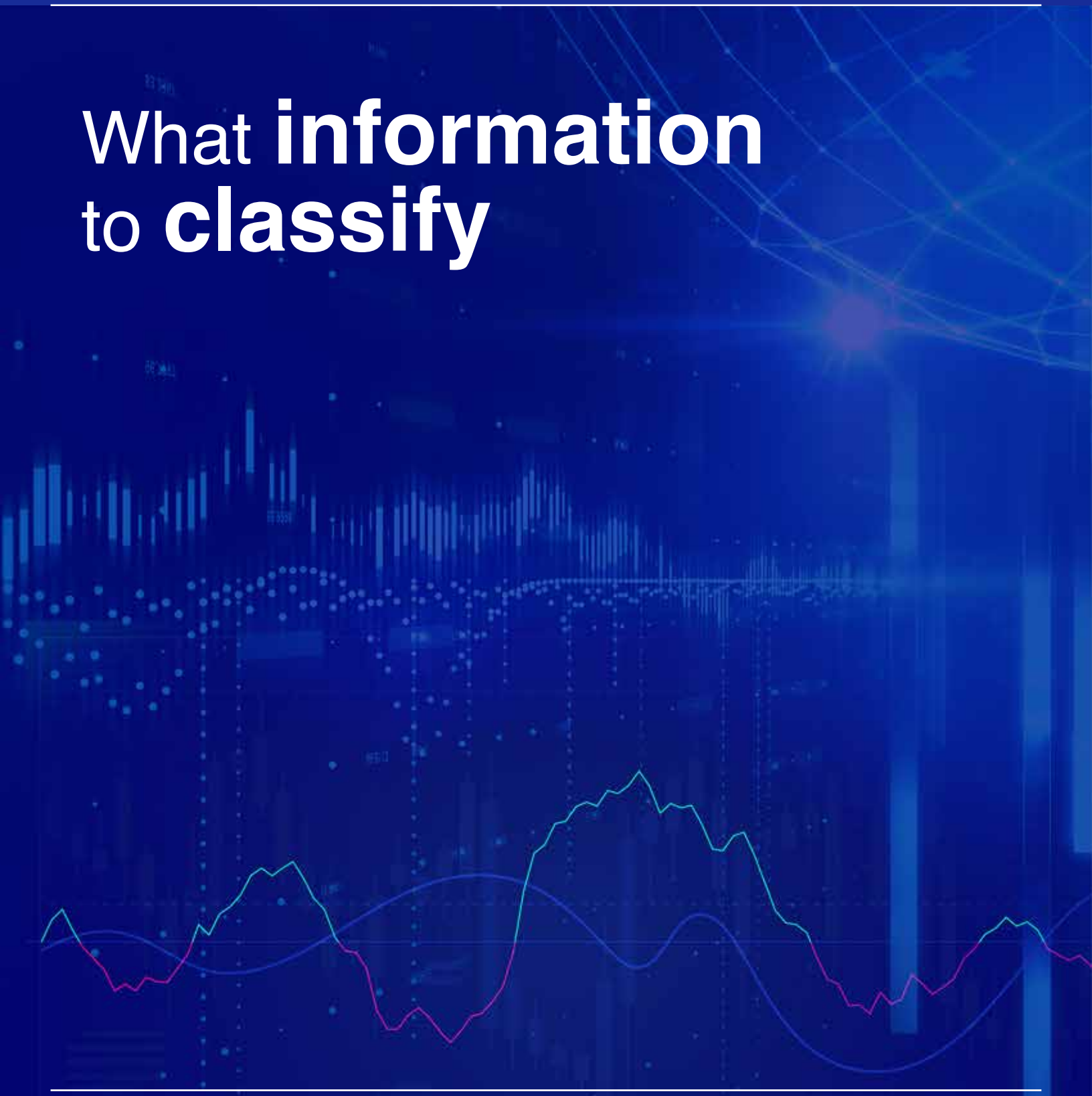
kriptos® Intelligent Solutions for Data

What **information** to **classify**



Contáctanos | **www.kriptos.io**

kriptos Intelligent Solutions for Data

424 Johnson St Unit A, Sausalito,
CA 94965, USA
info@kriptos.io

Structured vs Unstructured information

Information is typically defined as data with context. Data like a sequence of numbers is just a sequence of numbers until you get the context that it is a credit card or a bank account number. The added context transforms the data to information. With that broad definition there are many ways to start categorizing information. In the IT world the information is often first categorized into structured and unstructured data

.

Structured data means that the data is structured in a relational database with rows and columns in tables. Structuring data this way means that it is easy to query the data. While it may not be easy to classify structured data, with a little effort it is a feasible task.
Tables can be classified depending on the columns of data they contain. Tables that contain sensitive data like for example credit card information can be classified as confidential.

Unstructured data and information on the other are everything that is not structured in a relational database.
This includes but is not limited to:

- Text documents
- Word documents
- PDF documents
- Spreadsheets
- Emails and messages
- Social media data (Facebook, Twitter, WhatsApp, etc)
- Server, application and website files and logs
- Online newspapers and blogs
- Images
- Audio and video files

Querying unstructured data is not feasible. You can search through unstructured data looking for key words, but it is not the same as querying structured data. This also complicates the work of classifying the data as you need to analyze each document to know its content before you can classify.

Although structured data is different from unstructured data, both types of data should be classified according to the sensitivity level of the data.

kriptos Intelligent Solutions for Data

424 Johnson St Unit A, Sausalito, CA 94965, USA
info@kriptos.io

Types of information

Information or data can be divided into four types of information. These four types are:

Facts – Facts are data that is true and is not influenced by opinion or open for discussion.

Fictional – Fictional information is information that has been created by someone for the purpose of entertainment or persuasion.

Subjective – Subjective information is information that is influenced by someone´s opinion or point of view.

Objective – Objective information unlike subjective information attempts to be unbiased and present multiple points of view.

These four types of information often co-mingle and can be found together in the same document. A research paper can have both facts and objective analysis of those facts. A blog post or a news article can contain all of the above as it attempts to entertain and enlighten at the same time.

Of these four types of information, the facts are the most critical type when it comes to classification. Someone´s opinion can be sensitive and there may be good reason to protect an objective analysis, the facts should normally come first. This is where you can find most of the important information that is protected by laws and regulations.

Category of information

Another way to categorize information is by category of information or data. This is based on the function that the information has or in what area it is used. Types of information or data can include but is not limited to:

- **Financial data**
- **Research and development data**
- **Education and training information**
- **Personal identifiable information**
- **Health data**
- **Entertainment information**
- **Location data**
- **Communication**

Sometimes the same data can be put in several categories like for example a person´s health records can be considered health data and personal identifiable data.

While all these categories of information should be classified, it is critical to identify and classify information that needs to be protected according to law or regulation. This usually includes financial data, personal identifiable data and health data. These categories are again quite wide and include a lot of different data. The personal identifiable information category alone can contain any information relating to an identified or identifiable natural person.

This can include name, address, identification numbers, location data, online identifiers, physical traits, physiological, genetic, mental, economic, cultural or social identity.

Conclusion

There are many ways to divide and categorize information, but this separation should not be performed with a goal to know which information to classify. It is by classifying all information that you will be able to protect that which is important.