Article Type: Cybertrust

Changes in the Privacy Landscape in Recent Years: Analysis from AI and non-AI Contexts

Abdul Majeed and Seong Oun Hwang, *Department of Computer Engineering, Gachon University, Korea*

Abstract—In recent years, the privacy landscape has changed significantly due to stringent privacy and utility requirements. We discuss changes in the privacy landscape for both AI and non-AI contexts by highlighting the latest technical solutions.

Privacy preservation (PP) has become an integral part of modern society owing to the rapid rise in technical solutions that collect, store, process, and disseminate huge amounts of personal data. Due to this collection and processing of a broad array of personal data, privacy risks have increased in manifold ways ranging from unique identifications to disclosures of intimate details of private lives. Recently, the proliferation of AI tools/models has made PP even more challenging because such tools can uncover and memorize intrusive private information from large and complex datasets. To that end, many PP techniques that have been developed do not let an AI model learn unnecessary information that can lead to private information prediction/estimation [1]. Conversely, PP has garnered significant attention in conventional data-sharing scenarios (e.g., when data in modified form is outsourced from owner environments to public domains for analysis). In this regard, many PP techniques have been developed that allow knowledge discovery while preventing the disclosure of personal information [2]. Based on the above analysis, the landscape of PP methods has changed in two distinct settings: AI and non-AI. In this article, how the privacy landscape has changed for both settings is analyzed, providing insights into the technologies that caused these changes and their impact on PP for diverse data types.

Privacy landscape change in AI context

Table 1 presents an overview of 10 major aspects that lead to privacy landscape changes in the AI context. Next, we concisely discuss each aspect.

TABLE 1. Major aspects that contributed to the change in privacy landscape in AI context.

Sr. #	Concise description of the aspect
1.	Emergence of methods that transfer compute to the data (algorithms → data)
2.	Data analytics from distributed parties without physically moving the data
3.	Training AI models in split without looking at data from any client that holds data
4.	Use of synthetic data (generative AI models) as a privacy-enabling technology with AI models
5.	Privacy for AI (privacy-aware training and inference)
6.	Integration of privacy-by-design and privacy engineering solutions
7.	Emergence of unlearning paradigms to prevent sensitive data memorization
8.	Coupling privacy with other dimensions of trustworthy AI (e.g., fairness/accuracy)
9.	Privacy preservation in multi-model settings and multi-modality data fusion
10.	Privacy guardrails (or PP) for LLMs

1) Before 2016, centralized learning (where data from multiple parties is first aggregated, and AI models are trained on the aggregated data to extract meaningful knowledge) proved to be a privacy-endangering paradigm. This paradigm transfers data to algorithms (data → algorithms) and is riskier in terms of privacy breaches. However, in 2016, the advent of federated learning (FL) reversed the notion. In FL, data is not aggregated in a central place, but AI models can still be trained on scattered data—computing started transferring toward data (algorithms → data). FL changed the privacy landscape in AI environments by limiting the amount of data moving to the central server, thereby effectively protecting privacy. Of late, FL has become a leading PP method in AI settings because it does not move data from data owners' environments to train AI models.

2) Google Research coined the term federated an-

XXXX-XXX © 2025 IEEE
Digital Object Identifier 10.1109/XXX.0000.0000000

This article has been accepted for publication in Computer. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MC.2025.3588418

alytics[1](FA) in 2020. In FA, analytical results are drawn from various distributed sources while keeping the data itself confidential. FA is a PP way to execute data science and evaluation queries—counts, heavy-hitter discoveries, histograms, model-quality metrics, etc.—directly on users' devices (a.k.a. data silos). FA has the lowest computing overhead because the learning itself is not included; instead, only the results are shared with the server/coordinator. FA brought a huge change to the privacy landscape because data remains local while analytical results are aggregated from diverse sources [3]. FA offers robust privacy guarantees while drawing conclusions/statistics from the scattered data.

