Certification – User Level – Data Governance according to Anjana Data

2021

Challenges to become Data-Driven

Understanding Data Governance

The importance of data sharing

Challenges to become Data Driven

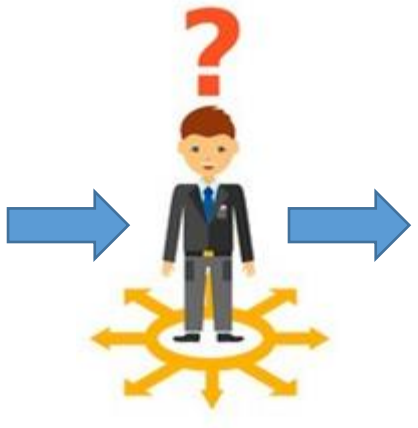
Understanding Data Governance

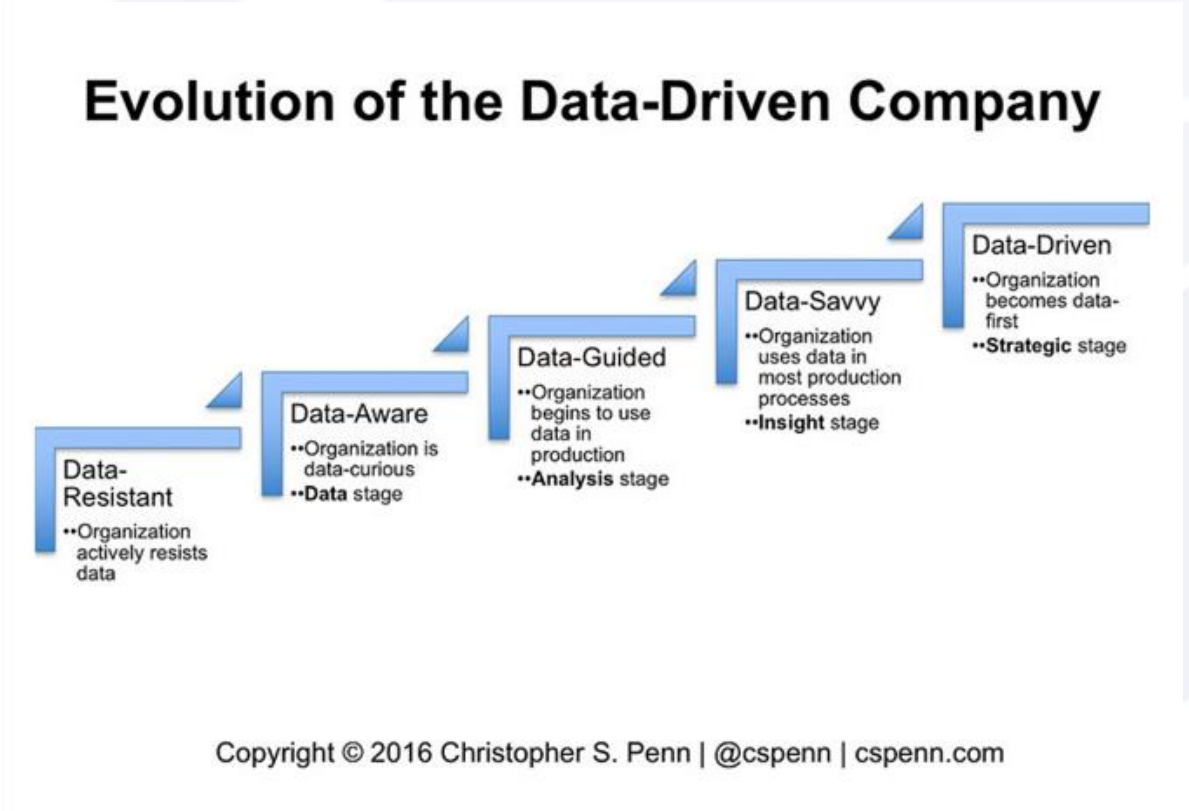
The importance of data sharing

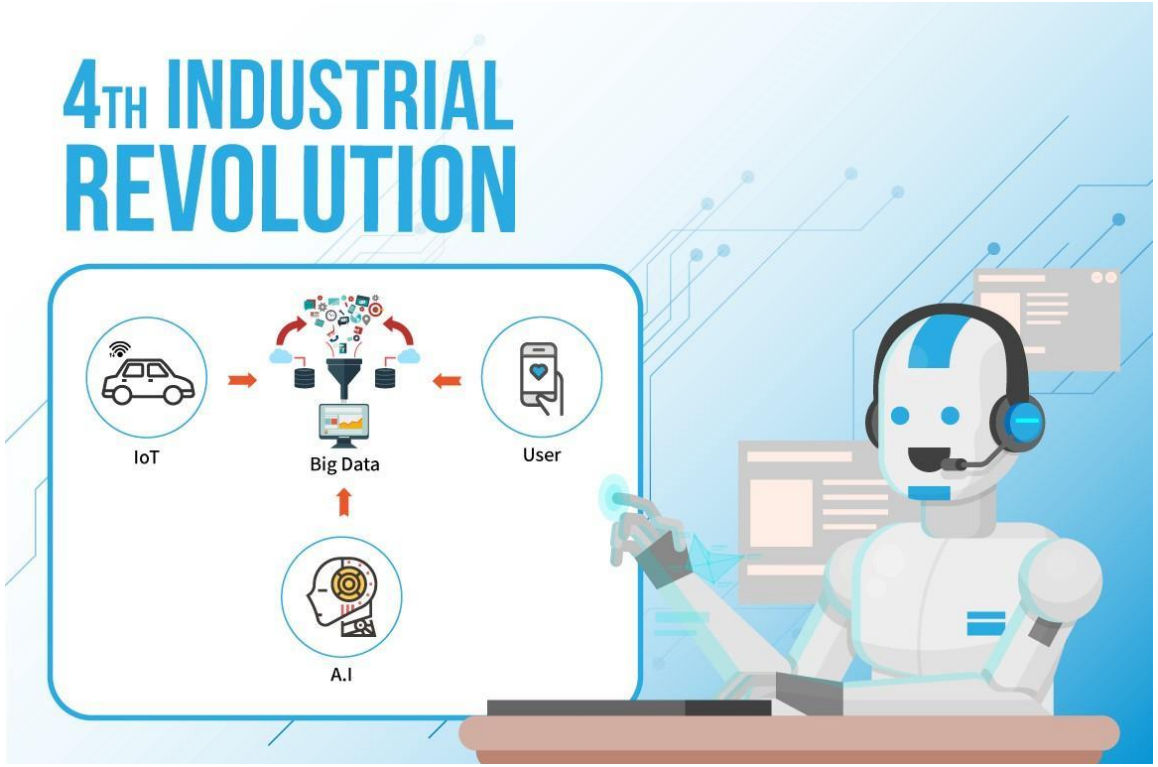
What does it mean to be “Data-Driven”?

A data driven organization is one that understands data as a strategic asset and bases its decisions in data analysis and interpretation

DATA



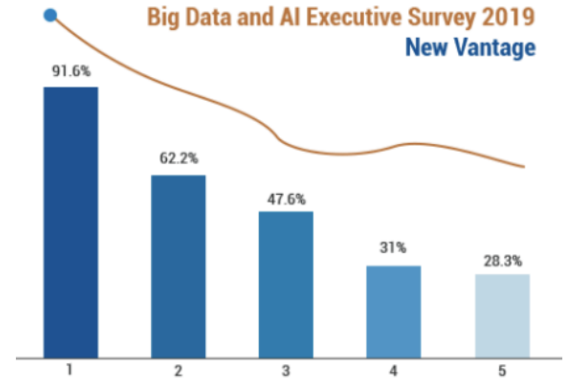




"Data is a new class of economic asset, like currency or gold"

-- World Economic Forum (2012)

1. **91.6%** of the leading organizations are Investing in Big Data and AI.
2. Only **62.2%** are reaching positive measurable results from their investments.
3. **47.6%** say that they compete with their data and analytics offerings.
4. Barely **31%** achieved to create a data-driven company.
5. Only **28.3%** have managed to create a data culture.

