Engineering 52 (2025) 34–39

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Views & Comments

Data Security and Privacy for AI-Enabled Smart Manufacturing

Xin Wang [a,b,d], Daniel E. Quevedo [c], Dongrun Li [a,b], Peng Cheng [d,*],
Jiming Chen [d], Youxian Sun [d]

[a] Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China
[b] Shandong Provincial Key Laboratory of Industrial Network and Information System Security, Shandong Fundamental Research Center for Computer Science, Jinan 250014, China
[c] School of Electrical and Computer Engineering, The University of Sydney, Sydney, NSW 2006, Australia
[d] State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China

1. Data security in smart manufacturing

The global manufacturing sector is undergoing a digital transformation as traditional systems—reliant on physical assets such as raw materials and labor—struggle to meet demands for greater flexibility and efficiency. The integration of advanced information technology facilitates smart manufacturing (SM), which optimizes production, management, and supply chains [1]. In conventional computer-integrated SM, data processing and decision-making typically rely on predefined rules and algorithms. In contrast, artificial intelligence (AI)-enabled SM leverages complex machine learning and deep learning techniques, enabling systems to iteratively learn from data, adapt dynamically, and improve operational flexibility in response to changing conditions. Central to this paradigm is the synergy between AI and data-driven design, where data serves as a core element of production. By employing advanced sensing, computation, and modeling, AI-enabled SM trains models on diverse data sources, enabling the models to improve scheduling, quality control, and predictive maintenance [2]. In this way, data directly influences the effectiveness and adaptability of AI in manufacturing.

However, the openness of AI technologies in SM has raised significant data security concerns. Large-scale data collection and real-time transmission increase the risk of breaches, manipulation, and cyberattacks [3]. First, physical-layer data used for AI training can expose sensitive information about manufacturing processes, such as device status and configurations, potentially breaching commercial confidentiality. In addition, as manufacturing networks become more interconnected, they present more opportunities for cyberattacks. A prominent example is the 2017 WannaCry ransomware attack [4], which infected over 200 000 computers across 150 countries, resulting in 4 billion USD in damages. The attack disrupted manufacturing operations, causing significant financial losses and widespread supply chain disruptions.

Frequent data exchanges make global supply chains vulnerable, while differing security practices across enterprises, especially in data-sharing environments, complicate data security management. Traditional security measures, such as encryption and access control for static data, may be inadequate for protecting dynamic real-time data flows in AI-enabled SM. Due to the physics of feedback actuation, data flows exhibit temporal correlations and are often interdependent. This opens up novel solutions for detecting and counteracting data security and privacy attacks.

Cyber–physical integration is a defining feature of SM, yet it also introduces new challenges in data security and privacy (Fig. 1). On the one hand, the exchange and fusion of heterogeneous manufacturing data in the cyber layer enhances AI model generalization, thereby improving decision-making and production efficiency in the physical layer. On the other hand, the AI-driven nature necessitates large volumes of diverse, real-time data for model training, which expands the attack surface for potential adversaries. Specifically, malicious actors could exploit AI models to infer sensitive information through joint inference attacks. Additionally, the interconnectedness between cyber–physical layers and across manufacturing systems creates opportunities for stealthy data-/model-poisoning attacks, where subtle manipulations of input data can significantly degrade model performance. Moreover, the fusion of low-quality and multimodal data may introduce misalignment and inconsistencies, compromising model accuracy and increasing error rates. Collectively, these challenges—exacerbated by cyber–physical integration and AI deployment—can lead to severe operational consequences, including equipment downtime, system instability, and diminished operational efficiency. Table 1 [5–19] systematically summarizes the key data security and privacy challenges in AI-enabled SM alongside their corresponding representative solutions.

