

What is a Semantic Layer?

There are several reasons why the notion of semantic layers has reached the forefront of today's data management conversations. The analyst community is championing the data fabric tenet. The data mesh and data lake house architectures are gaining traction. Data lakes are widely deployed. Even architectural-agnostic business intelligence tooling seeks to harmonize data across sources.

Each of these frameworks requires a semantic layer to ascribe business meaning to data – via metadata – so end users can understand data for their purposes and streamline data integration. This layer sits between users and sources, so the former can comprehend data without knowing the underlying data formats.

What are the advantages of a semantic layer in your data infrastructure?

A semantic layer is an intermediate layer in an Enterprise architecture that sits between the data sources and the applications that use the data. It provides a number of advantages in terms of data management, integration, and accessibility. Some specific advantages of a semantic layer include:

- 1. Improved data integration:** A semantic layer can help to integrate data from multiple sources by providing a common data model and set of APIs that can be used to access the data. This makes it easier to build applications that work with data from multiple sources.
- 2. Enhanced data accessibility:** A semantic layer can provide a higher level of abstraction over the data sources, making it easier for users to access and work with the data. This can be particularly useful for users who are

not technical experts or who do not have in-depth knowledge of the underlying data sources.

3. Better data governance: A semantic layer can help to enforce data governance policies by providing a centralized point of control for data access and management. This can help to ensure that data is used in a consistent and controlled manner.

4. Upgraded data quality: A semantic layer can help to improve the quality of the data by providing tools and processes for data cleansing, validation, and transformation. This can help to ensure that the data is accurate and consistent.

5. Advanced data security: A semantic layer can provide an additional layer of security by controlling access to the data sources and enforcing security policies. This can help to protect sensitive data and ensure that it is only accessed by authorized users.

Overall, a semantic layer can provide a number of benefits in terms of data integration, accessibility, governance, quality, and security, making it a valuable component of a data infrastructure.

Semantic Layers with W3C's Semantic Technologies

Semantic Technology refers to a set of tools and technologies that are used to represent, store, and manipulate data in a way that allows it to be understood and interpreted by computers. Some examples of semantic technology include graph databases, ontologies, and semantic web standards such as RDF and OWL.

While semantic technology is the preferred way to implement a semantic layer, some other players have tried other technologies including traditional relational databases, data warehousing tools, or even flat files. The key is to provide a

common data model and set of APIs that can be used to access the data in a consistent and predictable manner.

That being said, Standards based W3C Semantic Technology, like that offered by AllegroGraph, has a huge advantage when it comes to implementing a semantic layer. In particular, Semantic Technology is well-suited for representing complex, interconnected data relationships, and it can provide a high level of flexibility and adaptability when it comes to working with different data sources and structures. As such, semantic technology can be a particularly useful choice for organizations that need to integrate and work with large volumes of complex data.

There are rare cases where a proprietary semantic layer may work and the organization might not mind getting locked into the ecosystem of a vendor for their metadata management needs. But for the majority of use cases, the best way to future-proof the enterprise is to adopt a standardized semantic layer with semantic technologies. This method provides a seamless business understanding of data that complements any current or future IT needs, while reinforcing data integration, analytics, and data governance.

AllegroGraph Named “2023 – Trend Setting Product” by Database Trends and Applications

Franz Inc., is proud to announce it has been named a “2023 – Trend Setting Product” by Database Trends and Applications.

According to Database Trends and Applications, today's data environments are highly diverse—residing on many platforms and requiring a variety of approaches to ensure data resiliency and availability. Delivering technology alone will not be enough in 2023. According to Gartner, investing in sustainable technology will give companies a leading edge as we move forward into a new year.

Sustainable technology is comprised of a framework of solutions that increases the energy and efficiency of IT services; enables enterprise sustainability through technologies like traceability, analytics, emissions management software, and AI; and helps customers achieve their own sustainability objectives.

Investments in sustainable technology also have the potential to create greater operational resiliency and financial performance, while providing new avenues for growth.

