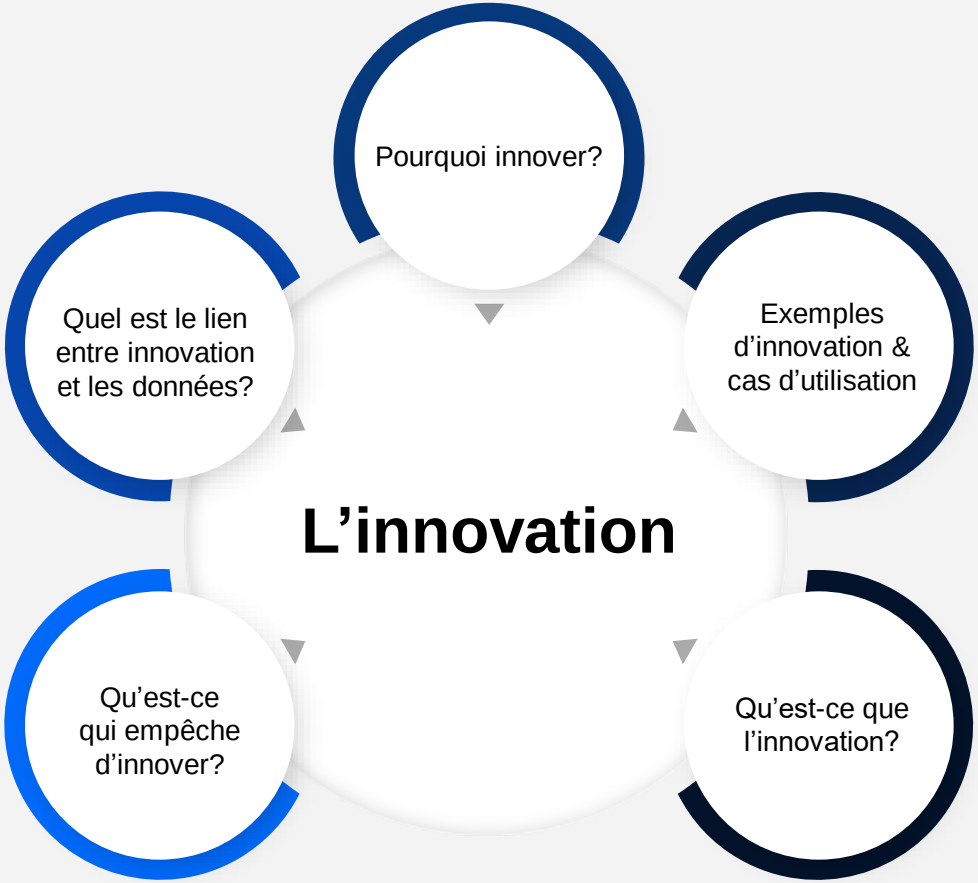
KPI DIGITAL

La stratégie de données
Un catalyseur pour l’innovation!

Richard Langlois, Vice-président données et analytiques

Contenu



Pourquoi innover ?

INCREASE REVENUE

REDUCE COSTS & OPTIMIZE HEADCOUNT

IMPROVE CUSTOMER EXPERIENCE

IMPROVE EMPLOYEE EXPERIENCE & TEAM EFFICIENCY

OPTIMIZE PROCESSES

GAIN COMPETITIVE EDGE

Pourquoi innover ?

PROVIDE SPEED TO VALUE

IMPROVE DECISION-MAKING & DISCOVER NEW INSIGHTS

ACHIEVE ENVIRONMENTAL SUSTAINABILITY

ENHANCE EXISTING PRODUCTS & SERVICES

ENABLE NEW BUSINESS MODELS & ANTICIPATE COST OF DISRUPTION

MITIGATE RISK OF NON-COMPLIANCE

And while nearly 90% of executives surveyed say innovation is a top-three priority, fewer than 10% are satisfied with their organizations' performance on that front.

– McKinsey

Exemples d’innovation

SAIC Maxus

At SAIC Maxus in China, the company developed a web application to enable customers to customize and place orders and then track the production status.

They use **3D and digital twin design**, which uses a digital version to represent the physical asset to configure and produce the cars.

Then they applied an **automated smart engineering system and an AI quality assurance tool** to check and identify errors.

Outcome:
This has led to shorter time to market 35% and increased accuracy 99.8% – Forbes

Exemples d’innovation

Unilever

A local entrepreneurial team has **digitally enabled end-to-end quality management** at a **Unilever** factory in Dubai, UAE.

Outcome:
A reduction of material waste by **more than 40%** with limited investment and in a short time period. – *Forbes*

Unilever with eCommerce giant, Alibaba, have created recycling machines that use artificial intelligence (AI) to automatically identify and sort plastics for recycling in China.

Exemples d'innovation

Henkel in Düsseldorf, Germany, has developed a **unique cloud-based data platform** that connects more than 30 sites and more than 10 distribution centers in real time

This helps to:

- **meet growing customer and consumer expectations on service and sustainability.**
- **simultaneously achieve double-digit cost and inventory reductions - Forbes**

Adhesive Technologies Markets

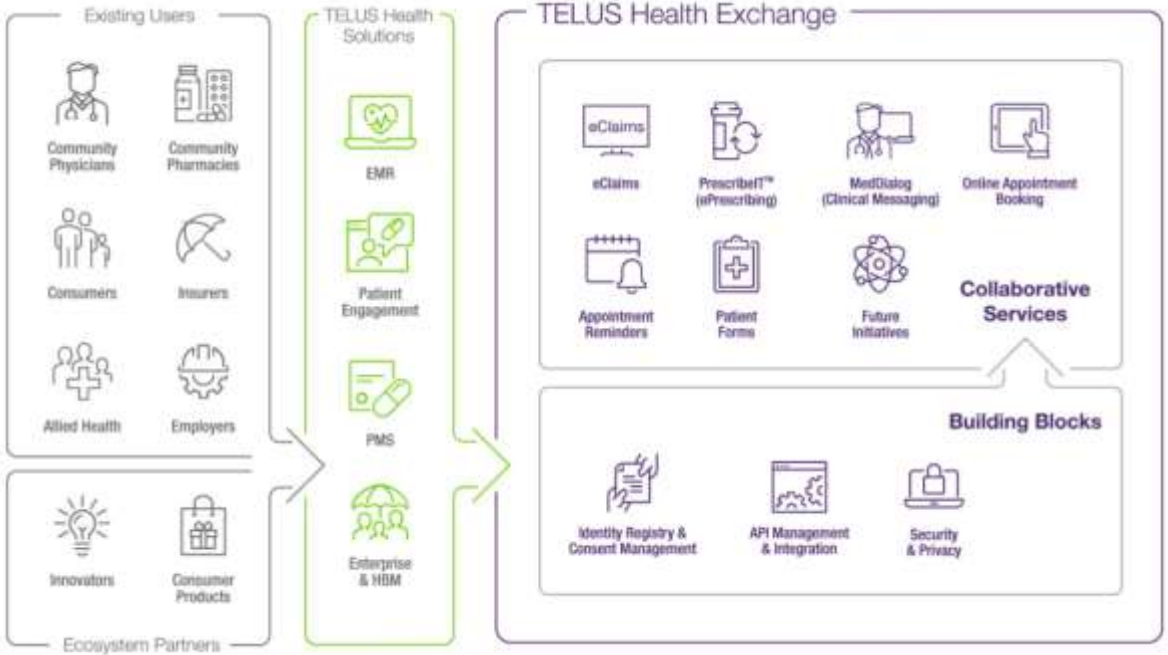
Study: Nearly **$1 trillion wasted** in U.S. healthcare annually - Journal of the American Medical Association

Exemples d'innovation

Telus Health Exchange (THX)

Better for care providers
We're changing the way healthcare providers interact with patients. THX incorporates cutting-edge technology to improve communication between doctors and patients—like virtual visits, electronic prescription services, online appointment booking, and real-time messaging services.

Better for patients
Patient-centric care is our priority. By integrating digital-forward services offered by our partners into the TELUS Health Exchange, we can help improve the patient experience, facilitate faster access to healthcare information and services, and empower Canadians to manage their health.



Driving Strategic Business Objectives & Value.

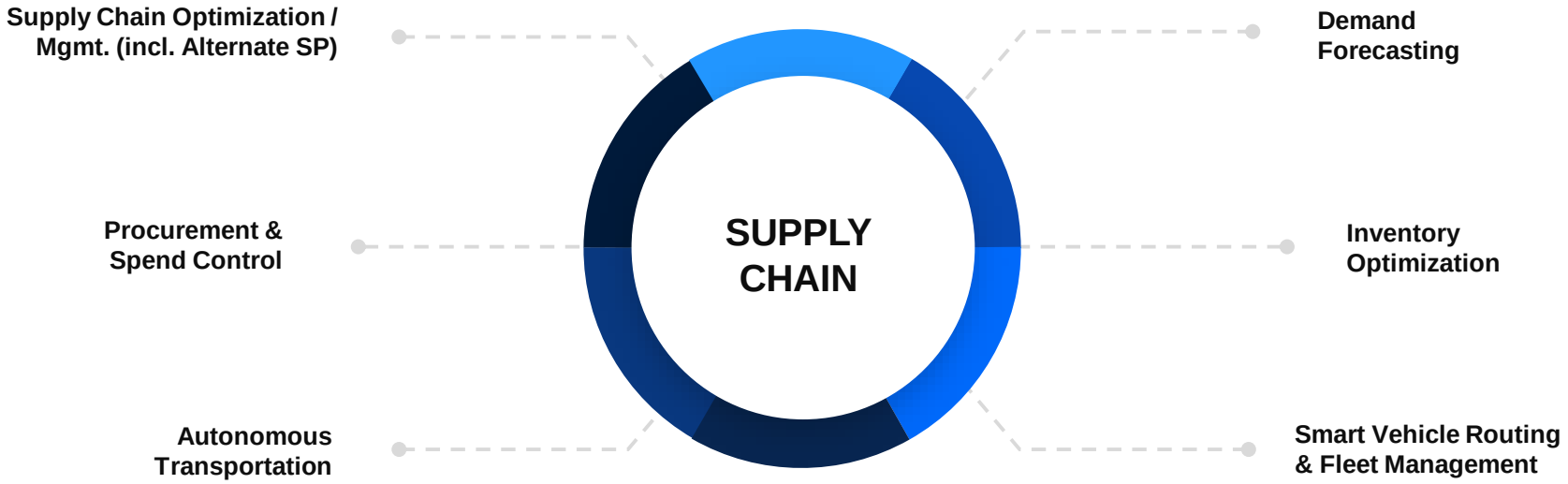
Use cases across the enterprise

Driving Strategic Business Objectives & Value.

