2020: The Future is Data

Michael Dell
Senior Vice President
Aerospace, Defence and Maritime
IHS

Kilobytes to Zetabytes in One Generation

- 1,024 Kilobytes = One Megabyte
  - 1,024 Megabytes = One Gigabyte
    - 1,024 Gigabytes = One Terabyte
      - 1,024 Terabytes = One Petabyte
        - 1,024 Petabytes = One Exabyte
          - 1,024 Exabyte = One Zetabyte

Kilobytes to Zetabytes in One Generation

What is Big Data
Volume – Data is being created at an unprecedented level
Variety – Many sources of information from different sources and formats
Velocity – Data is being produced at a massive speed, decision windows are shrinking

- 235 Terabytes of data collected by the U.S. Library of Congress by April 2011[1]

- 2.7 Zetabytes of data exist in the digital universe at the start of 2012[2]

- 40% projected growth in global data per year[3]

Projected Information Growth 2011-2021
Equivalent Multiples of Libraries of Congress

[1] - US Library of Congress
[2] - IDC Predictions 2012: Competing for 2020, Dec 2011
[3] - Manyika, et. al. McKinsey Global Institute, June 2011.

Social, Consumer, Business Data Overload

- **Social Media Data**
  - 5 billion mobile phones in use
  - 30 billion pieces of content shared on Facebook per month
  - 175 million tweets per day
  - You Tube uploads >24 hours of video per minute
- **Consumer Data**
  - Walmart already doing 1 million transactions per hour
  - 450 billion daily transactions (for money) on the internet by 2020
- **Business to Business Data**
  - A Boeing 737 generates as much data in one flight as collected by the US Library Congress from 1800-2011
  - IHS Maritime: 2% of our storage for first 250 years; 98% for the future

More Raw Data Today than at Any Previous Time in History

- "We create as much information in **two days now** as we did from the dawn of man through 2003." – Eric Schmidt, CEO Google

- "Information overload is a **headache** for individuals and a **huge challenge** for businesses. Companies are swimming, if not **drowning, in wave after wave of data** — from increasingly sophisticated computer tracking of shipments, sales, suppliers and customers, as well as e-mail, Web traffic and social-network comments." – NY Times, 23 April, 2011

All This is Making us Worry

Customer concerns about data silos, lack of skillsets to analyze, and lack of time predominate

Question: What are your organization's three biggest impediments to using big data for effective decision-making [Select up to three options]

Rank	Answer	Response
1	Too many "silos" - data is not pooled for the benefit of the entire organization	56%
2	Shortage of skilled people to analyze the data properly	51%
3	The time taken to analyze large data sets	44%
4	Unstructured content in big data is too difficult to interpret	42%
5	Big data is not viewed sufficiently strategically by senior management	35%
6	The high cost of storing and manipulating large data sets	17%
7	Big data sets are too complex to collect and store	15%
8	Something not on this list (please specify)	4%

But Analytical and Decision-Making Processes Remain the Same

The process hasn't changed….

| Physical World | Raw Data | Operable Data Sources | Analytical Process | Analytical Outcome | Action |

…but the size, pace and variety of the inputs has…

…the tools to manage the inputs has improved….

…as has the ability to automate data governance….

…the power of the analytics has massively improved…

…the outcome can be much more predictive….

…and the action can be automated….

So What is Possible?

"Big Data" Can:

- Acquire and integrate LOTS of data from LOTS of multi-varied sources in LOTS of different formats
- Analyze relationships / patterns QUICKLY in PARALLEL
- Use the analysis to target segments of users from the individual to well defined groups
- Make predictions from historical trends and current relationships – predictive analytics

"Big Data" cannot:

- Fully interpret the subjective… yet….ever?
- Understand and analyze that which it does not collect
  - E.g. In the heads of people; data that is not yet public
- Anticipate step changes
- Identify all implicit / explicit assumptions (hidden variables)
- TAKE DECISIONS

Here's How We Think About It

Predictive Analytics — Predictive modeling using statistical and machine learning techniques

Prescriptive Analytics — Recommend decisions using optimization, simulation, action oriented

Descriptive Analytics — Gain insight from historical data with reporting, scorecards, clustering etc.

