A Data-rich World for a Better World: From Sensors to Sense-Making

Kirk Borne
Principal Data Scientist, Booz Allen Hamilton
A Data-rich World for a Better World: From Sensors to Sense-Making

Kirk Borne

Principal Data Scientist
Booz Allen Hamilton

http://www.boozallen.com/datascience
• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
OUTLINE

• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
Ever since we first explored our world…

...We have asked questions about everything around us.
So, we have collected evidence (data) to answer our questions, which leads to more questions, which leads to more data collection, which leads to more questions, which leads to **BIG DATA!**

\[ y \sim x! \approx x^x \rightarrow \text{Combinatorial Growth!} \]  
(all possible interconnections, linkages, and interactions)

\[ y \sim 2^x \text{ (exponential)} \]

\[ y \sim 2 \times x \text{ (linear growth)} \]
Making Sense of the World with **Smart Data**

**Semantic, Meaning-filled Data:**
- Ontologies (formal)
- Folksonomies (informal)
- Tagging / Annotation
  - Automated (Machine Learning)
  - Crowdsourced
  - “Breadcrumbs” (user trails)

**Broad, Enriched Data:**
- Linked Data (RDF)
  - All of those combinations!
- Graph Databases
- Machine Learning
- Cognitive Analytics
- Context
- The 360° view

*The Human Connectome Project: mapping and linking the major pathways in the brain.*

http://www.humanconnectomeproject.org/
OUTLINE

• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
Everything Interconnected

Internet of Things (IoT):
Deploying intelligence at the point of data collection!
(Machine Learning at the edge of the network = Edge Analytics!)

The Internet of Things (IoT) will be an interconnected universe of Dynamic Data-Driven Application Systems (dddas.org) =>

Combinatorial Explosive Growth of Smart Data!

https://www.linkedin.com/pulse/can-elephant-bigdata-dance-iot-internet-things-tune-lambba

Ubiquitous Smart Data in the IoT:
Applications and Use Cases are everywhere...
... that demands dynamic action = Edge Analytics

- Smart Apps
- Predictive Retail
- Precision Marketing
- Smart Highways
- Precision Traffic
- Smart Cities
- Predictive Law Enforcement
- Personalized Learning
- Smart Health
- Precision Medicine
- Precision Farming
- Personalized Financial Services
- Smart Organizations
- Predictive Maintenance
- Prescriptive Maintenance
- Smart Grid

Smart => Predictive, Precision, Personalized!
OUTLINE

• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
What is Self-Driving? ... ~Autonomous! 
(e.g., cars, cities, deep space probes, organizations)

<table>
<thead>
<tr>
<th>Defining Characteristics</th>
<th>Data Analytics Maturity</th>
<th>Organizational Posture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing, Streaming</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning, Agile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contextual, Optimizing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deciding, Acting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** additional KEY CHARACTERISTICS = Stateful, Secure, Compliant, “Embedded Reporting” = explosive growth of “in situ” data that delivers CONTEXT and SENSE-MAKING
What is Self-Driving? ... ~Autonomous! (e.g., cars, cities, deep space probes, organizations)

<table>
<thead>
<tr>
<th>Defining Characteristics</th>
<th>Data Analytics Maturity</th>
<th>Organizational Posture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing, Streaming</td>
<td>Descriptive</td>
<td></td>
</tr>
<tr>
<td>Responsive</td>
<td>Diagnostic</td>
<td></td>
</tr>
<tr>
<td>Learning, Agile</td>
<td>Predictive</td>
<td></td>
</tr>
<tr>
<td>Contextual, Optimizing</td>
<td>Prescriptive</td>
<td></td>
</tr>
<tr>
<td>Deciding, Acting</td>
<td>Cognitive</td>
<td></td>
</tr>
</tbody>
</table>

Note: additional KEY CHARACTERISTICS = Stateful, Secure, Compliant, “Embedded Reporting” = explosive growth of “in situ” data that delivers CONTEXT and SENSE-MAKING
**What is Self-Driving? ... ~Autonomous!**
*(e.g., cars, cities, deep space probes, organizations)*

