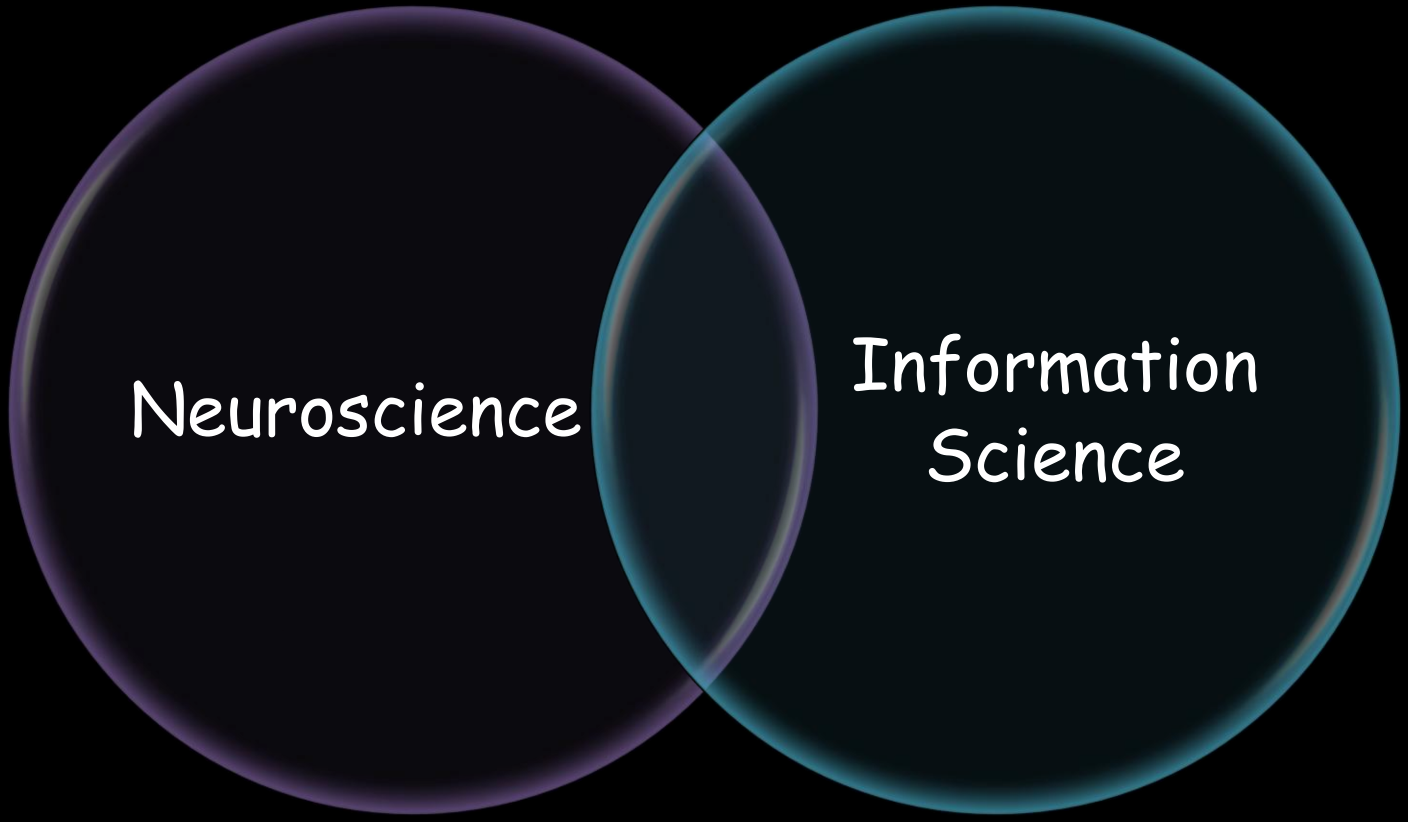
Neuroinformatics



Rationale for Sharing

Sharing Principles

Data Curation

Ontology

XML

Data Repositories

Atlases

Smart Metadata Search

Neuroinformatics fosters collaborative neuroscience research by providing:

-principles,

-vocabulary &

-tools

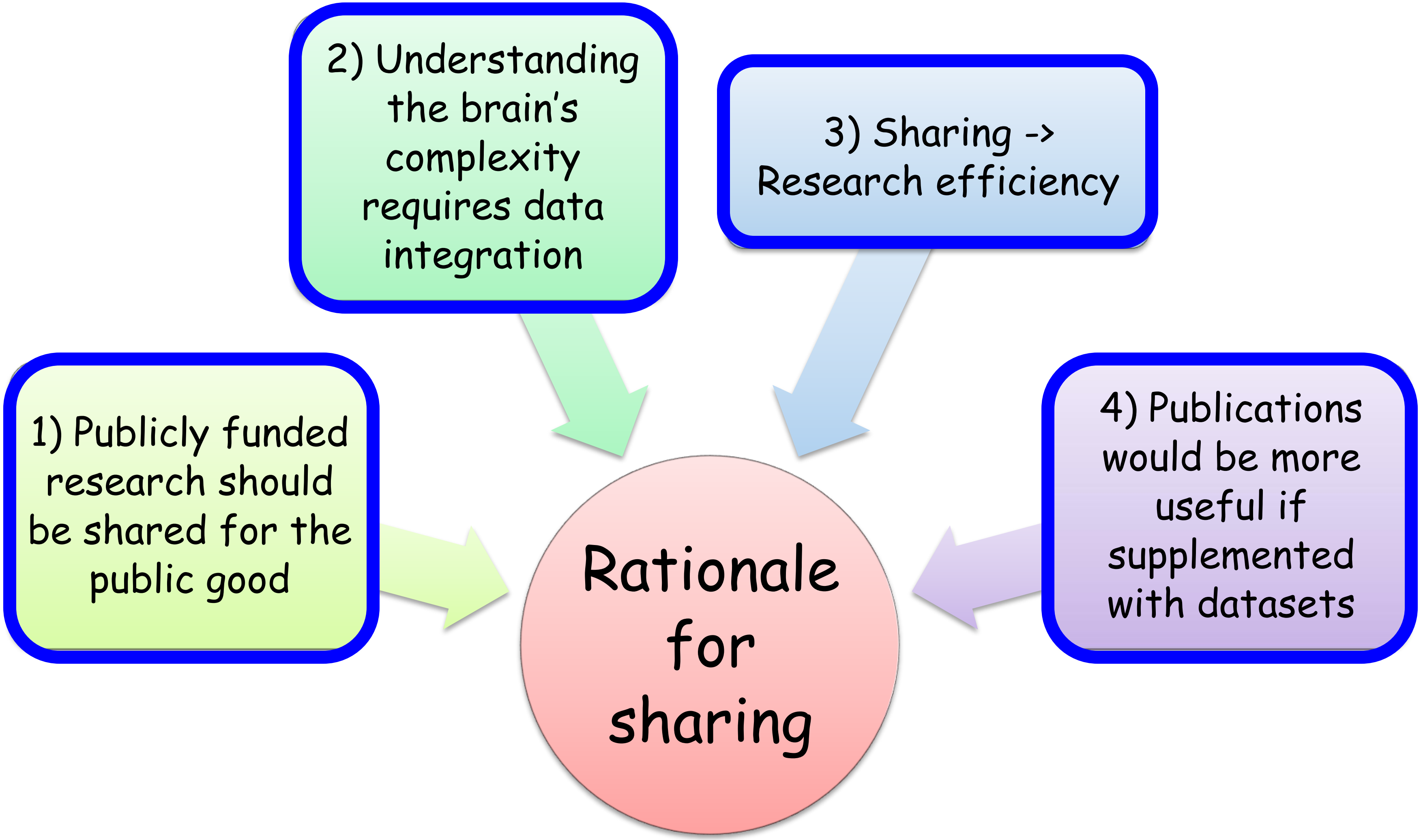
for sharing neuroscience data.

My focus will be on how neuroinformatics impacts neuroimaging.

Neuroinformatics began with the Human Brain Project (NIMH) in 1993, and grew to involve lots of significant groups and projects, e.g.,....

Neuroinformatics efforts are now international.

The **H**uman **B**rain **P**roject explained the rationale for sharing:


2) Understanding the brain's complexity requires data integration
3) Sharing -> Research efficiency
1) Publicly funded research should be shared for the public good
Rationale for sharing
4) Publications would be more useful if supplemented with datasets

The Human Brain Project also developed data sharing principles…

so that no one would get burned by the process of sharing.

Sharing should not imply relinquishing.

The data originator should be consulted & included in any data re-use.

The originator should retain ownership of the data.

Sharing requires extra subject protection

We need to use HIPPAA compliant de-identification

Include information about archiving and sharing in our consents.

