G-Node German Neuroinformatics Node

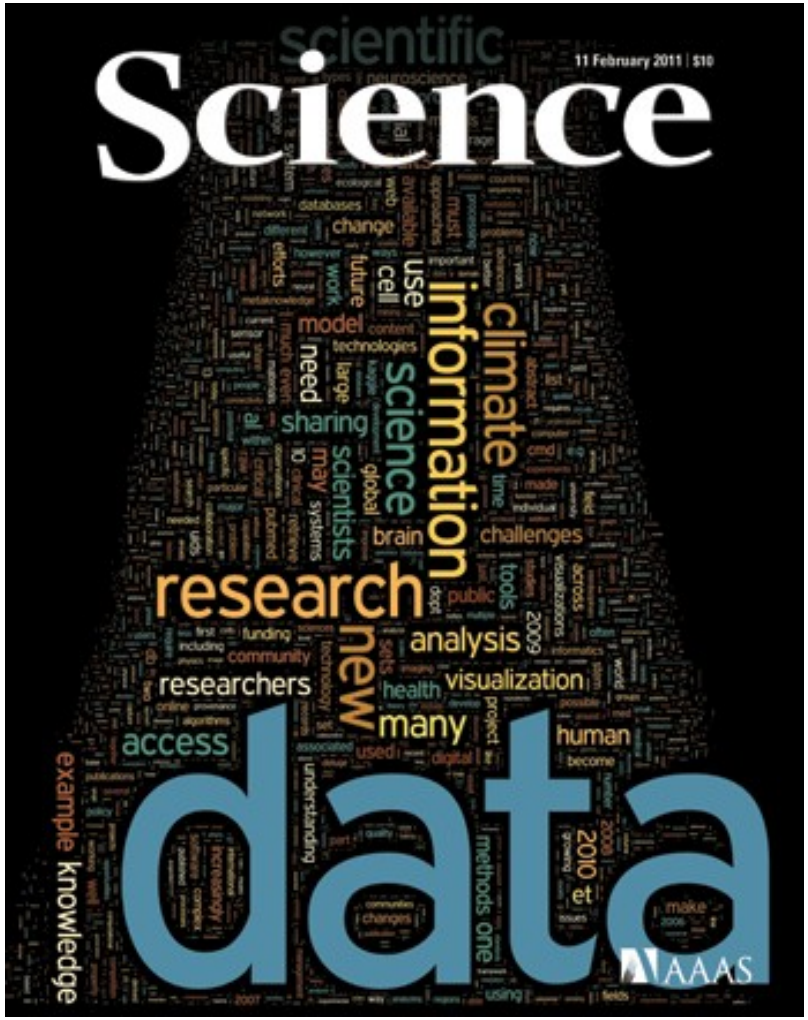
incf International Neuroinformatics Coordinating Facility

The German Neuroinformatics Node - Database tools for data management and collaboration in neurophysiology

Thomas Wachtler, Philipp Rautenberg, Andrey Sobolev, Andreas Herz

German Neuroinformatics Node
Ludwig-Maximilians-Universität München

Federal Ministry of Education and Research



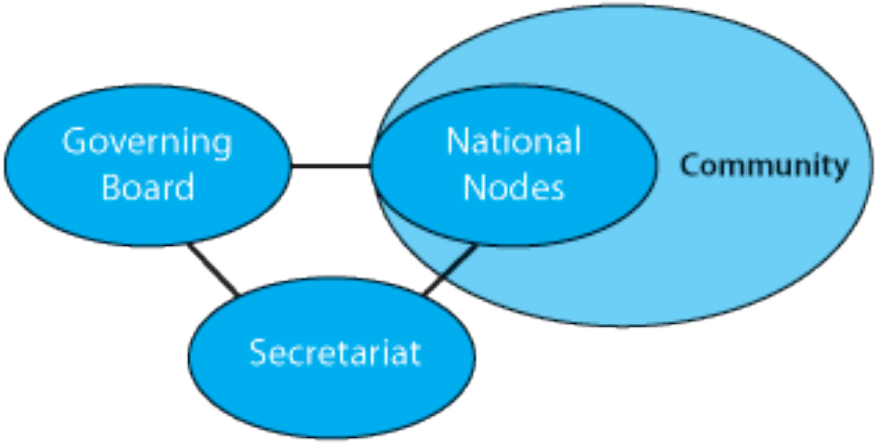
Science, Feb 11, 2011

International Neuroinformatics Coordinating Facility (INCF)

OECD Global Science Forum
Working Group on Neuroinformatics (1996-2002)

Recommendations for international development of neuroinformatics

➔ **International Neuroinformatics Coordination Facility** (INCF)
- established 2005
- INCF Secretariat in Stockholm
- National Nodes



www.incf.org

INCF Activities: Programs

INCF Programs address important scientific issues in specific areas of neuroinformatics. In these programs, INCF initiates and coordinates **international groups of scientists** to develop solutions in the form of **standards, guidelines, products or services**.

Current Programs:

- Digital Brain Atlasing
- Ontologies of Neural Structures
- Multi-Scale Modeling
- Standards for Data Sharing

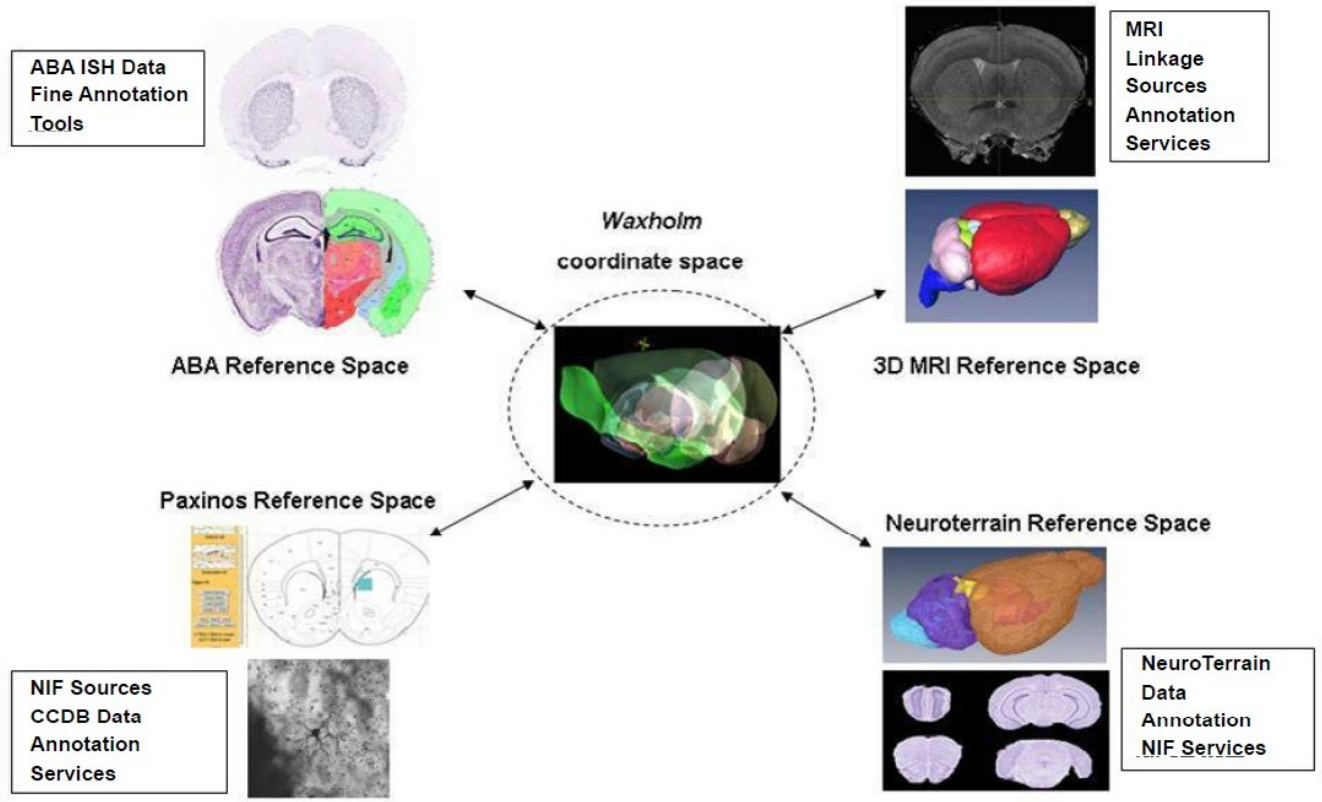
INCF Programs (I)