(1) **Challenge 1: joint inference risks for sensitive information.** Data circulation and sharing are vital for AI-driven improvements in manufacturing, but they can expose sensitive information, such as system operations and human-related data. Manufacturing data often involves strong spatiotemporal correlations [20], making single-point protection methods insufficient to prevent adversaries from exploiting these correlations to jointly infer sensitive information. In addition, privacy-preserving

---

* Corresponding author.
E-mail address: lunarheart@zju.edu.cn (P. Cheng).

https://doi.org/10.1016/j.eng.2025.08.008

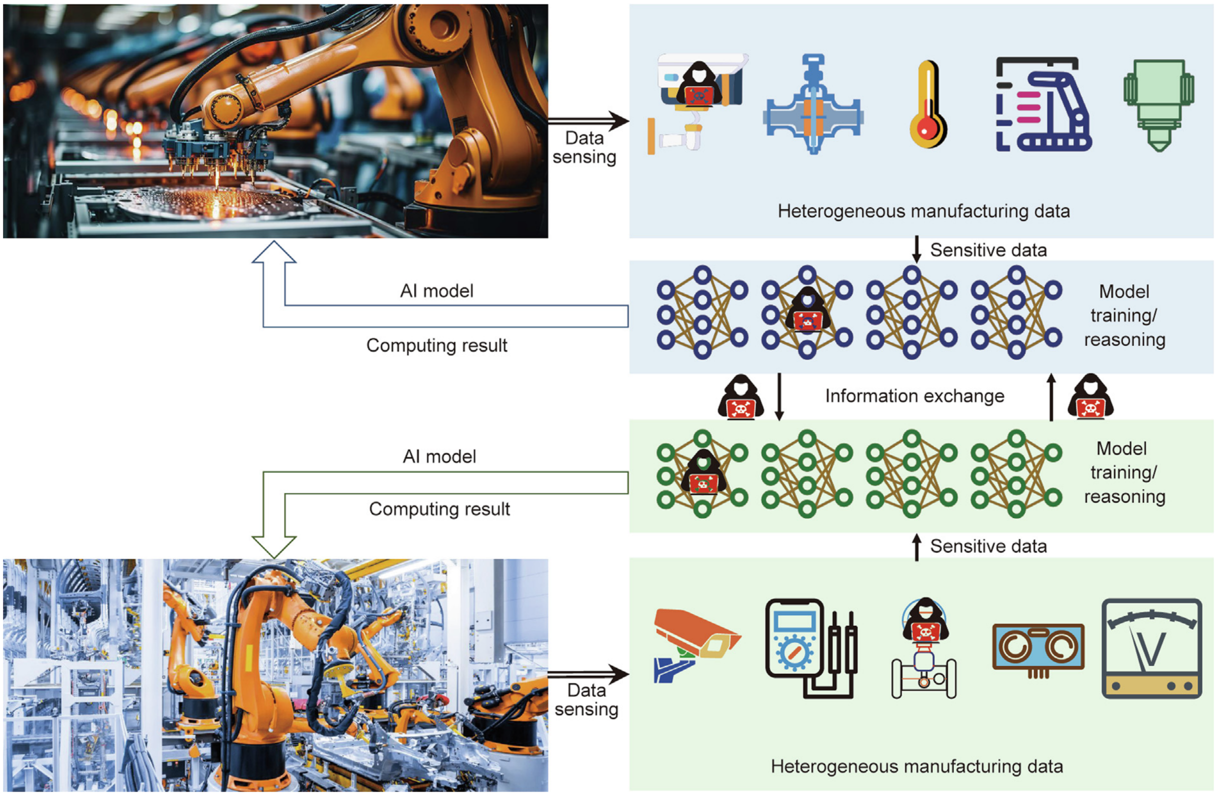

Fig. 1. Data security and privacy challenges in SM.

Table 1
Data security/privacy challenges and existing solutions for AI-enabled SM.

