



## Neuro-imaging Informatics:

**Neil Killeen**

The Centre for Neuroscience  
University of Melbourne,  
Victoria 3010, Australia

## Talk Outline


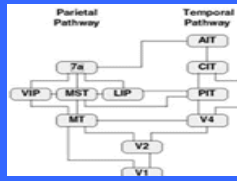
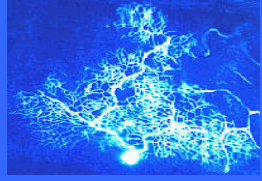
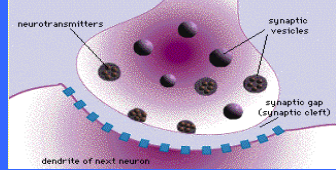
- Neuro-Informatics/Neuro-imaging Informatics
- Components of System and Workflow
- Deployment
- Summary and Future Work
- Collaborators
  - Gary Egan (Howard Florey Institute), Iven Mareels (University of Melbourne)
  - Wilson Liu, Wee Siong Soh (Howard Florey Institute)
  - Jason Lohrey (Arcitecta)
  - Steve Melnikoff (Victorian E-Research Strategic Initiative; VeRSI)

## Neuro-Informatics Technology to Understand the Brain

*“... Combines neuroscience and informatics research to develop and apply the advanced tools and approaches that are essential for major advances in understanding the structure and function of the brain.”*

OECD Global Science Forum Neuroinformatics Working Group, 2002

### Neuroinformatics

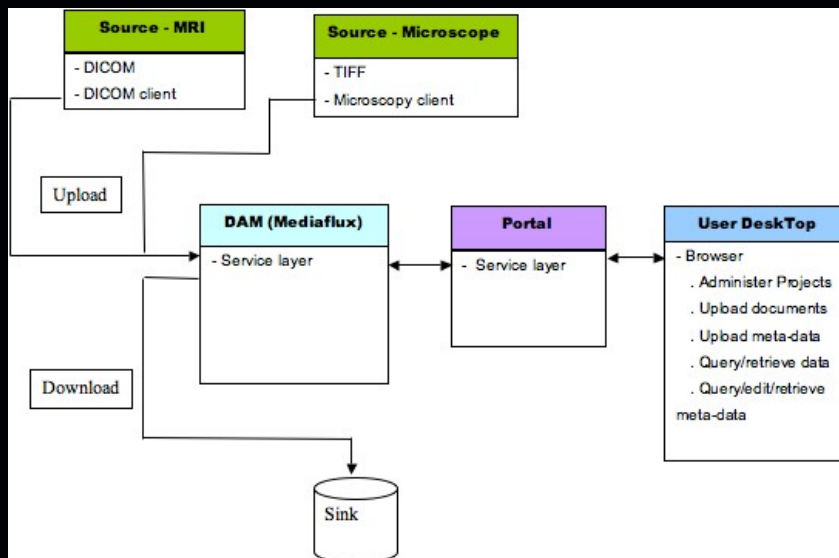
m	CNS	
10 cm	Systems	
1 cm	Maps	
1 cm	Networks	
100 μm	Neurons	
μm	Synapses	
Å	Molecules	

From David Willshaw, Edinburgh (2002)

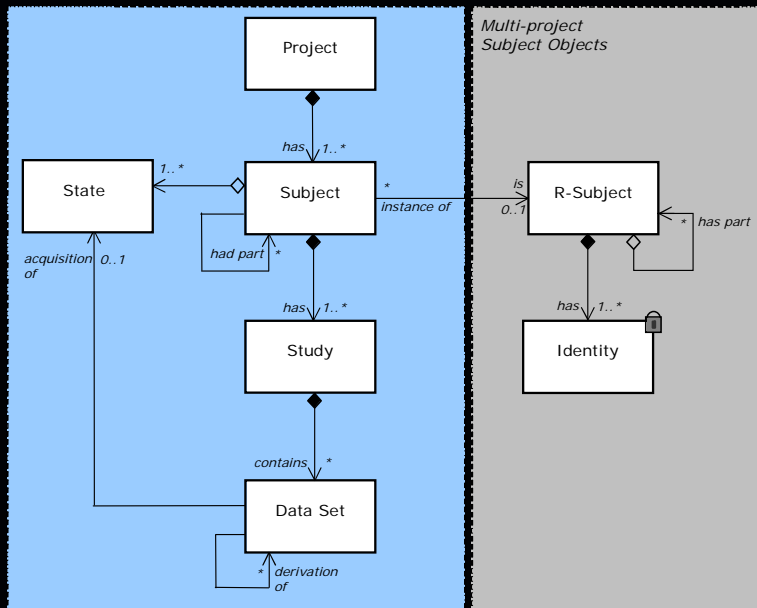
## Neuro-Imaging Informatics Requirements