Use cases across the enterprise

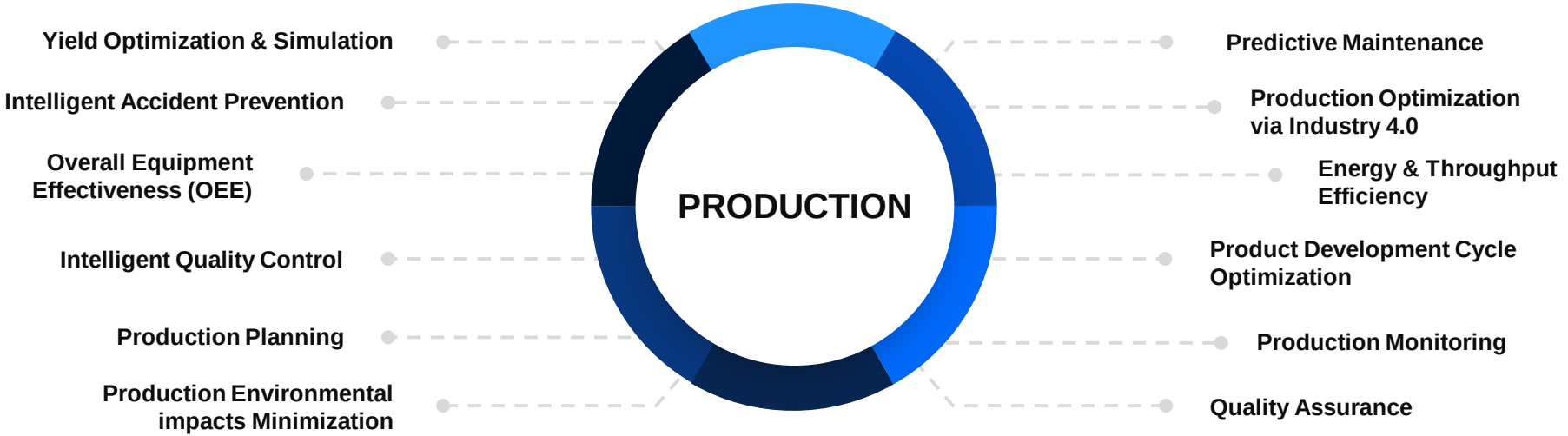
Driving Strategic Business Objectives & Value.

Use cases across the enterprise



Driving Strategic Business Objectives & Value.

Use cases across the enterprise



Driving Strategic Business Objectives & Value.

Use cases across the enterprise

Driving Strategic Business Objectives & Value.

Use cases across the enterprise

Qu’est-ce que l’innovation?

Innovation

[inəˈvāSH(ə)n] **noun.**

‘Something (process, product or service) original or improved which creates value’

‘Execute an idea which address a specific challenge and that produces value for the enterprise and the client’

Substantial.

Scalable.

Repeatable.

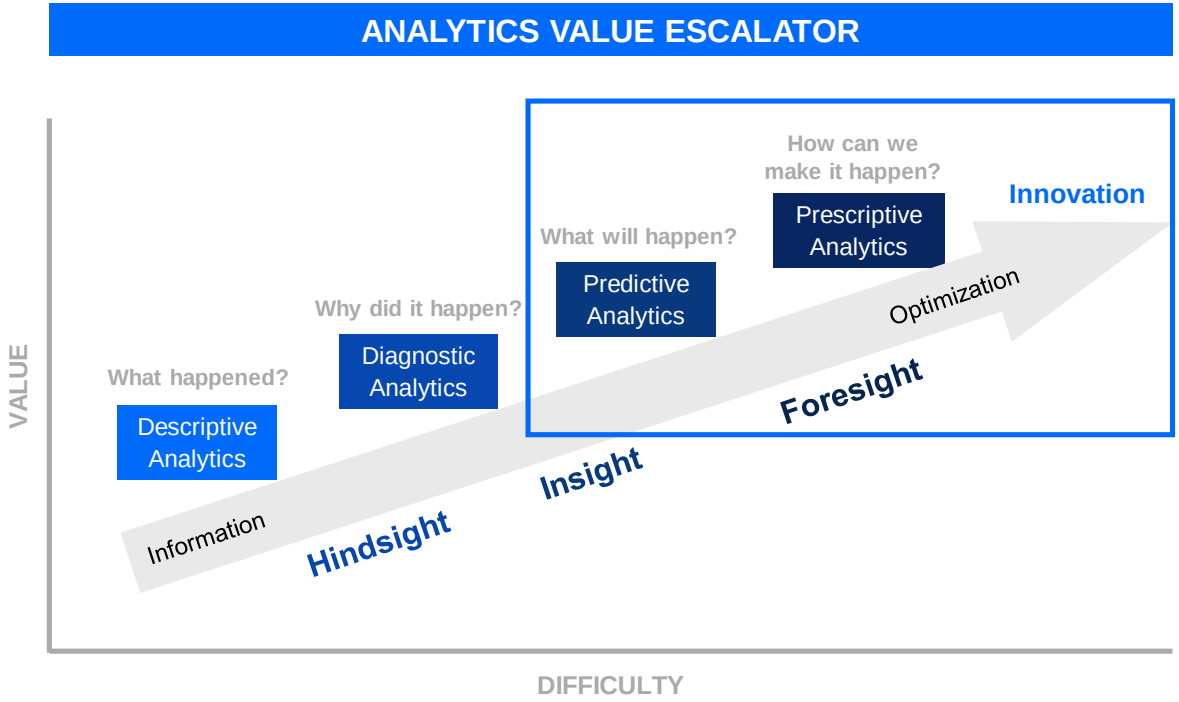
Incremental.

Disruptive.

Sustainable.

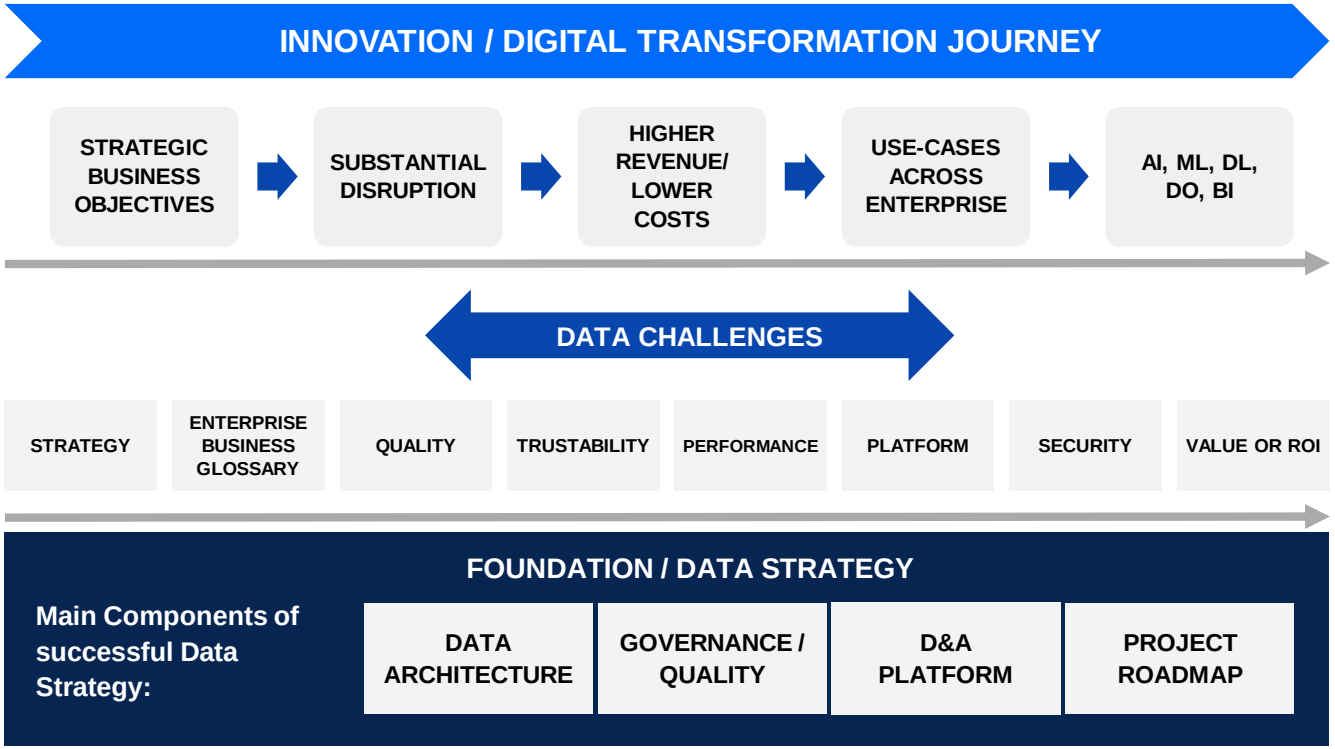
Qu’est-ce que l’innovation ?

Via l’intelligence artificielle (IA)

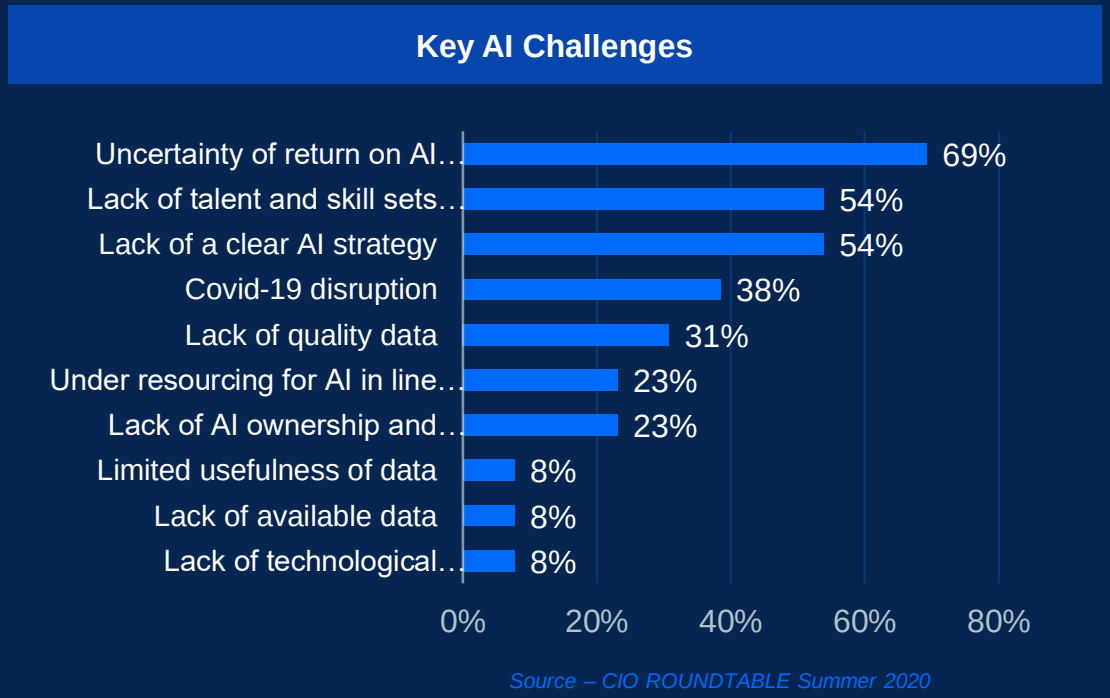


Qu’est-ce qui empêche d’innover ?

