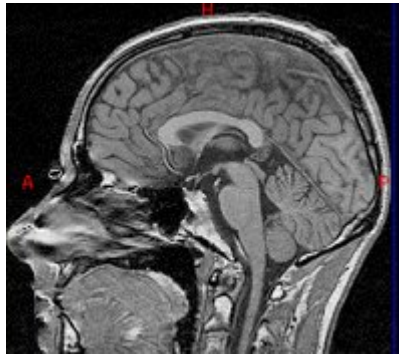
Neurobase: Sharing data and image processing tools in neuroimaging

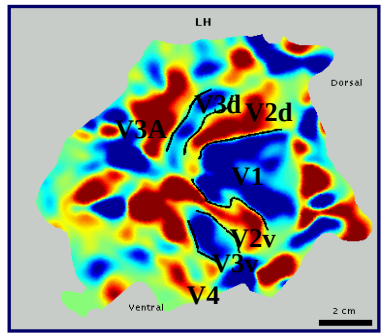
Bernard Gibaud[1], Christian Barillot[1], Eric Simon[2]
Florent Aubry[3], Habib Benali[4], Olivier Dameron[5], Michel Dojat[6], Alban Gaignard[1], Serge Kinkingnéhun[4], Jean-Pierre Matsumoto[2], Mélanie Pélégrini-Issac[4], Lynda Témal[1], Romain Valabregue[1]

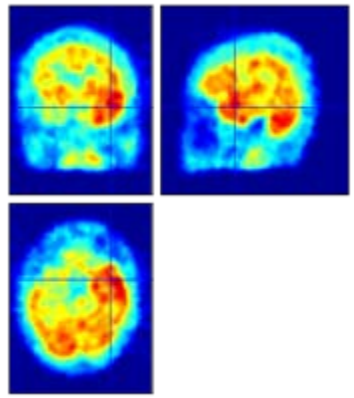
[1]VisAGeS, U746 Inserm/INRIA, IRISA, Rennes
[2]Business Objects & Medience SA
[3]U455 Inserm, Toulouse
[4]IFR 49, Paris et Orsay
[5]Ex-IDM, Inserm ERI1, Rennes
[6]U594 Inserm, Grenoble

Context: research in neuroimaging

- Neuroimaging: to highlight the brain's morphology, functionality, physiology, metabolism, under normal and pathologic conditions
- Application
  - Basic research
    - Mapping of brain functions
    - Modeling of cognitive functions, e.g.: vision, motor, language, memory, etc.
  - Clinical applications
    - Multiple sclerosis, epilepsy, etc
    - Dementia and neuro-degenerative pathologies : Alzheimer, Parkinson, Stroke, etc.







Context: ressources produced by research in neuroimaging

- Data, denoting knowledge about brain
  - Functional maps
  - Morphological and physiological abnormalities related to the various brain diseases
  - Behavioral data
- Know-how
  - Exploration methods: paradigms, imaging techniques e.g. specific MR sequences…
  - Image processing tools
    - Segmentation, registration, quantification, etc.
    - Statistical analysis
  - Image processing pipelines
    - Suitable for a specific problem

Neurobase : general objectives

- Optimise collaborative work in neuroimaging, through the sharing of data and image processing tools
- in order to:
  - carry out large scale experiments
  - re-use existing image processing tools
  - validate new image processing tools
    - Access to validation data sets
    - Comparison to existing processing tools

Contraints

- Legitimate need for autonomy of the collaborating centres, regarding
  - Local organisation of the data
  - Sharing policy
- Data confidentiality
  - Compliance to existing regulation