<table>
<thead>
<tr>
<th>Defining Characteristics</th>
<th>Data Analytics Maturity</th>
<th>Organizational Posture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing, Streaming</td>
<td>Descriptive</td>
<td>Passive</td>
</tr>
<tr>
<td>Responsive</td>
<td>Diagnostic</td>
<td>Reactive</td>
</tr>
<tr>
<td>Learning, Agile</td>
<td>Predictive</td>
<td>Proactive</td>
</tr>
<tr>
<td>Contextual, Optimizing</td>
<td>Prescriptive</td>
<td>Predictive Reactive</td>
</tr>
<tr>
<td>Deciding, Acting</td>
<td>Cognitive</td>
<td>Sense-making</td>
</tr>
</tbody>
</table>

**Note:** additional KEY CHARACTERISTICS = Stateful, Secure, Compliant, “Embedded Reporting”

= explosive growth of “in situ” data that delivers CONTEXT and SENSE-MAKING
OUTLINE

• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
# Levels of Maturity in the Self-Driving Enterprise

<table>
<thead>
<tr>
<th>Defining Characteristics</th>
<th>Data Analytics Maturity</th>
<th>Organizational Posture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing, Streaming</td>
<td>Descriptive</td>
<td>Passive</td>
</tr>
<tr>
<td>Responsive</td>
<td>Diagnostic</td>
<td>Reactive</td>
</tr>
<tr>
<td>Learning, Agile</td>
<td>Predictive</td>
<td>Proactive</td>
</tr>
<tr>
<td>Contextual, Optimizing</td>
<td>Prescriptive</td>
<td>Predictive Reactive</td>
</tr>
<tr>
<td>Deciding, Acting</td>
<td>Cognitive</td>
<td>Sense-making</td>
</tr>
</tbody>
</table>
5 Levels of Analytics Maturity

1) Descriptive Analytics
   - Hindsight (What happened?)

2) Diagnostic Analytics
   - Oversight (real-time / What is happening? Why did it happen?)

3) Predictive Analytics
   - Foresight (What will happen?)

4) Prescriptive Analytics
   - Insight (How can we optimize what happens?)

5) Cognitive Analytics
   - Right Sight (the 360 view, what is the right question to ask for this set of data in this context = Game of Jeopardy)
   - Finds the right insight, the right action, the right decision, ... right now!
   - Moves beyond simply providing answers, to generating new questions and new hypotheses, for your “next-best move”

@KirkDBorne ---- disummit, March 2017 ---- http://www.boozallen.com/datascience
PREDICTIVE Analytics

Find a function (i.e., the model) \( f(d,t) \) that predicts the value of some predictive variable \( y = f(d,t) \) at a future time \( t \), given the set of conditions found in the training data \( \{d\} \).

=> Given \( \{d\} \), find \( y \).

PRESCRIPTIVE Analytics

Find the conditions \( \{d'\} \) that will produce a prescribed (desired, optimum) value \( y \) at a future time \( t \), using the previously learned conditional dependencies among the variables in the predictive function \( f(d,t) \).

=> Given \( y \), find \( \{d'\} \).
Context is King!

“It’s not what you look at that matters – it’s what you see.” – Henry David Thoreau

- Context is “other data” about your data = *i.e.*, Metadata!
- The 3 most important things in your data are: Metadata, Metadata, Metadata!
- Metadata are...
  - Data that describes Data
  - Other Data that describes Your Data
  - Your Data that describes Other Data
- *e.g.*, Connected “Smart” Cars = that car that is braking 3 vehicles ahead of you = informs your vehicle to brake now!
- The Smart Business = predictive maintenance algorithm alerts the corresponding asset, the right skilled technician, & the right tool to converge at right place at the right time!
- IoT sensor data can provide a lot of context data (metadata!)
OUTLINE

• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
Case Study: Mars Rovers (metaphor)

@KirkDBorne ---- #disummit, March 2017 ---- http://www.boozallen.com/datascience
Mars Rover: intelligent data-gatherer, mobile data mining agent, autonomous decision-support system, and self-driving!

