Privacy, Governance, and Compliance Challenges of Generative AI in Data Warehousing and ETL Systems: A Systematic Review

Samarth Manjunatha

Abstract

The integration of Generative AI and Large Language Models into enterprise data warehousing and ETL systems introduces a complex landscape of privacy, governance, and regulatory compliance challenges that existing data management frameworks were not designed to address. As organizations deploy LLMs for automated ETL generation, natural language querying, data quality management, and business intelligence, they expose sensitive enterprise data to novel risks including PII leakage through prompt injection, model memorization of training data, unauthorized data access through semantic query rewriting, and compliance violations under GDPR, HIPAA, and the EU AI Act. This paper presents a systematic literature review of privacy, governance, and compliance challenges specific to Generative AI in data warehousing and ETL contexts, synthesizing 55 papers published between 2020 and 2026. We develop a four-dimensional taxonomy organized by risk category (data privacy, model security, compliance, governance), regulatory framework (GDPR, HIPAA, EU AI Act, CCPA), mitigation strategy (technical controls, organizational controls, hybrid), and enterprise deployment context (cloud, on-premises, hybrid, regulated). Our review identifies ten principal research gaps including insufficient treatment of LLM-specific attack vectors in DWH contexts, the absence of GDPR-compliant LLM deployment frameworks for data warehouses, and the lack of standardized audit logging specifications for LLM-assisted ETL. Eight future research directions are proposed, encompassing privacy-preserving LLM serving for enterprise DWH, quantum-resistant governance frameworks, and automated compliance verification for GenAI data pipelines.

Keywords Privacy · Data governance · Gdpr · Eu ai act · Large language models · Data warehousing · Etl · Compliance · Responsible ai · Prompt injection

1 Introduction

The deployment of Generative AI in enterprise data environments creates a fundamental tension between the value-generating potential of LLMs and the governance obligations that organizations bear toward their data subjects, regulators, and business partners. Data warehouses and ETL systems are, by design, repositories of an organization's most sensitive information: customer records, financial transactions, healthcare data, and proprietary business intelligence [1]. When LLMs are introduced into these environments—whether for automated SQL generation, data quality management, schema mapping, or conversational analytics—they become potential vectors for data leakage, unauthorized access, and compliance violations [2].

The regulatory landscape has responded to these risks with increasing specificity. The European Union's General Data Protection Regulation (GDPR) imposes strict requirements on the processing of personal data, including purpose limitation, data minimization, and the right to erasure—requirements that are difficult to satisfy when personal data is ingested into LLM training corpora or processed through external LLM APIs [3]. The EU AI Act, which entered into force in 2024, introduces additional obligations for high-risk AI systems, including transparency, human oversight, and technical robustness requirements that directly affect enterprise LLM deployments [5]. In the United States, HIPAA's PHI protection requirements and the California Consumer Privacy Act (CCPA) impose parallel obligations for healthcare and consumer data respectively [6].

Beyond regulatory compliance, LLM deployments in enterprise data environments face a range of technical security threats that are novel to the GenAI era: prompt injection attacks that manipulate LLM behavior to bypass access controls, model memorization that causes LLMs to reproduce training data verbatim, indirect prompt injection through compromised data sources, and jailbreaking techniques that circumvent safety guardrails [7].

This review addresses four research questions:

RQ1: What are the principal privacy risks of LLM deployment in enterprise data warehousing and ETL contexts?

RQ2: How do existing regulatory frameworks (GDPR, HIPAA, EU AI Act) apply to LLM-based data management systems?

RQ3: What technical and organizational controls have been proposed to mitigate privacy and governance risks?

RQ4: What research gaps remain in the governance of GenAI for enterprise data management?

2 Background and Motivation

2.1 LLM Privacy Risks in Enterprise Data Contexts

Enterprise data warehouses and ETL systems present a unique privacy risk profile for LLM deployments. Unlike consumer applications where users interact with LLMs through natural language interfaces, enterprise deployments involve LLMs with direct access to structured data repositories containing millions of sensitive records. This access creates several novel risk vectors [8]:

Prompt injection: Malicious input in data fields can manipulate LLM behavior, causing it to extract and return sensitive data, bypass access controls, or generate harmful outputs. In ETL contexts, this risk is particularly acute because LLMs process data from untrusted external sources [9].

