# Aligning Your Strategic Initiatives with a Realistic Big Data Analytics Roadmap

3 key strategic advantages, and a realistic roadmap for what you really need, and when

Enterprise Information Management

Cognizant

Topics to be discussed

1 Effective exploration of your data to identify your next Analytics opportunity for innovation, and competitive advantage?

2 Key tools and partners to align your strategic goals to big data KPI's & metrics?

The Answer?

3 Leverage successful patterns for big data insight and avoid re-invention of technology

4 Optimized HW/SW Platforms to reduce time, cost, risk, and waste

5 Use Cloud, SAAS, and external data platforms to enrich and extend my on-premises IT Investments

Cognizant
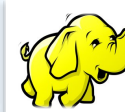
Why is Big Data Multiplying the Effectiveness of the BI / Analytics Investment?
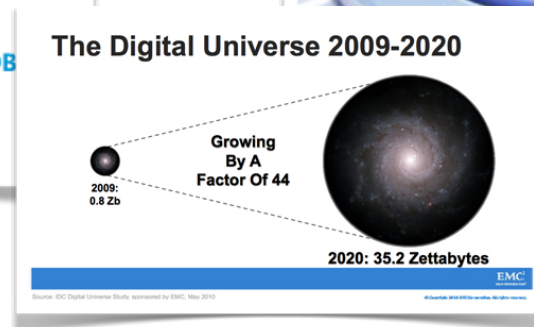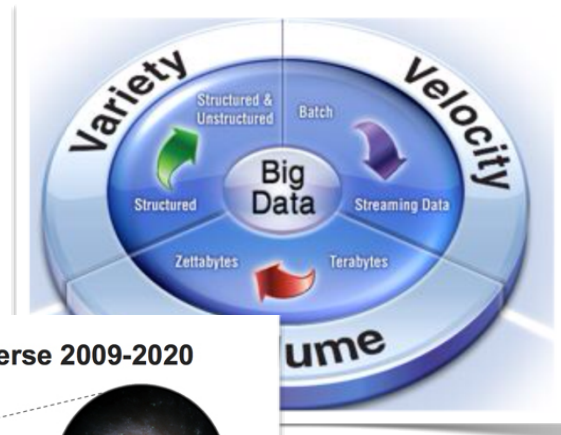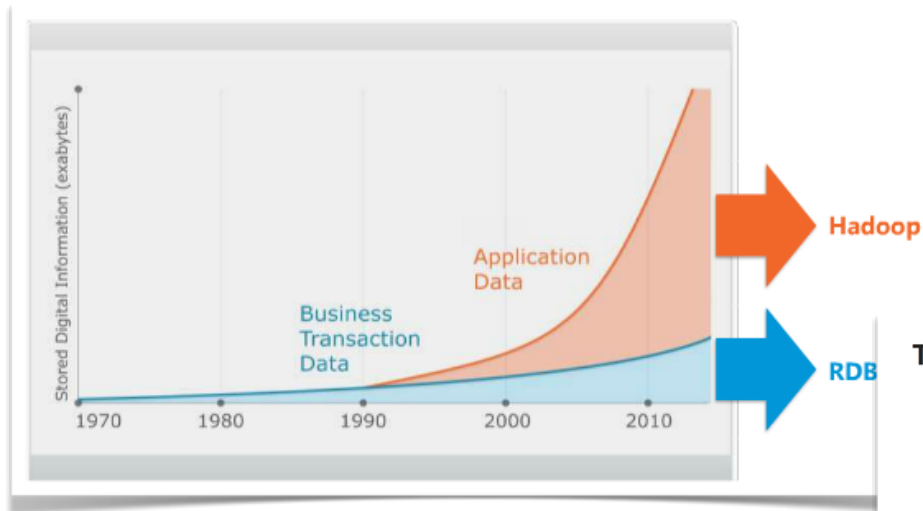
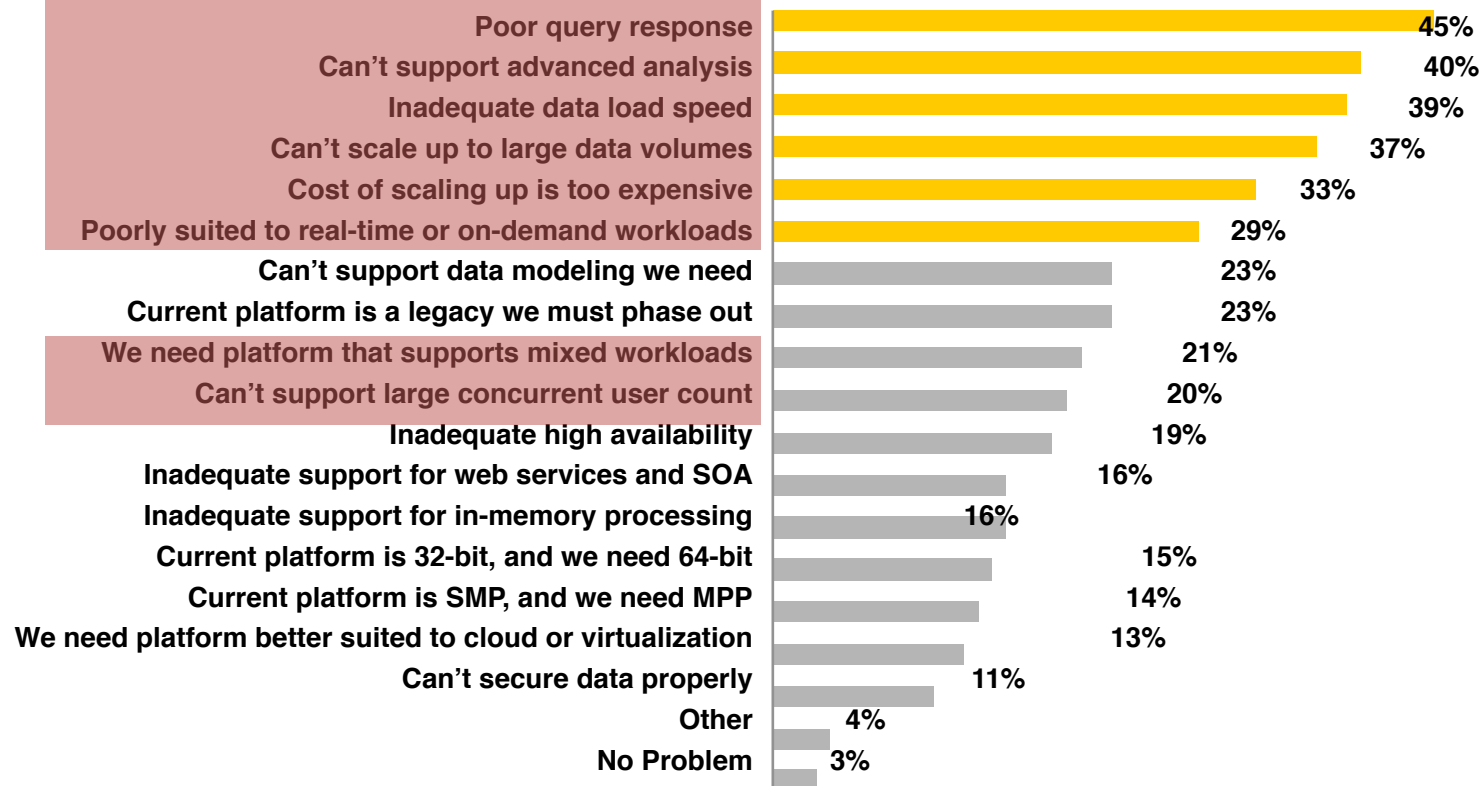| VOLUME		VARIETY		VELOCITY
Desire to **run analytics on entire data sets**, not sample sizes–better accuracy		Rapid data generation coming from **diverse sources**. New sources of data is unstructured / semi-structured.		Demand for **near real-time processing** of new data (for customer targeting, loyalty, fraud, etc.)
–"By 2014, 85% of the **data warehouses currently deployed <u>will fail to scale</u>** to meet new information volume and complexity demands." Source: Gartner		–"Through 2015, enterprises integrating high-value and diverse new information types and sources into a coherent information management infrastructure, will financially **outperform their industry peers** by more than 20%." Source: Gartner		

DW/BI & Analytics Pain TODAY

Traditional Systems Not Built for Large Data Volumes and Advanced Analytics

What Problems Are Driving You to Look at New Data Management Solutions?

| Problem	Percentage
Poor query response	45%
Can't support advanced analysis	40%
Inadequate data load speed	39%
Can't scale up to large data volumes	37%
Cost of scaling up is too expensive	33%
Poorly suited to real-time or on-demand workloads	29%
Can't support data modeling we need	23%
Current platform is a legacy we must phase out	23%
We need platform that supports mixed workloads	21%
Can't support large concurrent user count	20%
Inadequate high availability	19%
Inadequate support for web services and SOA	16%
Inadequate support for in-memory processing	16%
Current platform is 32-bit, and we need 64-bit	15%
Current platform is SMP, and we need MPP	14%
We need platform better suited to cloud or virtualization	13%
Can't secure data properly	11%
Other	4%
No Problem	3%

Cognizant

How can I ensure Big Data has Value to the Business?

