

5 Graph Databases to Consider

Of the major categories of NoSQL databases – document-oriented databases, key-value stores and graph databases – we've given the least attention to graph databases on this blog. That's a shame, because as many have pointed out it may become the most significant category.

Graph databases apply [graph theory](#) to the storage of information about the relationships between entries. The relationships between people in social networks is the most obvious example. The relationships between items and attributes in recommendation engines is another. Yes, it has been noted by many that it's ironic that relational databases aren't good for storing relationship data. Adam Wiggins from Heroku has a lucid explanation of why that is [here](#). Short version: among other things, relationship queries in RDBSes can be complex, slow and unpredictable. Since graph databases are designed for this sort of thing, the queries are more reliable.

Google has its own graph computing system called [Pregel](#) (you can find the paper on the subject [here](#)), but there are several commercial and open source graph databases available. Let's look at a few.

Neo4j

This is one of the most popular databases in the category, and one of the only open source options. It's the product of the company Neo Technologies, which recently moved the community edition of Neo4j from the AGPL license to the GPL license (see our coverage [here](#)). However, its enterprise edition is still proprietary AGPL. Neo4j is ACID compliant. It's Java based but has bindings for other languages, including Ruby and Python.

Neo Technologies cites several customers, though none of them

are household names.

Here's a fun illustration of how relationship data in graph databases works, from an [InfoQ article](#) by Neo Technologies COO Peter Neubauer:

FlockDB

FlockDB was created by Twitter for relationship related analytics. Twitter's Kevin Weil talked about the creation of the database, along with Twitter's use of other NoSQL databases, at [Strange Loop](#) last year. You can find our coverage [here](#).

There is no stable release of FlockDB, and there's some controversy as to whether it can be truly referred to as a graph database. In [a DevWebPro article](#) Michael Marr wrote: This lead [MyNoSQL](#) blogger Alex Popescu to write: "Without traversals it is only a persisted graph. But not a graph database."

The biggest difference between FlockDB and other graph databases like Neo4j and OrientDB is graph traversal. Twitter's model has no need for traversing the social graph. Instead, Twitter is only concerned about the direct edges (relationships) on a given node (account). For example, Twitter doesn't want to know who follows a person you follow. Instead, it is only interested in the people you follow. By trimming off graph traversal functions, FlockDB is able to allocate resources elsewhere.

However, because it's in use at one of the largest sites in the world, and because it may be simpler than other graph DBs, it's worth a look.

AllegroGraph

[AllegroGraph](#) is a graph database built around the W3C spec for the Resource Description Framework. It's designed for handling [Linked Data](#) and the [Semantic Web](#), subjects we've written about often. It supports SPARQL, RDFS++, and Prolog.

AllegroGraph is a proprietary product of [Franz Inc.](#), which markets a number of Semantic Web products – including its flagship set of LISP-based development tools. The company [claims](#) Pfizer, Ford, Kodak, NASA and the Department of Defense among its AllegroGraph customers.

GraphDB

GraphDB is graph database built in .NET by the German company [sones](#). sones was founded in 2007 and [received a new round of funding](#) earlier this year, said to be a “couple million” Euros. The community edition is available under an APL 2 license, while the enterprise edition is commercial and proprietary. It's available as a cloud-service through Amazon S3 or Microsoft Azure.

InfiniteGraph

[InfiniteGraph](#)

is a proprietary graph database from Objectivity, the company behind the object database of the same name. Its goal is to create a graph database with “virtually unlimited scalability.”

According to [Gavin Clarke at The Register](#): “InfiniteGraph map is already being used by the CIA and Department of Defense running on top of the existing Objectivity/DB database and analysis engine.”