A graph-based database of hierarchical brain features

Noah Lee and **Arno Klein**

Columbia University, New York, NY

incf Neuro Informatics 2011

Sept. 2011, Boston, USA

Summary

We are building a publicly available graph-based database of anatomic brain structures, quantitative characterization of their geometric, shape, and spectral features, and relationships among structures and label regions. We have selected a noSQL (not only SQL) graph database as opposed to a standard relational database for the following reasons:

(1) Our data model should be flexible enough to accommodate changing data and knowledge representations.

(2) We want to naturally represent and store connected and semi-structured data with a rich relationship model.

(3) It must support large-scale big data storage, efficient querying, and deep inferences by exploiting the power of graph algorithms to perform comparative feature analysis in populations.

We present the first prototype of our interactive Mindboggle-db web interface together with its infrastructure.

Our graph-based database enables us to build a rich and flexible relationship model of hierarchical brain structures, their features, and label regions. The web-based interface makes our data model available to a wide audience without any software installations or plugins. Ongoing and future work is devoted to the integration of other publicly available datasets and the interfacing with existing neuroinformatics platforms.

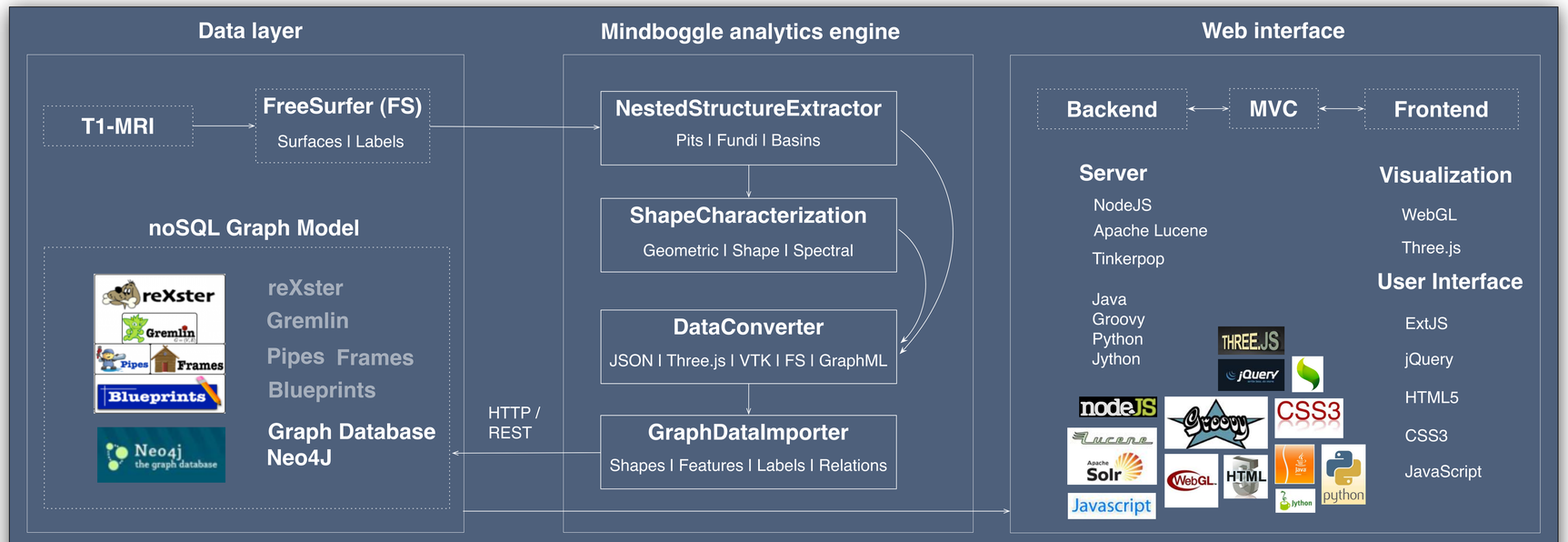


Fig. 1. The Mindboggle graph-based database infrastructure showing the individual components of the data layer, the analytics engine, and the web interface.

noSQL graph model of the Mindboggle database

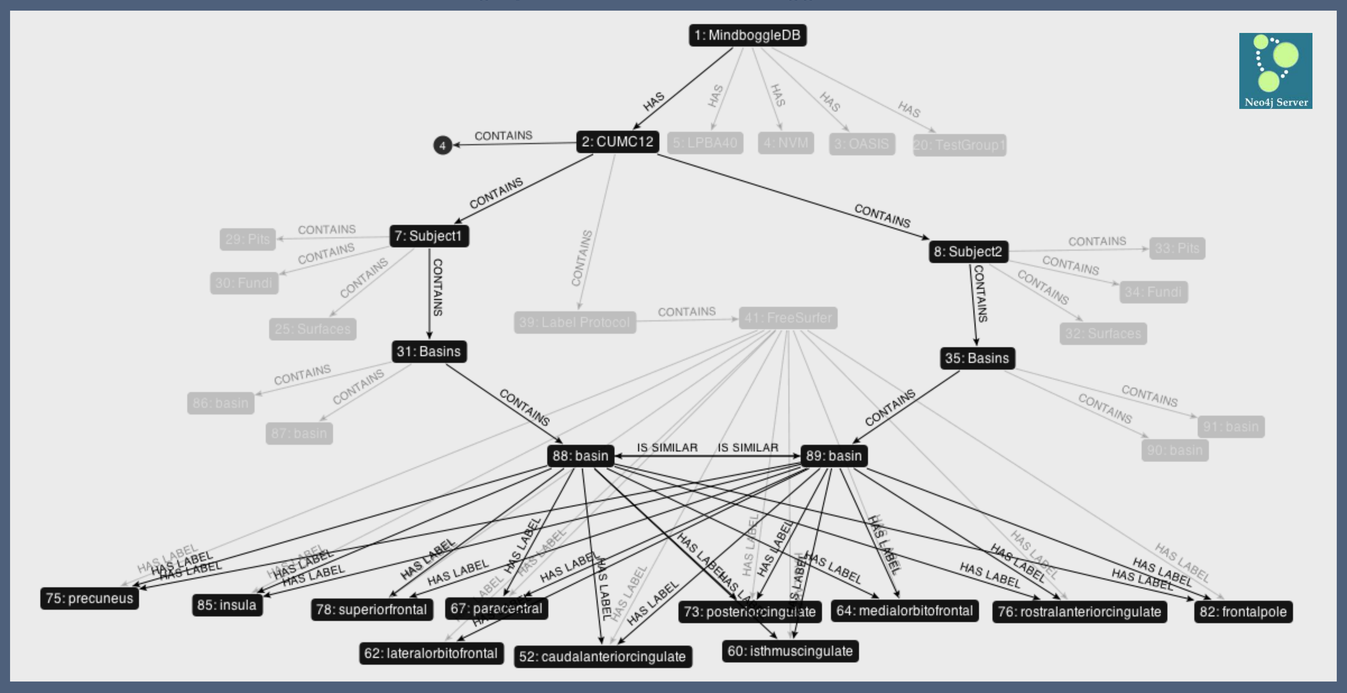
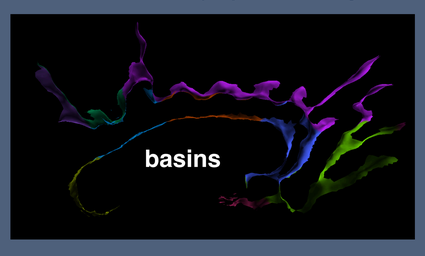


Fig. 2. The Mindboggle[1,2] graph-based data model exemplified on the CUMC12 dataset[3]. The graph model shows an example of how relations between basins and their labeled regions are represented for two subjects. Note that basin 88 and 89 (also shown below in Fig. 3) share similar labels.
Structure similarity can be computed based on geometric, shape, and spectral features.

Basin structure (88) from subject 1



Basin structure (89) from subject 2

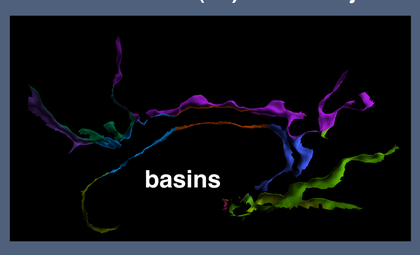


Fig. 3. Extracted basin structures using the NestedStructureExtractor. The different colors correspond to the label names in Fig. 2.

How to query the graph model?

