The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

CHATGPT AND DATA BREACHES: LEGAL LIABILITY FOR LLM - INDUCED THIRD-PARTY PRIVACY VIOLATIONS UNDER THE DPDP ACT, 2023

~ Ms. Aashi Lalchand Khanchandani

ABSTRACT

Generative large language models, such as ChatGPT, are trained and continue to generate substantial amounts of personal information about individuals who have not consented to the model's developer. This structural feature exposes a gap in India's Digital Personal Data Protection Act, 2023 which relies on a fiduciary processor framework that presumes a bilateral relationship between a data principal and a data fiduciary. This article examines whether the DPDP Act can effectively assign liability when a Large Language Model discloses memorised personal data or fabricates false information about a named third party who has never interacted with the system. Drawing on global enforcement experience, including Italy's since-annulled fine against OpenAI, multiple complaints filed by the privacy group noyb under the General Data Protection Regulation, the Samsung trade-secret disclosure, and the Georgia defamation suit Walters v. OpenAI, this article argues that the DPDP Act's consent-centric design, ambiguous breach-notification requirements, and absence of data accuracy provisions leave third-party victims of LLM-induced disclosure largely unprotected. The analysis compares the DPDP Act with the GDPR and the EU Artificial Intelligence Act's data-governance provisions, and proposes targeted statutory and regulatory reforms, such as activating the Significant Data Fiduciary mechanism for foundation-model providers and introducing a limited third-party accuracy right.

Keywords: *Digital Personal Data Protection Act, 2023, Large Language Models, Data Fiduciary, AI Hallucination, Third-Party Privacy.*

I. INTRODUCTION

When a user in Mumbai submits a query to ChatGPT, the response may incorporate, reproduce, or fabricate personal information about individuals who have never accessed the application, reviewed its privacy notice, or provided consent.[1] Large Language Models *(hereinafter referred to as "LLM")* are trained on datasets collected from the internet and refined through extensive user interactions. This process frequently embeds the names, biographies, and private details of identifiable individuals within the model parameters. Three categories of incidents illustrate the resulting risks to third parties. *First*, in 2023, employees of Samsung's semiconductor division entered confidential source code and meeting transcripts into ChatGPT, transferring proprietary and personal information beyond their employer's control and into a system that, by default, retained these inputs for further training.[2] *Second*, the European privacy group noyb filed multiple complaints against OpenAI before Austrian, Polish, and Norwegian data protection authorities after ChatGPT fabricated a criminal conviction, an inaccurate birth date, and other false biographical details about named individuals who had never used the service.[3] *Third,* an American radio host initiated a defamation lawsuit against OpenAI after ChatGPT generated a false embezzlement allegation. This litigation tested and, under Georgia law, ultimately rejected the proposition that an AI developer can be held liable for fabricated model output.[4]

---

[1] noyb, *ChatGPT Provides False Information About People, and OpenAI Can't Correct It* (Apr. 29, 2024), https://noyb.eu/en/chatgpt-provides-false-information-about-people-and-openai-cant-correct-it (last visited June 18, 2026)

[2] *See* Samsung ChatGPT Data Leak (2023): What Was Leaked & How, *Authentech* (2026), https://authentech.ai/shadow-ai/samsung-chatgpt-incident/ (last visited June 18, 2026); Lewis Maddison, *Samsung Workers Made a Major Error by Using ChatGPT*, *TechRadar* (Apr. 4, 2023), https://www.techradar.com/news/samsung-workers-leaked-company-secrets-by-using-chatgpt (last visited June 18, 2026).

[3] *See* noyb, *AI Hallucinations: ChatGPT Created a Fake Child Murderer* (Mar. 20, 2025), https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer (last visited June 18, 2026); Natasha Lomas, *ChatGPT Hit with Privacy Complaint over Defamatory Hallucinations*, *TechCrunch* (Mar. 20, 2025), https://techcrunch.com/2025/03/19/chatgpt-hit-with-privacy-complaint-over-defamatory-hallucinations/ (last visited June 18, 2026).

