Table 1
Existing research works on blockchain security and privacy.

Schemes	Privacy	Security	Focus on
[27]	×	√	Security risks, real attacks, solutions, and future directions in blockchain.
[38]	√	√	Blockchain technology to improve the interoperability and security of health management system.
[41]	×	√	Security supervision and research prospects in blockchain.
[21]	√	√	Security and Privacy in different consensus algorithms of blockchain.
[1]	√	×	Privacy protection mechanism and application scenarios in blockchain.
[22]	√	√	The core concept, working principle and research direction of blockchain.
[17]	√	√	Theoretical models, analytical models and experimental tools of blockchain.
[28]	√	√	Security, privacy, and trust of blockchain technology in crowdsourcing services.
[26]	√	√	Challenges and solutions for blockchain applications in data-driven networks.
[36]	√	√	Blockchain security and privacy in electronic health record systems.

immature technologies and the imperfect management of blockchains, they face many problems. In terms of security, there are still many attacks against blockchain consensus mechanisms, smart contracts, data content, etc., such as 51% attacks, smart contract vulnerabilities and denial of service attacks [4]. Moreover, a set of formal regulatory rules and mechanisms have not yet been formed for illegal data on blockchains. In terms of privacy protection, blockchains need to disclose the transaction information for an entire network to allow the nodes in the network to reach a consensus, which creates a risk of privacy leakage.

Zubaydi, Varga, and Molnár [50] investigated and analyzed related research of blockchains combined with IoT. They mainly discuss the security and privacy problems that exist in a combination of blockchain and IoT. First, they introduce basic blockchain and IoT concepts, principles and architectures and then compare the related literature from the perspectives of application scenarios and technology selection. Finally, prospects for the combination of blockchain and IoT to address security and privacy issues are discussed. Mohanta et al. [30] analyzed the technical implementation of blockchains in different application fields from an academic perspective. It also discusses the progress of the application of blockchains in different fields by different organizations and expounded upon blockchain's security and privacy problems. However, most existing articles only focus on one aspect of blockchain privacy or security, and there are few papers that can cover security supervision, privacy protection and data exchange. Thus, this paper aims to classify and analyze the security, privacy and regulation attributes of blockchain and discuss the security solutions to implement these attributes.

1.1. Related works

In blockchain applications, security and privacy have always been research focus areas. Existing works on blockchain security and privacy are compared in Table 1. Li et al. [27] studied the security threats faced by blockchains and analyzed the corresponding attacks and security enhancement possibilities. Villarreal et al. [38] conducted a systematic literature review to investigate the potential of blockchain technology in enhancing the interoperability and security of health care systems. Moreover, they proposed a high-level architecture and validated its feasibility through an experimental approach, utilizing a model-driven engineering methodology for smart contracts. The study's outcomes offer valuable insights into the challenges and opportunities of applying blockchain in health care systems, highlighting the need for a balanced approach to the security and interoperability concerns in designing blockchain-based solutions. Focusing on the security supervision problems existing in blockchains, Wang et al. [41] conducted a systematic overview and explained the security supervision problems. At the same time, the challenges of supervision and possible future research directions were proposed. Joshi, Han and Wang [21] found that some current blockchain studies mainly focus on applying blockchain to various applications, while a comprehensive investigation on the application prospects and technology of blockchain has not yet been completed. As a result, they provided a comprehensive analysis of blockchain from the security and privacy perspective and discussed the challenges and opportunities for different consensus algorithms.

In terms of blockchain privacy protection, Bernabe et al. [1] investigated the current status of blockchain privacy protection technologies and analyzed the common privacy protection mechanisms and blockchain platforms. Additionally, they analyzed the main blockchain application scenarios for privacy technologies in e-government, e-health, smart cities, etc. Kolb et al. [22] elaborated upon the blockchain working mechanisms and principles. The main challenges faced by current blockchains were analyzed, and the