In general,

Publication is an optimal time to share the data

Data sharing & mining should be rewarded

Finally…

Sharing infrastructure & costs should be supported

Summary:

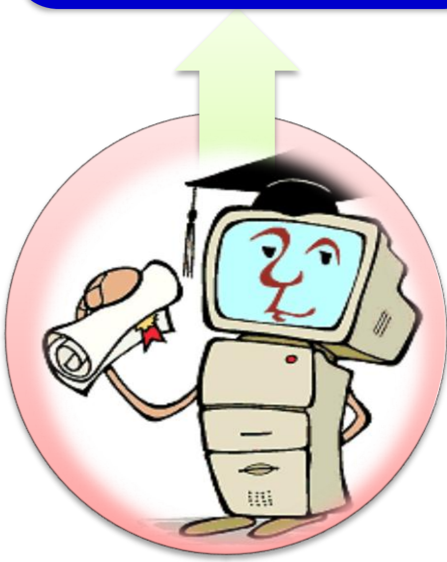
The Human Brain Project kick started neuroinformatics by:

sponsoring lots of projects,

and defining the rationale and rules for sharing.

The rationale and the principles sound good, but does this really work? And, what's in it for the sharer?

Data set reanalysis



Van Horn et al. (2007) describe sharing data from their object recognition study. It was subsequently used in:

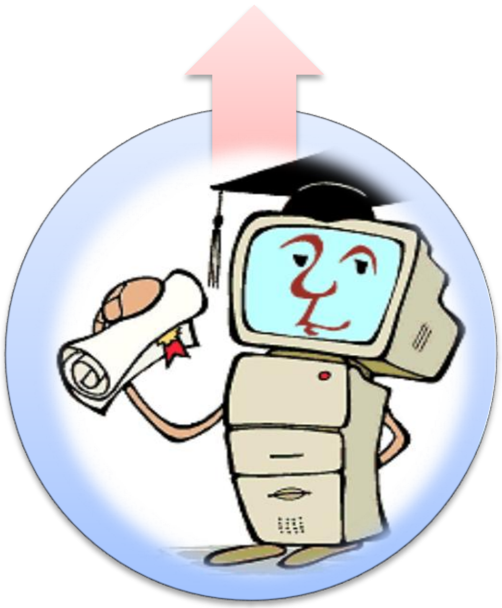
-a visual imagery study

-an object classification study,

-to test a theory of consciousness, &

-to do systems modeling of the brain.

Testing new analytic tools



They also report their dataset being used:

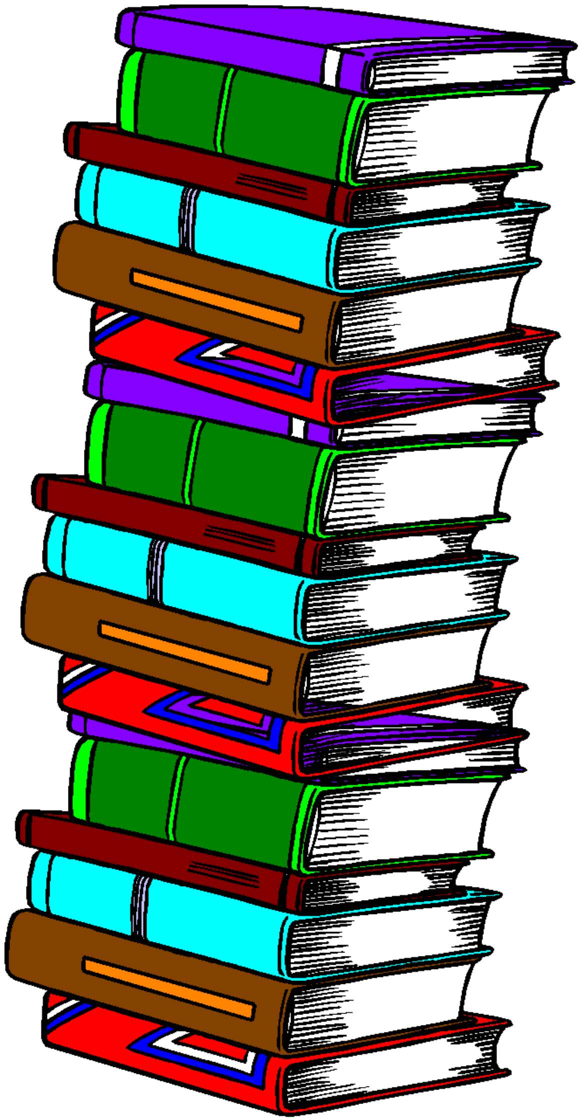
-to deploy new analytic tools,

-do Dynamic Causal Modeling &

-to explore the nature of reproducibility.

(Van Horn and Ishai, 2007)

So, shared data grows into more publications and citations for you,



and the scientific impact of your data is increased when you share it.

The most tangible outcome of the neuroinformatics effort is the excellent sharing infrastructure:

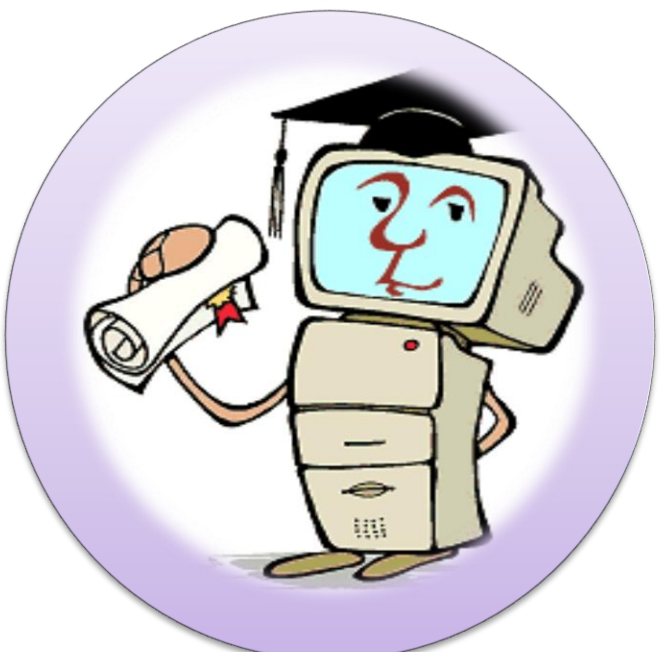
- ontologies,
- standard data formats,
- repositories,
- atlases, and
- tools.

The infrastructure is good because it adheres to the principles of data curation:

That is, sharing is only going to work if everyone can:

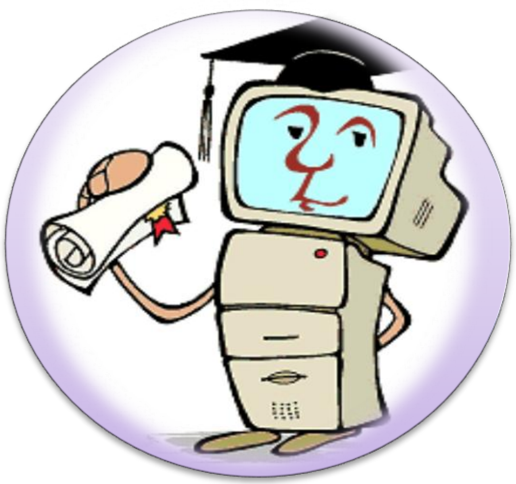
1) Read the data,
2) Understand how it was processed, and
3) Discuss experimental goals and results in the same terms.

We need shared data formats, optimized for neuroimaging



- Formats should be supported by the neuroimaging community,
- tailored for neuroimaging needs,
- documented,
- and extensible

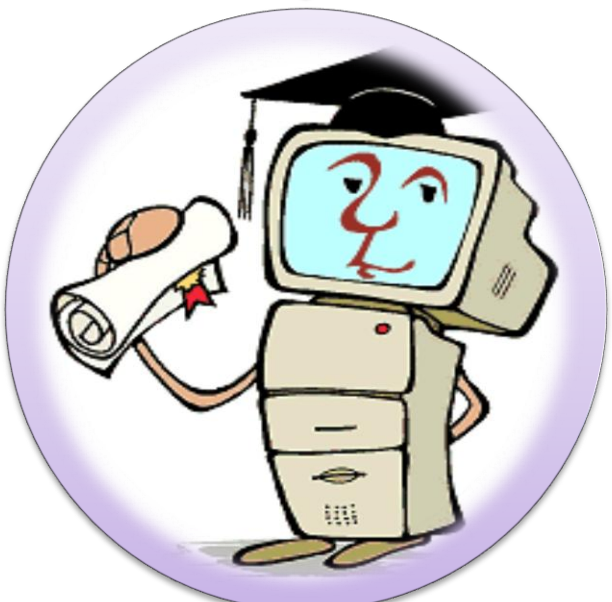
We need shared, transparent, extensible tools



Tools should be

- available to everyone,
- extensible,
- and transparent:
  - We should have access to source code* (*more on this later)
  - The tools should log their activities.

