DATA & ANALYTICS HANDBOOK

Prepared by the Data & Analytics Working Group

Canadian Electricity Association | Association canadienne de l'électricité

Canadian Electricity Association | Association canadienne de l'électricité

Data & Analytics Handbook

Version Control

Version	Date	Description	Author(s)
1.0	August 2020	Initial Members-Only Release – full edition	Data and Analytics Working Group
2.0	January 2021	Public Version – redacted	Data and Analytics Working Group

About the Canadian Electricity Association

Founded in 1891, the Canadian Electricity Association (CEA) is the national forum and voice of the evolving electricity business in Canada. The Association contributes to the regional, national and international success of its members through the delivery of quality value-added services and by promoting electricity as a key social, economic and environmental enabler that is essential to Canada's prosperity.

CEA members generate, transmit, and distribute electrical energy to industrial, commercial, residential, and institutional customers across Canada every day. Members include integrated electric utilities, independent power producers, transmission and distribution companies, power marketers, and the manufacturers that keep the industry running smoothly.

Purpose

This document is provided to the Data Analytics Working Group (DAWG) community for educational purposes. CEA nor DAWG warrant that it is suitable for any other purpose and makes no expressed or implied warranty of any kind and assumes no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information contained herein.

The CEA DAWG may be updating this document from time to time to ensure that it is kept up to date and that it remains a valuable reference for the future.

Confidentiality

This handbook contains utility processes and operating procedures. This public version of the handbook has been prepared for general consumption. If more details are required, a membership to the CEA is also required. To protect the confidentiality of members specific cases will use common language such as ‘company’ rather than identifying the specific utility.

For further inquiries on this document please contact info@electricity.ca

Canadian Electricity Association | Association canadienne de l'électricité

Contributing Utilities

Canadian Electricity Association | Association canadienne de l'électricité

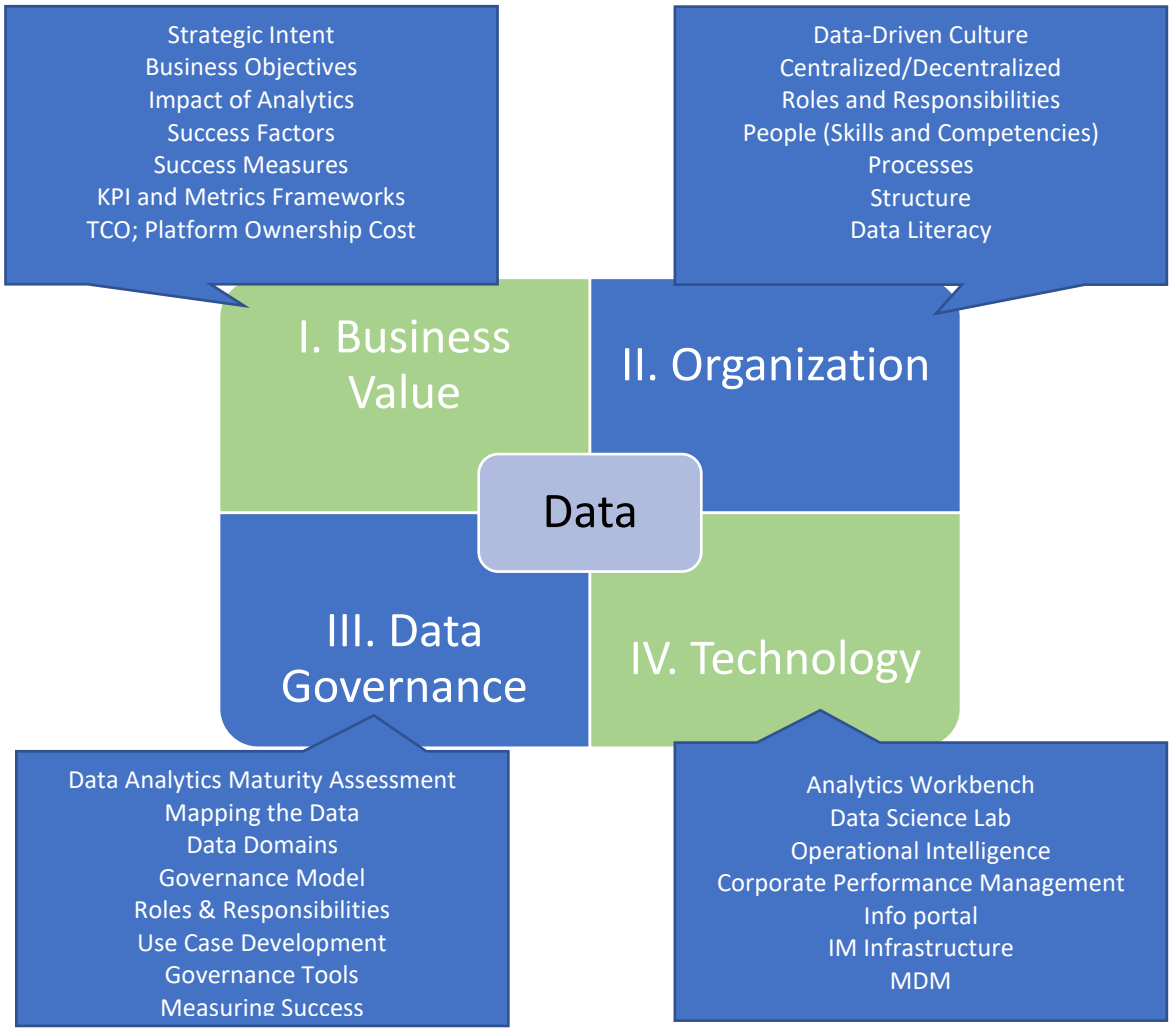
Contents

Acknowledgements

The CEA's Data and Analytics Working Group ("DAWG") is a collaborative and information exchange forum for data strategies, data insights & analytics and the enabling technologies. The group searches, shares and generates ideas, best practices and approaches around implementing effective organizational structures, tools and policies, effective governance models, and priorities to determine the highest value return from analytics for CEA's member utilities.

In 2019, the DAWG has embarked on an exploration of ways to build data strategies for our data-intensive organizations to improve data & analytics maturity scores, capabilities, tools and overall value extracted from enterprise-wide data.

Following the group discussions, it was decided that the DAWG's work in 2019/2020 would be dedicated to four distinct areas that are critical to the success of an organizational data strategy.



I. Business Value

The first question we ask when presented with a proposal to develop a data strategy for our organization is **WHY?**

Understanding the importance of demonstrating the value proposition in the most tangible ways, the members of the DAWG have formulated a representation of a) **key use cases**, along with a framework for prioritizing these use cases, and b) assembled a **use case repository** based on the experience of their own organizations.

The summary and examples provided below showcase the broad range of applications for strategic data use across all areas of the organizational value chain.

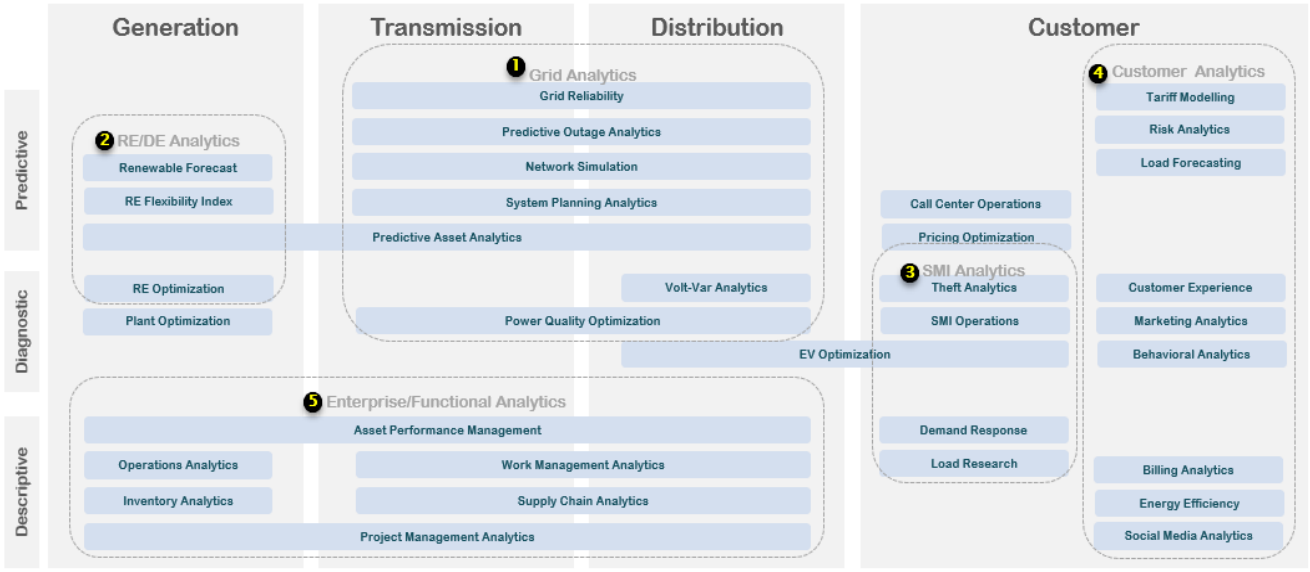
a) Key Use Cases

Major categories of utility analytics use cases

#	Use Case*	Definition
1	Grid Analytics	Use cases directly impacting transmission and distribution network operations including asset management.
2	RE/DE Analytics	Use cases supporting virtual power plant (VPP) and nontraditional generation source integration to the grid.
3	SMI/AMI Analytics	Use cases leveraging smart meter/ sensors data to improve AMI operations.
4	Customer Analytics	Use cases directly impacting customer processes, experience and perceived customer value to the utility.
5	Functional/Enterprise Analytics	Use cases to improve corporate operations like supply chain management, field crew performance, work management, project management etc.

** In the Technology section, under Reference architecture, we will present the activation of at least one-use case from each domain.*

Analytical Use Cases & Value Towers



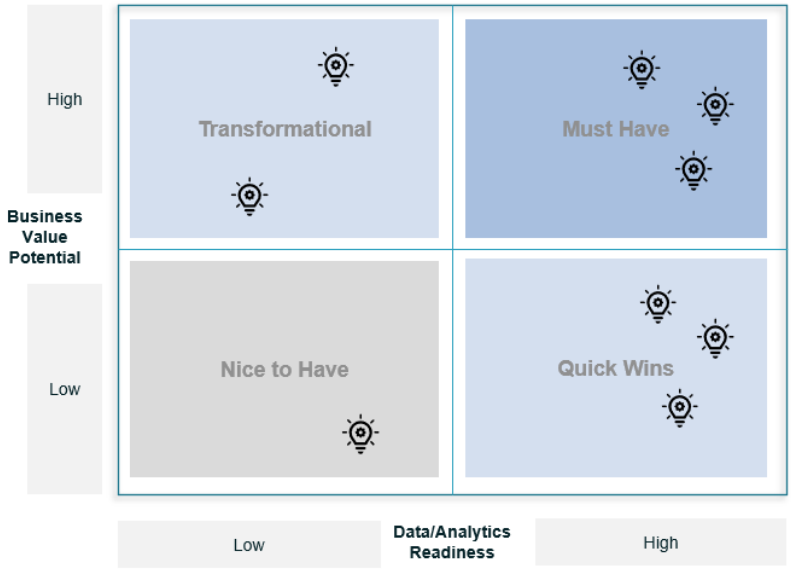
b) Use-Case Repository

Understanding the importance of demonstrating the value proposition in the most tangible ways, the members of the DAWG have assembled a list of use case based on the experience of their own organizations, that showcase the broad range of applications for strategic data use across all areas of the organizational value chain.

- This library of use cases will be kept as current as possible, as new cases are being added or existing cases finalized. The use-case repository is for active members of the committee and are not provided in this version of the handbook. Use-cases are categorized in the following functional area.
    - Customer
    - Distribution Grid
    - Transmission and Power Generation
    - Metering
    - Finance
    - Marketing
    - Operation and Information Technology
    - Human Resources

A Framework for Prioritizing Use-Cases

- **Must have:** Easy-to-build analytics use cases (e.g., those that require low effort/incremental investment/little or no data cleansing), which bring large business benefits to the utility.
- **Quick wins:** Easy-to-build use cases with low to medium benefits and which need minimal efforts for improving data/analytics readiness.
- **Transformational:** Use cases that are difficult to build but bring large business benefits to the utility (e.g., significant effort is required to improve data/analytics readiness).
- **Nice to have:** Difficult-to-build use cases that bring low to medium business benefits to utilities that need data/analytics readiness improvement



Other Reflections and Examples:

Example: By the end of 2019 an organization had already established its data governance framework in 5 key business units with well-defined data stewardship responsibilities the effort continued by adding 3 more business units in 2020, and more after, until the entire organization adopts the chosen data governance model.

- Each business unit data steward (DS) can apply the data governance framework to up to 10 initiatives (use-cases) at the same time. The CDS will then prioritize the initiatives for their respective business unit. And with a clear direction towards the digital transformation, analytics initiative such as algorithms, KPIs, machine learning, get to sit higher on the priority stack of the many organizational priorities.

Example: A company has been strategically assessing the highest value use cases that would demonstrate "big wins" in a relatively short period of time and create a healthy data governance conversation across the organization. Here are some examples:

- Enhancing the load forecasting model by adding to the granularity and currency of data inputs
- Refreshing the main data warehouse, a system that touches a significant number of units and supports critical functions for the organization's reliability mandate.
- Improving the external reporting capabilities and interfaces - the organization needs to better understand the needs and wants of the external users and define a strategy for external data presentment.

II. Organization

Business Intelligence and Analytics Enterprise Operating Models

- An operational model defines how a Business Intelligence and Analytics (BIA) team delivers services to the business, in collaboration with IT and business stakeholders, and how BIA efforts are to be governed.
- Deploying the right operational model is key in allowing a BIA team to deliver high-quality solutions as efficiently as possible, to ensure that the team is always working on initiatives of the highest priority, and to strike the right balance between autonomy and control in the development of analytic applications.
- There is no universal 'right' operational model. An organization must choose a model, which may evolve over time, based on its BIA maturity and culture.

Operating Models include: Centralized, Consulting, Functional, Decentralized (Dispersed) and Centre of Excellence :

Operating Model	Description	Specifics	Pros	Cons
Centralized	Activity is delivered from a central group. Businesses do not have access to data, tools, or resources.	A group of analysts, acting as a core unit or business function, serve the entire company Usually implemented by large, single-business organizations with a high need for analytics applications that cross functional boundaries Businesses do not have access to data, tools or resources The analytics function can report to IT but increasingly is commanding its own C-suite leadership under a Chief Analytics Officer	Central management and strategic deployment of resources	Potential for gaps in understanding of the business
Consulting	Teams are centralized and are assigned to work on projects throughout the organization through a chargeback system.	A variation of the centralized model Analysts are centralized but assigned to projects throughout the organization and charge for their time	Clear tracking and measurement of outcomes	Results and focus is on who can pay vs. strategic needs
Functional	Teams are dispersed across the organization, and a small center of insights within an analytics dominant function supports teams.	Commonly adopted by organizations at an early stage of maturity with scarce BIA resources Analysts are placed in the business function(s) that dominate analytical activity of the company. These functions then act as pioneers to develop common standards and instill an enterprise-wide analytics culture.	Resources are concentrated where they are needed the most	Limited ability to expand and scale in other functions
Decentralized (Dispersed)	Activities conducted in business units with no centralized coordination.	BIA resources are spread throughout the organization No formal mechanism for coordination or collaboration Normally the result of business units attempting to create their own BIA function independently	Functions have complete control over needs	Limited enterprise-wide collaboration

Center of Excellence (Hybrid)	Teams exist primarily in business units, but their activities are coordinated by a centralized group	A hybrid of the centralized and decentralized models Analysts are based primarily in business functions and units, but activities are coordinated by a small centralized group Center of Excellence (CoE) is responsible for: training, adoption of analytical tools, innovation, facilitating communication among analysts, building enterprise competencies, providing overall program definition and operation Most appropriate for large, diverse businesses with a variety of analytical needs because it mirrors the structure of the company	Resources remain close to business and are coordinated by a center	CoE may not be able to control/oversee decentralized staff

Key Takeaways:

- Select a model that maximizes the enablement of business value
- It is important to be able to shift between models when it is realized business needs are not being met
- "A hybrid model is essential" – Gartner
- Totally centralized tends to be too slow and lacking in business domain expertise
- Totally decentralized struggles to deliver consistency and achieve cross business unit collaboration
- Co-locate teams with business and technical expertise where possible
- Senior management must be 100% committed for the transition to a data-driven organization to succeed
- Transition to a data-driven organization is not an overnight process

Example1: Business Intelligence and Analytics Operating Model

WHAT IS AN OPERATING MODEL?

- A operational model defines how a BI and Analytics team delivers services to the business, in collaboration with IT and business stakeholders, and how BIA efforts are to be governed.
- Deploying the right operational model is key in allowing a BIA team to deliver high-quality solutions as efficiently as possible, to ensure that the team is always working on initiatives of the highest priority, and to strike the right balance between autonomy and control in the development of analytic applications.
- There is no universal 'right' operational model. An organization must choose a model, which may evolve over time, based on its BIA maturity and culture.

WHAT IS AN OPERATING MODEL?

- A operational model defines how a BI and Analytics team delivers services to the business, in collaboration with IT and business stakeholders, and how BIA efforts are to be governed.
- Deploying the right operational model is key in allowing a BIA team to deliver high-quality solutions as efficiently as possible, to ensure that the team is always working on initiatives of the highest priority, and to strike the right balance between autonomy and control in the development of analytic applications.
- There is no universal 'right' operational model. An organization must choose a model, which may evolve over time, based on its BIA maturity and culture.

OTHER DEFINITIONS

- **Mission**: Describes why an organization (in this case the BIA team) exists – answers the question: what is our purpose?
- **Capability**: What an organization does to achieve its mission.
- **Process**: How a capability is delivered.
- **Service**: A capability delivered directly to an organization's customers .

