Open and Big Data for Life Imaging
March 30, 2015

Technologies for Data Sharing in Neuroimaging

David N. Kennedy, PhD.

Child and Adolescent NeuroDevelopment Initiative
Department of Psychiatry
University of Massachusetts Medical School

University of Massachusetts Medical School
UMASS
CANDI
incf
NITRC

Overview

- **Why Share?**
- **Framework for Sharing**
- **Existing Solutions**
- **Main Problems**

Data Sharing is Key to Scientific Reproducibility

Reproducibility, in general, has a number of manifestations:

- **Publication-level Replication**. Take any given publication and cast it in a reproducible fashion. This is important due to the impact academia places on the publication as the principle currency of scientific productivity.
  - Eliminate errors
  - Make it complete (data, workflow, execution, results)
  - 3rd party certifiably re-executable
- **Generalizable Reproducibility** across publications. Huge problem since we dis-incentivize publication of replication studies as 'not novel'
  - Rerun same analysis on different 'similar' data (subjects and acquisition variables)
  - Rerun 'comparable' analysis on same data (methods variables)
  - Rerun same analysis on expanded 'similar' data (increase nominal power)

The Cost of Irreproducibility

- Irreproducibility costs money. Because of variability of results, more grants get funded to keep attempting to 'finally' resolve the mounting conundrum.
- Irreproducibility costs trust. Variability of results is part of why progress in neuropsychiatric disorders has been excruciatingly slow. Akin to NIMH's Research Domain Criteria (RDoC), we advocate RAnC (Reliable Analysis Criteria).
- Irreproducibility costs lives. Any delay in turning findings into treatment or prevention has a tangible impact on the lives of millions of people worldwide.

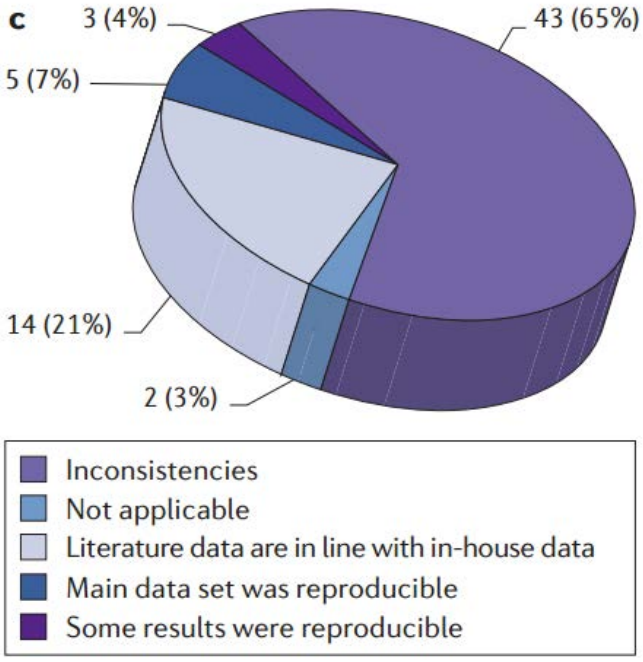


Chart from Prinz, et al. Nature Reviews Drug Discovery 10, 712 (September 2011) Bayer Healthcare

Causes of Reproducibility/ Generalizability Issues

- Low power
- Mistakes
- Ineffective Data Sharing
- Methodological Variance
- Sampling
- P – hacking (P-phishing)
- Publication Bias
  - Only reporting potentially significant results, and not the rest

P-Curve: A Key to the File-Drawer

Uri Simonsohn
University of Pennsylvania

Leif D. Nelson
University of California, Berkeley

Joseph P. Simmons
University of Pennsylvania

Because scientists tend to report only studies (publication bias) or analyses (*p*-hacking) that "work," readers must ask, "Are these effects true, or do they merely reflect selective reporting?" We introduce *p*-curve as a way to answer this question. *P*-curve is the distribution of statistically significant *p* values for a set of studies (*ps* < .05). Because only true effects are expected to generate right-skewed *p*-curves—containing more low (.01s) than high (.04s) significant *p* values—only right-skewed *p*-curves are diagnostic of evidential value. By telling us whether we can rule out selective reporting as the sole explanation for a set of findings, *p*-curve offers a solution to the age-old inferential problems caused by file-drawers of failed studies and analyses.

Journal of Experimental Psychology: 2014, Vol. 143, No. 2, 534–5

NATURE REVIEWS | NEUROSCIENCE
VOLUME 14 | MAY 2013 | 365

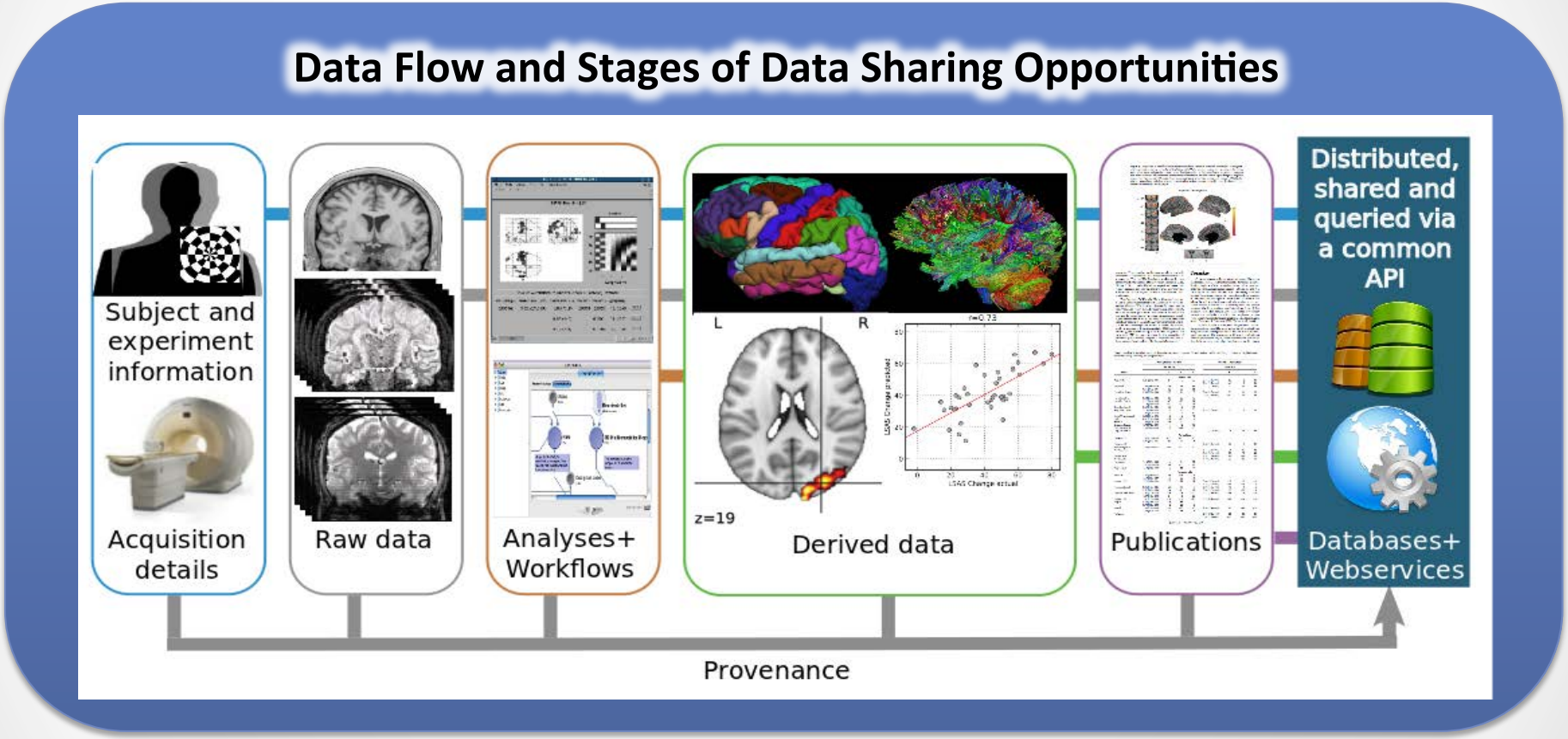
Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button[1,2], John P. A. Ioannidis[3], Claire Mokrysz[1], Brian A. Nosek[4], Jonathan Flint[5], Emma S. J. Robinson[6] and Marcus R. Munafò[1]

FAIR Data

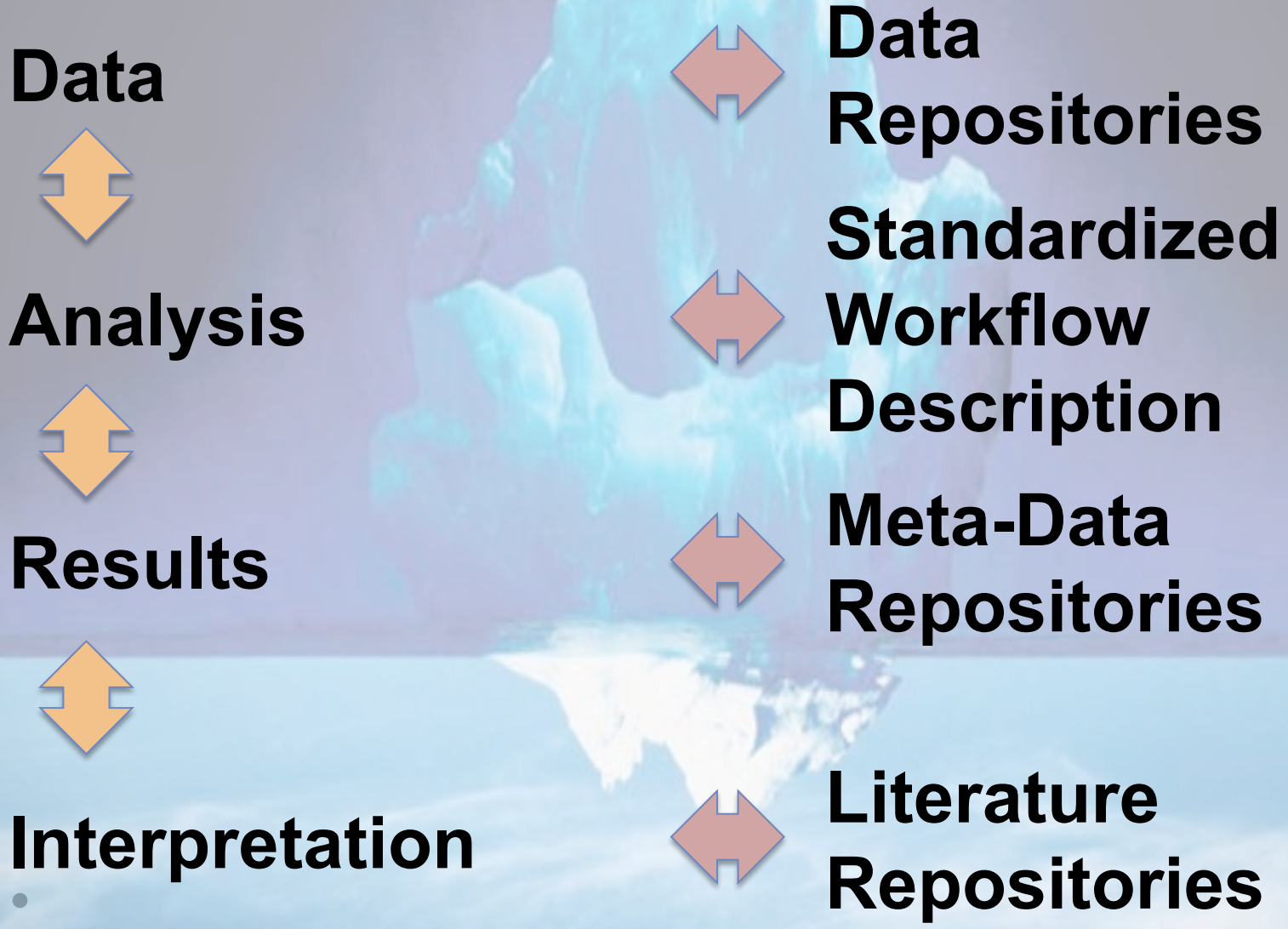
- Data FAIRport (http://www.datafairport.org/): valuable scientific data should be 'FAIR':
  Findable, Accessible, Interoperable and Re-usable.
- Provenance: provenance-aware data capture, analysis workflows, and results markup.
- Publication should include the necessary supplemental access to the requisite data, workflows and results to maintain maximum reproducibility of the scientific literature,
- Why? These principles are critical to the protection of the public investment in the scientific enterprise and to maximize the potential impact of this research in return for our participants' involvement in this project.
- How?