1) Drive around surface of Mars – take samples of items (Mars rocks):
   • experimental data type = mass spectroscopy = data histogram = feature vector

2) Perform Intelligent Data Operations (Data Mining in Action):
   • Supervised Learning
     – Search for items with known characteristics, and assign items to known classes
   • Unsupervised Learning
     – Discover what types of items are present, without preconceived biases; find the set of unique classes of items; discover unusual associations; discover trends and new directions (new leads) to follow
   • Semi-supervised Learning
     – Find the rare, one-of-kind, most interesting items, behaviors, contexts, or events

3) Enact On-board Intelligent Data Understanding & Decision Support
   = Science Goal (or Business Goal) Monitoring:
     – “stay here and do more”; or else “follow trend to a more interesting location”
     – “send discoveries to human analysts immediately”; or else “send results later”
Smart Sensors & Sentinels for Data-Driven Sense-Making and Decision Support

From Sensors to Sentinels to Sense

(for any application domain with streaming data from sensors)

• New knowledge and insights are acquired by monitoring and mining actionable data from all digital inputs (Sensors!)

• Alerts are triggered autonomously, without intervention (if permitted), applying machine learning and actionable business decision rules for pattern detection and diagnosis. (Sentinels! = embedded machine learning / data science algorithms, at the point of data collection = trained to minimize False Positives and “Alarm Fatigue”)

• “Smart Sensors” (powered by Machine Learning-enabled sentinels) will therefore deliver actionable intelligence (Sense!)
The MIPS Analytics Framework

for Dynamic Data-Driven Application Systems (DDDAS)

• **MIPS =**
  – Measurement – Inference – Prediction – Steering

• This applies to any Network of Sensors:

• **Machine Learning enables the “IP” part of MIPS:**
  – Autonomous (or semi-autonomous) Classification
  – Intelligent Data Understanding
  – Rule-based
  – Model-based
  – Neural Networks
  – Markov Models
  – Bayes Inference Engines

**Alert & Response systems:**
• Actionable insights from streaming sensor data
• Automation of any data-driven operational system

http://dddas.org
The MIPS Analytics Framework
for Dynamic Data-Driven Application Systems (DDDAS)

- **MIPS** =
  - Measurement – Inference – Prediction – Steering
- This applies to any Network of Sensors:
- **Machine Learning** enables the “IP” part of MIPS:
  - Autonomous (or semi-autonomous) Classification
  - Intelligent Data Understanding
  - Rule-based
  - Model-based
  - Neural Networks
  - Markov Models
  - Bayes Inference Engines

![From Sensors to Sentinels to Sense](http://dddas.org)

**Alert & Response systems**:  
- Actionable insights from streaming sensor data  
- Automation of any data-driven operational system
• First came Data, then Big Data, now Smart Data
• IoT: Dynamic Data-Driven Application Systems
• The Self-Driving Enterprise
• Steps to Discovery: Data Science & Analytics
• Case Study: From Sensors to Sense-Making
• Data Rich for a Better World
Big Data + IoT + Citizen Data Scientists = Partners in Sustainability

The Internet of Things (IoT):
Knowing the knowable via deep, wide, and fast data from ubiquitous sensors!

Big Data:
In the Big Data era, Everything is Quantified and Tracked:
– Populations & Persons
– Smart Cities, Farms, Highways
– Environmental Sensors
– IoE = Internet of Everything!

Discovery through Data Science:
– Class Discovery
– Correlation Discovery
– Novelty Discovery
– Association Discovery

17 SDGs are KPIs for the World! (currently, the SDGs have 229 key performance indicators)
Big Data + IoT + Citizen Data Scientists = Partners in Sustainability

The Internet of Things (IoT):
Knowing the knowable via deep, wide, and fast data from ubiquitous sensors!

Big Data:
In the Big Data era, Everything is Quantified and Tracked:

The Art & Science of being Data-Rich:
From Sensors to a Better World

17 SDGs are KPIs for the World!
(currently, the SDGs have 229 key performance indicators)