALUATING CHATGPT'S CLAIMS IN DATA PROTECTION AND LEGAL COMPLIANCE

OpenAI acknowledges the limited knowledge of the platform regarding events post-2021; hence, the technology can produce a biased response. The expansion of the platform has raised accuracy concerns along with issues regarding data protection. Under the current model, OpenAI's AI trainers monitor and review the user conversation to utilise such research and

[1] 'Introducing ChatGPT Plus' (*OpenAI*, 01 February 2023) <https://openai.com/blog/chatgpt-plus> accessed 08 July 2023

[2] Kevin Roose, 'How ChatGPT Kicked off an A.I. Arms Race' *The New York Times* (03 February 2023) <https://www.nytimes.com/2023/02/03/technology/chatgpt-openai-artificial-intelligence.html?searchResultPosition=1> accessed 08 July 2023

[3] 'OpenAI Platform' (*OpenAI*, 2023) <https://platform.openai.com/docs/models/overview> accessed 08 July 2023

[4] Catherine Thorbecke, 'OpenAI, Maker of Chatgpt, Hit with Proposed Class Action Lawsuit Alleging It Stole People's Data | CNN Business' *CNN* (28 June 2023) <https://edition.cnn.com/2023/06/28/tech/openai-chatgpt-microsoft-data-sued/index.html> accessed 08 July 2023

sustain futuristic goals that are aligned with the practical output. The data deletion process does not cater to the removal of specific prompts.[5]

ChatGPT's content (input and output) can be used for commercial purposes, including sale or publication, and the data given to it may be sensitive given the large-scale commercial expansion of this technology.[6] The data usage therefore can violate the regulations and privacy policies of such corporations. The matter becomes grave here and requires international consideration.

OpenAI classifies the information to be confidential once it qualifies the non-public criteria i.e. the information is not generally in the public realm. The definition of 'confidentiality' under their privacy policy is fairly narrow. The circumstances do not cover a wide range of possible scenarios. The terms do not explicitly address situations in which information is accidentally or unintentionally disclosed or situations in which information is shared within a restricted group under a non-disclosure agreement.

When using OpenAI's services to process personal data, users must follow applicable data protection laws, such as GDPR. This includes providing legally sufficient privacy notices and obtaining the necessary consent. For users processing personal data via the OpenAI API, OpenAI provides a Data Processing Addendum (DPA) that governs data processing activities.[7]

There is concern that the massive collection and processing of personal information to 'train' the algorithms on which the platform is based lacks a legal basis. The connected issues also highlight the absence of proper compliance mechanisms in the form of data security laws in India. India currently lacks comprehensive legislation for personal data protection on par with the GDPR. However, India must establish a strong data security regime and regulations to protect user privacy. This includes strict privacy policies for businesses, ethical codes of conduct for users, and severe penalties for violations. To address the global nature of cybercrime, international cooperation is required for information sharing and prosecuting transnational offenders. These

[5] Natalie, 'What Is Chatgpt?' (*OpenAI Help Center*) <https://help.openai.com/en/articles/6783457-what-is-chatgpt&gt> accessed 08 July 2023
[6] 'Terms of use' (*OpenAI*, 14 March 2023) <https://openai.com/policies/terms-of-use> accessed 08 July 2023
[7] *Ibid*

safeguards are critical for protecting personal information, ensuring user privacy, and effectively combating cyber threats.

UNDERSTANDING DATA PROTECTION IN LIGHT OF AI USAGE AND ITS IMPLICATIONS

Data protection laws typically focus on safeguarding natural persons' data, often overlooking the data protection needs of juristic persons. Furthermore, data processing across multiple locations raises jurisdictional concerns, necessitating an examination of extra-territorial applications. Legislation must ensure transparent data processing procedures and hold responsible entities or individuals accountable for data security. The laws should include the necessary mechanisms to address these issues and effectively uphold data protection standards.[8]

The framework should recognise and allow for the lawful processing of personal data based on legitimate interests, public interests, significant interests, and other relevant grounds. However, unless justified by these grounds, obtaining data subjects' consent is required for the processing of personal data, whether using AI or any other method. Individuals' explicit consent should be prioritised in the framework to ensure that their personal data is processed in a transparent and lawful manner. While ChatGPT recommends that children under 18 avoid using their interface without prior permission from a guardian, we witness that these guidelines are rarely followed. Hence, before the A.I. processes such data provided in such cases, the same must be evaluated and consented to by the children's guardian or parents. The legal framework also must acknowledge and critically engage with these circumstances. The government has also set up a task force under the National Mission for Artificial Intelligence. This project will therefore help in monitoring and coordinating all such technological developments that utilise AI by taking on the role of a nodal agency. As the AI mission expands AI deployment in India, it is critical to assess the proportionality of AI advancements and data protection requirements. The goal is to strike a balance between encouraging innovation and ensuring the legality of data protection