Program on Digital Brain Atlasing

Coordinate atlasing projects (rodent)

Standardization:
Waxholm Space

Establish Digital Atlasing Infrastructure



INCF Programs (II)

Program on Ontologies of Neural Structures

Establish platform for translation and clarification of terminologies

Integration of other ontology projects

Linked to Digital Atlasing program

Developments:

- **Structural Lexicon**
  Consistent definition of structural terms

- Neuronal Naming
  Neuron registry and convention for naming

INCF Programs (III)

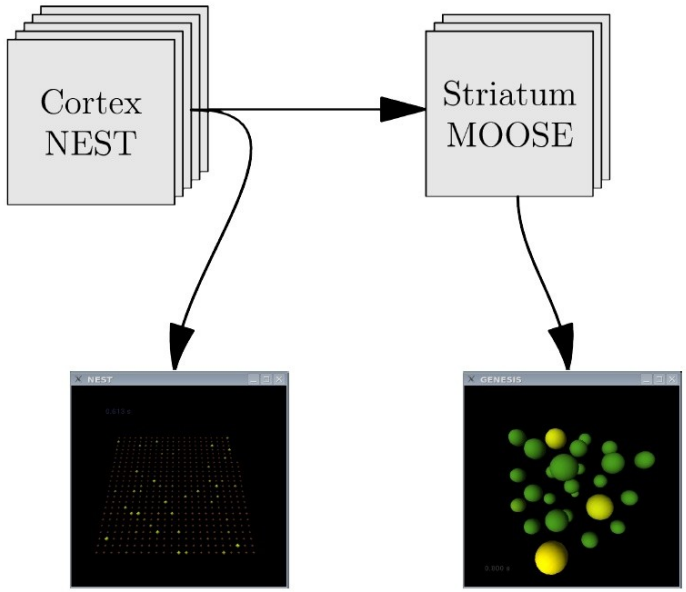
Program on Multi-Scale Modeling

Promote interoperability and sharing of model development and simulation tools

- Standards and Guidelines for Large-Scale Modeling:
  Language standard for spiking neuron models (**nineML**)

- **MUSIC Multi-Simulation Coordinator**
  Standardized interface for communication between neural network simulators

- Blue Gene/L Supercomputer Access

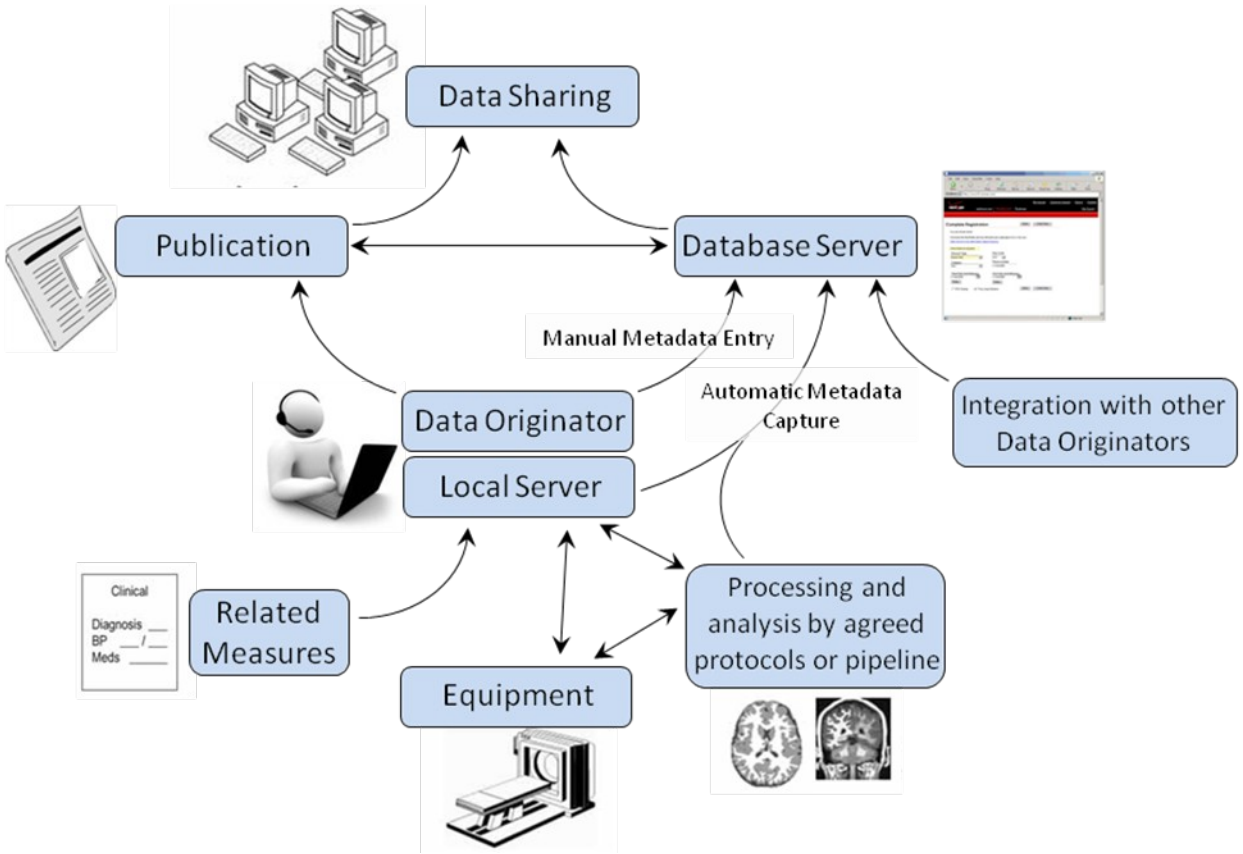


INCF Programs (IV)

Program on Standards for Data Sharing

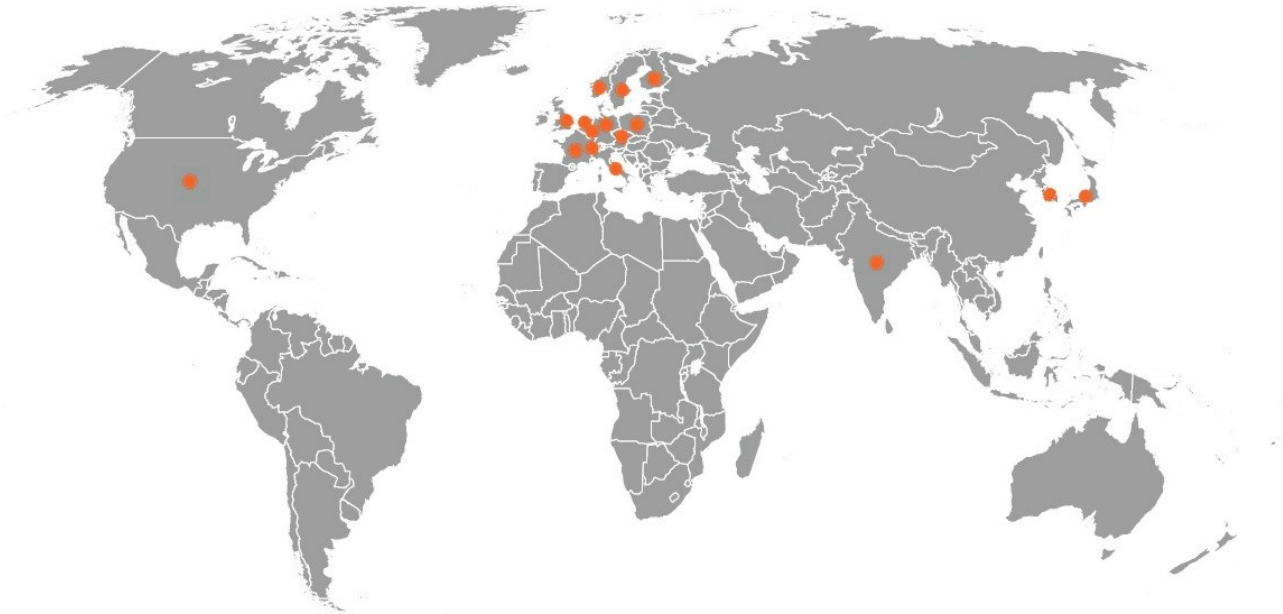
Development of methods for automated metadata acquisition and standards for interoperability between data management systems

