Exploring Data Privacy and Protection Mechanisms in AI-Driven Organizations: Challenges and Best Practices

---

Dr Aaftab Qureshi[1], Dr Farukh Khan[2], Dr Parag Pande[3]
School of Computers, IPS Academy Indore
Maharaja Ranjit Singh Institute of Professional Studies (MCA Department)
aftabqureshi@ipsacademy.org[1], farukhkhan@ipsacademy.org[2], paragpande28@icloud.com[3]
https://doi.org/10.64882/ijrt.v14.iS2.1312

Abstract

It emerged as the explosive use of artificial intelligence in all company's systems aggravated data privacy and protection problems, owing to bigdata processing at scale, autonomous learning and black-box decision-making. We review the main privacy threats, governance challenges and best practices in AI-based organizations through a SLR according to PRISMA guidelines, focused on papers published from 2020 to 2025. The results demonstrate that privacy risks in AI stem from a combination of technical vulnerabilities and human factors, such as data memorization, inference attacks, prompt injection, and the complexity of regulation. The findings imply that effective privacy could be achieved only in a stacked way, joining up technical solutions with ethical AI governance, rules and regulation compliance, as well inter-organizational awareness to ensure responsible and trusted application of AI.

Keywords: Data Privacy, AI-Driven Organizations, Data Protection, Security Mechanisms, Ethical AI, Legal Compliance, Data Security Challenges, Best Practices, Privacy Regulations, AI Ethics.

Introduction

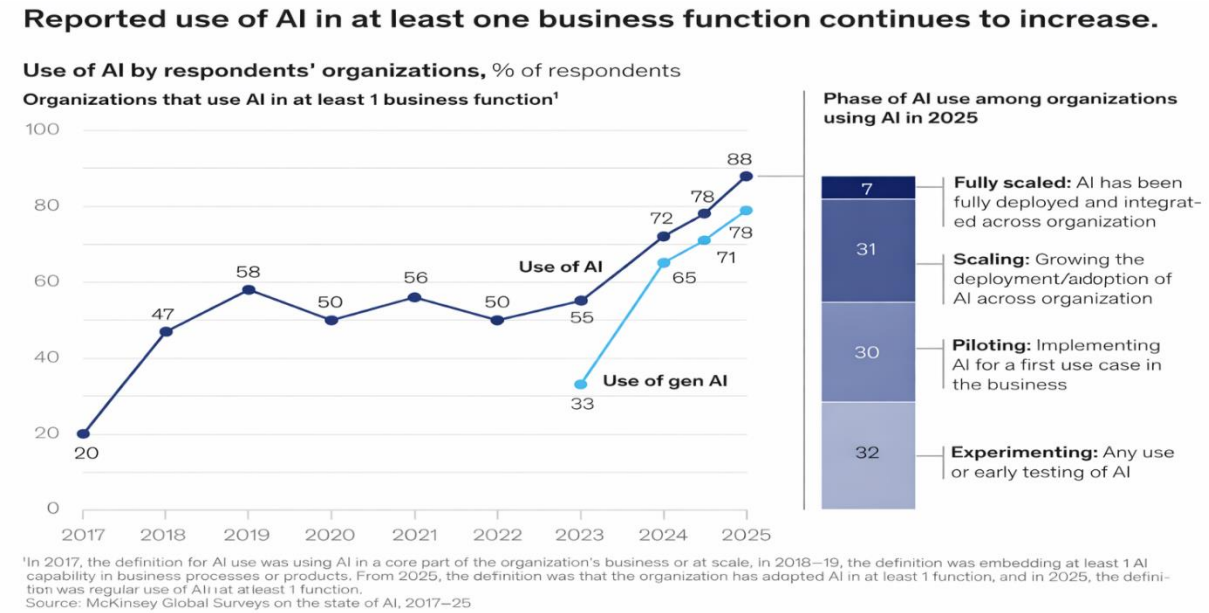
Organizations are leveraging AI to drive game-changing innovations across industries such as healthcare, finance, manufacturing and retail. Yet as AI technologies are used more and more in decision-making processes, they also raise deep challenges with privacy and data protection. AI-enabled enterprises analyse massive volumes of data, plenty of it personally and commercially sensitive hence the huge risk to privacy. Apprehension is compounded by the opaqueness of many artificial intelligence models, which are difficult to trace in order to justify decision making and hold people accountable. Also, those risks are further magnified by the biases that exist in training data and drive biased predictions as famously happened with hiring and lending. Regulatory requirements represent a further major challenge, with a widely varying landscape in legal regimes around the world that address AI. In order to tackle these challenges, companies that are powered by AI need to build strong data protection measures from the early stages of AI creation. Privacy best practices such as Privacy by Design, Data Minimization and Explicit Consent Mechanisms are the cornerstone of mitigating privacy risks. Enhanced privacy-enhancing technologies (PETs), e.g., differential privacy and federated