3) In 2017, a powerful alternative to FL emerged in the form of split learning (SL) where an AI model (say, a neural network) is split into multiple small segments distributed across different clients and servers, and each client/server pair trains/tests its respective model's portion without looking at the raw data to accomplish the full model training/testing. In FL, the server does only a lightweight job of aggregation or coordination. In contrast, the server in SL does the job of training/testing for an assigned portion of an AI model, addressing the computation skew issues in FL. It is worth noting that servers do not fully train an AI model, and the clients' data privacy is strictly maintained. SL complements FL in many ways, and is a mainstream PP solution in AI environments these days.
4) According to a Gartner Report[2], synthetic data (SD) will dominate real data (RD) in AI development by 2030. SD has revolutionized the privacy domain by being the coarse form of RD and is thereby legally compliant. SD has been used in AI environments to resolve many privacy issues, such as membership inference attacks [4]. SD has significantly changed the privacy landscape because access to real data is often limited, and most of the existing RD is often scarce. Since 2018, SD curation with generative models (GANs, diffusion models, etc.) has become one of the mainstream PP technologies. In recent years, it has been widely used in AI environments as a privacy-enabling technology to mitigate diverse privacy attacks (data reconstruction, sensitive information prediction/inference, etc.).
5) Most AI models can learn privacy-engendering patterns (or unnecessary information) from the data, which can bring unintended privacy risks. For example, different subgroups based on certain demographics (or other salient information) exist in training data, and the privacy of those subgroups can leak from an AI model at inference time. To thwart such risks, privacy for AI has garnered significant attention from the research community where AI model learning from private information is restricted. To this end, software (injecting noise into gradients during training, or performing inference or training directly on encrypted data), hardware (e.g., trusted execution environments), and special-purpose protocols (zero-knowledge proof, model compression, etc.) have emerged as remedies. Based on these developments, it is fair to say that PP for AI training/inference has attracted wide attention and has become a must-have feature in recent years.
6) Although the privacy-by-design (PbD) methods such as FL, SL, and FA have limited the mobility of data and offer more PP than other privacy engineering solutions (PES) like anonymization, masking, and pseudonymization, they face diverse privacy attacks (e.g., data leakage from gradients, and model stealing). To address these problems, researchers have explored ways to integrate PbD and PES to address privacy issues in AI settings. For instance, differential privacy (DP) and FL have been integrated into many contexts to preserve the privacy of gradients, local/global models, and model parameters, etc. Encryption/DP has been integrated with PbD methods to train models by using encrypted data (or ciphertexts) as well as by adding carefully crafted noise only to scaling factors, leading to robust privacy guarantees. In 2019, Google introduced TensorFlow Privacy with DP-FedAvg and FL + DP code examples to demonstrate the synergy of PbD and PES. In a nutshell, a drastic change in the privacy landscape occurred in 2017 when PbD methods and PES were combined into unified frameworks for PP in the AI environments for different use cases.
7) Recently, machine/deep/federated unlearning methods have garnered the attention of both academia and industry owing to legal restrictions (e.g., the EU's GDPR) as well as the principle of the right to be forgotten. According to the regulations, if users want to withdraw their data, it

[1] https://research.google/blog/federated-analytics-collaborative-data-science-without-data-collection/

[2] https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai

must be removed from all sources—including AI models. However, erasing data from AI models is difficult because they cannot easily remove a particular individual's data after it is used. To this end, the unlearning methods have brought a huge change in the privacy landscape by removing the contribution of an erased/deleted subset of the training data without sacrificing predictive performance. The unlearning concept has been investigated in ML and DL settings. Recently, this concept has been studied in FL [5], underscoring the change in the privacy landscape in AI environments. Recent research has been moving toward development of certified unlearning for large-scale generative/foundation models as well as conventional AI models in order to provide effective PP.

8) In recent years, a lot of effort has been made to make AI trustworthy, and PP is just one dimension of it. However, PP alone can seriously impact other dimensions of trustworthiness. For example, a privacy-preserved AI model might have poor accuracy in some downstream tasks. Similarly, an AI model with strong PP may yield unfair decisions that can impact specific groups. Therefore, the privacy landscape has changed significantly, and other dimensions of trustworthiness have been jointly investigated alongside privacy. Besides, an intricate relationship exists between privacy and other dimensions of trustworthy AI, and therefore, different types of trade-off have been studied in the recent literature [6]. Some privacy standards like the EU privacy guidelines have emphasized the need for a careful balance between privacy and fairness and multi-dimensional trade-offs across all pillars of trustworthy AI.

9) In the past, most AI models were able to process only one type of data: either images or text. However, recent AI models have strong capabilities to simultaneously process multiple types of data (e.g., text + images, image + text + tables, audio + images + text), each one having diverse privacy requirements. To this end, the privacy landscape has changed, and methods have been upgraded to provide privacy guarantees in multi-model as well as multi-modality data settings [7]. In some scenarios, multiple AI models are integrated into a pipeline to solve real-world problems. Some methods have been developed that guarantee data privacy across diverse AI models. However, it is still challenging to ensure robust privacy guarantees in these complex settings. Based on these developments, it is fair to say that the privacy landscape has changed from single-model/data to multi-model/multi-modality data.

