Agenda

- **All kinds of data**
  - Large volumes
  - Valuable insight, but difficult to extract
  - May be extremely time sensitive
- **Big Data is a Hot Topic Because Technology Makes it Possible to Analyze ALL Available Data**

*"**Big data** technologies describe a new generation of technologies and architectures, designed to **economically extract value** from very large **volumes** of a wide **variety** of data, by enabling high **velocity** capture, discovery and/or analysis."*

Source: Matt Eastwood, IDC

Data Is The New Oil

- **In its raw form, oil has little value.**
  - Once **processed and refined**, it helps power the world.

Forbes.com

"*big data has arrived at Seton Health Care Family, fortunately accompanied by an analytics tool that will help deal with the complexity of more than two million patient contacts a year…*"

FT FINANCIAL TIMES
World business newspaper

"*Increasingly, businesses are applying analytics to social media such as Facebook and Twitter, as well as to product review websites, to try to "understand where customers are, what makes them tick and what they want", says Deepak Advani, who heads IBM's predictive analytics group.*"

The New York Times

"*At the World Economic Forum last month in Davos, Switzerland, Big Data was a marquee topic. A report by the forum, "big data, Big Impact," declared data a new class of economic asset, like currency or gold.*"

CNBC

"*Data is the new oil.*"
Clive Humby

THE WALL STREET JOURNAL.

"*Companies are being inundated with data—from information on customer-buying habits to supply-chain efficiency. But many managers struggle to make sense of the numbers.*"
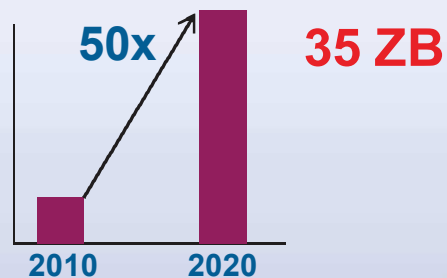
Forbes.com

"*…now Watson is being put to work digesting millions of pages of research, incorporating the best clinical practices and monitoring the outcomes to assist physicians in treating cancer patients.*"
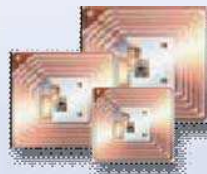
Characteristics of Big Data

- $V^4 =$ Volume Velocity Variety Veracity

Cost efficiently processing the growing Volume

50x   **35 ZB**

2010   2020

Responding to the increasing Velocity

30 Billion RFID sensors and counting

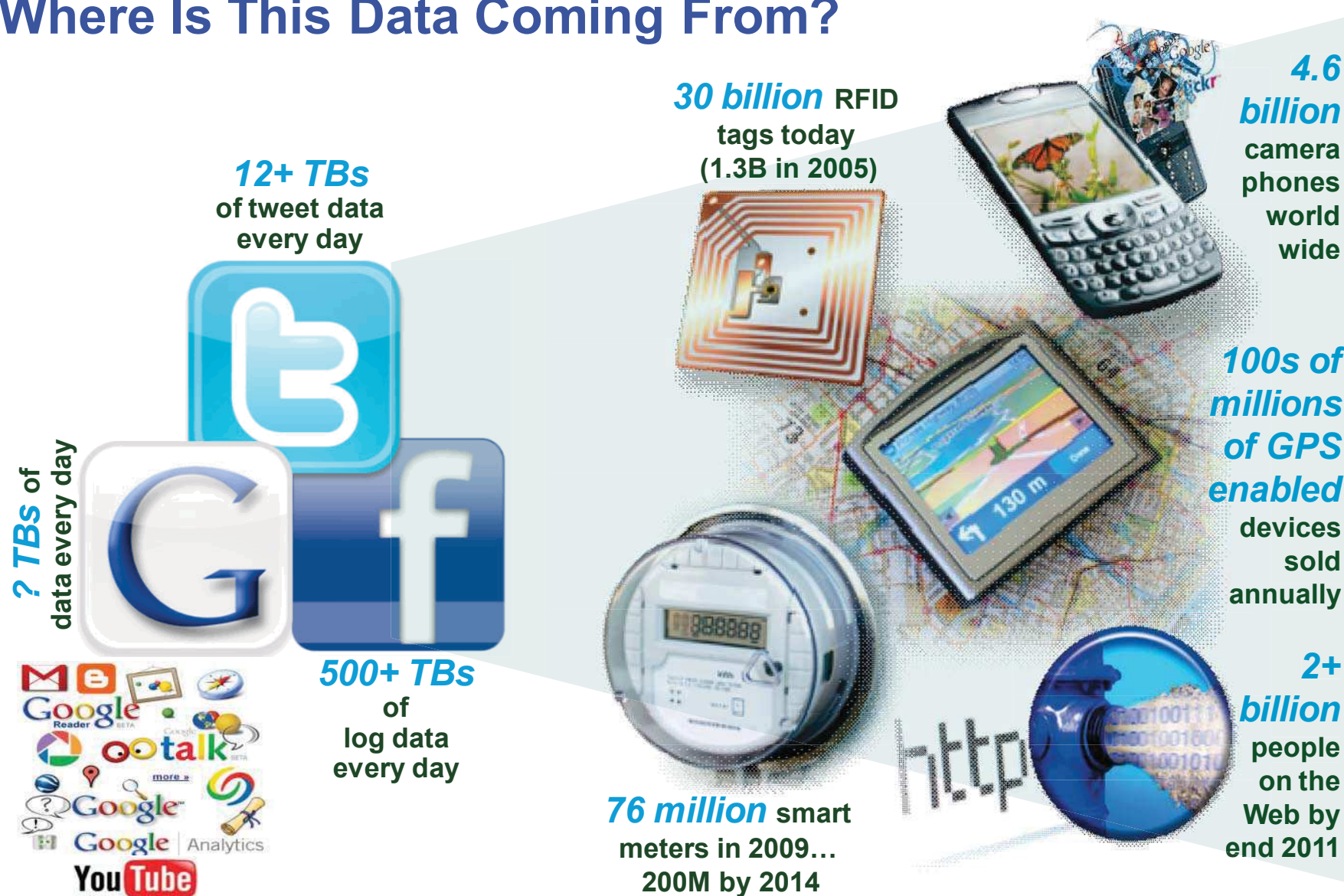**Collectively analyzing the broadening Variety**

