Big Data and Business Analytics: Accelerating Digital Transformation in Enterprises and Industries

James PANG, Ph.D.
Visiting Associate Professor
School of Computing
National University of Singapore

Jorge SANZ, Ph.D.
Director Business Analytics Center
School of Computing – Business School
National University of Singapore

Agenda

1 Business Analytics Solution Landscape

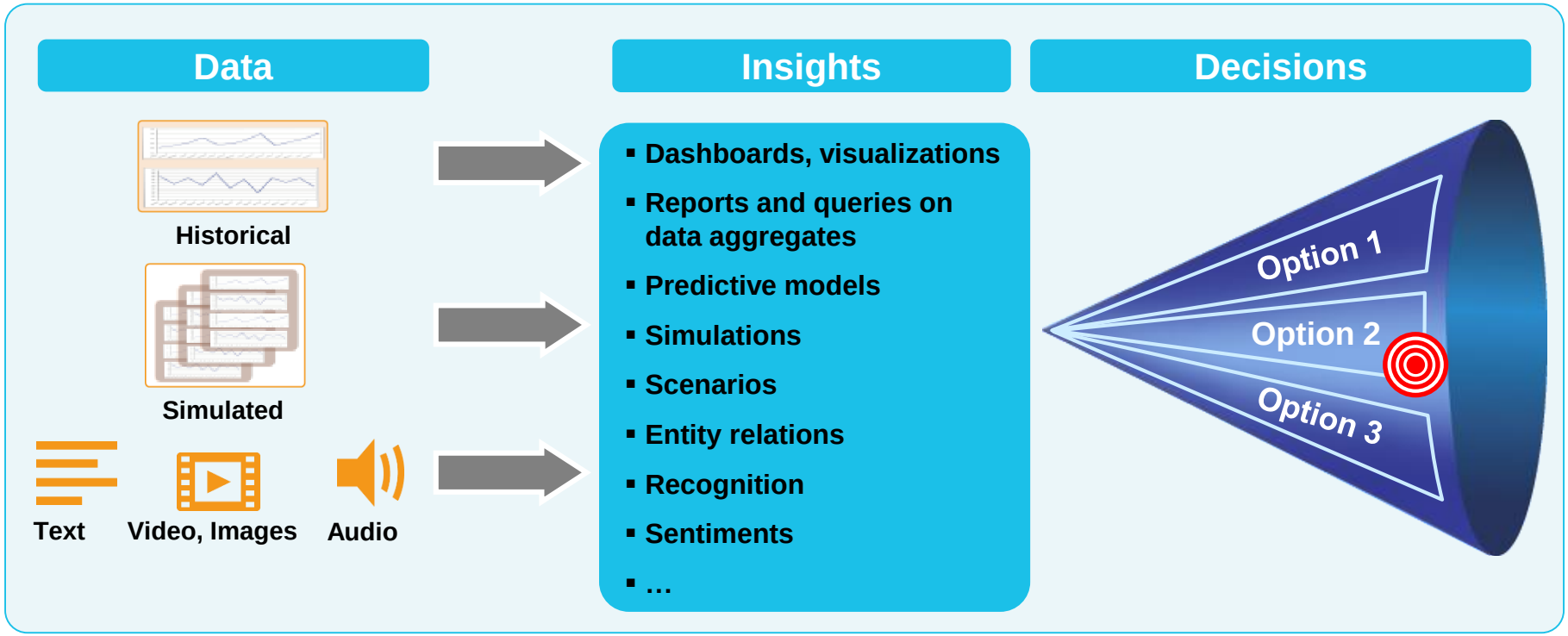
2 Big Data, Big Opportunity

3 Future Outlook

4 NUS Business Analytics Center Projects

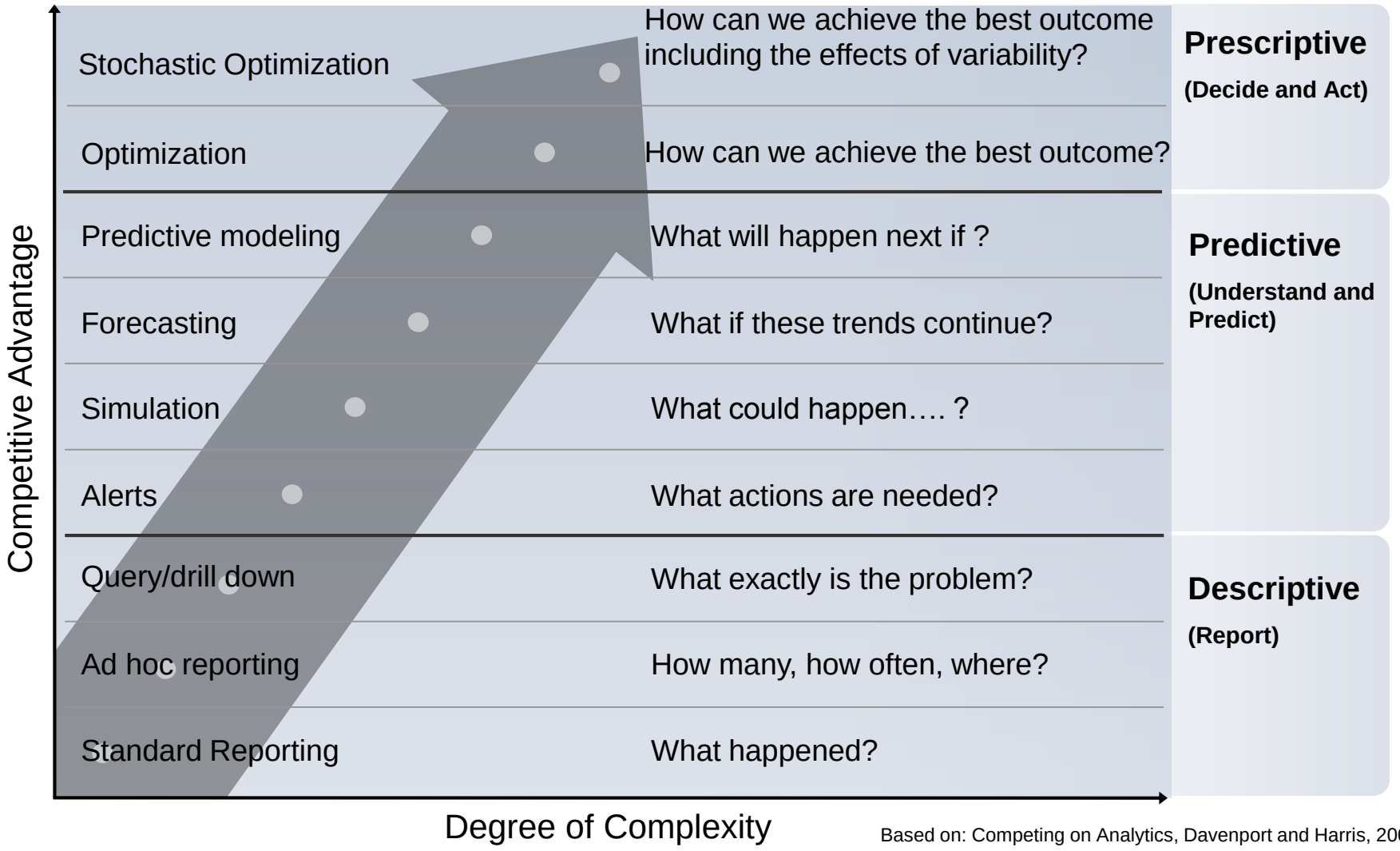
What is Analytics?

Analytics is broadly defined as the process of deriving **insight** from **data** in order to make better **decisions**



Use case	Data	Insight	Action/Decision
Pricing optimization	Past sales, price levels, etc	Predict sales level from price	Set prices in order to maximize profit/revenue
Outcome Based Management	Social data	Social context for each case	Select best social program

Traditional Analytics Landscape



Based on: Competing on Analytics, Davenport and Harris, 2007

Dimensions of Analytics

Data Sources
– Internal data
– External data

Maturity Level
– Descriptive, predictive, prescriptive, etc.

Insight
– Trends, patterns, sentiments, etc.

Decision Automation
– Decision support vs Decision automation

Analytics Process
– Data driven, human driven, event driven

Genericity
– Industry specific vs Cross industry

Solution Granularity
– Component, asset, product, solution

Usage
– Inward/IT vs Outward/LOB

Computation Complexity
– Simple (e.g. reporting), complex (predictive, optimization), very complex (NLP + machine learning + scoring, etc.)

Users
– Business user, Data scientist, Business Analyst/Consultant, Application developer

* NLP – Natural Language Processing
* LOB – Line of Business

Select the Right Analytics Techniques for the Use Cases

- The five common used business analytics techniques are show below:

Visualization

Allows users to visually gain insight from data and modeling results, e.g. data exploration, dashboard, BI reports

Advanced Statistics & Machine Learning

Transform complex data into business insight through identifying the data patterns and trends

Optimization

Mathematically finds the best solution to complex problems with many decision options and constraints

Simulation

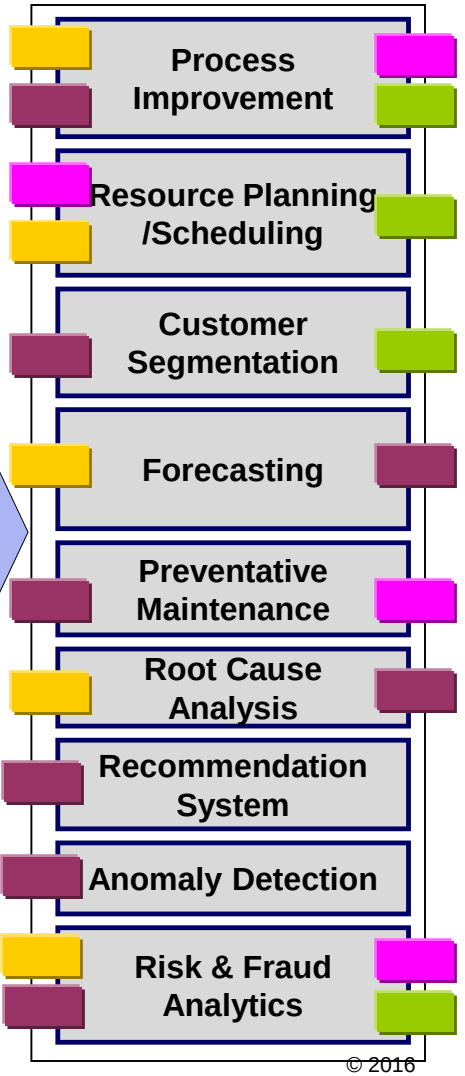
Simulate and analyze a system, process, behaviour, or business problem