learning, may also provide new methods to protect data but still work for AI effectively. Provided they take a layered, proactive stance, businesses can safeguard user data, guarantee adherence to legislation and engender trust among users and other stakeholders.

In this research we examine challenges that organizations relying on AI have when it comes to protecting the privacy of their data and practices for these organizations to deal with such challenges. The report looks at trends, regulations and technology developments to indicate practical actions firms should be considering as part of their efforts to bolster their data defenses in the increasingly transformational world of AI.

Evolution of Data Privacy Concerns in AI

The progression of data privacy issues in AI has narrowed down from typical data-protection concerns to more complex problems related to machine learning and the use of automated decision-making systems. It started, of course, with traditional threats such as unauthorized access to data. But recent conversations have drawn attention to some of the AI-specific vulnerabilities in multimodal AI systems that process sensitive information, for example regarding healthcare. This pivot has also raised ethical concerns, namely around deepfakes and Reidentification. For public health, adoption of AI has been tempered by worries over data breaches and privacy, such as the development of COVID-19 tracing apps; Unsingable. In industrial scenarios, like Industry 4.0, growing system interconnectivity extends the list of potential attack vectors; for example, the Not Petya ransomware attack that caused hundreds of millions in financial losses. As technology becomes increasingly complex, the need for privacy frameworks that accommodate both technical risk and human behavior is more important than ever before.



https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

This image provides a visual representation of the increasing adoption of artificial intelligence (AI) across businesses, as reported in a McKinsey Global Survey. The chart

illustrates the rising use of AI in at least one business function by organizations between 2017 and 2025, highlighting key trends and shifts in AI utilization.

Growth in AI Adoption (2017-2025)

The line chart to the left shows continual acceleration of AI adoption over time, but a particular jump from 2023–2025. Back in 2017, just twice as many organizations were using AI in at least one business function (20%) and that number increased to 47% before last year ended. That grew further to 58% in 2020 and 56% in 2021 but it then levelled off at 50% in both 2022 and now for next year. But the figure skyrockets in 2024, when AI usage in enterprises goes to 78%, according to the chart. This growth is accentuated in 2025 with the number rising to 88%, attesting to the expanding use of AI across industry. One of the most substantial trends is the ascent of generative AI (gen AI), which see a dramatic transition from 2023 and beyond. In 2025, 71% of businesses were using generative AI, double the number (33%) in 2023 and highlighting how quickly this technology is being harnessed as tool.

Phases of AI Use in Organizations (2025)

The stacked bar chart on the right presents the breakdown of AI implementation stages among organizations that are using AI in 2025. This categorization highlights the various phases of AI integration within organizations:

1. **Fully Scaled (7%)**: This segment represents organizations where AI has been fully deployed and integrated across the entire business, indicating a mature, well-established use of AI.
2. **Scaling (31%)**: These organizations are in the process of expanding their AI applications, growing their AI capabilities across multiple functions or scaling AI solutions within specific business areas.
3. **Piloting (30%)**: At this stage, businesses are testing AI for the first time in select use cases. These organizations are exploring AI applications to assess their feasibility and impact before a wider rollout.
4. **Experimenting (32%)**: This group is engaged in early testing or use of AI technologies, indicating that many organizations are in the early stages of adopting AI and experimenting with its potential applications.

Implications for AI Adoption in Business

A number of important trends can be deduced from the data in this image. For one, the adoption of AI is now out of its novelty stage and entering the mainstream as more companies implement it into their business. The explosion in generative AI usage speaks to the transformative nature of it; from content creation and automation through decision-making.

But there are differing levels of maturity across the stages of AI adoption. Although many businesses are still trying out or testing AI, a substantial number have already extended their deployment of AI-based solutions throughout the enterprise, signaling an increasing integration of AI into business strategy. Fully scaled organizations are leading the way in AI adoption, with advanced use of AI across all areas of business to aid operational efficiency, make better decisions, and improve customer experience.

