Unstructured Data

Definition and examples
What is unstructured data?

- It is information that is not organized in some type of data structure and has not a predefined data model.
- Unstructured data is not contained in a database. This makes it hard for computers to analyze unstructured information.
- Usually, unstructured data is in a text form. Only 20 percent of the data available to businesses is structured data, and the rest 80 percent is unstructured.
- While structured data is highly organized and is easy for analysis, unstructured information cause a lot of challenges to data analysis experts.
Among the most popular unstructured data we use every day for business:

- **Emails**

- **Text files**
  - Word processing, spreadsheets, PDF files, reports, graphics files, XML files and etc.

- **Websites**
  - YouTube, photo sharing sites, Instagram, Flickr, business websites, etc.

Examples of Unstructured Data:

- **Social Media**
  - The data generated from the social media platforms such as Facebook, Twitter, LinkedIn

- **Media**
  - All sorts of media such as digital images, audio, video, MP3, and etc.

- **Mobile data**
  - This includes information such as different types of text messages and location
Communications
All types of communication such as live chat, collaboration software, IM, etc.

Customer-generated content
Online comments, emails or calls to customer service department, etc.

Raw data from research
Surveys, focus groups, interviews are common unstructured data examples

Books, magazines, blogs
All of the writings (books, news, articles, press releases, and etc.)

Medical records
The significant list of healthcare records that create the whole “patient picture”

Scientific data
Oil gas exploration, seismic imagery, atmospheric data, space exploration and so on.