The Brede Wiki

A social neuroinformatics web-service with structured information from neuroscience

Finn Årup Nielsen

The Lundbeck Foundation Center for Integrated Molecular Brain Imaging;
Department of Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark;
Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark.

Summary

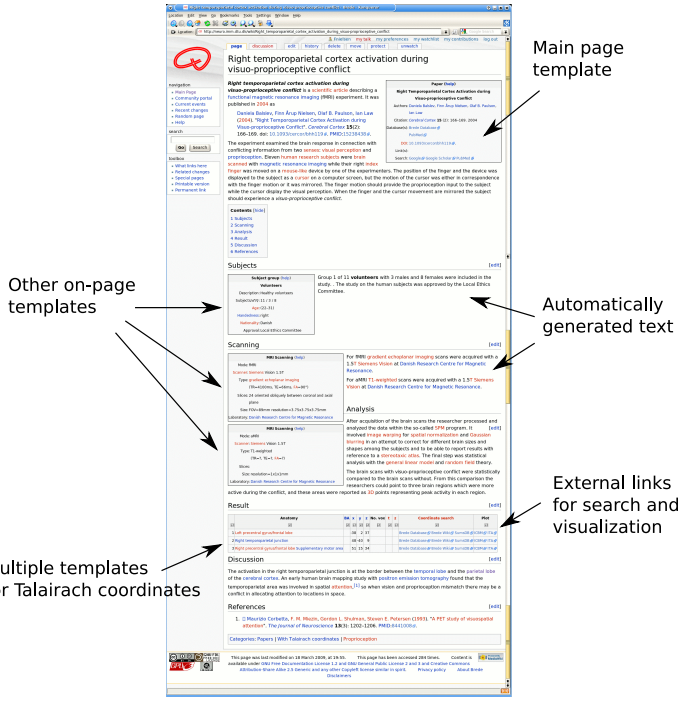
The Brede Wiki running on MediaWiki software represents data in templates. Data is from published peer-reviewed neuroscience articles. Further information is ontologies of, e.g., brain regions and brain functions. Since data in the templates is represented in a simple format all template data can be extracted and represented in SQL. From an SQL database specialized searches can be performed, e.g., search on Talairach coordinates.

Background

Several databases exist for storing Talairach coordinates,[1] but contributing data is not straightforward. For the Brede Database[2] contributors need to download a Matlab program,[3] understand it, enter data and submit it to the database curator which then adds it to the database. A wiki presents a much more direct way of editing. Wikis are often unstructured. However, some wikis, such as semantic wikis, are capable of labeling its content so data can be 'structured' and understood by a computer. Using the template functionality in MediaWiki—the Wikipedia software—is one simple approach for a structured wiki.

Data from MediaWiki templates have been extracted from Wikipedia and the large-scale DBpedia Web-based database have been built from the data.[4] Wikipedia templates can form the basis for descriptive as well as multivariate statistical analyses.[5,6] The Brede Wiki uses the MediaWiki templates to structure data.

Example page from the Brede Wiki



Brede Wiki templates

In the Brede Wiki the templates of MediaWiki store structured information. Template definitions construct infoboxes and some constructs schematic natural language.

Information in the templates includes ontologies for brain function (cognitive components), brain regions, organizations, events and software. Further templates describe researchers and papers. For result data `Talairach coordinate`, `Gene personality association` and `Brain volume` templates are defined.

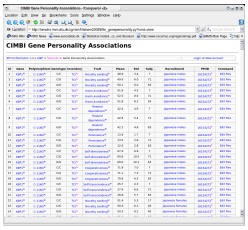
Database links

Paper, brain regions, cognitive components and researchers are linked to corresponding items in the Brede Database. Items in the Brede Wiki may also link to PubMed and other NLM services, Publisher's fulltext articles, Wikipedia, BrainInfo (NeuroNames), SumsDB, NeuroLex, The Internet Brain Volume Database (IBVD), SumsDB as well as Internet portals Neuroscience Information Framework (Neurogateway), I Do Imaging, Internet Analysis Tools Registry (IATR), Neuroscience Database Gateway (NDG).

The Brede Wiki may construct links for Web-based services such that individual Talairach coordinates can be searched and visualized, e.g., with the Brede Database, SumsDB or ICBM View. Text searches are provides to DOAJ, Open J-Gate and Google.

Data entry

Apart from standard wiki input Brede Wiki template coordinate data can be formated from an SPM extension. Another of our structured wikis (see figure) can output its personality genetics data in template format ready for inclusion in the Brede Wiki.



