

Using Collibra for Reference Data Management At Aspen Insurance

Reference data management - the management of code tables containing taxonomies and hierarchies - is a difficult area for many enterprises. There is a need to manage the actual data, principally the code and description values, but there is also a need to manage the meaning and context of both the tables and their content. This is a core Enterprise Data Management challenge since reference data can be anywhere from 20% to 50% of the tables in a database, and usually encompasses more than half of the business semantics of an enterprise (counting records in reference data tables as business concepts).

Introduction

Aspen Insurance is one enterprise that has recognized it must fully govern its reference data to realize its vision for enterprise data management. It is a leading insurance and reinsurance company with a global footprint. In recent years Aspen has grown in terms of diversity of business, products, and geographies, all of which has increased the demand for data integration to unlock value in the data and drive risk down.

"Reference data is easier to manage when you have single operational silos, but multiple platforms and a rising demand for analytics mean it cannot be ignored." Aspen CDM.

From Vision to Collibra

Reference data was not the only data governance need that Aspen had. Also in the mix were regulatory and data quality requirements. Rather than deal with all of these elements separately, Aspen took the time to formulate a vision for mature data management and then set about looking for a tool where it could be implemented. Aspen chose Collibra because of its semantic capabilities, collaborative nature, and ease of use and business friendliness, and immediately began to move reference data into it.

Collibra's open, business-facing interface was used to develop a portal where any Aspen user could look up the portfolio of reference data.

"Integrating knowledge management with data value management is often overlooked in reference data management, so it is something we have been careful to make available in Collibra." Felix van der Maele, CEO, Collibra.

As a result users could find a reference data table, find the data content inside it, and also acquire knowledge about the individual table and about the individual records in it. This solved the problem of the full business meaning of the codes being decoupled from the data of the tables.

Aspen also took advantage of Collibra's advanced data governance capabilities, including the ability to define and execute stewardship workflows. A problem with reference data is its diversity. For instance, an external table such as ISO 3166-1 Alpha-2 Country Codes needs a completely different maintenance approach, since it depends on an external authority, to Claims Type which is purely internal and owned by Aspen. Without Collibra every reference data table would have needed custom code development - which would have been impossible. Instead, workflows were configured in Collibra for each table, and immediately deployed within the tool to provide the required governance for the tables.

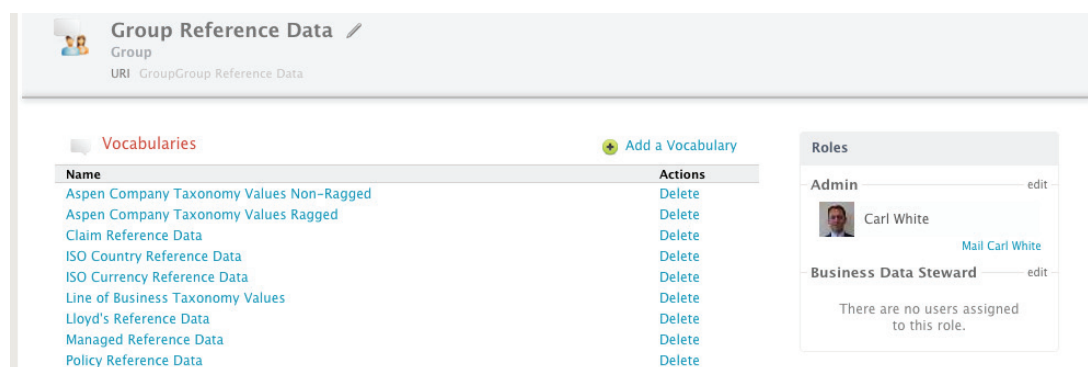


Figure 1: Single Reference Data Governance Platform in Collibra

Taming the complex business semantics inherent in reference data is only one challenge. Another is hierarchies. These often suffer from the fact that they can be implemented in a simple data model which then hides complexities that users need to rapidly find and understand.

Figure 2: Aspen Enterprise Hierarchy

Collibra has a visualization tool that immediately presents the actual hierarchy straight from the reference data table.

This materializes the reality of the hierarchy which otherwise would required a user to extract them on a case by case basis from a relational database table.



Another big issue with reference data management is getting it synchronized across the enterprise. Different applications have different interfaces and exist in diverse environments, making this task very hard. Yet, without synchronization of reference data, data integration is at risk.

Aspen solved this problem by using the governance features of Collibra to establish communities of interest for each reference data table. So, when a particular table is updated, emails are automatically sent to stewards in the community of interest for that table. Beyond this, reference data is fed into a central low-latency data environment that specifically exists for certain application to perform reference data lookups. Such caching components are becoming increasingly common in architectures, but, of course, they cannot manage any aspect of semantics – for which Collibra provides a user friendly front end.

Business-friendly user interface
Involve business through WIKI-style collaboration

Easy report building

Save custom views through a combination of filters

- Define custom roles and custom workflow (e.g. approval)

Compare and rollback between versions

Semantic modeling

Supports semantic relations and ontology standards

XML, UML, XMI, Excel, RDF/OWL, PDF, ...

java, javascript, REST, Relational Database, ...

Collibra's Business Semantics Glossary supports you in launching or maturing your Data Governance program. We help you define what your data means for your business, involving all stakeholders, while integrating with your existing IT infrastructure.

Collibra is the only platform that allows the business to take charge of the data in their organization.

For more information on the Business Semantics Glossary, you can contact Collibra via

@collibra (twitter)