Rule Engine

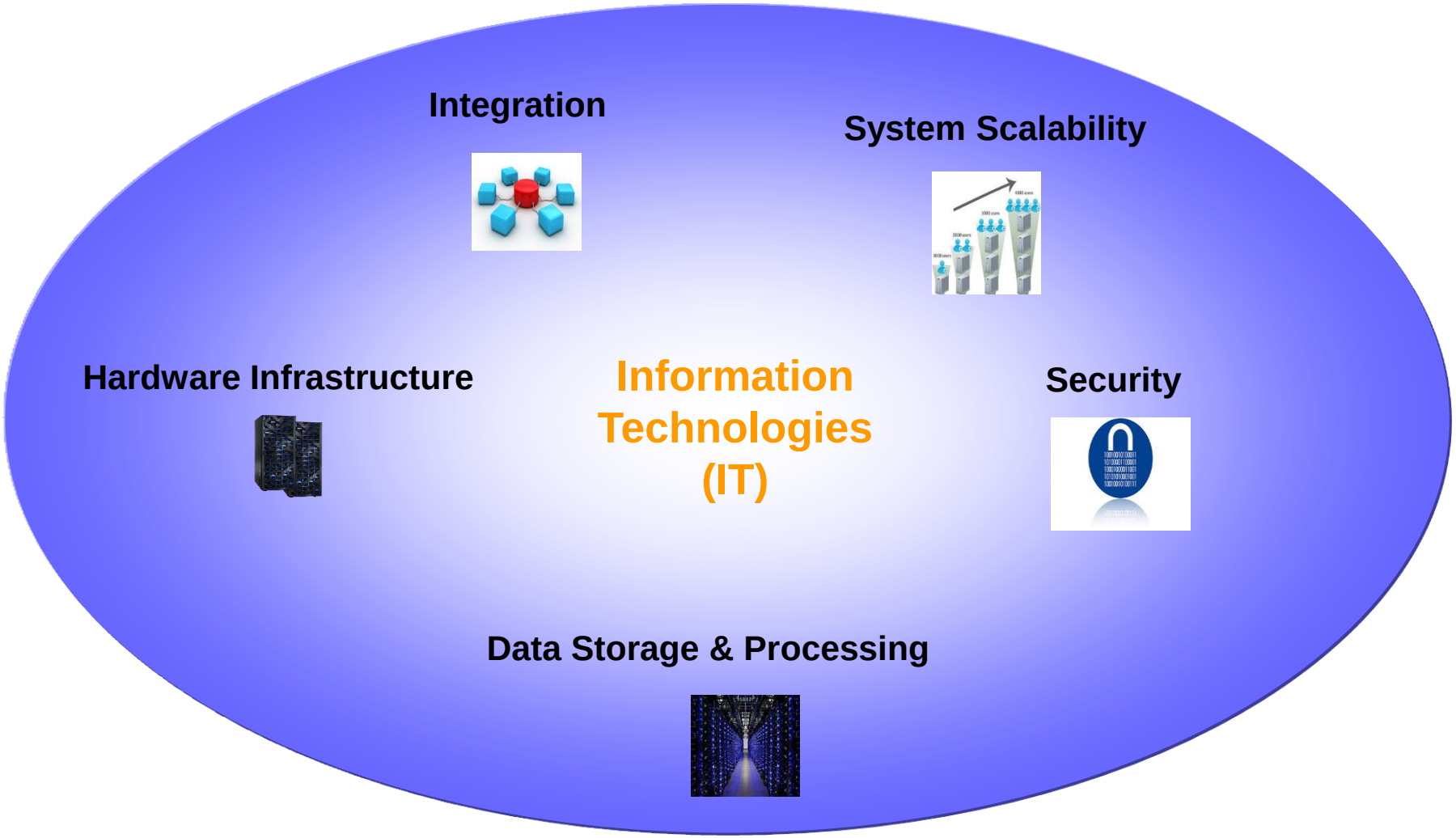
Creates, manages and executes business rules in robust repositories, for use across wide applications

Commonly used for following purposes

Typical Analytics Use Cases



Only Analytics Techniques are Not Enough



3 Key Elements of a Successful Business Analytics Solution

Industry Domain Experts

- Provide industry domain knowledge and experience to ensure the analytics models representing the real world requirement
- Analyze and confirm the analytics results

- Mathematical analysis expertise
- Transform business requirements into mathematical requirements
- Build the analytics models and algorithms

Analytics Experts

- IT Architecture design
- Applications development
- Data storage and processing

IT Experts

Build the Business Analytics Solution with 3 Key Elements

Industry Knowledge

+

Analytics

+

IT

Industry Analytics Applications

Finance / Banking | **Retail / E-Commerce** | **Healthcare** | **Telecom** | **Manufacturing** | **...**

Analytics Core Services

Advance Statistics and Machine Learning | Optimization | Simulation | Rule Engine | Visualization

Data Storage and Processing

Hadoop System 

Stream Computing 

Enterprise Data Warehouse 

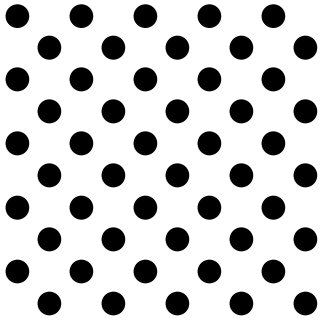
Information Integration, Security & Governance

Hardware Optimization
(Compute-Centric Model -> Data Centric Model)

A New Big Data Era

- **90%** of the world's data is created in the last two years
- **80%** of the world's data today is unstructured
- **1 Trillian** connected devices generate 2.5 quintillion bytes data / day

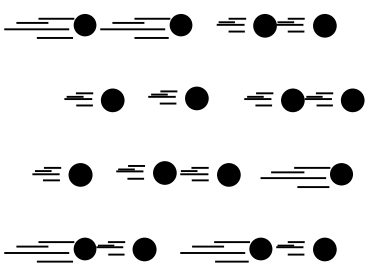
Volume



Data at Scale

Terabytes to exabytes of existing data to process (e.g. CRM, ERP data, etc.)

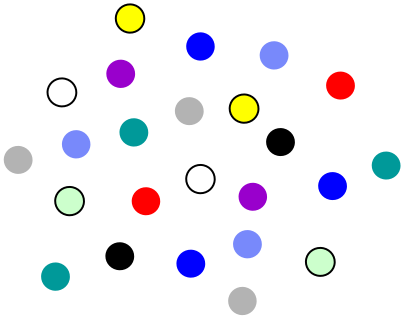
Velocity



Data in Motion

Streaming data, milliseconds to seconds to respond (e.g. data from smart sensors, mobile device, etc.)

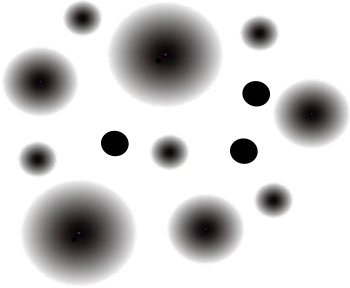
Variety



Data in Many Forms

Structured, unstructured, text, multimedia (e.g. rational DB, images, free text, video, etc.)

Veracity



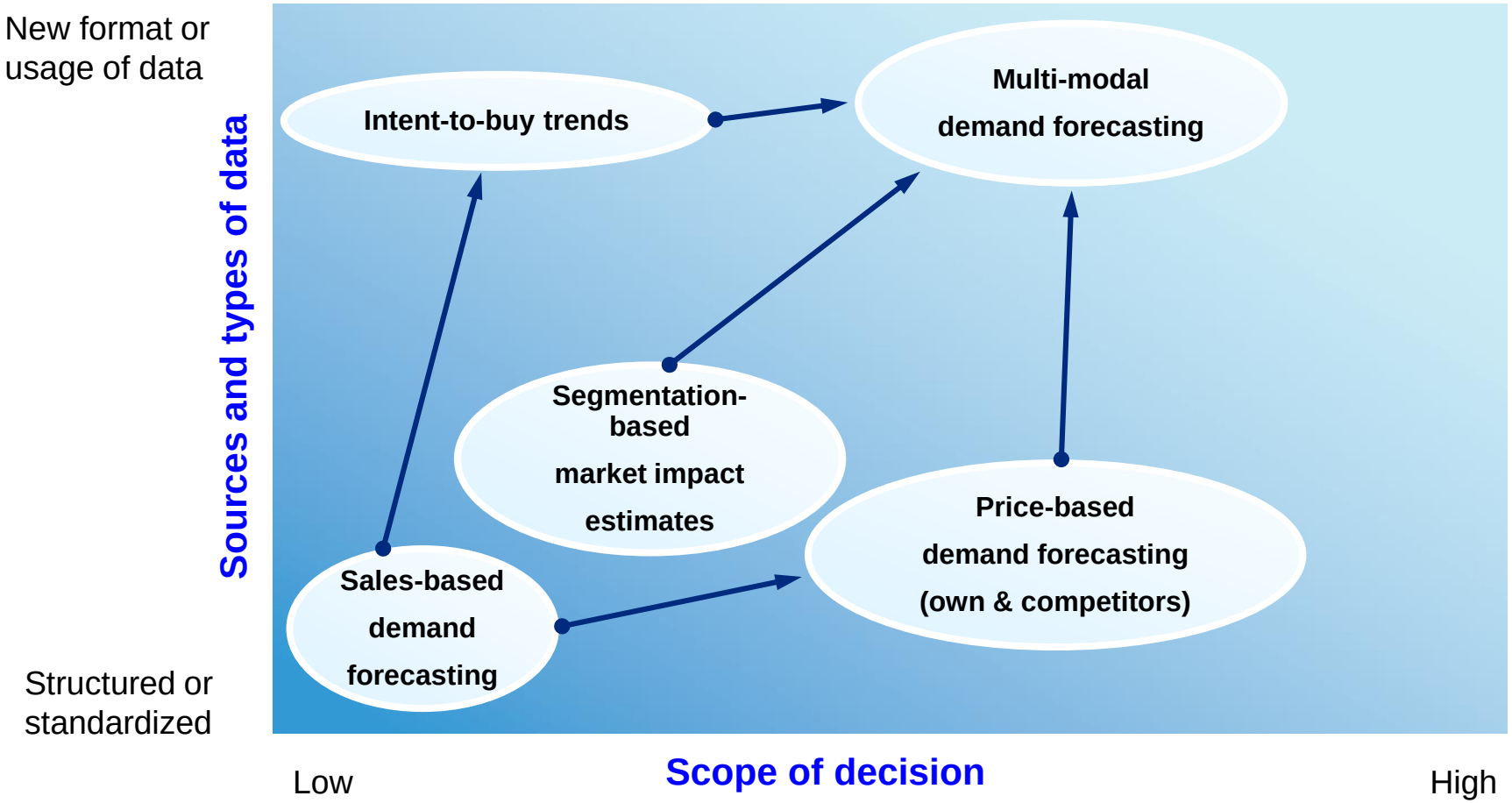
Data Uncertainty

Uncertainty due to data error, inconsistency & incompleteness, ambiguities, model approximations (e.g. manual errors, device errors, models errors, etc.)

