Computer Science Review 59 (2026) 100844

Contents lists available at ScienceDirect

Computer Science Review

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

Review article

Distributed edge artificial intelligence empowered data trading: Theories, algorithms, and applications

Bing Mi [a], Kongyang Chen [b,c,d,*]

[a] School of Public Finance and Taxation, Guangdong University of Finance and Economics, China
[b] School of Artificial Intelligence, Guangzhou University, China
[c] Yunnan Key Laboratory of Service Computing, Yunnan University of Finance and Economics, China
[d] Pazhou Lab, Guangzhou, China

ARTICLE INFO

Keywords:
Data trading
Edge artificial intelligence
Data pricing
Privacy computing
Data quality assessment
Incentive mechanism
Decentralized data transactions

ABSTRACT

As data emerges as a crucial production factor in the digital economy, data trading has become a key innovation across industrial applications. However, traditional data trading models face significant challenges, including privacy concerns, pricing mechanisms, transaction efficiency, and decentralization. Edge Artificial Intelligence (AI) offers a transformative approach to data trading by enabling local intelligence, privacy-preserving computation, and decentralized architectures. This paper explores the integration of Edge AI with data trading, analyzing its technical foundations, key challenges, and diverse applications. Edge AI enhances data trading by reducing data transmission costs, enabling privacy-preserving federated learning, and facilitating decentralized trust mechanisms through blockchain. Compared to centralized models, Edge AI-based data trading supports secure, efficient, and transparent transactions while mitigating the data misuse. This paradigm benefits various domains, including IoT-enabled smart factories and connected vehicles, financial risk assessment and credit scoring, secure electronic health record sharing, supply chain optimization, and personalized advertising. Despite its advantages, Edge AI-driven data trading still faces challenges in standardization, security, computational efficiency, data quality assessment, and regulatory compliance. Future research should focus on establishing global data trading protocols, optimizing privacy-preserving mechanisms, and improving decentralized transaction models to enhance security and efficiency. Edge AI-enabled data trading will play an increasingly vital role in facilitating trustworthy and efficient data markets across digital industries.

Contents

* Corresponding author at: School of Artificial Intelligence, Guangzhou University, China.
E-mail address: kychen@gzhu.edu.cn (K. Chen).

https://doi.org/10.1016/j.cosrev.2025.100844
Received 17 May 2025; Received in revised form 24 September 2025; Accepted 12 October 2025
Available online 16 October 2025

1. Introduction

In the digital economy area, data has become a key production factor, providing core support for artificial intelligence, Internet of Things (IoT), smart cities, fintech, and healthcare. The global data market is experiencing rapid growth, with businesses and institutions increasingly relying on data-driven decision-making. The rise of data trading platforms offers new possibilities for data transaction and data trading [1,2]. However, traditional data trading models often depend on centralized platforms, leading to issues like data silos, privacy violence, and abnormal pricing, which severely restrict data transaction and data value exploration.

Data trading demand is an critical issues in IoT, smart cities, finance, and healthcare [3–5]. For instance, in IoT environments, large volumes of data generated by industrial sensors and smart devices need to be shared among entities to optimize production processes and improve efficiency. In smart cities, data sharing in areas like traffic and environmental monitoring can enhance urban management. Financial institutions rely on large-scale data analysis for risk control and credit evaluation, while the circulation of medical data is crucial for Artificial Intelligence (AI)-assisted diagnosis and precision medicine.

However, several key challenges remain in practical applications, including: *(1) Data privacy and security* [6–8]: Data trading involves the transmission of sensitive information, and it is a critical issue to ensuring data privacy while preventing data misuse or data leakage . Traditional data trading models, which rely on third-party platforms for matching and management, are vulnerable to data leakage and unauthorized use. Ensuring security in a decentralized environment is a major research focus. *(2) Data pricing problem* [9,10]: The value of data varies greatly depending on its source, quality, and the specific context. It is a complex challenge to establish scientifically sound pricing mechanism that ensures fair compensation for data providers while ensuring an appropriate return on investment for consumers. Traditional pricing models based on market supply and demand or cost-plus methods often fail to reflect the true value of data. *(3) Data trading efficiency* [11]: The computational and communication overhead in data trading is a key factor affecting market efficiency. It is crucial to enhance transaction efficiency in the data market, especially including reducing computation costs, optimizing data transmission paths, and improving transaction confirmation speeds in large-scale distributed data trading. Traditional platforms often require significant computational and storage resources, while decentralized architectures may introduce additional delays and overhead. *(4) Decentralized trading* [12,13]: Most current data trading platforms still rely on centralized entities for data storage and certification, leading to high risks of single points failure and data monopolies. Moreover, centralized platforms are vulnerable and non-transparency in the data trading process. Building decentralized trading platforms based on blockchain and smart contract mechanisms to enhance fairness and transparency in data transactions is a promising research direction.

Edge AI deploys computation locally on devices or edge servers to make data processing more efficient and secure, providing a novel solution for data trading [14]. The main roles of Edge AI in data trading include: *(1) Local intelligent computation to reduce data transmission and enhance privacy protection* [15,16]: Traditional data trading models often require data to be uploaded to the cloud for processing, leading to high transmission costs and privacy risks. Edge AI enables data processing on edge devices, with only model updates or encrypted summaries uploaded, thus reducing data transmission and enhancing privacy protection. This is especially important in fields like IoT devices, smart traffic, and medical diagnostics. *(2) Federated learning and privacy-preserving computing to enable secure data sharing* [17,18]: Federated learning, a decentralized machine learning technique, allows multiple data providers to collaborate on model training without sharing raw data, making it ideal for privacy-sensitive data trading scenarios. Additionally, privacy-preserving techniques like differential privacy and homomorphic encryption enable secure computation and data transaction without exposing the data itself. *(3) Distributed architecture to improve fairness and transparency in data trading* [19]: By combining blockchain and smart contracts, Edge AI can build decentralized data trading platforms that automate and audit the data trading process. Smart contracts can support data trading agreements, ensuring fair dealings for both parties, while the distributed ledger records transaction history to improve market transparency. This mechanism helps reduce the risk of monopoly by centralized entities and enhances fairness in data trading.

1.1. Comparison with existing works

While previous works have provided valuable insights into the application of Edge AI in specific domains such as financial technology and smart grids, they are often limited in scope. For instance, Andronie et al. [20] discuss the potential of generative AI integrated with blockchain in fintech management, while Zahid et al. [24] review smart grid architectures combining AI, blockchain, and digital twins. Zheng et al. [22] focus specifically on federated learning techniques for privacy preservation in energy systems, and Ali et al. [23] examine privacy-preserving machine learning approaches in IoT-based smart grids.

Table 1
Comparison with existing works.

Year	Reference	Contribution	Data processing	Data transactions	Data Pricing	Privacy Protection	Data Quality Assessment
2024	[20]	Discusses generative AI in IoT-blockchain fintech management	yes	yes	yes	no	no
2024	[21]	Examines ML in financial services including trading predictions	yes	no	no	no	no
2024	[22]	Focuses on privacy-preserving federated learning in power systems	yes	yes	no	yes	yes
2025	[23]	Review of privacy-preserving ML for IoT-smart grids	yes	yes	no	yes	yes
2025	[24]	Review of smart grid tech with AI, blockchain, digital twins	yes	yes	yes	yes	no
2025	[25]	Overview of high-frequency trading using Edge AI and fintech	yes	no	yes	no	no
2025	Our paper	Explores the integration of Edge AI with data trading, analyzing its technical foundations, key challenges, and diverse applications	yes	yes	yes	yes	yes

In contrast, as listed in Table 1, our paper provides several distinct advantages: *(1) Greater system scope:* We establish a comprehensive multi-dimensional framework that analyzes Edge AI-driven data trading from system architecture, trading mechanisms, incentive design, to privacy and quality assurance. *(2) Broader theme coverage:* Unlike studies that focus solely on privacy [23] or trading speed [25], our paper spans all five core dimensions: data processing, decentralized transactions, pricing strategies, privacy protection, and data quality assessment. *(3) Cross-domain perspective:* Our paper covers diverse application domains including finance, energy, supply chain, healthcare, and IoT, offering a horizontal comparative view. *(4) Dual-layered analysis:* We not only assess technical enablers but also investigate mechanism-level factors such as fairness, transparency, and regulatory compliance in data trading markets, which are largely underexplored in prior literature. Therefore, this work fills a critical gap in the literature by offering a holistic and integrative review of Edge AI-enabled data trading and provides a foundation for future theoretical development and practical deployment.

1.2. Our contributions

The general framework of our paper is illustrated in Fig. 1. It focuses on Edge AI-enabled data trading, addressing the following core topics: *(1) Fundamental concepts of data trading:* Analyzing the components of existing data trading systems and the advantages and disadvantages of centralized versus decentralized architectures. *(2) Core & supporting technologies:* Including data privacy protection, smart contracts, federated learning, blockchain, and how these technologies drive data trading. *(3) Edge AI's role in data trading:* Detailed discussion on the applications of Edge AI in data processing, privacy protection, and transaction optimization. *(4) Practical application scenarios of edge AI-enabled data trading:* Covering IoT, smart cities, finance, healthcare, and supply chain management data trading models. *(5) Challenges and future directions:* Exploring the trends in data trading and challenges such as standardization, privacy protection, computational efficiency, data quality assessment, and legal compliance.

The remainder of this paper is organized as follows: Section 2 introduces the fundamental concepts of data trading, Section 3 discusses the core technologies, Section 4 outlines the methods of Edge AI-enabled data trading, Section 5 presents the application scenarios, Section 6 covers research challenges and future directions, and finally, Section 7 provides the conclusion.

2. Fundamental concepts of data trading

2.1. Definition of data trading

Data trading refers to the process in which data owners or providers sell or share data through market mechanisms to data consumers, in exchange for economic or other compensation [26,27]. The data market serves as the primary platform for data trading, connecting both the supply and demand sides of data, while offering functionalities such as pricing, quality assessment, and privacy protection. It contains various stages including data collection, storage, processing, analysis, and trading, involving multiple stakeholders such as enterprises, governments, and research institutions. Early visions of data markets emphasizes the role of cloud platforms to facilitate data discovery, data pricing, data quality evaluation, and data integration [28]. These data marketplaces are further categorized with their business models and practical limitations [29]. Data Markets for machine learning training data are also proposed by [30], providing auction and matching mechanisms to handle data replicability and verification issues. Specifically, the theory of data pricing has been extensively explored prior to 2020. In early work, data privacy compensation has been considered as a key trading issue, with an auction mechanism under differential privacy guarantees [31]. Private data pricing is also integrated with the query accuracy to allocate micro-payments to data owners [32]. Further studies developed arbitrage-free pricing schemes for general queries, highlighting inherent trade-offs among universality, fairness, and no-arbitrage conditions [33].

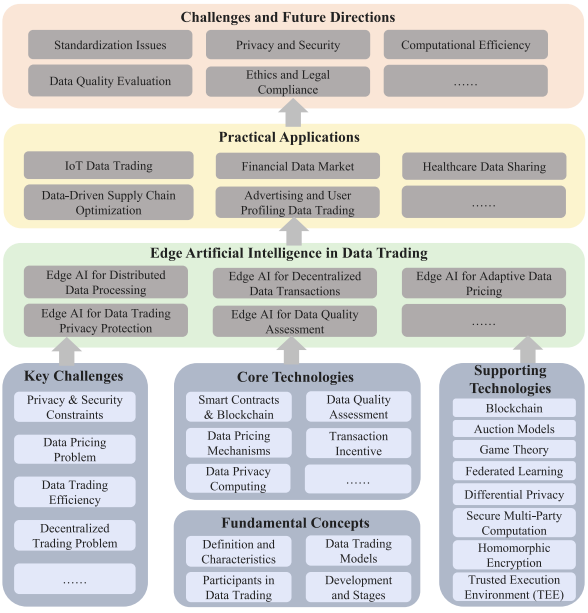

Fig. 1. The general framework of our paper.

In addition, data trading also involves challenges such as data governance, standardization, and compliance. As data flows across borders increase, the differences in laws and regulations between different countries and regions affect the feasibility of data trading. For example, the EU's GDPR [34] and China's Data Security Law present different compliance requirements for data trading. Establishing a unified global standard for data trading has become an important research direction. The literature [26] analyzes the legal challenges and difficulties in Chinese data trading markets, highlighting the importance of developing data trading models and implementing targeted legal regulations.

2.2. Characteristics of data trading

Data trading refers to a series of economic activities where data is exchanged between data suppliers and data consumers. Compared to traditional commercial trading, data trading has the following distinct characteristics:

(1) Intangibility and infinite replicability: Data has no physical form, and its value does not directly decrease with repeated use. On the contrary, it may generate new value through continuous integration, cleaning, and deep mining. As a result, data trading not only involves one-time transactions but also covers a long-term value transfer mechanism. The literature [35] identifies the basic characteristics of data as a production factor, including intangibility, replicability, and derivability. Due to its intangibility, data is difficult to physically control and transfer in the transaction process, unlike traditional tangible assets. At the same time, the infinite replicability of data means it can be copied and used multiple times without depleting the original data, presenting new challenges and opportunities for data trading. The literature [36] further points out that the intangibility and replicability of data make ownership definition difficult. Since data can be infinitely replicated, its ownership is hard to clearly define as that of traditional assets.

(2) Diversity and non-standardization: Due to the wide variety of data types and formats, standardization and quality assessment are often required during trading to ensure that both parties have a sufficient consensus on the data's content and value. The literature [37] focuses on the key concepts and mechanisms of data trading, highlighting the issues of diversity and non-standardization. With data coming from diverse sources and formats, data trading requires the establishment of unified standards and norms to ensure data quality and interoperability. The literature [38] explores the standardization of open data, noting that the diversity and non-standardization of data are major challenges faced by open data platforms. To address these challenges, it further proposes several solutions for data standardization to improve data usability and interoperability, including data quality assurance, metadata and traceability tracking, and data governance.

(3) Complexity of ownership: The issues of data ownership and intellectual property rights are particularly important in data trading. Data trading requires the supplier to clarify data ownership, usage rights, and revenue rights, and also addresses how to protect personal privacy and trade secrets during the transaction process. Thus, data trading is closely related to data ownership clarification, data quality certification, and legal compliance. The literature [39] identifies three major challenges in data transactions: technology, standards, and law, with ownership complexity being a core issue. It proposes a threefold governance system, consisting of technological methods, standardized norms, and legal protections, to effectively address the ownership complexity in data trading. The literature [26] analyzes the complexity of ownership in data trading, pointing out that the current legal framework is still underdeveloped, leading to difficulties in clarifying data ownership. It emphasizes that the legal regulation of data trading should consider data's intangibility and replicability, and clarify the rights and obligations in data transactions with reasonable legal arrangements.

2.3. The development of data trading

With the continuous advancement of information technology, data markets gradually emerged toward standard and scalable data trading [40] (see Table 2).

Phase 1 (1990s – early 2000s): Data collection and targeted market: In the early stages, data trading focused mainly on the collection of user data and targeted advertising. Companies use online behavioral data to optimize advertisement placements. Major players include search engine companies like Google and Yahoo, which provides advertising services based on users' search history.

Phase 2 (2000–2010): The rise of big data and data platform: With the development of cloud computing and big data technologies, data storage and analysis capabilities significantly improve, further enhancing the commercial value of data. Companies begin utilizing data platforms for trading, such as Data Management Platforms (DMP) and Customer Relationship Management (CRM) systems.

Phase 3 (2010–2020): Data market with privacy protection: As various data trading platforms emerged, data privacy protection regulations are introduced such as the EU's GDPR [34] and California's CCPA [41]. New technologies such as blockchain and privacy computing are applied in data trading to improve data security and traceability.

Phase 4 (2020 – present): Decentralized data trading and the AI data ecosystem: Data trading platforms have increasingly shifted toward decentralization, with blockchain technology widely adopted to ensure data traceability and security. The rapid development of AI has driven the transformation of the data market, leading to the emergence of AI training data markets and Data-as-a-Service (DaaS).

2.4. Data trading models

Depending on the trading entities, transaction methods, and trading environment, data trading models can be classified into the following main types:

(1) Centralized trading model: In the centralized trading model, transactions typically rely on specialized data trading platforms or exchanges. These platforms or exchanges are responsible for the full process management, including listing, auditing, pricing, matching, and settlement of data products. Centralized platforms such as AWS Data Exchange and Google Cloud Public Datasets offer efficient data storage and computing capabilities, ensuring data quality and security. Centralized trading models typically rely on a platform-mediated marketplace where a central entity provides catalogs, pricing, and value-added services [28]. Centralized trading markets usually maintain efficient data matching and pricing mechanisms [30]. However, centralized trading models also face several issues, including data monopolies, non-transparent pricing, and privacy leakage risks. The literature [26] points out that the lack of trust and robustness in centralized platforms in China's data trading market suppresses the free flow of data. The literature [40] also highlights that traditional centralized data trading platforms face trust and robustness issues in real-world applications such as industrial IoT.

(2) Decentralized trading model: With the development of technologies like blockchain and smart contracts, decentralized data trading models have gradually emerged. In this model, data transactions do not rely on a single centralized institution. Instead, it uses distributed ledger technology to record transaction information with a blockchain system, ensuring transparency, security, and openness in the transaction process. Decentralized platforms such as Ocean Protocol and SingularityNET leverage blockchain and distributed storage technologies to enable peer-to-peer data trading, enhancing transparency and data ownership control. Decentralized models emphasize distributed repositories and blockchain-based smart contracts to ensure trust and interoperability [42]. Practical applications in smart cities also demonstrate how decentralized marketplaces can address challenges in trust, interoperability, and incentives for data sharing [43]. The literature [12]

Table 2
The development of data trading.

Time period	Main tasks	Representative companies or products
1990s - early 2000s	Data collection and targeted marketing	Google, Yahoo
2000–2010	Big data and data platform	data management platforms, customer relationship management systems
2010–2020	Data market with privacy protection	data trading platforms
2020 - present	Decentralized data trading and AI data ecosystem	AI-enabled data markets and Data-as-a-Service

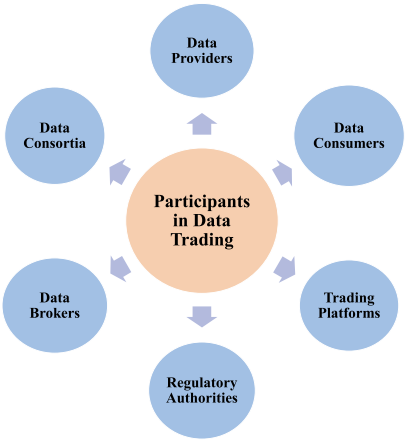

Fig. 2. Participants in data trading.

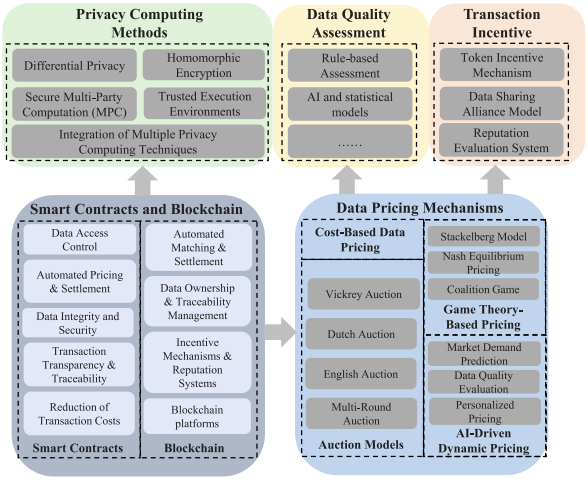

Fig. 3. Core technologies of data trading.

proposes a blockchain-based decentralized data trading model. It uses smart contracts to achieve secure storage and access of data, addressing the trust issues in traditional centralized data trading and improving transparency and security. The literature [44] introduces a decentralized data trading mechanism combining Hashed Timelock Contracts (HTLC) with blockchain technology, demonstrating its effectiveness in improving transaction security and efficiency through simulation experiments. However, it still face issues such as slower transaction speeds, smart contract vulnerabilities, and difficulty in ensuring data quality.

(3) Hybrid trading model: Some markets adopt a hybrid trading model to balance the advantages of centralized management and decentralization. This model introduces some decentralized technologies into centralized data trading platforms to automate and enhance transparency in part of the trading process, while still retaining centralized control over data quality auditing and risk management. The hybrid model can be flexibly adjusted to meet specific trading scenarios, meeting high-frequency, low-cost transaction needs while ensuring compliance and security during the trading process. Hybrid trading models combine centralized coordination with decentralized execution. For example, the literature [42] presents a smart contracts-based hybrid architecture with a central directory and decentralized data repositories. More practical hybrid configurations are for different market places are exploited in [29].

2.5. Participants in data trading

Data trading involves several key participants [12,40], as shown in Fig. 2, mainly including the following categories:

(1) Data providers: Data providers are entities such as enterprises, research institutions, and governments that hold data and sell data for profit. Data providers are responsible for ensuring the quality, compliance, and privacy protection of the data to meet transaction requirements.

(2) Data consumers: Data consumers are enterprises, researchers, developers, and other entities that need data for analysis, modeling, or decision optimization. They purchase data to support decision-making, improve product or service effectiveness, or conduct scientific research.

(3) Trading platforms: Trading platforms provide services such as data storage, transaction matching, pricing, and privacy protection, creating a secure and transparent environment for both data suppliers and consumers. Typical trading platforms include AWS Data Exchange and Ocean Protocol, which also offer data quality assurance and compliance auditing services.

(4) Regulatory authorities: Regulatory authorities are responsible for formulating rules for data trading, ensuring data security, privacy, and legality. International regulations such as the GDPR [34] and the CCPA [41] have strict requirements on the compliance and privacy protection aspects of data transactions.

(5) Data brokers: Data brokers are intermediaries responsible for facilitating the matching of data supply and demand. They often provide value-added services such as data cleaning, quality assessment, and privacy protection, helping to enhance the market liquidity and usability of data.

(6) Data consortia: A data consortium is a new cooperative model where multiple organizations jointly establish data-sharing platforms to reduce data silos and promote the exchange and sharing of data resources. Data consortia help gather diverse data from different departments, enhancing data variety and application scenarios.

3. Core technologies of data trading

In this section, we will explore the key technologies in the field of data trading (see Fig. 3).

3.1. Smart contracts and Blockchain technology

Traditional data transactions rely on centralized platforms for matching and supervision, but these platforms may suffer from issues such as data leakage, non-transparent pricing, and high intermediary fees. Blockchain technology addresses these issues by recording all transaction data in a distributed ledger, ensuring that every data transaction is traceable and immutable, thereby resolving issues of data ownership and traceability [52,53]. By leveraging blockchain, data trading platforms can transparently display the transaction process, reducing the

Table 3
Smart contracts and blockchain.

Technology	Highlight
Smart contract automation [45]	Proposes a technical framework integrating blockchain and big data to enhance data transaction accuracy, security, and scalability
Ethereum transaction behavior analysis [27]	Analyzes Ethereum transaction patterns to explore user behavior insights in the context of data trading
Automated matching and settlement [12]	Demonstrates how smart contracts ensure decentralized access and transparency in data trading systems
Data ownership and traceability management [44]	Uses hashed timelock contracts to enhance the trust, transparency, and traceability of data trading
Incentive and reputation mechanisms [13]	Explores blockchain's potential to build incentive-based and trustworthy data trading systems through DLT
Ocean protocol platform [46,47]	Offers a token-incentivized blockchain-based data marketplace with privacy-preserving smart contract governance
Datum platform [48]	Provides decentralized encrypted storage and user-controlled data authorization for secure data transactions
Streamr platform [49]	Supports real-time data streaming markets with blockchain-enhanced secure and scalable infrastructure
SingularityNET platform [50,51]	Facilitates decentralized AI-driven data trading using smart contracts to ensure fairness and optimize matching

risk of information asymmetry. Smart contracts, which are executable programmatic protocols on the blockchain, allow parties to automatically execute transactions once certain predefined conditions are met, without the need for intermediary intervention. This automation significantly reduces transaction time and costs while minimizing the risk of human error and fraud.

Thus, smart contracts combined with blockchain technology can create a trustless data trading mechanism that addresses these pain points [54,55,55] (see Table 3). Their main applications include: *(1) Data access control:* Smart contracts can be used to manage data access rights, ensuring that data is only used under authorized conditions. For instance, data providers can set access rules in the smart contract, allowing only those who meet payment conditions or other agreed terms to access the data. *(2) Automated pricing and settlement:* Smart contracts support predefined data pricing mechanisms and can automatically execute payments once a transaction is completed, eliminating the subjectivity and non-transparent in manual pricing. Furthermore, after contract execution, transaction records are automatically stored on the blockchain, ensuring traceability. *(3) Data integrity and security:* In traditional data trading, data providers often have to deliver data first and wait for payment, which exposes them to risks of misuse or theft. Smart contracts, combined with privacy-enhancing technologies like zero-knowledge proofs, can complete transactions without revealing the original data to ensure data security. *(4) Transaction transparency and traceability:* All transactions based on smart contracts are recorded on the blockchain, allowing any party to verify the transaction history of the data, preventing duplicate sales or malicious tampering. *(5) Reduction of transaction costs:* Traditional data trading typically requires third-party intermediaries, such as data trading platforms or data hosting agencies, which increase transaction costs. Smart contracts reduce intermediary steps through automated execution of trading agreements, improving transaction efficiency. For example, the literature [45] proposes a comprehensive technical solution combining blockchain and big data to support data transaction accuracy, security, and scalability. It contains five algorithms: decentralized storage solutions, smart contract automation, proof-of-work consensus mechanisms, sharding for scalability, and data encryption for privacy protection. The literature [27] focuses on analyzing Ethereum transaction behaviors, indirectly shedding light on the concept of data trading, where transaction data is used to analyze user behavior, providing a basis for behavior analysis in data transactions.

With blockchain's distributed ledger and immutability features, smart contracts in data trading are applied primarily in the following areas: *(1) Automated matching and settlement:* Smart contracts can match buyers and sellers in real-time and automatically complete payment and settlement processes, reducing the risks of human intervention and information asymmetry. For example, the literature [12] highlights how blockchain technology brings decentralization, transparency, and immutability to data trading. By using smart contracts to enable decentralized data access, it solves the trust and security issues in traditional data trading, enhancing transaction transparency and reliability. *(2) Data ownership and traceability management:* Blockchain records information on data generation, transmission, and transaction processes, ensuring transparency in data sources and ownership, providing both legal and technical guarantees for data trading. For example, the literature [44] emphasizes the importance of data trading in distributed networks, pointing out that traditional centralized methods have security vulnerabilities and lack trust among participants. By combining hashed timelock contracts with blockchain, the security and transparency of data trading can be enhanced, solving the reliability and trust issues in data trading. *(3) Incentive mechanisms and reputation systems:* Blockchain's transparent features allow for the construction of a transaction system based on credit and incentives, supervising and motivating the behaviors of data suppliers and consumers. For instance, the literature [13] discusses how blockchain shows practical service reliability in data trading and other fields.

Currently, several blockchain platforms are dedicated to building secure and efficient decentralized data trading markets: *(1) Ocean protocol platform [46,47]:* This platform offers a blockchain-based data marketplace, allowing data providers and consumers to trade directly. It uses smart contracts to manage data access rights and ensure data privacy, while token incentives encourage data flow and increase market activity. *(2) Datum platform [48]:* This platform ensures data security and traceability using blockchain technology. Users have control over their data and can authorize transactions through encryption. It uses decentralized storage to enhance data security. *(3) Streamr platform [49]:* This platform provides a real-time data streaming marketplace suitable for IoT, finance, and other scenarios. By combining blockchain and decentralized computing, it enables secure data sharing. The distributed network reduces dependency on centralized servers to enhance scalability. *(4) SingularityNET platform [50,51]:* This platform focuses on AI data trading, enabling decentralized AI service transactions through blockchain. It uses smart contracts to ensure fair transactions between data providers and consumers, optimizing the data trading matching mechanism with AI models.

3.2. Data pricing mechanisms

Data pricing is a core element of data trading, as a reasonable pricing mechanism guarantees fair transactions for data consumers. Due to the intangible nature of data, its infinite replicability, and its dynamic value changes, traditional pricing methods are often inadequate. Current research on data pricing primarily focuses on the following approaches (see Table 4):

Cost-Based Approaches: This method considers the costs incurred during data collection, cleaning, storage, and processing as the fundamental basis for pricing. While cost-based pricing provides a lower limit

Table 4
Data pricing mechanisms.

Technology	Highlight
Vickrey auction [56]	Establishes the second-price auction mechanism where bidders are incentivized to bid their true value
Dutch auction [57]	Describes a descending price auction suitable for real-time or perishable data transactions
English auction [58]	Analyzes the efficiency and welfare outcomes of ascending-bid auctions in competitive markets
Multi-Round auction [59]	Introduces multi-round bidding to improve fairness and flexibility in data transactions
Stackelberg model [60,61]	Applies a leader–follower game model where data providers set prices and consumers respond strategically
Nash Equilibrium pricing [62]	Determines prices based on mutual best responses of buyers and sellers with no incentive to deviate
Coalition game [63]	Models cooperative pricing strategies among data providers using value-sharing principles
Data quality evaluation [64]	Uses ML to assess data integrity, timeliness, and accuracy for automated pricing decisions
Market demand prediction [65]	Predicts future data demand using AI to dynamically adjust data prices
Personalized pricing [60]	Implements deep reinforcement learning to customize prices based on buyer profile and demand
Blockchain-based pricing model [60]	Combines Stackelberg game and DDQN reinforcement learning for secure and incentive-compatible pricing
Negotiation-based dynamic pricing [65]	Proposes a smart contract-driven pricing scheme responsive to requester needs and market conditions
Three-party pricing framework [61]	Combines Stackelberg and bargaining games to support dynamic customized data pricing
Learning-based auction mechanism [66]	Uses deep learning to solve optimal multi-item auction designs, improving pricing mechanism flexibility

for the data price, it does not account for the added value that data may generate during its use.

Auction Models: Auction-based pricing is a common approach in data markets, especially those with high competition [9]. Common auction models include the following: *(1) Vickrey auction (also known as a second-price auction) [56]:* In this model, buyers bid, and the highest bidder wins, but pays the second-highest bid. This mechanism encourages buyers to bid based on the true value of the data, minimizing speculative bidding. Bidders are motivated to bid according to their true valuation to maximize their utility under this model. *(2) Dutch auction [57]:* The seller starts with a high price and gradually lowers it until a buyer accepts. This model is suited for time-sensitive data, such as real-time sensor data. *(3) English auction [58]:* A traditional ascending-bid auction suitable for highly competitive data markets. The literature [58] analyzes the efficiency of English auctions in markets with a large number of participants and shows that they can achieve near-optimal social welfare under certain conditions. *(4) Multi-round auction [59]:* This model incorporates multiple bidding rounds, enabling data buyers and sellers to engage in fairer transactions.

Game Theory-Based Pricing: Game theory provides a market equilibrium-based approach to pricing in data markets [10]. Several game theory models include: *(1) Stackelberg model [60,61]:* This model is applicable to monopolistic data markets where the data provider sets the price, and the data consumers decide whether to accept the price. *(2) Nash equilibrium pricing [62]:* Pricing is determined based on supply and demand, where both the data buyer and seller have no incentive to change their strategies unilaterally. *(3) Coalition game [63]:* This model applies to scenarios where multiple data providers collaborate to set prices, such as in data alliances or federated learning settings.

