

## Examrace

### Big Data System Part 2 Computer Science YouTube Lecture Handouts

Get unlimited access to the best preparation resource for competitive exams : [get questions, notes, tests, video lectures and more](#)- for all subjects of your exam.

Topics:

1. What is data?
2. Types of big data
3. Structure data
4. Unstructured data
5. Semi structure

#### Introduction of Big Data

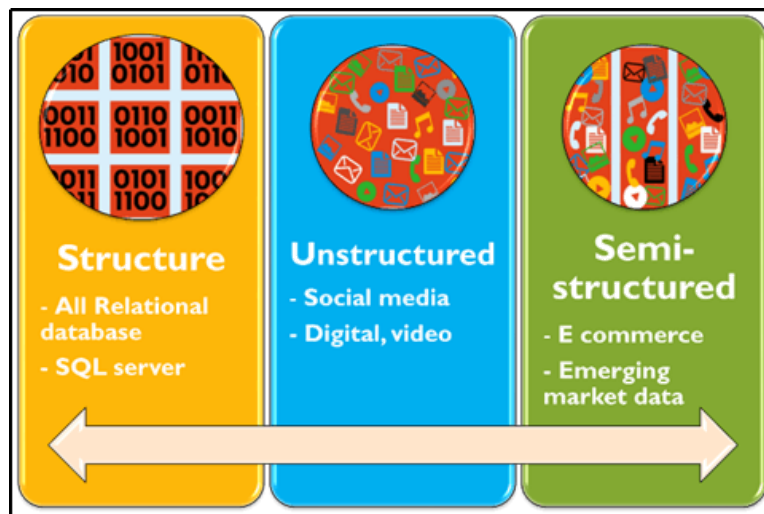
- Big data is also a data which are very large in size. Big data means huge data that increases over time.
- Big data term is used to describe a large amount of data.
- Usually we work on megabytes and gigabytes size but the size of big data is in terabytes of Petabytes or Exabytes or even more.
- Big data is in different formats that cannot be handled by traditional tools and applications and constantly increasing the size of this data.

Example

Facebook's database generates more than 500 terabytes of data every day, this data is generated mainly by commenting photos and videos upload messages, etc.

#### Types of Big Data

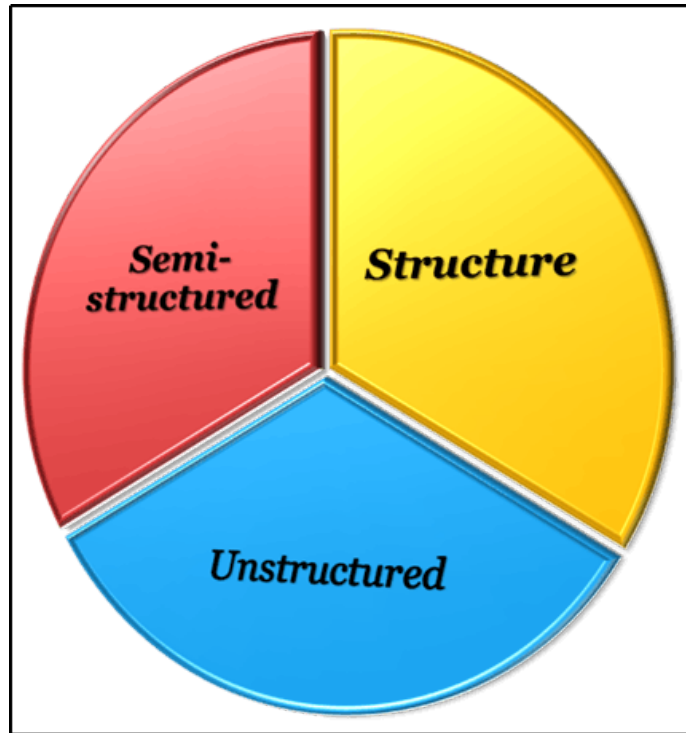
**Big Data ' could be found in three forms:**



