INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 14 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

RE-IMAGINING DATA PRIVACY IN THE AGE OF GENERATIVE AI: CHALLENGES, RISKS, AND REGULATORY GAPS

AUTHOR - DR. PURANJAN PRASAD PAUL* & MISS MADHURIKA DEY**

* ASSISTANT PROFESSOR AT THE ICFAI UNIVERSITY TRIPURA

** RESEARCH SCHOLAR AT THE ICFAI UNIVERSITY TRIPURA

BEST CITATION - DR. PURANJAN PRASAD PAUL & MISS MADHURIKA DEY, RE-IMAGINING DATA PRIVACY IN THE AGE OF GENERATIVE AI: CHALLENGES, RISKS, AND REGULATORY GAPS, INDIAN *JOURNAL OF LEGAL REVIEW (IJLR),* 5 (14) OF 2025, PG. 1073-1078, APIS – 3920 – 0001 & ISSN - 2583-2344.

Abstract

Generative AI, especially large language models and multimodal generative systems, has transformed content creation, research, and enterprise workflows. Nonetheless, these systems present new and heightened risks to personal data and privacy. This paper synthesises technical privacy harms, including memorising and data leakage, inference, and model inversion, as well as dataset provenance issues. This section evaluates the effectiveness of current legal frameworks, such as the GDPR, U.S. sectoral enforcement, and the EU AI Act/EDPB guidance, in addressing these harms or their shortcomings in doing so. Regulatory gaps persist. The paper, drawing on recent technical and policy literature, argues that conventional data protection regulations and new AI frameworks fail to sufficiently tackle the accountability loop for generative systems. A hybrid strategy is proposed to address this issue. This strategy integrates technical mitigations, including differential privacy, provenance and data lineage, robust access controls, audit logging, and watermarking of synthetic outputs, alongside regulatory reforms. The reforms encompass clarified responsibilities for modellers and data controllers/processors, mandatory dataset provenance, requirements for model transparency, cooperation in cross-border enforcement, and established liability rules. The objective is to develop a regulatory-technical framework that safeguards individuals while maintaining advantageous innovation.

Keywords: Generative AI, Privacy Risks, U.S.Enforcement

1. Introduction

By analysing large datasets and understanding statistical patterns, generative artificial intelligence models are able to generate fluent text, graphics, codes, and other forms of media. By reproducing verbatim training examples, by allowing attribute inference about individuals, or by enabling re-identification when outputs are combined with external information, for example, the same training and deployment practices that enable these capabilities can expose personal and sensitive data in unexpected ways. Such breaches can happen for a number of reasons. Recent technological research has documented the process of memorisation and leaking from LLMs. At the same time, legislative activity all around the world, most notably the European AI Act and EDPB advice, is attempting to mitigate these dangers. At the same time, there are significant gaps between what models do, what the existing privacy law requires, and how regulators might hold model makers accountable for their actions

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 14 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

2. Privacy Risks of Generative AI

There are three characteristics of contemporary general artificial intelligence systems that make privacy issues more likely to occur and more damaging:

2.1 Scale and heterogeneity of training data: Web crawls, proprietary corpora, and federated enterprise data are used to train models through the training process. The likelihood that personal data will be included in the training set and that models will internalise and possibly recreate it is increased when scale is taken into consideration.

2.2 Memorisation and verbatim leakage: It is possible for LLMs to memorise uncommon training instances and then output them in their entirety when they are triggered. These behaviours can result in sensitive snippets being leaked even when they were not intended to be released. Such behaviour is a phenomenon that can be measured, according to empirical research.

2.3 New inference capabilities: Generative models can create realistic personal profiles from fragmented inputs. Attackers can exploit the outputs of these models to infer private traits, reassemble datasets, or carry out model-inversion assaults. The use of statistical regularities rather than direct retrieval is what these assaults are based on, which makes detection and rehabilitation more difficult. Given the combination of these characteristics, it may be concluded that traditional privacy safeguards, such as consent, deletion requests, and data minimisation during collection, are essential but not sufficient.

3. Classification of Generative AI privacy harms

3.1 Memory leakage

The most obvious concern is the reproduction of sensitive training examples verbatim, which includes things like private messages, credentials, and health information. Prompts or "jailbreaks" can trigger sequences copied from training data. Technical audits have revealed that LLMs, particularly for uncommon and distinctive sentences, may produce memorised sequences.

3.2 Re-identification and inferences to be drawn

Models are able to synthesise or infer private qualities about an identified or identifiable person (for example, political affiliation or health status) from aggregated patterns in the training data, even if the outputs are not perfect replicas of the originals. When these outputs are paired with those from other sources, they can make re-identification possible.

3.3 Uncertainty regarding the origin of the data's permission

It is common practice for training pipelines to combine datasets that have varying consent terms. Users frequently can't determine whether or not their content was beneficial to the training of a model, and platforms frequently do not have robust provenance records to verify the legal basis for processing. Data-subject rights such as access, deletion, and objection are threatened as a result of this scenario.