AI-Driven dynamic pricing: AI plays an important role in dynamic and precise data pricing. The main AI-driven pricing methods include: *(1) Market demand prediction [65]:* AI analyzes historical transaction data to predict future data demand and adjust prices dynamically. *(2) Data quality evaluation [64]:* Machine learning models assess factors like data integrity, accuracy, and timeliness to automatically adjust prices. The method proposed in [64] evaluates these data factors and provides a basis for pricing. *(3) Personalized pricing [60]:* Using deep reinforcement learning, this method adjusts prices based on the buyer's needs and payment capacity, enabling customized transactions.

In practice, a combination of multiple pricing models is often used. For example, the literature [60] proposes a blockchain-based IoT data trading pricing mechanism. To encourage better data circulation and maximize data value, the mechanism builds a pricing model between data owners and data requesters. This model uses blockchain architecture to ensure transaction data security and meet arbitration needs, while the pricing scheme based on data quality lays the foundation for future dynamic pricing. It also introduces a Stackelberg game-based pricing model to incentivize both parties to participate actively and maximize utility. Additionally, the literature incorporates an improved Double Deep Q-Network (DDQN) reinforcement learning algorithm to optimize the pricing strategy for data transactions. The literature [65] discusses the shortcomings of existing pricing mechanisms in data markets and proposes a new pricing scheme based on data requester demand. Existing pricing models are divided into static and dynamic pricing. Static pricing lacks adaptability and cannot respond flexibly to market changes and demand fluctuations, while dynamic pricing is primarily driven by data owners, often overlooking the interests of data requesters. The new pricing scheme involves three entities: Data Owners (DO), Data Requesters (DR), and Smart Contracts (SC). Through smart contracts, DOs and DRs can negotiate and reach a mutually satisfying pricing agreement, which is designed to set prices according to DR demand while ensuring DO interests, thereby improving transaction success rates and market competitiveness. Furthermore, the literature [61] proposes a three-party dynamic pricing mechanism for customized data trading. By combining the Stackelberg game and bargaining game theories, the framework is designed to address the complexity of customized data product transactions. The platform guides the bargaining process between buyers and sellers to determine the final transaction price. The literature [66] models auctions using a multi-layer neural network and treats optimal auction design as a constrained learning problem, solved using a standard machine learning pipeline. This approach allows for highly accurate recovery of analytical solutions for all known multi-item auction settings, thus presenting new mechanisms for pricing optimization.

3.3. Data privacy computing methods

With the rapid development of the data trading market, the issue of data privacy protection has become increasingly important. Traditional data trading models often involve centralized storage and sharing of data, which not only increases the risk of data leakage but also limits the willingness of data providers to share their data. Privacy computing technologies provides a more secure and efficient solution for data trading, allowing parties to compute and trade data without exposing the raw data itself [7]. The core goal of privacy computing is to maximize the protection of data owners' privacy while ensuring that the value of the data can still be utilized. In the context of data trading, privacy computing plays a critical role in enabling the secure sharing and computation of data, ensuring data confidentiality during transmission while maintaining the usability of the data for analysis. By using advanced cryptographic methods and secure hardware, privacy computing technologies allow data buyers and sellers to reach agreements in an environment of mutual distrust without exposing the original data [83]. Privacy computing encompasses key technologies

Table 5
Data privacy computing.

Technology	Highlight
Differential privacy [67–70]	Adds noise to query results to prevent identification of individuals, enabling secure statistical data analysis
Secure multi-party computation [71–73]	Enables joint computation across multiple parties without revealing their individual data
Homomorphic encryption [74]	Allows computations on encrypted data to yield encrypted results without decryption
Encrypted data pricing [75,76]	Supports privacy-preserving dynamic pricing and evaluation of encrypted data
Trusted execution environment [77–79]	Provides hardware-based isolation for secure data and computation even if OS is compromised
Integrated privacy technologies [80–82]	Combines privacy computing, blockchain, and federated learning to build secure, efficient data trading ecosystems

such as differential privacy, secure multi-party computation, homomorphic encryption, and trusted execution environments [84] (see Table 5).

Differential Privacy (DP): Differential privacy is a mathematical framework to provide privacy protection during data analysis, making it difficult for attackers to infer individual information from the analysis results [67,68]. The core idea is to add random noise to the query results, ensuring that even if a single data point is removed, the overall analysis results will not significantly change. Differential privacy has a wide range of applications in data trading [69,70]. For example, data providers can apply differential privacy techniques to data before selling it, allowing data consumers to perform statistical analysis without accessing the raw data. Government agencies can also use differential privacy to anonymize demographic data, balancing data sharing with privacy protection. Furthermore, online advertising recommendation systems can use local differential privacy to collect user behavior data while preventing privacy leaks.

Secure Multi-Party Computation (MPC): MPC is a cryptographic method that allows multiple data holders to collaboratively compute a function's output without exposing their private data. Unlike traditional data sharing methods, MPC allows for joint analysis and modeling without data leaving local storage [71]. In data trading scenarios, MPC can be used by multiple data providers to jointly train machine learning models without actually sharing their data [72,73]. For instance, hospitals can use MPC to collaboratively train disease prediction models while ensuring patient data privacy. Similarly, financial institutions can apply MPC for joint risk control analysis, enhancing fraud detection accuracy while safeguarding data privacy. MPC can also be used in auction scenarios within data trading markets, allowing parties to engage in fair transactions without knowing the bids of other participants.

Homomorphic Encryption (HE): Homomorphic encryption is a specialized cryptographic technique that allows computations to be performed directly on encrypted data, enabling correct results to be obtained without decrypting the data. This feature enables data owners to provide encrypted data to data consumers, who can perform analysis tasks without decrypting the data [74]. In data trading, homomorphic encryption can support privacy-preserving data pricing and computation [75,76]. For example, in a dynamic pricing model based on data quality, data consumers can evaluate the quality of data while it remains encrypted, preventing the exposure of raw data. Moreover, cloud computing environments can leverage homomorphic encryption for data analysis tasks, where data owners upload encrypted data to the cloud, and the cloud server performs computations on the encrypted data and returns encrypted results, ensuring data remains secure.

Trusted Execution Environment (TEE): A TEE is a hardware-based security technology that provides a secure execution area, allowing sensitive data and computational tasks to be processed in a protected environment. The key advantage of TEE is its hardware-level security protection, ensuring that data and computations remain secure even if the operating system or other software is compromised [77]. In the data trading market, TEE can be used to create trusted data markets, where data providers allow data analysis to be executed only within the TEE, without exposing the raw data [78,79]. For instance, on a data-sharing platform, data providers can store encrypted data in the cloud, and data consumers can only run pre-configured analysis tasks within the TEE and receive the results without accessing the raw data. This mechanism not only ensures data security but also enhances the transparency and auditability of data trading.

Integration of Multiple Privacy Computing Techniques: The application of privacy computing in data trading faces several challenges. First, homomorphic encryption has high computational overhead, requiring further algorithm optimization to improve efficiency. Second, MPC incurs high communication costs when multiple participants are involved, calling for efficient protocols to reduce computational overhead. Additionally, the noise mechanism in differential privacy requires more precise parameter optimization to balance between privacy protection and data usability. In the future, privacy computing development may include combining with blockchain technology to build decentralized, trusted data markets, enhancing the transparency and security of data trading [80]. Furthermore, integrating federated learning technologies can further improve privacy-preserving machine learning capabilities, enabling multiple data holders to perform collaborative modeling without sharing data [81,82]. These emerging technologies are expected to drive the data trading market toward a more secure and efficient future.

3.4. Data quality assessment

Data quality directly influences the value and trust between parties in a data transaction. In the process of data exchange, data products are typically evaluated based on accuracy, completeness, consistency, and timeliness. To ensure that listed data truly reflects its utility, data trading platforms must implement a comprehensive data quality assessment and standardization system. The rapid development of the data trading market calls for a more precise and intelligent approach to data quality evaluation. Traditional methods of data quality assessment often rely on human expertise or simple rules, but modern AI and statistical techniques offer more automated, flexible, and efficient assessment methods. Particularly when dealing with large-scale, complex, and diverse datasets, traditional approaches often fall short in meeting the demands for both timeliness and accuracy. Therefore, the use of AI and statistical models to perform reliable data assessments has become one of the core technologies for ensuring the secure and efficient operation of data trading platforms.

For example, the literature [85] reviews the current state and challenges of data quality assessment, emphasizing its critical role in data trading. It proposes a comprehensive framework for evaluating data quality across multiple dimensions, such as accuracy, completeness, consistency, timeliness, and reliability. It also discusses how machine learning and statistical methods can be employed to automate the data quality evaluation process. The literature [86] addresses a significant issue in data quality: data imbalance. It introduces various techniques to handle imbalanced datasets, aiming to improve the performance of machine learning models. These techniques include resampling, cost-sensitive learning, and ensemble methods. The literature also explores how data imbalance can be assessed and managed in the context of data trading. Data provenance, a key component of data quality assessment, is reviewed in [87]. Data provenance helps users understand the origins and processing history of data, allowing for a better evaluation of its trustworthiness and quality. The literature discusses various applications of data provenance, including its role in data quality assessment, data governance, and data trading. The literature [88] discusses data

quality evaluation and record linkage techniques. It examines multiple facets of data quality, including data cleaning, standardization, and integration, and describe how record linkage can enhance the completeness and accuracy of data. The impact of data cleaning on machine learning classification tasks is investigated in [89]. The literature proposes a method for evaluating the effectiveness of data cleaning to improve data quality, and discusses how these techniques can be applied in data trading. Furthermore, the issue of query pricing on incomplete data is explored in [89]. The authors introduce a data-quality-based pricing model that accounts for the completeness and accuracy of data. They also examine how data quality assessment can optimize query pricing strategies, which is essential for pricing mechanisms in data trading. A review of data management issues in machine learning, including data quality assessment, preprocessing, and optimization, is provided in [90]. The authors highlight the impact of data quality on machine learning model performance and propose various methods for managing data quality. The literature also explores how these methods can be applied in data trading to enhance data quality. In [91], the authors describe Google's data lake management system, Goods, emphasizing the significance of data quality management in large-scale data storage and processing. The literature discusses best practices in data quality assessment, cleaning, and optimization, offering valuable insights for data quality evaluation in data trading. The study in [92] focuses on data and lifecycle management issues in deep learning, with particular emphasis on data quality assessment. The authors propose a comprehensive framework for evaluating and optimizing data quality to enhance the performance of deep learning models. The literature also addresses the application of these techniques in the context of data trading. Finally, the review in [64] explores techniques for optimizing deep learning on GPUs, with a strong focus on the impact of data quality on model performance. The literature discusses data preprocessing, augmentation, and cleaning methods, all of which are critical for improving data quality. It also examines how these methods can be applied to optimize data quality in data trading.

3.5. Transaction incentive mechanism

The successful operation of the data trading market relies on active participation and a high level of trust between transaction parties. Therefore, platforms often need to design scientific and reasonable incentive mechanisms and trust systems to encourage data providers to continuously improve data quality, while ensuring the transaction rights of data users. The main components of these mechanisms are as follows:

Token Incentive Mechanism: A token incentive mechanism refers to the use of blockchain-based tokenization to motivate data providers, consumers, and platforms to participate in data trading. Tokens not only serve as a medium for transactions but also incentivize market participants to provide high-quality data, engage in data transactions, and promote the long-term development of the platform. This mechanism offers economic incentives through digital currencies (or tokens), rewarding participants for their involvement in data trading. As a special type of digital asset, tokens can be used as a medium of exchange during transactions, as well as a component of participation and reward mechanisms, facilitating the healthy operation of the market. The token incentive mechanism comprises three parts: (1) *Incentives for data providers:* Data providers often face issues such as data privacy leakage and misuse when uploading their data to data markets. The token incentive mechanism, leveraging encryption and smart contract technologies, provides data providers with a transparent and secure transaction environment. By publishing data and engaging in transactions, providers can earn token rewards, motivating them to actively participate in data sharing. (2) *Incentives for data consumers:* In data transactions, tokens serve as a payment method, reducing the trust cost of the transaction. Consumers can pay tokens based on factors such as data quality and value, while the data quality assessment mechanism ensures that consumers receive valuable data. (3) *Incentives for platforms:* As intermediaries, data trading platforms are responsible not only for providing transaction and data storage infrastructure but also for maintaining platform operations, data quality, and data transmission. The token incentive mechanism encourages platform operators to improve service quality and ensure fairness and transparency in data trading.

For example, the literature [93] presents a blockchain-based secure data trading ecosystem. This system ensures the transparency and security of data transactions through smart contracts and blockchain technology. The literature discusses how the token incentive mechanism facilitates transactions between data providers and consumers, enhancing market activity. The literature [1] investigates how to achieve data authenticity and privacy protection in data markets. The authors propose a blockchain-based incentive mechanism that rewards data providers with tokens for supplying high-quality data, while employing privacy protection techniques to safeguard data. In [94], the authors explore the application of blockchain and federated learning in industrial IoT, particularly for data sharing and privacy protection. They propose a blockchain-based incentive mechanism that rewards devices with tokens for providing data, while using federated learning techniques to protect privacy. The research in [95] examines the optimal contract mechanism in online data trading markets. The literature introduces a blockchain-based incentive mechanism that rewards data providers and consumers with tokens through smart contracts. It also discusses how contractual mechanisms can optimize data transaction efficiency and returns. Finally, the literature [96] studies how data authenticity and privacy protection can be achieved in data markets. The authors propose a blockchain-based incentive mechanism that encourages data providers to supply high-quality data through token rewards and incorporates privacy protection techniques to safeguard the data.

Data Sharing Alliance Model: The data sharing alliance model is an incentive mechanism based on collaboration and sharing. It involves cooperation between different data providers, consumers, and platforms to form a collective entity that aims to maximize data sharing and value creation. Unlike traditional centralized data exchange models, it emphasizes decentralized governance and the principles of the sharing economy. Its goal is to enhance the efficiency and transparency of data transactions through win-win cooperation. In this model, multiple data providers and consumers form an alliance based on common interests and objectives. While sharing data, the members of the alliance regulate data exchange and value distribution through internal mechanisms. Through this approach, alliance members can minimize the costs of data transactions while ensuring data privacy and security.

The core characteristics of the data sharing alliance model include the following: *(1) Decentralized management:* The operation of a data sharing alliance typically employs decentralized management, reducing dependence on a single platform or intermediary. Alliance members manage autonomously through technologies such as smart contracts and blockchain, ensuring transparency, fairness, and efficiency in data transactions. *(2) Collaborative sharing:* Data providers within the alliance can share their data, while consumers can purchase data or participate in data exchanges based on their needs. By establishing shared agreements, the alliance facilitates the efficient flow of data and value creation. *(3) Data governance and compliance:* The data sharing alliance not only focuses on data exchange but also on data governance and compliance management. Alliance members must adhere to privacy protection policies and relevant laws and regulations to ensure compliance in data transactions. *(4) Incentive mechanisms:* Similar to token incentive mechanisms, the data sharing alliance employs various incentive measures (such as token rewards and data contribution rewards) to encourage data sharing and transaction participation. Members of the alliance can earn rewards through activities such as sharing data and ensuring data quality.

