EMERGING TECHNOLOGIES

KINGSFORD KISSI MIREKU Ph. D.

CENTRAL UNIVERSITY

FAITH • INTEGRITY • EXCELLENCE

OBJECTIVES

1. Understand the impact that emerging technologies will have in the future.
2. Define the use of emerging technologies –
   a) Cloud Computing,
   b) Big data concept and analytics,
   c) Data Centre and virtualization,
   d) internet of things,
   e) 4G/5g Mobile technologies,
   f) optical computing,
   g) quantum computing and quantum cryptography,
   h) virtual reality and wearable computing.
4. Define the ways in which certain technologies will impact homes of the future.
5. Describe some emerging technologies and their uses that are extreme.

MODULE 2

BIG DATA CONCEPT AND ANALYTICS

CENTRAL UNIVERSITY

FAITH • INTEGRITY • EXCELLENCE

MODULE 2

What is Big Data?

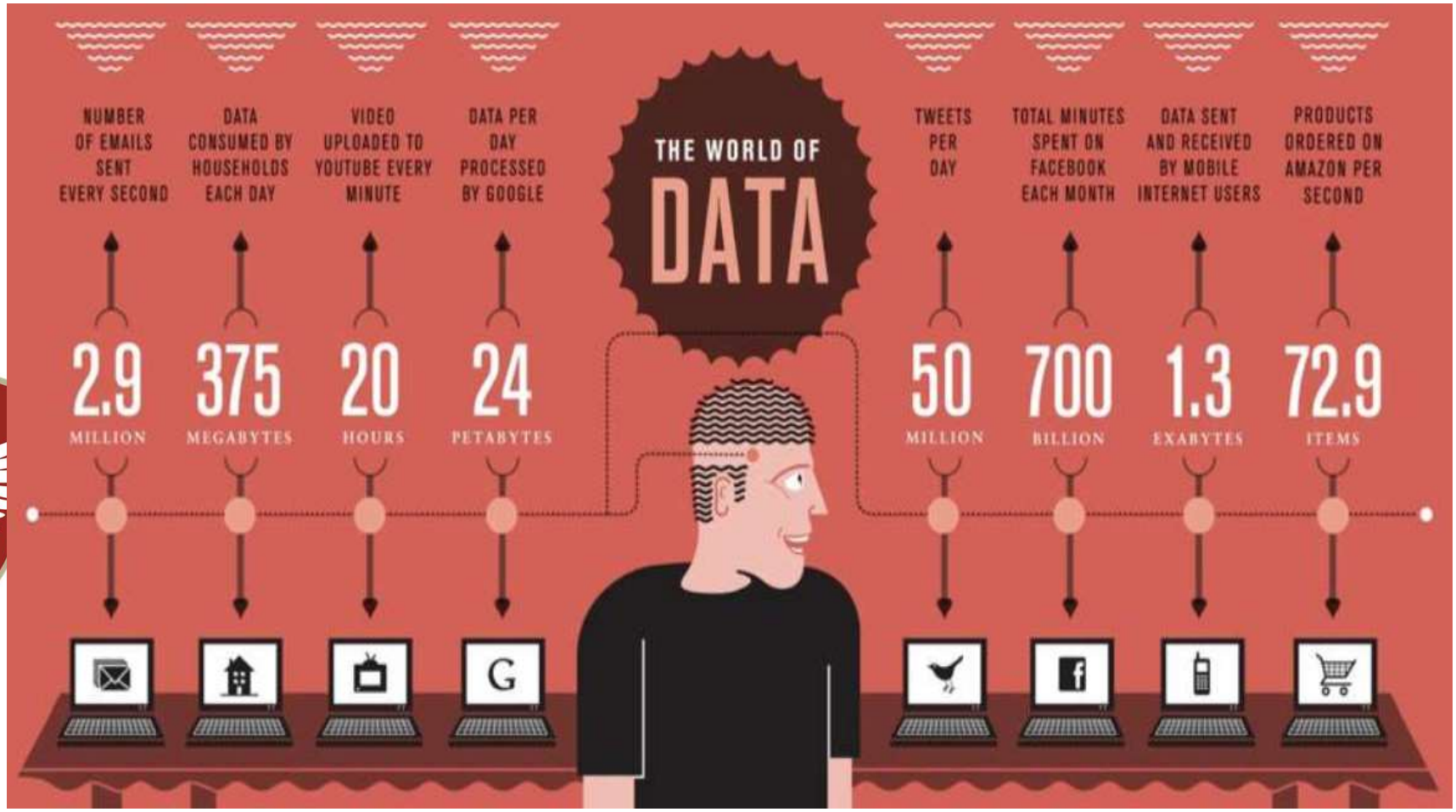
CENTRAL UNIVERSITY

FAITH • INTEGRITY • EXCELLENCE

What is Big Data?

- Massive sets of unstructured/semi-structured data from Web traffic, social media, sensors, etc
  - *Petabytes, exabytes of data*
    - Volumes too great for typical DBMS
  - *Information from multiple internal and external sources:*
    - Transactions
    - Social media
    - Enterprise content
    - Sensors
    - Mobile devices

What is Big Data?