- **Neuroimaging techniques (data sources)**
  - Magnetic resonance (MR) imaging
  - Microscopy (Optical and Electron)
  - Positron emission tomography (PET), Electro-encephalography (EEG)
- **Neuroimaging science databases**
  - Structural MR - human neonatal, adult psychiatric (schizophrenia)
  - Functional MR - normal adults
- **Requirements**
  - Support geographically dispersed teams & work-flows, multi-mode/location data acquisitions
  - Direct upload of data into system from data sources
  - Store, organize and manage information
  - Query and download of data
  - Handle evolving meta-data
  - Processed data products and application integration
  - Federate and inter-operate
  - Authenticated access via Portal and identity-based authorization

## Components



## DAM: PSSD Subject-Centric Information Model



## DAM : Object Identifiers

- Need to provide identifiers for objects
  - To provide citable IDs for external reference and data federation
  - To provide a practical means for humans to refer to asset collections and augment purely science meta-data based queries
- We allocate integer identifiers in a generic, extensible hierarchical scheme (N.N.N...). Applied to our object model:
  - Project.Subject.Study.DataSet (P.S.S.D)**. E.g. 10.23.17 would be study 17 of subject 23 of project 10
- Further prefixed by root: **T.P.S.S.D** so that different citable hierarchies can be created (e.g. Projects, Organizations, Methods)
- Further prefixed by a naming authority **N.T.P.S.S.D** so that uniqueness within a specified scope is achieved (onion model)
- Register IDs with a service like HANDLE or PILIN for global uniqueness: **NA.ORG/N.T.P.S.S.S**

## DAM : Authorization

Access to information and services is provided through role-based access. Each user is assigned roles that determines their access to assets and services. We define the following **Project-based** roles:

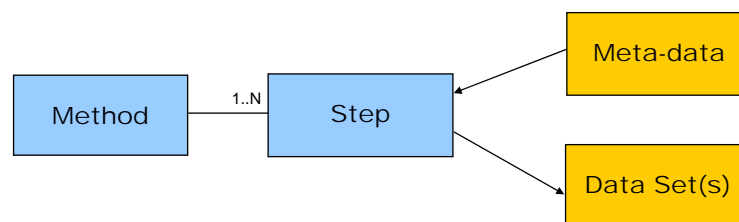
Role	Description
<i>ProjectAdmin_&lt;ID&gt;</i>	Super-user permissions; can modify project team/roles
<i>SubjectAdmin_&lt;ID&gt;</i>	Can create and administer subjects (identity)
<i>User_&lt;ID&gt;</i>	Access to all research data (but not subject identity)
<i>Guest_&lt;ID&gt;</i>	Access to searchable meta-data only

<ID> is the identifier (N.T.P) of the Project; so roles are e.g *ProjectAdmin\_1.1.10* There are also non-project based roles such as

Role	Description
<i>ProjectGenerator</i>	Able to create a Project
<i>User</i>	Basic access to system

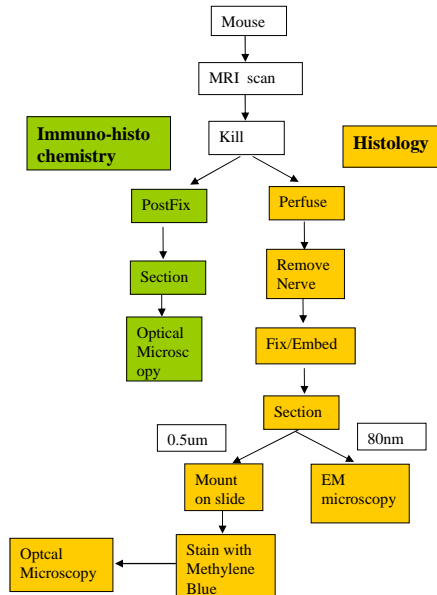
## DAM : Meta-data, Methods and State (1)