Data Challenges



Qu’est-ce qui empêche d’innover ?



Source – CIO ROUNDTABLE Summer 2020

Typical Challenges

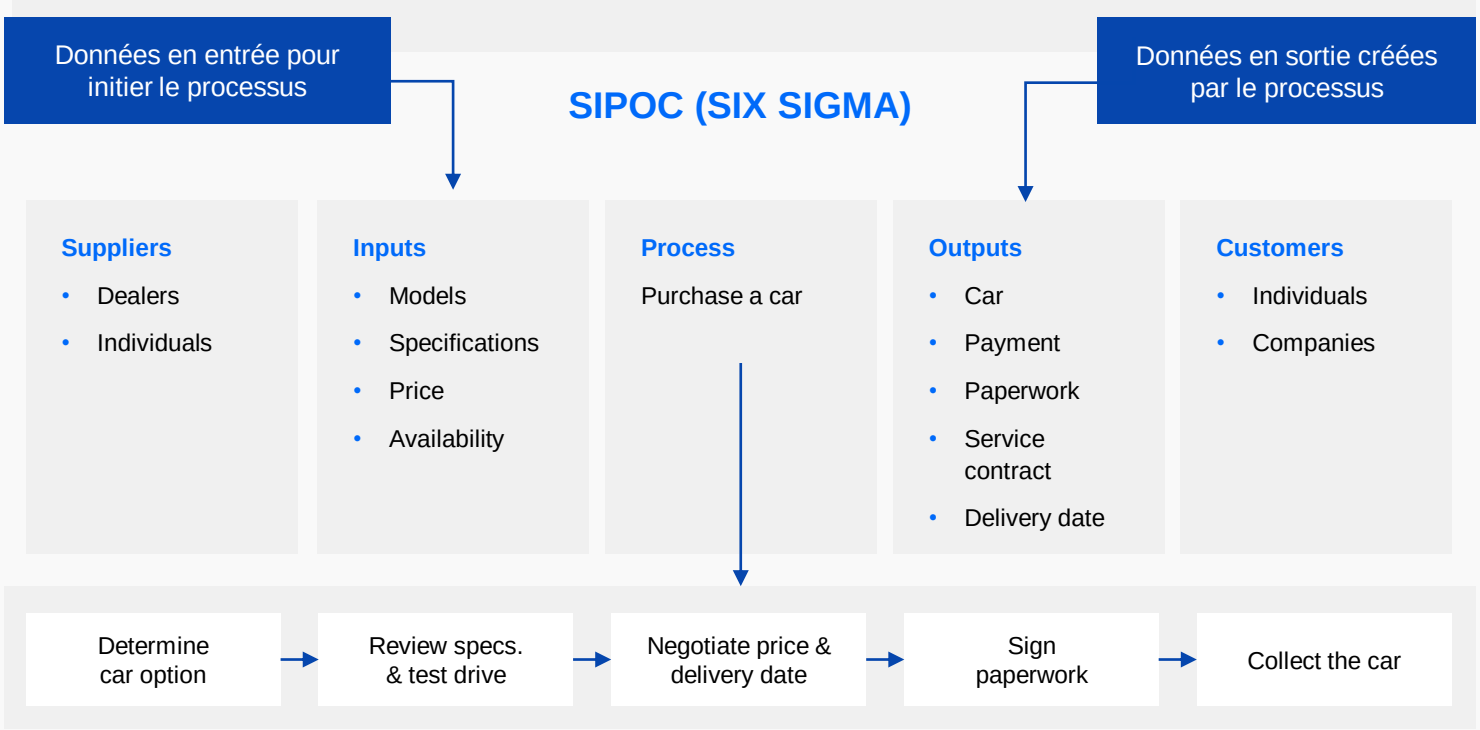
- Lack of clear AI Strategy
- Multiple data sources, lacking proper governance
- Lack of proper architecture and data & analytics platform to support AI
- Lack of AI Leadership
- Lack of AI talents

Qu’est-ce qui empêche d’innover?

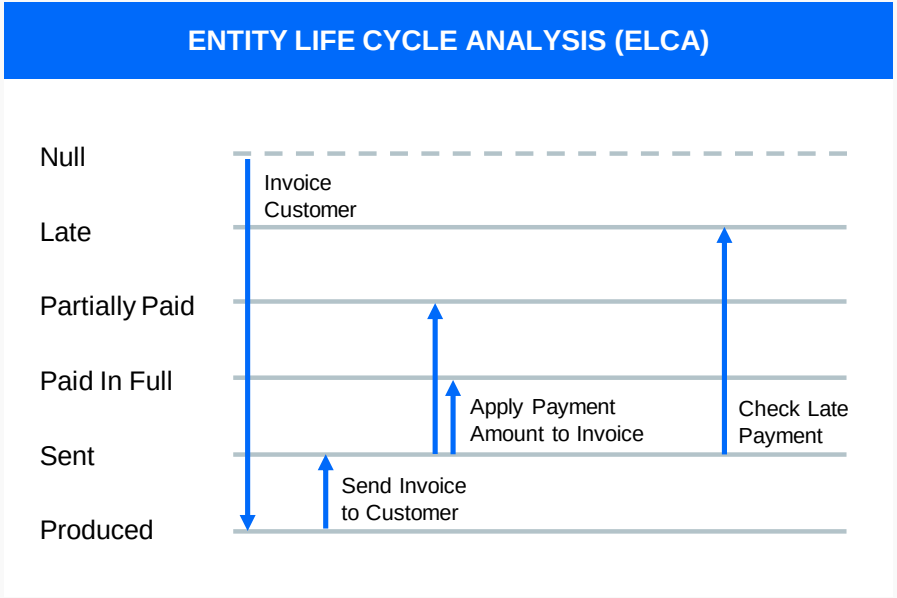
Certaines pratiques de gestion.

- Peu de projets innovants et challengeant sont définis (étant plus risqués…)
- Capitulation architecturale / historique
- Rigidité des processus de gestion de projets et budgétaire
- Manque de KPI sur l’innovation

Quel est le lien entre “quelque chose – processus” et les données ?



Quel est le lien entre “quelque chose – processus” et les données ?



Il y également d’autres artefacts qui établissent des liens entre processus et données, par exemple: les matrices CRUD, Event-Driven Process Chain (EPC), Business Process Modeling (BPM)

Data is embedded in every aspect of innovation.
By 2025, every decision, interaction & process will come from data and most employees will use data to optimize nearly every aspect of their work.

– IBM

STRATÉGIE DE DONNÉES GAGNANTES.

- Pourquoi une stratégie de données?
- Les éléments clés d’une stratégie de données gagnantes
- Les MVx (Minimum Viables x)
- Comment faire des Roadmaps?
- Comment vendre les projets innovants à l’interne?

Pourquoi une stratégie de données ?

How can data help you achieve your organizational objectives and deliver your business strategy?

To inform decision-making

To understand customers and trends

To provide smarter services and products

To comply with regulations and maintain reputation and trust

To create additional revenue/monetize

To improve supply-chains

To improve operations

To achieve environmental sustainability

*“The data strategy helps you **drill down to your core business needs and create an achievable plan for the future (roadmap).”***

– Forbes

A McKinsey Analytics survey of **1,000 companies with > $1 billion in revenue** and spanning 13 sectors and 12 geographies, has identified an elite group of companies (the breakaway companies) that is achieving the elusive goal of **analytics at scale.**

CEDM
Conceptual Enterprise Data Model

MDM
Master Data Management

DG
Data Governance

Modern
Data & Analytics Platform

Les éléments clés d’une stratégie de données gagnantes.

Most successful data strategies include 4 key elements.
– McKinsey

The KPI Data & Analytics Offerings.

1 Vision & Strategy

- **1.1 Global Data Strategy (incl. roadmaps)**
- **1.2 DL and DW Unification – Vision**
- 1.3 D&A initiatives Business Cases

2 Governance & Architecture

- **2.1 Data Governance Strategy**
- **2.2 Data Management (incl. Data Quality)**
- **2.3 Conceptual Enterprise Data Model (CEDM)**
- **2.4 MDM - Architecture & Governance**
- 2.5 D&A Platforms – Architecture & Governance
- 2.6 DL & DW Unification – Architecture & Governance

3 Data & Analytics Platform

- 3.1 Recommendations
- 3.2 Solution / Technical Architecture
- 3.3 Sizing
- **3.4 Trusted, bimodal, multi- usage D&A platform Implementation**

Implement:

- 3.5 Information Gov. Catalog
- 3.6 Data Science Frameworks
- 3.7 Visualisation tools
- 3.8 Optimisation tools
- 3.9 Cloud Pak for Data

4 Data And Analytics Value Chain

Manage Master Data	Collect for Analytics	Organise For Analytics	Analyse ML/DL/NLP/AI	Deploy / Infuse	Monitor
4.1 MDM - Executing	4.2 IoT & Other RT Ingestion 4.3 ETL / ELT 4.4 ETL Migration to modern D&A platform 4.5 Pattern based ETL	4.6 LakeHouse Architecture 4.7 Dimensional Enterprise Data Model (DEDM) 4.8 Physical Data Model (PDM) 4.9 DW & DM Architecture 4.10 Information Catalog / Glossary	4.11 Reports & Dashboards 4.12 OLAP 4.13 ML, DL, NLP/U models 4.14 DO models	4.15 Dashboards & Reports 4.16 Predictive Models & Algorithms deployment 4.17 DO deployment 4.18 Callable AI	4.19 Data Ingestion and Integration Monitoring 4.20 Model Performance Monitoring 4.21 Ingestion / ETL / ELT / Wrangling Monitoring

5 Pre-Built AI Services

- 5.1 Chatbot
- 5.2 Document Analysis
- 5.3 Image processing
- 5.4 Speech to Text
- 5.5 Forecast

