

Introduction to

Big Data

Data ???

Anything and everything is Data



Where it comes from?

- Everywhere

We speak , we move, sensors , computers , documents etc..

Classification of Data

- Structured
- Unstructured

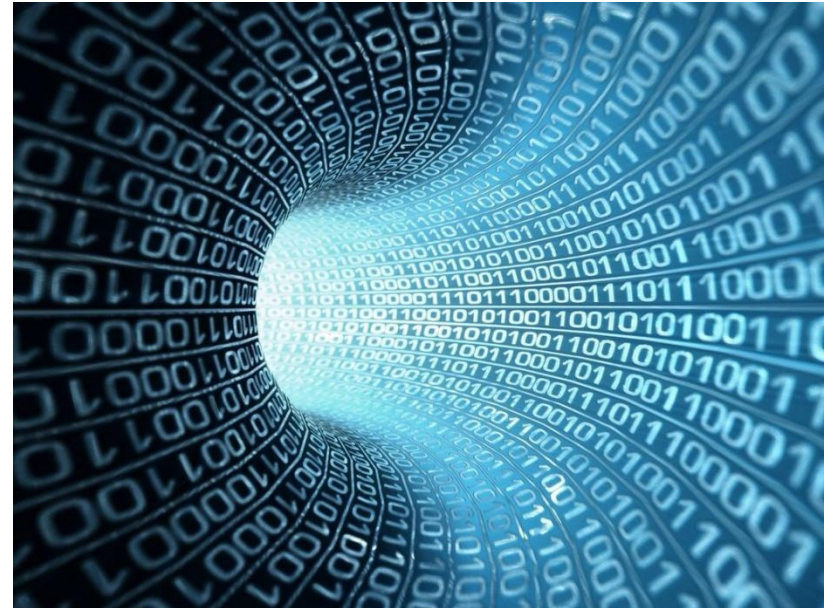
Human Generated Data – email , blogs , videos, pictures, etc.

Machine Generated Data – Automatics alerts , logs, sensor data

What is Big Data?

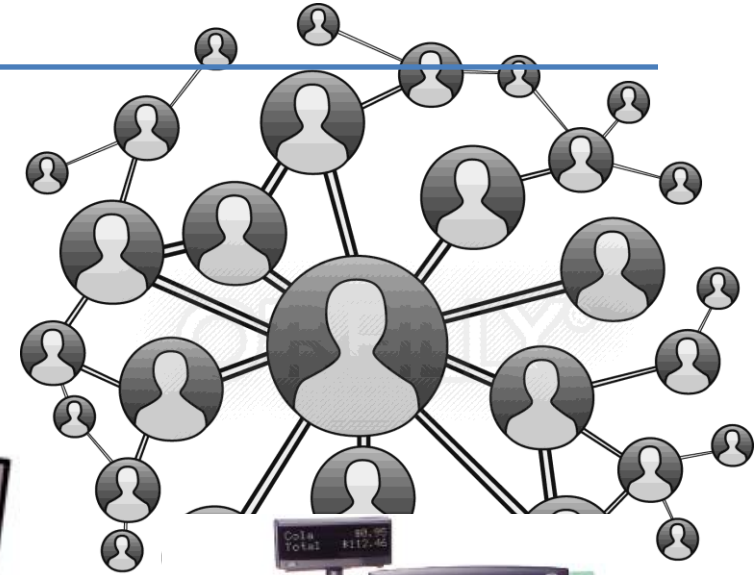
Big data is the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications.

src:wikipedia



Big Data is Everywhere

- Experiments ...Research...
- Photo Sharing
- CCTVs
- Social Network
- e-Commerce
- Point of Sales
- Mobile
-