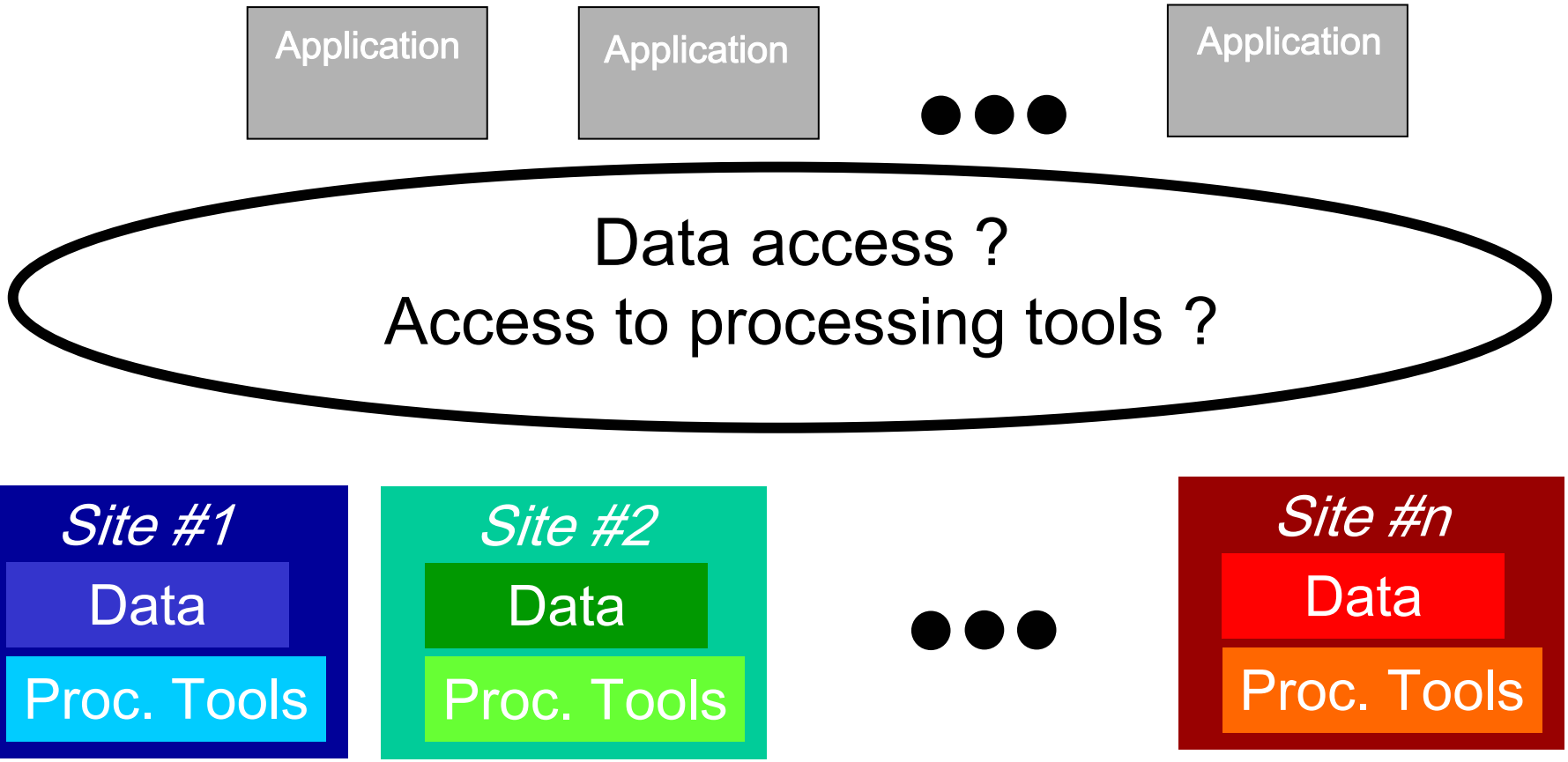
Neurobase
2002-2005: exploratory phase

- Creation of a suitable architecture
  - to access distributed & heterogeneous data
  - to integrate distributed components into specific pipelines
- With two specific objectives
  - Creation of a domain ontology (*semantic reference*)
  - Implementation of a demonstrator (to share data and image processing tools)