For instance, the literature [11] provides a comprehensive overview of data markets, discussing various aspects such as data search, productization, trading, pricing, revenue distribution, and privacy, security, and trust issues. It explores the data market policies and industry statuses in different countries and sectors, highlighting the challenges and future directions of data markets. It emphasizes the potential of the data sharing alliance model in enhancing the efficiency and transparency of data transactions. In the literature [97], a secure decentralized data trading alliance (SDTA) is proposed for the sharing and trading of electronic medical data. SDTA ensures the security and transparency of data transactions through blockchain technology, preventing data storage redundancy and recording data summaries and authentic transactions via a chain network. This mechanism is particularly suitable for privacy protection and efficient sharing of medical data. The study in [98] examines the challenges and research directions in data trading and monetization, with a particular focus on the technical and organizational challenges faced by data markets. Despite the important role of data markets in data transactions, data privacy protection and logistics issues remain major obstacles. The literature also discusses how the data sharing alliance model can resolve these challenges through decentralized mechanisms. In [99], the authors explore the interactions of data, organizations, and market factors in data sharing practices. The literature analyzes the applications of the data sharing alliance model in different scenarios, including open data, reciprocal sharing within ecosystems, data trading through data markets, and bilateral cooperation for specific purposes such as reducing carbon emissions. The study underscores the advantages of the data sharing alliance model in enhancing data transaction efficiency and transparency. The literature [100] presents a data exchange framework based on fairness and incentive compatibility, emphasizing the theoretical foundations of the data sharing alliance model in the data economy. Through game theory and mechanism design, the literature explores how decentralized governance mechanisms can achieve efficient data sharing and value maximization. Finally, the literature [2] explores the concept of federated data marketplaces, emphasizing their application in privacy-preserving data sharing. Federated data marketplaces allow organizations to collaborate on data analysis without sharing raw data, leveraging federated learning or other privacy-preserving technologies for efficient data sharing. The literature also discusses how the data sharing alliance model can enhance the efficiency and transparency of data transactions within federated data markets.

Reputation Evaluation System: A reputation evaluation system assesses market participants through a multi-dimensional factors, including historical transaction behavior, user reviews, and data quality. It helps reduce information asymmetry and transaction risks, thereby promoting efficient market operation. The reputation evaluation not only reflects the quality of traders' behavior but also provides a transparent and fair trading environment for market participants and strong support for platform risk management. The reputation evaluation system primarily quantifies the behavior of market participants, establishing a scoring system based on indicators such as historical transactions, user reviews, and data quality. By scoring the credibility of both parties in a transaction, the system creates a transparent credit profile that reduces information asymmetry in transactions, enhances transaction efficiency, and boosts market activity. The core goal of the reputation evaluation system is to establish a credit rating for market participants, reducing platform operation risks, increasing trust in the market, and encouraging positive trading behavior through appropriate incentives. The platform can offer participants preferential matchmaking and discount pricing based on their reputation scores, motivating them to improve service quality and enhance the platform's appeal.

The reputation evaluation system typically includes the following key indicators: *(1) Historical transaction record:* The historical transaction data of traders forms the foundation of the reputation evaluation system. By evaluating indicators such as transaction success rate, transaction amount, and frequency, the system can objectively reflect the trader's activity level and reliability. *(2) User reviews:* Buyers and sellers within the platform can rate or review transactions. These ratings influence the trader's reputation score and include feedback on data quality, delivery time, and transaction smoothness. *(3) Data quality indicators:* Data quality is a core factor in data transaction markets. The platform can quantify data quality based on factors such as completeness, accuracy, and timeliness, helping to assess the credibility of data providers. *(4) Risk management indicators:* The reputation evaluation system can also consider the platform's risk management needs, such as user default rates and the frequency of transaction disputes. These help the platform identify potential high-risk traders.

When building the reputation evaluation system, platforms typically assign weighted scores to different dimensions to create a comprehensive rating system. Specifically, the platform calculates the reputation score based on each participant's transaction history, user reviews, and data quality scores. For example, in [101], a fair and secure data transaction scheme is proposed, suitable for untrusted data buyers and sellers. The literature proposes a new concept based on probabilistic commutative encryption, allowing the decryption sequence of probabilistic and deterministic encryption algorithms to be interchangeable. This scheme meets both parties' security requirements under malicious models, ensuring fairness for both data buyers and sellers. Moreover, it allows partial data transactions to be outsourced to semi-honest cloud storage platforms. In [11], a comprehensive review of data markets is conducted, discussing various aspects of data transactions, including data search, productization, pricing, revenue distribution, and privacy, security, and trust issues. The literature also studies the data market policies and industry status in different countries and sectors, highlighting the importance of reputation evaluation systems in reducing information asymmetry and transaction risks. In [102], a blockchain-based data security trading method in distributed computing is proposed. The literature discusses how blockchain technology can ensure the security and privacy of data transactions while using zero-knowledge proof technology to guarantee transaction transparency and credibility. It also introduces a reputation evaluation mechanism based on smart contracts, which records historical behavior and user reviews to reduce transaction risks. In [103], a multi-level data sharing mechanism based on blockchain and smart contracts is explored. The literature presents a framework that uses blockchain for data sharing transparency and security, with smart contracts automating data transactions. The literature also discusses how the reputation evaluation system records the historical behavior of both parties in transactions, providing a transparent and fair trading environment for market participants. In [104], it presents a blockchain-based privacy protection and incentive mechanism for private data sharing in IoT. The literature discusses how blockchain technology can ensure the security and privacy of data sharing while using token incentives to encourage data providers to offer high-quality data. It also presents a reputation evaluation system to reduce transaction risks by recording historical behavior and data quality. In [105], an accountable and efficient data sharing scheme based on blockchain for industrial IoT is proposed. The literature explores how blockchain technology ensures data sharing transparency and security, with smart contracts automating data transactions. Additionally, it introduces a reputation evaluation system to record historical behavior and data quality, offering a transparent and fair trading environment for market participants.

4. Edge artificial intelligence in data trading

With the rapid growth of data volume and the fast development of the data trading market, efficient data processing and utilization have become critical issues in data transactions. Traditional data center architectures face computational bottlenecks when handling large-scale data, especially in scenarios requiring high real-time performance, bandwidth efficiency, and low latency. The centralized computing model of central servers and cloud computing platforms struggles to meet these demands [15]. Consequently, edge AI have emerged as promising solutions, offering new possibilities for data processing in data transactions [106] (see Fig. 4).

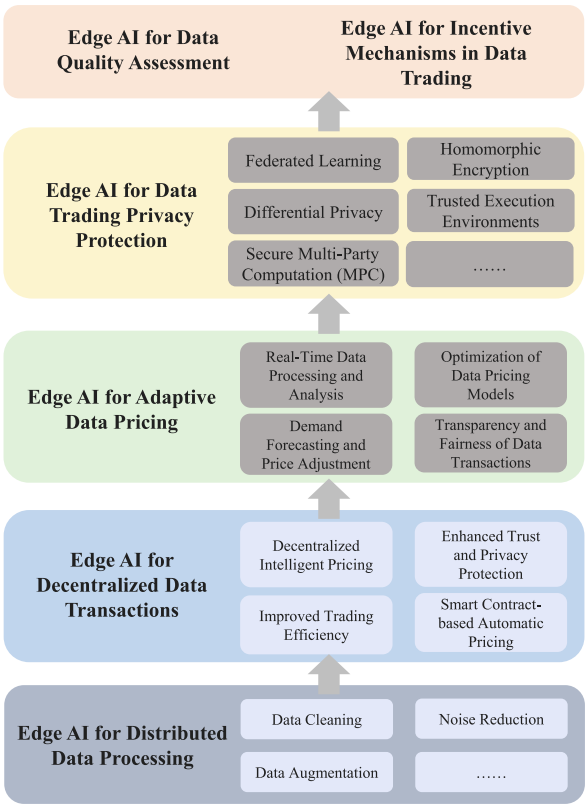

Fig. 4. Edge artificial intelligence in data trading.

4.1. The development of edge AI

Edge computing can be traced back to early visions that attempted to bridge the gap between mobile devices and centralized cloud services. The literature [107] introduces the *cloudlet* concept by deploying small-scale data centers close to end devices to reduce latency. Around the same time, the literature [108] extends cloud capabilities to the network edge and emphasizing low-latency, location-aware, and large-scale distributed support for the IoT. In industrial area, ETSI MEC white paper [109] provides the first standardized framework for mobile edge computing and defines its architecture, use cases, and deployment scenarios. The literature [110] further offers a comprehensive vision of edge computing, outlining its applications and research challenges such as a bandwidth bottlenecks, energy constraints, and privacy issues. The literature [111,112] exploit practical design challenges such as offloading, resource allocation, and quality of service on mobile edge computing, respectively. The literature [113] introduces *opportunistic edge computing* where end-user devices contribute resources to form scalable edge infrastructures. Other open problems in resource orchestration, security, and scalability are also discussed in [114,115]. More recently, the concept of *edge intelligence* has integrated artificial intelligence capabilities into edge nodes for on-device learning, inference, and adaptive services [116], enabling intelligent, secure, and AI-empowered edge ecosystems [117].

4.2. Edge AI for distributed data processing

Edge computing shifts data processing from the cloud to the network edge, enabling computations and analysis closer to the data source. This significantly reduces data transmission requirements while enhancing real-time processing and efficiency. Building upon this foundation, Edge AI integrates artificial intelligence into edge computing, allowing local devices to analyze data and make decisions through intelligent algorithms and models. This approach not only improves processing capabilities and accuracy but also alleviates transmission pressure, optimizing data utilization and supporting the efficient operation of data trading platforms. A key advantage of Edge AI in data trading is its ability to enhance critical pre-processing steps such as data cleaning, noise reduction, and data augmentation. By performing these tasks locally, Edge AI minimizes the amount of raw data that needs to be transmitted, thereby improving system efficiency. The literature [118] provides a comprehensive review of edge computing and Edge AI, highlighting the significance of data pre-processing, noise reduction, and enhancement. The study emphasizes that executing these processes on edge devices can significantly reduce data transmission overhead and enhance processing efficiency. Additionally, it explores feature compression techniques that help reduce data size and lower model resource requirements, which is particularly crucial for resource-constrained edge devices. To address data cleaning challenges in edge intelligence scenarios, the literature [119] introduces FedClean, a federated data cleaning protocol designed to preserve data privacy while filtering out anomalous data entries. Similarly, the literature [120] investigates feature extraction and compression techniques for Edge AI, demonstrating how these methods can optimize data processing workflows. By designing transmission-friendly and accuracy-preserving feature compression mechanisms, it ensures efficient data analysis even on resource-limited edge devices. Further emphasizing real-time data processing, the literature [121] explores how Edge AI reduces data transmission latency while improving processing efficiency. This study discusses how intelligent algorithms and models enable local data analysis and decision-making, streamlining data pre-processing, noise reduction, and augmentation. The literature [122] examines methods for enhancing data quality through Edge AI. By designing multi-precision AIGC services on edge servers, it supports adaptive image quality optimization under varying network conditions and dynamic tasks, ultimately improving data accuracy and usability for high-quality data trading.

4.3. Edge AI for decentralized data transactions

4.3.1. Edge AI and blockchain-based decentralized data trading mechanisms

Edge AI, by performing data processing and real-time analysis on edge devices, effectively reduces latency and improves data utilization efficiency. Blockchain, on the other hand, ensures the transparency, verifiability, and security of data transactions, addressing trust issues in the market. By combining Edge AI with blockchain technology, we can leverage the strengths of both to create a more trustworthy and efficient data trading market. Specifically, the combination of the two offers the following advantages: *(1) Decentralized intelligent pricing:* Edge AI can process data and make pricing decisions in real-time on edge devices, while blockchain ensures the transparency and immutability of the transaction process. This enables the data trading market to achieve decentralized intelligent pricing, avoiding the risks of centralized platforms. *(2) Improved trading efficiency:* Edge AI reduces data transmission latency and improves data processing efficiency, while blockchain ensures fast transaction confirmation and secure storage of transaction records. This enables data trading to take place in a highly efficient and secure environment. *(3) Enhanced trust and privacy protection:* Blockchain's trustless mechanism and encryption techniques effectively address privacy protection issues in data trading. Edge AI can process sensitive data locally, avoiding exposure to external servers, further enhancing data privacy. *(4) Smart contract-based automatic pricing:* Smart contracts in blockchain can automatically execute pre-defined contract terms. In data trading, smart contracts can adjust prices based on Edge AI's pricing results and execute transactions when specific conditions are met.

In recent years, many studies have explored the feasibility and optimization strategies of integrating Edge AI with blockchain for

data trading mechanisms. The literature [123] proposes a blockchain-based edge computing data trading model to improve the efficiency and security of data trading. The literature introduces a data relay and trading model involving data producers, relays, and consumers. It also introduces a new consensus mechanism, Proof-of-Data-Trading (PoDT), which combines Proof-of-Work (PoW) and Proof-of-Stake (PoS) to reduce the energy consumption of edge devices. The literature [19] discusses blockchain and edge AI integration for trustworthy and secure consumer applications. This study proposes a blockchain-based edge intelligence system using both public and private blockchains to address the shortcomings of traditional systems. Public blockchains ensure the privacy and security of consumer electronic device (CED) data communications, while private blockchains ensure secure communication between edge intelligence servers. The literature [124] proposes a blockchain-based machine learning framework for edge services in Industrial IoT. This framework encourages multi-party participation in edge services by constructing new smart contracts to improve data processing efficiency. The literature [125] reviews the terminal-edge-cloud collaborative computing (TECC) paradigm based on blockchain. It discusses how to alleviate edge resource pressure through lightweight blockchains and optimized resource allocation strategies. Additionally, it explores how smart contracts, identity authentication, and encryption techniques can enhance the security of the blockchain-based TECC architecture. The literature [126] investigates the integration of edge computing and blockchain technology in digital transportation, particularly for autonomous vehicle networks. It proposes a framework that includes the deployment of edge computing nodes along key traffic routes and utilizes blockchain for secure data exchange and smart contract-based automatic transactions. The literature also discusses challenges such as scalability, data synchronization, and network security, and examines the regulatory impact of these innovations. The literature [127] proposes a real-time IoT data processing framework called VCD-TSNet, which combines blockchain and edge computing technologies. This framework integrates deep learning models (such as VGG, ConvLSTM, and DNN) for spatial feature extraction, temporal modeling, and decision-making, while using blockchain to ensure data integrity and privacy. The literature [128] explores the application of blockchain in collaborative edge intelligence, ensuring fair incentives for data providers, optimizing data quality, and improving system efficiency, thereby promoting collaboration and data flow in edge intelligence systems. Specifically, devices improve service quality by sharing resources and knowledge. To prevent free-riding attacks (where devices benefit from the system without contributing), the system uses blockchain to record device contributions and designs a decentralized valuation protocol to incentivize honest contributors while penalizing free riders, thus enhancing collaboration efficiency. This mechanism, within the context of data trading, ensures that each data contributor receives fair compensation based on their contribution.

4.3.2. Practical application cases

The integration of Edge AI and blockchain in data trading systems has already been applied in several practical scenarios, such as the intelligent energy data market and intelligent healthcare data trading platforms.

(1) Intelligent energy data market: In intelligent energy systems, real-time collection and analysis of energy data are crucial for price settings. By deploying Edge AI technology, energy data can be processed directly on edge devices, enabling real-time analysis of energy demand, consumption, and supply capabilities, which in turn allows for adaptive pricing. Blockchain technology ensures the transparency and immutability of data transactions, guaranteeing fairness in the process. For instance, an intelligent grid system can utilize Edge AI to monitor energy consumption data in real time and dynamically adjust electricity prices based on market demand and supply conditions. Blockchain records each transaction of energy data on the distributed ledger, ensuring the authenticity of the data and the transparency of the transaction process. The literature [129] explores the application of blockchain and AI technologies in intelligent grids, particularly how they support decentralized energy trading and decision-making processes. It highlights that blockchain technology ensures transparency and security in energy transactions, while AI optimizes energy management and forecasting. By combining these two technologies, prosumers in the smart grid can participate more efficiently in the energy market, facilitating the efficient distribution and utilization of energy. Similarly, the literature [130] proposes a blockchain-based privacy-preserving data aggregation scheme that ensures the confidentiality and integrity of data, further enhancing the security and robustness of the smart grid system.