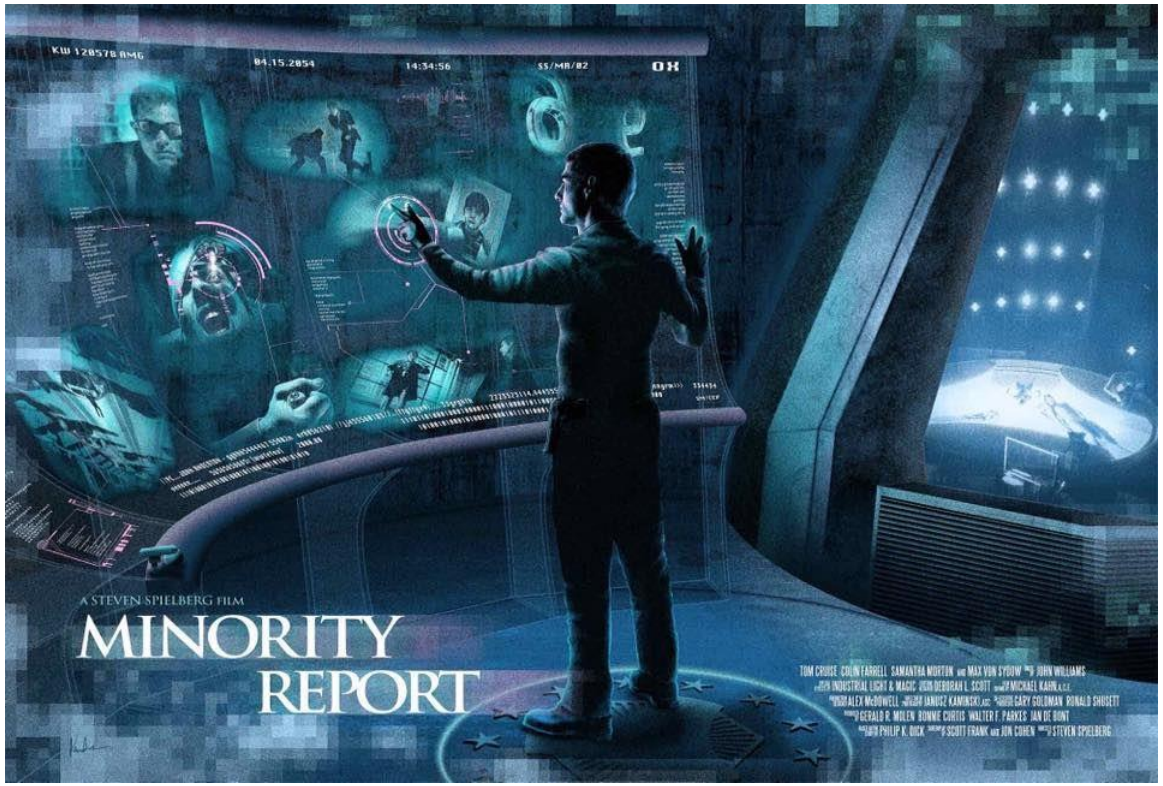


By 2022, 90% of corporate strategies will explicitly consider information as a critical business asset and data analysis as an essential competency"

-- Gartner Research (2020)

What does it look like in organizations?

When you order it on internet…



When it comes home…

"By 2022, less than 5% of data analytics initiatives will correctly identify trusted data and locate reliable data sources."

-- Gartner Research (2020)

4 What's going wrong? (I)

People: 62.5% Processes: 30% Technology: 7.5%

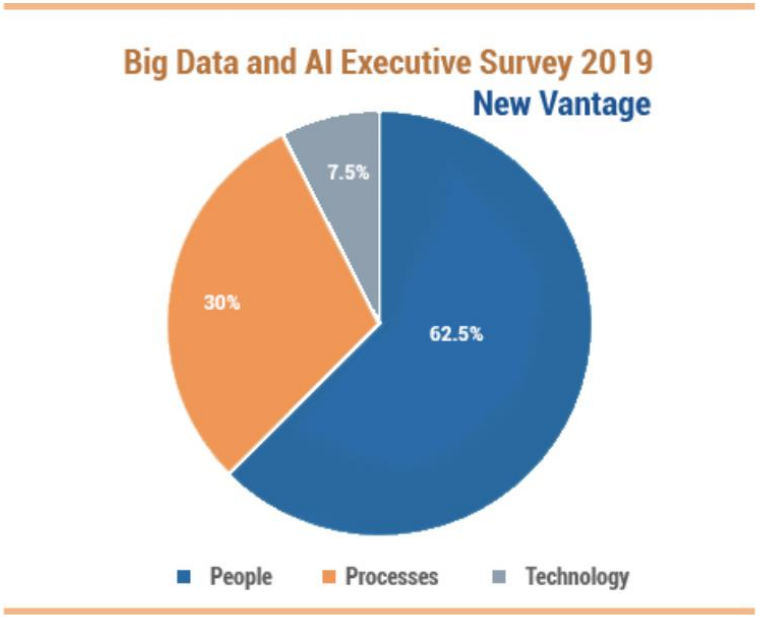


Image 2: Causes of failure in Data Governance initiatives.

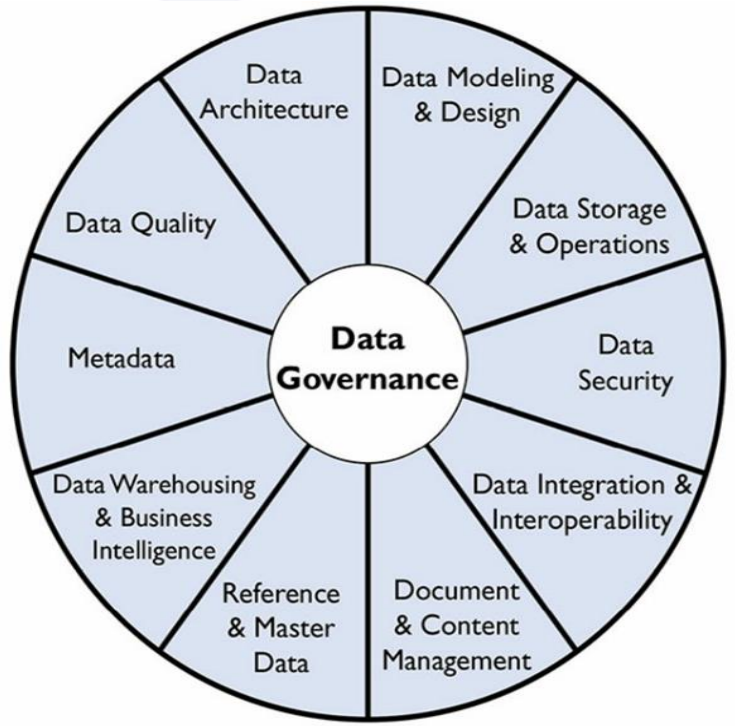


Figure 5 The DAMA-DMBOK2 Data Management Framework (The DAMA Wheel)

"Through 2025, 80% of organizations seeking to scale digital business will fail because they don't take a modern approach to data and analytics governance"

-- Gartner Research (2020)

What's going wrong? (II)

1. The initiative is leaded by IT only from a purely technical point of view
2. Data Governance is a time limited project instead of a long-term strategic initiative
3. There is no earlier knowledge of the data panorama inside the organization
4. Using "big bang" approaches rather than an incremental proposition
5. Thinking that purchasing a tool is the solution
6. The organization tries to afront Data Governance without being prepared for having the necessary maturity
7. The data strategy and data governance are not aligned with the global corporate strategy
8. The defined organization framework does not reach all the necessary levels inside the enterprise
9. The objective defined consists only in reaching minimums to cover certain regulation or compliance
10. Building Great Wall between business and IT likely growing isolation between them
11. Not involving all stakeholders, from developers to business users

Askham, Nicola: "9 biggest mistakes companies make when implementing data governance"

New objectives, new challenges

Turn data into a strategic asset!!

