The Justin Company

Business Framework for the Governance and Management of Enterprise Data

The Justin Company

July 2017

Table of Contents

Introduction

Research suggests that 90% of the world's data was generated in the past 2 years. This is indicative of the world's growing reliance on data. The Big Data phenomenon coupled with savvy Business Intelligence tools have enabled businesses including The Justin Company to track Key Performance Indicators instantaneously. Like any valuable asset, special care must be taken with data. If not governed properly, data can impede growth, mislead decision makers and present security concerns; hence the need for a governance framework.

A Data Governance Framework is a set of controls enforced by any organisation:

- To ensure integrity of data
- To permit optimal transformation of data from source to business use
- To avoid chaos
- To safeguard against malicious intent.

Gwen Thomas from the Data Governance Institute described it as:

"a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods."

LEVELS OF DATA CLASSIFICATION

RESTRICTED – Level 1

Employee social security, Customer Credit Card numbers, etc.

CONFIDENTIAL/SENSITIVE – Level 2

Personal data from private clients, contracts with suppliers, etc.

INTERNAL ONLY – Level 3

Details of a sales campaign, organization charts, etc.

PUBLIC – Level 4

Generally in the form of a publication (annual report, synopsis of a recently concluded project).

In order to create an effective framework, one must be cognisant of the principles of data in a data system. An important principle is the classification of the data on hand. Above are a few examples from a 4 tier classification system. Data classified under public is usually available to any interested party; both internal and external. The number of viewers decrease as we move up the levels to restricted; only designated personnel are able to access this data (in some cases a non-disclosure agreement may be requested).

In the data system diagram below, the classification determines the final storage location of the data in the Enterprise Data Warehouse (EDW) to permit access controls.

Big Data experts commonly describe a data system using the 3 Vs of data. These are: the **volume** of data being collected and stored, the **variety** of data formats and the **velocity** at which data is generated and processed.

Further, a key principle is how to measure the quality of data. Later in chapter 2 we'll discuss the seven characteristics of good data and expound on the other underlying principles previously mentioned.

The data governance framework will encapsulate these principles and thus understanding them will aid in gaining an appreciation for the framework.

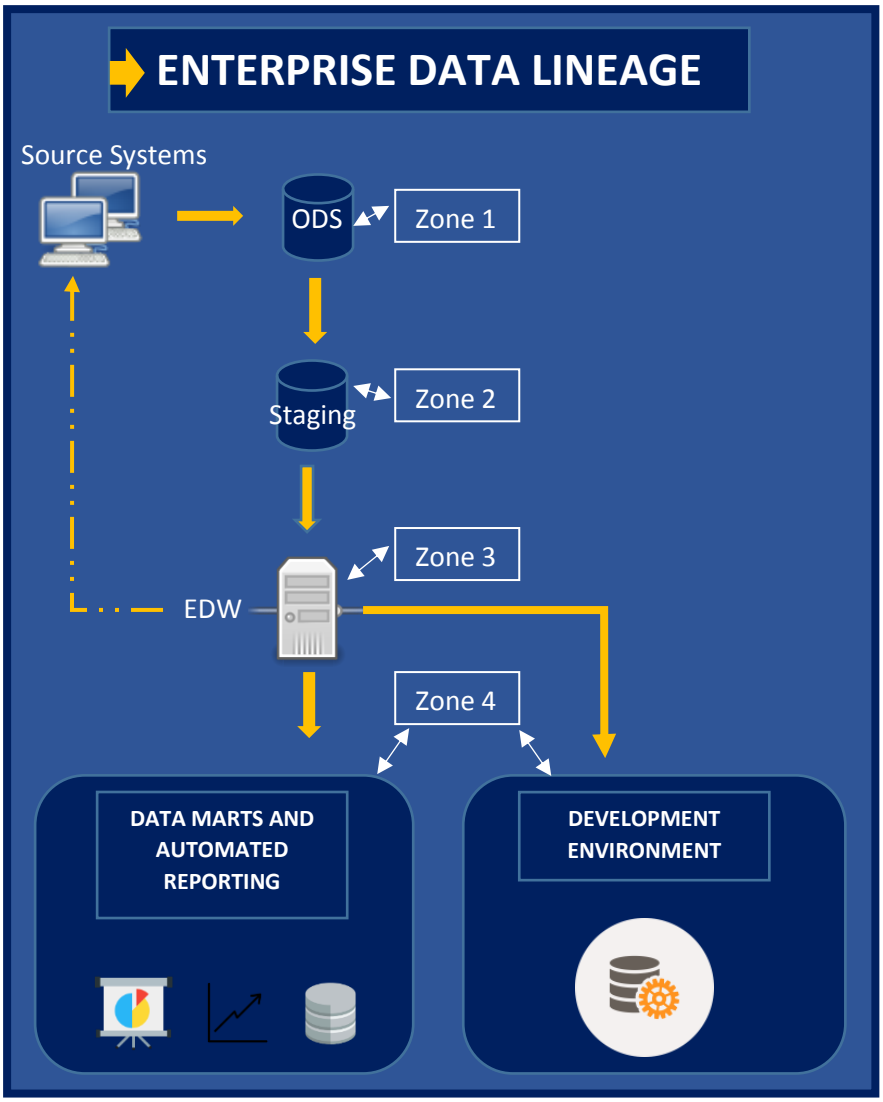

Figure 1: Enterprise Data Lineage Diagram

Chapter 3 of the document takes a critical tour through the existing data management controls and documents the vulnerabilities. We'll identify areas of:

- Non-standard data definitions
- Manual adjustments and data manipulations
- No clearly defined data ownership and accountability
- Lack of user understanding

The data system in *figure 1 (overleaf)* begins by collecting data from the source systems' Operational Data Store (ODS) and enriching that raw data as it moves through the zones to eventual usage. The framework is introduced in chapter 4 and advises on policies as data flows through the data system such as zonal-entry validations (e.g. checksum), access control (e.g. data masking). The framework establishes a foundation for the successful management of data assets. It defines data management roles and responsibilities, establishes policies and standards, and defines processes and procedures in order to better manage the enterprise data and build confidence in the accuracy and reliability of the data.

IT governance has a pivotal role to play given the impact that the tools could have on the data. Therefore we'll also outline the standardizing of software applications.

In many cases regulators and investors are the driving force behind data governance. Regulations like FATCA, IFRS9 and BASEL II have put financial institutions under increased pressure to prove data integrity and deliver quickly on risk reporting. Consequently, frameworks are often seen as defensive mechanisms. For this reason, we'll incorporate a data strategy which is a more offensive mechanism as it is used to harness data to attain the enterprise mission (e.g. considerations for cloud computing). In this manner, the framework seeks to create optimal value by maintaining a balance between realising benefits and mitigating risk.

Finally, in chapters 5 and 6 we'll discuss the steps that will be taken for a smooth transition and the mechanisms to continually make the process more transparent and reduce operational friction.