(2) Intelligent healthcare data trading platform: The intelligent healthcare data market requires real-time evaluation of data quality and scarcity, with dynamic price adjustments based on demand. Edge AI technology can process patient data, clinical records, and other healthcare information in real time on medical devices, allowing for the evaluation of data quality and pricing. Blockchain ensures data privacy and the immutability of transactions, safeguarding patient data. In healthcare data trading platforms, smart contracts can automatically set prices and execute transactions based on data quality, scarcity, and market demand. Blockchain guarantees the transparency and traceability of every transaction, preventing data misuse or tampering. For instance, the literature [131] discusses the application of blockchain-powered federated learning in intelligent healthcare data trading platforms. Blockchain technology ensures data privacy and security during the federated learning process while enabling distributed data sharing and model training. By integrating blockchain with federated learning, intelligent healthcare data platforms can achieve more efficient data transactions and more accurate medical diagnoses (see Table 6).

4.4. Edge AI for adaptive data pricing

With the rapid development of Edge Computing and AI, the data pricing model is also facing a significant transformation. Edge AI, which combines edge computing and AI technologies, enables efficient data processing and intelligent analysis directly on local devices, providing strong support for adaptive data pricing. By real-time collection and analysis of factors such as market demand, data characteristics, and transaction environment, Edge AI can offer dynamic pricing solutions for data transactions, optimizing resource allocation and improving market efficiency.

Adaptive Data Pricing: Traditional data pricing models often rely on static pricing rules, such as fixed prices or pricing based on data types [139]. These models overlook the volatility of data demand, the impact of data characteristics on pricing, and the real-time needs of both buyers and sellers. Thus, such pricing models fail to address the complexities of the dynamic market environment, leading to inefficiencies and an inability to maximize the value of data resources. Adaptive data pricing, on the other hand, is a strategy that adjusts prices in real time based on market demand and data characteristics. Unlike traditional static pricing models, adaptive pricing dynamically adjusts according to changes in the market, data variations, and user requirements, ensuring optimal pricing under real-time conditions [132,133]. Key components of adaptive data pricing models include: *(1) Market demand analysis*: Real-time monitoring of market demand allows for an understanding of the intensity of data needs from buyers and sellers. For example, during peak demand periods in a specific industry, data prices will automatically increase. *(2) Data characteristics assessment*: Pricing is based on the quality, timeliness, and scarcity of the data, ensuring that prices match the value of the data. *(3) Intelligent decision systems*: Adaptive pricing relies on intelligent decision systems that use AI algorithms to analyze and model data, automatically adjusting pricing strategies according to market fluctuations. *(4) Real-time feedback mechanisms*: Real-time monitoring of transaction situations and

Table 6
Edge AI for adaptive data pricing.

Technology	Highlight
Adaptive pricing fundamentals [132,133]	Describes real-time adaptive pricing that adjusts based on market demand and data characteristics
Edge AI for supply–demand pricing [134]	Introduces a Q-learning approach for real-time price adjustments in mobile networks based on supply–demand dynamics
Dual-layer pricing strategy [9]	Designs a two-layer settlement mechanism based on dynamic buyer bidding for fairness and adaptivity in transactions
Federated Pricing Optimization [135]	Proposes a Rubinstein bargaining-based bilateral pricing model to improve edge learning task efficiency and accuracy
RL-based competitive market pricing [136]	Employs reinforcement learning to build a fast, fair, and adaptive pricing mechanism for competitive data trading environments
Intelligent transportation data pricing [137]	Utilizes Edge AI for real-time traffic data analysis and dynamic pricing based on congestion and incident patterns
Energy market data pricing [138]	Applies Edge AI to monitor energy usage and adjust prices dynamically based on consumption and production
Energy forecast pricing [132]	Uses AI prediction models to stabilize pricing in response to electricity demand surges

user feedback ensures that pricing strategies are adjusted to maintain fairness and transparency.

Edge AI for Adaptive Data Pricing: Edge AI plays a crucial role in adaptive data pricing by enabling intelligent analysis and processing directly at the data source, allowing for a fast response to market demand and data characteristic changes. Specifically, Edge AI contributes to adaptive data pricing in the following ways: *(1) Real-time data processing and analysis*: By pushing data processing to devices closer to the data source, edge computing reduces data transmission delays. During adaptive data pricing, Edge AI can analyze data characteristics and market demand in real time, ensuring timely and accurate pricing. For example, the literature [134] employs a Q-learning algorithm based on supply–demand relationships, allowing mobile terminal users to make optimal offloading decisions in response to channel fading and resource price fluctuations, with servers dynamically adjusting pricing based on supply–demand dynamics. *(2) Demand forecasting and price adjustment*: Edge AI can use machine learning and deep learning techniques to predict market demand changes by analyzing historical data, market trends, and user behavior. When market demand increases, Edge AI can dynamically adjust data prices and decrease them when demand drops. This allows for precise pricing based on market fluctuations. For example, the literature [9] designs a dual-layer pricing strategy for both parties in a transaction, dynamically adjusting settlement prices based on buyer bids and ensuring fair transaction prices. *(3) Optimization of data pricing models*: Adaptive data pricing relies on complex algorithmic models, which Edge AI can quickly optimize on local devices, adjusting pricing models autonomously. Through iterative optimization, Edge AI can adapt pricing strategies based on market feedback and data characteristics, improving pricing accuracy and market adaptability. For example, in federated learning on edge networks, existing methods overlook communication efficiency and the interactions between mobile device clusters, leading to network congestion and imbalanced data distribution. The literature [135] proposes a bilateral pricing mechanism based on the Rubinstein bargaining model, utilizing pricing allocation, partitioning, and matching rules to enhance task participation and model accuracy. *(4) Transparency and fairness of data transactions*: Edge AI ensures the transparency and fairness of data transactions. By tracking market changes in real time through intelligent decision systems, Edge AI guarantees the fairness of pricing strategies, preventing unfair pricing caused by information asymmetry or market manipulation. Additionally, Edge AI can adjust based on the behavior and feedback of transaction participants, ensuring fairness in the transaction process. The literature [136] introduces a reinforcement learning-based pricing mechanism that enhances data adaptability and rapid pricing response in competitive markets with two or more competitors.

As a core technology for adaptive data pricing, Edge AI leverages its advantages in real-time data processing and intelligent analysis to effectively support the optimization of pricing models and the forecasting of market demand. It finds widespread application in scenarios such as intelligent transportation data trading and energy data trading. *(1) Intelligent transportation data trading:* Intelligent transportation systems generate vast amounts of traffic data, including traffic flow, vehicle speed, and road conditions [137,140]. By utilizing Edge AI, these data can be analyzed in real time on local devices to predict traffic flow and congestion, adjusting data prices based on real-time demand. For example, when a traffic accident occurs in a particular area, the demand for traffic data increases, and Edge AI can quickly raise data prices through intelligent pricing models to balance supply and demand. *(2) Energy data trading:* In the energy industry, the timeliness and quality of data are critical factors for pricing. Through Edge AI, energy companies can use AI models to monitor energy usage in real time, detecting anomalies promptly [138], and adjust prices based on real-time energy production and consumption data. When electricity demand surges, Edge AI can adjust data prices through predictive models to maintain market stability and fairness [132].

4.5. Edge AI for data trading privacy protection

Traditional data privacy protection techniques, such as data encryption and access control, while offering certain security guarantees, often require substantial computational resources and high transmission costs. These demands can conflict with the real-time and efficiency requirements of data trading platforms. Consequently, integrating Edge AI, especially privacy-preserving techniques, has become a key research directions in the field of data trading.

Federated Learning: Federated Learning is a distributed machine learning method that allows multiple participants to train models locally without exchanging data. This approach enables each participant to train a model using their own data and share only model parameters rather than raw data, thus mitigating the risk of data leakage [141]. In the context of Edge computing, federated learning can further optimize privacy protection by processing data locally on edge devices. Model parameters are sent to a central server for aggregation, creating a global model while preserving data privacy and reducing transmission bandwidth demands. For instance, the literature [17] designs a new data trading market framework combining blockchain and Federated Learning, allowing data to be exchanged in a *usable but invisible* manner. This approach enables joint model training without revealing raw data, offering a decentralized, secure, and trusted data trading platform. Similarly, the literature [142] proposes a local privacy model market, where data providers train models using local data and upload encrypted model parameters. An auction mechanism is used to evaluate and select the best models, integrating them into the global model while maintaining privacy through local differential privacy protections. Another example is presented by [10], which introduces a federated learning-based trading framework, incorporating a seller selection algorithm and an incentive mechanism. This system uses gradient similarity and the Shapley value to assess data value and

contribution, with an enhanced Upper Confidence Bound (UCB) algorithm selecting the best sellers based on historical contributions and data quality. The literature [143] presents a federated learning and reinforcement learning-based data auction mechanism. This system uses reinforcement learning to simulate auction mechanisms and dynamically adjust the selection and reward strategies for edge devices participating in federated learning based on market feedback and environmental conditions. The literature [144] proposes a federated learning-based data trading market. In this model, data consumers submitted their data requirements, and the platform searched for suitable data providers based on those needs. The data providers used their local data to train models and uploaded relevant information. The platform incorporates mechanisms such as local model privacy assessments, model aggregation to detect malicious intent, and a fair billing system to ensure the fairness and sustainability of data transactions.

Differential Privacy: Differential Privacy plays a crucial role in protecting the privacy of participants during the statistical analysis of data in data trading contexts. To safeguard privacy, data providers typically apply differential privacy techniques, adding noise to the data analysis process to prevent the disclosure of individual participants' information. This method ensures that even if the data is analyzed by third parties, sensitive individual data remains protected. For example, the literature [9] combines AI and differential privacy to intelligently analyze anonymized data, ensuring that no sensitive information could be inferred even after the data is processed by third parties. The literature [145] proposes an AI and differential privacy-based approach to address privacy concerns during data sales and mitigate high costs for data consumers. Furthermore, the literature [146] discusses a data trading scheme that balances data privacy and utility, using contracts to ensure fair compensation for data providers while maintaining data quality for buyers.

Secure Multi-Party Computation: The integration of Edge AI with secure MPC offers an efficient and privacy-preserving solution for data trading. Edge AI performs distributed intelligent computations at the data source, reducing the need for data transmission and enhancing computational efficiency, while secure MPC leverages cryptographic techniques to enable multiple parties to collaboratively compute results without revealing their private data. This combination enhances the credibility of data trading and fosters cross-organizational and cross-industry data sharing and collaboration, empowering privacy-preserving computations in data markets. For example, in the vehicular network data trading and authentication system, issues such as inherent errors in biometric feature extraction and privacy leakage are addressed by [147], which proposes a biometric identity authentication solution based on fuzzy, certificate-less signatures using RSA variants and elliptic curve discrete logarithms. This solution ensures the forgery resistance of the system through formal security proofs. The literature [148] focuses on improving the efficiency of message verification and addressing the high real-time requirements in vehicular networks. It introduces a collaborative approach between Roadside Units (RSUs) and vehicles to execute distributed message authentication, minimizing total verification delay while ensuring security and privacy.

Homomorphic Encryption: The combination of Edge AI and homomorphic encryption provides an innovative solution for data trading that balances privacy protection with computational capability. Edge AI performs intelligent computations at the data source, reducing transmission overhead and improving response times, while homomorphic encryption allows computations to be carried out on encrypted data, ensuring its confidentiality throughout the transaction process. This integration enables data providers to authorize computations without decrypting the data, allowing secure data analysis and value exchange while preserving privacy. This approach enables collaboration across organizations, promotes the flow of data under privacy protection, and provides a secure, efficient technological foundation for the development of data markets. For example, the literature [149] highlights the security and privacy challenges in Edge AI. The literature [150] proposes a blockchain-based privacy-preserving scheme tailored for smart grid applications, utilizing data segmentation, consortium blockchain, and smart contracts to enhance system security and decentralization while protecting the privacy of both data owners and recipients. The literature [151] introduces a trusted IoT data sharing method where reliable third-party blockchain services are deployed at edge nodes to register data computation relationships into blockchain smart contracts, ensuring verifiable and secure data sharing.

Trusted Execution Environments: The integration of Edge AI and TEEs offers a highly efficient and secure solution for data trading. Leveraging distributed intelligent computation, Edge AI processes data locally at the source, reducing data transmission demands and improving computational efficiency. Through hardware-level isolation mechanisms, TEEs ensure that data remains protected from data tampering or data eavesdropping during execution, enabling secure and trusted computation. The combination of these technologies allows data to be processed and traded in trusted environments, ensuring both privacy and data integrity while enhancing computational performance. For instance, the literature [152] proposes an edge computing solution based on TEEs, integrating security modules across different architectures (ARM/x86) to effectively detect and measure security risks in the edge computing environment, thereby ensuring the security of data transactions. The literature [153] introduces a software-based method for constructing secure trusted execution areas within TEEs, effectively protecting data confidentiality and integrity, particularly in privacy-preserving and secure computations in Edge AI and data trading contexts. Furthermore, the literature [154] presents a security architecture for deep neural network accelerators, effectively preventing model inversion attacks and ensuring model privacy in Edge AI applications.

Integration of Multiple Privacy-Preserving Techniques: The ultimate goal of combining Edge AI with privacy-preserving technologies is to create a model that can effectively protect privacy while facilitating efficient data trading. In such a model, both data providers and consumers can engage in secure and efficient transactions while safeguarding data privacy. The core of this privacy-preserving data trading model lies in reducing the need for data transmission through local data processing and intelligent algorithms, while ensuring privacy through encryption, differential privacy, and federated learning techniques during the trading process. A potential implementation flow could include: *(1) Local data processing:* Data providers preprocess the data at the edge device, performing tasks such as denoising and format conversion, and conducting local intelligent analysis. *(2) Privacy protection:* Techniques like homomorphic encryption, differential privacy, or federated learning are applied to protect the privacy of the data. This approach eliminates the need for data transmission to cloud or central platforms, thereby reducing the risk of privacy leakage. *(3) Intelligent matching:* Edge AI performs data analysis and model training to ensure the data trading platform can intelligently match data providers with data consumers based on their specific needs. *(4) Efficient trading:* Data trading occurs through encrypted data, with privacy-preserving mechanisms in place on the platform to ensure both the security and efficiency of transactions.

4.6. Edge AI for data quality assessment

The integration of Edge AI with data quality assessment offers an intelligent solution for ensuring data quality in data trading. Edge AI enables real-time computation and analysis at the data source, automatically identifying and correcting defects in data, such as missing values, outliers, and duplicates, ensuring that data meets high-quality standards before trading. Additionally, Edge AI continuously monitors and evaluates data quality using deep learning algorithms, dynamically adjusting data cleaning and optimization strategies. This synergy enables data trading platforms to ensure the accuracy, completeness,

and consistency of trading data, thus enhancing the credibility and value of data transactions and promoting the healthy and efficient development of the data market. For instance, the literature [155] provides a comprehensive analysis of the relationship between edge computing system performance and data quality, offering insights into the importance of data quality assessment. The literature [156] reviews the current research on multi-dimensional service quality evaluation and optimization in IoT, providing valuable reference and guidance for edge computing data evaluation and optimization research. The literature [157] further discusses security and privacy protection issues in edge computing, which are closely related to data quality assessment. The literature [158] proposed a blockchain-based framework for data quality assessment, where edge computing devices collaboratively process task data, integrating reinforcement learning and an improved delegated reputation proof mechanism to assess data quality while ensuring privacy.

