Databasing the Brain — Neuroinformatics

Finn Årup Nielsen

Neurobiology Research Unit
Copenhagen University Hospital Rigshospitalet
og
Informatics and Mathematical Modelling
Technical University of Denmark

October 24, 2005

Why database?

Bring order to data: Organize data for the individual study or for a whole range of studies.

Make search easy: PubMed and Google are examples on easy search and retrieval on text. They fail to search on specific neuroscience data, e.g., activation in basal ganglia.

Automate analysis: E.g. construct consensus across studies; compare a new study to the existing body of work.

Develop new tools: Neuroscience makes interesting heterogeneous data which enforce development of new tools.

Information increase

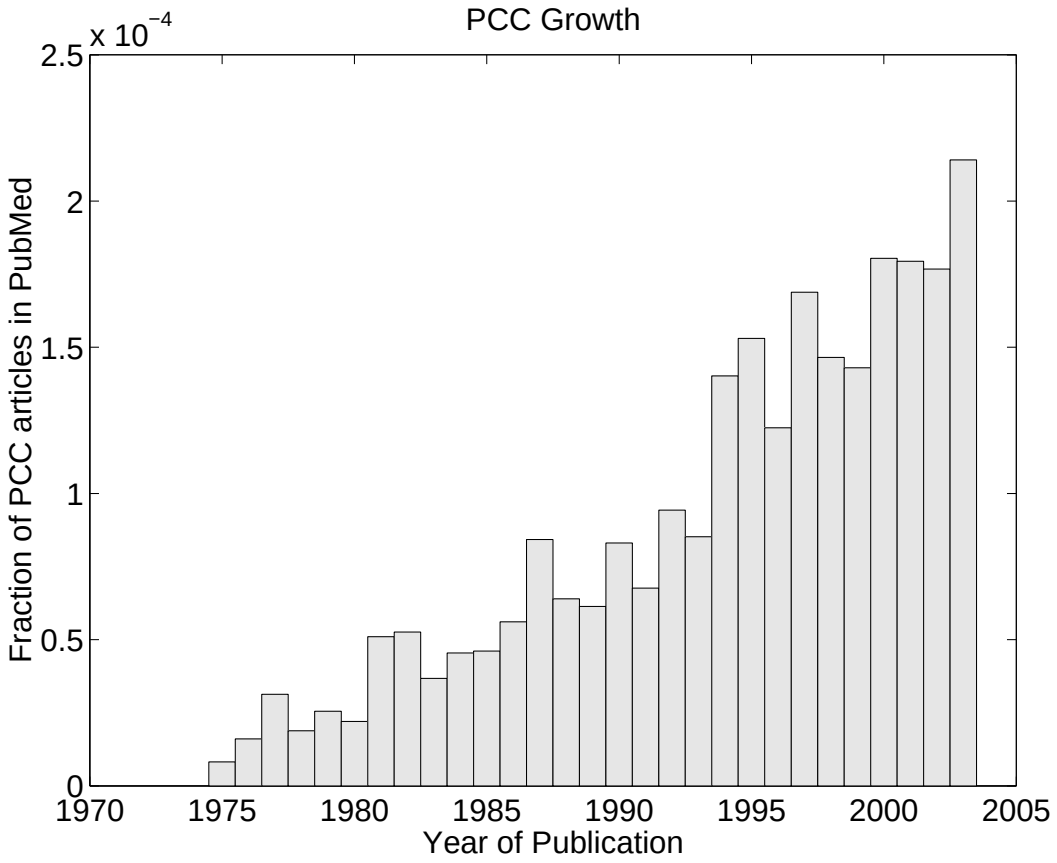


Figure 1: Increase in the number of articles in PubMed which are returned after searching on "Posterior cingulate".

The number of articles increases.

Can databases and computer-based methods help to oganize the large amount of new data?

How should data be represented? How can they be entered into a database? Which data mining methods can be developed? Internet services like bioinformatics?

Functional human brain mapping

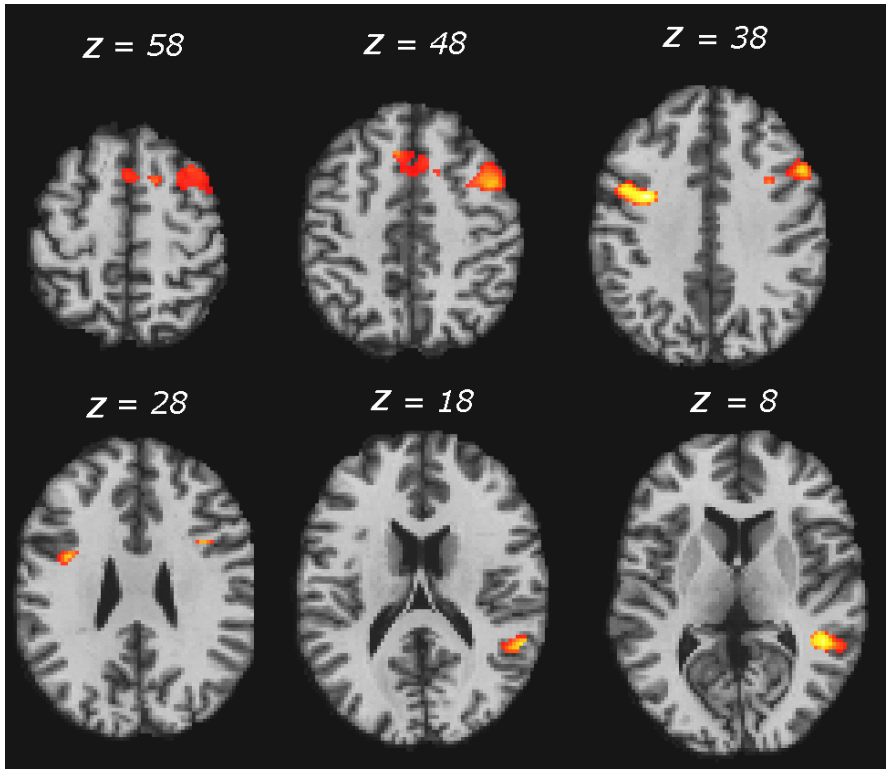


Figure 2: Figure from Balslev et al. (2005).

"Activation studies" or patient-control comparisons with PET, fMRI or SPECT. Lesionsstudies with MRI.