Cognizant

The Difference Between BI & Analytics
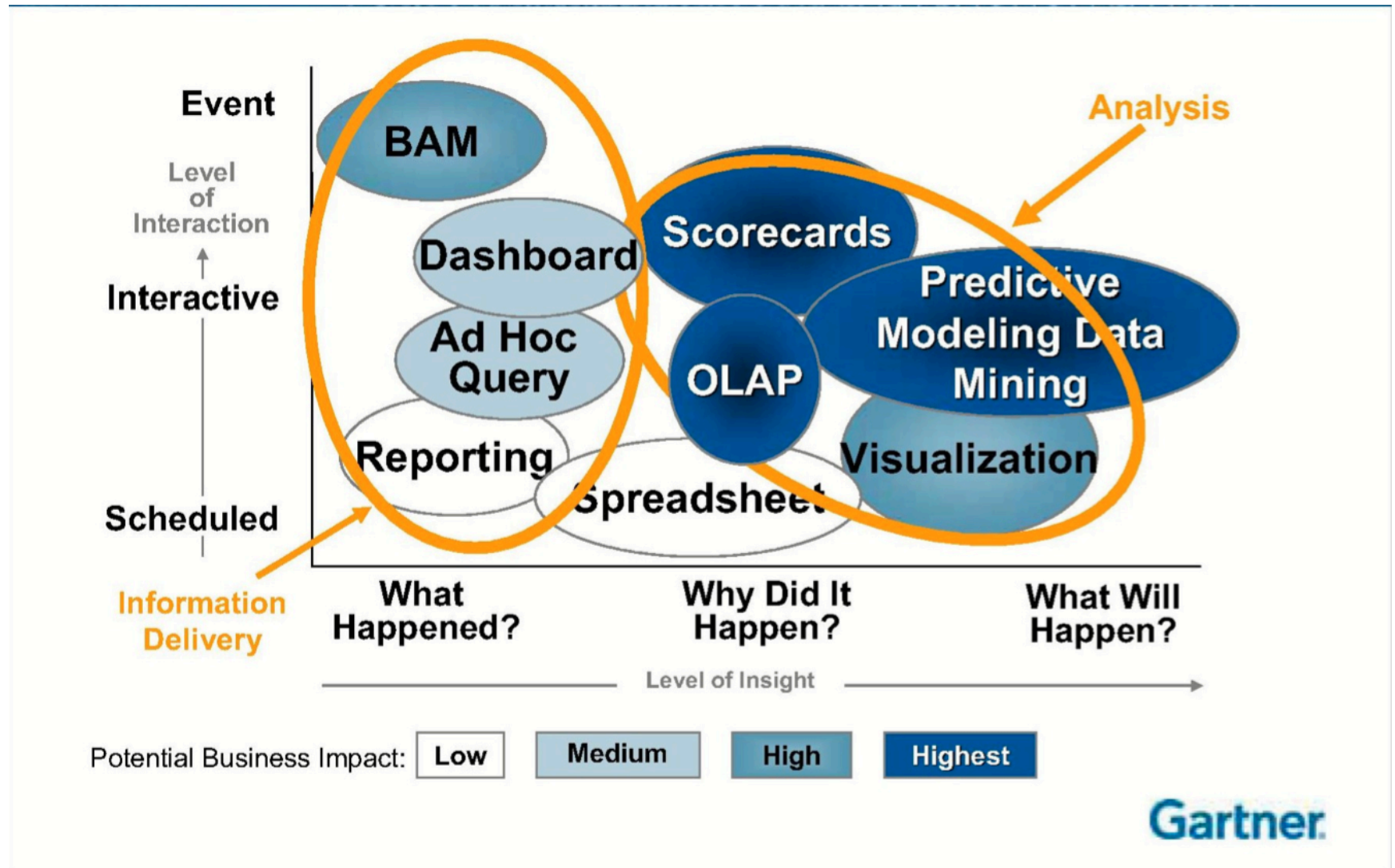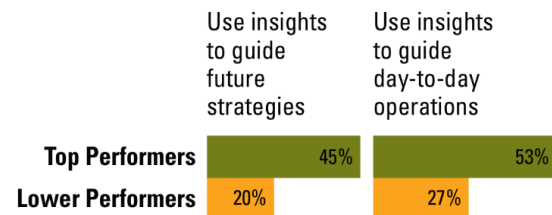
What Investments Does Big Data Touch?
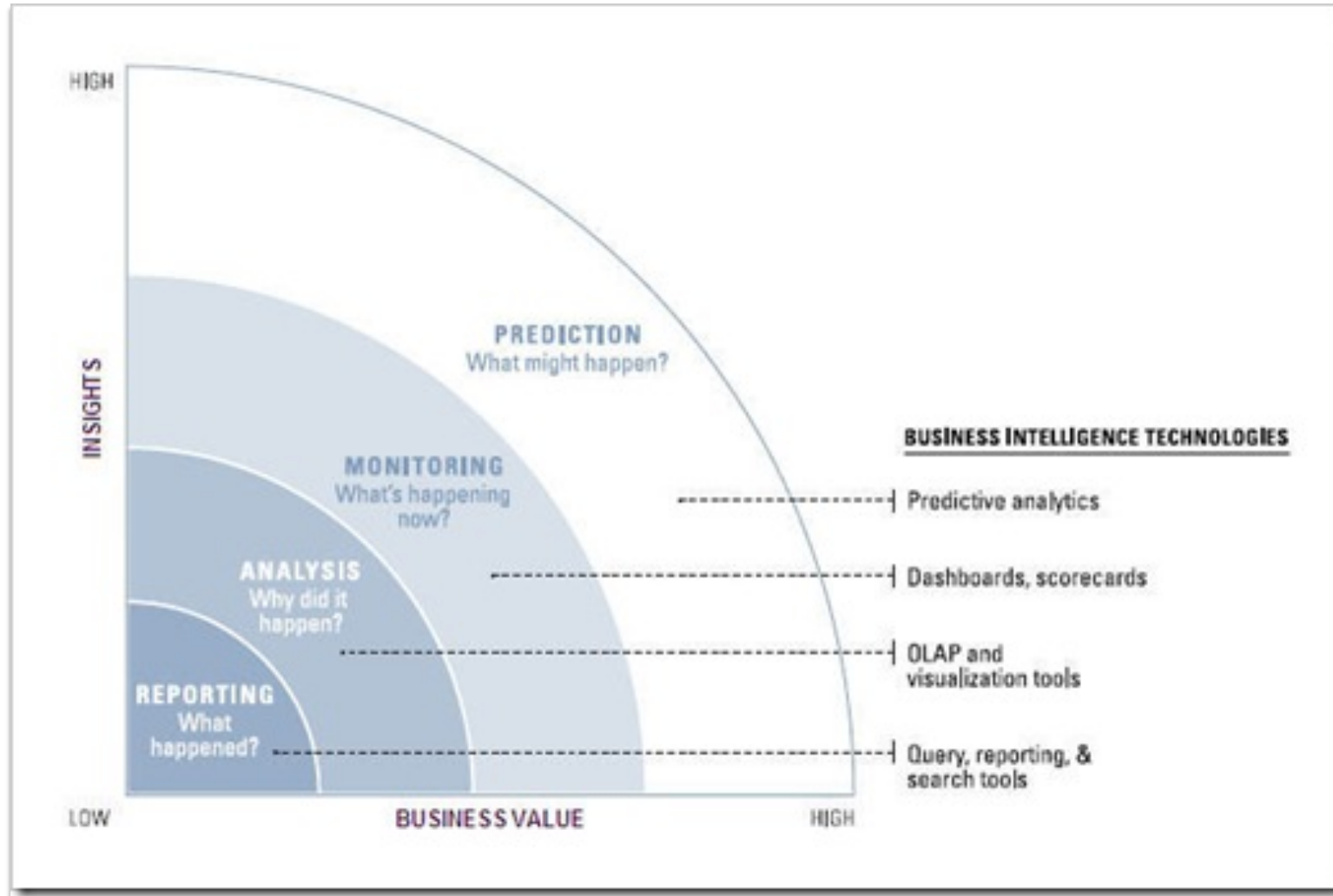
MITSloan
Management Review

Big Data, Analytics and the Path From Insights to Value
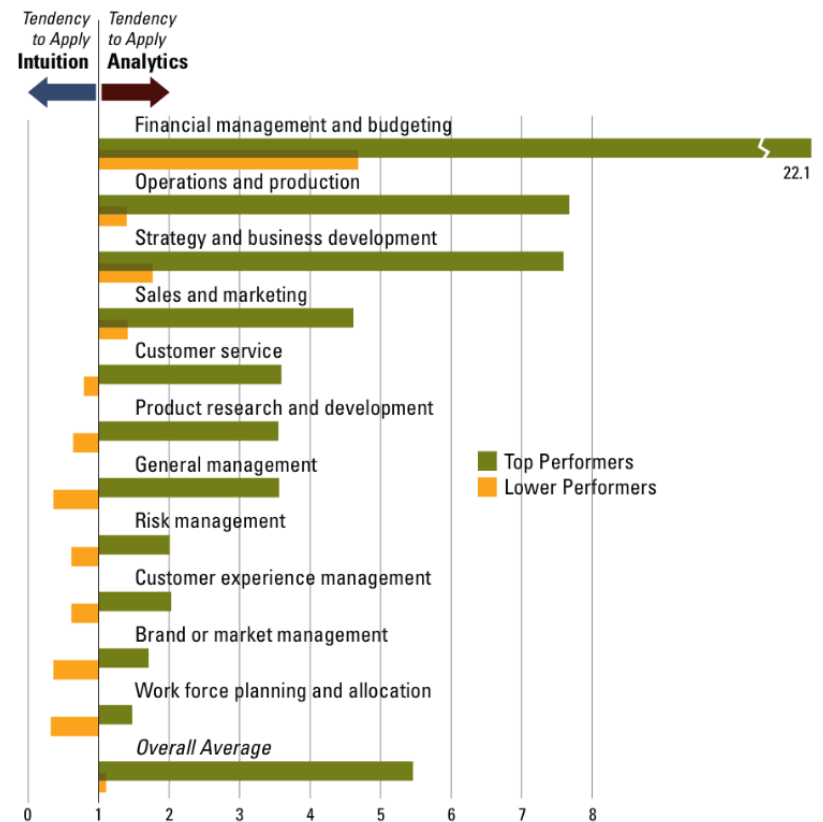
THE ANALYTICS HABITS OF TOP PERFORMERS

Top-performing organizations were twice as likely to use analytics to guide day-to-day operations and future strategies as lower performers.

