

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

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.

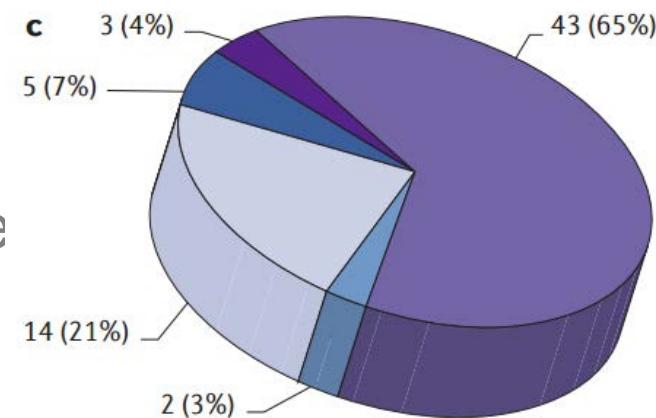


Chart from Prinz, et al. *Nature Reviews Drug Discovery* 10, 712 (September 2011)
Bayer Healthcare

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 ($p < .05$). 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

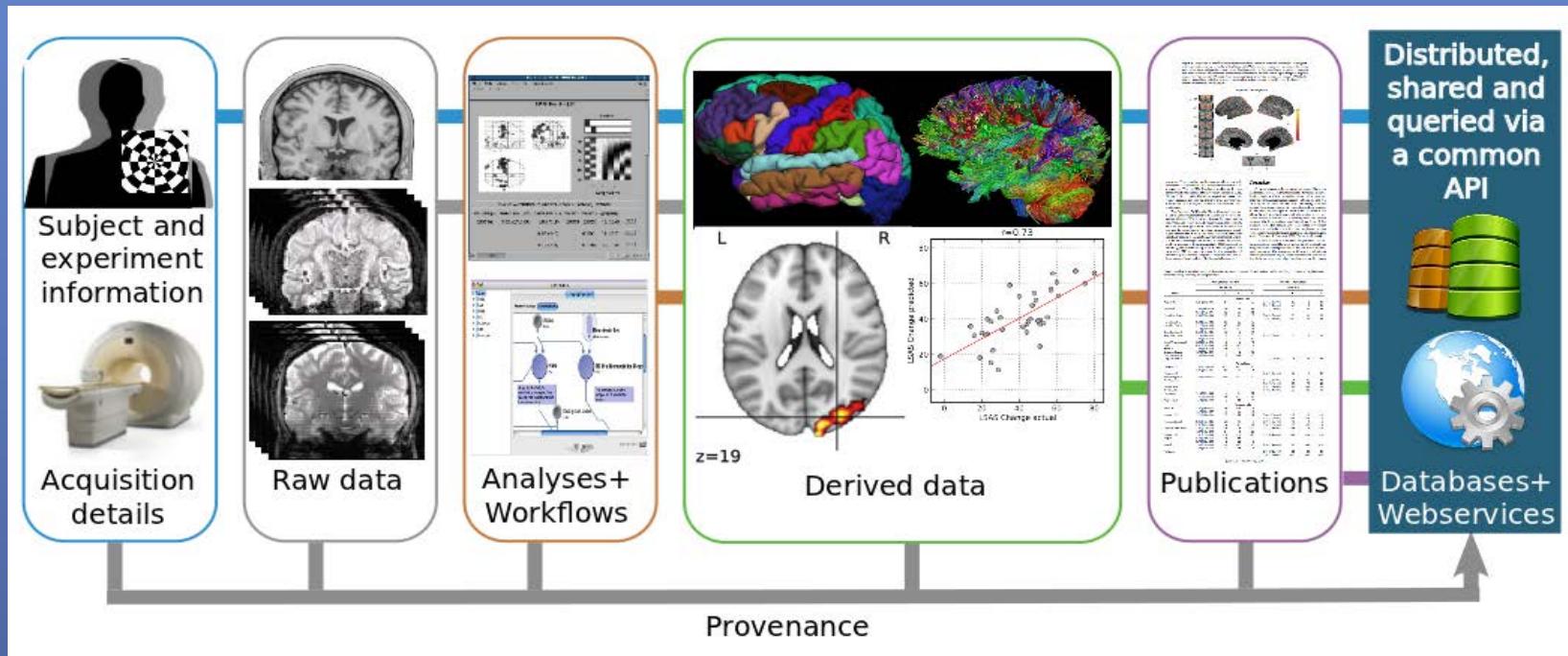
Power failure: why small sample size undermines the reliability of neuroscience

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



Stages of Information Sharing (towards scientific reproducibility)

Data



Analysis



Results



Interpretation

**Data
Repositories**

**Standardized
Workflow
Description**

**Meta-Data
Repositories**

**Literature
Repositories**



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



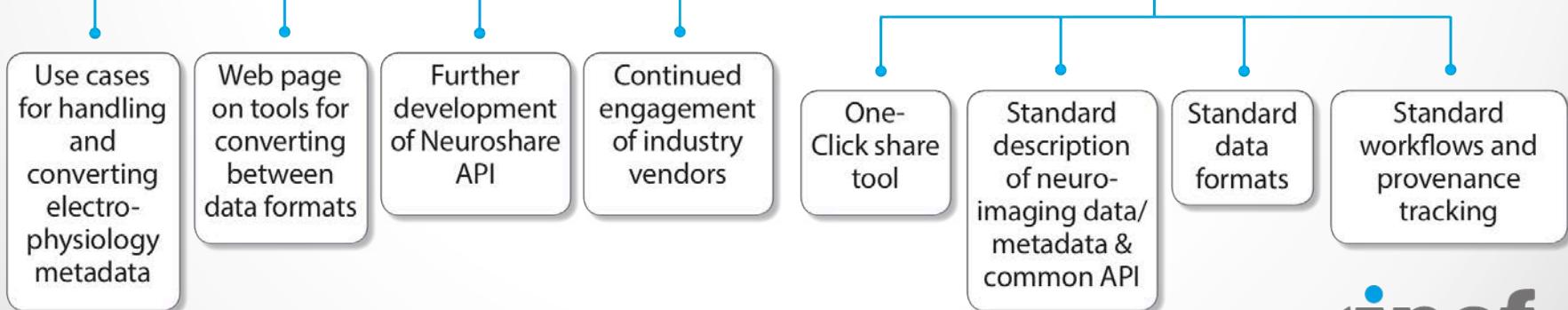
Electrophysiology Task Force

Focuses on electrophysiological data and databases

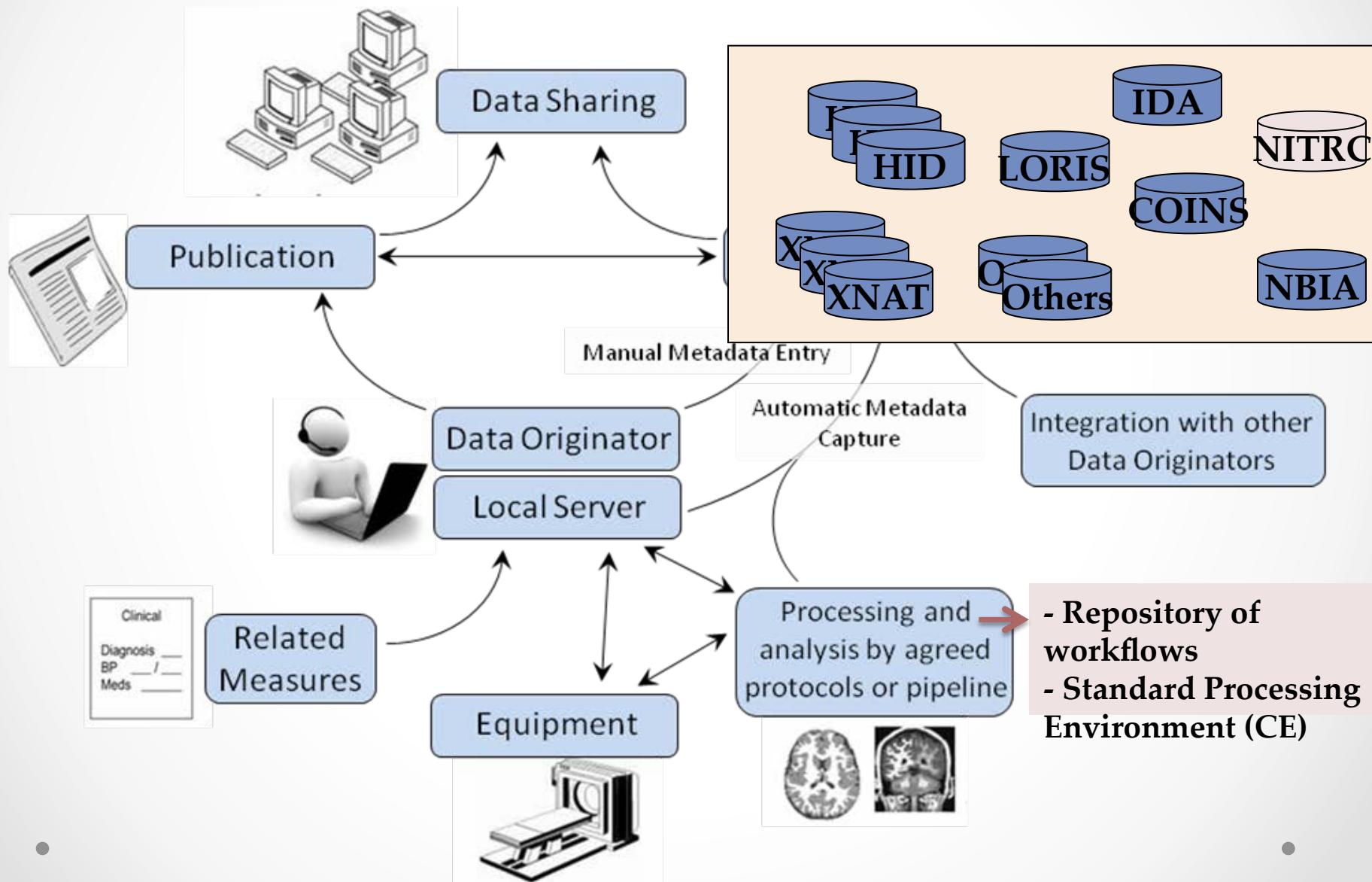


Neuroimaging Task Force

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



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 is out there.