Results often represented in the literature as 3-dimensional coordinates wrt. a standardized stereotaxic system ("Talairach")

(x, y, z)	z-score
$-38, 0, 40$	4.91
$48, -42, 8$	4.66
$52, 14, 38$	4.07

BrainMap database

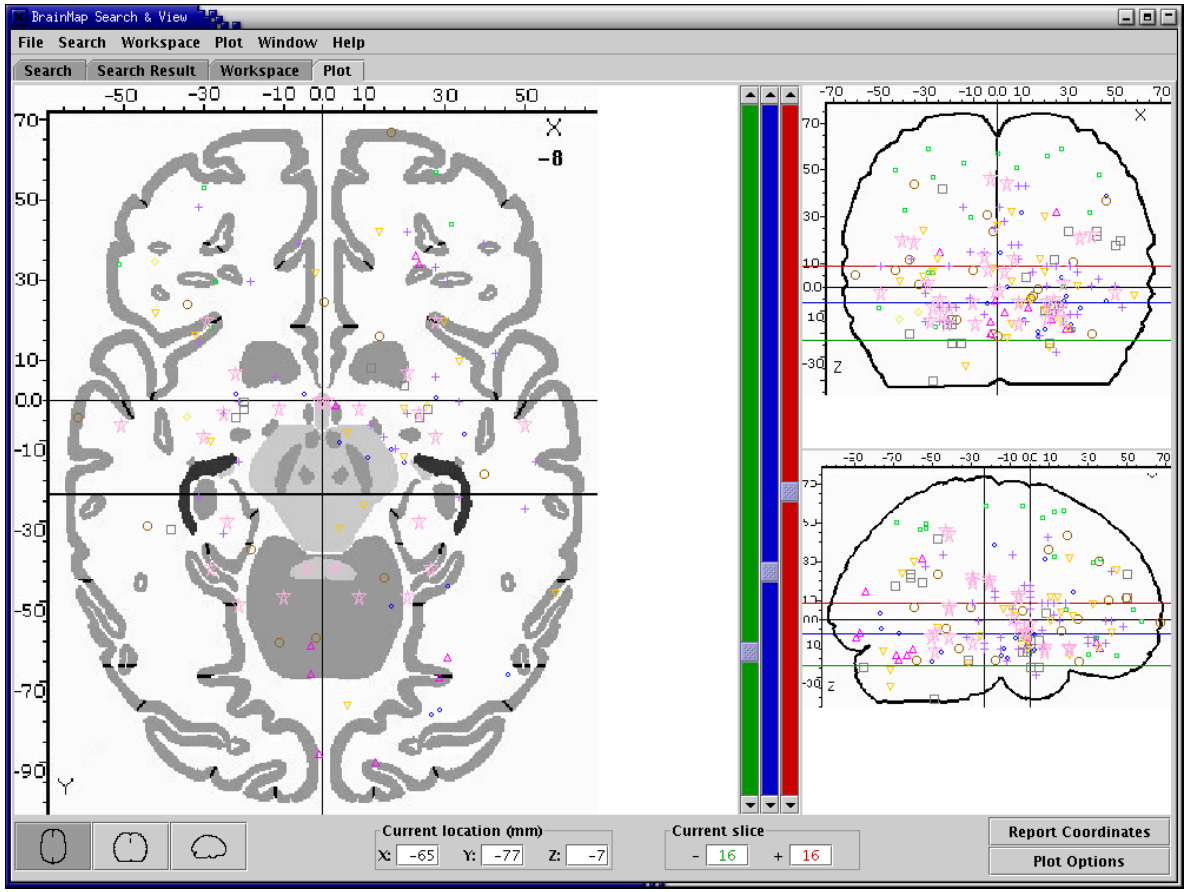


Figure 3: Screen shot of a graphical user interface to the BrainMap database with Talairach coordinates plotted after a search for experiments on olfaction.

One of the first and most comprehensive databases (Fox et al., 1994; Fox and Lancaster, 2002)

Presently 26678 locations from 765 papers

Graphical web-interface with search facilities, e.g., on author, 3D coordinate, . . .

Also possible to submit new studies

Brede Database

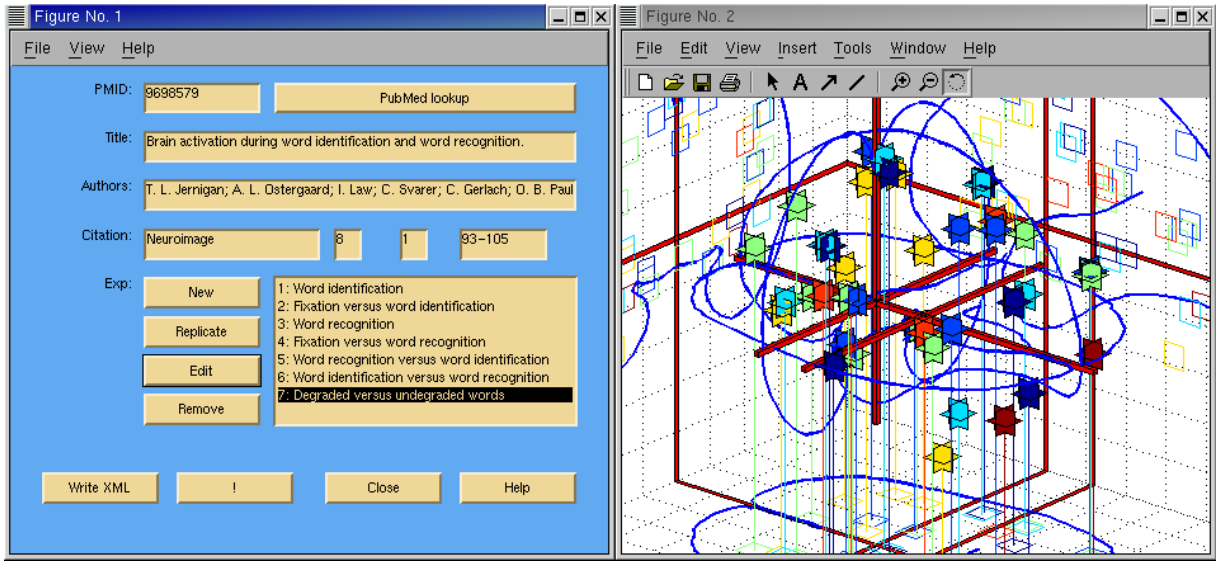


Figure 4: Screenshot of a program for entering data. Here with a study of Jernigan et al. (1998).

Smaller Brede Database similiar to BrainMap

Every studie saves, e.g., author, article title, abstract, scanner, number of subjects, coordinates, anatomical names, topic under study.

Taxonomy for brain regions and topics

XML “Lowtech” storage

```
...
<brainTemplate>SPM95</brainTemplate>
<behavioralDomain>Motion,Execution - Saccades</behavioralDomain>
<woext>57</woext>
<analysisSoftware>SPM95</analysisSoftware>
<analysisSoftware>AIR</analysisSoftware>
<analysisSoftware>AMIR</analysisSoftware>
<Loc>
  <type>loc</type>
  <functionalArea>Left frontal eye field</functionalArea>
  <brodmann></brodmann>
  <zScore>4.82</zScore>
  <coordReported>-0.050000 -0.002000 0.036000</coordReported>
...
```

Searching on Talairach coordinate

14 14 9 Location search e.g., 14 -9 -15

#	Distance	x	y	z	WOBIB	Description
1	0.5	13	14	8	76	Right caudate nucleus - Correlated with pain ratings in hot pain on right hand in rest, mental imagery and hypnosis (WOEXP: 238)
2	3.5	15	11	6	92	Caudate nucleus - Mildly depressed cancer patients (WOEXP: 293)
3	4.9	15	12	13	180	Right caudate nucleus - Intelligence and gray matter volume negative correlation (WOEXP: 564)
4	7.0	12	8	12	178	Right caudate - Semantic versus case (WOEXP: 550)
5	7.5	10	8	11	178	Right caudate - Semantic versus syllable counting via case judgment (WOEXP: 558)
6	8.1	5	14	8	76	Right caudate nucleus - Hot pain on right hand during hypnosis (WOEXP: 235)
7	8.3	13	7	3	171	Right striatum - Rhyme judgement and nonlinear ('convex') response in rapid auditory processing (WOEXP: 526)
8	8.3	13	7	3	171	Right striatum - Rhyme judgement and linear increase response in rapid auditory processing (WOEXP: 527)

Result after search for nearest coordinates to (14, 14, 9). Similar searches possible in xBrain and Antonia Hamilton's AMAT programs.