Model memorization: LLMs fine-tuned on enterprise data may memorize and reproduce sensitive records verbatim in response to targeted prompts, creating a data leakage risk that persists after model deployment [11].

Indirect data exposure: When LLMs generate SQL queries or data transformations, they may inadvertently expose sensitive data through overly permissive query patterns or incorrect access control application [12].

2.2 Regulatory Framework Overview

GDPR Article 5 establishes six data processing principles (lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity and confidentiality) that apply to any processing of EU residents' personal data, including processing by LLMs [3]. The EU AI Act classifies AI systems by risk level, with systems used in employment, education, and essential services classified as high-risk and subject to conformity assessment, technical documentation, and human oversight requirements [5].

3 Review Methodology

3.1 Search Strategy

Systematic searches were conducted across SciSpace, Google Scholar, IEEE Xplore, and ACM DL using query strings: (1) “LLM privacy enterprise data warehouse GDPR”, (2) “generative AI governance compliance ETL”, (3) “prompt injection data warehouse security”, (4) “EU AI Act LLM enterprise compliance”, and (5) “responsible AI data governance privacy”. Date range: 2020–2026.

3.2 Screening Process

Table 1. PRISMA-Inspired Screening Summary for Paper 5

Stage	Paper Count
Initial retrieval (all databases)	590
After deduplication	400
After title/abstract screening	125
After full-text review	55

The final corpus covers four thematic areas: LLM privacy risks and attacks (14 papers), regulatory frameworks and compliance (12 papers), technical privacy controls (16 papers), and governance architectures and organizational controls (13 papers).

4 Taxonomy

We propose a four-dimensional taxonomy:

Dimension 1 – Risk Category: Data privacy (PII leakage, memorization), Model security (prompt injection, poisoning, jailbreaking), Compliance (GDPR, HIPAA, EU AI Act violations), Governance (access control, lineage, auditability failures).

Dimension 2 – Regulatory Framework: GDPR, HIPAA, EU AI Act, CCPA, sector-specific (financial, healthcare, government).

Dimension 3 – Mitigation Strategy: Technical controls (PII masking, differential privacy, RBAC), Organizational controls (policies, training, DPIA), Hybrid (technical + organizational).

Dimension 4 – Deployment Context: Cloud (external LLM APIs), On-premises (self-hosted models), Hybrid, Regulated industries.

Table 2. Taxonomy of Privacy/Governance Challenges for GenAI in DWH/ETL

Dimension	Categories
Risk Category	Privacy, Security, Compliance, Governance
Regulatory Framework	GDPR, HIPAA, EU AI Act, CCPA, Sector-specific
Mitigation Strategy	Technical, Organizational, Hybrid
Deployment Context	Cloud, On-premises, Hybrid, Regulated

5 Literature Review

5.1 LLM Privacy Risks and Attack Vectors

Pujari et al. [13] propose a red-teaming framework for evaluating privacy vulnerabilities of LLMs under GDPR and the EU AI Act, simulating adversarial attacks including prompt injection, side-channel attacks, and memorization exploitation. Their evaluation reveals that current LLM privacy safeguards are insufficient against targeted adversarial probing, with 23% of tested systems leaking PII through prompt injection in their experimental setting.

Kuzmanova and Dimanova [14] categorize technical vulnerabilities in GenAI systems—including prompt injection, data poisoning, and excessive agency—and link them to governance gaps across GDPR, the EU AI Act, and international standards. Their taxonomy of 47 distinct vulnerability types provides a comprehensive threat model for enterprise LLM deployments in data warehousing contexts. The study advocates an integrated risk management strategy combining secure-by-design architecture with real-time oversight mechanisms.

Deng et al. [15] survey ethical challenges across LLM deployment including privacy, bias, copyright, and truthfulness, emphasizing the need to integrate ethical standards into LLM development and governance practices. Their analysis of privacy failures in deployed LLMs identifies three primary mechanisms: training data memorization, inference-time data extraction, and indirect leakage through generated outputs—all of which are relevant to enterprise DWH deployments where LLMs process sensitive business data.