4.7. Edge AI for incentive mechanisms in data trading

The combination of Edge AI and incentive mechanisms introduces intelligent regulation and rewards into data trading. Edge AI analyzes data value, quality, and demand in real-time, enabling the design of personalized incentive strategies for both data providers and consumers. These strategies can be dynamically adjusted based on factors such as data contribution, usage frequency, and quality ratings, ensuring fair returns for all parties involved in the transaction process. Through intelligent incentive mechanisms, platforms can promote data sharing by encouraging data providers to contribute high-quality data and motivating data consumers to make accurate data demand predictions and purchasing decisions. This integration not only increases the activity level of data transactions but also enhances the sustainability and efficiency of the entire data trading ecosystem. For example, the literature [159] studies an incentive mechanism based on edge computing and blockchain networks, specifically how to incentivize miners to purchase computational resources. This study introduces a two-stage Stackelberg game model to analyze the optimal strategies between miners and edge service providers. The literature [160] explores resource allocation issues in blockchain networks supported by edge computing and proposed a contract-based incentive mechanism that encourages edge service providers to offer computational services to blockchain miners. The literature [161] examines profit distribution between data producers and intermediaries in data transactions within ubiquitous edge computing environments, proposing a game-theory-based profit-sharing mechanism to incentivize data producers and intermediaries to actively engage in data trading. The literature [162] presents a blockchain-based incentive mechanism design for edge-assisted crowdsensing, where smart contracts are used to incentivize participants to provide high-quality data while preserving data privacy.

5. Applications of edge AI-empowered data trading

As the development of Edge AI, the data trading market will increasingly deployed in various industries, revealing significant potential across several key areas. In the future, the integration of edge computing and blockchain will enhance the trustworthiness and security of data transactions, promoting global data sharing and collaboration. At the same time, the optimization of data pricing models, privacy protection technologies, and market incentive mechanisms will further accelerate the data trading ecosystem.

5.1. IoT data trading

IoT data trading is becoming critical applications in smart factories, smart cities, and vehicular networks (V2X). Smart factories rely on sensors, automated equipment, and intelligent algorithms to facilitate efficient data flow. By leveraging the real-time processing capabilities of Edge AI, these systems reduce dependency on cloud computing, thereby improving production efficiency and lowering data transmission costs [3]. In manufacturing, for instance, edge nodes on machines can collect and process production data in real-time, optimize production scheduling, enhance equipment utilization, and reduce latency while improving anomaly detection capabilities [163, 164]. Furthermore, IoT data trading platforms enable smart factories to share production data with supply chain partners, thus improving overall supply chain efficiency. In smart cities, distributed computing optimizes data exchange for applications such as traffic management, environmental monitoring, and public safety, ensuring privacy while enhancing urban operational efficiency. For example, intelligent traffic management systems use Edge AI to analyze real-time data from vehicle sensors and roadside devices, dynamically adjusting traffic lights to reduce congestion and improve flow. In vehicular networks, the vast amount of data generated by vehicles must be processed efficiently at the edge to support autonomous driving, smart navigation, and in-car communication. Edge AI enables local preprocessing and analysis of this data, enhancing driving safety and system responsiveness. For example, autonomous vehicles can run AI models on edge computing devices to perform real-time object detection and path planning, without the need to upload data to remote clouds, thus reducing communication latency and bandwidth consumption.

5.2. Financial data market

The financial data market is increasingly being enhanced through Edge AI, improving the accuracy and security of data transactions [4]. Risk control, credit assessment, and quantitative trading are key applications in financial data trading [165]. Edge AI, with its distributed computing capabilities, can analyze financial behavior data locally, minimizing the risk of data leakage and ensuring regulatory compliance. Moreover, smart risk management systems leverage Edge AI to continuously monitor unusual trading patterns, further enhancing financial security. In credit assessment, Edge AI can perform real-time analysis of multi-source data, with user consent, providing more accurate risk evaluations for loans, insurance, and other financial services. In quantitative trading, which relies on high-frequency data processing, Edge AI enables trading algorithms to operate locally, significantly reducing latency and improving decision-making accuracy. For instance, high-frequency trading firms can use Edge AI to process market data and execute trading orders based on AI prediction models, thus enhancing market responsiveness.

5.3. Healthcare data sharing

Healthcare data sharing is one of the most sensitive areas of data trading. Through privacy-preserving computation techniques, Edge AI promotes the development of electronic health record (EHR) transactions and AI-assisted diagnostics [5,166]. In EHR transactions, Edge AI enables data to be processed locally while participating in model training and analysis without revealing personal information [167]. For example, hospitals can process patient medical records locally and share the trained AI model parameters with distributed networks for cross-institutional federated learning, improving disease prediction and diagnosis while preventing data leakage. In AI-assisted diagnostics, Edge AI enables medical devices to perform intelligent analysis at the edge, improving diagnostic efficiency, reducing reliance on remote cloud services, and enhancing healthcare response times, especially in remote or resource-limited areas.

Fig. 5. Challenges and future directions.

5.4. Data-driven supply chain optimization

In the supply chain and logistics sectors, data-driven optimization is becoming a crucial approach for enhancing operational efficiency [168,169]. Traditional supply chains often suffer from information asymmetry, data delays, and inaccurate prediction. The integration of Edge AI with data trading models significantly mitigates these challenges. Through edge computing, various stages of the supply chain, such as raw material procurement, production planning, inventory management, and logistics, can make more accurate decisions. For example, in smart warehouse systems, raw data can be processed at the edge, allowing dynamic adjustments based on market demand changes, thereby reducing excess inventory and improving supply chain responsiveness. In logistics, Edge AI, when combined with real-time traffic data and historical transportation records, can optimize route planning, reduce transportation costs and improve delivery efficiency.

5.5. Advertising and user profiling data trading

Advertising and user profiling data trading is another key area where Edge AI plays a transformative role. Precision marketing and recommendation systems rely on high-quality user data, while privacy protection remains a significant challenge in this field [170,171]. Edge AI allows user behavior data to be processed locally, uploading only anonymized or encrypted feature data, thereby protecting user privacy while improving the accuracy of ad targeting. For instance, e-commerce platforms can use Edge AI to process user browsing and purchase data to generate personalized recommendations without transmitting sensitive information. Advertisers can then use user feature data, generated by Edge AI, to participate in real-time bidding, increasing the return on investment (ROI) for advertisement. Additionally, by integrating blockchain technology, data trading platforms can enable controlled sharing of user data, giving users greater autonomy over their data while ensuring its authenticity and security during transactions.

6. Challenges and future directions

As the data trading ecosystem continues to evolve rapidly, Edge AI-enabled data trading still faces several challenges, including standardization, privacy and security, computational efficiency, data quality evaluation, and legal compliance, as shown in Fig. 5.

Standardization Issues: One of the most significant challenges in the global development of data trading is the lack of standardization. Currently, the data trading market lacks unified protocols and standards, which limits interoperability between different platforms [11]. To establish a globally accepted data trading protocol, it is essential to develop standardized metadata descriptions, data pricing models, and data exchange formats. Additionally, the differences in regulatory frameworks across countries further complicate data trading. Therefore, a key research focus moving forward will be how to coordinate international interests and promote the standardization of data trading.

Privacy and Security: Another major challenge in Edge AI-enabled data trading is privacy and security [9,172]. As the value of data increases, so does the risk of data leakage and data misuse. While existing privacy protection mechanisms, such as homomorphic encryption, federated learning, and differential privacy, help mitigate data privacy concerns, their high computational overheads and practical limitations still lead to significant challenges [173,174]. Future research should focus on optimizing these technologies to balance data security with computational efficiency.

Computational Efficiency: Regarding computational efficiency, Edge AI relies on distributed computing resources, but the associated overhead can lead to performance bottlenecks. For instance, in IoT environments, where terminal devices have limited computational power, reducing the computational demands for model training and inference is crucial to enable efficient data trading in resource-constrained settings [10]. Future research should focus on lightweight neural networks, edge computing optimization strategies, and adaptive computing architectures to enable low-power, high-performance data trading.

Data Quality Evaluation: Data quality evaluation is another critical factor in data trading. In the data market, the authenticity and integrity of data directly affect its trade value. However, given the diverse data sources and the varying quality levels, assessing the reliability and authenticity of data remains a significant challenge [64]. Future research could leverage blockchain technology, combined with smart contracts and traceability mechanisms, to ensure the verifiability of data sources. Furthermore, Edge AI can be employed to perform real-time data quality assessments, such as anomaly detection and data cleaning, which would enhance the credibility of data transactions.

Ethics and Legal Compliance: Finally, ethics and legal compliance are critical considerations. As the scale of data trading grows, ensuring compliance while protecting privacy has become a central concern for global regulatory bodies. For example, regulations such as GDPR [34] and CCPA [41] leads to strict limitations on personal data usage. Future data trading systems will need to ensure compliance at the technical architecture level. Additionally, addressing fairness in data trading, preventing data monopolies, and safeguarding the rights of data contributors will be essential policy directions for the future.

7. Conclusion

As data becomes the core component of the digital economy, data trading is evolving into a key element supporting innovation and growth across industries. However, traditional data trading models still face challenges such as privacy protection, data pricing, trading efficiency, and decentralized trust mechanisms. Edge AI, as a distributed intelligent computing paradigm, is reshaping the technological architecture and market model of data trading.

This paper discusses the fundamental concepts, key technologies, and core challenges of data trading, with a focus on the role of Edge AI. By leveraging local computing, privacy-preserving techniques (such as federated learning, homomorphic encryption, and differential privacy), and blockchain, Edge AI offers safer, more efficient, and transparent solutions. Unlike traditional centralized data markets, Edge AI enables decentralized data processing and trading, making transactions fairer, reducing the risk of data misuse, and improving market liquidity and data pricing. Edge AI-enabled data trading shows great potential in various fields, including IoT, financial markets, healthcare, supply chain management, and precision marketing. It enhances real-time data processing in IoT, optimizes risk control and credit assessment in finance, ensures privacy protection in healthcare data sharing, and improves market efficiency in supply chain and advertising. Despite

its potential, several challenges remain, including data standardization, privacy and security mechanisms, computational efficiency, data quality assessment, and compliance. Future research should focus on developing global data trading standards, more efficient privacy-preserving methods, and decentralized trading architectures to foster the healthy development of the data trading market.

In conclusion, Edge AI is reshaping the data trading ecosystem, accelerating the development of new data markets. With continuous optimization of technologies and market mechanisms, data trading will become more secure, efficient, and transparent, emerging as a vital infrastructure for the digital economy.

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

This work was supported by Guangdong Regional Joint Fund Project (No. 2022A1515110157), Guangdong Basic and Applied Basic Research Project (No. 2025A1515012874), Research Project of Pazhou Lab for Excellent Young Scholars (No. PZL2021KF0024), Foundation of Yunnan Key Laboratory of Service Computing (No. YNSC24115), University Research Project of Guangzhou Education Bureau (No. 2024312189), and Guangzhou Basic and Applied Basic Research Project (No. SL2024A03J00397).

Data availability

Data will be made available on request.

References

[1] C. Niu, Z. Zheng, F. Wu, X. Gao, G. Chen, Achieving data truthfulness and privacy preservation in data markets, IEEE Trans. Knowl. Data Eng. 31 (1) (2019) 105–119.

[2] Y. Liu, P. Liu, W. Jing, H.H. Song, Pd2s: A privacy-preserving differentiated data sharing scheme based on blockchain and federated learning, IEEE Internet Things J. 10 (24) (2023) 21489–21501.

[3] B. Fenelon, S. Enayati, H. Pishro-Nik, Private uav-assisted iot data collection: An energy-privacy trade-off, in: 20th Annual International Conference on Privacy, Security and Trust, PST 2023, Copenhagen, Denmark, August (2023) 21-23, IEEE, 2023, pp. 1–5.

[4] Q. Zhang, Y. Zhang, X. Yao, S. Li, C. Zhang, P. Liu, A dynamic attributes-driven graph attention network modeling on behavioral finance for stock prediction, ACM Trans. Knowl. Discov. Data 18 (1) (2024) 16:1–16:29..

[5] H. Yang, Z. Yi, X.A. Wang, Y. Su, Z. Tu, X. Yang, Improved lightweight cloud storage auditing protocol for shared medical data, Wirel. Commun. Mob. Comput. 2021 (2021) 8886763:1–8886763:13..

[6] K. Chen, Y. Lin, H. Luo, B. Mi, Y. Xiao, C. Ma, J.S. Silva, Edgeleakage: Membership information leakage in distributed edge intelligence systems, 2024, CoRR abs/2404.16851.

[7] Z. Chen, J. Huang, S. Liu, H. Long, A blockchain and A-DCNN integrated framework for privacy protection and intrusion detection of industrial iot, Computing 107 (1) (2025) 18.

[8] K. Chen, H. Zhang, X. Feng, X. Zhang, B. Mi, Z. Jin, Backdoor attacks against distributed swarm learning, ISA Trans. 141 (2023) 59–72.

[9] K. Chen, Z. Xu, B. Mi, Data trading combination auction mechanism based on the exponential mechanism, 2024, CoRR abs/2405.07336.

[10] K. Chen, Z. Xu, Federated learning for data market: Shapley-ucb for seller selection and incentives, 2024, CoRR abs/2410.09107.

[11] J. Zhang, Y. Bi, A survey on data markets, 2024, arXiv preprint arXiv:2411.07267.

[12] M.S. Chishti, F. Sufyan, A. Banerjee, Decentralized on-chain data access via smart contracts in ethereum blockchain, IEEE Trans. Netw. Serv. Manag. 19 (1) (2022) 174–187.

[13] F. Douglis, A. Stavrou, Distributed ledger technologies, IEEE Internet Comput. 24 (3) (2020) 5–6.

[14] W. Lan, K. Chen, J. Cao, Y. Li, N. Li, Q. Chen, Y. Sahni, Security-sensitive task offloading in integrated satellite-terrestrial networks, IEEE Trans. Mob. Comput. 24 (3) (2025) 2220–2233.

[15] Z. Liao, X. Han, X. Tang, C. Feng, An adaptable pricing-based resource allocation scheme considering user offloading needs in edge computing, IEEE Internet Things J. 12 (1) (2025) 582–594.

[16] W. Lan, K. Chen, Y. Li, J. Cao, Y. Sahni, Deep reinforcement learning for privacy-preserving task offloading in integrated satellite-terrestrial networks, IEEE Trans. Mob. Comput. 23 (10) (2024) 9678–9691.

[17] C. Li, Y. Yuan, F.-Y. Wang, A novel framework for data trading markets based on blockchain-enabled federated learning, in: 2022 IEEE 25th International Conference on Intelligent Transportation Systems, ITSC, 2022, pp. 3392–3397.

[18] K. Chen, W. Li, J. Cao, B. Mi, J. Shen, Optimizing federated incremental learning: Efficient malicious data removal for big data analytics, Tsinghua Sci. Technol. (2025).

[19] R. Gupta, D. Reebadiya, S. Tanwar, N. Kumar, M. Guizani, When blockchain meets edge intelligence: Trusted and security solutions for consumers, IEEE Netw. 35 (5) (2021) 272–278.

[20] M. Andronie, R. Blazek, M. Iatagan, R. Skypalova, C. Uta, Generative artificial intelligence algorithms in internet of things blockchain-based fintech management, Oeconomia Copernic. 15 (4) (2024) 1349–1381.

[21] K.J. Olowe, N.L. Edoh, S.J.C. Zouo, J. Olamijuwon, Review of predictive modeling and machine learning applications in financial service analysis, Comput. Sci. IT Res. J. 5 (11) (2024) 2609–2626.

[22] R. Zheng, A. Sumper, M. Aragues-Penalba, S. Galceran-Arellano, Advancing power system services with privacy-preserving federated learning techniques: A review, IEEE Access 12 (2024) 76753–76780.

[23] M. Ali, M. Suchismita, S.S. Ali, B.-J. Choi, Privacy-preserving machine learning for iot-integrated smart grids: Recent advances, opportunities, and challenges, Energies 18 (10) (2025) 2515.

[24] H. Zahid, A. Zulfiqar, M. Adnan, M.S. Iqbal, A. Shah, Transforming nano grids to smart grid 3.0: Ai, digital twins, blockchain, and the metaverse revolutionizing the energy ecosystem, 2025, TechRxiv Preprints.