Seaching on experiments

+1: 1.00000 **Mentalizing versus rule solving**. *Playing a computer-based version of "stone, paper, scissor" while believing the opponent was an other human versus playing while believing the opponent was a computer with a fixed rule-based algorithm.* WOEXP: 218.
Helen L. Gallagher; Anthony I. Jack; Andreas Roepstorff; Christopher D. Frith. *Imaging the intentional stance in a competitive game.* *NeuroImage* **16**(3 Pt 1):814-21, 2002. PMID: 12169265. WOBIB: 70.

+2: 0.68676 **Posttraumatic stress disorder**. *Benzodiazepine binding in posttraumatic stress disorder versus binding in normal subjects.* WOEXP: 206.
J. D. Bremner; R. B. Innis; S. M. Southwick; L. Staib; S. Zoghbi; D. S. Charney. *Decreased benzodiazepine receptor binding in prefrontal cortex in combat-related posttraumatic stress disorder.* *American Journal of Psychiatry* **157**(7):1120-1126, 2000. PMID: 10873921. WOBIB: 67.

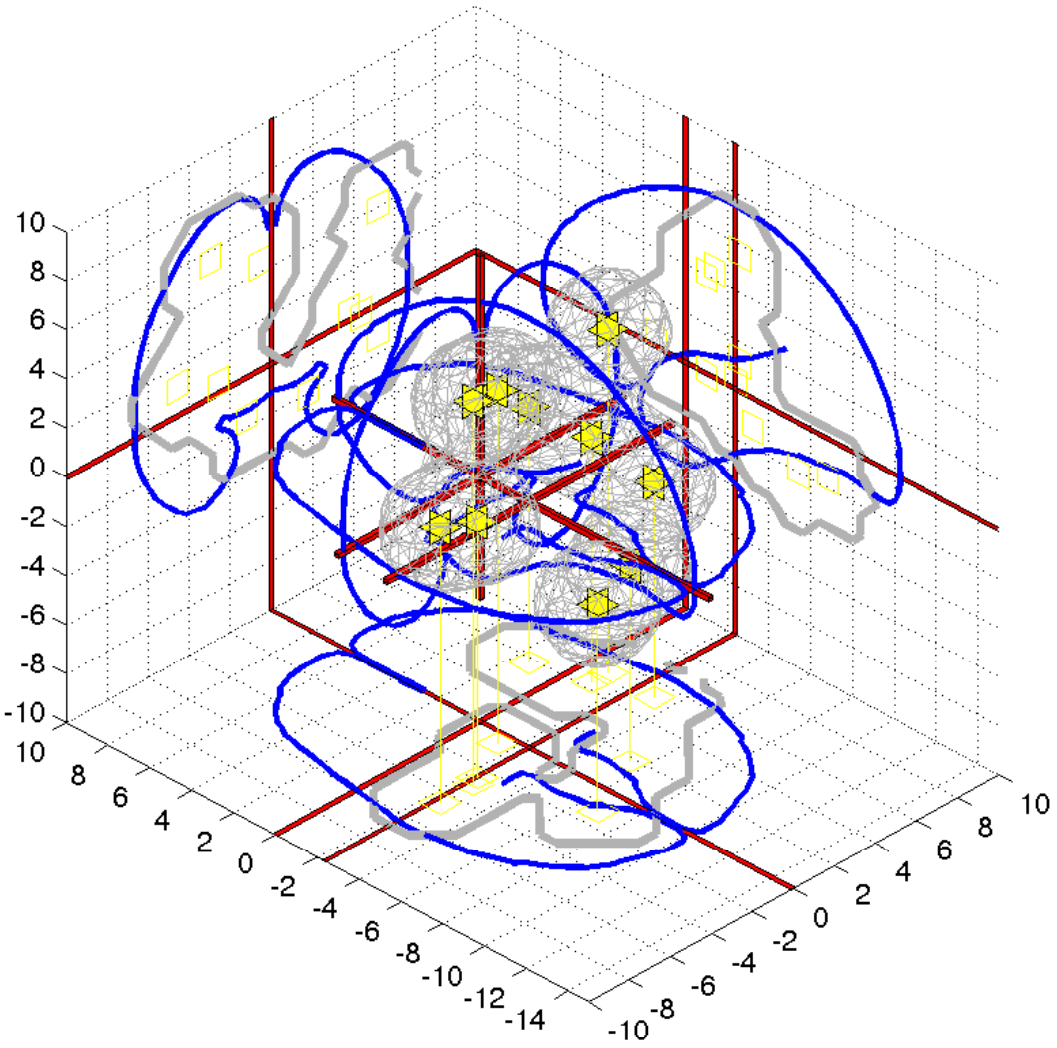
+3: 0.67565 **Forgiveness judgements**. *Judgements of visually displayed sentences about forgiveness situations with button press versus judgement involving social reasoning.* WOEXP: 451.
T. F. Farrow; Y. Zheng; I. D. Wilkinson; S. A. Spence; J. F. Deakin; N. Tarrier; P. D. Griffiths; P. W. Woodruff. *Investigating the functional anatomy of empathy and forgiveness.* *NeuroReport* **12**(11):2433-2438, 2001. PMID: 11496124. FMRIDCID: . WOBIB: 147.

+4: 0.64805 **Case judgment versus syllable counting**. *Case judgment of letters from visually presented words with button press versus counting the number of syllables in a visually presented word.* WOEXP: 553.
Russell A. Poldrack; Anthony D. Wagner; Matthew W. Prull; John E. Desmond; Gary H. Glover; John D. E. Gabrieli. *Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex.* *NeuroImage* **10**(1):15-35, 1999. PMID: 10385578. DOI: 10.1006/nimg.1999.0441. FMRIDCID: . WOBIB: 178.

+5: 0.60237 **Subject 3: Answering self-reflective questions versus answering semantic questions**. *Self-reflective and semantic yes/no questions posed through headphones were answered with button press.* WOEXP: 56.
Sterling C. Johnson; Leslie C. Baxter; Lana S. Wilder; James G. Pipe; Joseph E. Heiserman; George P. Prigatano. *Neural correlates of self-reflection.* *Brain* **125**(Pt 8):1808-14, 2002. PMID: 12135971. WOBIB: 20.

List with results after searching experiments that report similar activations as a "mentalizing" experiment of Gallagher et al. (2002).