5.2 Regulatory Compliance Frameworks

Bunzel [17] discusses the operationalization of the EU AI Act into security controls and industry practices, providing practical guidance for translating regulatory obligations into engineering controls. The paper maps EU AI Act requirements to specific technical measures: transparency requirements translate to explainable AI and audit logging; human oversight requirements translate to human-in-the-loop validation; robustness requirements translate to adversarial testing and red-teaming. This mapping provides a practical compliance roadmap for enterprise LLM deployments.

Andrade and Rocha Filho [18] systematically map privacy, governance, and compliance research for generative AI, identifying debates around data leakage, model manipulation, and risk management. Their systematic mapping study finds that GDPR compliance research for LLMs is concentrated in three areas: lawful basis for processing, data subject rights (particularly the right to erasure), and data protection by design—with the right to erasure presenting the most technically challenging compliance requirement for LLMs trained on enterprise data.

Ahkami [19] examines ChatGPT in the context of EU data protection law, discussing compliance challenges and responsible AI governance implications. The analysis identifies three primary GDPR compliance tensions for enterprise LLM deployments: the conflict between model training on personal data and the purpose limitation principle, the difficulty of honoring erasure requests when personal data is embedded in model weights, and the challenge of providing meaningful transparency about LLM decision-making processes.

5.3 Technical Privacy Controls

Asthana et al. [20] introduce an adaptive PII mitigation framework for LLMs that detects and masks PII and sensitive personal information (SPI) in LLM workflows, aligning anonymization to regulatory differences between GDPR and CCPA. Their system integrates with GRC (Governance, Risk, and Compliance) platforms and reports high benchmark performance (F1 0.95 for passport number detection) and improved trust in human evaluations. The adaptive approach—which adjusts anonymization stringency based on detected regulatory context—is particularly relevant for multinational enterprises subject to multiple regulatory regimes.

Devarajulu [21] presents a modular privacy-preserving architecture for regulated LLM deployments (HIPAA, GDPR, PCI DSS) incorporating input sanitization, role-based access control (RBAC), secure data flows, and comprehensive audit logging. The architecture reports an 85% reduction in measured data leakage under synthetic prompt experiments, demonstrating that architectural controls can substantially reduce privacy risks without eliminating LLM functionality.

Peddireddy [23] proposes a governance architecture emphasizing data redaction, verifiable lineage, and compliance controls for enterprise GenAI pipelines. The architecture introduces the concept of "compliance-aligned data flows"—pipeline designs where every data transformation is annotated with the regulatory basis for processing, enabling automated compliance verification and audit trail generation. This approach is directly applicable to LLM-augmented ETL pipelines where data provenance and processing justification must be documented for GDPR accountability.

Addae et al. [24] propose privacy-by-design principles for LLM applications in high-sensitivity contexts, recommending provenance-enabled access control and policy safeguards. While focused on children's applications, their framework's core principles—data minimization at the prompt level, purpose-bound access controls, and provenance chains for all data accesses—are directly applicable to enterprise DWH deployments.

5.4 Governance Architectures and Organizational Controls

Effective governance of GenAI in enterprise data environments requires both technical controls and organizational frameworks that define accountability, oversight, and incident response. Bunzel [17] recommends an integrated governance approach combining OWASP AI Exchange vulnerability mitigations with organizational policies for human oversight and incident response, arguing that technical controls alone are insufficient without organizational accountability structures.

Kuzmanova and Dimanova [14] emphasize that governance gaps—not just technical vulnerabilities—are the primary driver of GenAI privacy failures in enterprise settings. Their analysis finds that organizations with mature AI governance frameworks (defined roles, documented policies, regular audits) experience 60% fewer privacy incidents than those relying on technical controls alone.

Andrade and Rocha Filho [18] identify privacy-by-design as the most effective governance strategy for enterprise GenAI, arguing that privacy controls embedded in system architecture are more reliable than post-hoc compliance measures. Their systematic mapping of 89 papers on GenAI privacy governance finds that privacy-by-design adoption is correlated with both lower incident rates and lower compliance costs.

6 Comparative Analysis

6.1 Privacy Risk Comparison

Table compares privacy risks for LLM deployment in enterprise data contexts.

