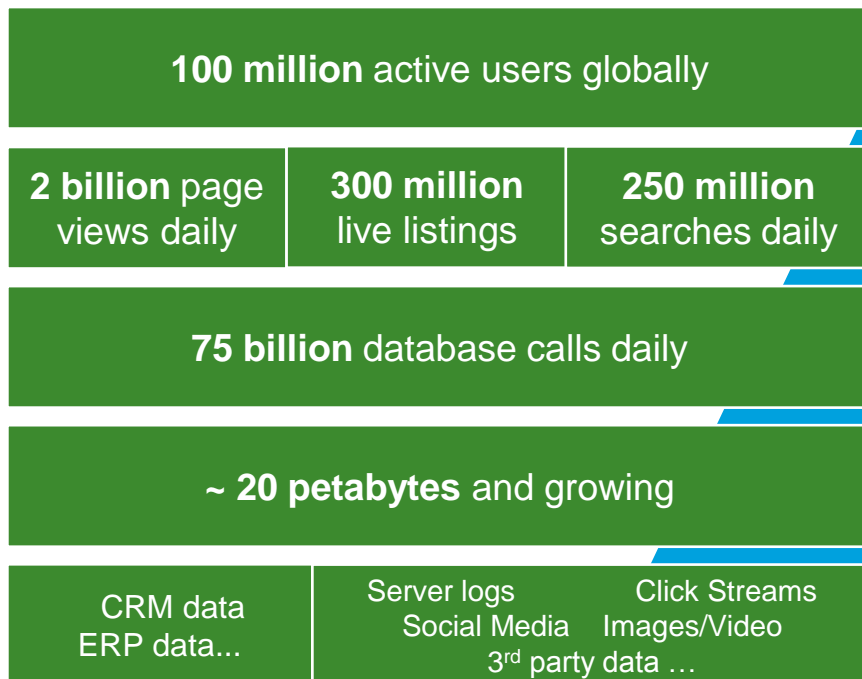


Big Data and its Dimensions

Big Data refers to internal and external data that is multi-structured, generated from diverse sources in near real-time and in large volumes making it beyond the ability of traditional technology to capture, manage and process within a tolerable amount of elapsed time.

eBay Big Data Environment



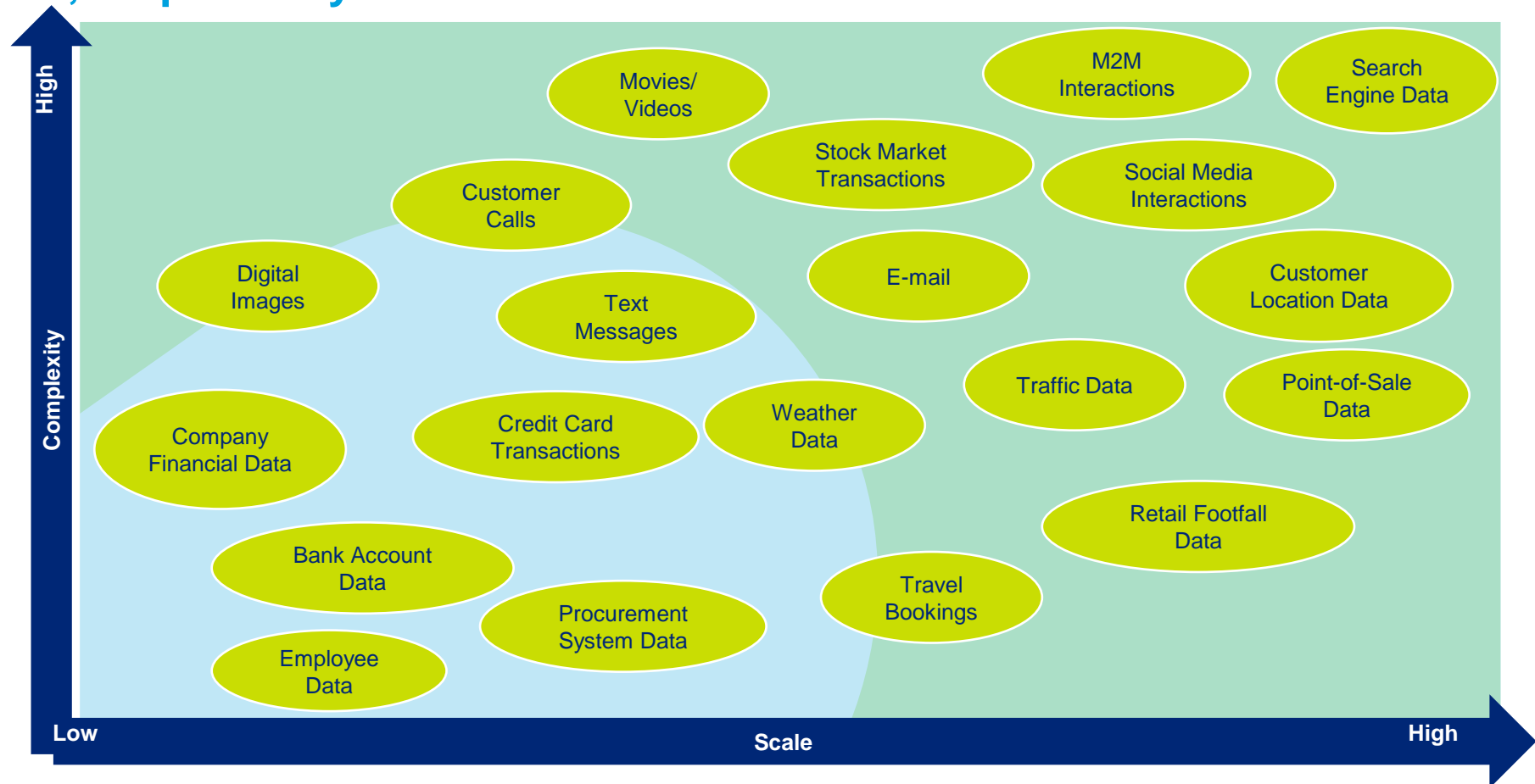
Big Data Dimensions

Volume	<ul style="list-style-type: none"> The sheer size of data in organizations is exploding from TB to PB
Velocity	<ul style="list-style-type: none"> The pace at which data is being generated today is significant
Variety	<ul style="list-style-type: none"> The data formats, structures and semantics are more diverse and inconsistent

Source: - <http://gigaom.com/cloud/under-the-covers-of-ebays-big-data-operation/>

Big Data Entails More than Just Growth in Data Volume

Big Data refers to enterprise data that is unstructured, generated from non-traditional sources, and/or real-time – in addition to being large in volume. Enterprises face the challenge and opportunity of storing and analyzing Big Data, respectively.



Note:

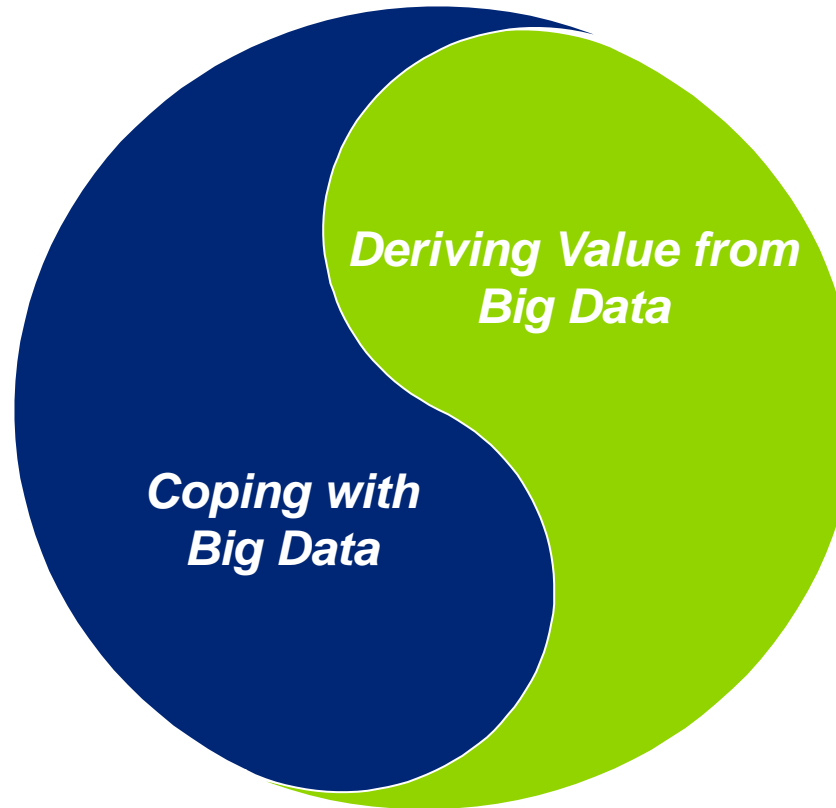
- Complexity refers to the different sources and formats of data.
- Scale refers to the volume and velocity of data.

 Denotes Big Data category

Big Data Entails More than Just Growth in Data Volume (2)

What are the Implications of Big Data for Enterprises?

- Enterprises cannot ignore the influx of data, mostly unstructured, and will need to increasingly invest in storage technology.
 - Gartner projects enterprise data to grow more than 650 percent over 2010-14. Moreover, eighty percent of the data will be unstructured.²



- Enterprises can leverage the data influx to glean new insights – Big Data represents a largely untapped source of customer, product, and market intelligence.
 - According to the 2011 IBM Global CIO Study, 83 percent of CIOs cited business intelligence and analytics as top priorities for their businesses over 2011-15.⁴

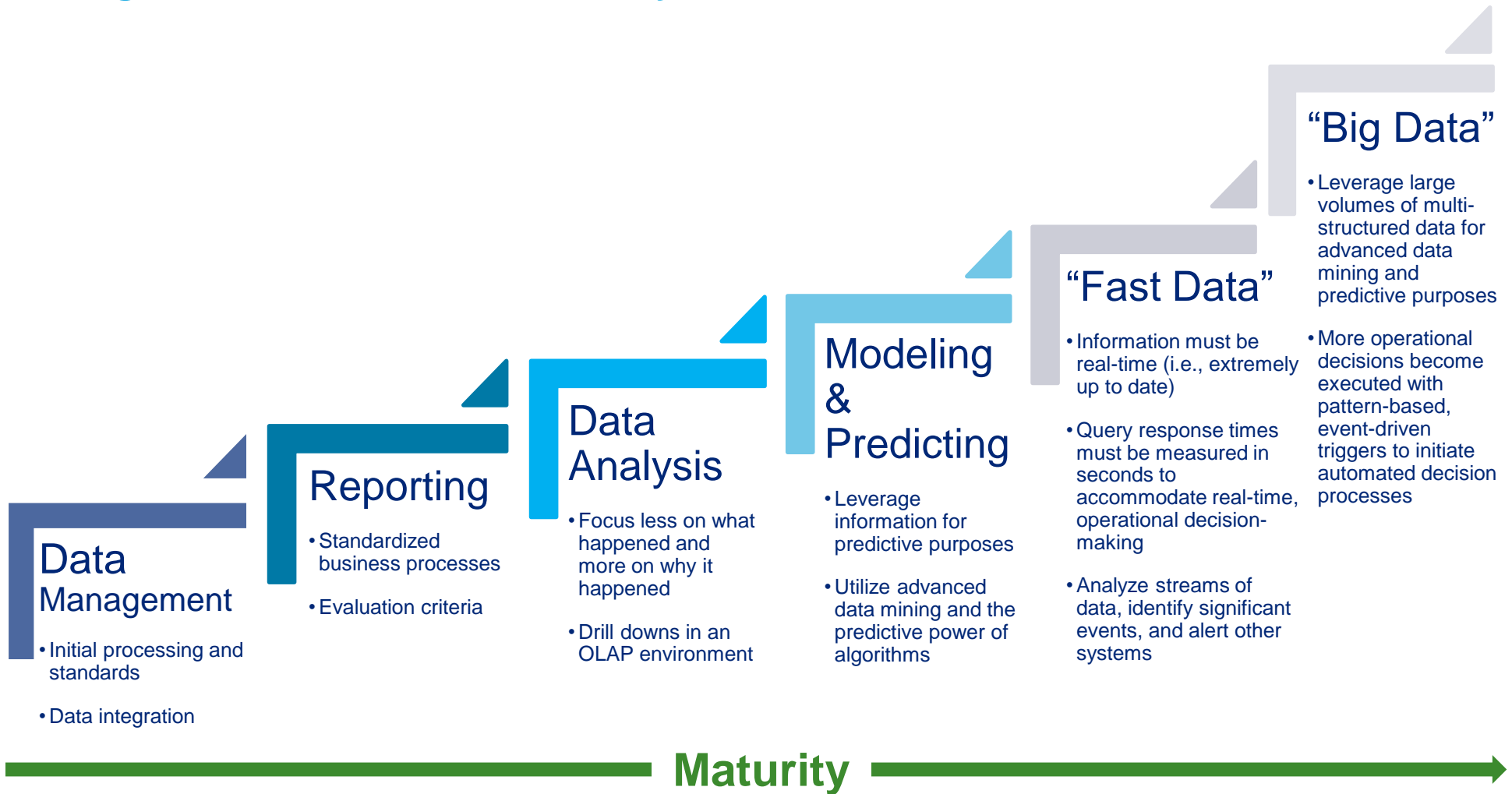
“Think of a stack of DVDs reaching from earth to the moon and back.”

– David Thompson, CIO, Symantec³

“The promise and scope of Big Data is that within all that data lies the answer to just about everything.” – Vivek Ranadivé, CEO, TIBCO³

Big Data, Information Management and Analytics

Big Data is the next step in the evolutionary path of Information Management and Advanced Analytics.



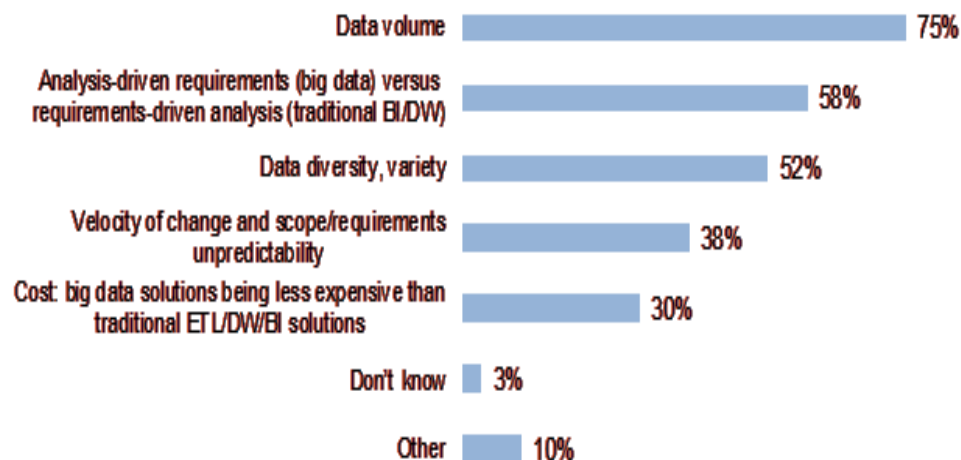
Source: TDWI, HighPoint Solutions, DSSResources.com, Credit Suisse, Deloitte

Current State of Big Data Initiatives

Big Data provides new tools and technologies that address problems in multiple functional areas across the enterprise.

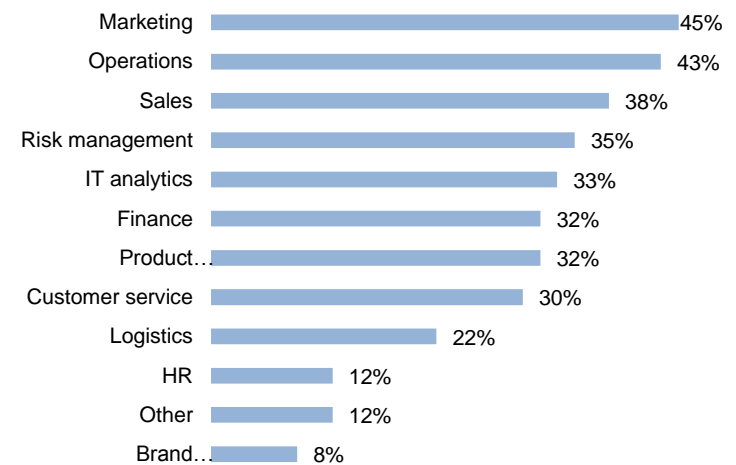
1. What are the main business requirements or inadequacies of earlier-generation BI/DW/ETL technologies, applications, and architecture that are causing you to consider or implement big data?

! In traditional BI and DW applications, requirements drive applications. Big data turns this model upside down

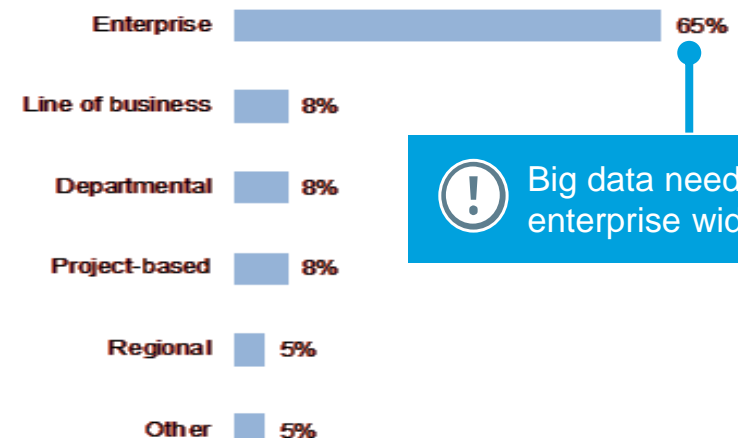


Source: Deloitte Research, Forrester

2. What enterprise areas does your big data initiative address?



3. What is the scope of your big data initiative?



! Big data need is enterprise wide

Current Big Data Challenges

Big Data provides opportunities however there are challenges that need to be addressed and overcome.

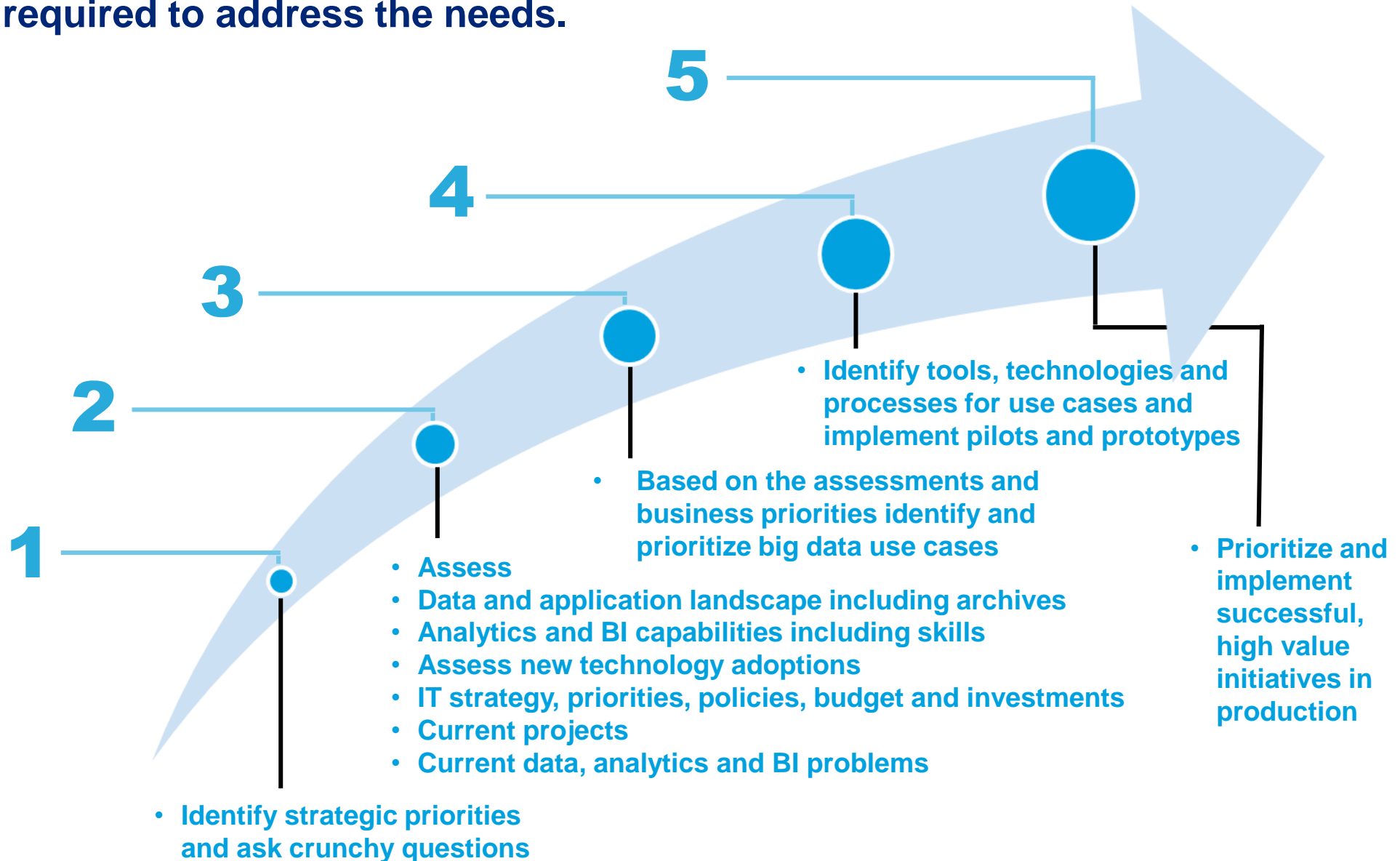
	Issue	Primary Challenge
Technology	Scalability	<ul style="list-style-type: none"> Flexibility of infrastructure to interact with extreme volume using a variety of data formats Cost and effort associated with scalability
	Integration	<ul style="list-style-type: none"> Increasing data volume, variety, and complexity results in increased time and monetary investments to remove barriers to compiling, managing and leveraging data across multiple platforms and systems Implications of standards and master data management
	Deployment	<ul style="list-style-type: none"> Identifying the best software and hardware solutions and determining the best overall infrastructure solution; internally, externally or using a combination Transitioning from legacy systems to newer technology
	Analytics	<ul style="list-style-type: none"> Considerable time and money invested to create algorithms that scale to big data volume and variety and improve user experience
Data	Data Quality	<ul style="list-style-type: none"> Compromise of quality due to volume and variety of data Cost of maintaining all data quality dimensions: <ul style="list-style-type: none"> Completeness, Validity, Integrity, Consistency, Timeliness, and Accuracy
	Governance	<ul style="list-style-type: none"> Identifying relevant data protection requirements and developing an appropriate governance strategy Reevaluation of internal and external data policies and regulatory environment
	Privacy	<ul style="list-style-type: none"> Privacy issues related to direct and indirect use of big data sources Evolving security implications of big data
People	Talent	<ul style="list-style-type: none"> Identifying and acquire the skill sets required to understand and leverage Big Data to add value

