



JUN  
12

**Big data, machine learning og chatbots |  
Gratis frokostmøte**

by Forum for kunstig intelligens og robotikk Norge

Free

**BHUVAN UNHELKAR, PHD, FACS**  
SENIOR CONSULTANT METHODSCIENCE;  
PROFESSOR, UNIV. OF SOUTH FLORIDA, SARASOTA-MANATEE  
[BHUVAN@METHODSCIENCE.COM](mailto:BHUVAN@METHODSCIENCE.COM); CELL: +1(941) 404-4406



**12<sup>th</sup> June, 2018,  
Oslo, Norway**

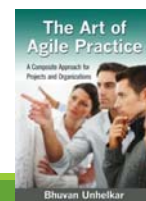
**“Big” is a Relative Term  
(Context dependent)  
[IBM 5MB Hard Drive 1956](#)**



## Presenter Profile - Dr. Bhuvan Unhelkar

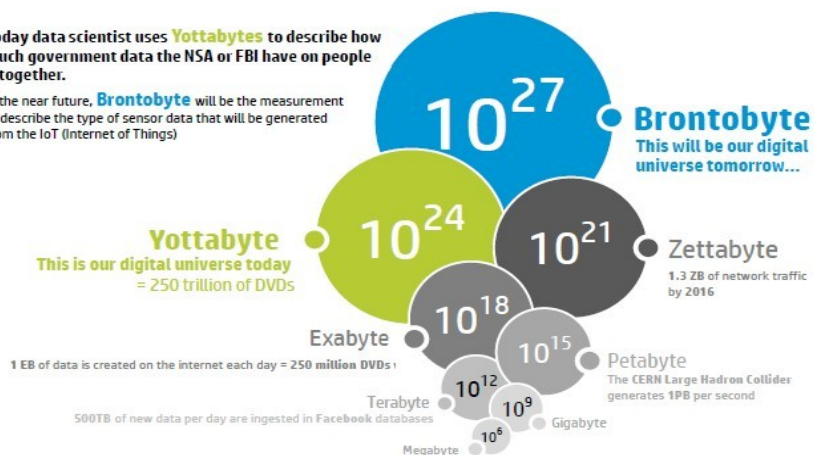
(BE, MBA, PhD, FACS, CBAP®, PSM)

- Professor and Lead IT Faculty, Univ. of South Florida, Sarasota-Manatee; Founder, MethodScience.com; PlatiFi.com;
- Courses: UML, Agile PM, Big Data; Designed and Presenter Australian Computer Society's Business Analysis (BAS) and Agile PM Online
- Visiting Faculty at UTS (Agile SW Modeling) and WSU (UWS)
- Author: 20 Books (Including Big Data Strategies for Agile Business)
- Supervisor: 7 PhD Completions;
- Fellow of the Australian Computer Society; IEEE Sr. Member; Life Member, Computer Society of India & BMA
- President – Rotary Club of Sarasota Sunrise, USA (Paul Harris Fellow+4; AG); TiE ex-Mentor Director;
- [www.unhelkar.com](http://www.unhelkar.com) & [www.methodscience.com](http://www.methodscience.com)



Today data scientist uses **Yottabytes** to describe how much government data the NSA or FBI have on people altogether.

In the near future, **Brontobyte** will be the measurement to describe the type of sensor data that will be generated from the IoT (Internet of Things)



[www.theregister.co.uk](http://www.theregister.co.uk)



Agile Product & Project Management  
Executive Update Vol. 15, No. 5

### Lean-Agile Tautology

by Bhuvan Unhelkar, Senior Consultant, Cutter Consortium

of the Lean-based Toyota Production System (TPS)?  
Another example from Lean-Six Sigma is the technique of asking "why" five times by a business analyst undertaking root-cause analysis of a system problem. And the works of Mary and Tom Poppendieck in the area of Lean software development provides a shining example of the application of Lean in the software arena!

# Update

## CUTTER CONSORTIUM



Agile Product & Project Management  
Vol. 15, No. 5

### Agile Outsourcing: Cross-Cultural, Cross-Regional Perspectives

by Bhuvan Unhelkar, Senior Consultant, Cutter Consortium

The Executive Report discusses the challenges and the value of using Agile methods in outsourced projects. The use of Agile ... will be ...

## Agile Business Analysis: Part II — Organizational Adoption with Centers of Excellence



Agile Product & Project Management  
Vol. 14, No. 5

### The Psychology of Agile: Fundamentals Beyond the Manifesto

by Bhuvan Unhelkar, Senior Consultant, Cutter Consortium

The destiny of Agile depends on contributions from many different disciplines and thought processes. The psychology of Agile aspires to make one such contribution by discussing the impacts and influences of

Senior Consultant,

Report series on agile business analysis es relating to organizational adoption. This eskers who may not necessarily be hands-on the organizational-wide responsibility for d project-level application, the report frameworks, BA centers of excellence

# <https://www.whizlabs.com/blog/big-data-trends-in-2018/>



CLOUD COMPUTING ▾ BIG DATA ▾ PROJECT MANA

Blog » Big Data » Big Data Trends in 2018



## Big Data Trends in 2018

By **Amit Verma** - January 22, 2018 BIG DATA [f](#) [t](#) [G+](#) [in](#) [p](#)

Big data is no longer just considered as a big or large set of data. Instead, it is more about business-driven data in the use of analytics capabilities for long-term business value. However, the massive increase in the data volume, varieties, and flow of data for analytics showcases that big data trends in 2018 will change a worth amount.

McKinsey Global Institute



May 2011

## Big data: The next frontier for innovation, competition, and productivity

[http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)

Big data can unlock significant value by making information transparent and usable at much higher frequency.

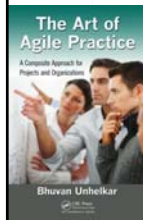
As organizations create and store more transactional data in digital form, they can collect more accurate and detailed performance information

Big data allows ever-narrower segmentation of customers and therefore much more precisely tailored products or services.

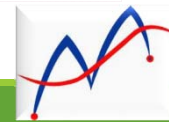
Sophisticated analytics can substantially improve decision-making.

Big data can be used to improve the development of the next generation of products and services.

# POSITIONING BIG DATA STRATEGIES



TRANSCENDING ANALYTICS AND  
TECHNOLOGIES;  
MAPPING TO BUSINESS AGILITY



© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

## What is Big Data?

‘Big Data’ is still ‘Data’

- But difficult to process with the given tools and technologies
- Could be made up of much wider ‘types’ or ranges of data



Handling such Data requires Latest:

- Techniques, tools, and architecture
- Frameworks to guide adoption and reduce Risks

With an aim to:

- Solve new challenges in Business
- Improving the Value proposition

## What is Big Data?

(from An Executive Primer on Big Data – Thornton May, Cutter IT Journal, 2012 pp 14-18, Survey based on [www.allanalytics.com](http://www.allanalytics.com))

could dance on the head of a pin. Rob Das, cotounder of Big Data pioneer Splunk, begins most presentations sharing results of a national poll asking people to respond to the question, “What is Big Data?” Respondents gave the following answers:<sup>5</sup>

- A legitimate problem stemming from the growth of unstructured data (51%)
- A new catchphrase for an old data management challenge (23%)
- A meaningless marketing catch-all (9%)
- Another way to say “data warehouse” (5%)
- Another way to say “Hadoop” (3%)

## What has Enabled Big Data to Happen?

Key enablers for the appearance and growth of Big Data are

- Increase in storage capacities in devices & on the Cloud
- Increase of processing power (CPU)
- Increase in Generation (and Availability) of data by businesses, individuals and machine sensors
- (Every day we create 2.5 quintillion bytes of data; 90% of the data in the world today has been created in the last two years alone)

## Some interesting Facts About Big Data: (Not that it matters too much!)

100 hours of video uploaded to YouTube every minutes

Can Take 15 years to watch all those Videos

NSA analyze 1.6% of all global internet traffic around 30 petabytes or 30 million gigabytes every day

Retailer could increase profit by 60% through the full exploitation of big data analytics

(SeSense.com) > Nearly 2.3 trillion gigabytes of data created every single day and the data universe doubling every two years