3.4 Misuse and reuse of downstream resources can occur

High-stakes situations can utilise the model's outputs for surveillance, targeted manipulation, or automated decision-making techniques. Although downstream applications may breach expectations and regulatory rights, the initial training may have been conducted on authorised data.

4. Regulatory and normative landscape

4.1 General Data Protection Regulation and other comparable data protection regimes

Lawfulness, purpose limitation, data minimisation, openness, and data subject rights (access, deletion, and portability) are some of the broad concepts that the General Data Protection Regulation (GDPR) stipulates. When these are used for GenAI, however, they raise several problems that are not commonly asked: who is the "data controller" when a foundation

model is trained on heterogeneous datasets? How should deletion be operationalised when data already contributes to the parameters of the model? What constitutes "personal data" for a model output if it either contains or permits re-identification? Regulators have published interpretive guidance to address these concerns, but practical issues persist.

4.2 EU AI Act and EDPB guidance

Recent advice has addressed the topics of general-purpose artificial intelligence (GPAI) models and data protection interactions. The EU AI Act, which has been operative since 2024 and has tiered implementation, provides duties for AI systems based on specific risk categories. The supplementary opinion of the EDPB and the guidelines of the Commission provide clarification on anonymity and legitimate interest, although they do not answer all practical concerns (for example, the provenance of model training and the usability of deletion rights).

4.3 United States and other jurisdictions

The United States does not have comprehensive federal privacy legislation at the moment; rather, it relies mostly on sectoral laws and enforcement organisations (FTC). Because of this, there is a patchwork of state laws and regulatory measures that may penalise unfair or misleading applications; however, there are no consistent guidelines addressing AI-specific issues. Laws pertaining to artificial intelligence and data protection are being combined in several places on a trial basis.

5. Key regulatory gaps

Based on the existing legal and technical literature as well as the policy activities, the following gaps are extremely important:

5.1 Ambiguity over roles and responsibilities

Data collectors, labellers, model developers, and deployers are those who are involved in modern machine learning supply chains. Currently available frameworks have difficulty assigning "controller" vs "processor" roles and responsibilities to these actors, which results in a reduction in accountability.

5.2 Ineffectiveness of deletion and “right to be forgotten”

The erasure right provided by GDPR assumes that data stored can be deleted. In the case of models in which information is absorbed into parameters, the existing legal framework does not adequately provide explicit procedures for assuring successful erasure or technical alternatives (for example, "machine unlearning").

5.3 Insufficient provenance and documentation requirements

Often, obtained consent records and provenance metadata are absent from training datasets. There are no globally accepted standards for dataset lineage, model cards, or dataset registries that can feed enforcement or compliance, despite the fact that regulators want transparency.

5.4 Poor alignment between privacy principles and AI risk frameworks

The focus of certain artificial intelligence legislation is on safety and dangers, such as disinformation and vital infrastructure; however, these regulations do not completely incorporate data-subject-centric privacy duties, such as consent and purpose limitation, in a manner that is model-tailored. While the risk categories established by the AI Act are helpful, the operational bearings on the loss of personal data are still insufficient.

5.5 Enforcement fragmentation and cross-border challenges

Models trained on global data and hosted across multiple jurisdictions complicate enforcement. There may be difficulties for regulators in gaining access to evidence, requiring source documentation, or coordinating remedies across international borders.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 14 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

6. Technical mitigations, strengths and limits

The risk to one's privacy can be mitigated by a collection of technical solutions; nevertheless, none of these techniques can solve the problem on their own.

6.1 Differential privacy

This technique involves the addition of calibrated noise during the training or aggregation process to set provable boundaries on the amount of information that may be leaked about a single training example. When applied to certain models, it is feasible; nevertheless, it frequently diminishes utility and can be challenging to calibrate when applied to big generative models. For the purpose of making practical tradeoffs, DP can be paired with other approaches.

6.2 Machine unlearning and targeted removal

Algorithms for "unlearning" aim to eliminate the impact of specific records from a previously trained model. Currently available techniques can be computationally intensive and may not ensure complete erasure across complicated structures. This issue is a research subject that is currently being actively investigated.

6.3 Data provenance, labeling, and consent metadata

Robust metadata standards, such as dataset manifests, consent flags, and approved uses, enable the ability to conduct downstream compliance checks and audits. They require industry adoption and interoperability standards, although they are relatively low-tech in comparison to model innovation.

6.4 Output watermarking and traceability

Through the use of techniques such as digital fingerprints and hidden signals, which are used to watermark model outputs, attribution and detection of model-generated information can be supported. These techniques also assist in identifying instances of misuse and linking outputs back to a model and its governance settings. Adoption in action is still in its infancy.

6.5 Access control, shadow mode testing, and monitoring

The attack surface can be reduced, and suspect extraction attempts can be identified by restricting model access through the use of gated application programming interfaces (APIs), rate limitations, and query logging. However, logging raises concerns around data retention and privacy, despite the fact that it is necessary for forensic and regulatory evidence.

Techniques can lessen the risk, but they cannot completely remove it. In addition, they involve tradeoffs in terms of utility, cost, and interoperability; hence, they need to be accompanied by additional legal requirements and incentives.