[4] *Walters v. OpenAI, L.L.C.*, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025); *see also Walters v. OpenAI, L.L.C.*, *Loeb & Loeb LLP* (May 30, 2025), https://www.loeb.com/en/insights/publications/2025/05/walters-v-openai-llc (last visited June 18, 2026).

Recent scholarship examining the intersection of artificial intelligence and data protection has predominantly focused on lawful bases for processing, transparency obligations, and the use of personal data in model training. However, comparatively little attention has been devoted to a distinct category of privacy harm: situations in which generative AI systems disclose, reconstruct, or fabricate personal information about individuals who have not interacted with the system. The Digital Personal Data Protection Act, 2023 *(hereinafter referred to as "DPDP" Act)* establishes a consent-centric framework that presumes a direct relationship between the data principal and the data fiduciary. [5]This article argues that such a presumption is increasingly unsustainable in the context of large language models capable of generating information about non-users.

India's DPDP Act was developed during the rapid proliferation of generative AI. However, its provisions reflect an earlier paradigm in which a data fiduciary collects personal data directly from, and owes notice and consent obligations to, an identifiable data principal.6 This examines whether this legal framework adequately addresses the unique ways in which large language models (LLMs) affect the privacy of third parties. It also considers potential reforms to address the resulting liability gap.

II. RESEARCH QUESTIONS

This article examines three questions.

First, does an LLM developer such as OpenAI qualify as a "Data Fiduciary", a "Data Processor", or neither when training a foundation model on personal data from individuals who have not provided consent?[7]

Second, do the Act's security-safeguard and breach-notification duties apply to disclosures generated by a model's authorised output, including memorisation of training data or hallucination of false personal details, or only to unauthorised-access incidents?[8]

[5] Usha Tandon & Neeraj Kumar Gupta, *Informational Privacy in the Age of Artificial Intelligence: A Critical Analysis of India's DPDP Act, 2023*, 2 L. & Digital Age 87 (2025); *see also Data Protection and AI Under the Digital Personal Data Protection Act, 2023: An Indian Perspective* (2025).
[6] Digital Personal Data Protection Act, No. 22 of 2023, §§ 2(i), 2(k), 5–6, India Code (2023)
[7] DPDP Act § 2(i).
[8] DPDP Act § 8(5)–(6).

Third, can the Act's fixed-penalty enforcement regime effectively deter, or remedy harm caused by foundation models whose training and inference primarily occur outside India?[9]

III. STATUTORY STRUCTURE AND KEY PROVISIONS OF THE DPDP ACT, 2023

The DPDP Act establishes liability for two primary categories of actors. A "Data Fiduciary" refers to any person who, individually or jointly, determines the purpose and means of processing personal data and remains responsible for compliance, irrespective of any contrary agreement, including those with a "Data Processor". A "Data Processor" is defined as an entity that processes personal data on behalf of the fiduciary.[10] The Act does not impose direct statutory obligations on processors; instead, accountability is routed exclusively through the fiduciary, who must ensure compliance via contractual arrangements.[11] Processing is permitted primarily based on consent that is free, specific, informed, unambiguous, and demonstrated by a clear affirmative act, following an itemised notice describing the data and its intended use.

The Chapter VIII establishes the Data Protection Board of India, which is empowered, following inquiry, to impose monetary penalties as specified in the Schedule. Penalties are capped at ₹250 crore for failure to implement reasonable security safeguards and ₹200 crore for failure to notify the Board and affected data principals of a personal data breach. Appeals may be made to the Telecom Disputes Settlement and Appellate Tribunal.[12] Section 10 authorizes the Central Government to designate certain fiduciaries as "Significant Data Fiduciaries", subjecting them to enhanced obligations, including the appointment of a resident Data Protection Officer and the conduct of periodic impact assessments.

