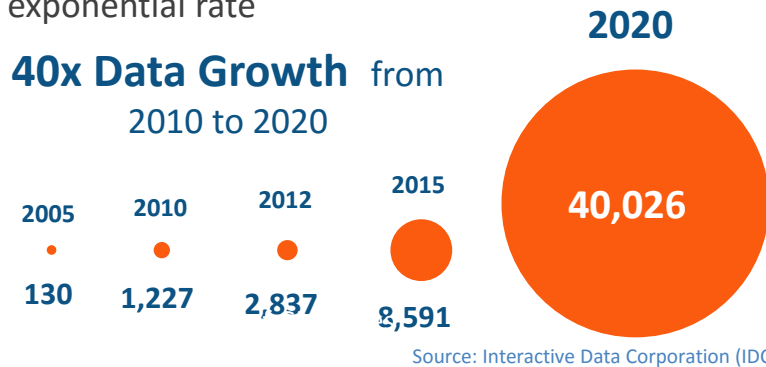# DScVR: Data Science Value Realization

November 2016

Despite increased opportunities to capitalize on data and analytics, clients struggle with developing their data science functions

Exponential Data Growth

Generation and collection of data is growing at an exponential rate

40x Data Growth from 2010 to 2020

2020

Year	Value
2005	130
2010	1,227
2012	2,837
2015	8,591
2020	40,026

Source: Interactive Data Corporation (IDC)

New Tools and Approaches

New open source technologies enable companies efficient access to and analysis of more data than ever before

GridGain
Zettaset
hortonworks
ORACLE
APACHE DRILL
calpont

Analytics-Driven Insights

Analytic insights are **driving action across industries** increasing revenue, minimizing risk, and decreasing costs

Healthcare Providers   Services
Utilities
Government   Education
Insurance   **Retail**
Pharmaceuticals and Life Sciences

The Struggle Is Real

85% of Fortune 500 companies will be unable to exploit big data for competitive advantage by 2016
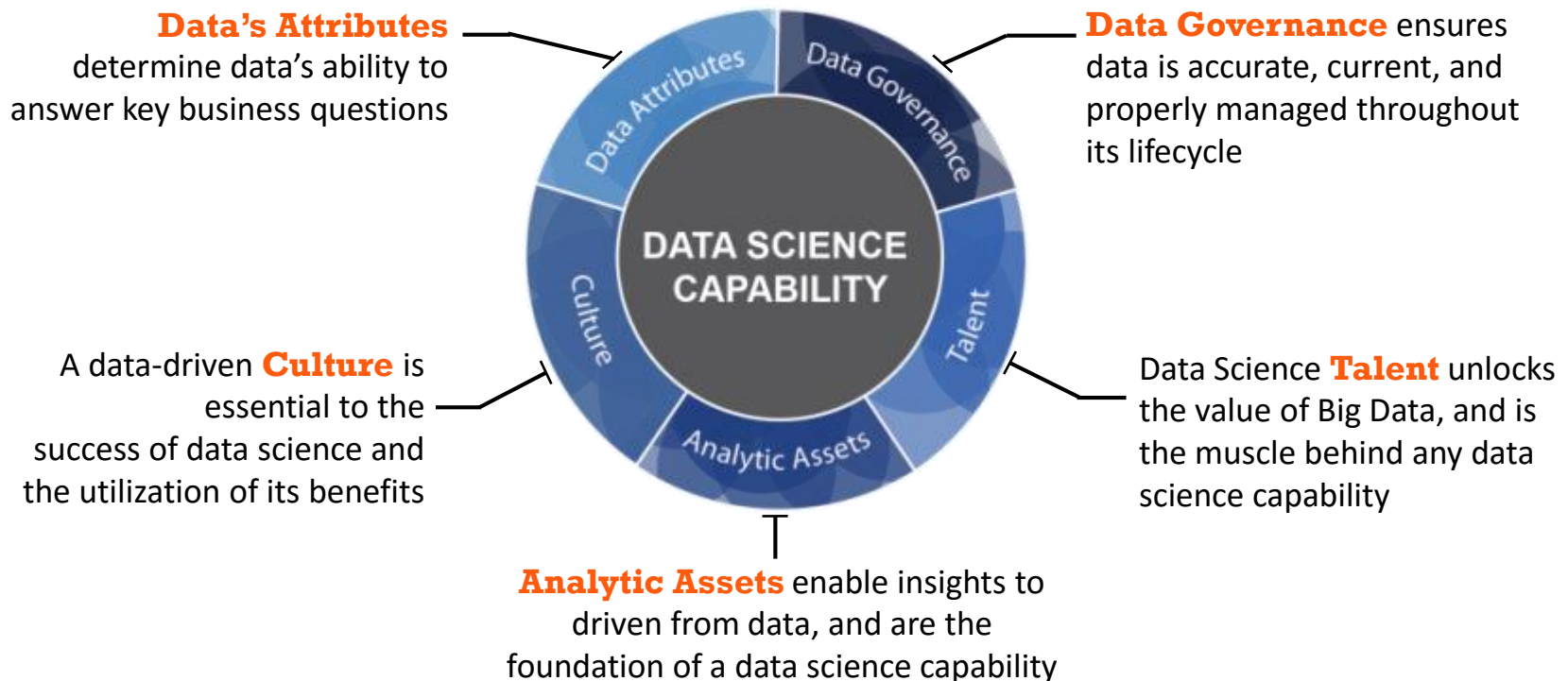
90% Information assets from advanced analytics that are siloed and unused across business processes

25% Big data implementations that will fail to deliver value resulting from inadequate infrastructure
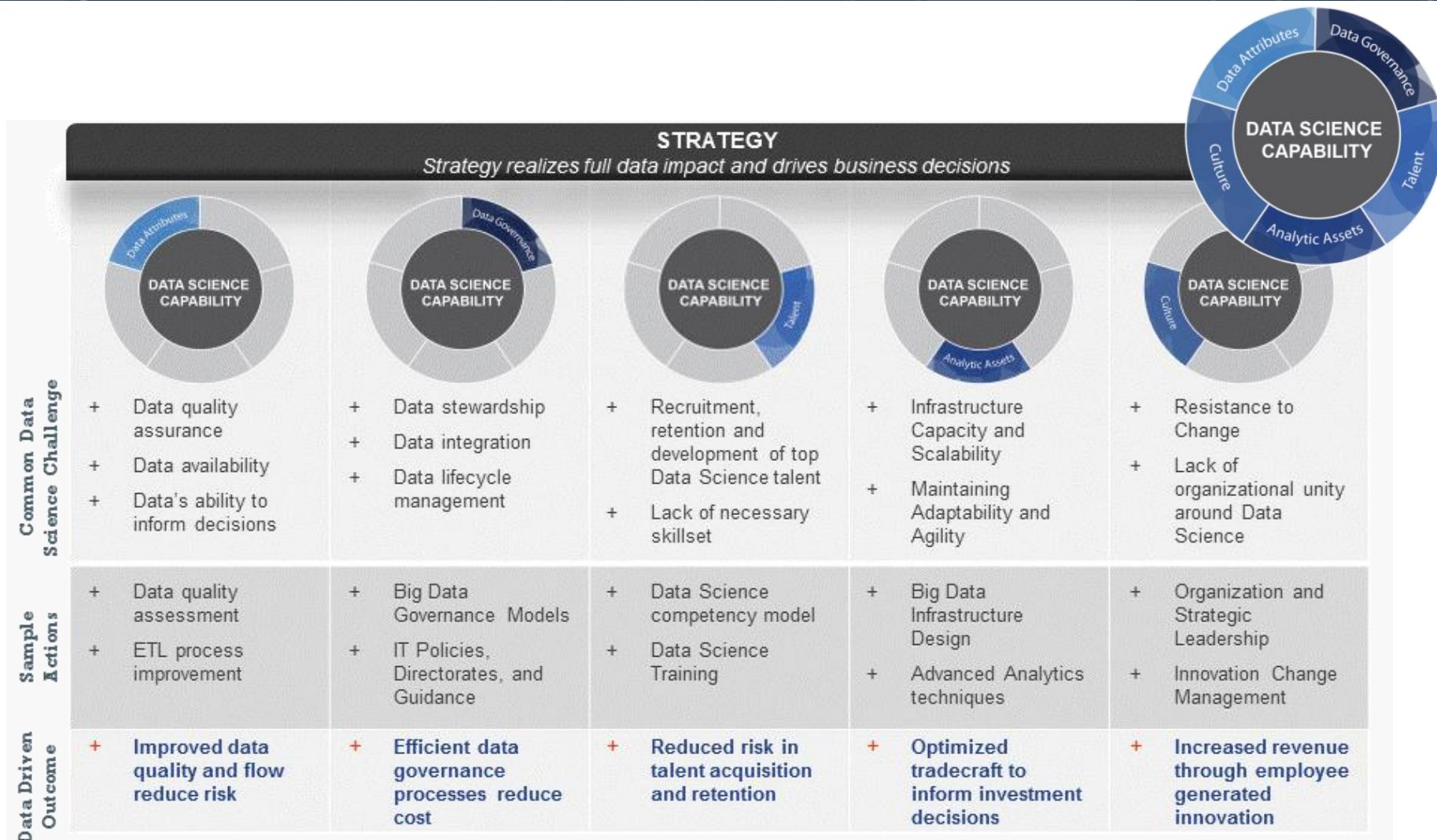