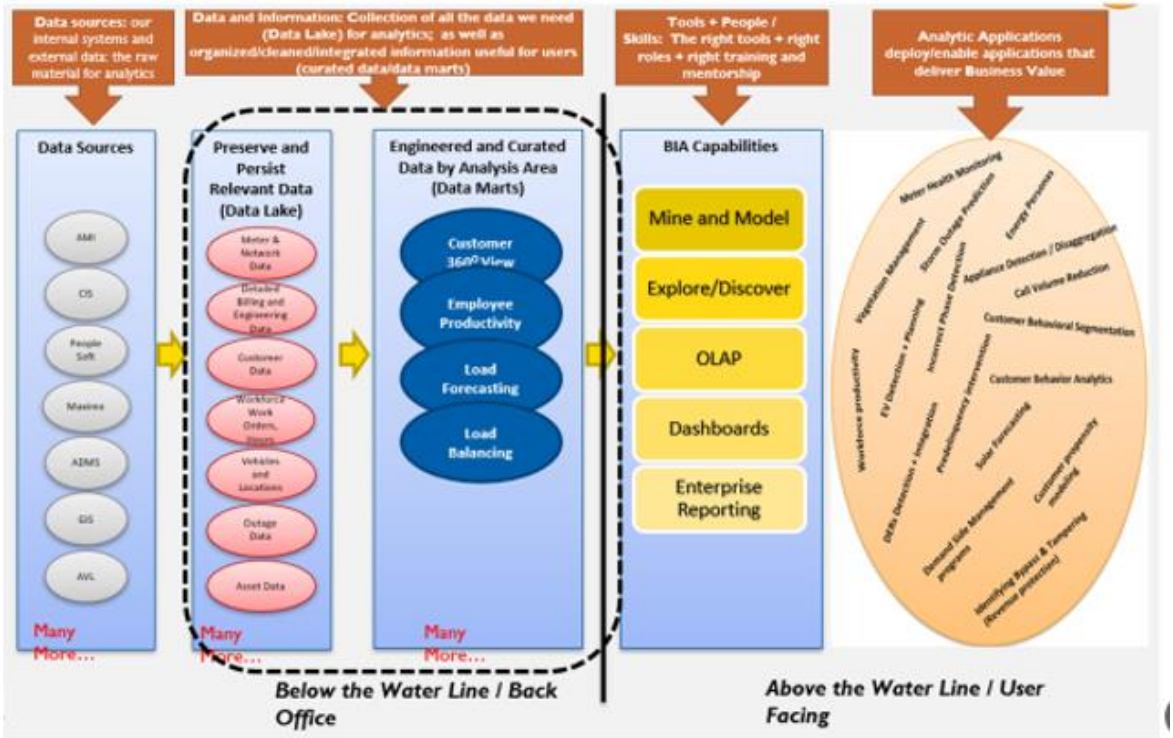
APPROACH FOR MODEL DEVELOPMENT

1. Define BIA mission
2. Evaluate and select an operational model
3. Create a model overview – 'business model canvas'
4. Define core services
5. Develop a scenario-based 'interaction model' for service delivery
6. Define capabilities
7. For each capability: define process/RACI, artifacts, standards and policies
8. Define organization structure, roles and skills

BIA MISSION

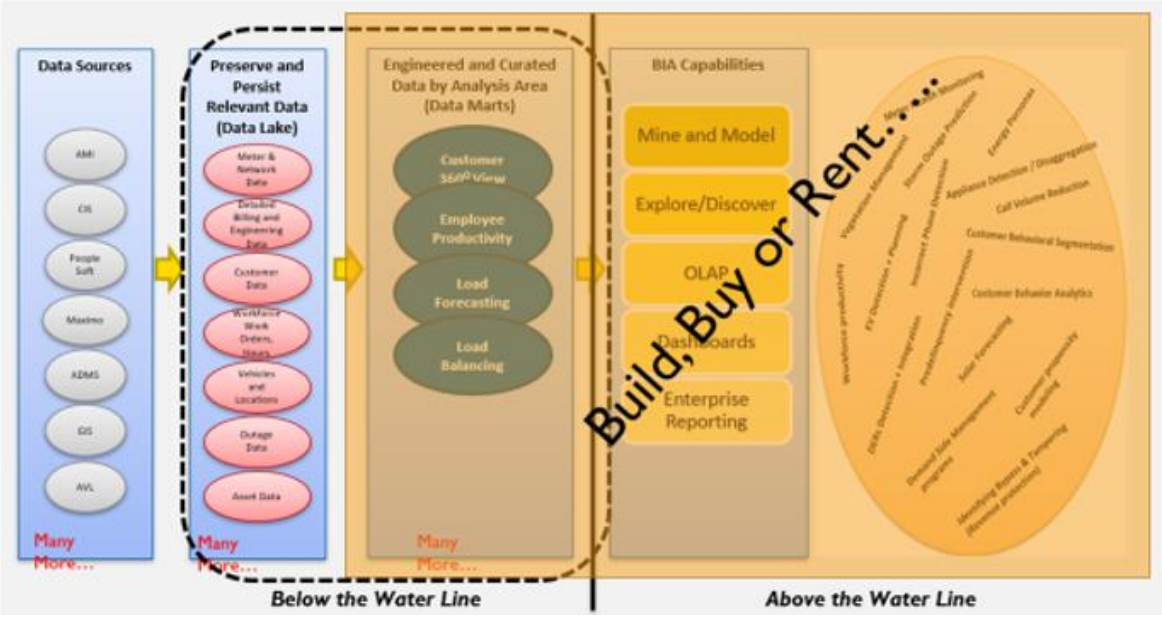
- Enable business-driven analytics via a self-serve model
- Provide world-class data management services
- Offer different tiers of support based on business unit analytics maturity/requirements
- Create an enterprise asset focused on delivery of business value

MISSION: DELIVER BIA ENTERPRISE ASSET



MISSION: DELIVER BIA ENTERPRISE ASSET

1

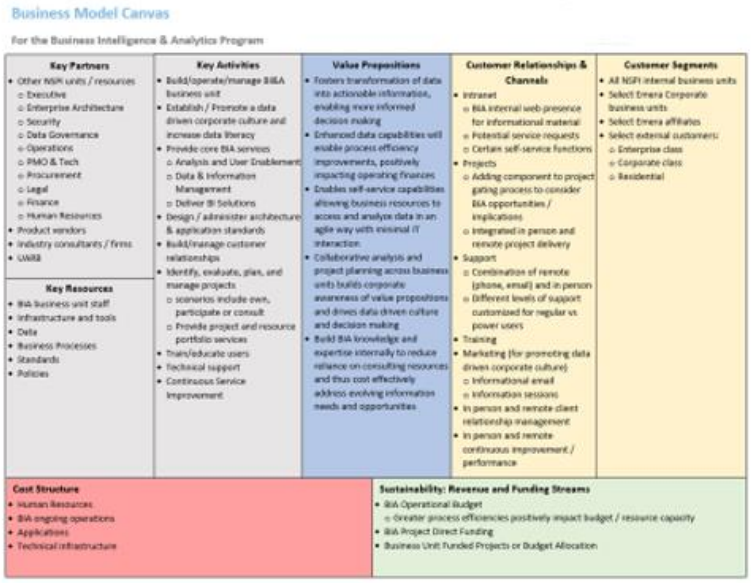


TARGET OPERATIONAL MODEL

2

- "Centre of Excellence" model (Gartner)
- Centralize data management functions
- Centralize governance
- 'Hybrid' application delivery:
  - Some centralized application development resources (analysts, report developers, data scientists)
  - Some BIA application development resources embedded in business units (report to BIA)
  - Some BIA application development resources will be 'power users' who do not report into the BIA team. We will offer explicit power user support (education, training, ongoing mentorship) as a service.

3 DEVELOP BUSINESS MODEL CANVAS



Business Model Canvas

For the Business Intelligence & Analytics Program

Key Partners	Key Activities	Value Propositions	Customer Relationships & Channels	Customer Segments
• Other NSPI units / resources ○ Executive ○ Enterprise Architecture ○ Security ○ Data Governance ○ Operations ○ PMO & Tech ○ Procurement ○ Legal ○ Finance ○ Human Resources • Product vendors • Industry consultants / firms • UWB	• Build/operate/manage BI&A business unit • Establish / Promote a data driven corporate culture and increase data literacy • Provide core BIA services ○ Analysis and User Enablement ○ Data & Information Management ○ Deliver BI Solutions • Design / administer architecture & application standards • Build/manage customer relationships • Identify, evaluate, plan, and manage projects ○ scenarios include own, participate or consult ○ Provide project and resource portfolio services • Train/educate users • Technical support • Continuous Service Improvement	• Fosters transformation of data into actionable information, enabling more informed decision making • Enhanced data capabilities will enable process efficiency improvements, positively impacting operating finances • Enables self-service capabilities allowing business resources to access and analyze data in an agile way with minimal IT interaction • Collaborative analysis and project planning across business units builds corporate awareness of value propositions and drives data driven culture and decision making • Build BIA knowledge and expertise internally to reduce reliance on consulting resources and thus cost effectively address evolving information needs and opportunities	• Intranet ○ BIA internal web presence for informational material ○ Potential service requests ○ Certain self-service functions • Projects ○ Adding component to project gating process to consider BIA opportunities / implications ○ Integrated in person and remote project delivery • Support ○ Combination of remote (phone, email) and in person ○ Different levels of support customized for regular vs. power users • Training • Marketing (for promoting data driven corporate culture) ○ Informational email ○ Information sessions • In person and remote client relationship management • In person and remote continuous improvement / performance	• All NSPI internal business units • Select Emera Corporate business units • Select Emera affiliates • Select external customers: ○ Enterprise class ○ Corporate class ○ Residential
Key Resources • BIA business unit staff • Infrastructure and tools • Data • Business Processes • Standards • Policies				
Cost Structure • Human Resources • BIA ongoing operations • Applications • Technical infrastructure			**Sustainability: Revenue and Funding Streams** • BIA Operational Budget ○ Greater process efficiencies positively impact budget / resource capacity • BIA Project Direct Funding • Business Unit Funded Projects or Budget Allocation	

The *business model canvas* explains BIA mission and operational model at "elevator pitch" level. Use to socialize and gain consensus on direction.

11

4 DEFINE CORE SERVICES TO BE OFFERED TO THE BUSINESS

Analysis and User Enablement		
Service	Definition	Accountable
Identify and understand business information needs	Work with business users to elicit, understand, and document information requirements.	BIA Application Manager
Provide self-service	Develop ability for users to leverage BIA tools to self-serve information needs.	BIA Application Manager
Support users	Perform intake, triage, and resolution / fulfillment activities for ad hoc requests. Provide user self-service support. Interpret information provided via analytic products. Explain/provide knowledge transfer regarding the meaning and quality of NSPI data.	BIA Application Manager
Train users	Educate users on how to effectively and efficiently use developed / provided BIA solutions.	BIA Application Manager
Educate users on data driven culture	Facilitate a deeper corporate understanding of the value of greater BIA adoption in corporate decision making.	BIA Director & BIA Application Manager
Advise on opportunities / initiatives	Provide BI & data management related advice on all appropriate opportunities and projects.	BIA Director

12

5 INTERACTION MODEL: HOW DO WE COLLABORATE TO DELIVER BIA CORE SERVICES?

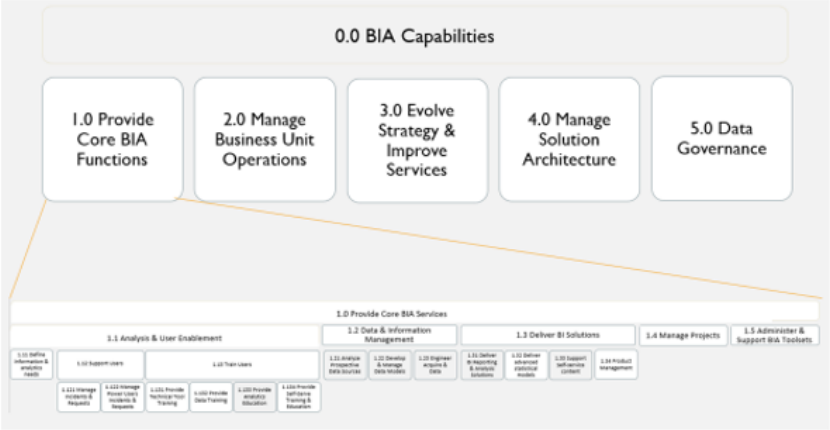
Interaction model answers the question: what capabilities does BIA need to 'consume' from partners for different service delivery scenarios?

Scenario-based:
- Opportunity assessment
- "Full stack" development
- User Support

This is critical for change management in introducing a new service.

13

6 DEFINE CAPABILITIES REQUIRED TO BOTH ENABLE AND DELIVER SERVICES

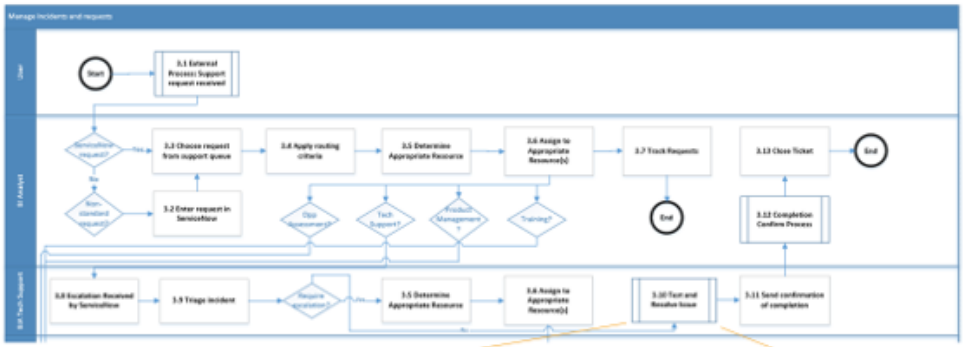


Need to decompose capabil ties to 'elementary' level to facilitate definition of processes.

14

7

DEFINE PROCESSES



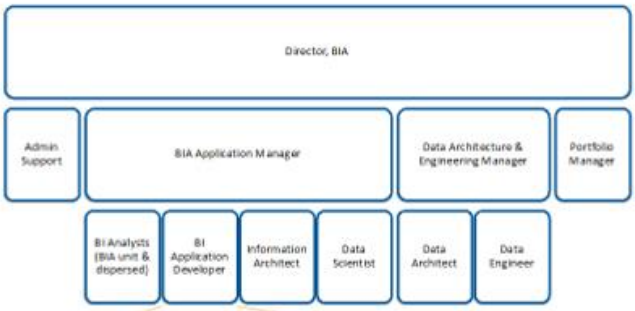
Description	Artifact Support	Applicable Policies / Standards	Dependency	Output(s)	Responsible	Accountible	Consult	Inform
Appropriate resource remedies issues and confirms fi nctionality	None				Resource assigned to resolution	BIA Manager		

15

8

DEFINE ORGANIZATIONAL STRUCTURE, POSITIONS/ROLES AND SKILLS

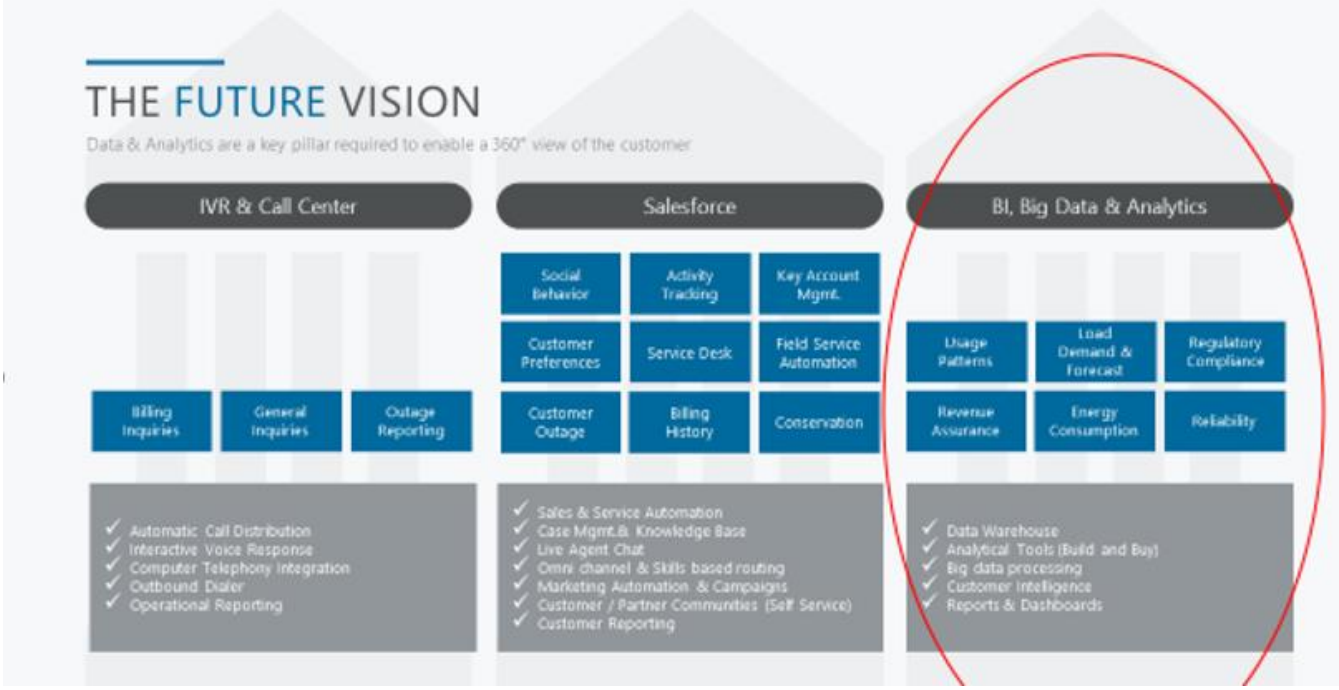
Note: this does not represent roles sourced from other teams (e.g., DBA, PM, Tech Admin, etc.)



