The National Consortium on Alcohol and Neurodevelopment in Adolescence (NCANDA): A Framework Supporting Neuroimaging Data Integration and Analysis

Nolan Nichols, Ph.D.[1,2]
Weiwei Chu, M.S.[1]
Kilian Pohl, Ph.D.[1,2]

[1]Center for Health Sciences, SRI International

[2]Department of Psychiatry & Behavioral Sciences, Stanford

The Challenge

National Consortium

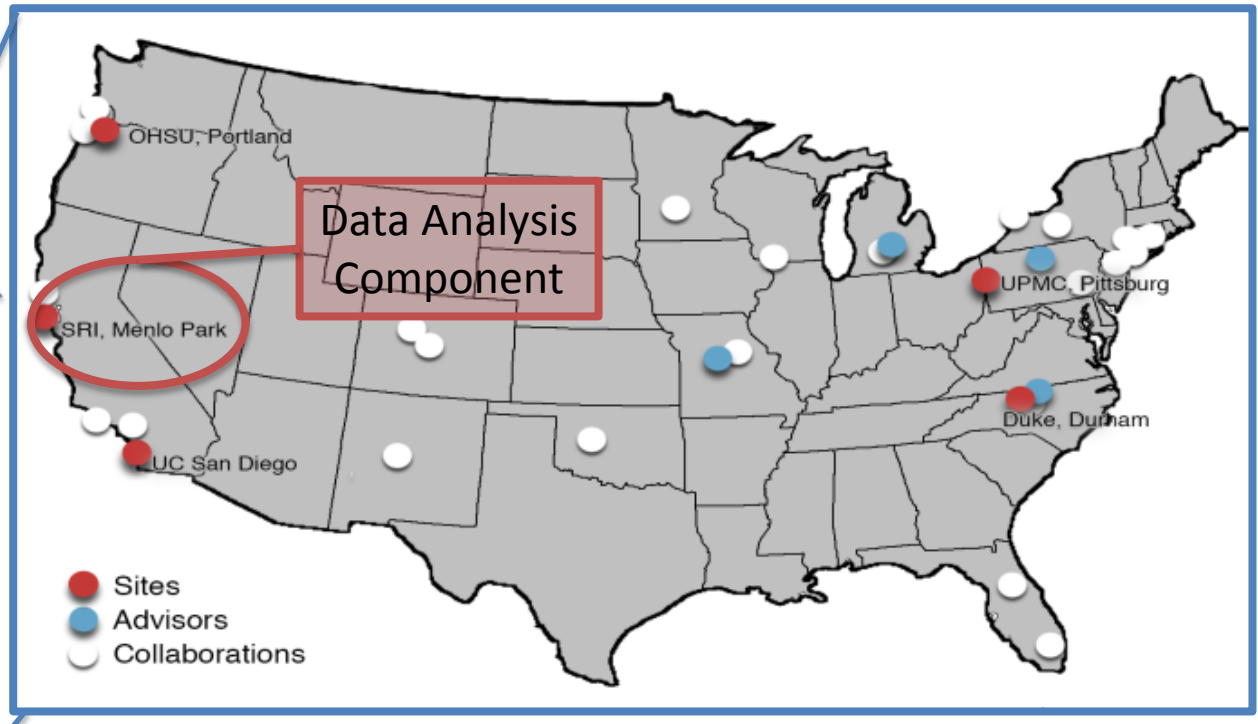
Traditional Research Lab

Scaling up

Multi-Site

Longitudinal

Multi-Modal

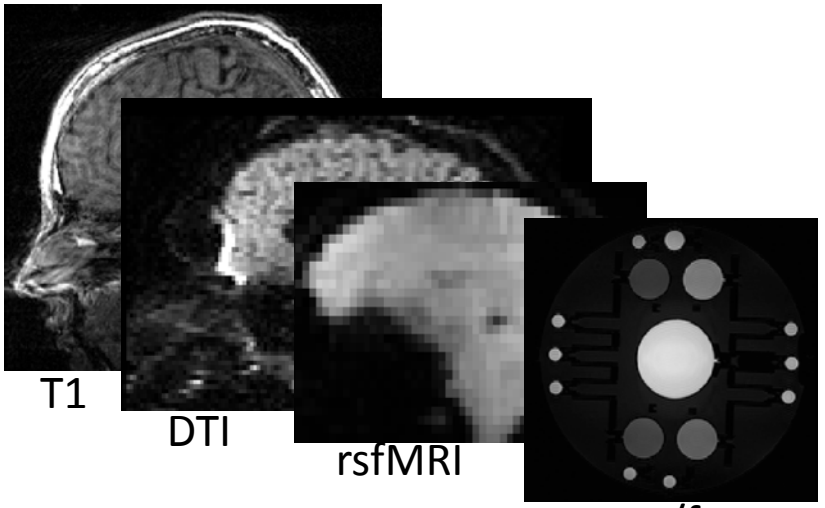


- Five sites across the US
- Total of 808 enrolled participants
- Baseline, 1 year, and 2 year follow-ups
- Data collection initiated in 6 months
- Limited resources for development
- What are the system requirements?

Requirements

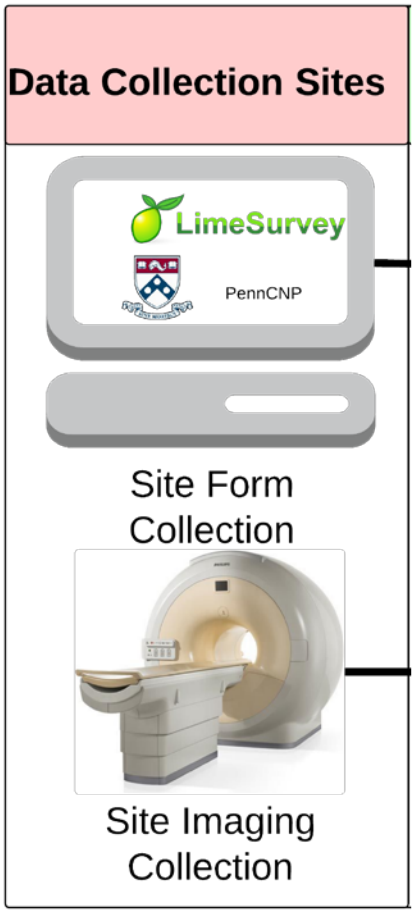
Clinical and Neuropsychological Instruments

Multimodal Imaging and Phantoms



- Accommodate heterogeneous instruments
- Validate data capture protocols
- Maintain ongoing data quality
- Ensure longitudinal visit time windows
- Automate as much as possible
- What Neuroinformatics resources to reuse, circa 2012?

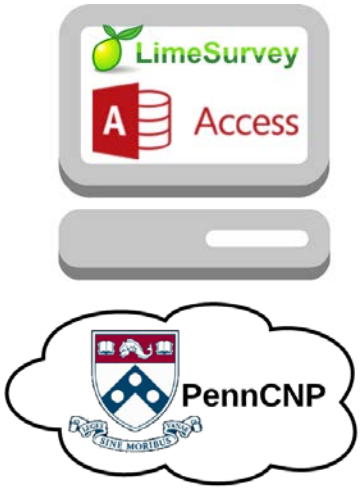
Overview and Approach



Sites collect:

- Demographic Information
- Clinical Data
- Neuropsychological Test Scores
- MRI
  - Anatomical
  - Diffusion
  - Functional

Clinical and Neuropsych Assessments



Extract

Transform

Load

REDCap Harmonization

Record ID	Visit Information	Delayed Discounting, $1000	Delayed Discounting, $100	Paced Auditory Serial Addition Test (PASAT)
[redacted] 2013-03-13				
[redacted] 2013-02-07				
[redacted] 2013-02-26				
[redacted] 2013-02-18				
[redacted] 2013-02-21				
[redacted] 2013-02-22				
[redacted] 2013-04-01				
[redacted] 2013-03-15				
[redacted] 2013-03-18				
[redacted] 2013-07-02				

Cross-sectional REDCap Project

Subject ID	Basic Demographics Baseline visit (Arm 1: Standard Protocol)	Visit Date and Notes Baseline visit (Arm 1: Standard Protocol)	Visit Date and Notes 6-month follow-up (Arm 1: Standard Protocol)	Visit Date and Notes 1y visit (Arm 1: Standard Protocol)	Delayed Discounting, $1000 Baseline visit (Arm 1: Standard Protocol)	Delayed Discounting, $1000 1y visit (Arm 1: Standard Protocol)	Delayed Discounting, $100 Baseline visit (Arm 1: Standard Protocol)	Delayed Discounting, $100 1y visit (Arm 1: Standard Protocol)
[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								