Percent of respondents whose organizations do these activities well or very well.

	Use insights to guide future strategies	Use insights to guide day-to-day operations
Top Performers	45%	53%
Lower Performers	20%	27%

ANALYTICS TRUMPS INTUITION

The tendency for top-performing organizations to apply analytics to particular activities across the organization compared with lower performers. A likelihood of 1.0 indicates an equal likelihood that the organizations will use either analytics or intuition.

Tendency to Apply **Intuition**   *Tendency to Apply* **Analytics**

Financial management and budgeting — 22.1

Operations and production

Strategy and business development

Sales and marketing

Customer service

Product research and development

General management

Risk management

Customer experience management

Brand or market management

Work force planning and allocation

Overall Average

■ Top Performers
■ Lower Performers

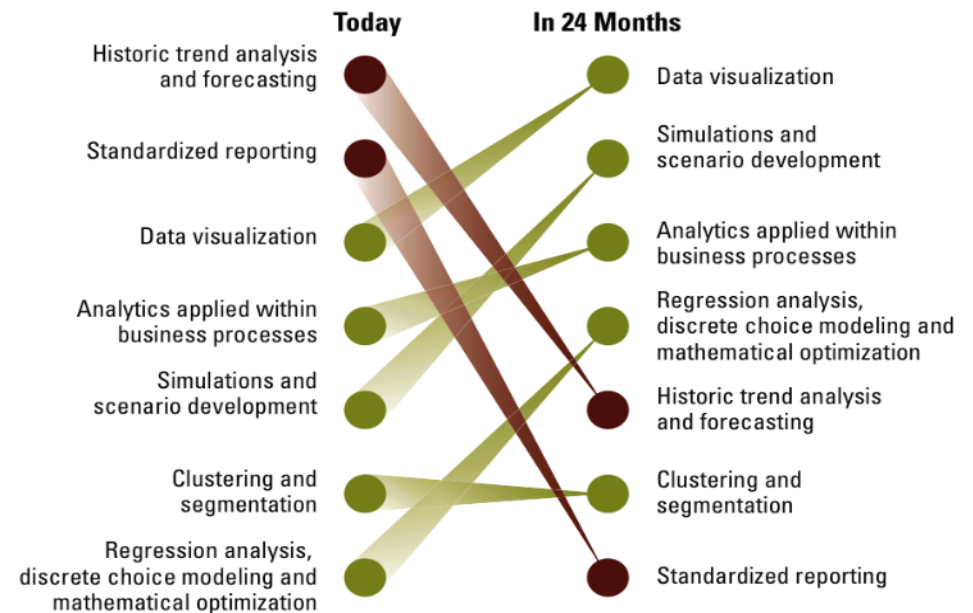0 1 2 3 4 5 6 7 8

Extreme analytics
Allow Us to Find:

• Patterns

• Trends

• Predictions Never
Before Feasible

MITSloan Management Review — Big Data, Analytics and the Path From Insights to Value

WHERE ARE DATA-DRIVEN MANAGERS HEADED?

Organizations expect that the ability to visualize data differently will be the most valuable technique in two years. Other techniques and activities that are currently delivering the most value today will still be done, but will be of less value.

Today
- Historic trend analysis and forecasting
- Standardized reporting
- Data visualization
- Analytics applied within business processes
- Simulations and scenario development
- Clustering and segmentation
- Regression analysis, discrete choice modeling and mathematical optimization

In 24 Months
- Data visualization
- Simulations and scenario development
- Analytics applied within business processes
- Regression analysis, discrete choice modeling and mathematical optimization
- Historic trend analysis and forecasting
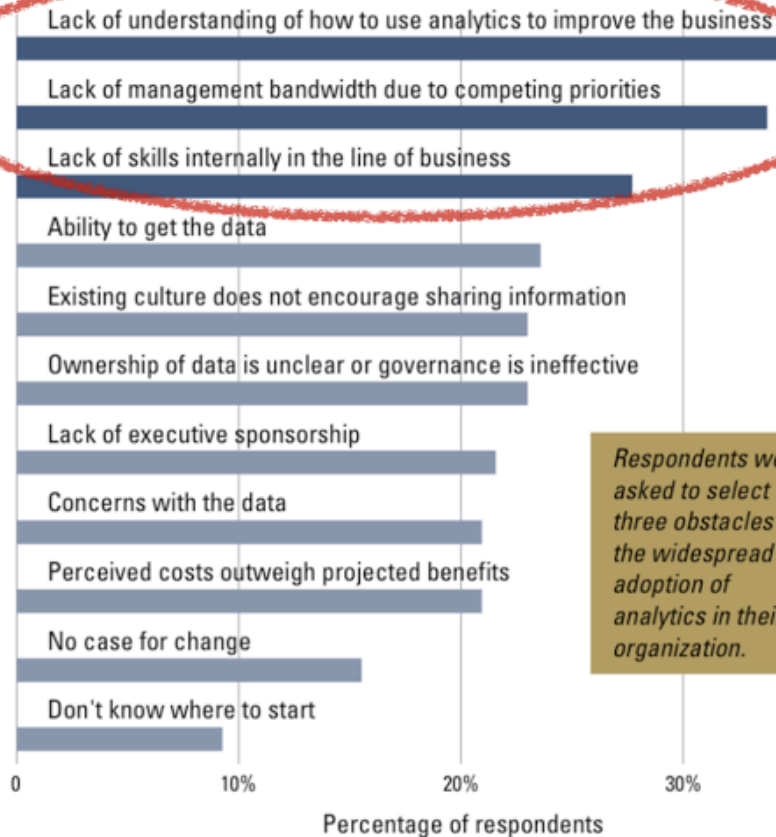- Clustering and segmentation
- Standardized reporting

Respondents were asked to identify the top three analytic techniques creating value for the organization, and predict which three would be creating the most value in 24 months.

Analytic "Coaches" are Key to Market Performers

THE IMPEDIMENTS TO BECOMING MORE DATA DRIVEN

The adoption barriers organizations face most are managerial and cultural rather than related to data and technology.

Lack of understanding of how to use analytics to improve the business

Lack of management bandwidth due to competing priorities

Lack of skills internally in the line of business

Ability to get the data

Existing culture does not encourage sharing information

Ownership of data is unclear or governance is ineffective

Lack of executive sponsorship

Concerns with the data

Perceived costs outweigh projected benefits

No case for change

Don't know where to start

0   10%   20%   30%

Percentage of respondents

Respondents were asked to select three obstacles to the widespread adoption of analytics in their organization.

1. How to Use analytics to Improve the Business

2. Lack of Management Bandwidth / Priorities

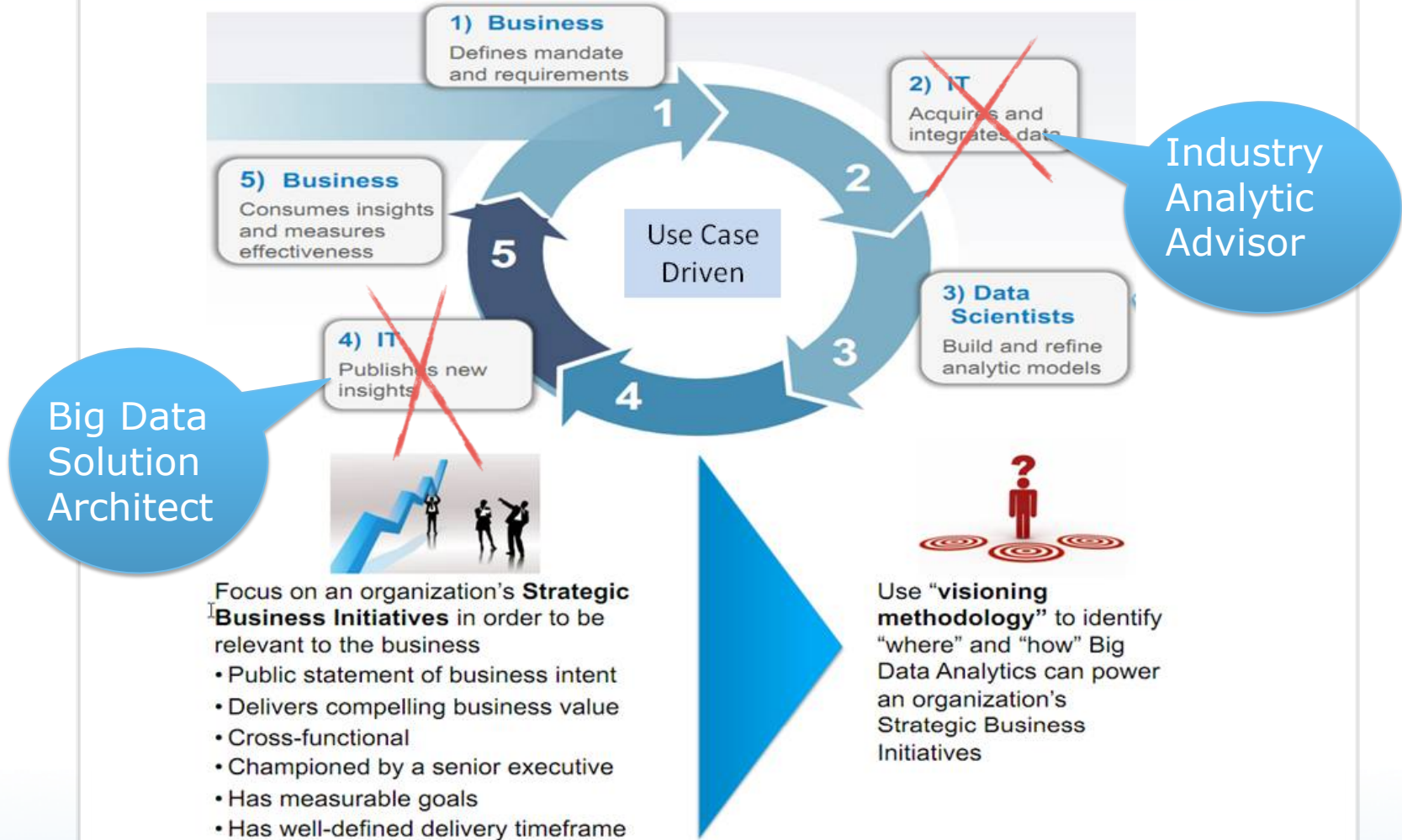- Lack of Analytics Skills at LOB

Cognizant

Every Level Can Benefit from an Analytics Advisor:

- I don't know what I want

- I don't know how

- I know, but need help

- I know, and I can do, but want to optimize
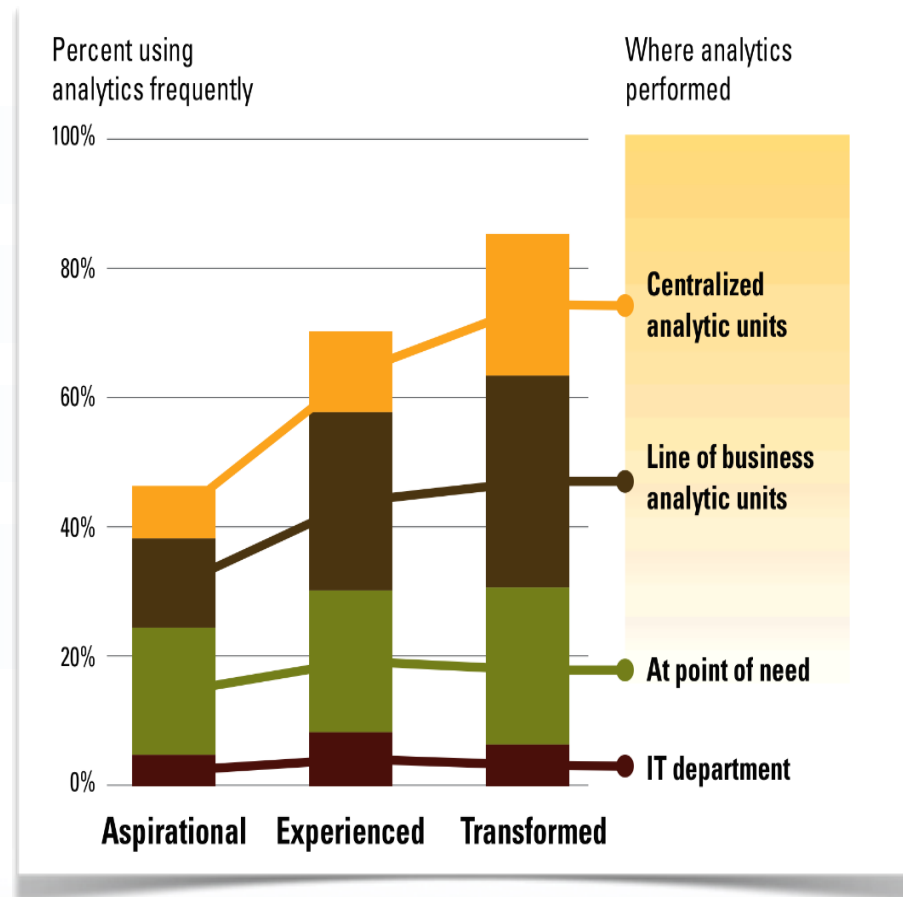
Where: Industry Analytic Advisors Provide a Critical Relationship Between Executives & Solution Makers...

Use Case: focusing on a compelling business opportunity

1) **Business**
Defines mandate and requirements

2) **IT**
Acquires and integrates data

3) **Data Scientists**
Build and refine analytic models

4) **IT**
Publishes new insights

5) **Business**
Consumes insights and measures effectiveness

Use Case Driven

Industry Analytic Advisor

Big Data Solution Architect

Focus on an organization's **Strategic Business Initiatives** in order to be relevant to the business
• Public statement of business intent
• Delivers compelling business value
• Cross-functional
• Championed by a senior executive
• Has measurable goals
• Has well-defined delivery timeframe

Use "**visioning methodology**" to identify "where" and "how" Big Data Analytics can power an organization's Strategic Business Initiatives

Cognizant

Most Companies Try to use IT to Hand off Key Insights...

Where are you in your Big Data Analytics Journey?

Cognizant

I Know I Need an Analytical Advantage...
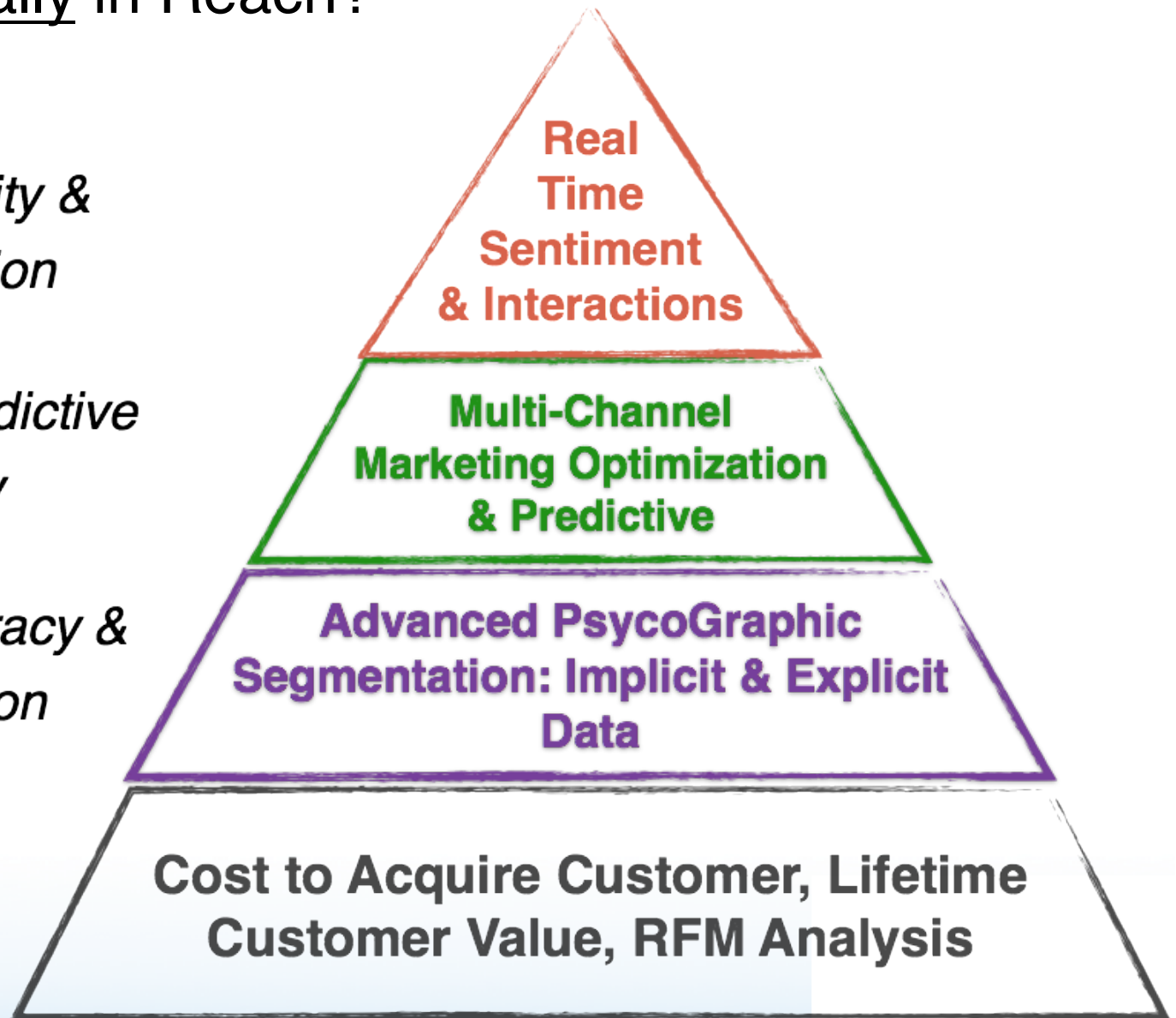
But What's <u>Really</u> in Reach?