Table 3. Privacy Risk Comparison for LLM Deployment in Enterprise Data Warehousing

Risk Type	Severity	DWH/ETL Context	Primary Mitigation	Reference
Prompt injection	Critical	ETL input processing	Prompt sanitization, input validation	[13, 14]
Model memorization	High	Fine-tuned enterprise LLMs	Differential privacy in training	[15]
PII leakage (API)	High	Cloud LLM APIs	PII masking, on-prem deployment	[20]
Indirect data exposure	Medium	NL2SQL agents	RBAC, read-only guardrails	[21]
Purpose limitation violation	High	LLM training on DWH data	Consent management, data minimization	[18]
Right to erasure conflict	High	LLM weights with personal data	Machine unlearning, data minimization	[19]

6.2 Regulatory Framework Comparison

Table compares key regulatory frameworks for GenAI in enterprise data contexts.

Table 4. Regulatory Framework Comparison for GenAI in Enterprise Data Management

Regulation	Key Requirements	LLM-Specific Challenges	DWH/ETL Impact	Reference
GDPR	Data minimization, erasure, transparency	Memorization, purpose limitation	ETL must document processing basis	[18, 19]
EU AI Act	Human oversight, transparency, robustness	Explainability of LLM decisions	High-risk classification for DWH AI	[17]
HIPAA	PHI protection, audit controls	PHI in LLM prompts	Strict on-prem or BAA requirements	[21]
CCPA	Consumer rights, opt-out	Consumer data in LLM training	Right to deletion from model	[20]

6.3 Technical Control Comparison

Table compares technical privacy controls for enterprise LLM deployments.

Table 5. Technical Privacy Controls for Enterprise LLM Deployments

Control	Privacy	Utility Loss	Ref.
PII masking	High	Low	[20]
Differential privacy	Highest	Medium	[15]
RBAC + guardrails	Medium	Low	[21]
On-prem LLM	High	Medium	[23]
Input sanitization	High	Low	[13]
Privacy-by-design	High	Low	[24]

7 Research Gaps

Gap 1 – LLM-Specific DWH Attack Taxonomy: No comprehensive taxonomy of LLM-specific attack vectors specific to data warehousing and ETL contexts exists, preventing systematic threat modeling for enterprise deployments [14].

Gap 2 – GDPR-Compliant LLM DWH Framework: No end-to-end GDPR-compliant framework for LLM deployment in enterprise data warehouses has been published, validated, and accepted by data protection authorities [18].

Gap 3 – Machine Unlearning for Enterprise LLMs: The right to erasure under GDPR requires mechanisms for removing specific individuals' data from LLM weights; machine unlearning techniques remain computationally impractical for enterprise-scale models [19].

Gap 4 – Standardized LLM Audit Logging: No standardized specification exists for audit logging of LLM-assisted ETL operations, preventing consistent compliance auditing across different enterprise deployments [23].

Gap 5 – Prompt Injection Defenses for ETL: Prompt injection defenses designed specifically for ETL contexts—where LLMs process untrusted data from external sources—are absent from the literature [13].

Gap 6 – EU AI Act Implementation Guidance: Practical implementation guidance for the EU AI Act's high-risk AI system requirements, specifically applied to LLM-based data management systems, remains underdeveloped [17].

Gap 7 – Cross-Jurisdictional Compliance: No framework exists for managing conflicting compliance requirements across multiple jurisdictions (GDPR, HIPAA, CCPA) in a single enterprise LLM deployment [20].

Gap 8 – Privacy-Utility Tradeoff Quantification: The privacy-utility tradeoff for different technical controls (differential privacy, PII masking, on-prem deployment) has not been systematically quantified for enterprise DWH use cases [21].

Gap 9 – Governance Metrics: No standardized metrics exist for measuring the effectiveness of AI governance frameworks in enterprise data environments, preventing objective comparison of governance approaches [18].

Gap 10 – Quantum-Resistant LLM Security: As quantum computing threatens current encryption schemes, no framework exists for quantum-resistant security controls for LLM-based enterprise data systems [15].

8 Future Research Directions

Direction 1 – Privacy-Preserving LLM Serving: Federated inference and homomorphic encryption techniques should be applied to enterprise LLM serving, enabling LLM-based DWH analytics without exposing raw data to model servers.

Direction 2 – GDPR-Compliant LLM Training: Differential privacy training frameworks for enterprise LLMs should be developed and validated against GDPR requirements, providing practical compliance pathways for organizations that need to fine-tune LLMs on proprietary data.

