# Construction of Data Security System

Bingbing Yu[1](✉) and Jiefan Hu[2]

[1] Chief Information Security Officer, Financial Industry, Beijing, China
margie2002@sina.com
[2] Data Governance Senior Manager, Financial Industry, Shanghai, China

**Abstract.** With the progress of science and technology, the development of new technologies, the in-depth application of big data, artificial intelligence, and cloud computing, data has gradually been transformed from information assets to production factors. Data breach, abuse, tampering and other security issues will cause great harm to enterprises, and even affect national security, societal order, public interests and market stability. Therefore, on the basis of meeting the fundamental business needs of enterprises, carrying out business and daily operation management, promoting the application and sharing of data, mining and realizing data value, strengthening data protection capabilities, and ensuring the safe flow of data are also the key points of data management work.

**Keywords:** Data Security · Data Classification · Data Lifecycle

## 1 Introduction

Information technology is constantly developing, and the basic business, core processes, inter industry transactions and activities of enterprises have all been run on information support carriers. The information generated by production and operation is gradually transformed into digital assets in different forms and circulated in information systems. With the progressing of science, the development of new technologies, and the in-depth implementation of big data, artificial intelligence, and cloud computing, data has gradually evolved from an information asset to a factor of production. Data breach, abuse, tampering and other security issues will cause significant harm to the enterprise, and even have an impact on national security, societal order, public interests and market stability. Therefore, on top of meeting the basic business needs of an enterprise, carrying out the business, and managing the day-to-day operation, it is important to incorporate promoting the utilization and sharing of data, mining and realizing data value, strengthening data protection capabilities, and ensuring the safe flow of data into the data management work.

In IT domain, the meaning of the word "security" has changed, from the security of the carriers of information systems to the security of data, from the technical security to the effective protection of and the legitimate utilization of data. Data security refers to taking necessary measures to ensure that data is effectively protected and legitimately utilized, as well as having the ability to ensure the "security" continuously.

S. Zhang et al. (Eds.): BigData 2023, LNCS 14203, pp. 218–229, 2023.
https://doi.org/10.1007/978-3-031-44725-9_16

# 2 Data Security Framework Based on Data Lifecycle

The data lifecycle includes data collection, data transmission, data storage, data usage, data deletion, and data destruction.

Data Security focuses on protecting data against unauthorized access and corruption during the whole life cycle of data. It covers a set of relevant standards, technologies, frameworks and processes.

Data Security includes the planning, development, and execution of data security policies and procedures, which provide proper authentication, authorization, access, and audit on data and information assets [1].

So, enterprises should develop a data security framework based on the security of the entire data lifecycle. This framework consists of three parts: security management, process and mechanism, and technologies and tools. Below is an example of data security framework (Fig. 1).

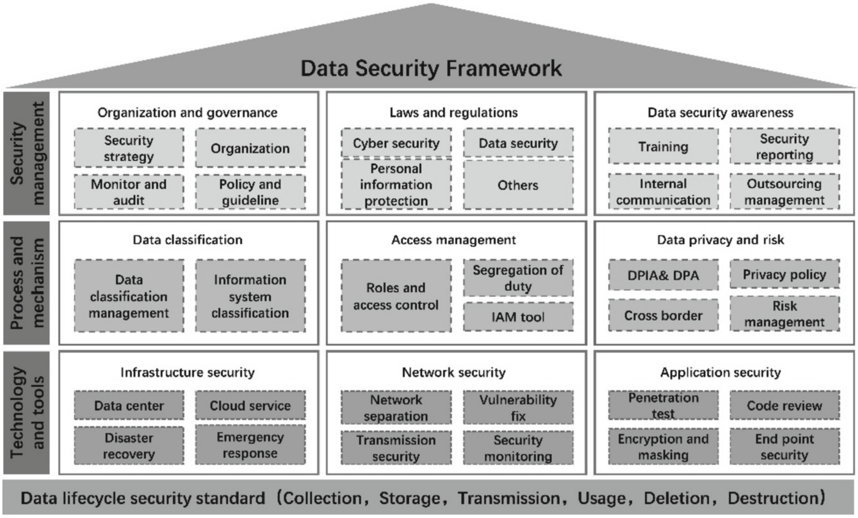


**Fig. 1.** Data Security Framework Example

## 2.1 Security Management

### Principles

To ensure data security, the following principles shall be followed when processing data:

*Legality and Compliance:* It is necessary to ensure the data activities are legitimate and compliant to regulations throughout the entire data lifecycle.