1. Enterprise Agility & Market Interaction

2. Strategic & Predictive Planning Clarity

3. Decision Accuracy & Insight Innovation
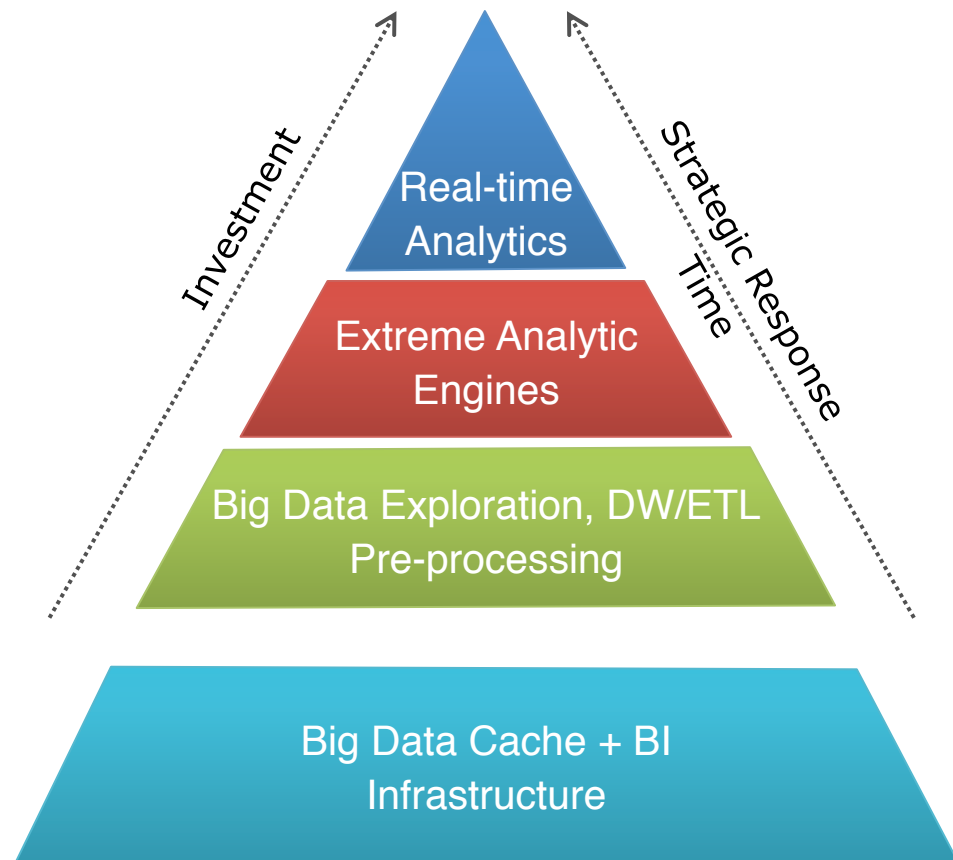
4. Performance Measurement

Real Time Sentiment & Interactions

Multi-Channel Marketing Optimization & Predictive

Advanced PsycoGraphic Segmentation: Implicit & Explicit Data

Cost to Acquire Customer, Lifetime Customer Value, RFM Analysis

3 Key Advantages: Now Within Reach

What do I **Really** need?

1 Enterprise Agility & Market Interaction

2 Strategic & Predictive Planning Clarity

3 Decision Accuracy & Insight Innovation

4 Performance Measurement

Pyramid diagram with axes labeled "Investment" (left) and "Strategic Response Time" (right):
- Real-time Analytics
- Extreme Analytic Engines
- Big Data Exploration, DW/ETL Pre-processing
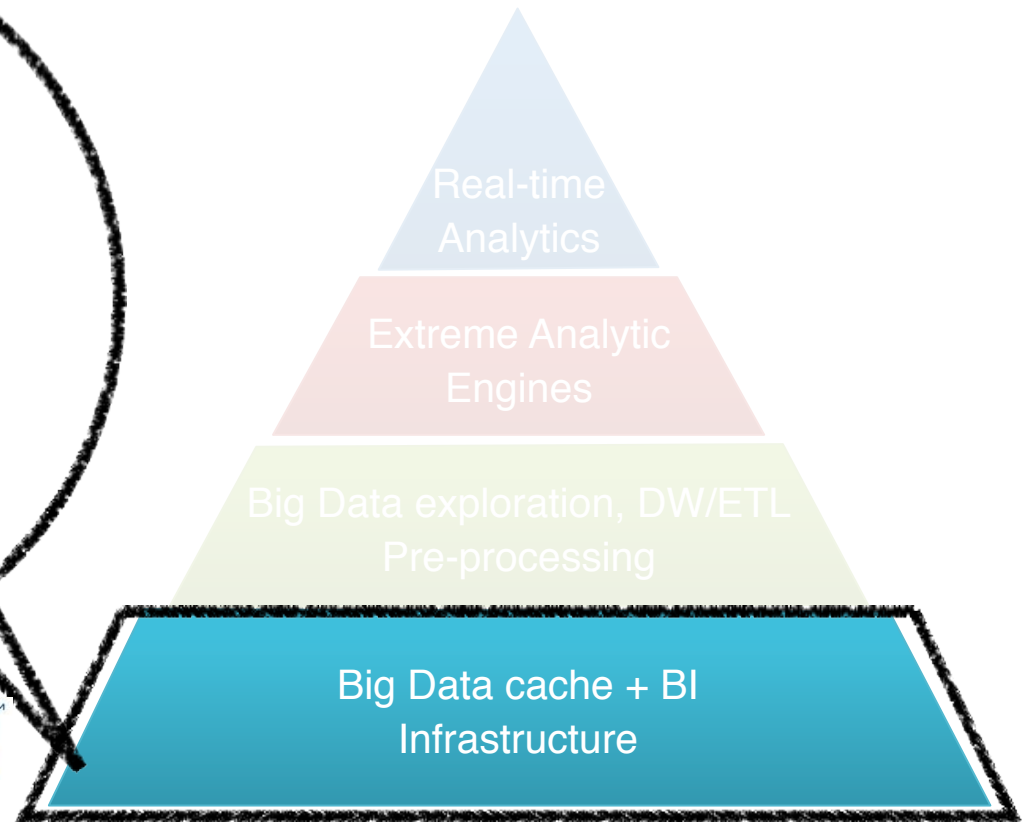- Big Data Cache + BI Infrastructure

Cognizant

Not Usually a Strategic Advantage: But Big Cost Advantages

Performance Management is most often used in an Operational Fashion, for Optimization

Data Cache or "Data Lake" allows for a Massive Data SandBox

(At 1/10 the cost of In-DW Storage?)

Real-time Analytics

Extreme Analytic Engines

Big Data exploration, DW/ETL Pre-processing
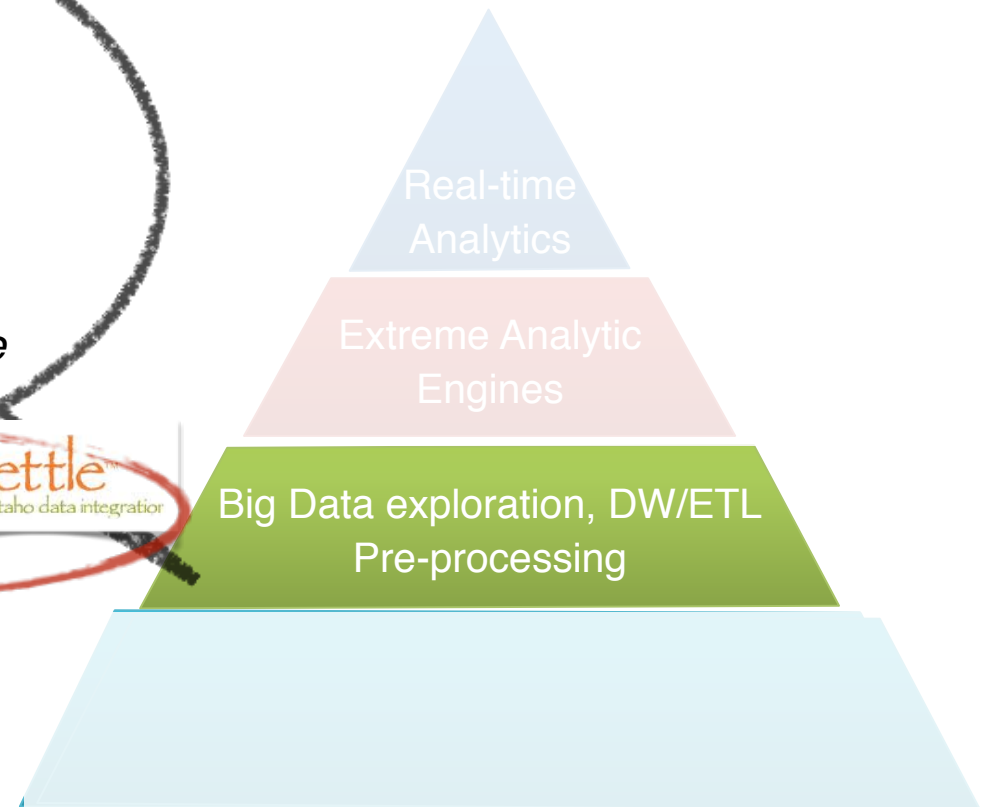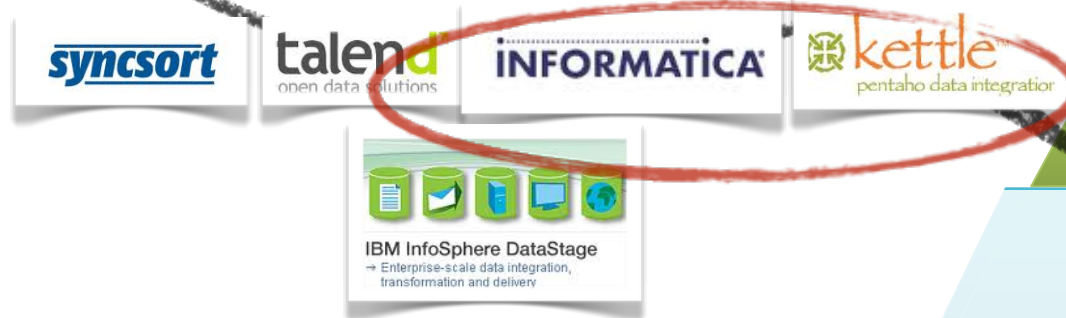
Big Data cache + BI Infrastructure