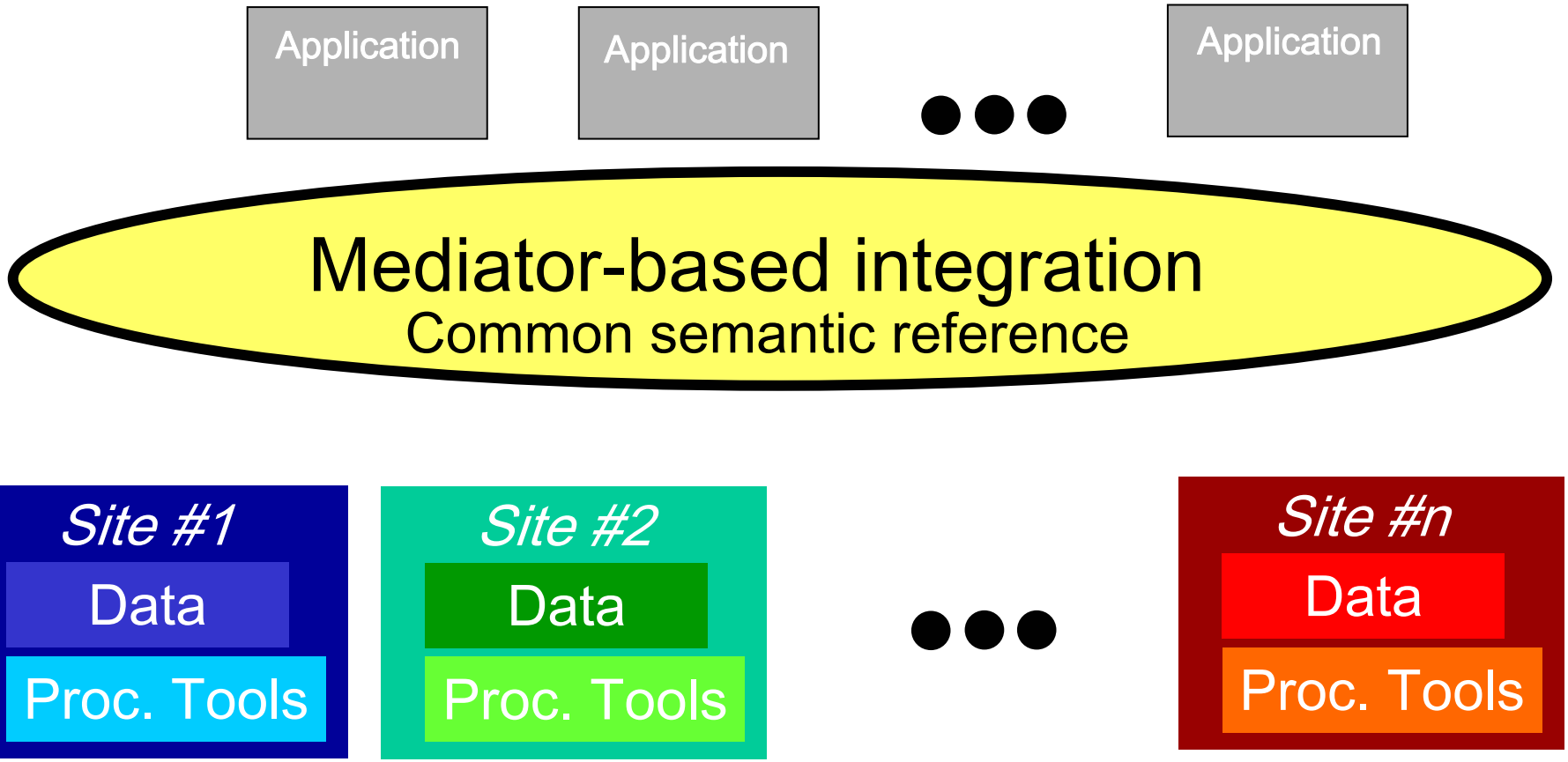
Partners

- VisAGeS et TeXMex projects, IRISA, Rennes
- IFR 49 « Functional Neuroimaging » Paris & Orsay (CEA-SHFJ, INSERM U678, CHR Pitié Salpétrière)
- CARAVEL Project, INRIA Rocquencourt / Medience SA
- INSERM U594, Grenoble
- IDM Lab., Fac of Medicine, University of Rennes I
- TIMC Lab. (SIC team), Grenoble
- EPIDAURE project, INRIA Sophia-Antipolis

