Brainomics: A management system for exploring and merging heterogeneous brain mapping data
Vincent Michel, Yannick Schwartz, Philippe Pinel, Olivier Cayrol, Antonio Moreno, Jean-Baptiste Poline, Vincent Frouin, Dimitri Papadopoulos Orfanos

To cite this version:

HAL Id:cea-00904768
https://hal-cea.archives-ouvertes.fr/cea-00904768
Submitted on 15 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Brainomics - A management system for exploring and merging heterogeneous brain mapping data

V. Michel¹, Y. Schwartz², P. Pinel³, O. Cayrol¹, A. Moreno³, JB. Poline⁴, V. Frouin⁴, D. Papadopoulos Orfanos⁴

¹Logilab, Paris, France, ²INRIA Saclay Parietal Team, Paris, France, ³Unicog, INSERM-CEA, Neurospin, Gif-sur-Yvette, France, ⁴CEA, Neurospin, Gif-sur-Yvette, France

Introduction

- Number of large datasets for brain mapping have been released [1, 2].
- Neuroimaging datasets more routinely include clinical data or genetics data.  
  - Exploitation requires
    - An efficient organization to integrate all the measures
    - An easy access to the relevant information.

Conclusion

- Open source solution to manage brain imaging datasets and associated meta data.
- Powerful querying and reporting tool, customized for emerging imaging-genetics field.

This work was supported by grants from the French National Research Agency (ANR GENIM; ANR-10-BLAN-0128) and (ANR IA BRAINOMICS; ANR-10-BINF-04).

http://www.brainomics.net/demo/

- Brings together brain imaging and genetics data.
- Relies on a high-level query language (RQL).
- Solution based on CubicWeb, a semantic framework.
- Visualizing / exporting data in several formats.

Data model

- Described in Python, using reusable modules called “cubes”.
- Modelisation of Scans, Questionnaires, Genomics results, Behavioral results, Subjects and Studies information.
- Data model optimized for large volumes (> 2000 subjects).
- Tested with several publicly available datasets [1, 2].
- Well established core technologies: SQL, Python, HTML5, Javascript.
- Licensed under LGPL since 2008.
- Used in production environments since 2005.
- Fine-grained security system coupled to the data model definition.
- Migration mechanisms controls model version / ensures data integrity.

Query using RQL

- Similar to the W3C’s SPARQL [6].
- Supports the basic operations (selection, insertion, etc.).
- Subquerying, ordering, couting, ...

Query all the female subjects of the database, with an age greater than 30

Any S WHERE S is Subject, S age > 30, S gender "female"

Query all the Cmap scans of subjects with an age greater than 25, and that have a score greater than 4.0 for the “algebre” question

Any SA WHERE S is Subject, S age > 25, X is QuestionnaireRun, X concerns S, A is Answer, A questionnaire_run X, A question Q, Q text "algebre", A value > 4, SA is Scan, SA concerns S, SA type "c map"

Views

- Each query result can be seen using different views.
- HTML pages, ZIP files, spreadsheets, XCEDE XML, ...
- May include processing (stat. maps computed on the fly).