*Clear Purpose:* A data security protection strategy should be developed to clarify the security protection goals and requirements for each stage of the data lifecycle.

*Controllable Whole Process.* Data security control mechanism and technical measures that match the security level of data should be adopted to ensure the confidentiality, integrity, and availability of data at all stages of the its lifecycle, and to avoid unauthorized access, destruction, tampering, leakage, or loss of data throughout the entire data lifecycle.

*Dynamic Control.* The security control strategy and security protection measures of data should not be one-off or static, but be able to adjust in real-time and dynamically to the factors such as business requirements, security environment attributes, and behaviors of system users.

*Consistency of Rights and Responsibilities.* The relevant departments and their responsibilities for data security protection should be clearly defined in organizations. The departments and its staff should actively implement the required measures and take their responsibilities of data security protection [2].

#### Security Management

The security management incorporate dimensions like data security organization and governance, laws and regulations, and data security awareness.

*Data Security Organization and Governance.* In addition, data compliance is also part of the data security work, and it important to abide the data laws and regulations. When implementing the specific data security work, national and industrial standards have provided significant guidance.

*Laws and Regulation.* In addition, data compliance has also been included in the work of data security, so the three major data laws and regulations should be abided. In order to better implement specific work, national and industry standards have provided significant guidance for the implementation of data security work.

*Security Awareness.* With the promulgation of laws and regulations, enhancing data security awareness is crucial for better data security work. Data security is not the job for a single person or department, it concerns everyone in the organization who have been involved in business and operations. The improvement of data security awareness comes from not only the training and communication to internal employees, but also the management of data security awareness of the third parties and the outsourcing counter parties. At the same time, regular audits and monitoring audits should be conducted on routine works to ensure that data security work is effectively implemented according to the designed security plan and process.

### 2.2 Process and Mechanism

The second layer of the data security framework, process and mechanism, includes data classification, access management, and privacy protection.

#### Data Classification

Data classification is the first step in data security governance. The hierarchical classification of data covers not only structured data, but also unstructured data. It includes data stored in information systems and data existing in business processes and business documents as well.

**Data Access**
The core of data security is the management of data access. Accesses should be managed following the principles of minimum fit-for-purpose and role-based access control.

**Privacy Protection**
Privacy protection is an important part of data security. Except for meeting the internal data security requirements, it is also necessary to meet all the requirements under applicable laws and regulations.

### 2.3 Technology and Tools

Data security cannot be separated from the security of data carrier, it requires the careful design from infrastructure security, to network security, then to application system security. There are a number of technologies and tools for data security, such as encryption, identity authentication and access control, data leakage prevention tools, monitoring logs, etc.

## 3 Data Security Work Practice Based on Risk Priority

### 3.1 Data Security Practice Steps

After determining the data security framework, the enterprise need to put the data security work into practice. It is known to all that data security is a seamless and endless task, and like a circle, it has no starting or ending point. One can never be never secured enough, but can always be more secured. Back to reality, it is not feasible to invest in security endlessly. So, we need to put the data security work into practice based on the risk priorities.

Although security is a common requirement of enterprises, the risk faced by each enterprise is different due to their natures of business and operations, and the subjective regulations.

In order to find a balance between security and cost, and to more accurately implement the data security work, the following four-step approach can be adopted to determine the direction and projects of data security. Literally, the four steps are data security requirements collection, data security gap analysis, data security improvement plan development, and rigorous implementation of security projects (Fig. 2).

**Needs Collection**
When collecting data security requirements, it is necessary to consider meeting the requirements of laws and regulations, national standards, and industry standards, as well as meeting the needs of business development to reduce the risks faced by data usage.