Clinical and Translational CTSAs are funded by the medical discovery to in

- [Brain-CODE](#) at [Harvard Catalyst](#)
- [ICTS](#) at the Univ
- [ICTS](#) at Washington
- [Laboratory of Cll](#)

The Human Connectome Project
XNAT is the foundation from 1,200 subjects at the backbone of the public!

Institutional Operations
XNAT is in use in neuro

- [Biomedical Imaging](#)
- [The Brain Imaging](#)
- [Centre of Excellence in Valencia](#)
- [Centre of Medical Imaging](#)
- [Cornell University](#)
- [Emory University](#)
- [The Genes, Cog](#)
- [Haglund Brain IR](#)
- [Institute of Cancer Research](#)
- [Mayo Clinic Neuroimaging](#)
- [Medical-Image.R](#)
- [New York State I](#)
- [Northwestern Ur](#)
- [Psychiatric Neuro](#)
- [The Rotman Rese](#)
- [University of Wash](#)
- [Vanderbilt Unive](#)
- [Wake Forest Uni](#)

Multi-Site Studies

- [Cardiovascular F](#)
- [IMAGEN Study](#)
- [Dominantly Inher](#)
- [The Jackson Hole](#)
- [PREDICT-HD st](#)
- [The Injury and T](#)
- [International Neu](#)
- [GENetic Frontot](#)
- [Prospective Urb](#)
- [The Virtual Brain](#)
- [CENTER-TBI: la](#)
- [National Consort](#)

Clinical / Pharmacology

- [PharmaCog](#) – P
- [early clinical dev](#)

Data Sharing

- [OASIS Brains](#)
- [mBIRN Morphor](#)
- [XNAT Central](#)
- [Minimal Interval](#)
- [SchizConnect: S](#)

The Qualitative Imaging

The [Quantitative Imaging](#) methods for the measurement of brain structures facilitating clinical decisions

- [PET-MRI for Ass](#)
- [Quantitative MR](#)
- [Quantitative MR](#)
- [Quantitative Im](#)

RESEARCH

COLLABORATIVE INFORMATICS

COINS

NEUROIMAGING SUITE

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 Collaborative Informatics and Neuroimaging Suite (COINS) has been created to facilitate communication.

LOG IN



LONI

Image Data Archive

IDA HOME

ABOUT

NEWS

DOCUMENTATION

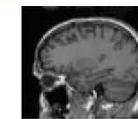
SOFTWARE

DATA

The LONI Image Data Archive (IDA) is a user-friendly environment for archiving, searching, sharing, tracking and disseminating neuroimaging clinical data. The IDA is utilized for dozens of neuroimaging research projects across North America and Europe and accommodates MRI, PET, fMRI, SPECT, and other imaging modalities. A flexible data de-identification engine and encrypted file transmission help ensure compliance with patient-privacy laws. All 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 with an Internet connection and web browser software. The IDA automatically extracts relevant metadata from the de-identified image files allowing data to be searched and retrieved for archival. Once archived, data may be downloaded and/or streamed into the LONI Pipeline processing environment. Integration of the LONI Data Archiving System (DAS) 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 Debab
- Inspector

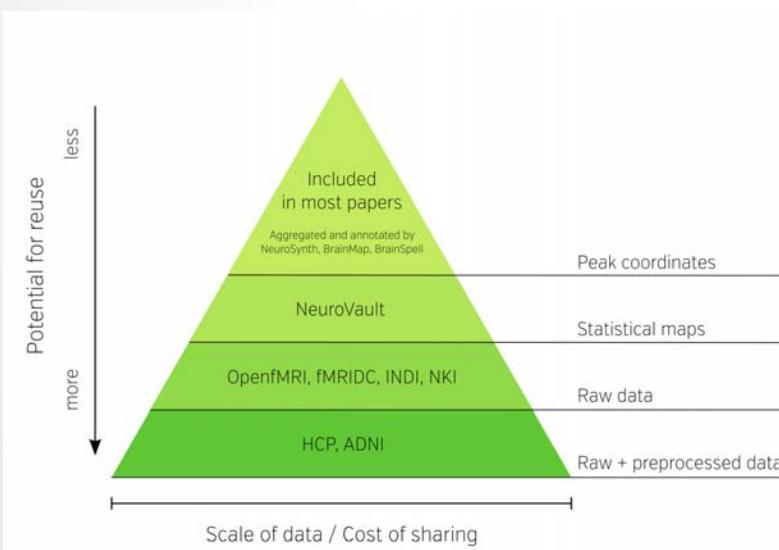
DATA

- Available
- Download

© 2015 LONI. All rights reserved.

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

[Home](#) [Res](#)



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

[Get started and upload an image!](#)

Supported by



Latest collections of images

Name

[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

Number
of
images

Integrated Views

PING Data Portal - DataExploration

Multimodal Imaging Laboratory, UCSD
This project is funded by the NIH/NIDA

PING

Projects Applications Documents Forum Help My account Log out

Compute ROI Statistics Compute Vertex Statistics

MRI_cort_area.ctx.total Area expansion factor

Covariates Predictor Variable s(Age) Interaction by Gender

Dev SES GAF

Data Version: latest Color plot by: Interaction Search: lookup name Help: Ontology ROI

Get Permalink To This Page

Open Expert Mode

PING data
Source: MMIL

Total Cortical Area in mm²-Whole Brain

Age at MRI

Highcharts.com

R^2 (adj.): 0.3962025

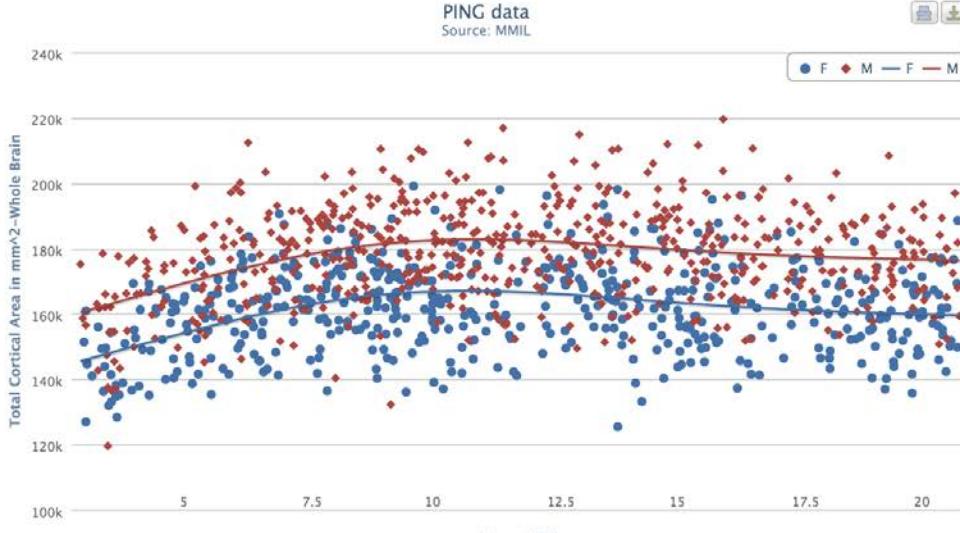


Figure 1: The coefficient of determination R^2 (adjusted) is 0.3962025 ($n = 1099$). The AIC score of the full model is 24022.45.

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 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>

Amazon Web Services Home

Sign in or Create a new account Your Account | Help | Sell on AWS Marketplace

ce GO Your Software

NITRC

Visit the NITRC Website

NITRC Products (2)

 **NITRC Computational Environment**
Free Tier Eligible ★★★★★ (3) | Version v0.30-all regions | Sold by [NITRC](#)
\$0.00/hr for software + AWS usage fees
NITRC-CE is a virtual computing platform pre-configured with many neuroimaging data analysis applications. NITRC-CE joins the family of successful NITRC services starting ...
Linux/Unix, Ubuntu 12.04 | 64-bit Amazon Machine Image (AMI)

 **NITRC Computational Environment for Cluster Compute Instances**
Version v0.30 CC | Sold by [NITRC](#)
\$0.00/hr for software + AWS usage fees
NITRC-CE for Cluster Compute Instances provides the same virtual computing platform for neuroimaging data analysis as NITRC-CE but on high performance computing machines ...
Linux/Unix, Ubuntu 12.04 | 64-bit Amazon Machine Image (AMI)

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



Amazon Web Services Home

Your Account | Help | Sell on AWS Marketplace

GO Your Software

Shop All Categories ▾ Search AWS Marketplace

NITRC Computational Environment

Sold by: NITRC

NITRC-CE is a virtual computing platform pre-configured with many neuroimaging data analysis applications. NITRC-CE joins the family of successful NITRC services starting with the flagship, NITRC-Resources, the "go to" place for neuroimaging tools and resources. NITRC Image Repository offers a select set of community-generated neuroimaging data sets, while this service, NITRC Computational Environment, offers the convenience of cloud-based computing against NITRC-IR data sets or your data sets. We welcome any suggestions on how to improve this service to make it a user friendly tool for ... [Read more](#)

Customer Rating (3 Customer Reviews)

Latest Version v0.30-all regions (Other available versions)

Base Operating System Linux/Unix, Ubuntu 12.04

Delivery Method 64-bit Amazon Machine Image (AMI) (Learn more)

Support See details below

AWS Services Required Amazon EC2, Amazon EBS

Highlights

- Need resources on demand to compute against your neuroimaging data? Tired of fighting for institutional compute resources and just need to get the compute done? Use NITRC-CE!
- Need access to the most popular neuroimaging analysis tools? Each release has more of the most popular neuroimaging tools, check our User Guide for a complete listing of installed packages. Use these resources separately, or pipeline them; we're agnostic!
- Need access to the most popular community-generated and curated neuroimaging analysis data sets? Access

Continue

You will have an opportunity to review your order before launching or being charged.