Source: IBM GTO 2012

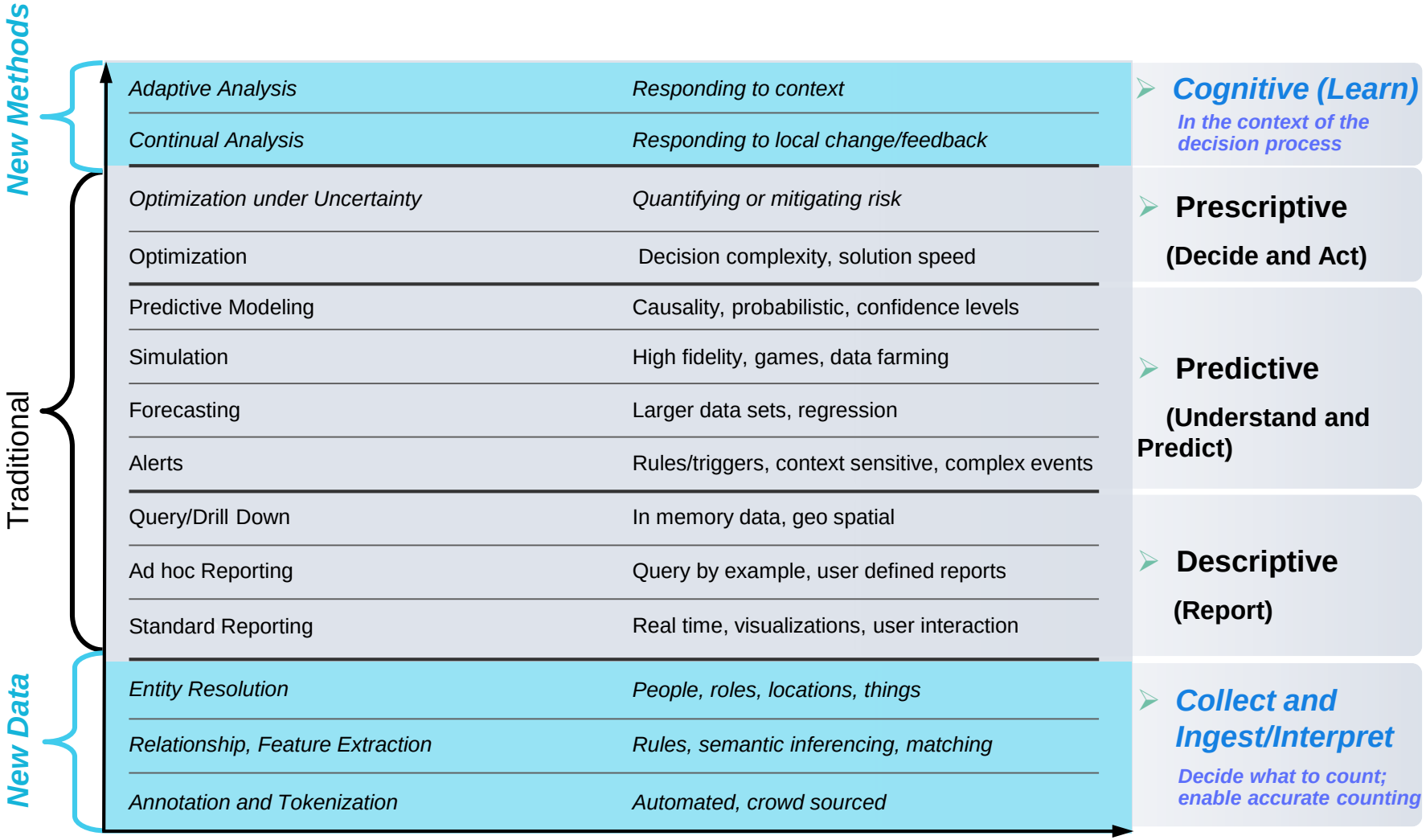
Big Data Brings New Opportunity and Value

The value of analytics grows by incorporating **new sources of data**, composing a variety of **analytic techniques**, spanning organizational silos, and enabling iterative, user-driven interaction



Source: IBM GTO 2012

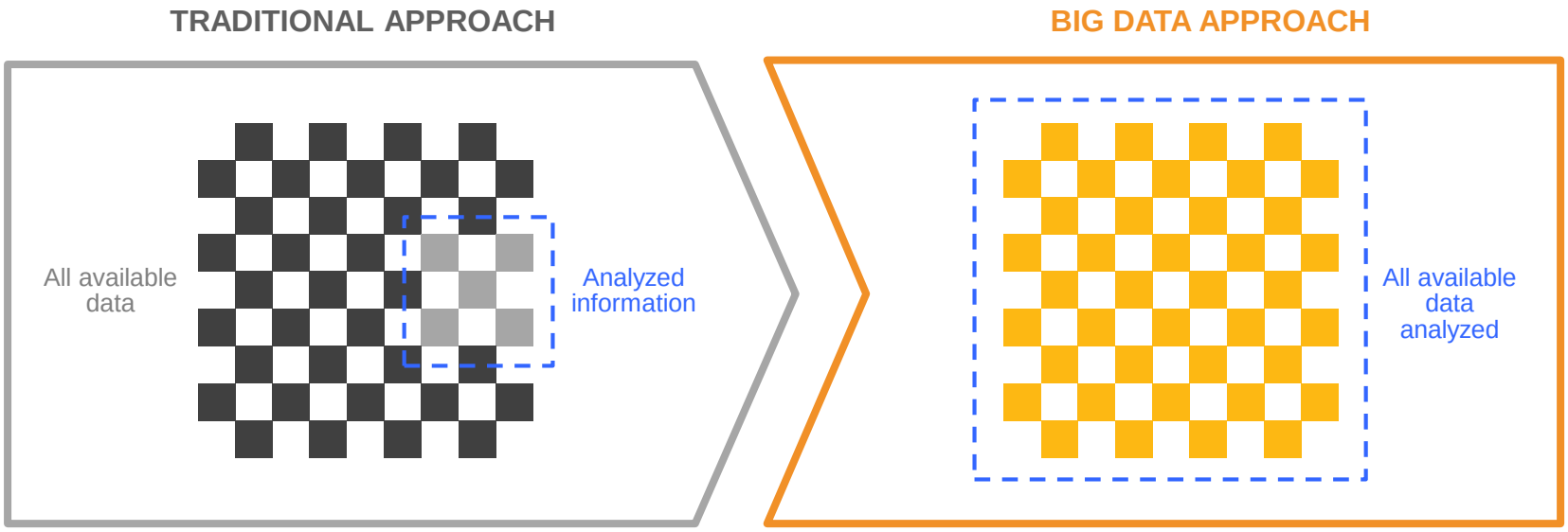
Big Data Requires to Expand the Analytics Landscape

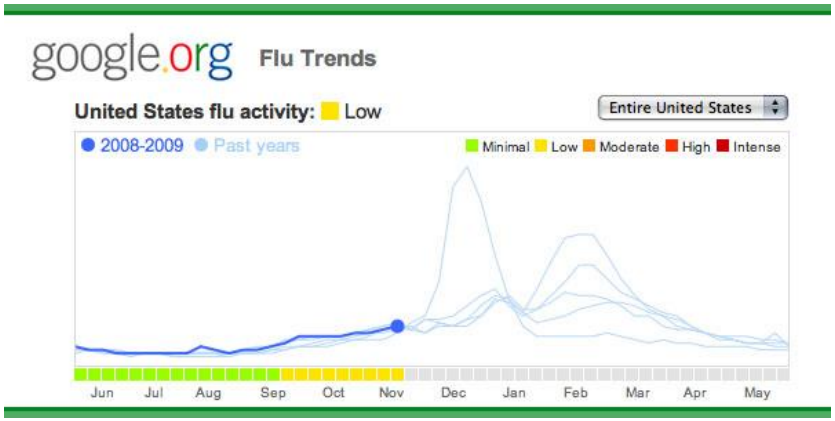


Source: IBM GTO 2012

Analytics Paradigm Shifts Enabled by Big Data

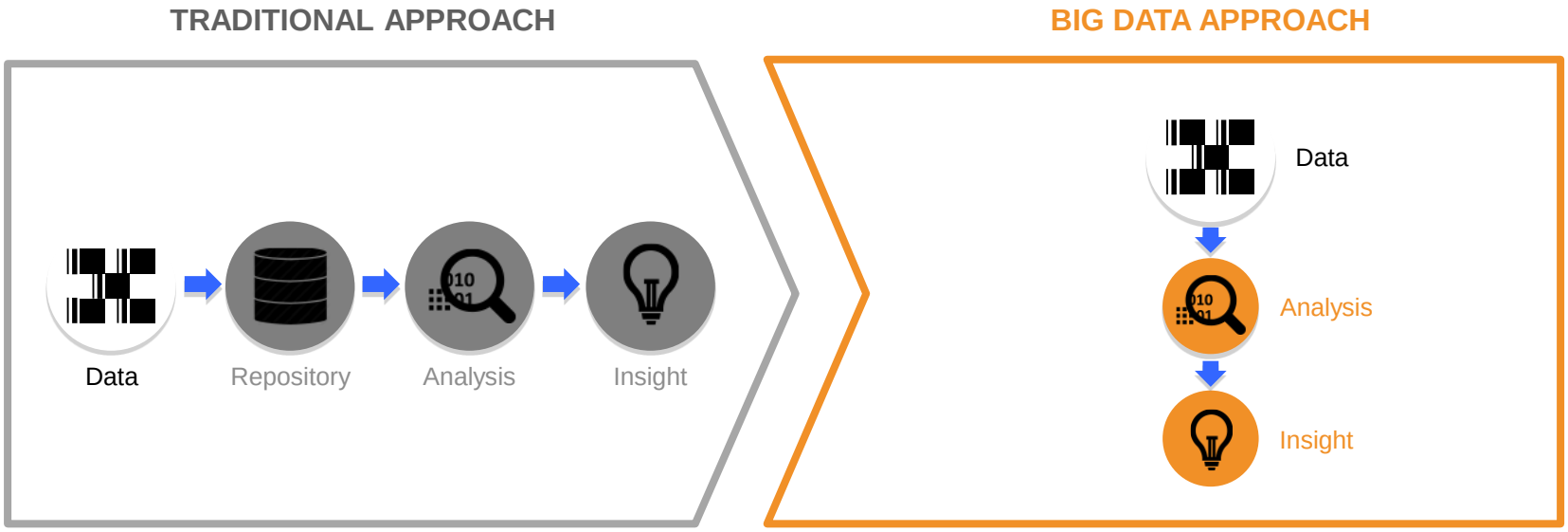
Leverage more of the data being captured