And we need a consistent, unambiguous vocabulary

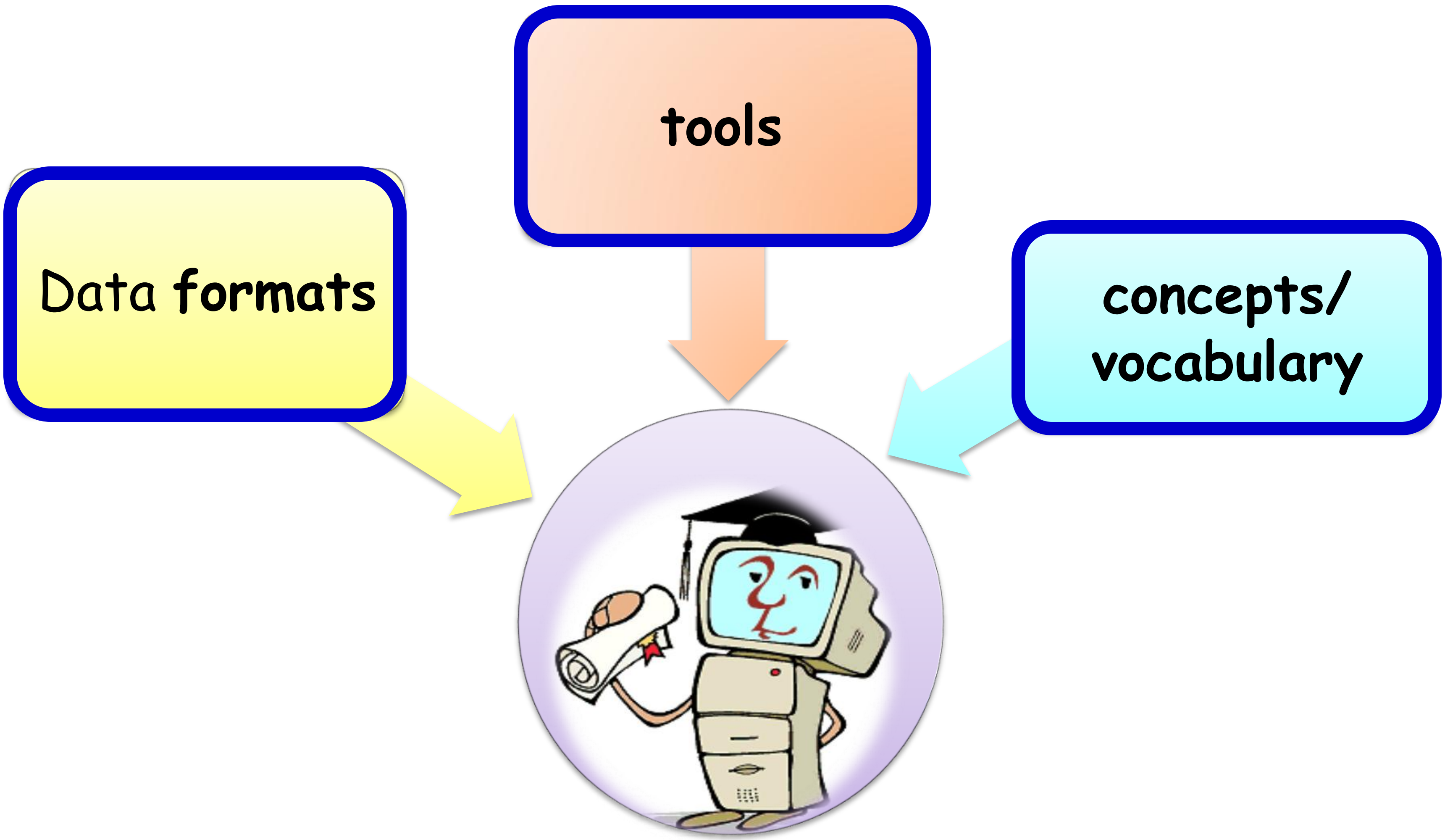


A shared vocabulary for

- experiments,
- data,
- analyses &
- neuroanatomy

will help with search and organization.

In summary, good data curation requires sharing 3 things:



To control tool costs, maintain transparency, and encourage enhancements, neuroinformatics groups have embraced open source tools.

Linux, the GNU tools and most of the image processing programs we depend on are open source.

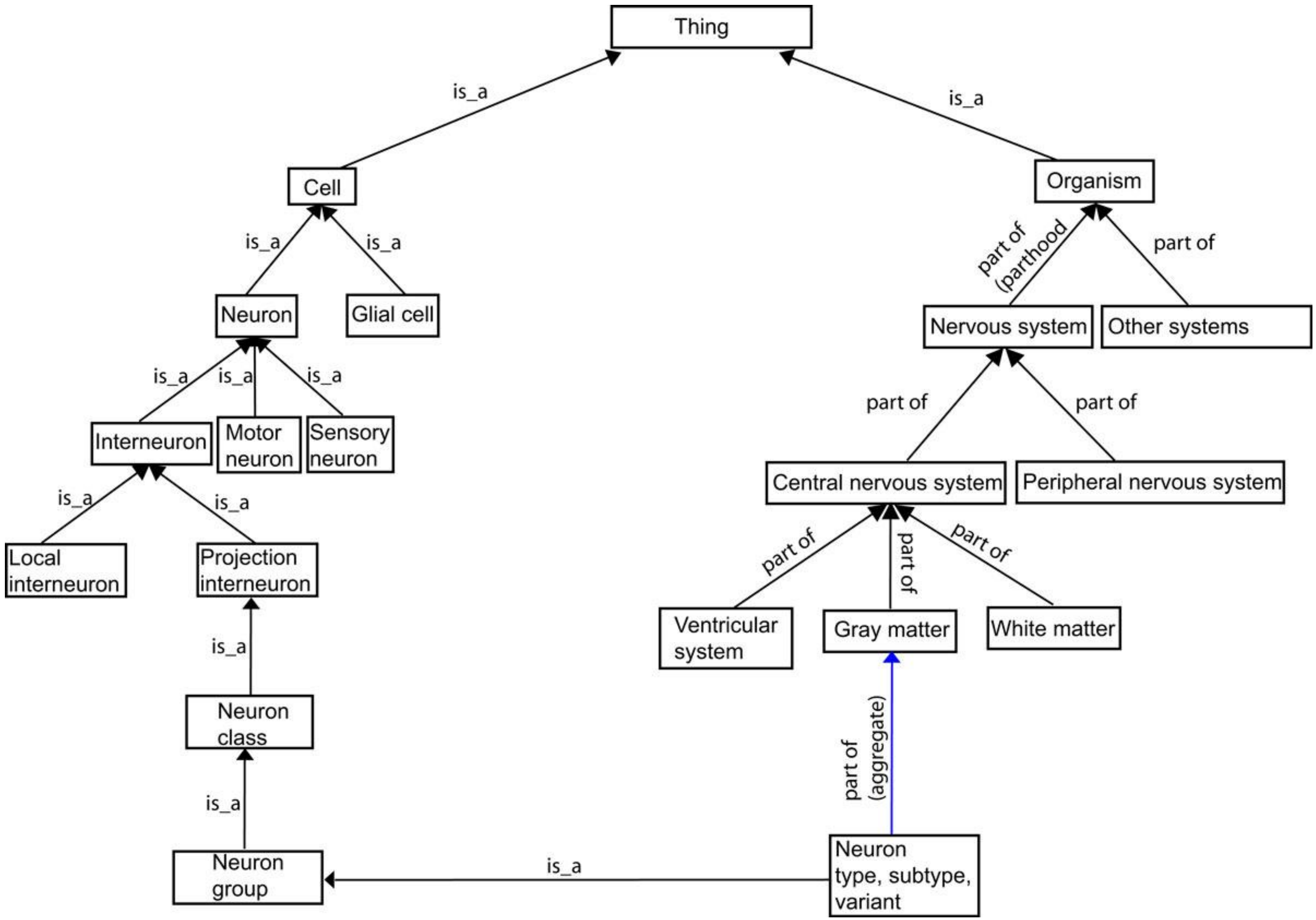
By using open source we guarantee that the software is free, and that the source code can be examined.

But, we need more than open source to make data sharing possible.