- 204 million emails sent
- 61,000 hours of music listened to on Pandora
- 20 million photo views
- 100,000 tweets
- 6 million views and 277,000 Facebook Logins
- 2+ million Google searches
- 3 million uploads on Flickr

What is Big Data? continued

- Companies leverage data to adapt products and services to:
  - Meet customer needs
  - Optimize operations
  - Optimize infrastructure
  - Find new sources of revenue
  - Can reveal more patterns and anomalies
- IBM estimates that by 2015, 4.4 million jobs will be created globally to support big data
  - 1.9 million of these jobs will be in the United States

DATA *VERSES* BIG DATA

Data

- *Data, the plural form of **Datum** means raw facts.*
- *Unorganized and unmeaningful*
- *They come in bits (1s and 0s)*

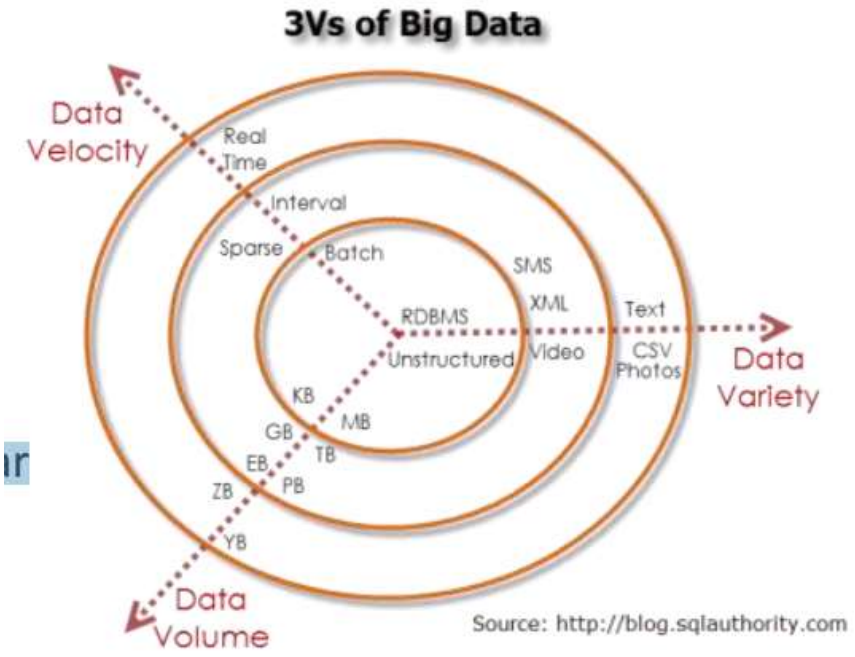
Big Data

- **Big data is just with:**
  - *More volume*
  - *Faster data generation (velocity)*
  - *Multiple data format (variety)*

DATA *VERSES* BIG DATA

Big data is just data with:

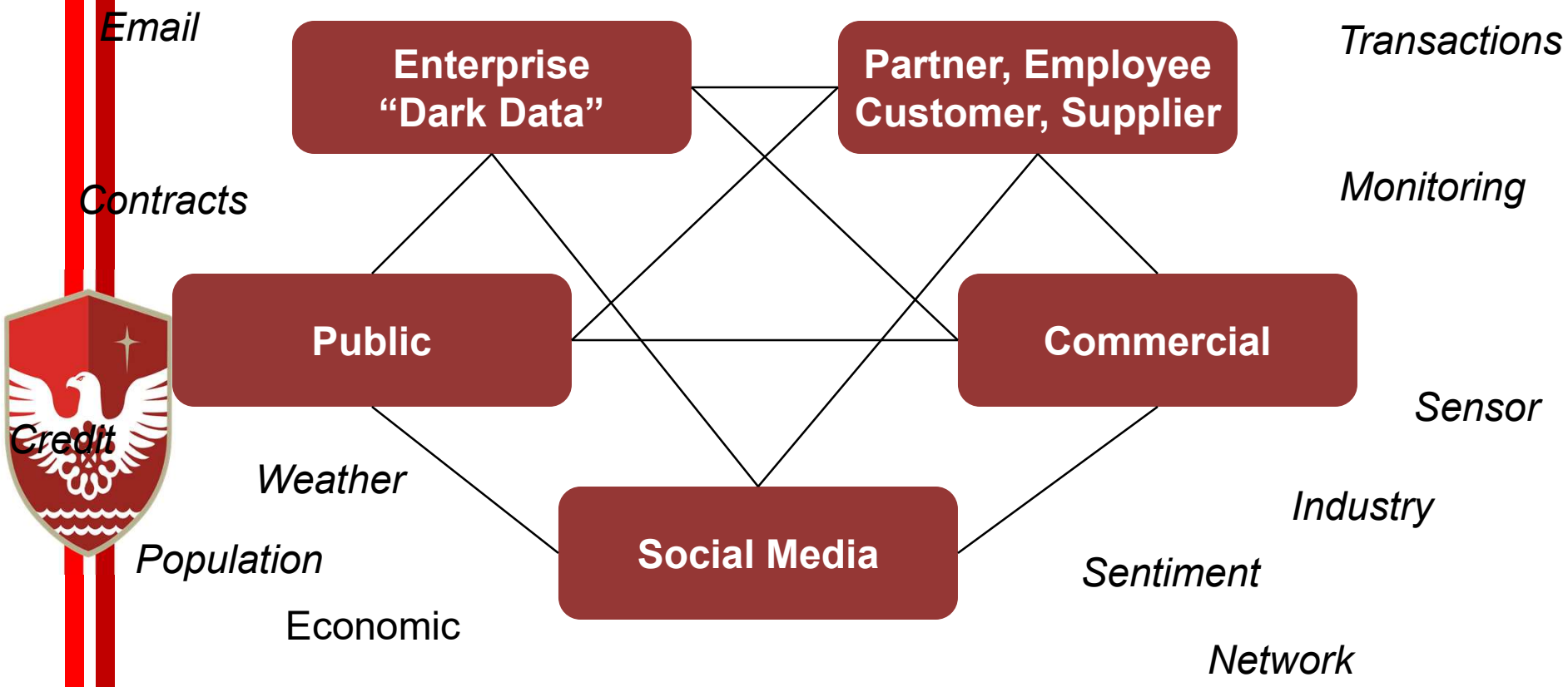
- *More volume*
- *Faster data generation (velocity)*
- *Multiple data format (variety)*
- *World's data volume to grow 40% per year*
- *& 50 times by 2020 [1]*
- *Data coming from various human & machine*
- *activity*