Free up your DW to focus on what it does best:

Analyze Structured Data, and Build Repeatable metrics

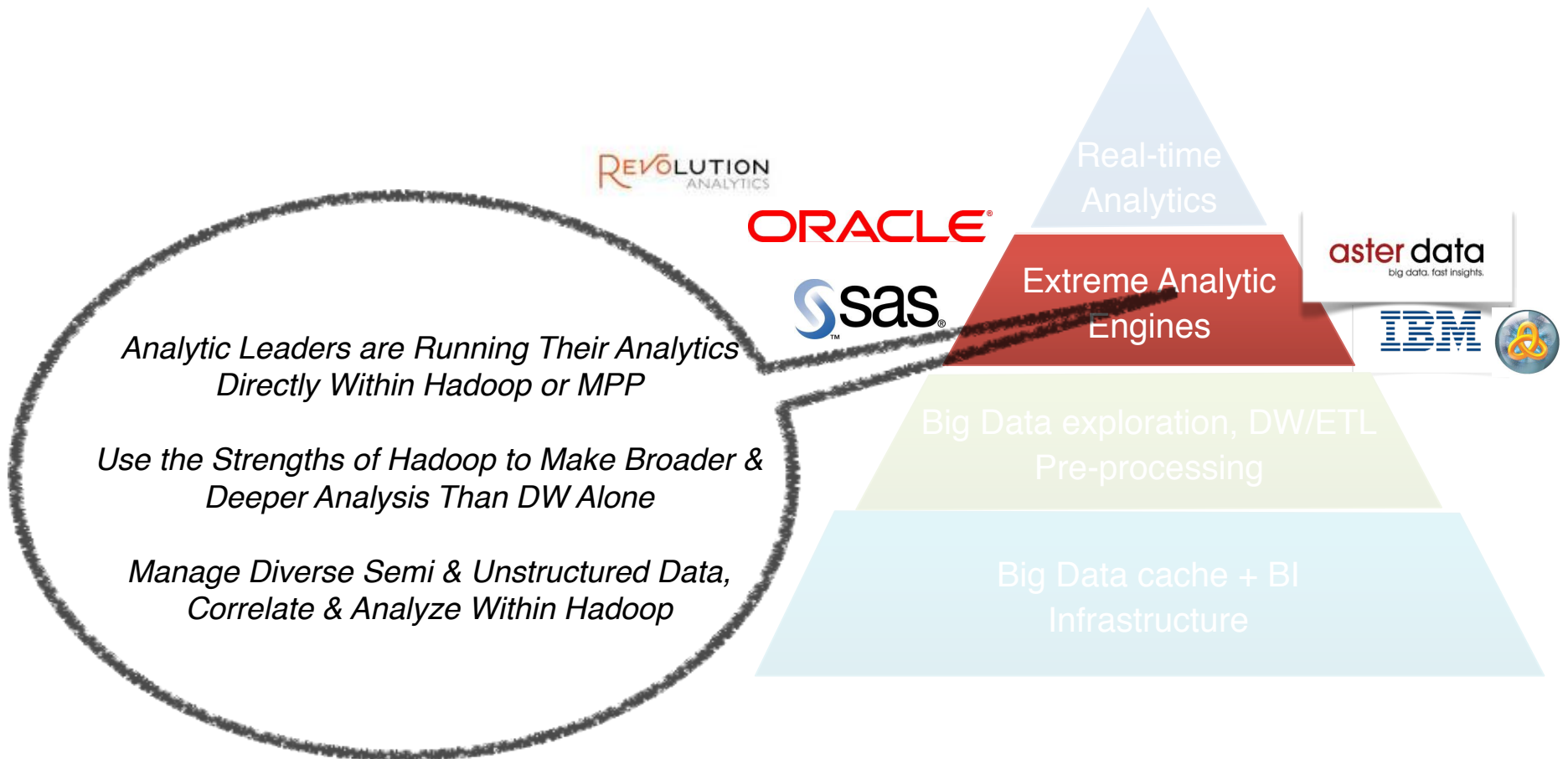*A Data Staging Area at Pennies on the Dollar*

Use the Horsepower of Hadoop to Transform Your Data

Scale from 4 to 4000 Nodes - Incrementally, with Fairly Predictable Results

syncsort

talend open data solutions

INFORMATICA

kettle pentaho data integration

IBM InfoSphere DataStage
→ Enterprise-scale data integration, transformation and delivery

Real-time Analytics

Extreme Analytic Engines

Big Data exploration, DW/ETL Pre-processing

Cognizant

Scale your Metrics & Algorithms to Broader Parameters

Data Mining Algorithms Haven't Changed That Much in 20 Years - But Experts Agree That More Data Increases Accuracy & Insight

REVOLUTION ANALYTICS

ORACLE®

§sas®

aster data
big data. fast insights.

IBM

Real-time Analytics

Extreme Analytic Engines

Big Data exploration, DW/ETL Pre-processing

Big Data cache + BI Infrastructure

Analytic Leaders are Running Their Analytics Directly Within Hadoop or MPP

Use the Strengths of Hadoop to Make Broader & Deeper Analysis Than DW Alone

Manage Diverse Semi & Unstructured Data, Correlate & Analyze Within Hadoop
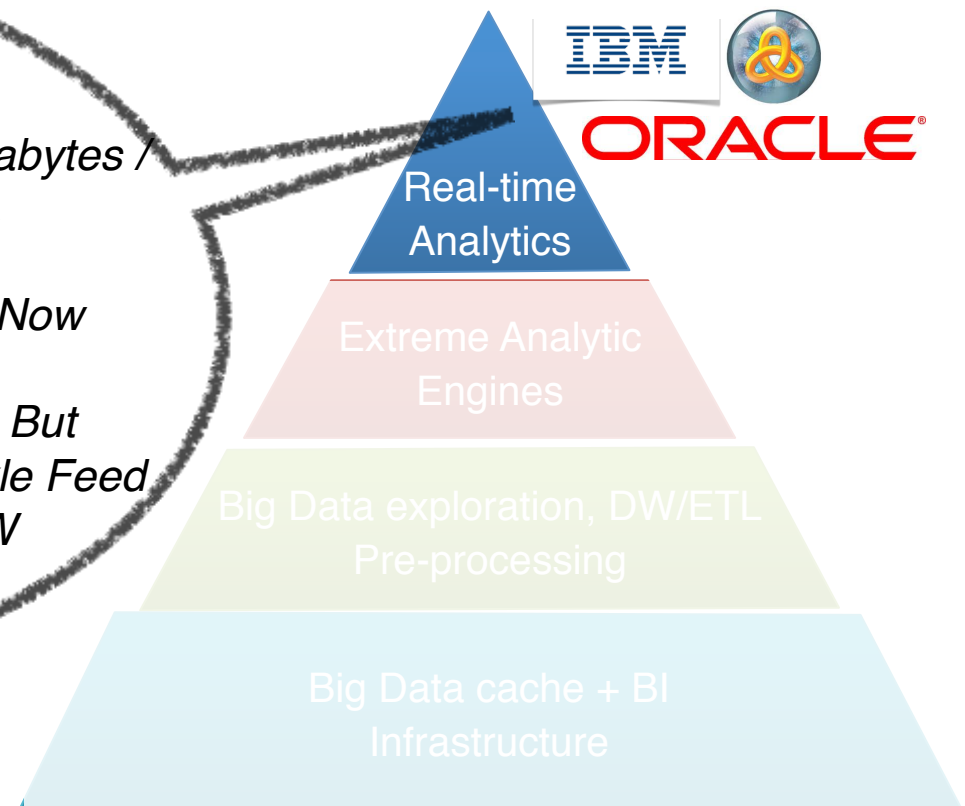
Cognizant

Just In Time: Worth More Than Decades of Hindsight

Much of Traditional BI and Some Analytics are Obsolete Upon Completion.

Streaming Data is Approaching Tera-Petabytes / Day in Telecom, Web, Media, etc.