Position	Overview	Responsibilities	Skills and Knowledge
BIA Application Developer	Responsible for developing and testing reports and analytic solutions	• Responsible for analytic application development (reports, dashboards, cockpits, etc.) • Responsible for custom development (SDK) for integration of reports and analytics into operational environments • Responsible for unit testing • Participates in requirements analysis Participates in report design • Participates in production support (second level)	• Organization specific business knowledge • BIA-specific Requirements gathering methods • Deep BIA tool knowledge – reporting and analysis, tuning • Report testing knowledge • Report design techniques • Solid knowledge of SQL

16

Example2: Business Intelligence and Analytics Operating Model



ANALYTICS OPERATING MODEL

[redacted] utilizes a "functional" operating model in support of its data and analytics program.

Analytics resources are dispersed across the organization, with a small analytics & data mgmt. team centralized in IT.

The program aims to kick-start an organizational shift towards a data-driven culture led by a program lead residing in the Information Technology team.

The program is just beginning and in the forming stages. It will be a marathon – not a sprint.

OPERATING MODEL INFLUENCERS

Key influencers driving model adoption

ANALYTICS VISION

[redacted] endeavor to support a self service BI model to gain agility and reduce IT demand

REVENUE & HEADCOUNT

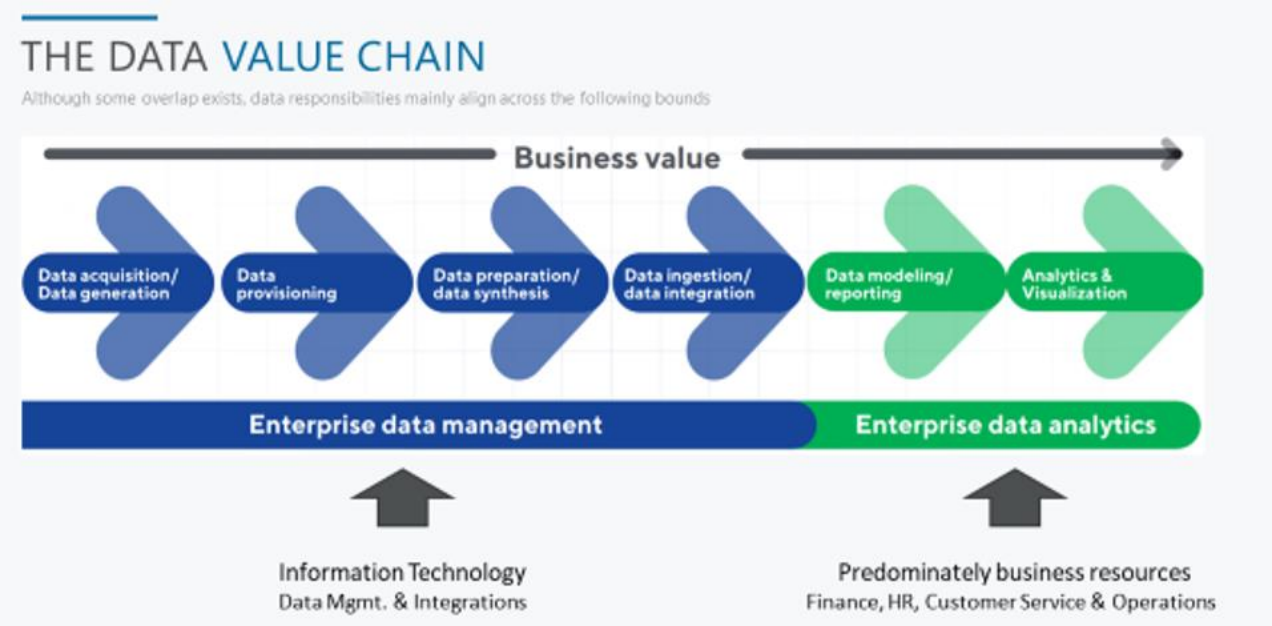
Utilities are under increasing pressure to minimize costs [redacted]

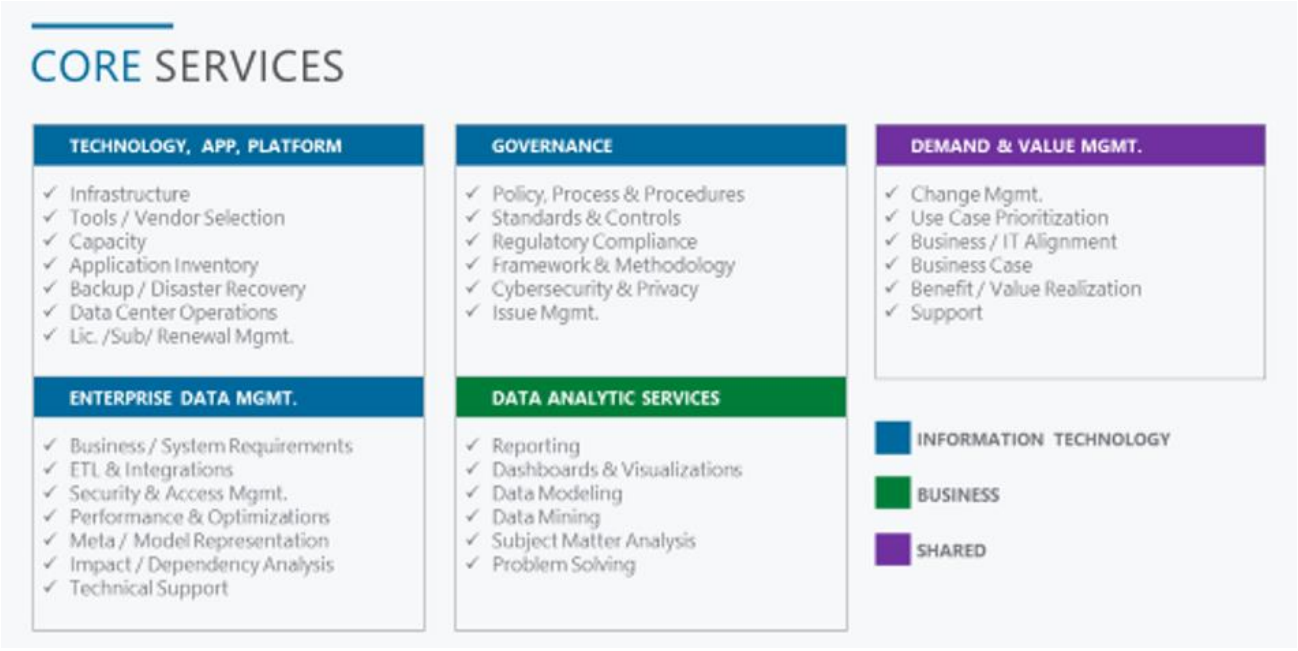
ORGANIZATION

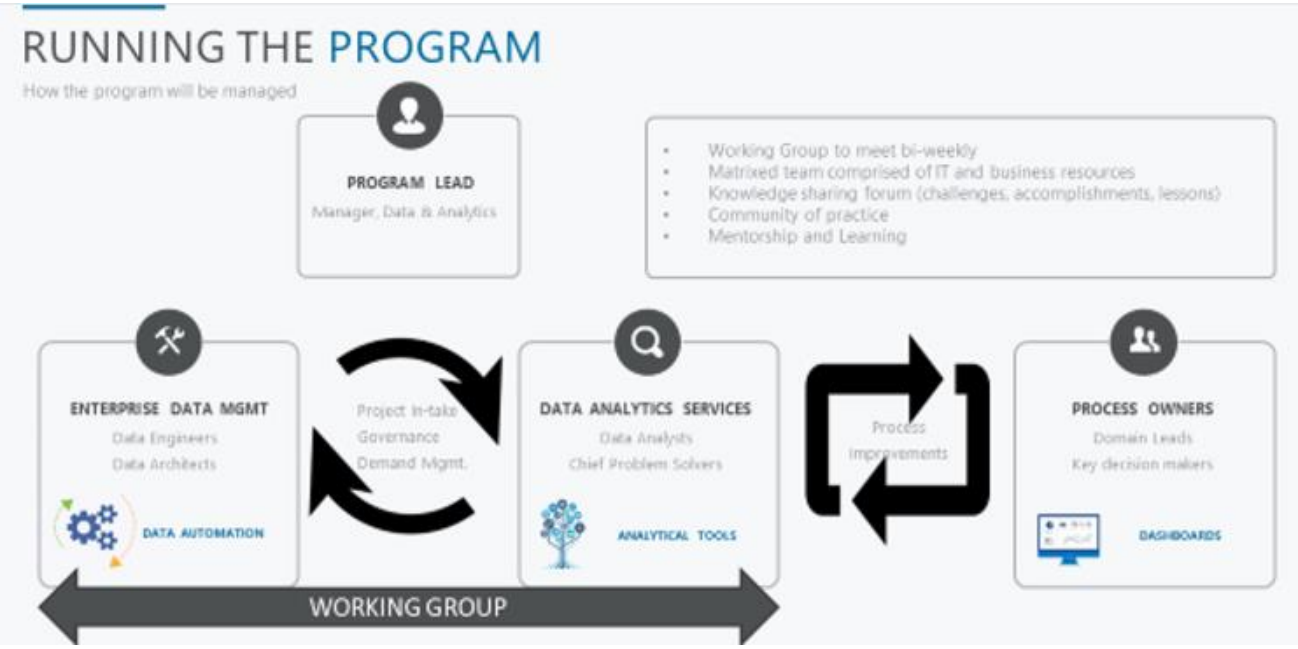
The organization is functionally driven today. Subject matter expertise rests within the business

DATA MATURITY

Our data maturity is low and requires a crawl, walk, run approach to improve


THE DATA VALUE CHAIN
Although some overlap exists, data responsibilities mainly align across the following bounds
Business value
Data acquisition/ Data generation
Data provisioning
Data preparation/ data synthesis
Data ingestion/ data integration
Data modeling/ reporting
Analytics & Visualization
Enterprise data management
Enterprise data analytics
Information Technology
Data Mgmt. & Integrations
Predominately business resources
Finance, HR, Customer Service & Operations

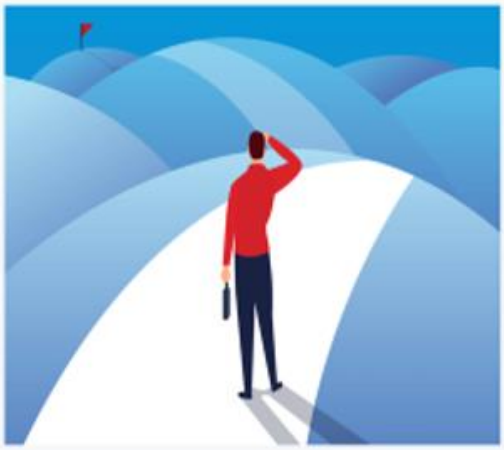

CORE SERVICES
TECHNOLOGY, APP, PLATFORM
✓ Infrastructure
✓ Tools / Vendor Selection
✓ Capacity
✓ Application Inventory
✓ Backup / Disaster Recovery
✓ Data Center Operations
✓ Lic. /Sub/ Renewal Mgmt.
GOVERNANCE
✓ Policy, Process & Procedures
✓ Standards & Controls
✓ Regulatory Compliance
✓ Framework & Methodology
✓ Cybersecurity & Privacy
✓ Issue Mgmt.
DEMAND & VALUE MGMT.
✓ Change Mgmt.
✓ Use Case Prioritization
✓ Business / IT Alignment
✓ Business Case
✓ Benefit / Value Realization
✓ Support
ENTERPRISE DATA MGMT.
✓ Business / System Requirements
✓ ETL & Integrations
✓ Security & Access Mgmt.
✓ Performance & Optimizations
✓ Meta / Model Representation
✓ Impact / Dependency Analysis
✓ Technical Support
DATA ANALYTIC SERVICES
✓ Reporting
✓ Dashboards & Visualizations
✓ Data Modeling
✓ Data Mining
✓ Subject Matter Analysis
✓ Problem Solving
INFORMATION TECHNOLOGY
BUSINESS
SHARED


RUNNING THE PROGRAM
How the program will be managed
PROGRAM LEAD
Manager, Data & Analytics
Working Group to meet bi-weekly
Matrixed team comprised of IT and business resources
Knowledge sharing forum (challenges, accomplishments, lessons)
Community of practice
Mentorship and Learning
ENTERPRISE DATA MGMT
Data Engineers
Data Architects
DATA AUTOMATION
Project In-take
Governance
Demand Mgmt.
DATA ANALYTICS SERVICES
Data Analysts
Chief Problem Solvers
ANALYTICAL TOOLS
Process Improvements
PROCESS OWNERS
Domain Leads
Key decision makers
DASHBOARDS
WORKING GROUP


IMPROVING DATA LITERACY
A shared responsibility and a key focus of the working group
Connect
We are connecting IT and business, data engineers and analysts, the data curious and the data ready across the organization
Learn
The program will focus on skills building via education & training, hands on experience along with continuous engagement and collaboration
Together we will create a base vocabulary, common standards, and interesting stories to tell
Create
We will assess our strengths, weaknesses, successes and failures that inform our future decisions
Measure & Impact

OUR PROGRAM GOALS

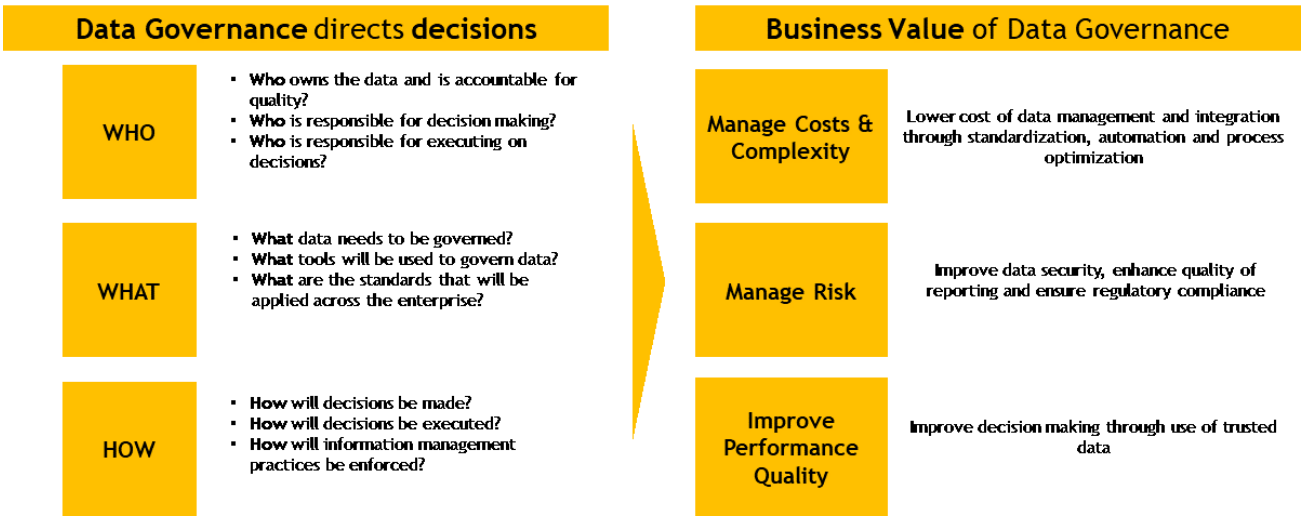
- ❑ Build a data rich culture
- ❑ Develop analytical skillsets across the organization from within
- ❑ Simplify and demystify our data, more analysis vs curation
- ❑ Foster collaboration, creativity and innovation
- ❑ Apply data and innovation into strategy based on business objectives
- ❑ Improve our customer experience, better manage our business
- ❑ To protect our data with appropriate safeguards
- ❑ Improve evidence-based decision making



III. Data Governance

"Data and analytics leaders can exploit business opportunities and challenges only if they have established the right foundations for data and analytics governance" – Gartner 2019

- At the highest level, a Data Governance program enables our businesses to break deficiencies associated with organizational silos and to create additional enterprise-wide value from the organizational data assets, while minimizing risks and transforming the workplace culture.
- Below are the WHO / WHAT / HOW questions that are being answered by Data Governance. By having answers to these questions, organizations drive significant Business Value from their Data Governance activities.
- As each organization embarks on this journey, we need to take some time to reflect on the questions below. If there are not clear answers to these questions, chances are that are we might have some foundational gaps that will require a systematic, structured program to address how data is governed in your organizations.

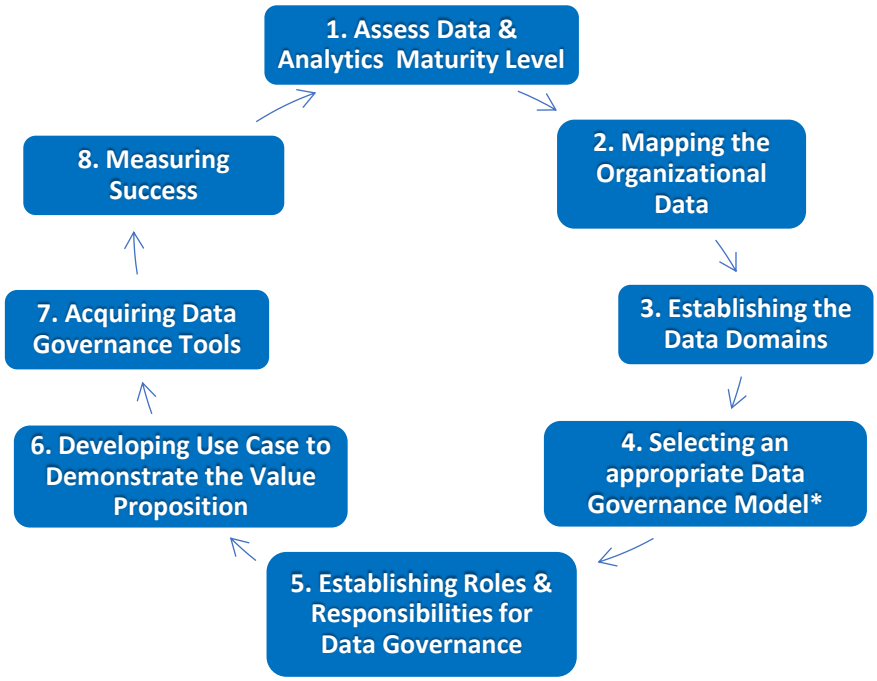


When building a business case for instituting a Data Governance framework, the language that seems to resonate the most with Executives centers around **Risk Management** and **Value Creation**. Whether it is about:

- ***compliance with privacy laws*** - which requires a comprehensive enterprise-wide data catalogue and rigorous data classifications
- ***safeguarding reputational value*** - ensuring that no questionable data or reports are being posted to various users, or
- ***effective and efficient business decision making*** - ensuring that business units across the organization have quick access to the best data for the best decisions

Building a Data Governance Organization

The following chapters describe the key steps that typical organizations take in their journeys towards building a data governance organization. The narrative is based on established theory from various expert resources, coupled with the real-life experiences of several organizations within CEA's DAWG, that are already on their way to establish and maintain good data governance programs.

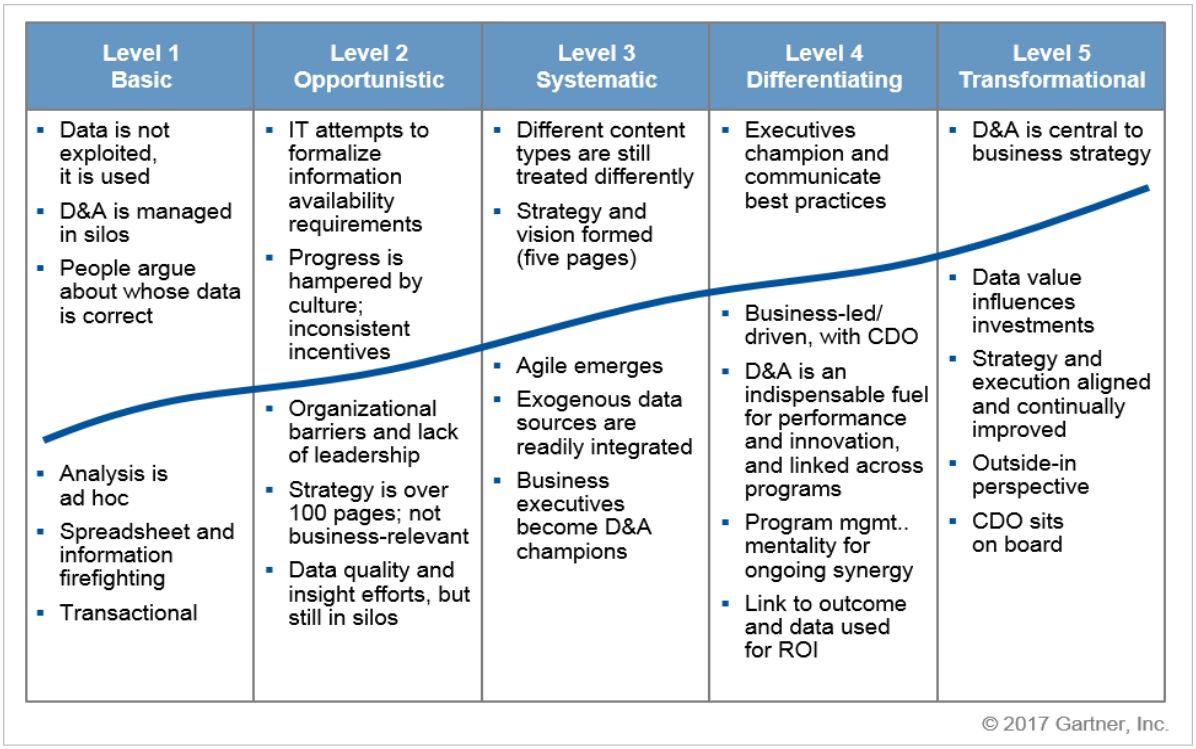


*please also refer to Section II. of this Handbook, under Organization.

STEP 1: Assessing the Maturity Level of the Data & Analytics Functions

- Typically, the first step that an organization takes in building their Data Strategy is an assessment of the current organizational status as far as maturity and capabilities for data & analytics.

- There are a number of models currently available that can support an organization in mapping their current status, such as Gartner's Data & Analytics Maturity model, or the Utility Analytics Institute model. This Gartner instrument can also be used as an initial engagement mechanism as it offers a cross-functional collaboration tool for a number of business units across the organization (IT & BU's) to jointly contribute to the development of an enterprise Data maturity model.



Data & Analytics Maturity Model, Gartner 2017

- The end product, if one chooses to go through the detailed analysis, is a comprehensive report that outlines overall and distinct maturity scores under a broad number of dimensions (vision, strategy, people, processes, tools, etc) with recommendations for how to address various areas of strength and weaknesses.

- Most of the DAWG members have completed the Gartner Data & Analytics maturity model assessments, to certain degrees of details, with results presented and discussed during the in-person meeting in Ottawa (June 2019). **The conclusion was that most organizations had at the time overall scores between 1's and 2's, with some exceptions scores in the low 3's.**

- Having said that, most organizations also recognize that despite these relatively low overall maturity scores, they also all have "pockets of excellence" in various business areas, where abilities and practices for data governance are much more evolved than the rest of the organization.

STEP 2: Mapping the Organizational Data

- This is a critical step in the development of a functional data governance framework as an organization needs to know in detail where its data is, the status of the data (eg. quality), who owns what, data flows and business transformations, interactions between various systems, use of data by business units, nomenclature and terminology, processes & standards, key pain points, etc.

- The outcome of this step will be a comprehensive collection of artifacts that will allow an organization to articulate its key data issues and opportunities and create long term roadmaps for their data strategies.

- Use of consultants and internal resources are of help as this is a laborious, time consuming exercise that requires structure and an iterative approach to the information collection.

- A gradual approach in mapping the data is recommended, such as starting with smaller but critical data sets or where the most visible issues are.

Example: In this business case, the initial focus of the data mapping was centered around core operational grid systems and external reporting:

a) 22 Core Operational Systems that are critical to grid operations

- **Dataset Blueprint:** Outlines the various datasets within categorized data domains.
- **Dataset Dictionary:** Describes the attributes, data owners, and system of record of each dataset.
- **CRUD Matrix:** Displays where the datasets are being created, updated and consumed.
- **Data Flows:** Illustrates the flow and interactions of datasets between applications for core processes.

b) Analysis of external reporting

1. **Cost and Value Drivers:** Without a well-defined understanding of the usage and value of reports, it is difficult to make informed decisions on reporting initiatives and rationalization priorities.
2. **Operating Model:** Establish business unit accountability and support collaborative working processes for data and reporting activities.
3. **Information Management:** Define and enforce information management standards to ensure knowledge transfer, reduce miscommunication, and provide a trusted source of truth.
4. **User Experience:** Data provisioning activities are continuously evolving in response to consumer demands, but there are opportunities for improvement such as providing more customizability.
5. **Privacy and Data Protection:** Lead in cybersecurity, there are opportunities to revisit internal policies to ensure maximum data privacy and confidentiality.

Note: To really understand data, we must also understand and document **metadata.** Metadata management is at the foundation of a data governance framework, as it captures both **Business Metadata** (governed by business people, such as name, description, format, size, subject to PI, NERC – CIP, etc) as well as **IT Metadata** (governed by IT people, such as application table, column, etc).

STEP 3 – Establishing the Data Domains

- Armed with a good understanding of the overall data landscape and key opportunities, organizations can approach the next step of establishing their Data Domains - this is **a prerequisite** to defining the data governance processes and to share accountability of data across the enterprise. A few models are available for this exercise, as shown below, each with its pro's and con's:

	Subject Area	Technology	Organization chart
Description	Define data domains based on subject areas	Define data domains based on the platforms or systems hosting data	Define data domains aligned to different enterprise functions
Pros	• Creates common business language to collaborate around data. • Domains cross enterprise functions which allows visibility on data	• Easier to delimit domains. • Simpler governance tooling that do not go beyond selected platform for the domain.	• Easier to assign owners to the domain. • Have a higher understanding of the business needs and requirements.
Cons	• Requires heavily on leadership buy-in to assign owners.	• Similar data can often be governed by different owners leading to inconsistencies and creates complexity in the governance program. • Complex governance processes	• Deepening data silos across the organization limiting collaboration. • Similar data might exist across multiple lines of business and managed differently.

Example: The approach in establishing the data domains, by Subject Area (HR, Finance, Contracts, Market Data, etc), is a model that fits the philosophy of creating a shared approach for data governance, with partnership between Business and IT. This model is also agnostic to organizational / business processes changes, stays true to the core organizational mandate and can be scaled up and down as required.

Example: Another company had a similar Subject Area approach in establishing their data domains which have evolved over time to a more comprehensive enterprise-wide view.

STEP 4: Selecting an appropriate Data Governance Model

- With a good understanding of where data is, how it is currently governed, and a general awareness across the organization of the importance of data governance, the next step is in choosing a more structured, well defined model for building a Data Governance framework.

- As previously presented in the **II. Organization** chapter of this Handbook, data governance models range from Decentralized, to Hybrid, to Centralized solutions all with their advantages and challenges - which makes the solution of choosing the initial model specific to each organization based on the industry, organization size and tenure, business objectives, organizational structure and culture.

- Given the relatively low overall data & analytics maturity, many organizations these days are opting for a **Hybrid Model** so they can establish Data Ownership throughout the entire organization with consistent definitions and standards, while allowing each business unit to be in charge of the use of their data and thus promoting the best data uses and practices across the organization.

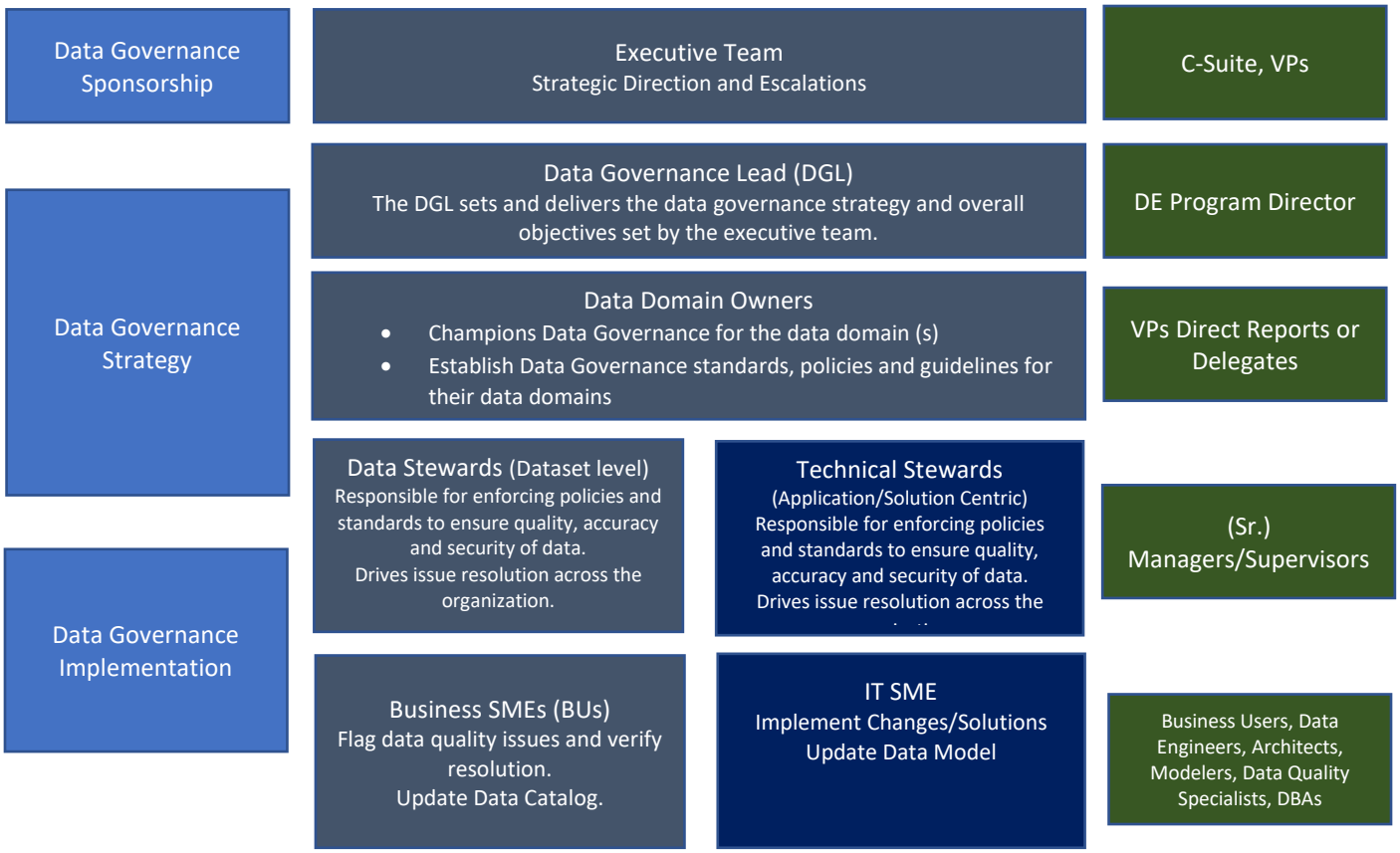
    - **Centralized Models** are best suitable for younger and smaller organizations, where the issue of control does not impede with the business speed as much as it would in an already more established organization where siloes and different practices have already formed and solidified. This isn't likely the case for the organizations that are members of this CEA working group.

    - As organizations matures in their data governance practices, the model can also shift more towards a **Decentralized Model**, whereby the business units have acquired enough knowledge and practices that they can manifest an enterprise-wide vision for Data Governance - in such organizations though the role of a CDO becomes even more critical so that practices can evolve and can be enforced in a consistent fashion across the entire organization, while forward-looking decisions can be made with a birds-eye view on the strategic business priorities.

STEP 5 – Establishing Roles and Responsibilities for Data Governance

The chosen data governance framework will dictate the target operating model that each organization must choose for implementation and maintenance.

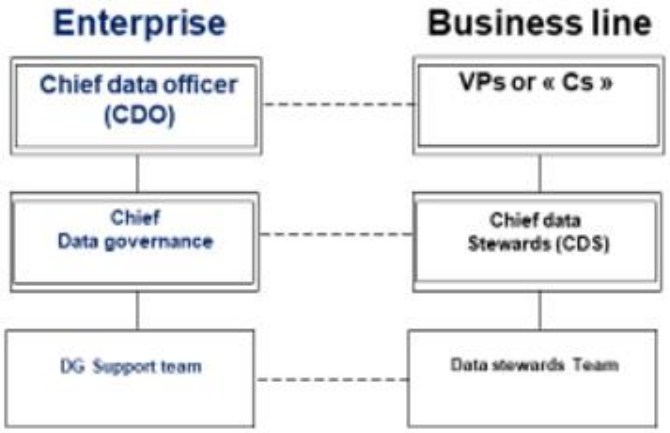
Example1: One way to unfold the Hybrid Model and how this model translates into roles and responsibilities for ownership and stewardship between the business units and IT - please note different organizations may choose different nuances of establishing these roles depending on their specific organizational structure.



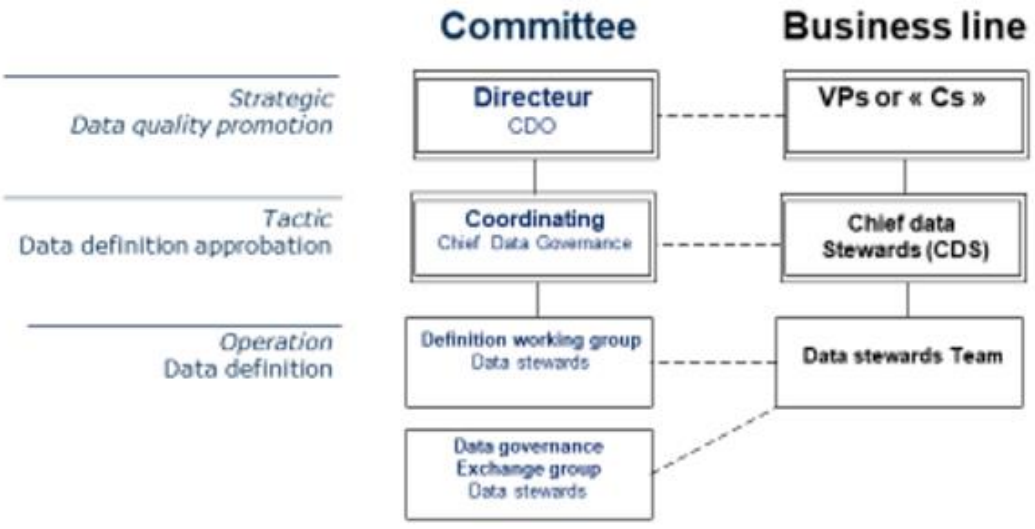
Example2: Using a similar model as in the example above, another organization has had a few more years of efforts in establishing their data governance framework and functions now at a higher-level maturity, with key structures, policies, metrics and KPI's already implemented in many business units

- Interesting to note is that this organization has evolved to a stage where they are no longer using the term "ownership" when it comes to data. The organization sees a company as the only "data owner" with business units having ***fiduciary*** responsibilities for the data (for example, HR has a fiduciary role for the employees' information)
- As data governance is clearly defined as a business activity, the leadership and stewardship of the data are now already distributed within many business units across the organization. There is also a centralized team to implement and support the data governance, under an established Chief Data Office (CDO) role and which leads the various committees.
- The CDO role is key in maintaining the data management strategy (along with its' Data Governance body) with all business units, who each has a Chief Data Steward (CDS), along with several business stewards. The relationship between the CDO and the CDS's ensures that the data is properly governed, valued and leveraged within a harmonious matrix structure.

