



# The Promise of Industrial Big Data

---

## Big Data Real-Time Analytics

**Katherine Butler**  
**1<sup>st</sup> Annual Digital Economy Congress**  
**San Diego, CA**  
**Nov 14<sup>th</sup> – 15<sup>th</sup>, 2013**



imagination at work

© General Electric Company, 2013

# Individual vs. Ecosystem

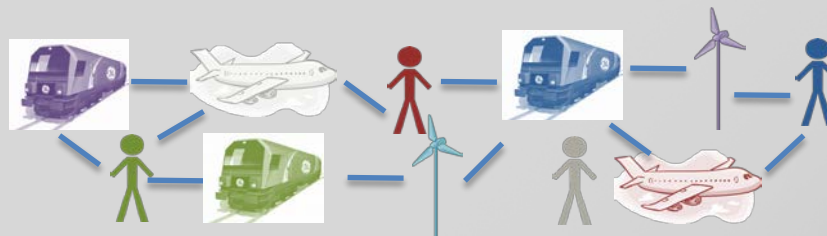
What Happened When **1B** People Became Connected?



- Entertainment is Digitized
- Social Marketing Emerged
- Communications Mobilized
- IT Architecture Virtualized
- Retail & Ad Transformed

**Consumer Internet**

What Happens When **50B** Machines & People Become Connected?



- Remote monitoring
- Predictive Analytics
- Virtualized OT
- Machine Learning & Automation
- Employee Productivity

**Industrial Internet**

50 Billion Smart Devices collect, store & use data while connecting with other devices



imagination at work

© General Electric Company, 2013

# A Brave New Digital World

**30,000** jet engines in the next 15 years will go into service

As many as **50 billion** machines connected by 2020

**500 gigabytes** of data from single wind turbine blade

**2 1/2 zettabyte** of data produced in a day vs. all of 2008

**40% projected growth** in global data generated per year vs. 5% growth in IT spending

**1.5 million** more data-savvy scientists needed to take full advantage of big data in the US

## It's All About Real Time Analytics



imagination at work



# Industrial Internet

***“The Industrial Internet represents an open, global ecosystem of highly intelligent machines that connect, communicate and cooperate with each other and people, leading to breakthrough levels of efficiency and productivity - a productivity revolution.”***

Customers' demand for productivity combined with technical trends (Mobile, Smart Devices, Cloud Computing etc.) are dramatically impacting how we connect and how we perceive and deliver value

Customers' focus is on optimizing assets and discovering meaningful insights from massive volumes of data

Expectation is for excellent, high-tech, real-time user experiences

The Industrial Internet is already here and growing fast



imagination at work

# Promise of Industrial Big Data

Big Data is driving a revolution in thinking and processes. Entire organizations and industries and society itself are being transformed.

Yet according to IDC, only 0.5 percent of data collected currently is being processed.

Why is the promise of Big Data so compelling?  
What's different?

- Sheer quantities of data being digitized (e.g., streaming or sensor generated data doesn't fit in traditional databases)
- More devices are connected to the Internet today than there are people to use them
- Devices are growing increasingly intelligent
- Real-Time and even Predictive Analytics are emerging as key requirement in multiple industries



imagination at work

# Promise of Industrial Big Data (cont.)

Benefits to industries and domains such as Medical Research, Supply Chain Management and Energy Conservation are incredibly compelling

Need certainty and clear rules and governance in order to realize such benefits

Industrial Data needs to move seamlessly across country borders

Following planes, trains and automobiles

Global industries need global data flow

Consumer privacy hugely important but cannot solely define the dialog

Industrial Big Data is available to mine now

But, Sensors not solely for machines; wearable sensors increasingly used (sports/leisure)

Collision between Big Data benefits and need for clear and certain privacy guidelines is on us



imagination at work

# Re-Imagine



## Outcomes that Matter

Increasing Freight Utilization  
Rail



## Power of 1

\$27B industry value by  
reducing system  
inefficiency

Predictive Maintenance  
Healthcare



\$63B industry value by  
reducing process  
inefficiency

Predictive Diagnostics  
Power



\$66B industry value with  
efficiency improvements in  
gas-fired power plant fleets



imagination at work

© General Electric Company, 2013

*\*Over 15 years, nominal US dollars*



# GE's Installed Wind Fleet

~21,000 WTGs, ~32GW+ installed ... 23 countries

Over 425 million MWhRs produced



24X7 support

Centers - Germany & NY – Hot Swap Capability

7,000+ turbines controlled remotely ... 98%+ availability

Fleet analysis operating rhythm with engineering-based fault resolution, continuously evolved