Here's What We Do Today

Predictive Analytics

✓ Energy Costing and Valuation Analytics

✓ Predicting customer buying and revenue trends by industry

Prescriptive Analytics

✓ Supplier Strategic Sourcing Risk

✓ EHS&S Operational Risk & Preventive Maintenance

✓ Unconventional Well Real Time Analysis and Benchmarking

Descriptive Analytics

✓ Global Commodity Tracking

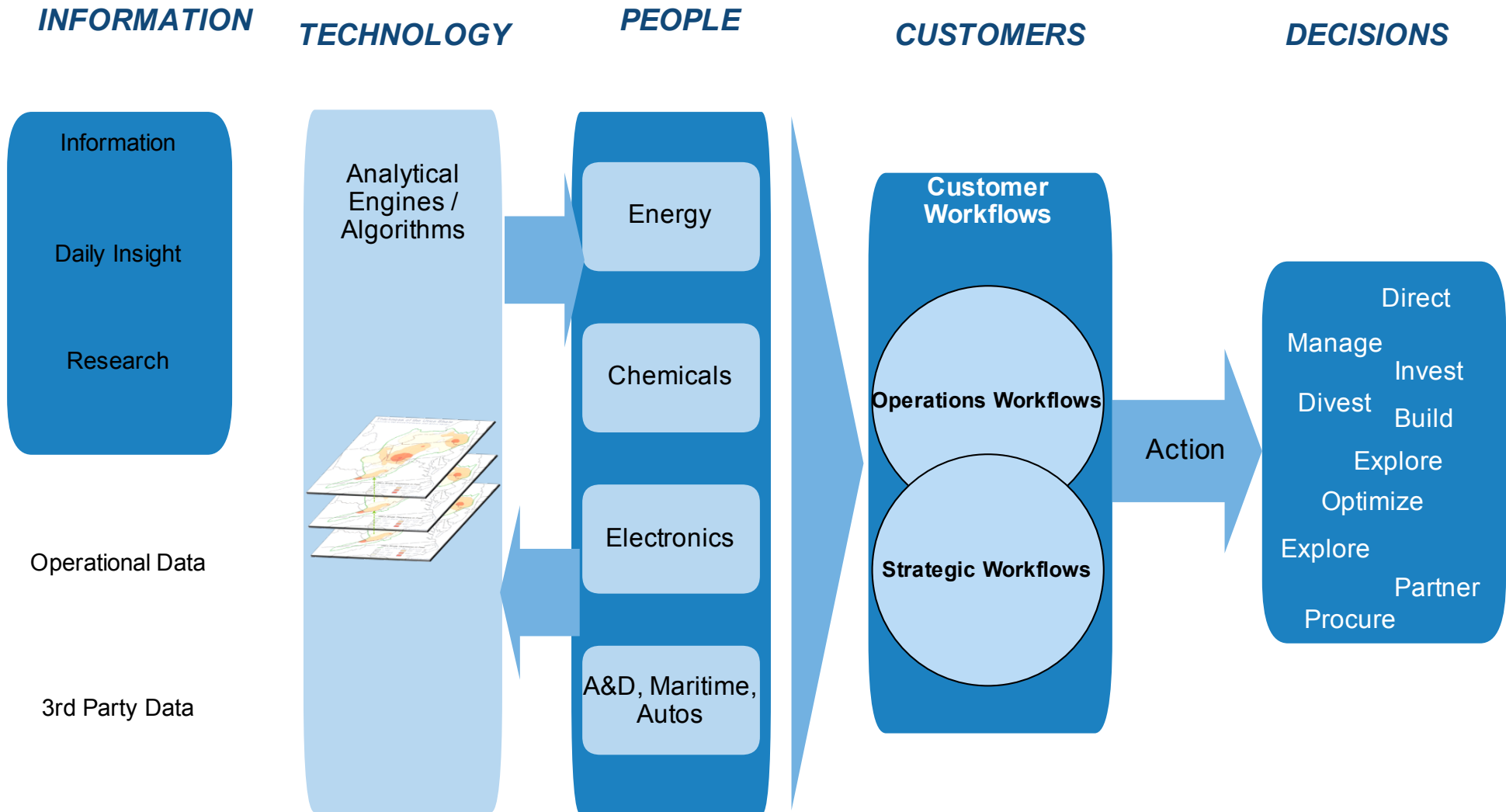✓ Autos OEM Industry and Company Performance Benchmarking

✓ Tracking Fracking

Starting with our existing maritime data….

… Here's Just One Place We're Going….

Deep real-time and historical analysis of the world's major cargo flows by ship and cargo type

… Building Upon our Information, Technology and People …

INFORMATION   **TECHNOLOGY**   **PEOPLE**   **CUSTOMERS**   **DECISIONS**

Information

Daily Insight

Research

Operational Data

3rd Party Data

Analytical Engines / Algorithms

Energy

Chemicals

Electronics

A&D, Maritime, Autos

Customer Workflows

Operations Workflows

Strategic Workflows

Action

Direct

Manage

Invest

Divest

Build

Explore

Optimize

Explore

Partner

Procure