Data storage measurement units

Unit	Value	Size
bit (b)	0 or 1	1/8 of a byte
byte (B)	8 bits	1 byte
kilobyte (KB)	1000^1bytes	1,000 bytes
megabyte (MB)	1000^2bytes	1,000,000 bytes
gigabyte (GB)	1000^3bytes	1,000,000,000 bytes
terabyte (TB)	1000^4bytes	1,000,000,000,000 bytes
petabyte (PB)	1000^5bytes	1,000,000,000,000,000 bytes
exabyte (EB)	1000^6bytes	1,000,000,000,000,000,000 bytes
zettabyte (ZB)	1000^7bytes	1,000,000,000,000,000,000,000 bytes
yottabyte (YB)	1000^8bytes	1,000,000,000,000,000,000,000,000 bytes

Where does Big Data come from?



TYPES OF DATA

- When collecting or gathering data we collect data from individuals cases on particular variables.
- A ***variable*** is a unit of data collection whose value can vary.
- Variables can be defined into *types* according to the level of mathematical scaling that can be carried out on the data.
- There are four types of data or levels of measurement:
  - Categorical (Nominal)
  - Ordinal
  - Interval
  - Ratio

CATEGORICAL (NOMINAL) DATA

- **Nominal or categorical** data is data that comprises of categories that *cannot* be rank-ordered – each category is just different.
- The categories available **cannot be placed in any order** and no judgement can be made about the relative size or distance from one category to another.
- Categories bear no quantitative relationship to one another. Examples:
  - *customer's location (America, Europe, Asia)*
  - *employee classification (manager, supervisor, associate)*
- This means **"No mathematical operations can be performed on the data relative to each other"**.
- Therefore, nominal data reflect **qualitative differences** rather than **quantitative ones**.

CATEGORICAL (NOMINAL) DATA

- Examples:

What is your gender?
(please tick)

Male	☐
Female	☐

Did you enjoy the film?
(please tick)

Yes	☐
No	☐

- Systems for measuring nominal data must ensure that each category is **mutually exclusive** and the system of measurement needs to be **exhaustive.**

  - ***Mutually exclusive****: each observation (person, case, score) cannot fall into more than one category.*

  - ***Exhaustive:*** *the system of categories system should have enough categories for all the observations*

ORDINAL DATA

- Example:

How satisfied are you with the level of service you have received? *(please tick)*	
Very satisfied	
Somewhat satisfied	
Neutral	
Somewhat dissatisfied	
Very dissatisfied	

- Ordinal data is data that **comprises of categories that *<u>can</u>* be ranked in an orderly manner.**

Similarly with nominal data the distance between each category cannot be calculated but the **categories can be ranked above or below each other.**

- *No fixed units of measurement*
- *Examples: - college football rankings*

- survey responses (poor, average, good, very good, excellent)

INTERVAL DATA

- Ordinal data but with constant differences between observations
- Ratios are not meaningful
- Examples:
  - ***Time*** *– moves along a continuous measure or seconds, minutes and so on and is without a zero point of time.*
  - ***Temperature*** *– moves along a continuous measure of degrees and is without a true zero.*
  - ***SAT scores***

RATIO DATA

- Ratio data measured on a ***continuous*** **scale** and ***does*** **have a natural zero point.**
- Ratios are meaningful

Examples:

- *Monthly sales*
- *Weight*
- *Height*
- *Age*

Data for Business Analytics

(continued)

Classifying Data Elements in a Purchasing Database

	A	B	C	D	E	F	G	H	I	J
1	**Purchase Orders**									
2										
3	**Supplier**	**Order No**	**Item No.**	**Item Description**	**Item Cost**	**Quantity**	**Cost per order**	**A/P Terms (Months**	**Order Date**	**Arrival Date**
4	Spacetime Technologies	A0111	6489	O-Ring	$ 3.00	900	$ 2,700.00	25	10/10/11	10/18/11
5	Steelpin Inc.	A0115	5319	Shielded Cable/ft.	$ 1.10	17,500	$ 19,250.00	30	08/20/11	08/31/11
6	Steelpin Inc.	A0123	4312	Bolt-nut package	$ 3.75	4,250	$ 15,937.50	30	08/25/11	09/01/11
7	Steelpin Inc.	A0204	5319	Shielded Cable/ft.	$ 1.10	16,500	$ 18,150.00	30	09/15/11	10/05/11
8	Steelpin Inc.	A0205	5677	Side Panel	$195.00	120	$ 23,400.00	30	11/02/11	11/13/11
9	Steelpin Inc.	A0207	4312	Bolt-nut package	$ 3.75	4,200	$ 15,750.00	30	09/01/11	09/10/11
10	Alum Sheeting	A0223	4224	Bolt-nut package	$ 3.95	4,500	$ 17,775.00	30	10/15/11	10/20/11
11	Alum Sheeting	A0433	5417	Control Panel	$255.00	500	$ 127,500.00	30	10/20/11	10/27/11
12	Alum Sheeting	A0443	1243	Airframe fasteners	$ 4.25	10,000	$ 42,500.00	30	08/08/11	08/14/11
13	Alum Sheeting	A0446	5417	Control Panel	$255.00	406	$ 103,530.00	30	09/01/11	09/10/11
14	Spacetime Technologies	A0533	9752	Gasket	$ 4.05	1,500	$ 6,075.00	25	09/20/11	09/25/11
15	Spacetime Technologies	A0555	6489	O-Ring	$ 3.00	1,100	$ 3,300.00	25	10/05/11	10/10/11