10) The advent of the large language model (LLM) has widened the spectrum of data privacy, and there is a growing risk of privacy breaches from such systems. In LLMs, there is a risk of privacy leakage from two ends: the LLM leaks private information from the training data, or users themselves expose secret information about their acquaintances/organizations to the LLM. Due to privacy concerns, some recent launches of LLMs (e.g., China's DeepSeek) were banned in many countries. To protect the privacy of sensitive data, many guardrails have been developed recently and integrated into famous LLMs like ChatGPT. These guardrails ensure that LLMs do not leak sensitive information (e.g., political issues) or give noisy answers when sensitive information is requested. Besides, conventional approaches such as DP, unlearning, etc., have been used in LLMs for PP [8]. However, by using clever prompting techniques or humorous engagement users can easily bypass these guardrails, so additional technical solutions are required. To that end, the privacy landscape has changed significantly, and many policy-driven or technical solutions are being developed for PP in LLMs.

Privacy landscape change in non-AI context

As stated earlier, the privacy landscape has also changed in non-AI contexts like data sharing, processing, and fusing. Table 2 presents a concise overview of 10 aspects that highlight the privacy landscape change in the non-AI context, and we discuss each aspect along with suitable examples.

1) According to MIT Technology Review, DP is considered a mainstream technology with a big impact on human life, particularly from the perspective of building trust and addressing the privacy paradox[9]. Owing to its robustness against adversaries and its theoretical guarantees, DP has become a gold standard in the privacy community. Many famous companies like Google, Microsoft, Apple, YouTube, etc., are using DP to protect the privacy of their affiliates [9]. In recent years, DP has been adopted for privacy protection in diverse types of data. It has been widely adopted in static scenarios (data sharing, query-based systems, etc.) as well as dynamic scenarios (client/server systems, distributed systems, etc.), which indicates a landscape change. It has

This article has been accepted for publication in Computer. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MC.2025.3588418

TABLE 2. Major aspects that contributed to the change in privacy landscape in non-AI context (general scenarios).

Sr. #	Concise description of the aspect
1.	DP adoption for diverse data modalities and heterogeneous settings
2.	AI for privacy: integrating AI with traditional anonymity/DP methods for PP
3.	Transitions from ad-hoc anonymity/DP methods to personalized methods
4.	Synergies between semantic and syntactic methods for effective PP
5.	Synthetic data (or dummy data) aided anonymization for PP in data sharing
6.	Shallow anonymization by skipping some attributes from noise/anonymity
7.	Development of attack-specific privacy methods in data sharing/processing
8.	Investigation and mitigation of invisible risks from data anonymization/noising
9.	Downstream task-aware data noising or sanitization for utility enhancement
10.	LLM-powered privacy disclosures, and countermeasures for effective PP

also undergone rigorous enhancement in terms of noise optimization, the theoretical foundation, etc., indicating a vital change in the privacy area.

2) Recently, a lot of AI techniques (particularly machine learning) have been integrated to address performance deficiencies (or other performance bottlenecks) in traditional anonymization/DP methods. For example, feature selection methods have been integrated to prune the less relevant attributes and reduce the computing burden. Methods such as clustering have been adopted to reduce information loss from data anonymization. Similarly, AI methods have been adopted to determine the optimal value of ϵ for a given dataset. Based on these developments, it is fair to say that AI synergy with traditional anonymization/DP-based approaches has brought a change in the privacy landscape. In the future, more AI methods are expected to optimize conventional workflows of PP methods to effectively protect privacy and enhance utility.
3) In the past, most PP methods were ad-hoc, meaning that privacy/utility parameters were determined by data owners without consulting data providers. Recently, this notion has changed, and some approaches let users define the sensitivity of their data, or choose the amount of noise injected into their data [9], indicating a drastic change in the privacy landscape. Methods that involve users handling their own data are referred to as personalized ones. These methods are vital to provide more tailored privacy guarantees as well as robust utility for downstream tasks. Future PP methods are expected to be more user-oriented than ad-hoc ones, which can lead to a reduction in data-specific biases in data-driven applications/products.
4) In the privacy literature, particularly in the context of data sharing, there are two broad categories of PP. Syntactic methods (k-anonymity, ℓ-diversity, t-closeness) alter data based on generalization hierarchies. In contrast, semantic methods ensure privacy guarantees by adding noise to the data. Recently, methods from both these categories have been integrated to optimize performance in terms of both privacy and utility. For example, t-closeness (a syntactic method) has been integrated with DP (a semantic method) to overcome performance issues. Analysis of the intricate relationship between different privacy models, and the development of hybrid PP models, underscores a privacy landscape change in the big data era.
5) As stated earlier, SD has become a vital component for AI development, leading to highly accurate and more generalizable model developments. Similarly, SD has been used recently for data protection when synthetic records are added to RD to provide robust privacy guarantees. SD has been integrated with RD to guarantee privacy from models like ℓ-diversity, which is sometimes difficult to accomplish due to data imbalance. Recently, SD has become a promising privacy-enabling technology, outperforming both syntactic and semantic privacy methods [10]. The integration of SD with anonymization/DP methods indicates a drastic change in the privacy landscape, and effectively contributes to accomplishing a privacy/utility trade-off in diverse use cases.
6) In the past, most methods anonymized entire segments of the data, leading to greater information loss and poor utility. In recent years, methods have been developed that apply anonymization to selective or vulnerable attributes/values by considering their availability from external sources [11]. Some methods allow users to provide additional information regarding the sensitivity of attributes, and privacy protection is accordingly applied [9]. Moving away from anonymizing/noising entire segments of the data and applying DP/anonymity to selective parts indicates a change in the privacy landscape. User feedback acquisition or classification of data based