**Gap Analysis**
Based on the requirements and the understanding of existing data security measures of

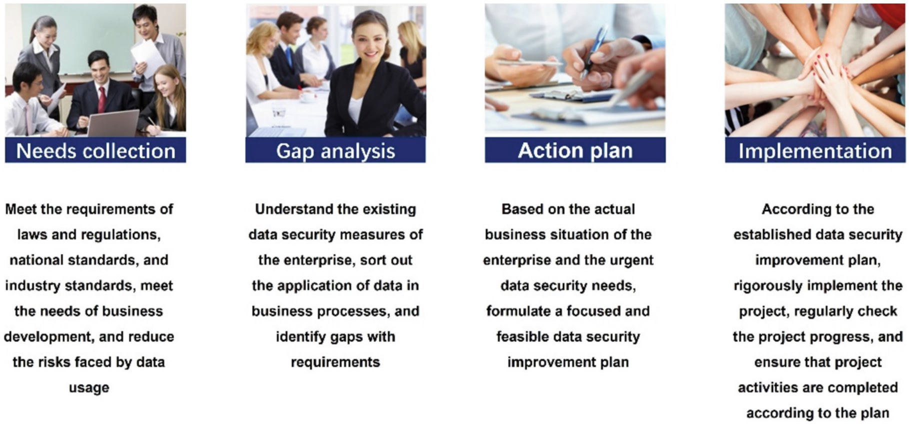

Needs collection
Gap analysis
Action plan
Implementation
Meet the requirements of laws and regulations, national standards, and industry standards, meet the needs of business development, and reduce the risks faced by data usage
Understand the existing data security measures of the enterprise, sort out the application of data in business processes, and identify gaps with requirements
Based on the actual business situation of the enterprise and the urgent data security needs, formulate a focused and feasible data security improvement plan
According to the established data security improvement plan, rigorously implement the project, regularly check the project progress, and ensure that project activities are completed according to the plan

**Fig. 2.** Data Security Work Practice Based on Risk Priority

the enterprise, analyze the usage of data in business processes and identify the gap with the requirements.

**Action Plan**
Based on the gap and the actual enterprise operation situation, considering the critical data security needs, a focused and feasible data security improvement plan shall be worked out.

**Implementation**
Finally, based on the established data security improvement plan, implement the project rigorously, check the project progress regularly, and ensure that project activities are completed in accordance to the plan.

### 3.2 Content of Data Security Practice

Once the data security risks are confirmed and the direction of work is determined for the enterprise, the data security work should be implemented. The following are several important tasks.

**Data Classification**
The core of data security lies in the management of data security throughout the entire data lifecycle. Implementing data lifecycle security management can further clarify the data protection requirements for each stage of the data lifecycle. It helps to allocate data protection resources and costs reasonably and establish a comprehensive data lifecycle protection mechanism.

In order to reasonably allocate data protection resources and costs, it is necessary to implement data classification. This will further clarify the targets of data protection and help enterprises to allocate theta protection resources and costs reasonably and to implement data security work in a focused manner, and lead to satisfactory returns. At the same time, a standardized data classification management system can enhance the safety of data sharing between organizations or industries, which is conducive to mining and realizing data value.

Data classification is the first step of data security work. The hierarchical classification of data covers not only the structured data, but also the unstructured data, which includes the data stored in information systems, and the data exists in business workflows and documents as well.

Data classification can be developed according to integrity, confidentiality, and availability. It is also a good practice to classify data based on its confidentiality level, like basic protection, special protection, and high protection. With the data classification, the corresponding protection measures should be set accordingly. The following are example (Table 1) of data classification and the corresponding data security requirements.

**Table 1.** Examples of Data Security Classification and Protection Measures

| Confidential level | Data security requirements |
|---|---|
| Public | There are no specific IT security requirements |
| Internal | User authentication<br>System accesses should be managed with permission control tools<br>The privileged access of data containing IT assets must be managed with privileged access management tools<br>The information system must be connected to the security monitoring tool before carrying out any day-to-day user activity<br>The information system must be connected to and maintained in the vulnerability management tool<br>Confidentiality labels must be in place<br>There must be the function to prevent data loss and there-fore restricts the data flows outside the organization<br>Data sharing with anybody outside of the organization should be limited and it must be explicitly approved by the data owner |
| Restricted | On top of the above<br>User authentication based on multi factor authentication<br>The information system must be connected to the security monitoring tool before conducting any detailed user activity<br>There must be the function to protect data loss and there-fore restricts the data flows inside and outside the organization<br>It is necessary to encrypt the static data inside the organization environment and the data in transit (external/internal)<br>Data can only be shared with a limited number of staff who have signed the Non-disclosure agreements and a limited number of vendors. The data sharing should be limited to the data owner or their designated personnel |

(*continued*)

**Table 1.** (*continued*)

| Confidential level | Data security requirements |
|---|---|
| Confidential | On top of the above<br>Clearly manage the approvals of the access granted<br>The information system must be connected to security tools to obtain complete user activity and security logs<br>Data sharing is strictly limited to the data owner or personnel approved by the data owner |

### Data Encryption

Data encryption is a long-history technology that converts plaintext into cipher-text through encryption algorithms and encryption keys, while decryption involves restoring ciphertext to plaintext through decryption algorithms and decryption keys. The core of it is Cryptology. Data encryption is still the most reliable way for computer systems to protect information. It uses cipher graph to encrypt information, to achieve information concealment, and thus to protect information security [3].