Gremlin scripts

Method: POST http://192.168.23.219:7474/db/data/ext/GremlinPlugin/graphdb/execute_script Send

```
{
"script": "g.v(88).outE"
}
```

JSON response

```
[{
 "start" : "http://192.168.23.219:7474/db/data/node/88",
 "data" : {
 },
 "type" : "IS SIMILAR",
 "extensions" : {
 },
 "end" : "http://192.168.23.219:7474/db/data/node/89"
}, {
 "start" : "http://192.168.23.219:7474/db/data/node/88",
 "data" : {
 },
  "type" : "HAS LABEL",
 "extensions" : {
 },
 "end" : "http://192.168.23.219:7474/db/data/node/60"
},
```

Fig. 4. The graph model can be queried via REST, Cypher, reXster, or Gremlin. This example shows how to query all outgoing relations from basin 88.
Much more powerful graph traversal algorithms are possible.

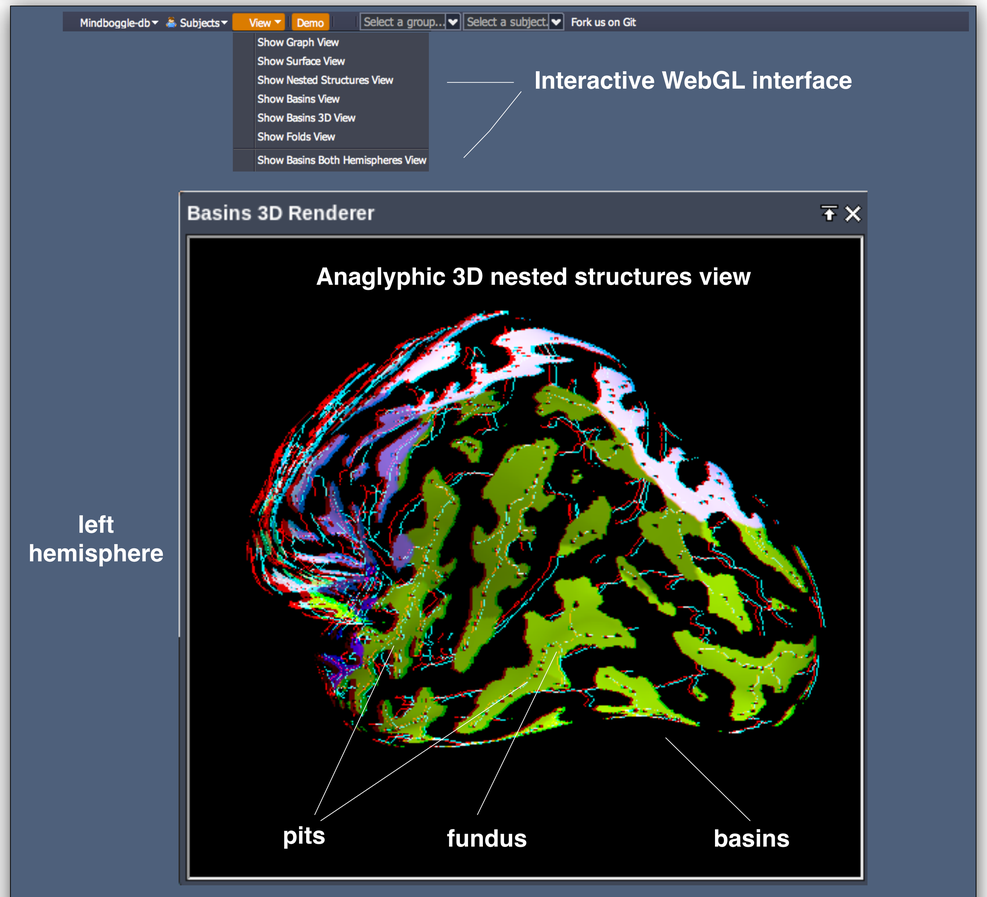


Fig. 5. The interactive WebGL interface provides a flexible way to visualize and explore the database.

[1] Klein, A, Hirsch, J. 2005. Mindboggle: A scatterbrained approach to automate brain labeling. *NeuroImage.* 24(2): 261-280. PMID: 15627570 (http://www.mindboggle.info)
[2] Klein, A, Mensh, B, Ghosh, S, Tourville, J, Hirsch, J. 2005. Mindboggle: Automated brain labeling with multiple atlases. *BMC Medical Imaging.* 5:7. PMCID: PMC1283974
[3] Klein, A., et al. 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage,* 46(3):786–802. PMCID: PMC2747506

This work was funded by NIMH R01 grant MH084029.