

# Harmonizing big data with an enterprise knowledge graph

Franz's CEO, Jans Aasman, recently wrote the following article for InfoWorld.



*In addition to streamlining how users retrieve diverse data via automation capabilities, a knowledge graph standardizes those data according to relevant business terms and models*

*One of the most significant results of the big data era is the broadening diversity of data types required to solidify data as an enterprise asset. The maturation of technologies addressing scale and speed has done little to decrease the difficulties associated with complexity, schema transformation and integration of data necessary for informed action.*

*The influence of cloud computing, mobile technologies, and distributed computing environments contribute to today's variegated IT landscape for big data. Conventional approaches to master data management and data lakes lack critical requirements to unite data—regardless of location—across the enterprise for singular control over multiple sources.*

*The enterprise knowledge graph concept directly addresses these limitations, heralding an evolutionary leap forward in big data management. It provides singular access for data across the enterprise in any form, harmonizes those data in a standardized format, and assists with the facilitation of action required to repeatedly leverage them for use cases spanning organizations and verticals.*

# AllegroGraph v6.4 – Now Available

Release 6.4.0 is a major release with significant new features.

The most important and far-reaching change is support for **multi-master replication**.

AllegroGraph has long supported single-master replication, where several AllegroGraph instances share data in a repository, but only one of them can make changes (adding or deleting triples).

In multi-master replication, even though one instance is identified as the controlling instance, any instance can add or delete triples, with the remainder catching up with those changes while perhaps making other changes of their own. Single-master replication is still supported and is described in the Replication document. The new multi-master replication facility is described in Multi-master Replication.

AllegroGraph Multi-master Replication is a real-time transactionally consistent data replication solution. It allows businesses to move and synchronize their semantic data across the enterprise. This facilitates real-time reporting, load balancing, and disaster recovery.

Single repositories can be replicated as desired. The replicas each run in an AllegroGraph server. A single server can serve multiple replicas of the same repository (this is not typical

for production work but might be common in testing). Note if there are multiple replicas in a single server, each replica must either be in a different catalog or must have a different name.

The collection of servers with replicas of a particular repository is called a *replication cluster* (or just *cluster* below in this document). Each repository in the cluster is called an *instance*. One instance is designated as the *controlling instance*, which will be described in more details below.

Each instance in the cluster can add or delete triples and these additions and deletions are passed to all other instances in the cluster. How long it takes for instances to synchronize depends on factors external to AllegroGraph (such as network availability and speed and whether the other servers are even available) but given time and assuming all instances are accessible, after a period of no activity (no additions or deletions) all instances will become synchronized.

---

## Gruff v7.0 – Time Machine Tutorial

Here is an example for trying out the new time slider in Gruff's graph view. It uses triples from crunchbase.com that contain a history of corporate acquisitions and funding events over several years. Gruff's time bar allows you to examine those events chronologically, and also to display only the nodes that have events within a specified date range.

\* Download the Crunchbase triples from the bottom of the Gruff

download page at <https://franz.com/agraph/gruff/download/>.

- \* Create a new triple-store and used “File | Load Triples | Load N-Triples” to load that triples file into the new triple-store. Use “File | Commit” to ensure that the loaded triples get saved.

- \* Select “Visual Graph Options | Time Bar | Momentary Time Predicates” and paste the following five predicate IRIs into the dialog that appears. The time bar will then work with the date properties that are provided by these predicates, whenever you are browsing this particular triple-store.

[http://www.franz.com/hasfunded\\_at](http://www.franz.com/hasfunded_at)  
[http://www.franz.com/hasfirst\\_funding\\_at](http://www.franz.com/hasfirst_funding_at)  
[http://www.franz.com/hasfounded\\_at](http://www.franz.com/hasfounded_at)  
[http://www.franz.com/haslast\\_funding\\_at](http://www.franz.com/haslast_funding_at)  
[http://www.franz.com/hasacquired\\_at](http://www.franz.com/hasacquired_at)

- \* Select “View | Optional Graph View Panes | Show Time Bar” to reveal the time bar at the bottom of the graph view. The keyboard shortcut for this command is Shift+A to allow quickly toggling the time bar on and off.