Big data industry is 54.3 billion in 2017

## Dimensions of Big Data

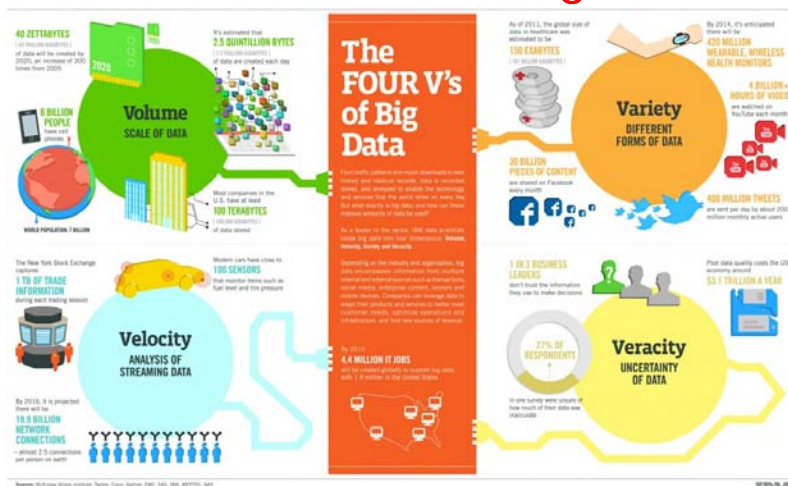
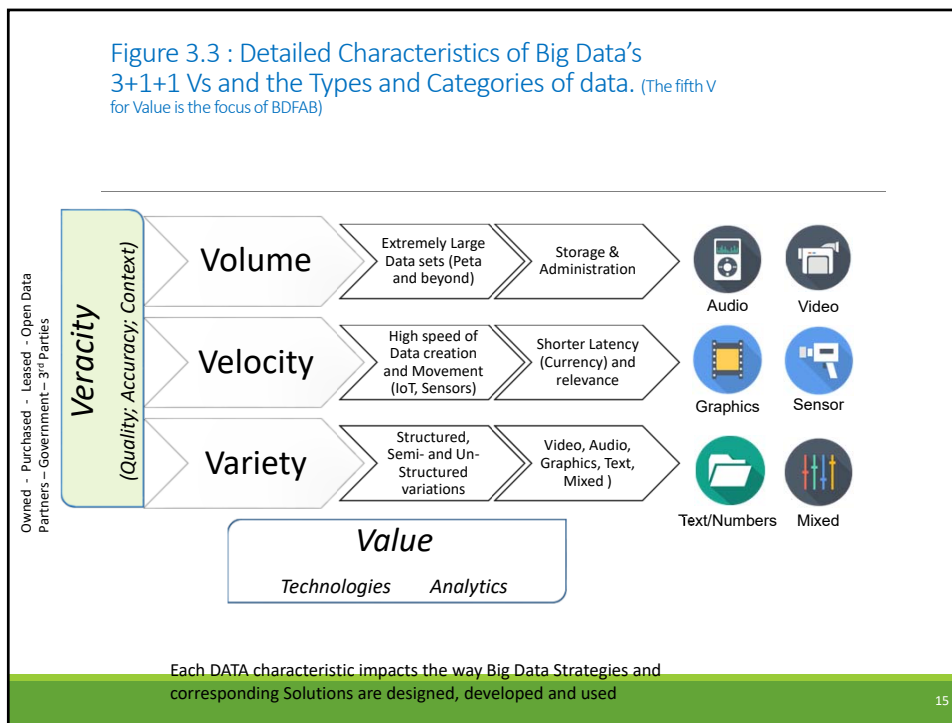


Figure 3.3 : Detailed Characteristics of Big Data's  
3+1+1 Vs and the Types and Categories of data. (The fifth V  
for Value is the focus of BDFAB)



## The Key Business Challenge

Big or Small – Business is NOT interested in Data

Business interest are as follows:

- Can Data be used to “Make sense” of a situation?
- Can Data enhance Decision Making?
- Can Data be Managed? (i.e. Security & Privacy)
- Is there a Business “Value” in Data?



Executive  
UPDATE

**CUTTER CONSORTIUM**  
●●● Access to the Experts

DATA ANALYTICS & DIGITAL TECHNOLOGIES

## Extracting the 5th “V” for Value in Big Data Strategies

by [Bhuvan Unhelkar](#), Senior Consultant, Cutter Consortium

[Early literature](#) on big data discusses the “three Vs”: volume, velocity, and variety. Adding the “fourth V” — veracity (or quality) — creates the [contemporary description](#) of big data. These characteristics of big data are made possible due to Hadoop-based technologies and corresponding sophisticated analytical algorithms. Big data technologies (or analytics) in themselves are not capable of providing strategic value to business. In fact, big data in itself [does not guarantee](#) business success and, as highlighted in my last *Executive Update*, can be described as something I call the “[big data strategy lacuna](#).”

Figure 1.1 Big Data Strategies – Transcending Analytics and Technologies

**Strategies (High Level):**

- Business/ End-User Value;
- Make sense of existing data ;
- Spotting Patterns
- Identifying new data
- Mapping (Unstructured) to a Structure d
- Optimizing Business Processes
- Contextualizing Enterprise Architecture
- Sustainability Value
- Balancing the ROI

### Strategies

Decision Maker;  
End-User; Manager;  
Owner; Worker.  
[Government;  
Community;  
Environment]

ROI; Values

Algorithms

Sentiment,  
Predictive,  
Prescriptive,  
(NPS, Text Mining)

↕

Agile  
Business

↕

Application in Business

Building the Capabilities

↕

Hadoop/  
HDFS Spark  
NoSQL  
Databases;  
MapReduce

Architecture

### Analytics

↔

### Technologies

Storage &  
Backups; Sharing;  
Interfacing;  
Security and  
Privacy;

Management

↕

SMAC-stack

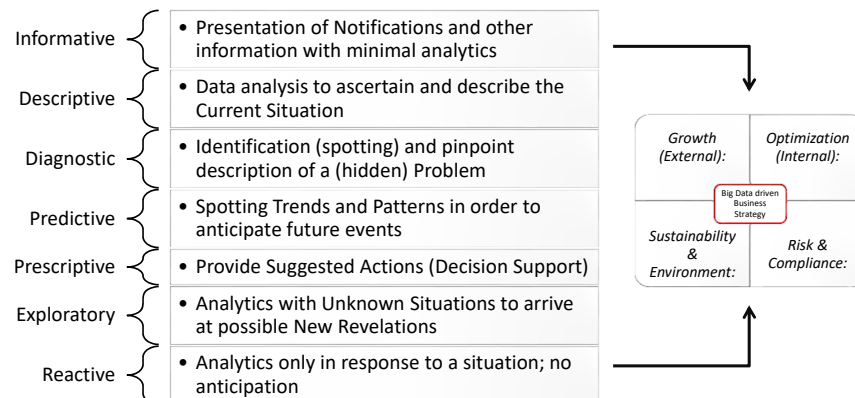
9

# ANALYTICS & TECHNOLOGIES

MACHINE LEARNING;  
ARTIFICIAL INTELLIGENCE

© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

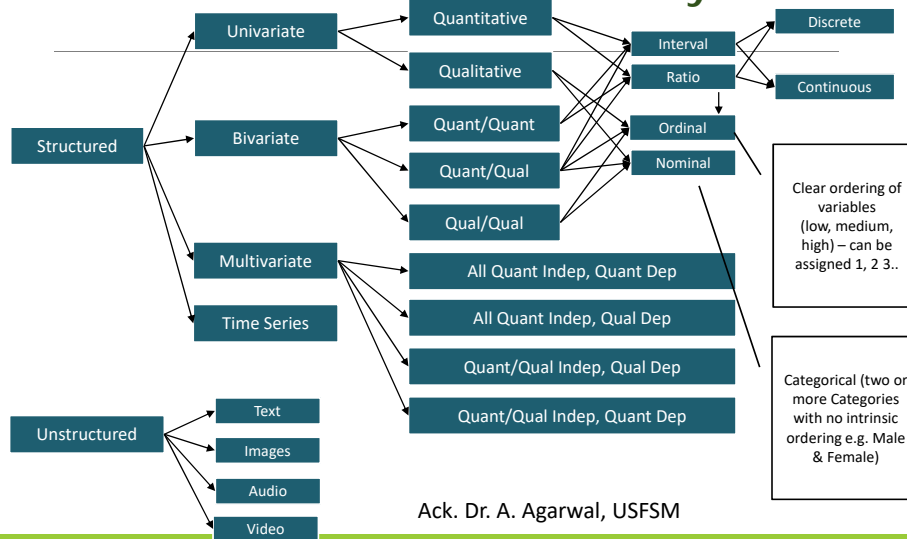
## Categories providing Agile Business Values (and form basis of Business Strategies)