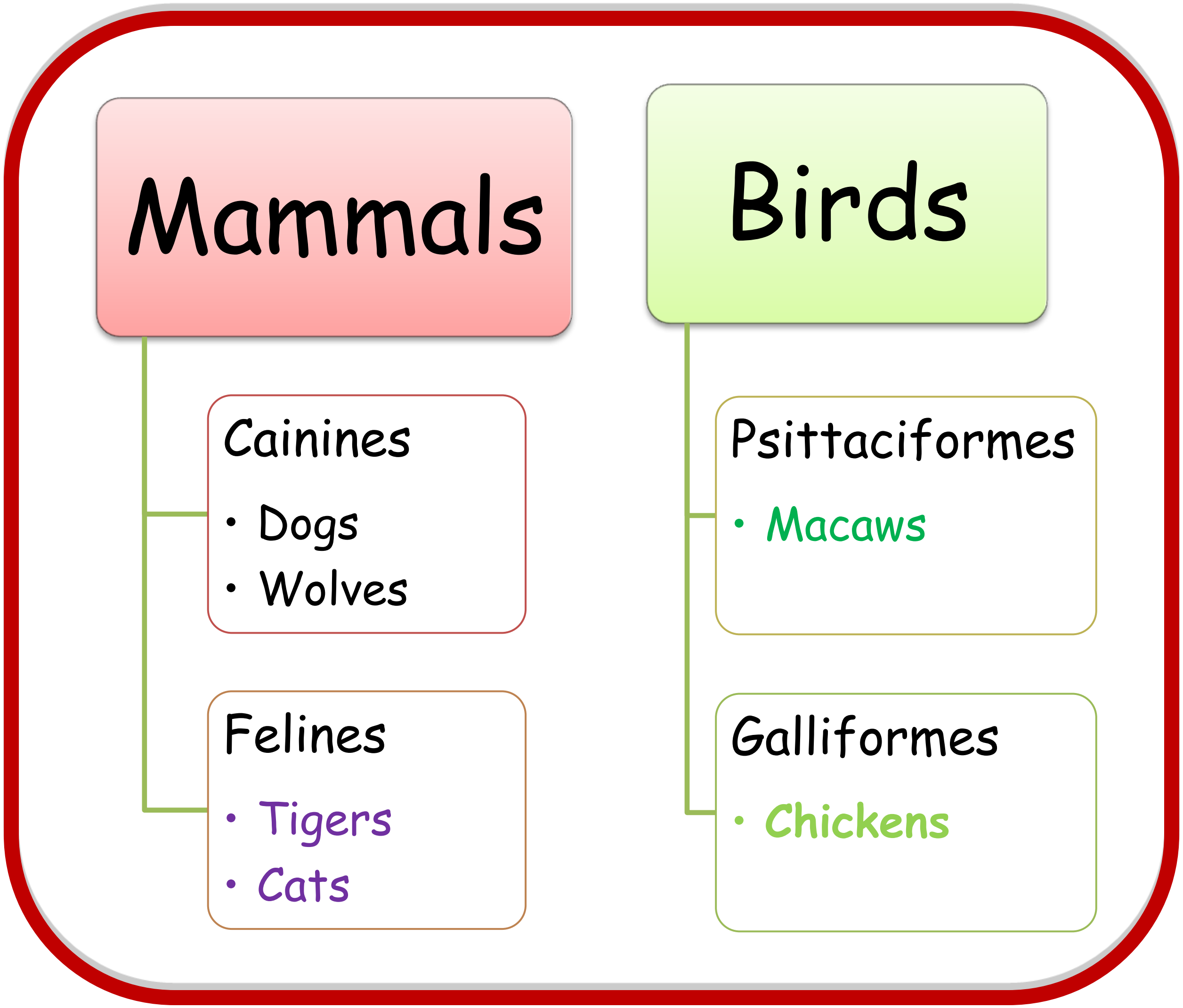
We need
a **shared vocabulary**
across
repositories,
atlases,
search tools and
image processing tools.

We need shared ontologies.

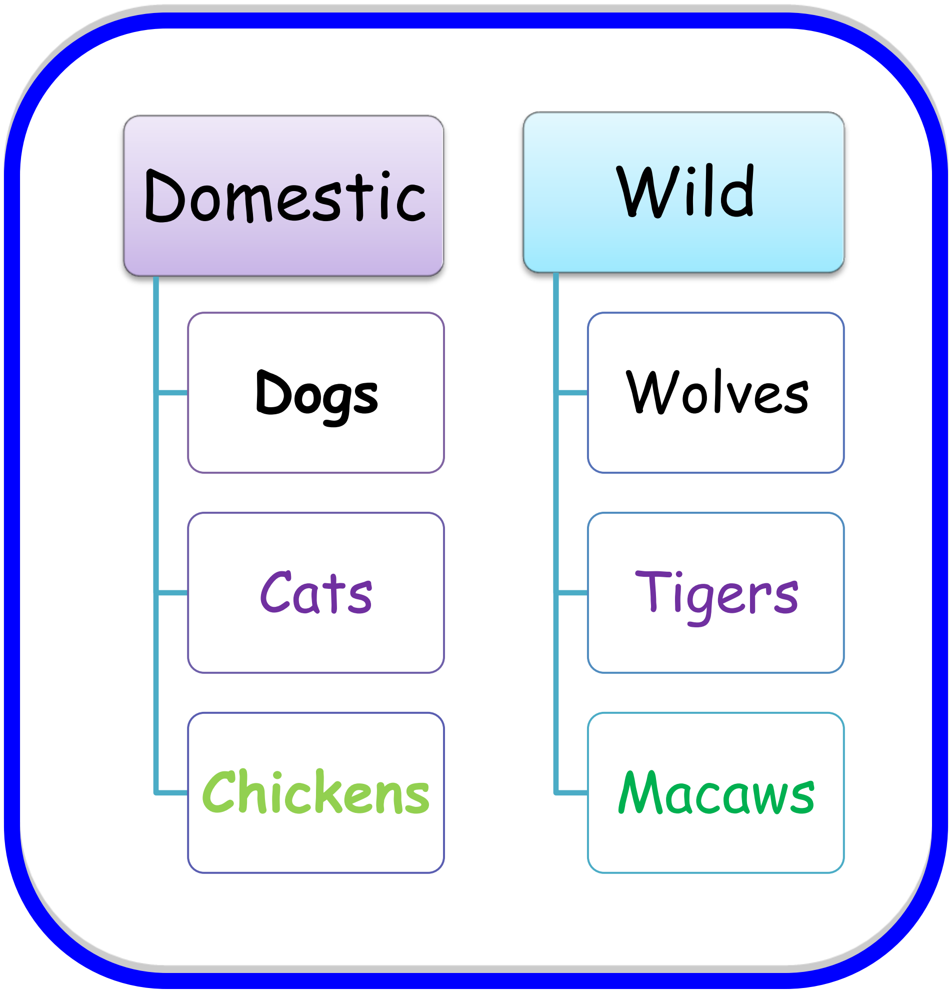
An ontology is a representation of concepts and their relationships.



Here's one possible ontology for some common animals:



And here's another possible ontology for the same animals:



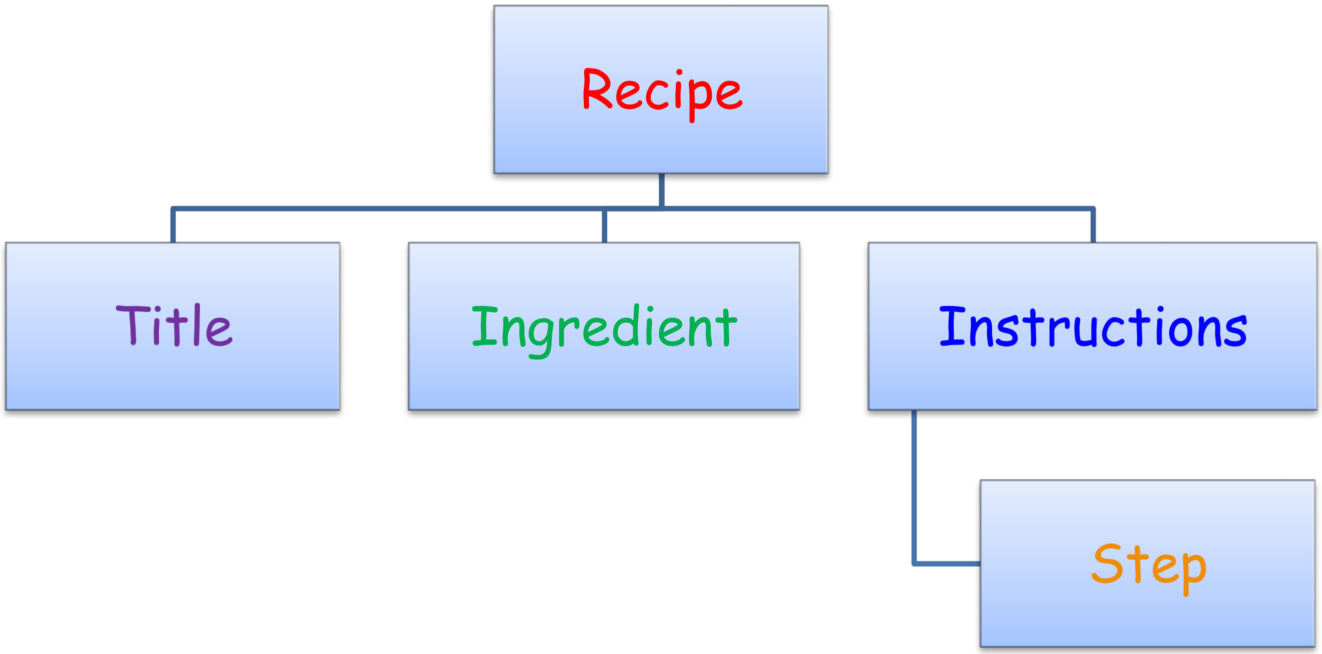
So, you see that ontologies can carve up the world in different ways. But, to unambiguously describe data, neuroscientists need **SHARED** ontologies.

