How Clear is Your Data? Practical Advice for Proper Levels of Abstraction

by JP Romero and Gaurav Arora

With most data governance initiatives, companies focus on making sure their data assets are findable and trustable. But they frequently neglect an equally important aspect of their data: its clearness.

In any enterprise, technical and business people can speak entirely different languages. So why do we assume that a single data depiction will be sufficient for the whole organization?

The most successful data governance programs span the organization, and always involve a plethora of users. So ideally, companies should bring visibility and enforce governance across all kinds of people[1]. This means describing data in digestible metadata blocks that can be easily understood by different user groups. The goal is to make everyone's lives easier by avoiding unnecessary details and providing easily findable, meaningful answers and insights.

Levels of abstraction are important and exist in many different disciplines. People use diagrams with boxes and arrows to convey different messages, whether it's for a business process or a database schema. Companies should favor data governance models that support this.

Here's an example. Say a company creates a *Business Glossary* to provide metadata at a business level of abstraction. For this context, it could contain *Business Terms*, and then link the terms that represent a *Data Asset* to a separate construct. Users should still be able to navigate to the *Data Asset*. But in this example, it doesn't make sense to store technical information such as *Column Data Precision* in a *Business Term*, because it would destroy the business level of abstraction.

It is important to clearly understand and establish an organization's asset types and what their characteristics are to understand the context of each asset[2]. This helps prevent those excruciating long discussions where everyone disagrees because they are comparing apples to oranges. What might be an acceptable description of a Business Term might not be as helpful in the case of a metric or KPI.

There are many ways to capture these levels of abstraction and their context, but there is no golden catch-all solution. Each company's modelling strategy will depend on governance maturity, organizational hierarchy, current operation, data management strategy and technology choice, among other factors.

Learn More

Learn about the different approaches to Enterprise Information Management

References

[1] Firican, George. (2018) Understanding the different types of a data steward http://www.lightsondata.com/understanding-data-stewardship-types/

[2] Romero, JP. (2018) A No-Fluff Primer on Data Governance. http://viewpoints.io/entry/a-no-fluff-primer-on-data-governance

Originally published on September 4th, 2019

What's your view? Add your question or comment

About the Authors

JP Romero

jp.romero@kalypso.com

JP is a Senior Technical Consultant at Kalypso. He loves designing and implementing EIM solutions with a wide range of clients across different sectors, such as financial services, retail, and travel & tourism.

Gaurav Arora

gaurav.arora@kalypso.com

Gaurav is a partner at Kalypso.