This shift in AI adoption means that companies will have to increasingly focus on the strategic planning, ethical competencies and regulatory actions required for maximizing the benefits of AI not just simply integrating more use-cases at a technical level. With AI deployment on the rise, companies by necessity will need to solve for data privacy, security issues and evolving workforce needs in addition to fueling AI deployments and ethical use of AI.

Literature Review

OWASP (2025), GenAI systems present their own particular set of risks, as a result of using large data sets, non-transparent training processes and complex deployment settings. These vulnerabilities range from accidental data exposure, model inversion attacks and prompt injection to unintended data retention and are often poorly handled by conventional cybersecurity models. The OWASP white paper calls out the need for custom security controls specifically adapted for LLM architectures, such as safely handling data during model training and inferences; enforcing solid access control policies; encrypting data at rest and in motion; and ongoing monitoring of how a trained model behaves. Finally, it illustrates the regulatory compliance issues in particular related to data protection laws like GDPR and without underestimating the necessity of minimization, anonymization and auditability when deploying GenAI workflows. We consider our study as a step forward toward the emerging area of literature by listing best practices in a systematic manner so that they can be implemented to get benefit from AI innovation while keeping data governed from security perspectives. The work as a whole highlight's critical needs for proactive and context-specific security strategies in order to responsibly deploy GenAI technologies at scale.

Papagiannidis et al. (2024) offer a systematic scoping review to aggregate the empirical contributions on responsible AI to overcome the fragmentation and ambiguousness stemming from earlier literature. The authors make a clear distinction between the responsible Building upon a critical analysis, the paper pinpoints structural, relational and procedural practices as key dimensions of responsible AI governance. This multi-dimensional approach contributes to the literature not only by abandoning normative considerations but also by proposing a governance-oriented perspective on how responsible AI can be incorporated into organizational processes and decisions making, monitoring and evaluation multidimensionally. In addition, the authors uncover implicit assumptions in the extant literature, contributing to a structured research agenda for the Information Systems (IS) field and calling attention to empirical testing as well as context dependent analysis. Towards a structured approach for AI governance Published in: 2020 IEEE International Conference on Engineering, Technology and Innovation (Ice/Technology and Innovation) Overall, this work makes very important contributions to the expanding dialogue on regulations related to AI by creating conceptual clarity and pragmatic path forward in terms of potential future research.

Hussain (2025) examines the privacy risks associated with AI-enabled systems, including unauthorized data access, inference attacks, and unethical data usage across sectors such as healthcare and finance. The study emphasizes that conventional data protection mechanisms are insufficient to address the complexity and scale of AI-driven privacy threats.

To mitigate these risks, the literature highlights the effectiveness of advanced privacy-preserving techniques, including differential privacy, homomorphic encryption, federated learning, and adversarial training.

Javed (2025) examines how the integration of large language models with cloud-native architectures enables highly personalized customer interactions while simultaneously introducing complex security vulnerabilities. The study identifies emerging risks such as prompt injection attacks, content safety issues, data privacy breaches, and regulatory compliance challenges that extend beyond traditional cybersecurity frameworks.

Billiris, Gill, and Bandara (2025) tackle above challenges through building a systematic taxonomy of data privacy risks specialized for AI systems. From a comprehensive review of 45 works, the authors outline 19 specific privacy risks that can be classified into dataset-level, model-level, infrastructure-level and insider threat domains, indicating the complexity of AI related privacy threats. One of the contributions of this study is the focus on human and organisational factors where both, in their own way, contributes to causing errors that lead to privacy risks. These findings challenge current security strategies primarily addressing technical controls, emphasizing on the necessity of holistic privacy governance where a combination of technical, behavioural and organizational aspects should be considered. More broadly, the taxonomy presents an organized view of AI privacy risk and contributes to our knowledge on trustworthy and responsible development of AI.

Tanisha et al. (2024) analyze these challenges from an AI-driven perspective and illustrate how intelligent are more and more deployed for managing, processing, and securing large volumes of industrial data. The research stresses that, even as AI is enabling analytics, automation and predictive maintenance 4 to become more efficient it also simultaneously carries the potential for misuse: the trigger effect. The literature also calls for combining regulatory frameworks with privacy-preserving techniques based on AI, to cope with such threats successfully. Its integration of technical, regulatory and ethical aspects into a holistic framework makes the study's findings a useful resource while adopting AI responsibly in Industry 4.0 and offers a comprehensive guide for researchers and policymakers dealing with data privacy in advanced industrial environments.