These designations are determined by factors such as the volume and sensitivity of data processed and the risk to data principals' rights. To date, no foundation-model provider has been

[9] DPDP Act § 33 & sched.
[10] DPDP Act §§ 2(i), 2(k), 8(1).
[11] *India's Digital Personal Data Protection Act 2023 Brought into Force*, *Hogan Lovells* (2025), https://www.hoganlovells.com/en/publications/indias-digital-personal-data-protection-act-2023-brought-into-force- (last visited June 18, 2026).
[12] DPDP Act §§ 5–6.

designated under this mechanism.[13] Section 3(b) extends the Act's extraterritorial application to processing outside India that relates to offering goods or services to data principals within India. This provision, in principle, applies to a foreign large language model developer offering a chatbot to Indian users. However, commentators note the absence of an operational enforcement mechanism, such as a mandatory local representative, to ensure compliance in practice.[14] The Act does not define a standard for anonymisation and does not grant individuals the right to demand correction of inaccurate personal data generated by a system with which they have no fiduciary relationship. The statutory correction right exists only between a fiduciary and its own data principal.[15] This omission is especially consequential in the context of generative artificial intelligence systems. In contrast to traditional databases, large language models can generate personal information about individuals who have not directly provided data to the system. As a result, such cases may fall outside the conventional fiduciary–principal framework contemplated by the DPDP Act.

IV. THE LLM LIABILITY GAP

Applying this framework to LLMs reveals three fundamental structural mismatches

A. Absence of a Fiduciary Relationship with Third-Party Victims

The DPDP Act assumes that the individual harmed by unlawful processing is the same person whose personal data was collected, processed, or disclosed. This assumption creates substantial challenges in the context of generative artificial intelligence. Large language models can generate false personal information about individuals who have never interacted with the platform and, consequently, have neither received notice nor provided consent. Complaints submitted by noyb against OpenAI illustrate this problem. For example, ChatGPT allegedly identified an individual as a convicted child murderer, despite the absence of any such criminal

[13] DPDP Act § 33 & sched.; *see also DPDP Act Penalties: Powers of the Data Protection Board*, *KSandK* (Oct. 14, 2025), https://ksandk.com/data-protection-and-data-privacy/penalties-adjudication-under-indias-dpdp-act-2023/ (last visited June 18, 2026).
[14] DPDP Act § 10; *see also Digital Personal Data Protection Act, 2023*, *Legal Service India* (Apr. 5, 2026), https://www.legalserviceindia.com/Legal-Articles/digital-personal-data-protection-act-2023/ (last visited June 18, 2026).
[15] DPDP Act § 3(b); *see also Extraterritorial Application in Data Privacy: Lessons for India's DPDP Act*, *CyberPeace Found.* (Mar. 6, 2025), https://cyberpeace.org/resources/blogs/extraterritorial-application-in-data-privacy-lessons-for-indias-dpdp-act (last visited June 18, 2026).

history. Because the affected individual did not establish a fiduciary relationship with OpenAI, DPDPAct's remedial architecture is limited in its ability to provide effective redress.[16]

B. Hallucinations and the Limits of Breach Notification

The Act conceptualizes a personal data breach primarily as incidents involving unauthorized access, disclosure, or acquisition of personal data. However, a large language model (LLM) may disclose personal information through its intended operation rather than through external compromise. This distinction is important. For example, while the March 2023 ChatGPT incident involved a software vulnerability that exposed user information, a hallucinated output does not entail unauthorized access. Nevertheless, the harm to individuals resulting from such outputs may be equally severe. Therefore, the statutory criteria for breach notification do not adequately address harms arising from model inference.[17]

C. Absence of an Accuracy Principle

A key distinction between the DPDP Act and the General Data Protection Regulation *(hereinafter referred to as "GDPR")* concerns data accuracy. Article 5(1)(d) of the GDPR requires that personal data be accurate and, where necessary, kept up to date. This provision has served as the legal basis for complaints related to AI-generated misinformation. In contrast, the DPDP Act does not impose a similar obligation. As a result, individuals affected by hallucinated outputs lack a clear statutory mechanism to request correction when the generated information concerns them but was not collected directly from them.[18]

---

[16] See noyb, *AI Hallucinations: ChatGPT Created a Fake Child Murderer* (Mar. 20, 2025), https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer (last visited June 19, 2026).; Natasha Lomas, *ChatGPT Hit with Privacy Complaint over Defamatory Hallucinations*, TECHCRUNCH (Mar. 20, 2025), https://techcrunch.com/2025/03/19/chatgpt-hit-with-privacy-complaint-over-defamatory-hallucinations/ (last visited June 19, 2026).