Data security/privacy challenge	Response strategy	Representative solutions
Joint inference risk for sensitive information	Optimizing data circulation and blocking potential attack pathways	Data synthesis [5] Bit/domain embedding [6] Differential privacy [7–9] Incentive mechanism [10]
Stealthy data-/model-poisoning attack	Identifying/tracking stealthy dynamic poisoning attacks and ensuring algorithmic resilience	Digital watermarking [11] Variational autoencoders [12] Smart contract detection [13] Unlearning verification [14] Resilient loss [15]
Fusion discrepancy of low-quality and multimodal data	Enhancing subpar data, maximizing cross-modality information utilization, and aligning local models and representations	Diffusion model [16] Multi-task learning [17] Graph convolutional networks [18] Local-global alignment [19]

techniques have their own limitations: Federated learning (FL) enables cross-system collaborative learning without raw data sharing but is vulnerable to model inversion attacks that can reveal private information [21]. Cryptographic methods such as homomorphic encryption are resource-intensive, making them impractical for field sensors and devices with limited resources. Non-cryptographic approaches, such as noise perturbation, may degrade model performance [22]. Hence, developing protection strategies that balance privacy, resource cost, and model performance is critical to mitigate joint inference risks.

(2) **Challenge 2: stealthy data-/model-poisoning attack.** The integration of AI in manufacturing is challenged by frequent data exchanges, dynamic network environments, and the lack of centralized regulation, creating opportunities for malicious adversaries to launch data- or model-poisoning attacks. Attackers can use both short-term static and long-term dynamic methods to manipulate data or model parameters while evading detection. Additionally, the uncertainty introduced by privacy-protection methods can further aid attackers in bypassing detection systems [23]. Thus, the key challenge in identifying and defending against poisoning attacks in a fine-grained manner is to develop effective strategies that leverage multidimensional information.

(3) **Challenge 3: fusion discrepancy of low-quality and multimodal data.** Reliable data fusion is essential for unlocking the full value of manufacturing data, but the diversity of equipment and environments in the industry introduces significant challenges. First, many data-collection and sensing devices operate in dynamic industrial settings, resulting in noisy and biased real-world datasets. Second, differences in data formats across enterprises and a lack of expert knowledge can lead to missing key data blocks, such as labels, resulting in incomplete data modalities. Additionally, structured, semi-structured, and unstructured data vary significantly in type and size, while semantic mismatches and imbalanced modalities further complicate the fusion process [24]. In such cases, significant discrepancies in data and improper fusion may yield only marginal improvements in generalization while severely compromising cross-task adaptability. Worse, such data can effectively function as malicious inputs, degrading overall model performance. More critically, adversaries could exploit these mismatches to launch stealthier poisoning attacks, as systems must employ broader detection thresholds to avoid excessive

false alarms. Addressing these fusion discrepancies is critical in improving the robustness and reliability of AI models when processing low-quality, multimodal data across platforms and enterprises—a key data security challenge for AI-enabled SM.

2. Existing solutions for challenge 1: Privacy leakage mitigation

To mitigate privacy leakage and sensitive information inference risks in AI-enabled SM, as highlighted in challenge 1, researchers have developed methods to optimize data circulation and block potential attack pathways.

2.1. Privacy metric selection

Selecting an appropriate privacy metric is essential in order to design effective protection schemes. Differential privacy (DP) [25] is a commonly used non-cryptographic method that ensures the outputs from two adjacent inputs remain indistinguishable through noise injection. A key challenge in DP is calculating the sensitivity of the original data processing function, which often requires simplifying assumptions. To address this, Kullback–Leibler DP [10] has been introduced, which combines the degree of sensitivity and privacy protection by using relative entropy to analyze the similarity between output distributions. This approach moves away from analyzing output differences from a probabilistic perspective, making it more suitable for the cross-system collaborative algorithms used for data circulation.

