

A Tale of Two Semantic Layers



Helyx Horwitz • June 26, 2024

In the rapidly evolving landscape of information and data management, the term Semantic Layer is increasingly utilized to describe two different types of 'layers' in the modern data stack. The first Semantic Layer is focused on enabling centralized analytics within a company, which we will refer to as a **'metrics layer'** due to its frequent conflation with that term. The metrics layer is an evolution of the **original use of the term** 'Semantic Layer' and was first coined by Business Objects in the 1990s to describe a product that enabled end users to query databases without knowing SQL or the underlying structure of the databases.

The second Semantic Layer is an abstracted sense-making layer that brings all the data and information managed by a company into context, which we will refer to directly as the **'Semantic Layer'**. The Semantic Layer focuses on linking information structures and is more directly related to the frameworks behind the **semantic web**. The Semantic Layer prioritizes interoperability and extensibility to form a layer of abstraction that enables understanding across the various components of an enterprise's domain.

Ensuring that the right kind of Semantic Layer is in place for the job at hand is essential for building an enterprise infrastructure that not only meets today's needs but can also be extended and adapted to encompass future needs.

The Layer at Issue

Both versions of the term Semantic Layer are positioned between the content and data produced by an organization and the applications that utilize and present that data and content. Semantic Layers enable a universal understanding of enterprise-based knowledge, consisting of an abstraction layer that brings context and standardized definitions to individual data points. In the metrics layer, this happens by defining core measurements and repeated aggregations so that the analytics are repeatedly and consistently interpreted and implemented. In a **Semantic Layer**, not only are metrics defined, but also the data concepts at the entity level. The relationships along with implicit understandings within the entirety of

an organization are made explicit through the use of an ontological data model. This includes establishing relationships between content and the data itself across a variety of sources and types. Furthermore, informed by a core component of the semantic web, Semantic Layers are frequently built on widely adopted standards that make them interoperable and extensible in a variety of ways.

Semantic Layer Tools

A quick search for the phrase ‘Semantic Layer tool’ will bring up numerous platforms, think-pieces, and reports about products promoting **all-in-one** (AIO) solutions that enable ‘self service metrics’. Purportedly, these tools will break down organizational silos and enable data-driven decision making. These AIO tools paint a compelling picture of seamlessly integrated metrics stores that enable business users to query and understand critical data without needing to know SQL or the structure of the data storage underneath it. This vision is both exciting and very real; these metrics layers enable businesses to increase the accessibility and use of their data assets. That said, the one-stop-shop Semantic Layer they promote is not as powerful or even as all inclusive as they claim, and they do not fulfill the true potential of a Semantic Layer.

Current platforms that market themselves as ‘AIO Semantic Layer’ solutions predominately fall into the category of metrics layers. These metrics layer tools focus on providing the end user with better and more consistently defined ways to utilize analytics and quantitative measurements across the organization. While this function is an important component of a metrics layer, it does not in and of itself bring forth the vision of a Semantic Layer.

Additionally, these solutions tend to employ proprietary frameworks, potentially creating vendor-lock and limiting interoperability with other platforms. At EK, we have worked with many clients to build **semantic solutions that enable far more than analytics alone**.

Some of the critical features and enablements of these types of Semantic Layers are detailed below:

Feature	Purpose	Metrics Layer	Semantic Layer
Makes multiple formats of data available for both humans and machines	Translates implicit human understanding into a machine readable format	✗	✓
Reflects human understanding of processes, relationships, and domains	Enables contextual understanding and action	✗	✓
Captures and connects content and data based on business or domain meaning and value	Enables domain specific understanding and interpretation	✓	✓
Represents the whole of the business/domain model	Facilitates enterprise level contextual understanding	✗	✓
Ability to connect unstructured data with structured and semi-structured data sources	Enables aggregation and unification of varied data formats and sources	✗	✓
Enriches data in source systems, enabling cyclical data quality improvement from source to downstream consumption	Improves data quality, decreases ongoing maintenance	✗	✓
Enables logical inference and semantic reasoning	Enables novel insight generation	✗	✓
Enables statistical inference	Enables analytical insight generation	✓	✓
Enables centralized governance of semantic assets	Support centrally managed governance	✗	✓
Conforms to open standards such as semantic web standards	Promotes interoperability, long term archiving, and decreases risk of vendor lock	✗	✓
Represents data independent of the underlying storage/format	Enables aggregation across platforms and data formats	✓	✓
Enables data federation and virtualization	Enables more performant and flexible processing	✓	✓