Analytics Paradigm Shifts Enabled by Big Data

Leverage data as it is captured



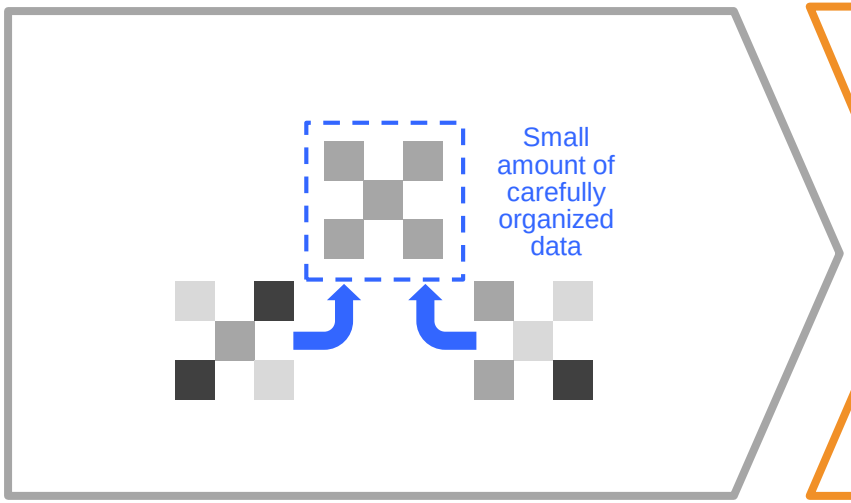
Analyze data ***after*** it's been processed and landed in a data warehouse or data mart

Analyze data ***in motion*** as it's generated, in real-time

Analytics Paradigm Shifts Enabled by Big Data

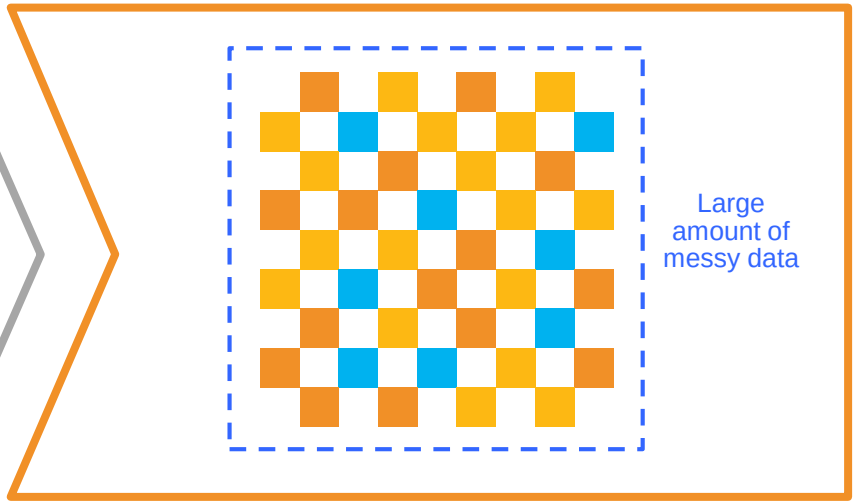
Reduce effort required to leverage data

TRADITIONAL APPROACH

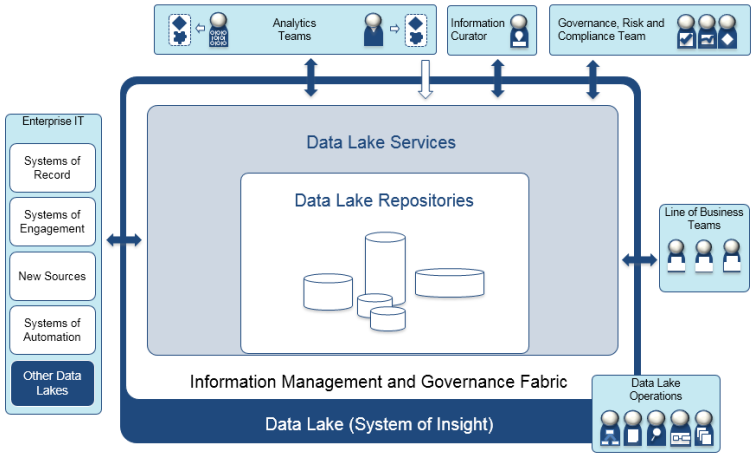


Carefully cleanse data ***before*** any analysis

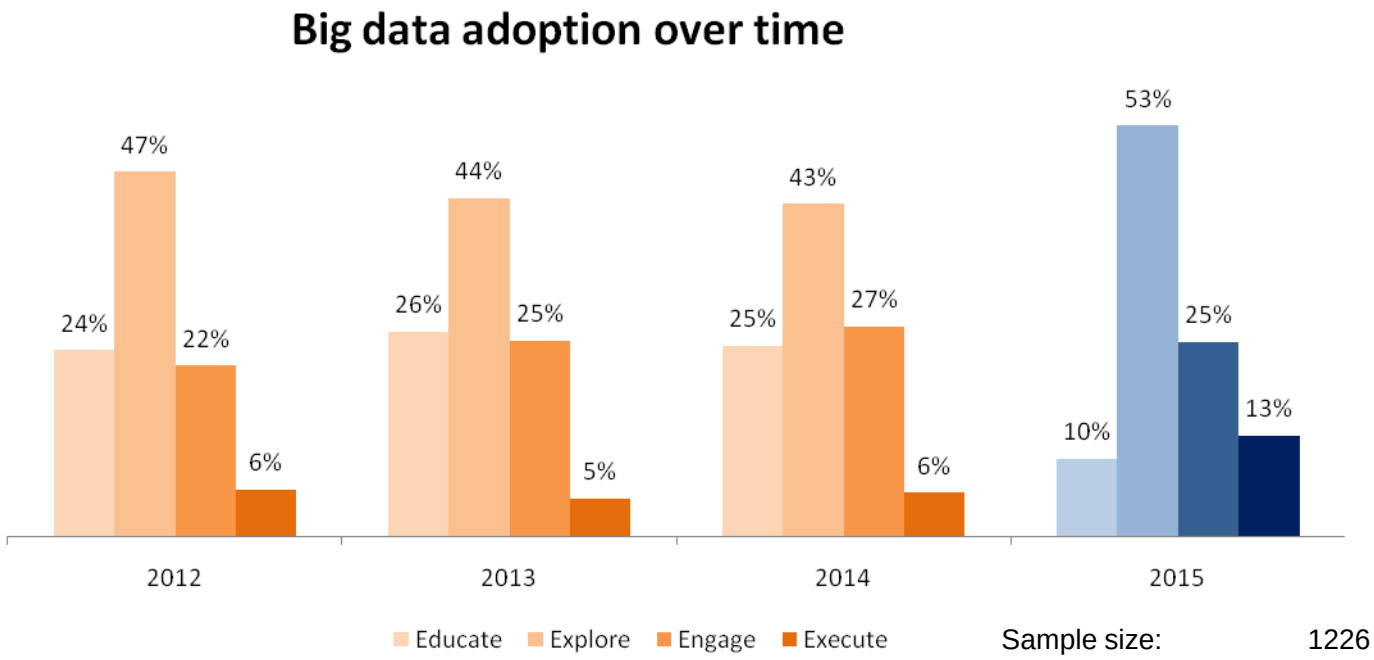
BIG DATA APPROACH



Analyze data as is, cleanse as needed



Organizational Adoption of Big Data and Business Analytics Technologies Has Picked Up Pace …



Educate:
Learning about big data capabilities

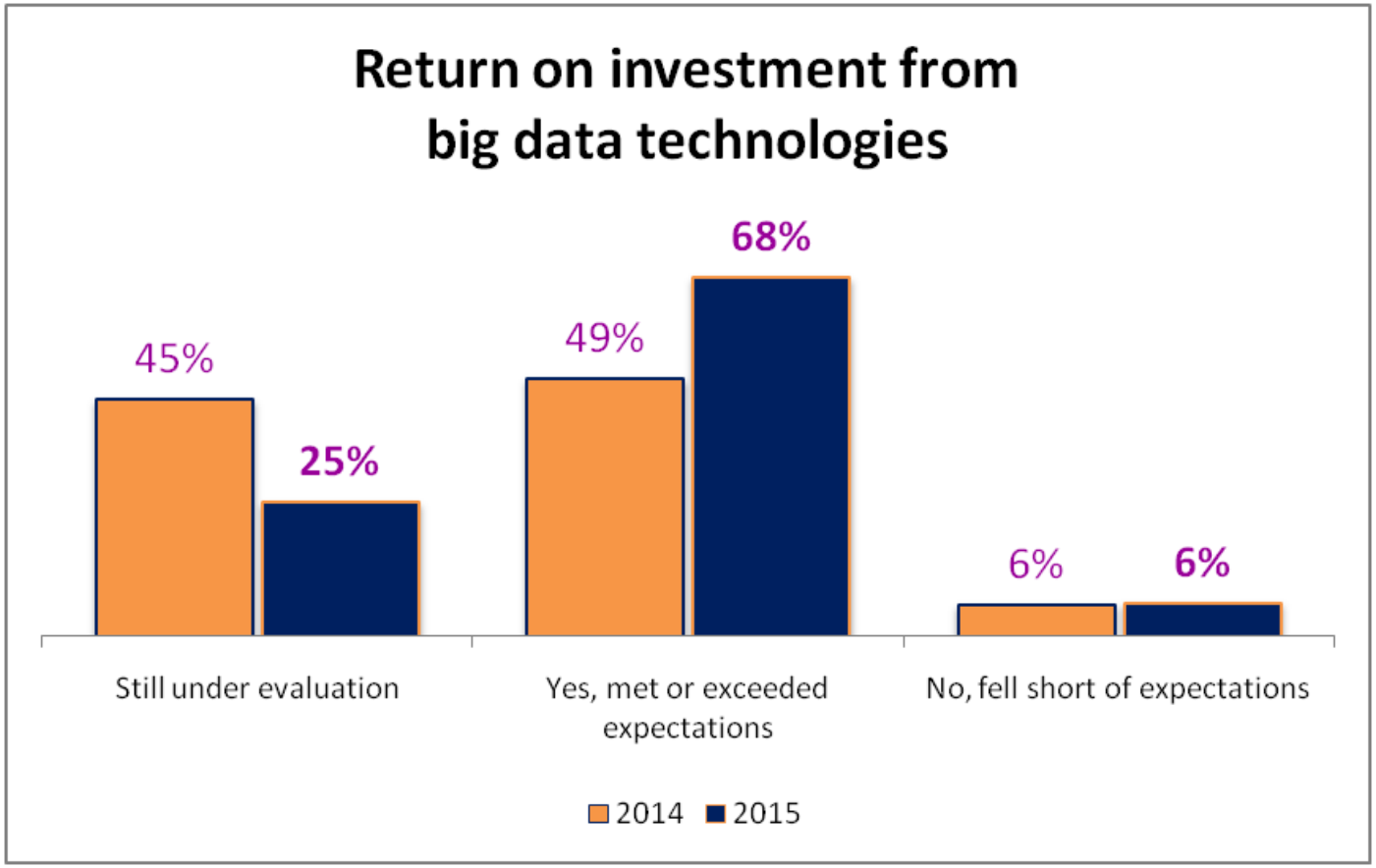
Explore:
Exploring internal use cases and developing a strategy