People & Processes

- CDO and Data Office against the world
- Data governance is usually led by IT
- High walls between Business and IT
- Understood as a limited and finite project
- Low level of data culture and literacy
- Business users need to spot value
- Leveraged by regulation compliance
- Data governance means bureaucracy
- Plenty of silos between departments
- Needs and requirements evolve over time

Technology

- Tangled and assorted IT ecosystems
- Outburst of new technologies
- New added-value data architectures
- Interoperability and no vendor lock-in

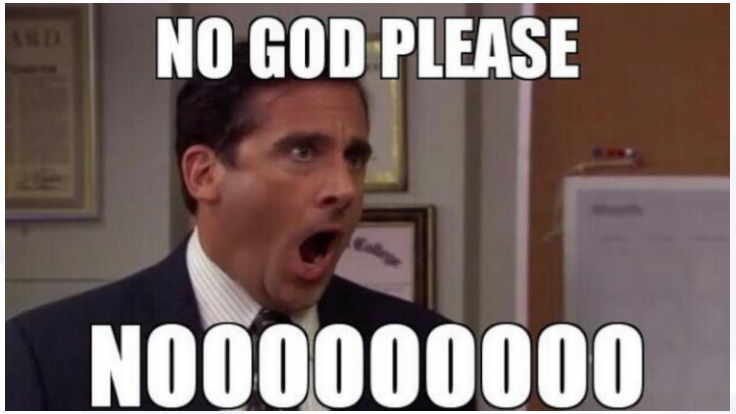
Market solutions

- Legacy data management technologies do not offer high business value
- IT-focused hard to implement solutions
- High initial investment and pricing based on users not on the use

Challenges to become Data Driven

Understanding Data Governance

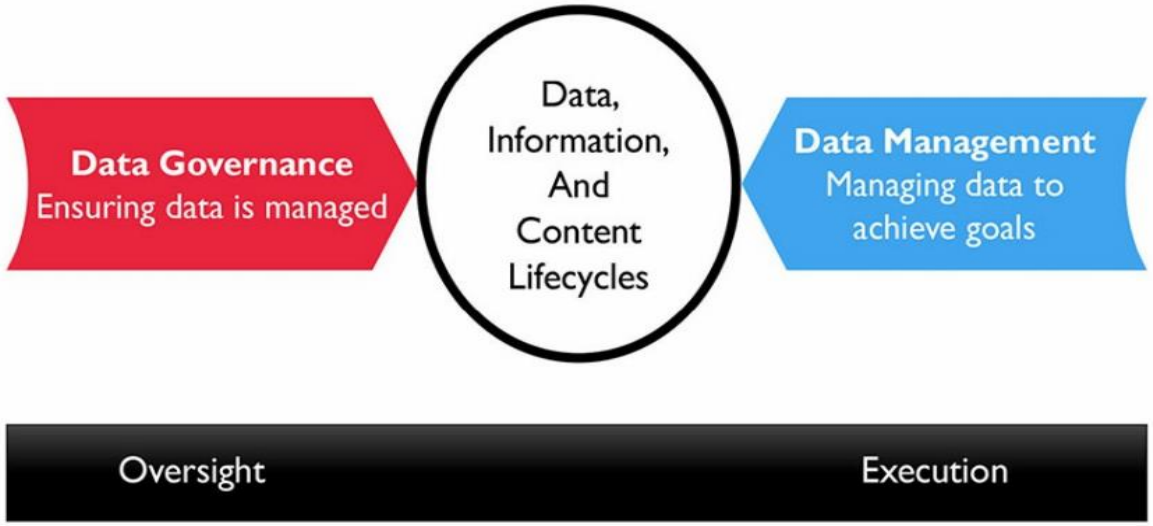
The importance of data sharing



Data Governance is NOT…

- **A transient trend**
- **A one-time thing nor a finite project**
- **An exclusive responsibility of IT**
- **Just about complying with regulations**
- **Just about identifying Data Owners and Data Stewards or creating Data Committees**
- **The same as building a new Data Lake nor a centralized repository of data**
- **About bureaucreacy nor unuseful formalisms which will slow and complicate data processes**
- **A new problem for data consumers (developers, engineers, analysts, scientists, business users)**
- **The same as Data Quality, Security, Privacy, Storage, Integration nor Data Management**

Data Governance vs Data Management

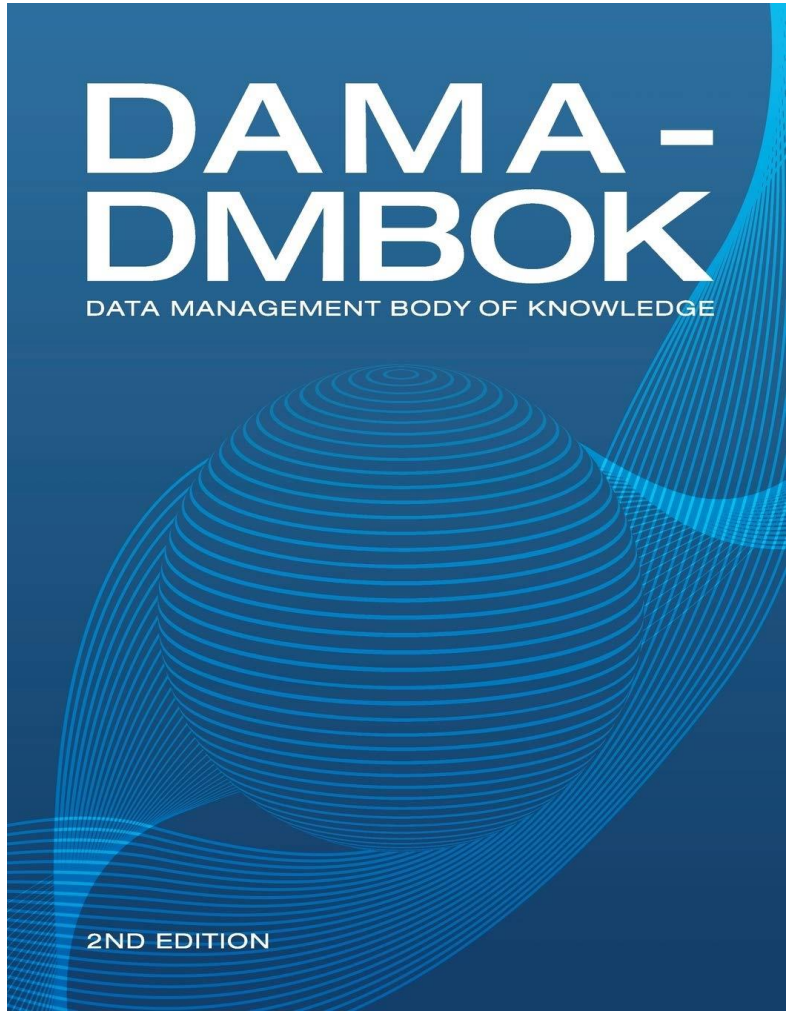


- Data Governance function **guides** the rest of Data Management functions.
- The purpose of Data Governance is to assure that the data are properly managed, according to the **policies and best practices defined**.
- While the overall objective of data management is to ensure that an organization obtains value from its data, the Data Governance focuses on how **decisions** about data are made and how **people** and **processes** are expected to behave in relation to data.