6 Other Services

- 6.1 Executive Advisory Services
- 6.2 Outsourcing
- 6.3 Audit
- 6.4 Training

Many D&A projects also include:

- 7.1 Bus. Requirements / Epics / Stories
- 7.2 Business Processes / Objectives
- 7.3 Project Plan, Status and Sprints
- 7.4 Innovation Programs Grants

Les MVx (Minimum Viable x)

Key Deliverables and the MVx…

No MVx on these, since they depend on MVC

- Dimensional Enterprise Data Model (DEDM)
- Messages
- Specifications
- Tests Cases

MVP

Minimum Viable Product

- Wireframes & Dashboards
- AI Use Cases

MVF

Minimum Viable Foundation

- Modern Data & Analytics Platform on Cloud
- Templates
- DevOps
- Data Pipelines
- Cloud Configurations
- Secure Cloud environment (landing zone)

MVC

Minimum Viable Product

- Conceptual Enterprise Data Model (CEDM)
  - CEDM Taxonomy
  - CRUD matrix
- Business Requirements Document (BRD) / Stories

MVA

Minimum Viable Agile

- Project planning, project reporting, Agile and forecasting

MVM

Minimum Viable Capabilities Map

- Capability Map
- 'To Be' Process Model (BPMN)
- SIPOC

MVR

Minimum Viable Roadmap

MVG

Minimum Viable Product

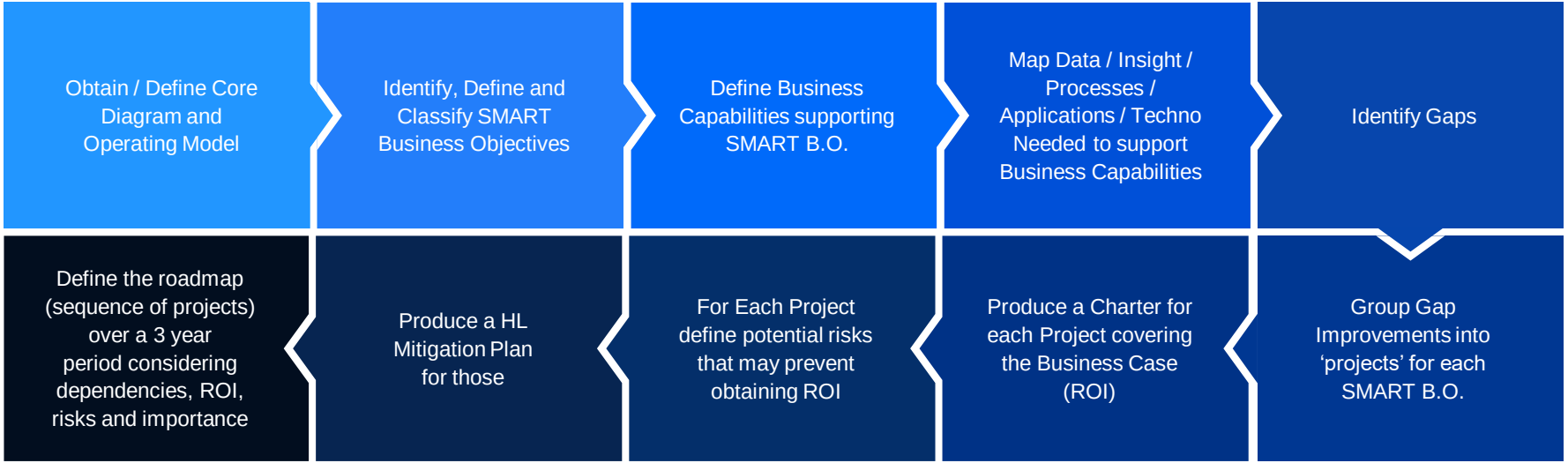
- Data Governance
- Master Data Management (MDM)

MVS

Minimum Viable Solution Architecture

Global Data Strategy – How to Build Roadmaps.

KPI Methodology (for all projects including AI)



- Core Diagrams are useful but not mandatory, Operating Model understanding is mandatory
- SMART = Specific, Measurable, Achievable, Relevant, Timely
- Data/Insights/Processes/Applications/Techno should be considered from the following angles:
  - Suitability to address the Business Capabilities to support the SMART B.O.
  - Manageability (Maintainability/Transferability/Evolution)
  - TCO (Total Costs of Ownership)
  - Solid Data Architecture (allowing mapping to CEDM)
  - Governance & Data Management (incl. security, privacy, compliance)

- Business Cases should involved the Sponsors (BUs), Finance, IT
- AI risk mitigation plan may involve AB testing, simulations and more
- Roadmaps are more precise and of higher quality when done after having defined a CEDM
- Some Business Capabilities may support more than 1 SMART B.O.

Use Cases of interest (an example)

Comment vendre les projets innovants à l'interne ?

Upsell and cross-sell of seasonal products, which includes:

Customers segmentation

Campaigns (promoting products) for specific customers segments

Seasonal Products Bundle Offerings

Product Recommendations at time of ordering

Sophisticated Sales Dashboards

Complete view of the Customers, the Products & Services, which includes:

DG & MDM on Customers

DG & MDM on Products & Services

DG & MDM on relevant reference concepts such as Market, Territory, BU, and so on

Sales Forecasts

Pricing Optimization per Market

Expanding Service Offerings Catalog by combining Seasonal Services & Seasonal Products

ANNEXES.

- Le CEDM (Conceptual Entreprise Data Model)
- La gouvernance de données (un cadre de référence pratique)
- Le MDM (Master Data Management)
- Les plateformes D&A modernes permettant de faire tous les types d’analytiques (ML,DL,DO) pour l’ensemble de l’entreprise

Conceptual Enterprise Data Model (CEDM).

Used for:

- CEDM is the source for all definitions of an Organization Business Dictionary, allowing to anchor a Data Governance program
- The Business Dictionary could be completed with a Data Management framework
- Guide the creation of Logical Models and Physical Databases (integration zone)
- Ideal input to create a Dimensional Enterprise Data Model (DEDM)
- An attributed and typed CEDM is your canonical data model
- Basis for micro-services strategy
- CRUD and SIPOC to integrate data with business processes
- Defining the wireframes structures
- Since business concepts are defined with proper naming, can reuse in BRD and stories (Agile)
- Since business concepts are defined, they are being reused for MDM projects
- Backbone for all your data related deliverables (next slide)

CEDM is the Backbone of Multiple Deliverables.

CEDM Characteristics:
- Enterprise wide and using business language
- Organized in Data Domains and Sub Domains (Taxonomy)

Dashboards / Reports

CEDM (Business Concepts)

BRD, Epics/Stories | **Data Governance** | **MDM** | **Business Processes** | **DEDM** | **Wireframes** | **Business Glossary**

- **Pipeline from Integration Zone to Consumption Zone**
- **Replicating Consumption Zone to Serving Layer**

- **Messages from source systems**
- **Pushing messages to Raw Zone**
- **ETL to Integration Zone**

Data & Analytics Application Architecture
Cloud Secure Environment

These 2 deliverables are independent of the CEDM

Management typically focused on the dashboards/reports, the tip of the iceberg, but great dashboards depends on business understanding (the CEDM) and having clean data

CEDM Benefits.

BENEFITS – BUSINESS RELATED

- A CEDM allows usage of CRUD and SIPOC to integrate data with business processes. This allows identifying missing and/or incomplete business processes. Value could be quite significant.
- A better understanding of a business lead to both incremental improvements and sometimes radical improvements. Value could be quite significant.

BENEFITS – OTHERS

- All projects requiring a data architecture could reuse areas of the CEDM.

BENEFITS – DELIVERABLES RELATED

+80% savings	A well done CEDM allows production of a complete Business Glossary (only cost for the Business Glossary consists of putting in place a procedure to generate the Business Glossary from the CEDM).
+50% savings	Attributing the CEDM allows creation of the Integration Zone of the Data Lake.
+80% savings	Ideal input to create a Dimensional Enterprise Data Model (DEDM).
+80% savings	Defining the wireframes structures.
+50% savings	Since business concepts are defined with proper naming, they are being reused as is in both the BDR and stories (Agile).
+10 - 50% savings	Since business concepts are defined, they are being reused for MDM projects.

CEDM – TAXONOMY.

It takes more time to organize in 3 levels the Data Subjects. But ultimately, you need them. A flat model with hundreds of entities is difficult to use across the enterprise and with multiple projects.

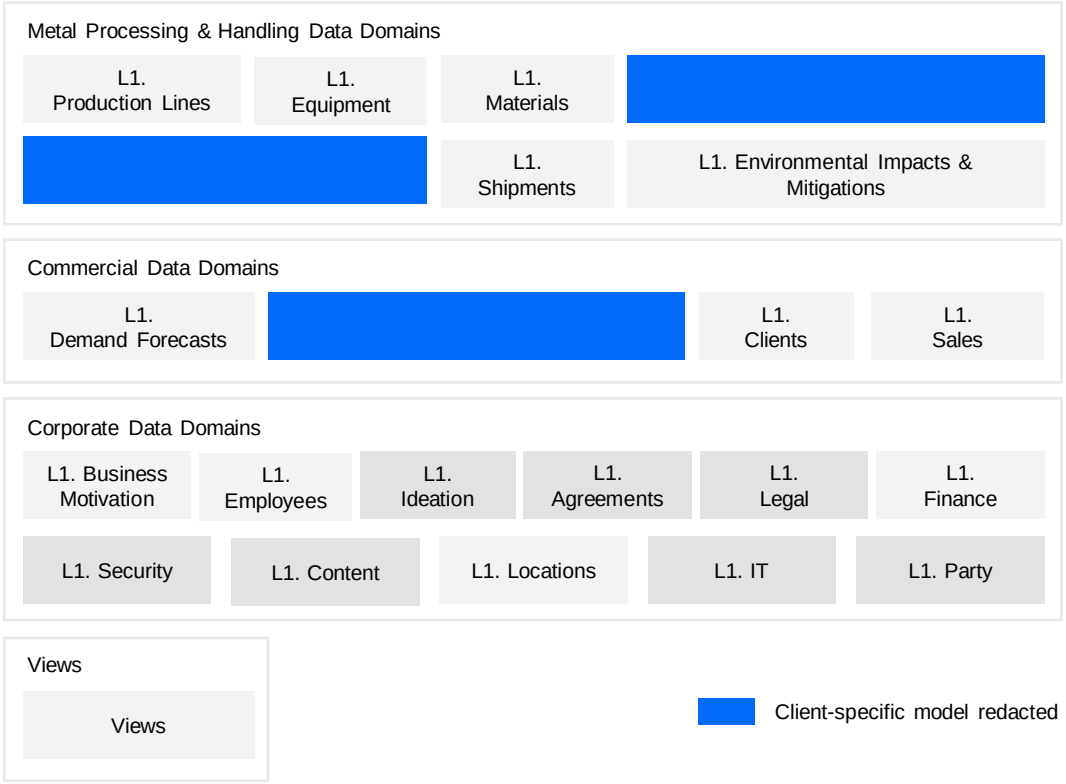
So we build the CEDM initially without Data Subjects and then come back to organize it when not in project critical path.