Focuses on
- imaging data
- electrophysiology



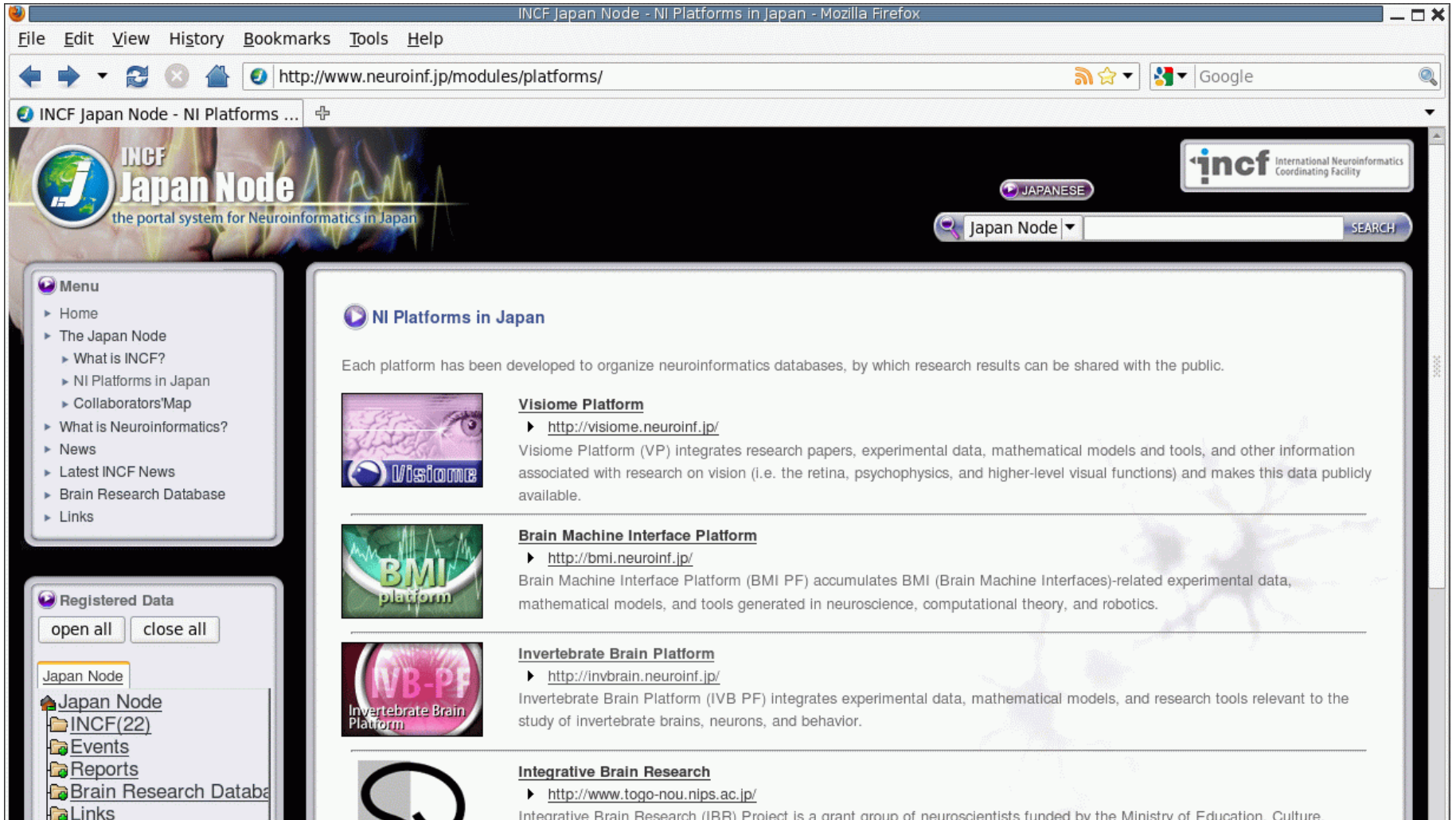
INCF National Nodes

Belgium
Czech Republic
Finland
France
Germany
India
Italy
Japan
Netherlands
Norway
Poland
South Korea
Sweden
Switzerland
United Kingdom
United States