Pricing Details

For region [US East \(Virginia\)](#)

Free Tier Eligible

This product can be used for free on a Micro instance for up to 750 hours per month if you qualify. See details.

Hourly Fees

Total hourly fees will vary by instance type and EC2 region.

EC2 Instance Type	Software	EC2	Total
Standard Micro (t1.micro)	\$0.00/hr	\$0.02/hr	\$0.02/hr
Standard Small (m1.small)	\$0.00/hr	\$0.08/hr	\$0.08/hr
Standard Medium (m1.medium)	\$0.00/hr	\$0.12/hr	\$0.12/hr
Standard Large (m1.large)	\$0.00/hr	\$0.24/hr	\$0.24/hr
Standard XL (m1.xlarge)	\$0.00/hr	\$0.48/hr	\$0.48/hr
High-Memory XL (m2.xlarge)	\$0.00/hr	\$0.41/hr	\$0.41/hr
High-Memory 2XL (m2.2xlarge)	\$0.00/hr	\$0.82/hr	\$0.82/hr
High-Memory 4XL (m2.4xlarge)	\$0.00/hr	\$1.64/hr	\$1.64/hr

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





Home Tools & Resources Community Support About NITRC

SEARCH Tools/Resources ▾

Help | + Share | [f](#) [e](#) [g](#) [GO](#)

Instructions

From this screen you can start a new NITRC-CE session if one is not already running, or you can connect to or end a NITRC-CE session that is already running. If you have any unsaved work in a session that is ended, it will be lost. If you are having problems connecting, [try these instructions](#) for connecting to NITRC-CE through your organization's firewall.

Start Session - Start a VNC session

Logout - Log out from NITRC-CE. (Note: This will not close your Amazon EC2 instance.)

Software License Status

AFNI Status: [Installed](#) - [Freeware](#)

Ants Status: [Installed](#) - [BSD License](#)

DTIPrep Status: [Installed](#) - [BSD License](#)

FreeSurfer Status: [License not installed](#) - [FreeSurfer License](#)
To update your FreeSurfer License, please [click here](#).

FSL Status: [Installed](#) - [FSL License](#)

ITK-SNAP Status: [Installed](#) - [GNU General Public License \(GPL\)](#)

LONI Pipeline Server Status: [Installed](#) - [LONI Software License](#)
Launch a [Pipeline web start client](#) configured to use this server: [Launch](#)

MRtrix Status: [Installed](#) - [GNU General Public License \(GPL\)](#)

R Status: [Installed](#) - [GNU General Public License V2 \(GPL V2\)](#)

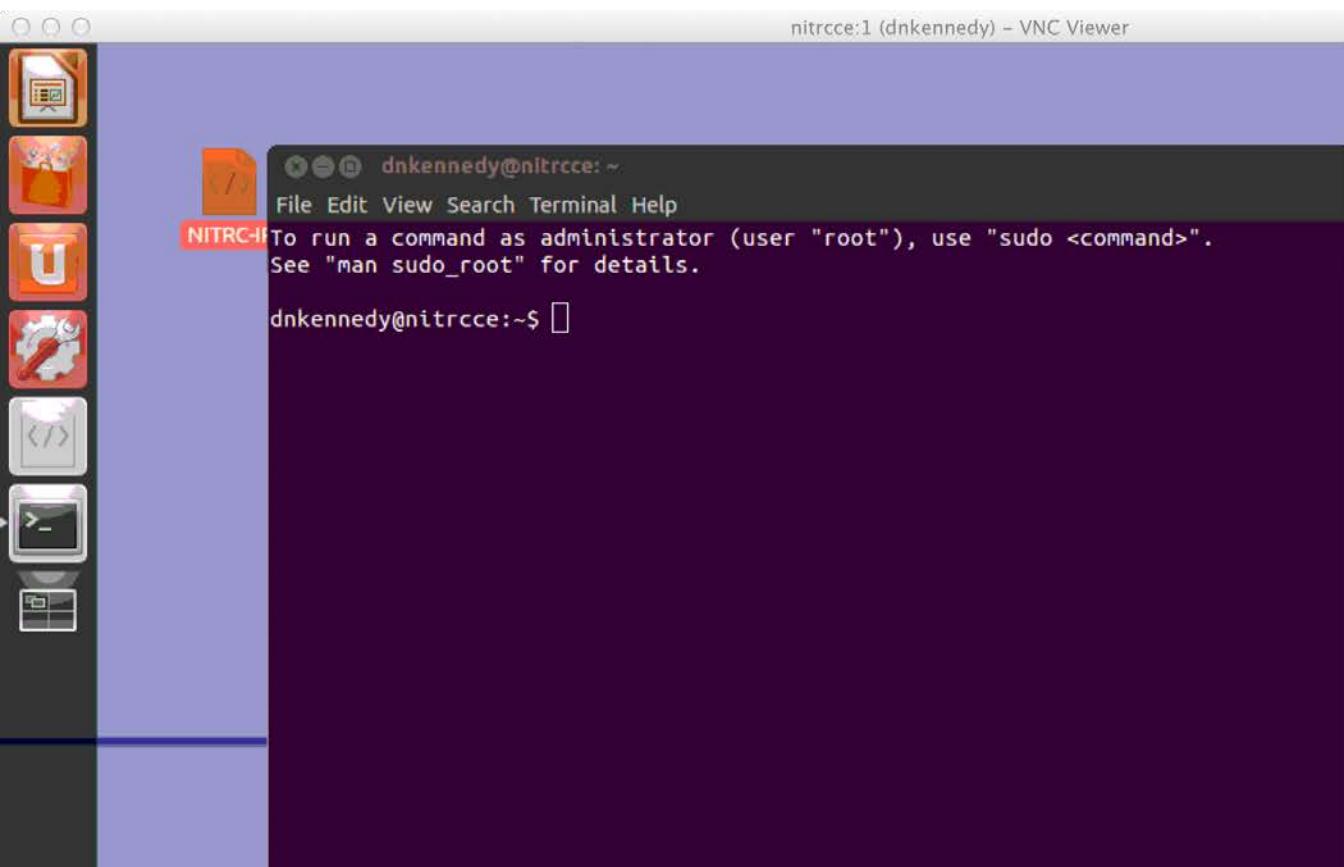
For a complete list, please go to: [User Guide - NITRC-CE Installed Packages](#)

[Feedback](#) | [Privacy](#) | [Accessibility](#) | [Copyright](#) | [Glossary](#) | [Site Map](#)

Registered with **NIF** NATIONAL INSTITUTE OF

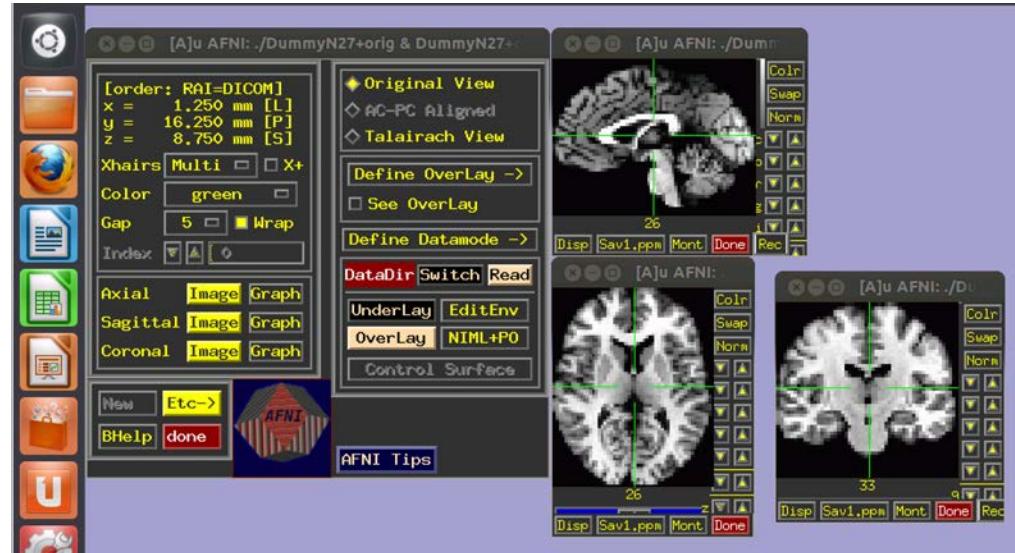
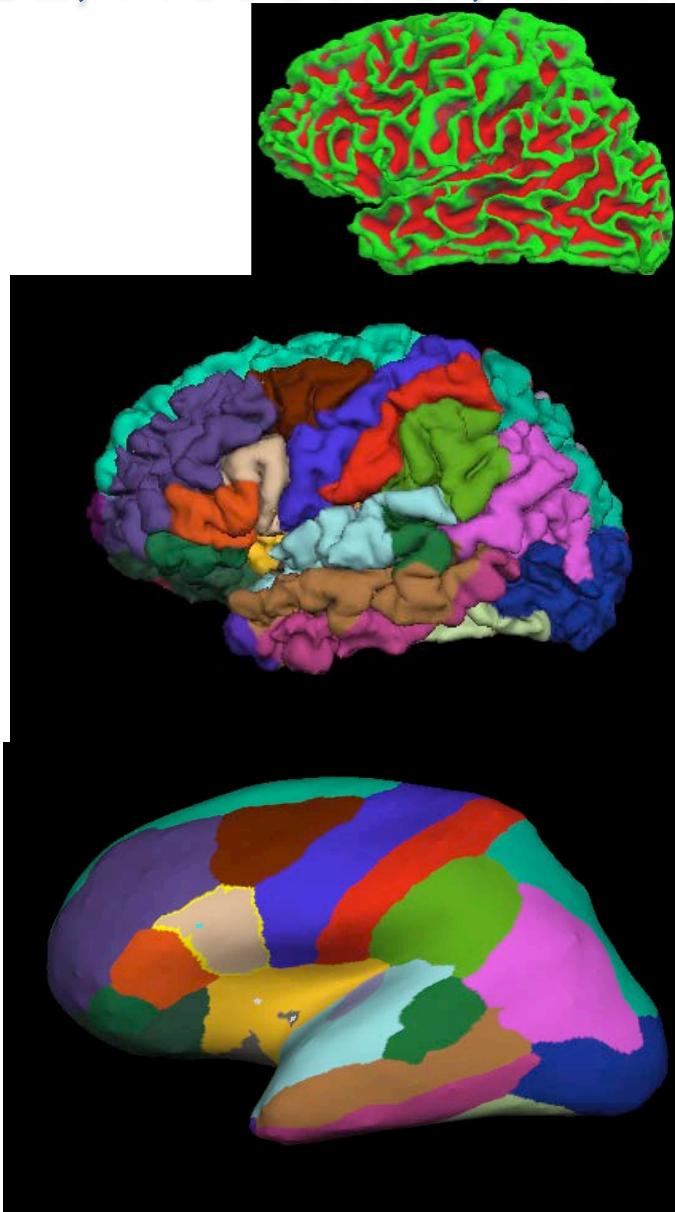
for Neuroscience Research