There's a number of data encryption methodologies. The proper methodology should be selected for different scenarios. The below Table 2 listed the data encryption methodologies in different scenarios.

**Table 2.** Data Encryption Methods and Scenario Examples

| Encryption level | Prevention scope | Applicable scenarios |
|---|---|---|
| Application level | Malicious direct access to databases | Must be applicable to restricted and confidential data |
| Database level | Malicious copying of database files (Invalid for authorized database users) | Must be applied in databases containing restricted data and confidential data. It can be superseded by the upper-level encryption |
| File level | Malicious copying of files<br>(Not suitable for file system administrators and authorized file users) | Must be applied to files containing restricted data and trade confidential data. It can be superseded by the upper-level encryptions |
| Disk level | Physical loss of disk | Must be applied to the disks out of the physical control of the company and the disks containing restricted data and confidential data. It can be superseded by upper-level encryption |

### Data Access Control

Data security focuses on protecting data from unauthorized access and corruption throughout the entire data lifecycle. The control of data access is the restriction on the usage of data. Firstly, data can only be used by authorized users, and the unauthorized users cannot use it; Secondly, the authorized users can only use data within the granted scope of permission, and any handling of data beyond the scope cannot be executed.

Data access control can be achieved in the following ways:

*Permission Control.* Access control is to restrict the users' usage of data. Data can only be used by authorized users, and the unauthorized users cannot use it. In addition, authorized users can only use data within the permitted scope, while not that exceeding the scope.

*System Access Control.* System access control provides the first level of security protection for the information system. Unauthorized personnel are unable to open the information system through authentication, which means they cannot access data or operate on it.

*Data access Frequency Control.* For the real users who have already been granted the relevant data access, if they access to the data more frequently than that matches their actual work requires, a rate limiting process should be performed. The common rate limiting algorithms include token bucket algorithm, leaky bucket algorithm, and counter algorithm.

### Dealing with External Threats

As a part of network security, data security has always been one of the key targets of network attacks, and the increasing number of external attack methods has brought great challenges and risks to data security and privacy protection. Enterprises should adopt machine learning algorithms to detect anomalies and identify threats as soon as possible with the analysis and warning based on the monitoring of log. The following are some examples of external threats and the measures to address them.

*Distributed Denial of Service Attack (DdoS).* It refers to the situation that multiple attackers in different locations attack one or more targets simultaneously, or one attacker that controls multiple machines in different locations attacks the target simultaneously using these machines. Due to the fact that the attack targets are distributed in different locations, this type of attack is called a distributed denial of service attack, and there could be multiple attackers [4]. Once a company is caught up in DDoS attacks, the servers will fall into access delay, access failure, or even server unavailability. To reduce the risk, enterprises can consider deploying cloud-based security protection software such as high defense IP.

*Phishing Fraud.* Phishing attack usually contains malicious attachments. Once being clicked and opened, an enterprise's devices would be attacked, leading to severe incidents of data theft or leakage. So, enterprises should use tools to block phishing emails and use anti-phishing engines. It is also very important to provide anti phishing trainings to employees on a regular basis. In addition, enterprises can also run phishing email exercises to raise the vigilance of the staff. For example, send phishing emails regularly or irregularly to designated or undesignated addresses, find potential weaknesses of security awareness based on the analysis of employees' reading and clicking behaviors, and then deliver the security trainings according to the findings.

*Ransomware.* Ransomware is a prevalent Trojan virus. It blocks the normal usage of data assets or computing resources with methods such as harassment, intimidation, or even kidnapping user files and extorts money from the users taking this condition. This type of user data assets includes documents, emails, databases, source codes, images, and compressed files, etc. The forms of ransom include real money, Bitcoins, or other virtual currencies [5]. Ransomware is usually combined with data theft. In ransom incidents, hackers may threaten to release the data if ransom is not paid. Ransomware has become the most common security threat targeting various industries and one of the important underground black industries of the Internet. Both enterprises and individuals can become the targets of ransomware attacks and extorting [6].