Example

Just as an auditor controls financial processes but doesn't execute financial management, data governance ensures that data is properly managed without directly executing data management activities

Data Governance according to DAMA-DMBOK



- Data Governance is defined as the **exercise of authority and control** (planning, monitoring and implementation) about the management of data assets.

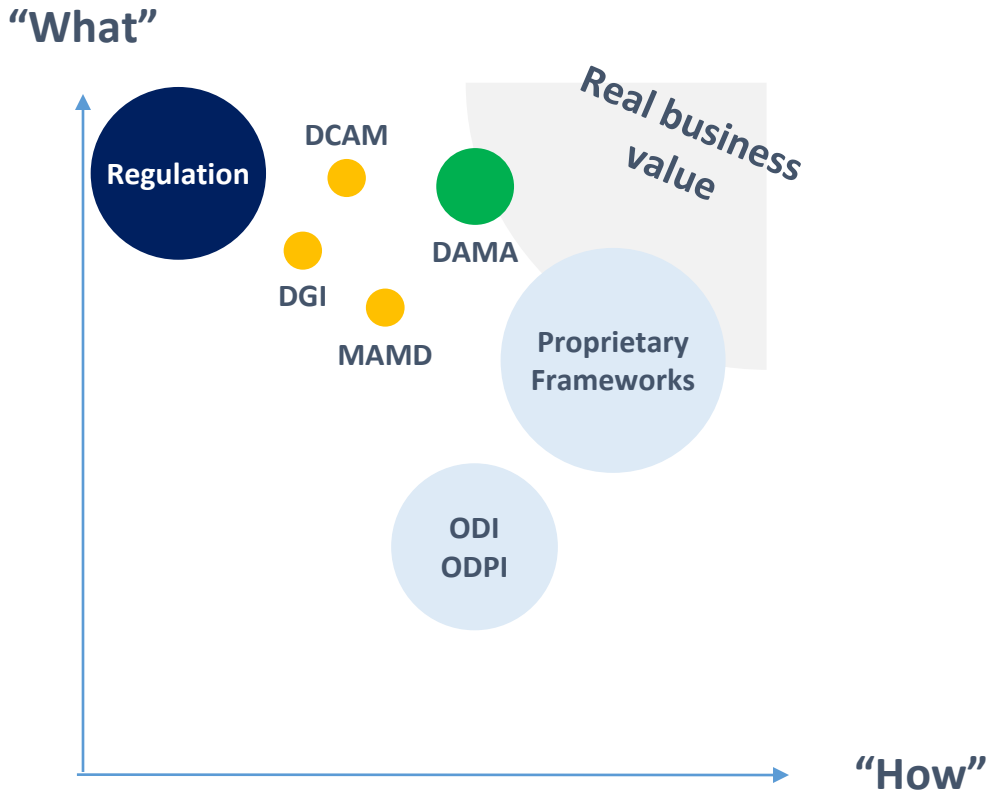
Most of the programs include...

- Strategy
- Policies
- Standards and quality
- Oversight
- Compliance
- Incidents management
- Data Management projects
- Valuation of data assets

- Data Governance requires **organizational change management** support as well as **C-level management** sponsorship.

REQUIRES A CULTURAL CHANGE

Data Governance. From “What” to “How” (I)



Many of us know the **"What"**
But what’s really complicated is the **"How**

- Organization size
- Activity sector
- Digitization level
- Culture and philosophy
- Organization and geography

Data Governance. From “What” to “How” (II)

“WHAT”

- Strategy, framework elements and governance model
- Pillars and bases for the operativization of data governance
- Data culture, communication and change management
- Training, talent and technology as enabler

“HOW”

- Clear and defined objectives in the short, mid and long term
- Business-led, aligned with Technology
- Operational, effective and efficient governance model
- Incremental and iterative approach based on use cases

Who is in charge of setting this up?

Board Management

Security

Data Science

Regulation

Chief Data Officer

Data Office

Business

Technology

Departments

"By 2021, in 75% of large enterprises, the office of the CDO will be seen as a mission-critical function comparable to IT, business operations, HR and finance"

-- Gartner Research (2020)

6 Who else intervenes and how?

Consumer's perspective

data analysts
data engineers
data scientists

Oversight perspective

C-level
Data Office
Audit
Legal
Compliance

Producer's perspective

data owners
data stewards
data engineers

- Approach with support for different types of operating model
- Roles with functions and responsibilities clearly defined
- Hierarchical Organizational Units according to information domains

Senior Management
Support and promote Data Culture

Owners & Stewards
Data controllers and providers

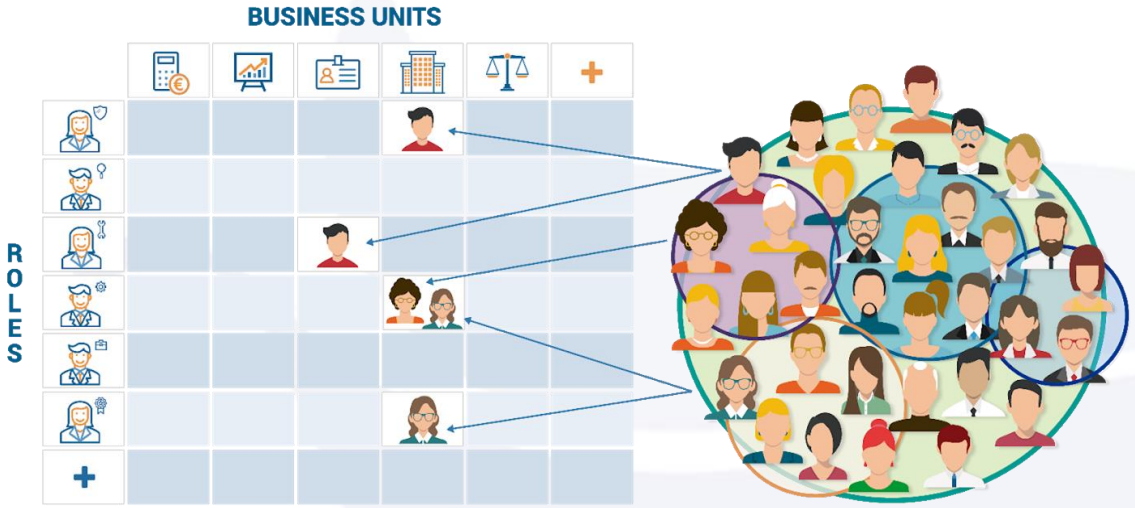
CDO & Architecture
Data strategy and governance model

Systems
Manage and evolve technical infrastructure

Legal & Audit
Use Audit and Data processig

Developers
Metadata use and process development

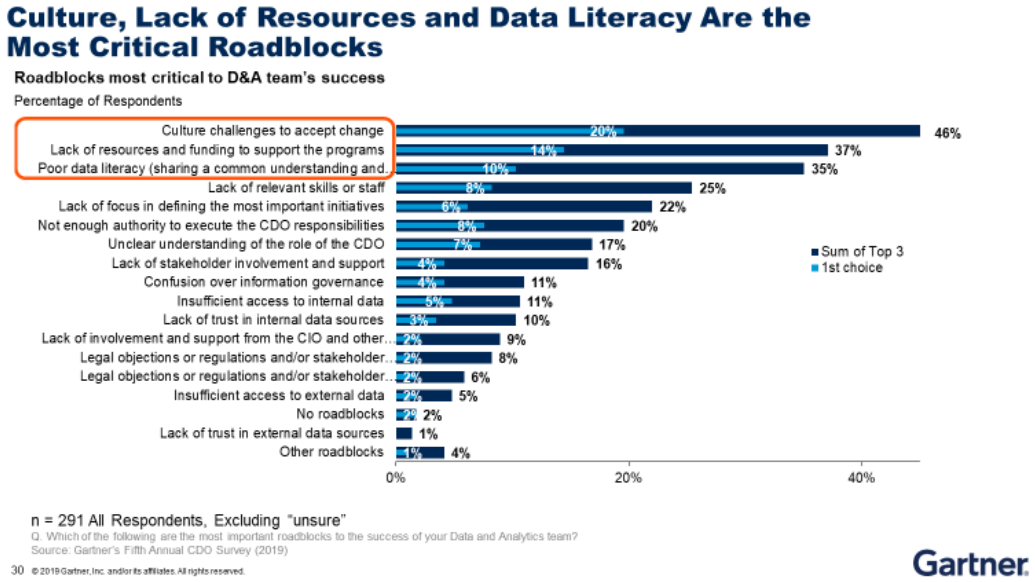
Business Users
Fast and agile access to data