This article has been accepted for publication in Computer. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MC.2025.3588418

on sensitivity as specified by users requires modifications/optimizations in the design of parameters for the underlying method, necessitating upgrades to many existing PP methods.

7) In the past, most PP methods were intended to provide safeguards against three main types of privacy attack: identity, sensitive information/attribute, and membership disclosures. However, due to the rapid proliferation of digital tools, the nature of privacy attacks changed from the above three disclosures to more sophisticated attacks like spatial-temporal activity disclosure, intent disclosure, sensitive preference disclosure, and data reconstruction. As a result, many attack-specific PP methods have been developed to safeguard against these attacks. The development of attack-specific PP methods indicates a change in the privacy landscape that advances the privacy field from more tailored and robust protection. This trend occurred in the AI context as well, where PP methods have been designed to thwart a specific type of attack (e.g., the model inversion attack).
8) Although most PP methods offer safeguards against diverse types of attack, and assist in resolving trade-offs between privacy and utility, they still have some invisible risks. For example, analysts cannot directly utilize anonymized/noised data in the knowledge discovery process, and some post-processing is imperative. In this case, overly anonymized/noised data cannot easily revert to real data; therefore, some work has been done to overcome these difficulties in the use of anonymized/noised data. In addition, anonymized/noised data can induce bias or fairness issues, leading to data-specific biases. In some cases, anonymity or noise cannot be directly added to the data owing to quality-related issues (missing values, outliers, etc.). To address these issues, some methods have been developed that investigate and resolve some of these issues, highlighting a much-needed change in the privacy landscape.
9) In recent years, anonymized/noised data has been widely used for diverse types of downstream tasks such as training AI models, pattern mining, rules extraction, etc. However, most PP methods do not consider the tasks for which the anonymized/noised data will be used, seriously impacting the utility of the data in the respective tasks. Of late, some methods have been developed that consider the feature/characteristics of downstream tasks to provide better utility [12]. The development of task-specific or downstream task-aware PP methods highlights a change in the privacy landscape in the current big data era.
10) Recently, LLMs have started contributing to many aspects of PP, such as optimizing noise for certain datasets, generating SD, providing a defense against certain attacks, and devising new types of attack. To this end, some LLM-powered PP methods have been developed that can offer a better defense, compared to conventional methods, against certain types of attack. However, some LLM-powered methods help attackers extract personal information from diverse sources, underscoring the misuse of LLMs [13]. Recently, a significant change occurred in conventional LLM privacy settings to prevent privacy disclosures and provide protection/countermeasures.

Lessons learned and ways forward

The privacy landscape has significantly evolved in the past few years, driven by strict privacy and utility requirements, legal measures, advancements in AI models, and the emergence of LLM. In the AI context, PP solutions tend to limit data mobility to avoid privacy issues. In contrast, PP methods in non-AI settings resolve the privacy/utility trade-off through algorithmic enhancements (or user involvement). Numerous developments in both contexts indicate an ever-increasing trend in the importance of privacy. With the advent of LLMs, privacy has garnered even further attention from academia, industry, and policymakers.