Federated data governance

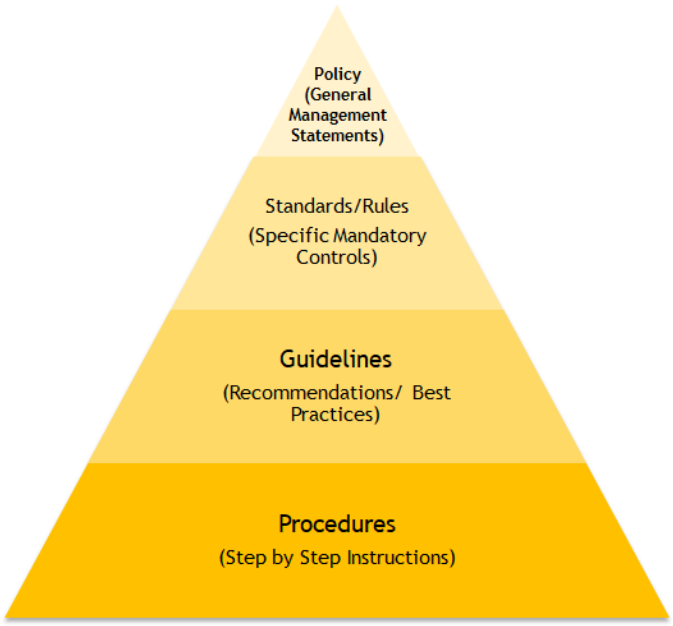


Data governance Committees



Data Governance Controls

The implementation of a successful data governance framework depends on the structure and rigor put around managing the data that drives business capability.



- Policies outline roles and responsibilities, defines the scope of data to be protected, and provides a high level description of the controls that must be in place to protect information. In addition, it should make references to the standards and guidelines that support it.
- Standards/Rules consist of specific low level mandatory controls that help enforce and support the data policy. Standards help to ensure consistency across the business and usually contain controls relating to the implementation of processes or specific technology.
- Guidelines consist of recommended, non-mandatory controls that help support standards or serve as a reference when no applicable standard is in place. Guidelines should be viewed as best practices that are not usually requirements, but are strongly recommended.
- Procedures and Workflows consist of step by step instructions to assist workers in implementing the various policies, standards and guidelines.

In summary, there are some critical factors to a successful Data Governance implementation:

Senior Level Sponsorship	A top down model will ensure that data governance is an enterprise wide objective and warrant the necessary behaviors and responsibilities for each individual.
Updating Job Descriptions	All roles with governance duties must contain specific governance tasks as a part of the job description to ensure employees will carry out desired activities. Specific KPIs need to be set to measure performance and individuals must be rewarded according to their performance.
Use Case Alignment	Data governance must be planned and implemented relative to specific use cases. All use cases must be prioritized and evaluated to determine when and how to apply a well suited governance model.
Value Add Approach	Governance must add value to business activities rather than being treated as an auditing exercise. Individuals must see this as an improvement and benefit to their current role within the organization.

STEP 6: Developing Use-Cases to Demonstrate the Value Proposition

- The implementation of a Data Governance framework cannot succeed without the demonstration of its value proposition through specific use cases that will put the theory at work and will demonstrate to the business units the rewards associated with having a sound data governance practice.

- The first chapter of this Handbook, I. **Business Value,** has been dedicated to use-cases development, in recognition of their paramount to demonstrating the value proposition of a data strategy program.

STEP 7: Acquire Essential Data Governance Tools

- Chances are most of our organizations have been historically using rather basic tools to keep track of our data assets, which is not a sustainable practice in the long run to develop and implement a lasting data governance organization. As such, technology components are essential these day to support the vast amount of data that need to be governed in our organizations.

Here are some examples of data governance capabilities and some options for some of the more common technologies:

DATA GOVERNANCE CAPABILITY	CAPABILITY DESCRIPTION	OPTIONS (NON-EXHAUSTIVE)
Automated Data Governance	Data Governance requires a cross-organizational solution designed to provide robust tools for the management of governance and stewardship.	Collibra, IBM Information Server, Talend, Informatica
Master Data Management	Master Data Management (MDM) is the collection of people, process and technology components working together to ensure Master Data is coordinated across the enterprise.	Informatica MDM, IBM InfoSphere MDM, Talend MDM, SAP Master Data Governance, VisionWare MDM, Oracle Product Hub, Orchestra Networks EBX, SAS MDM, Information Builders Omni-Gen, Informatica Integration Hub
Metadata Management	Metadata Management is the automation of information regarding the characteristics of any artifact, such as its name, location perceived importance, quality or value to the enterprise, and its relationships to other artifacts that an enterprise has deemed worth managing.	Collibra, Enterprise Metadata Management (Oracle), InfoSphere, ASG Data Intelligence
Data Quality Management	Data Quality is often measured by the following dimensions and components. The Data Quality Dimensions are based on summary definitions of data quality and the components provide additional detail about the dimensions, through automation. • Accuracy: Degree to which the information contains errors and meets business rules for the data • Completeness: Degree to which the complete data set is available or the fields are populated • Consistency: Degree to which data is the same in its definition, business rules, format and value across systems	Information Steward (SAP), IBM InfoSphere QualityStage, Informatica IDG. Talend Data Quality, SAS Data Quality,
Data Movement Management	Data Movement technology allows the extraction, transformation, loading and manipulation of data from one or more sources into a destination system and the capture of associated metadata.	IBM InfoSphere DataStage, Informatica PowerCenter, Talend Open Studio

- The market for data governance solutions is mature with multiple vendors offering solutions with overlapping capabilities, which fall on a spectrum, based on degree of specialization such as: Data Catalogue (to primarily store and manage technical metadata), Business Glossary (provides business terms and definitions of what the data means, how it is calculated and used), Workflows (the orchestration of business processes between business participants), Data Quality (to assess the reliability of the data) and Security Information Management Policy (access rights and management by roles and groups).

- At minimum, organizations embarking on a data journey should at least look at procuring a data catalogue to support storing and managing technical metadata; however other solutions that include business glossaries, workflows, data quality and security functions should be incorporated as well.

More details on various data technologies are provided in a dedicated chapter in this Handbook, under Technology.

STEP 8: Measuring Success

- Measuring success is critical to maintaining the attention of the organization on the Data Governance program and ensuring that the controls are working and producing the desired value.

Here are a couple of examples from the member organizations:

Success Factors: The following success factors were recommended in order to drive the implementation of a data governance program:

a) **Governance Organization:** Establish an operating model to build and evolve data standards across the organization.
b) **Process Standardization:** Reduce complexity, redundancy and inefficiencies by introducing standardized processes data structures, and technologies.
c) **Data Quality:** Increase inefficiencies and data confidence by introducing common data quality process, metrics and technologies.
d) **Data as an Asset:** Properly communicate that data is a business asset for the organization and must be managed accordingly.
e) **Data Source Rationalization:** Rationalize its data sources and applications.
f) **Single Source of Truth:** Establish a single source of truth for all master data, managed at a corporate level.

- Each and every one of the above factors can be distilled into KPI's that would be reportable at the highest accountability level (executive teams, boards) as the measurements of the program success.
- These KPI's can be developed based on the specific insights gathered in the initial current state assessment process which can act as a baseline.
- Similar mechanisms for assessment can be used in later stages of program development to maintain consistency of the measurement instrument and showcase progress against the baseline.

Sample Company: The company has a number of measures (KPI's) to define the success of their data governance program, such as:

- Number of deployed business units – with business stewards assigned and with prioritized initiatives that apply the DG Framework.
- Participation of the business unit in the DG committees
- Number of defined business terms
- Silo elimination

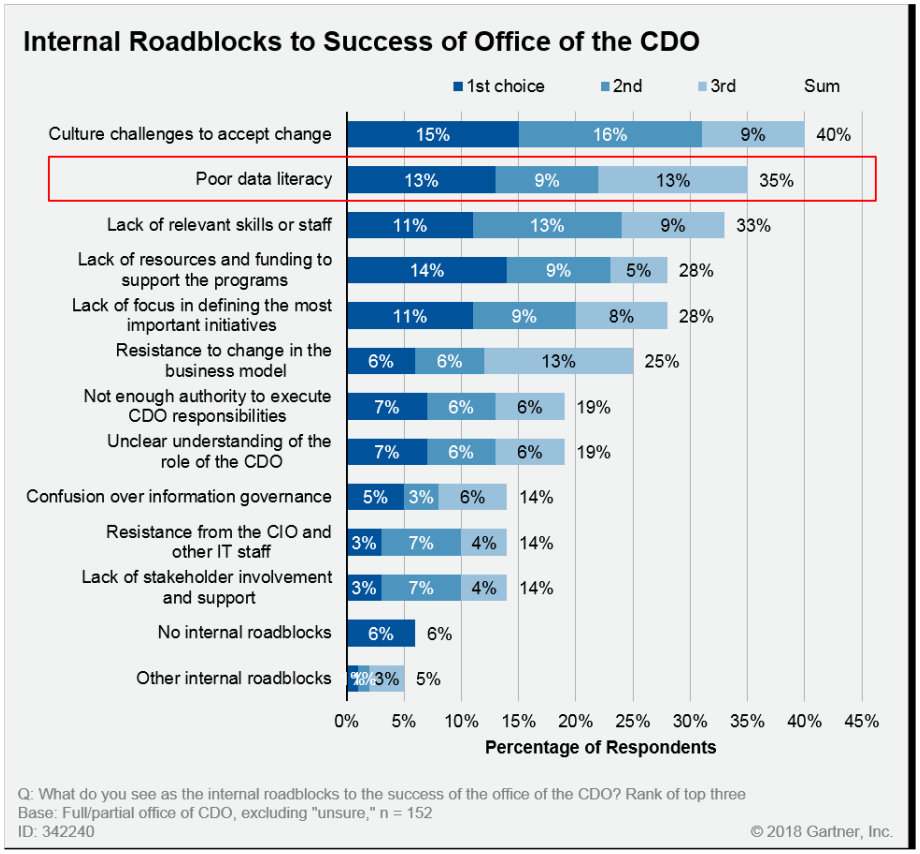
Other Strategic Considerations

I. The Sustainability of Efforts

- At some point in the implementation of a data governance framework, certain organizations reach a ceiling that it is become harder and harder to break.
- It is not unusual, especially after going through the hard work of putting such an ambitious program in place, to lose some of the initial focus and motivation to sustain the effort.
- Staying **data-fit** requires ongoing commitment and effort on a daily basis. It is up to the data champions in our organizations to rekindle the spirit and the interest of our organizations through compelling use case and creative means of using the data and building employee and executive satisfaction.

II. How does Data Literacy fit into Data Governance

- Data literacy, as defined by Gartner is the ***"ability to write, read or interpret data in context".***
- A major problem in our organizations are "the silos" who sustain separate systems, cultures and practices of data. A good data governance model helps break the silos and emphasizes the collaboration between Business Units towards shared business goals. Which in turn requires a certain level of shared data literacy across the organization, anchored in common business term definitions and with alignment from all business units.
- With digital transformation, data literacy becomes very important especially as most enterprises don't currently ***master*** their data. There are a number of challenges to the success of a data driven organization, and poor data literacy ranks second only after a non-open mindset to change.



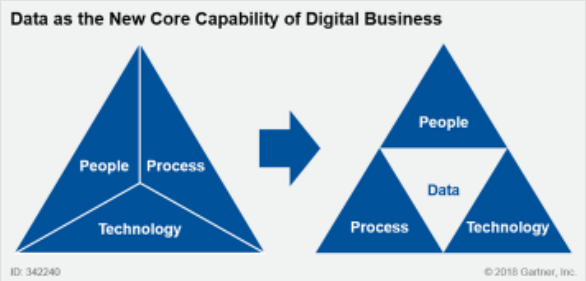
"If we don't give everyone the ability to simply read and write, we aren't giving everyone a chance to succeed." — Barbara Bush (1925-2018), former First Lady of the United States and founder of the Family Literacy Foundation.

THE DATA LITERACY IMPERATIVE

- Gartner defines Data Literacy as **"the ability to read, write and communicate data in context,** including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case, the application and resulting value."
- With the emergence of data, analytics, artificial intelligence (AI) and machine learning as the new core elements of digital business and society, the ability for creators and consumers of solutions built on these elements to "speak data" in a common way has never been greater. Data and analytics leaders must treat information as a **second language** and data literacy as a core element of digital transformation.
- The changes to business will be profound. Gartner expects that, by 2020, 80% of organizations will initiate deliberate competency development in the field of data literacy, acknowledging their extreme deficiency. In addition, by 2020, 50% of organizations will lack sufficient AI and data literacy skills to achieve business value. **Let's all be part of the other 50%!**

2

THE DATA LITERACY IMPERATIVE

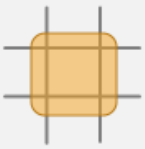


- People, process and technology. These are the three elements common to all business change. But now, any organization undergoing a digital transformation must factor in a fourth key element — data!

Not only must organizations take steps to educate professionals who are involved in crafting data-driven solutions, products and services, they must also ensure those steps achieve the goal of teaching all relevant employees to speak data as their new second language, as well as developing and nurturing communities in which the language will flourish

3

BUILDING THE CASE FOR DATA LITERACY

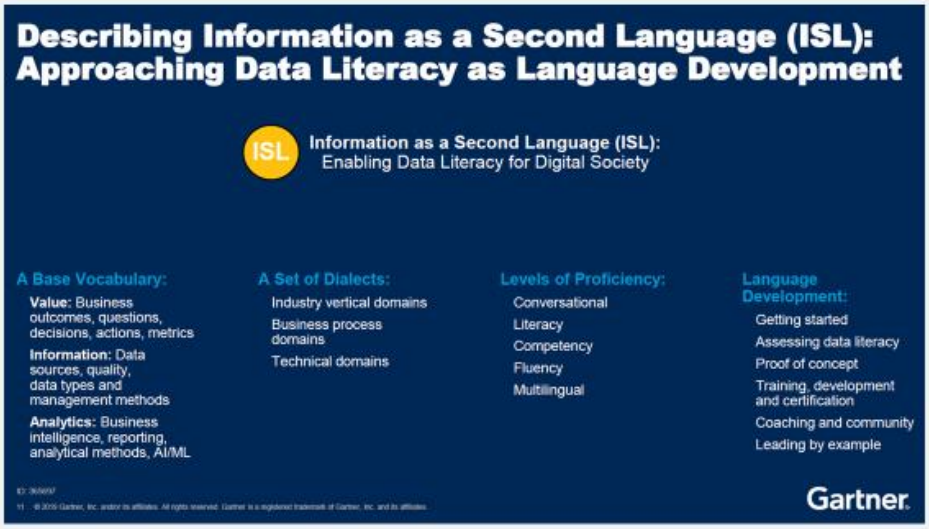


- Gartner has a very useful, practical tool to assess an initial organizational assessment for measurement of data literacy to target areas for development and baseline starting levels.
- In Gartner, download the Toolkit PowerPoint file for an initial gauge of overall data literacy with 25 questions. Organized across five themes to highlight areas of relative strength/weakness (general, business/value, data, analytics, culture).
- Use initially to establish a baseline and then measure periodically to assess improvement.
- Can be applied with teams of:
  - Data creators, or those who create/produce data and analytical solutions
    - Examples: CDO, data engineer, data scientist, program manager
  - Data consumers, or those who are users of data and analytical solutions
    - Examples: an executive, business analyst, analytics power user, citizen data scientist, front-line worker

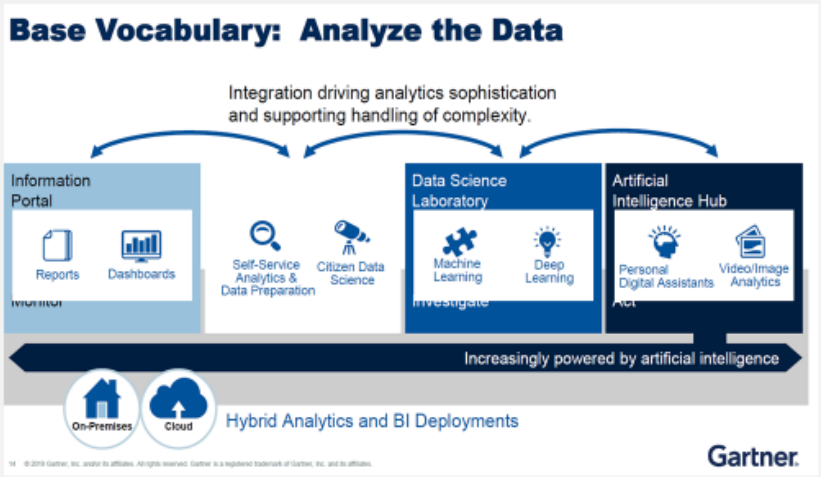
https://www.gartner.com/document/3983896?ref=cust_reco_sdemail&docType=RESEARCH

BUILDING THE CASE FOR DATA LITERACY

DEMONSTRATING THE CASE FOR DATA LITERACY



DEMONSTRATING THE CASE FOR DATA LITERACY

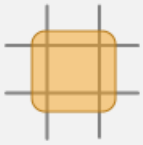


DEMONSTRATING THE CASE FOR DATA LITERACY

Any data literacy program should be planned within the overall context of an explicitly defined and communicated data and analytics strategy, in support of the organizational business and digital strategy

1) Ensure leadership commitment and staff involvement to champion data literacy awareness on an ongoing basis and craft a clear case for change before pursuing the tactical implementations of specific delivery initiatives.
2) Start by implementing a limited-scope and highly targeted pilot program for data literacy skills training.
3) Work with a group of stakeholders that already has enthusiasm and appetite for data and analytics, and that recognize that improving data literacy is a necessary factor for success.
4) Choose to run the pilot in a business area where there is high likelihood of achieving measurable business outcomes.
5) Incorporate a general high-level awareness campaign to raise overall basic understanding of the importance and value of data throughout the organization.
6) Scale the data literacy program and extend the curriculum to a broader range of courses and classes, based on growing demand and identified areas of business impact.
7) Consider people's current learning levels, competencies and desire to participate when inducting them onto the training.

