DengFeng Qiao Yi He MingYi Wang

II. Data protection

What is data privacy?

Data privacy refers to the control that individuals or organizations have over the collection, storage, use, and disclosure of their data. It ensures that personal and business data can only be accessed, used, and disseminated with authorization.

Key elements for organizations to maintain data privacy compliance

Practical aspect: Comprehensive data privacy training for employees is crucial. Give employees a deep understanding of data handling norms and responsibilities, and understand what is allowed and what is not. At the same time, regular data privacy audits are conducted to carefully check whether the data processing process complies with the privacy policy, so as to detect and correct potential problems in time.

Rules: Organizations need to have clear and unambiguous rules for data collection, use, and storage. For example, only collect data that is relevant to the business, and avoid excessive collection of personal information. At the same time, the retention period of data should be clearly stipulated, and the data beyond the period should be properly processed.

Guidelines: Publish detailed data privacy protection guidelines, e.g. when dealing with sensitive data, encryption must be used to ensure data is secure in transit and storage.

Tooling: Secure your data with advanced data encryption tools, access control tools, etc. For example, encrypting software is used to encrypt stored sensitive data, and authentication systems are used to severely restrict access to data.

The importance and difference of data privacy for individuals and businesses

For individuals, personal privacy data (such as personal identity information, financial information, etc.) is extremely important. If these data are leaked, it may lead to serious consequences such as loss of personal property and damage to reputation. Individuals are more concerned about the security of personal information in various fields such as the Internet, financial institutions, and medical institutions. For example, when shopping online, individuals worry that their personal information will be illegally shared by merchants to third parties, and thus they will suffer from the harassment of spam SMS and scam phone calls.

For enterprises, enterprises have a large amount of business data (such as customer information, trade secrets, etc.). Data breaches may cause companies to lose the trust of customers, face legal actions, and suffer huge financial losses. In addition to the security of internal data, enterprises also pay special attention to the security of data transmission between partners and third-party service providers, as well as strict compliance with relevant laws and regulations to avoid penalties due to data privacy issues.

3. Database security

As an important carrier of data storage, database may face a variety of security threats, such as malicious network attacks, software vulnerabilities, intrusions, misuse or care-ness and damage.

In order to ensure the security of the database, we need to establish a complete set of tools, processes and methods to ensure the security of the database system, which includes both the security of the data stored in the database, and the security of the data management system itself.

In terms of preventing malicious network attacks, network security devices such as firewalls, intrusion detection systems (IDS) and intrusion prevention systems (IPS) can be used. A firewall is like a solid wall that filters network traffic according to preset rules to prevent unauthorized access. IDS and IPS can monitor abnormal activities in the network in real time and take corresponding defense measures.

For software vulnerability issues, patches of database management systems and related software should be updated in time to fix known vulnerabilities. At the same time, security vulnerabilities should be scanned regularly to actively find potential security problems, so as to prevent them before they occur.

In terms of preventing intrusion, misuse or careless and damage, establishing a strict user rights management mechanism is the key. Different database access privileges are assigned to users according to their roles and responsibilities. For example, ordinary employees can only query and modify data related to their own work, while administrators have higher privileges. Detailed logging of database operations is performed so that abnormal operations can be found in time. For example, the user login time, operation content and other information are recorded, and if abnormal login or data modification behavior is found, it can be investigated in time. In addition, database security training is carried out for employees to improve their security awareness and reduce security problems caused by misuse or care-ness.

4. Characteristics and Applications of Big Data

Big data has the following remarkable characteristics:

Data volume: The amount of data generated and stored in the system is huge, covering a huge amount of information resources.

The speed of data processing: Data is generated, collected, and analyzed at an extremely fast rate, reflecting the dynamics of things.

Data Type aspect: Ability to handle diverse data types from multiple sources, including structured, semi-structured and unstructured data.

Data quality aspects: The quality and accuracy of the data are important considerations for big data, and only accurate data can lead to reliable conclusions.

There are several efficient ways to store data:

One is a framework that enables distributed processing of large, complex datasets, scales to thousands of nodes, each providing local computation and storage, and is open source. The other is an open source distributed NoSQL database modeled after Google's Bigtable, which can be used with HBase or as a supplement to Hadoop with a column-oriented store. Such databases can manage large tables with billions of rows and millions of columns. There is also a cross-cloud platform that is capable of providing fast, reliable and secure access to data in various formats, while enabling fast and reliable processing and querying, and supporting secure collaboration.

Big data analytics is the process of extracting meaningful insights from large amounts of data, such as:

Through analysis, we can understand customer behavior and optimize customer experience. Predicting future trends to help companies make better business decisions; Improve your marketing campaigns by understanding what works and what doesn't Improve operational efficiency by identifying bottlenecks in the operation process; Reduce losses by detecting fraud and other forms of misuse earlier.

In summary, every link of data, from its basic concept to protection measures, to its application in the field of big data, is crucial. We must deeply understand the nature of data, pay attention to data protection, and make full use of the advantages of big data, in order to move forward steadily in the wave of digitalization.