

Single source of truth

In information systems design and theory, **single source of truth** (**SSOT**) is the practice of structuring information models and associated data schema such that every data element is stored exactly once. Any possible linkages to this data element (possibly in other areas of the relational schema or even in distant federated databases) are by reference only. Because all other locations of the data just refer back to the primary "source of truth" location, updates to the data element in the primary location propagate to the entire system without the possibility of a duplicate value somewhere being forgotten.

Deployment of an SSOT architecture is becoming increasingly important in enterprise settings where incorrectly linked duplicate or de-normalized data elements (a direct consequence of intentional or unintentional denormalization of any explicit data model) pose a risk for retrieval of outdated, and therefore incorrect, information. A common example would be the electronic health record, where it is imperative to accurately validate patient identity against a single referential repository, which serves as the SSOT. Duplicate representations of data within the enterprise would be implemented by the use of pointers rather than duplicate database tables, rows, or cells. This ensures that data updates to elements in the authoritative location are comprehensively distributed to all federated database constituencies in the larger overall enterprise architecture.

SSOT systems provide data that are authentic, relevant, and referable.^[1]

Contents

Implementation

- Enterprise service bus (ESB)
- Master data management (MDM)
- Data warehouse (DW)
- SOLID & Source Code

See also

References

External links

Implementation

The "ideal" implementation of SSOT as described above is rarely possible in most enterprises. This is because many organisations have multiple information systems, each of which needs access to data relating to the same entities (e.g., customer). Often these systems are purchased "off-the-shelf" from vendors and cannot be modified in non-trivial ways. Each of these various systems therefore needs to store its own version of common data or entities, and therefore each system must retain its own copy of a record (hence immediately violating the SSOT approach defined above). For example, an ERP (enterprise resource planning) system (such as SAP or Oracle e-Business Suite) may store a customer record; the CRM (customer relationship management) system also needs a copy of the customer record (or part of it) and the warehouse dispatch system might also need a copy of some or all of the customer data (e.g., shipping address). In cases where vendors do not support such modifications, it is not always possible to replace these records with pointers to the SSOT.

For organisations (with more than one information system) wishing to implement a Single Source of Truth (without modifying all but one master system to store pointers to other systems for all entities), three supporting technologies are commonly used:

- Enterprise service bus (ESB)
- Master data management (MDM)
- Data warehouse (DW)

Enterprise service bus (ESB)

An enterprise service bus (ESB) allows any number of systems in an organisation to receive updates of data that has changed in another system. To implement a Single Source of Truth, a single source system of correct data for any entity must be identified. Changes to this entity (creates, updates, and deletes) are then published via the ESB; other systems which need to retain a copy of that data subscribe to this update, and update their own records accordingly. For any given entity, the master source must be identified (sometimes called the Golden Record). It should be noted that any given system could publish (be the source of truth for) information on a particular entity (e.g., customer) and also subscribe to updates from another system for information on some other entity (e.g., product).

An alternative approach is point-to-point data updates, but these become exponentially more expensive to maintain as the number of systems increases, and this approach is increasingly out of favour as an IT architecture.

Master data management (MDM)

An MDM system can act as the source of truth for any given entity that might not necessarily have an alternative "source of truth" in another system. Typically the MDM acts as a hub for multiple systems, many of which could allow (be the source of truth for) updates to different aspects of information on a given entity. For example, the CRM system may be the "source of truth" for most aspects of the customer, and is updated by a call centre operator. However, a customer may (for example) also update their address via a customer service web site, with a different back-end database from the CRM system. The MDM application receives updates from multiple sources, acts as a broker to determine which updates are to be regarded as authoritative (the Golden Record) and then syndicates this updated data to all subscribing systems. The MDM application normally requires an ESB to syndicate its data to multiple subscribing systems.^[2] Customer Data Integration (CDI), as a common application of Master Data Management, is sometimes abbreviated CDI-MDM.

Data warehouse (DW)

While the primary purpose of a data warehouse is to support reporting and analysis of data that has been combined from multiple sources, the fact that such data has been combined (according to business logic embedded in the data transformation and integration processes) means that the data warehouse is often used as a *de facto* SSOT. Generally, however, the data available from the data warehouse are not used to update other systems; rather the DW becomes the "single source of truth" for reporting to multiple stakeholders. In this context, the Data Warehouse is more correctly referred to as a "single version of the truth" since other versions of the truth exist in its operational data sources (no data originates in the DW; it is simply a reporting mechanism for data loaded from operational systems).

SOLID & Source Code

In software design, the same schema, business logic and other components are often repeated in multiple different contexts, while each version refers to itself as "Source Code". To address this problem, the concepts of SSoT can also be applied to software development principals using processes like recursive transcompiling to iteratively turn a single source of truth into many different kinds of source code, which will match each other structurally because they are all derived from the same SSoT.

See also

- Don't repeat yourself (DRY)
- SOLID (object-oriented design)
- Database normalization
- Single version of the truth
- System of record

References

1. IBM Smarter Planet - Operational risk management for financial services^[1] (http://www.ibm.com/smarterplanet/za/en/banking_technology/nextsteps/solution/Z017038Z16405R75.html)
2. BAYT Job Site - June 2014^[2] (<http://www.bayt.com/en/specialties/q/7370/what-are-the-top-business-processes-and-applications-that-need-master-data-management/>)

External links

Retrieved from "https://en.wikipedia.org/w/index.php?title=Single_source_of_truth&oldid=844176395"

This page was last edited on 3 June 2018, at 04:35 (UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.