[25] K. Balaji, Revolutionizing high-frequency trading: The impacts of financial technology and data science innovations, in: Machine Learning and Modeling Techniques in Financial Engineering, IGI Global, 2025.

[26] W. Liu, X. Lin, Data transaction mode and its legal regulation in the context of market-oriented allocation of data elements, in: G. Kanaparan W. Hong (Ed.), Computer Science and Education. Computer Science and Technology, Springer Nature Singapore, Singapore, 2024, pp. 42–51.

[27] M.S. Bhargavi, S.M. Katti, M. Shilpa, V.P. Kulkarni, S. Prasad, Transactional data analytics for inferring behavioural traits in ethereum blockchain network, 2020, pp. 485–490.

[28] M. Balazinska, B. Howe, D. Suciu, Data markets in the cloud: An opportunity for the database community, PVLDB, in: Proceedings of the VLDB Endowment, vol. 4, 2011, pp. 1482–1485.

[29] F. Stahl, F. Schomm, G. Vossen, The Data Marketplace Survey Revisited, Tech. rep., European Research Center for Information Systems (ERCIS), University of Münster, 2014.

[30] A. Agarwal, M.A. Dahleh, T. Sarkar, A marketplace for data: An algorithmic solution, in: ACM Conference on Economics and Computation, pp. 701–726.

[31] A. Ghosh, A. Roth, Selling privacy at auction, Games Econom. Behav. 91 (2015) 334–346.

[32] C. Li, D.Y. Li, G. Miklau, D. Suciu, A theory of pricing private data, ACM Trans. Database Syst. 39 (4) (2014) 34:1–34:28..

[33] B.-R. Lin, D. Kifer, On arbitrage-free pricing for general data queries, Proc. VLDB Endow. (PVLDB) 7 (2014) 757–768.

[34] General data protection regulation (gdpr), 2025, https://gdpr-info.eu/, (Accessed 08 February 2025).

[35] S. Xu, The construction of data property right systemunder market-oriented allocation of factors, J. Chongqing Univ. 29 (1) (2023) 255–267.

[36] Y. Li, J. Hu, Data transactions in China: Legislation, promotion, and regulation, Int. J. Digit. Law Gov. 1 (2) (2024) 413–437.

[37] Trusted data transaction pre-standardization workshop, 2025, https://www.trusted-data-transaction.org/en/, (Accessed 08 February 2025).

[38] B. Hyseni, L. Abazi Bexheti, Synchronization and standardization of open data platforms: A systematic literature review, TEM J. 13 (2) (2024) 1265–1276.

[39] K. Xu, Ternary governance of data transaction circulation:technology, standard and law, J. Jishou Univeristy Soc. Sci. 43 (1) (2022) 96–105.

[40] Y. Jiang, Y. Zhong, X. Ge, Smart contract-based data commodity transactions for industrial internet of things, IEEE Access 7 (2019) 180856–180866.

[41] California consumer privacy act (ccpa), 2025, https://www.oag.ca.gov/privacy/ccpa/, (Accessed 08 February 2025).

[42] B.-P. Ivanschitz, T.J. Lampoltshammer, V. Mireles, A. Revenko, S. Schlarb, L. Thurnay, A data market with decentralized repositories, in: Proceedings of the 2nd Workshop on Decentralizing the Semantic Web Co-Located with the 17th International Semantic Web Conference, DeSemWeb@ISWC 2018, Monterey, California, USA, October 8 2018, 2165 of CEUR Workshop Proceedings, CEUR-WS.org, 2018.

[43] G.S. Ramachandran, R. Radhakrishnan, B. Krishnamachari, Towards a decentralized data marketplace for smart cities, in: IEEE International Smart Cities Conference, ISC2, 2018, pp. 1–8.
[44] B. Yu, Y. Guan, S. Geng, L. Miao, Y. Zhang, Y. Gong, A blockchain-enhanced secure and reliable data transaction scheme in mas via htlc, 2024, pp. 494–499.
[45] M.H. Krishna, H. Raju, A. Jain, T.M.K. Al-Rubaye, N. Thandra, P. Tewari, Synergizing blockchain with big data for improved data integrity and transparency, in: 2024 7th International Conference on Contemporary Computing and Informatics, IC3I, vol. 7, 2024, pp. 680–685.
[46] O. Protocol, What is ocean?, 2025, https://docs.oceanprotocol.com/discover/what-is-ocean, (Accessed 12 January 2025).
[47] Metaschool, Building on ocean protocol–a developer's guide, 2025, https://metaschool.so/articles/build-on-ocean-protocol, (Accessed 12 January 2025).
[48] Datum, Blockchain data storage and monetization, 2025, https://datum.org/, (Accessed 20 January 2025).
[49] S. Network, The decentralized data network, 2025, https://streamr.network/, (Accessed 20 January 2025).
[50] SingularityNET, Decentralized marketplace for ai services, 2025, https://www.singularitynet.com/, (Accessed 03 February 2025).
[51] SingularityNET, Crypto-powered ai agents: The next evolution in web3, 2025, https://blog.spheron.network/crypto-powered-ai-agents-the-next-evolution-in-web3, (Accessed 03 February 2025).
[52] S. Jiang, J. Cao, H. Wu, K. Chen, X. Liu, Privacy-preserving and efficient data sharing for blockchain-based intelligent transportation systems, Inf. Sci. 635 (2023) 72–85.
[53] A.S.V. Koe, S. Ai, Q. Chen, J. Tang, K. Chen, S. Zhang, X. Li, Hieraledger: Towards malicious gateways in appendable-block blockchain constructions for iot, Inf. Sci. 632 (2023) 87–104.
[54] Y. Wang, C.-H. Hsieh, C. Li, Research and analysis on the distributed database of blockchain and non- blockchain, in: 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics, ICCCBDA, 2020, pp. 307–313.
[55] S. Rouhani, R. Deters, Data trust framework using blockchain technology and adaptive transaction validation, IEEE Access 9 (2021) 90379–90391.
[56] W. Vickrey, Counterspeculation, auctions, and competitive sealed tenders, J. Financ. 16 (1) (1961) 8–37.
[57] P. Milgrom, R.J. Weber, A theory of auctions and competitive bidding, Econometrica 50 (5) (1982) 1089–1122.
[58] J.M. Swinkels, Efficiency of large private value auctions, Econometrica 69 (1) (2001) 37–68.
[59] L. Anton, S. Tuomas, Approximating revenue-maximizing combinatorial auctions, in: AAAI'05: Proceedings of the 20th National Conference on Artificial Intelligence, 2005, pp. 267–273.
[60] S. Ding, S. Guo, C. Liu, Z. Wang, Blockchain-based pricing mechanism research on iot data transactions, in: 2023 IEEE 29th International Conference on Parallel and Distributed Systems, ICPADS, 2023, pp. 631–638.
[61] Y. Chen, M. Liu, J. Zhang, A. Tan, W. Yan, A three-party dynamic pricing mechanism for customized data transactions, Symmetry 16 (11) (2024) 1540.
[62] J.F. Nash, Equilibrium points in n-person games, Proc. Natl. Acad. Sci. 36 (1) (1950) 48–49.
[63] L.S. Shapley, A value for n-person games, Contrib. Theory Games 2 (1953) 307–317.
[64] Y. Gong, G. Liu, Y. Xue, R. Li, L. Meng, A survey on dataset quality in machine learning, Inf. Softw. Technol. 162 (2023) 107268.
[65] Y. Lu, J. Wang, L. Liu, H. Yang, Get by how much you pay: A novel data pricing scheme for data transactions, Inf. Syst. 61 (6) (2024) 103849.
[66] D. Paul, F. Zhe, N. Harikrishna, Optimal auctions through deep learning, Commun. ACM 64 (8) (2021) 109–116.
[67] J. Yan, Y. Zheng, X. Yang, C. Chen, X. Guan, Privacy-preserving localization for underwater acoustic sensor networks: A differential privacy-based deep learning approach, IEEE Trans. Inf. Forensics Secur. 20 (2025) 737–752.
[68] S. Wang, Y. Li, Y. Zhong, K. Chen, X. Wang, Z. Zhou, F. Peng, Y. Qian, J. Du, W. Yang, Locally private set-valued data analyses: Distribution and heavy hitters estimation, IEEE Trans. Mob. Comput. 23 (8) (2024) 8050–8065.
[69] N. Wang, S. Wang, M. Li, L. Wu, Z. Zhang, Z. Guan, L. Zhu, Balancing differential privacy and utility: A relevance-based adaptive private fine-tuning framework for language models, IEEE Trans. Inf. Forensics Secur. 20 (2025) 207–220.
[70] W. Chen, H. Chen, T. Han, W. Tong, S. Zhong, Secure two-party frequent itemset mining with guaranteeing differential privacy, IEEE Trans. Mob. Comput. 24 (1) (2025) 276–292.
[71] L. Chen, D. Xiao, X. Xiao, Y. Zhang, Secure and efficient federated learning via novel authenticable multi-party computation and compressed sensing, IEEE Trans. Inf. Forensics Secur. 19 (2024) 10141–10156.
[72] T. Peng, W. Zhong, G. Wang, E. Luo, S. Yu, Y. Liu, Y. Yang, X. Zhang, Privacy-preserving truth discovery based on secure multi-party computation in vehicle-based mobile crowdsensing, IEEE Trans. Intell. Transp. Syst. 25 (7) (2024) 7767–7779.
[73] X. Li, H. Wang, Z. Li, L. Wu, X. Wei, Y. Su, R. Lu, Publicly verifiable secure multi-party computation framework based on bulletin board, IEEE Trans. Serv. Comput. 17 (4) (2024) 1698–1711.
[74] J. Takeshita, D. Reis, T. Gong, M.T. Niemier, X.S. Hu, T. Jung, Accelerating finite-field and torus fully homomorphic encryption via compute-enabled (S)RAM, IEEE Trans. Comput. 73 (10) (2024) 2449–2462.
[75] H. Li, F. Wan, M. Gong, A.K. Qin, Y. Wu, L. Xing, Privacy-enhanced multitasking particle swarm optimization based on homomorphic encryption, IEEE Trans. Evol. Comput. 28 (5) (2024) 1336–1350.
[76] L. Sun, H. Li, Y. Peng, J. Cui, Efficient secure CNN inference: A multi-server framework based on conditional separable and homomorphic encryption, IEEE Trans. Cloud Comput. 12 (4) (2024) 1116–1130.
[77] N. Yang, L. Yang, X. Du, X. Guo, F. Meng, Y. Zhang, Blockchain based trusted execution environment architecture analysis for multi - source data fusion scenario, J. Cloud Comput. 12 (1) (2023) 122.
[78] S. Queyrut, V. Schiavoni, P. Felber, Mitigating adversarial attacks in federated learning with trusted execution environments, in: 43rd IEEE International Conference on Distributed Computing Systems, ICDCS 2023, Hong Kong, July (2023) 18-21, IEEE, 2023, pp. 626–637.
[79] A.P. Kalapaaking, I. Khalil, M.S. Rahman, M. Atiquzzaman, X. Yi, M. Almashor, Blockchain-based federated learning with secure aggregation in trusted execution environment for internet-of-things, IEEE Trans. Ind. Inform. 19 (2) (2023) 1703–1714.
[80] C. Li, Y. Yuan, F. Wang, A novel framework for data trading markets based on blockchain-enabled federated learning, in: 25th IEEE International Conference on Intelligent Transportation Systems, ITSC 2022, Macau, China, October (2022) 8-12, IEEE, 2022, pp. 3392–3397.
[81] Y. Zhao, Z. Liu, C. Qiu, X. Wang, F.R. Yu, V.C.M. Leung, An incentive mechanism for big data trading in end-edge-cloud hierarchical federated learning, in: IEEE Global Communications Conference, GLOBECOM 2021, Madrid, Spain, December (2021) 7-11, IEEE, 2021, pp. 1–6.
[82] S. Zheng, Y. Cao, M. Yoshikawa, H. Li, Q. Yan, Fl-market: Trading private models in federated learning, in: S. Tsumoto, Y. Ohsawa, L. Chen, D.V. den Poel, X. Hu, Y. Motomura, T. Takagi, L. Wu, Y. Xie, A. Abe, V. Raghavan (Eds.), IEEE International Conference on Big Data, Big Data 2022, Osaka, Japan, December (2022) 17-20, IEEE, 2022, pp. 1525–1534.
[83] J. Liang, Y. Gong, Security and privacy protection protocol based on edge computing in smart campus, Expert. Syst. J. Knowl. Eng. 42 (1) (2025).
[84] M. Odema, M.A.A. Faruque, Privynas: Privacy-aware neural architecture search for split computing in edge-cloud systems, IEEE Internet Things J. 11 (4) (2024) 6638–6651.
[85] M. Sedir, H. Hazar, N. Felix, S. Divesh, Data quality assessment: Challenges and opportunities, 2024, arXiv preprint arXiv:2403.00526.
[86] H. He, E.A. Garcia, Learning from imbalanced data, IEEE Trans. Knowl. Data Eng. 21 (9) (2009) 1263–1284.
[87] M. Herschel, R. Diestelkamper, H.B. Lahmar, A survey on provenance: What for? what form? what from? VLDB J. 26 (6) (2017) 881–906.
[88] T.N. Herzog, F. Scheuren, W.E. Winkler, Data Quality and Record Linkage Techniques, Springer, 2007.
[89] X. Miao, Y. Gao, L. Chen, H. Peng, J. Yin, Q. Li, Towards query pricing on incomplete data, IEEE Trans. Knowl. Data Eng. 34 (8) (2022) 4024–4036.
[90] C. Chai, J. Wang, Y. Luo, Data management for machine learning: A survey, IEEE Trans. Knowl. Data Eng. 23 (5) (2023) 4046–4667.
[91] C. Olston, F. Korn, N. Noy, N. Polyzotis, S. Whang, S. Roy, Managing google's data lake: An overview of the goods system, IEEE Data Eng. Bull. 39 (3) (2016) 5.
[92] H. Miao, A. Li, L.S. Davis, A. Deshpande, Modelhub: Deep learning lifecycle management, in: 2017 IEEE 33rd International Conference on Data Engineering, ICDE, 2017, pp. 1393–1394.
[93] W. Dai, C. Dai, K.-K.R. Choo, C. Cui, D. Zou, H. Jin, Sdte: A secure blockchain-based data trading ecosystem, IEEE Trans. Inf. Forensics Secur. 15 (2020) 725–737.
[94] Y. Lu, X. Huang, Y. Dai, S. Maharjan, Y. Zhang, Blockchain and federated learning for privacy-preserved data sharing in industrial iot, IEEE Trans. Ind. Inform. 16 (6) (2020) 4177–4186.
[95] L. Tian, J. Li, W. Li, B. Ramesh, Z. Cai, Optimal contract-based mechanisms for online data trading markets, IEEE Internet Things J. 6 (5) (2019) 7800–7810.
[96] C. Niu, Z. Zheng, F. Wu, X. Gao, G. Chen, Trading data in good faith: Integrating truthfulness and privacy preservation in data markets, 2017, pp. 223–226.
[97] X. Zhang, Y. Su, J. Qin, J. Sun, Sdta: Secure decentralized trading alliance for electronic medical data, Comput. J. 67 (2024) 2573–2585.
[98] Q. Ramadan, Z. Boukhers, Data trading and monetization: Challenges and open research directions, in: The International Conference on Future Networks and Distributed Systems, ICFNDS'23, 2023, pp. 1–8.
[99] M. Fassnacht, J. Leimstoll, C. Benz, D. Heinz, Data sharing practices: The interplay of data, organizational structures, and network dynamics, Electron. Mark. 34 (47) (2024).
[100] H. Akrami, B.R. Chaudhury, J. Garg, A. Murhekar, On the theoretical foundations of data exchange economies, 2024, arXiv preprint arXiv:2412.01968.
[101] W. Yuan, Fair data transactions across private databases, IEEE Access 8 (2020) 53720–53732.