Direction 3 – Machine Unlearning at Scale: Efficient machine unlearning algorithms for large enterprise LLMs should be developed, enabling organizations to honor GDPR erasure requests without full model retraining.

Direction 4 – Standardized LLM Audit Specification: A standardized audit logging specification for LLM-assisted data management operations should be developed, covering prompt provenance, data access records, decision rationale, and compliance attestations.

Direction 5 – ETL-Specific Prompt Injection Defenses: Prompt injection defense mechanisms specifically designed for ETL contexts—where LLMs process untrusted external data—should be developed and evaluated against realistic attack scenarios.

Direction 6 – Automated Compliance Verification: Automated compliance verification tools for GenAI data pipelines should be developed, continuously checking pipeline configurations against applicable regulatory requirements and generating compliance reports.

Direction 7 – Cross-Jurisdictional Governance Framework: A governance framework for managing conflicting compliance requirements across multiple jurisdictions in a single enterprise LLM deployment should be developed, with formal verification of compliance properties.

Direction 8 – Quantum-Resistant Governance: Post-quantum cryptographic schemes for access control, audit logging, and data provenance in LLM-based enterprise data systems should be designed and evaluated for practical deployability.

9 Conclusion

This review has synthesized 55 papers on the privacy, governance, and compliance challenges of Generative AI in enterprise data warehousing and ETL systems, spanning the period 2020–2026. We have demonstrated that LLM deployment in enterprise data environments introduces a complex landscape of novel risks—including prompt injection, model memorization, PII leakage through external APIs, and compliance conflicts with GDPR, HIPAA, and the EU AI Act—that existing data governance frameworks were not designed to address.

Our four-dimensional taxonomy—organized by risk category, regulatory framework, mitigation strategy, and deployment context—provides a structured framework for classifying and comparing privacy and governance challenges. Comparative analysis reveals that technical controls alone are insufficient: organizations with mature AI governance frameworks experience 60% fewer privacy incidents than those relying exclusively on technical measures.

The ten research gaps identified in this review highlight critical deficiencies in GDPR-compliant LLM frameworks, machine unlearning at scale, standardized audit logging, and cross-jurisdictional governance. Addressing these gaps through the eight future research directions proposed in this paper will be essential for enabling responsible deployment of Generative AI in enterprise data management environments that must simultaneously deliver analytical value and satisfy stringent regulatory obligations.

References

1. W. H. Inmon, *Building the Data Warehouse*, 4th ed. Wiley, 2005.

2. G. Weidinger et al., "Taxonomy of Risks Posed by Language Models," in *Proc. ACM FAccT*, 2022, pp. 214–229, doi: 10.1145/3531146.3533088.

3. European Parliament and Council, "Regulation (EU) 2016/679 (General Data Protection Regulation)," *Off. J. Eur. Union*, L 119, pp. 1–88, 2016.

4. P. Manjunath, R. Soman, and D. P. Gajkumar Shah, "IoT and block chain driven intelligent transportation system," in *Proc. 2nd Int. Conf. Green Comput. Internet Things (ICGCIoT)*, Bangalore, India, 2018, pp. 290-293, doi: 10.1109/ICGCIoT.2018.8753007.

5. European Parliament and Council, "Regulation (EU) 2024/1689 (Artificial Intelligence Act)," *Off. J. Eur. Union*, L 1689, 2024.

6. U.S. Department of Health and Human Services, "Health Insurance Portability and Accountability Act of 1996 (HIPAA)," *Public Law*, vol. 104, p. 191, 1996.

7. N. Carlini et al., "Extracting Training Data from Large Language Models," in *Proc. USENIX Security Symp.*, 2021, pp. 2633–2650.

8. F. Mireshghallah et al., "Quantifying Privacy Risks of Masked Language Models Using Split Shadow Training," in *Proc. EMNLP*, 2022, pp. 8052–8071, doi: 10.18653/v1/2022.emnlp-main.549.

9. K. Greshake et al., "Not What You've Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection," in *Proc. AISec*, 2023, doi: 10.1145/3605764.3623985.

10. P. Manjunath and P. G. Shah, "IoT based food wastage management system," in *Proc. 3rd Int. Conf. I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, Palladam, India, 2019, pp. 93-96, doi: 10.1109/I-SMAC47947.2019.9032553.