2.2. Privacy protection mechanisms

To meet the specific privacy protection needs of manufacturing enterprises, protection mechanisms have been designed to ensure that participants are motivated to share data while preserving privacy. Various privacy-preserving solutions cater to the heterogeneous protection requirements of users in different settings. For trajectory data analysis—an essential task for the path planning of industrial robots—a novel algorithm has been introduced for generating differentially private synthetic trajectory data using Markov models of varying orders [7]. The algorithm employs an adaptive model-selection strategy to enhance data utility, selecting high-quality models while reducing noise. Similarly, the work in Ref. [8] introduces a DP solution for the privacy-aware range query problem. For distributed machine learning, the scheme in Ref. [9] integrates label randomization and combined noise addition without a central server, offering heterogeneous privacy guarantees based on data sensitivity and participant trust levels. These privacy-aware approaches address diverse computing and learning requirements across manufacturing systems and enterprises while facilitating secure data sharing.

2.3. Inference attacks defense

To counter inference attacks, researchers have developed methods based on data generation and embedding techniques. First, synthetic datasets can be utilized to replicate the statistical properties of sensitive manufacturing data [5]. These datasets preserve the overall patterns while omitting specific operational details, allowing insights to be shared without disclosing proprietary information. By maintaining statistical integrity and removing identifiable links, the risk of re-identification or fusion attacks is reduced. Additionally, various embedding techniques, such as least-significant bit embedding, frequency domain embedding, and spatial domain embedding [6], integrate sensitive data into transmitted data while maintaining the appearance and structure of the original. These methods effectively conceal sensitive information, significantly increasing the difficulty for adversaries attempting to exploit private data.

2.4. Privacy and performance balance

Balancing privacy protection and algorithm performance is vital for the wide-spread adoption of perturbation-based protection methods in AI-enabled SM, as computed results or trained AI models directly impact the optimization of the manufacturing system. To address this challenge, two incentive mechanisms have been designed to encourage participants to contribute data or intermediate information with reduced noise perturbation [10]. The first mechanism, based on contract theory, offers a monetized incentive that achieves optimal computation accuracy within a given budget. The second, a non-monetized counterpart, introduces computation accuracy as a key optimization objective for participants. By minimizing an objective function that includes both privacy loss and computation deviation, participants can self-select an optimal noise perturbation level. Both incentive mechanisms establish optimal privacy settings that effectively balance privacy protection with performance trade-offs.

2.5. Limitations and future challenges

Although existing methods have achieved a partial balance between privacy protection, data sharing, and algorithm optimization, they still face challenges: ① inevitable data utility loss from privacy-preserving noise injection; ② limited adaptability of privacy protection in diverse manufacturing scenarios and evolving threats; ③ difficulties in enabling cross-domain data sharing due to heterogeneous systems and policies; and ④ computational complexity that hinders real-time deployment. Addressing these issues requires developing adaptive privacy frameworks that can accommodate varying sensitivity levels while maintaining practical utility, as well as establishing efficient cross-system collaboration mechanisms for AI-enabled SM environments.

3. Existing solutions for challenge 2: Poisoning attack detection and defense

Efforts to solve challenge 2 have led to the development of several advanced defense methods, particularly in identifying and tracking stealthy dynamic poisoning attacks. When poisoning attacks evade detection, ensuring algorithmic resilience while maintaining high performance also remains a critical research focus.

3.1. Attack detection mechanisms

The core principle of poisoning attack detection lies in identifying their malicious characteristics. One approach, model watermarking, embeds imperceptible markers into a model during training to verify its authenticity [11]. These markers remain intact during normal usage but are distorted if the model is tampered with or stolen. Such distortions enable the detection of adversarial modifications and help trace malicious activity. Another technique leverages variational autoencoders, which learn the distribution of benign training data by encoding it into a latent space and reconstructing it [12]. Reconstruction errors are computed, with large errors indicating deviations from the normal pattern, thereby permitting anomaly detection. Malicious attacks can be identified by flagging data points that deviate significantly from the learned distribution. Additionally, blockchain-based detection employs smart contracts to monitor the model's training process in real time [13]. These contracts automatically trigger countermeasures upon detecting malicious alterations. By integrating these methods, systems can detect complex and continuously evolving poisoning attacks, ensuring stable manufacturing operations despite escalating security threats.