Asking Relevant Questions to Solve the Challenges

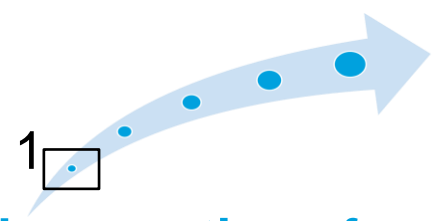
Area	Questions
Scalability	<ul style="list-style-type: none"> • What levels of availability and reliability are possible in mission-critical applications with large data volumes ? • Who can provide the best real-time scalability and storage? • Is the cost to value justified for building the infrastructure to handle Big Data ?
Integration	<ul style="list-style-type: none"> • With the increasing volume of data, what is the best infrastructure and integration strategy to keep up with the increasing demands from volume and processing speed. • How can non-traditional unstructured data be integrated with data stored in traditional transactional systems?
Deployment	<ul style="list-style-type: none"> • With an ever increasingly complex environment, how do we develop a strategy for integration and deployment? • Given the specialized nature of processing needed, is cloud computing an appropriate platform choice, and if so, what variant of cloud computing (public, private, hybrid) is needed?
Analytics	<ul style="list-style-type: none"> • With so much data, what are the right questions to gain additional insight that adds value to the enterprise? To our customers? • How can we tailor the user experience for analytic insights? • How can decision-makers comprehend the results of analyzing so much data quickly enough to act?
Data Quality	<ul style="list-style-type: none"> • With increasing volume of data, what data do we keep? And where do we focus on quality with the greatest returns? • How do we insure the quality unstructured data? • How do we measure the quality of extraction and processing procedures used with Big Data?
Governance	<ul style="list-style-type: none"> • What data governance is appropriate when analysis is distributed, needs change and data definitions and schemas evolve over time? • What intellectual property, licensing, and data protection considerations apply when Big Data environments are distributed across organizational and national boundaries? • How are regulations around audit trails and data destruction to be interpreted in a Big Data environment?
Privacy	<ul style="list-style-type: none"> • How does the company plan to address Personally Identifiable Information? • If externally hosted, what level of confirmation can we get that the provider will keep our customer data/information private? • How can security and privacy concerns be factored into the design of a Big Data environment to reduce vulnerability to external and internal threats?
Talent	<ul style="list-style-type: none"> • Does your organization have the skills sets necessary to leverage Big Data? • How can current IT skill sets best be leveraged in evolving the infrastructure to include Big Data?

Big Data Suggested Roadmap

A Big Data Roadmap begins with the decision makers and their crunchy questions and then proceeds to the data sources and technologies that are required to address the needs.



Identify Opportunities



Identifying strategic opportunities starts with asking crunchy questions for “sticky” business issues. This process is independent of the underlying data (volume, variety and velocity) and therefore applicable to both traditional and big data analytics.