[17] Digital Personal Data Protection Act, No. 22 of 2023, § 8(5)–(6), India Code (2023); *see also* A Bug Revealed ChatGPT Users' Chat History, Personal and Billing Data, *Help Net Security* (Mar. 27, 2023), https://www.helpnetsecurity.com/2023/03/27/chatgpt-data-leak/ (last visited June 19, 2026).; ChatGPT Temporarily Banned by the Italian Garante, *Datenschutz-Notizen* (Apr. 4, 2023), https://www.datenschutz-notizen.de/chatgpt-temporarily-banned-by-the-italian-garante-3041656/ (last visited June 19, 2026).

[18] Regulation (EU) 2016/679, art. 5(1)(d), 2016 O.J. (L 119) 1, 36 [hereinafter GDPR]; *see also* noyb, *AI Hallucinations: ChatGPT Created a Fake Child Murderer* (Mar. 20, 2025), https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer (last visited June 19, 2026).; Natasha Lomas, *ChatGPT Hit with Privacy Complaint over Defamatory Hallucinations*, *TechCrunch* (Mar. 20, 2025), https://techcrunch.com/2025/03/19/chatgpt-hit-with-privacy-complaint-over-defamatory-hallucinations/ (last visited June 19, 2026).

The case of *Walters v. OpenAI, L.L.C.* exemplifies the limits of current legal remedies, as a radio host alleged ChatGPT falsely accused him of embezzlement. Although the claim was dismissed, the litigation highlighted challenges in attributing legal responsibility for AI-generated misinformation under traditional defamation frameworks.[19]

The significance of *Walters v. OpenAI* goes beyond dismissing an individual defamation claim. The decision shows the practical challenges of imposing liability for probabilistic outputs from large language models. The Georgia Superior Court ruled that a reasonable reader would not see ChatGPT's output as a reliable factual statement, especially given OpenAI's repeated warnings about inaccuracies and model limits. This case suggests traditional tort doctrines may offer limited remedies for harms from AI-generated misinformation. Consequently, data-protection frameworks may become more important for addressing reputational and privacy harms linked to hallucinated outputs.

V. COMPARATIVE ANALYSIS OF THE GDPR AND THE EU AI ACT

The European context is particularly instructive, as it highlights the limitations of even a highly developed regulatory regime. The GDPR defines personal data broadly as information relating to an identified or identifiable natural person, with its recitals clarifying that the key criterion is identifiability rather than accuracy. This interpretation enables regulators to classify hallucinated yet identifying outputs as personal data, thereby subjecting them to the regulation's rectification and erasure rights.[20] The GDPR's accuracy principle served as the foundation for noyb's actions against OpenAI in Austria and Norway.[21] However, enforcement has been inconsistent. For example, Italy's Garante fined OpenAI €15 million in December 2024 for processing personal data without a sufficient legal basis and for failing to report the March 2023 breach. This fine was subsequently annulled by an Italian court in 2026. Additionally, broader

[19] *Walters v. OpenAI, L.L.C.*, No. 23-A-04860-2 (Ga. Super. Ct. May 19, 2025); *see also* Georgia Court Dismisses Defamation Lawsuit Against OpenAI over ChatGPT Output, *Loeb & Loeb LLP* (May 30, 2025), https://www.loeb.com/en/insights/publications/2025/05/walters-v-openai-llc (last visited June 19, 2026).
[20] Regulation (EU) 2016/679, art. 5(1)(d), 2016 O.J. (L 119) 1, 36; *see also Walters*, supra note 4.
[21] GDPR art. 4(1), 2016 O.J. (L 119) at 33; GDPR recital 26.

taskforce investigations initiated across multiple member states in 2023 have not resulted in further final decisions.[22]

The EU AI Act establishes an additional regulatory layer specific to artificial intelligence, which is largely absent from the DPDP Act. Article 10 requires that training, validation, and testing datasets for high-risk systems be relevant, representative, and free of errors, and mandates documented governance practices to address data provenance and gaps. These obligations are intended to complement, rather than replace, GDPR compliance.[23] In contrast, India currently lacks a comparable data-governance mandate specifically tailored to AI training data. Consequently, the DPDP Act addresses generative AI, if at all, only through its general fiduciary obligations.