3.2. Attack-resilient defense

Ensuring the resilience of learning algorithms against stealthy poisoning attacks is critical for deploying intelligent models in real-world manufacturing systems. To enable secure cross-system collaboration under adversarial conditions, Wang et al. [15] proposed FENDER, an alternating direction method of multipliers (ADMM)-based resilient defense scheme against label-flipping attacks. This scheme integrates a resilient loss function designed to handle unknown and asymmetric flipping rates. This scheme ensures consistent classification probabilities, regardless of flipped or true labels, and converges at a linear rate. To mitigate vulnerabilities in existing unlearning verification mechanisms (e.g., susceptibility to data breaches and verification manipulation), Gao et al. [14] introduced VeriFi, a structured local verification framework for departing participants. This framework operates in two stages, making and checking, ensuring that data from departing participants is effectively erased once the model stabilizes. By integrating watermarking and fingerprinting modules, along with Byzantine-robust aggregation rules, the framework increases the security and robustness of the verification process. These defense schemes are tailored to align local models with the unique characteristics of devices and production environments, ensuring robust operation under diverse adversarial scenarios.

3.3. Limitations and future challenges

Despite recent advances, existing defense mechanisms exhibit several critical limitations. First, high rates of false positives and false negatives hinder reliable discrimination between benign system variations and poisoning attacks. Second, many approaches impose substantial computational overhead, compromising their suitability for real-time manufacturing applications. Current solutions also demonstrate limited adaptability to increasingly sophisticated and heterogeneous attack strategies, rendering them ineffective against dynamically evolving threats. Additionally, scalability remains a key challenge—some methods exhibit performance degradation when handling high-dimensional data or operating in large-scale distributed environments, significantly reducing overall system efficiency.

4. Existing solutions for challenge 3: Fusion discrepancy alignment

To address challenge 3, existing studies have focused on enhancing subpar data and maximizing cross-modality information utilization to reliably fuse low-quality and multimodal manufacturing data. A critical aspect of this challenge is achieving alignment between local models and representations to facilitate robust multimodal information exchange across enterprises, a topic gaining attention in FL-enabled SM applications.

4.1. Robust fusion for low-quality and multimodal data

To handle low-quality data, diffusion models employ robust denoising and data-recovery techniques, mitigating fusion discrepancies in complex manufacturing environments [16]. These models extract meaningful signals from noisy inputs, generate high-fidelity samples, and reduce the biases induced by environmental fluctuations. For multimodal data fusion, multi-task learning trains a single model to perform multiple related tasks simultaneously [17], promoting knowledge sharing across tasks. This approach also enables the model to better interpret and manage the complexities of manufacturing environments. Additionally, graph convolutional networks (GCNs) [18] leverage structural relationships between manufacturing system components to reduce fusion inconsistencies, enhancing the model's ability to extract meaningful patterns from multimodal data.

4.2. Model alignment for heterogeneous data

To ensure alignment between local models (maintained by individual enterprises) and a global model, recent work introduced a secure FL framework based on model similarity optimization [19]. This approach combines a federated contrastive learning framework with a triple Siamese network for local training with a dual aggregation mechanism for global model generation. In particular, it uses model similarity to adjust the update directions of local models and aggregation weights for the global model. Furthermore, applying a stop-gradient scheme creates an adversarial training effect, ensuring the global model continuously updates toward the local model during the local training process, maintaining a balance between both models. The dual aggregation mechanism dynamically aggregates local models, enabling the global model to incorporate more personalized knowledge. As a result, this framework enables local models to retain enterprise-specific knowledge while benefiting from the generalization capabilities of the global model, achieving effective local–global alignment.

4.3. Limitations and future challenges

While current methods have advanced low-quality and multimodal data fusion, they still face several critical challenges. Diffusion models and multi-task learning require substantial computational resources, leading to long training times, especially when handling large-scale datasets. GCNs are limited by their dependence on carefully constructed graph representations, which may not always be available or reliable in manufacturing settings. Furthermore, existing local–global collaborative optimization approaches might remain vulnerable to privacy breaches through inference attacks shown in challenge 1, necessitating the integration of enhanced defense mechanisms like adaptive DP or multi-level secure aggregation.