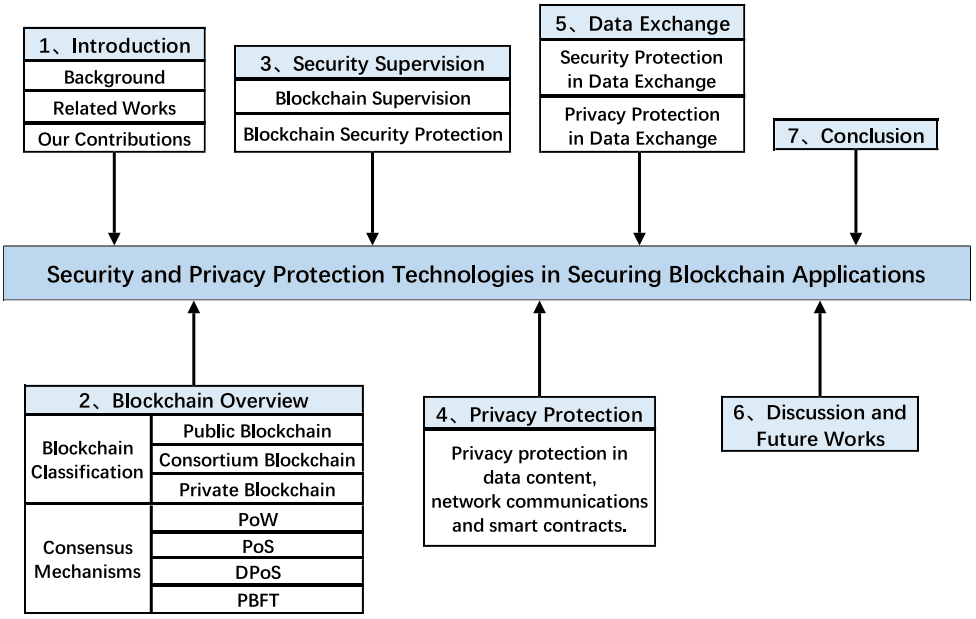

Fig. 1. The structure of this article.

research and technology to solve these challenges were summarized. By classifying the basic blockchain theories and mechanisms, Huang et al. [17] conducted a comprehensive investigation and analysis of the latest research results, providing a guide to the development of the blockchain frontier for theories, modeling and tools.

In specific application scenarios, Ma et al. [28] conducted research on security, privacy, trust and other technologies in crowdsourcing services and summarized blockchain's existing advantages and challenges. Li et al. [26] provided an overview of the existing works of blockchain technologies in computer network applications and identified the blockchain challenges and solutions in data-driven network applications. Shi et al. [36] analyzed and compared the security and privacy of electronic health record systems based on blockchains and discussed the future development direction of electronic health records.

1.2. Our contributions

The research contributions of this article are as follows.

- This paper presents basic blockchain knowledge, including blockchain's classification and consensus mechanisms.
- This paper conducts a comprehensive review and analysis of blockchain in terms of privacy protection, security supervision and data exchange. It also summarizes the existing solutions and the remaining deficiencies. Through a comparative analysis of the literature, the current situation is made more intuitive, which is helpful for us to explore blockchain's future development direction.
- Finally, some directions and problems to be solved in future research on blockchain are proposed. This contributes to better research and application of blockchain.

1.3. Paper organization

The structure of this article is shown in Fig. 1 and is organized as follows. Section 2 introduces the preliminaries of blockchain. In sections 3, 4, and 5, blockchain security supervision, privacy protection and data exchange are analyzed, respectively. Section 6 presents a discussion and future works. Finally, Section 7 concludes the article.

2. Blockchain overview

2.1. Blockchain classification

Based on the difference in the read and write permissions and the management permissions of data, blockchains can be divided into public blockchains, consortium blockchains and private blockchains, which is also the most common classification method. The features of the public blockchain, consortium blockchain and private blockchain are introduced in detail below.

Public blockchain: Public blockchains, also known as permissionless blockchains, do not contain a centralized authority. The associated data do not include read and write access sets. Any node in the network can be accessed at any time, and all nodes can freely participate in transactions on the blockchain. They are generally considered to be blockchains with complete decentralization, high anonymity and tamper resistance. Public blockchains are at the center of many popular applications, such as Bitcoin.

Table 2
Consensus mechanism.