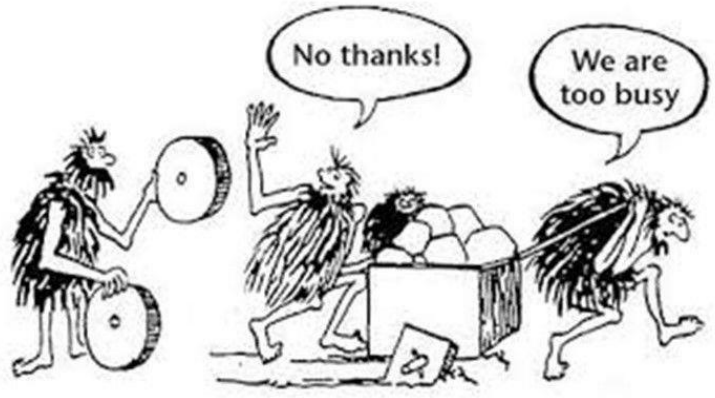


Operating Model

- When defining the operating model for Data Governance, the following aspects of the organization must be taken into account:
  - Value of data
  - Business model
  - Cultural factors
  - Impact of regulation
- The policies define the game rules according to the governance principles and the procedures establish how the different roles must behave according to those policies.
- Governance processes need to be automated.

Data Culture & Change Management





Planning

Training and Coaching

Influencing systems development

Policy implementation

Communication Plan

Promote the value of data assets with announcements, workshops, circulars, ...

Capture continuous feedback about Data Governance activities

Implementing **Data Governance capabilities and skills** through training and coaching

Measure the effects of change management with surveys, metrics and KPIs

Incorporate incentives in relation to data management best practices

Metadata-centric approach

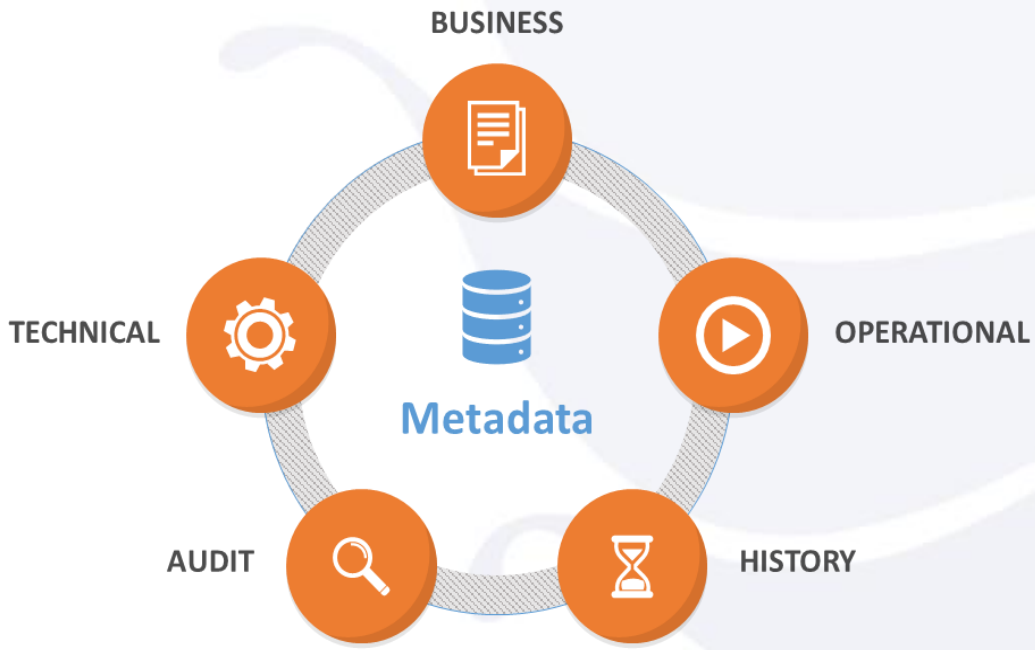
> *"An organization can't be Data-Driven without being Metadata-Driven"*
>
> **DAMA-DMBOK**

DECLARATIVE METADATA

Metadata declared by end-users, passively extracted from technical platforms as built-in configuration or inferred from analytical algorithms

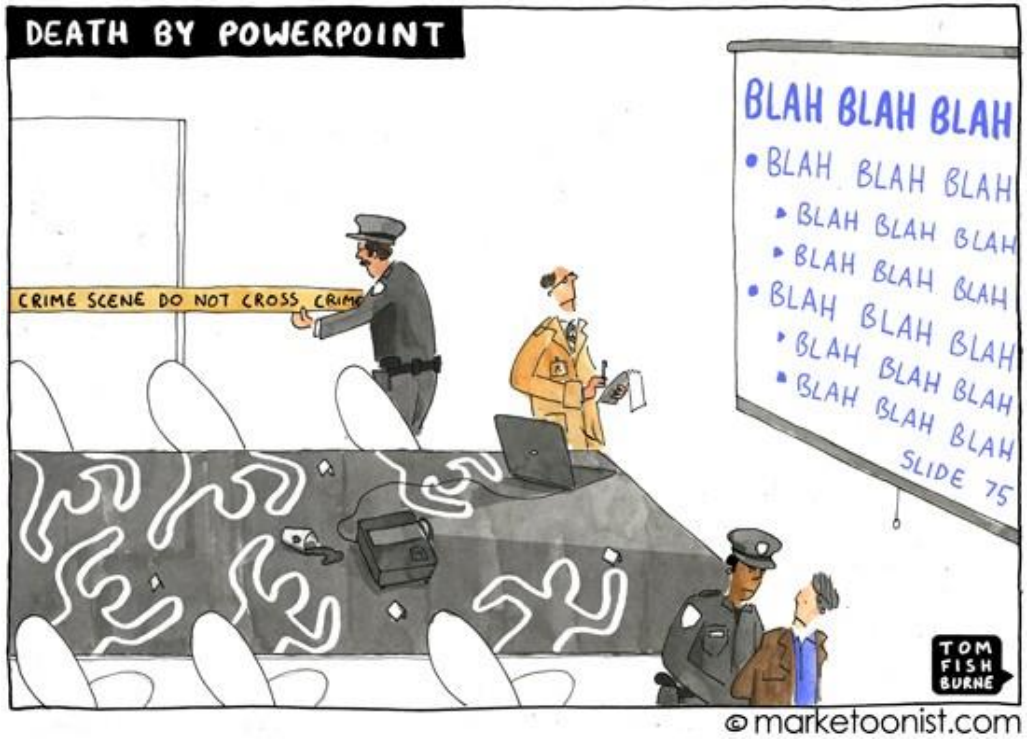
EXECUTION METADATA

Metadata actively obtained after the processes execution over the technical platforms or generated from the users activity and interaction



Metadata-centric approach allows **technical abstraction** in data management and the **decoupling** between data governance and the technologies for data capture, storage, processing and exploitation thanks to a **centralized metadata management**

Operativization. Proactive and preventive approach



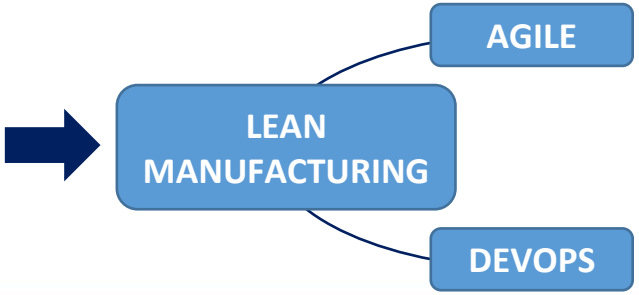
Data Governance PROACTIVE & PREVENTIVE

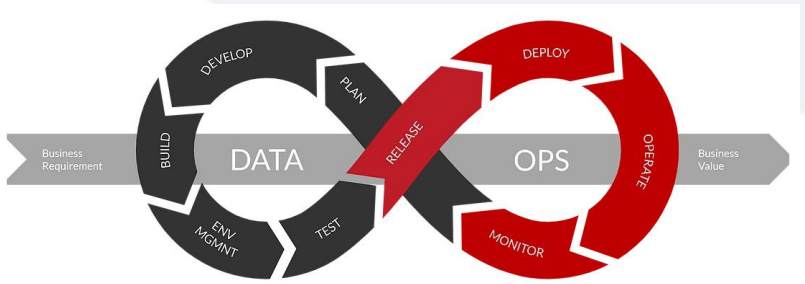
- Putting Data Governance at the forefront of **the data value chain**
- **Collaborative** approach empowering different stakeholders

- Integration with **demand management**
- **Integration** with other technologies and pieces
- **Automation** of common technical processes
- Incremental and iterative approach by **use cases**
- **Democratisation** and governed self-service
- **Monitoring** for continuous improvement

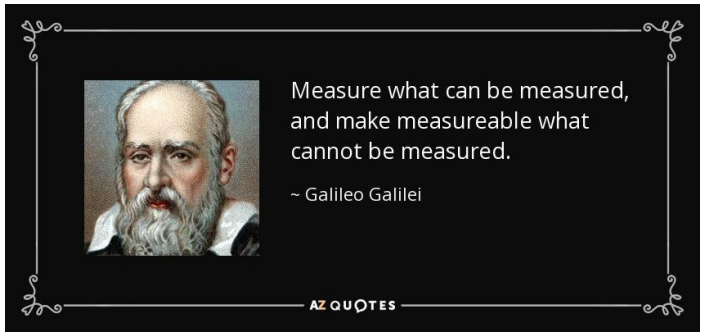
- **Metadata repository** at the heart of the data ecosystem
- Abstracting **data management** from technologies and platforms

*"86% of the organizations that launch specific initiatives of Data Governance **don't manage to operate it** in the day to day"*

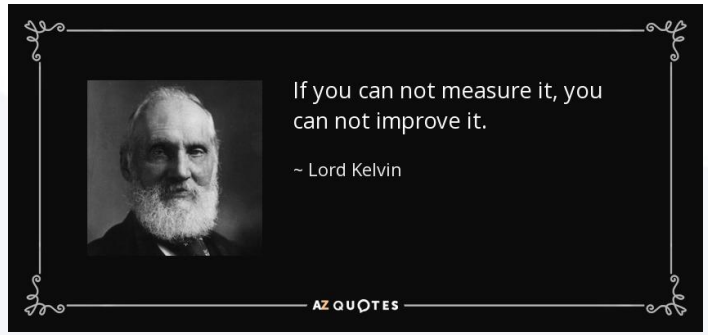




Progress and success have to be measured



"42% of the organizations that launch specific initiatives for Data Governance ***DO NOT analyze, measure or monitor it"***



Tracking metrics

Value

- Contribution to company's goals
- Risk reduction
- Improved efficiency of operations

Effectiveness

- Achievement of goals and objectives
- Use of tools
- Effectiveness of communication
- Effectiveness of training
- Speed in the adoption of changes

Sustainability

- Performance of policies and processes: Are the defined policies and procedures working properly?
- Compliance with rules and procedures: Is the staff following the guidelines and changing their behaviour as it is needed?

Success metrics

- Re-use of data and maximization of synergies between initiatives for data usage, processing and exploitation.
- Reduction of Time-to-market and Time-to-value in data analytics projects and/or data-driven products and services
- Satisfaction of data producers and consumers
- Costs and time reduction and increase of efficiency and productivity in:
  - Infrastructure management, maintenance and systems licensing
  - Development of projects for data capture, storage, handling, processing, exploitation and usage.
  - Operational and manual tasks for data management (preparation, standardization, cleansing, cataloguing, delivery, ...)
  - Compliance with regulatory and normative requirements
- Risk mitigation: operational risk, errors in production, reputational risk, regulatory non-compliance, poor data quality, ...

Challenges to become Data Driven

Understanding Data Governance

The importance of data sharing

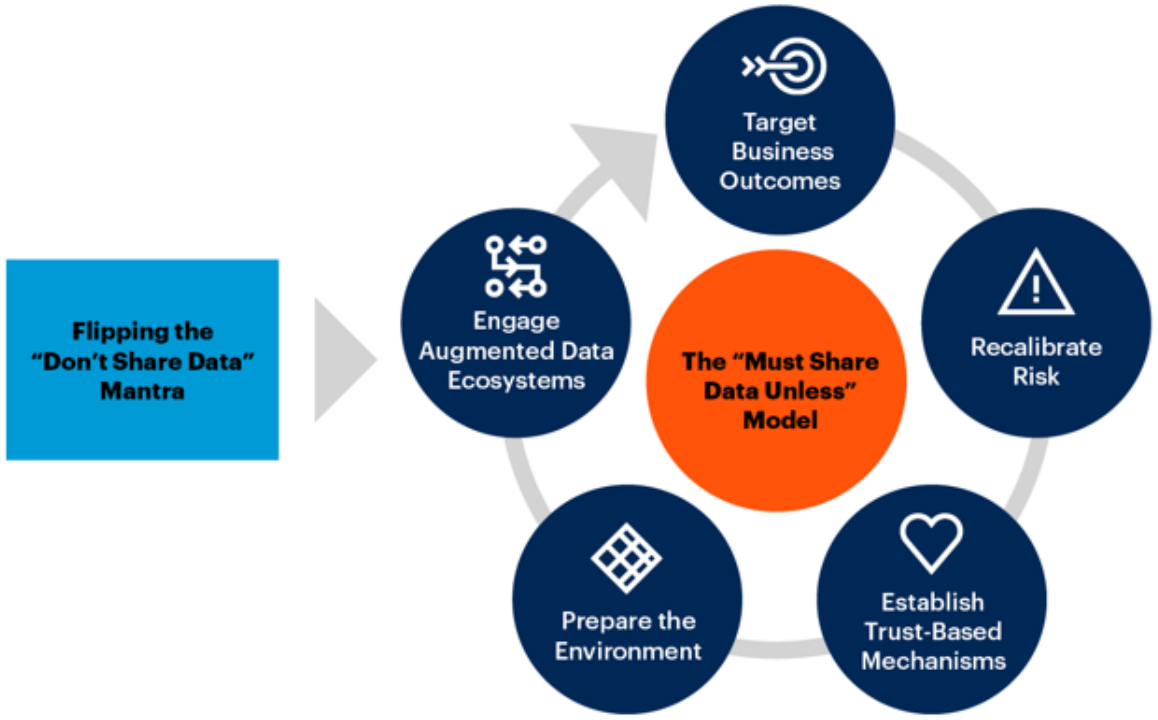
Data sharing is vital to beingData-Driven