56% Organizations that report not knowing how to get value from big data
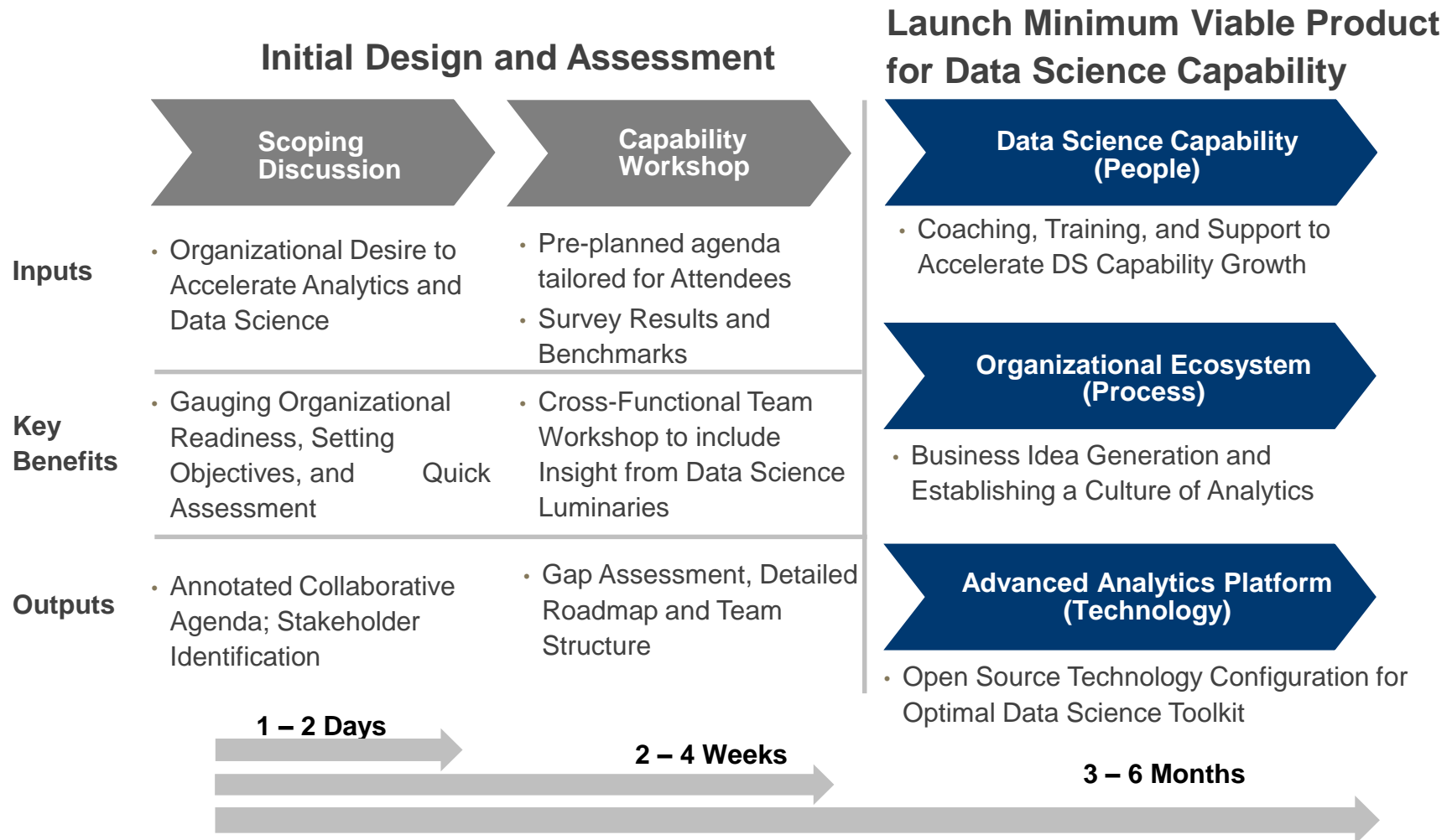
Booz | Allen | Hamilton

