Structured data exists in a format created to be captured, stored, organized and analyzed. Unstructured data is the data which does not conforms to a data model and has no easily identifiable structure such that it cannot be used by a computer program easily. Structured data consists of numbers and values, whereas unstructured data consists of sensors, text files, audio and video files, etc. Structured data has a predefined data model and is formatted to a set data structure before being placed in data storage, whereas unstructured data is stored in its native format and not processed until it is used. Structured data is used in machine learning (ML) and drives its algorithms, whereas unstructured data is used in natural language processing (NLP) and text mining.

Structured data is sourced from GPS sensors, online forms, network logs, web server logs, OLTP systems, etc., whereas unstructured data sources include email messages, word-processing documents, PDF files, etc.