The Road to AmiGO 2

Seth Carbon

Berkeley BOP (http://berkeleybop.org),
Lawrence Berkeley National Lab

04 October 2012

Away from MySQL

We have been using an increasing number of tricks, extensions, and caches to keep the performance at an acceptable level. Things have to change...

- Complicated queries: enrichment, subsets, search, reports, etc.
- Data: ~1,500,000 -> ~13,000,000 -> ~80,000,000 -> ???
- Providing services

Into Solr

Introducing Solr

- A specialized HTTP server over the Lucene document store

- A common document-oriented approach (Lucene; similar to QuickGO)

- All of your data with just a URL

Why You Might Care

AmiGO 2 has greatly increased in flexibility, speed, and development turnaround time over SQL.

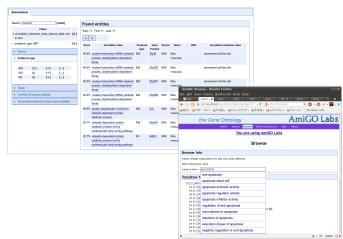For example: a deep text search from ~30s. down to ~0.3s.

It has also made things that were previously not feasible easy to do. For example:

- Very fast searches, even for previously impossible topics
- Widgets for search, graphs, etc.
- Easier direct calculations (counts, information content)
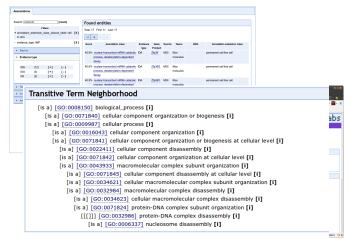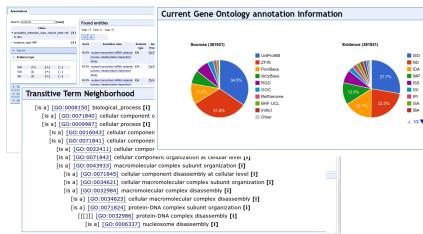- Easy third-party use (data and clients)

AmiGO 2

Looking good!

AmiGO 2

Looking good!

AmiGO 2

Looking good!

AmiGO 2

Looking good!

Secondary Tools

- Term Enrichment and Mapper/Slimmer
  - Being moved to Galaxy
- The Birds
  - GOOSE (SQL environment) not going anywhere for a while
  - Gannet (Solr environment) now included in AmiGO 2.
- New reporting tools
  - Scripting and remote use now possible

What We Have to Do

- Finishing up. . .

- Figuring out deployment

- Education and beta

Information and Demo

AmiGO 2 wiki and demo site