# Big Data Introduction

The term "Big Data" is used to describe the **collection of complex and large data sets** such that it's difficult to capture, process, store, search and analyze this kind of data using conventional data base management tools and traditional databases management systems.

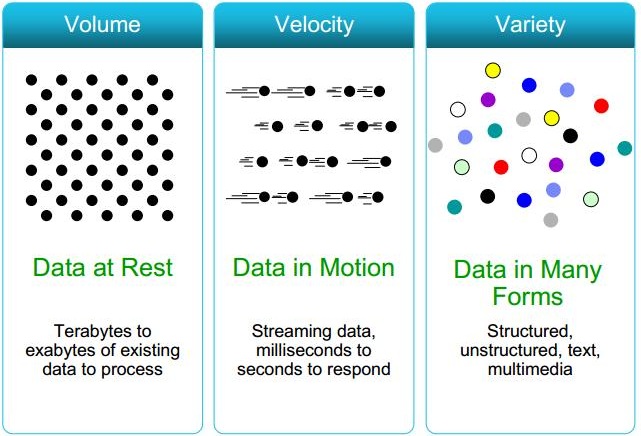
## Where does Big Data come from?

Basically the data coming from everywhere like

* + - sensors used to gather climate information
    - posts to social media sites
    - digital pictures and videos
    - software logs, cameras
    - microphones
    - scans of government documents
    - GPS trails
    - purchase transaction records
    - cell phone GPS signals
    - traffic

All these together constitute Big Data.

Characterization of Big Data - Volume, Velocity and Variety (3Vs):



* 1. **Volume:** The benefit gained from the ability to process large amounts of information is the main attraction of big data analytics. More data leads to more accurate analysis.   
     **Example:** If you could run that forecast taking into account 300 factors rather than 6, could you predict demand better?   
     Turn 12 terabytes of Tweets created each day into improved product sentiment analysis.   
     Convert 350 billion annual meter readings to better predict power consumption
  2. **Velocity:** Sometimes 2 minutes is too late. For time critical applications where Time is the core factor such that catching the frauds, catching the hackers, running status of train, big data must be used as it streams into your enterprise in order to maximize its value. Not only is the volume of data large, it is arriving ever more rapidly.   
     **Example:** Scrutinize 5 million trade events created each day to identify potential fraud.
  3. **Variety:** Big data includes both structured and unstructured data such as text, sensor data, audio, video, click streams, log files and more. New insights are found when analyzing these data types together.  
     **Example:** Monitor 100's of live video feeds from surveillance cameras to target points of interest Exploit the 80% data growth in images, video and documents to improve customer satisfaction.

All these types of data can have a significant effect on a business. Finding out quickly what the data means and understanding its importance provides a business with an ongoing advantage as well as the opportunity to realize competitive benefits.

4th Challenge - Veracity:

