Blog

# DATA GOVERNANCE



May 10, 2023

2023

# What exactly is Data Governance?

Data Governance is establishing standards and policies for how efficiently information can be used by various teams within an organization. Data Governance includes roles, processes, and metrics to be defined and implemented in order to optimize data operations, comply with the data regulations, and achieve organizational goals.

## Why is Data Governance important?

Data Governance is vital in today's modern enterprise architecture, but why is Data Governance so complex for enterprises? In today's digital economy, enterprises are introducing data initiatives to support and drive data modernization, and decision-making by improving efficiency through automation. These initiatives reside under the umbrella of an enterprise, and therefore a central governance program is a core to the success of these initiatives.



# The Top 3 Reasons Why Data Governance is Extremely Important:

#### Ever-Increasing Data Volume

Enterprises architecture is spreading wider and therefore the amount of data produced has increased rapidly. Complimenting this, similar data growth has been witnessed with the applications residing outside the enterprise.

\*Study suggests that the volume of the data is doubling every two years.

#### 🖌 Multi-Dimensional Data Needs

Modern enterprises are data-driven and increasingly introducing data science initiatives to support decision-making. More effective decision-making leads to higher levels of performance. However, the decision outcomes of data science initiatives are not always highly effective due to uncertainty about the quality of the data. To improve the effectiveness of the decision-making process, an enterprise needs multi-dimensional data from all possible sources.

#### Federated Teams

 $(\checkmark)$ 

The enterprise architecture is ever-evolving with the advent of the modern age. To conquer Business Agility and DevOps, enterprises are adopting federated architecture, therefore allowing teams to interoperate efficiently and allowing business imperatives/functions to get the highest precedence over. Federated teams in the enterprise architecture lead to the establishment of a common body/structure to orchestrate and govern the overall data movement.



## Ever-Increasing Data Volume

For an organization, to implement Data Governance effectively, the following key foundation blocks are essential:

#### 1. Strategic Planning –

Ensures that the data policies are defined at the enterprise level, promotes data compliance across the enterprise, and has a governance program established within the enterprise.

#### 2. Multi-Dimensional Data Needs -

The most critical block of Data Governance is to ensure how the enterprise metadata is created, captured, cataloged, and maintained. The Data Governance Program also creates the Data Governance Process, and it is implemented during this stage.

#### 3. Data Architecture -

The Data Architecture block ensures that the Enterprise Data Model is created with compliant Data Governance processes. This stage also assesses and creates a uniform process for application integration and interfaces.



#### 4. Data Quality -

Data Quality is an important block for Data Governance Frameworks from a business assurance perspective. It starts with an assessment of the state of the Data Quality and matures towards the uniform implementation of Data Quality rules that ensures validity, consistency, completeness, and accuracy.

#### 5. Data Security -

The main purpose of Data Security is to protect the data and stay compliant with industry and government regulations. The block encompasses the techniques and technologies that drive the protection of digital information from any unauthorized access, corruption, modification, or disclosure.

#### 6. Data Stewardship -

A comprehensive approach to Data Management to ensure the quality, integrity, accessibility, and security of the data. Data Stewards (primarily data managers and administrators) are responsible for implementing the data governance policies and standards and maintaining data quality and security as mentioned in the previous points.





# **Data Governance Framework**



While the building block described in the above section are in place for an organization, the following framework depicts how Data Governance can be implemented for both data states –

#### 

#### **Data at Rest**

Primarily data residing in a data warehouse, data lake, databases, and tables.

#### $\bigcirc$

#### Data in Motion:

During data movement (data pipelines), accessing the data within the application integration layer/interfacing stage.



## **Benefits of Data Governance:**

Albeit Data Governance isn't optional for an organization, it brings many more benefits to an organization.

#### Compliance with Data Regulations -

Albeit Data Governance isn't optional for an organization, it brings many more benefits to an organization.

#### Easier, Faster, and Secured Data Access –

With the implementation of Data Governance, it brings the easiest and fastest access to the data for an organization. App owners within the organization/outside organizations have a standard way of operations with a non-optional security advantage.

#### 🧭 Improved quality of Data

It is vital that Data Cleaning and Data Quality rules are implemented while implementing Data Governance. With the help of improved quality of data (org-wide standards) it ensures that data accuracy, uniformity, completeness, and consistency to the data consumers.



# **Data Governance Accelerators**

Hit the ground running with PurpleCube, the Modern Data Management Platform. The tool comes with pre-built Data Cleansing rules, Data Quality, and Data Security rules that help enterprises accelerate Data Governance implementations.





# **Business Objective**

A European Union's leading international bank, namely in the Americas and in the Asia-Pacific region.

#### They had 76 Source Systems that were:

- Non-availability of attribute relationships
- Diverse data sources leading to inconsistent views and challenges in data cataloging
- Lack of standard data access policies leading to inconsistent access management
- Challenges in detecting data quality on which business rules had to applied
- Diverse and non-standard toolset leading to lack of uniformity in data governance

## **Solution**

- Create "Data Sampling" for identifying the "Data Quality"
- Import "Metadata" from relevant sources & build data lineage
- Create "Business Glossary" for all the assets
- Create the "Data Stewardship" process to ensure the Assets are approved or rejected by the "Analysts"

### **Key Successes**

- Data owners were successfully able to establish data lineage
- Users were able to detect data quality rules and issues by looking into the sample data provided.
- A self-service portal was created for data stewards instead of using email communication





1390 Market Street, Suite 200, San Francisco, California 94102, US

www.purplecube.ai



© COPYRIGHT 2023 PurpleCube, Inc. ALL RIGHTS RESERVED