

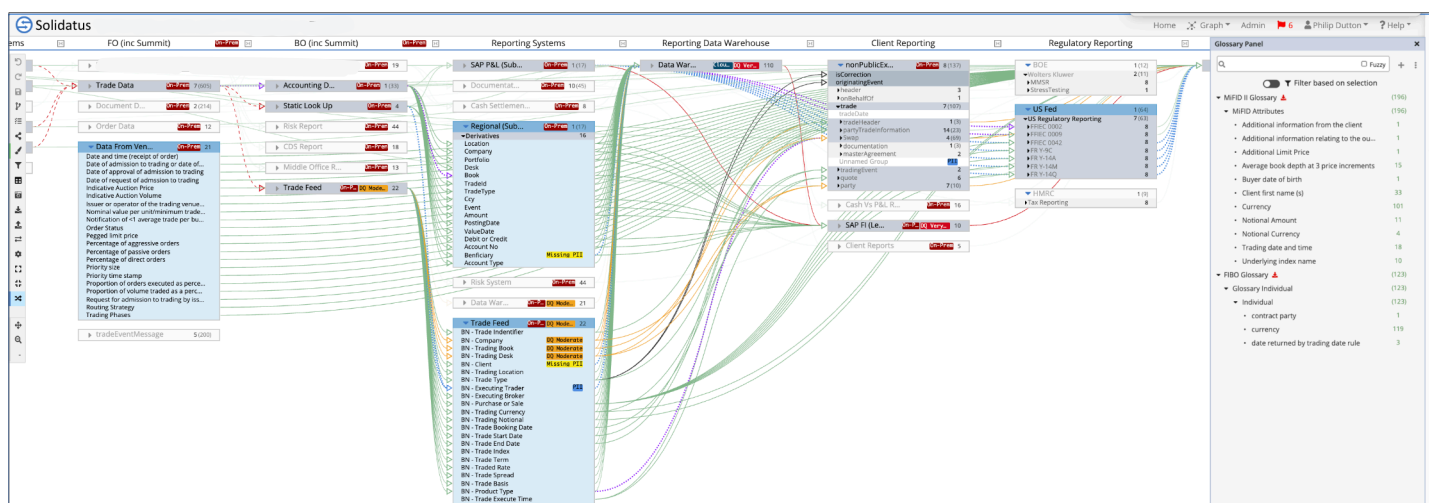
Cataloging Data: A foundation of Data Management

Driven principally by ever-growing regulatory burdens, enterprises have become increasingly aware of the data that they both generate and acquire in order to be able to run their businesses. The role of Chief Information Officer, almost unknown 10 years ago but now essential for global and major companies, indicates the emphasis placed on data in commerce today. Data, it is said, is 'the New Oil' – yet how can this invaluable (but intangible) resource be managed, optimised, and exploited for competitive advantage? The key to achieving these objectives is robust, sustainable Data Governance and the foundations of this are laid by Data Cataloging.

KEY PRINCIPLES AND CONSIDERATIONS

A company's data constitutes the modern equivalent of the historic 'books and records'. It may also be its stock in trade: a credit card company, for example, may create nothing physical (increasingly, monthly statements to customers and even the card itself exist only in electronic forms within our computers and mobile devices) but all of the purchases of goods and services it enables are real-world events that depend on the data that company manages. It is imperative that the data is of high quality, by which we mean it is complete, accurate, up-to-date and secure. Ensuring that these criteria are maintained is an essential part of Data Management.

In order to achieve high-quality data, an organisation needs to develop or employ a number of related data catalog artefacts, including [Data Dictionaries](#), [Business Glossaries](#) and [Data Taxonomies](#). Each of these plays a different role in comprehending, managing and optimising an enterprise's data estate so let's examine each in more detail and see how Solidatus is uniquely able to support not only their development within a company but also to allow the knowledge held within them to be shared and deployed across the organisation.



SOLIDATUS FOR DATA CATALOGING

Data Cataloging is, fundamentally, the understanding and documentation of the meaning of data, established through the process of identifying, validating and adding metadata to the logical and physical data entities of the enterprise. The greater the sharing of this understanding across an organisation, the greater are the business benefits and competitive advantage which accrue from it.