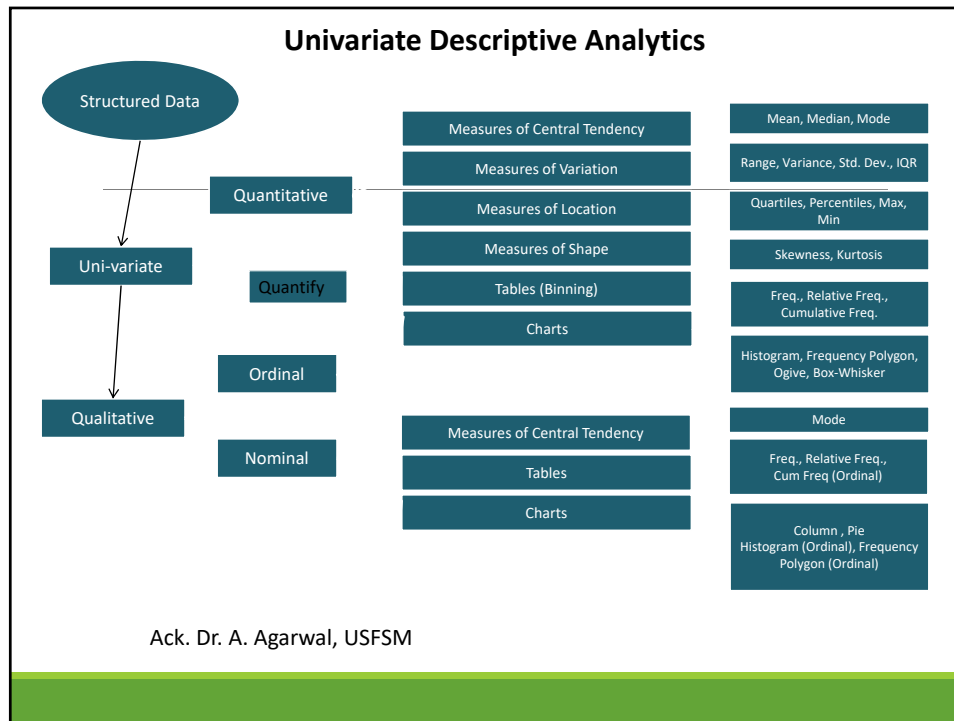


**Figure 3.10: Identifying Associations and mapping Clusters**  
 (Based on Agarwal, A., "Predictive Analytics and Text Mining") – An increasingly finer granular Analytics (Amazon's Long-tail Marketing : Products (A,B,C..) and Customers (X1, X2..))

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
A				1		1				
B		1						1	1	1
C			1		1		1		1	
D				1		1				
E	1		1				1			1
F			1				1	1		
G				1		1				
H		1			1				1	
I				1		1				1
J	1		1		1					
K		1						1	1	


### A Framework for Descriptive Analytics






## Machine Learning...

1. Consider what a Machine learns.....  $A + B = C$
2. We then add conditions: If  $A \geq 5$  THEN.....  $A + B = C$
3. Machine Algorithm: What are my options when  $A \geq 5$ ; and when  $A < 5$ ;
4. What had happened in the past? (Examining Historical Patterns)
5. When  $A \geq 5$  and  $B = 'n'$ , (where 'n' ranged from  $-10$  to  $+25$ ),  $C$  was Positive for 1 BILLION records



**Cutter  
IT Journal**

The Journal of  
Information Technology Management




---

Vol. 29, No. 4  
April 2018

**"The sheer variety, volume, and velocity of data the IoT generates create a major obstacle to realizing its full potential. This calls for a strategic approach to data management."**

— Bhuvan Unhelkar and San Murugesan,  
Guest Editors

**NOT FOR DISTRIBUTION**  
For authorized use, contact  
Cutter Consortium:  
+1 781 643 8700  
info@cutter.com



## IoT Data Management and Analytics: Realizing Value from Connected Devices

**Opening Statement**  
by Bhuvan Unhelkar and San Murugesan ..... 3

**Solving the Jigsaw Puzzle: An Analytics Framework for Context Awareness in the Internet of Things**  
by Anurag Agarwal, Ramakrishna Govindu, Sunita Lodwig, and Ravi T. Ngo ..... 6

**IoT Data Management Challenges: A View from the Trenches**  
by Prasad Shah, Suman Dutta, Reha Varkeyanban, Sathakara Pejaraj, and Vijay Navkar ..... 12

**Exploratory, Embedded, Enabled: The Three E's of IoT Data Management Maturity**  
by Donald S. Wynn, II and Renee M.E. Hunt ..... 17

**Leveraging EA and IoT Synergy for Digital Transformation**  
by Nishu K. Hazra ..... 24

**A Common Thread: Applying Hex Elementization in IoT Data Analytics**  
by Gishai Nair and Yi-Chen Lan ..... 31



LET'S TALK ABOUT SIX

## A Common Thread: Applying Hex Elementization in IoT Data Analytics

by Gishai Nair and Yi-Chen Lan

---

From the Stone Age to the Renaissance, from the Industrial Revolution to the Digital Age, there have always been upheavals and paradigm shifts in the way we humans have conducted ourselves in any field or phase of society. These revolutions helped us break through the boundaries we set up from our prior set of knowledge. The Internet of Things (IoT) is the latest focus of society's participants, consumers and producers alike. What is exciting is the vast array of IoT applications. IoT can be extended to the machine-to-machine (M2M), transportation, healthcare, nanotechnology, and utility domains. It can be used for vertical and horizontal IoT solutions. This disruptive technology has major implications for e-commerce and the industrial Internet.<sup>1</sup> IoT development will also accelerate the adoption of big data. In fact, the big data flow from IoT will expand to the extent that managing it smartly will be paramount.

IoT at its core is all about integration and marriage of "things" which can encompass anything from objects and machines to software and human beings. The question is how IoT will exist among a myriad of data types, information exchange systems, applications, and hardware and communication systems. Integration is not just confined to different kinds of devices and their mediums, but must also be undertaken as a business strategy.<sup>2</sup> The Digital Age seems to challenge long-held economic principles; abundance of data, not scarcity, will become the norm.<sup>3</sup> As IoT becomes ubiquitous on a lot of things (IoT), it will entail a lot of Un-Interpreted Data (AUID). The challenge is to ensure that IoT is interconnected in an increasingly meaningful way. However, with so many companies working on so many services, which will communicate in so many different ways and via so many different protocols, how are we going to find a common thread among them?

"Hex elementization" is a concept conceived to deal with this. Hex elementization involves breaking a source of information, irrespective of its form, into

"hex elements" with six pieces of information and then integrating them. The source of information here can be a set of unstructured data, structured data, microphone or digital signals, sensor data, machine language, a language used by robots to communicate with each other, or any other future way of communication. The intention is to break the information sources down into the simplest form while embedding within them enough attributes to help them automatically seek and match themselves with information seen in other unrelated pieces of hex elements. Consider, for example, the hundreds of x-ray photos, CT scans, ultrasound scans, blood reports, and other pieces of data sitting in cupboards and digital storage in a hospital. A heuristic, architecture-based approach can enable them to find meaningful information. An architecture into which you can throw anything and get unexpected and interesting relationships and information snippets is what we are theorizing in hex elementization.

**HEX ELEMENTIZATION EXPLAINED**

Hex elementization aims to create a platform — an architecture, an environment — encompassing end-to-end integration. The traditional way of thinking about data is to find its (bits) is unutilized if we want to create a common "program" for IoT. No matter how large or diverse the data is, it needs to be broken down into smaller chunks that will enable ease of interaction. These smaller chunks (i.e., hex elements) can be combined and related into meaningful information.

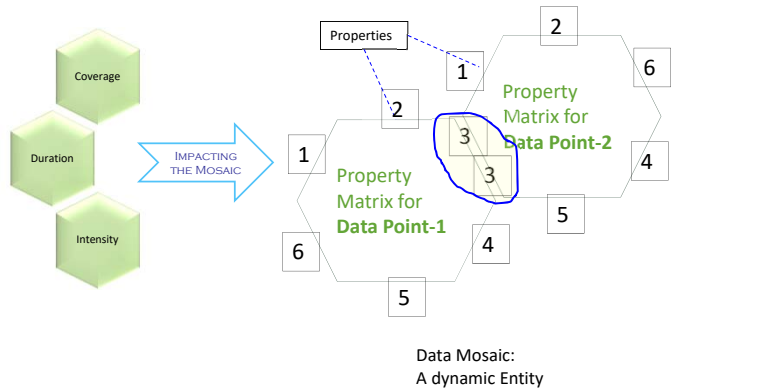
Why "hex" elements? Because the hexagon is a shape that can be found in the most basic of structures. It is found widely in nature — for example, honeycombs, ice crystals, and even the security discovered grapes!<sup>4</sup> In other words, the hexagon is a shape that occurs naturally and is trusted by nature. Nature finds this structure helpful, effective, and efficient. Figure 1 shows how data from disparate sources (IoT devices) can be broken down into six-sided hex elements.

