

# Big Data, Integration & Governance

# Harnessing the Power of Big Data

Stewart Hanna – Worldwide Big Data 15 – 20 November | Canberra – Melbourne - Sydney



# "Data is the new Oil"

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

# **Forbes**

"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..."

> "Companies are being inundated with data—from information on customerbuying habits to supply-chain efficiency. But many managers struggle to make sense of the numbers."



"Data is the new oil." Clive Humby

#### 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.



"...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." 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

#### Los Angeles Times

The Oscar Senti-meter — a tool developed by the L.A. Times, IBM and the USC Annenberg Innovation Lab — analyzes opinions about the Academy Awards race shared in millions of public messages on Twitter."



# An Explosion of Data





# By 2016, annual Internet traffic will reach **1.3 Zettabytes**



#### Google processes > 24 Petabytes of data in a single day



# Facebook processes 500+ Terabytes of data daily



Twitter processes **12 Terabytes** of data daily



**150 Exabytes** global size of "Big Data" in Healthcare, growing between 1.2 and 2.4 EX / year



Hadron Collider at CERN generates **40 Terabytes** of usable data / day



AT&T transfers about **30 Petabytes** of data through its network daily



For every session, NY Stock Exchange captures **1** Terabyte of trade information

# **Applications for Big Data Analytics**





In Order to Realize New Opportunities, You Need to Think Beyond Traditional Sources of Data

# Transactional and Application Data



- Volume
- Structured
- Throughput

## **Machine Data**



- Velocity
- Semi-structured
- Ingestion



**Social Data** 

- Variety
- Highly unstructured
- Veracity

#### Enterprise Content



- Variety
- Highly unstructured
- Volume



# Getting to the value requires exploiting the unique characteristics of big data





\* Truthfulness, accuracy or precision, correctness

# Why Didn't We Use All of the Big Data Before?







# Big Data represents a New Approach

**Leverage** analytics as a game-changer

**Embrace** big data to maximize your business outcomes

Get **inspired** and take action





Organizations drive transformation by starting with one of these four high-value initiatives





# IBM Big Data Platform...



New analytic applications drive requirements for a big data platform



- Integrate and manage the full variety, velocity and volume of data
- Apply advanced analytics to information in its native form
- Visualize all available data for ad-hoc analysis
- Development environment for building new analytic applications
- Workload optimization and scheduling
- Security and Governance





Reduced Complexity, Faster Time to Value

# Customizable, pre-built business logic, data processing, and visualization

#### Social Data Analytics



- Sentiment analytics
- Intent to purchase
- Gain competitive intelligence

## Machine Data Analytics



- Log processing
- Pattern analysis
- Real time visibility into operations

## Telco Event Analytics



- CDR analytics
- Customer event analytics
- Real-time processing

University of Ontario Institute of Technology (UOIT) uses big data to improve quality of care for neonatal babies

Need

- Performing real-time analytics using physiological data from neonatal babies
- Continuously correlates data from medical monitors to detect subtle changes and alert hospital staff sooner
- Early warning gives caregivers the ability to proactively deal with complications

#### Benefits:

- Detecting life threatening conditions 24 hours sooner than symptoms exhibited
- Lower morbidity and improved patient care

Major U.S. Global Institutional Investment Firm to Improve Risk Management of Client Portfolios

- Repository of all holdings and benchmarks information in BigInsights Distribute File System for fault-tolerance on commodity hardware
- Identify client investment exposures based on traded security, issuer, country, currency, etc.
- Compare exposures to benchmarks for compliance and active bet exposure
- Improve current manual process



Marine Institute Ireland, monitors buoy sensor data to detect floods sooner Need

- To make better use of the environmental data it collected
- Compiling this data to perform trend analyses or draw conclusions, however, was particularly complicated

## Benefits

- Uses maritime data to increase the accuracy of flood predictions and support automated alarm systems, improving public safety
- Fosters new businesses and overall growth by providing seafood, shipping and monitoring operations with streaming maritime data



Major US Oil and Gas Exploration and Production Company

- Ice floe interrupts daily operations and inhibits year round operation
- Using IBMs Big Data Platform to track, forecast and provide early warning of ice floe in the Arctic Circle
- Analyzing in realtime Satellite imagery, Wind & Sonar data and MetOcean weather forecast and other sources yet to be engineered
- Achieving savings in the 100's of Millions and Billions



# TerraEchos for Low Latency Surveillance Data Analysis

# Capabilities Utilized:

## Stream Computing

- Deployed security surveillance system to detect, classify, locate, and track potential threats at highly sensitive national lab
- Stream computing collects and analyzes acoustic data from fiber-optic sensor arrays
- Analyzed acoustic data fed into TerraEchos intelligence platform for threat detection, classification, prediction & communication

## Significant benefits:

- Enables Terraechos solution to analyze and classify streaming acoustic data in real-time
- Provides lab & security staff with holistic view of potential threats & non-issues
- Enables a faster and more intelligent response to any threat

# 1111

# Major U.S. Wireless Telco Increases Revenue and Improves Customer Satisfaction

- Deployed real-time CDR analysis solution to handle exploding data volume growth and performance requirements
- Analyzes call, internet usage, and text records in real-time to identify and address poorly performing cells
- Significant Benefits:
  - Over 90% reduction in time to merge/load call record data
  - Over 90% reduction in storage
  - Increased network quality, improved customer satisfaction, reduced churn

## Retail Point-of-Sales, Marketing, and Inventory Analysis

- Enable better understanding of shoppers by analyzing *new data sources* such as social media together with transaction histories, web logs, etc
  - Luxury good analysis group of stores for stocking luxury brand
  - Internet vs Store purchase analysis
  - Gift card analysis
  - Sentiments on products and brands
  - Reconcile what business know about a customer's behavior
- Improve gross margin, inventory management, vendor negotiation, etc, by
  - Customer segmentation, cross-selling,
  - Market, product, price analysis
  - Targeted promotions, etc





# Big Data, Integration & Governance

15 – 20 November | Canberra – Melbourne - Sydney