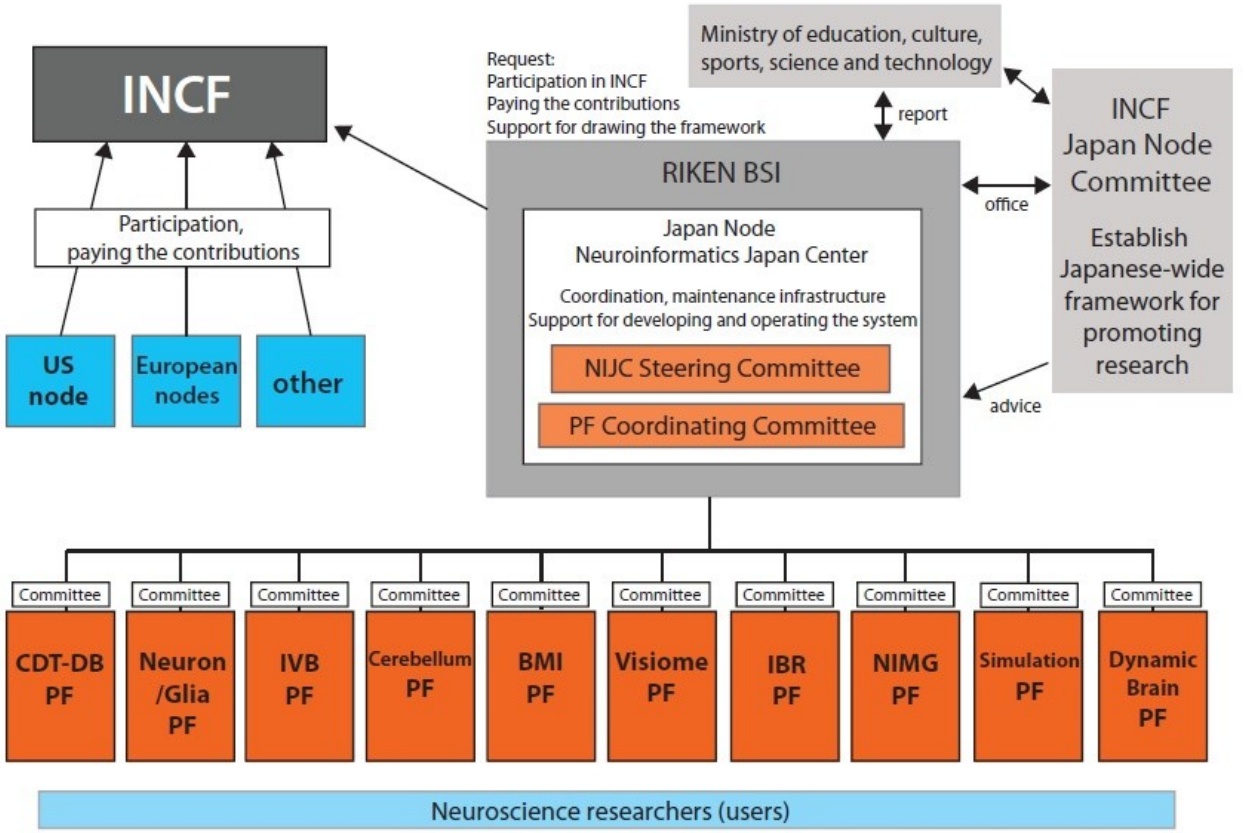


http://www.incf.org/about/organization/nodes

J-Node: Japanese Node of the INCF



J-Node: Japanese Node of the INCF

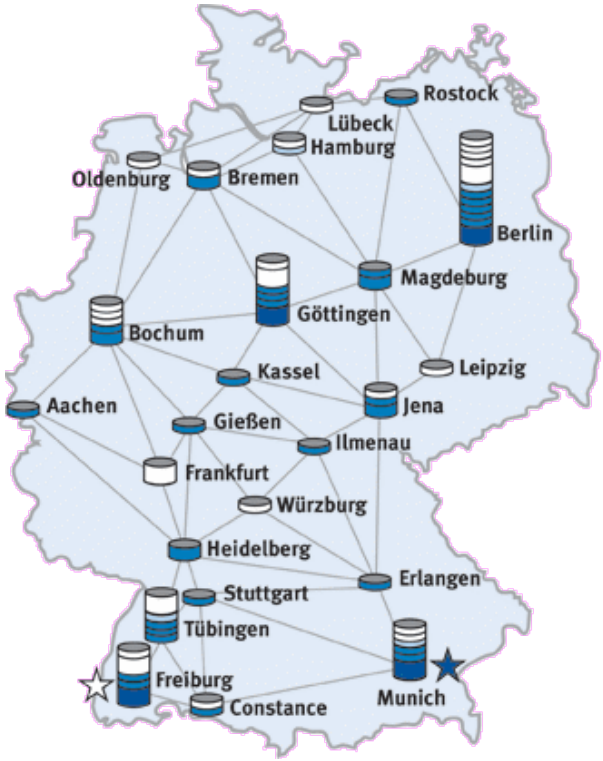


The structure and function of the J-Node

G-Node: German Node of the INCF

www.g-node.org

- Integral part of the National Network Computational Neuroscience
- Focus on cellular and systems neurophysiology
- Development of tools and infrastructure for data management and analysis, to facilitate collaboration and scientific progress
- Teaching & Training in neuroinformatics



Teaching & Training

Training in neuroinformatics for students and postdocs

- Courses on advanced data analysis methods
  next:
  3rd G-Node Winter Course on Neural Data Analysis
  March 7-11, 2011, Munich, Germany
- Courses on scientific programming techniques
  next:
  Advanced Scientific Programming in Python
  Sep 11-16, St Andrews, UK
  http://python.g-node.org
- Online teaching material

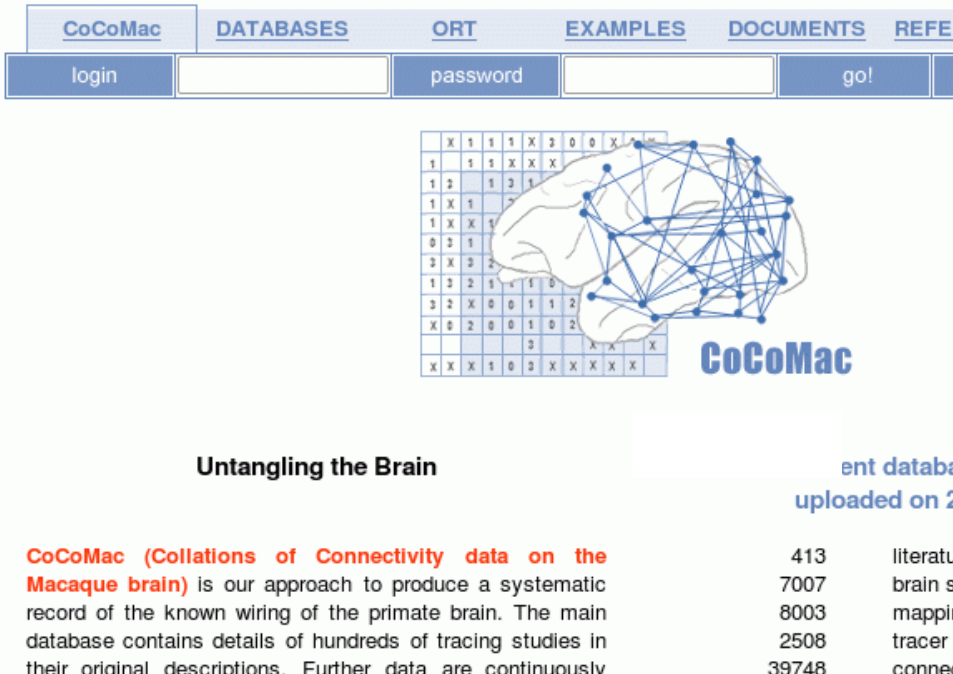
Services

Infrastructural and development support of neuroinformatics resources for cellular and systems neuroscience

Example: New version of **CoCoMac** database

(in collaboration with Rembrandt Bakker, Nijmegen and Markus Diesmann, Jülich)

- Web based access for automated queries by other applications
- Developing advanced user interface for community-driven development



Development of Tools for Efficient Data Management

Why do we care about data management?

- Progress in neuroscience increasingly depends on collaborative efforts, exchange of data, re-analysis of data
- Advances in technology and methodology dramatically increase volume and complexity of the data recorded
- Typically, each lab develops methods for managing and analyzing the data.
  - Duplicate work
  - Collaborative work or re-use of data is hampered by the effort it takes to understand how to access the data

Development of Tools for Efficient Data Management

What is needed?

- Tools for organizing and exchanging data that can be used by different laboratories

  ➔ save time and resources

- Standardized methods to specify the metadata that are necessary to access and analyze data

  ➔ ensure future access to data and facilitate collaborative work

- Integrated organization of data and metadata to facilitate data management and data analysis

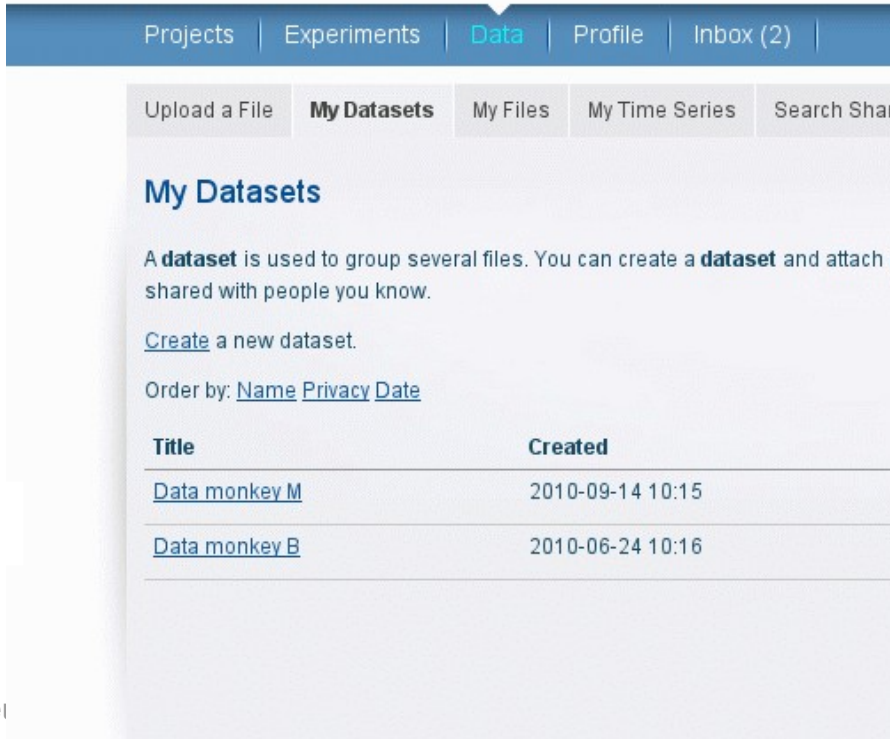
1 - Web-based Platform for Data Management and Data Sharing

(Django Web Framework, Python, postgreSQL Database)

Provides secure, web-based access to your data:

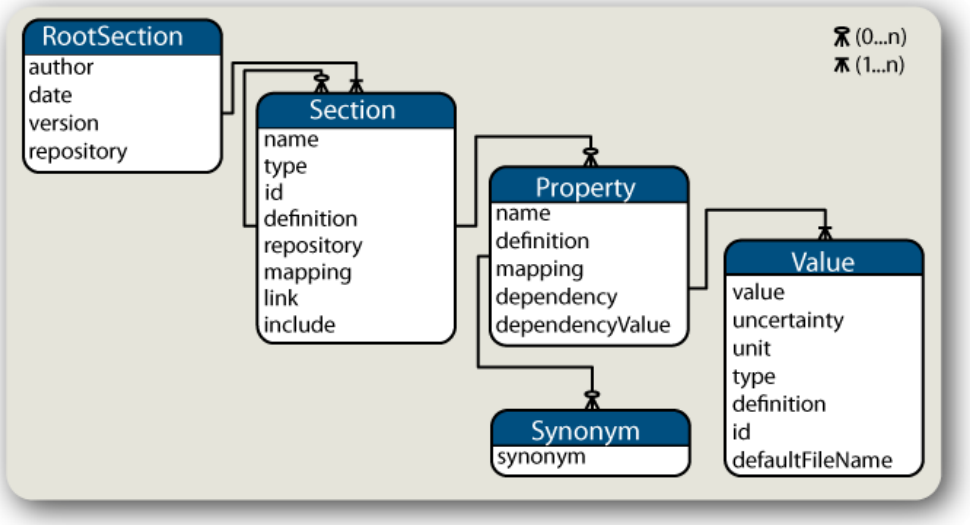
- Data Storage and Retrieval
- Data Organization
  - Projects - Experiments - Datasets
- Data Sharing
  - Private - Shared - Public

http://portal.g-node.org/data



2 - Method for Metadata Specification and Exchange

- Format: **metadata properties** are described in a **hierarchical structure of key-value pairs**, e.g., in XML

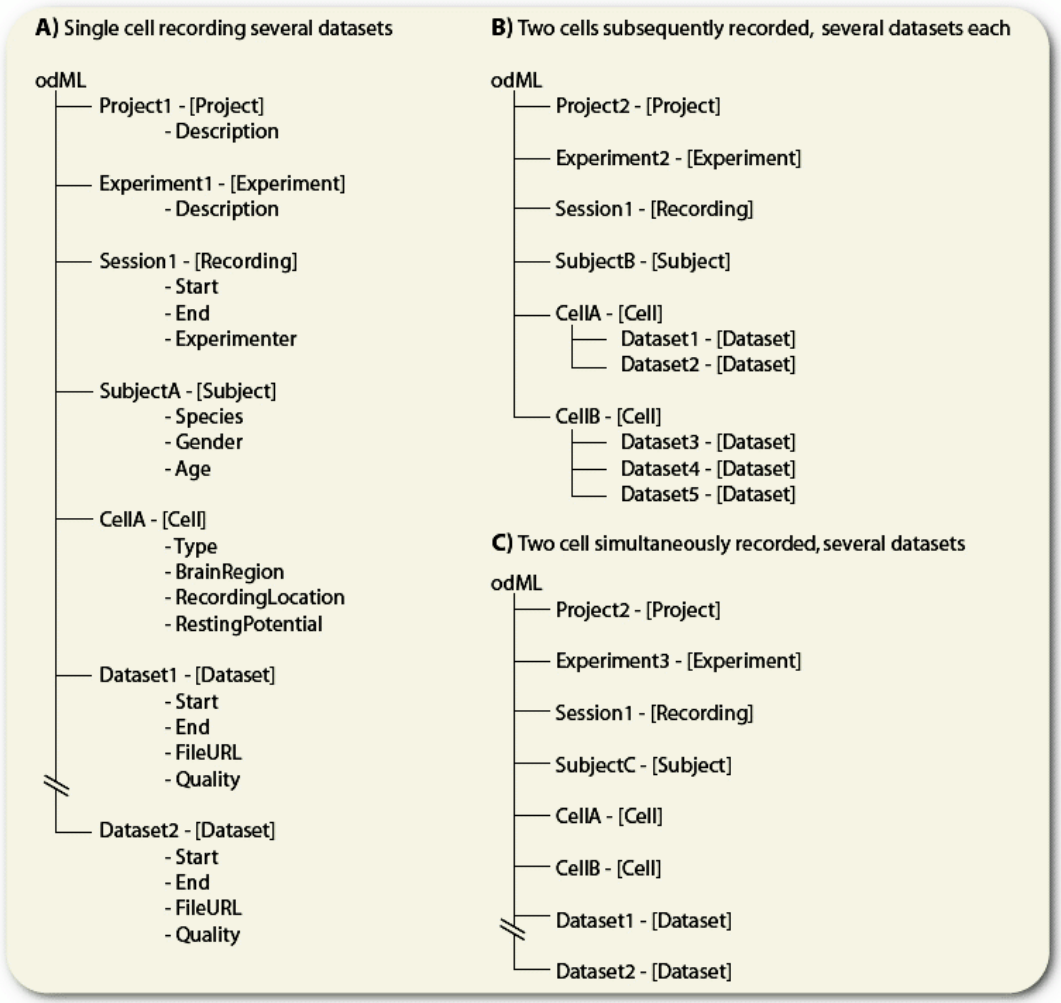


- Content: **recommended terminologies**

http://www.g-node.org/odml

Grewe et al, submitted

odML: Representing The Structure of Metadata



- format is **simple, flexible, inherently extensible**
  ➔ can be adapted to the specifics of the lab or experiment
- can carry **any metadata**
  ➔ no information is lost
- enables **automated collection of metadata** starting in the laboratory

http://www.g-node.org/odml

odML: Terminologies for Neurophysiology

Terminologies (available at www.g-node.org/odml)

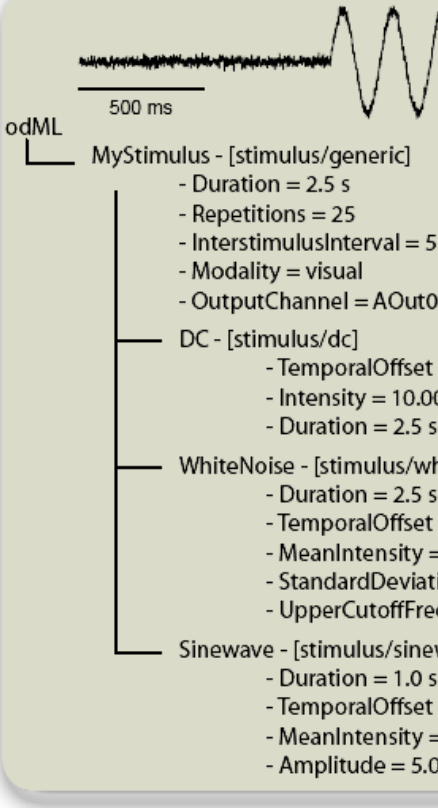
name	type	description
Analysis	analysis/generic	Descriptions of an analysis.
PSTH	analysis/psth	Properties to describe a peri stimulus time histogram.
PowerSpectrum		
Coherence		
Cell		
Dataset		
Electrode		
Event		
EventList		
Experiment		
Behavior		
Electrophys		
Imaging		
Psychophys		
Hardware		
Amplifier		
Attenuator		
CameraObje		
DataAcquisi		
Eyetracker		
Filter		
FilterSet		
ImageAcqui		
Light Source		
Microscope		
Microscope-		
Objective		
Scanner		
StimulusIsol		
HardwareProperties		
HardwareSettings		
Person		
Preparation		
Project		

Stimulus Terminology

name: Stimulus
type: stimulus/generic
baseURL: http://g-node.org/odml/terminologies/v1.0/
definition: Section to describe a generic stimulus. This section is basis of various related sections that specify more specific stimuli.