80% of the worlds data is unstructured

Establishing the Veracity of big data sources

1 in 3 business leaders don't trust the information they use to make decisions

Where Is This Data Coming From?

12+ TBs
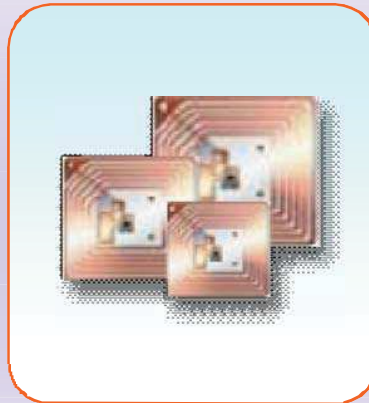of tweet data
every day

? TBs of
data every day

500+ TBs
of
log data
every day

30 billion RFID
tags today
(1.3B in 2005)

76 million smart
meters in 2009…
200M by 2014

*4.6
billion*
camera
phones
world
wide

*100s of
millions
of GPS
enabled* devices
sold
annually

*2+
billion*
people
on the
Web by
end 2011

Realize New Opportunities

- **Think Beyond Traditional Sources of Data**

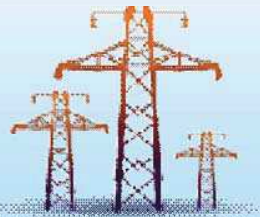| Transactional and Application Data	Machine Data	Social Data	Enterprise Content
▪ Volume ▪ Structured ▪ Throughput	▪ Velocity ▪ Semi-structured ▪ Ingestion	▪ Variety ▪ Highly unstructured ▪ Veracity	▪ Variety ▪ Highly unstructured ▪ Volume

Imagine the Possibilities of Harnessing *Your* Data Resources

- **Big data challenges exist in every organization today**

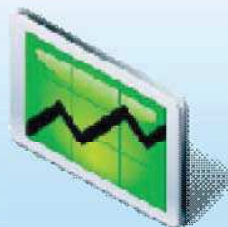|		
Government cuts acoustic analysis from hours to 70 Milliseconds	**Utility avoids power failures by analyzing 10 PB of data in minutes**	**Hospital analyses streaming vitals to detect illness 24 hours earlier**
Retailer reduces time to run queries by 80% to optimize inventory	**Stock Exchange cuts queries from 26 hours to 2 minutes on 2 PB**	**Telco analyses streaming network data to reduce hardware costs by 90%**

Leveraging Big Data Requires Multiple Platform Capabilities

Understand and navigate federated big data sources		Federated Discovery and Navigation
Manage & store huge volume of any data		Hadoop File System MapReduce
Structure and control data		Data Warehousing
Manage streaming data		Stream Computing
Analyze unstructured data		Text Analytics Engine
Integrate and govern all data sources		Integration, Data Quality, Security, Lifecycle Management, MDM

Use Cases for a Big Data Platform

Know Everything about your Customer
- Social media customer sentiment analysis
- Promotion optimization
- Segmentation
- Customer profitability
- Click-stream analysis
- CDR processing
- Multi-channel interaction analysis
- Loyalty program analytics
- Churn prediction

Innovate New Products at Speed and Scale
- Social Media - Product/brand Sentiment analysis
- Brand strategy
- Market analysis
- RFID tracking & analysis
- Transaction analysis to create insight-based product/service offerings

Run Zero Latency Operations
- Smart Grid/meter management
- Distribution load forecasting
- Sales reporting
- Inventory & merchandising optimization
- Options trading
- ICU patient monitoring
- Disease surveillance
- Transportation network optimization
- Store performance
- Environmental analysis
- Experimental research

Instant Awareness of Risk and Fraud
- Multimodal surveillance
- Cyber security
- Fraud modeling & detection
- Risk modeling & management
- Regulatory reporting

Exploit Instrumented Assets
- Network analytics
- Asset management and predictive issue resolution
- Website analytics
- IT log analysis