Booz Allen unraveled this problem, by studying and defining the make-up of a high-performing data science capability

Data's Attributes determine data's ability to answer key business questions

Data Governance ensures data is accurate, current, and properly managed throughout its lifecycle

A data-driven **Culture** is essential to the success of data science and the utilization of its benefits

Data Science **Talent** unlocks the value of Big Data, and is the muscle behind any data science capability

Analytic Assets enable insights to driven from data, and are the foundation of a data science capability

Booz | Allen | Hamilton

Our Data Science Value Realization (DScVR) solution helps clients understand their capacity to generate value from data

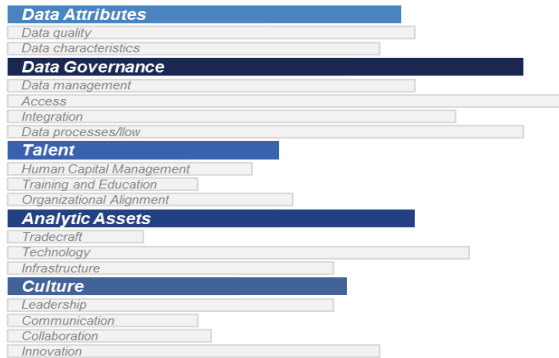|	STRATEGY — Strategy realizes full data impact and drives business decisions				
Common Data Science Challenge	+ Data quality assurance + Data availability + Data's ability to inform decisions	+ Data stewardship + Data integration + Data lifecycle management	+ Recruitment, retention and development of top Data Science talent + Lack of necessary skillset	+ Infrastructure Capacity and Scalability + Maintaining Adaptability and Agility	+ Resistance to Change + Lack of organizational unity around Data Science
Sample Actions	+ Data quality assessment + ETL process improvement	+ Big Data Governance Models + IT Policies, Directorates, and Guidance	+ Data Science competency model + Data Science Training	+ Big Data Infrastructure Design + Advanced Analytics techniques	+ Organization and Strategic Leadership + Innovation Change Management
Data Driven Outcome	+ **Improved data quality and flow reduce risk**	+ **Efficient data governance processes reduce cost**	+ **Reduced risk in talent acquisition and retention**	+ **Optimized tradecraft to inform investment decisions**	+ **Increased revenue through employee generated innovation**

Booz | Allen | Hamilton

We apply DScVR to accelerate your approach to Data Science and hone investments to critical areas

Initial Design and Assessment

Launch Minimum Viable Product for Data Science Capability

	Scoping Discussion	**Capability Workshop**	
Inputs	• Organizational Desire to Accelerate Analytics and Data Science	• Pre-planned agenda tailored for Attendees • Survey Results and Benchmarks	**Data Science Capability (People)** • Coaching, Training, and Support to Accelerate DS Capability Growth
Key Benefits	• Gauging Organizational Readiness, Setting Objectives, and Quick Assessment	• Cross-Functional Team Workshop to include Insight from Data Science Luminaries	**Organizational Ecosystem (Process)** • Business Idea Generation and Establishing a Culture of Analytics
Outputs	• Annotated Collaborative Agenda; Stakeholder Identification	• Gap Assessment, Detailed Roadmap and Team Structure	**Advanced Analytics Platform (Technology)** • Open Source Technology Configuration for Optimal Data Science Toolkit

1 – 2 Days

2 – 4 Weeks

3 – 6 Months

Booz | Allen | Hamilton

DScVR products include an assessment of clients' current capability, tailored recommendations, and an implementation roadmap

Capability Scorecard

Organizational Network Map

Focused Recommendations

1-2 Year Roadmap

Booz | Allen | Hamilton
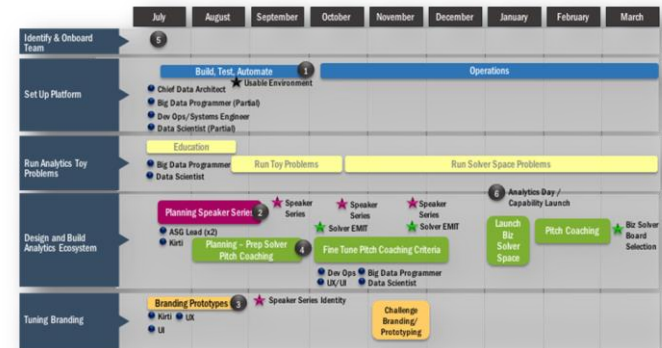
Reach out to our DScVR Points of Contact for more information