Coordinates-to-volume transformation

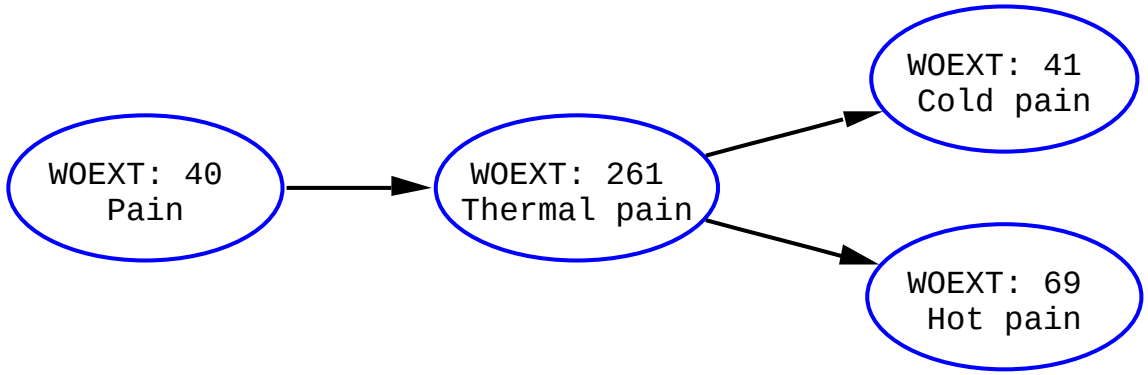


Coordinates in an article converted to volume-data by filtering each point (kernel density estimation) (Nielsen and Hansen, 2002; Turkeltaub et al., 2002)

One volume for each article

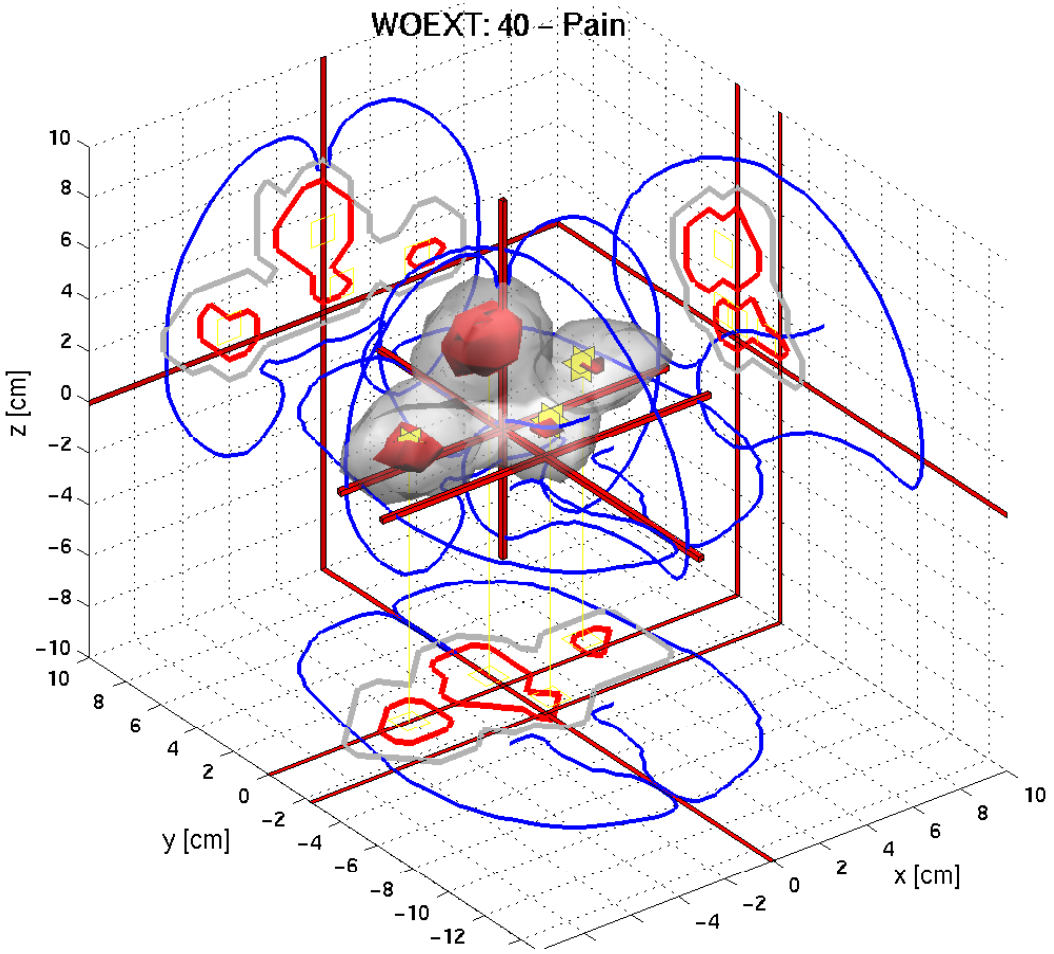
Yellow coordinates from a study by Blinkenberg et al. (1996), with grey wireframe indicating the isosurface in the generated volume

Taxonomy for cognitive components, . . .



Memory, episodic memory, episodic memory retrieval, empathy, disgust, 5-HT2A receptor, . . .

Supervised datamining

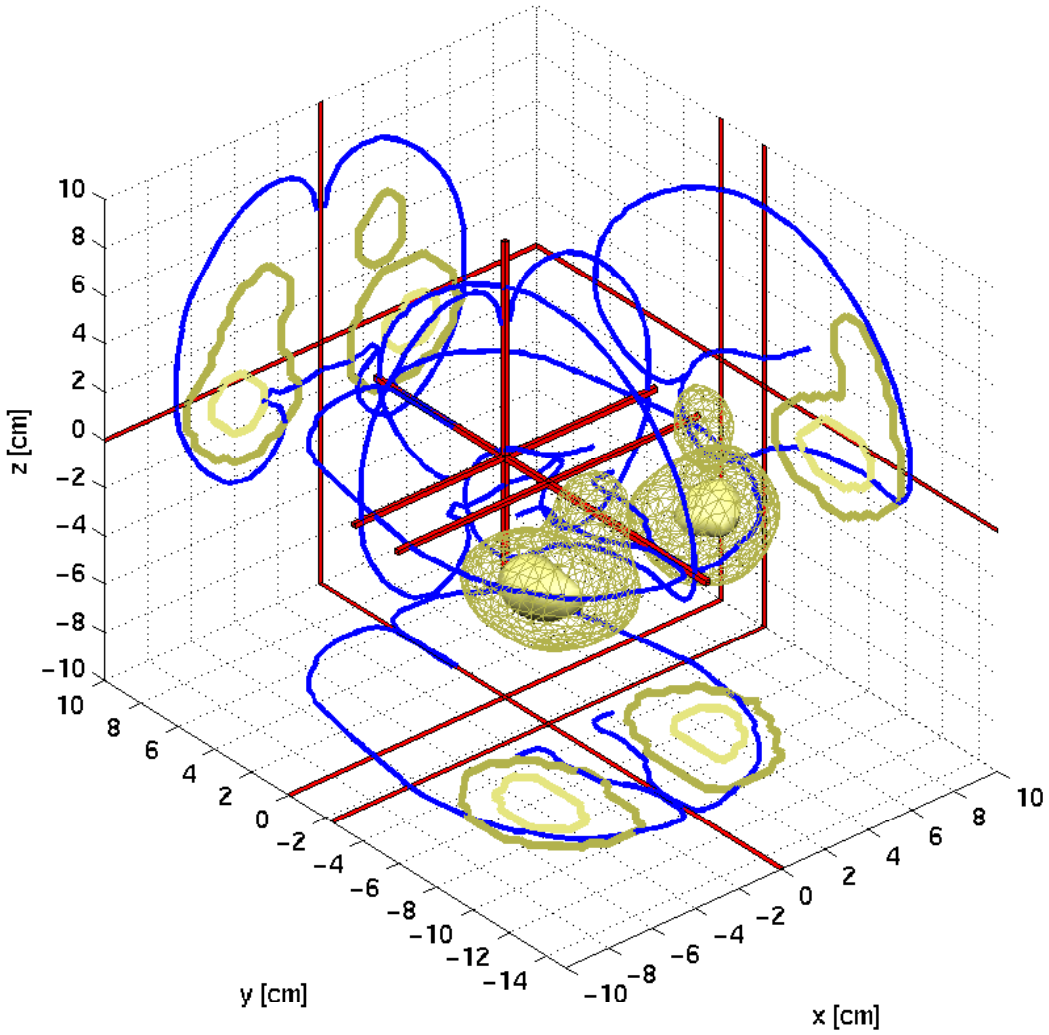


Volume for a specific taxonomic component: "Pain"