Properties:

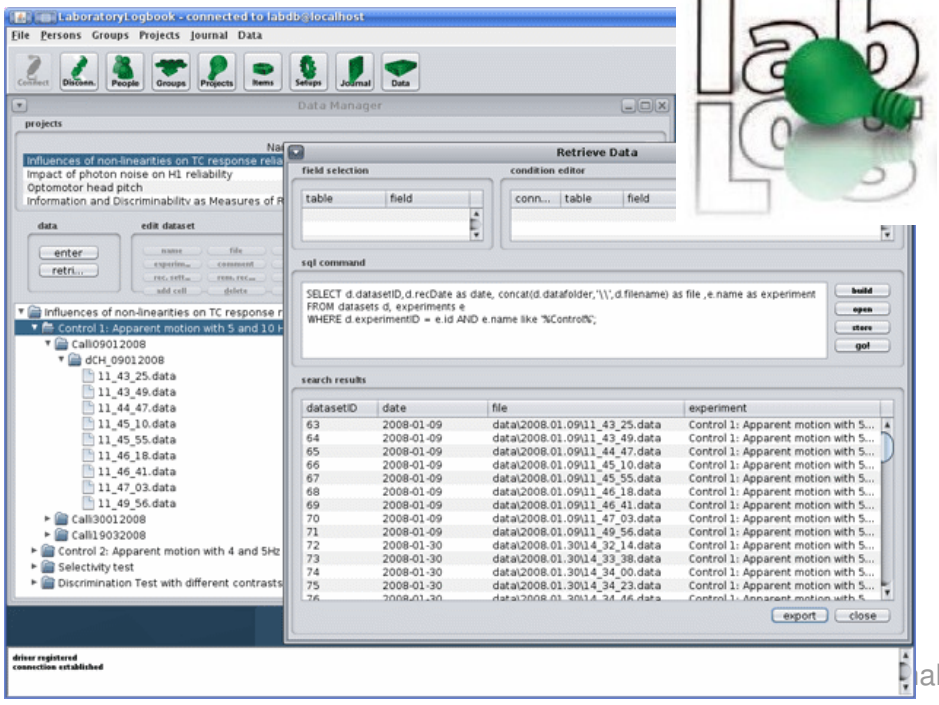
name	type	unit	definition
Author	string	-	The author of this stimulus.
Description	string	-	A textual description of the stimulus.
Duration	float	s	The duration of the stimulus in seconds.
StartTime	time	-	The time the stimulus started.
EndTime	time	-	The time the stimulus ended.
Interstimulus-Interval	float	s	The time between successive stimulus presentations.
Repetitions	int	-	The number of stimulus presentations.
Intensity	float	-	The intensity of the stimulus.
Modality	string	-	The stimulus modality, e.g. visual, auditory, etc.
OutputChannel	string	-	The physical output device (e.g. an analog output channel, a monitor screen, a loudspeaker, etc.). This information can be used to explicitely express that several stimuli (that share the same OutputChannel) are jointly presented.
StimulusFile	binary	-	The stimulus file used. Including it will increase the size of the metadata file and we recommend to rather point to a file using an URL.
StimulusFileURL	URL	-	The URL of the applied stimulus file.



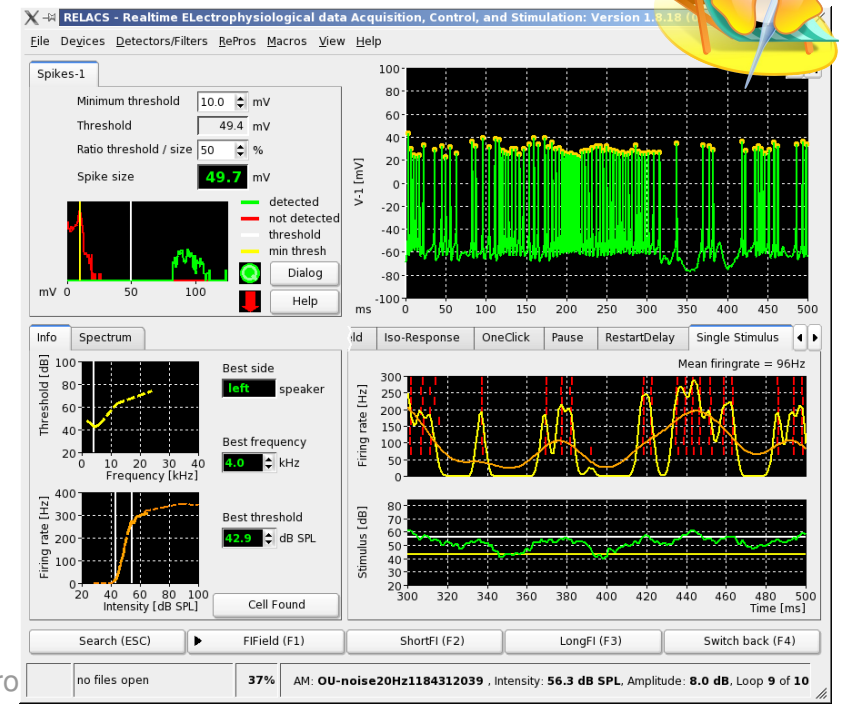
odML: Automated Collection of Metadata

via libraries provided (Java, C++, Python), odML can be integrated into data acquisition, stimulation, and analysis tools, e.g.:

LabLogBook (Jan Grewe, Munich)
http://lablog.sourceforge.net

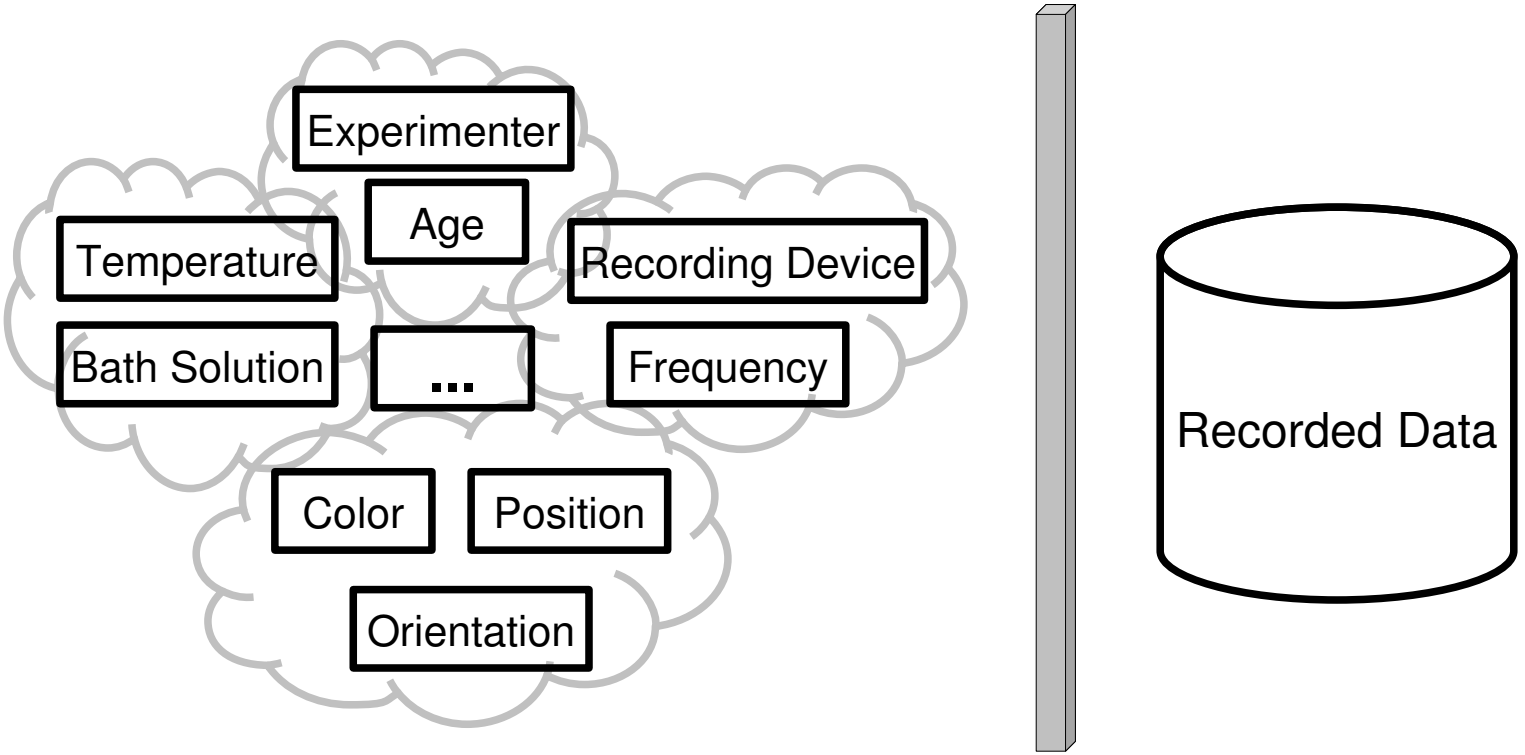


Relacs (Jan Benda, Munich)
http://www.relacs.net



3 - Development of Integrated Database System for Data Management

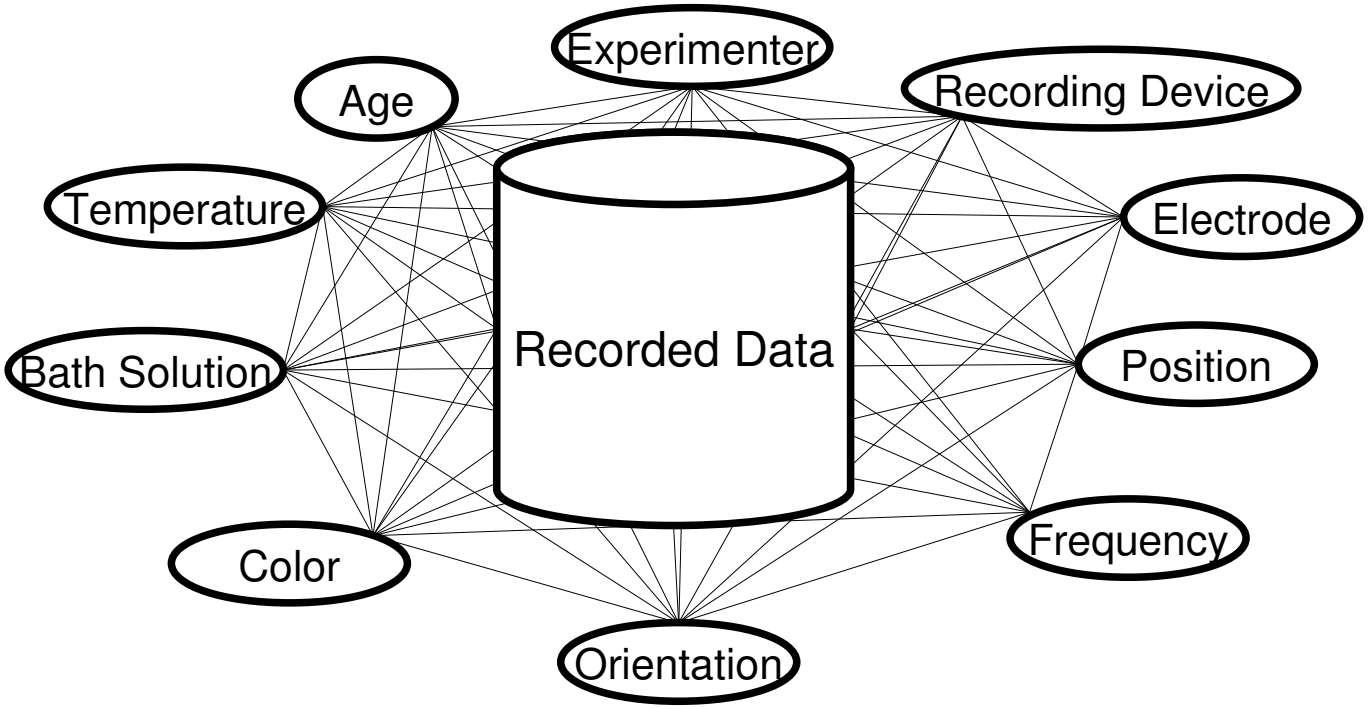
How to organize your data?



- The answer depends on how you want to control your data -

Development of Integrated Database System for Data Management

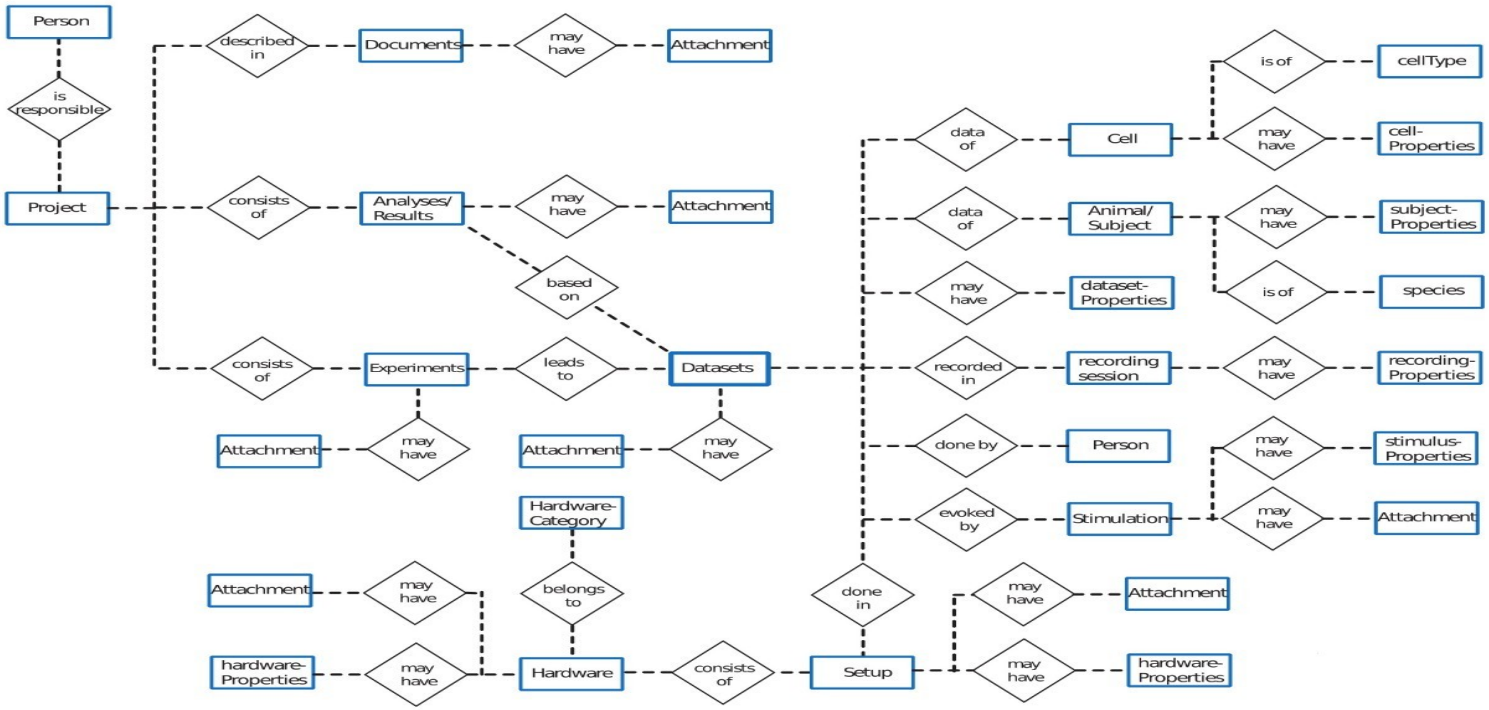
How to organize your data?



Organization of data and metadata should represent the structure of the experiment

Development of Integrated Database System for Data Management

Goal: Database schema for neurophysiology experiments



What is the benefit?