Jaiya (2024) emphasizes the dual role of AI, as a tool for enhancing data protection but also as a source of increased privacy risks. Although these enable greater efficiency and governance with respect to regulation, they also present an escrow of issues like bias in systems, transparency about how data are being utilized, potential misuse, and vulnerability against adversary attack. The study calls for explainable AI, ethical governance and harmonized regulatory standards to drive legitimate accountability and trusted adoption of AI in support of responsible and sustainable data privacy.

Research Methodology

In this paper we employ a SLR approach to investigate data privacy and protection solutions for AI-based enterprises. The study complies with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta- Analyses) protocol, to ensure a transparent, systematic and reproducible process. Literature including scholarly publications, conference papers,

industry reports and standards white papers from 2020 to 2025 were retrieved from well-known academic databases and organizational sources. Inclusion and exclusion criteria effectively identify high-quality and pertinent studies related to AI-related privacy risks, governance mechanisms, and mitigation approaches. Thematic analysis of the chosen literature was conducted with respect to the classifications of privacy as well as new best practices/governance standards that are becoming apparent from experimental practice.

Findings

The results suggest that organizations using AI will have to manage multifactorial data privacy risks which include technical and human weaknesses. Critical risks including data recall, inferring attacks, prompt injection, unauthorized access and non-compliance to regulations. The review reveals that there are privacy risks that distribute uniformly inside the data sets, models, infrastructure and organizational procedure. Several privacy-preserving techniques include but are not limited to differential privacy, federated learning, homomorphic encryption as well as explainable AI and Mops secure pipelines are intense hot topics being discussed for how they can effectively address these types of attacks. But the evidence in the literature suggests that technical controls alone are not enough without robust governance and organizational awareness.

Conclusion

The researchers recommend a comprehensive and multi-layered framework for privacy in AI-centric enterprises. Current data protection standards are not well-suited for the unique features of AI systems, including their capacity for autonomous learning and operation in black-box mode. Backed by robust technical measures, effective privacy management requires a combination of ethical AI governance, regulatory compliance, and human-centered organizational behaviors. The connection between mature forms of IG such as minimization, accountability and transparency, to the emerging AI governance frameworks is crucial for trusted and responsible AI systems.

Recommendations

Based on the findings, the study recommends that organizations:

- Implement privacy-by-design and security-by-design principles throughout the AI lifecycle.
- Adopt multi-layered privacy-preserving techniques, including differential privacy, federated learning, and secure ML Ops.
- Strengthen AI governance frameworks by integrating ethical guidelines, regulatory requirements, and risk assessment tools.
- Invest in employee training and awareness to reduce privacy risks caused by human error.
- Conduct continuous security evaluation and auditing to adapt to evolving AI threats.

Future Scope

Future research should focus on developing scalable and computationally efficient privacy-preserving AI techniques suitable for real-world deployment. Greater emphasis is needed on explainable AI (XAI) to enhance transparency and regulatory compliance. The

transition toward Industry 5.0, with its human-centric approach, presents opportunities to study collaborative AI governance models. Additionally, future studies may work toward creating standardized AI privacy risk assessment frameworks and exploring cross-jurisdictional regulatory harmonization to support global AI deployment.