DEMONSTRATING THE CASE FOR DATA LITERACY



8) Design analytics training programs and embed on-the-job learning experiences based on the change readiness of users and their and ability to improve data and analytics competency over time, and as their skills and needs evolve.

9) Work in partnership with your human resources function (HR) and line-of-business leaders to identify the various job roles/personas within your organization that require a given level of data literacy, and then define the learning goals and outcomes required by each role.

10) Map the required levels of data literacy training in the relevant topics and competencies to each job role or persona (for example a retail sales clerk would need a different training plan from a supply chain performance improvement manager).

11) Collaborate with HR and line-of-business leaders to assess skills and training requirements, design upskilling roadmaps, create a curriculum, and determine training performance metrics.

12) Customize courses and classes - Some roles will have a study program that is fully mandatory, some have a mix of mandatory and electives, some may be elective only. Some roles will only need "101" level literacy, some should be all the way at "401."

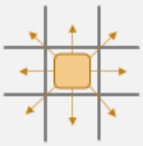
13) Assemble the curriculum into modules for each area of study that are appropriate to various delivery methods. For example, self-learning reading, online computer-based training, informal lunch-and-learn sessions, classroom-based learning, or on-the-job coaching.

14) Augment your capability to develop and deliver the program content by leveraging additional third-party resources, organizations and education institutions that offer solutions, services and courses of study related to data literacy and data-driven business.

15) Monitor the results of improved data literacy within the workforce by using data literacy assessments and by measuring associated improvements to data-driven business outcomes.

(see "Use the Gartner Data and Analytics Compass to Drive Strategy" and "Build Your Digital Business Platform Around Data and Analytics")

DEMONSTRATING THE CASE FOR DATA LITERACY



- To overcome barriers and be successful, data and analytics leaders must plan a change management program that includes deliberate training, coaching and awareness for data literacy skills across the workforce.

- Gartner provides a Toolkit that defines an overarching program for a range of data and analytics topics that are all potentially appropriate to data literacy and are broken out into a series of related courses and classes. This overarching curriculum is then configurable, just as any academic curriculum for a program of study would be defined (please see below)

Suggested Courses and Classes

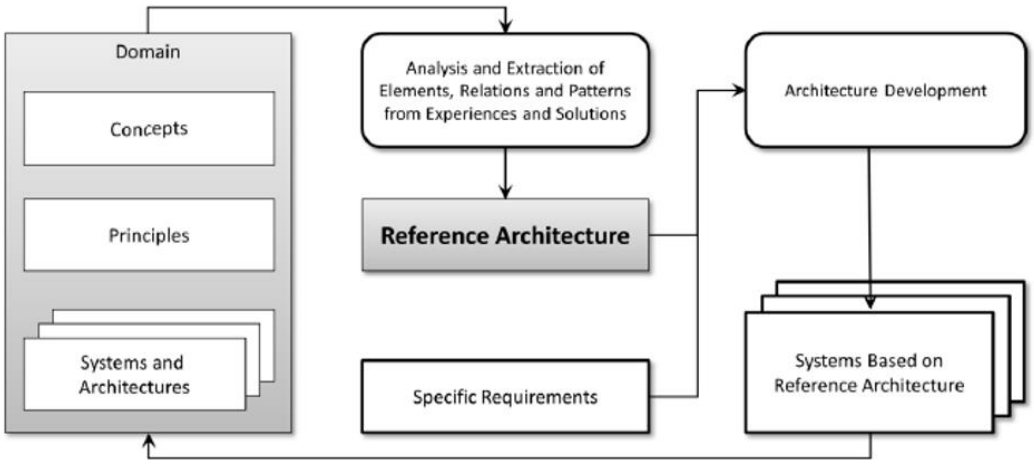
D. Technology

Purpose and Scope of This Section

Purpose	**To propose a standard data management and analytics technology architecture model aligned to business outcomes and use cases common to utilities.**
In Scope	• Conceptual and logical reference architecture for data & analytics. • Covers a subset of analytic centric utility use cases. • Covers end to end data analytics life cycle (i.e. data sources to insights & actions) • Covers common data types & integration patterns.
Out of Scope	• Organizational structure and operating model for delivering data management & analytics. • Approaches to data management & analytics governance. • Detailed solution designs. • Physical architecture and deployment models. • Risk & cost treatments of architecture approaches and building blocks

I. Reference Architecture

Definition and Purpose



- An architecture model comprising of a set of conceptual building blocks for data management & analytics.
- Covers a wide range of analytic centric utility use cases.
- Covers end to end data management & analytics life cycle (i.e. data sources to insights & actions)
- Covers common data types & integration patterns.
- Provides best practices and guidance to develop concrete analytical solutions.

Guiding Principles

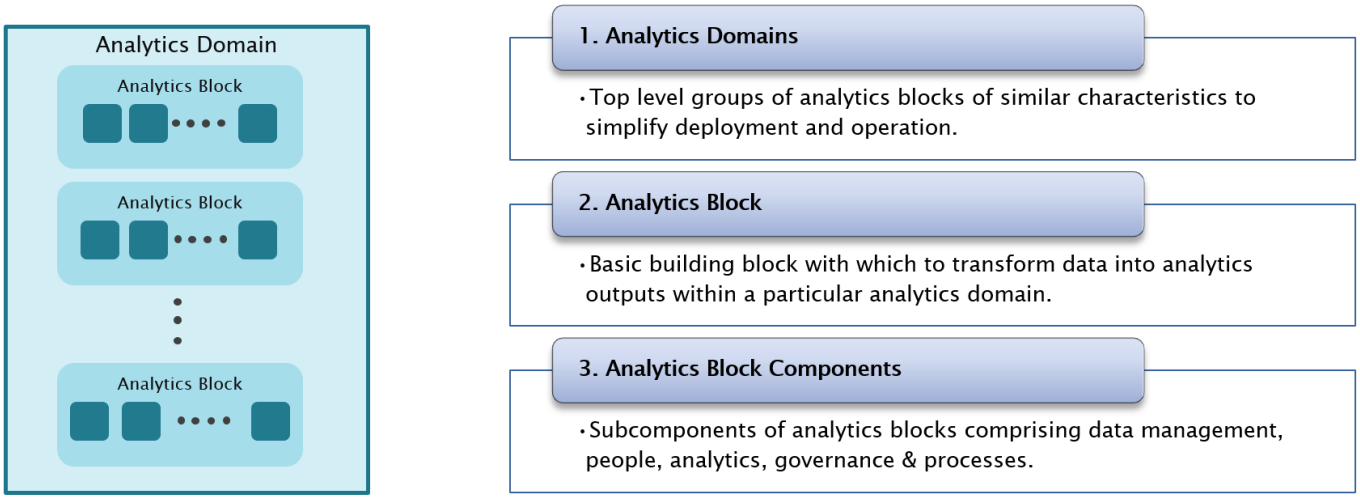
- Reference architecture must be driven by business outcomes and associated analytical use cases.
- Reference architecture building blocks must be able to realize different types of analytical solutions across a utility's functional areas.
- Reference architecture must be realizable through commercially available technology products (solution blue prints).
- Reference architecture should be flexible enough to be implemented incrementally.
- Reference architecture must be able to functionally and technically scale out in line with business needs.

Key Features of a Reference Architecture

- Secured access to accurate data any time/ anywhere.
- Less data preparation lead time.
- Improve data timeliness.
- Keeping up with market trends on data visualization and analytical tools.

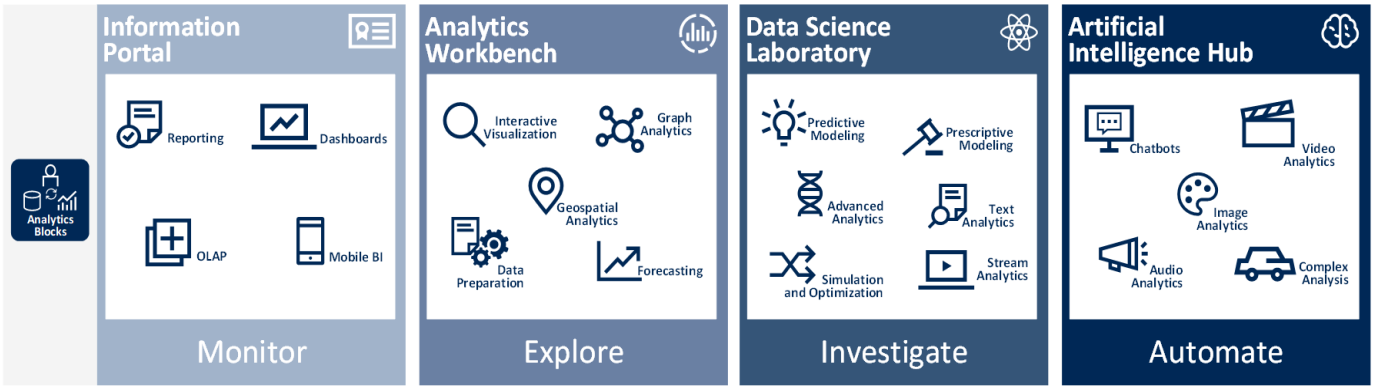
- Retention of historical analytical outputs.
- Ability to aggregate internal data sources with external data sources or cloud based data sources.
- Consistent data access mechanism.
- Reduction of data movement and multiple data copies.
- Visibility on enterprise wide data usage and data access control (optimum level of data security control)
- Scalable data access architecture (supporting variations in workload types & volumes)
- Supports and enforces good data & analytics governance practice.
- Supports enterprise views of data.

Building Blocks of Data Analytics Reference Architecture



Building Block 1 – Analytics Domains

Analytics blocks are diversified but can be grouped into clusters of similar characteristics called analytics domains to simplify deployment and operation.



Fast, easy to consume, trusted access to reports and dashboards

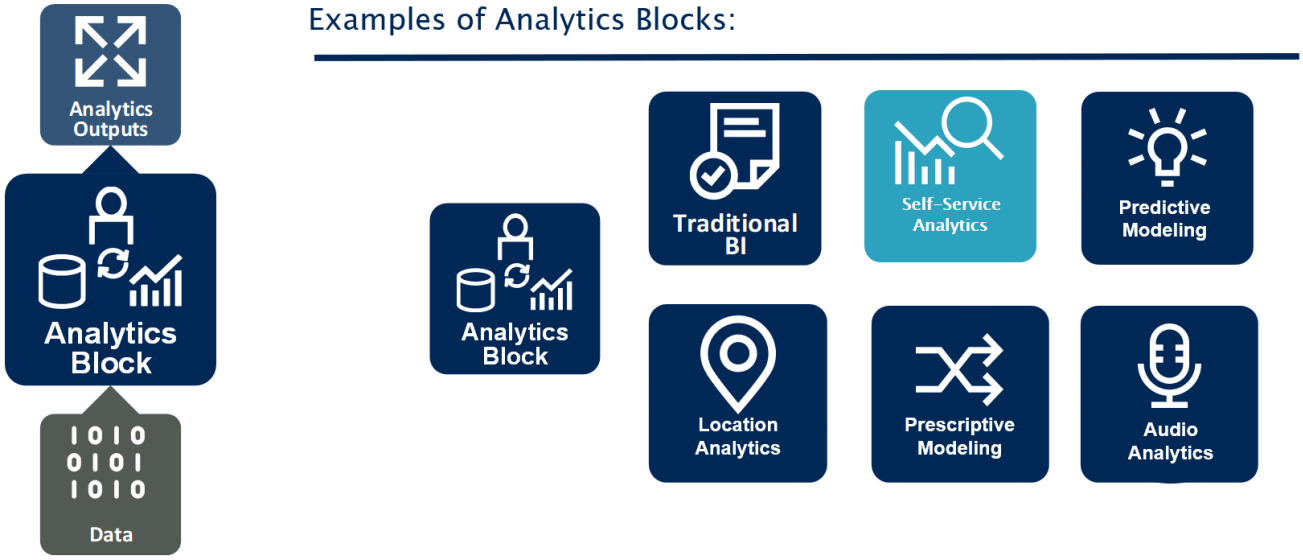
Tools for business users to prepare analytics data and author content

Tools for data scientists (and CDS's) to deliver ML and data science

Packaged AI Services that deliver NLP and computer vision based capabilities

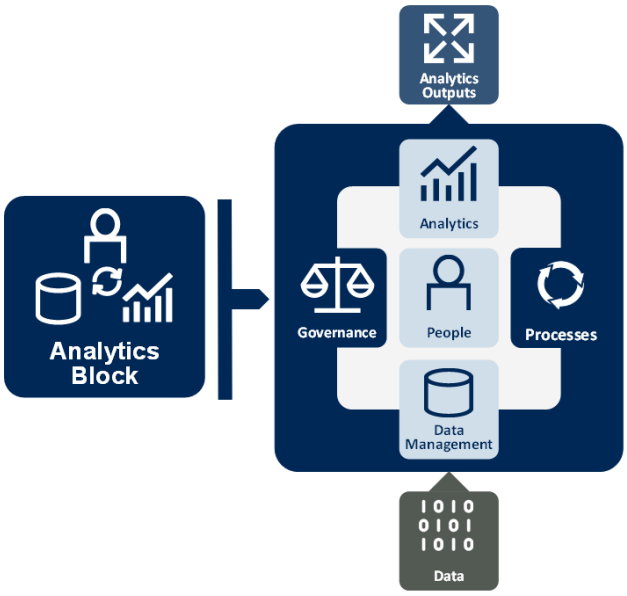
Building Block 2 – Analytics Blocks

Analytics Blocks are granular analytics functions, and their technical and organizational components are able to transform data into analytics outputs to support business outcomes.



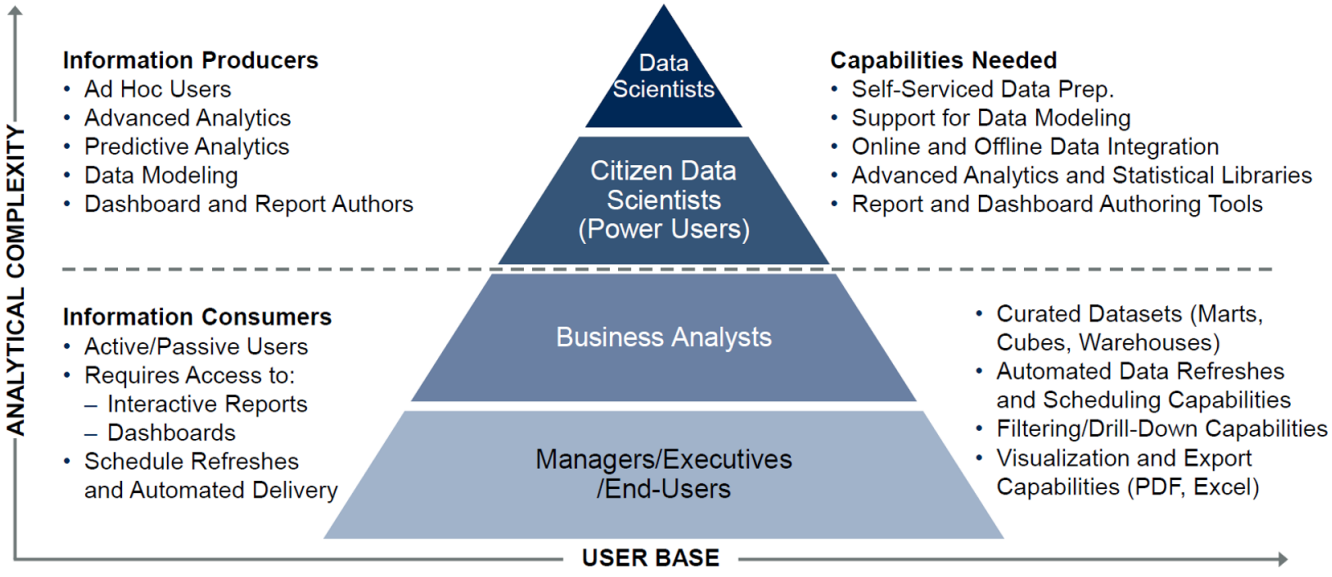
Building Block 3 – Analytics Block Components

Analytics blocks are composed of several components that are required for their deployment and operation. Misuse, lack of alignment or gaps in these components will usually lead to sub optimized use of analytics blocks.