Longitudinal REDCap Project

Visit Classification & Scoring

Import to Central Repository

Multi-Modal Imaging

Upload

XNAT

Protocol

Export

REDCap™

Longitudinal Project

Sites Scans
T1, T2, DTI, rsfMRI,
ADNI/fBIRN Phantoms

Reading

DateToDVD	2013-07-11
FindingsDate	2013-08-13
Findings	normal

QA Flags (TO BE SET ONLY BY QA STAFF)

Scans

Scan	Type	Series Desc	Usability
1	ncanda-localizer-v1	ncanda-localizer-v1	unknown
2	ncanda-calibration-v1	ncanda-calibration-v1	unknown
3	ncanda-t1spgr-v1	ncanda-t1spgr-v1	usable
4	ncanda-t2fse-v1	ncanda-t2fse-v1	usable
6	ncanda-dti6b500pepolar-v1	ncanda-dtib500pepolar-v1	usable
7	ncanda-dti60b1000-v1	ncanda-dtib1000-v1	usable
8	ncanda-grefieldmap-v1	ncanda-grefieldmap-v1	usable
9	ncanda-rsfmri-v1	ncanda-rsfmri-v1	usable

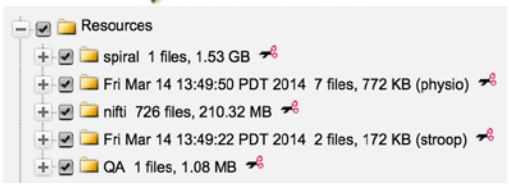
T1w MRI Scan Date	2015-05-28 Today Y-M-D	Date of T1w image acquisiton
Age at T1w MRI Scan * must provide value	21.9416167911	Age at T1w image acquisition
DTI Scan Date	2015-05-28 Today Y-M-D	Date of DTI image acquisiton
Age at DTI MRI Scan * must provide value	21.9416167911	Age at DTI image acquisition
Rs-fMRI Scan Date	2015-05-28 Today Y-M-D	Date of rs-fMRI image acquisiton
Age at rs-fMRI Scan * must provide value	21.9416167911	Age at rs-fMRI image acquisition

- Resources
  - spiral 1 files, 1.53 GB
  - Fri Mar 14 13:49:50 PDT 2014 7 files, 772 KB (physio)
  - nifti 726 files, 210.32 MB
  - Fri Mar 14 13:49:22 PDT 2014 2 files, 172 KB (stroop)
  - QA 1 files, 1.08 MB

NIfTI
Conversion

- Semi-automatic QA
- Neuroradiologist readings
- Hourly/nightly QA reports
- Event-based workflow to populate image processing pipeline