Issue	**DPDP Act**	**GDPR**	**EU AI Act**
Accuracy Principle	No express provision	Art. 5(1)(d)	Indirectly supported
Third-Party Correction Right	No	Limited	Limited
Training Data Governance	No	Limited	Article 10
AI-Specific Obligations	No	No	Yes
Local Representative Requirement	No	Yes	Yes

VI. RECOMMENDATIONS

This analysis shows the main challenge is not a lack of privacy regulation but that current rules do not fully address harms caused by foundation-model outputs. The recommendations below aim to close this gap while keeping the DPDP Act's core structure.

[22] *See supra* note 3.
[23] *Lewis Silkin*, supra note 17; Theodore Christakis, *Generative AI and GDPR Enforcement in Europe: A Lot of Noise, One Fine, Zero Survivors*, *Cross-Border Data Forum* (Mar. 19, 2026), https://www.crossborderdataforum.org/generative-ai-and-gdpr-enforcement-in-europe-a-lot-of-noise-one-fine-zero-survivors/ (last visited June 19, 2026).

1. Clarify, through rulemaking or Board guidance, that a "personal data breach" encompasses instances where a model memorizes or generates identifiable personal data, in addition to traditional unauthorized system access. The notification requirement should be adapted to reflect the operational characteristics of foundation models.

2. Employ the Significant Data Fiduciary mechanism to identify major foundation-model providers serving Indian users. Require these entities to maintain accountability within India and to regularly evaluate the provenance of their training data, in accordance with the data-governance standards outlined in Article 10 of the EU AI Act.

3. Create a specific right for people to correct or challenge inaccurate personal data produced by AI, even if they are not the original data principal. This right should apply when a generative system fabricates or misattributes personal data, and complaints can be handled by the Data Protection Board's current process.

4. Use rulemaking to clearly define who is responsible for what in layered AI systems. The foundation-model developer should be liable for base-model training, while the business that uses the model in its own service should be responsible for its deployment.

5. Make foreign AI developers who serve a certain number of Indian users appoint a representative based in India. This will help enforce the Act's rules beyond India's borders, like the GDPR's representative requirement.[24]

VII. CONCLUSION

The DPDP Act's fiduciary processor framework effectively governs bilateral data relationships but is less suited to address an LLM's ability to disclose, retain, or generate personal information about individuals outside any consent process. Incidents such as the Samsung disclosure, noyb's complaints about hallucinations, and the *Walters* and Garante proceedings show this is a recurring issue in generative AI deployment. Three principal shortcomings in the existing statutory framework have been identified. First, the DPDP Act presumes a direct fiduciary

[24] Regulation (EU) 2024/1689, art. 10, 2024 O.J. (L 1689); *see also Article 10: Data and Data Governance*, *Securiti* (Feb. 25, 2025), https://securiti.ai/eu-ai-act/article-10/ (last visited June 19, 2026).

relationship between the entity processing personal data and the individual affected. However, large language models challenge this by generating information about individuals who may have had no prior interaction with the system. Second, the Act's breach-notification provisions focus mainly on incidents involving unauthorized access, not disclosures from model inference. Third, unlike the GDPR, the Act lacks an independent accuracy principle to address hallucinated personal information. Addressing this gap will require the Data Protection Board to interpret the Act’s breach and accuracy provisions with intent, and for the Central Government to apply the Significant Data Fiduciary mechanism proactively. Otherwise, the DPDP Act risks regulating outdated data practices while neglecting those shaping the current landscape. If India’s privacy laws continue to view data harms only through the lens of direct relationships, they may become less relevant to generative AI. In these systems, people affected by data disclosures can include those who never interacted with the technology.