

The Foundation of Data Fabrics: Semantic Knowledge Graphs

By Dr. Jans Aasman, CEO, Franz Inc.:



Data management agility has become of key importance to organizations as the amount and complexity of data continues to increase, along with the desire to avoid creating new data silos. The concept of creating a 'data fabric' as an agile design concept has been proposed by leading analysts, such as Mark Beyer, Distinguished VP Analyst at Gartner. "The emerging design concept called 'data fabric' can be a robust solution to ever present-day management challenges, such as the high-cost and low-value of data integration cycles, frequent maintenance of earlier integrations, the rising demand for real-time and event-driven data sharing, and more," says Mark Beyer.

As a data fabric readily connects and provides singular access to all data sources distributed throughout the enterprise, semantic knowledge graphs provide the foundation that makes this design possible. Semantic knowledge graphs and aspects of AI are necessary for the data fabric architecture to work. According to Gartner, "The semantic layer of the knowledge graph makes it more intuitive and easy to interpret, making

the analysis easy for D&A leaders. It adds depth and meaning to the data usage and content graph, allowing AI/ML algorithms to use the information for analytics and other operational use cases.” In this respect, graph applications are the enabler of both data fabrics and the AI that supports them.

Data fabrics involve additional tooling like respective layers for data integration and run-time orchestration, in addition to active metadata management. Nonetheless, these capabilities would fail to properly function without the semantic layer, and data cataloging value, of semantic knowledge graphs that are foundational to realizing this grand data management vision.

Semantic Curation

Semantic knowledge graphs are the underlying framework for the ability to seamlessly connect to, access, and query all data sources relevant to the enterprise. This capability includes sources internal and external to organizations, in any type of cloud setting, on-premises, or at the cloud’s edge. The first way semantic knowledge graphs enable a uniform fabric across each of these environments, tools, and technologies is by furnishing a layer harmonizing the semantics between them.



The numerous resources joined together in a comprehensive fabric involve data of different structure variations (structured, unstructured, and semi-structured), terminology, schema, taxonomies, business units, and storage formats.

Semantic knowledge graphs specialize in harmonizing data with these and any other type of distinction via standardized data models and uniform taxonomies. Moreover, they do so in business-friendly terminology as opposed to arcane IT code. Thus, end-users from data scientists to sales personnel can readily understand what any of the data in a data fabric means, as well as how they may relate to his or her business goals.

Architectural Gains

The second capability of semantic knowledge graphs that's indispensable to the data fabric precept outlined by Gartner is connecting the metadata together from the array of sources involved. The assortment of metadata represented via this paradigm is considerable and includes business, technical, and operational metadata, the final of which pertains to application composition, execution results, and runtime environments. Granted, data cataloging capabilities are required to tag that metadata, classify it, and add tools for data lineage and for exchanging this information between users. Still, this metadata should ideally be represented in a semantic knowledge graph.

Another area of specialization of these graph applications is their ability to link together data of any variation. They do so in a manner that focuses on the relationships between data (or metadata), providing a vital element of contextualized understanding of how data pertains to each other that other approaches overlook. Such metadata is essential for identifying best practices and techniques for integrating specific datasets, orchestrating various applications, and selecting the most appropriate source for any particular business task. It's also the basis for automating aspects of any of these needs via AI.

Read the full article at [Data Science Central](#).