Framework


Data Flow and Stages of Data Sharing Opportunities
Subject and experiment information
Acquisition details
Raw data
Analyses+ Workflows
L
R
z=19
Derived data
Publications
Distributed, shared and queried via a common API
Databases+ Webservices
Provenance

Stages of Information Sharing

(towards scientific reproducibility)



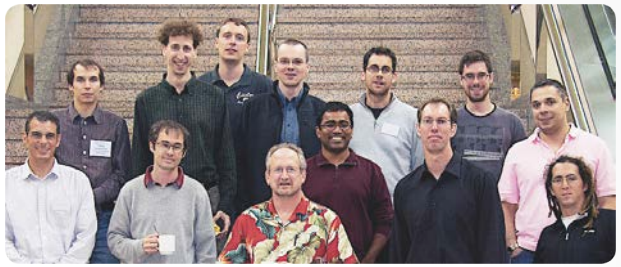
incf
Program on Standards for Datasharing

Mission: to develop metadata and data standards for reproducible research; to develop standards for archiving, storing, sharing, and re-using neuroscience data and databases

Electrophysiology Task Force

Focuses on electrophysiological data and databases

- Use cases for handling and converting electro-physiology metadata
- Web page on tools for converting between data formats
- Further development of Neuroshare API
- Continued engagement of industry vendors



Neuroimaging Task Force

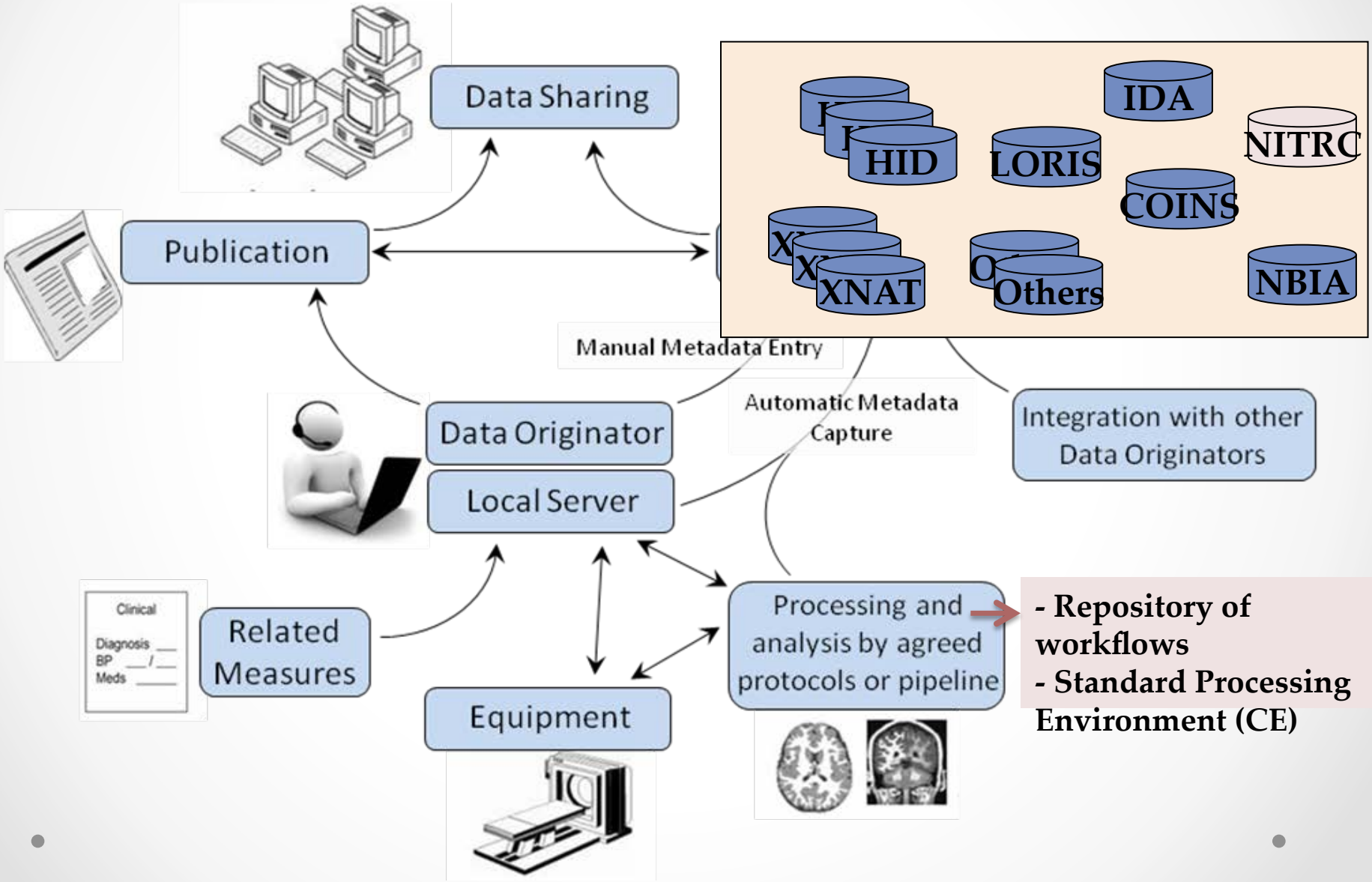
Focuses on neuroimaging data and databases of neurological/psychiatric disorders and cognitive function

- One-Click share tool
- Standard description of neuro-imaging data/ metadata & common API
- Standard data formats
- Standard workflows and provenance tracking

incf

www.incf.org/about/programs/datasharing

Data-sharing infrastructure



Many Existing Resources

- Image (Raw Data) Level
  - Required
    - ADNI (IDA), NDAR (NIMH), HCP (XNAT), etc.
  - Optional
    - NITRC (XNAT), COINS, LORIS, OpenfMRI
- Derived Data
  - Statistical Maps
    - NeuroVault
  - Activation Foci
    - BrainMap, SuMSDB, etc.

XNAT Impl

XNAT has become the world. The following lis out there.

Clinical and Translati
CTSAs are funded by t medical discovery to im

- Brain-CODE at t
- Harvard Catalyst
- ICTS at the Univ
- ICTS at Washing
- Laboratory of Cli

The Human Connecto
XNAT is the foundation from 1,200 subjects at backbone of the publicl

Institutional Operatio
XNAT is in use in neurc

- Biomedical Imag
- The Brain Imagir
- Centre of Excelle Valencia
- Centre of Medica
- Cornell Universit
- Emory University
- The Genes, Cog
- Hoglund Brain In
- Institute of Canc
- Mayo Clinic Neu
- Medical-image A
- Neuroimaging R
- New York State I
- Northwestern Ur
- Psychiatric Neur
- The Rotman Res
- University of Iow
- University of Wa
- Vanderbilt Unive
- Wake Forest Uni

Multi-Site Studies

- Cardiovascular F
- IMAGEN Study
- Dominantly Inhe
- The Jackson He
- PREDICT-HD st
- The Injury and T
- International Ne
- GENetic Frontot
- Prospective Urb
- The Virtual Brain
- CENTER-TBI: la
- National Consor

Clinical / Pharmacolo

- PharmaCog – Pr early clinical dev

Data Sharing

- OASIS Brains
- mBIRN Morphor
- XNAT Central
- Minimal Interval
- SchizConnect: S

The Qualitative Imagi

The Quantitative Imagi methods for the measu facilitating clinical decis

- PET-MRI for Ass
- Quantitative MR
- Quantitative MR
- Quantitative Ima

RESEARC
- Technolo
- Methodo
- Applicati

COLLABORATIVEINFORMATICS INS NEUROIMAGINGSUITE

HOME | DATA EXCHANGE | PARTICIPANT PORTAL | COINS TEAM | PUBLICATIONS | FAQ | LIVE DEMO | CONTACT | USER MAP

As neuroimaging research continues to grow, dynamic neuroinformatics systems are necessary to store, retrieve, mine, and share the massive amounts of data. The **C**ollaborative **I**nformatics and **N**euroimaging **S**uite (**COINS**) has been created to facilitate communication

LORIS

LOGIN

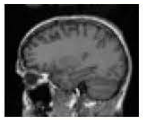
LONI
Image Data Archive

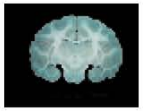
IDA HOME | ABOUT | NEWS | DOCUMENTATION | SOFTWARE | DATA

The LONI Image Data Archive (IDA) is a user-friendly environment for archiving, searching, sharing, tracking and disseminating neuroimagin clinical data. The IDA is utilized for dozens of neuroimaging research projects across North America and Europe and accommodates MRI, P other imaging modalities. A flexible data de-identification engine and encrypted file transmission help ensure compliance with patient-privacy data are stored on redundant servers with daily and weekly on- and off-site backups.

Archiving data in the IDA is simple, secure and requires no specialized hardware, software or personnel. All that is required is a computer wi and web browser software. The IDA automatically extracts relevant metadata from the de-identified image files allowing data to be searched archival. Once archived, data may be downloaded and/or streamed into the LONI Pipeline processing environment. Integration of the LONI D format translation engine allows users to download image data in a number of file formats in addition to the original file format.

Image Data Storage, Protection & Sharing





ABOUT
- Overview

NEWS
- New Features
- Announcements

DOCUMENTATION
- Instructions
- Collaboration
- FAQ

SOFTWARE
- Debabeler
- De-identification Debablet
- Inspector

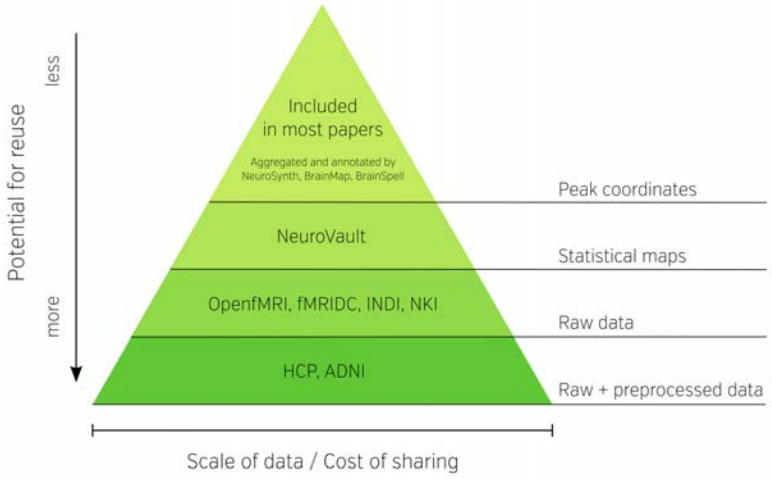
DATA
- Availab

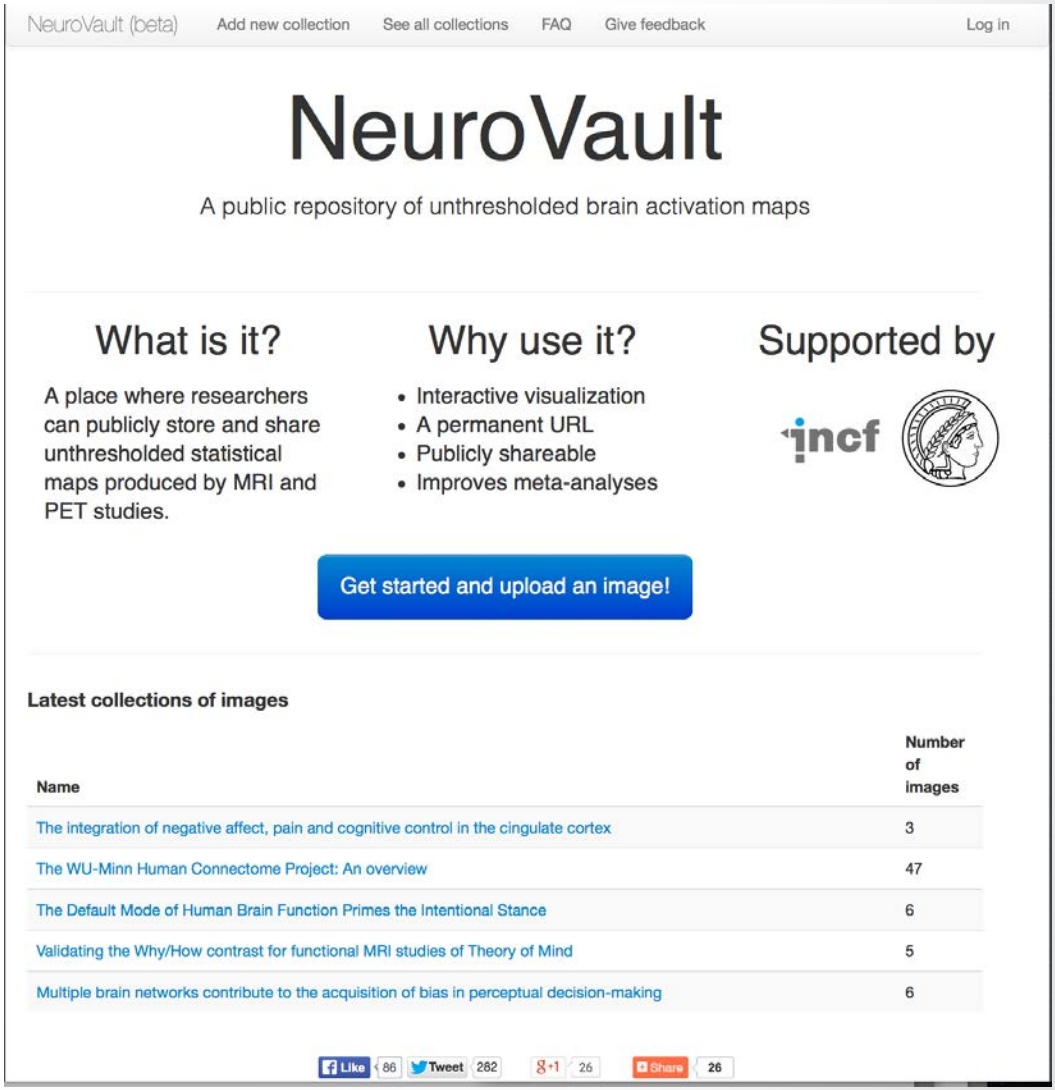
Down

Transferring data from ida.loni.usc.edu...