[8] D Majumdar and HK Chattopadhyay, 'Emergence of AI and Its Implication towards Data Privacy: From Indian Legal Perspective' (2020) 3(4) International Journal of Law Management & Humanities <https://www.ijlmh.com/emergence-of-ai-and-its-implication-towards-data-privacy-from-indian-legal-perspective/> accessed 08 July 2023

measures. As AI technologies are used in a variety of industries, this includes putting in place a comprehensive legal framework to protect personal data.[9]

Data protection and privacy rules are designed to give individuals the freedom to choose how their personal information is collected, used, and disclosed. While consent is the essence of data protection here, there may be grounds under which data processing is carried out irrespective of the person's consent. Further, modern-day data protection is also governed by the FIPPS (Fair Information Practice Principles). The FIPPS principles include that personal data recording systems should not be kept secret in order to ensure transparency; individuals should be able to access and comprehend the information stored in records about them; individuals should have control over how their information is used and shared, as well as the ability to prevent unauthorised use for a variety of purposes; and individuals should be able to correct or update their personally identifiable information in records.[10] OECD guidelines also harmonised the understanding of privacy eventually, and they were modified to include data security breach notification, thereby enhancing accountability regarding personal data.[11] These developments have however been considered incompatible in light of modern AI developments.[12]

EXAMINING THE LIMITATIONS AND THE NEED FOR LEGAL REFORM

The EU GDPR (European Union General Data Protection Regulation 2016) was developed as a data protection legal framework in response to the rapidly evolving data landscape. It is regarded as one of the world's strictest data protection regulations.

Data protection is carried out via a right-based approach here. An extensive set of individual participation rights are implemented by the EU General Data Protection Regulation (GDPR),

---

[9] 'National Mission for Artificial Intelligence' (*India Science, Technology & Innovation*, 2020) <https://www.indiascienceandtechnology.gov.in/st-visions/national-mission/national-mission-artificial-intelligence> accessed 08 July 2023

[10] Pam Dixon, 'A brief introduction to fair information practice principles' (*World Privacy Forum*, 05 June 2006) <https://www.worldprivacyforum.org/2008/01/report-a-brief-introduction-to-fair-information-practices/> accessed 08 July 2023

[11] 'Thirty Years After: The OECD Privacy Guidelines' (*OECD*, 2011) <http://www.oecd.org/sti/ieconomy/49710223.pdf> accessed 08 July 2023

[12] 'Big Data, Artificial Intelligence, Machine Learning and Data Protection' (*Information Commissioner's Office*) <https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf> accessed 08 July 2023

enabling people to retain control over their personal data even after it has been gathered. These rights give individuals the right to find out if personal data is being collected, access it, correct any errors, transfer it to another service provider, restrict processing, request erasure, object to processing, object to direct marketing, and object to automated judgments, as well as the right to object to processing in general. The GDPR creates an independent supervising authority with a variety of duties and powers to monitor compliance and defend basic rights. This organization ensures the free flow of data while defending the rights of individuals by monitoring and enforcing compliance with it. Additionally, the regulator has the authority to issue sanctions to guarantee efficient compliance. Many nations that have recently passed data protection laws now embrace the GDPR as a model. Variations of this concept, known as 'co-regulatory models', involve the collaboration of businesses and the government in the protection of personal information in nations like Australia and Canada.

It may be an important point of reference while taking into account how the EU GDPR has affected Indian legislation pertaining to the protection of personal data. The concepts and rights-based approach of the GDPR, which emphasises the significance of individual rights, data control, and the establishment of an independent regulatory authority, can influence the development of laws with a similar focus in India. Similar standards can be adopted by India to harmonise its data protection legislation with global best practices and improve the security of its residents' personal information. OpenAI was recently ordered by the Garante, the Italian data protection body, to stop processing the personal data of Italian users who use the chat service. This choice is motivated by worries about adhering to the EU General Data Protection Regulation (EU GDPR), and it has consequences that go beyond Italy, possibly affecting the usage of ChatGPT and related technologies in the EU and elsewhere.