Categorical Categorical Categorical Categorical Ratio Ratio Ratio Ratio Interval Interval

Figure 1.2

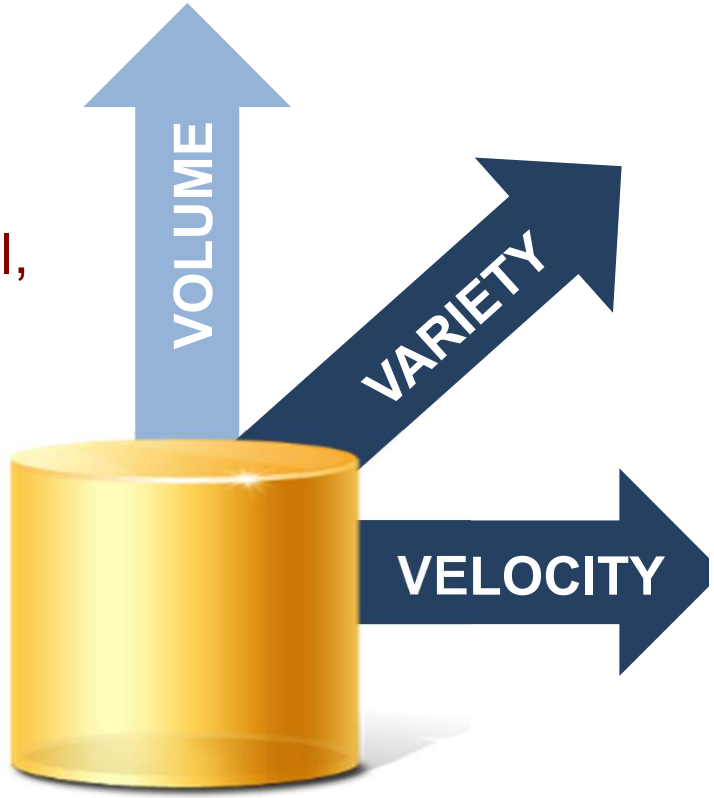
If there was field (column) for **Supplier Rating** (*Excellent, Good, Acceptable, Bad*), that data would be classified as **Ordinal**

BIG DATA CHARACTERISTICS

Growing quantity of data
e.g. social media, behavioral, video

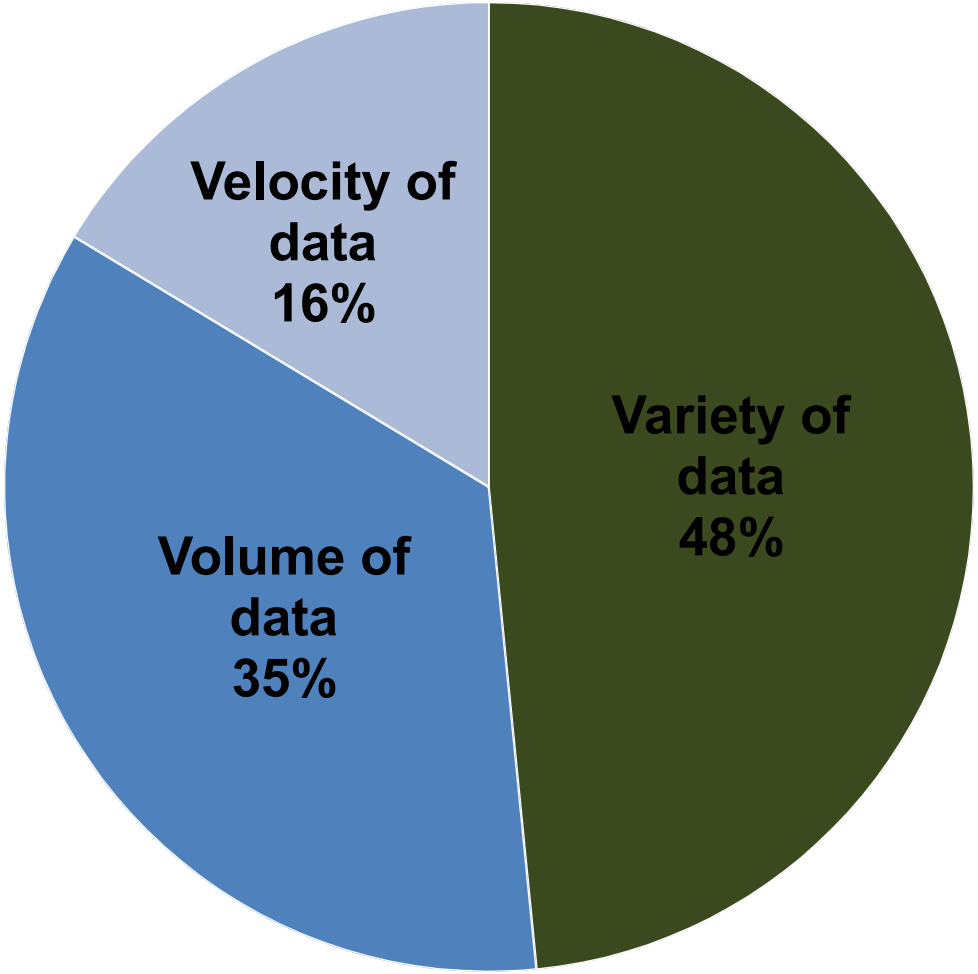
Quickening speed of data
e.g. smart meters, process monitoring