©Examrace. Report @violations @<https://tips.fbi.gov/>

## Structured

- Structure data is data that is well-organized and through which we can easily get information. Big data contains a large amount of data, out of which the data is well organized and can be easily stored, processed and analyse.
- So structure data refers to highly organized information that can be used easily.
- **Structure data can be stored in tabular column such as relational database.**



[@Examrace. Report @violations @https://tips.fbi.gov/](https://tips.fbi.gov/)

For example:

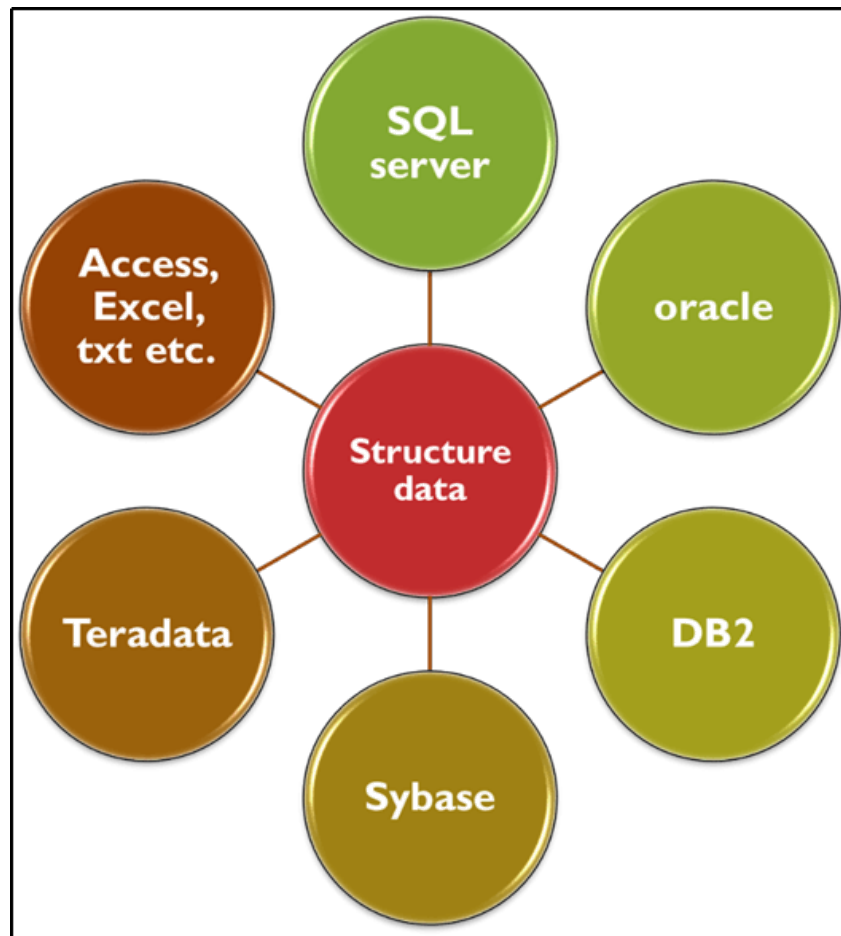
There are 5000 employees in an organization, maintaining data of all these employees, storing them, processing and analysing them so that we can easily get any information and take any decision. Because structure data is limited data and does not produce any confusion.

There are two source of structure data

Machine-generated structured data	Human-generated structured data
<ul style="list-style-type: none"><li>• All data received from sensors, weblogs, and financial systems for ex. (medical devices, GPS data, )</li></ul>	<ul style="list-style-type: none"><li>• all the data a human input into a computer for ex.(name and other personal details)</li></ul>