There can be multiple ontologies, as long as we know which one applies at any given time.

We need a tool that can translate our theoretical ontologies into something the computer can interpret and use to parse documents.

The Extensible Markup Language (XML) is that tool. XML is a general-purpose language for structuring data into hierarchies.

So, here's a simple recipe ontology:



We'll turn our recipe ontology into an "XML Schema" (a.k.a "grammar")

and call it "RecipeML"
(Recipe Markup Language)

RecipeML Schema

```
<?xml version="1.0" encoding="utf-8"?>
<xsd: schema xmlns:xsd=http://www.w3.org/2001/XMLSchema>
   <xsd: complexType name="recipe" />
     <xsd: element name="Title" type="xsd:string" />
     <xsd: element name="Ingredient" type="xsd:string" />
     <xsd:complexType name="Instructions" />
       <xsd: element name="Step" type="xsd:string" />
    </xsd: complexType >
</xsd: schema >
```

And we use RecipeML to mark up a document:

```xml
<recipe name="bread" prep_time="5 mins" cook_time="3 hours">
  <title>Basic bread</title>
  <ingredient amount="8" unit="dL">Flour</ingredient>
  <ingredient amount="10" unit="grams">Yeast</ingredient>
  <ingredient amount="4" unit="dL" state="warm">Water</ingredient>
  <ingredient amount="1" unit="teaspoon">Salt</ingredient>
  <instructions>
    <step>Mix all ingredients together.</step>
    <step>Knead thoroughly.</step>
    <step>Place in a bread baking tin.</step>
    <step>Bake in the oven at 180 C for 30 minutes.</step>
  </instructions>
</recipe>
```

XML thus facilitates data exchange via the Internet by encoding document structure and content.

So XML is the perfect tool for representing an ontology and encoding semantic knowledge in our documents.

If you've heard of the semantic web or Web 2.0, this is what it is all about.

For example:

Today, <macaw>Fred</macaw> ate mixed veggies for the first time. He's getting to be a big boy!

In conjunction with this ontology:

- **Mammals**
  - Cainines
    - Dogs
    - Wolves
  - Felines
    - Tigers
    - Cats
- **Birds**
  - Psittaciformes
    - Macaws
  - Galliformes
    - Chickens

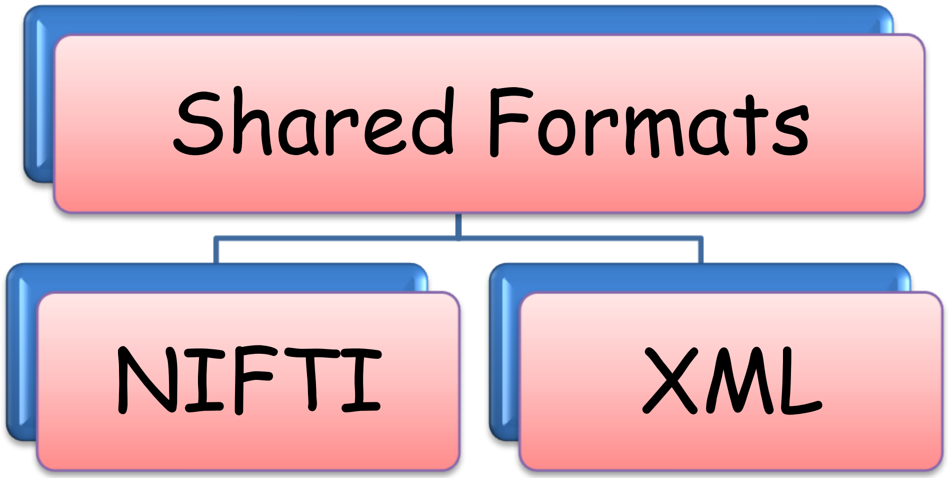
the XML tag <macaw> tells us that Fred is a Psittaciformes AND a bird.

XML is the basis for a number of Markup Languages used by neuroscience:

- BrainML is Brain Markup Language for exchanging neuroscience data.
- MML is Medical Markup Language.
- VRML is Virtual Reality Markup Language.
- MRML is Medical Reality Markup Language (MML+VRML)

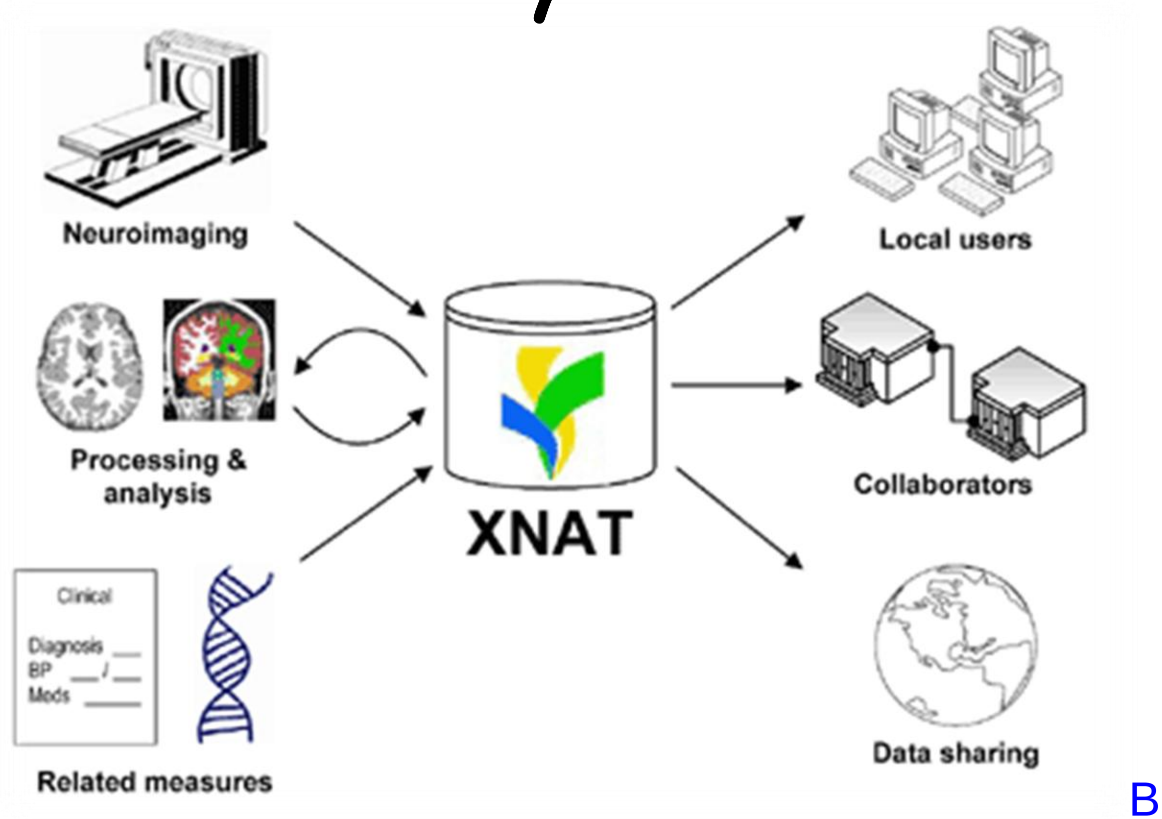
So, using open source, good ontologies and XML we build smart repositories of neuroimaging data and articles on the web.

Now, lets take a quick look at some of the infrastructure projects that have emerged from the neuroinformatics push.


Shared Formats
NIFTI
XML

XML is ubiquitous, extensible and accessible. It is being used to store everything from databases to scene files, and much more.