For enterprises, in order to avoid losses of interests due to ransomware, data security leaders need be proactive to deploy data protection strategies and technologies and to improve monitoring and detection capabilities. Nowadays, most ransomware technology would not be applied for attacks independently. Generally, it would be applied in combination with advanced network attacks, which increases the difficulty of defending against complex attacks. In these circumstances, a multi-layer network security protection system should be built combining a variety of protection methods, such as advanced threat protection, gateway antivirus, intrusion prevention, and other network-based security protection methods.

Network isolation has always been a good technology for network security protection, and it can also be used to defend against ransomware. Besides, it is necessary to backup and restore data on regular basis, since the reliable backup of data can minimize the losses caused by ransomware. But, at the same time, it is also necessary to have security protection for these data backups to avoid infection and damage of data. System updates and security patches, endpoint protection, network segmentation, security software websites, application whitelists, response plan development, and data security backup are all security protection measures. Other measures include terminal protection and monitoring of encrypted network traffic [7].

## 4 Construction of Data Security System

### 4.1 Data Security Evaluation

Before developing the data security system, an enterprise should be clear about the data security goals and how to evaluate data security. The security evaluation of data should be conducted from the following aspects: [8].

**Compliance.** The data is handled on the basis of meeting the requirements of laws and regulations throughout its entire lifecycle.

**Reasonable Access Control.** Ensure that only authorized personnel can access data.

**Data Encryption.** Encrypt data according to different data classification.

**Data Backup and Recovery.** Backup data regularly to ensure that no data loss is caused by accident.

**Data Security Audit.** Enterprises should make sure that internal audits is conducted on data security and regulation compliance at least once a year. Internal auditors must sit in an independent department to avoid any conflict of interest. Conduct an external audit at least every three years.

**Data Classification.** Proceed data classification in accordance with the relevant regulations of national information security level protection in a diligent manner.

**Employee Training.** All new employees should receive data security training within six months of employment. At least two data security related training sessions per year should be taken by all employees.

### 4.2 Data Security Standards

Laws and regulations provide direction for data security work, while numerous national and industry standards provide specific guidance for the implementation of security work. In order to design data security work and plans for enterprises, data security can be evaluated based on national and industry standards that are suitable for the enterprise.

The enterprise data security department shall conduct comprehensive research on the existing data security work of the enterprise, and identify the corresponding working areas. By conducting preliminary risk assessment and investment in improving security work, the of improvement can be determined.

After setting goals, enterprises can start building a data security system from the following four dimensions: organizational structure, process systems, technical tools, and personnel management (Fig. 3).

Organization

Establish a top-down data security management system that covers four levels: decision-making, management, execution, and supervision, and clarify organizational structure and job settings.

Process

Establish a unified data security management process system, clarify the responsibilities of data security work at all levels of departments and relevant positions, and standardize work processes.

Technology and tools

Adopting appropriate technical tools to strengthen security management: encryption, desensitization, authorization management, monitoring logs, user entity analysis, data leakage prevention.

Personnel management

Data security awareness education and training. Conduct safety audits on important positions, establish dedicated personnel and separate responsibilities, and if necessary, establish dual personnel and dual positions.

**Fig. 3.** Four dimensions of data security management

**Organizational Structure**

Data security governance needs to be carried out in a top-down manner, so it is very important to establish a data security management system that covers decision-making, management, execution, and supervision from top level to bottom level, as well as a clear organizational structure and job settings. Consider appointing a dedicated data security officer in the enterprise, who is fully responsible for the organizational construction of data security. The organizational structure can include the CEO and senior leaders reporting directly to the CEO as the data governance steering committee for the decision-making level. The management team consists of the Chief Information Security Officer, Data Protection Officer, Business Unit Heads, IT Heads, Compliance and Risk Management, and other key management positions. In the execution layer, dedicated team members will take the responsibilities to implement the day-to-day data security work. And the auditors will conduct audit review and supervision independently on specific data security works [9].

**Process**

At the process level, establish a unified data security management process, clarify the responsibilities of data security work at all levels of departments and relevant positions, and standardize work processes.