Notification of a personal data breach involving ChatGPT led to Garante's inquiry into the site. They discovered various issues while evaluating GDPR compliance. It concluded that OpenAI had not effectively informed data subjects about the gathering of their personal information as required by GDPR Articles 13 and 14. Furthermore, the lack of age verification features in the ChatGPT signup procedure sparked worries about subjecting kids to objectionable material. Additionally, the Garante identified possible inconsistencies in the processed personal data and

emphasised the need for a distinct and well-defined legal foundation for data collection and processing by OpenAI. Garante's case shows the growing scrutiny that EU data protection authorities are giving to digital technology, particularly with regard to their adherence to GDPR rules. It emphasises how crucial it is to make sure that AI-based systems like ChatGPT comply with GDPR regulations. OpenAI has been given a deadline in which to address the identified GDPR violations. The Garante issued a second decision outlining specific requirements that OpenAI must meet in order for the injunction to be lifted and operations to resume in Italy. However, it is important to note that other regulators, both within and outside the EU, may take similar actions based on data subject rights concerns and other legal considerations. Companies face challenges due to the complex and fragmented nature of global regulations surrounding ChatGPT and similar technologies. Data protection requirements may differ across jurisdictions, necessitating careful navigation of diverse regulatory landscapes.[13]

Companies planning to use ChatGPT or similar AI technologies should carefully consider their data privacy obligations, update their privacy policies, conduct impact assessments, and stay up-to-date on evolving legal requirements. Compliance with GDPR and other data protection regulations should be a top priority for businesses considering using AI technologies such as ChatGPT. ChatGPT's compliance challenges in adhering to EU GDPR regulations highlight the importance of paying close attention to data protection obligations. Following GDPR requirements and staying up-to-date on regulatory developments will be critical for ensuring the responsible and legal use of AI technologies in the evolving digital landscape.[14]

The ChatGPT privacy policy outlines the different types of personal information that OpenAI collects. Account information, user content, communication information, and social media information are all included. It also explains how certain information, such as log data, usage data, and device information, is automatically collected. Users gain a better understanding of what information is being obtained by providing a clear overview of the data collected. The

[13] Rosa Barcelo et al., 'ChatGPT: A GDPR-Ready Path Forward?' (*Lexology*, 21 April 2023) <https://www.lexology.com/library/detail.aspx?g=2eaa58be-0fa0-471f-bbd2-1b4499f555f1> accessed 08 July 2023

[14] *Ibid*

usage is stated to analyse services, improve services, communicate with users, prevent fraud, and comply with legal obligations.

The policy then addresses the disclosure of personal information. Personal information may be shared with vendors, service providers, and affiliates to help with business operations and service delivery. Personal information may also be disclosed in the context of business transfers or as required by law. This disclosure information assists users in understanding under which scenario their personal information may be shared with third parties.

In India, the right to privacy includes the concept of 'informational privacy', which acknowledges the importance of protecting personal information from both state and non-state actors. Individuals have the right to informational privacy in order to protect their personal information and prevent its dissemination. It is important to note, however, that the right to privacy is not absolute and can be subject to reasonable limitations.[15]

The Indian court has established a test to determine the permissibility of state actions in order to limit the state's discretion in matters of privacy. According to this test, the action must be authorised by law and necessary to achieve a legitimate state goal; the extent of state interference must be proportionate to the need for such interference; and procedural safeguards must be in place to prevent state abuse of power. The court has identified the protection of national security, the prevention and investigation of crime, the encouragement of innovation and knowledge dissemination, and the prevention of the dissipation of social welfare benefits as legitimate aims of the state. These principles serve as the foundation for understanding data protection in India. In India, data protection laws and regulations seek to strike a balance between the individual's right to privacy and the legitimate interests of the state. Data protection measures must be implemented to ensure that personal data is collected, processed, and stored securely, with appropriate safeguards in place to prevent unauthorized access or misuse.[16]

---

[15] 'White paper of the committee of experts on a data protection framework for India' (*Ministry of Electronics & Information Technology*) <https://www.meity.gov.in/writereaddata/files/white_paper_on_data_protection_in_india_18122017_final_v2.1.pdf> accessed 08 July 2023

[16] *Ibid*

To effectively navigate the legal and regulatory framework and ensure compliance with applicable data protection laws, it is critical to consider these principles as well as the evolving landscape of data protection in India. The issue of confidentiality, as stated in the 'Terms and Conditions' of OpenAI's ChatGPT, raises concerns about the possibility of data protection violations as well as the broader implications of using such data as input. While the terms of OpenAI recognise the need for confidentiality, there are underlying issues and ambiguities that can arise in practice. One of the difficulties is that data entered as input into ChatGPT may be used to violate a third party's data protection rights. This raises concerns about how personal data is handled and protected, particularly when it is shared within the platform. Looking at the wide range of data collected and processed by ChatGPT, privacy and data protection principles must be carefully considered. Furthermore, the 'Terms and Conditions' under OpenAI state that information already in the public domain, regardless of how sensitive it is, is not confidential.[17]