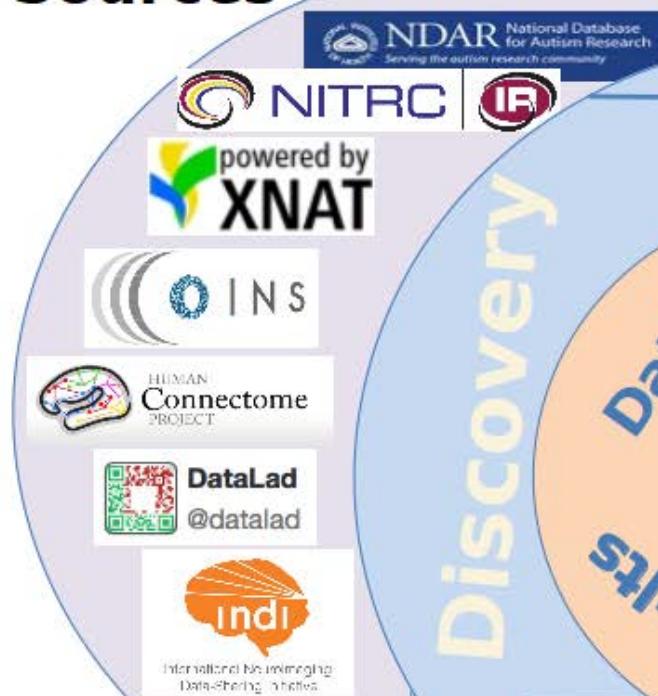
NITRC



Powered By
NeuroDebian



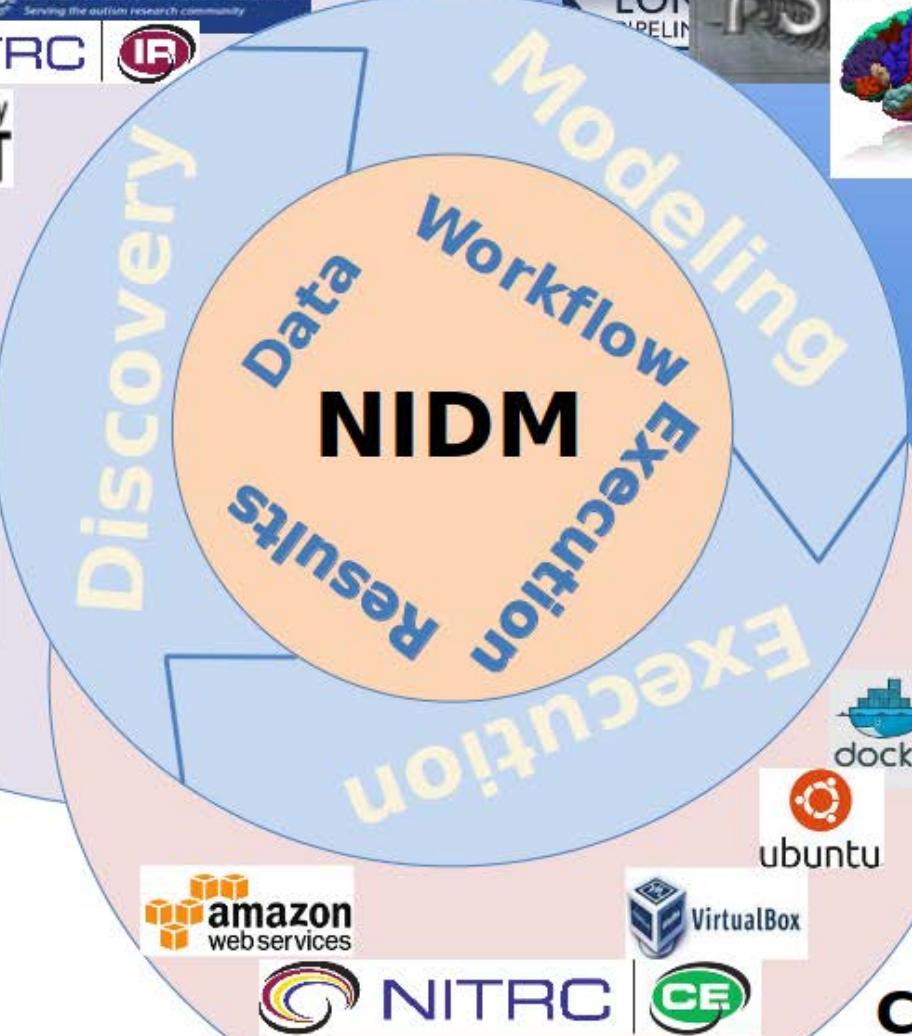
Data Sources



Software & Workflows



Compute Environments





Provenance and NI-DM:

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



www.incf.org/about/programs/datassharing





Provenance and NI-DM:

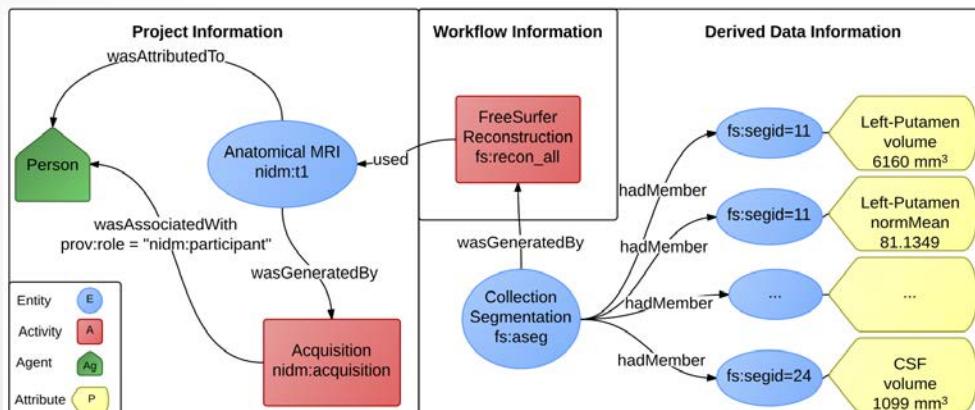
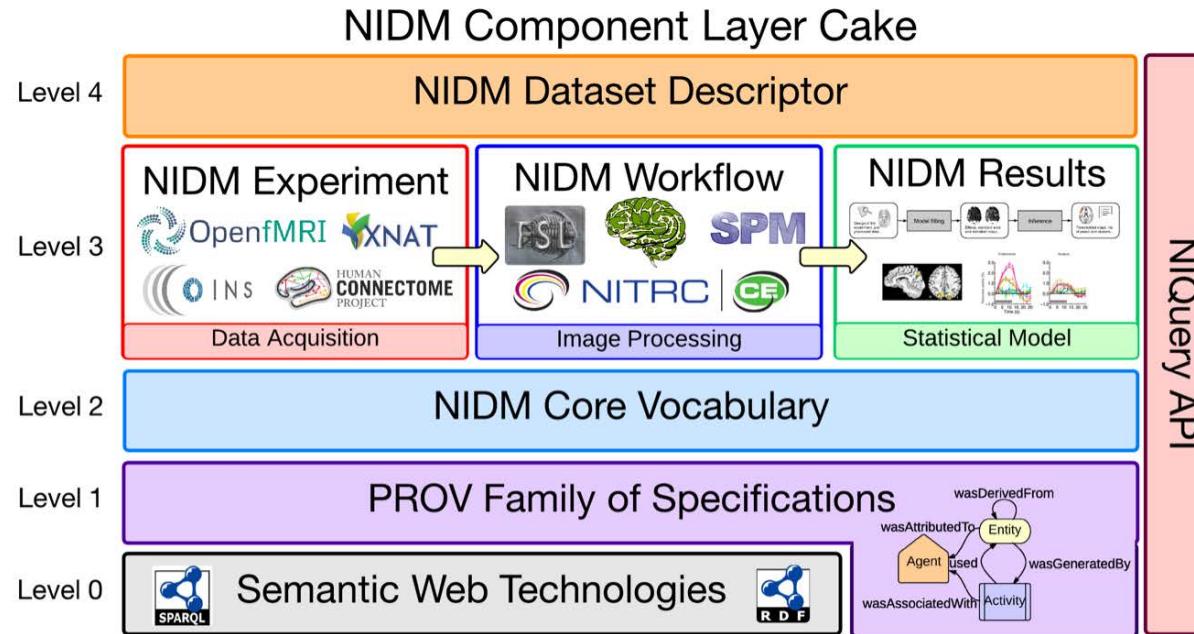
- **The NI-DM data model**