References

1. Billiris, G., Gill, A., & Bandara, M. (2025). *Privacy in the age of AI: A taxonomy of data privacy risks*. In *Proceedings of the Australasian Conference on Information Systems (ACIS 2025)*. University of the Sunshine Coast & Australasian Association for Information Systems.
2. Billiris, G., Gill, A., & Bandara, M. (2025). *Privacy in the age of AI: A taxonomy of data privacy risks*. In *Proceedings of the Australasian Conference on Information Systems (ACIS 2025)*. University of the Sunshine Coast & Australasian Association for Information Systems.
3. https://adversa.ai/blog/adversa-ai-unveils-explosive-2025-ai-security-incidents-report-revealing-how-generative-and-agentic-ai-are-already-under-attack/.
4. https://blog.qasource.com/data-privacy-in-ai-testing.
5. https://community.trustcloud.ai/docs/grc-launchpad/grc-101/governance/data-privacy-and-ai-ethical-considerations-and-best-practices/.
6. https://data.folio3.com/blog/data-privacy-stats/.
7. https://hai.stanford.edu/ai-index/2025-ai-index-report
8. https://hai.stanford.edu/news/privacy-ai-era-how-do-we-protect-our-personal-information.
9. https://news.stanford.edu/stories/2025/10/ai-chatbot-privacy-concerns-risks-research.
10. https://ovic.vic.gov.au/privacy/resources-for-organisations/artificial-intelligence-and-privacy-issues-and-challenges/.
11. https://trustarc.com/resource/midyear-momentum-data-privacy-trends-2025/
12. https://www.ai21.com/knowledge/ai-data-privacy/.
13. https://www.aidataanalytics.network/data-governance/articles/7-trends-shaping-data-privacy-in-2025.
14. https://www.axiomlaw.com/blog/artificial-intelligence-data-privacy-challenges
15. https://www.blackfog.com/5-enterprise-use-cases-ai-privacy-concerns/.
16. https://www.crowdstrike.com/en-us/press-releases/ransomware-report-ai-attacks-outpacing-defenses/.
17. https://www.dataguard.com/blog/growing-data-privacy-concerns-ai/.
18. https://www.eff.org/deeplinks/2025/12/breachies-2025-worst-weirdest-most-impactful-data-breaches-year.
19. https://www.ey.com/en_lu/insights/ai/data-protection-in-the-ai-driven-era
20. https://www.f5.com/company/blog/top-ai-and-data-privacy-concerns.
21. https://www.fortra.com/blog/ai-data-privacy-challenges-and-solutions.
22. https://www.huntress.com/blog/biggest-data-breaches.
23. https://www.ibm.com/reports/data-breach.

24. https://www.ibm.com/think/insights/ai-privacy.
25. https://www.ibm.com/think/insights/ai-privacy.
26. https://www.jacksonlewis.com/insights/year-ahead-2025-tech-talk-ai-regulations-data-privacy
27. https://www.kiteworks.com/cybersecurity-risk-management/ai-data-privacy-risks-stanford-index-report-2025/.
28. https://www.kiteworks.com/cybersecurity-risk-management/ai-data-privacy-risks-stanford-index-report-2025/.
29. https://www.l-ten.org/Web/Web/News-Insights/focus-articles/Data-Protection-and-Privacy-in-AI-Based-Learning-Systems.aspx.
30. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai.
31. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work.
32. https://www.netfriends.com/blog-posts/5-data-privacy-best-practices-for-ai-users.
33. https://www.nsa.gov/Press-Room/Press-Releases-Statements/Press-Release-View/Article/4192332/nsas-aisc-releases-joint-guidance-on-the-risks-and-best-practices-in-ai-data-se/.
34. https://www.sentinelone.com/cybersecurity-101/data-and-ai/ai-data-security/.
35. https://www.thinkbrg.com/thinkset/ai-and-data-protection-in-2025-everything-that-rises-must-converge/
36. https://www.traverselegal.com/blog/ai-data-privacy-compliance/.
37. https://www.tredence.com/blog/ai-privacy.
38. https://www.trustcloud.ai/ai/boost-trust-with-powerful-ethical-ai-and-data-privacy-practices/
39. Jaiya, H. (2024). *Harnessing AI for data privacy: Examining risks, opportunities and strategic future directions*. *International Journal of Science and Research Archive, 13*(02), 2878–2892. https://doi.org/10.30574/ijsra.2024.13.2.2510.
40. Javed, A. (2025). *Data privacy and security in AI-driven customer platforms: A cloud computing perspective*. *European Journal of Computer Science and Information Technology, 13*(44), 84–95. https://doi.org/10.37745/ejcsit.2013/vol13n448495.
41. OWASP. (2025, February 13). *LLM and Gen AI data security best practices*. OWASP GenAI Security Project. https://genai.owasp.org/resource/llm-and-gen-ai-data-security-best-practices/.
42. Papagiannidis, E., Mikalef, P., & Conboy, K. (2024). *Responsible artificial intelligence governance: A review and research framework*. *The Journal of Strategic Information Systems, 33*(1), 101885. https://doi.org/10.1016/j.jsis.2024.101885.
43. Tanisha, J., Pillai, A. R., Roy, G. S., Koshy, A., Kothari, S., & Ajitha, D. (2024). *Privacy and data protection challenges in Industry 4.0: An AI-driven perspective*. *World Journal of Advanced Engineering Technology and Sciences, 12*(2), 064–089. https://doi.org/10.30574/wjaets.2024.12.2.0287.