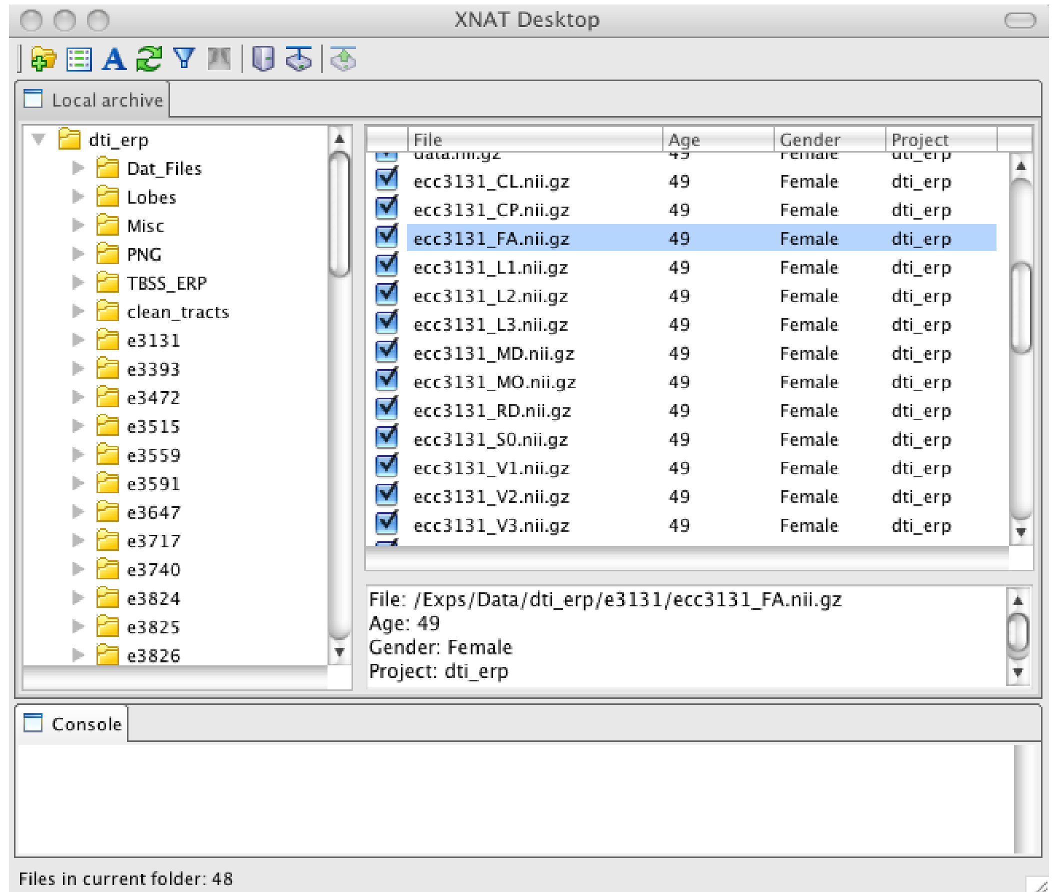
XNAT is an open source XML database used to store collaborative data at BIRN and beyond



Back to Pictures

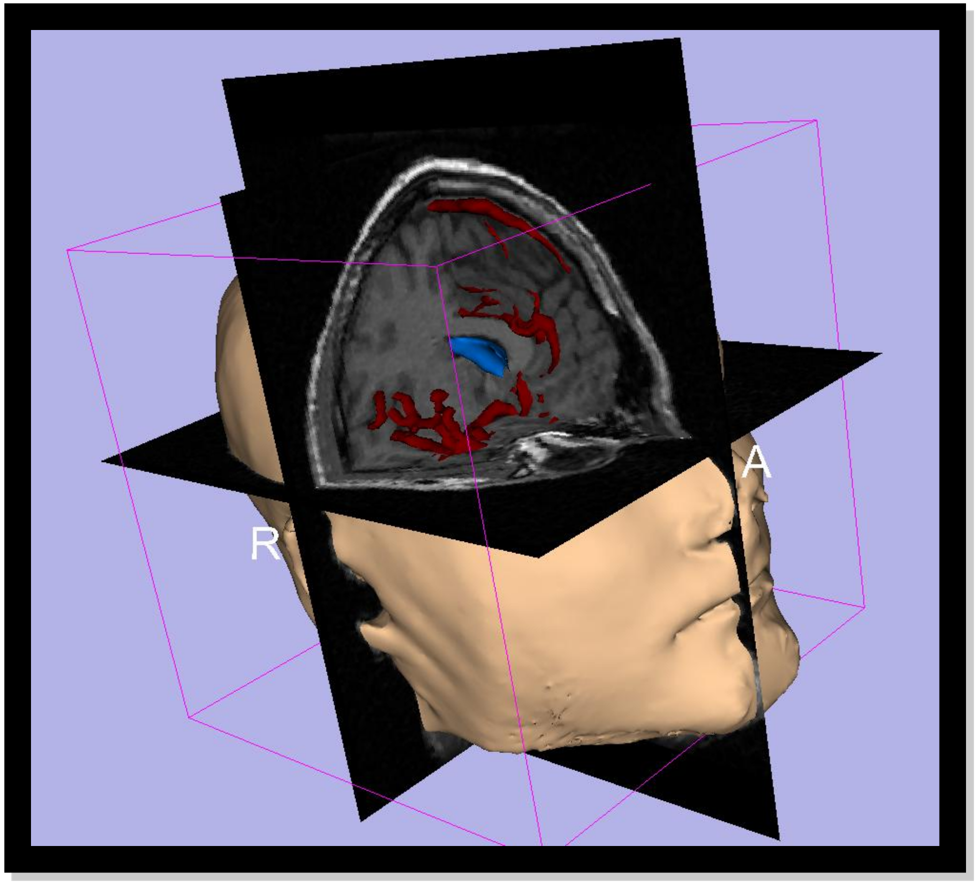
XNAT Desktop

- An database for Mac, Windows or Linux.
- Allows you to load different XML schemas (ontologies).
- Compatible with *THE* XNAT database (and allows uploads to it).



3DSlicer

-Open Source tool for image visualization and processing.
-Uses MRML (XML) to save complex scenes.
-Built-in XNAT upload compatibility.



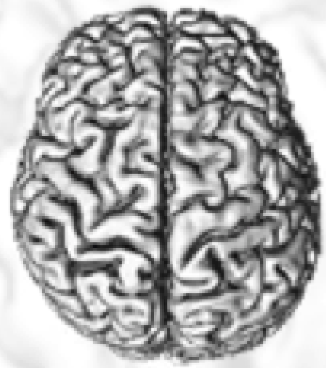
In addition to developing shared formats
like BrainML, MRML and NIFTI,
the neuroinformatics effort has provided
a variety of web resources and tools:

Data Repositories

The fMRI Data Center

fMRIDC

SEARCH fMRIDC Database FOR SUBMIT



A public repository of peer-reviewed fMRI studies and their underlying data.

Funded By
The National Science Foundation
The W. M. Keck Foundation
The National Institutes of Mental Health
A Sun Center of Excellence for Neuroscience

InfoWorld 100

INFORMATION

How do I get started?
Answers to questions commonly posed by first-time visitors.

Q&A about fMRIDC
A comprehensive list of frequently asked questions about the fMRIDC.

Available Datasets
A list of datasets currently available.

Information for Authors
How to submit your imaging data to the Data Center.

fMRIDC NEWS

fMRIDC now shipping data
November 27, 2007 - Datacenter project relocates to UCSB

fMRIDC Moving to UCSB
June 2, 2006 - The fMRI Data Center will not be accepting new data submissions until further notice while we prepare for a transition to UCSB

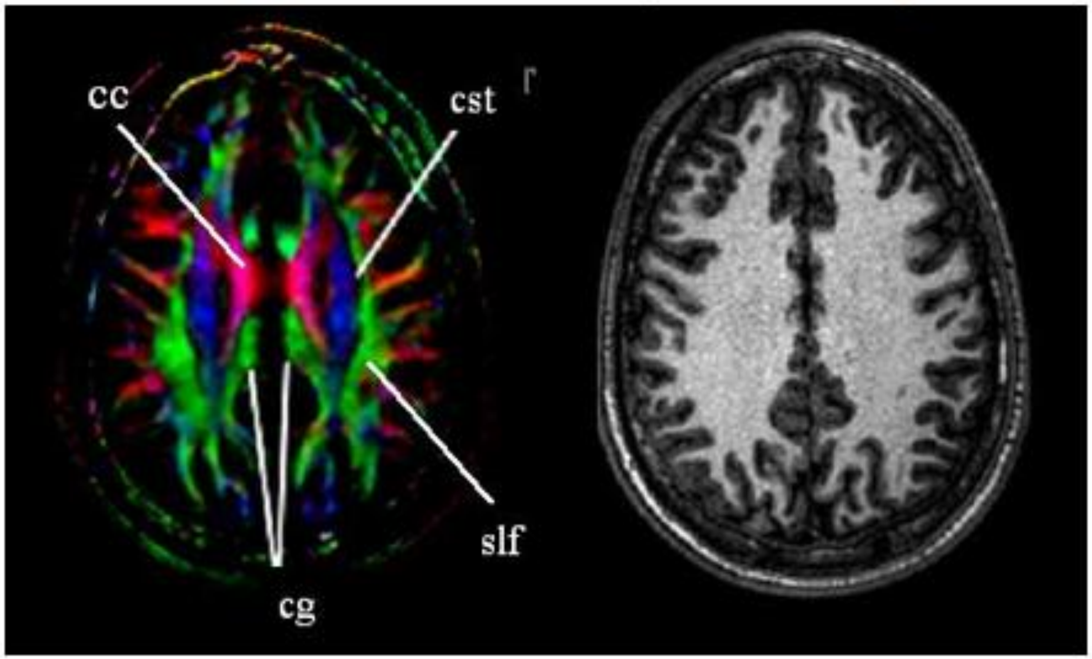
fMRIDC Summer Workshop to be Delayed
February 23, 2006 -

Continuing Progress in Neuroinformatics
January 13, 2006 - A letter in tody's *Science* encourages Federal funding for continued adavances in Neuroinformatics

More news items...