Consensus mechanism	Byzantine fault tolerance	Throughput	Confirm speed	Energy consumption
PoW	$\surd$	$\approx$ 7TPS	Slow	Many
PoS	$\surd$	$\geq$ 25TPS	Slow	More
DPoS	$\surd$	$\geq$ 300TPS	Slow	Few
PBFT	$\surd$	$\geq$ 10000TPS	Quick	Few

Consortium blockchain: Consortium blockchains, also known as the permission blockchains, are between the public blockchains and the private blockchains. They adopt a "partial decentralization" method in structure and are jointly managed by several enterprises or institutions [20]. The participants need to register and authenticate in advance, and the joining of a node requires the permission of other members of the consortium. Compared with public blockchains, the consortium blockchains include fewer participating nodes. The data are recorded and maintained by the authenticated participants who have the right to read the data.

Private blockchain: Private blockchains are controlled by a single organization or individual whose write permissions are limited to themselves. They have strict standards for access nodes, so they are characterized by fast transactions and heightened privacy. Compared with public blockchains, the private blockchains have higher security, but the degree of decentralization is greatly weakened.

2.2. Consensus mechanism

The consensus mechanism plays an important role in maintaining the stable operation of a system and mutual trust between the nodes. The consensus mechanism refers to the process of making nodes agree on the transaction information [24]. Nodes use the consensus mechanism to make consistency judgments on transactions, thereby weakening the function of a centralized supervision system. Typical consensus mechanisms include proof of work (PoW), proof of stake (PoS), delegated proof of stake (DPoS) and practical Byzantine fault tolerant (PBFT).

PoW: The PoW consensus mechanism was first proposed by Satoshi Nakamoto in a Bitcoin system. Each node determines the right to account by computing random numbers to ensure the consistency of the data. In PoW, the contest for bookkeeping rights can only be compared with arithmetic power and is completely decentralized. At the same time, it is also destructive, and the data in the block can only be tampered with when the computing power exceeds 50%. However, the contention of nodes for computing power consumes huge resources in PoW. The large mining pools formed by competing for computing power lead to alternative centralization, which deviates from the original intention of centralization. Restricted by internet bandwidth, each block can accommodate a limited number of transactions.

PoS: PoS solves the problem of PoW consuming too much energy. It gives corresponding interest based on the number and time of the holding tokens. In this way, nodes with more rights will maintain the security of a blockchain to reach a consensus. PoS reduces the resource consumption of the accounting rights. It requires low node performance and shortens the time to reach consensus. However, the more tokens a node has, the greater the probability of obtaining the accounting right. This will form a strong Matthew effect, and fairness cannot be guaranteed. Moreover, the attack cost for a saboteur on the network is low, and the security remains to be verified.

DPoS: DPoS votes through the proportion of assets, and nodes will choose relatively reliable nodes to maximize their own interests. Each node elects an accounting node, and the accounting node generates blocks in sequence. When no valid block is generated within the specified time, the other accounting nodes are responsible for generating the block. Nodes can change accounting nodes by voting at any time. DPoS takes a short time to verify blocks, greatly improves the efficiency of block generation and greatly increases the number of transactions that can be accommodated. However, the degree of decentralization of DPoS is weak. If the number of nodes is small, the accounting nodes are not fair and representative.

PBFT: A PBFT consensus mechanism was proposed to solve the general Byzantine problem. It reduced the operating complexity of Byzantine protocols from exponential to polynomial. Assuming that the number of nodes in the system is N and the number of nodes with Byzantine errors is f, as long as $f \leq (N-1)/3$, the system will keep working normally. PBFT has high consensus efficiency and transaction frequency. However, when more than $1/3$ of the nodes are Byzantine problem nodes, no consensus can be reached. A comparison of common consensus mechanisms is shown in Table 2.