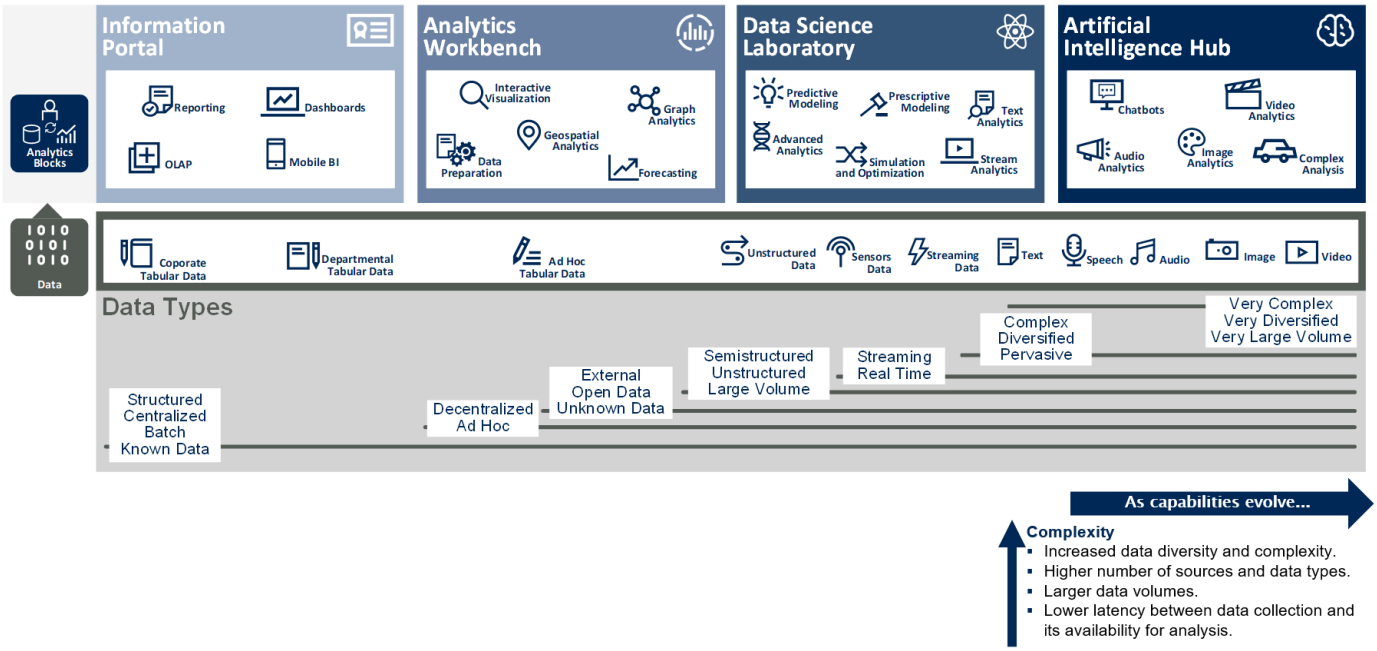
Analytics Outputs	Results generated by analytics processes; can interface directly with users, be operationalized through integration into systems or used as inputs for other analytics processes.
Processes	Series of steps or actions required to deploy and operate the components on the analytics block and generate the expected analytics outputs.
Analytics	Analytics technologies and techniques used to transform data into insights or decisions.
People	Set of data and analytics roles needed to support or execute the processes on the analytics block; can also be called "people capabilities" or "skills."
Data Management	Technologies to collect, process, organize and make data available to analytics processes and users. The main components are often data repositories.
Governance	Set of rules that guarantee the integrity, security, privacy, availability and usability of data for the analytics processes.
Data	Diverse types of inputs for the analytics blocks operation, generated by entities such as internal business applications, external ecosystems, connected "things," customers or users.

Aligning Analytical Capabilities to End User Needs



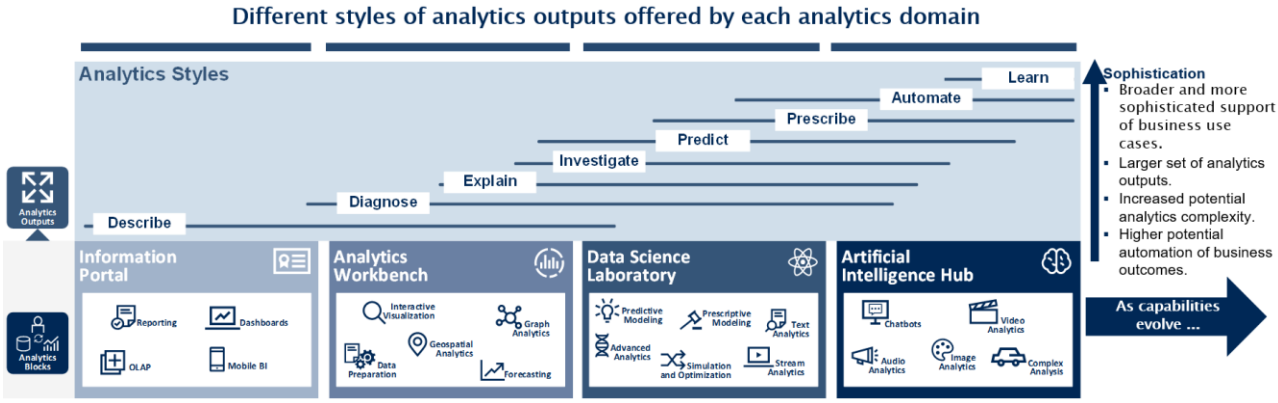
Analytics Domains Data Requirements

As more analytics capabilities are added across the different domains, or incorporated from analytics silos in the organization, the data requirements will become more challenging.



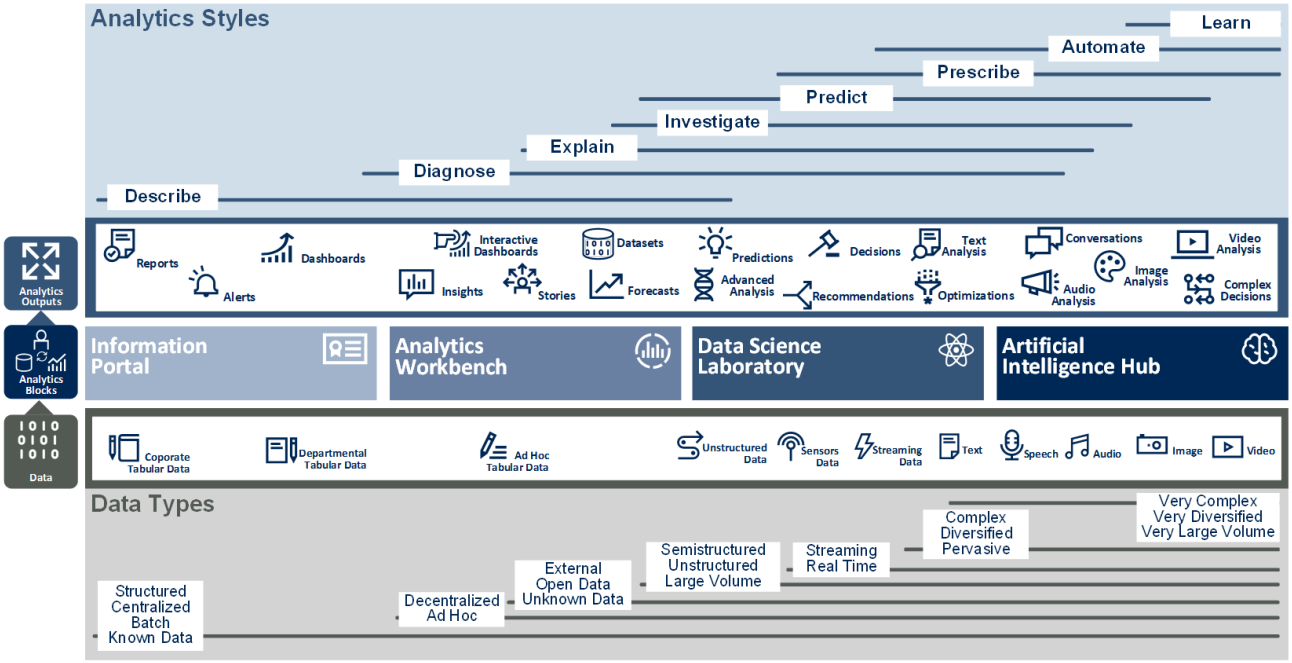
Analytics Domains Outputs

As more capabilities are added across the different domains, the potential styles of analytics become more diversified and sophisticated. It will be possible to solve more complex problems and address a broader range of use cases.



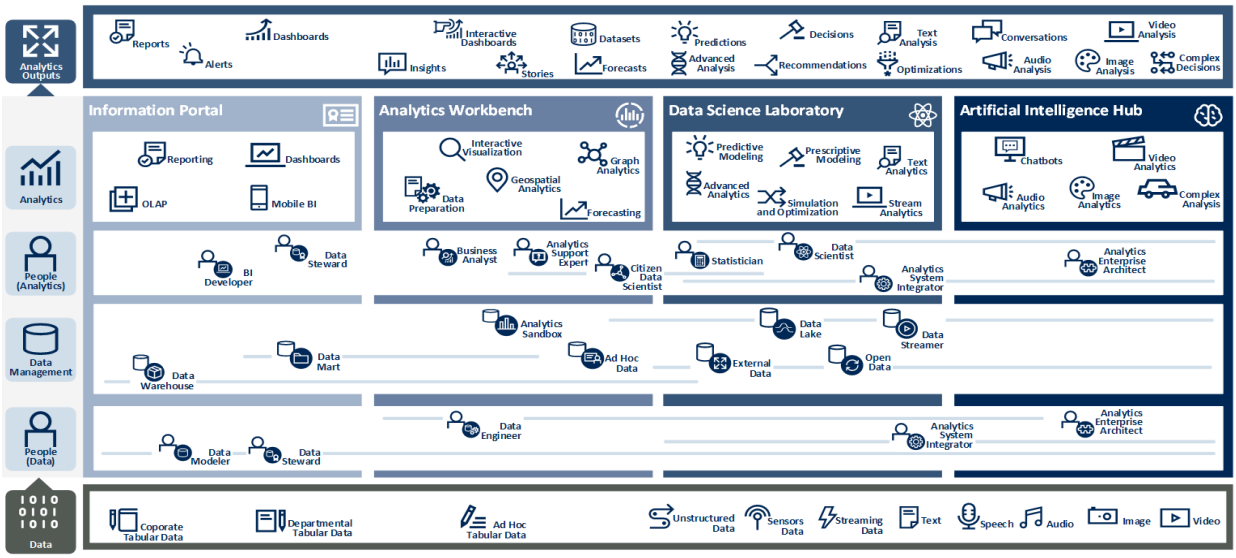
Analytics Domains Data Requirements and Analytics Outputs

The analytics domains will have different requirements for data and produce different types of analytics outputs.

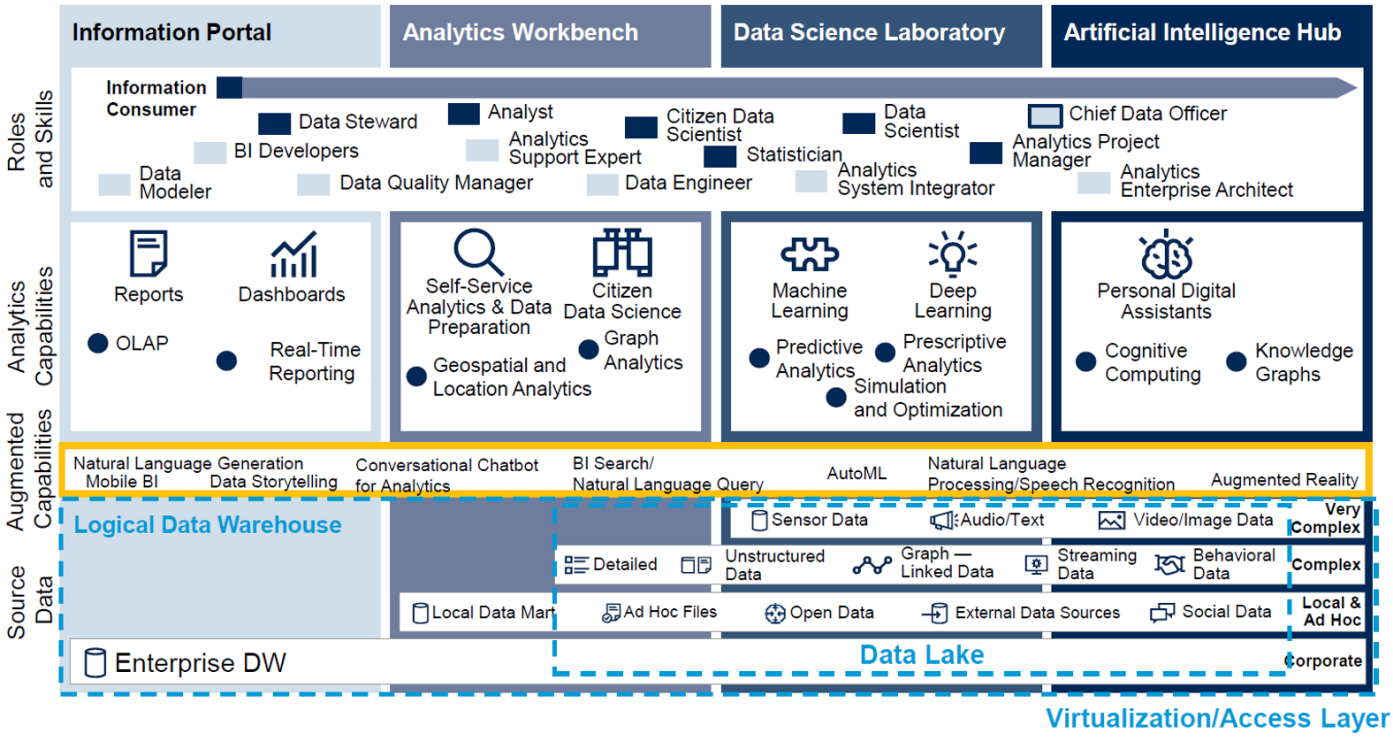


The Reference Data Analytics Architecture

The reference data analytics architecture is a high-level representation of how to plan, deploy and operate a comprehensive data analytics portfolio. It includes data inputs, the required capabilities for data management, people and analytics, and the potential analytics outputs.



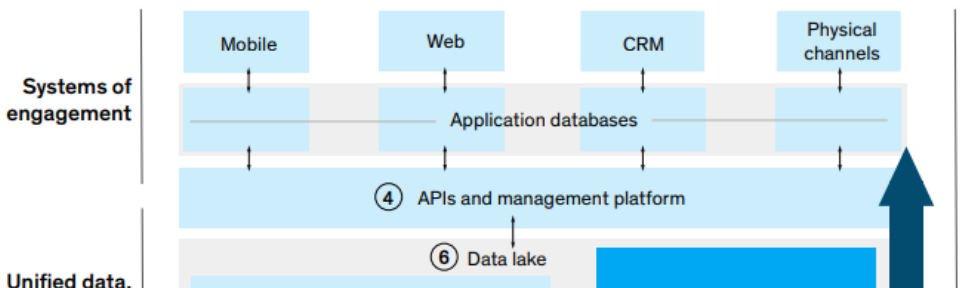
Data Virtualization and Access Layer



* See appendix for details

Emerging Trends in Data Architecture

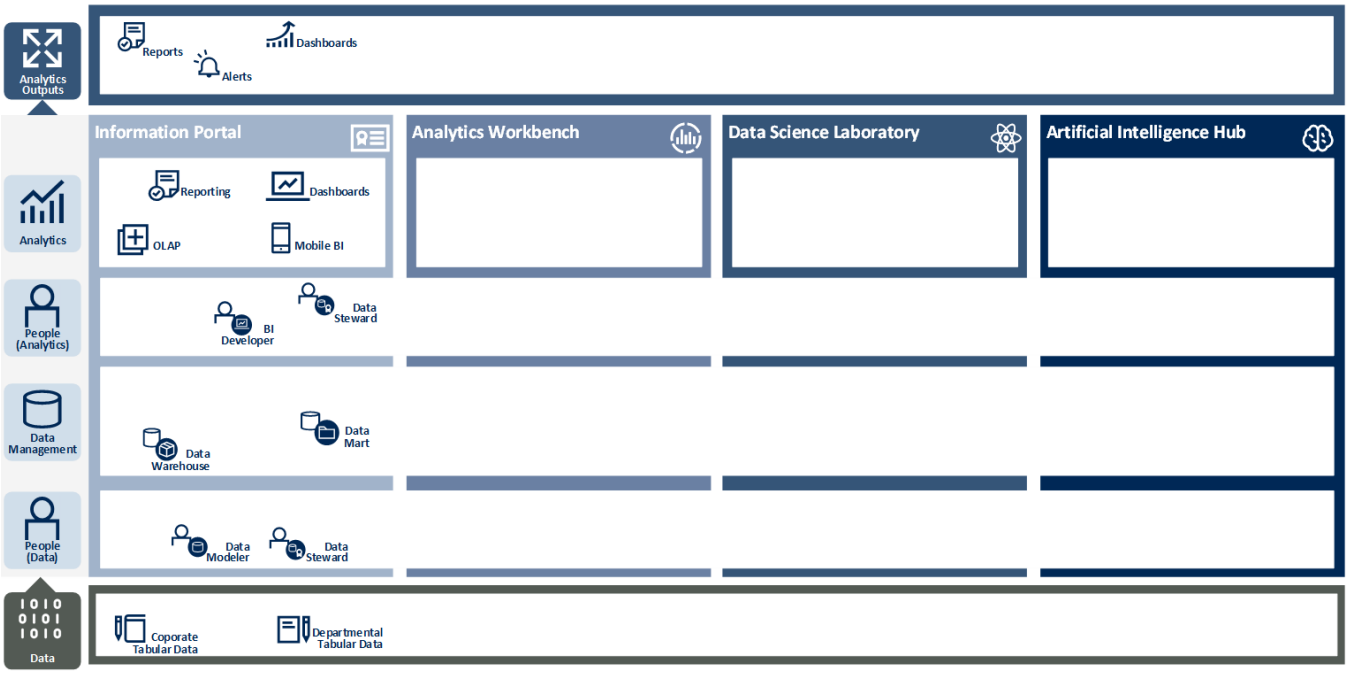
1. Favoring cloud based data platforms over on premise deployment.
2. Shifting more towards real time data processing compared to batch.
3. Favoring 'best of breed', modular data management products over monolithic.
4. Favoring data access through simple, secure and standard integration technologies over point-to-point techniques.
5. Favoring 'data domain based' architecture to a central enterprise data warehouse.
6. Favoring flexible data schemas to rigid data models.