Volume threshold at statistical values determined by resampling statistics (Nielsen, 2005).

Red areas are the most significant areas: Anterior cingulate, anterior insula, thalamus. In agreement with "human" reviewer (Ingvar, 1999).

Unsupervised datamining



Construction of a matrix $\mathbf{X}$(papers $\times$ voxels)

Decomposition of this matrix by multivariate analysis, e.g., principal component analysis, clustering, independent component analysis

Left image: non-negative matrix factorization with components weighting for (perhaps) face recognition (Nielsen et al., 2004)

Other technique: Replicator dynamics (Neumann et al., 2005).

Text representation: a "bag-of-words"

	'memory'	'visual'	'motor'	'time'	'retrieval'	...
Fujii	6	0	1	0	4	...
Maddock	5	0	0	0	0	...
Tsukiura	0	0	4	0	0	...
Belin	0	0	0	0	0	...
Ellerman	0	0	0	5	0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋱

Representation of the abstract of the articles in "bag-of-word". Table counts how often a word occurs

Exclusion of "stop words": common words (the, a, of, ...), words for brain anatomy, and a large number of common words that appear in abstracts. Mostly words for brain function are left.

Grouping of words from articles

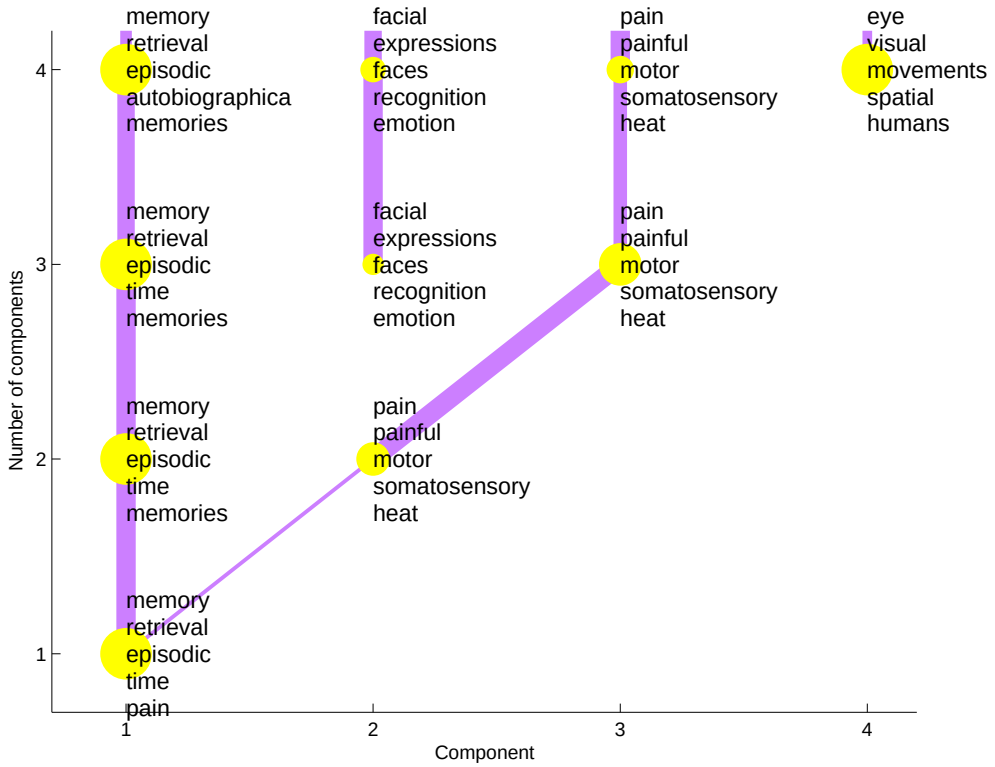


Figure 5: Grouped words.

Multivariate analysis of the text in *posterior cingulate* articles to find "themes", which can be represented with weights over words and articles.

Most dominating words: memory, retrieval, episodic

pain, painful, motor, somatosensory

facial, expressions, faces,

eye, visual, movements

Text and volume: Functional atlas

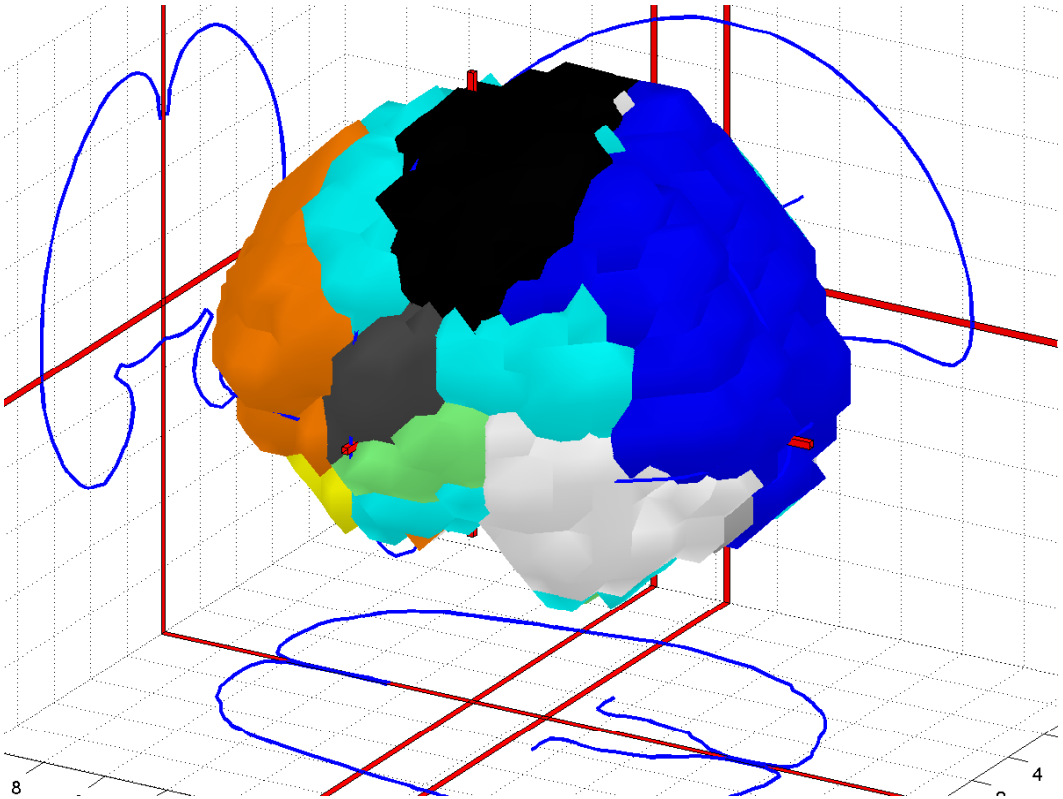


Figure 6: Functional atlas in 3D visualization.