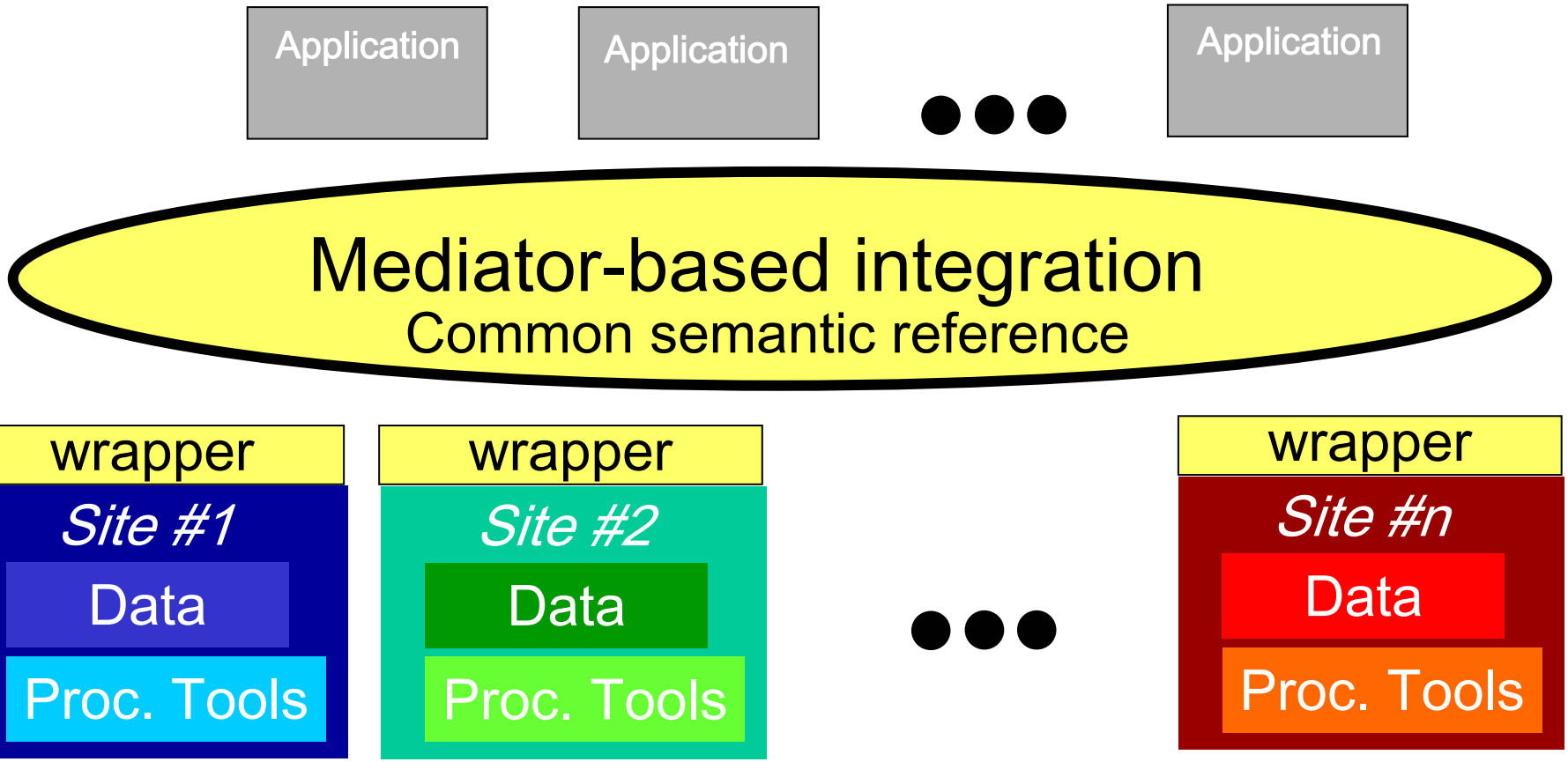
Federated system



Approach


Application
Application
Application
Mediator-based integration
Common semantic reference
Site #1
Data
Proc. Tools
Site #2
Data
Proc. Tools
Site #n
Data
Proc. Tools

Approach


Application
Application
Application
Mediator-based integration
Common semantic reference
wrapper
Site #1
Data
Proc. Tools
wrapper
Site #2
Data
Proc. Tools
wrapper
Site #n
Data
Proc. Tools

Common ontology

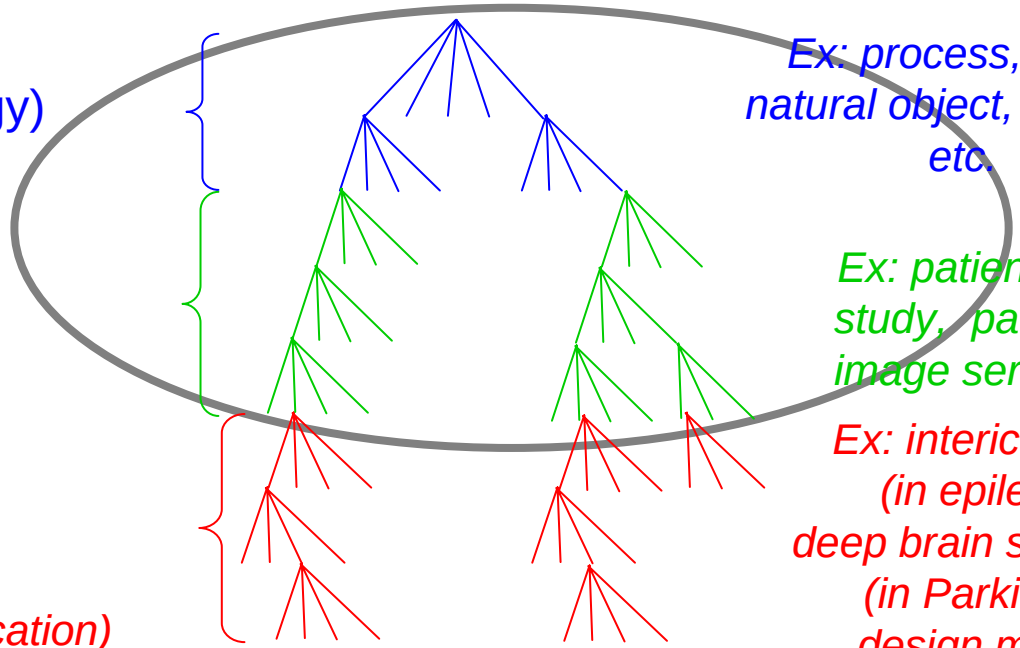
- « a formal, explicit specification of a shared conceptualization » (Gruber 1993)
  - Necessary to write applications and wrappers (entities, range of values)

Ontology: scope

General concepts
(upper level ontology)
(non specific)

General concepts
of domain of interest
(domain-specific)

All other concepts
of domain of interest
(i.e. to deploy in a real life application)



Ex: process, state, natural object, artefact, etc.