“Today's data environments are highly diverse—residing on many platforms and requiring a variety of approaches to ensure data resiliency and availability,” said Tom Hogan, Group Publisher, Database Trends and Applications. “To help make the process of identifying useful products and services easier, each year, DBTA presents a list of ‘Trend-Setting Products.’ These products, platforms, and services range from long-established offerings that are evolving to meet the needs of their loyal constituents, to breakthrough technologies that may only be in the early stages of adoption.”

The Hype Around Semantic

Layers: How Important Are Standards?

There are several reasons why the notion of semantic layers has reached the forefront of today's data management conversations. The analyst community is championing the data fabric tenet. The data mesh and data lake house architectures are gaining traction. Data lakes are widely deployed. Even architectural-agnostic business intelligence tooling seeks to harmonize data across sources.



Each of these frameworks requires a semantic layer to ascribe business meaning to data – via metadata – so end users can understand data for their purposes and streamline data integration. This layer sits between users and sources, so the former can comprehend data without knowing the underlying data

formats.

Additionally, a semantic layer must incorporate a digital asset knowledge graph for a unified description of data assets in all sources – like those feeding data lakes and data lakehouses. This catalog is especially important for identifying what data is in unstructured data sources, relational databases, streaming data, document stores, and other sources for data fabric or data mesh deployments.

Some “semantic layers” use non-standard, proprietary technologies to store metadata. This approach prevents the use of industry-wide ontologies like FIBO, (financial services), SNOMED (medical), SCONT0 (supply chain), OBML (life sciences), CDM-Core (manufacturing), GoodRelations (e-commerce), or SWIM (aviation). It also complicates future data

integration and reinforces vendor lock-in.

Conversely, semantic layers implemented with W3C's Semantic Technologies are based on open-source standards that complement an organization's existing IT infrastructure. They future-proof the enterprise, prevent vendor lock-in, and provide a uniform view of all data (regardless of differences in formatting, types, and structure) that's optimal for data integration, data governance, and monetization opportunities.

Read the Full Article at Dataversity.

Using Ansible for AllegroGraph multi-server installation

Introduction

Visit our Github example page for more details on creating Ansible playbooks for installing, starting and stopping an AllegroGraph server on one or more machines.

You must edit three files to personalize the configuration and then you can use the Makefile to install, start and stop the AllegroGraph servers on one or more machines.

Configuration

There are a vast number of server configuration parameters for AllegroGraph, far more than you would want to express as arguments to a configuration function.

In this directory there is a file `agraph.cfg-template` that you should edit to add or modify the configuration options (Link Here) you wish to set.

The only options you should *not* specify in `agraph.cfg-template` are

Port
SSLPort

as these will be added to the final `agraph.cfg` file based on values you put in `vars.yaml`.

basedir

One important variable in `vars.yaml` is `basedir`. The server will be installed in a newly created directory that is the value of `basedir`. Also a

BaseDir

directive will be put in the `agraph.cfg` specifying this value. This means that inside `agraph.cfg` you can (and should) use relative pathnames to refer to directories and files inside this directory tree.

settings

This line is always in the `agraph.cfg`

SettingsDirectory settings

It places the settings directory as a subdirectory of the `basedir`. Do not change this line as the settings directory has to be here in order for the super user password to be installed correctly.

Installation

Before starting the installation edit the following files:

inventory.txt

Insert the name of the machines on which you want the AllegroGraph server to be installed. Replace the sample machine names already in the file with the names of your machines. After you edit the inventory.txt file you can type

```
% make
```

to see if the machines you specified are reachable by ansible.

vars.yaml

That file contains descriptions of the variables to be set as well as some sample values that you'll need to change.

agraph.cfg-template

This is the file that will be modified to create the agraph.cfg that will be installed with AllegroGraph. You should review the server settings document to see which additional configuration parameters you wish to specify.

make install

The command

```
% make install
```

will run through the installation steps to install the server. It will do a superseding install meaning it will overwrite the server executables but it will not remove any repositories. However it is best to backup your installation before doing the install in case something unexpected happens and repos are lost.

make clean-install

If you wish to completely remove an installed AllegroGraph server so that make install gives you a totally fresh directory then

```
% make clean-install
```

will delete everything including repos that are found in subdirectories of the installation.

make start

To start the server on all machines do

```
% make start
```

make stop

To stop the server on all machines do

```
% make stop
```

If you are going to make install be sure to stop all servers before doing so.