Automatic construction of functional atlas, where words for function become associated with brain areas

Blue area: visual, eye, time

Black: motor, movements, hand

White: faces, perceptual, face

Green: auditory, spatial, neglect, awareness, langauge

Orange: semantiv, phonological, cognitive, decision

Funktional atlas — medial view

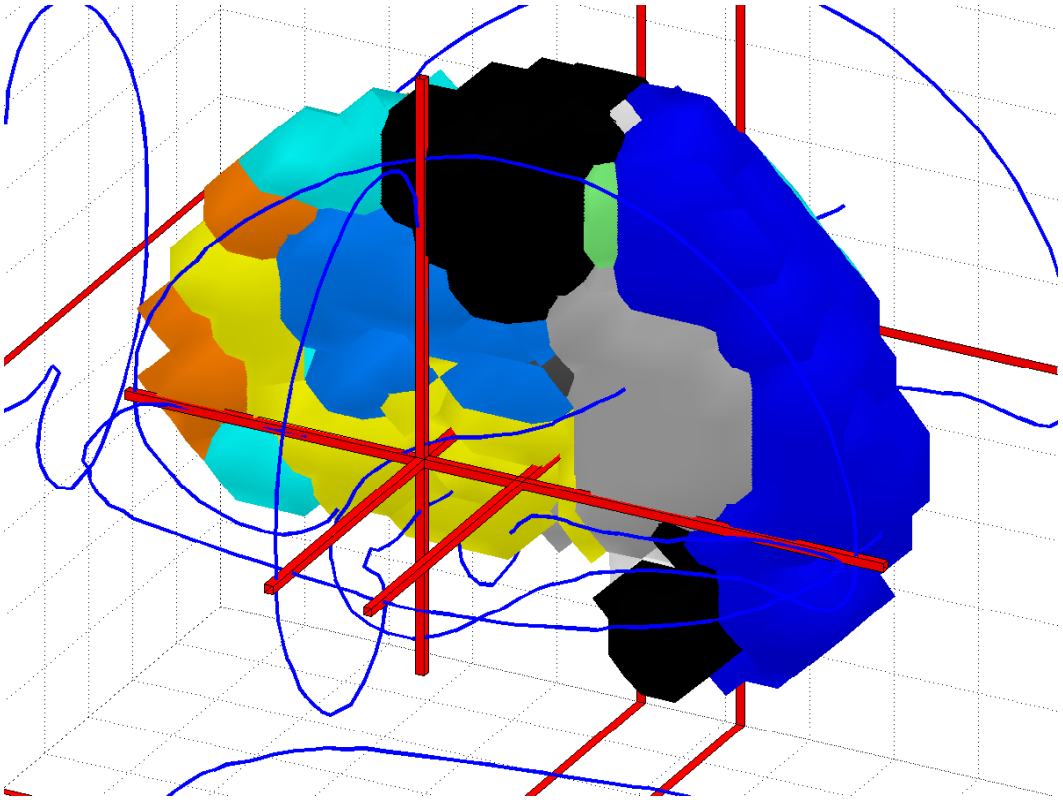


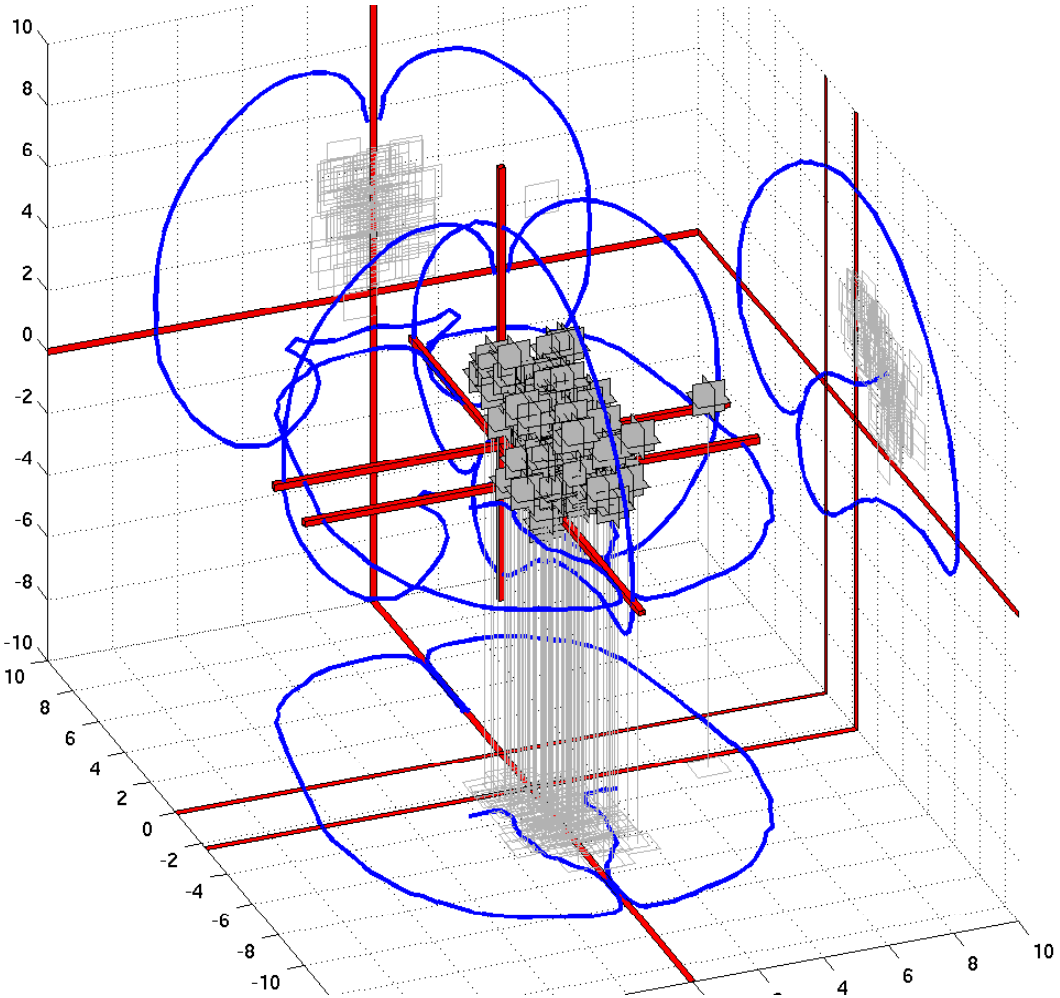
Figure 7: Visualization of the medial area.

Grey area: retrieval, neutral, words, encoding.