Engage:
Implementing infrastructure and running pilot activities

Execute:
Using big data and analytics pervasively across the enterprise

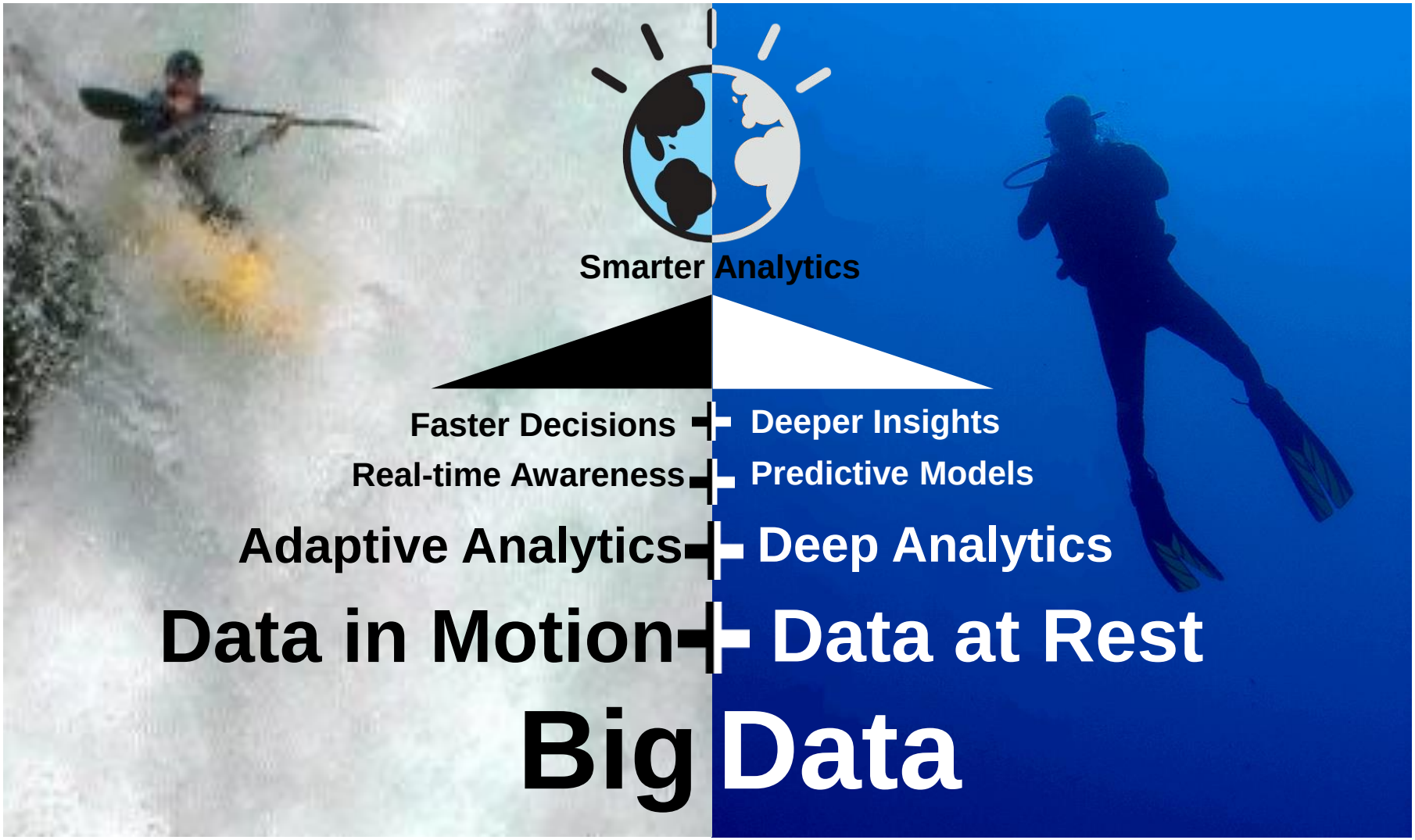
Source: Analytics: The upside of disruption. IBM Institute for Business Value 2015 Analytics research study

… and Once Implemented, Big Data and Business Analytics Technologies Are Paying Off

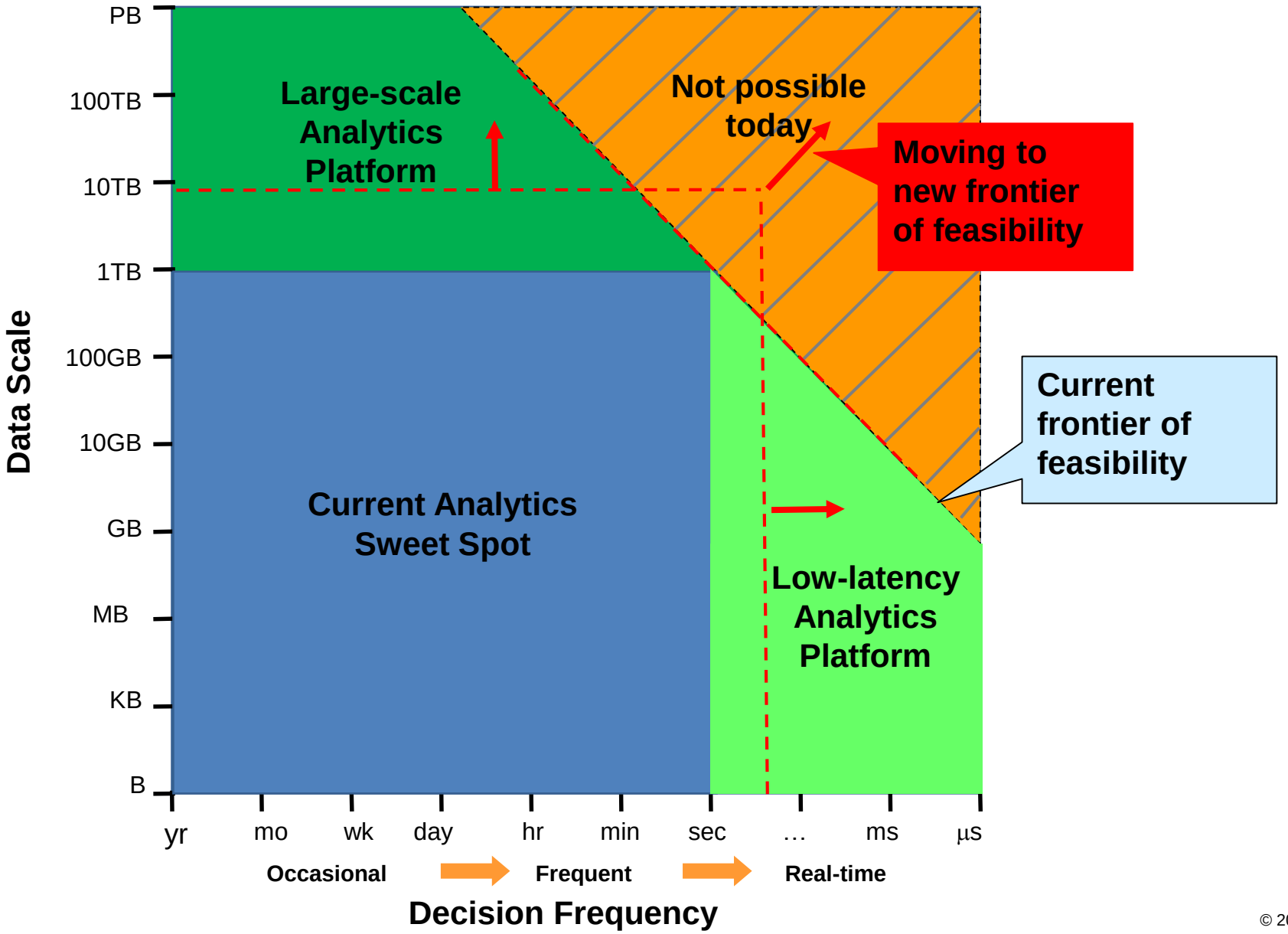


Source: Analytics: The upside of disruption. IBM Institute for Business Value 2015 Analytics research study