Your Competitors Are Investigating It Now

Data Does Not Have to Be Retained, But Business Rules Can Take Action or Trickle Feed Key Insights to DashBoards or DW

IBM

ORACLE®

Real-time Analytics

Extreme Analytic Engines

Big Data exploration, DW/ETL Pre-processing

Big Data cache + BI Infrastructure

Cognizant

Why You Need a Partner with Credibility and Methodology for Big Data Analytics

Physical security of data center

- Operational & Application security
- Cluster security and user authorization
- Effective disaster recovery plan

Integration with existing IT Systems

Integration strategy with Hadoop cluster

- Control access groups

Hardware Requirements

Information Security & Governance

Key Considerations

Requirement based configuration

- Light Processing config
- Balanced Compute config
- Storage heavy config
- Compute intensive config

Monitoring and alerting mechanism

- Library set up for various projects
- Data access frameworks to be set up(HBASE,PIG,HIVE)
- Hadoop Core setup (Common Java runtime for HDFS)

Software Requirements

Cognizant

Cognizant's integrated Social Media Analysis & Reporting Tool (iSMART)

Listener	Calls social media web services (Facebook, Twitter, Blogs, E-Mails, Foursquare, Youtube etc) and crawl through social media websites
Integrate	Houses social media data and associated BI/analytics structures. Integrates with the warehouse and data mart to bring the 360 degree view
Analyze	Measures brand's sentiment on Social media sites using Cognizant's proprietary Sentiment Analysis algorithm
Track	Mobile BI ready dashboards and reports Brand Sentiment Analytics Competitive Intelligence & Customer Insights Social Media ROI Calculation Marketing Campaign Effectiveness
Act	Make informed decision Take action, and Track the results

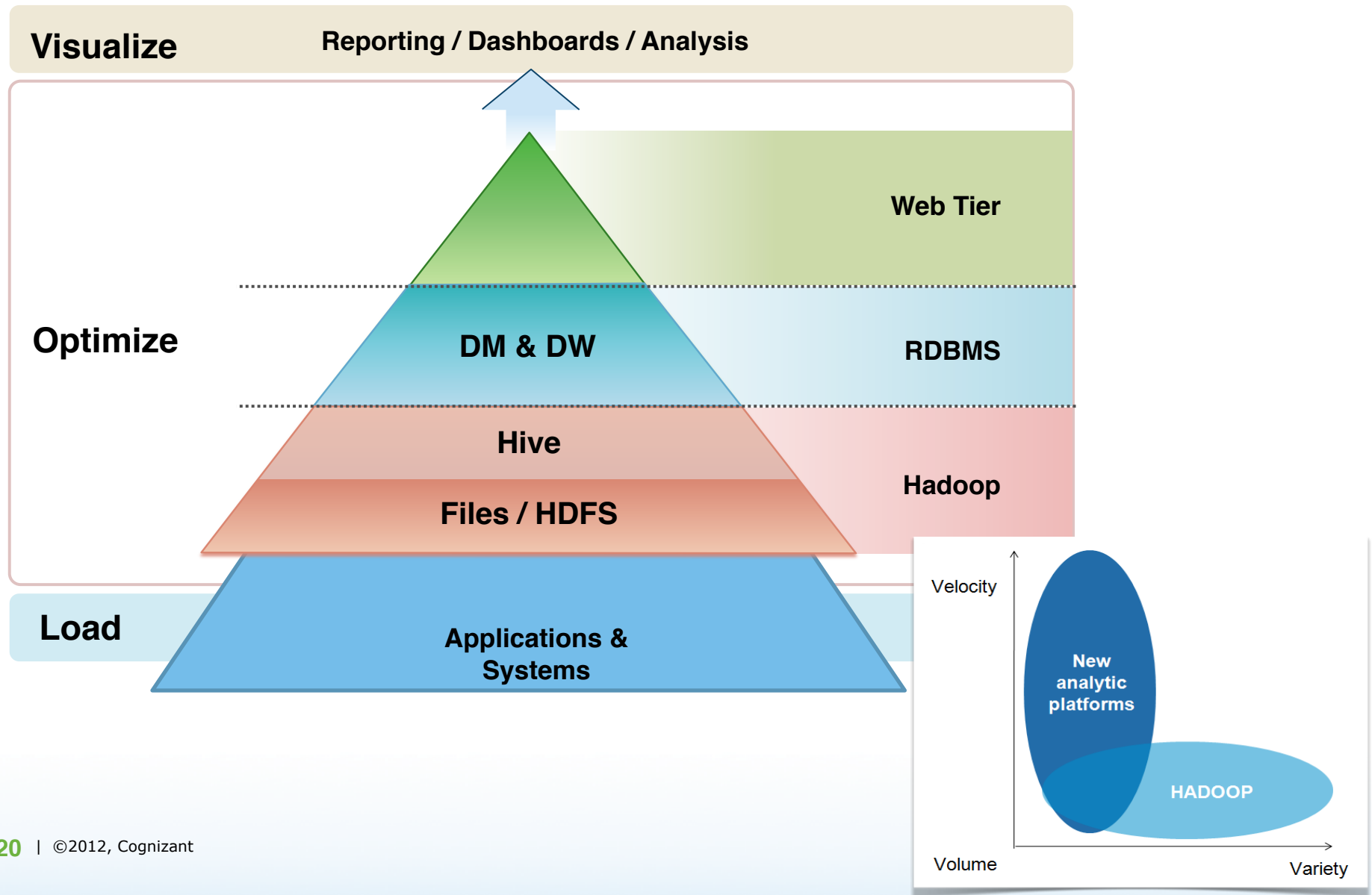## Predictable Big Data Frameworks for Your Vertical Innovation

Cognizant

Offerings for your Big Data Analytics Journey Must Be Aligned & Integrated

Implementations and Architecture

Integrations and Migrations from existing applications

Platform specific Benchmarking services

Analytics and Analytical Lab

Reporting Services

Staff Augmentation

Monitoring and Administration

Cognizant

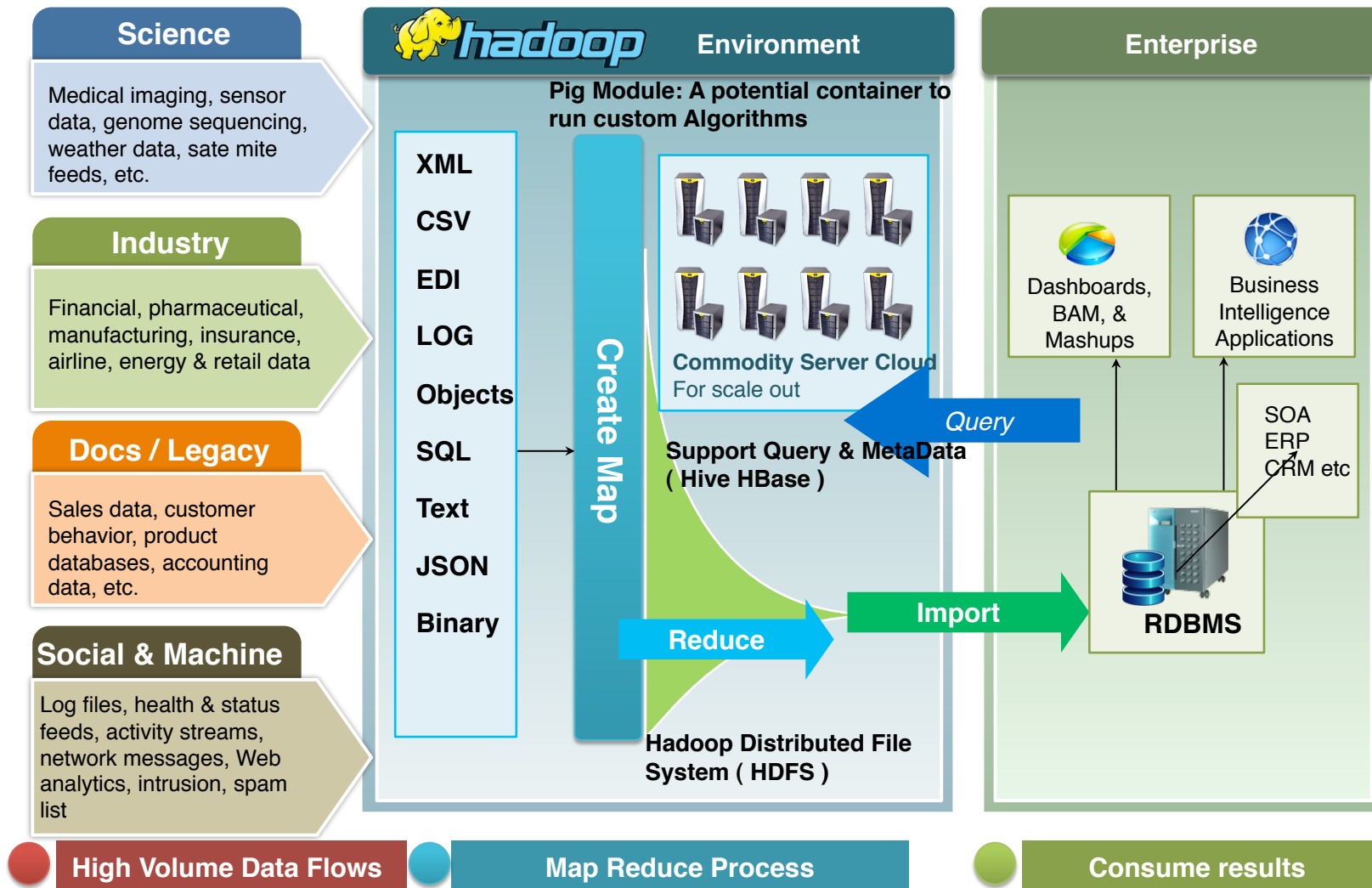# Backup: Why You Need Big Data Solution Frameworks