Asking Crunchy Questions

Customers and social media

- What’s the buzz about your company online—and how could it impact sales forecasts?
- What are analysts saying about your organization? What about
- Customers and online influencers?
- Who are the next 1,000 customers you’ll lose—and why?
- Which trade promotion programs have the highest impact on
- profitability?
- What factors most influence customer loyalty? Why?
- How do factors such as politics and demographics affect the price your customers are willing to pay?
- Which factors have the most adverse effects on customer satisfaction?

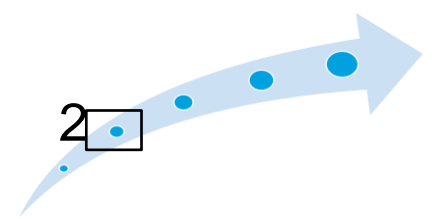
Sustainability and supply chain

- Which facilities are using more energy than they should?
- Which suppliers are at risk of going out of business?
- What is the impact of shipping costs on pricing?
- Which locations offer the best options for setting up your next distribution center?

Employees and risk

- Which new-hire characteristics best reflect your organization’s risk intelligence profile?
- Which are most likely to steal from you?
- Why do high-potential employees leave your company? What would cause them to stay?

Assess Current Capabilities



Assessment of current capabilities is crucial for identifying the scope and requirements for Big Data initiatives. Data and Application assessment can provide information about the data assets that can be leveraged. Current Analytics and BI capabilities assessment can provide information on the tools that can be integrated with Big Data technologies.

Data and Application Assessment

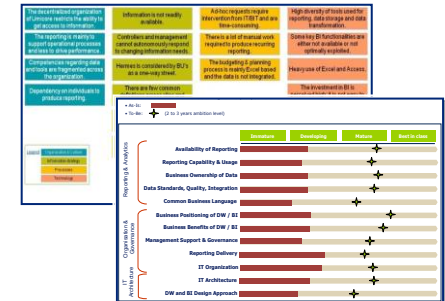
Current Data and Application landscape assessment should reveal –

- All data assets in the organization including archives
- All source systems in the organization and their dependency
- Associated data governance practices
- Data Standards, Quality and integration information
- Information about Master Data Management and maturity

Analytics and BI Capabilities

Current Analytics and BI capabilities assessment should reveal –

- All Analytics and BI capabilities and maturity
- Ownership and usage of Analytics and BI
- Related skill sets
- Analytics and BI practices and policies