Get The Cutter Edge Now: [www.cutter.com](http://www.cutter.com)

**NOT FOR DISTRIBUTION** For authorized use, contact  
Cutter Consortium: +1 781 643 8700 • [info@cutter.com](mailto:info@cutter.com)

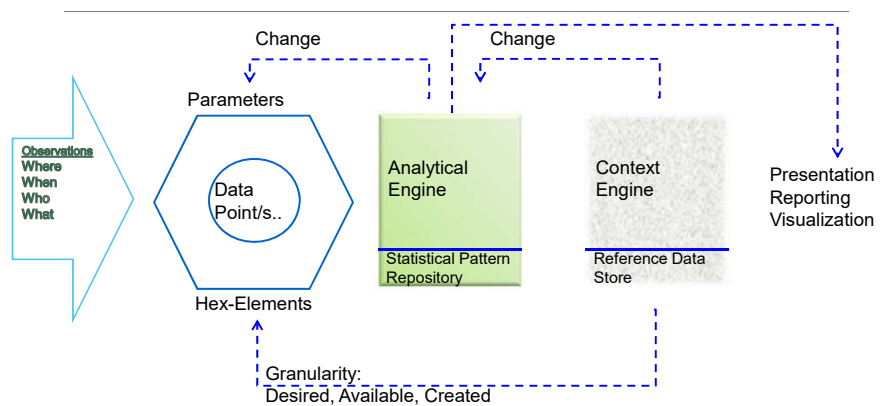
Vol. 29, No. 4 CUTTER IT JOURNAL 31

Figure 3.6: Hex-Elementization as a mechanism for Context of a Data Point  
(Based on research conducted by Girish Nair, WSU, Australia)

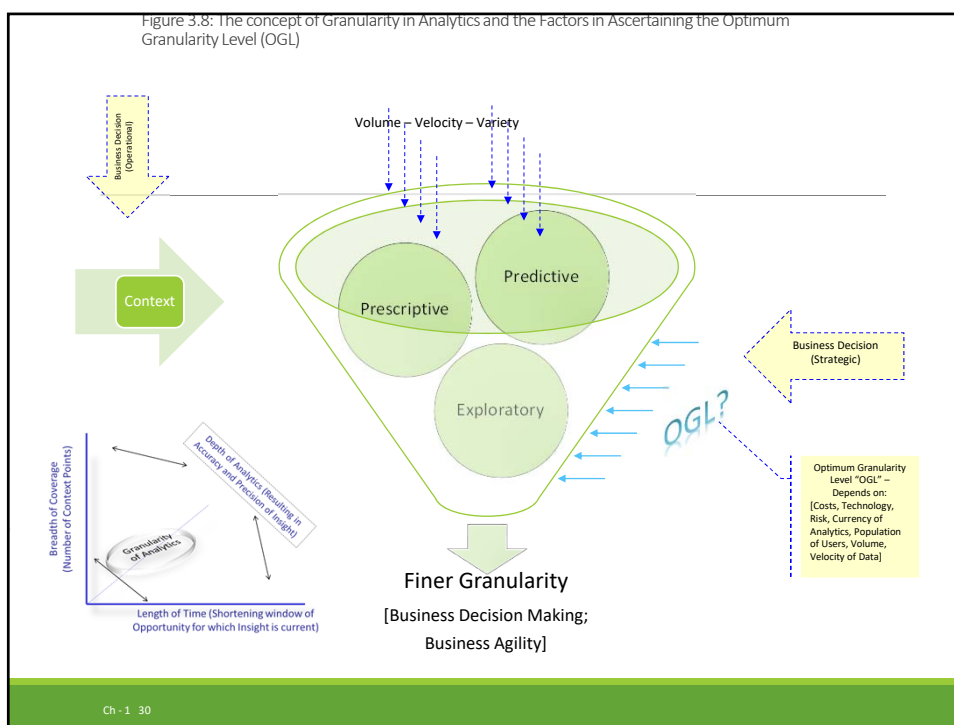
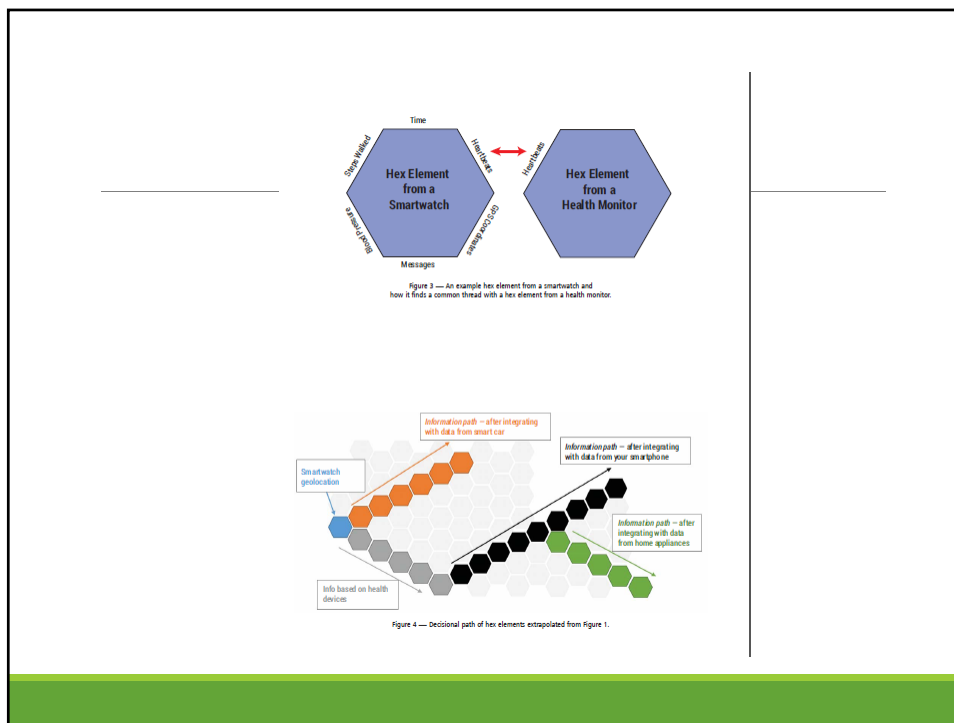


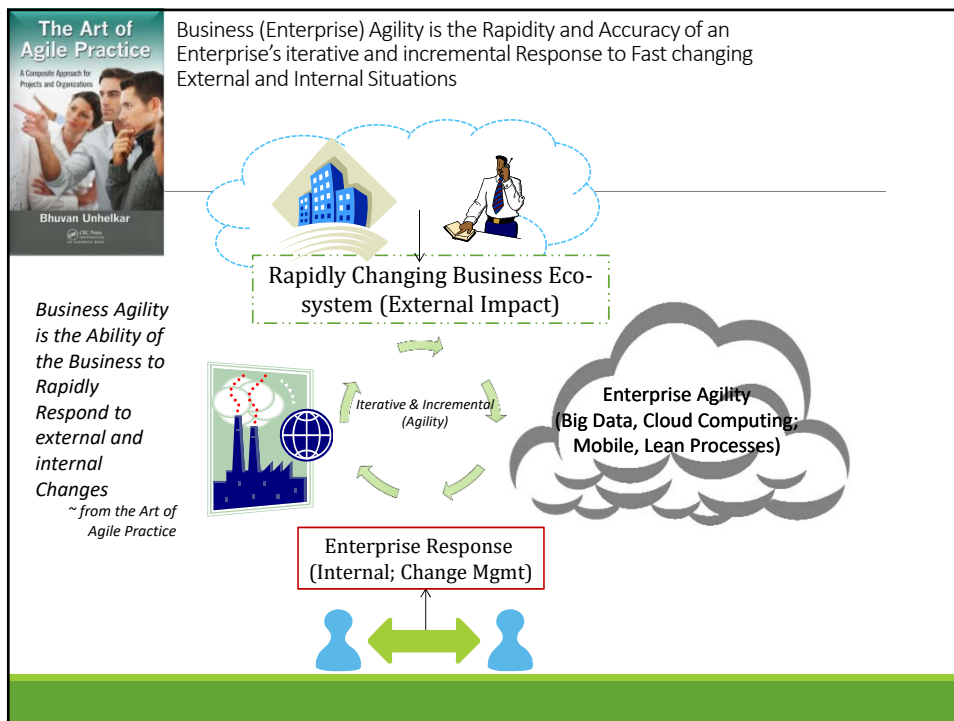
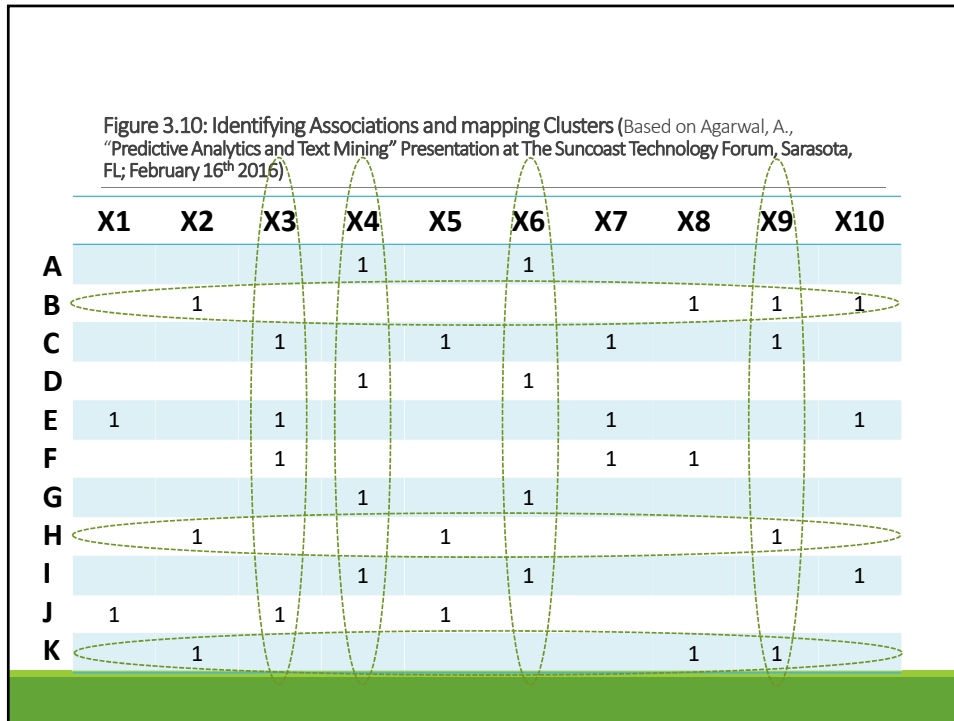
Ch - 3 27