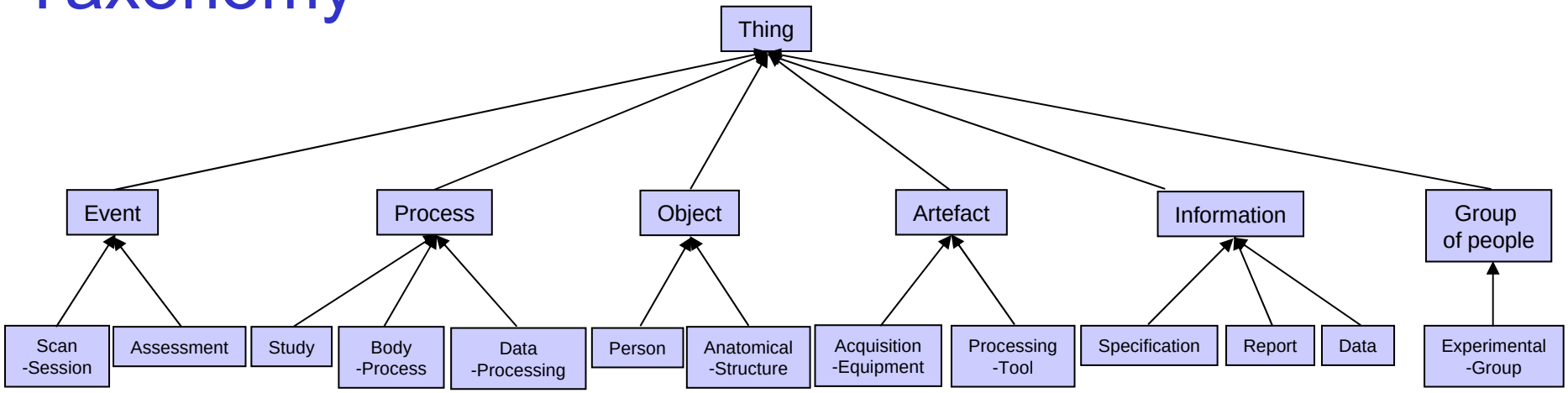
Ex: patient, scan, study, pathology, image series, etc.

Ex: interictal state (in epilepsy), deep brain stimulation (in Parkinson), design matrices (fMRI), etc.