DScVR Points of Contact

EZMERALDA KHALIL
Principal
khalil_ezmeralda@bah.com

CUTTER BRENTON
Lead Associate
brenton_cutter@bah.com

SUSAN MICHENER
Lead Associate
michener_susan@bah.com

ARI HAMALIAN
Lead Associate
hamalian_ari@bah.com

ABOUT BOOZ ALLEN

Booz Allen brings its pioneering work in advanced analytics—and the industry-leading expertise of its more than 500-member data science team—to transform our clients' data into actions that keep them competitive in today's data-driven economy. The first ever National Data Science Bowl, along with Booz Allen's recently launched Explore Data Science training program and Field Guide to Data Science is part of the firm's ongoing commitment to supporting data science education and awareness.

Booz | Allen | Hamilton

Appendix

Our framework first evaluates attributes of internal and external data sets to analyze their potential to drive key business decisions
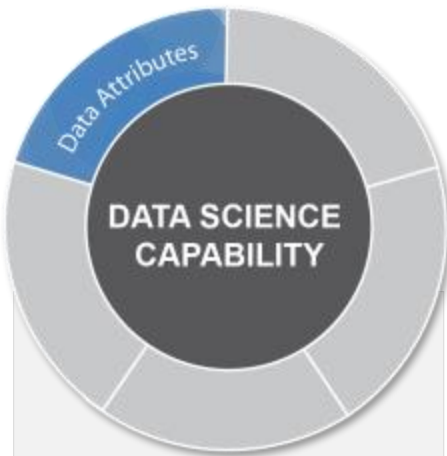
Data Attributes determine data's ability to answer key business questions

DATA ATTRIBUTES

DATA QUALITY

DATA CHARACTERISTICS

DEFINITION

- Data's ability to meet its intended purpose in the decision making process

- The kind of data an organization has available
- Data size, speed, movement, features, etc.

PURPOSE

- Determine the current value of an organization's data through a quality assessment

- Establish the types of data relevant to the organization
- Develop an understanding of the balance between data confidentiality, integrity and availability

Booz | Allen | Hamilton

We then review existing data governance structures to understand the influence of data access and flow within the lifecycle

DATA GOVERNANCE

DATA SCIENCE CAPABILITY
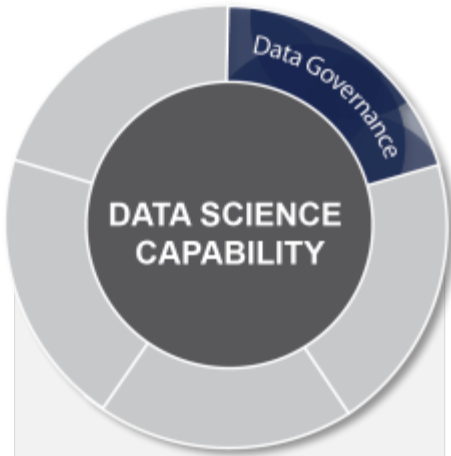
Data Governance

Data Governance ensures data are accurate, current, and properly managed throughout their lifecycle

	DATA MANAGEMENT	ACCESS	INTEGRATION	FLOW
DEFINITION	+ The responsibility of organizations to proactively maintaining data throughout its lifecycle	+ How organizations access and interact with data and resulting analytic insights	+ How disparate data sets are combined and used together	+ How data moves throughout the organization
PURPOSE	+ Understand data management principles and processes currently utilized by the organization	+ Understand how data access is tied to organizational hierarchy + Understand how data access is influenced by data characteristics	+ Identify processes that integrate and aggregate data + Understand how data is localized prior to integration	+ Understand the processes that promote data flow, both within the organization and between business units + Understand the data lifecycle and associated business rules

Booz | Allen | Hamilton

Our framework also assists in gauging whether an organization has properly aligned its talent and data strategies
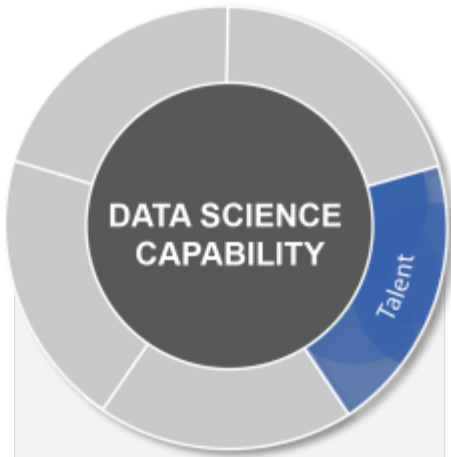