Transformative Semantic Layer Use Cases

Measuring Metrics

Metrics layers as well as Semantic Layers, can be utilized to better understand, make available, and generate consistency regarding key analytics and measurements across an enterprise. A primary use case for metrics layers is the concept of self-serve analytics, or

analytic measures that can be easily accessed and understood by users across the organization. Key factors that enable this are concepts like business glossaries which provide standardized definitions to alleviate variance in interpretation. The synthesis provided by standardized definitions, parameters for metrics, and integration with downstream tools enable analytic measurements to be more widely accessible and usable across a company, even by those without a data science background or understanding. While this is the primary focus and promotion behind the metrics layer, it is also one potential product of a Semantic Layer and the semantic frameworks that comprise it. For instance in a past project, **EK worked with a large financial services firm that wanted to enable navigating customer analytics alongside product information**. The semantic structure of an ontology helped link and make those metrics from a variety of formats available in one place. Enabling easily usable and understandable metrics is one use case for the Semantic Layer, but a true Semantic Layer goes beyond simply making metrics available for business users.

Information Discovery

To be truly transformative, a Semantic Layer must go beyond focusing solely on the possibilities of self-serve analytics. When humans process new information, we generally use layers of experience, contextual insights, and an understanding of the multiple, complex relationships between concepts to inform our comprehension. Similarly, graph data formats focus on establishing context and relationships between concepts. One of the critical enablers behind the approach to a Semantic Layer is the movement into a graph data format. This divorces data structures from the limits of relational data models and more closely resembles the way that humans think, via multi-dimensional connections.

EK worked with a bioscience technology provider who faced challenges connecting data across a large digital ecosystem, and this had a particularly difficult impact on their e-commerce website. The search experience was driven by keywords, and it did not adequately integrate all available data sources, which were spread across an array of platforms and formats. EK's solution included a semantic data model that utilized a knowledge graph to bring the diversity of information sources into an actionable context. The use of a graph data framework enabled similarity indexing to create recommendations that were grounded in contextual relevance based on a multitude of information facets. Additionally, a link classifier was used to create recommendations based on a user's engagement across products and content. In this instance, graph data science was an essential component that enabled recommendations around complex data assets without manually creating explicit relationships for the over 70,000 product items and content assets already in place.

By enabling new and multi-dimensional connections between data and information, a Semantic Layer can bring about new insights. Semantic layers can help connect the dots without knowing the question that you are asking. Most importantly, the structure is not

dependent on human assertions or insights. Semantic Layers tie together the structures and relationships across a domain to not only organize and standardize the presentation of data, but also to enable new insights.

Aggregation with 360 Views

A Semantic Layer can encompass an enterprise's entire scope of knowledge, providing a layer of aggregation that brings together information from across formats and systems. This serves as a connection not only between data points but also between content and data, both structured and unstructured. The ability of the **Semantic Layer to account for content assets** and metadata unlocks a wider range of use cases including recommendation engines, semantic search, and 360-degree views.

At EK, a common client request is the enablement of **360-degree views**. Most enterprise organizations have the information needed for this aggregation spread across multiple data formats and sources. Product information might be stored in a Product Information Management (PIM) platform, while the active sales data is pushed into the Enterprise Revenue Platform (ERP), and the images of a product and a PDF user manual are stored in the Digital Asset Management (DAM) system.