Yellow: emotion, emotions, disgust, sadness, happiness

Light blue: pain, noxious, verbal, unpleasantness, hot

Searching on a specific area

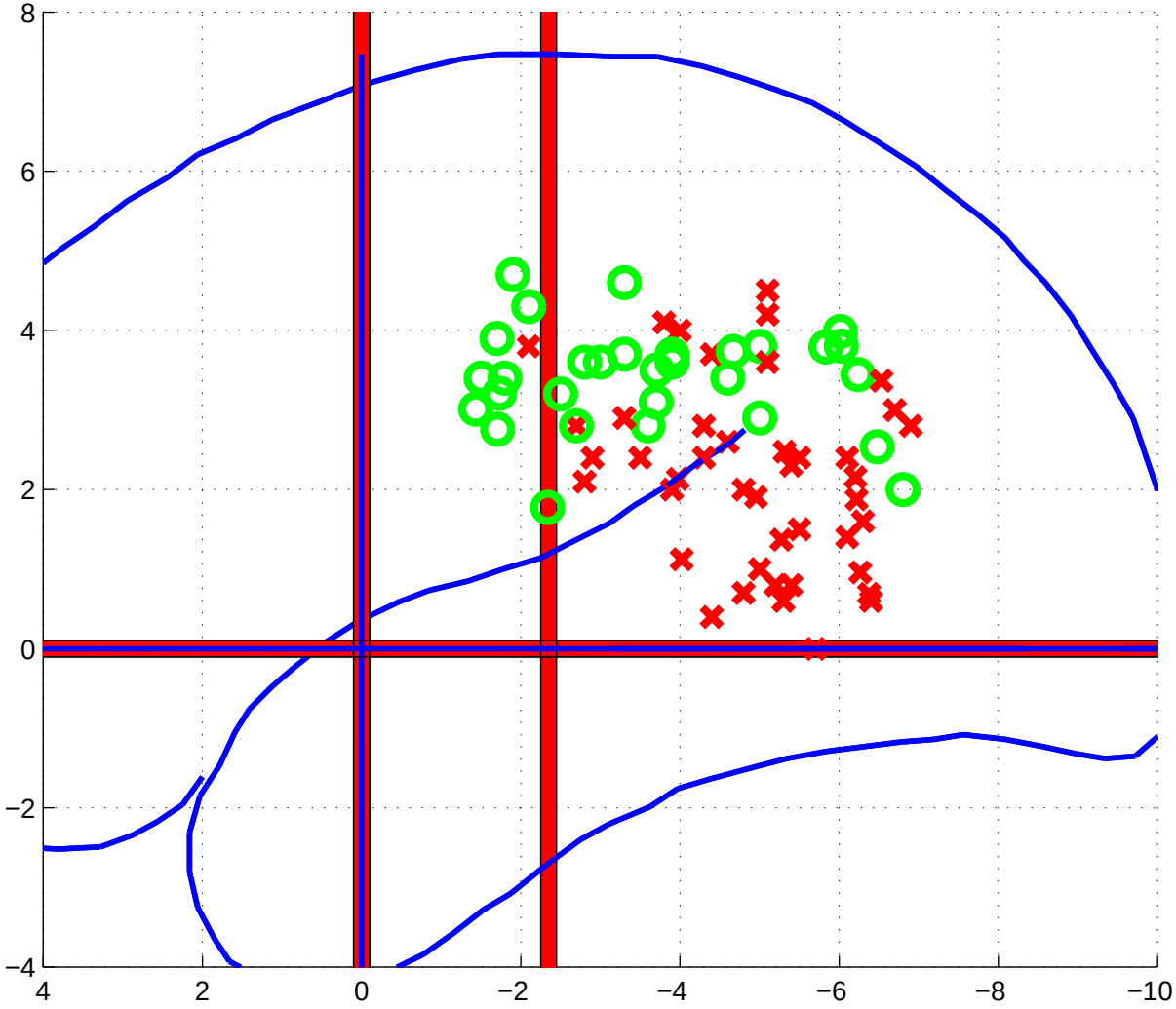


Searching for all coordinates labeled as "posterior cingulate": Here 116 "posterior cingulate" coordinates.

One outlier: "Right postcentral gyrus/posterior cingulate gyrus" from (Jernigan et al., 1998).

Possible to find the corresponding articles for the coordinates — and cluster these articles

Memory and pain



Is there a different between how memory and pain coordinates distribute in posterior cingulate?

Sagittal plot of memory (red x) and pain (green circles).

Apparently the memory coordinates have a tendency to lie in the posterior/inferior part for posterior cingulate.

Imaging databases

fMRIDC: fMRI Data Center stores scanning data from fMRI studies. With Internet-based search.

Neurogenerator: Storing, information retrieval and visualization of imaging data.

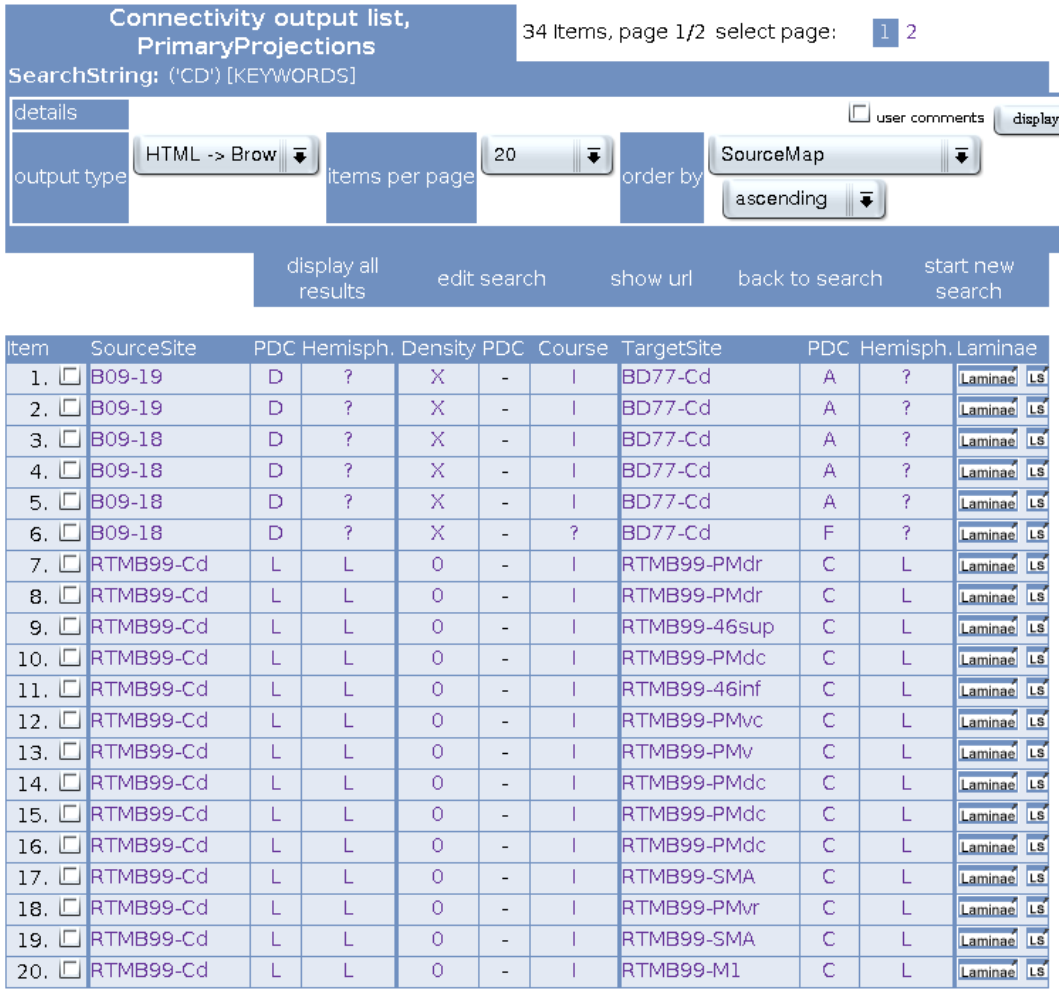
SumsDB: Cortex surface-based database.