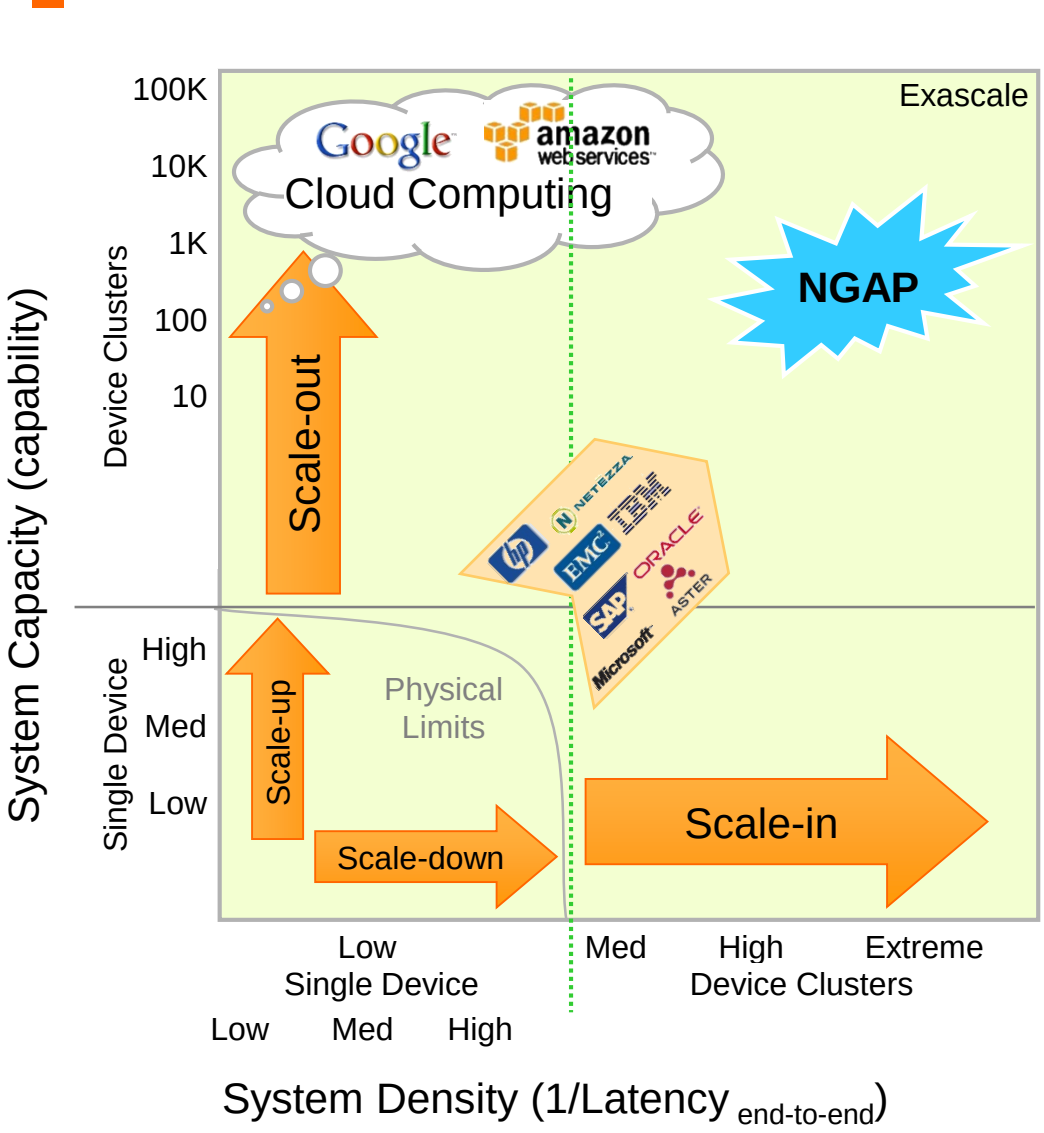
Trend 1: Faster and Deeper Analytics



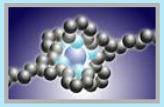
Future Shift of Analytics Frontier



Next Generation Analytics Systems (NGAS)

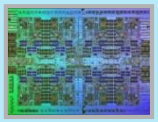


Scale-down
Maximize feature density


Atom Transistor


Atom Storage

Scale-up
Maximize device capacity


Gigabyte HDD

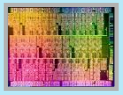

Terabyte HDD

Scale-out
Maximize system capacity


Blade Server


NAS

Scale-in
Maximize system density
Minimize end-to-end latency

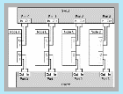

Manycore


FPGA

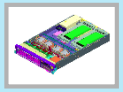

3D Chips


BPRAM/SCM


FLASH SSD


Interconnect


In-mem DB


DAS

Source: IBM GTO Study

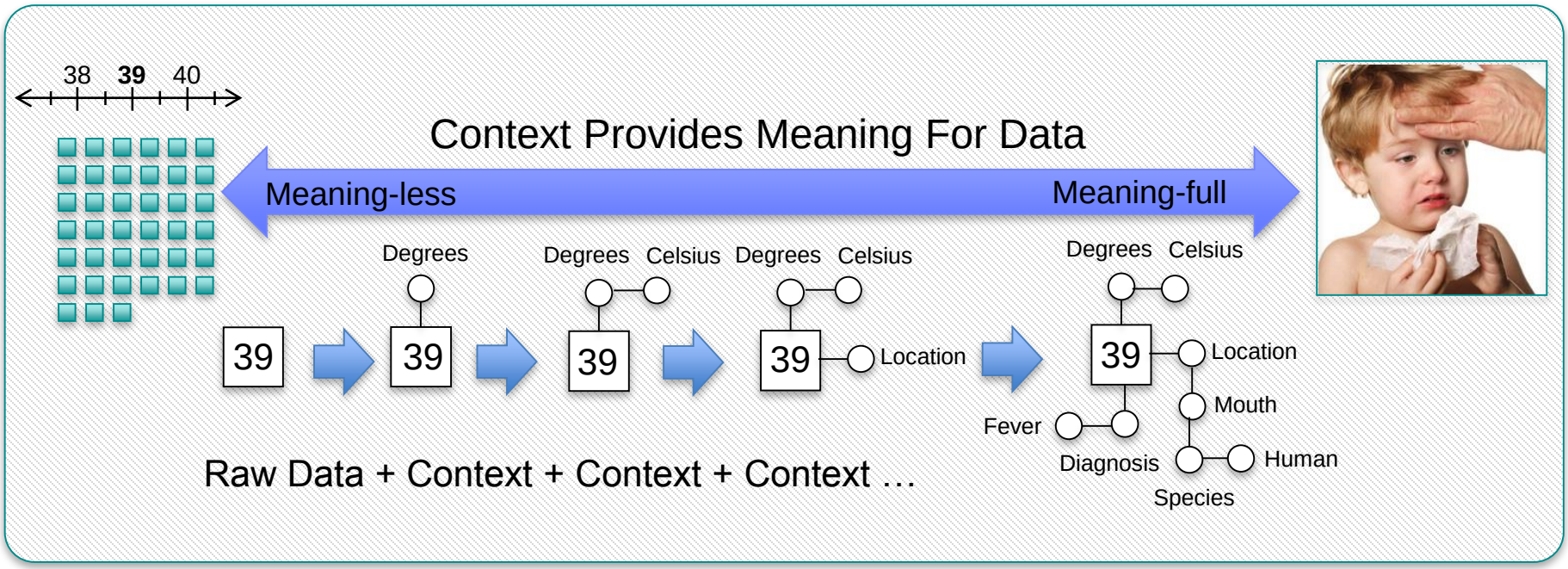
Trend 2: Analytics Responding to Context of Data

Context is what gives Meaning to Data.

Data without Context is Meaningless.

Meaningless Data is Worthless.

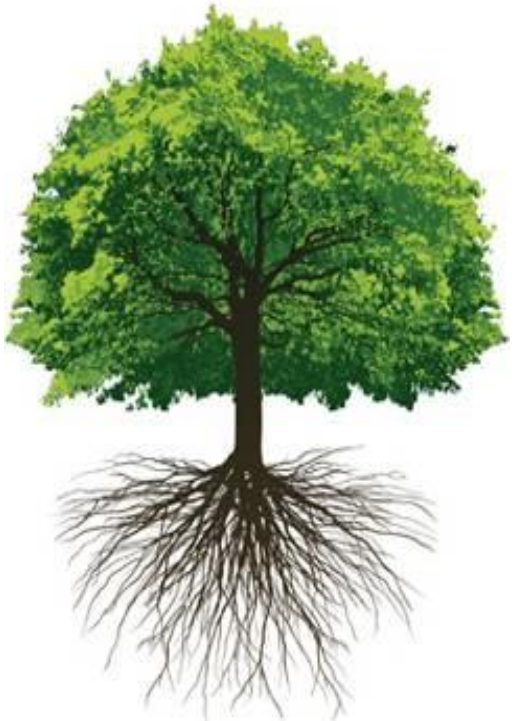
The more Context, the higher the Value.



What does Context Mean for the Future of the Enterprise?

A simple analogy for a complex concept

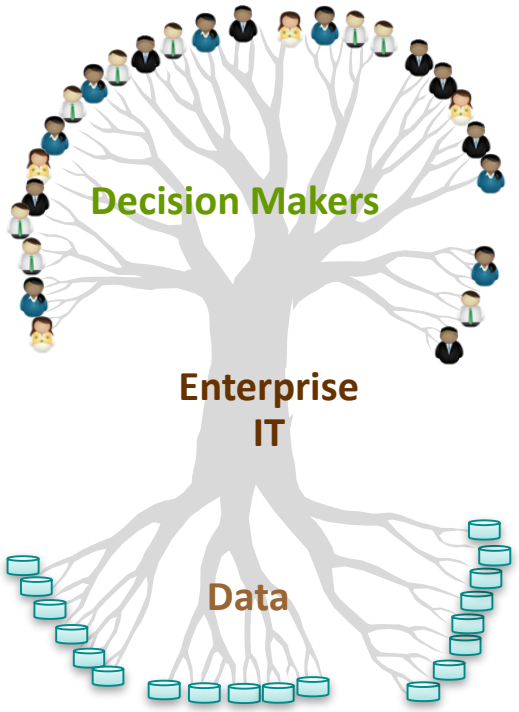
Leaves convert solar energy into sugar to power the organism, driving growth and productivity



Trunks connect and integrate the entire organism

Roots provide water and nutrients required for operation and growth as well as a stable foundation

People power the enterprise by converting business situations into communications, decisions and actions, resulting in **products** and **services**