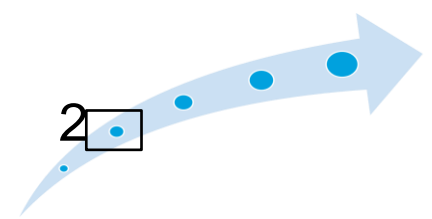


Other Considerations

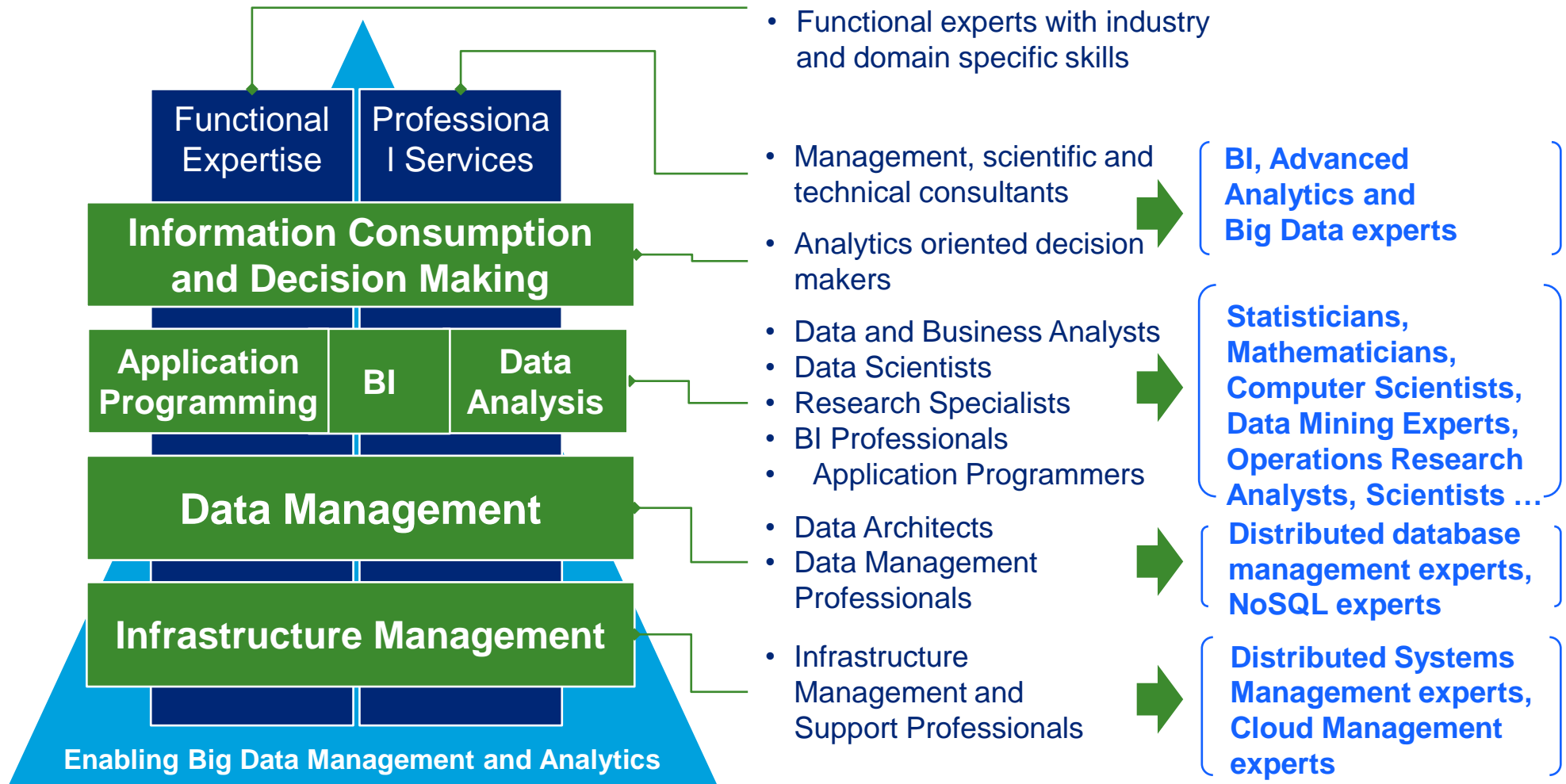
Assessment phase should consider –

- Analytics and BI pain points assessment
- IT strategy, priorities, policies, budget and investments
- Current Analytics and BI related project
- Other potential projects where Big Data technologies can be leveraged

Assess Big Data Analytic Skill Sets

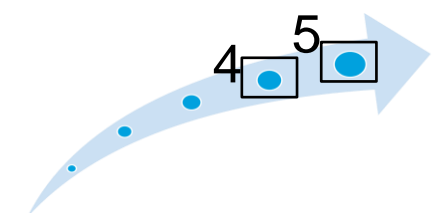


The drive to find and employ new skills sets comes with the need to manage rapidly increasing quantities of data and the desire to extract valuable insights from it.

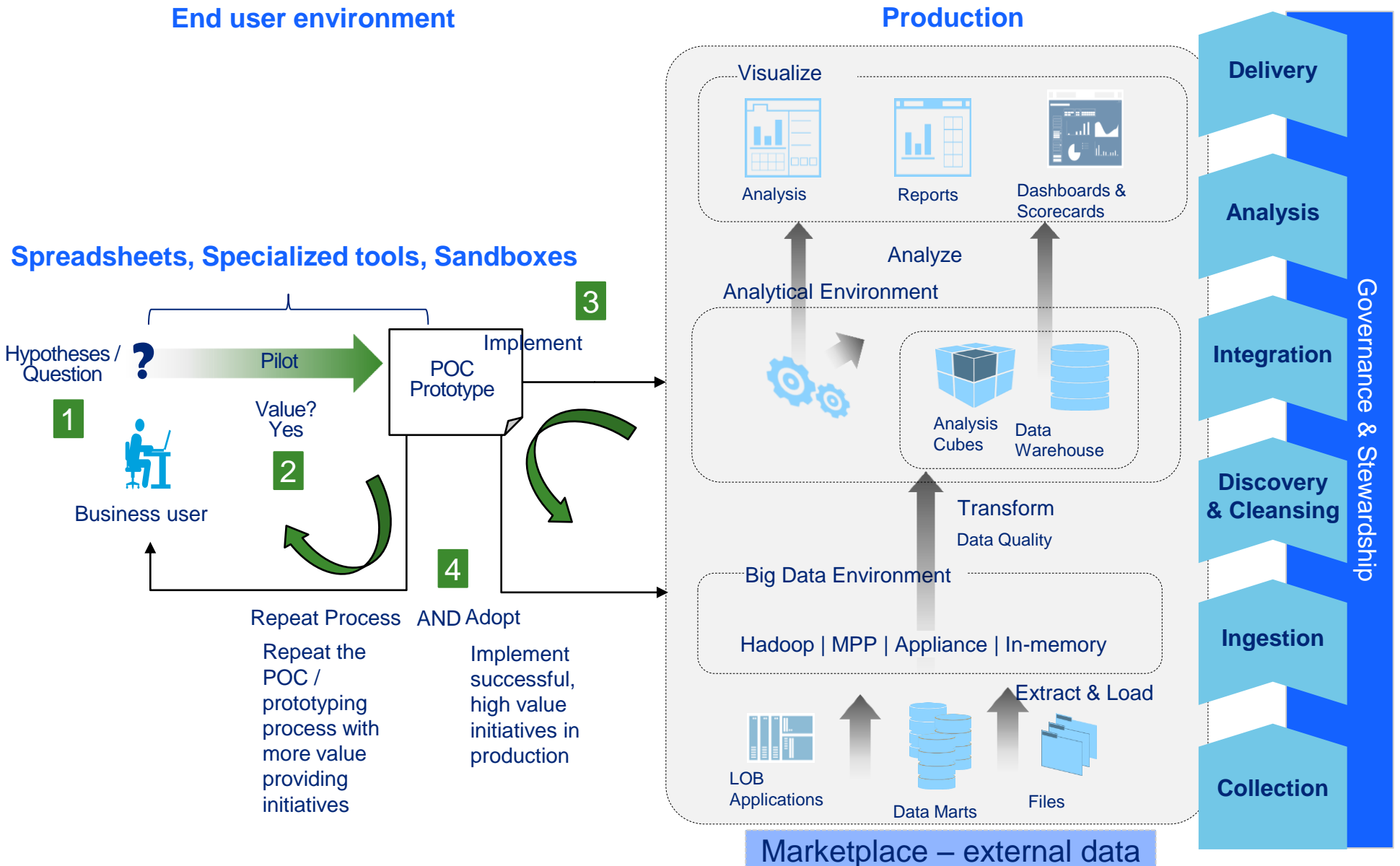


Besides Data Growth, Technology Advances Are Driving The Competitive Landscape for New Talent!