Personality genetics wiki.[7]

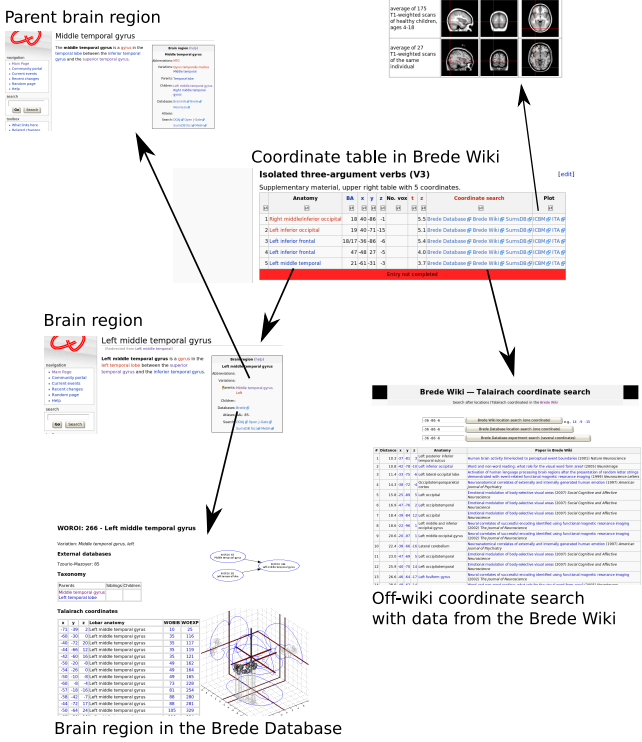
Extraction of data

Extraction of template data from Wikipedia requires extra processing due to wiki markup and variations in data entry.[5,8] Extraction from templates in the Brede Wiki is simpler since the template are non-nesting, they do not contain wiki markup and they have only lower-case names. This means the all data from the templates can be extracted and represented directly in an SQL database. Two types of SQL tables are built from the template data:[9] A master table with information from all tables, and secondary tables which only contain data from each specific template.

The extracted data are made available on the Web as SQL and SQLite files from the Brede Wiki homepage. Also MediaWiki dumps and ontologies in the SKOS format[10] are downloadable.

A Web script uses the data from the SQLite file enabling search for nearby coordinates.[11,12]

Automated linking and querying for coordinate data



Off-wiki coordinate search with data from the Brede Wiki

Brain region in the Brede Database

Data

The Brede Wiki currently holds around 700 pages. Of these pages are 115 for brain regions, 84 for papers with 47 containing Talairach coordinates, 66 for researchers, 9 for organization and 25 for software descriptions. These numbers could increase if data from other databases were automatically added to the wiki with programs such as the bots on Wikipedia that have built, e.g., the 'Gene Wiki'.[13]

Acknowledgment

Thanks to the Lundbeck Foundation for funding and Lars Kai Hansen and Daniela Balslev for discussions.

References

[1] Derrfuss J and Mar RA. Lost in localization: The need for a universal coordinate database. NeuroImage, doi:10.1016/j.neuroimage.2009.01.053, 2009.

[2] Nielsen FÅ. The Brede database: a small database for functional neuroimaging. *NeuroImage*, 2003;19. Presented at the 9th International Conference on Functional Mapping of the Human Brain, June 19–22, 2003, New York, NY. Available on CD-Rom.

[3] Nielsen FÅ and Hansen LK. Experiences with Matlab and VRML in functional neuroimaging visualizations. In Klasky S and Thorpe S, eds., *VDE2000 - Visualization Development Environments, Workshop Proceedings, Princeton, New Jersey, USA, April 27–28, 2000*. Princeton Plasma Physics Laboratory, Princeton, New Jersey, 2000; pages 76–81.

[4] Auer S, et al. DBpedia: A nucleus for a web of open data. In *The Semantic Web*, vol. 4825 of *Lecture Notes in Computer Science*. Springer, Heidelberg/Berlin, 2008; pages 722–735.

[5] Nielsen FÅ. Scientific citations in *Wikipedia*. *First Monday*, 2007;12.

[6] Nielsen FÅ. Clustering of scientific citations in Wikipedia. In *Wikimania*. 2008;

[7] Nielsen FÅ. A small wiki for personality genetics. 37th Annual Meeting on Biochemistry and Molecular Biology: Frontiers in Genomics, 2008.

[8] Auer S and Lehmann J. What have Innsbruck and Leipzig in common? extracting semantics from Wiki content. In *The Semantic Web: Research and Applications*, vol. 4519 of *Lecture Notes in Computer Science*. Springer, Berlin/Heidelberg, 2007; pages 503–517.

[9] Nielsen FÅ. Brede Wiki: Neuroscience data structured in a wiki. In Lange C, ed., *Proceedings of the Fourth Workshop on Semantic Wikis — The Semantic Wiki Web*. CEUR-WS.org, 2009; pages 129–133.

[10] Miles A and Bechhofer S. SKOS Simple Knowledge Organization System Reference. W3c candidate recommendation, W3C, MIT, 2009.

[11] Szewczyk MM. *Databases for neuroscience*. Master's thesis, Technical University of Denmark, Kongens Lyngby, Denmark, 2008. IMM-MSC-2008-92.

[12] Wilkowski B, et al. Coordinate-based meta-analytic search for the SPM neuroimaging pipeline. In *Proceedings of the Second International Conference on Health Informatics*. INSTICC Press, 2009; pages 11–17.

[13] Huss, III JW, et al. A gene wiki for community annotation of gene function. *PLoS Biology*, 2008;6:e175.