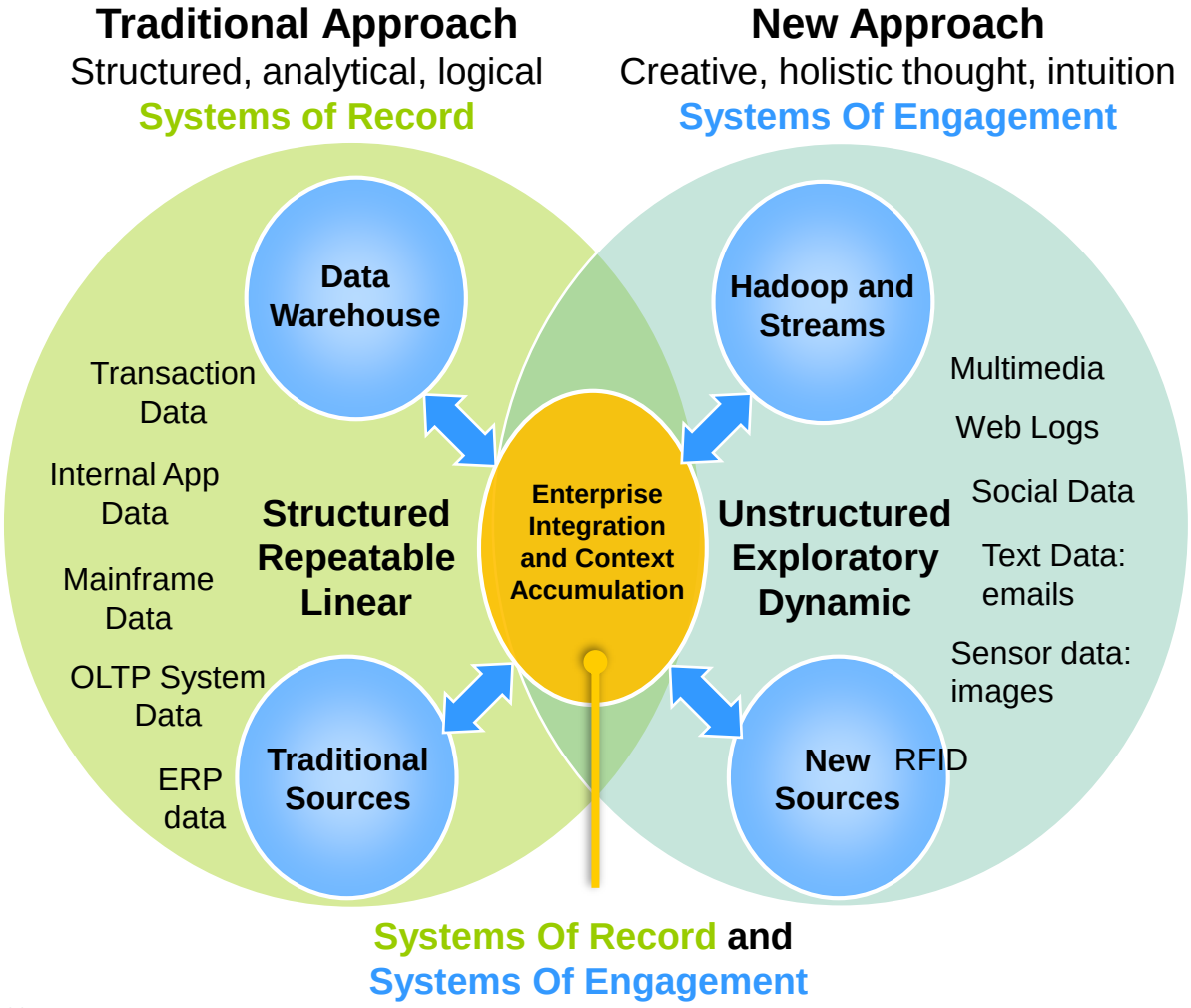


Information Technology connects and integrates the entire enterprise

Data provides the raw materials for insight, and 'grounding' in the reality of the enterprise's environment

The Future Contextual Enterprise

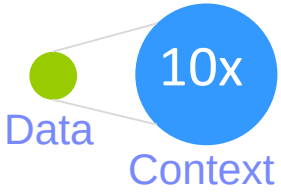
The ability to dynamically build and accumulate context at scale from data sources to deliver new business value



- Big Data and NoSQL / NewSQL tools help to realize business value without the limitation of fixed relational schemas that make existing infrastructures too brittle.
- Big Data enables extracting value and information from things in unanticipated ways.
- Capture relationships fit to purpose for industry domains

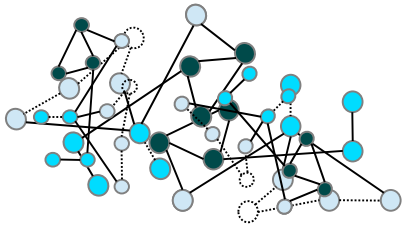
Context at Scale brings Technical Challenges

- Fully contextualized information will require at least **10x the storage** of raw data.



New Storage Requirements

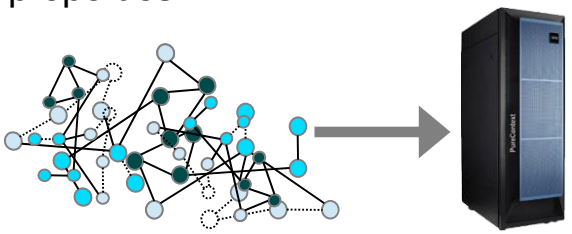
- Continual ingestion and curation will require **continual deep analytics** to discover new insights.
- **Dynamic schema** requirements and temporality will drive new database requirements.



Unique Data Requirements

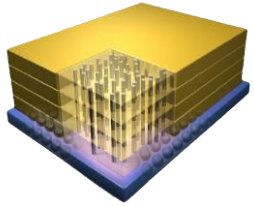
- **High dimensional data** with dynamic content and relations lead to irregular graphs that are notoriously hard to partition and favor large in-memory systems.

- Context never rests but data grows continuously with **streaming high velocity input**. Highly dynamic graphs will require ultra scalable data-structures, that support local graph traversal and at the same time representation/querying of global properties.



Vertically Integrated Systems

- Context size, dynamics, and access patterns will require **data-centric, scale-in, highly integrated systems**.

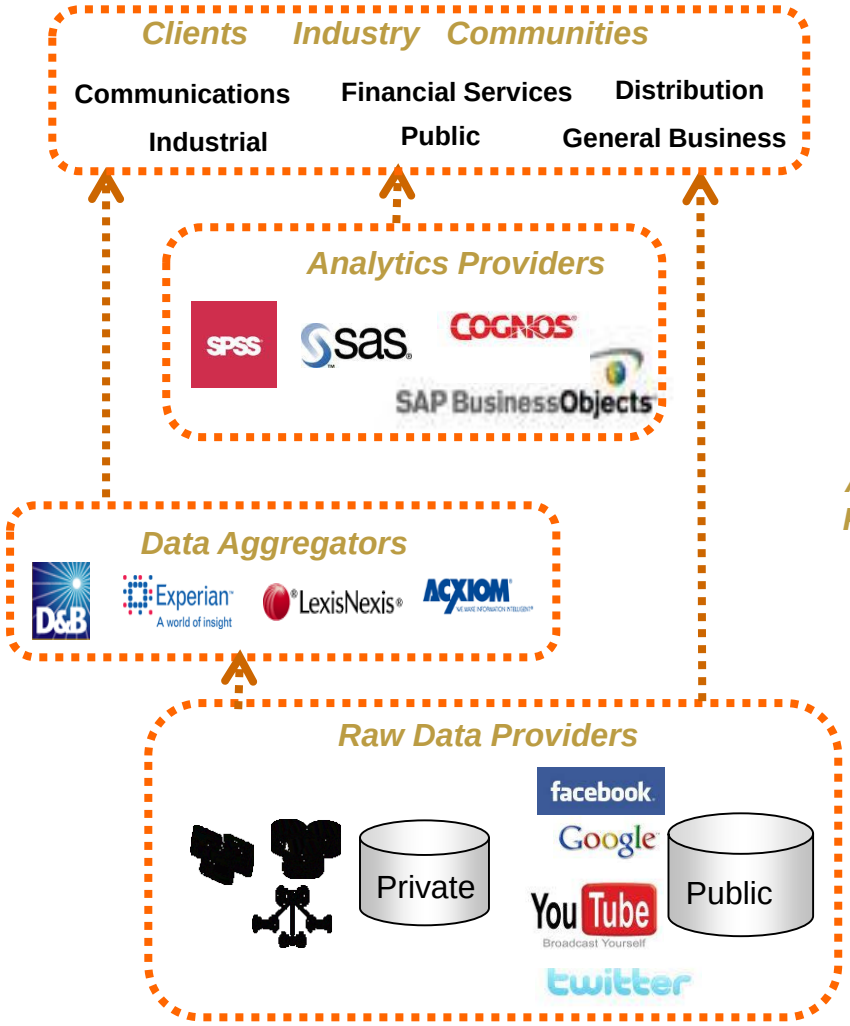


Hybrid Memory Cube

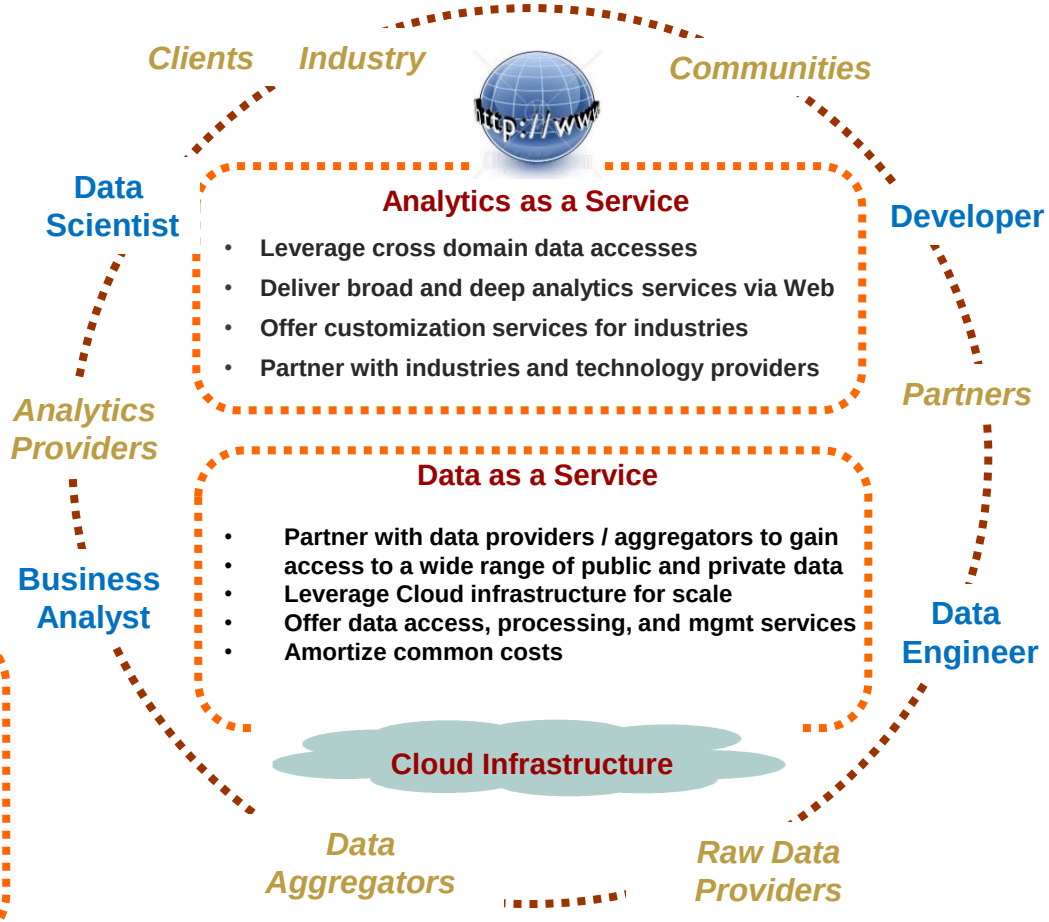
- **Hybrid memory cubes** and other disruptive technologies will enable large-scale, real-time, contextual processing.