Increase in types of data
e.g. app data, unstructured data



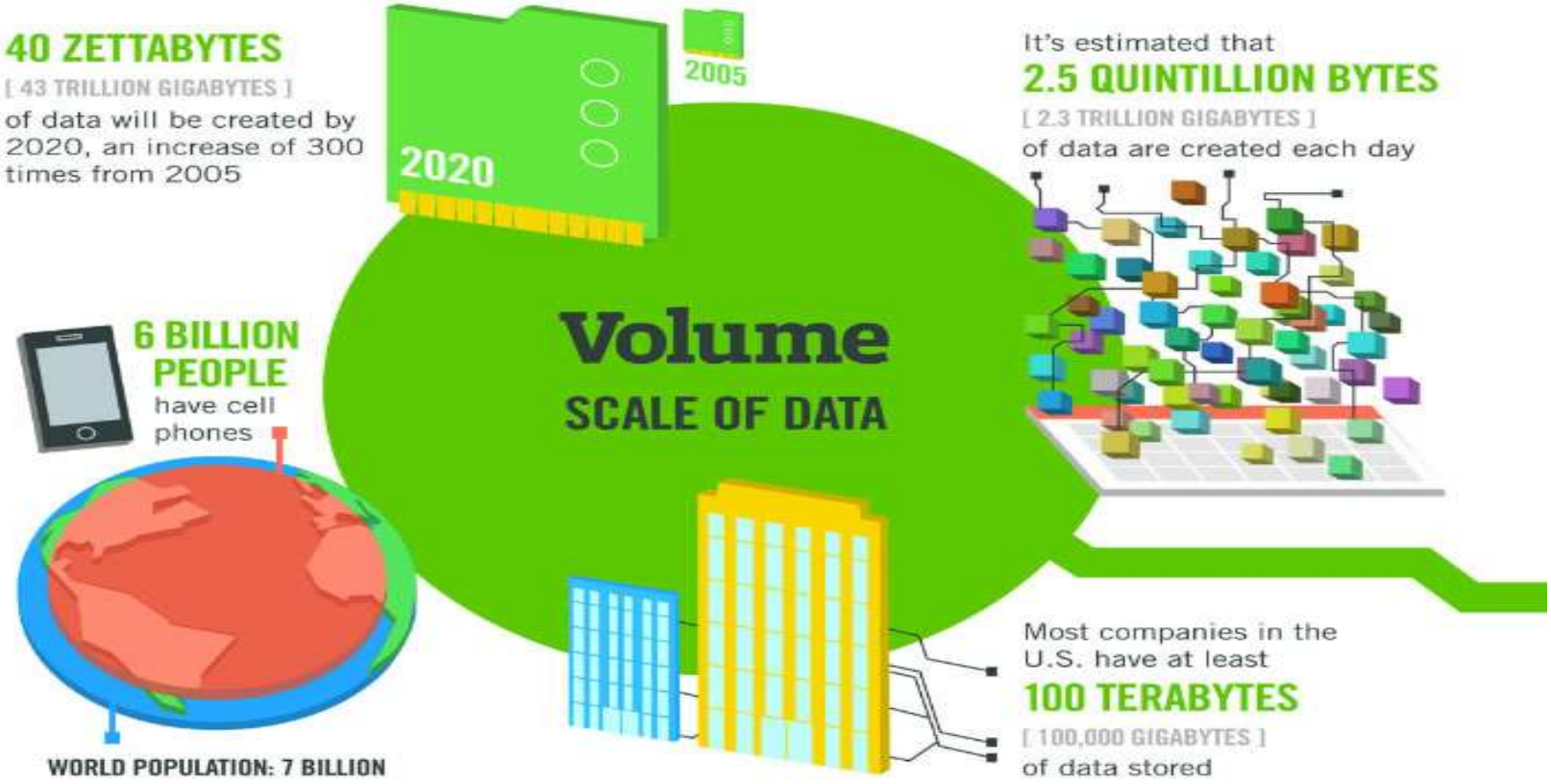
Gartner, Feb 2001

Which Big Data characteristic is the biggest issue for your organization?