Development of Integrated Database System for Data Management

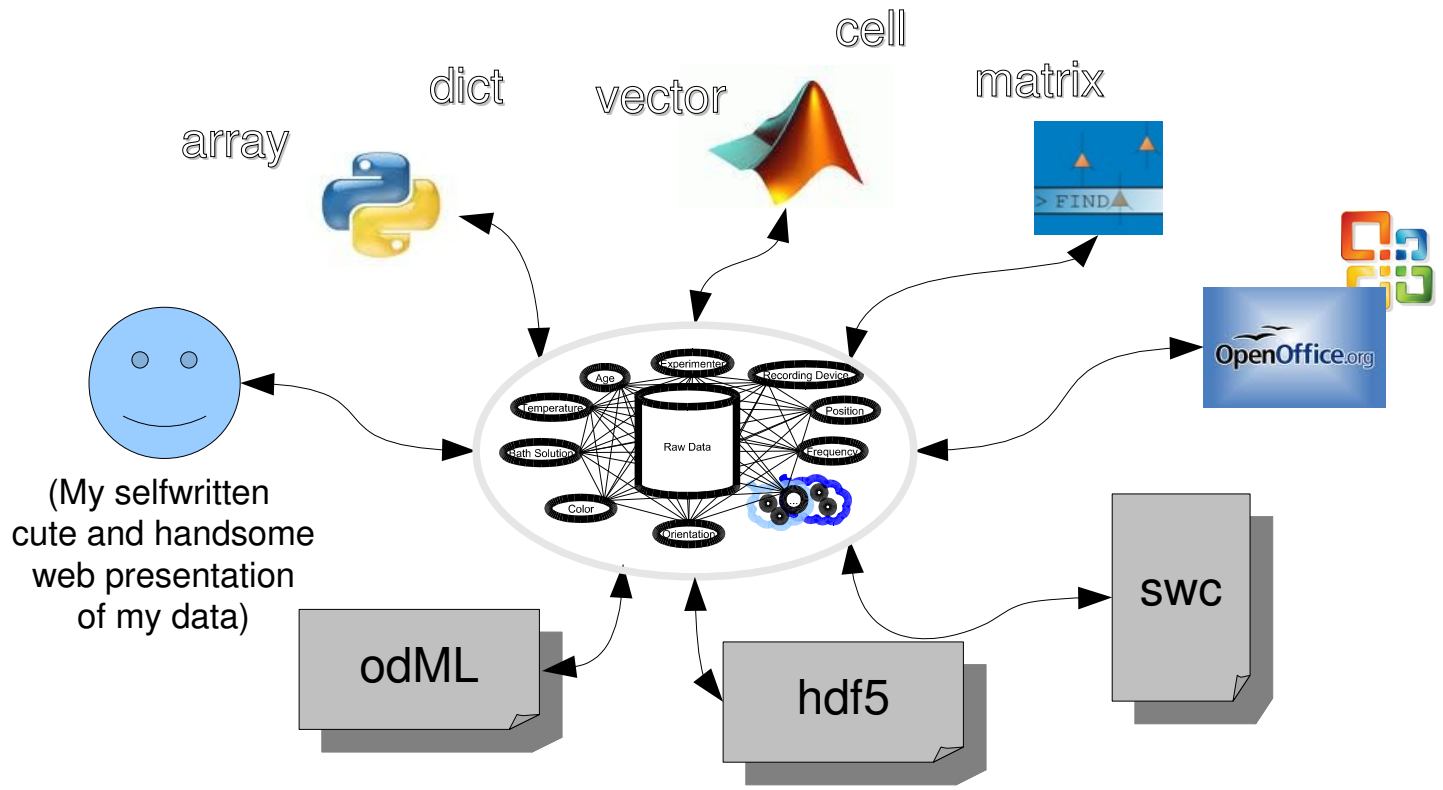
Don't search - ask!

„Give me all spike trains where

- age=5 AND
- stimulus color=red ... AND ...
- ... OR ...“

Integrated organization of data and metadata in database enables **efficient access and automated selection** of data for analysis

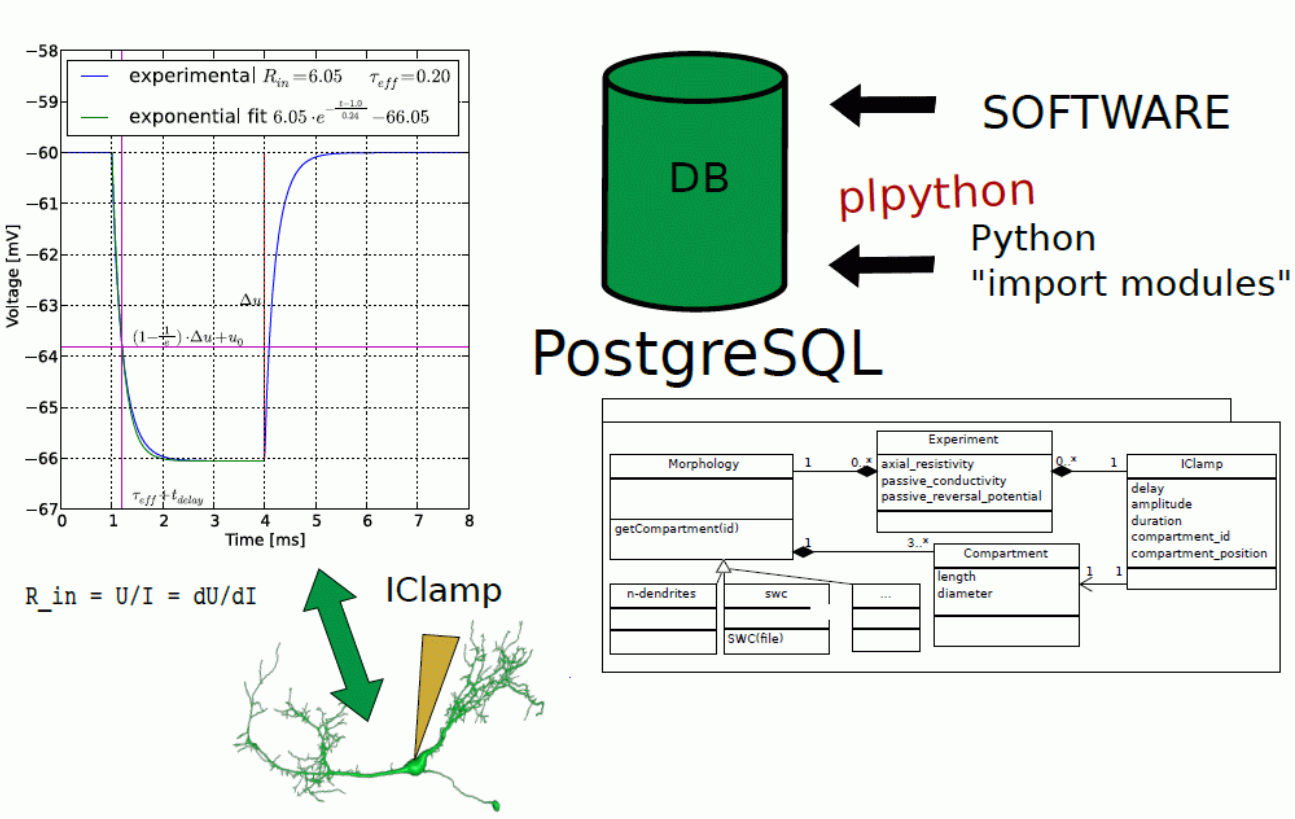
Development of Integrated Database System for Data Management



Interfaces to various applications, languages and formats enable **automated selection of data** and **integration with data analysis** workflow

Development of Integrated Database System for Data Management

Extension: Integration of morphological data and simulation tools



see Poster: Rautenberg et al

Summary

- Key to efficient data management is the unified access to data and metadata
- Using database technology, we develop methods for neurophysiology data management that facilitate data access and data exchange.
- Interfaces to common software tools and programming languages will enable automated data selection and integration in the data analysis work flow

Acknowledgments

Hiroyuki Ai, Rembrandt Bakker, Jan Benda, Philip Berens, Holger Blasum, Andrew Davison, Piotr Durka, Alexander Ecker, Felix Franke, Alexander Gail, Christian Garbers, Samuel Garcia, Jan Grewe, Nina Hubig, Hidetoshi Ikeno, Christian Kellner, Christian Klaes, Philipp Meier, Eilif Müller, Matthias Munk, Roman Moucek, Martin Nawrot, Sebastian Philipp, Raphael Ritz, Jürgen Rybak, Willi Schiegel, Michael Schmuker, Christine Seitz, Adrian Stoewer, Michael Stransky, Zbysek Szmek, Christian Tatarau, Alvaro Tejero Cantero, Tiziano Zito

Federal Ministry of Education and Research