By eliminating confusion and misunderstandings over the nature of the data, organisations get to iterate faster and management gains the confidence to implement robust, appropriate, fit-for-purpose Data Governance that supports this accelerated cadence. The Solidatus Data Catalog is able to resolve terminology conflict and inconsistent nomenclature so all users share a common language. Users understand exactly what the metadata refers to, even if they or others have previously used different terms or colloquialisms.

DELIVERING THE FOUNDATIONS TO DATA MANAGEMENT

DATA DICTIONARY

A Data Dictionary is a description of the constituents of a physical data set, for example, a table and its columns. A Data Dictionary may, at its simplest, be a list of the column headings in a table, together with a description of the data that is entered into them. Usually, however, they include other metadata, such as the data type, any size limits, whether the field forms part of a key, standards to be used (eg, ISO 4217 for currency codes) cardinality and optionality of content. Typically, Data Dictionaries are kept by the support team maintaining a system. This limits sharing of information and standards across the company, leading to inconsistencies and undue reconciliations. Designed from its inception to foster collaboration, Solidatus offers the ideal platform on which to build, store, collate, maintain and share Data Dictionaries. Its ability to hold unlimited metadata and to display this simply and clearly means better understanding of the data and, if held as a Glossary model, can be linked to Solidatus lineage models to highlight inconsistencies and possible data breaks across the systems estate.

BUSINESS GLOSSARY

A Business Glossary contains the words or phrases (terms) used by the business to conduct the business, with their meanings, within the business context. A Business Glossary refers not to physical data, as with a Data Dictionary, but to logical, or conceptual, data. Business Glossaries are necessary because every company uses terms in a way that is unique to that company, although they may not necessarily be consistently applied across the company. For example, terms like 'Client', 'Customer', 'Counterparty' and 'Account' may all be used interchangeably but have very different meanings, say, in Sales and the Back Office. A Business Glossary, therefore, provides meaning and also context for that meaning, resolving ambiguity, cataloging synonyms and identifying context-dependent usages. Solidatus Glossary models allow subject matter experts (SMEs) across the organisation to record the terms their area employs: these can then be shared, centrally-collated and disseminated across the business to ensure consistency of understanding and eliminate confusion.

DATA TAXONOMY

Organisations wanting to understand their data in its entirety may build a Data Taxonomy, which identifies the main categories of data that they have and then, with increasing granularity, the data structures within them, down to an indivisible attribute level. As with the Business Glossary, this describes logical data. A Business Glossary is a prerequisite for developing a Data Taxonomy: without one, the enterprise-wide nature of a Data Taxonomy means that the synonyms and context-dependent usages described above will preclude an accurate understanding of the underlying data. The benefits to an organisation of a Data Taxonomy are considerable: rationalisation of data acquisition and storage, elimination of duplication, determination of appropriate data ownership and identification of gaps within the physical data assets across the business. The Solidatus Glossary model supports multi-level hierarchical domain structures and is an ideal repository for a Data Taxonomy. Through its in-built federated workflow it enables the apportioning of discovery and documentation required to build enterprise-scale results in a realistic and achievable time frame.

ABOUT SOLIDATUS

Award-winning Solidatus, the leader in metadata management, enables the world's largest data-rich and regulated organisations to effectively manage their data, people and processes, reducing complexity and risk through transparency, automation and collaboration. We provide organisations with a solution that allows them to fundamentally redesign their organisational data culture and capabilities by enabling the creation of a holistic organisation-wide digital map that details all the relationships that interact and impact their data, accelerating modernisation and transformation. The Solidatus methodology for digitally transforming organisations to be data-centric and lineage-enabled is changing the way organisations manage their data. Quickly being adopted by organisations across the globe, including top-tier global financial, pharmaceutical, utility and infrastructure firms and has been implemented by leading consulting and technology firms.

KEY BENEFITS

- ▶ Reduced maintenance overheads through better understanding of system data usage
- ▶ Consistent terminology, leading to better understanding of business aims and processes
- ▶ Improved adherence to standards, reducing breaks and reconciliation requirements
- ▶ Rationalisation of data acquisition and storage
- ▶ Facilitated adoption of industry best-practice
- ▶ Easier, efficient, and more widespread re-use of data assets
- ▶ Identification and resolution of data ownership questions
- ▶ Establishment and enforcement of enterprise-wide data quality standards