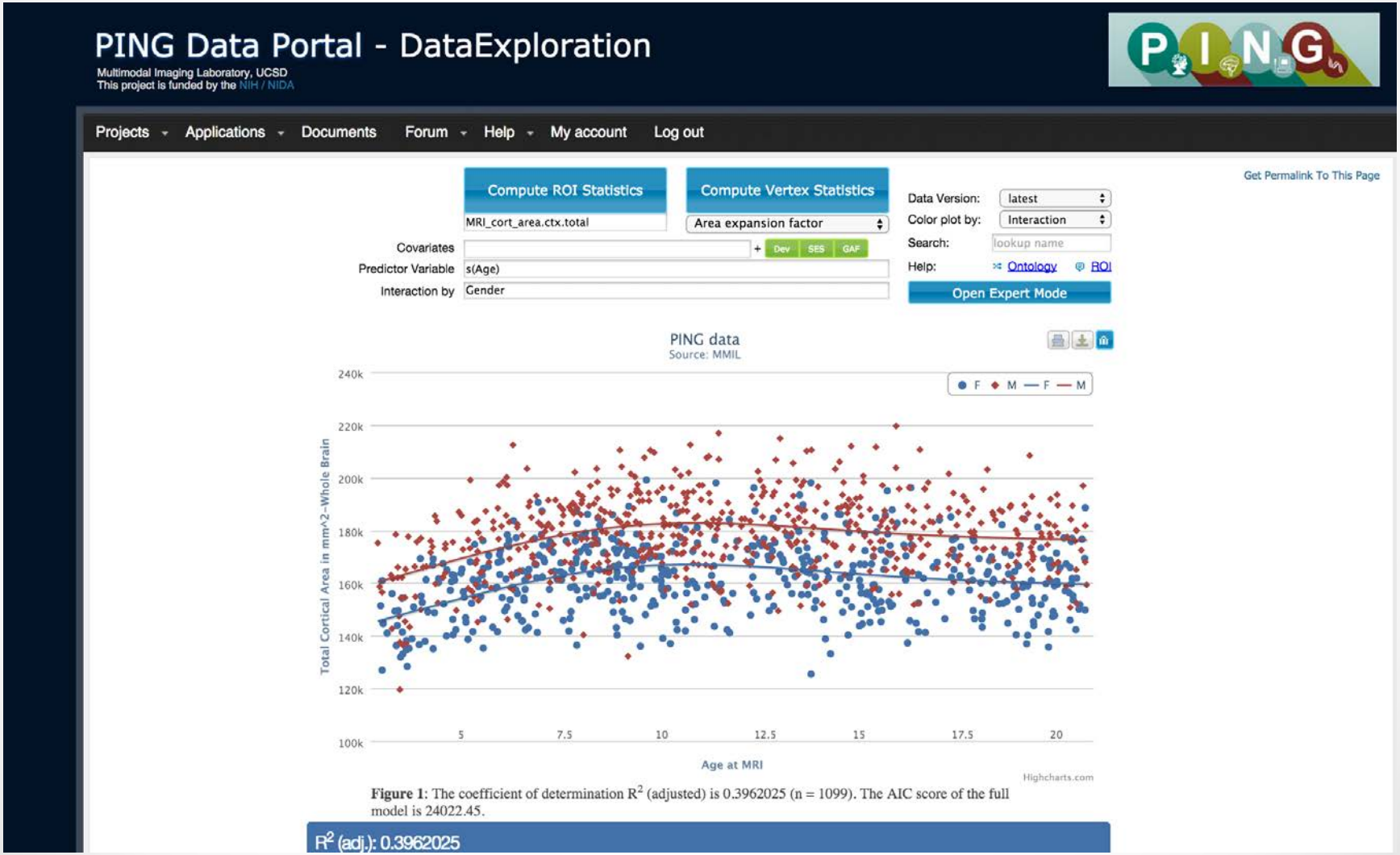
Home Re

Centre de neuroimagerie McConnell Brain Imaging Center


Potential for reuse
less
more
Included
in most papers
Aggregated and annotated by
NeuroSynth, BrainMap, BrainSpell
Peak coordinates
NeuroVault
Statistical maps
OpenfMRI, fMRIDC, INDI, NKI
Raw data
HCP, ADNI
Raw + preprocessed data
Scale of data / Cost of sharing


NeuroVault (beta)
Add new collection
See all collections
FAQ
Give feedback
Log in
NeuroVault
A public repository of unthresholded brain activation maps
What is it?
A place where researchers can publicly store and share unthresholded statistical maps produced by MRI and PET studies.
Why use it?
Interactive visualization
A permanent URL
Publicly shareable
Improves meta-analyses
Supported by
incf
Get started and upload an image!
Latest collections of images
Name
Number of images
The integration of negative affect, pain and cognitive control in the cingulate cortex
3
The WU-Minn Human Connectome Project: An overview
47
The Default Mode of Human Brain Function Primes the Intentional Stance
6
Validating the Why/How contrast for functional MRI studies of Theory of Mind
5
Multiple brain networks contribute to the acquisition of bias in perceptual decision-making
6
Like 86
Tweet 282
+1 26
Share 26

Integrated Views



Discovery and Action

The Three NITRC's

Neuroimaging Informatics Tools and Resources Clearinghouse
nitrc.org

- NITRC Resource Registry (NITRC-R)
- NITRC Image Repository (NITRC-IR)
- NITRC Computational Environment (NITRC-CE)

NITRC | IR

NITRC | CE

powered by XNAT

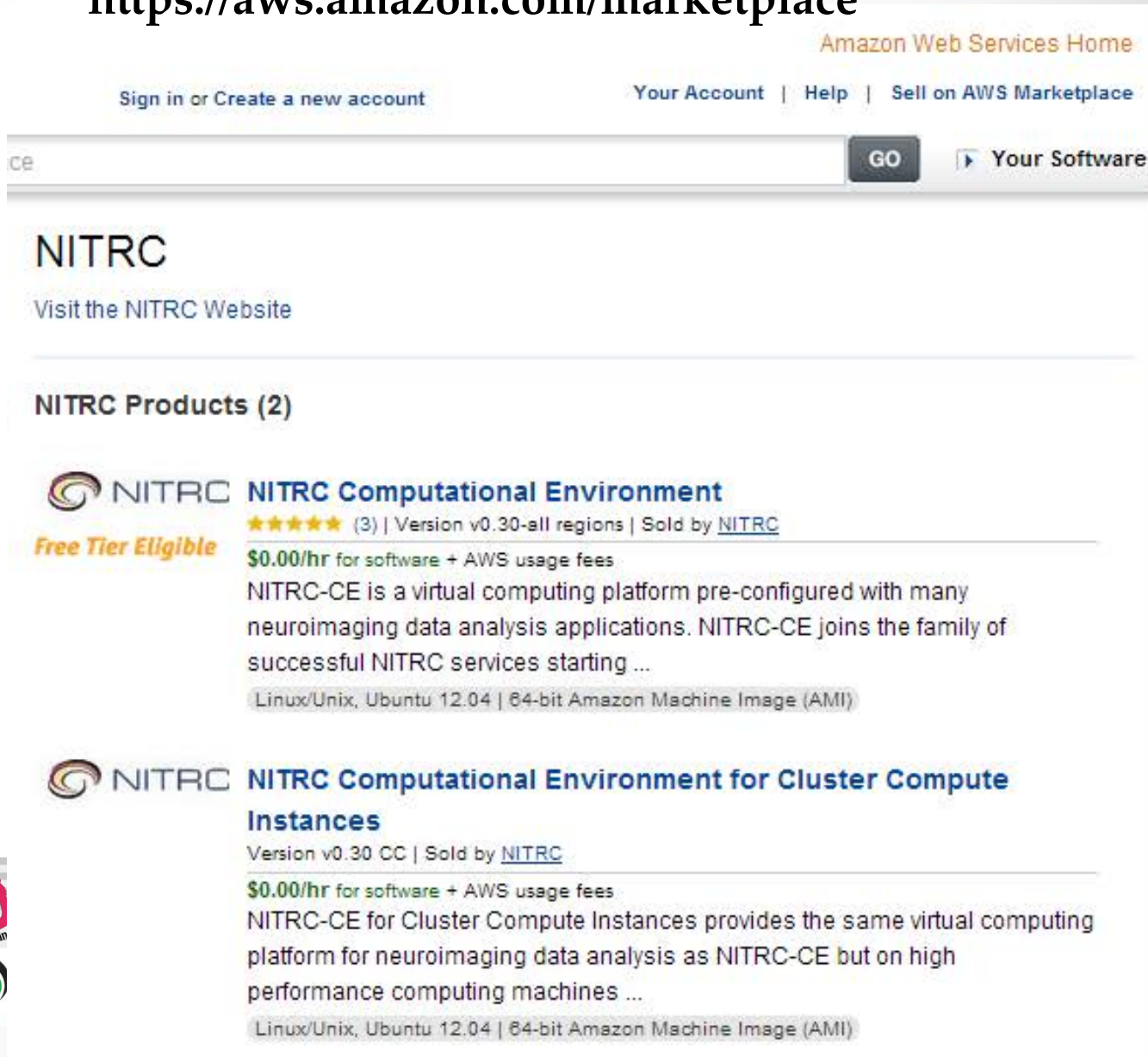
NITRC

debian

NITRC Computational Environment (NITRC-CE)

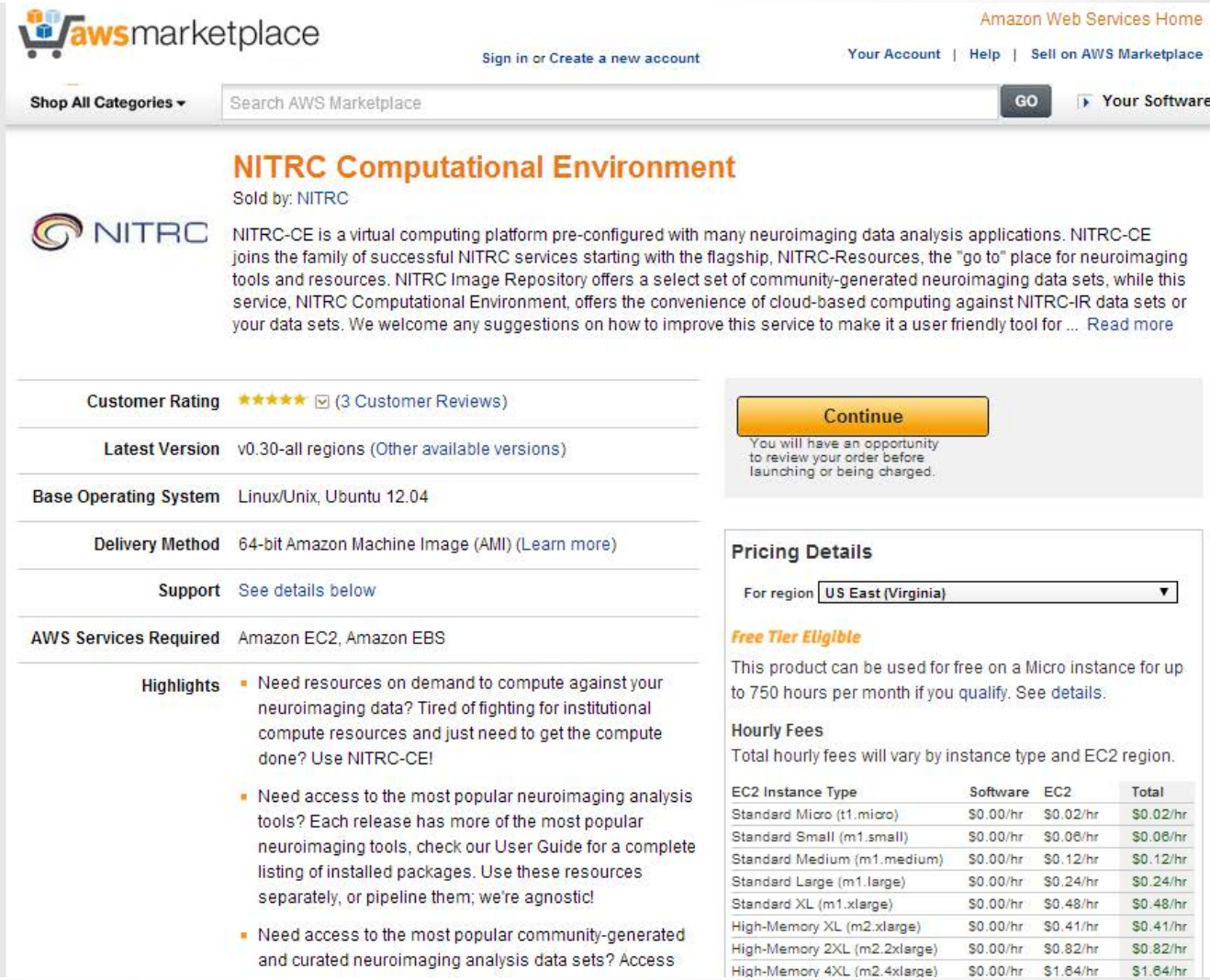
- Powered by NeuroDebian
- Dynamic, 'summonable', cloud-based (Amazon, Azure, local VM) computational platform
- Scalable performance (1-16+ cores)
- FSL, AFNI, FreeSurfer, and many more

https://aws.amazon.com/marketplace



NITRC Computational Environment (NITRC-CE)

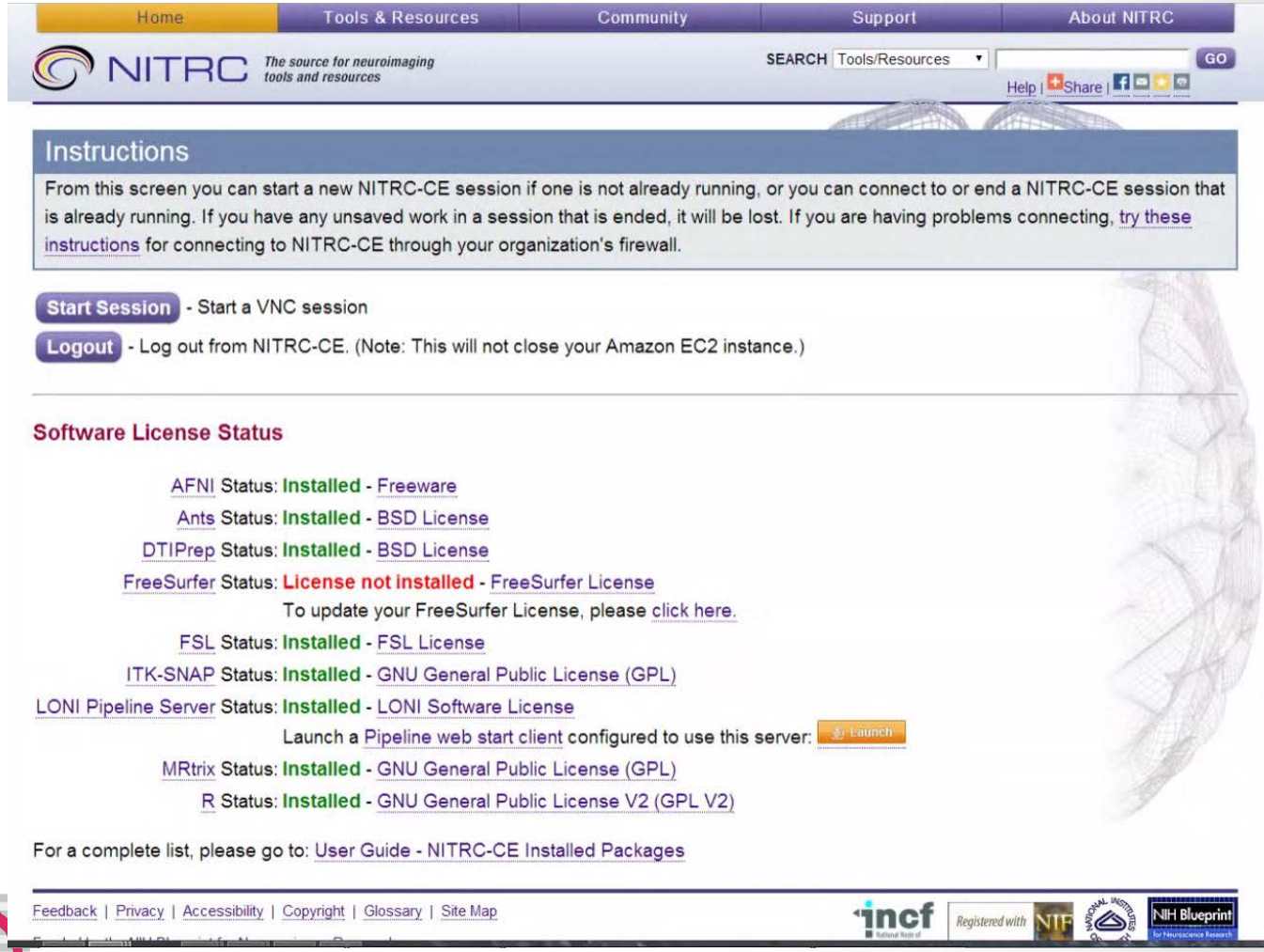
- Powered by NeuroDebian
- Dynamic, 'summonable', cloud-based (Amazon, Azure, local VM) computational platform
- Scalable performance (1-16+ cores)
- FSL, AFNI, FreeSurfer, and many more