This brings up an important point about the growing use of generative AI technologies and their potential consequences. If one platform uses a third party's personal data, the broad privacy policies common to such platforms may classify that information as non-confidential, even if it contains sensitive personal data. This interpretation opens the door to a cascade of potential loopholes that could be exploited to gain access to and misuse sensitive personal data. These issues highlight the importance of strong data protection mechanisms and clear guidelines in the context of generative AI technologies like ChatGPT to ensure the confidentiality and privacy of personal data. It advocates for a thorough understanding and implementation of data protection laws and regulations, as well as open and user-friendly privacy policies. To address these concerns, a multifaceted approach involving both technological solutions and regulatory frameworks is required. To protect individuals' privacy rights and prevent the exploitation of personal data, it is critical to strike a balance between innovation and data protection. It advocates for ongoing dialogue, collaboration, and a collaborative effort on the part of technology providers, policymakers, and users to establish effective safeguards and mitigate the potential risks associated with the use of personal data in generative AI platforms. However, it

[17] Terms of use (n 6)

is highlighted in the latest draught of the Personal Data Protection (PDP) Bill 2023 that generative artificial intelligence (AI) platforms such as ChatGPT or Google's Bard may be unable to process the personal data of Indians available in the public domain.[18]

The transfer of data will be governed by Sections 33 and 34 under Chapter VII of the Personal Data Protection Bill. This provision places a prohibition on the processing of sensitive personal data and critical personal data beyond the Indian Territory, even if such data is transferred beyond the Indian Territory. The processing of such data, however, is subject to the consent of the individual whose data is in question. However, certain sets of data that are considered critical personal data by the central government will only be processed within Indian borders. While the central government has a lot of control over the entire mechanism, the Justice B.N. Srikrishna Committee[19] also highlighted that certain amendments are required within the present draft. There is also a need to incorporate the proportionality test under the provisions prescribing personal data processing, and the same must not be driven by mere necessity when the data is processed without the consent of the individual, as stated even in the joint committee report on the bill.[20]

CONCLUSION AND ENSURING DATA SAFEGUARDS IN INDIA'S AI LANDSCAPE

In conclusion, the rapid growth and adoption of AI technologies such as ChatGPT in India have resulted in significant advancements in a variety of sectors. However, the proliferation of these technologies has raised concerns in India about data protection and legal compliance. The absence of comprehensive personal data protection legislation in India comparable to the EU GDPR highlights the need for a strong data security regime and regulations to protect user privacy. The most recent draft of the Personal Data Protection (PDP) Bill 2023 addresses some of these concerns by imposing limitations on the processing of personal data, particularly sensitive and critical personal data, even if it is publicly available. However, amendments are

[18] Suraksha P, 'Personal Data of Indians in Public Domain May Get Shielded from AI' *The Economic Times* (17 July 2023) <https://economictimes.indiatimes.com/tech/technology/personal-data-of-indians-in-public-domain-may-get-shielded-from-ai/articleshow/101805843.cms> accessed 08 July 2023

[19] Committee of Experts under the Chairmanship of Justice B.N. Srikrishna, *Report on A Free and Fair Digital Economy Protecting Privacy, Empowering Indians* (2018)

[20] Joint Committee Report on the Personal Data Protection Bill 2019

still required to ensure that the provisions adhere to the proportionality and consent principles emphasised by the Justice B.N. Srikrishna Committee.

Furthermore, initiatives such as the National Mission for Artificial Intelligence and the formation of a task force by the Indian government demonstrate recognition of the importance of monitoring and coordinating technological developments that use AI. These efforts should include a focus on harmonising data protection legislation, promoting strict privacy policies for businesses, governing the terms and policies of artificial intelligence technology such as generative AI and ensuring severe penalties for violations. International cooperation is also essential in dealing with the global nature of cybercrime and enabling effective information sharing and prosecution of transnational offenders. India's legislation in such matters must strike a balance between encouraging innovation and safeguarding individual privacy rights. To establish a strong legal framework and implement effective data protection mechanisms, technology providers, policymakers, and users must work together. By adopting best practices from global standards such as the EU GDPR and addressing the specific challenges and requirements of the Indian context, India can ensure the responsible and lawful use of AI technologies while protecting its residents' privacy and data security.