

Metadata is as important as the data itself

Data analytics, by design, is a field that thrives on collecting and organizing data. In this reading, you are going to learn about how to analyze and thoroughly understand every aspect of your data.



Take a look at any data you find. What is it? Where did it come from? Is it useful? How do you know? This is where metadata comes in to provide a deeper understanding of the data. To put it simply, **metadata** is data about data. In database management, it provides information about other data and helps data analysts interpret the contents of the data within a database.

Regardless of whether you are working with a large or small quantity of data, metadata is the mark of a knowledgeable analytics team, helping to communicate about data across the business and making it easier to reuse data. In essence, metadata tells the who, what, when, where, which, how, and why of data.

Elements of metadata

Before looking at metadata examples, it is important to understand what type of information metadata typically provides.

Title and description

What is the name of the file or website you are examining? What type of content does it contain?

Tags and categories

What is the general overview of the data that you have? Is the data indexed or described in a specific way?

Who created it and when

Where did the data come from, and when was it created? Is it recent, or has it existed for a long time?

Who last modified it and when

Were any changes made to the data? If yes, were the modifications recent?

Who can access or update it

Is this dataset public? Are special permissions needed to customize or modify the dataset?

Examples of metadata

In today's digital world, metadata is everywhere, and it is becoming a more common practice to provide metadata on a lot of media and information you interact with. Here are some real-world examples of where to find metadata:

Photos

Whenever a photo is captured with a camera, metadata such as camera filename, date, time, and geolocation are gathered and saved with it.

Emails

When an email is sent or received, there is lots of visible metadata such as subject line, the sender, the recipient and date and time sent. There is also hidden metadata that includes server names, IP addresses, HTML format, and software details.

Spreadsheets and documents

Spreadsheets and documents are already filled with a considerable amount of data so it is no surprise that metadata would also accompany them. Titles, author, creation date, number of pages, user comments as well as names of tabs, tables, and columns are all metadata that one can find in spreadsheets and documents.

Websites

Every web page has a number of standard metadata fields, such as tags and categories, site creator's name, web page title and description, time of creation and any iconography.

Digital files

Usually, if you right click on any computer file, you will see its metadata. This could consist of file name, file size, date of creation and modification, and type of file.

Books

Metadata is not only digital. Every book has a number of standard metadata on the covers and inside that will inform you of its title, author's name, a table of contents, publisher information, copyright description, index, and a brief description of the book's contents.

Data as you know it

Knowing the content and context of your data, as well as how it is structured, is very valuable in your career as a data analyst. When analyzing data, it is important to always understand the full picture. It is not just about the data you are viewing, but how that data comes together. Metadata ensures that you are able to find, use, preserve, and reuse data in the future. Remember, it will be your responsibility to manage and make use of data in its entirety; metadata is as important as the data itself.