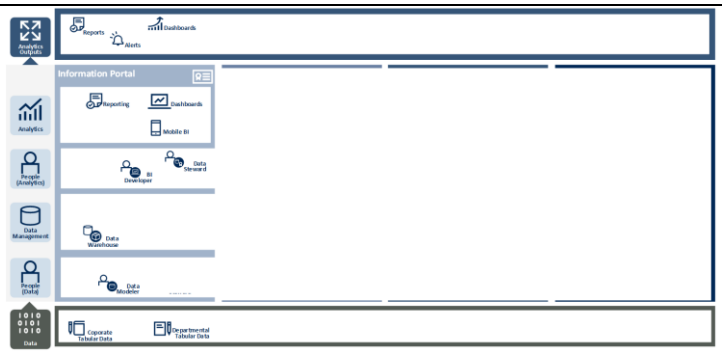
Enabling Reference Architecture

- ### Analytic Domain: Information Portal

The information portal is centralized and favors governance and trust in information over agility and user autonomy. Data and analytics teams play a major role in the development of structured data repositories and the design of BI content.

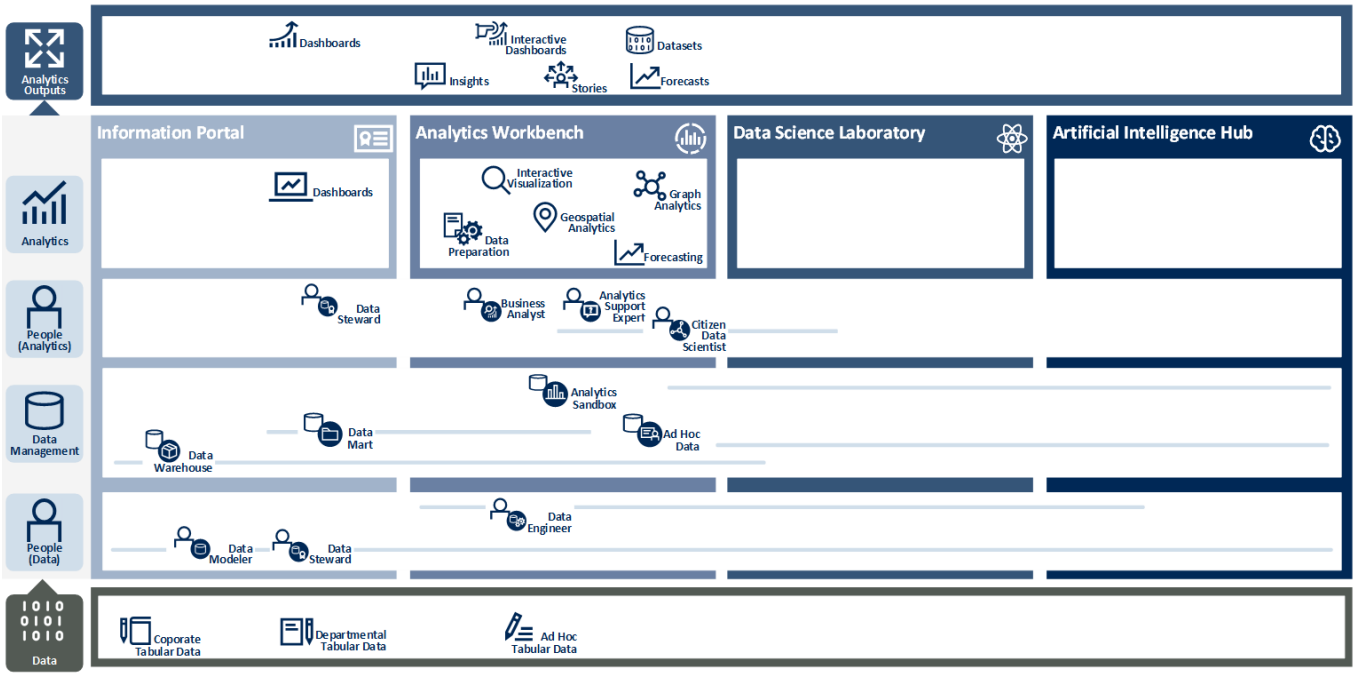


- ### Sample Information Portal Use Case

Analytics Use Case	Supply Chain Analytics
Functional Domain	Enterprise Analytics
Analytic Type	Descriptive
Description	Purchase order analysis | Stock in transit | Inventory turn over
Associated KPI	Fill rate | Lead time | Quality
Key Business Outcomes	Improve operational efficiency | Reduce cost
Analytic Domain	Information Portal
Analytic Block	Dashboard | Reporting
Key Data Sources	ERP | EDW
Solution Components	

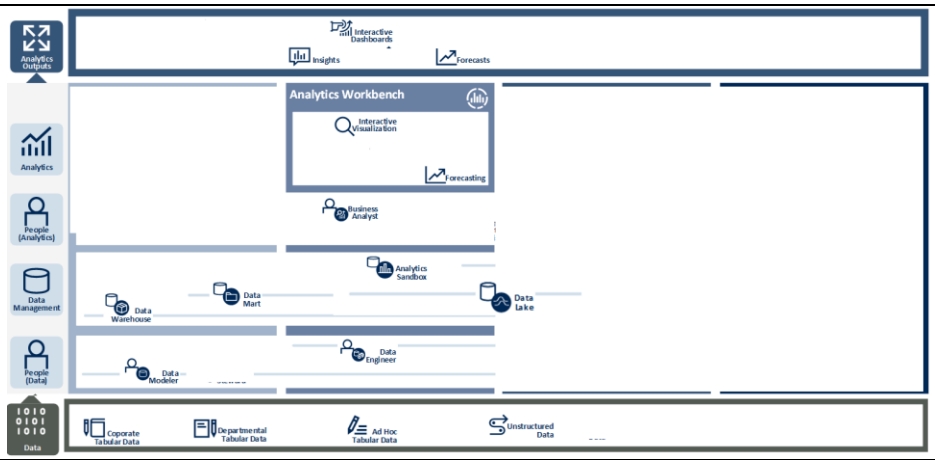
- **Analytic Domain: Analytics Workbench**

User empowerment (self-serve analytics), quick access to data, and easy-to-use visual interfaces are the top characteristics of an effective analytics workbench. The data and analytics team support business users with data preparation, easy and secure access to data, autonomous exploration of data and management of data exploration platforms.



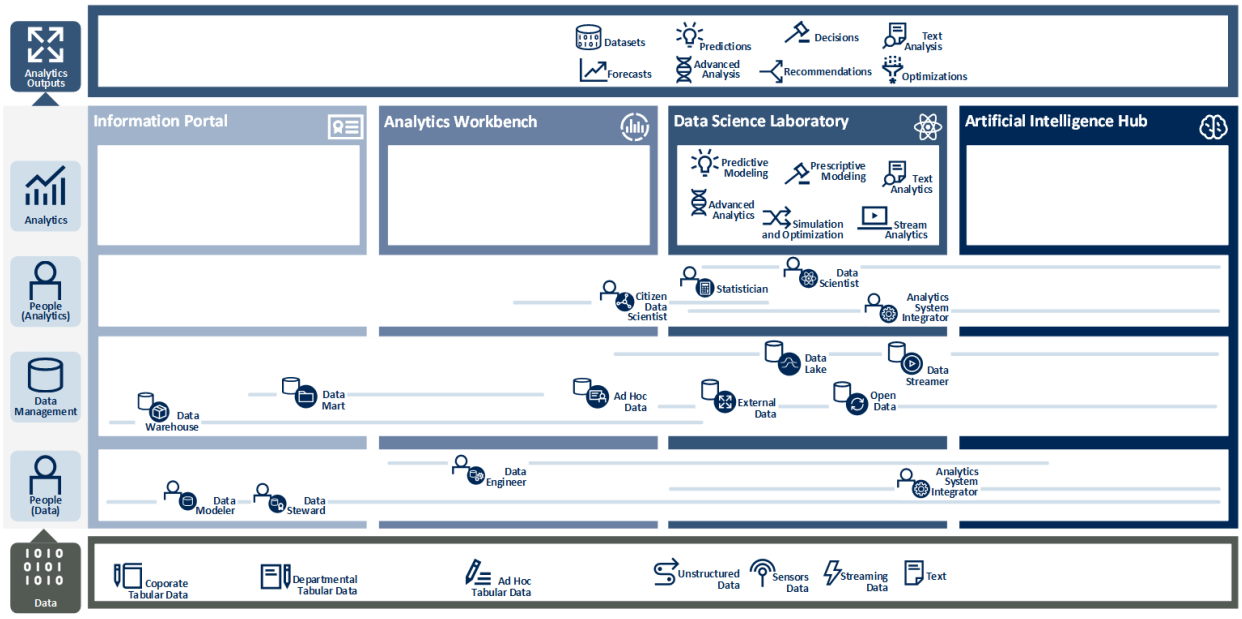
- **Sample Use Case for Analytics Workbench**

Analytic Use case	SMI Theft Analytics
Functional Domain	SMI/AMI Analytics
Analytic Type	Diagnostic
Description	Theft detection | Temper detection
Associated KPI	% Reduction in energy theft | Revenue recovery/year
Key Business Outcomes	Reduce electricity theft | Recover revenue| Reduce cost of energy
Analytic Domain	Analytics Workbench
Analytic Block	Interactive Visualization | Forecasting
Key Data Sources	MDMS | ERP | GIS
Data Management	EDW | MDMS | Energy data mart

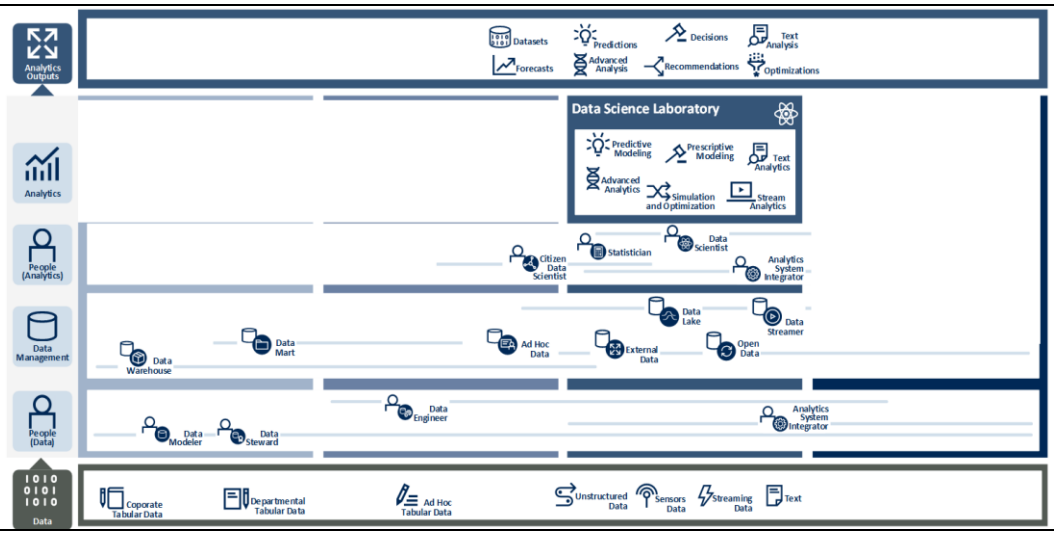
Solution Components	

- **Analytic Domain: Data Science Laboratory**

The data science laboratory relies on experts (with deep business, analytics and technical knowledge) to explore data with a diverse set of advanced tools and methods. The potential range of data source types is broad and can be used to generate analysis with high business impact.

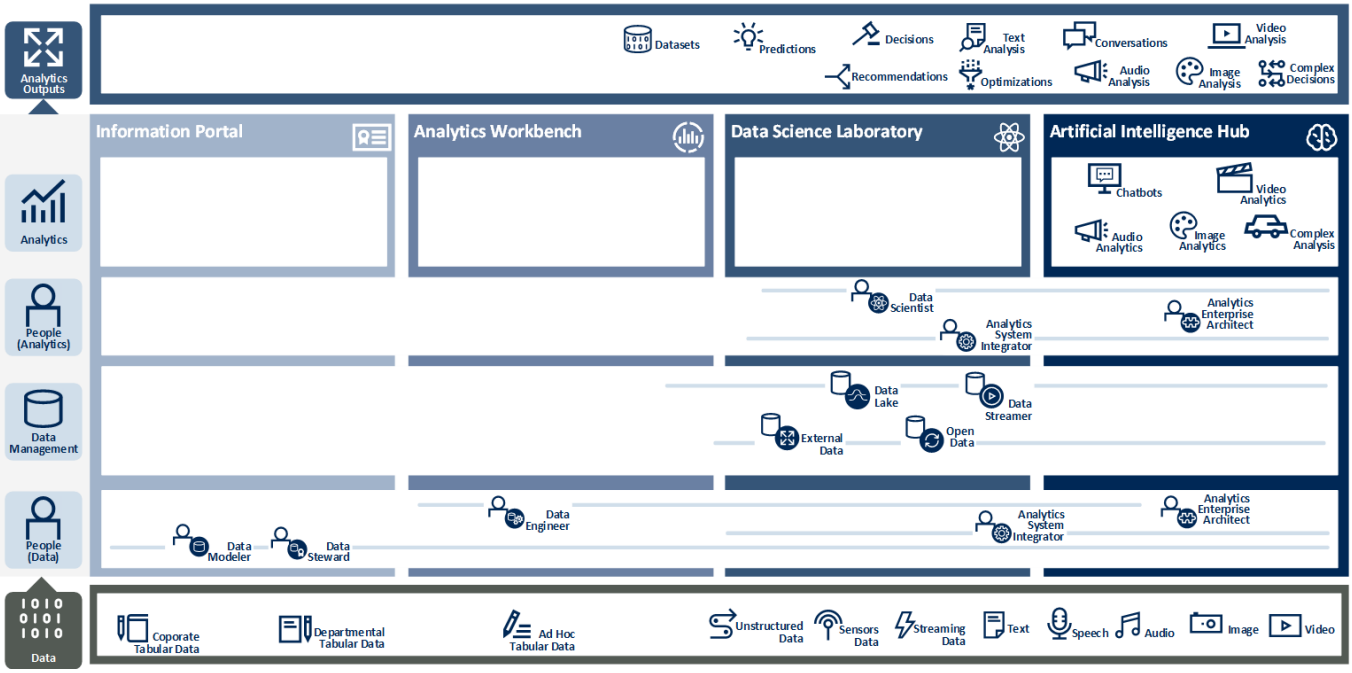


- ## Sample Use Case for Data Science Laboratory

Analytic Use case	Predictive Asset Analytics
Functional Domain	Grid Analytics | RE/DE Analytics
Analytic Type	Predictive
Description	Asset investment planning | Asset performance management | Predictive reliability
Associated KPI	SAIDI-SAIFI-MAIFI | % At-risk assets | % Unplanned expenditures
Key Business Outcomes	Improve system reliability | Reduce asset TCO | Improve safety & compliance.
Analytic Domain	Data Science Laboratory
Analytic Block	Advanced analytics | Predictive modelling | Stream analytics
Key Data Sources	ERP (EAM) | GIS | SCADA | IoT Sensors
Solution components	

- ## Analytic Domain: Artificial Intelligence Hub

The artificial intelligence hub uses sophisticated algorithms trained by data scientists to learn and perform complex tasks. It can automate processes that would normally require human intervention, such as understanding an audio recording of a customer request or operating a machine.



Additional References, Bibliography and Resources:

- The Data Governance Institute (DGI) - http://www.datagovernance.com
- The Data Management Association (DAMA) - https://dama.org/content/what-data-governance
- Gartner – https://www.gartner.com/en
- Utility Analytics Institute (UAI) - https://utilityanalytics.com/
- Gartner toolkit for data and analytics architecture. Available at: https://www.gartner.com/en/documents/3986181/toolkit-creating-a-modern-data-and-analytics-strategy-an
- The Role of Big Data Analytics in Smart Grid Management. Available at: https://link.springer.com/chapter/10.1007/978-981-15-0135-7_38
- A practical approach for power utilities seeking to create sustaining business value, Cognizant. Available at: https://www.cognizant.com/whitepapers/a-practical-approach-for-power-utilities-seeking-to-create-sustaining-business-value-codex2826.pdf
- KPI Dashboards. Available at: https://kpidashboards.com/kpi/industry/utilities/
- https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/How%20to%20build%20a%20data%20architecture%20to%20drive%20innovation%20today%20and%20tomorrow/How-to-build-a-data-architecture-to-drive-innovation.pdf

Glossary

Word / Acronym	Definition
AI	Artificial Intelligence
AHI	Asset Health Index
AMI	Advanced Metering Infrastructure
AMS	Asset Management System
API	Application Programing Interface
AVL	Automatic Vehicle Location
BI	Business Intelligence
BIA	Business Intelligence and Analytics
CAO	Chief Analytics Officer
CDO	Chief Data Officer
CDS	Chief Data Steward
CIO	Chief Information Officer
CIS	Customer Information System
CRM	Customer Resource Management
DR	Demand Response
DW	Data Warehouse
DS	Data Steward
EAS	Energy Analytics System
EDW	Enterprise Data Warehouse
ETL	Extract, Transform, Load
EV	Electric Vehicle
GIS	Geographic Information System
IoT	Internet of Things
IIoT	Industrial Internet of Things
JIT	Just-In-Time
KPI	Key Performance Indicator
MDM	Master Data Management
ML	Machine Learning
NLP	Natural Language Processing
NTL	Non-Technical Loss
OMS	Outage Management System
PMU	Phasor Measurement Unit
RE/DE	Renewable /Distributed Energy
SCADA	Supervisory Control and Data Acquisition
TDD	Test-Driven Deployment
VPP	Virtual Power Plant

electricity.ca | electricite.ca