DengFeng Qiao Yi He MingYi Wang

I. Data and related concepts

What is data?

Data is an objective record of facts, observations, or commands, and it can take the form of numbers, words, images, sounds, and so on. For example, the daily temperature values, the text content of an article, a beautiful painting image, and the audio of a beautiful song are all data. Data is like an unprocessed treasure, containing a wealth of information resources.

What is information?

Information is the product of data processing, organization and interpretation, and it has a clear meaning for the receiver. For example, the temperature data in a period of time is statistically analyzed, and the trend of temperature change is obtained, which is the information. Summarizing the text content of an article, the core point extracted is also information.

The difference between data and information

Data is often raw, isolated elements, like pearls scattered on the ground. And the information is to connect these pearls in series, after processing and integration to form a fine necklace. The representation form of data is relatively single, and information is often the comprehensive result of multiple data correlation, combination and analysis. For example, a single sales order amount is the data, and the trend of sales growth after analyzing multiple order amounts is the information.

What is metadata?

Metadata can be understood as data that describes data. It contains details about when the data was created, who created it, the format of the data, where the data came from, and more. For example, for a photo, the metadata might include the date it was taken, the camera model used, the resolution of the photo, and so on.

Why do we need metadata?

From the point of view of data management, metadata can let us quickly understand the basic situation of the data, and quickly locate the information we need in the massive data. For example, in a large filesystem, metadata can be used to easily find files of a particular format or when they were created. In the aspect of data sharing, metadata can make different users or systems better understand the meaning and usage of data, thus promoting the circulation and sharing of data. For the long-term preservation of data, metadata can ensure the integrity and understandability of data during the process of data migration or storage medium replacement.