3. Security supervision

Blockchain is an innovative technology that integrates research results in many fields, such as computer science, mathematics and economics. It has the potential to provide transformative opportunities for many industries, such as finance, government affairs, property rights and supply chains. Blockchain has become a hot topic of social concern. However, the difficulty of blockchain supervision often leads to the occurrence of security incidents, resulting in fewer practical applications of blockchain. In recent years, technical research on blockchain security supervision has become one of the most important research directions. This section explores blockchain's problems and technologies from a supervision perspective and then analyzes blockchain security protection from three aspects, the network, consensus mechanisms and smart contracts. Blockchain supervision can be used as a way to protect

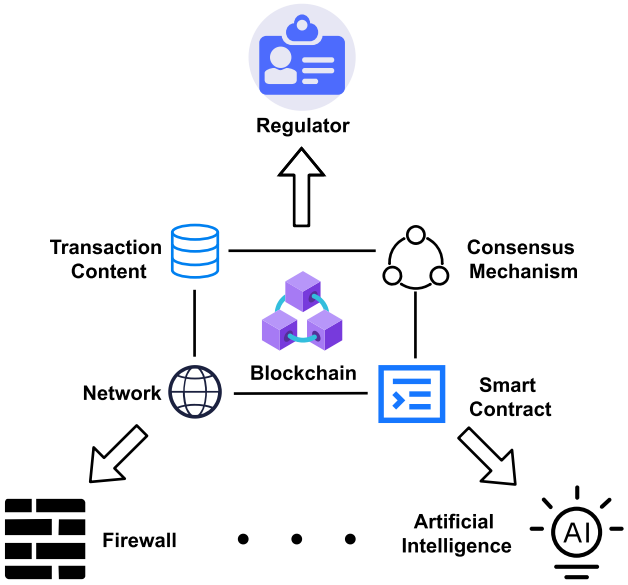

Fig. 2. The objects and common methods of blockchain security supervision.

blockchains, and security protection, and also to assist in the supervision of blockchains. The objects and common methods of blockchain security supervision are shown in Fig. 2.

3.1. Blockchain supervision

Many experts have performed research on blockchain supervision. Yong et al. [44] utilized machine learning, consensus mechanisms and smart contracts to monitor vaccine supply. Smart contracts are used to query and track vaccine records. Scheme [44] effectively established a vaccine accountability system and also uses regulatory agencies to manage expired vaccines. However, only enterprises, wholesale agencies and centers for disease control and prevention are considered in the vaccine blockchain system, and some other participants may be involved in the actual scenarios. Peng et al. [32] designed a double-level blockchain structure to implement the supervision of vaccine production. The first layer stores the private data of the production company, including the production records and hash values. The second layer is public information, which mainly contains production record hashes and vaccine information. Wen et al. [43] proposed a blockchain regulatory framework that uses attribute-based encryption (ABE) to establish privacy protection in the supervision process and effectively weighs supervision and privacy protection. Yin et al. [37] designed a method to complete supervision by using desymbolization, which was implemented with supervisory machine learning. They used 957 entity-authenticated data as samples to build a classifier that could distinguish 12 categories. Román-Martínez et al. [33] proposed a service-oriented architecture that leverages blockchain technology to facilitate the proper execution of consent in health information exchanges. Furthermore, the use of blockchain technology allows for secure maintenance of access control events and subject of care consents, promoting trusted collaboration among organizations, supervisory authorities and individuals. Overall, the proposed architecture is a promising solution to address the issue of subject of care consent in controlling personal health data in distributed and federated environments. It also ensures compliance with regulations and promotes interoperability. Table 3 shows a comparison of blockchain supervision technologies.

3.2. Blockchain security protection

Blockchain security threats are mainly divided into three aspects, networks, consensus mechanisms and smart contracts. In terms of networks, blockchains mainly use a P2P network to achieve block synchronization. However, because the P2P network is deficient in network security management and authentication, it is easy for attackers to attack the topology and routing protocols. A solar eclipse attack restricts the target node's interaction with other nodes by controlling the data transmission of the target node. A cloning attack can create multiple identities to partition the network so that conflicting transactions or double payments can be verified in separate partitions. For routing protocols, an attacker can partition the network, which would prevent effective communication between the networks within different partitions. It can also slow down the transmission speed of the blocks to other nodes or a specified node set while ensuring that the connection is not interrupted. This can waste the computing power of the miners who dig new blocks and increase the competitiveness of other miners.

In terms of the consensus mechanism, there are security threats, such as 51% attacks and selfish mining strategies that are hidden after calculating new blocks and announced at the right time. Wang et al. [42] introduced the concept of mining pool closure. The pool refuses for more nodes to join when certain conditions are met. In addition, the profitability of the different strategies in the

Table 3
Blockchain supervision technologies.