NITRC | CE

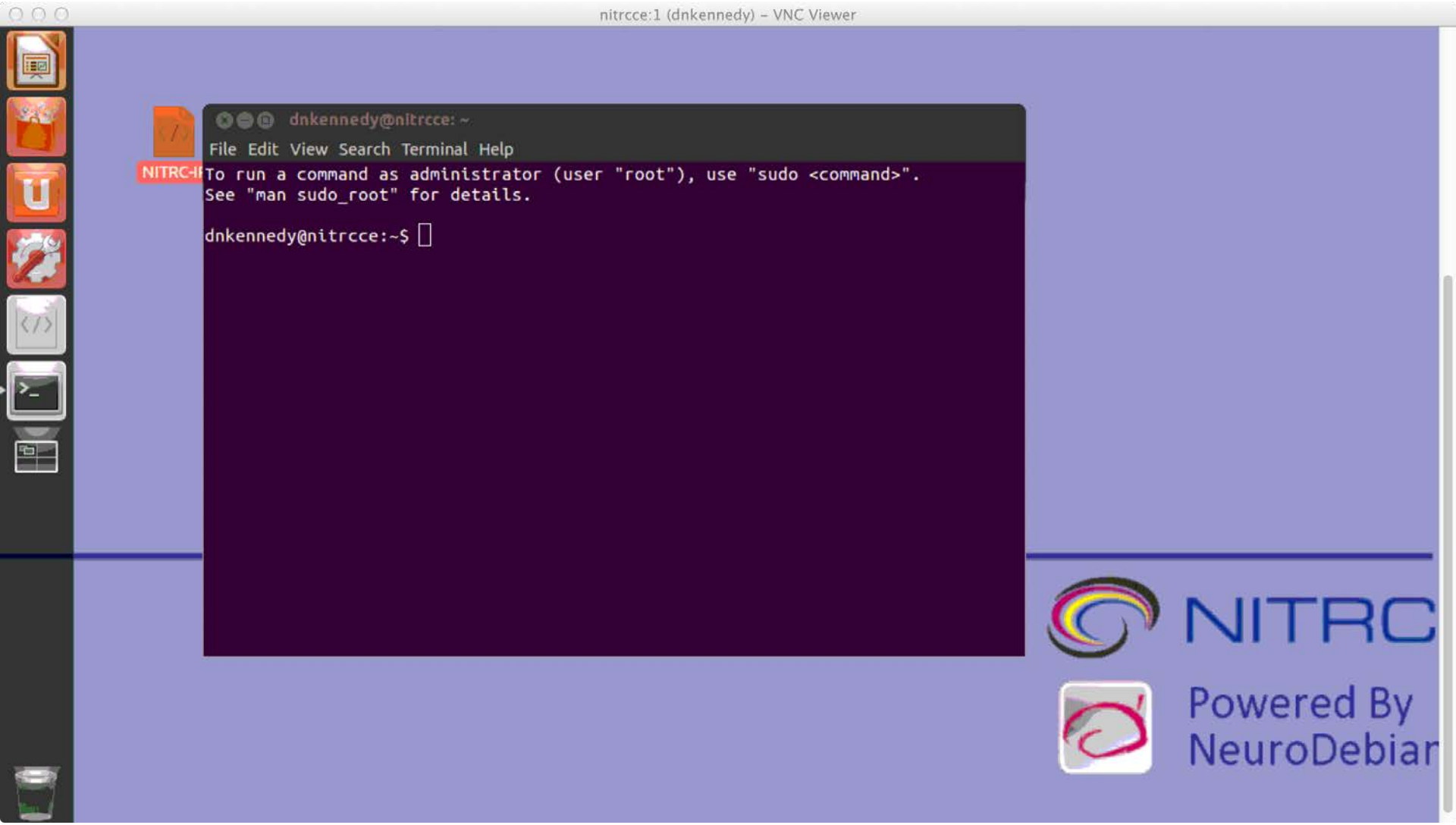
NITRC Computational Environment (NITRC-CE)

- Powered by NeuroDebian
- Dynamic, 'summonable', cloud-based (Amazon, Azure, local VM) computational platform
- Scalable performance (1-16+ cores)
- FSL, AFNI, FreeSurfer, and many more

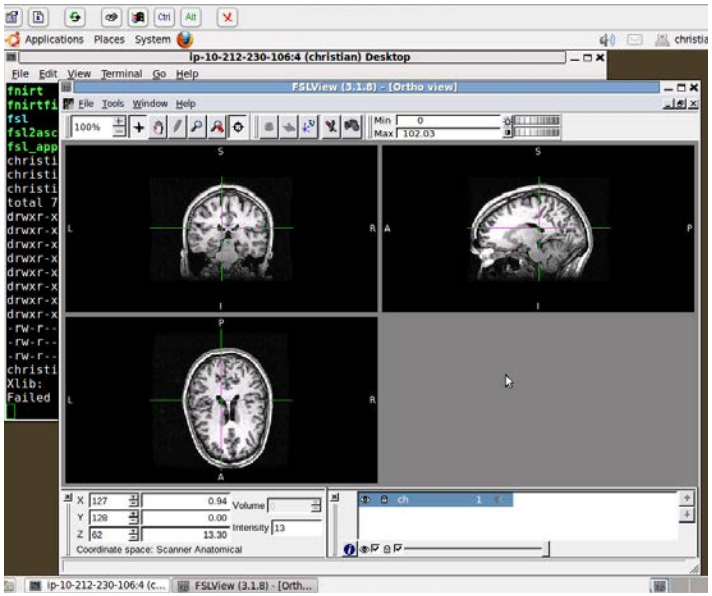


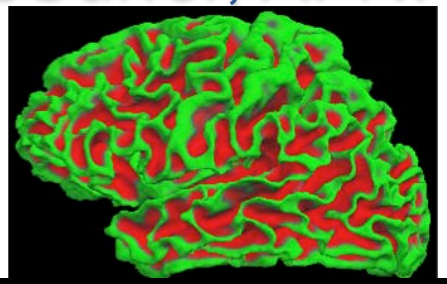
NITRC | CE

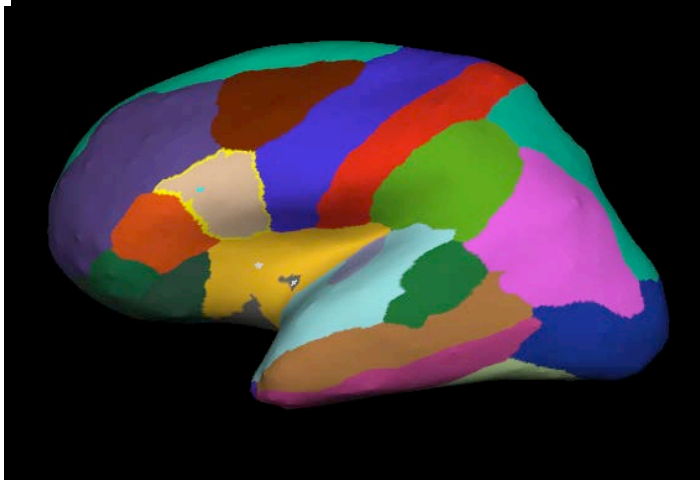
Sample Views: CE Desktop

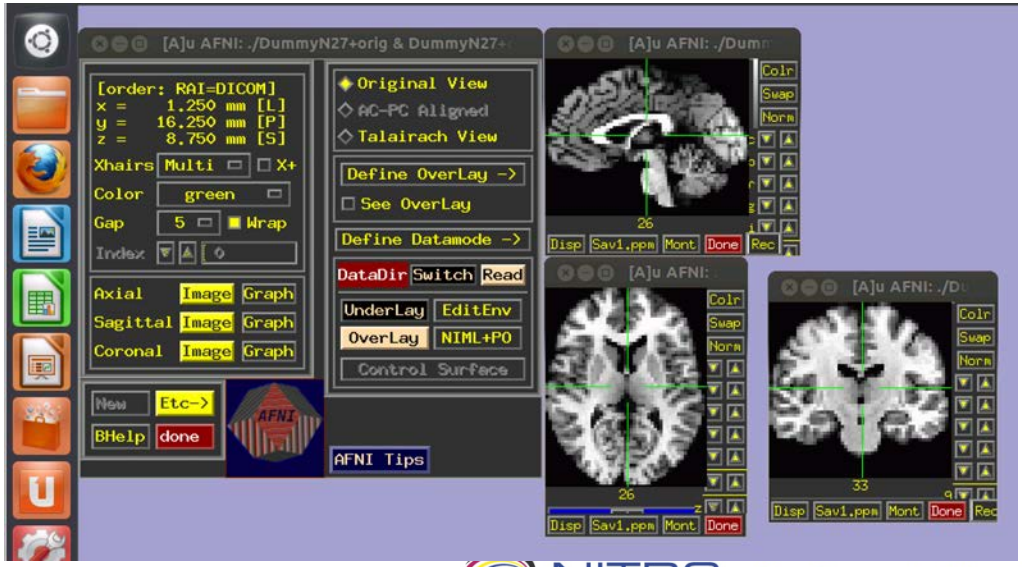


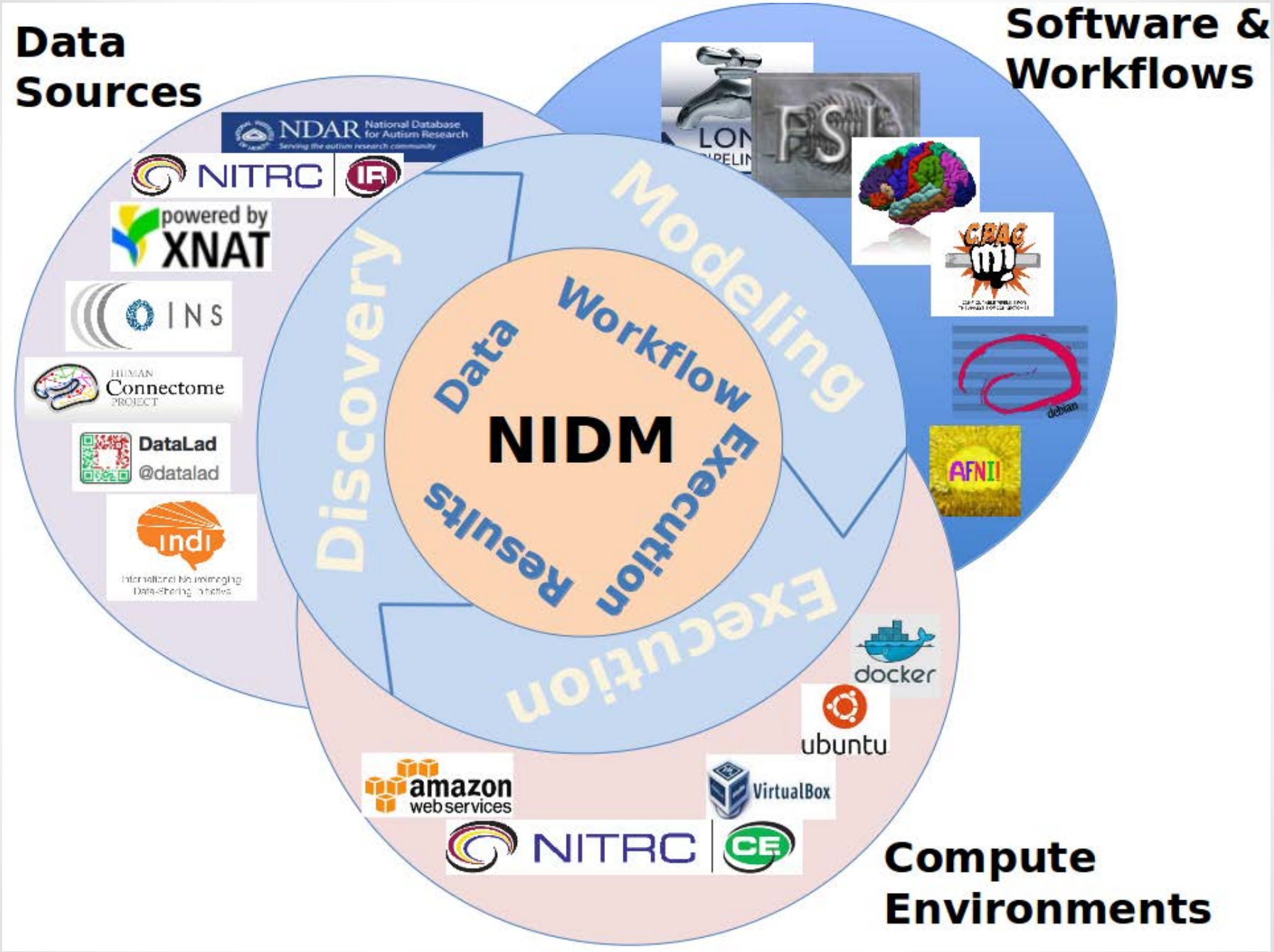
Sample Views: FSL, FreeSurfer, AFNI










Data Sources
Software & Workflows
Compute Environments
Discovery
Modeling
Execution
Execution
Data
Workflow
Execution
Results
NIDM
NDAR National Database for Autism Research
NITRC
IR
powered by XNAT
INS
HUMAN Connectome PROJECT
DataLad @datalad
indi
FSL
CBAC
debian
AFNI
docker
ubuntu
amazon web services
VirtualBox
NITRC
CE

incf Program on Standards for Datasharing

Provenance and NI-DM:

- The NI-DM data model
- Capture content from data and workflows (provenance)
- Apps for storing and querying NI-DM stores

incf Program on Standards for Datasharing

Provenance and NI-DM:

- **The NI-DM data model**

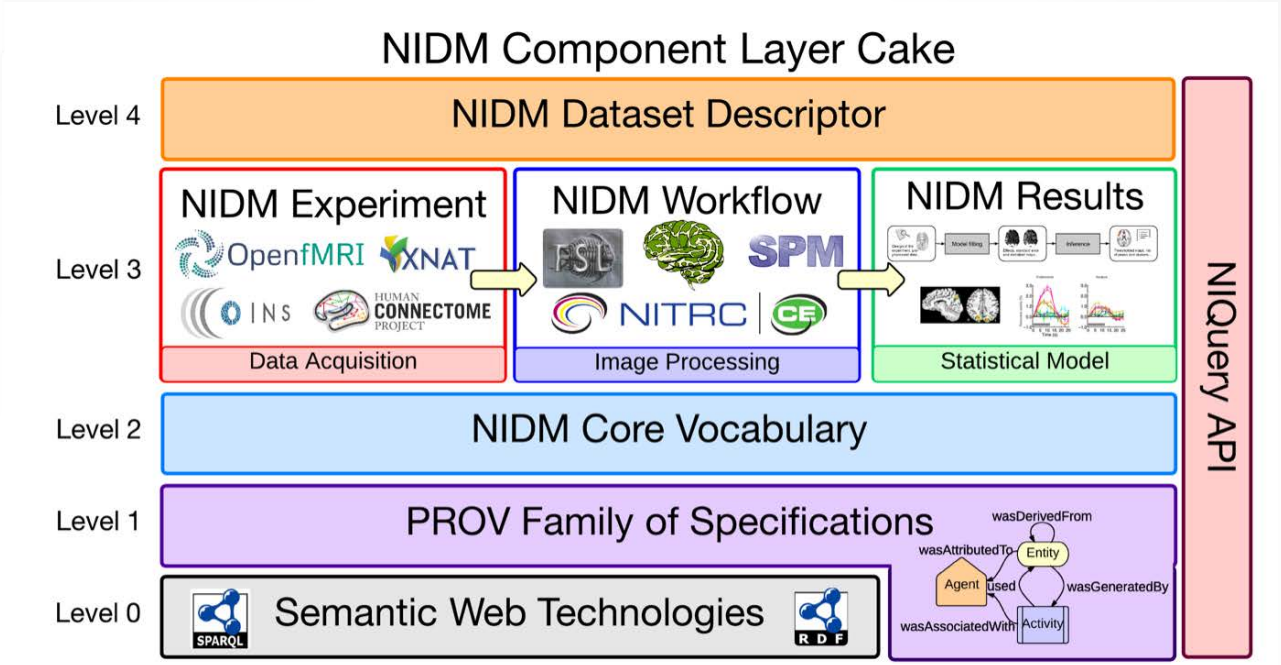
Neuroimaging Data Model (NI-DM) (Keator et al., 2013)

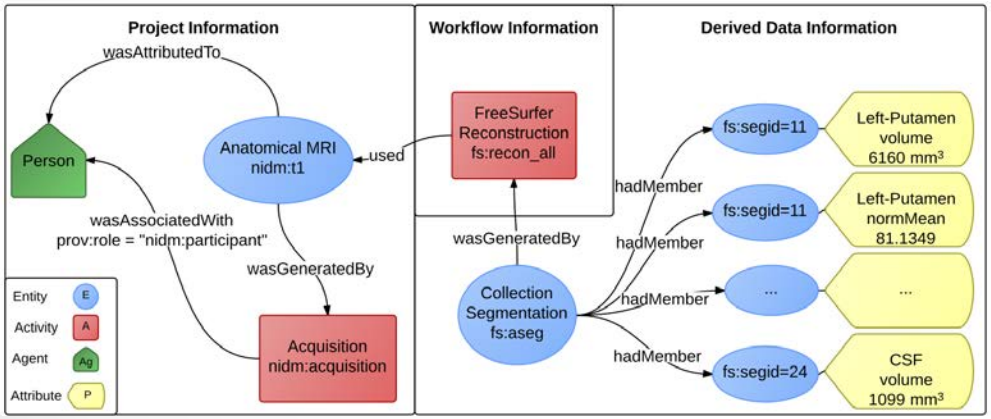
- Based on PROV-DM hence borrows PROV ontology (PROV-O)
- Structured information encoding
- Consistent vocabulary
- Metadata standards via domain specific object models

- Terms
  A lexicon of all things brain imaging. (e.g., DICOM terms, software specific terms, statistic terms, paradigm terms)
- Object Models
  Structured information in brain imaging (e.g., directory structures, CSV/Tab delimited files, brain imaging file formats)
- Integrated provenance
  How are entities generated or derived and by what or who?

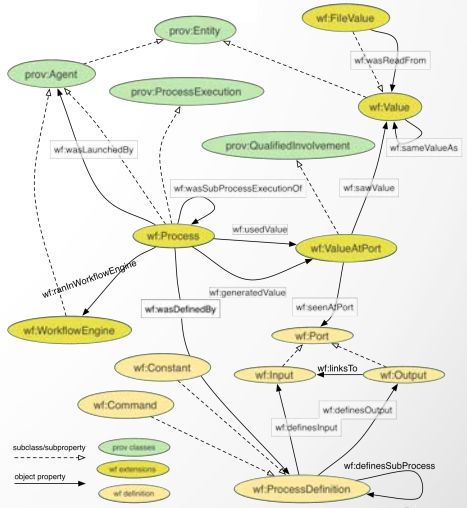
incf

Provenance and NI-DM

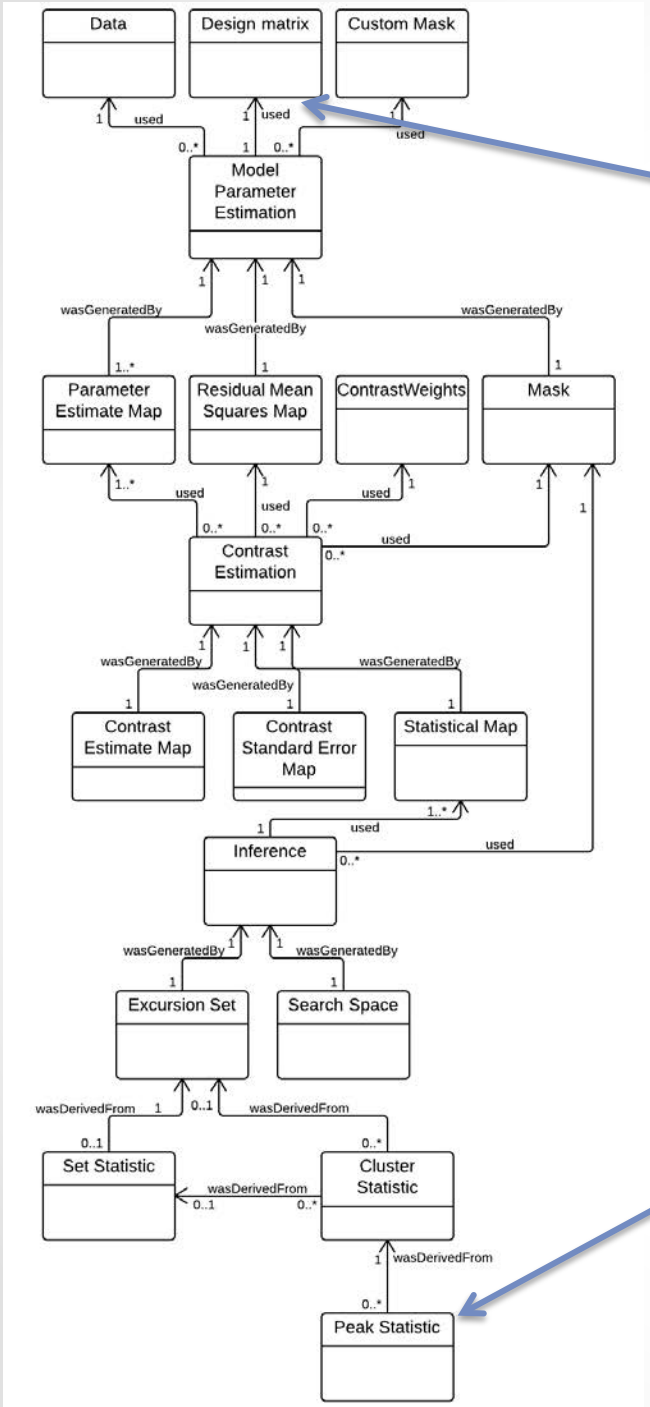




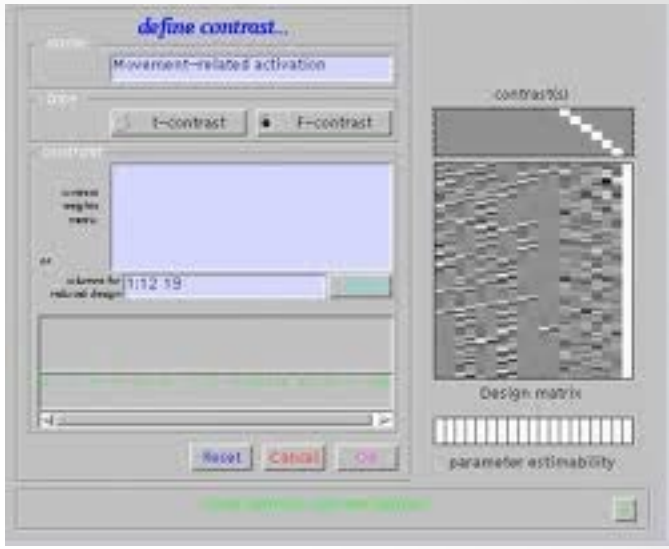
NI-DM can bridge information dimensions across Project, Workflow, and Derived Data

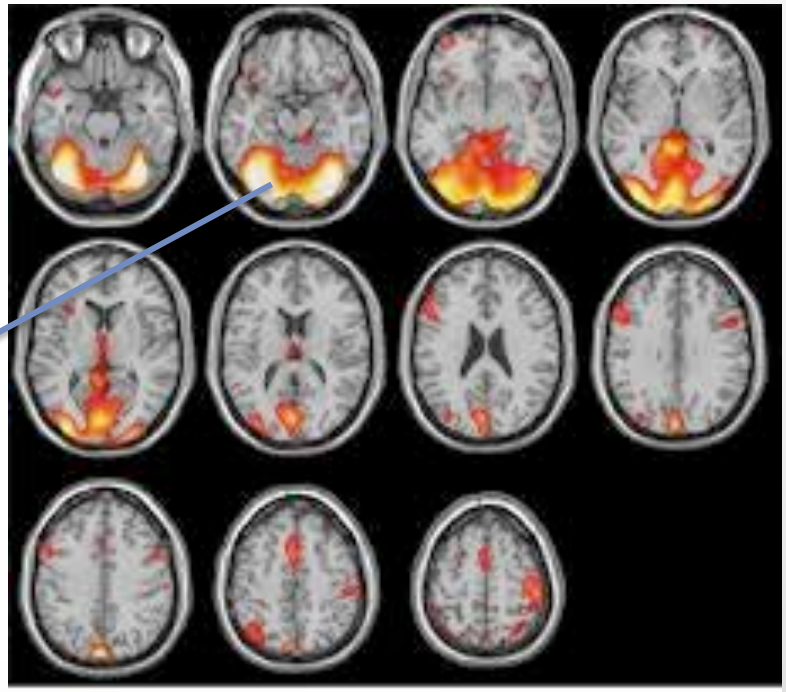


Workflow Ontology (WFO) extensions for PROV-DM



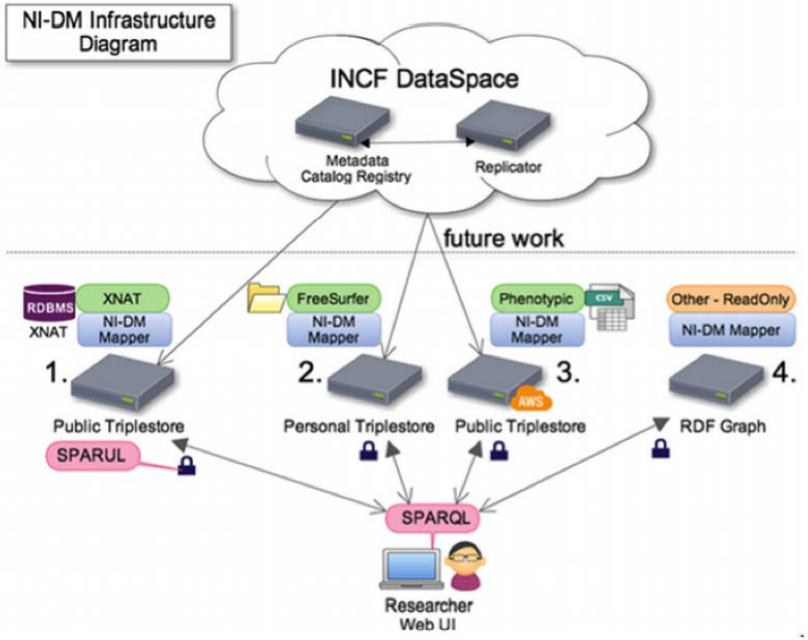
NI-DM in task activation





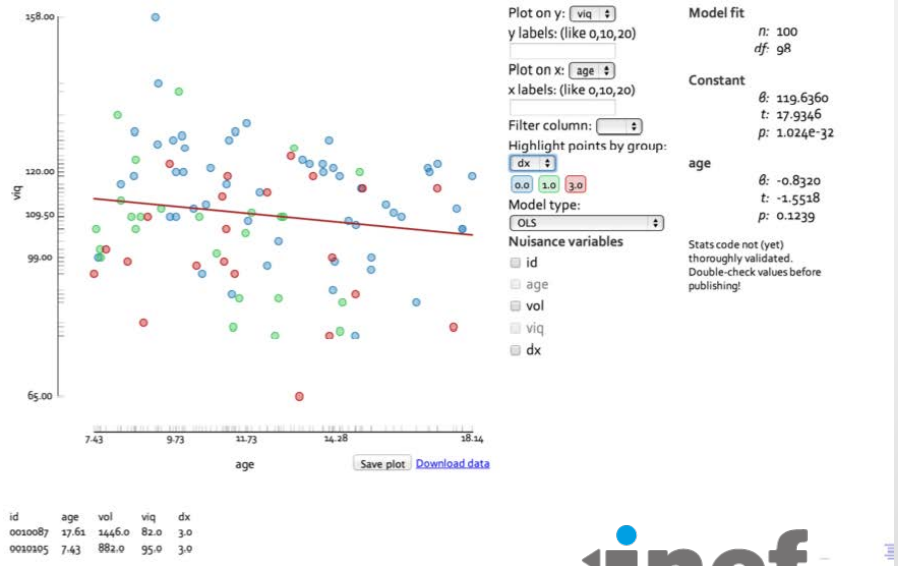
Provenance and NI-DM:

- Capture content from data and workflows (provenance)
- Apps for storing and querying NI-DM stores



Federated Query

```
App call: http://localhost:5000/u?url=http://bit.ly/1atAL00
Scatterize: https://github.com/njvack/scatterize
```



Problems & Barriers

- **Incentive**
- **Cost**
- **Credit**
- **Ethics**
- **Harmonization**
- **Reproducibility**

Incentive Barrier

- Funders & Publishers beginning to require sharing
- PubMed Central ‘success story’ for open publications
- Education & Simplification needed

The ‘Cost’ Barrier

The ~~Sharing~~ Tax

(aka Data Persistence Insurance)

- Sharing costs money: someone pays for it
- Well, not just ‘someone’, but rather the researcher themselves, regardless of:
  - If centrally supported
  - Imposed on the researchers themselves
  - Mandated in the costs of a specific program
- Given a finite budget, the cost of the sharing tax comes out of the number of subjects from whom data is acquired

~~The Sharing Tax~~
Data Persistence Insurance, continued

- What is the magnitude of the ~~sharing tax~~ persistence insurance?
  - ~5-10% of the cost of acquisition
- What return do we need on this?
  - ~10-20%, at least, in improved scientific ‘value’
- Now we have to document the real world benefit of sharing
- Or else…

One modest proposal:

- Funding agencies could set a future ‘market’ for ‘commercial’ data archival (10 years) at a target price.
  - i.e. for neuroimaging, a target cost equal to or less than 5% of the acquisition cost: is a typical hour of MRI scanning costs $500, then this would be $25 for archival of an typical 400MB scanning session
- The magnitude of this ‘potential future market’ per year can be estimated by the number of funded MRI sessions by the funding agency multiplied by the target price
  - This is a pretty big market (~$25-50M), should entice commercial or other vendors
- This ‘costs’ the funding agency nothing
  - i.e. they pay the same as they would otherwise; cost would be 5% reduction in subjects, in exchange for guaranteed future data accessibility

Neuroinform (2014) 12:361–363
DOI 10.1007/s12021-014-9239-0

EDITORIAL

Data Persistence Insurance

David N. Kennedy

Credit Barrier

Nested DOIs

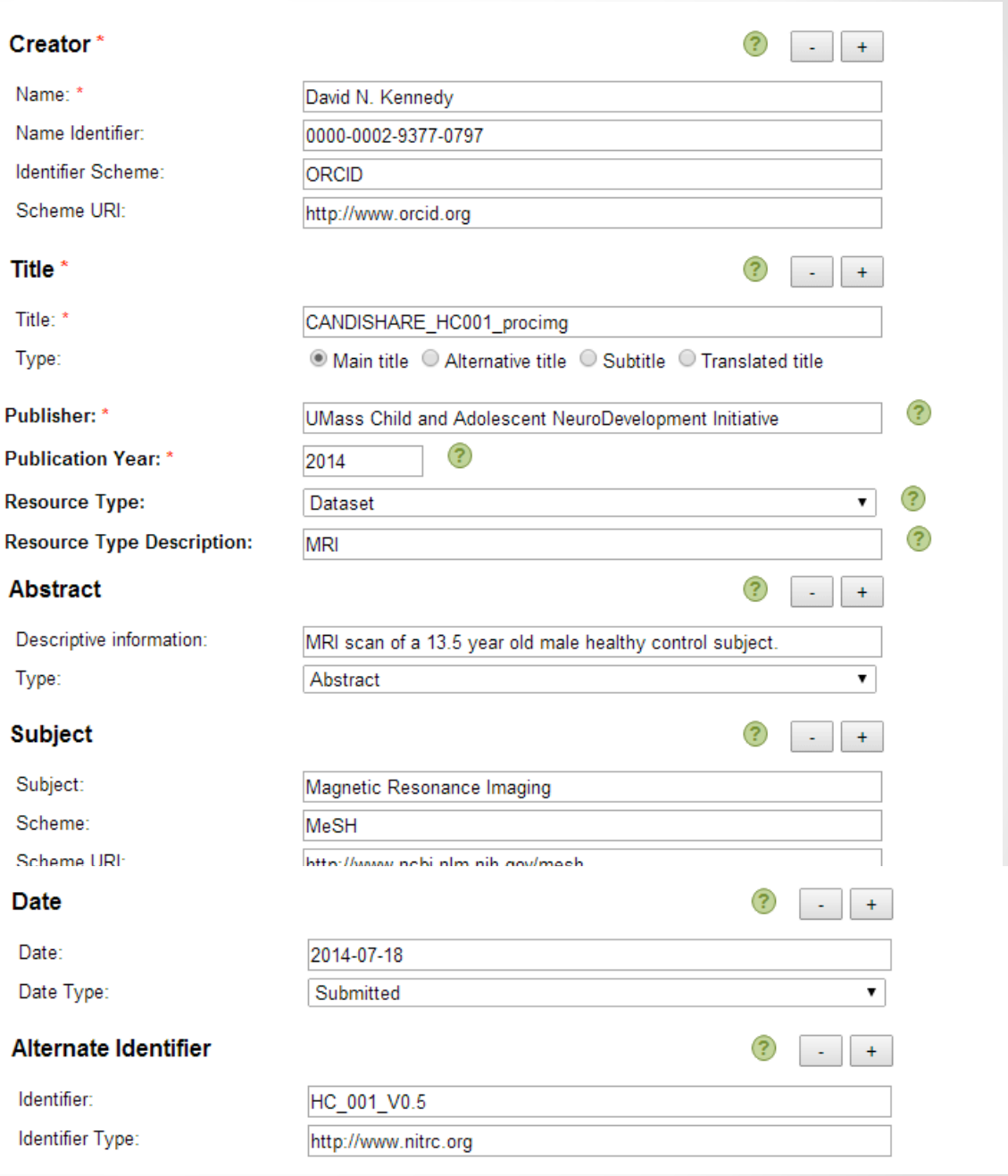
- Publication DOI
  - DataSet DOI
    - Individual DataElement DOIs

New publications using shared data cite the original Individual DataElement DOIs in the new DataSet DOI as part of the New Publication DOI

Credit where credit is due: Through the citation of the original Individual DataElement DOIs in new publication datasets, original data providers and original publications are cited

Data DOI's

- For tracking pooled use credit
- Documentation of of novel cohorts
- Cohort DOIs
- Derived data DOIs (or handles)



Creator *

Name: * David N. Kennedy
Name Identifier: 0000-0002-9377-0797
Identifier Scheme: ORCID
Scheme URI: http://www.orcid.org

Title *

Title: * CANDISHARE_HC001_procimg
Type: Main title / Alternative title / Subtitle / Translated title

Publisher: * UMass Child and Adolescent NeuroDevelopment Initiative

Publication Year: * 2014

Resource Type: Dataset

Resource Type Description: MRI

Abstract

Descriptive information: MRI scan of a 13.5 year old male healthy control subject.
Type: Abstract

Subject

Subject: Magnetic Resonance Imaging
Scheme: MeSH
Scheme URI: http://www.ncbi.nlm.nih.gov/mesh