- PSDD model is generic (has basic required meta-data). Becomes domain specific with content & meta-data.
- Science experiments are processes with several steps. Each step might be described by meta-data and involve a data acquisition. Encapsulate this in a *Method*:
  1. What the steps in the experiment are.
  2. What meta-data and content are captured for each step.
  3. On which objects the meta-data and content reside.
  4. What steps define a new Subject state



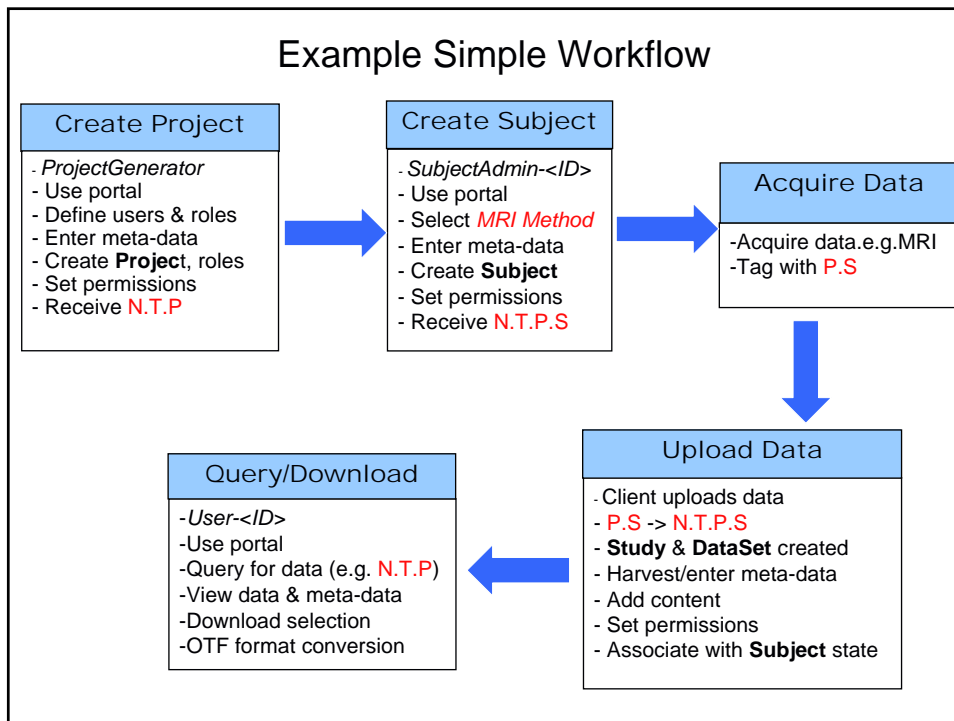
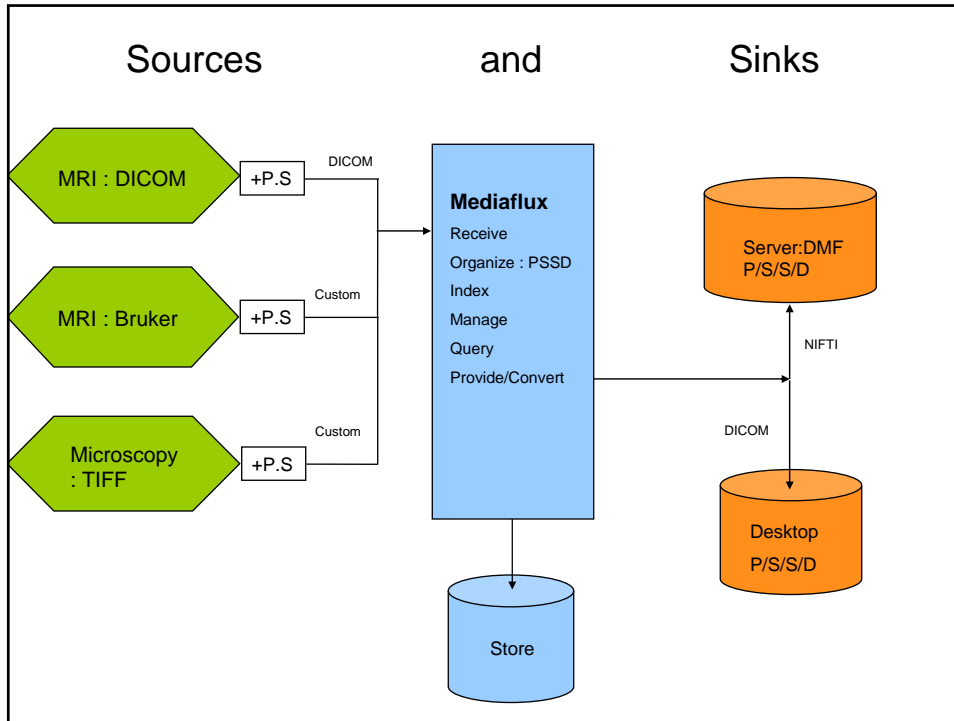
## DAM : Methods, meta-data and state (2)

- Simple methods (e.g. Human MRI) and complex, prescriptive methods (e.g. combine MRI and Microscopy in workflow)
- Avoids need for generality (but allows it & use of standards)
- Retain flexibility for repeat and ad-hoc measurements
- Build Methods from reusable components
- Methods are citable !



## Digital Asset Management System - Mediaflux

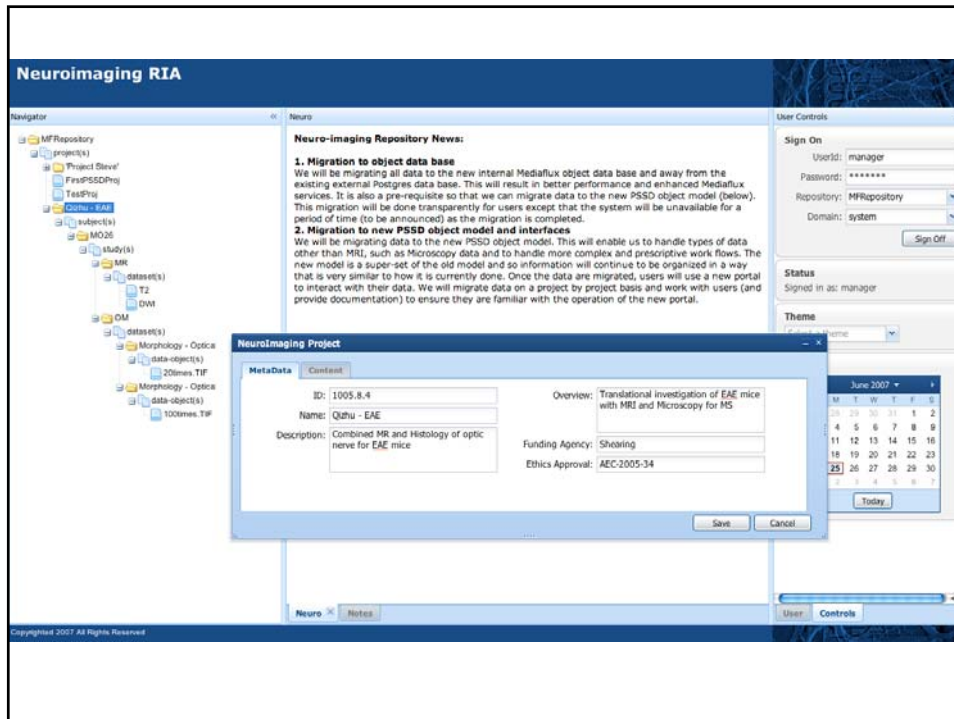
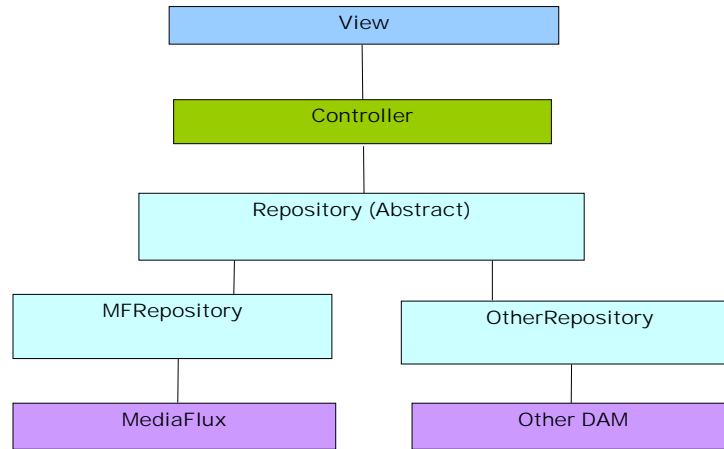
- **Server Features**
  - Robust, extensible (plugin) service-oriented, multi-platform Java framework
  - Support for remote communications (TCP/IP, HTTP)
  - Multiple interfaces (Java, HTTP, web-services, aterm, command line, Tcl)
  - Hierarchical role-based authorization model
  - Built-in DICOM server component
  - Supports replication, federation and distributed queries
- **Asset Management Features**
  - Meta-data managed with XML
  - Incremental schema evolution
  - Powerful query language
  - Scalable object data base
- **Workflow Engine for codifying short- or long-term processes**
- **Extensibility**
  - Plug-in Java framework
  - Extensible content-conversion framework
  - Extensible content store framework
- **Administration** : Rapid install, archive services, on-line upgrading
- **Security** : authenticated (local/LDAP)/authorized access, auditing, Transport Layer Security
- **Integrated Web-server**
- Strong customer engagement, excellent responsiveness to enhancements & defects. Capability, price & engagement -> cost-effective solution



# Graphical Interfaces

## Principles

1. Drive the interface from the object model (“data driven”)
2. Separate the View from the Business logic
3. Use best practices and standards
4. Web and thick Java client as appropriate



## Deployment

- Support multi-institute teams of Neuro-scientists at CfN, HFI, MHRI, MNC & receives MRI data from the RCH (human) & HFI's aMRIF (animal).
- Operated with subset of the object model for 6 months, managing MRI data in ~20 projects (~1TB). System operates with Gridsphere and MF portals
- Mediaflux operates in an SGI Altix shared-memory (256GB), 64P single-image Linux system with 60-TB data store. Access via VPN
- The integration of better data management and high-end (sometimes parallel) computation has resulted in substantial productivity gains
- The new PSSD object model is implemented as a package in Mediaflux & we have successfully uploaded MRI and Microscopy data to it.
- The new web-based, data-driven interface is being built as a rich web application. We have successfully implemented the full path from View to Repository for a subset of functionality
- A thick Java-client is being developed to handle microscopy upload/state

## Summary and Future Work

### Summary

- Better data management and an integrated high-end computational environment (Altix) -> significant productivity gains
- New Neuro-imaging projects exploring brain disease & function have increasingly complex workflows with multi-mode & distributed data acquisitions
- We have developed a generic subject-centric Object Model that handles arbitrary data types, complex & derived DataSets & Subject state which when combined with the experimental *Method* provides a platform to meet these challenges
- Interfaces are being implemented with a generic object-model driven approach, state handling & an abstraction of the underlying DAM/Repository

### Future Work

- Processed Data in Repository (history)
- Other related Neuro-science domains (e.g. MRS, EEG)
- Integrate with other objects (e.g. Attachments, Atlases)
- External PSSD registration, federation, inter-operability (possible now with MF)
- Application integration
- Meta-data standards for Neuro-imaging

## Present

- Neuro-informatics emerged in the 90s as data bases and tools emerged (mainly via Human Brain project)
- Awareness still poor in early 2000s
- Focii on imaging (BIRN), genes (Allen atlas) and cells (SenseLab)
- Subsequent activity/funding/groups led to information portals and efforts towards developing standards
- Growing collections of data (with degrees of accessibility)
- Just starting to integrate applications
- International Neuro-informatics Coordinating Facility (12 nations)
- Needed now
  - Repositories (with federation)
  - Terminology and Languages (confront Neuro-babble)
  - Community acceptance and need