Schemes	Challenge	Implementation	Field	Blockchain type	Advantages
[44]	Vaccine records are prone to falsification.	Machine learning	Vaccine supply	×	Vaccine traceability, evaluation and demand forecasting
[32]	Production records are easily forged and modified.	Double blockchain structure	Vaccine production	×	Data recoverable, small waste of time and space resources
[43]	Traditional supervision is difficult to apply to the consortium blockchain.	ABE	×	Consortium blockchain	Access control, privacy protection
[37]	Bitcoin anonymity leads to illegal activity.	Machine learning	Bitcoin blockchain	Public blockchain	De-anonymity, type prediction
[33]	Healthcare data is difficult to share and supervise.	Blockchain	Healthcare	Consortium blockchain	Secure sharing, access control

Table 4
Security protection technologies.

Schemes	Challenge	Object of protection	Protectionmethod	Advantages
[42]	Block withholding attack	Consensus mechanism	Mining pool closed	Weighing the efficiency and vulnerability of mining pools
[34]	Selfish mining	Consensus mechanism	Block height difference	Low cost
[7]	Selfish mining	Consensus mechanism	Bifurcation length difference	Detection heuristic
[3]	Replay attack	Smart contract	Bytecode language Move	Resource security, semantic guarantees
[2]	Smart contract simplified	Smart contract	Intermediate language Albert	Ease the certification of compilers

consensus process is also a hot research direction. Saad et al. [34] judged the possibility of selfish mining by the difference in block height, that is, by checking the height of the budget block and the actual area. Chicarino et al. [7] used forks to determine whether there is selfish mining. If the block height difference between the forks is greater than two blocks, selfish mining is considered.

In the process of writing smart contracts, the design defects of the programming language itself and the developers' understanding of deviations in the programming language may also lead to loopholes in the smart contract. Most developers use Go or Java for contract development and testing. Some scholars have also proposed a language specifically for smart contracts. Blackshear et al. [3] provided a bytecode language for resources. When a resource's memory address changes, the original memory address is set to be invalid. The type of resource can only be modified by its creator; thus, resource security can be guaranteed. Albert [2] is involved in the Tezos project. As an intermediate language, Albert can conduct a formal analysis of contracts. The comparisons are shown in Table 4.

4. Privacy protection

With the development of blockchains, the privacy issues have become increasingly obvious. Because of the decentralized and open characteristics of blockchains, they not only create advantages but also create data and privacy leakage problems. In the financial field, if an attacker obtains specific transaction data, he or she can analyze the data and speculate on the user's transaction rules to conduct illegal activities. At the same time, the attacker can also analyze the macrofinancial trends, which affects the normal operation of the financial industry. Therefore, data privacy is an important issue that needs to be resolved in blockchains. Privacy protection for blockchains can be divided into privacy protection for data content, network communication, and smart contracts. Therefore, we introduce and analyze blockchain privacy protection technologies from these three aspects.

In terms of data content, Guo et al. [15] proposed an efficient, decentralized and fair marginal transaction system by using blockchain and differential privacy. The system uses blockchains to prevent malicious nodes from tampering with transactions and contracts, and introduces an index mechanism to ensure the authenticity and privacy of the transactions. To solve the security and

credit conflict problems in distributed cross-domain recommendation systems, Wang et al. [11] proposed a set of trusted cross-domain recommendation models by using blockchain and multifeature knowledge maps. The model uses federated learning and blockchains to ensure the security and credibility of the system, and feature encryptions with statistical weights are used to establish data privacy. Zhang et al. [45] proposed a cloud mining pool selection system that supports privacy protection and verification. Time-locked puzzles and the additive homomorphism EIGamal are used to ensure the authenticity of the mine pool and the privacy of the data. At the same time, the system uses semihonest cloud servers to cooperate with each other to improve the computing efficiency of the IoT devices.