- \* Select “Display | Display Some Sample Triples” to do just that. The time bar will now display a vertical line for each of the requested date properties of the displayed nodes. Moving the mouse cursor

over these “date property markers” will display more information about those events.

\* Click down on the yellow-orange rectangle at the right end of the time bar and drag it to the left. This will make the “time filter range” smaller, and nodes that have date properties that are no longer in this range will temporarily disappear from the display. They will reappear if you drag the slider back to the right or toggle the time bar back off.

For more information, the full time bar introduction is in the Gruff documentation under the command “View | Optional Graph View Panes | Show Time Bar”.

Check out the “Chart Widget” for showing date properties of the visible nodes.

---

## **Semantic Computing, Predictive Analytics Need Reliable Metadata**



**Our Healthcare Partners at Montefiore were interviewed at Health Analytics:**

*Reliable metadata is the key to leveraging semantic computing and predictive analytics for healthcare applications, such as population health management and crisis care.*

*As the healthcare industry reaches the saturation point of electronic health record adoption, and slowly moves past the pain of the implementation process, it may seem like the right time to stop thinking so much about hammering home basic data governance principles for staff members and start looking at the next phase of health IT implementation: the big data analytics environment.*

*After all, most providers are now sitting on an enormous nest egg of patient data, which may be just clean, complete, and standardized enough to start experimenting with population health management, operational analytics, or even a bit of predictive risk stratification. Many healthcare organizations are experimenting with these advanced analytics projects in an effort to prepare themselves for the financial storm that is approaching with the advent of value-based care.*

*The immense pressure to cut costs, meet quality benchmarks, shoulder financial risk, and improve patient outcomes is causing no small degree of anxiety for providers, who are racing to batten down the hatches before the typhoon overtakes them.*

*While it may be tempting to jump into quick-win analytics that use “good enough” datasets to solve a specific pressing use case, providers may be at risk of repeating the same mistakes they made with slapdash EHR implementations: creating data siloes, orphaned reports, and poor quality datasets that cannot be used in a reliable, repeatable way for meaningful quality improvements.*

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

---

# New York Times Article – Is There a Smarter Path to Artificial Intelligence?

From the New York Times – June 20, 2018

**This article caught our attention because they featured a startup that was using Prolog for AI. We have been strong proponents of Prolog for Semantic Graph solutions for many years.**

*For the past five years, the hottest thing in artificial intelligence has been a branch known as deep learning. The grandly named statistical technique, put simply, gives computers a way to learn by processing vast amounts of data. Thanks to deep learning, computers can easily identify faces and recognize spoken words, making other forms of humanlike intelligence suddenly seem within reach.*

*Companies like Google, Facebook and Microsoft have poured money into deep learning. Start-ups pursuing everything from cancer cures to back-office automation trumpet their deep learning expertise. And the technology's perception and pattern-matching abilities are being applied to improve progress in fields such as drug discovery and self-driving cars.*

*But now some scientists are asking whether deep learning is really so deep after all.....*

*.....Those other, non-deep learning tools are often old*

*techniques employed in new ways. At Kyndi, a Silicon Valley start-up, computer scientists are writing code in Prolog, a programming language that dates to the 1970s. It was designed for the reasoning and knowledge representation side of A.I., which processes facts and concepts, and tries to complete tasks that are not always well defined. Deep learning comes from the statistical side of A.I. known as machine learning.*

**Our Tweet with links to AllegroGraph Prolog documenation and the full article:**

**nytimestech** “computer scientists are writing code in **#Prolog**... It was designed for the reasoning and knowledge representation side of **#AI** ...” <https://buff.ly/2lmYwkv> – **#AllegroGraph** is the only **#GraphDatabase** to include **#Prolog** for your AI apps. <https://buff.ly/2yv0IzF>