Source: Getting Value from Big Data, Gartner Webinar, May 2012

Volume



- Volume
  - Petabytes, exabytes of data
  - Volumes too great for typical DBMS

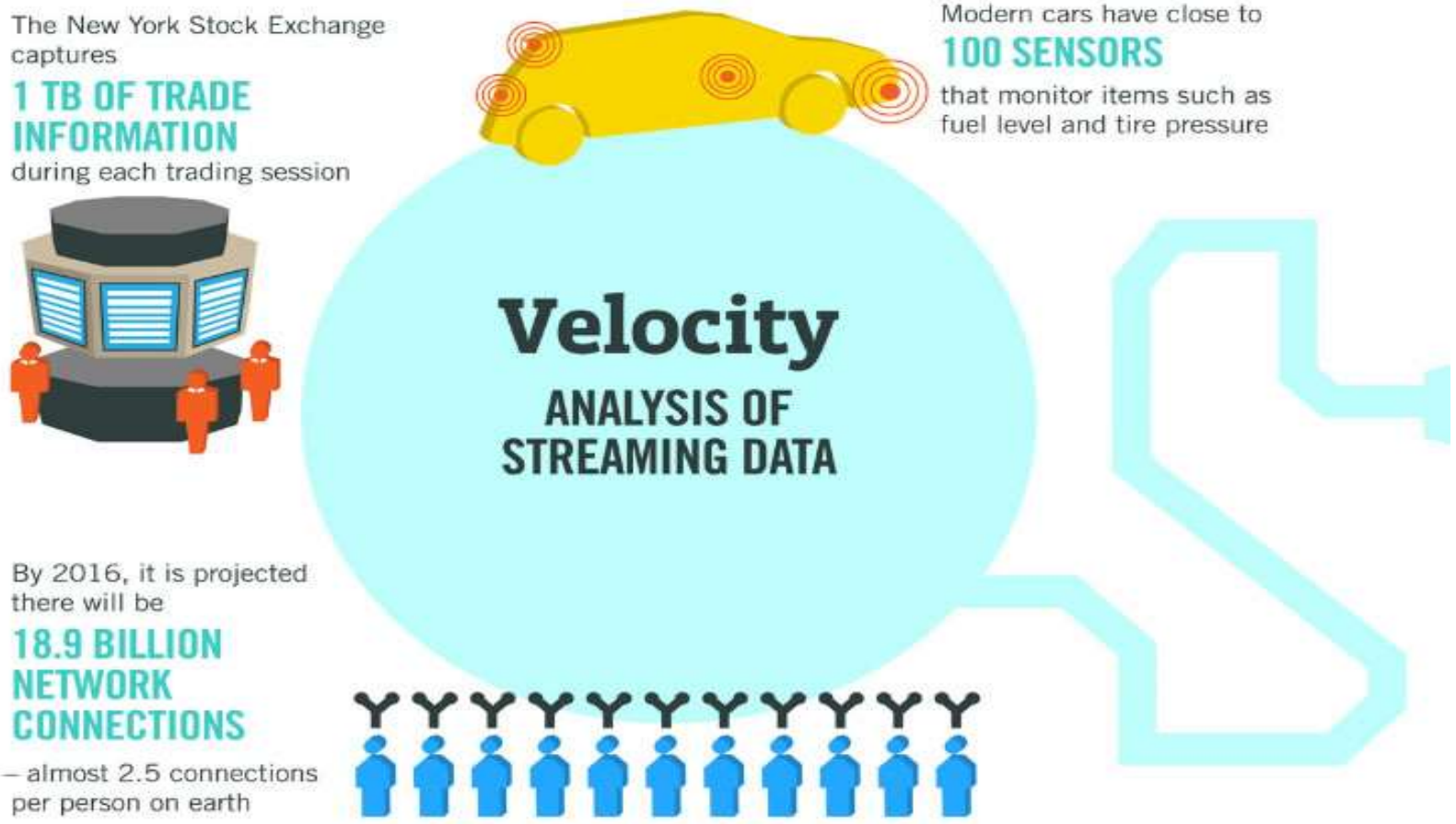
Volume - Bytes Defined

	Managerial Definition	Exact Amount	To Put It in Perspective
1 Terabyte (TB)	One trillion bytes	2^{40} bytes	Printed collection of the Library of Congress = 20 TB
1 Petabyte (PB)	One quadrillion bytes	2^{50} bytes	eBay data warehouse (2010) = 10 PB eBay will increase this 2.5 times by 2011 Teradata > 10 PB
1 Exabyte (EB)	One quintillion bytes	2^{60} bytes	
1 Zettabyte (ZB)	One sextillion bytes	2^{70} bytes	Amount of data consumed by U.S. households in 2008 = 3.6 ZB

Megabyte: 2^{20} bytes or, loosely, one million bytes

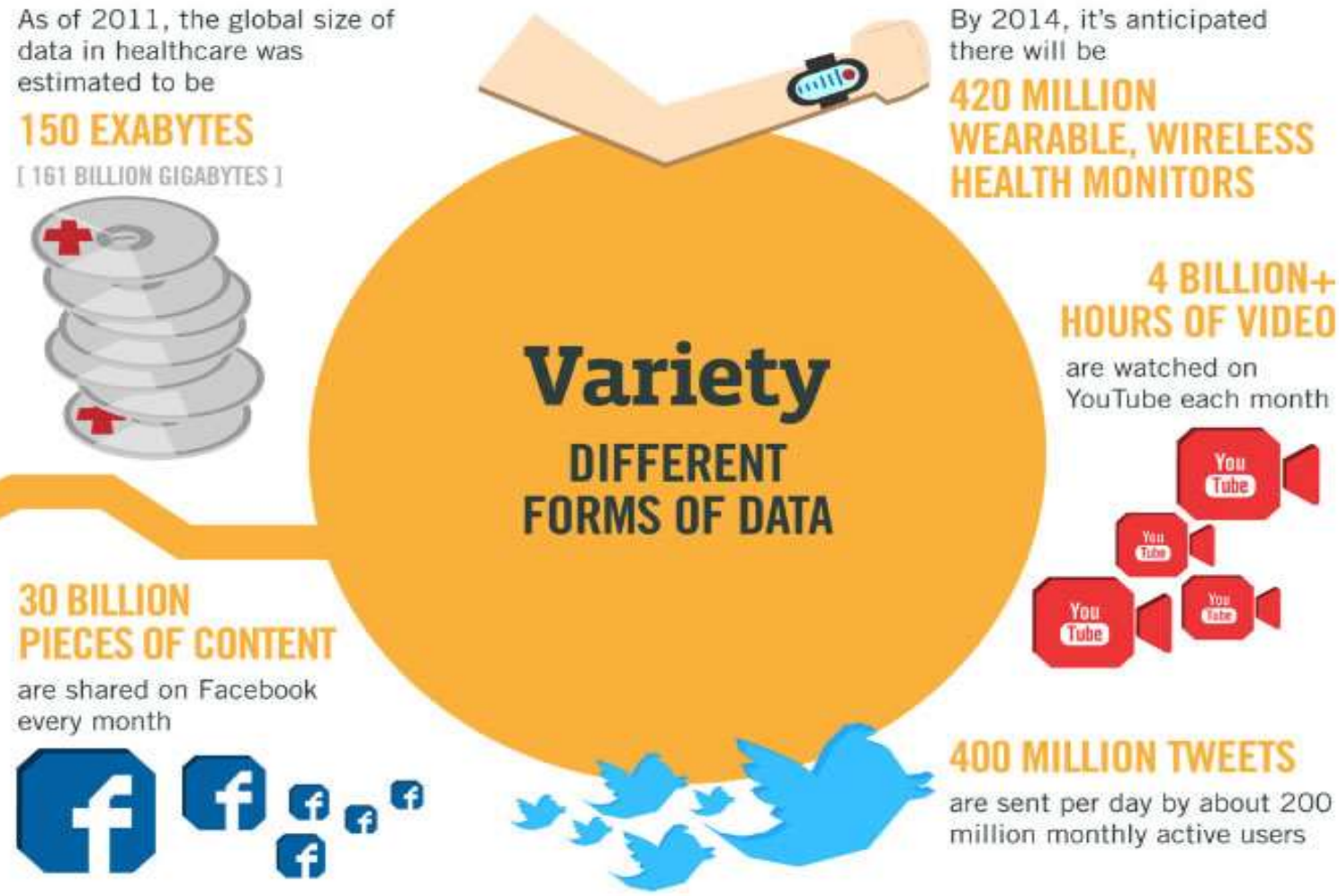
Gigabyte: 2^{30} bytes or, loosely one billion bytes