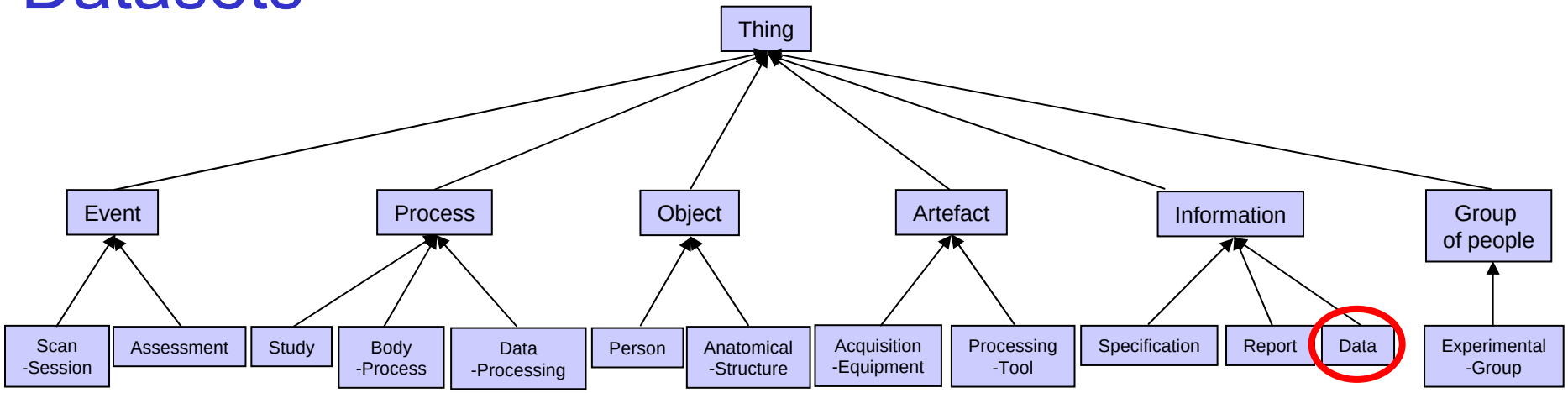
Neurobase ontology

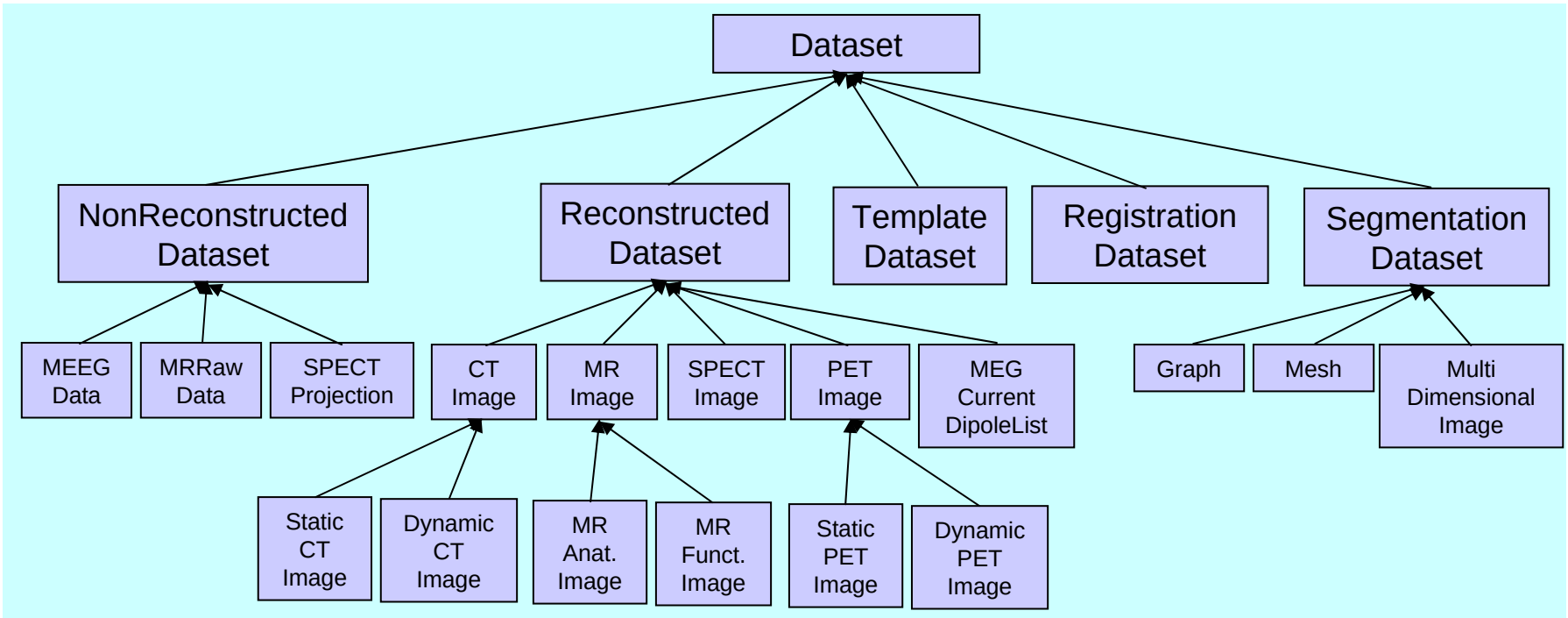
- Scope
  - Studies (subjects, experimental context, clinical aspects, etc.)
  - Datasets and description of their content (images, ROI, registration data, etc.)
  - Image processing (processing tools, processing, etc.)
- Method
  - Integration of multiple sources (fMRIDC, DICOM, Neurobase partners' experience)
  - Representation : UML, then Protégé

Taxonomy


Thing
Event
Process
Object
Artefact
Information
Group of people
Scan -Session
Assessment
Study
Body -Process
Data -Processing
Person
Anatomical -Structure
Acquisition -Equipment
Processing -Tool
Specification
Report
Data
Experimental -Group

Datasets


Thing
Event
Process
Object
Artefact
Information
Group of people
Scan -Session
Assessment
Study
Body -Process
Data -Processing
Person
Anatomical -Structure
Acquisition -Equipment
Processing -Tool
Specification
Report
Data
Experimental -Group


Dataset
NonReconstructed Dataset
Reconstructed Dataset
Template Dataset
Registration Dataset
Segmentation Dataset
MEEG Data
MRRaw Data
SPECT Projection
CT Image
MR Image
SPECT Image
PET Image
MEG Current DipoleList
Graph
Mesh
Multi Dimensional Image
Static CT Image
Dynamic CT Image
MR Anat. Image
MR Funct. Image
Static PET Image
Dynamic PET Image

Demonstrator

- Web Application
  - Dataset selection based on user-defined criteria (subjects, studies, datasets)
  - Execution of « dataflows »
  - Display of results