JOHNS HOPKINS MEDICAL INSTITUTE
LABORATORY OF BRAIN ANATOMICAL MRI

Human White Matter Anatomy and Development



A repository of normal human dtis: childhood thru adulthood

Home page | Participating centers | Sponsoring institutes | Image gallery | Database login →

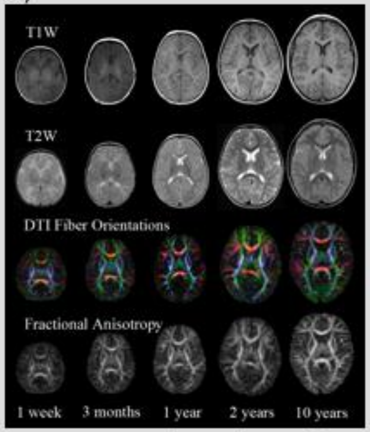
The NIH MRI Study of Normal Brain Development

A project sponsored by the National Institutes of Health

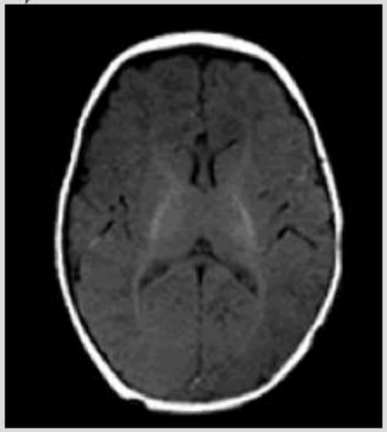
Image gallery

Below are some images from our gallery. Please click for full view. Image A) Images of T1W, T2W, DTI Fiber Orientations, Fractional Anisotropy at various stages of development. Image B) Animation of a T1W image from 3 months to 11 months. C) Cortical thickness output

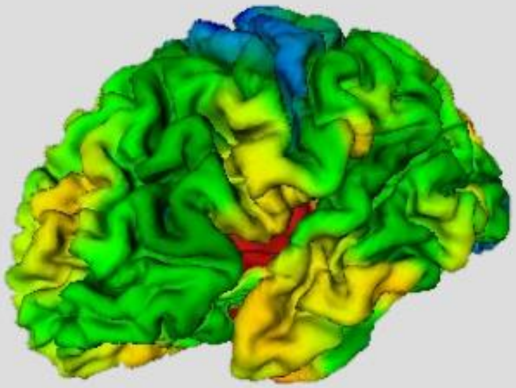
A)



B)



C)

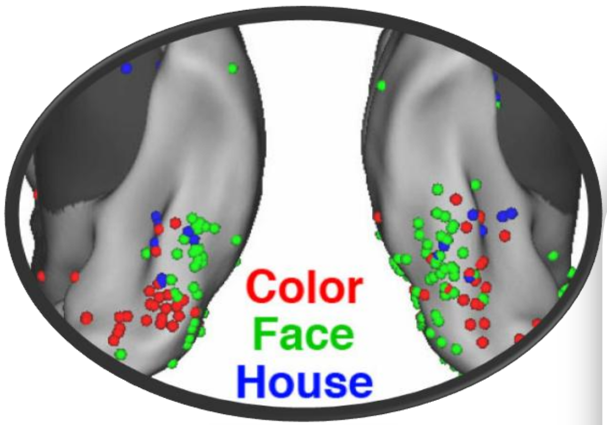


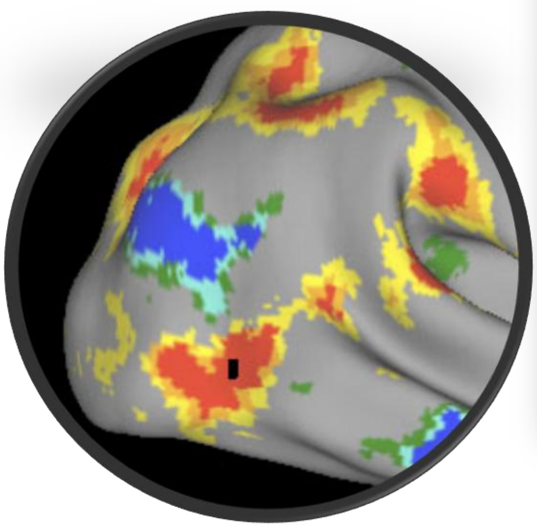
Study design overview | Data user documentation | FAQs | Site map | Privacy policy | Contact us

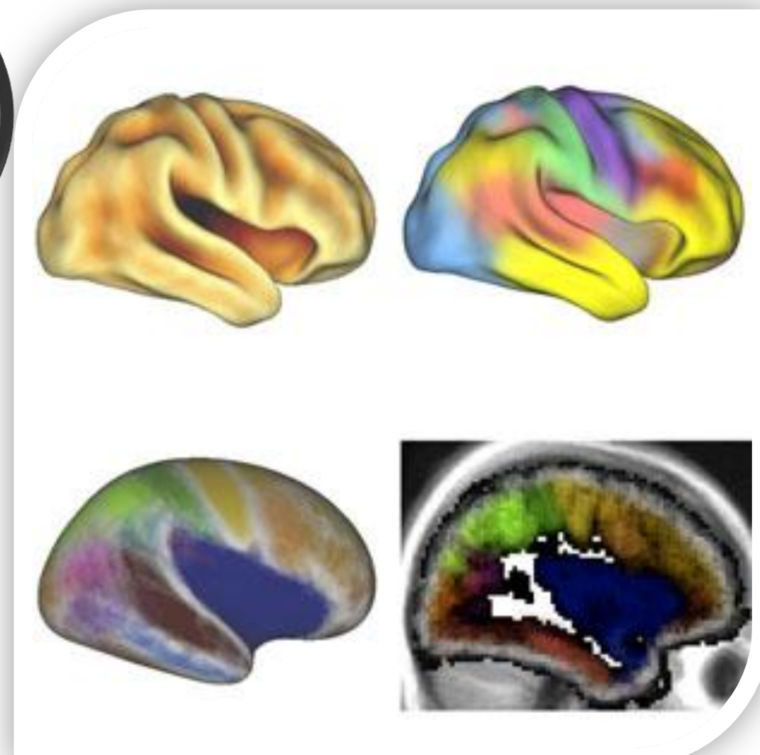
provides shared developmental structural and dti data

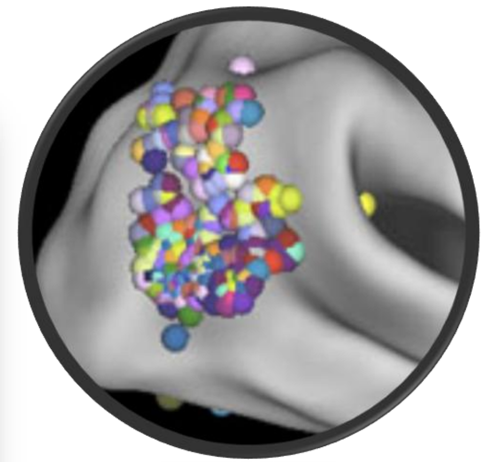
SumsDB

(System Management Surface Database): *A* web-acessible repository of structural and functional data