Velocity



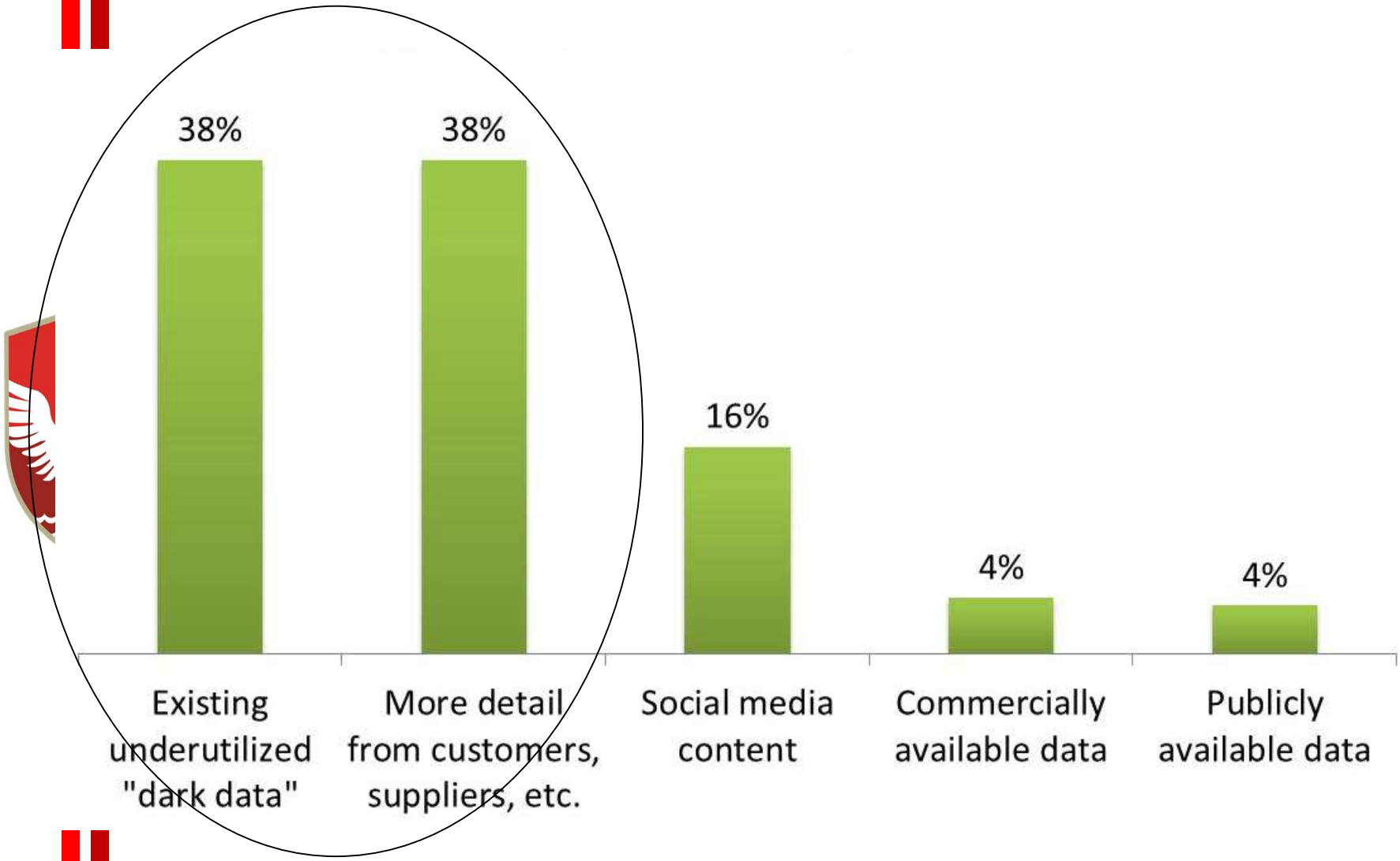
- Velocity
  - Massive amount of streaming data

Variety



- Variety
  - Massive sets of unstructured/semi-structured data from Web traffic, social media, sensors, and so on

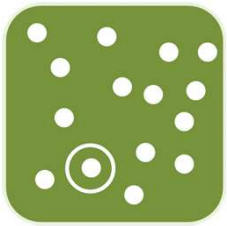
Which source of data represents the most immediate opportunity?