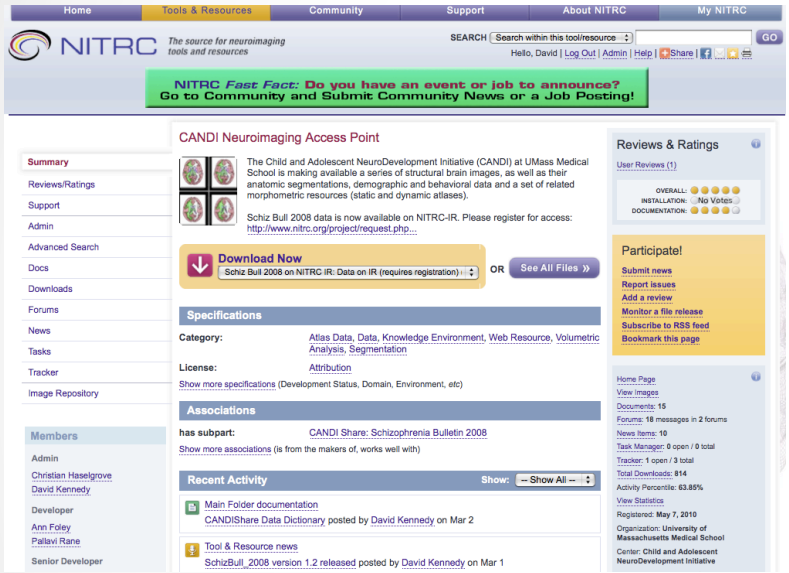
Date

Date: 2014-07-18
Date Type: Submitted

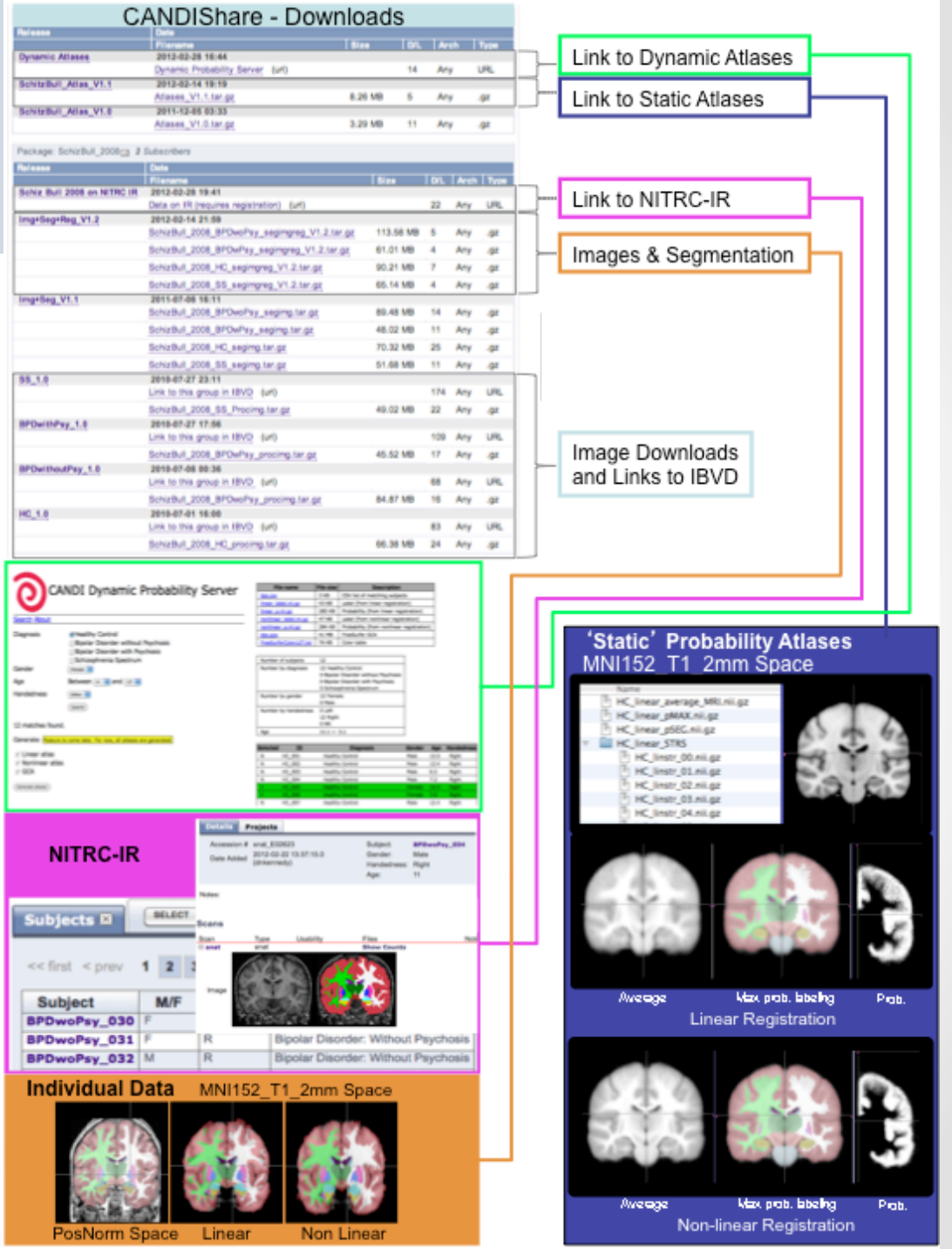
Alternate Identifier

Identifier: HC_001_V0.5
Identifier Type: http://www.nitrc.org

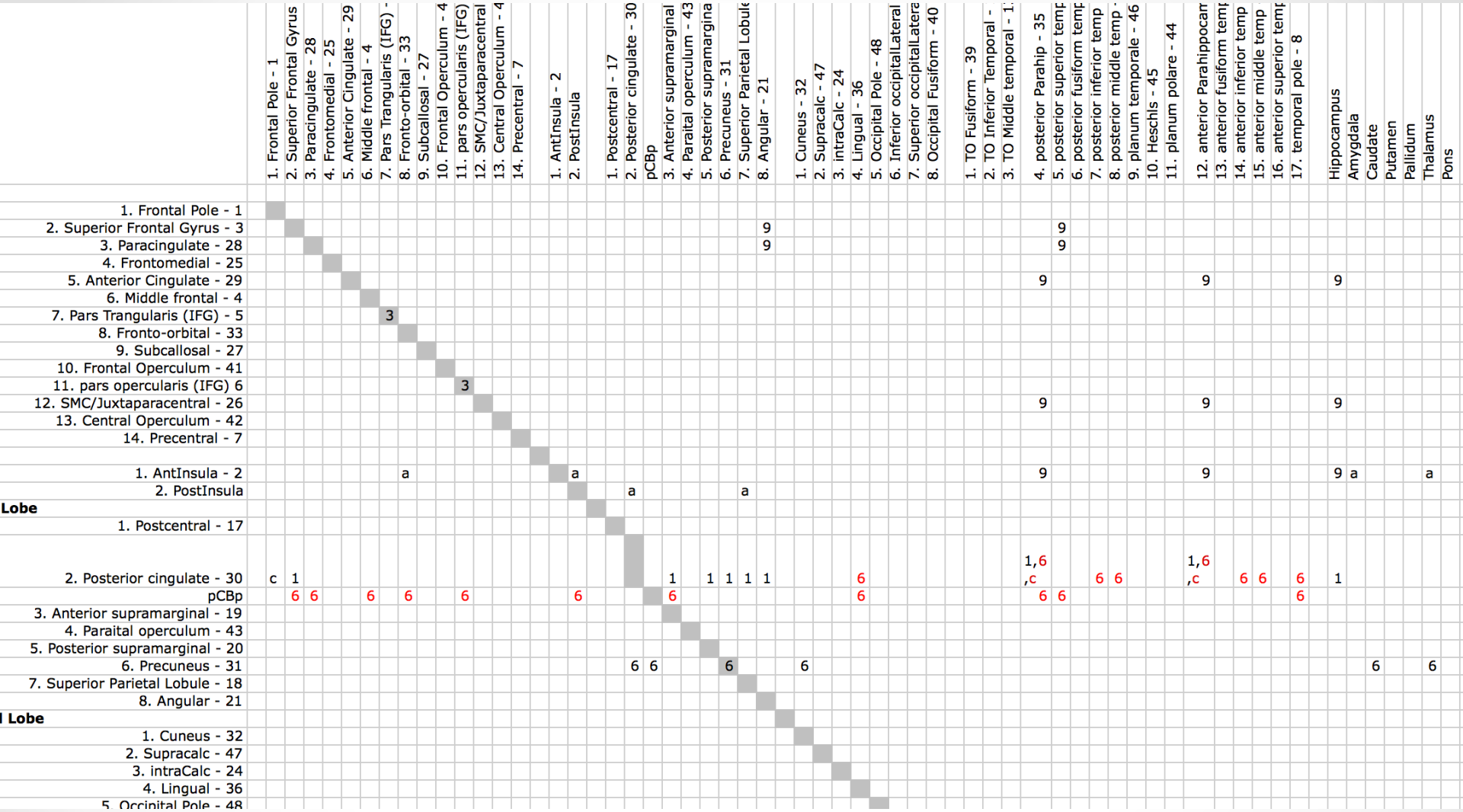
Integrated Publication



http://www.nitrc.org/projects/candi_share/

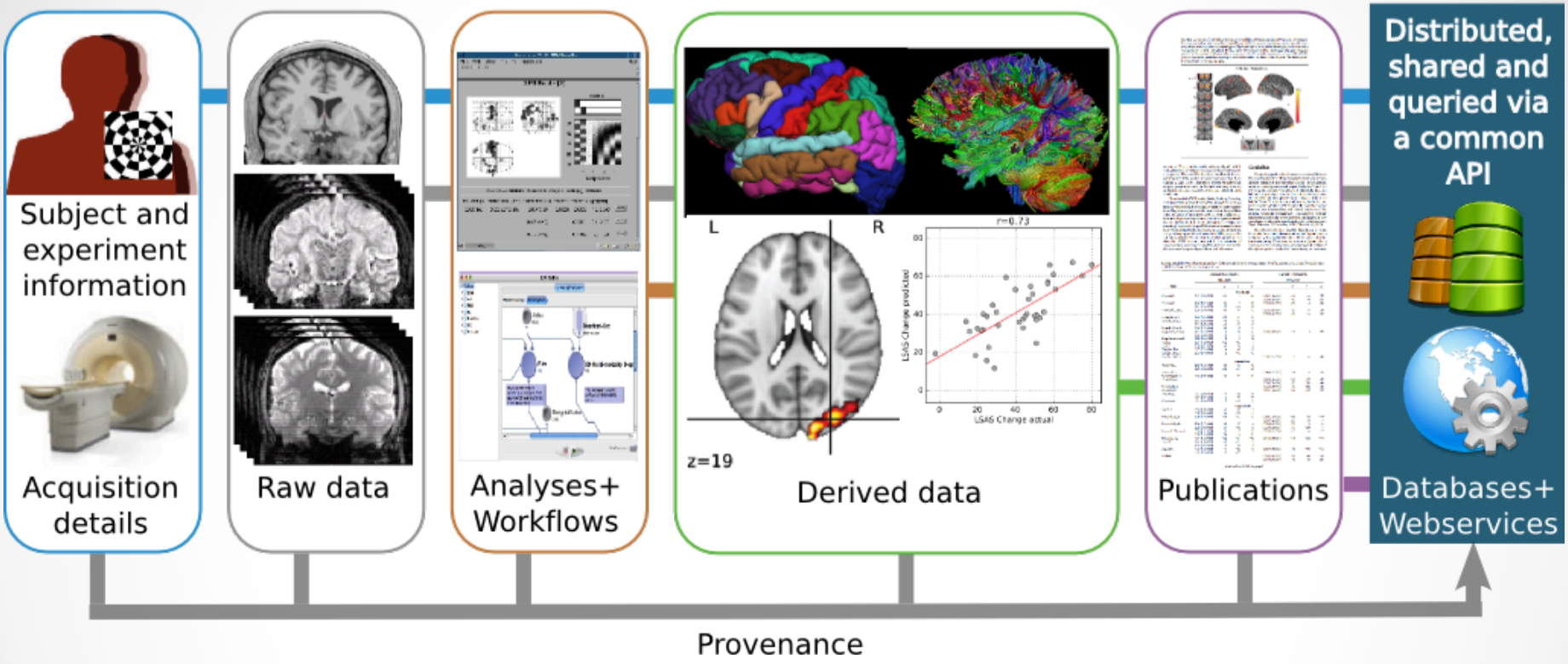


The ‘Reproducibility’ Barrier



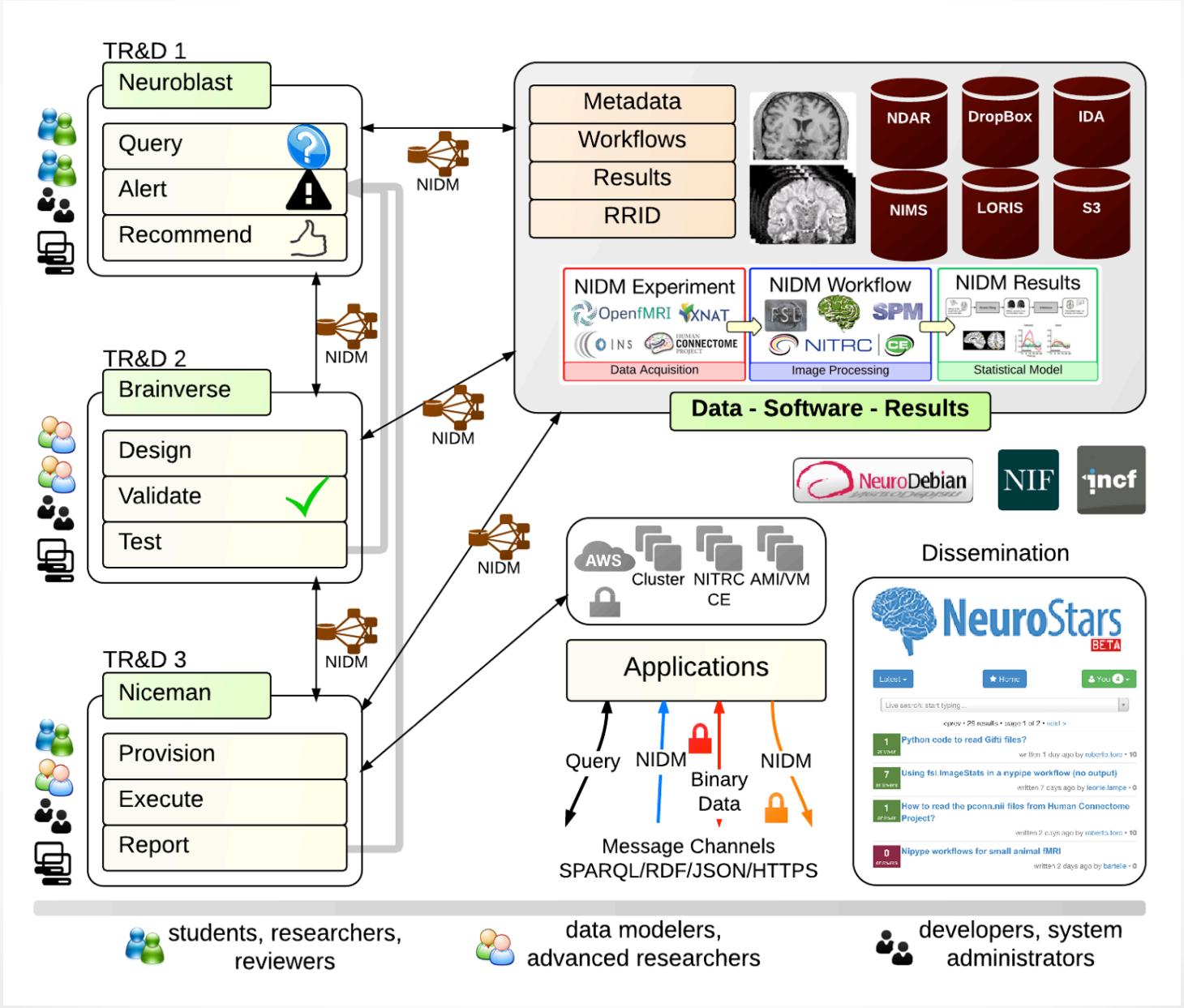
In case you didn't think there was a problem with the current publication system… Multiple small N studies that themselves are not repeatable will not likely lead to a consistent view of a complex disease process…

Framework - Revisited



Demographics	**MRI**	**SPM**	**Contrasts**	**Journals**	**XNAT**
Stimuli	**- T1**	**FSL**	**Correlation**	**Preprint**	**IDA**
Scanner	**- T2**	**FreeSurfer**	**Prediction**	**archives**	**LORIS**
Sequence	**dMRI**	**AFNI**		**Websites**	**COINS**
	fMRI	**...**		**Blogs**	**Amazon S3**
Assays	**- Task**	**Nipype**			**DropBox**
- Clinical	**- Resting**	**LONI**			**Globus**
- Neuropsych	**EEG, MEG,**	**...**			**Neurovault**
- Behavioral	**PET, MRS**	**QA/QC**			**Neurosynth**

A User-centric Model



Part 1:

- User 1 writes a paper (structural MRI analysis) on 103 subjects (mixed gender, w/ and w/o a specific diagnosis, age range 5-18, etc.)
- Data is shared (as per publication requirement) in NITRC.
- 'Data' includes raw MRI scans and the results of specific segmentation workflow.
- Publication1 gets a doi (Pub1doi)

- Shared Data:

U1Sub1raw.nii	U1Sub1proc.nii
U1Sub2raw.nii	U1Sub2proc.nii
▪	
▪	
▪	
U1Sub103raw.nii	U1Sub103proc.nii

Schizophrenia Bulletin vol. 34 no. 1 pp. 37–46, 2008
doi:10.1093/schbul/sbm120
Advance Access publication on November 13, 2007

Diagnostic and Sex Effects on Limbic Volumes in Early-Onset Bipolar Disorder and Schizophrenia

Jean A. Frazier[1–4], Steven M. Hodge[5], Janis L. Breeze[2,3], Anthony J. Giuliano[2,3,6], Janine E. Terry[3], Constance M. Moore[2,3,7], David N. Kennedy[3,5,8], Melissa P. Lopez-Larson[2–4], Verne S. Caviness[5,8,9], Larry J. Seidman[2,6,10], Benjamin Zablotsky[3], and Nikos Makris[5,11]

[2]Department of Psychiatry, Harvard Medical School; [3]Child and Adolescent Neuropsychiatric Research Program, Cambridge Health Alliance; [4]Department of Psychiatry, McLean Hospital; [5]Center for Morphometric Analysis, Massachusetts General Hospital; [6]Department of Psychiatry, Massachusetts Mental Health Center at the Beth Israel Medical Center; [7]Brain Imaging Center, McLean Hospital; [8]Department of Neurology, Harvard Medical School; [9]Department of Pediatric Neurology, Massachusetts General Hospital; [10]Department of Psychiatry, Massachusetts General Hospital; [11]Department of Neurology, Massachusetts General Hospital

Objective: **The limbic structures in early-onset schizophrenia-spectrum illness (SZ) and bipolar disorder (BPD) were studied to discern patterns associated with diagnosis and sex.** ***Methods:*** **Thirty-five youths with DSM-IV BPD with-hormone receptors. In addition, smaller thalamus was associated with SZ while larger right NA volumes were most related to BPD. This study underscores the importance of assessing diagnostic effects and sex effects on the brain in future studies and provides evidence that boys and girls with SZ and BPD may have differential patterns of neuropathology associated with disease expression.**

Key words: mood disorders/psychosis/brain imaging technique/child psychiatry

Introduction

Schizophrenia (SZ) and bipolar disorder (BPD) both typically onset in late adolescence or early adulthood.[1–4] Neurodevelopmental models of both disorders posit subtle disease processes that affect critical circuits in the brain early in development that then manifest in disease around the time of puberty or shortly thereafter.[4–6] Not only is there increasing support for the neurodevelop-

Description: Version 1.1 of the CANDI Share Schizophrenia Bulletin 2008 data.