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

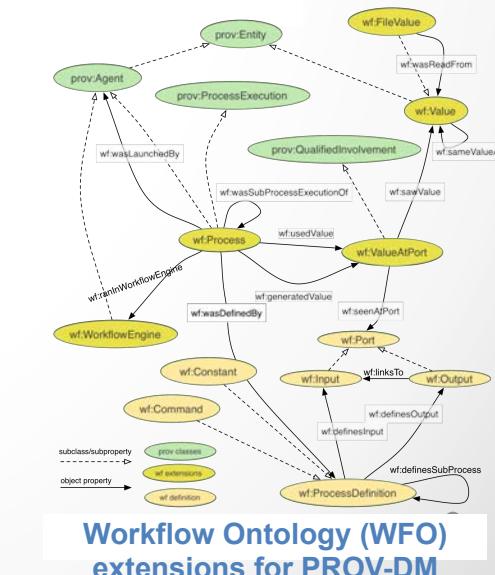
- Based on PROV-DM hence borrows PROV ontology (PROV-O)
- Structured information encoding
- Consistent vocabulary
- Metadata standards via domain specific object models

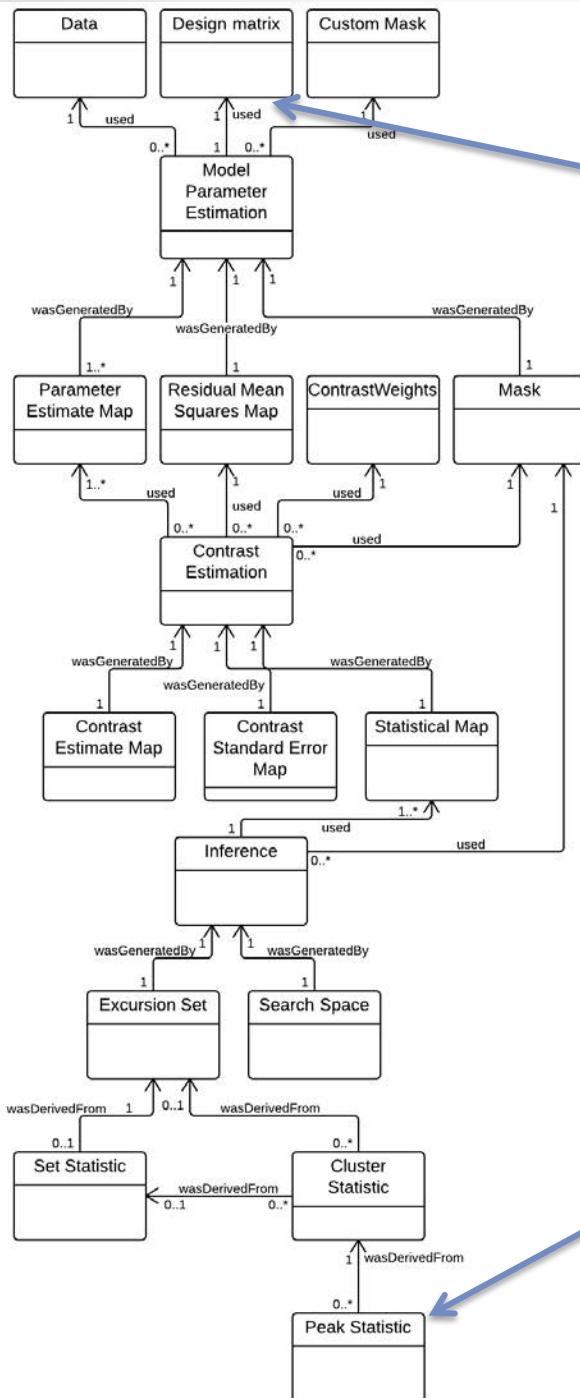
- Terms
 - A lexicon of all things brain imaging. (e.g., DICOM terms, software specific terms, statistic terms, paradigm terms)
- Object Models
 - Structured information in brain imaging (e.g., directory structures, CSV/Tab delimited files, brain imaging file formats)
- Integrated provenance
 - How are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

Provenance and NI-DM

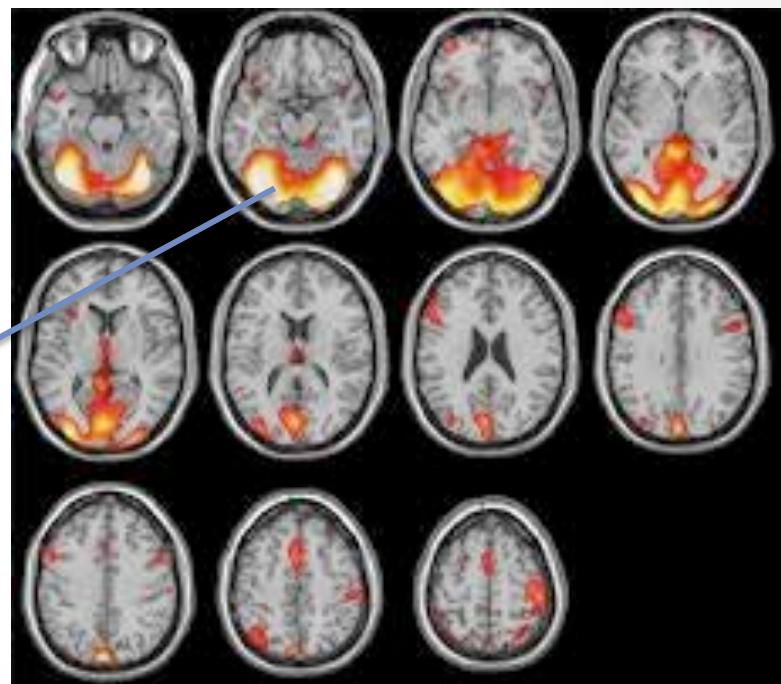


NI-DM can bridge information dimensions across Project, Workflow, and Derived Data





NI-DM in task activation

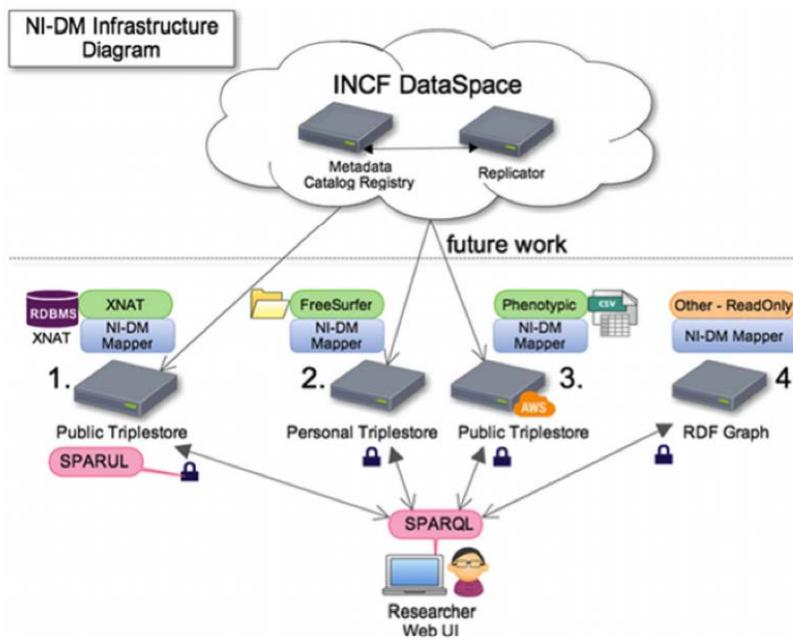


SPM 2.6



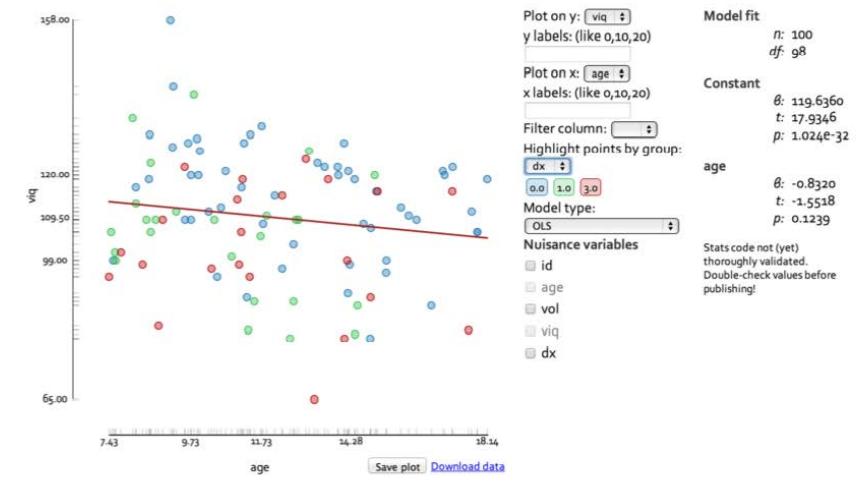
Provenance and NI-DM:

- 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...