Trend 3: Analytics Ecosystem Transformation

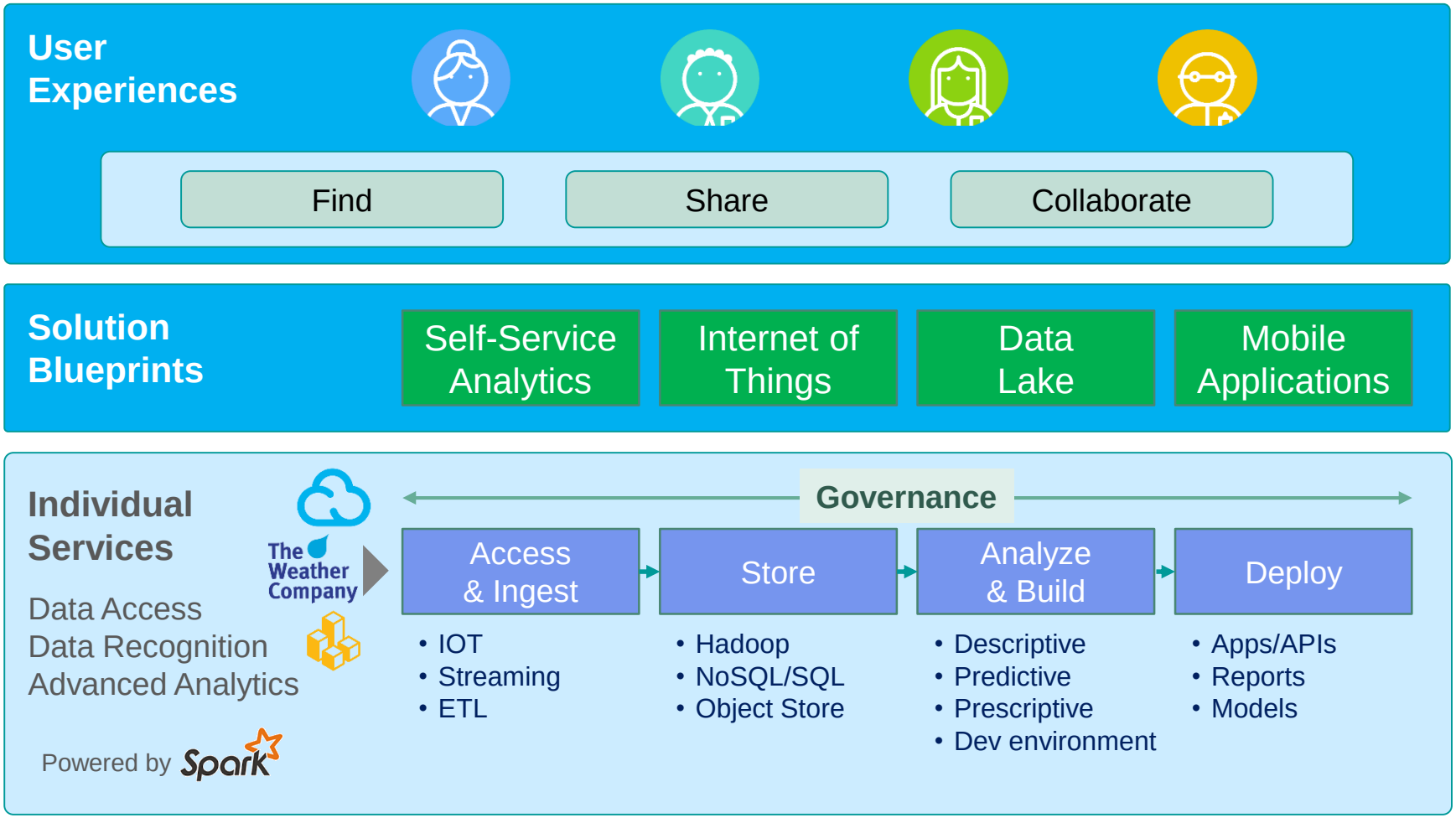
Traditional Analytics Ecosystem
A siloed, linear Model



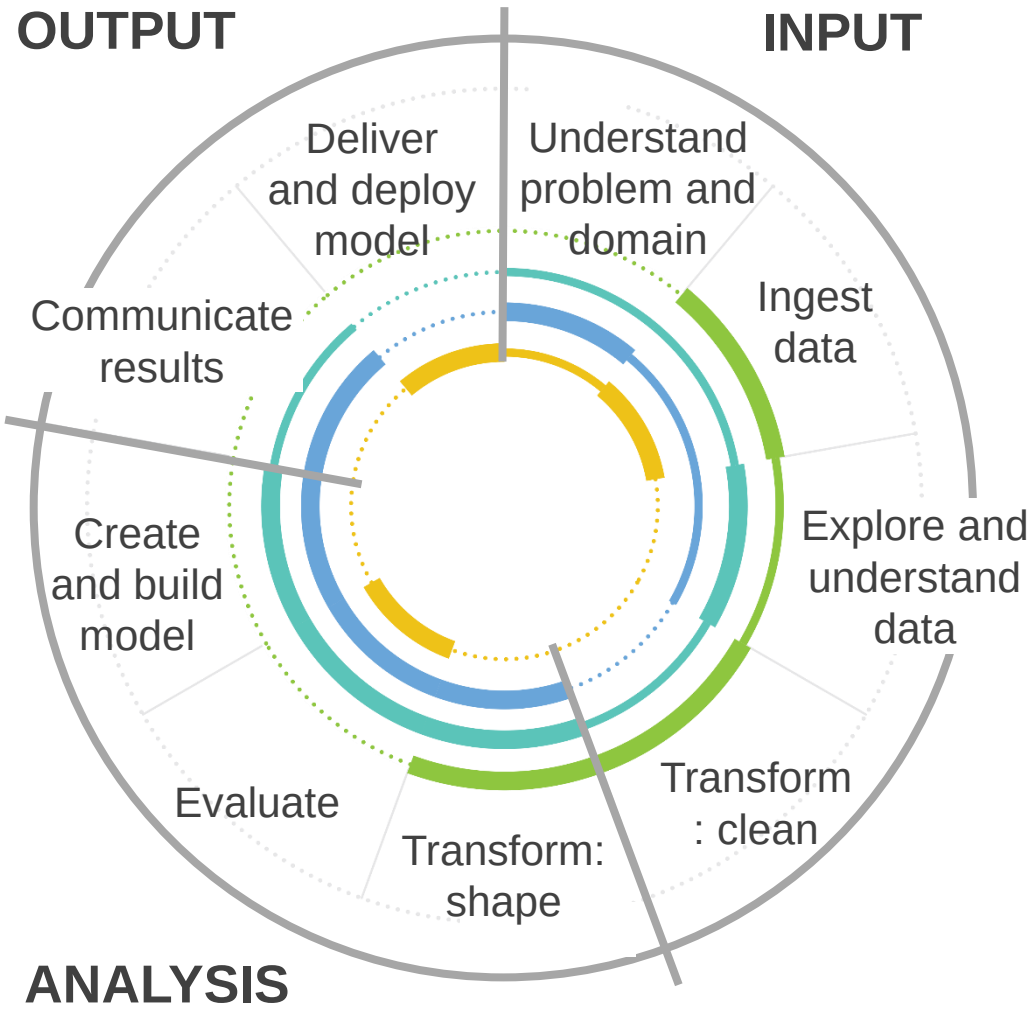
Emerging Analytics Ecosystem
A collaborative, co-developing, and value co-creation model



IBM DataWorks Platform for Analytics Ecosystem Transformation



For Users in the Ecosystem to Collaborate Together



Data Engineer
Architects how data is organized & ensures operability
DataWorks Forge

Data Scientist
Gets deep into the data to draw hidden insights for the business
Data Science Experience

Business Analyst
Works with data to apply insights to the business strategy
Watson Analytics

App Developer
Plugs into data and models & writes code to build apps
Bluemix

Other Important Business Analytics Trends

Cognitive Computing, Machine Learning, Artificial Intelligence

Open source analytics platform

Self-service analytics

NUS Business Analytics Center: Vision and Mission

- Support the goal of making Singapore the regional hub for Business Analytics
  - Build long-term, world-class capabilities to support Singapore's strategic economic focus on business analytics
  - Develop a center of excellence that equips local companies and organizations with talented skills and leaders in business analytics
    - Deep relevance to different industries through direct engagements with companies and enterprises
    - Hands-on practical experience and PoCs with external organizations
- Build a Center with world-class academic and professional reputation
  - Differentiated study program that combines business and computing into analytics curriculum and projects
  - Bring academics and practitioners from different NUS departments and industry to work together with students

NUS Business Analytics Center Projects, 2016 (I)

Finance

OCBC Bank

- ALM Roll-Tagging Prediction
- Risk Classification Model
- Analysing High Risk Segments in Auto Loan Portfolio

HSBC

- Analysis of Customer Queuing Time & Headcount planning

UBS

- CRM/ Markets Analytics

The Center of Property Management Solutions
acgmanagement

- Text Mining

PayPal

- Customer-Money Life Cycle
- Marketplace analysis

DBS

- PnL Analytics
- Cyber Security
- Anti Money Laundering
- Fraud analytics
- Risk assessment model for investigation program

MUFG

- Developing an accurate model to provide estimates on how long a job should take given the characteristics of the job

Insurance

Swiss Re

- Analyzing cancer claims for policy holders

AXA redefining / insurance

- Motor Pricing KPIs
- Travel Pricing KPIs and exploratory analysis
- - Cross-sell and up-sell in insurance

Allianz

- Social Network and Geospatial Analytics in context of Insurance

WAREES

- Optimising Endowment Portfolio Performance

NUS Business Analytics Center Projects, 2016 (II)

Healthcare

- Frequent Attenders to the Emergency Department

- Deriving Insights from NEHR (National Electronic Health Record)

- Emergency Medical Service (EMS) Ambulance Demand Analytics & Prediction

- Breakout detection for Hep C patients.

- Understanding Family Attitudes and Social Support Networks through Analytics

- Case Study on Global FP&A Transformation
- Balance Sheet Forecasting
- IT Tools Comparison
- Case Study on Global FP&A Transformation
- Sales Forecasting and Tools for Predictive Analytics

Information Technology

- Sales Management Analytics
- HR Analytics
- Pricing Assesment Tool based on Analytics

- BlueMix & Watson Analytics

- Online Analytics

- An Analytics Approach to Improve Subscription Rate for Nursing Course (prelim title)

- Global logistics cost optimisation
- Global project - Social Media

- Predicting High Risk Churn Segments Via Product Usage Data

NUS Business Analytics Center Projects, 2016 (III)

Logistics & Manufacturing

- POS transactional data

- Analysis on overtime cost

- Customer credit risk analysis

- IoT / Event Analytics in Manufacturing
- Data Lake architecture to deliver a virtualization layer for disparate data sources.

Retail

- Supply Chain Network Optimization

- News recommendation engine for high net worth customers
- Economic Scenario Stress Testing

- Determining optimal level of markdowns through customer segmentation for revenue maximization

- Market Research
- Social Media /Digital Marketing/PR

Government

- Optimising maintenance schedule for fleet management

- Automatic Rostering System

Thank You !

Dr. PANG Yan, James
Email: jamespang@nus.edu.sg

Dr. Jorge SANZ
Email: jorges@nus.edu.sg