Subjects ☒ SELECT

<< first < prev 1 2 3 4 5 next > last >> 20 1 of 6 Pgs (103 Rows)

Subject	M/F	Hand	YOB	MR Sessions
BPDwoPsy_030	F	R		1
BPDwoPsy_031	F	R		1
BPDwoPsy_032	M	R		1
BPDwoPsy_033	M	R		1
BPDwoPsy_034	M	R		1
BPDwoPsy_035	M	R		1
BPDwoPsy_036	M	R		1
BPDwoPsy_037	M	R		1
BPDwoPsy_038	M	R		1
BPDwoPsy_039	F	R		1

MR Session: BPDwoPsy_031_MR

Details | Projects

Accession # xnat_E02617
Date Added 2012-02-22 13:37:09.0 (dnkennedy)

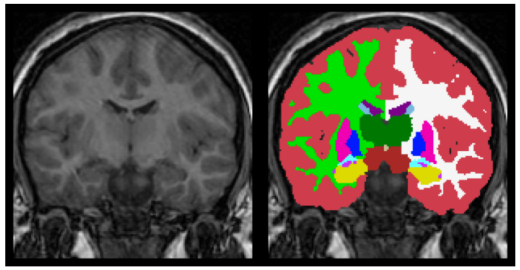
Subject: BPDwoPsy_031
Gender: Female
Handedness: Right
Age: 15

Notes:

Scans

Scan	Type	Usability	Files
anat	anat		Show Counts

Image



Vox. Res. 0.9375, 0.9375, 1.5
Matrix 256 x 256

Part 2:

- Users 2-4 also write papers as above, share data. We have publication doi's for Pub2-4
- We have more shared data:

U2Sub1raw.nii U2Sub1proc.nii

U2Sub2raw.nii U2Sub2proc.nii

▪

▪

▪

U4Sub103raw.nii U4Sub103proc.nii

Part 3:

- User 5 wants to use the shared data to do a new publication, looking at gender differences of 10 year olds in the healthy population. A query on the above shared data identifies 80 of the 270 subjects. The users obtains the raw data for these specific cases and performs another analysis workflow. A paper is written (Pub5doi)
- Data for the paper:

U1Sub2raw.nii	aka U5Sub1raw.nii	U5Sub1proc.nii
U1Sub7raw.nii	aka U5Sub2raw.nii	U5Sub2proc.nii
U1Sub27raw.nii	aka U5Sub3raw.nii	U5Sub3proc.nii
U1Sub31raw.nii	aka U5Sub4raw.nii	U5Sub4proc.nii
U1Sub34raw.nii	aka U5Sub5raw.nii	U5Sub5proc.nii
U2Sub7raw.nii	aka U5Sub6raw.nii	U5Sub6proc.nii
U2Sub14raw.nii	aka U5Sub7raw.nii	U5Sub7proc.nii
.		
.		
.		
U4Sub37raw.nii	aka U5Sub270raw.nii	U5Sub80proc.nii

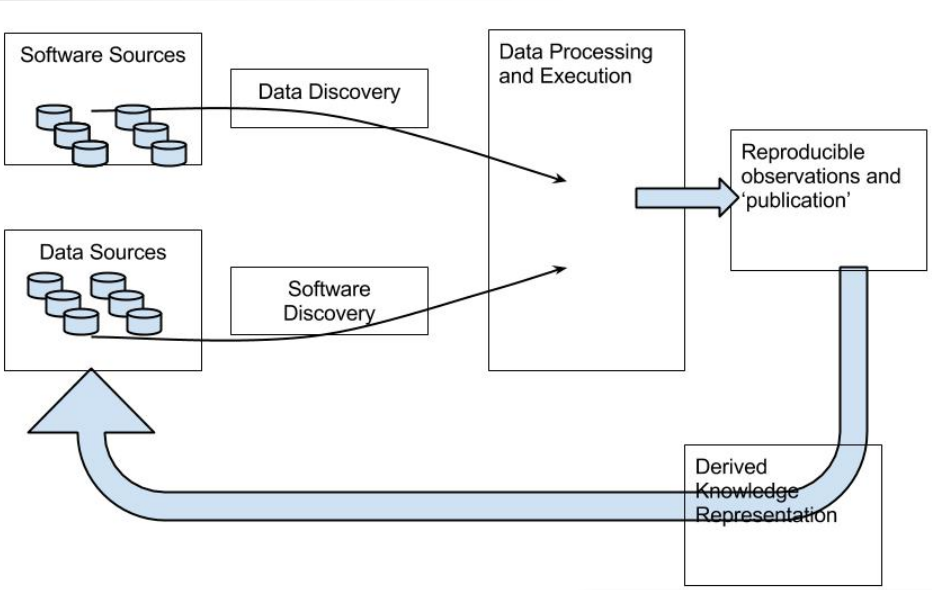
Filter(s): Age = 10

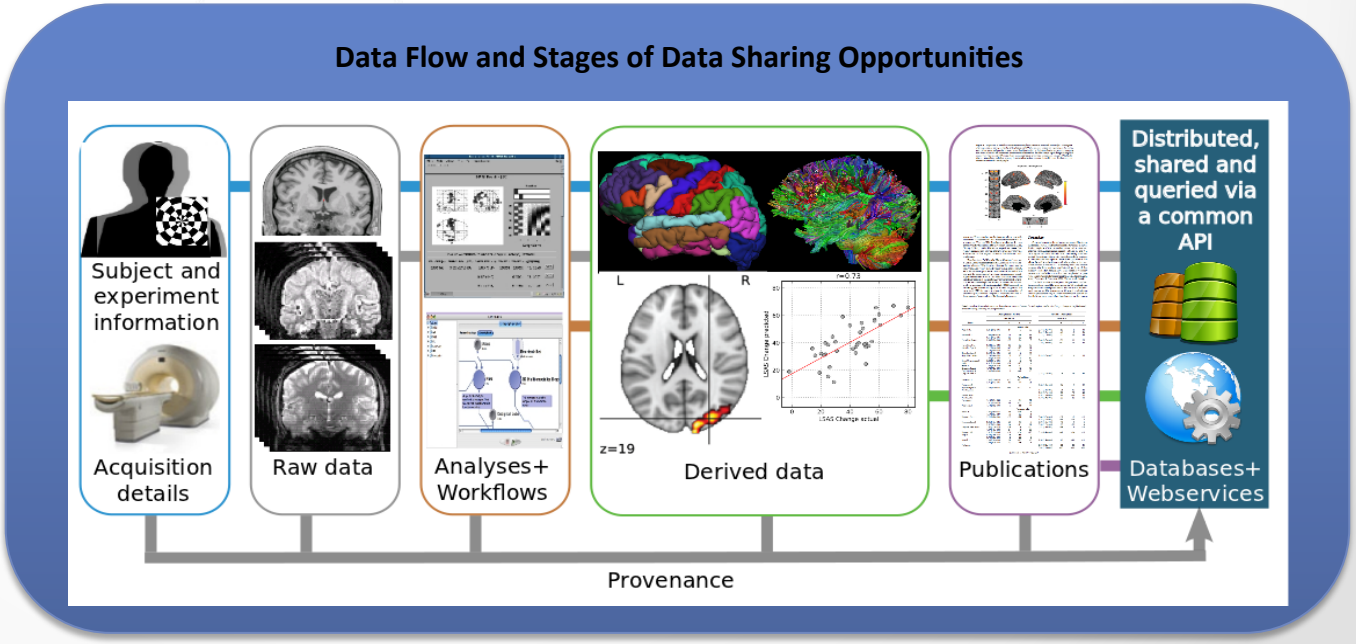
Label	Project	Subject	M/F	Age	Scans	Field Strength	Resting TR
BPDwoPsy_034_MR	cs_schizbull08	BPDwoPsy_034	M	10	anat(1)	1.5T	
BPDwoPsy_049_MR	cs_schizbull08	BPDwoPsy_049	M	10	anat(1)	1.5T	
BPDwoPsy_053_MR	cs_schizbull08	BPDwoPsy_053	F	10	anat(1)	1.5T	
BPDwoPsy_055_MR	cs_schizbull08	BPDwoPsy_055	M	10	anat(1)	1.5T	
BPDwPsy_075_MR	cs_schizbull08	BPDwPsy_075	M	10	anat(1)	1.5T	
BPDwPsy_081_MR	cs_schizbull08	BPDwPsy_081	F	10	anat(1)	1.5T	
HC_005_MR	cs_schizbull08	HC_005	F	10	anat(1)	1.5T	
HC_018_MR	cs_schizbull08	HC_018	F	10	anat(1)	1.5T	
HC_020_MR	cs_schizbull08	HC_020	M	10	anat(1)	1.5T	
HC_023_MR	cs_schizbull08	HC_023	M	10	anat(1)	1.5T	
IPCAS_26058_baseline	corr	IPCAS_26058	M	10	anat(1), rest(2)	3T	2500.0
IPCAS_26062_baseline	corr	IPCAS_26062	M	10	anat(1), rest(2)	3T	2500.0
IPCAS_26064_baseline	corr	IPCAS_26064	F	10	anat(1), rest(2)	3T	2500.0
IPCAS_26065_baseline	corr	IPCAS_26065	M	10	anat(1), rest(2)	3T	2500.0
IPCAS_26074_baseline	corr	IPCAS_26074	M	10	anat(1), rest(2)	3T	2500.0
KKI_1638334_1	adhd_200	KKI_1638334	F	10	anat(1), rest(1)		2500.0
KKI_1652369_1	adhd_200	KKI_1652369	F	10	anat(1), rest(1)		2500.0
KKI_1692275_1	adhd_200	KKI_1692275	F	10	anat(1), rest(1)		2500.0
KKI_1779922_1	adhd_200	KKI_1779922	M	10	anat(1), rest(1)		2500.0

Overview

- **Framework for Sharing**
- **Existing Solutions**
  - **Mandated**
    - **ADNI (IDA)**
  - **Grassroots**
    - **NITRC (Fcon 1000, etc)**
- **Main Problems**
  - **Credit**
  - **Ethics**
  - **Cost**
  - **Harmonization**
  - **Reproducibility**

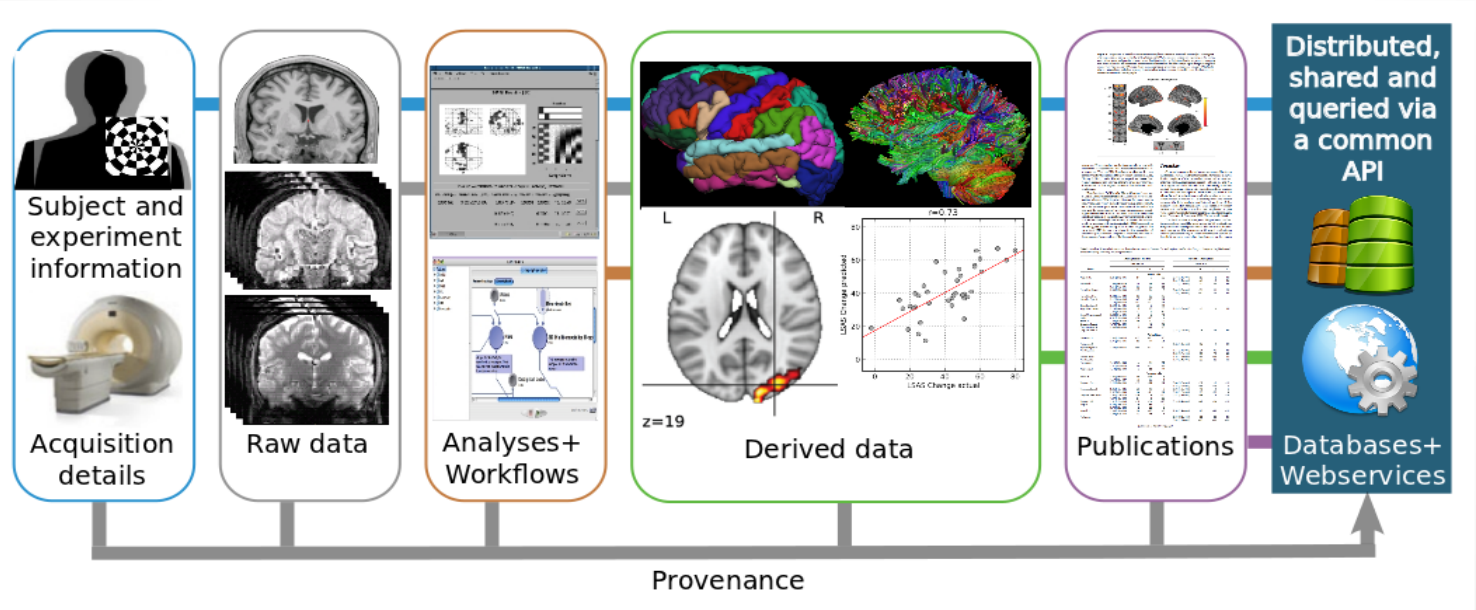
The Research Framework





Resource Sharing: Beyond Data Sharing

- While we are huge proponents of resource sharing overall, in this Center we're not addressing the ’raw data sharing' topic; we are, rather, addressing the sharing of the 'public domain' aspects of publications that have been overlooked: the data descriptors, the workflow descriptors, the execution descriptors and the results descriptors. All of these elements are non-PHI completely sharable elements.



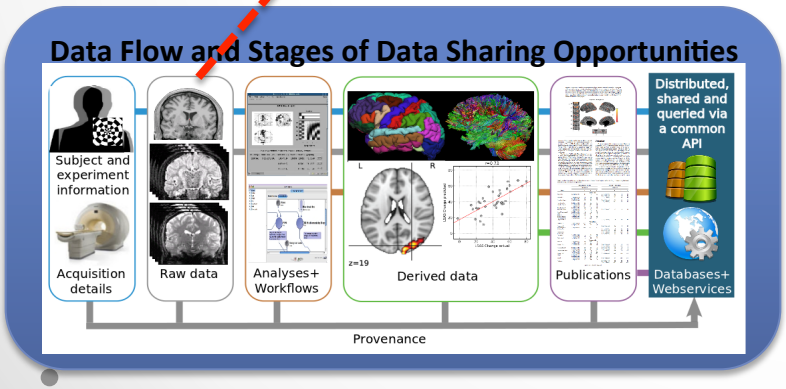
- All data will be marked up via NIDM (http://nidm.nidash.org/) and provided via a NIH-funded NITRC (Neuroimaging Informatics Tools and Resources Clearinghouse - nitrc.org) project. NIDM is the NeuroImage Data Model, an instance of the W3C PROV specification.

One-Click Data Share

Barrier: "Data sharing is too hard…"

Solution: One-Click Data Sharing Tool

Data is anonymized and pushed to the INCF XNAT server. User is prompted for missing metadata. QA is performed and the results are available on line. The user is notified by e-mail.



Planned extensions:

- Promote to community
- Extend QC for structural & diffusion datasets
- Improvements to the user interface & ease of installation
- NIFTI support
- Process large databases to establish distribution of QC across scanners & populations

See http://xnat.incf.org/ to access this resource.

One-Click Data Share

- Drag-n-Drop DICOM directory onto 'INCF Push' app