Rodent databases: NeSys (projections), Mouse brain library: Nissl-stained

BrainInfo (NeuroNames): Database of brain structures.

Connectivity databases: CoCoMac, CoCoDat, BAMS, XANAT, ...

CoCoMac connectivity database



Connectivity output list, PrimaryProjections — 34 Items, page 1/2 select page: 1 2

SearchString: ('CD') [KEYWORDS]

details | user comments | display

output type: HTML -> Brow | items per page: 20 | order by: SourceMap | ascending

display all results | edit search | show url | back to search | start new search

Item	SourceSite	PDC	Hemisph.	Density	PDC	Course	TargetSite	PDC	Hemisph.	Laminae
1.	B09-19	D	?	X	-	I	BD77-Cd	A	?	Laminae LS
2.	B09-19	D	?	X	-	I	BD77-Cd	A	?	Laminae LS
3.	B09-18	D	?	X	-	I	BD77-Cd	A	?	Laminae LS
4.	B09-18	D	?	X	-	I	BD77-Cd	A	?	Laminae LS
5.	B09-18	D	?	X	-	I	BD77-Cd	A	?	Laminae LS
6.	B09-18	D	?	X	-	?	BD77-Cd	F	?	Laminae LS
7.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdr	C	L	Laminae LS
8.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdr	C	L	Laminae LS
9.	RTMB99-Cd	L	L	0	-	I	RTMB99-46sup	C	L	Laminae LS
10.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdc	C	L	Laminae LS
11.	RTMB99-Cd	L	L	0	-	I	RTMB99-46inf	C	L	Laminae LS
12.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMvc	C	L	Laminae LS
13.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMv	C	L	Laminae LS
14.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdc	C	L	Laminae LS
15.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdc	C	L	Laminae LS
16.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMdc	C	L	Laminae LS
17.	RTMB99-Cd	L	L	0	-	I	RTMB99-SMA	C	L	Laminae LS
18.	RTMB99-Cd	L	L	0	-	I	RTMB99-PMvr	C	L	Laminae LS
19.	RTMB99-Cd	L	L	0	-	I	RTMB99-SMA	C	L	Laminae LS
20.	RTMB99-Cd	L	L	0	-	I	RTMB99-M1	C	L	Laminae LS

CoCoMac records anatomical connectivity in the Macaque brain with data from presently 395 papers.

Brain region ontology (Stephan et al., 2000).

Stores “from”, “to” and how strong the link is, what tracer, etc.

Visualization of connectivity, analysis of, e.g., small-worldness (Sporns et al., 2004)

More information

Bibliography on Neuroinformatics

http://www.imm.dtu.dk/~fn/bib/Nielsen2001Bib/

References

Balslev, D., Nielsen, F. Å., Paulson, O. B., and Law, I. (2005). Right temporoparietal cortex activation during visuo-proprioceptive conflict. *Cerebral Cortex*, 15(2):166–169. PMID: 152384438. WOBIB: 128. http://cercor.oupjournals.org/cgi/content/abstract/15/2/166?etoc.

Blinkenberg, M., Bonde, C., Holm, S., Svarer, C., Andersen, J., Paulson, O. B., and Law, I. (1996). Rate dependence of regional cerebral activation during performance of a repetitive motor task: a PET study. *Journal of Cerebral Blood Flow and Metabolism*, 16(5):794–803. PMID: 878424. WOBIB: 166.

Fox, P. T. and Lancaster, J. L. (2002). Mapping context and content: the BrainMap model. *Nature Reviews Neuroscience*, 3(4):319–321. http://www.brainmapdbj.org/Fox01context.pdf. Describes the philosophy behind the (new) BrainMap functional brain imaging database with "BrainMap Experiment Coding Scheme" and tables of activation foci. Furthermore discusses financial problems and quality control of data.

Fox, P. T., Mikiten, S., Davis, G., and Lancaster, J. L. (1994). BrainMap: A database of human function brain mapping. In Thatcher, R. W., Hallett, M., Zeffiro, T., John, E. R., and Huerta, M., editors, *Functional Neuroimaging Technical Foundations*, pages 95–105. Academic Press, San Diego, California.

Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage*, 16(3 Part 1):814–821. PMID: 12169265. ISSN 1053-8119.

Ingvar, M. (1999). Pain and functional imaging. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 354(1387):1347–1358. PMID: 10466155.

Jernigan, T. L., Ostergaard, A. L., Law, I., Svarer, C., Gerlach, C., and Paulson, O. B. (1998). Brain activation during word identification and word recognition. *NeuroImage*, 8(1):93–105. PMID: 9698579. WOBIB: 35.

Neumann, J., Lohmann, G., Derrfuss, J., and von Cramon, D. Y. (2005). Meta-analysis of functional imaging data using replicator dynamics. *Human Brain Mapping*, 25(1):165–173. http://www3.interscience.wiley.com/cgi-bin/abstract/110474181/. ISSN 1065-9471.

Nielsen, F. Å. (2005). Mass meta-analysis in Talairach space. In Saul, L. K., Weiss, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 17*, pages 985–992, Cambridge, MA. MIT Press. http://books.nips.cc/papers/files/nips17/NIPS2004_0511.pdf.

Nielsen, F. Å. and Hansen, L. K. (2002). Modeling of activation data in the BrainMap™ database: Detection of outliers. *Human Brain Mapping*, 15(3):146–156. DOI: 10.1002/hbm.10012. http://www3.interscience.wiley.com/cgi-bin/abstract/89013001/. CiteSeer: http://citeseer.ist.psu.com/nielsen02modeling.html.

Nielsen, F. Å., Hansen, L. K., and Balslev, D. (2004). Mining for associations between text and brain activation in a functional neuroimaging database. *Neuroinformatics*, 2(4):369–380. http://www2.imm.dtu.dk/~fn/ps/Nielsen2004Mining_submitted.pdf.

Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9):418–425.

Stephan, K. E., Zilles, K., and Kötter, R. (2000). Coordinate-independent mapping of structural and functional data by objective relational transformation (ORT). *Philosophical Transactions of the Royal Society, London, Series B, Biological Sciences*, 355(1393):37–54. PMID: 10703043.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., and Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *NeuroImage*, 16(3 part 1):765–780. PMID: 12169260. DOI: 10.1006/nimg.2002.1131. http://www.sciencedirect.com/science/article/-B6WNP-46HDMPV-N/2/xb87ce95b60732a8f0c917e288efe59004.