Figure 3.7: Journey of a Data Point via Context and Analytical Engines and the Granularity-driven Feedback Loop



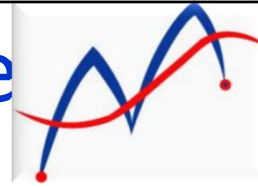
Ch - 1 28







# Sub-Module



## EXAMPLES (USE CASES) IN PRACTICE

TOWARDS STRATEGIC USE OF BIG DATA

© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

### Farecast by Oren Etzioni

(Part of Bing now)

Crunches 200-billion price records

Saves \$50 per ticket on average

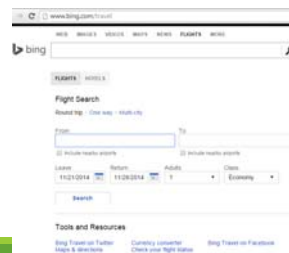
- 75% of the time

Entirely based on Big Data technologies in the background

- Was first Major attempt at use of “Big Data” in 2004; Technology Followed Business
- Sophistication grew with Hadoop

Acquired by Microsoft for \$110 M

*Shift in Mindset in how Data is Used*



34

## 23andMe.com (DNA exploration)

Gene sequencing as an Industry

- From a few thousand \$ to \$ 99

Hints at Predicting cancer, heart etc.

Need to sequence Billions of base-pair DNAs

- Not many are possible for this cost
- Steve Job's sequenced entire Gene set ( $N=All$ )



35

## Inrix.com (Traffic Scorecard)

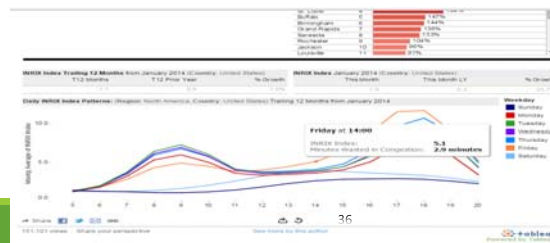
Data is collected from 100 Million Vehicles

Real time Geo positions

- BMW, Ford and Toyota
- Collaborations for Value
- Fleet vehicles (Taxis)

Smartphone Apps

- Driver permits data collection



36

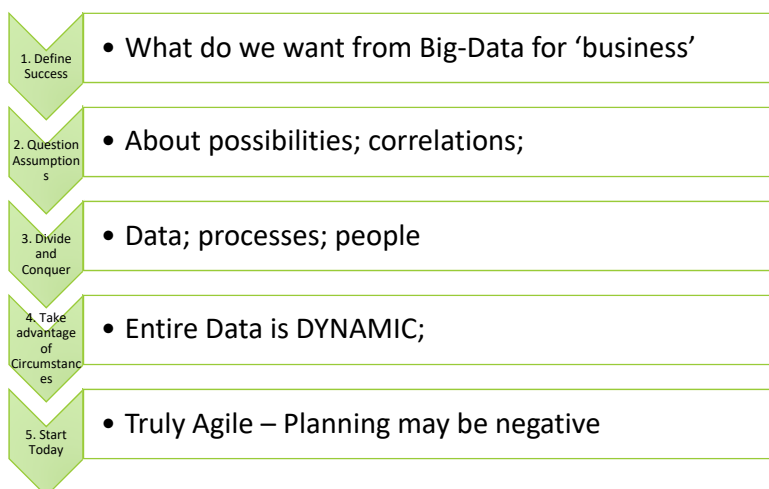
## Predicting Crimes and Terrorist events

Big data has been used to predict crimes & disaster before it happens (e.g.

- Predictive policing; Enabling Evasive Actions;



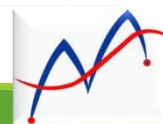
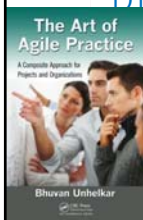
Campaign Manager – 2008 elections  
(Dan Siroker - Optimizely) 5 part strategy:  
*For Success with Big Data*



# BIG DATA FRAMEWORK OF AGILE BUSINESS (BDFAB v2.5)



## DEVELOPMENT & USE OF THE FRAMEWORK FOR BIG DATA ADOPTION



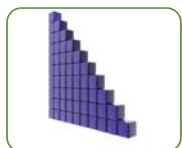
© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

### Examples of Technical, Analytical and Strategic Decisions related to Big Data



How many Nodes on a Cluster? How to distribute 1 Petabyte?

- MapReduce / HDFS / Spark (In-Memory)
- NoSQL (MongoDB) for non-Transactional Unstructured



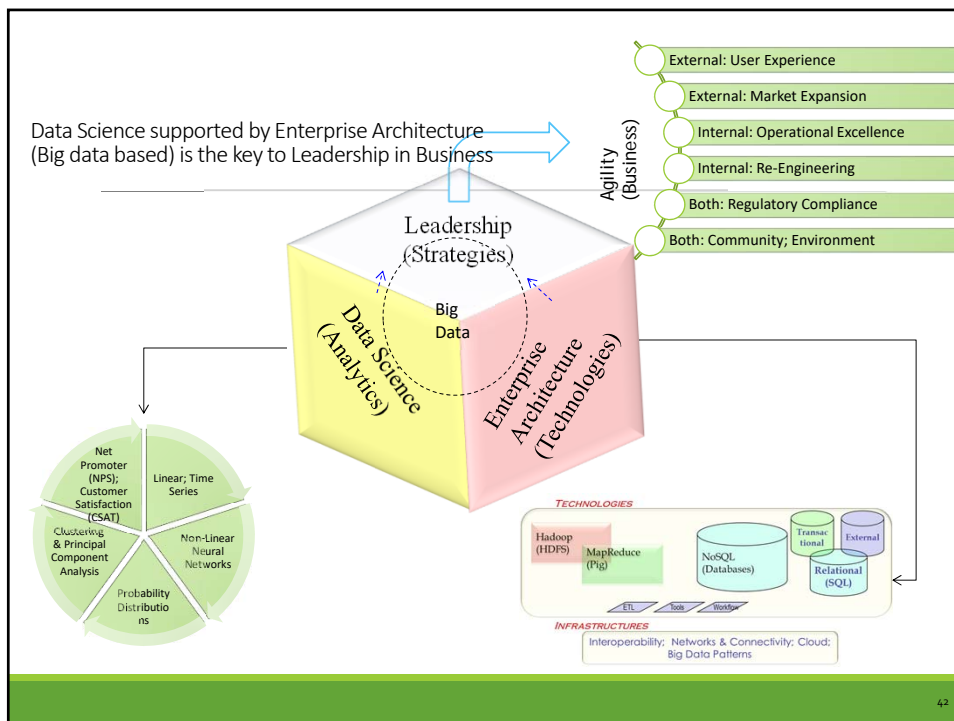
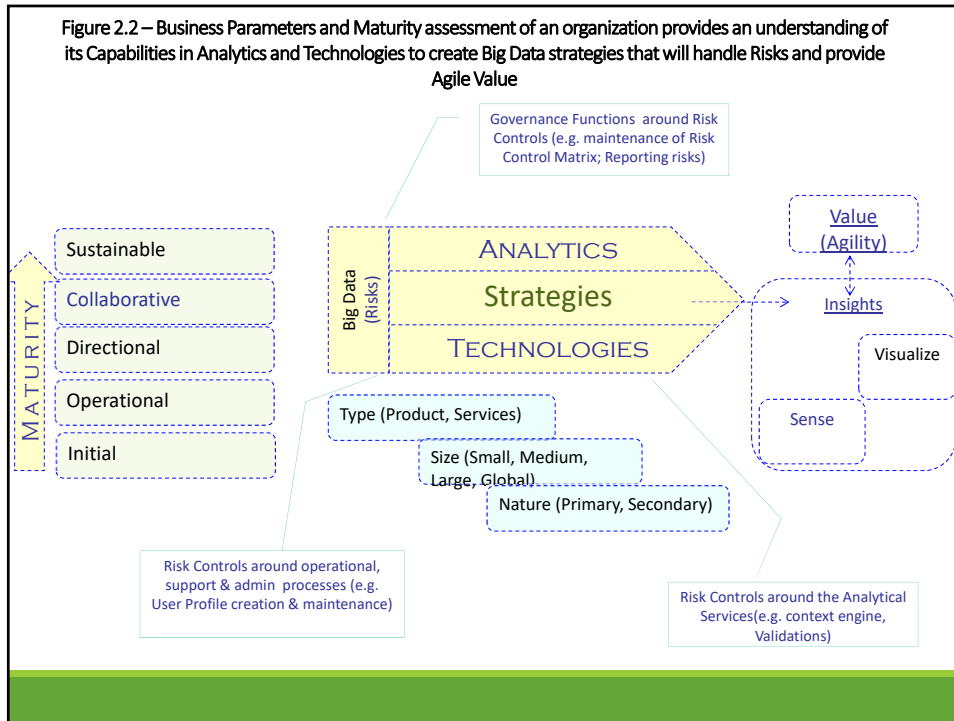
What is the NPS of the new product launched last month?