Project Benefits:

- Much faster, so the team can start other deliverables
- Attribution get done right after per areas that reach stability in order to create the Integration Zone of the Cold Layer (see the D&A Platform Architecture section)

A CEDM IS ORGANIZED BY DATA SUBJECTS

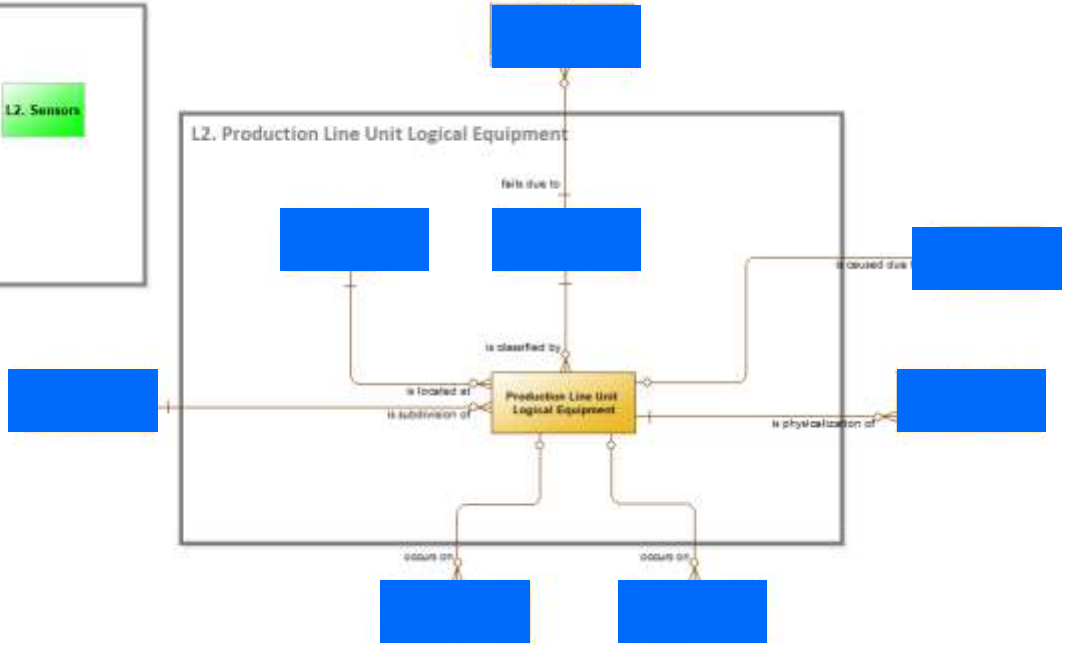
(Shown here are Level 1 Data Subjects)



CEDM – TAXONOMY.

A CEDM IS ORGANIZED BY DATA SUBJECTS
(Shown here are Level 2 Data Subjects which defines the Level 1 – Equipment)

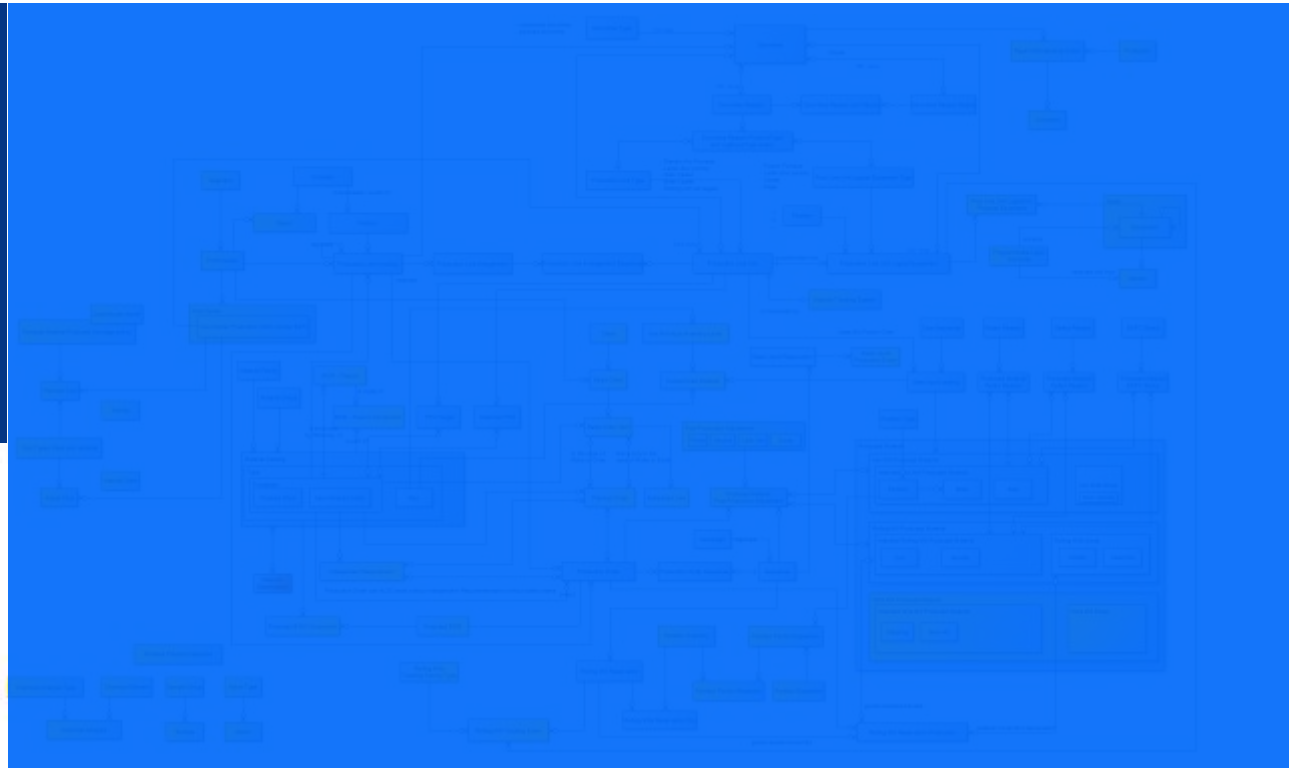
A CEDM IS ORGANIZED BY DATA SUBJECTS
(Shown here is a Level 3 which is the expansion of the Data Subjects L2 – Equipment)



We use PowerDesigner to capture the CEDM (Taxonomy)

CEDM – WHITEBOARDING.

- The process of creating a CEDM is done by whiteboarding with SMEs
- We use YeD software for this purpose
- We model in real-time while facilitating the sessions with the SMEs



In yellow, out of Phase x scope

Data Architecture & Governance Framework.

BUSINESS FUNCTIONS GOVERNANCE

- Financial governance
- HR governance
- Legal Governance
- Ethics Governance
- Engineering Governance
- …

BUSINESS OBJECTIVES
(Smarter Objectives)

BUSINESS OPERATIONS
(processes/capabilities map)

DATA POLICIES, MANAGEMENT & CONTROLS

- Data Governance Organization (Governors, Stewards, Custodians, …)
- Data Policies (includes coverage of regulations)
- Data Management Framework (data classification, security, privacy, data quality (DQ), …

OPERATIONAL DATA GOVERNANCE

CEDM – Conceptual Enterprise Data Model and Business Glossary

Process & Data alignment (CRUD and SIPOC)	Allocating Data Subjects to Gov.	Classifying Data Sensitivity	Physical mapping to CEDM	Implement DM processes, rules and controls
Setting DQ Objectives to reach Bus Objectives	Include DQ projects into Roadmap (+ financing)	Master Data Management (MDM)	Creation of Microservices & monitoring	Monitor DQ, security and privacy progress

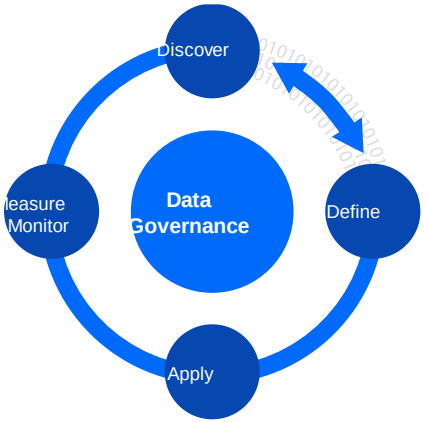
BI GOVERNANCE

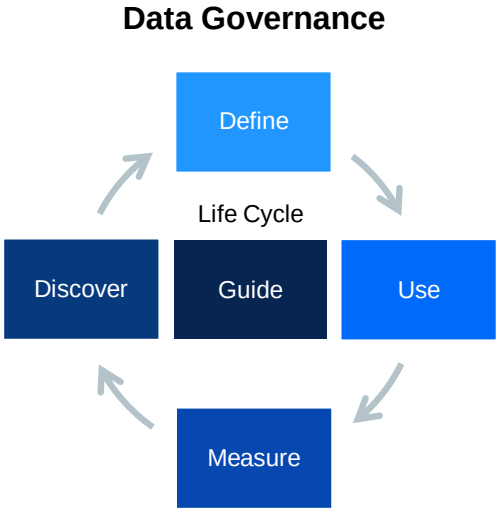
- DEDM – Dimensional Enterprise Data Model (1st step to Dimensional Conformity)
- Base metrics and derived metrics common metadata (definitions, calculations, lineage)
- Access Rights Management (ARM)
- **BI INFORMATION SERVICES & MONITORING**

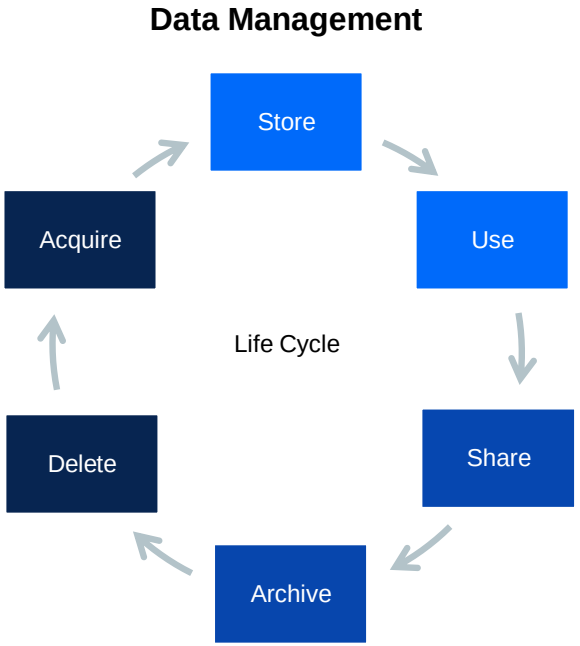
AI GOVERNANCE

- **Responsible AI**
  - Explainable AI
  - Interpretable Machine Learning Technology
  - Ethical AI
  - Secure AI (including ARM)
  - Human-Centered AI
  - Compliance
- **AI SERVICES & MONITORING**