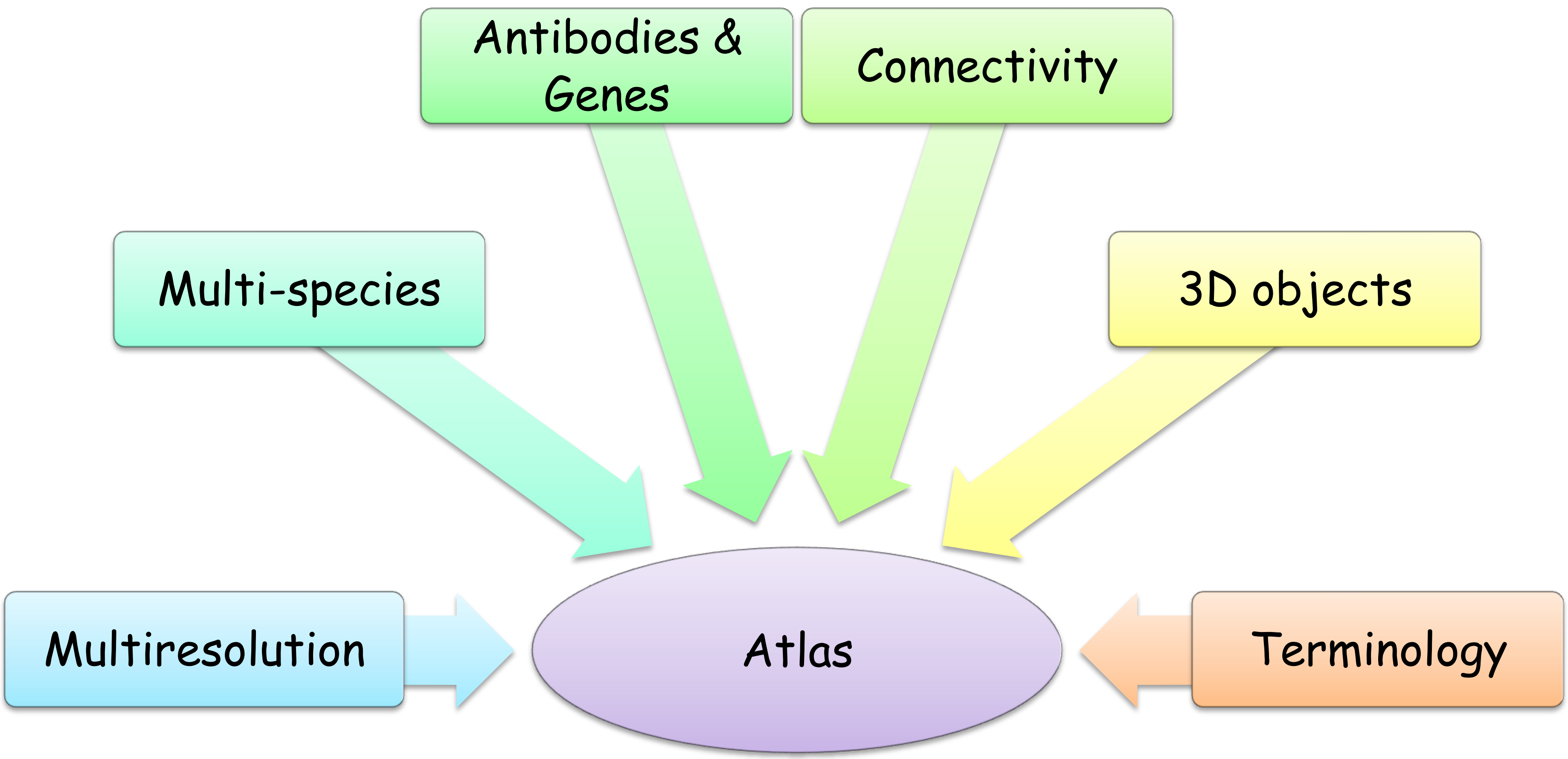


IBSR provides manually-guided expert segmentation results and MRIs to encourage the evaluation and development of segmentation methods.

"Next Generation" Brain Atlases

Back to Concepts

BrainMaps.org



BrainInfo

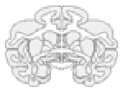
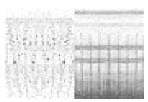
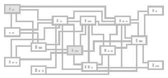
area 16 of Brodmann

Acronym: **16**

Search by Name of Brain Structure

Search by Browsing the Brain Atlas

Tools and Methods

What, Where and How Big is It? 	What is Written about It? PubMed
Show It! 	Other Names for It: Encephalon 脳 МОЗГ BRAIN cerebro cerveau Gehirn
Internal Structure 	Which Species Have It? 
What Cells Does It Have? 	Connectivity? 
Genes Expressed There 	

Metadata Search

Back to Concepts

Smart search tools use neuroscience ontologies to explore the neuroscience literature:

brainmap.org

home forum icbm2tal publications credits contact

Announcements - February 2009

Recently, the BrainMap database has been experiencing intermittent outages when using the Sleuth software. This most frequently occurs when downloading a large number of papers to a workspace, but has also been known to occur during the initial search step. We are working hard to determine the source of the problem and will try to get things fixed as soon as possible. If you see this error, try to simplify your search statement. In the meantime, we apologize for any inconvenience that this may cause.

What is BrainMap?

BrainMap is an online database of published functional neuroimaging (fMRI and PET) experiments with coordinate-based *(x,y,z)* activation locations in Talairach space. The goal of BrainMap is to provide a vehicle to share methods and results of studies in specific research domains, such as language, memory, attention, emotion, and perception. BrainMap can also be used to perform meta-analyses of similar research studies.

BrainMap was created and developed at the Research Imaging Center of the University of Texas Health Science Center San Antonio (UTHSCSA).

BrainMap Software

There are 3 different applications in the BrainMap software suite:

Quick Author Search

Want to check if a paper is already in the BrainMap database? Just type in the author's last name below:

Search

Current Database Status

Papers: 1711
Experiments: 7920
Paradigm Classes: 80
Locations: 64135

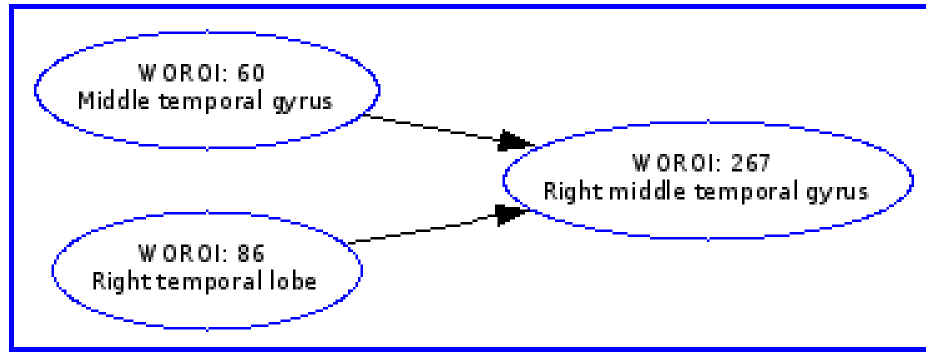
BrainMap Project Funding

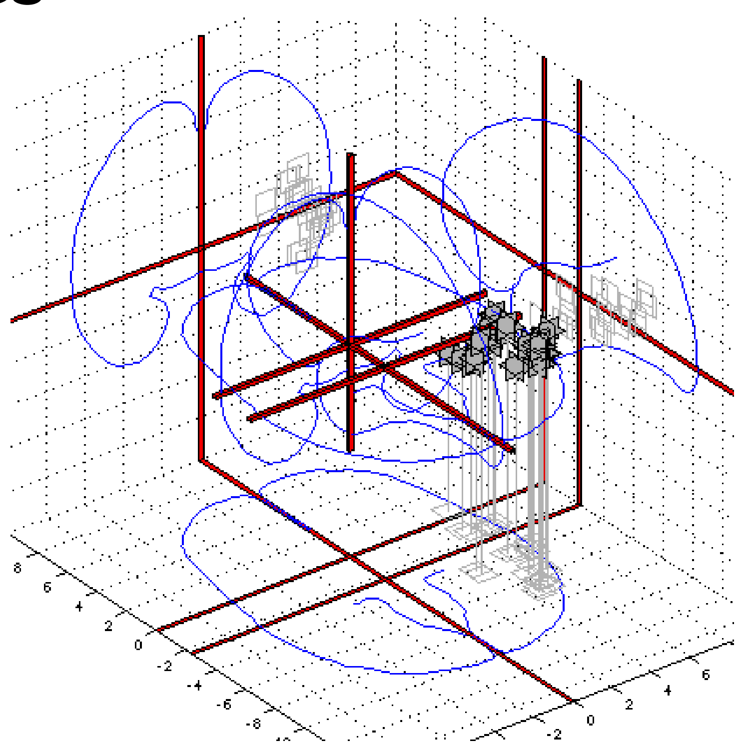
BrainMap is currently funded by the Human Brain Project of the National Institute of Mental Health (NIMH).

Current Software Versions

Brede Database: Functional Imaging Studies

A product of the Human Brain Project:
-Look up brain area->
1) Relevant passages
2) Related areas diagram
3) Activation loci.





Neurolex provides an online extensible lexicon and ontologies for neuroscience:

NeuroLex, the Neuroscience Lexicon

A dynamic lexicon of **7,148** neuroscience concepts supported

by The Neuroscience Information Framework

Hierarchies: Behavioral Activity • Behavioral Paradigms • Brain Regions • Cells • Diseases • Molecules • Nervous System Function • Subcellular Components • Information Resources • Resource Types • Qualities

Tables: Behavioral Activity • Brain Regions • Cell Types • Diseases • Molecules • Nervous System Function • Organism Synonyms • Resources and Information Entities • Qualities

All Categories A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

About • FAQs • Ontology Browser • NIFSTD ontologies • **How to Contribute**

Textpresso for Neuroscience

Search engine/database that uses neuroscience ontologies to extract information from articles
(You can search for concepts, not just keywords)

And organized lists of tools

NITRC
The Source for Neuroimaging Tools and Resources

A listing of neuroimaging software, links and ratings

Ratings & Reviews User Reviews (2)

Overall: **Installation:** **Documentation:**

- Category: Other Information Resource, Algorithm or Reusable Library
- Development Status: 5 - Production/Stable
- Intended Audience: Developers
- Natural Language: English
- Programming Language: C, Java, MATLAB
- Supported Data Format: NIfTI-1

Registered: 2007-05-17 11:54
Activity Percentile: 80.55%
View tool/resource activity statistics.
View a list of RSS feeds available for this tool/resource.
View images available for this tool/resource.

Summary

So, why should we care about neuroinformatics and data sharing?

We make sure data is protected, used wisely, and not lost or forgotten.

We get more citations, publications and collaborators.

We make meta-analyses possible

We provide data to develop the next generation of tools

The community get better atlases

- In sum, neuroinformatics offers a powerful, fiscally responsible way forward for neuroscience.

Rationale for Sharing

Sharing Principles

Data Curation

Ontology

XML

Data Repositories

Atlases

Smart Metadata Search