The “Data” Rises

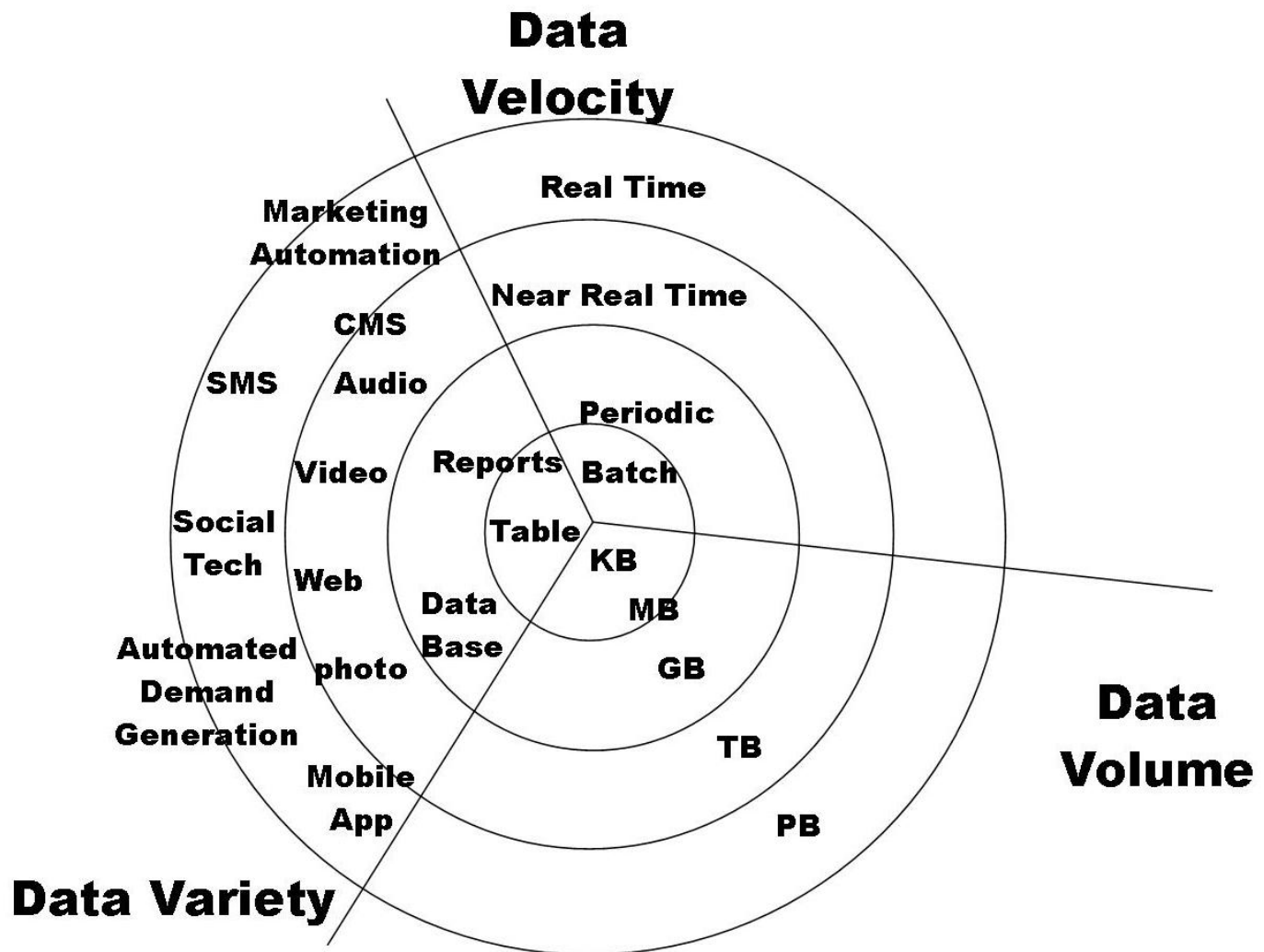
- 2005 – 130 Exabytes
- 2010 – 1227 Exabytes
- 2012 – 2837 Exabytes
- 2015 – 8591 Exabytes
- 2020 – 40026 Exabytes



Examples

- Google processes 20 PB a day (2008)
- Facebook has 2.5 PB of user data + 15 TB/day
- eBay has 6.5 PB of user data + 50 TB/day
- CERN's Large Hydron Collider (LHC) generates 15 PB a year

Characteristic of Big Data



Big Data Generators

Next Generation Web



Mobile Devices



Sensors

Facts

- Most of the data goes unprocessed till now.
- Approx 90% of data is unstructured and Enterprises was unable to use the data.
- Traditional Systems not able to store Big Data very well.