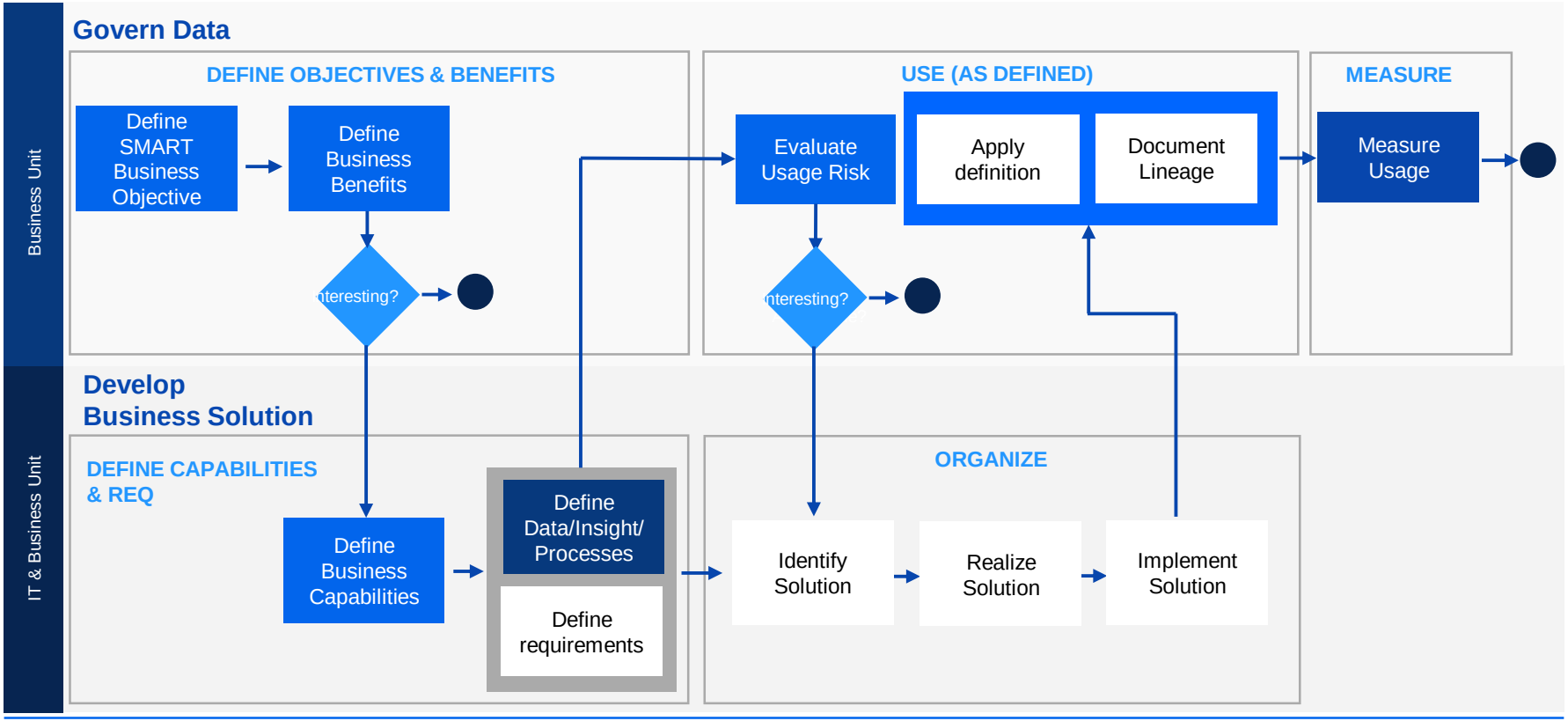
Reference – Life cycle





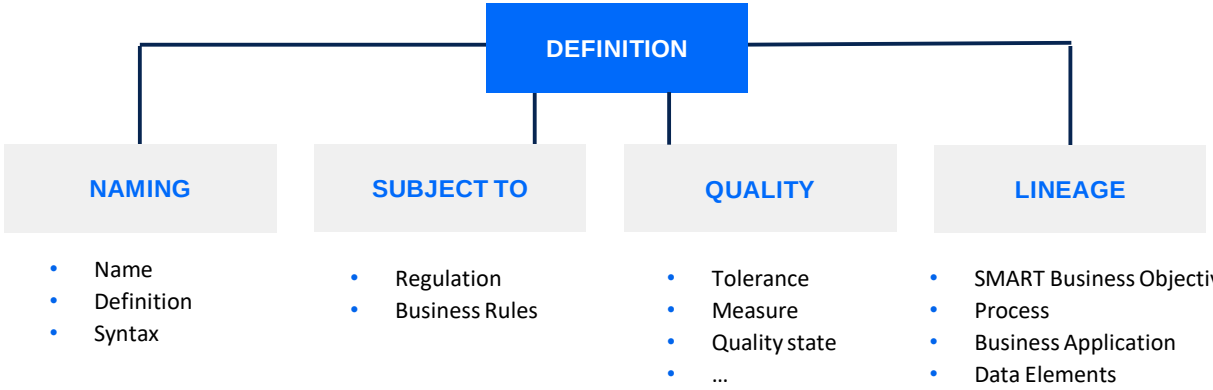


Data Governance & Solution Process



Business Data.

- Represent a concept (tangible or not) capturing an event, a transaction, a context, an idea, a plan (and more) used by business units to support their value chains and business processes. These concepts are made available via a Business Glossary which is an output of the CEDM (Conceptual Enterprise Data Model)
- Is defined via its business metadata
- Could be grouped in Data Subjects



About Policy

A **Policy** [(1)] is a Directive that is <u>not directly enforceable</u> whose purpose is to govern or guide the enterprise. Business Policies provide the basis for Business Rules. Business Policies also govern Business Processes. **Business Rules** provide specific, <u>practicable</u> governance or guidance to implement Business Policies

Examples of Business Policy vs Business Rules

- BP: Depreciation of rental cars must be minimized
- BR: The Car assigned to a Rental must be, at the time of assignment, the one with the lowest mileage of the available Cars in the requested Car Group
- BR: A Rental cannot be extended by phone if the Car's odometer reading is greater than (next service mileage — 500)

Organization Data Policies covers

- Security
  - Data access
  - Logging
  - Monitoring
  - …
- Privacy
- Compliance
- Quality
  - Availability
  - Accessibility
  - Validity
  - Integrity
  - Coverage
  - Timeliness
- Lifecycle
- Data principles in multiple geographies

Mapping data to support the organization's goals and objectives

Example of a Data Access Policy:

'The **<u>purpose</u>** of the data access policy is to ensure that employees have appropriate access to organizational data and information. While recognizing the company's responsibility for the security of data, the procedures established to protect that data **<u>must not interfere unduly</u>** with the efficient conduct of the organization's business. This **<u>policy applies to all business units and to all uses of company data</u>**, **<u>regardless of the offices or format in which the data reside</u>**.

The policy will **<u>protect its data assets through security measures</u>** that assure the proper use of the data when accessed. **<u>Every data item will be classified</u>** by the relevant data steward to have an appropriate access level. Data access will be conducted in accordance with the policies established by the organization.

Any employee or non-employee denied access **<u>may appeal the denial to the Data Governance Committee</u>**; decision is final from the DG Committee.'

Data management enforces Policies, Compliance, Principles and best practices within your organization. What does it cover?

Daily Operations	**Lifecycle Management**	**Back-up / Restore / DR / HA**
Security	**Privacy**	**Data Catalog** (implied by data access policy)

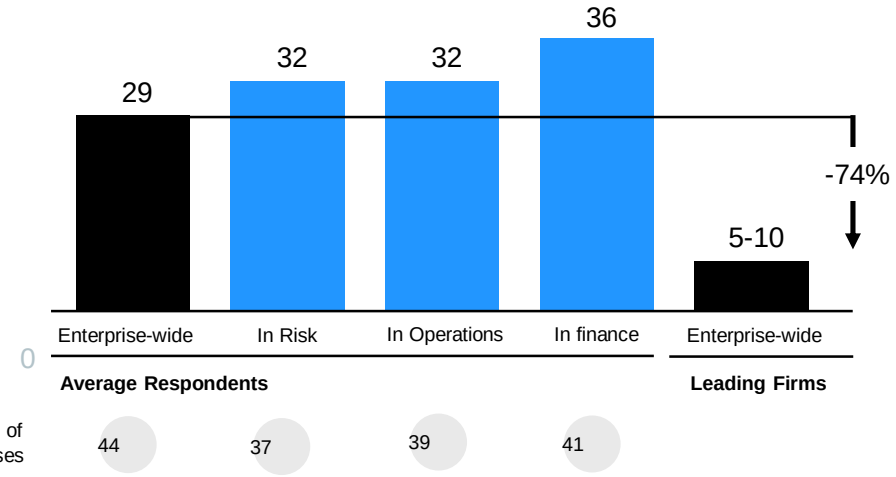
Data Quality

- DQ Targets and monitoring to support business objectives using **Quality Attributes** such as:
  - integrity, accuracy, availability, accessibility, coverage, compliance, timeliness …
- Implies that the level of DQ is directly related to its usage
- Master Data Management

- Organizations lose between $10 to $14 Million USD annually due to poor data - Gartner
- 80% of companies believe they lost revenue due to data challenges - CIO.com
- 40% of all leads have inaccurate data [re: Integrate]
- Employees spend half of their time coping with managing data quality tasks - MIT Sloan

What Impacts will Poor Data Governance Have on your Organization?

McKinsey [2020]: Time Spent on Non-Value Added Tasks Due to Poor DG



Lack of data quality and availability can cause employees to spend a significant amount of time on non-value-added tasks

Time spent on non-value-added tasks due to poor data quality and availability estimated % of total employee time

While it's challenging to directly attribute value to data governance, there are multiple examples of its significant indirect value. Leading firms have eliminated millions of dollars in cost from their data ecosystems and enabled digital and analytics use cases worth millions or even billions of dollars. **Data governance is one of the top three differences between firms that capture this value and firms that don't!**

Information Catalog / Glossary.