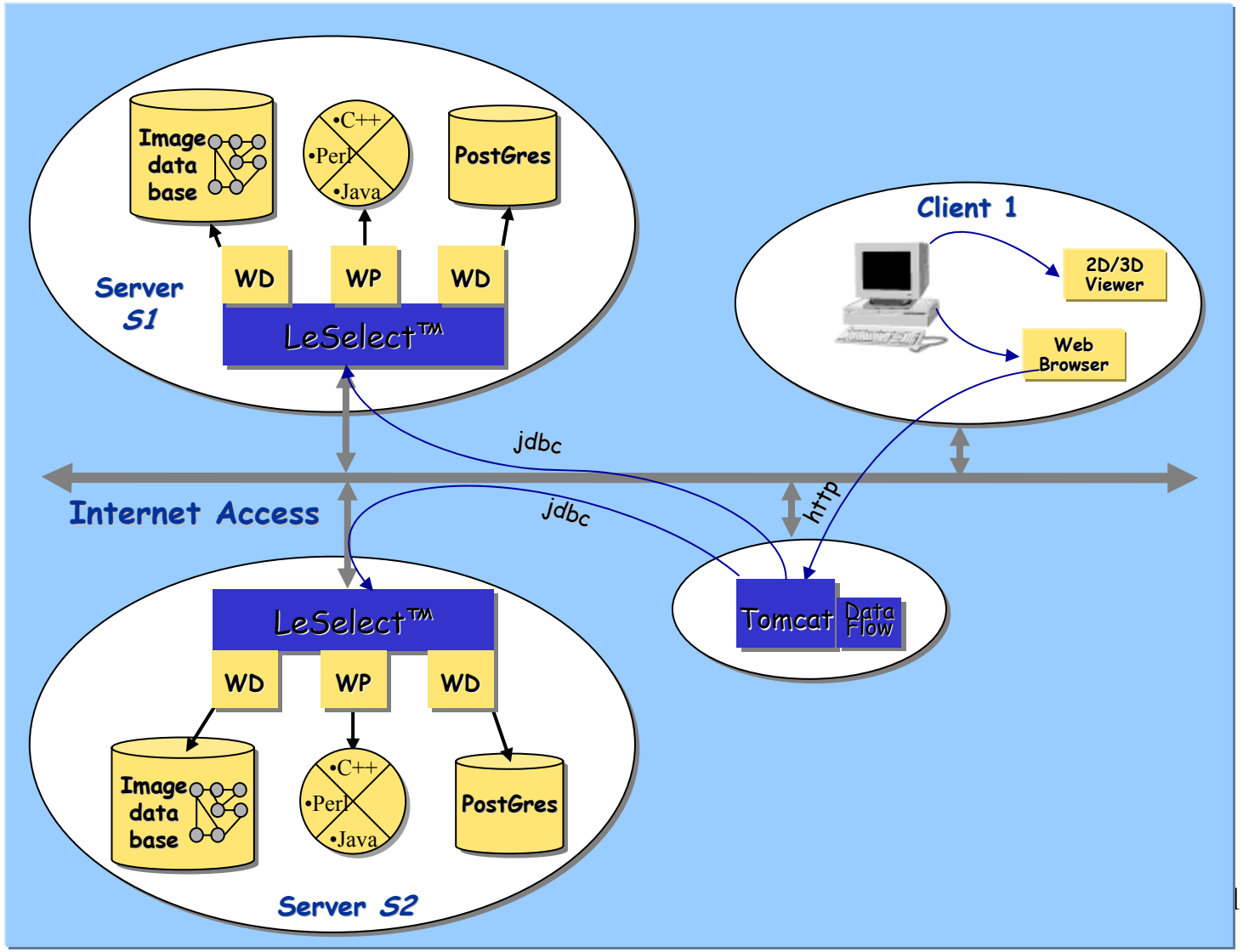
  → Software environment: Servlet container Tomcat
- Mediation system
  - Publish and access the data
  - Invoke the programs

  → based on « Le Select » (Medience SA)