IEEE – Entity Event Knowledge Graph for Powerful Health Informatics

As part of Franz's participation in the IEEE – ICHI conference, our paper has been published and is available from

the IEEE Website.



ICHHI 2022 is a premier community forum concerned with the application of computer science, information science, data science, and informatics principles, as well as information technology, and communication science and technology to address problems and support research in healthcare, medicine, life science, public health, and everyday wellness.

Franz Inc. presented on June 14th – **Entity Event Knowledge Graph for Powerful Health Informatics**

Download Franz's IEEE Publication – Entity Event Knowledge Graph for Powerful Health Informatics.

Conference Website

Franz Inc. Named a Big Data

50 Innovator

Franz Inc. has been named to the “Big Data 50: Companies Driving Innovation in 2022” by Database Trends and Applications.

AllegroGraph provides organizations with essential Knowledge Graph solutions, including Graph Neural Networks, Graph Virtualization, GraphQL, Apache Spark graph analytics, and Kafka streaming graph pipelines. These capabilities exemplify AllegroGraph’s leadership in empowering data analytics professionals to derive business value out of Knowledge Graphs.

“Data has only become more important as organizations look ahead to what a post-pandemic world could look like,” said Tom Hogan, Group Publisher, Big Data Quarterly. “To support organizations in navigating through new challenges and a rapidly evolving big data ecosystem, Big Data Quarterly presents 2022s ‘Big Data 50,’ a list of companies driving innovation and expanding what is possible in terms of collecting, storing, and extracting value from data.”

Some of the new approaches being embraced to help drive greater benefit from data are DevOps and DataOps, data quality and governance initiatives, hybrid and multi-cloud architectures, IoT and edge computing, and a range of next-gen databases.

According to ResearchAndMarkets.com, big data in business intelligence apps will reach \$54.9B by 2027, data integration and quality tools are projected to reach \$10.2B globally by 2027, and enterprise performance analytics will reach \$31.4 globally by 2027.

Industry verticals of various types have challenges in capturing, organizing, storing, searching, sharing, transferring, analyzing, and using data to improve business.

Big data is making a big impact in certain industries such as the healthcare, industrial, and retail sectors.

Another report by Quest Software titled, “The 2022 State of Data Governance and Empowerment Report,” found that data quality has overtaken data security as the top driver of data governance initiatives, with 41% of those surveyed agreeing that their business decision-making relies fundamentally on trustworthy, quality data.

At the same time, however, 45% of IT leaders say that data quality is the biggest detractor from ROI in data governance efforts. While they recognize its importance, they’re struggling to improve the quality of their data, and thus the ability to strategically and maximally leverage data in practice.

While the challenges of data visibility and observability differ across industries, DataOps was overwhelmingly recognized as the primary solution to drive forward data empowerment. Nine in 10 people surveyed agreed that strengthening DataOps capabilities improves data quality, visibility, and access issues across their businesses. The biggest opportunities to improve DataOps accuracy and efficiency lie in investing in automated technologies and deployment of time-saving tools, such as metadata management.

To support organizations in navigating through new challenges and a rapidly evolving big data ecosystem, Big Data Quarterly presents 2022s “Big Data 50,” a list of companies driving innovation and expanding what is possible in terms of collecting, storing, and extracting value from data.

Dr. Jans Aasman, Named Keynote Speaker for SEMANTiCS Conference 2022

Dr. Jans Aasman, CEO, of Franz Inc. will be delivering a Keynote presentation at the 2022 SEMANTiCS conference in Vienna, Austria.

Dr. Aasman's presentation, "The Role of Graphs in AI and Quantum Computing" will describe three emerging technology trends that will impact the Graph community and thought leadership opportunities for these technologies. Dr. Aasman's talk will cover Knowledge Graph's role in Natural Language Understanding, Graph Neural Networks (GNN) for predictive AI applications, and the convergence of Graph technologies and Quantum Computing.