5 reasons to have an enterprise data catalog

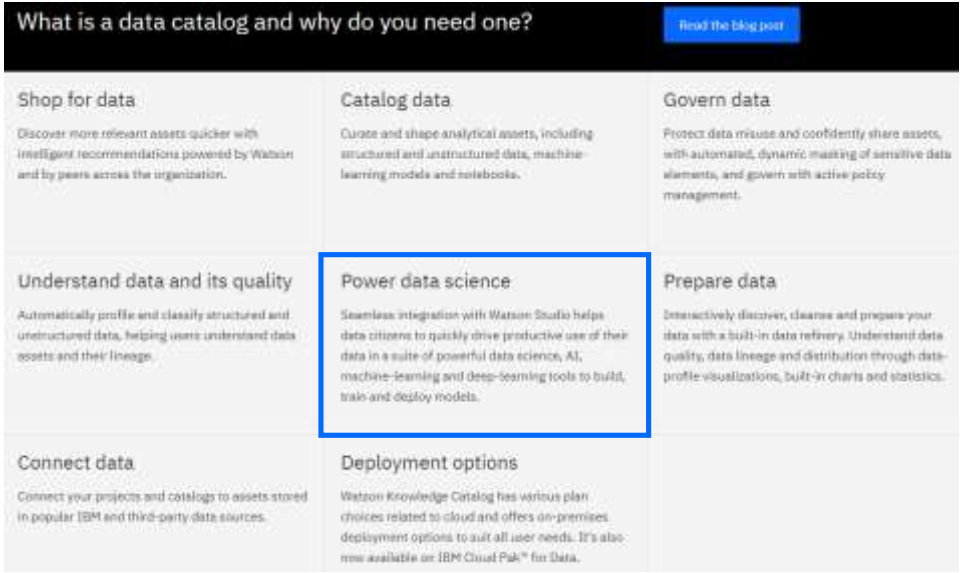
1. Speed and self-service. Rather than submitting requests to an IT group for data that will meet analysts' business needs, Analysts simply search through a data catalog themselves

2. Comprehensive search and access to relevant data.
An analyst will not know if they're missing relevant data or the most up to date asset unless they can search across all available data assets

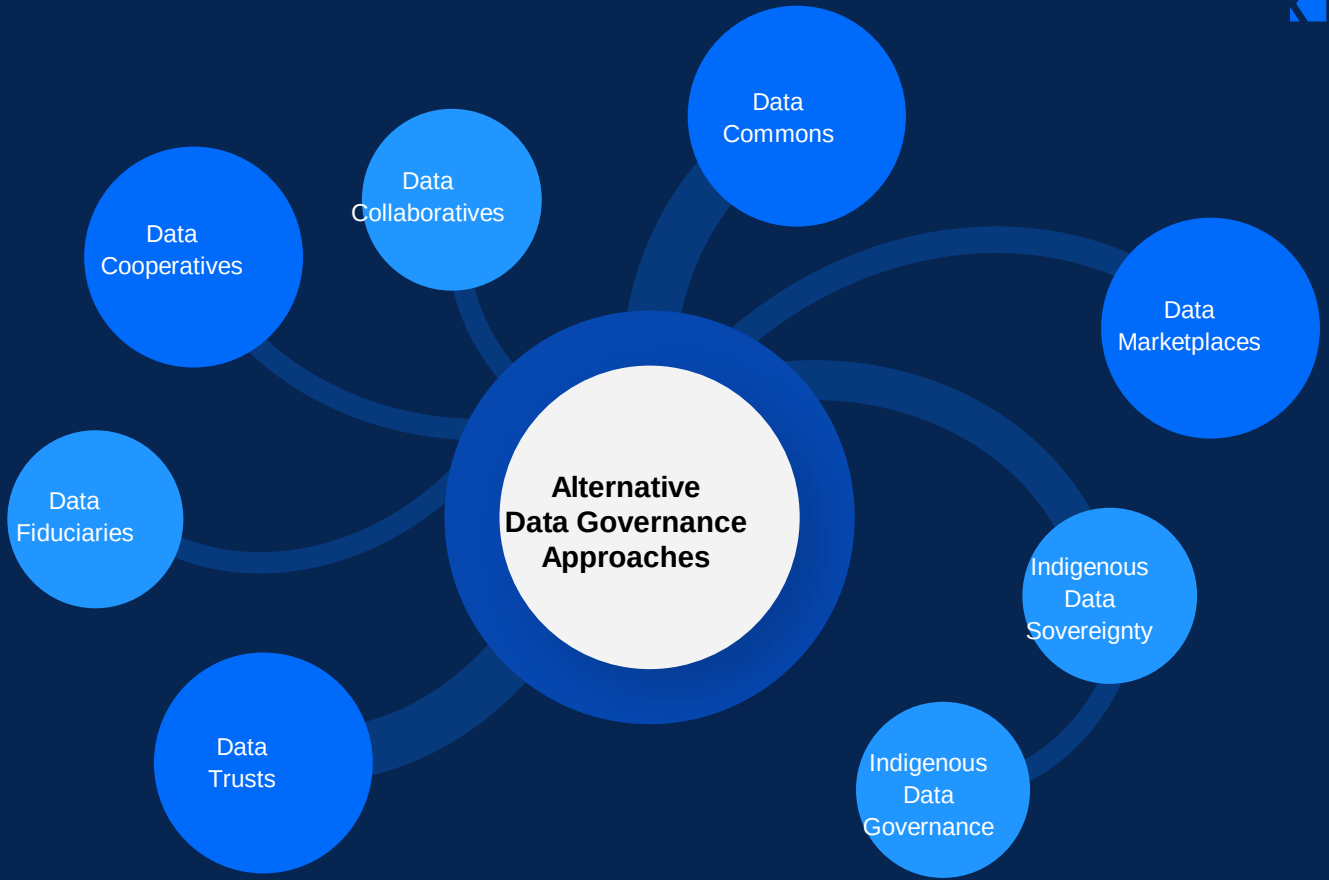
3. Meaningful context. When an analyst finds a data asset that would be useful to them, they can read a description, view business metadata and business term definitions, and read comments provided by others about the data

4. Improves trust and confidence in data. By previewing the data and profiling it, an analyst can very quickly see if certain fields have null or incorrect values and make comments

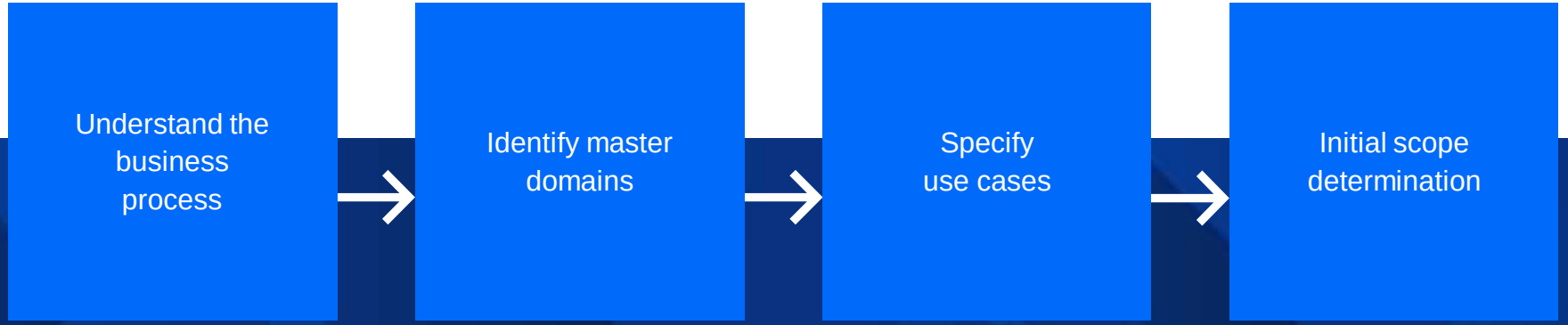
5. Protects data while staying compliant. Instead of having IT masking each column, data rules automatically run based on automatic classification of data



Global Data Governance.



Master Data Management (MDM)



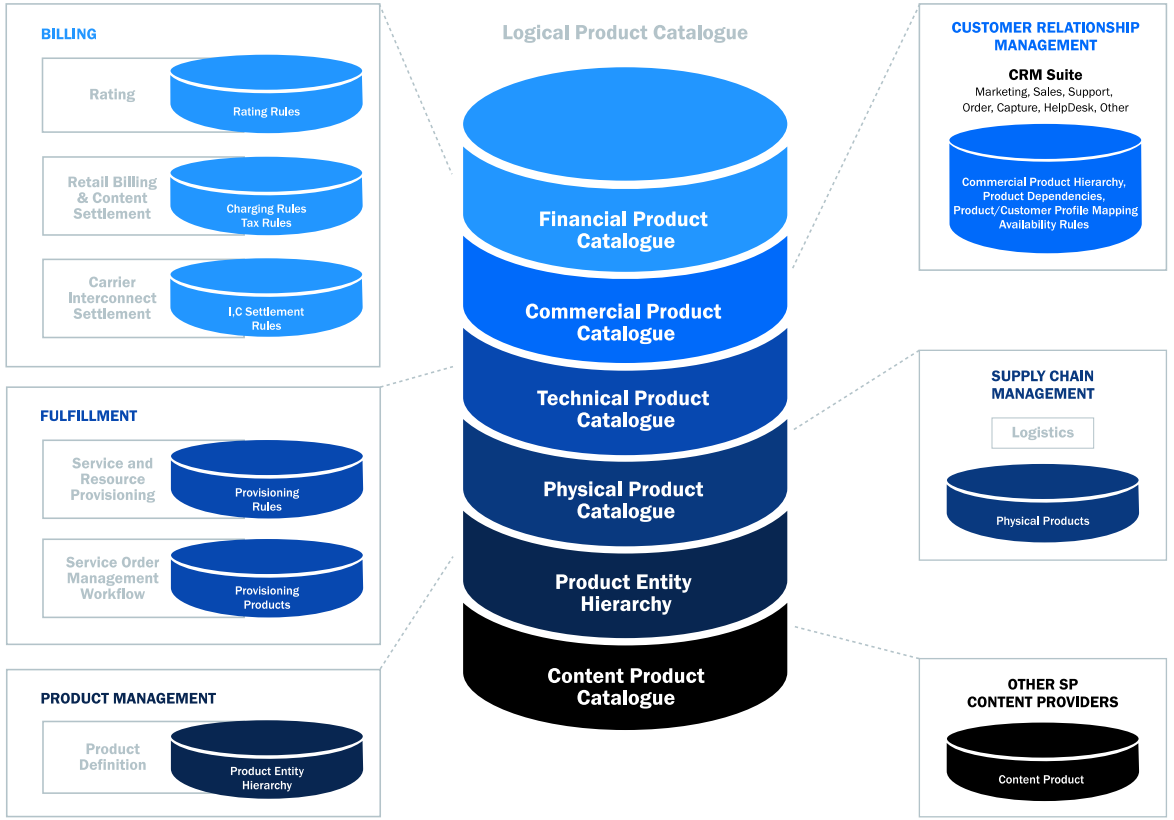
- Target level of accuracy, coverage, accessibility, availability, contextual richness, timely, atomicity, determined with stakeholders, for core data assets (customers, suppliers, products, etc.) with required processes and technologies
- Identification of organization's specific core data assets to master with contributing data sources
- Strategy of unique golden records or multiple perspectives given usage for each core data assets
- COO / CDO champion

PRODUCT MANAGEMENT
KEY DATA

Example of the need to master the Product core data asset in a Telecom.