©Examrace. Report @violations @<https://tips.fbi.gov/>

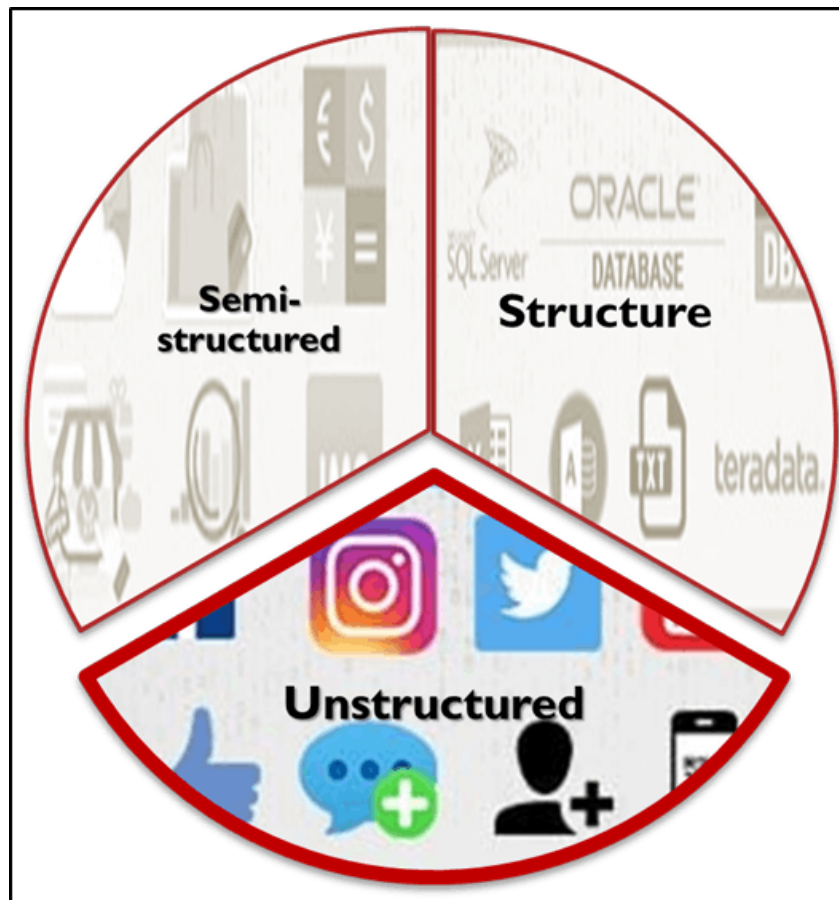
Structure data



©Examrace. Report @violations @<https://tips.fbi.gov/>

### Unstructured

- Unstructured data is opposite to the structured data. Unstructured data is un-organized data, which is very difficult to store, process and analyze.
- This data is very large and unlimited, which is not easy to store and process. So this is not a tabular database.
- Processing and analysing unstructured data is very difficult and time consuming.