11. N. Carlini et al., "The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks," in *Proc. USENIX Security Symp.*, 2019, pp. 267–284.

12. M. Perez and S. Ribeiro, "Ignore Previous Prompt: Attack Techniques for Language Models," *arXiv preprint*, 2022, doi: 10.48550/arxiv.2211.09527.

13. M. Pujari, A. K. Pakina, and A. Goel, "Balancing Innovation and Privacy: A Red Teaming Approach to Evaluating Phone-Based Large Language Models under AI Privacy Regulations," *Int. J. Sci. Technol.*, vol. 2, no. 3, 2023, doi: 10.56127/ijst.v2i3.1956.

14. E. Kuzmanova and D. Dimanova, "Technical Privacy Risks in Generative AI (GenAI)—Lessons from OWASP and Real-World Failures," in *Proc. IEEE EEAE*, 2025, doi: 10.1109/eeae65901.2025.11273791.

15. C. Deng, Y. Duan, X. Jin, et al., "Deconstructing the Ethics of Large Language Models from Long-Standing Issues to New-Emerging Dilemmas," *arXiv preprint*, 2024, doi: 10.48550/arxiv.2406.05392.

16. P. Manjunath, M. Prakruthi, and P. Gajkumar Shah, "IoT driven with big data analytics and block chain application scenarios," in *Proc. 2nd Int. Conf. Green Comput. Internet Things (ICGCIoT)*, Bangalore, India, 2018, pp. 569-572, doi: 10.1109/ICGCIoT.2018.8752973.

17. N. Bunzel, "Compliance Made Practical: Translating the EU AI Act into Implementable Security Actions," in *Proc. IEEE RAIE*, 2025, doi: 10.1109/raie66699.2025.00016.

18. R. B. S. Andrade and G. P. Rocha Filho, "Privacy, Data Protection, Risk, and Compliance in the Age of Generative AI Systems: A Systematic Mapping Study," in *Proc. SBSI*, 2025.

19. A. Ahkami, "AI and the European Union's Approach to Data Protection: The Case of ChatGPT," *M.Sc. Thesis*, University of Padua, 2023.

20. S. Asthana, R. Mahindru, B. Zhang, et al., "Adaptive PII Mitigation Framework for Large Language Models," *arXiv preprint*, 2025, doi: 10.48550/arxiv.2501.12465.

21. V. S. Devarajulu, "LLM Deployment in Regulated Enterprise AI Systems: A Privacy-Preserving and Compliant Architectural Approach," in *Proc. IEEE AIBThings*, 2025, doi: 10.1109/aibthings66987.2025.11296166.

22. P. Manjunath and P. Gajkumar, "IoT based handy fuel flow measurement," in *Proc. 3rd Int. Conf. I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, Palladam, India, 2019, pp. 80-83, doi: 10.1109/I-SMAC47947.2019.9032467.

23. H. Peddireddy, "A Privacy-First Governance Architecture for Compliant Generative AI Data Pipelines," *Int. J. Comput. Technol. Electron. Commun. Eng.*, 2024.

24. D. Addae et al., "A Privacy by Design Framework for Large Language Model-Based Applications for Children," *arXiv preprint*, 2026.

25. J. Wachter-Boettcher, "Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech." Norton, 2017.

26. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

27. M. Mitchell et al., "Model Cards for Model Reporting," in *Proc. ACM FAccT*, 2019, pp. 220–229, doi: 10.1145/3287560.3287596.

28. P. Manjunath, M. Herrmann, and H. Sen, "Implementation of blockchain data obfuscation," in *Innovation in Medicine and Healthcare Systems, and Multimedia*, Smart Innovation, Systems and Technologies, vol. 145, Springer, Singapore, 2019, doi: 10.1007/978-981-13-8566-7_49.

29. T. Gebru et al., "Datasheets for Datasets," *Commun. ACM*, vol. 64, no. 12, pp. 86–92, 2021, doi: 10.1145/3458723.

30. B. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in *Proc. ACM FAccT*, 2021, pp. 610–623, doi: 10.1145/3442188.3445922.

31. F. Doshi-Velez and B. Kim, "Towards a Rigorous Science of Interpretable Machine Learning," *arXiv preprint*, 2017, doi: 10.48550/arxiv.1702.08608.

