Bloor Research Covers AllegroGraph with FedShard

Written By: Daniel Howard
Published: 3rd June 2020

Link to Bloor Research Post on AllegroGraph 7

Graph database and *"knowledge graph solution"* vendor Franz Inc. has recently released its latest version of its graph offering, AllegroGraph, which as of the new release is in version 7.0. Alongside this, the company also released version 8.0 of Gruff, its visualisation and discovery engine for AllegroGraph and SPARQL endpoints. As of version 8, Gruff is now available via a web client as well as on the desktop.

For those of you who are unfamiliar with Franz and AllegroGraph, here is a brief description of the latter taken from our 2019 deep dive:

> *Franz AllegroGraph… is a semantic graph database focused on generating sophisticated semantic knowledge graphs, initially from your existing data. The graph database itself is an RDF-based quad store (in other words, a triple store where all the triples are named) with property graph support… The product's primary focus is on transactional processing; however, it is often used for analytics as well. Consequently, it is OLTP-enabled and fully ACID compliant, and additionally offers immediate consistency. The product is also highly secure, and supports the requirements for various*

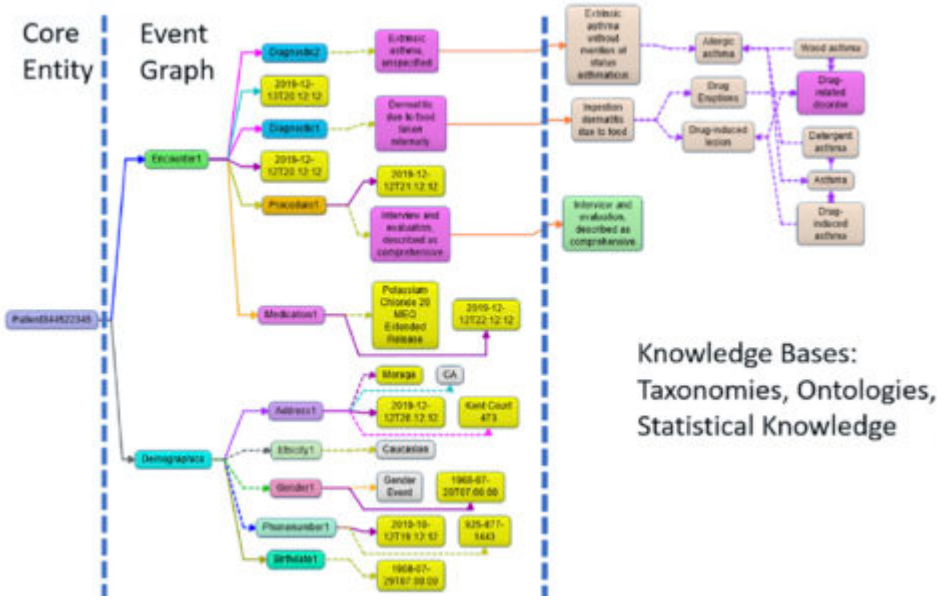
> *government security standards, including HIPAA…It includes a wealth of features, including distributed deployment and querying, multi-modal ingestion, multi-master replication, AI and machine learning, and natural language processing (NLP).”*

The full text of this report can be found here.

As for the version 7.0 release of AllegroGraph, arguably the most compelling new capability is its ability to create what Franz refers to as *“Entity-Event Knowledge Graphs”* (or EEKGs) via its patented FedShard technology, an example which can be seen below. These differ from regular graphs in that they are designed to capture a number of core entities (such as products, patients or customers) and any events (which may be time-stamped) relating to those entities within a hierarchical tree structure. Notice in the image below that each node – each event – *“branches”* into one or more additional events, and that these branches usually do not interact: this is characteristic of the tree model. Moreover, these events can terminate in a knowledge base, such as a taxonomy or an ontology. This allows you to bring a wealth of supporting information into your graph when that information may be relevant to your core entity (or entities). For example, suppose you have a core entity that is a hospital patient. In this case, you might use a publicly available catalogue of drug interactions as one of your knowledge bases. This all creates an intuitive way to visually represent your entities and the events related to them while bringing in outside knowledge where it is useful. In this context it is also worth noting that AllegroGraph has probabilistic capabilities (this is not a new feature) so that, for example, a physician can assign a probability to a diagnosis.

Notably, your EEKGs can be built incrementally, starting with a simple model and extending gradually and as needed, but without requiring you to actually alter any existing parts of your model. EEKGs also store provenance information for your

events, which captures where the initial data used to generate each of your events originally came from, and how its transformation into a graph object was achieved, thus providing data lineage.



The FedShard feature in the 7.0 release provides enhanced capabilities for horizontally distributed deployment, and EEKGs especially have been designed to take advantage of this. Other new features include improved JSON and JSON-LD document handling, Natural Language Processing (NLP) and speech recognition functionality. For the latter in particular, Franz has been able to leverage voice to text capabilities to extract conceptual meaning from real speech, then store that meaning in a graph and subsequently run analytics on it. This is exciting, because in effect it allows you to analyse recorded conversations. For organisations that collect and store a lot of such conversations – call centres, for instance – this could prove very useful.

Franz clearly considers this a major release for AllegroGraph. Certainly, the introduction of an explicit entity-event graph is not something I've seen before. The newly introduced text to speech capabilities also seem highly promising.