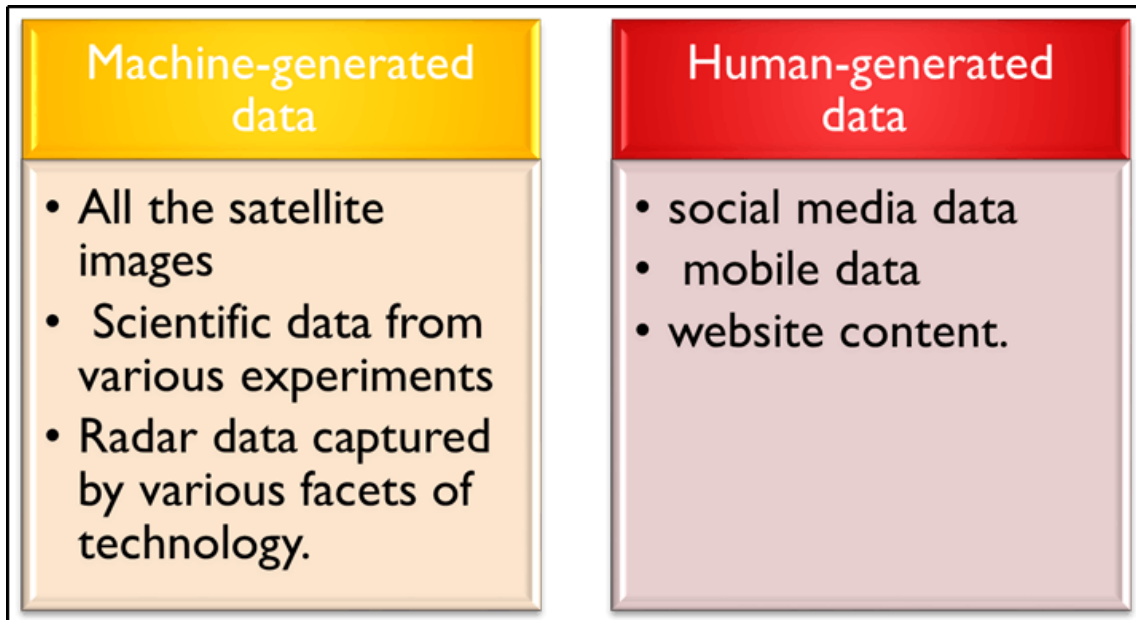


<https://tips.fbi.gov/>

For example:

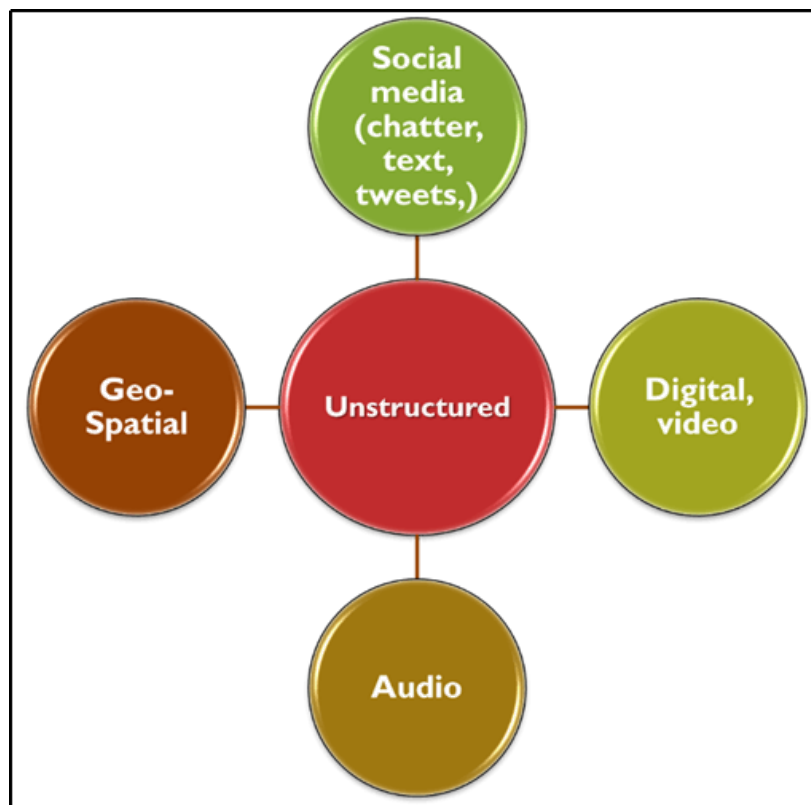
Photos videos that are uploaded to social media such as twitter, Facebook, Instagram, WhatsApp are unstructured data. These data are on a very large scale and in very large quantities, which is not so easy to store, process and maintain.