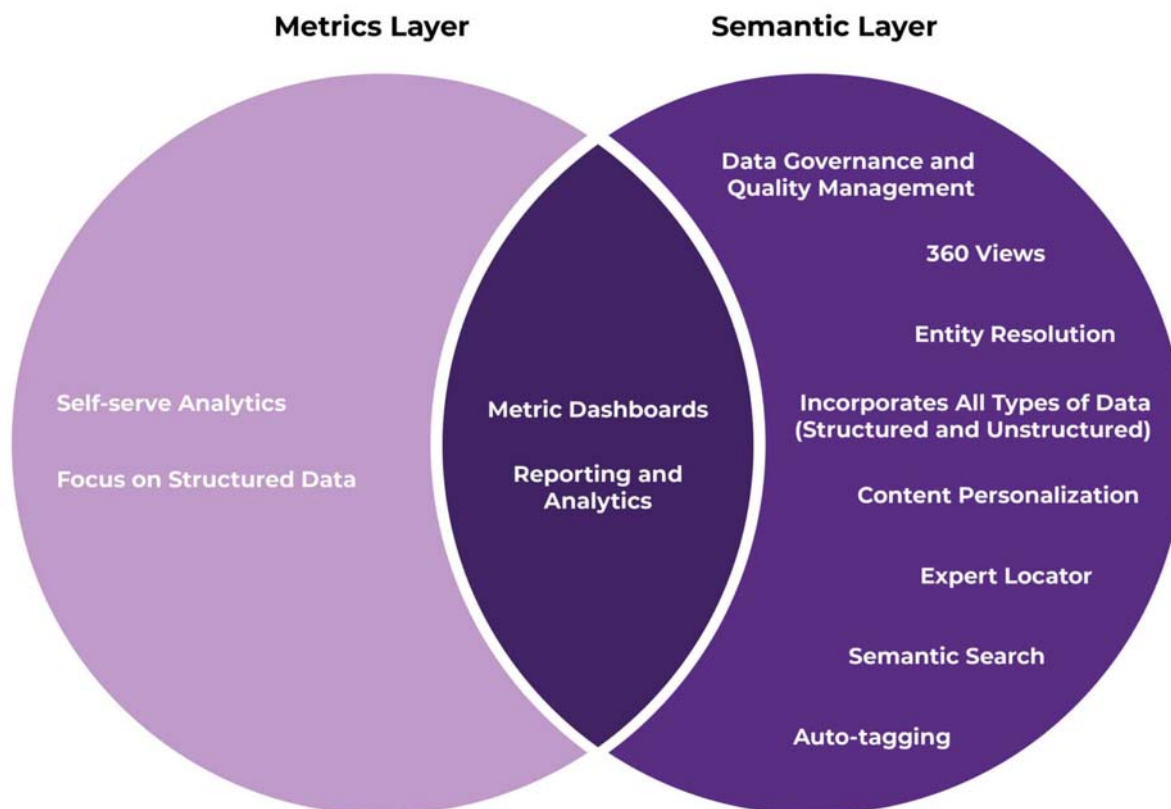
The ontology that informs a Semantic Layer can illustrate these relationships, establishing that the PIM contains relevant metadata and product SKU information, while the images associated with that product are located in the DAM, and current sales data should come from the ERP. The Semantic Layer establishes the connections between these disparate data points across differing structures in order to bring all the information together in a single view.

Source of Truth

One of the critical features that the Semantic Layer enables is a unified understanding of a business's content and data in a way that is contextual, clear, and consistent across the enterprise. Recently, Air Canada was forced to honor a refund policy that **was fabricated by its customer assistance chatbot**. The chatbot found bad information somewhere in the available Air Canada information repositories and passed along that inaccurate information to a customer. The risk of exposure of inaccurate, outdated, or simply incorrect information increases exponentially when technologies like chatbots are exposed to a large data repository without appropriate frameworks in place for fact-checking. The Semantic Layer is one possible solution to establish a single source of truth for a business. Through this source of truth, the reliability and consistency of downstream applications, such as chatbots, can be improved.

In a recent case study, **EK worked with a federal research and development center to utilize existing information repositories to establish sources of truth** (among other

things) as part of a Semantic Layer implementation. This resulted in more consistent and more accurate metadata across the repository of over 100,000 documents, along with better retrieval and understanding of documents overall. By establishing this organizationally specific source of truth, Semantic Layers enable domain-specific understanding that is unique to the context of the business in which the layer is embedded.



Conclusion

These two kinds of 'semantic' layers might be confusing for the layperson but they each support organizations in different ways and have divergent use cases when it comes to garnering value from the layer. When seeking a truly transformational Semantic Layer for an enterprise, it is important to seek a solution that goes beyond enabling democratized analytics and structured data to offer powerful insights across the entirety of a company's knowledge and information assets. This enables use cases and applications that go beyond metrics dashboards and support broad 360-degree views, accurate chatbots, and recommendation engines. While the promise of AIO metrics layer technologies is appealing, the reality is that most enterprises will need the breadth of a customized layer and the depth of integration and interoperability that a semantically informed Semantic Layer provides. If you would like to start or expand your organization's use of a Semantic Layer, **contact EK today**.