Moving forward, it is vital to develop novel PP methods for both AI and non-AI settings to resolve the privacy/utility trade-off at the least possible cost. It is important to link LLMs with PP methods such as DP, anonymization, FL, micro-aggregation, SD generation, etc., to safeguard personal data. It is imperative to develop LLM-powered PP pipelines tailored for fused data of different modalities. Also, the development of PP guardrails for LLMs to avoid disclosure of personal information is an urgent topic for research. Lastly, aligning PP methods with users' needs, and empowering users to make informed decisions about their data use is crucial for the future of data privacy.

Concluding remarks

The development of new PP methods, or upgrading existing methods to meet the growing privacy and utility requirements in both AI and non-AI contexts, has drastically changed the privacy landscape in the past few years. This growth was fueled by the need for more robust PP methods in both AI and non-AI areas, including LLMs and FL. The enclosed analysis of how the privacy landscape has changed in both contexts, along with the contributions of technical solutions, can pave the way to the development of robust next-generation PP methods.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (RS-2024-00340882). Seong Oun Hwang and Abdul Majeed are co-corresponding authors of this article.

REFERENCES

1. S. Yu, F. Carroll, and B. L. Bentley, “Insights into privacy protection research in ai,” *IEEE Access*, vol. 12, pp. 41 704–41 726, 2024.
2. R. Arshad and M. R. Asghar, “Characterisation and quantification of user privacy: Key challenges, regulations, and future directions,” *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2024.
3. D. Wang, S. Shi, Y. Zhu, and Z. Han, “Federated analytics: Opportunities and challenges,” *IEEE Network*, vol. 36, no. 1, pp. 151–158, 2021.
4. L. Hu, J. Li, G. Lin, S. Peng, Z. Zhang, Y. Zhang, and C. Dong, “Defending against membership inference attacks with high utility by gan,” *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 3, pp. 2144–2157, 2022.
5. Z. Liu, Y. Jiang, J. Shen, M. Peng, K.-Y. Lam, X. Yuan, and X. Liu, “A survey on federated unlearning: Challenges, methods, and future directions,” *ACM Computing Surveys*, vol. 57, no. 1, pp. 1–38, 2024.
6. S. Shaham, A. Hajisafi, M. K. Quan, D. C. Nguyen, B. Krishnamachari, C. Peris, G. Ghinita, C. Shahabi, and P. N. Pathirana, “Privacy and fairness in machine learning: A survey,” *IEEE Transactions on Artificial Intelligence*, 2025.
7. M. Alduniawi, K. Akkaya, and R. Sun, “Privacy-preserving oriented design for multi-modality models using fl,” in *2023 5th IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*. IEEE, 2023, pp. 163–168.
8. V. Rathod, S. Nabavirazavi, S. Zad, and S. S. Iyengar, “Privacy and security challenges in large language models,” in *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE, 2025, pp. 00 746–00 752.
9. J. Chen, C. Hu, Z. Liu, T. Xiang, P. Hu, and J. Yu, “Secret specification based personalized privacy-preserving analysis in big data,” *IEEE Transactions on Big Data*, 2024.
10. Q. Razi, S. Datta, V. Hassija, G. Chalapathi, and B. Sikdar, “Privacy utility tradeoff between pets: Differential privacy and synthetic data,” *IEEE Transactions on Computational Social Systems*, 2024.
11. A. Borrero-Foncubierta, M. Rodriguez-Garcia, A. Muñoz, and J. M. Dodero, “Protecting privacy in the age of big data: exploring data linking methods for quasi-identifier selection,” *International Journal of Information Security*, vol. 24, no. 1, pp. 1–14, 2025.
12. A. Utaliyeva, J. Shin, and Y.-H. Choi, “Task-specific adaptive differential privacy method for structured data,” *Sensors*, vol. 23, no. 4, p. 1980, 2023.
13. Y. Liu, Y. Jia, J. Jia, and N. Z. Gong, “Evaluating llm-based personal information extraction and countermeasures.”

Abdul Majeed is an Assistant Professor in the Department of Computer Engineering, Gachon University, Korea. He received his Ph.D. in Computer Information Systems & Networks from the Korea Aerospace University in 2021. His research interests include secure personal data publishing, data-centric AI, machine learning, and synthetic data. Contact him at ab09@gachon.ac.kr.

Seong Oun Hwang is a Professor in the Department of Computer Engineering, Gachon University, Korea. He received his Ph.D. in computer science from the Korea Advanced Institute of Science and Technology (KAIST) in 2004. He is a senior member of IEEE. His research interests include cryptography, data-centric AI, cybersecurity, and machine learning. Contact him at sohwang@gachon.ac.kr.