5. Conclusion and future opportunities

Data is a critical component in modern manufacturing industries. Due to the strong spatiotemporal correlations inherent in manufacturing processes, as well as the complex data flows between cyber–physical layers and across enterprises, ensuring reliable data security and privacy is essential to unlocking the full potential of data. This paper has systematically analyzed emerging data security threats in AI-enabled SM environments, including privacy leakage, poisoning attacks, and fusion discrepancies, while evaluating existing countermeasures. However, current solutions often address these challenges in isolation, whereas real-world manufacturing scenarios increasingly exhibit concurrent or interdependent threats. There remains an urgent need for an integrated security framework that holistically addresses these interconnected challenges while accounting for manufacturing-specific constraints.

Drawing insights from existing solutions, potential integrated security countermeasures for manufacturing systems could encompass both single- and cross-system environments. In single-system settings where manufacturers serve as both data generators and consumers, possible countermeasures might include: generating synthetic datasets with embedded watermarks alongside privacy-preserving data embedding techniques during production, and utilizing variational autoencoders for secure transmission combined with diffusion models and multi-task learning to mitigate fusion biases induced by low-quality multimodal data. For cross-system scenarios, a prospective framework could explore decentralized approaches featuring adaptive privacy guarantees through Kullback–Leibler DP optimization for local training, robust loss functions derived from benign data characteristics for model security, and watermarking/fingerprinting for parameter verification. These could potentially be integrated with a local–global collaborative training scheme enhanced by multi-level secure aggregation to balance security, generalization, and personalization in heterogeneous data fusion. By aligning security solutions with specific problem domains and operational scenarios, one can develop optimized lightweight methods for each system module. This targeted approach would significantly reduce the cumulative overhead of security measures while maintaining robust protection, thereby better meeting the practical constraints and performance requirements of manufacturing environments.

Beyond the integration and enhancement of existing security algorithms, several promising research avenues remain underexplored in data protection and privacy preservation for AI-enabled SM:

(1) **Matching mechanism of customized model setup and personalized privacy protection.** Customized models better address the specific needs of production equipment, improving predictions and optimization for manufacturing systems. However, this personalization increases the risk of privacy leakage due to a reliance on sensitive enterprise data [26]. By introducing personalized noise into customized models, adaptive DP can help strike a balance between achieving optimal customization and ensuring sufficient protection for sensitive data.

(2) **Collaborative design of large-scale model integration and lightweight model training.** A collaborative framework of large and small models addresses privacy and security risks in data sharing [27]. Lightweight models enable personalized training within enterprises, abstracting knowledge for secure transmission. Larger models enhance resource integration and information fusion. This large-small model collaboration paradigm ensures efficient cross-enterprise data sharing with strong privacy protection.

(3) **Dynamic optimization of computation and communication resources under efficient data flow.** Effective resource allocation is critical for both security enhancement and personalized model training. Reinforcement-learning-based strategies [28] dynamically assess and adapt computation and communication workloads, optimizing system efficiency while facilitating robust security countermeasures.

(4) **Effective AI-based defense against emerging data security threats.** Advanced threats, such as deepfake-based attacks, pose significant security risks in critical manufacturing domains, including product quality control and equipment monitoring [29]. By integrating diverse AI techniques, their complementary strengths can be leveraged to develop robust solutions against evolving threats. For example, combining generative adversarial networks with watermarking can simultaneously prevent sensitive data inference and trace malicious activities [30], offering dual protection through privacy preservation and attack detection.

In summary, the growing demand for data security and privacy, alongside rapid AI advancements, presents significant technical challenges for the manufacturing industry. By leveraging emerging technologies and addressing these challenges, a connected, efficient, and resilient manufacturing ecosystem can be built, fostering enterprise innovation while ensuring robust data security and privacy protection.