Pilot and Adopt



Valuable time and money can be saved by adopting a business user driven pilot/prototyping approach that targets value providing initiatives.



Bringing Big Data into the current Business Ecosystem

Big Data introduces new technologies and tools for coping with the volume, velocity, and variety that characterize data sources in current business ecosystem. The opportunities are exciting however a multitude of difficult questions need to be answered.

- What **data sources** should be collected and how can they be acquired efficiently?
- How is **data quality** managed across so many sources of data, many of which come from outside the organization, such as public social networks?
- What **structure** can be derived from non-traditional data sources (documents, Web logs, video streams, etc.) to make storage, analysis, and ultimately decision-making easier?
- How can non-traditional unstructured data be **integrated** with data stored in traditional transactional systems?
- How can decision-makers **comprehend the results** of analyzing so much data quickly enough to act?
- What **data governance** is appropriate when analysis is distributed, needs change and data definitions and schemas evolve over time?

Bringing Big Data into the current Business Ecosystem (2)

- What **architectures and algorithms** can be used to decompose problems and data for rapid execution in parallel environments?
- What levels of **availability and reliability** are possible in mission-critical applications, when data volumes are so large?
- Is specialized **hardware** required for a particular need, or can low-cost commodity hardware be leveraged to scale processing?
- Given the specialized nature of processing needed, is **cloud computing** an appropriate platform choice, and if so, what variant of cloud computing (public, private, hybrid) is needed?
- How can **security and privacy** concerns be factored into the design of a Big Data environment to reduce vulnerability to external and internal threats?
- How **are regulations around audit trails** and data destruction to be interpreted in a Big Data environment?
- What intellectual property, licensing, and data protection considerations apply when Big Data environments are distributed across organizational and national boundaries?
- How can current IT **skill** sets best be leveraged in evolving the infrastructure to include Big Data?

Big Data Analytic Skill Sets

The drive to find and employ new skills sets comes with the need to manage rapidly increasing quantities of data and the desire to extract valuable insights from it.

Title	Skill Set	Business Area of Expertise	Business Need
Data Analyst	Provide design and execution capabilities for the collection and analysis of data.	<ul style="list-style-type: none"> Data collection 	
Data Architect (Data Scientist)	Provide oversight and direction for monitoring and implementing the principles governing the design and evolution of systems including the data assets, components, relationships, and the fundamental organization of systems.	<ul style="list-style-type: none"> System Design Policy Guidance 	
Engineer (Research Analyst / R&D Specialist)	Provide expertise on the latest software engineering tools, oversees the construction of precise and large-scale data sets, and performs database research to recommend ways to improve the quality of data.	<ul style="list-style-type: none"> Software Engineering Tools Database Research Data Quality 	
BI Professionals (incl. Directors and Specialists)	Provide strategic design and implementation of BI software and systems, which include database and data warehouse integration. Provide implementation support for new enterprise systems and applications.	<ul style="list-style-type: none"> Systems Applications Life Cycle Development 	
Other (Marketers, Consultants, Statisticians, Data Governors, Risk Managers, etc.)	General operational workforce to supplement the business needs around data. Some additional skills sets provided by this group are planning data collection, data types and sizes, and identifying relationships and trends in data and factors that could affect the results of research.	<ul style="list-style-type: none"> Analyze and interpret statistical data Adapt statistical methods Design research projects 	
Other ???	Other skills sets required to augment the increased workforce? (Ex: HR?)	<ul style="list-style-type: none"> ??? 	

There is a point of view that...

- Virtually every industry will be transformed by analytics and data
- The transformation will require hard and soft capabilities
- Ultimately it's about making better decisions
- This is not business as usual – it's an historic opportunity to transform your business!



Questions to Consider...

There is a claim that 90% of all data has been produced in the past two years. Is this nonsense or do you see data sources increasing exponentially at this rate?

Are there new types of data, outside of the GL and other structured data, that you are being asked to capture for analysis?

Overall, is the business asking you to provide detailed and huge data volumes ready for analysis..... that can produce significant ROI? What are the hot areas? Are these requests outside of your current system capabilities?



Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/mx/aboutus for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

Deloitte provides audit, tax, consulting, and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and high-quality service to clients, delivering the insights they need to address their most complex business challenges. Deloitte has in the region of 200,000 professionals, all committed to becoming the standard of excellence.

As used in this document, "Deloitte" means Deloitte Consulting Group, S.C., which has the exclusive legal right to engage in, and limit its business to, providing consulting services (including information technology, strategy and operations, and human capital) in Mexico, under the name "Deloitte".

This publication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or their related entities (collectively the "Deloitte Network") is, by means of this publication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person or entity who relies on this publication.