7. Policy proposals: closing the gap

The technology specifications, documentation requirements, and enforceable duties in a pragmatic privacy governance architecture for generative artificial intelligence. The proposals that have been prioritised are listed below

7.1 Clarify legal roles across ML supply chains

The obligations of dataset curators, labelling providers, model trainers, and deployers should be defined by legislation or regulatory advice, and they should include clear liability for individuals who do not comply with the regulations. When this occurs, blame-shifting is reduced, and remedial avenues are improved.

7.2 Mandatory dataset provenance and consent records

At a minimum, a metadata manifest should be required for datasets that are used for training. This manifest should include information on the dataset's origin, licensing and consent terms, categories of personal data, retention policies, and a hashed audit trail. It is necessary for authorities and auditors to have the ability to examine these records while maintaining acceptable safeguards. This significantly aids both enforcement and subject requests.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 14 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

7.3 Practical data-subject remedies for models

To adapt erasure and access rights to model contexts, it is necessary to (a) mandate that providers accept verified targeted removal requests and (b) mandate disclosure of whether a person's data was used in training and offer pragmatic remediation (for example, model fine-tuning to reduce influence and model unlearning where it is feasible). It is necessary for regulatory guidelines to translate erasure into technical methods that can be implemented.

7.4 Risk-proportionate technical requirements

Regulators should impose minimum technological controls for high-risk models and contexts. These controls should include mandatory privacy risk assessments, the utilisation of data protection for sensitive datasets, the recording of training and inference data flows, and access limitations for models that have been trained on sensitive personal data. The European Union Artificial Intelligence Act, while providing a model for risk-based duties, should closely align with data protection regulations.

7.5 Transparency and standardised documentation

It should be mandatory to have standardised model cards and dataset statements that identify the provenance of training data, privacy risk mitigations (DP settings, redaction procedures), intended uses, and known constraints. These standard disclosures make it possible for regulators, auditors, and downstream users to make decisions based on accurate information.

7.6 Incentives for privacy-preserving innovation

Public grants, standardisation efforts, and procurement preferences can accelerate the use of the DP and unlearning approaches. The establishment of interoperability standards for metadata and dataset registries is something that regulators can do.

7.7 International coordination and enforcement cooperation

Regulatory bodies are required to develop cooperative mechanisms (such as mutual legal aid for investigations, shared compliance standards, and cross-recognition of audit reports) due to the fact that models and data are able to traverse international borders. Fragmented enforcement creates gaps that bad actors can exploit.

8. Implementation challenges and trade-offs

To effectively handle various tradeoffs, policymakers and practitioners must:

8.1 **Utility vs privacy:** It is possible for strong data protection or excessive redaction to have a negative impact on the performance of a model. Regulations must compensate for risks and facilitate flexible technical implementations, which audits validate.

8.2 **Transparency vs IP/security:** When it comes to intellectual property and security, full disclosure of datasets and model internals is beneficial for privacy supervision; yet, it may also expose trade secrets or make it easier for adversarial assaults to be waged. This can be balanced out with a tiered disclosure paradigm, which allows for confidential audit access under non-disclosure agreements.

8.3 **Enforcement feasibility:** Audits of large models and datasets necessitate substantial resources to ensure the viability of enforcement. Regulators require both technical capacity and cooperation from the commercial sector.

8.4 **Global competitiveness:** Companies operating in jurisdictions with less stringent regulations may be at a disadvantage if they are subject to excessively prescriptive regulations; consequently, worldwide alignment is essential to prevent regulatory arbitrage. Recent legislative actions demonstrate that different jurisdictions are striking a balance between safeguards and competitiveness in a variety of different ways.

Technical research has established that the use of targeted prompting during LLM memorisation can result in the verbatim leakage of sensitive training strings. This has prompted regulators and industry groups to offer recommendations for mitigations such as input filtering, DP training, and query monitoring. However, policy responses (EDPB views and EU AI Act guidelines) not only emphasise the necessity of documenting legitimate interests and reidentifiability analyses, but they also clarify that anonymous models fall outside the scope of the General Data Protection Regulation (GDPR) under specific circumstances. These interactions demonstrate that the existing tools and standards are complementary to one another, but they are not sufficient on their own since they lack better provenance and remedial methods.

10. Conclusion

Generative artificial intelligence confers transformative powers, but it also poses equally transformative concerns to privacy. The legal architecture that was intended for traditional data processing has a difficult time addressing leakage and inference concerns that are created by the distinctive behaviours of LLMs and multimodal generators. A combination of technical mitigations and regulatory clarity is required to close the gap. These include standardised dataset provenance, clear assignment of legal responsibility across machine learning supply chains, model-aware interpretations of data-subject rights, risk-based technical mandates, and international enforcement cooperation. It is recommended that policymakers steer clear of purely prescriptive technological rules and instead establish outcome-focused duties that encourage concrete privacy protections while still preserving innovation. When this is done, generative artificial intelligence will become more trustworthy and safer for both individuals and society.