ANGRY MOB FUN RUN



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: if a typical hour of MRI scanning costs \$500, then this would be \$25 for archival of a 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 novel cohorts
- Cohort DOIs
- Derived data DOIs (or handles)

Creator *

Name: * [?] - +

Name Identifier: [?] - +

Identifier Scheme: [?] - +

Scheme URI: [?] - +

Title *

Title: * [?] - +

Type: Main title Alternative title Subtitle Translated title

Publisher: *

[?] - +

Publication Year: *

[?] - +

Resource Type:

[?] - +

Resource Type Description:

[?] - +

Abstract

Descriptive information: [?] - +

Type: [?] - +

Subject

Subject: [?] - +

Scheme: [?] - +

Scheme URI: [?] - +

Date

Date: [?] - +

Date Type: [?] - +

Alternate Identifier

Identifier: [?] - +

Identifier Type: [?] - +

Integrated Publication

NITRC The source for neuroimaging tools and resources

NITRC Fast Fact: Do you have an event or job to announce? Go to Community and Submit Community News or a Job Posting!

CANDI Neuroimaging Access Point

Summary

Reviews/Ratings

Support

Admin

Advanced Search

Docs

Downloads

Forums

News

Tasks

Tracker

Image Repository

Members

Admin

Christian Haselgrave

David Kennedy

Developer

Ann Foley

Pallavi Rane

Senior Developer

Recent Activity

Show: [- Show All -](#)

Specifications

Category: Atlas Data, Data, Knowledge Environment, Web Resource, Volumetric Analysis, Segmentation

Licenses: Attribution

Show more specifications (Development Status, Domain, Environment, etc)

Associations

has subpart: CANDI Share: Schizophrenia Bulletin 2008

Show more associations (from the makers of, works well with)

Recent Activity

Show: [- Show All -](#)

Main Folder documentation

CANDIShare Data Dictionary posted by David Kennedy on Mar 2

Tool & Resource news

SchizBull_2008 version 1.2 released posted by David Kennedy on Mar 1

Reviews & Ratings

User Ranks (1)

OVERALL:

INSTALLATION:

DOCUMENTATION:

Participate!

Submit news

Report issues

Add a review

Monitor a file release

Subscribe to RSS feed

Bookmark this page

CANDI Dynamic Probability Server

Recent Data

Diagnosis:

Gender:

Age:

Handedness:

12 matches found:

NITRC-IR

Subjects

Projects

Individual Data MNI152_T1_2mm Space

CANDIShare - Downloads

Releases

Name	Date	File Name	Size	DL	Arch	Type
Dynamic Atlases	2012-02-28 16:44	DynProbServer (url)	14	Any	URL	
SchizBull_Atlas_V1.1	2012-02-14 19:19	Atlases_V1.1.tar.gz	8.26 MB	5	Any	gt
SchizBull_Atlas_V1.0	2011-12-05 03:33	Atlases_V1.0.tar.gz	3.29 MB	11	Any	gt

Package: SchizBull_2008 (2 Subscribers)

Releases

Name	Date	File Name	Size	DL	Arch	Type
SchizBull_2008 on NITRC-IR	2012-02-28 19:41	Data on IR (requires registration) (url)	22	Any	URL	
ImgSegReg_V1.2	2012-03-14 21:59	SchizBull_2008_BPDwoPsy_segmreg_V1.2.tar.gz	113.58 MB	5	Any	gt
SchizBull_2008_BPDwoPsy_segmreg_V1.2.tar.gz	61.01 MB	4	Any	gt		
SchizBull_2008_HD_segmreg_V1.2.tar.gz	90.21 MB	7	Any	gt		
SchizBull_2008_SS_segmreg_V1.2.tar.gz	65.14 MB	4	Any	gt		
ImgSeg_V1.1	2011-07-06 16:11	SchizBull_2008_BPDwoPsy_segmreg.tar.gz	89.48 MB	14	Any	gt
SchizBull_2008_BPDwoPsy_segmreg.tar.gz	48.02 MB	11	Any	gt		
SchizBull_2008_HD_segmreg.tar.gz	70.32 MB	25	Any	gt		
SchizBull_2008_SS_segmreg.tar.gz	51.68 MB	11	Any	gt		
SS_V1.0	2010-07-27 23:11	Link to this group in IBVD (url)	174	Any	URL	
SchizBull_2008_SS_Procimg.tar.gz	49.02 MB	22	Any	gt		
BDwithPsy_V1.0	2010-07-27 17:46	Link to this group in IBVD (url)	109	Any	URL	
SchizBull_2008_BPDwoPsy_procimg.tar.gz	45.52 MB	17	Any	gt		
BDwithoutPsy_V1.0	2010-07-08 00:36	Link to this group in IBVD (url)	68	Any	URL	
SchizBull_2008_BPDwoPsy_procimg.tar.gz	84.87 MB	16	Any	gt		
HD_V1.0	2010-07-01 16:00	Link to this group in IBVD (url)	83	Any	URL	
SchizBull_2008_HD_procimg.tar.gz	66.38 MB	24	Any	gt		