12,700+ turbines w/ data connections



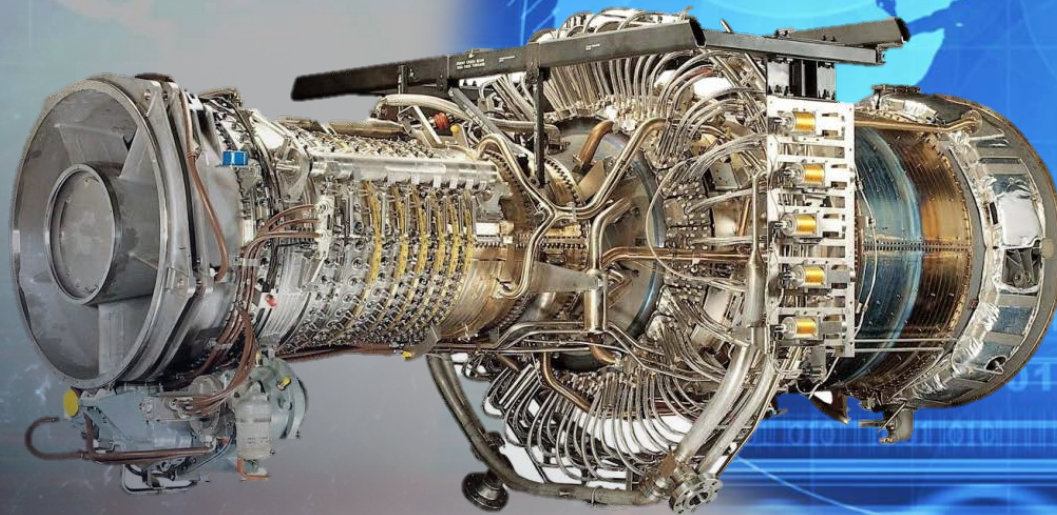
imagination at work

© General Electric Company, 2013



# Value of Big Data Analytics

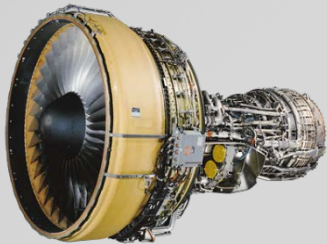
**1 gas turbine  
compressor blade  
monitoring  
potential: 500  
gigabytes per day**



imagination at work

© General Electric Company, 2013

# GE Intelligent Engines: 3<sup>rd</sup> Generation



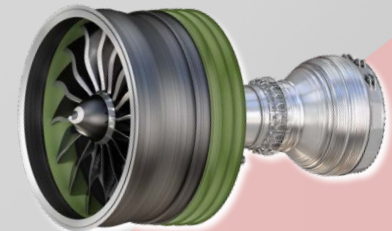
## Performance

- Few sensors, 100 parameters
- Flight snapshot



## Health

- 23 sensors, 1,000 parameters
- Continuous, full flight data
- Wi-fi connectivity



## Predictive

- +++ sensors and parameters
- Fault isolation
- Greater confidence

(per yr.) **Before 11GB**

**Today 1.4TB**

**GE9X 3-5TB**

**GENx generates ½ TB per flight / 5k data points analyzed per second**

# Big Data Legal Landscape

## Licensing

- Rights in original data and analyzed data

## Intellectual Property

- Tradeseecret

- Patent

- Copyright (compilation in US and EU96/9 Database Directive in EU)

## Competition Law/Antitrust

## Privacy and classification

## Data Protection

## Contract

- Liability for security breaches

- Liability for failures in analytics, particularly predictive (e.g., incomplete or inaccurate data, unpredictability of unreliable data such as weather)



# Big Data Regulatory

## Landscape

Different Jurisdictions Driving Different Approaches

### European Union

More Focus On Individual Data Protection And Privacy

### United States

Specific Regulations By Industry Such As HIPAA, Graham Leach, PCI DSS, FTC (Section 5)



# Big Data and Privacy Collision

## Privacy Law

vs.

## Big Data

Data Minimization

Notice & Consent Tied to Purpose

Data Destruction

Aggregation & Anonymization

Data Maximization

Not necessarily known at collection

Data Re-use

Re-identification Risk



# Evolving Regulatory Landscape

## EU: Right to be Forgotten

EU on verge of implementing a comprehensive EU Data Protection Regulation.

Violations of the regulation re processing and transfer of personal data could trigger a maximum fine of 100 M Euros or 5% of such company's annual revenues.

# Evolving Regulatory Landscape

**(cont.)**  
United States

Federal Trade Commission, under Section 5, is increasing scrutiny in area of Internet of Things

8/2013 Chairwoman, Edith Ramirez warned that the FTC intended to take a more active role in policing large collections of data. Focus on tech companies such as Google and Twitter.

11/2013 FTC will host Internet of Things workshop in DC to explore consumer privacy and security issues posed by the growing connectivity of devices.



# Evolving Regulatory Landscape (cont.)

## FTC Cited Privacy Concerns

Indiscriminate collection of data

Data Generation

Predictive Analytics

Indefinite Storage

Data Breaches

FTC Chairwoman, Edith Ramirez at Technology Institute Aspen Forum





# Corporate Governance

## Implement **Privacy by Design**

Ensure policies and practices re Big Data privacy and security are robust

Use an end to end data management program (from classification to retention and destruction)

Train, train and train – Clear Accountability

## **Privacy By Design**

“Building privacy in as products and services are being developed. To do that, companies need to perform risk assessments to lay bare vulnerabilities by asking tough questions - (for example) are security measures appropriate given the volume and sensitivity of the data?”

FTC Chair Edith Ramirez at the Technology Policy Institute Aspen Forum

# Corporate Governance (cont.)

Ownership

Use

Purpose

Collection

Custody

Protection

Aggregation/Anonymization

Destruction





# Aggregation and Anonymization

Help enhance consumer privacy, but there are no clear industry standards

Reduces the risk of re-identification and better enables organizations to respond to new regulatory requirements

De-identified data can be triangulated with other data sets to enable a mapping back to the original data (Netflix case)

Need to be very careful exposing anonymized data



# What does this all mean?

Disruption is occurring in every industry – Analog to Digital Industries

The potential benefits of intelligent systems are vast and the synergistic effects of widespread machine instrumentation can be realized across fleets and networks

Companies need to anticipate the digital revolution and prepare now to handle both the quantity and different types of data

What's needed is a clear set of laws and regulations that allow the free flow of industrial data across country borders while protecting the rights of individuals



imagination at work

© General Electric Company, 2013