Product Catalogue view for a Telecom company

- Data (information and knowledge) on products and services are extensive
- The diagram is at an application level which prevents to see data concepts such as specifications, characteristics, bundles, customer targets …
- The CEDM behind these applications cover all of this but is confidential, so we substitute the CEDM view with an application view



Why do Organizations Design & Build their New D&A Platforms to do so Little?

How about handling all data (sensors and business transactions) in real-time for all type of analytics (BI, AI and DO)?

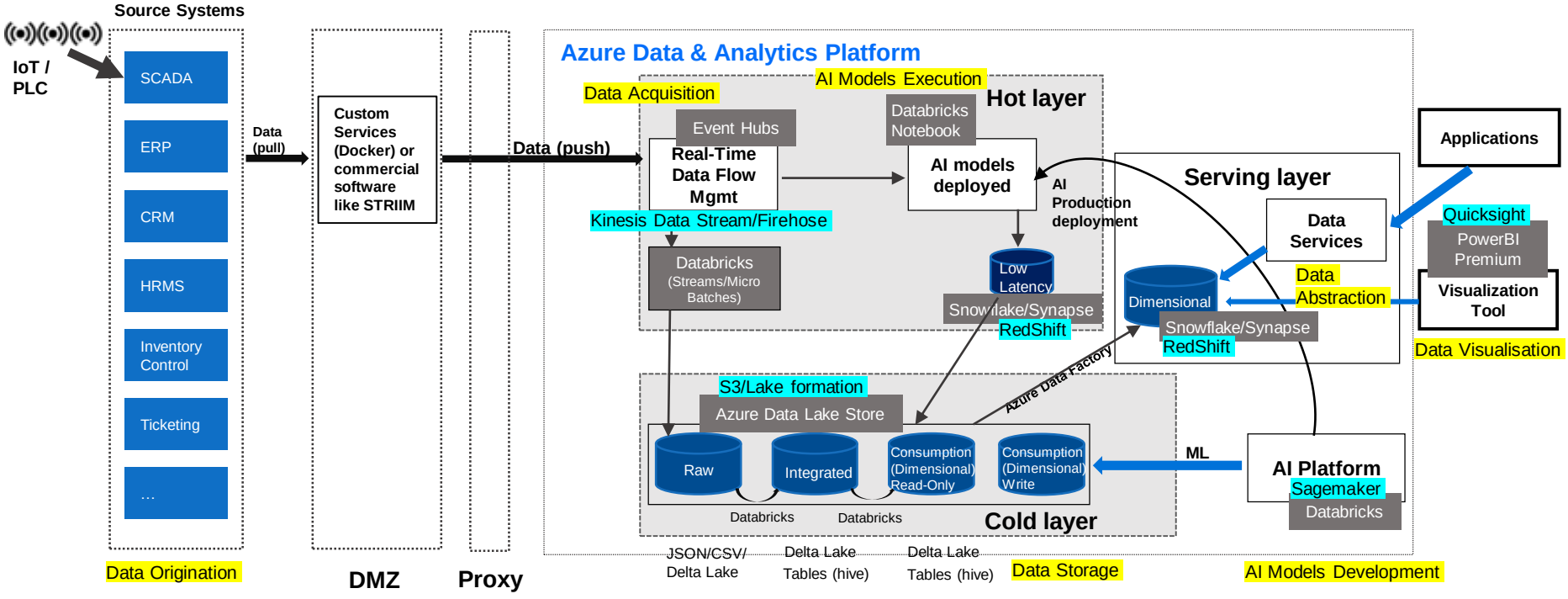
THE MODERN, TRUSTED, BIMODAL, MULTI-USAGES D&A PLATFORM

- **Supports Governance** (data/appl. access rights, privacy enforcements, …)
- **All data (sensors & business transactions) can be pushed** to the D&A platform. This push architecture uncouples the D&A platform from source systems allowing:
  - Increase resilience to source systems changes)
  - Easier integration of companies resulting from M&As
- The data is available right away **to execute an AI model in real-time in the cloud or on edge devices**
- **BI dashboards could be updated, in near real-time,** by having the consumption zone of the Data Lake replicated in the serving layer
- **The entire data lake (all Cold Layer zones) and the serving layer are refreshed in seconds**!
- **This improved Lambda architecture combined with using Delta Lake allow the cold layer to populate 100% of the dashboards**
  - Versus having the dashboards populated from the Hot and Cold layers
  - **Delta Lake also simplifies the management of slowly changing Type 2 dimensions**
- **AI models deployments are simple (no refactoring)**

- **Libraries of algorithms are available to be used for ML/DL/DO** using data persisted on the Data Lake and **processing cycles could be added on demand (elastic)**
- **Data scientists prefer using the consumption zones (less manipulations, less joins) to build AI models** as long as the enhanced dimensional models are built with rich contextual dimensions
- Bimodal considerations **allow data scientists to use data from all zones**
- Persist insights/recommendations made by AI in low latency DB and in Cold Layer **allowing using AI outputs as inputs to other processes & analytics**
- Creation of analytics services to simplify consumption and governance via services (API), metadata / abstraction layer and OLAP layer
- **Monitoring of all analytics assets** for performance, drift, biases and overall health of data pipelines
- Expose reasoning behind decisions taken by models and track degradation and biases of deployed models
- Simplification, for consumers since all analytics are handled from same D&A platform

Multiple Innovation programs exists that will help your organizations to finance a significant part of this!

The KPI Modern and Trusted Data & Analytics Platform on Azure - Could be AWS or others
Governed, Real-time, Bi-Modal, Cloud based (elastic) and support all types of data and analytics

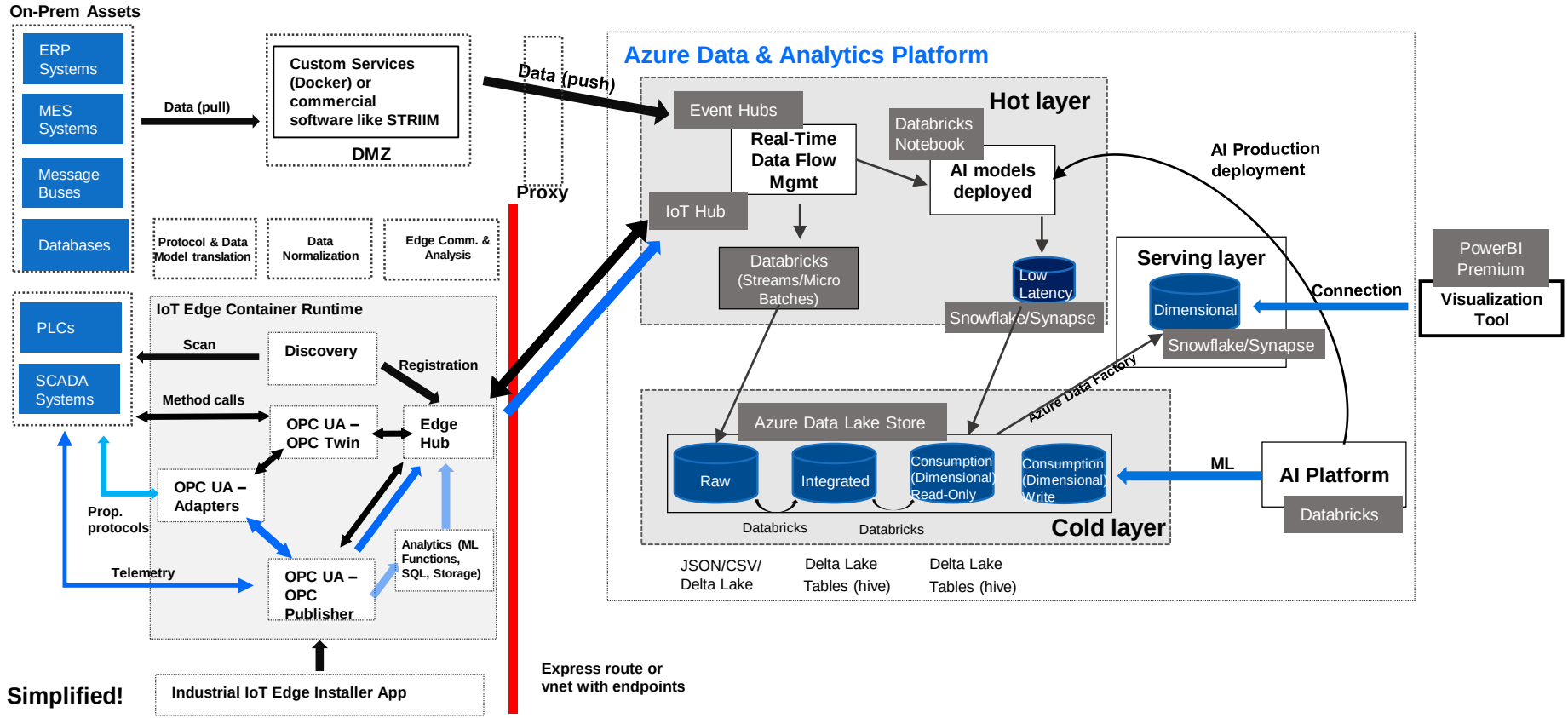


Simplified!

Improved Lambda architecture:
- Fast Cold layer (Streaming + Delta Lake)
- Serving layer mostly fed by cold layer (no synch issues with Hot Layer)
- Hot layer for real-time AI

The KPI Modern and Trusted Data & Analytics Platform on Azure –

Governed, Real-time, Bi-Modal, Cloud based (elastic) and support all types of data and analytics



Il est évident que les organisations veulent innover. Avez-vous la bonne stratégie de données ?

Propulsé par vos objectifs. Motivé par vos défis.

richard.langlois@kpidigital.com