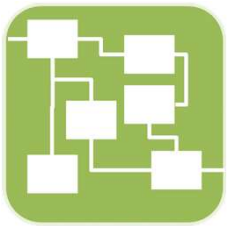
Source: Getting Value from Big Data, Gartner Webinar, May 2012

BIG DATA OPPORTUNITIES

Making better informed decisions
e.g. strategies, recommendations

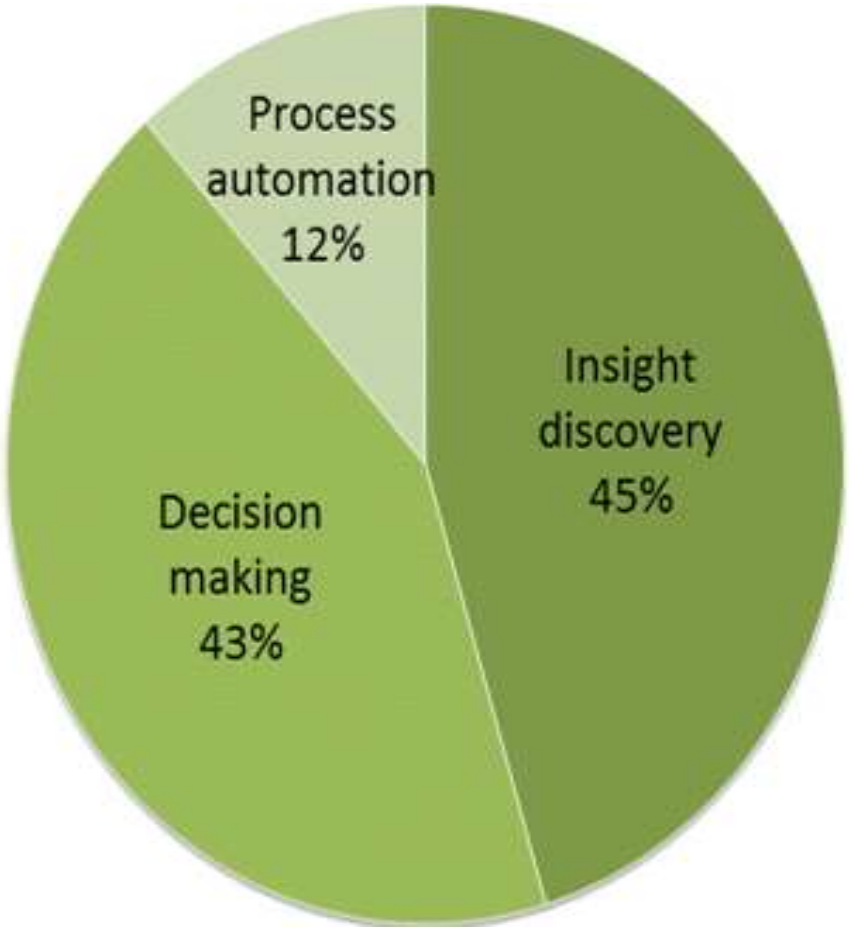


Discovering hidden insights
e.g. anomalies forensics, patterns, trends



Automating business processes
e.g. complex events, translation

Which is the biggest opportunity for Big Data in your organization?



Through 2017:

- 85% of Fortune 500 organizations will be unable to exploit big data for competitive advantage.
- Business analytics needs will drive 70% of investments in the expansion and modernization of information infrastructure.

Source: Getting Value from Big Data, Gartner Webinar, May 2012

IDENTIFYING INSURANCE FRAUD

- **Opportunity**
  - *Save and make money by reducing fraudulent auto insurance claims*
- **Data & Analytics**
  - *Predictive analytics against years of historical claims and coverage data*
  - *Text mining adjuster reports for hidden clues, e.g. missing facts, inconsistencies, changed stories*
- **Results**
  - *Improved success rate in pursuing fraudulent claims from 50% to 88%; reduced fraudulent claim investigation time by 95%*
  - *Marketing to individuals with low propensity for fraud*

Quality Improvement

- Opportunity
  - *Move from manual to automated inspection of burger bun production to ensure and improve quality*
- Data & Analytics
  - *Photo-analyze over 1000 buns-per-minute for color, shape and seed distribution*
  - *Continually adjust ovens and process automatically*
- Result
  - *Eliminate 1000s of pounds of wasted product per year; speed production; save energy; Reduce manual labor costs*

Is the company using all of its "senses" to observe, measure and optimize business processes?

Improving Corporate Image

- Opportunity
  - *Improve reputation, brand and buzz by tapping social media*
- Data & Analytics
  - *Continually scanning twitterverse for mentions of their business*
  - *Integrating tweeters with their robust customer management system*
- Results
  - *Saw tweet from a top customer lamenting late flight—no time to dine at Morton's*
  - *Tuxedo-clad waiter waiting for him when he landed with a bag containing his favorite steak, prepared the way he normally likes it with all the fixin's*

MORTON'S
THE STEAKHOUSE

How can the company listen, analyze and respond in real-time?

MODULE 3

Business Analytics

check it out in presentation 3

CENTRAL UNIVERSITY

FAITH • INTEGRITY • EXCELLENCE