CRediT authorship contribution statement

Xin Wang: Writing – original draft, Investigation, Methodology. **Daniel E. Quevedo:** Supervision, Writing – review & editing. **Dongrun Li:** Validation, Writing – review & editing. **Peng Cheng:** Writing – review & editing, Conceptualization, Investigation. **Jiming Chen:** Supervision, Project administration. **Youxian Sun:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported in part by the National Natural Science Foundation of China (62293511 and 62402256), in part by the Shandong Provincial Natural Science Foundation of China (ZR2024MF100), in part by the Taishan Scholars Program (tsqn202408239), and in part by the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (ICT2025B13).

References

[1] Wang B, Tao F, Fang X, Liu C, Liu Y, Freiheit T. Smart manufacturing and intelligent manufacturing: a comparative review. Engineering 2021;7(6):738–57.
[2] Sahoo S, Lo CY. Smart manufacturing powered by recent technological advancements: a review. J Manuf Syst 2022;64:236–50.
[3] Suvarna M, Yap KS, Yang W, Li J, Ng YT, Wang X. Cyber–physical production systems for data-driven, decentralized, and secure manufacturing—a perspective. Engineering 2021;7(9):1212–23.
[4] Li J, Sisodia D, Stafford S. On the detection of smart, self-propagating internet worms. IEEE Trans Depend Secure Comput 2023;20(4):3051–63.
[5] Cai Z, Xiong Z, Xu H, Wang P, Li W, Pan Y. Generative adversarial networks: a survey toward private and secure applications. ACM Comput Surv 2022;54(6):132.
[6] Kombrink MH, Geradts ZJMH, Worring M. Image steganography approaches and their detection strategies: a survey. ACM Comput Surv 2025;57(2):33.
[7] Wang H, Zhang Z, Wang T, He S, Backes M, Chen J, et al. PrivTrace: differentially private trajectory synthesis by adaptive Markov models. In: Proceedings of the 32nd USENIX Security Symposium; 2023 Aug 9–11; Anaheim, CA, USA. Red Hook: Curran Associates, Inc.; 2023. p. 1649–66.
[8] Du L, Zhang Z, Bai S, Liu C, Ji S, Cheng P, et al. AHEAD: adaptive hierarchical decomposition for range query under local differential privacy. In: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security; 2021 Nov 15–19; online conference. New York City: Association for Computing Machinery; 2021. p. 1266–88.
[9] Wang X, Ishii H, Du L, Cheng P, Chen J. Privacy-preserving distributed machine learning via local randomization and ADMM perturbation. IEEE Trans Signal Process 2020;68:4226–41.
[10] Wang X, Ishii H, He J, Cheng P. Dynamic privacy-aware collaborative schemes for average computation: a multi-time reporting case. IEEE Trans Inf Forensics Security 2021;16:3843–58.
[11] Tan J, Zhong N, Qian Z, Zhang X, Li S. Deep neural network watermarking against model extraction attack. In: Proceedings of the 31st ACM International Conference on Multimedia; 2023 Oct 29–Nov 3; Ottawa, ON, Canada. New York City: Association for Computing Machinery; 2023. p. 1588–97.
[12] Kieu T, Yang B, Guo C, Cirstea RG, Zhao Y, Song Y. Anomaly detection in time series with robust variational quasi-recurrent autoencoders. In: Proceedings of 2022 IEEE 38th International Conference on Data Engineering; 2022 May 9–12; Kuala Lumpur, Malaysia. Piscataway: IEEE; 2022. p. 1342–54.
[13] Chen D, Liao Z, Chen R, Wang H, Yu C, Zhang K, et al. Privacy-preserving anomaly detection of encrypted smart contract for blockchain-based data trading. IEEE Trans Dependable Secur Comput 2024;21(5):4510–25.
[14] Gao X, Ma X, Wang J, Sun Y, Li B, Ji S, et al. VeriFi: towards verifiable federated unlearning. IEEE Trans Depend Secure Comput 2024;21(6):5720–36.
[15] Wang X, Fang C, Yang M, Wu X, Zhang H, Cheng P. Resilient distributed classification learning against label flipping attack: an ADMM-based approach. IEEE Internet Things J 2023;10(17):15617–31.
[16] Carlini N, Hayes J, Nasr M, Jagielski M, Sehwag V, Tramèr F, et al. Extracting training data from diffusion models. In: Proceedings of the 32nd USENIX Security Symposium; 2023 Aug 9–11; Anaheim, CA, USA. Red Hook: Curran Associates, Inc.; 2023. p. 5253–70.

[17] Zhang Y, Yang Q. A survey on multi-task learning. IEEE Trans Knowl Data Eng 2022;34(12):5586–609.

[18] Ren H, Lu W, Xiao Y, Chang X, Wang X, Dong Z, et al. Graph convolutional networks in language and vision: a survey. Knowl-Based Syst 2022;251:109250.

[19] Wang X, Wang Y, Yang M, Li F, Wu X, Fan L, et al. FedSiam-DA: dual-aggregated federated learning via Siamese network for non-IID data. IEEE Trans Mobile Comput 2025;24(2):985–98.

[20] Liu Y, Wen R, He X, Salem A, Zhang Z, Backes M, et al. ML-Doctor: holistic risk assessment of inference attacks against machine learning models. In: Proceedings of the 31st USENIX Security Symposium; 2022 Aug 10–12; Boston, MA, USA. Red Hook: Curran Associates, Inc.; 2022. p. 4525–42.

[21] Zhu L, Liu Z, Han S. Deep leakage from gradients. In: Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, Garnett R, editors. Advances in neural information processing systems 32: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019); 2019 Dec 8–14; Vancouver, BC, Canada. Red Hook: Curran Associates, Inc.; 2019.

[22] Liu B, Ding M, Shaham S, Rahayu W, Farokhi F, Lin Z. When machine learning meets privacy: a survey and outlook. ACM Comput Surv 2022;54(2):31.

[23] Al-Rubaie M, Chang JM. Privacy-preserving machine learning: threats and solutions. IEEE Secur Priv 2019;17(2):49–58.

[24] Zhao F, Zhang C, Geng B. Deep multimodal data fusion. ACM Comput Surv 2024;56(9):216.

[25] Dwork C, Roth A. The algorithmic foundations of differential privacy. Found Trends Theor Comput Sci 2014;9(3–4):211–407.

[26] Nasr M, Shokri R, Houmansadr A. Comprehensive privacy analysis of deep learning: passive and active white-box inference attacks against centralized and federated learning. In: Proceedings of 2019 IEEE Symposium on Security and Privacy; 2019 May 19–23; San Francisco, CA, USA. Piscataway: IEEE; 2019. p. 739–53.

[27] Lu S, Guo L, Wang W, Zhao Z, Yue T, Liu J, et al. Collaborative training of tiny-large vision language models. In: Proceedings of the 32nd ACM International Conference on Multimedia; 2024 Oct 28–Nov 1; Melbourne, VIC, Australia. New York City: Association for Computing Machinery; 2024. P. 4928–37.

[28] Zabihi Z, Eftekhari Moghadam AM, Rezvani MH. Reinforcement learning methods for computation offloading: a systematic review. ACM Comput Surv 2023;56(1):1–41.

[29] Kaur A, Noori Hoshyar A, Saikrishna V, Firmin S, Xia F. Deepfake video detection: challenges and opportunities. Artif Intell Rev 2024;57(6):159.

[30] Du L, Chen M, Sun M, Ji S, Cheng P, Chen J, et al. ORL-AUDITOR: dataset auditing in offline deep reinforcement learning. In: Proceedings of Network and Distributed System Security Symposium 2024; 2024 Feb 26–Mar 1; San Diego, CA, USA; 2024. p. 184.