Gartner says…

- Data teams are **twice as likely** to generate measurable benefits from external data sharing and **3 times** more from internal sharing of data assets with partners and collaborators
- By 2022, **90% of corporate strategies** will explicitly consider information as a critical business asset and its analysis as a core competency
- By 2023, organizations that promote data **sharing will outperform their competitors** on the majority of business value metrics
- Data sharing is a **business necessity** to accelerate digital business (December 2020)
- Organizations that list data sharing as a business necessity rather than a data management function will perform better in Digital Business and **be more successful** than their competitors (February 2021)

Compartir datos tiene múltiples beneficios

The Gartner Data Sharing Model
To Accelerate Digital Business



Source: Gartner
727589_C

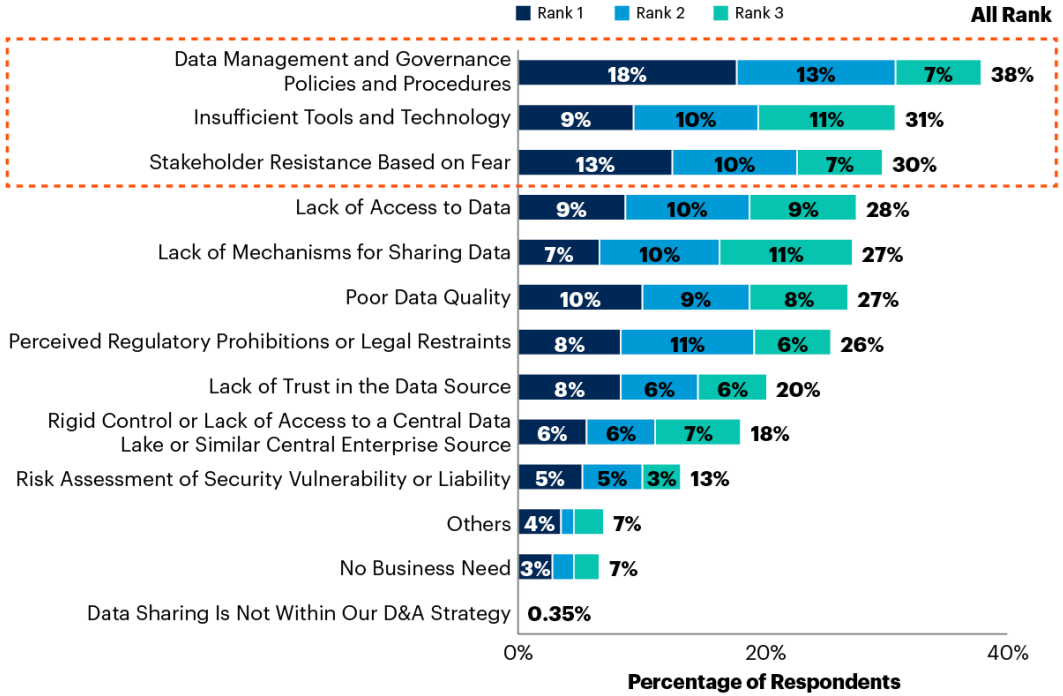
Gartner

Benefits

- Maximize **synergies** between different stakeholders to **improve** their day-to-day operation
- Improve the **quality** of critical data for better and faster **decision** making
- **Reuse** developments and works already done by increasing the **efficiency** of data projects
- **Saving** on operating costs and improving the **productivity** of professionals
- Improving **time-to-market** and **time-to-value** of new products and services
- **Help** data scientists develop better models of advanced analytics
- Bring **transparency** and **trust**

Data sharing isn't so simple

Roadblocks to Data Sharing Internally



n = 285 respondents who are sharing data internally (Q12), excluding "unsure"

Q: What do you see as the top challenges to sharing data internally at your organization? Please rank up to 3 challenges in order of importance.

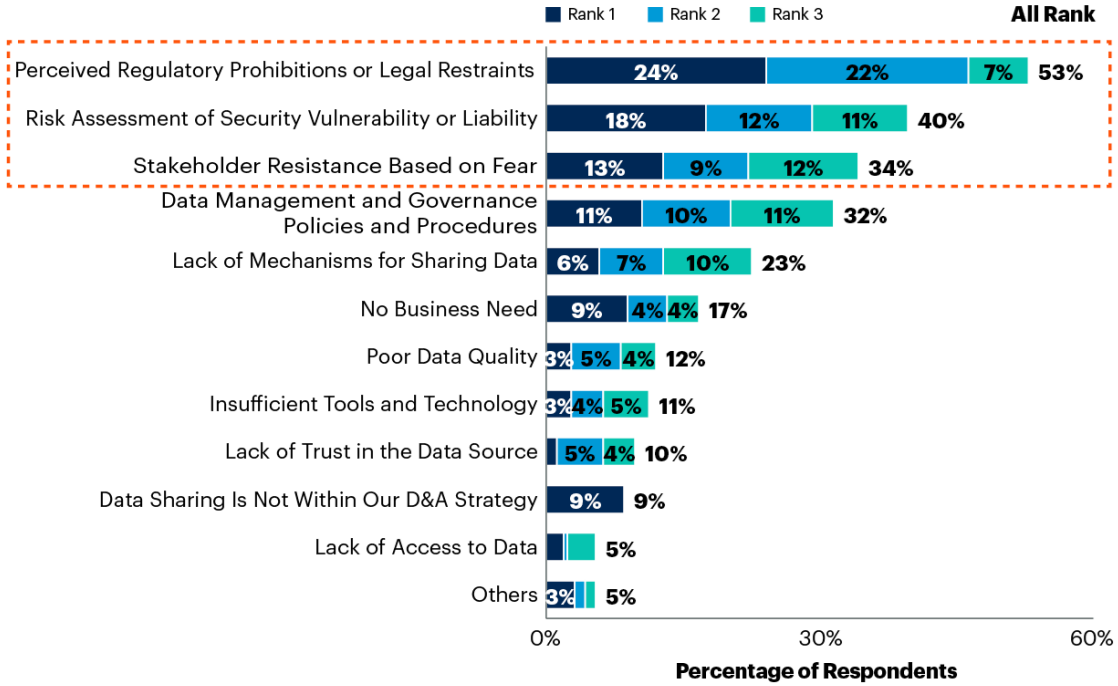
Source: Gartner's Fifth Annual CDO Survey (2019)

Note: Values less than 3% not shown.

727589_C

Gartner

Roadblocks to Data Sharing Externally



n = 257 respondents who are sharing or exchanging data externally (Q12), excluding "unsure"

Q: What do you see as the top challenges to sharing or exchanging externally?

Source: Gartner's Fifth Annual CDO Survey (2019)

Note: Values less than 3% not shown.