Data Analysis



Import to Pipeline

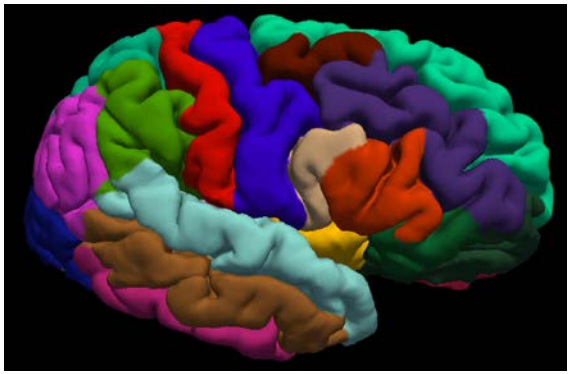
Light Weight Data Pipeline

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

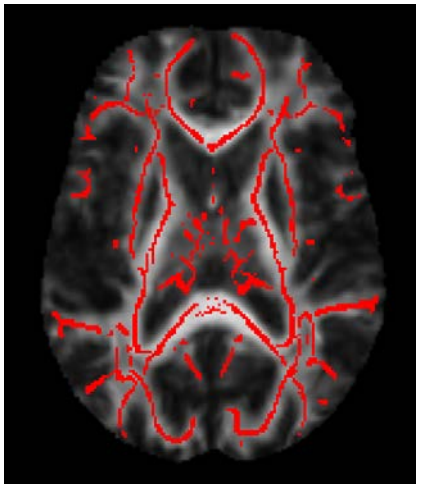
Anatomical

Diffusion

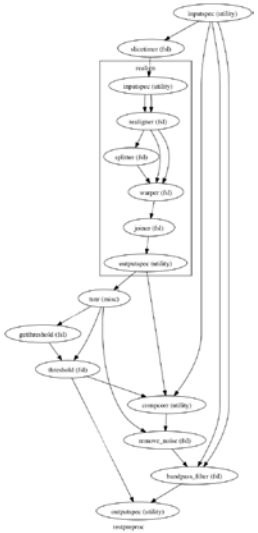
Resting State



FreeSurfer Parcellation and Segmentation

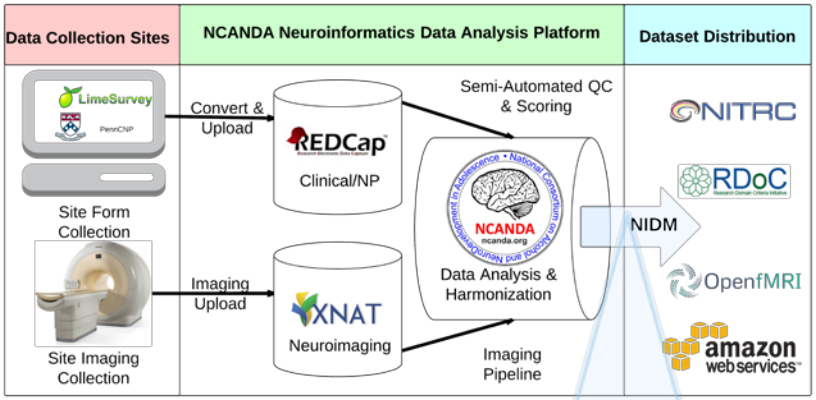


FSL Tract-Based Spatial Statistics

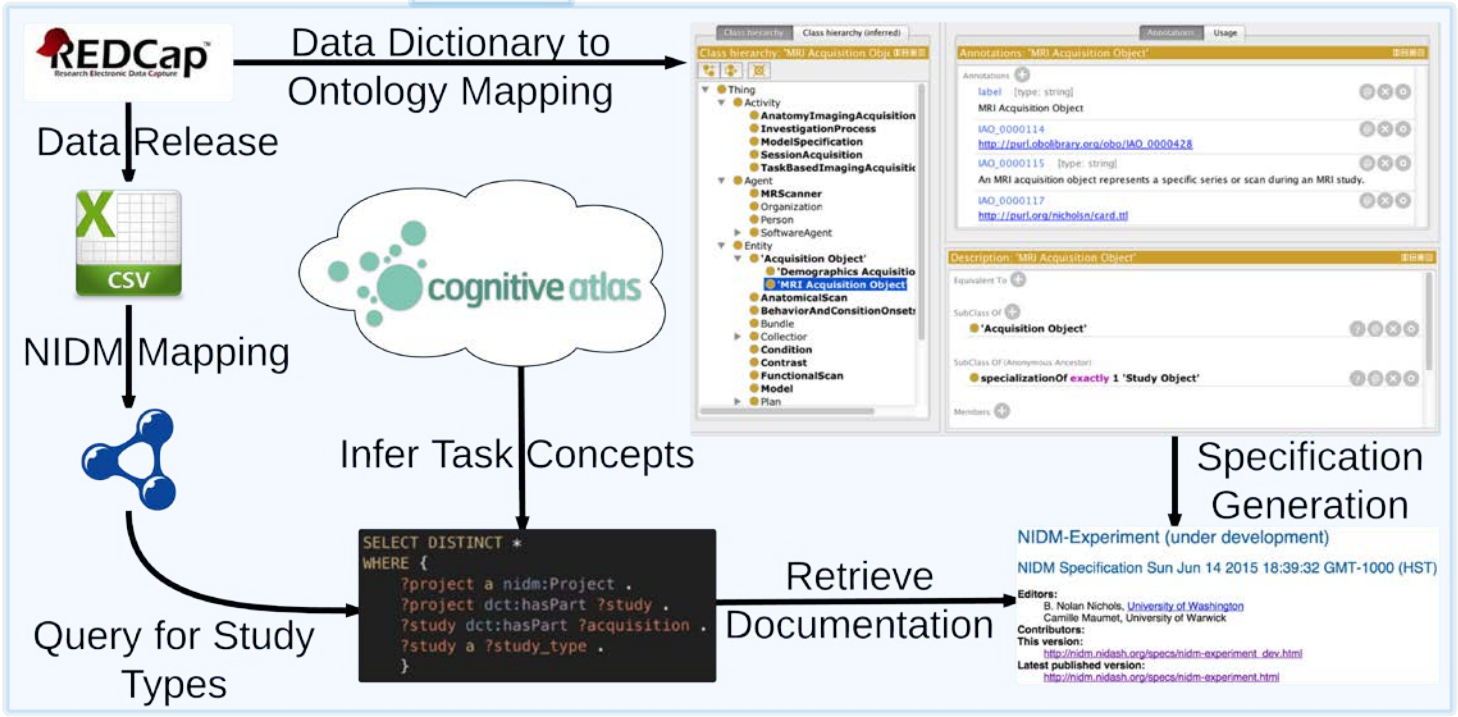


Nipype-based Preprocessing

Next Steps and Data Sharing

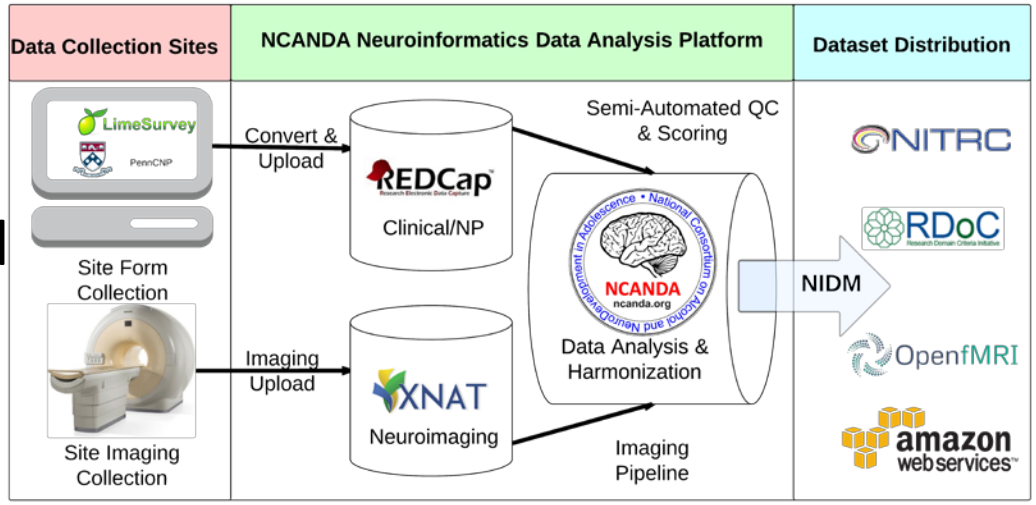


- Curated with REDCap data dictionaries
- Use Case for NIDM and BIDS Standards
- Adopt the NIDM process to curate data
- Develop object models of clinical, neuropsych, and imaging measures
- Demonstrate usage of NIDM for distributing NCANDA datasets



Conclusions and Lessons Learned

- The ecosystem of Neuroinformatics software for imaging studies is mature
- Neuroinformatics tools can be reused and extended to develop scalable Neuroinformatics platforms
- Early involvement of informaticians may be able to simplify system architecture using common platforms
- Version Control Systems are an innovative way to capture data asynchronously before a data management system implementation
- Demonstrated the reuse of neuroinformatics tools to provide data integration and analysis hub for a multi-site, longitudinal study on adolescent development



https://www.nitrc.org/projects/ncanda-datacore

Acknowledgements

National Institute on Alcohol Abuse and Alcoholism (NIAAA)
Data Analysis Component 5U01AA021697

Data Analysis Component

Weiwei Chu
Kevin Cummins
Dongjin Kwon
Dolf Pfefferbaum
Kilian Pohl
*Torsten Rohlfing
Yong Zhang

Data Sharing Task Force

Satra Ghosh
Chris Gorgolewski
David Keator
JB Poline
Jessica Turner
… and many others!