There are two source of unstructured data



©Examrace. Report @violations @<https://tips.fbi.gov/>

**Unstructured**

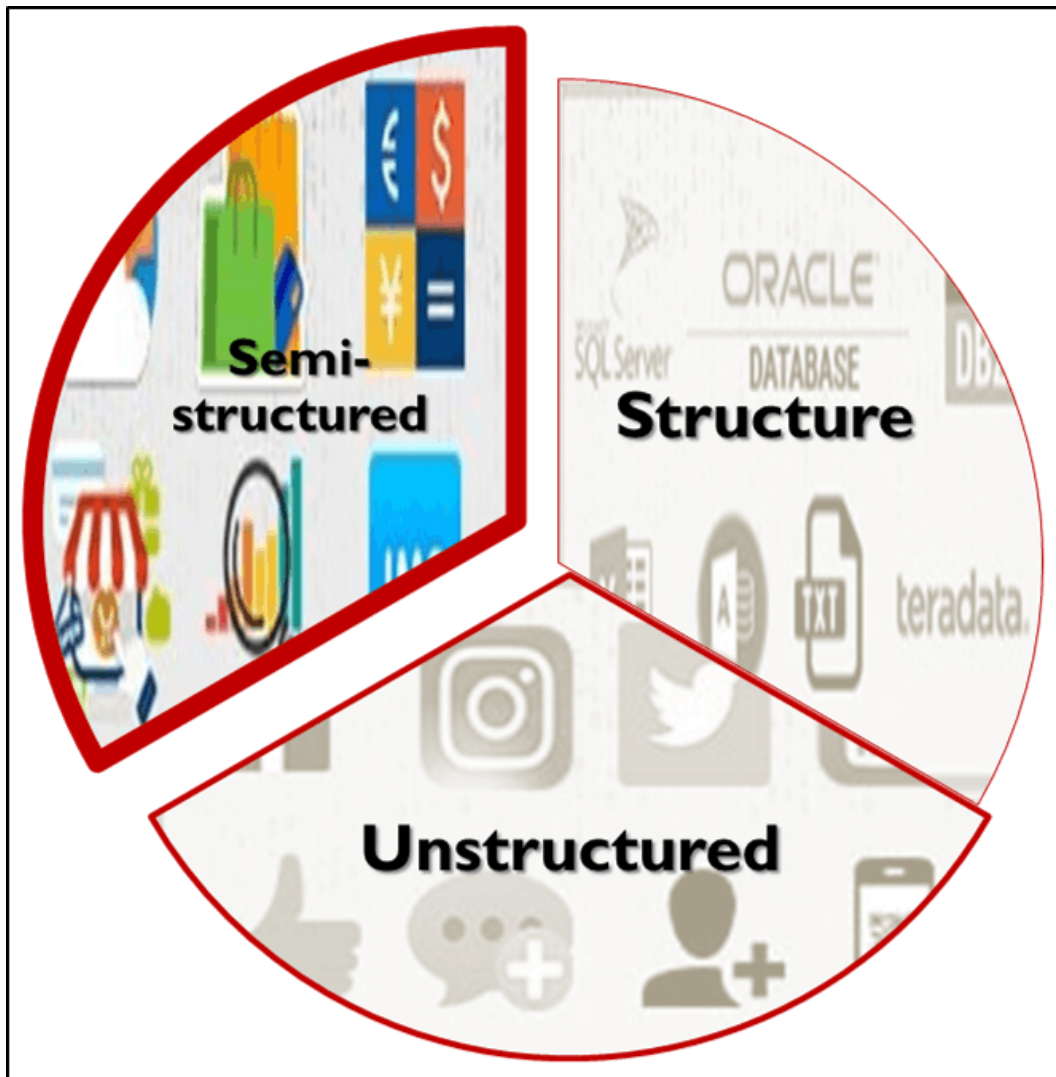


*©Examrace. Report @violations @<https://tips.fbi.gov/>*

### Semi- Structured

Semi-structured data is a combination of structure and unstructured data. This is called hybrid data. Hybrid data is the combination of organized and un-organized data. We cannot store of these type of data using traditional database format, but it contain some organizational properties.