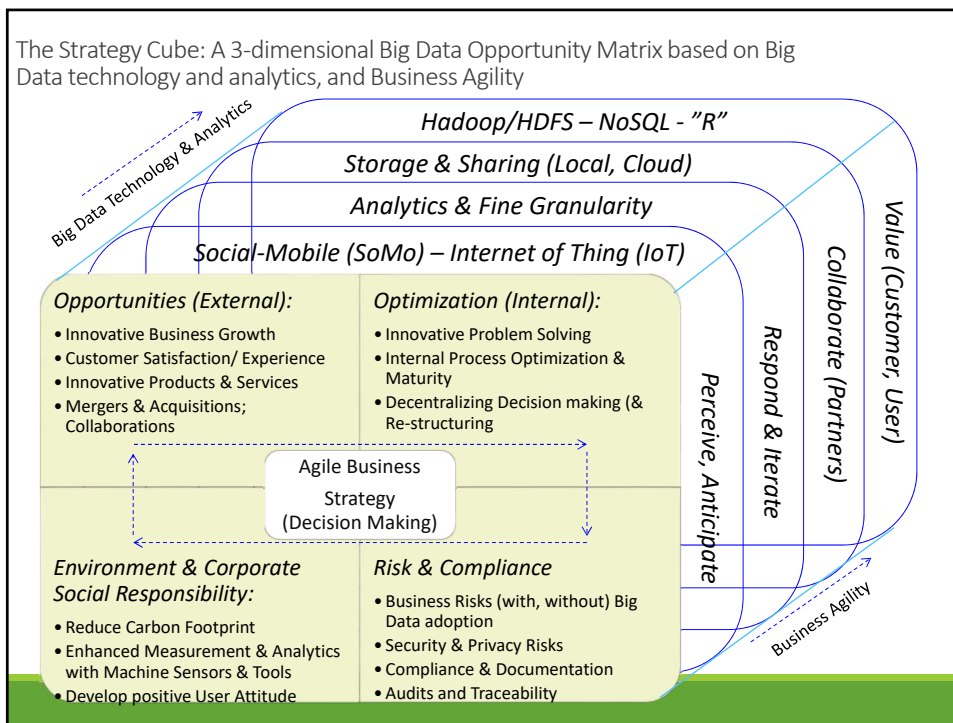
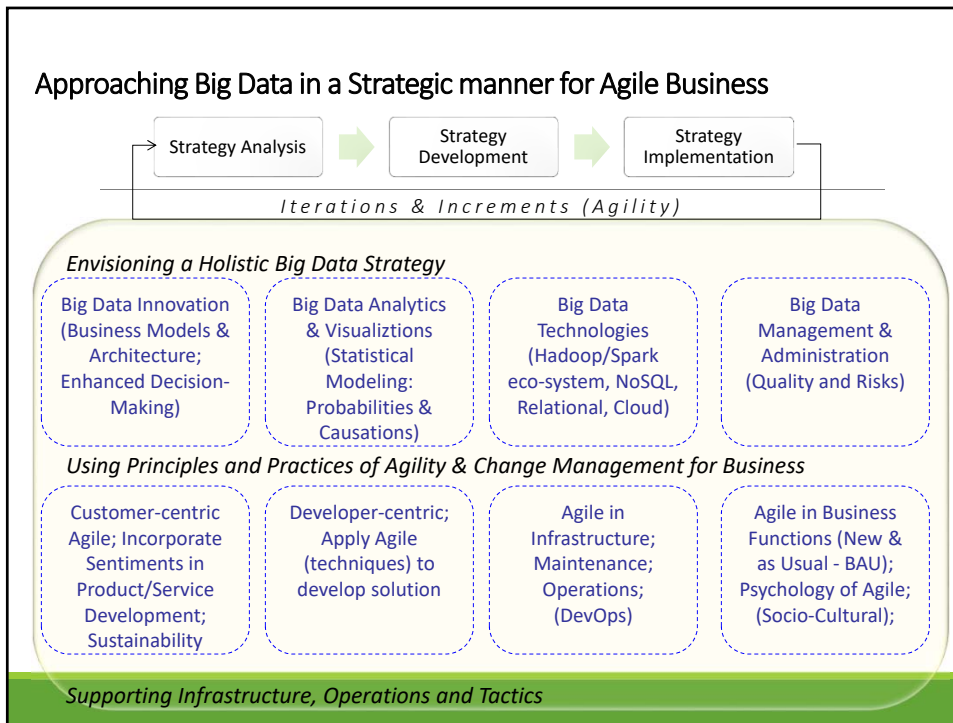
- Why is the NPS so low? And which product parameters need to be changed to improve the score?



Customer Service Excellence

- Undertaking Business Process Integration with a NoSQL DB to enable decentralization in Decision Making
- Enhance Knowledge Sharing in the process to enable Business Agility





## *Opportunities (External):*

Innovative Business Growth

Customer Satisfaction/ User Experience

Innovative Products & Services

Mergers & Acquisitions; Collaborations with External,  
Government, Third-party Organizations



## *Optimization (Internal):*

Innovative Problem Solving

Internal Process Optimization & Maturity

Decentralizing Decision making (& Re-structuring of  
the organization to a Process-oriented flat, non-  
hierarchical structure)



## *Environment & Corporate Social Responsibility:*



Reduce Carbon Footprint

Enhanced Measurement & Analytics with Machine Sensors & Tools to provide most current Data

Incorporate Automated feedback loops

Develop positive User Attitude

## *Risk & Compliance*

Business Risks (with, without) Big Data adoption

Security & Privacy Risks

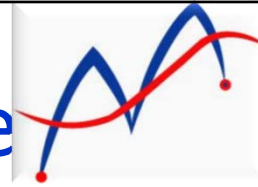
Compliance & Documentation

Audits and Traceability





# Sub-Module



## THE BIG DATA MANIFESTO

THINKING FOR OVERALL INDUSTRY  
RESPONSIBILITY

© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

## The Big Data Manifesto

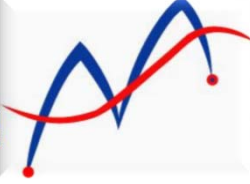
We will use Big Data Analytics and Big Data Technologies to provide Business Value to Users (Customers, Staff and Partners)

We will use Agile Principles to Dynamically create high Quality Business Processes that utilize Big Data and enable Business Agility

We will be Outcome-driven, Fail-fast, Collaborative, welcoming Change and Holistic in applying Big Data to Digital Business so as to Reduce Risks

We will use Big Data with due deference to Security, Privacy and Compliance needs of Individuals, Society and Government

We will embed Environmental and Social Responsibility in Big Data driven Digital Enterprise