Neuropsychology Protocol

8 RFA-required functional domains; valid across wide age range

Executive Function & Attention
Penn Continuous Performance Task
LeJuez Distress Tolerance: PASAT
SRI Stroop Match-to-Sample
Penn Conditional Exclusion Task
Penn Fractal N-back Task

Emotion Processing/Regulation
Penn Emotion Recognition Task
Penn Emotion Differentiation Test

Handedness & Dexterity
Edinburgh Handedness Inventory
Grooved Pegboard
Penn Motor Praxis Task

Intelligence
Penn Logical Reasoning test
Penn Matrix Analysis Test
WRAT-4 Vocabulary

Memory: Immediate-Delayed
Penn Visual Object Learning
Penn Word Memory
Penn Facial Memory

Reward Seeking & Learning
Stanger Delay Discounting Task

Visual Discrimination
Landolt C - Acuity
Ishihara – Color

Achievement
Penn Vocabulary Test
WRAT-4 Arithmetic

Classic Tests
Rey-Osterrieth Complex Figure
Ataxia – Walk-a-Line
Digit Symbol Substitution

Neuropsychology Test Battery

25 tests → ~250 primary variables from ~1000 measures

Test		
Ishihara Test	May be completed during clinical assessment.	20 min.
Landolt C		
Edinburgh Handedness Questionnaire		
WRAT- 4 Reading		5 min.
Penn WebCNP		
Motor Praxis Test	Administered on the MacBook Air. Must be completed in one session.	WebCNP 60 min.
Penn Facial Memory Test		
Penn Word Memory Test		
Penn Continuous Performance Test-Number Letter Version		
Short Fractal N- Back Test - 2 Back Version		
Penn Matrix Analysis Test		
Penn Facial Memory Test - Delayed		
Penn Word Memory Test - Delayed		
Break		
Penn Short Visual Object Learning Test	Administered on the MacBook Air. Must be completed in one session.	
Emotion Recognition Test		
Penn Conditional Exclusion Task		
Measured Emotion Differentiation		
Penn Vocabulary Test		
Penn Logical Reasoning		
Short Visual Object Learning Test - Delayed		
Stroop Match to Sample	Administered on the Dell laptop.	60 min.
Stanger Delay Discounting Task		
PASAT-C Lejuez Distress Tolerance		
Rey-O Copy	Must be completed in one session.	
Rey-O Immediate		
WRAT- 4 Arithmetic	Must be completed in ~30 minutes.	
Ataxia		
Grooved Pegboard		
Rey-O Delayed		
WAIS-IV Coding [Symbols can be distractors to the Rey-O]		

Baseline NP battery = 180 min.
Follow-up = 150 min.

References

1. L. D. Johnston, P. M. OMalley, R. A. Miech, J. G. Bachman, and J. E. Schulenberg, “Monitoring the Future national survey results on drug use: 1975-2014: Overview, key findings on adolescent drug use.” 2015.
2. T. Rohlfing, K. Cummins, T. Henthorn, W. Chu, and B. N. Nichols, “N-CANDA data integration: anatomy of an asynchronous infrastructure for multi-site, multi-instrument longitudinal data capture,” Journal of the American Medical Informatics Association, pp. amiajnl–2013–002367, 2013.
3. WebCNP: https://webcnp.med.upenn.edu/
4. LimeSurvey: http://www.limesurvey.org/
5. Blaise: http://www.blaise.com
6. ePrime: http://www.pstnet.com/eprime.cfm
7. Marcus, D., Olsen, T., Ramaratnam, M., & Buckner, R. (2007). The extensible neuroimaging archive toolkit. *Neuroinformatics*, *5*(1), 11–33.
8. Subversion: https://subversion.apache.org/
9. Keator, D. B., Helmer, K., Steffener, J., Turner, J. A., Van Erp, T. G., Gadde, S., et al. (2013). Towards structured sharing of raw and derived neuroimaging data across existing resources. *NeuroImage*, *82*, 647–661. http://doi.org/10.1016/j.neuroimage.2013.05.094