Data Science **Talent** unlocks the value of Big Data, and is the heart of any Data Science capability

TALENT

	HUMAN CAPITAL MANAGEMENT	TRAINING & EDUCATION	ORGANIZATIONAL ALIGNMENT
DEFINITION	+ The recruitment, retention and alignment of talent with the proper combination of skills	+ The cultivation of Data Science talent with the desired skill sets within an organization	+ Organizational structures, reporting relationships, roles and responsibilities, and processes that align data
PURPOSE	+ Assess current human capital management policies/strategies to perform root-cause analysis of acquisition, retention, and performance issues	+ Research the processes by which Data Science talent is developed within the firm + Evaluate the resources devoted to development of Data Science talent	+ Understand how Data Science talent is aligned within the organization

Booz | Allen | Hamilton

We then characterize an organization's analytic assets to assess the supporting IT infrastructure, analytics tools, and methodologies

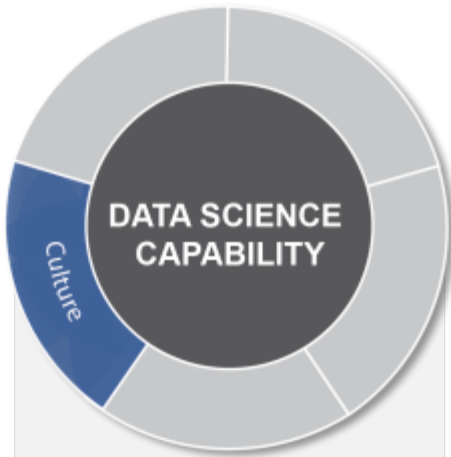**DATA SCIENCE CAPABILITY**
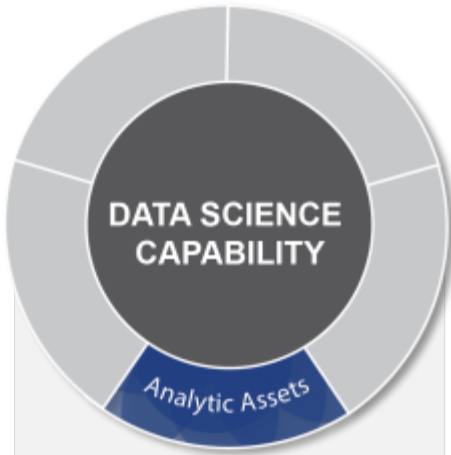
Analytic Assets

Analytic Assets, enable insights to be derived from data and are the foundation of a Data Science capability

ANALYTIC ASSETS

	TRADECRAFT	INFRASTRUCTURE	TECHNOLOGY
DEFINITION	+ The analytic methodologies employed by an organization	+ The IT systems and architectures employed by an organization to support data and data analysis	+ The tools utilized by an organization to generate and distribute analytic insights
PURPOSE	+ Develop a contextual understanding of analytic tradecraft + Research the organization's analytic history	+ Evaluate the organizations current infrastructure and its role in supporting analytics + Assess current infrastructure	+ Identify and assess technology used to facilitate data and analytic processes

Booz | Allen | Hamilton

Lastly, we identify ways to enhance underlying elements of a data-driven culture such as communication, collaboration, and innovation

CULTURE

DATA SCIENCE CAPABILITY

Culture

A data-driven **Culture** is essential to the success of any Data Science capability and the utilization of its benefits

	LEADERSHIP	COMMUNICATION	COLLABORATION	INNOVATION
DEFINITION	+ The level of devotion to analytically-driven decision making	+ The socialization of analytic insight and the transparency of decision making	+ The alignment of data and decision making	+ An organization's willingness to explore and its commitment to accepting change
PURPOSE	+ Develop an understanding of existing decision making processes + Evaluate the way in which leaders utilize analytics, and how they desire to use it in the future	+ Develop an understanding of how Data Science is perceived within the firm and if it differs from the Data Science strategy	+ Develop an understanding of if/how different business units collaborate on analytic activities	+ Baseline the organization's balance between open innovation and calculated risk + Gain insight on how innovation is fostered/driven within the organization