**Technical Tools**

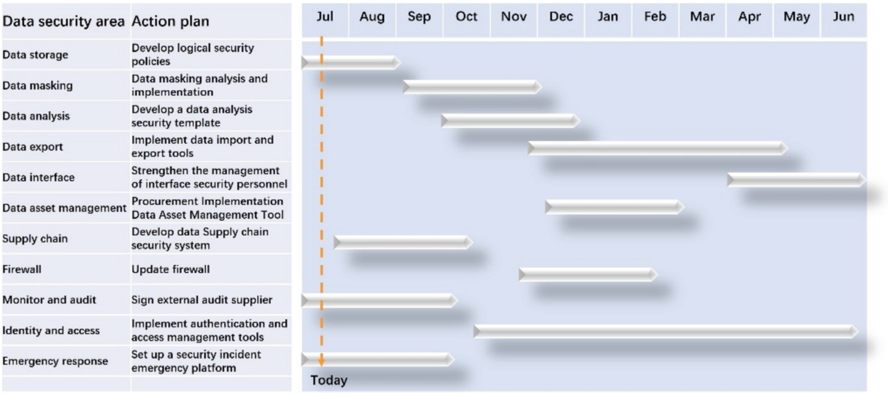
Strengthen security management by balancing risk and investment and adopting appropriate technical tools: to protect data using encryption and data masking technologies, to expend the coverage of existing authorization management systems, to reinforce the standards of log monitoring, to analyze the users and entities based on the logs, to make alerts for abnormal situations and handle them promptly. To prevent potential problems, enhance the usage of data leakage prevention tools, establish prevention and detection mechanisms, and assign dedicated personnel to monitor the alarms triggered by the tools on a timely basis.

**Personnel Management**

In terms of personnel management, it is necessary to enhance data security awareness education and training. Conduct safety examination on key positions, establish dedicated personnel and separate responsibilities. If necessary, establish dual personnel and dual positions [10].

## 4.3 Data Security System Construction Plan

Based on the evaluation and analysis and the goals set for data security, enterprises can develop rigorous improvement plans and implement it accordingly. Here is a simple example, and there are a number of data management process systems and documents behind it (Fig. 4).



| Data security area | Action plan |
|---|---|
| Data storage | Develop logical security policies |
| Data masking | Data masking analysis and implementation |
| Data analysis | Develop a data analysis security template |
| Data export | Implement data import and export tools |
| Data interface | Strengthen the management of interface security personnel |
| Data asset management | Procurement Implementation Data Asset Management Tool |
| Supply chain | Develop data Supply chain security system |
| Firewall | Update firewall |
| Monitor and audit | Sign external audit supplier |
| Identity and access | Implement authentication and access management tools |
| Emergency response | Set up a security incident emergency platform |

**Fig. 4.** Example: Developing an improvement plan and implementing it

After the analysis, the data security officer (dedicated personnel) of the enterprise can define the goals of the works following and determine the key points. Then the data security work can be carried out like clockwork.

## References

1. Debo, H., Sus, E., Lau, S., Ele, S., Ev, S.: DAMA-DMBOK Data Management Body of Knowledge, 2nd edn. DAMA International, pp. 217–220 (2017)
2. We, L., Liw, Ch., Li, J.F., Yac, Z., et al.: Data Lifecycle Security Specification, pp. 5–7 (2021)
3. Kepu, C., Kexu, B.: Data Encryption. Baidubaike. Kepu China. Scientific Encyclopedia Entry Writing and Application Work Project. https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E5%8A%A0%E5%AF%86/11048982?fr=aladdin. Accessed 20 Apr 2022
4. Ling, W.M.: Data service platform: distributed denial of service attacks and preventive measures. In: International Academic conference on Office Automation, p. 1, 20 November 2018
5. Kepu, C., Kexu, B.: Rasomware. Baidubaike. Kepu China. Scientific Encyclopedia Entry Writing and Application Work Project. https://baike.baidu.com/item/%E5%8B%92%E7%B4%A2%E8%BD%AF%E4%BB%B6/5243210?fr=ge_ala. Accessed 06 Jan 2023
6. Baijiahao. https://baijiahao.baidu.com/s?id=1769993208230818350&wfr=spider&for=pc. Accessed 07 Jan 2022
7. Lia, L.J., Bin, W.Y., Du, Q.J., Tian, W.Z,. Fe, M.: Data Security Practice Guide, pp. 155–158 (2022)
8. Ning, W.A., Ka., Y.: Data Security Area Guide, pp. 175–180 (2022)
9. Ladl, J.: Data Governance: How to design, deploy, and sustain an effective data governance program, p. 27 (2021)
10. Li, Z.: Data Governance and Data Security, pp. 105–108 (2019)