Jans Aasman is a Ph.D. psychologist and expert in Cognitive Science – as well as CEO of Franz Inc. As both a scientist and CEO, Dr. Aasman continues to break ground in the areas of Artificial Intelligence and Knowledge Graphs as he works hand-in-hand with numerous Fortune 500 organizations as well as government entities worldwide.

The SEMANTiCS conference is an annual gathering of technology professionals, industry experts, researchers and decision makers to share and learn about new technologies, innovations and enterprise implementations in the fields of Linked Data and Semantic AI. Since 2005, the conference series has focused on semantic and graph technologies, which are today together with other methodologies such as NLP and machine learning the core of intelligent systems.

Innovative knowledge-sharing tools elevate the modern workplace – Article

The most meaningful developments to the knowledge-sharing space—and, by extension, to that of knowledge management as a whole—do not pertain to specific tools, platforms, or technologies.



Instead, they pertain to the goals of knowledge sharing, which have been irrevocably shaped by numerous forces in the modern workplace to include everything from distributed paradigms for working remotely to increasingly low latency responses characteristic of the digital age in which we live.

As such, today's knowledge-sharing tools are designed for collaboration, engagement, interactivity, and crowdsourcing. The tools themselves have changed little over the past couple of years and still involve facets of data catalogs, taxonomies, search, text analytics, data discovery, and data governance.

What's evolved, however, is their features, which have been updated for the sort of real-time interactions that make knowledge more accessible, reliable, and utilitarian than ever before.

The point of cataloging enterprise knowledge is to provide a central place to steer users to information relevant to their particular needs. "If you have a metadata graph that links the domain objects in your enterprise to the data catalog, then

you can start doing recommendations,” said Jans Aasman, CEO of Franz. “Like, here’s all the databases that are used the most for when people want to do something like this.”

Read the Full Article at [KMWorld](#).

Graph Neural Networks in AllegroGraph – Video

Enterprises are subscribed to the power of modeling data as a graph and the importance of using Knowledge Graphs for customer 360 and beyond. The ability to explain the results of AI models, and produce consistent results from them, involves modeling real-world events with the adaptive schema consistently provided via Knowledge Graphs.

Probably the most important reason for building Knowledge Graphs has been to answer the age old question: “What is going to happen next?” Given the data, relationships, and timelines we know about a customer, patient, product, etc. (“The Entity of Interest”), how can we confidently predict the most likely next event.

For example, in healthcare, what is the outcome for this patient given the sequence of previous diseases, medications, and procedures. For manufacturers, what is going to require repair next in this aircraft or some other point in the supply chain.

Machine Learning and more recently, Graph Neural Networks (GNNs) have emerged as a mature AI approach used by companies for Knowledge Graph enrichment. GNNs enhance neural network methods by processing graph data through rounds of message

passing, as such, the nodes know more about their own features as well as neighbor nodes. This creates an even more accurate representation of the entire graph network.

In this presentation we describe how to use graph embeddings and regular recurrent neural networks to predict events via Graph Neural Networks. We will also demonstrate creating a GNN in the context of a Knowledge Graph for building event predictions.

For more info – <https://github.com/franzinc/agraph-examples>

Montefiore Semantic Data Lake Tackles Predictive Analytics

Montefiore Medical Center is preparing to launch a sophisticated predictive analytics program for crisis patients, which is rooted in its real-time semantic data lake technology.

Semantic computing is becoming a hot topic in the healthcare industry as the first wave of big data analytics leaders looks to move beyond the basics of population health management, predictive analytics, and risk stratification.

This new approach to analytics eschews the rigid, limited capabilities of the traditional relational database and

instead focuses on creating a fluid pool of standardized data elements that can be mixed and matched on the fly to answer a large number of unique queries.

Montefiore Medical Center, in partnership with Franz Inc., is among the first healthcare organizations to invest in a robust semantic data lake as the foundation for advanced clinical decision support and predictive analytics capabilities.

Read the full article at [Health IT Analytics](#)