Link to Dynamic Atlases

Link to Static Atlases

Link to NITRC-IR

Images & Segmentation

Image Downloads and Links to IBVD

'Static' Probability Atlases MNI152_T1_2mm Space

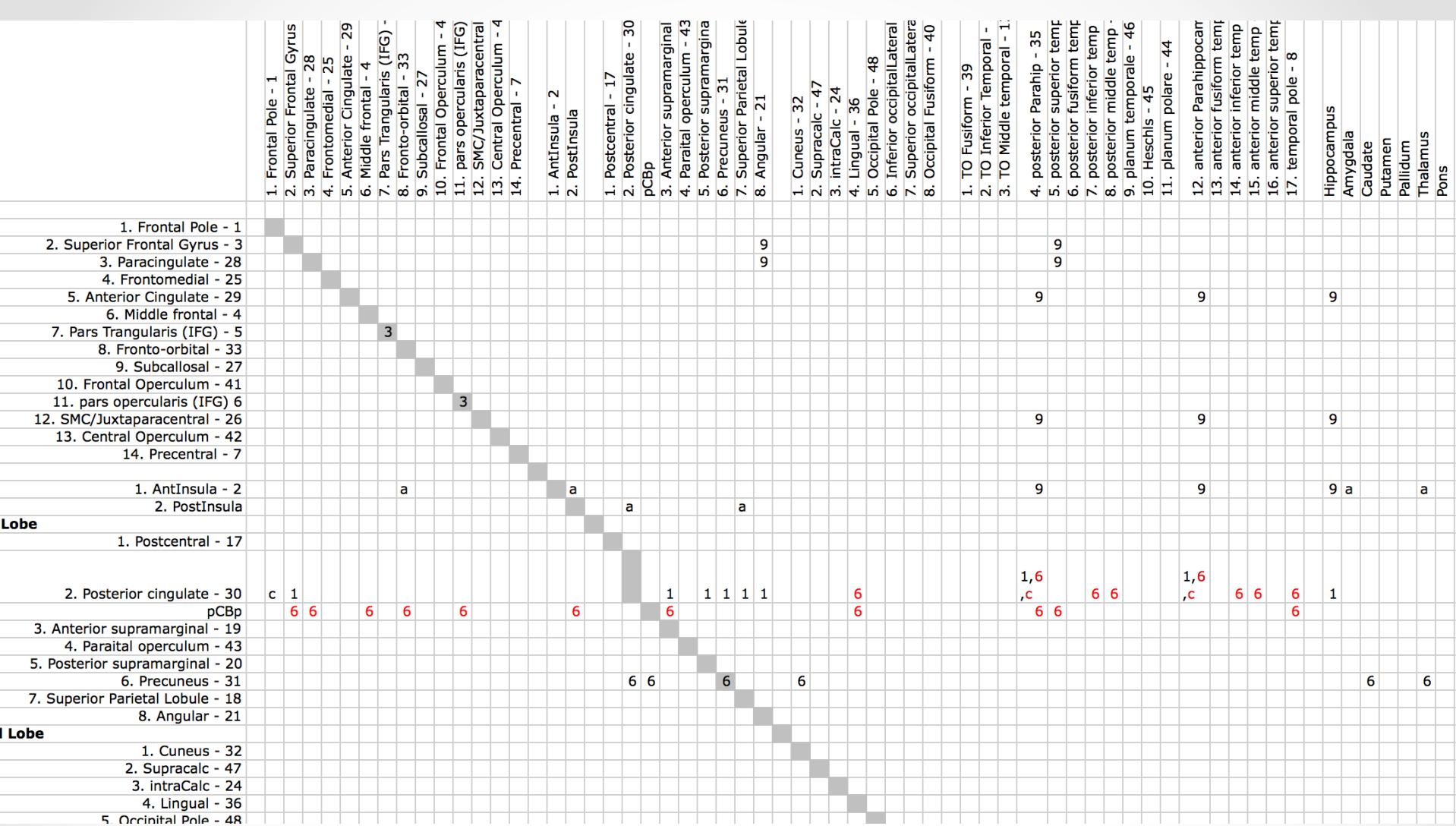
Individual Data MNI152_T1_2mm Space

Linear Registration

Non-linear Registration

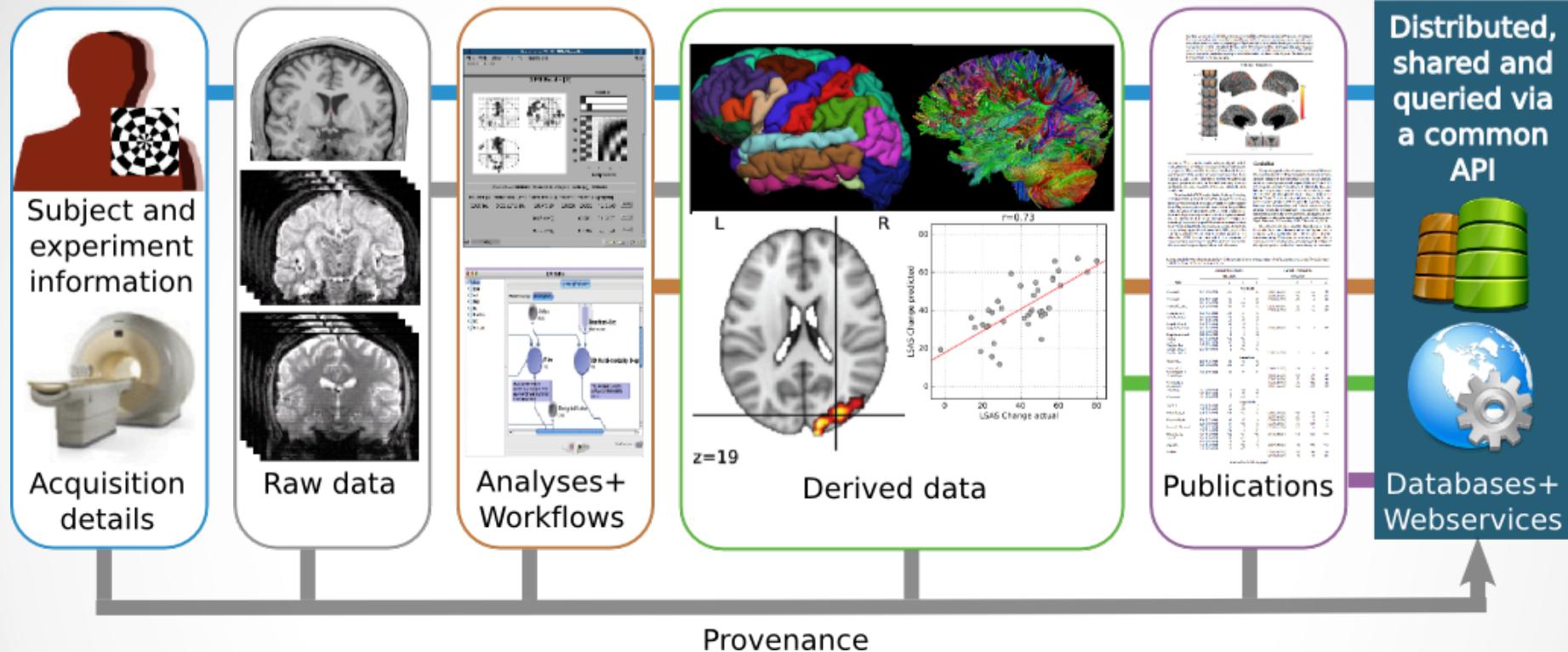
http://www.nitrc.org/projects/candi_share/

The ‘Reproducibility’ Barrier



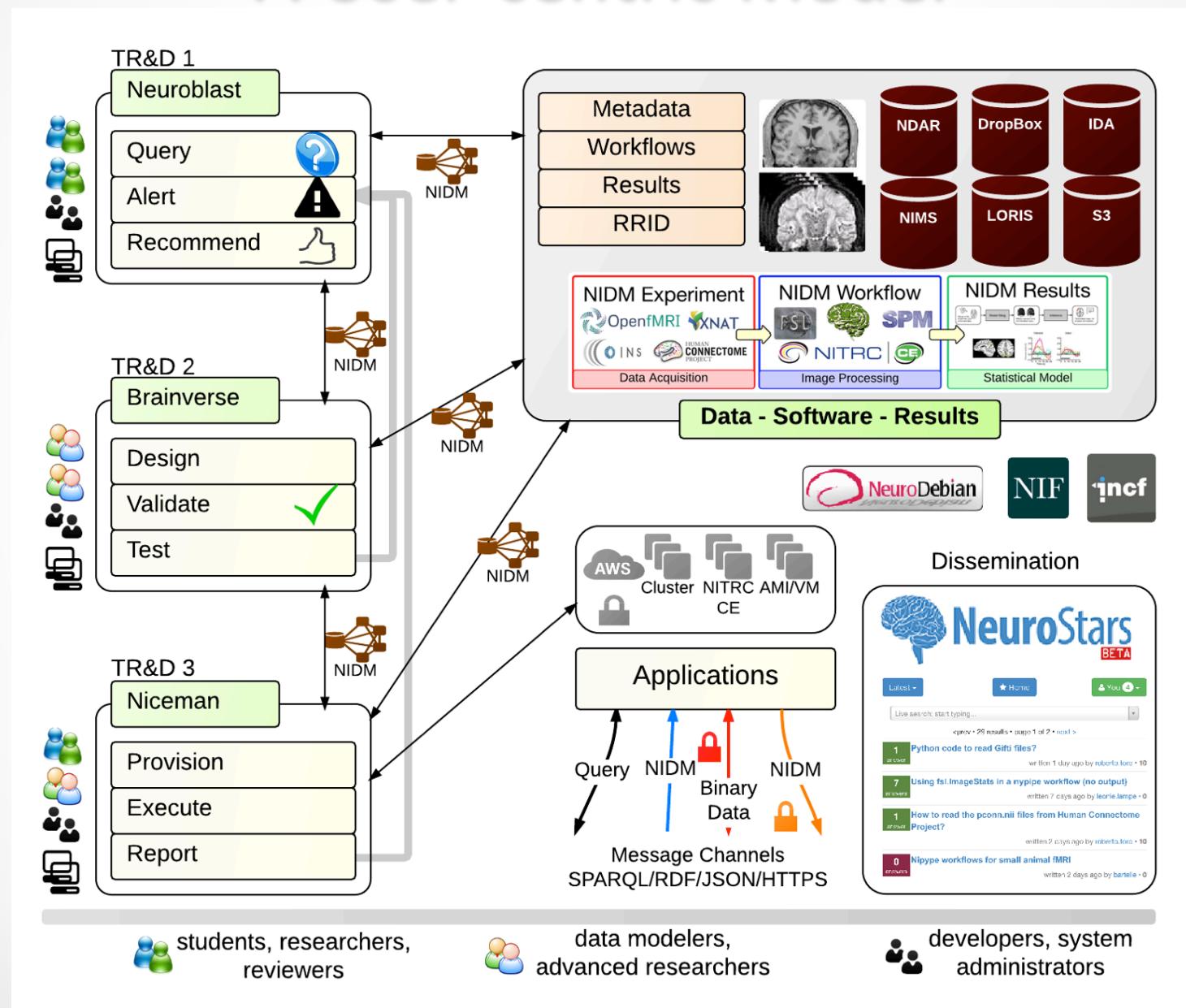
In case you didn't think there was a problem with the current publication system... Multiple small N studies that themselves are not repeatable will not likely lead to a consistent view of a complex disease process...

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

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

Jean A. Frazier^{1–4}, Steven M. Hodge⁵, Janis L. Breeze^{2,3},
Anthony J. Giuliano^{2,3,6}, Janine E. Terry³,
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 Zablotsky³, and
Nikos Makris^{5,11}

²Department of Psychiatry, Harvard Medical School; ³Child and Adolescent Neuropsychiatric Research Program, Cambridge Health Alliance; ⁴Department of Psychiatry, McLean Hospital; ⁵Center for Morphometric Analysis, Massachusetts General Hospital; ⁶Department of Psychiatry, Massachusetts Mental Health Center at the Beth Israel Medical Center; ⁷Brain Imaging Center, McLean Hospital; ⁸Department of Neurology, Harvard Medical School; ⁹Department of Pediatric Neurology, Massachusetts General Hospital; ¹⁰Department of Psychiatry, Massachusetts General Hospital; ¹¹Department 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 neurodevelopmental model for both disorders but there is also emer-

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

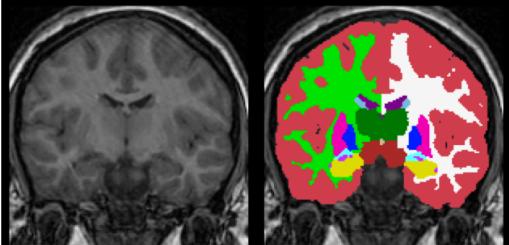
Accession # xnat_E02617	Subject: BPDwoPsy_031
Date Added 2012-02-22 13:37:09.0 (dnkennedy)	Gender: Female
	Handedness: Right
	Age: 15

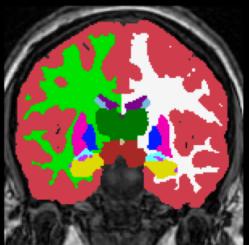
Notes:

Scans

Scan	Type	Usability	Files
<input type="checkbox"/> 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

U2Sub2raw.nii

■

■

■

U2Sub1proc.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 user 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
⋮	⋮	⋮

<< first < prev

1 2 3 4 5

next > last >>

20 ▾

1 of 14 Pgs (270 Rows)