727589_C

Gartner

The “Data sharing” dilemma

You must enhance the use and reuse of data throughout the organization and for that you have to democratize its access

You must have strict control over the use of data to ensure regulatory and normative compliance

Aurea mediocritas

"In the middle, there is virtue"

-- Aristóteles

Business & Regulation

Business & Technology

Self-service & Control

Traditional & Innovative

Un activo estratégico, diferentes visiones

Producer's perspective

data owners
data stewards
data engineers

- ✓ Control over the data production processes (data quality & availability)
- ✓ Data preparation and certification for data sharing and consumption by others
- ✓ Knowledge about which owned data is being used by who and for what purpose
- ✓ Full audit to support normative and regulatory compliance

Oversight perspective

C-level
Data Office
Audit
Legal
Compliance

- ✓ ROI maximization for data initiatives
- ✓ Productivity and efficiency boosting
- ✓ Processes and tasks automation for costs reduction and operational risk mitigation
- ✓ Unified and homogeneus vision of the company's data usage and consumption
- ✓ Normative and regulatory compliance
- ✓ Possibility to value and monetize data

Consumer's perspective

data analysts
data engineers
data scientists

- ✓ Gain deeper knowledge about available data along with its context and meaning
- ✓ Better decisions driven by the access to better-quality and well-managed data
- ✓ Agility in data consumption and exploitation leading to self-service
- ✓ Control over the data production processes (data quality & availability)

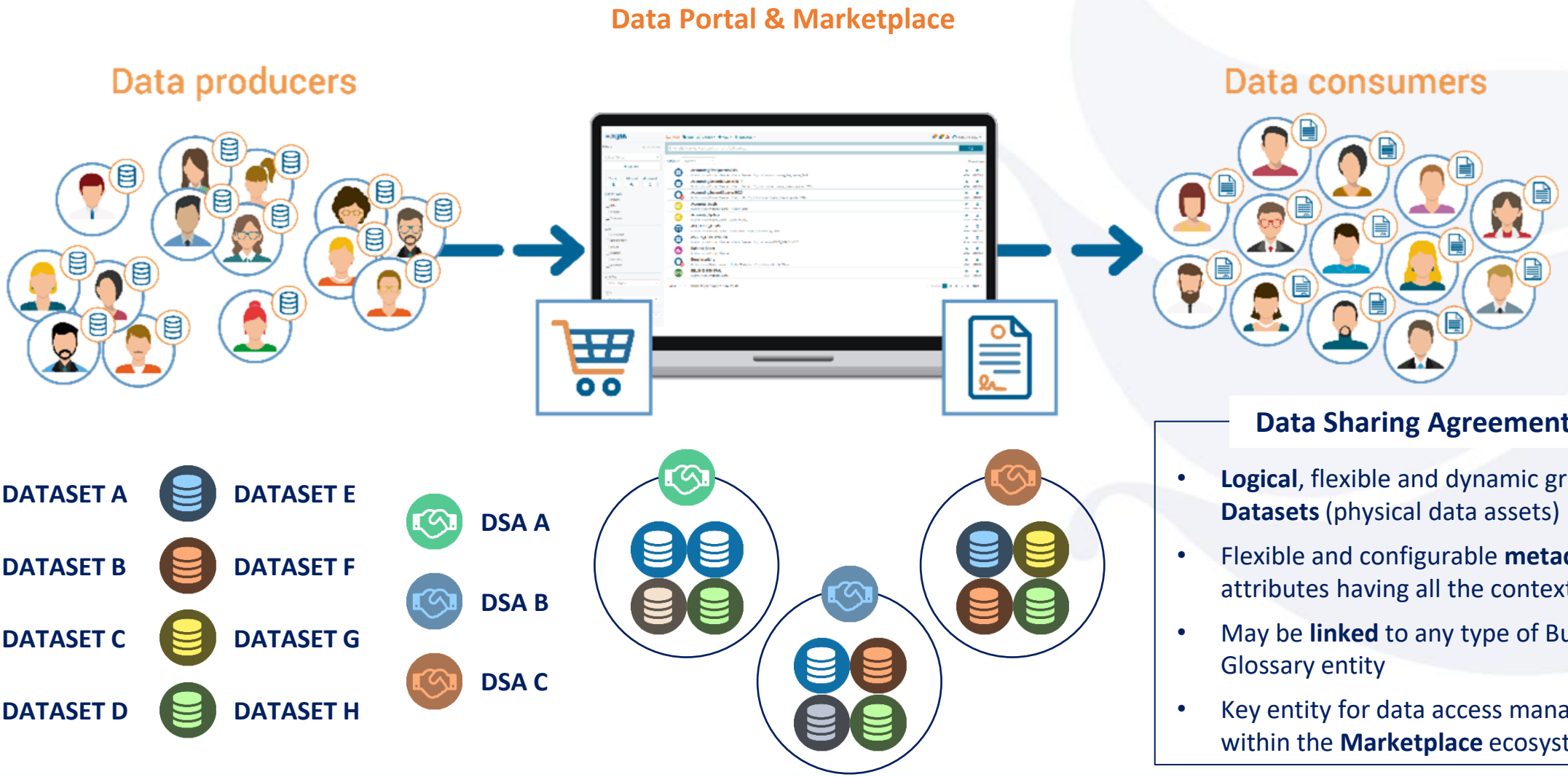
Wha is a Data Sharing Agreement

What exactly is a Data Sharing Agreement?

The key element to bridge the gap between business and IT worlds for data sharing and consumption management

- ✓ Mechanism to ease **regulatory and normative compliance** for data usage
- ✓ Promote the standardization of **data and information access management**
- ✓ Offer stakeholders **a new experience for data sharing**
- ✓ **Bring the data closer to the business users** abstracting them from the IT perspective
- ✓ Allow the definition, implementation and operativization of **Data Contracts**
- ✓ Provide the necessary flexibility for the enforcement of **data access policies**

Operationalizing DSAs (I). Availability of DSAs



Data Sharing Agreements

- **Logical**, flexible and dynamic groups of **Datasets** (physical data assets)
- Flexible and configurable **metadata** attributes having all the context
- May be **linked** to any type of Business Glossary entity
- Key entity for data access management within the **Marketplace** ecosystem

Operationalizing DSAs (II). DSAs Governance

Policies, Procedures & Data Contracts

Contract intermediary

Data owner of the DSA

Data providers

Data owners from included datasets

Data consumers

Business users adhered to the DSA

Governing Bodies

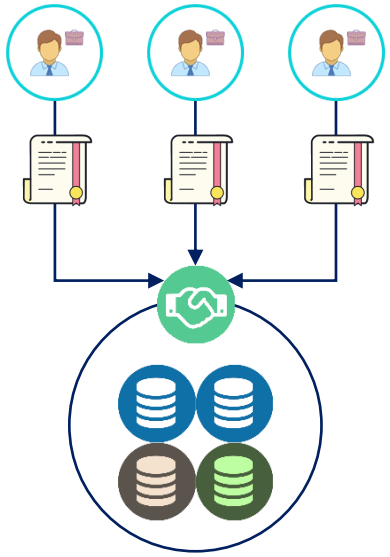
Cross-domain areas controlling data sharing and consumption

Establish a common and standard model for the **management of data access** through the implementation of **multi-stakeholder Data Contracts** between producers and consumers

Like any **data asset,** they must have their own **policies** and **procedures** within the data governance model defined and according to the **data domain** to which they belong

They can involve several **stakeholders**, define different **categories**, identify various levels of data **sensitivity** and **confidentiality**, support **versioning** and have **expiration** date

Operationalizing DSAs (III). Integration with data platforms



Physical Dataset

- Stored in multiple **repositories**
- Supported by different **technologies**
- Available in various **formats**

- Unambiguous identification of data **producers and consumers**
- **Data access policies** for specific use cases (encoding, encryption, masking, anonymization, filtering, ...)
- **Requirements for producers** for the correct consumption of the information contained in the Datasets (Quality, Availability, SLAs, ...)
- **Licensing terms and conditions of use for consumers** (associated costs, transferred rights and liabilities, additional clauses, ...)

Integration with Identity Management Systems and Technologies

Integration with Security and Privacy Systems and Technologies

Integration with Data Platforms and Data Services for Consumption

The future of DSAs

- ✓ DSAs as a key element in the implementation of **DataOps** to ease and promote the governed self-service of information
- ✓ DSAs as a **Best Practice** to ease data sharing by ensuring regulatory compliance in complex environments
- ✓ **Smarter** DSAs thanks to the incorporation of new Artificial Intelligence and Machine Learning algorithms
- ✓ DSAs **based** on semantics, taxonomies and use cases