# Sub-Module


## THE BDFAB v2.5


---


A HOLISTIC RISK REDUCTION FRAMEWORK


© MethodScience; Material Designed and Presented by Dr. B. Unhelkar


Values 


Roles 

Building Blocks 

Artefacts 


Conditions 

Practices 

Compendium 

*Agile Business Values [7]*

*Key roles (Technical, Business) in Big Data & Agile [10]*



*Inputs and Outputs (Iterative Deliverables) [10]*

*Business Parameters (type, size, maturity)*

*Composite Agile (CAMS) Practices [13]*

X 12 Laned Adoption (Transformation) Process

Big Data Strategy Cube (The Opportunity Matrix)

Big Data Manifesto: x 5 Statements

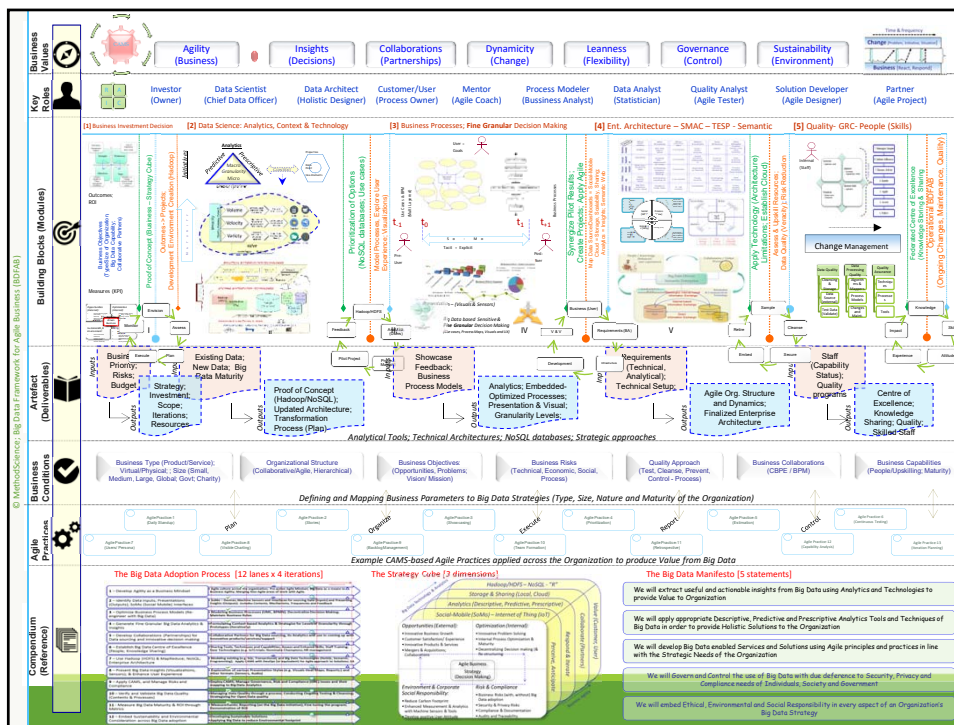
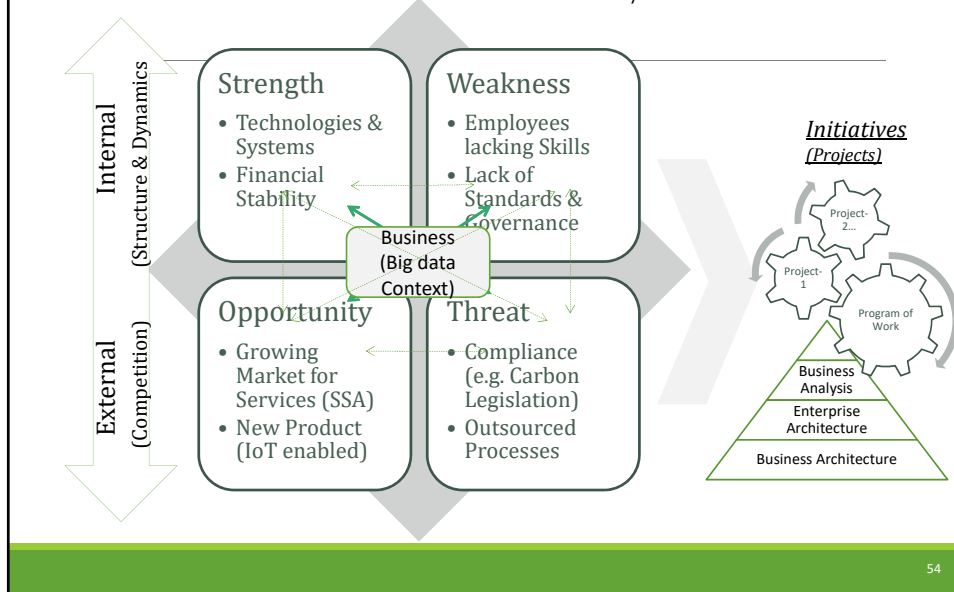


Figure 2.3: SWOT Analysis (Example) of a Business Organization in the context of Big Data – and resultant Projects (Activities) with the backdrop of Business Architecture, Enterprise Architecture and Business Analysis



**Figure 4.12: Transforming to Big Data Driven Agile Business: The BDFAB Adoption Roadmap - 12 Lanes and 4 Iterations**  
**[Editor – Landscape please]**

	Iteration-1	Iteration-2	Iteration-3	Iteration-4
<b>1 – Develop Agility as a Business Mindset</b>	Agile culture across the organization; Pro-active Agile Mindset; Big Data as a means to Business Agility; Merging Non-Agile areas of work with Agile.			
<b>2 – Identify Data Inputs &amp; Outputs; SoMo (Social Mobile) Interfaces</b>	SoMo – Devices, Machine Sensors and Interfaces for sourcing Data (Inputs) and Presenting insights (Outputs). Includes Contents, Mechanisms, Frequencies and Feedback			
<b>3 – Optimize Business Process Models (Re-engineer with Big Data)</b>	Modeling Business Processes (UML, BPMN); Decentralize Decision Making; Maintain Business Rules			
<b>4 – Generate Fine Granular Big Data Analytics &amp; Insights</b>	Formulating Context-based Analytics & Strategize for Levels of Granularity through Prototypes (iteratively)			
<b>5 – Develop Collaborations (Partnerships) for Data sourcing and innovative decision making</b>	Collaborative Partners for Big Data sourcing, its Analytics and use in coming up with Innovative products / services / support			
<b>6 – Establish Big Data Centre of Excellence (People; Knowledge Sharing)</b>	Sharing Tools, Techniques and Capabilities; Assess and Enhance Skills; Staff Training; New Technologies (e.g. IoT) trials; Nominate Champions; HR management			
<b>7 – Use Hadoop (HDFS) &amp; MapReduce; NoSQL; Enterprise Architecture</b>	Modeling existing (e.g. SQL, Transactional) and Big Data Technologies (NoSQL, Semantic, Programming); Apply CAMS with DevOps (or equivalent) for Agile approach to Solutions; EA			
<b>8 – Present Big Data Insights (Visualizations, Sensors); &amp; Enhance User Experience</b>	Exploration of various Presentation Styles (e.g. Visuals Heat Maps; Reports;) and other formats (Sensors, Audio)			
<b>9 – Apply CAMS, and Manage Risks and Compliance</b>	Deploy CAMS; Manage Governance, Risk and Compliance (GRC) issues and their mapping to Big Data Analytics			
<b>10 – Verify and Validate Big Data Quality (Contents &amp; Processes)</b>	Managing Data Quality through a process; Conducting Ongoing Testing & Cleansing; Strategizing for Open Data quality			
<b>11 – Measure Big Data Maturity &amp; ROI through Metrics</b>	Measurements; Reporting (on the Big Data initiative); Fine tuning the program; Demonstration of ROI			
<b>12 – Embed Sustainability and Environmental Consideration across Big Data adoption</b>	Developing Sustainable Solutions and Applying Big Data to reduce Environmental footprint			



# Sub-Module

## A LOOK AT THE TOOLS

---

### IN THE BIG DATA SPACE

© MethodScience; Material Designed and Presented by Dr. B. Unhelkar

## Types of tools used in Big-Data

Where processing is **hosted**?

- Distributed Servers / Cloud (e.g. Amazon EC2)

Where data is **stored**?

- Distributed Storage (e.g. Amazon S3)

What is the **programming model**?

- Distributed Processing (e.g. MapReduce)

How data is **stored & indexed**?

- High-performance schema-free aggregated No-SQL databases (e.g. MongoDB)

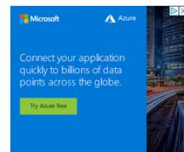
What operations are performed on data?

- Analytic / Semantic Processing

<https://www.guru99.com/big-data-tools.html>

GURU99 Home Testing SAP Web Must Learn Big Data

### Top 15 Big Data Tools in 2018



Today's market is flooded with an array of Big Data tools. They bring cost efficiency, better time management into the data analytical tasks. Here are some important big data tools with their key features and download links.