To improve the privacy and anonymity of the blockchain transaction system, Zheng et al. [47] designed a blockchain-based organizational-level privacy-preserving transaction system. Their system provides a balance between the privacy protection and security supervision of the blockchain content. Wang et al. [40] designed a transaction privacy protection framework that uses the Paillier cryptosystem to hide accounts in Bitcoin. The framework can resist active and passive attacks, effectively improving the privacy of the transactions. Zhong et al. [49] designed a blockchain-based lightweight payment system that supports privacy protection. Only the final settlement information of the transaction is stored in the blockchain, thus ensuring the privacy protection of the transaction content.

In terms of networks, the Tor network [10] uses onion routing to protect the communication between the nodes. However, simply combining the Tor network with blockchain poses security risks. An attacker can analyze the traffic changes in the network by monitoring the traffic sending, and receiving the status of each node in the Tor obfuscation network, and based on the flow in and out of each node, at the same time, judge the link relation between nodes and mine out the real sender and receiver address information.

The privacy protection of smart contracts has gradually become a research hotspot. Cheng et al. [5] designed a system, called Ekiden, that combines trusted hardware and blockchain, which separates execution and consensus to achieve high scalability and performance of the system. During the setup phase, smart contracts are verified, encrypted and stored. The public and private keys must be provided when invoking and obtaining smart contracts. The privacy of smart contracts is protected by storing encrypted contracts. The Origo protocol [12] focuses on smart contracts in Ethereum through zero-knowledge proof (ZKP) to achieve privacy protection of smart contracts. When the contract is initiated, the system requires the participants to pay the default currency, which will be confiscated if there is malicious behavior during the transaction.

5. Data exchange

Centralized data exchange platforms do not provide sufficient trust and cannot meet complex requirements. Therefore, blockchain has been used to implement data exchanges. However, there are also many security and privacy issues in blockchain-based data exchanges. For example, during a data exchange, the privacy of the user's identity and transaction data may be leaked, invalidated or illegal data may be uploaded, the exchange efficiency may be low during data transmission, and it is difficult to track illegal anonymous users. In the process of data exchange, not only should the authenticity and reliability of the data be guaranteed, but resolving the data privacy and tracking the users when malicious data is discovered should also be considered. Therefore, this section introduces and analyzes security protection, privacy protection and malicious user tracking in blockchain-based data exchanges. The general model of blockchain-based data exchange is shown in Fig. 3.

5.1. Security protection in data exchange

Trusted data exchange based on blockchain can establish private data protections. Table 5 compares the related technologies of security protection in data exchange. Zheng et al. [48] proposed a blockchain-based digital asset exchange mechanism. At the same time, a reward and punishment mechanism was designed to encourage data providers to provide more cost-efficient data services. Many studies have been conducted on data exchange security in real-life fields. In ecosystem data transactions, due to the existence of dishonest buyers and data brokers, there are many limitations in the traditional data transaction platforms. Dai et al. [8] proposed a data transaction ecosystem based on blockchain. Their scheme [8] assumes that others cannot access the seller's raw data but only the analytically processed results, reducing the challenge of protecting the dataset to the challenge of protecting the data processing parameters. Focusing on internet digital advertising (IDA), such as the proliferation of spam advertisements and the difficulty in data exchange, Ding et al. [9] proposed a digital advertising media promotion system based on the IDA media ecosystem by combining blockchain with IDA. The system in [9] uses a multichain structure to ensure the healthy development of the IDA market and to improve the quality and effectiveness of advertising.

The existing blockchain-based data sharing schemes only guarantee the exchange of data and rarely consider the efficiency of sharing. Chi et al. [6] considered the efficiency and security of data sharing and proposed a data sharing scheme based on blockchain. To ensure the security of the data exchanges, they designed a data sharing model based on hyperledger and identity authentication. A community detection algorithm was proposed in [6], which divides the clients into different data sharing communities based on the similarity of the tag data. In [6], selecting the scope of data sharing based on the community test results can effectively narrow the scope of a query and improve the efficiency of data sharing. To ensure the security of medical documents in data sharing, Sharma et al. [35] proposed a medical IoT system based on blockchain. Blockchain ensures the security of medical certificates, and smart contracts implement identity verification and access control. However, the processing of medical image data needs to be optimized.

To date, most schemes only consider data sharing within the same organization, and rarely consider data sharing between different groups of users. Huang, Chen and Wang [18] proposed a multigroup data sharing system with traceability and anonymity. Users can