[102] Z. Bowei, P. Heng, L. Kunyang, A blockchain and zero knowledge proof based data security transaction method in distributed computing, Electronics 13 (21) (2024) 4260.
[103] C. Rong, A. Chakravorty, Enabling multilevel data sharing based on blockchain and smart contracts, IEEE Blockchain Tech. Briefs (2018).
[104] T. Li, H. Wang, D. He, J. Yu, Blockchain-based privacy-preserving and rewarding private data sharing for iot, IEEE Internet Things J. 9 (16) (2022) 15138–15149.
[105] C. Huang, D. Liu, J. Ni, R. Lu, X. Shen, Achieving accountable and efficient data sharing in industrial internet of things, IEEE Trans. Ind. Inform. 17 (2) (2021) 1416–1427.
[106] A. Koukosias, C. Anagnostopoulos, K. Kolomvatsos, Task-aware data selectivity in pervasive edge computing environments, IEEE Trans. Knowl. Data Eng. 37 (1) (2025) 513–525.
[107] M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The case for vm-based cloudlets in mobile computing, IEEE Pervasive Comput. (2009) 14–23.
[108] F. Bonomi, R. Milito, J. Zhu, S. Addepalli, Fog computing and its role in the internet of things, in: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, 2012, pp. 13–16.
[109] ETSI, Mobile-Edge Computing — Introductory Technical White Paper, Tech. rep., ETSI Industry Specification Group (ISG) MEC, 2014, URL https://www.etsi.org/technologies/multi-access-edge-computing.
[110] W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, Edge computing: Vision and challenges, IEEE Internet Things J. 3 (5) (2016) 637–646.
[111] S. Yi, C. Li, Q. Li, A survey of fog computing: Concepts, applications and issues, in: Proceedings of the 2015 Workshop on Mobile Big Data, 2015, pp. 37–42.
[112] Y. Mao, C. You, J. Zhang, K. Huang, K.B. Letaief, A survey on mobile edge computing: The communication perspective, IEEE Commun. Surv. Tutor. 19 (4) (2017) 2322–2358.
[113] R. Olaniyan, O. Fadahunsi, M. Maheswaran, M.F. Zhani, Opportunistic edge computing: Concepts, opportunities and research challenges, Future Gener. Comput. Syst. 89 (2018) 633–645.
[114] C. Mouradian, D. Naboulsi, S. Yangui, R.H. Glitho, M.J. Morrow, P.A. Polakos, A comprehensive survey on fog computing: State-of-the-art and research challenges, IEEE Commun. Surv. & Tutorials 20 (1) (2017) 416–464.
[115] W.Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, A. Khan, Edge computing: A survey, Future Gener. Comput. Syst. 97 (2019) 219–235.
[116] D. Xu, T. Li, Y. Li, X. Su, S. Tarkoma, T. Jiang, J. Crowcroft, P. Hui, Edge intelligence: Empowering intelligence to the edge of network, Proc. IEEE 109 (11) (2021) 1778–1837.
[117] C. Kongyang, L. Guomin, Z. Hongfa, L. Waixi, S. Jiaxing, Resilient task offloading in integrated satellite-terrestrial networks with mobility-induced variability, Digit. Commun. Netw. (2025).
[118] X. Wang, W. Jia, Optimizing edge ai: A comprehensive survey on data, models, and systems, 2025, arXiv preprint arXiv:2501.03265.
[119] L. Ma, Q. Pei, L. Zhou, H. Zhu, L. Wang, Y. Ji, Federated data cleaning: Collaborative and privacy-preserving data cleaning for edge intelligence, IEEE Internet Things J. 8 (8) (2021) 6757–6770.
[120] C. Tian, Z. Li, H. Yuan, R. Hamzaoui, S.K. Liquan Shen, Feature compression for cloud–edge multimodal 3d object detection, 2024, arXiv preprint arXiv:2409.04123.
[121] P. Ye, W. Wang, B. Mi, K. Chen, Edgestreaming: Secure computation intelligence in distributed edge networks for streaming analytics, ACM Trans. Multimed. Comput. Commun. Appl. 21 (8) (2025) 222.
[122] K. Chen, Y. Li, W. Lan, B. Mi, S. Wang, Aigc-assisted digital watermark services in low-earth orbit satellite-terrestrial edge networks, 2024, CoRR abs/2407.01534.
[123] S. Yang, K. Qiu, F. Zhang, L. Cao, F. Li, M. Tang, L. Zhu, Efficient data trading and placement in blockchain-based edge computing systems, IEEE Open J. Commun. Soc. (2024).
[124] Y. Tian, T. Li, J. Xiong, M.Z.A. Bhuiyan, J. Ma, C. Peng, A blockchain-based machine learning framework for edge services in iiot, IEEE Trans. Ind. Inform. 18 (3) (2022) 1918–1929.
[125] L. Li, J. Li, R. Liu, Z. Li, Overview of blockchain-based terminal-edge-cloud collaborative computing paradigm, Comput. Electr. Eng. 120 (2024) 109737.
[126] E. Badidi, O.E. Harrouss, Integrating edge computing and blockchain for safer and more efficient digital transportation systems, Procedia Comput. Sci. 251 (2025) 273–280.
[127] Z. Gao, W. Yan, The real-time data processing framework for blockchain and edge computing, Alex. Eng. J. 120 (2025) 50–61.
[128] Y. Du, Z. Wang, C. Leung, V.C.M. Leung, Towards collaborative edge intelligence: Blockchain-based data valuation and scheduling for improved quality of service, Futur. Internet 16 (8) (2024).
[129] W. Hua, Y. Chen, M. Qadrdan, J. Jiang, H. Sun, J. Wu, Applications of blockchain and artificial intelligence technologies for enabling prosumers in smart grids: A review, Renew. Sustain. Energy Rev. 161 (2022) 1–15.
[130] C. Zhao, L. Wang, Z. Liu, K. Zhang, L. Wang, W. Li, K. Chen, Bprm: Blockchain-based privacy-preserving and robust data aggregation supporting multi-functionality for fog-assisted smart grid, IEEE Internet Things J. (2024) 1–1.
[131] J. Zhu, J. Cao, S. Divya, S. Jiang, Blockchain-empowered federated learning: Challenges, solutions, and future directions, ACM Comput. Surv. 55 (11) (2023) 1–31.
[132] L. Dierks, S. Seuken, The competitive effects of variance-based pricing, in: C. Bessiere (Ed.), Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, International Joint Conferences on Artificial Intelligence Organization, 2020, pp. 362–370.
[133] G. Yamuna, D. Paul Dhinakaran, C. Vijai, P.J. Kingsly, Raynukaazhakarsamy, S.R. Devi, Machine learning-based price optimization for dynamic pricing on online retail, in: 2024 Ninth International Conference on Science Technology Engineering and Mathematics, ICONSTEM, 2024, pp. 1–5.
[134] M. Tao, X. Li, K. Ota, M. Dong, Single-cell multiuser computation offloading in dynamic pricing-aided mobile edge computing, IEEE Trans. Comput. Soc. Syst. 11 (2) (2024) 3004–3014.
[135] B. Pan, J. Lu, S. Cao, J. Liu, W. Tian, M. Li, Bilateral pricing for dynamic association in federated edge learning, IEEE Trans. Mob. Comput. (2025) 1–14.
[136] J. Gao, Z. Wang, X. Wei, An adaptive pricing framework for real-time ai model service exchange, IEEE Trans. Netw. Sci. Eng. 11 (5) (2024) 5114–5129.
[137] S.S. Gill, M. Golec, J. Hu, M. Xu, Edge ai: A taxonomy, systematic review and future directions, 2024, arXiv preprint arXiv:2407.04265.
[138] Y. Himeur, A.N. Sayed, A. Alsalemi, F. Bensaali, A. Amira, Edge ai for internet of energy: Challenges and perspectives, 2023, ArXiv abs/2311.16851.
[139] V. Abhishek, I.A. Kash, P. Key, Fixed and market pricing for cloud services, in: 2012 Proceedings IEEE INFOCOM Workshops, 2012, pp. 157–162.
[140] K. Chen, G. Tan, M. Lu, J. Wu, CRSM: a practical crowdsourcing-based road surface monitoring system, Wirel. Netw. 22 (3) (2016) 765–779.
[141] K. Chen, X. Zhang, X. Zhou, B. Mi, Y. Xiao, L. Zhou, Z. Wu, L. Wu, X. Wang, Privacy preserving federated learning for full heterogeneity, ISA Trans. 141 (2023) 73–83.
[142] S. Zheng, Y. Cao, M. Yoshikawa, H. Li, Q. Yan, Fl-market: Trading private models in federated learning, in: 2022 IEEE International Conference on Big Data, Big Data, 2022, pp. 1525–1534.
[143] G. Li, J. Cai, J. Lu, H. Chen, Incentive mechanism design for cross-device federated learning: A reinforcement auction approach, IEEE Trans. Mob. Comput. (2024) 1–17.
[144] Q. Li, Z. Liu, K. Xu, Martfl: Enabling utility-driven data marketplace with a robust and verifiable federated learning architecture, in: Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security, 2023.
[145] F. Yang, X. Liao, X. Lei, N. Mu, D. Zhang, Towards privacy-preserving and practical data trading for aggregate statistic, IEEE Trans. Sustain. Comput. 9 (3) (2024) 452–463.
[146] M.M. Khalili, X. Zhang, M. Liu, Designing contracts for trading private and heterogeneous data using a biased differentially private algorithm, IEEE Access 9 (2021) 70732–70745.
[147] L. Wang, J. Xu, B. Qin, M. Wen, K. Chen, An efficient fuzzy certificateless signature-based authentication scheme using anonymous biometric identities for vanets, IEEE Trans. Dependable Secur. Comput. 22 (1) (2025) 292–307.
[148] J. Xu, L. Wang, M. Wen, Y. Long, K. Chen, Dpb-ma: Low-latency message authentication scheme based on distributed verification and priority in vehicular ad hoc network, IEEE Trans. Veh. Technol. 72 (4) (2023) 5152–5166.
[149] C. Wang, Z. Yuan, P. Zhou, Z. Xu, R. Li, D.O. Wu, The security and privacy of mobile-edge computing: An artificial intelligence perspective, IEEE Internet Things J. 10 (24) (2023) 22008–22032.
[150] Z. Guan, X. Zhou, P. Liu, L. Wu, W. Yang, A blockchain-based dual-side privacy-preserving multiparty computation scheme for edge-enabled smart grid, IEEE Internet Things J. 9 (16) (2022) 14287–14299.
[151] L. Ma, B. Duan, B. Zhang, Y. Li, Y. Fu, D. Ma, A trusted iot data sharing method based on secure multi-party computation, J. Cloud Comput. 13 (1) (2024) 138.
[152] Z. Ning, F. Zhang, W. Shi, A study of using tee on edge computing, J. Comput. Res. Dev. 56 (7) (2019) 1441–1453.
[153] S. Zhao, Q. Zhang, Y. Qin, et al., Sectee: A software-based approach to secure enclave architecture using tee, in: ACM SIGSAC Conference on Computer and Communications Security, 2019, pp. 1723–1740.
[154] X. Wang, R. Hou, Y. Zhu, Npufort: A secure architecture of dnn accelerator against model inversion attack, in: ACM International Conference on Computing Frontiers, 2019, pp. 190–196.
[155] H.-L. Truong, M. Karan, Analytics of performance and data quality for mobile edge cloud applications, in: 2018 IEEE 11th International Conference on Cloud Computing, CLOUD, 2018, pp. 660–667.
[156] J. Huang, F. Liu, J. Zhang, Multi-dimensional qos evaluation and optimization of mobile edge computing for iot: A survey, Chin. J. Electron. 33 (4) (2024) 859–874.
[157] G. Kaur, R.S. Batth, Edge computing: Classification, applications, and challenges, in: 2021 2nd International Conference on Intelligent Engineering and Management, ICIEM, 2021, pp. 254–259.
[158] J. An, S. Wu, X. Gui, X. He, X. Zhang, A blockchain-based framework for data quality in edge-computing-enabled crowdsensing, Front. Comput. Sci. 17 (4) (2023) 174503.

[159] Z. Chang, W. Guo, X. Guo, Z. Zhou, T. Ristaniemi, Incentive mechanism for edge-computing-based blockchain, IEEE Trans. Ind. Inform. 16 (11) (2020) 7105–7114.
[160] Z. Yu, Z. Chang, L. Wang, G. Min, Contract-based incentive design for resource allocation in edge computing-based blockchain, IEEE Trans. Netw. Sci. Eng. 11 (6) (2024) 6143–6156.
[161] Y. Huang, Y. Zeng, F. Ye, Y. Yang, Profit sharing for data producer and intermediate parties in data trading over pervasive edge computing environments, IEEE Trans. Mob. Comput. 22 (1) (2023) 429–442.
[162] C. Ying, H. Jin, J. Li, X. Si, Incentive mechanism design via smart contract in blockchain-based edge-assisted crowdsensing, Front. Comput. Sci. 19 (2024) 193802.
[163] A. Nawaz, J.P. Queralta, J. Guan, M. Awais, T.N. Gia, A.K. Bashir, H. Kan, T. Westerlund, Edge computing to secure iot data ownership and trade with the ethereum blockchain, Sensors 20 (14) (2020) 3965.
[164] Y. Nabil, H. ElSawy, S. Al-Dharrab, H. Mostafa, H. Attia, Data aggregation in regular large-scale iot networks: Granularity, reliability, and delay tradeoffs, IEEE Internet Things J. 9 (18) (2022) 17767–17784.
[165] K.F. Mojdehi, B. Amiri, A. Haddadi, A novel hybrid model for credit risk assessment of supply chain finance based on topological data analysis and graph neural network, IEEE Access 13 (2025) 13101–13127.
[166] S.M. Pournaghi, M. Bayat, Y. Farjami, Medsba: a novel and secure scheme to share medical data based on blockchain technology and attribute-based encryption, J. Ambient. Intell. Humaniz. Comput. 11 (11) (2020) 4613–4641.
[167] C. Gan, X. Xiao, Y. Zhang, Q. Zhu, J. Bi, D.K. Jain, A. Saini, An asynchronous federated learning-assisted data sharing method for medical blockchain, Appl. Intell. 55 (2) (2025) 208.
[168] Q. Wen, Y. Gao, Z. Chen, D. Wu, A blockchain-based data sharing scheme in the supply chain by iiot, in: IEEE International Conference on Industrial Cyber Physical Systems, ICPS 2019, Taipei, Taiwan, May (2019) 6-9, IEEE, 2019, pp. 695–700.
[169] W. Chen, T. Choi, A. Dolgui, D.A. Ivanov, E. Pesch, Digital manufacturing and supply chain: creating benefits through operations research and artificial intelligence, Ann. Oper. Res. 344 (2) (2025) 569–574.
[170] B. Xie, C. Hu, H. Huang, J. Yu, H. Xia, DCI-PFGL: decentralized cross-institutional personalized federated graph learning for iot service recommendation, IEEE Internet Things J. 11 (8) (2024) 13837–13850.
[171] Z. Xu, B. Li, W. Cao, Enhancing federated learning-based social recommendations with graph attention networks, Neurocomputing 617 (2025) 129045.
[172] K. Chen, D. Zhang, B. Mi, Private data leakage in federated human activity recognition for wearable healthcare devices, 2024, CoRR abs/2405.10979.
[173] K. Chen, D. Zhang, B. Mi, Y. Huang, Z. Li, Fast yet versatile machine unlearning for deep neural networks, Neural Netw. 190 (2025) 107648.
[174] K. Chen, Y. Huang, Y. Wang, X. Zhang, B. Mi, Y. Wang, Privacy preserving machine unlearning for smart cities, Ann. Telecommun. 79 (2024) 61–72.