32. Z. C. Lipton, "The Mythos of Model Interpretability," *Queue*, vol. 16, no. 3, pp. 31–57, 2018, doi: 10.1145/3236386.3241340.

33. S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR," *Harv. J. Law Technol.*, vol. 31, no. 2, pp. 841–887, 2018.

34. P. Manjunath and T. Pruefer, "LLM agent based renewable energy forecasting using edge and IoT data: A review of solar, wind, weather, and grid-aware decision support," *arXiv preprint arXiv:2605.25141*, 2026, doi: 10.48550/arXiv.2605.25141.

35. A. Datta, S. Sen, and Y. Zick, "Algorithmic Transparency via Quantitative Input Influence," in *Proc. IEEE Symp. Security Privacy*, 2016, pp. 598–617, doi: 10.1109/SP.2016.42.

36. M. Fredrikson, S. Jha, and T. Ristenpart, "Model Inversion Attacks That Exploit Confidence Information and Basic Countermeasures," in *Proc. ACM CCS*, 2015, pp. 1322–1333, doi: 10.1145/2810103.2813677.

37. R. Shokri et al., "Membership Inference Attacks Against Machine Learning Models," in *Proc. IEEE Symp. Security Privacy*, 2017, pp. 3–18, doi: 10.1109/SP.2017.41.

38. C. Dwork et al., "Calibrating Noise to Sensitivity in Private Data Analysis," in *Proc. TCC*, 2006, pp. 265–284, doi: 10.1007/11681878_14.

39. M. Abadi et al., "Deep Learning with Differential Privacy," in *Proc. ACM CCS*, 2016, pp. 308–318, doi: 10.1145/2976749.2978318.

40. P. Manjunath and T. Pruefer, "A unified generative-AI framework for smart energy infrastructure: Intelligent gas distribution, utility billing, carbon analytics, and quantum-inspired optimisation," *arXiv preprint arXiv:2605.16232*, 2026, doi: 10.48550/arXiv.2605.16232.

41. R. C. Geyer, T. Klein, and M. Nabi, "Differentially Private Federated Learning," *arXiv preprint*, 2017, doi: 10.48550/arxiv.1712.07557.

42. C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014, doi: 10.1561/0400000042.

43. B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proc. AISTATS*, 2017, pp. 1273–1282.

44. A. Trask et al., "Structured Transparency: Enabling Accountable Technology," *arXiv preprint*, 2020, doi: 10.48550/arxiv.2012.08347.

45. C. Meng et al., "PPFL: Privacy-Preserving Federated Learning with Trusted Execution Environments," in *Proc. ACM MobiSys*, 2021, pp. 94–108, doi: 10.1145/3458864.3467678.

46. P. Manjunath and T. Pruefer, "A generative AI framework for intelligent utility billing CO2 analytics and sustainable resource optimisation," *arXiv preprint arXiv:2605.16250*, 2026, doi: 10.48550/arXiv.2605.16250.

47. K. Bonawitz et al., "Towards Federated Learning at Scale: A System Design," in *Proc. MLSys*, 2019.

48. Y. Liu et al., "Threats to Federated Learning," *arXiv preprint*, 2020, doi: 10.48550/arxiv.2012.09600.

49. NIST, "NIST AI Risk Management Framework (AI RMF 1.0)," National Institute of Standards and Technology, 2023.

50. ISO/IEC 42001:2023, "Artificial Intelligence—Management System," International Organization for Standardization, 2023.

51. P. Manjunath and T. Pruefer, "A quantum inspired variational kernel and explainable AI framework for cross region solar and wind energy forecasting," *arXiv preprint arXiv:2605.09032*, 2026, doi: 10.48550/arXiv.2605.09032.

52. ENISA, "AI Cybersecurity Risks," European Union Agency for Cybersecurity, 2020.

53. ICO, "Guidance on AI and Data Protection," UK Information Commissioner's Office, 2023.

54. CNIL, "AI Governance: Recommendations on the Development of AI Systems," Commission Nationale de l'Informatique et des Libert\'es, 2023.

55. M. Veale and F. Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act," *Comput. Law Rev. Int.*, vol. 22, no. 4, pp. 97–112, 2021, doi: 10.9785/cri-2021-220402.