#### 1) Hadoop:



The **Apache Hadoop** software library is a big data framework. It allows distributed processing of large data sets across clusters of computers. It is designed to scale up from single servers to thousands of machines.

#### Features:

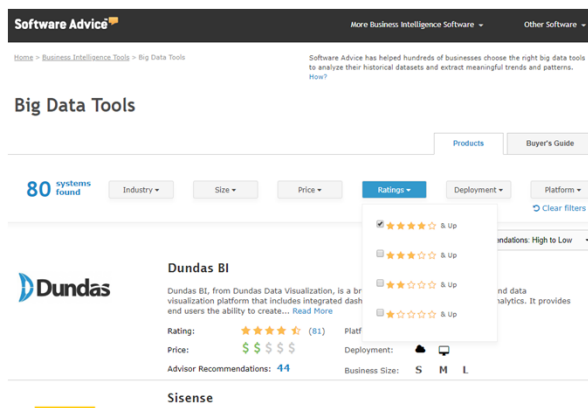
- Authentication improvements when using HTTP proxy server
- Specification for Hadoop Compatible Filesystem effort
- Support for POSIX-style filesystem extended attributes
- It offers robust ecosystem that is well suited to meet the analytical needs of developer
- It brings Flexibility In Data Processing
- It allows for faster data Processing

<https://www.whizlabs.com/blog/big-data-tools/>



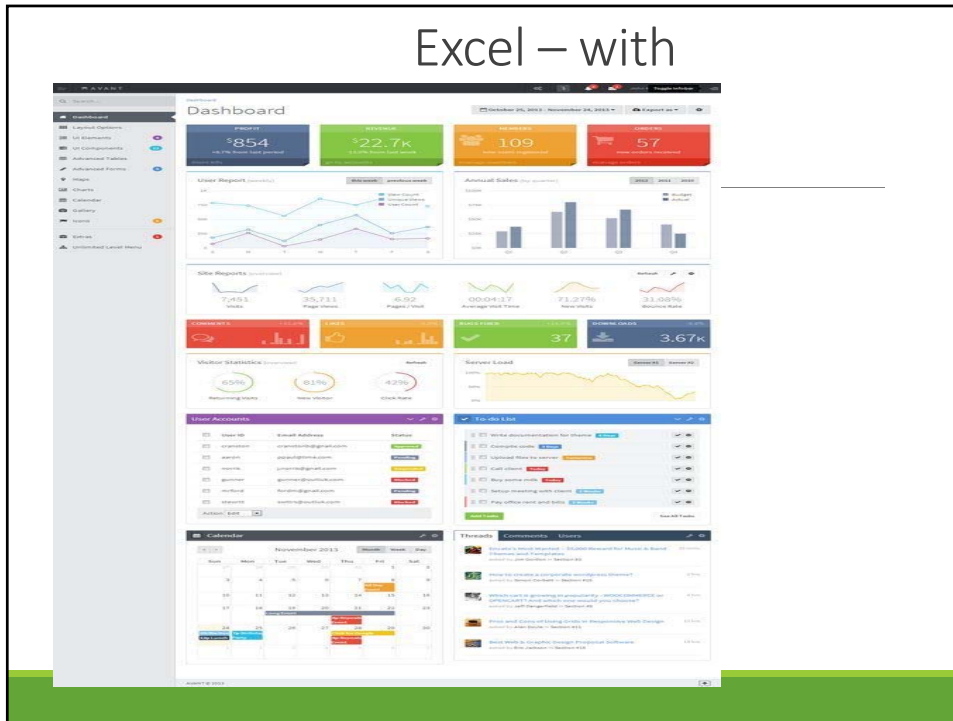
The screenshot shows a blog post from Whizlabs. The header includes the Whizlabs logo and navigation links for 'CLOUD COMPUTING', 'BIG DATA', and 'PROJECT MAN'. The main content features a large graphic with 'TOP 10' in a laurel wreath, followed by the title 'Open Source Big Data Tools in 2018'. Below the title, it says 'By Amit Verma - March 12, 2018' and includes social media icons for Facebook, Twitter, Google+, LinkedIn, and Pinterest. The introductory text states: 'Today almost every organization extensively uses big data to achieve the competitive edge in the market. With this in mind, open source big data tools for big data processing and analysis are the most useful choice of organizations considering the cost and other benefits. Hadoop is the top open source project and the big data bandwagon roller in the industry. However, it is not the end! There are plenty of other'.

<https://www.softwareadvice.com/bi/big-data-comparison/>

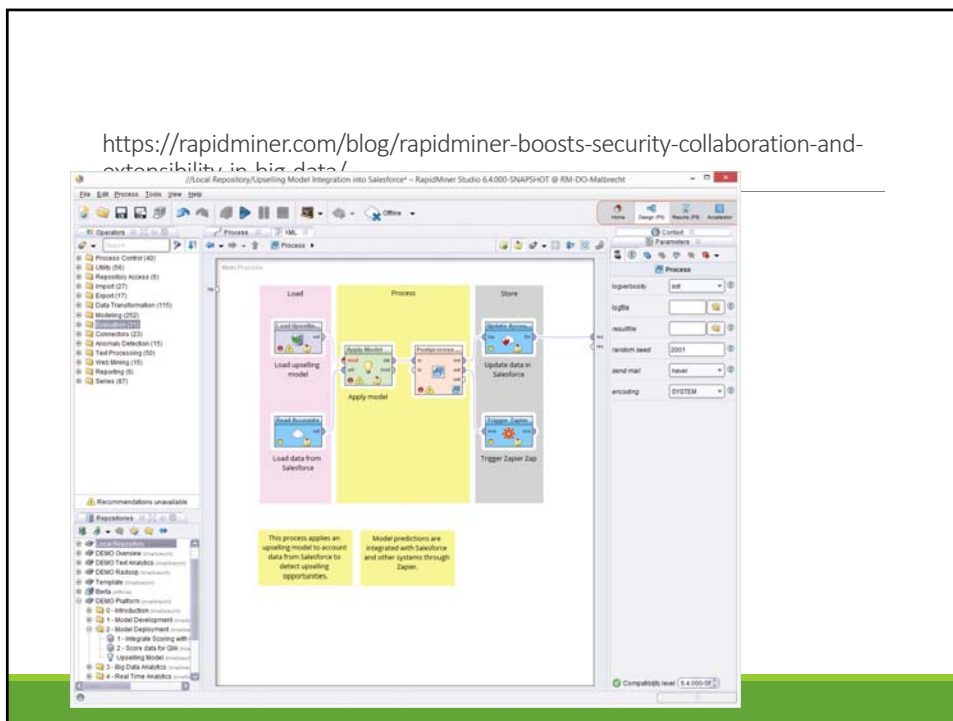


The screenshot displays the Software Advice website interface for 'Big Data Tools'. It features a navigation bar with 'Software Advice' and links for 'More Business Intelligence Software' and 'Other Software'. Below the navigation, there are filters for 'Industry', 'Size', 'Price', 'Ratings', 'Deployment', and 'Platform'. A search bar indicates '80 systems found'. The main content area shows a list of tools, with 'Dundas BI' highlighted. The Dundas BI entry includes a description: 'Dundas BI, from Dundas Data Visualization, is a business visualization platform that includes integrated dashboards and users the ability to create... Read More'. It also displays a rating of 4.5 stars (81 reviews), a price range of '\$ \$ \$ \$ \$', and deployment options for 'On-premise' and 'Cloud'. Below the Dundas BI entry, the 'Sisense' entry is partially visible.

# Excel – with




<https://rapidminer.com/blog/rapidminer-boosts-security-collaboration-and-extendibility-in-big-data/>



Method Science

Discussion, Q & A



A. The imperative in Big Data is not “How to” compute but “What to compute?”: Please Agree/Disagree; Let’s further add “Why to Analyze?”


B. Your views on Strategic use of Big Data!  
(Especially in the Context of your Organization)

The Art of Agile Practice  
A Research Report for Projects and Operations  
Bhuvan Umhalkar

Big Data Strategies for Agile Business  
BIG DATA  
IoT  
AGILE BUSINESS  
Bhuvan Umhalkar

## KEY POINTS OF THE SESSION

- UNDERSTANDING THE BIG DATA FRAMEWORK FOR AGILE BUSINESS (BDFAB V2.5)
  - As a Holistic Framework for Big Data adoption in a Business
  - As a Risk reduction mechanism
- CONSIDERING THE KEY ELEMENTS OF THE BDFAB:
  - Starting with VALUE





## Conclusions & Future Directions

- Focus of Big Data is Business VALUE!
- Business is interested in AGILITY
  - 
  - Accuracy and Rapidity of Decision making
- Refine BDFAB to 2.5 and make it tool-enabled



*Do you see what  
we see?*