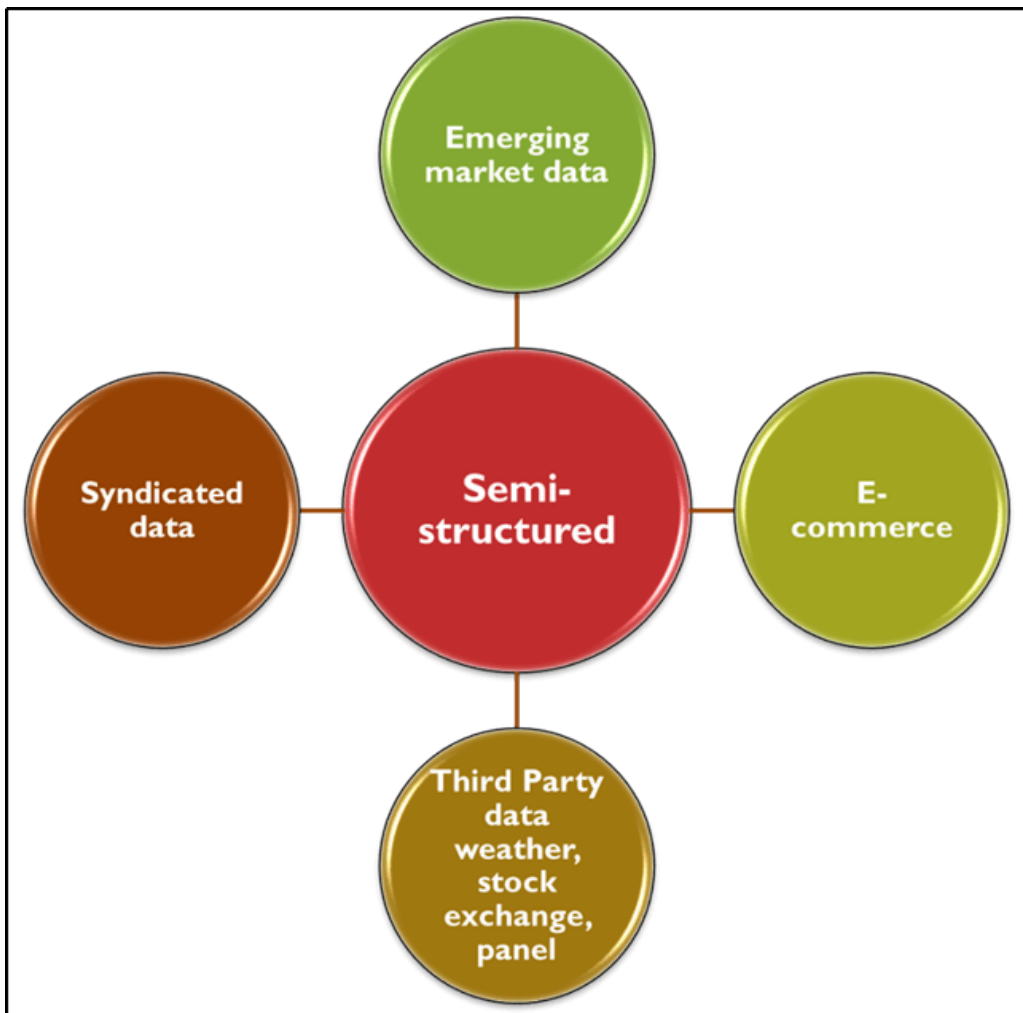




©Examrace. Report @violations @<https://tips.fbi.gov/>

For example:

- For example, NoSQL documents are considered to be semi-structured, since they contain keywords that can be used to process the document easily.
- **XML data, JSON files**, and others.



[@Examrace. Report @violations @https://tips.fbi.gov/](https://tips.fbi.gov/)

### MCQs

Q1. Which data is called hybrid data?

Options:

1. Organized data
2. Unorganized data
3. Combination of organized & unorganized data
4. None of this

Answer: 3

Q2 ... Data can be stored in tabular column such as relational database.

Options:

1. Structure data
2. Unstructured data
3. Semi structure
4. None of this

Answer: 1

✉ Mayank

Developed by: [Mindsprite Solutions](#)