Cognizant

Hadoop Creates New Possibility and Efficiency in the New Analytic Platform... But <u>MAJOR</u> Assembly is Required
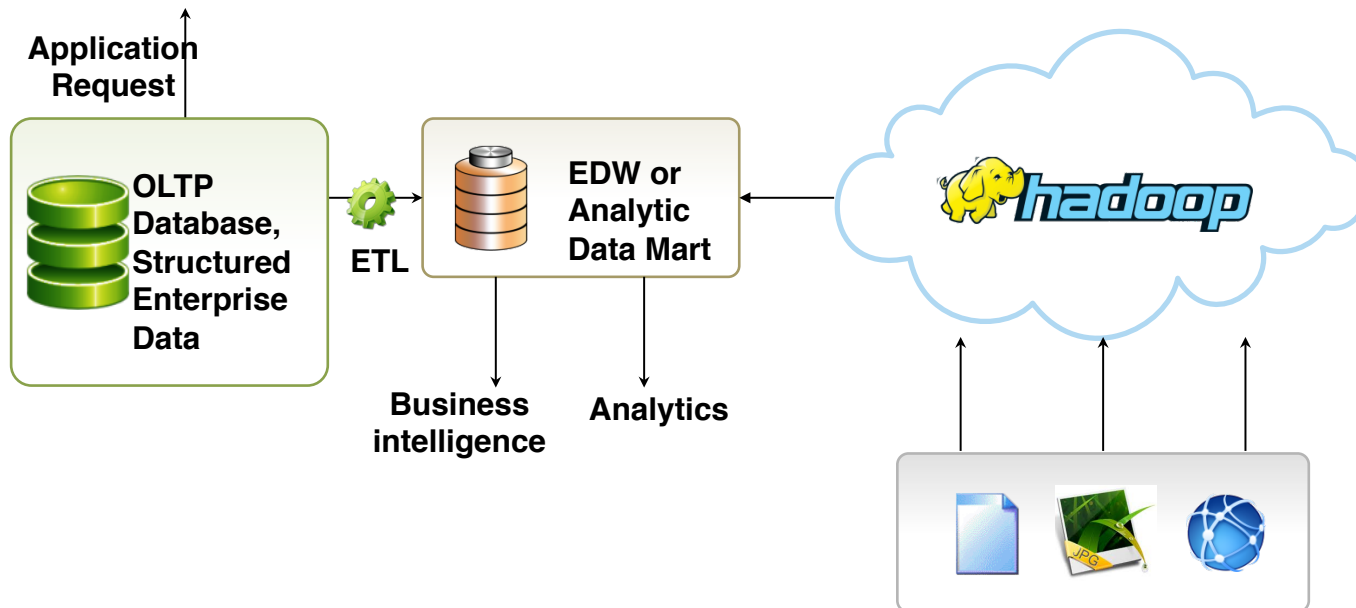
Big Data Analytics-as-a-Service

Science
Medical imaging, sensor data, genome sequencing, weather data, sate mite feeds, etc.

Industry
Financial, pharmaceutical, manufacturing, insurance, airline, energy & retail data

Docs / Legacy
Sales data, customer behavior, product databases, accounting data, etc.

Social & Machine
Log files, health & status feeds, activity streams, network messages, Web analytics, intrusion, spam list

hadoop Environment

Pig Module: A potential container to run custom Algorithms

- XML
- CSV
- EDI
- LOG
- Objects
- SQL
- Text
- JSON
- Binary

Create Map

Commodity Server Cloud
For scale out

Support Query & MetaData (Hive HBase)

Query

Reduce

Import

Hadoop Distributed File System (HDFS)

Enterprise

Dashboards, BAM, & Mashups

Business Intelligence Applications

SOA ERP CRM etc

RDBMS

| High Volume Data Flows | Map Reduce Process | Consume results |

"Integration is often 80% of the total cost in advanced analytics ."- Data Mining Expert Ronnie Kohavi, PHD
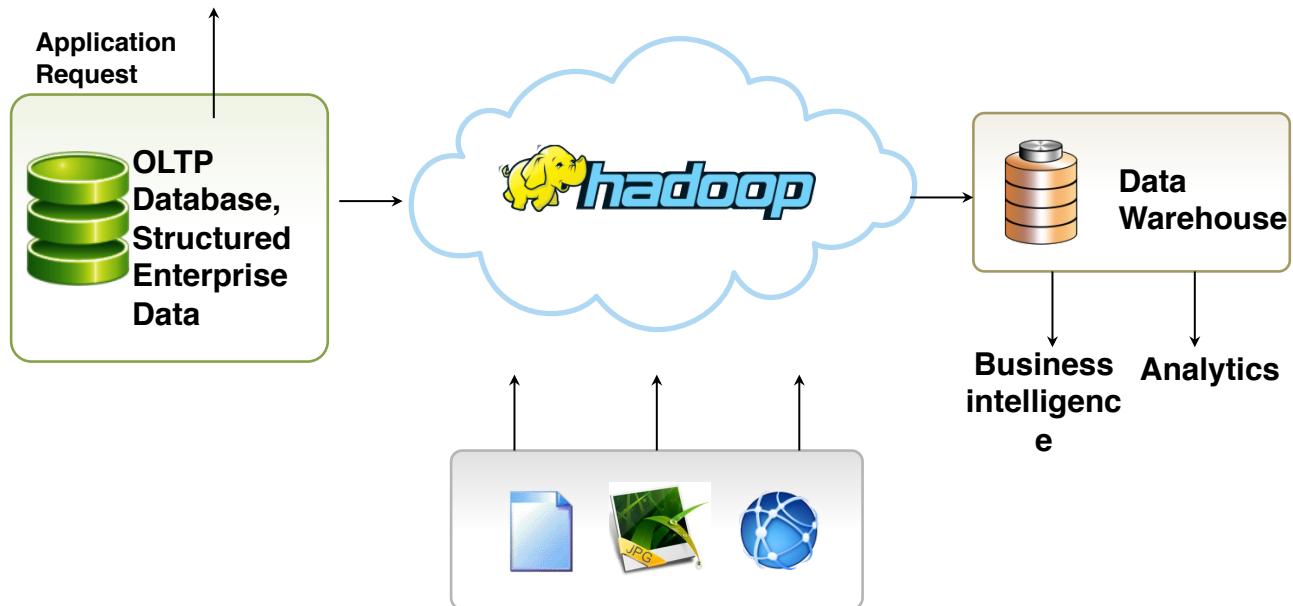
Cognizant

Hadoop as DW Augmentation

Stage 1 — Add unstructured Data

Application Request

OLTP Database, Structured Enterprise Data

ETL

EDW or Analytic Data Mart

Business intelligence

Analytics

hadoop

Cognizant

Hadoop as an ETL "Pre Processing" Perform

Hadoop as an ETL Perform
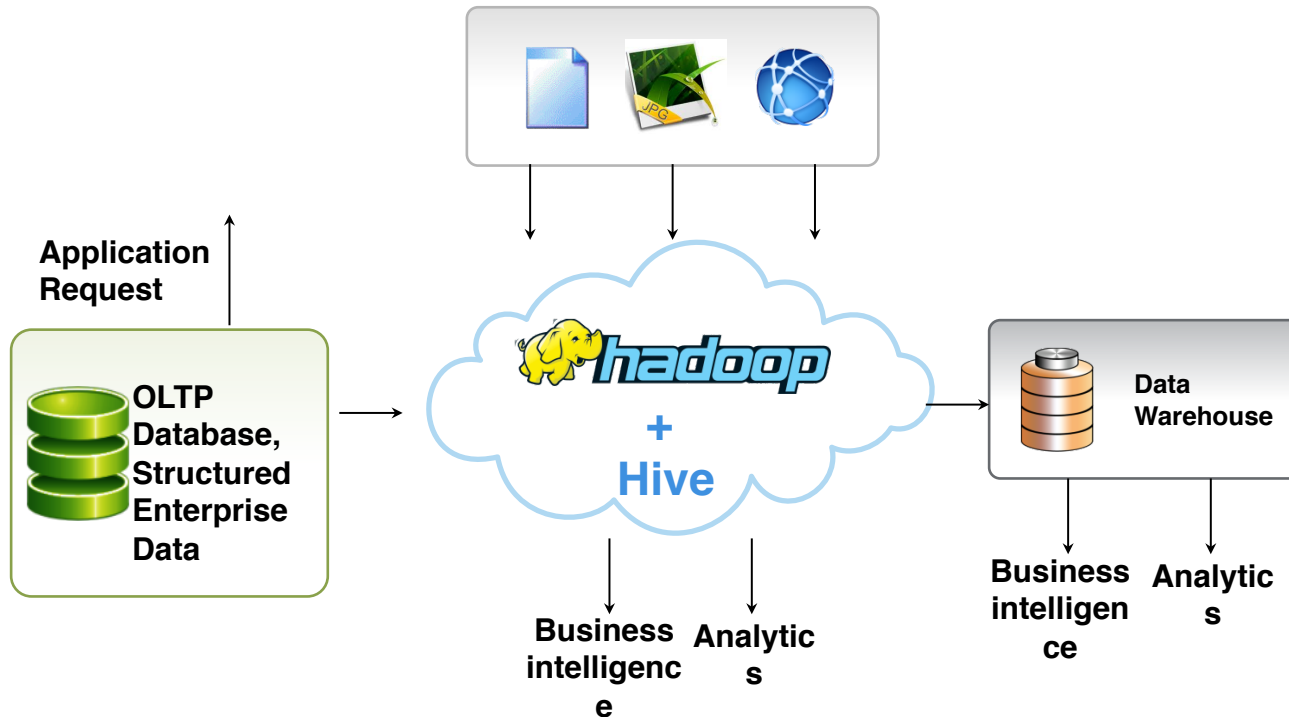
Optimized Hadoop & Big Data Environments: Use Proven Building Blocks for Success

	Title	Description	Reasoning	Known Customers & Markets
1	**Storage** (Hadoop + DBMS)	Integrated in the data processing pipeline Hadoop is used to pre-process unstructured data (i.e. Apache web server logs) before uploading the data in the Data Warehouse for advanced analytics	Hadoop licensing and support is 10x cheaper than DW	Sears (retail) Monsanto (life science) CBS (media) Home Depot (retail)
2	**Compute** (standalone Hadoop HDFS+MR)	Massively-parallel compute engine Hadoop is used to run complex compute-intensive analysis	In-house MR software skills Strong bias towards open-source	GE (broad variety) CODONiS (life science)
3	**Database** (Hadoop Hbase)	Massively-parallel NoSQL database Hadoop is used to run SQL queries against extremely large volumes of data	Leverage old HW In-house MR/SQL skills Strong bias towards open-source	Ricoh (content mgmt)

Cognizant

Thank you

Enterprise Information
Management

Cognizant