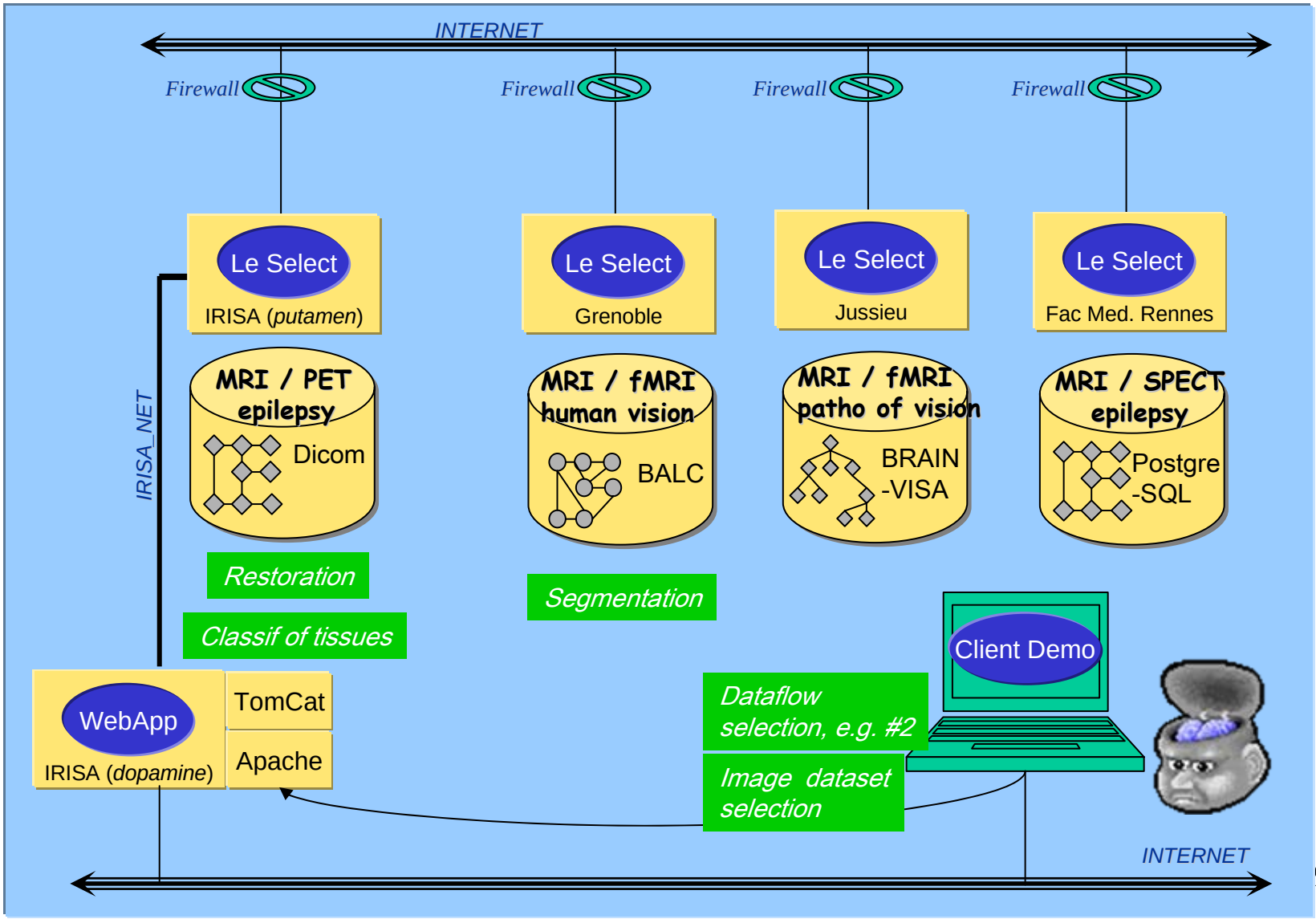
Le_Select (Medience SA)

- Initially developed in CARAVEL (Inria, Rocquencourt)
- Major features
  - Uniform access to distributed heterogeneous data
  - Application of transformations to data
  - Data « published » according to a relational model, and accessed in SQL (e.g. using JDBC)
  - Invocation of data processing programs on arbitrary datasets
  - Fully distributed

Demonstrator architecture



Demonstrator deployment: 4 sites



Discussion - ontology

- Need to evolve toward a formal ontology
  - i.e. expressed in a logical language (e.g. OWL)
  - Necessary for :
    - Management of « intelligent » queries
    - Wrappers
- Articulated to formal and consensual « upper layer ontologies »,
  - e.g. DOLCE (Wonderweb) or BFO (Barry Smith *et al.)*
  - interoperability with external terminology systems, e.g.:
    - Unified Medical Language System (UMLS, NLM)
    - Foundational Model of Anatomy (FMA, UW Seattle)
- Difficult trade-off
  - Complexity / practical usability

Discussion: demonstrator

- Deployment of a real-life application (e.g. epilepsy or multiple sclerosis) would be a difficult challenge
  - Significant implementation effort
- Mediation software
  - powerful, but limited
  - difficult to control (by a user community)
    - Need for close partnership with research/industry
    - same difficulties were reported in other projects, e.g. in the US, such as BIRN: Biomedical Informatics Research Network (NCRR, NIH)

Conclusion / perspectives

- Experience from exploratory phase
  - very positive
  - potential impact is important in neuroimaging, but also in other fields e.g. genomics, or cancer research (CaBIC)
- Work to be pursed
  - Ontology
  - Orchestration of dataflows
- Work on a real-life application necessary
  - Major question : what mediator ?
    - « Le Select » ?
    - or « general-purpose » ? such as a GRIDs Toolkit : Globus (GGF), SRB ?

Acknowledgements

- Ministry of research (ACI)
- Regional Council of Brittany
- All other Neurobase contributors