Reload

Option

Filter(s): Age = 10

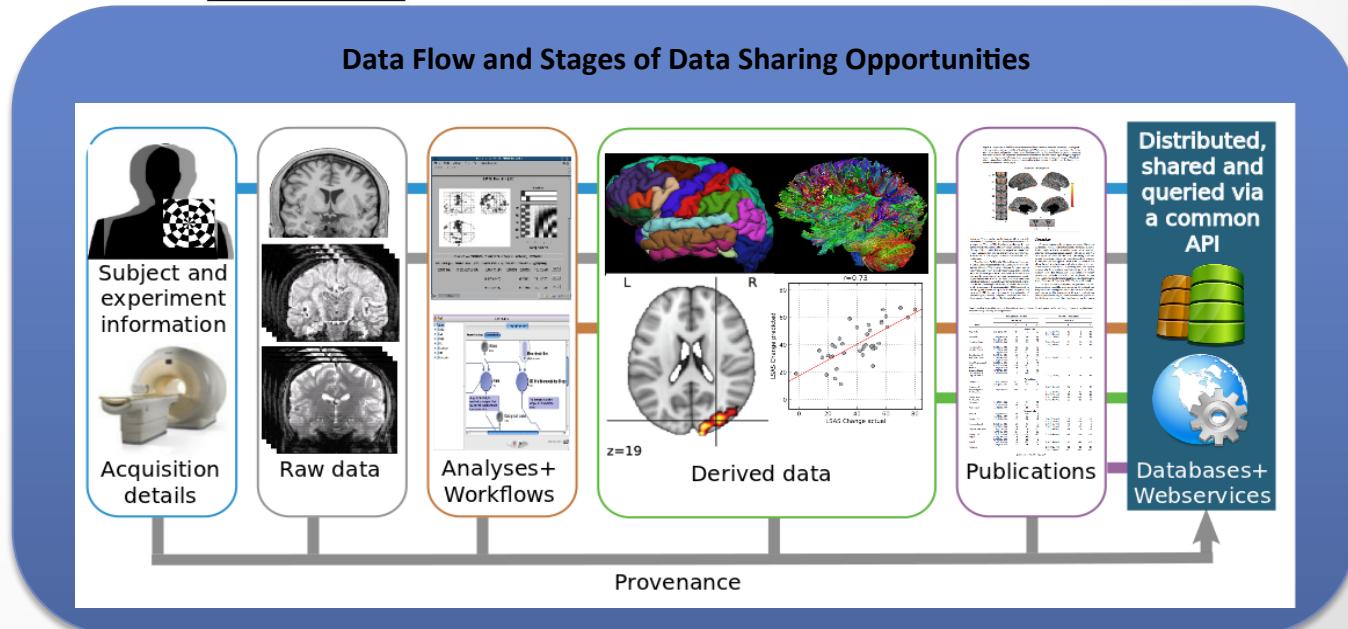
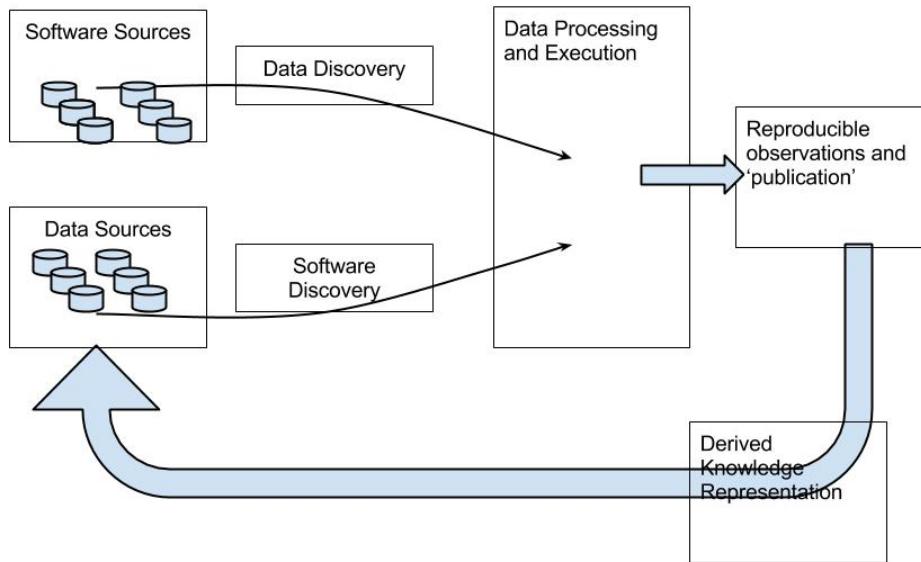
Label	Project	Subject	M/F	Age	Scans	Field Strength	Resting TR
BPDwoPsy_034_MR	cs_schizbull08	BPDwoPsy_034	M	10	anat(1)	1.5T	
BPDwoPsy_049_MR	cs_schizbull08	BPDwoPsy_049	M	10	anat(1)	1.5T	
BPDwoPsy_053_MR	cs_schizbull08	BPDwoPsy_053	F	10	anat(1)	1.5T	
BPDwoPsy_055_MR	cs_schizbull08	BPDwoPsy_055	M	10	anat(1)	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
KKI_1842810_1	adhd_200	KKI_1842810	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**

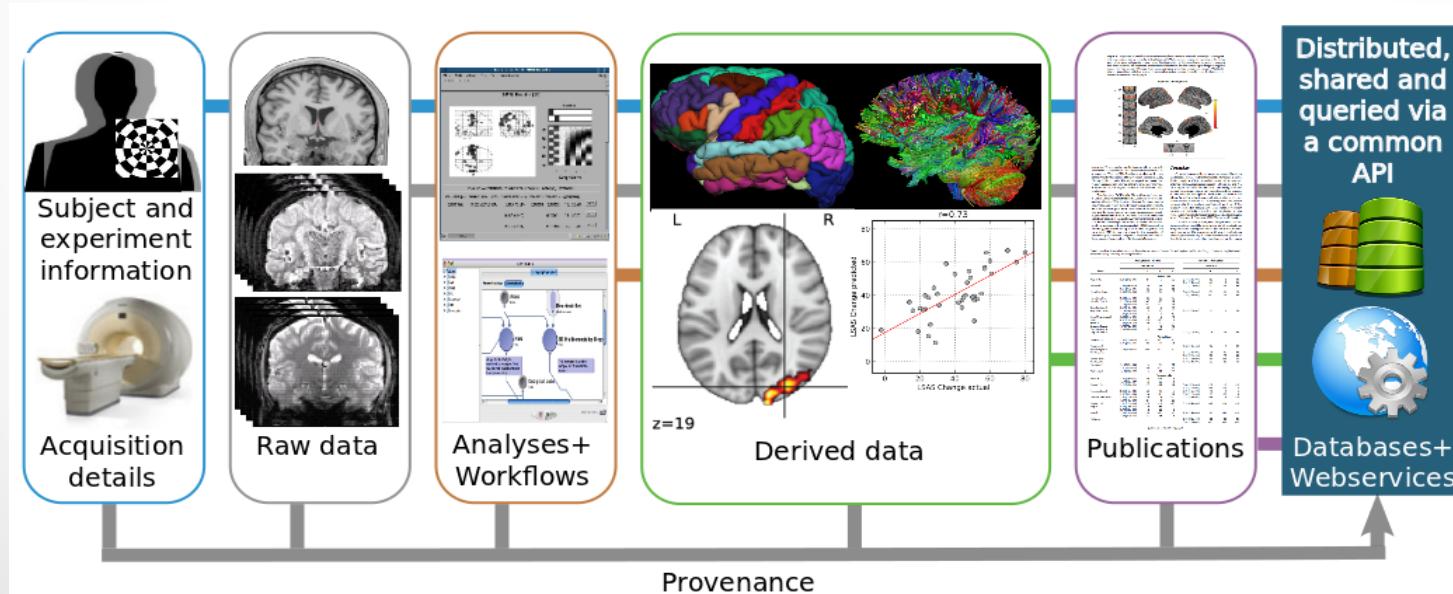


The Research Framework



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.



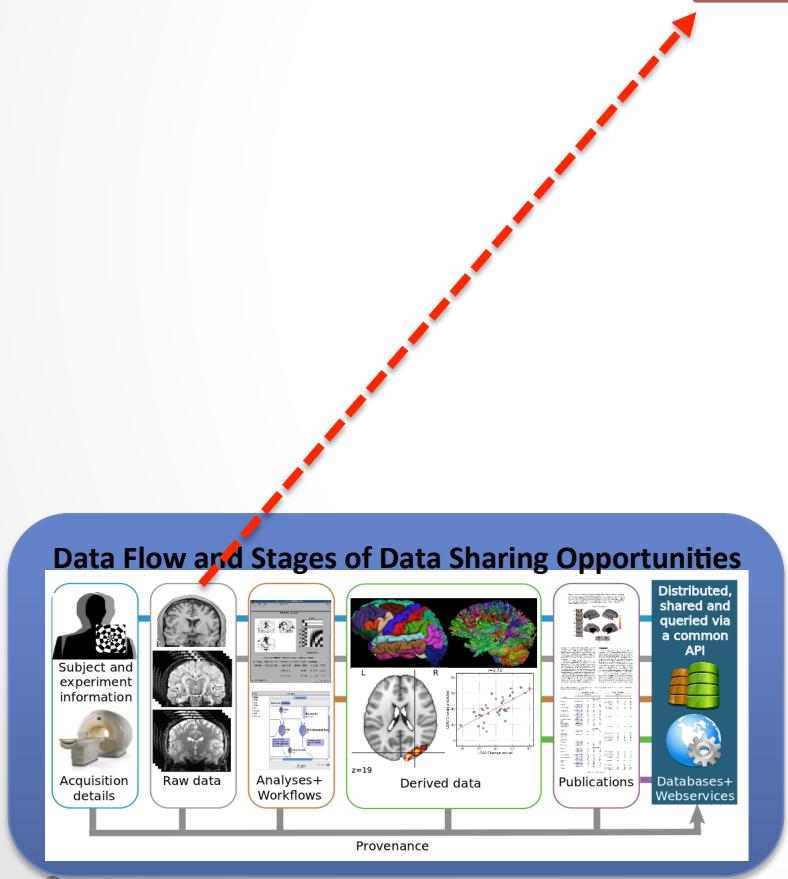
- All data will be marked up via NIDM (<http://nidm.nidash.org/>) and provided via a NIH-funded NITRC (Neuroimaging Informatics Tools and Resources Clearinghouse - nitrc.org) project. NIDM is the NeuroImage Data Model, an instance of the W3C PROV specification.

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

