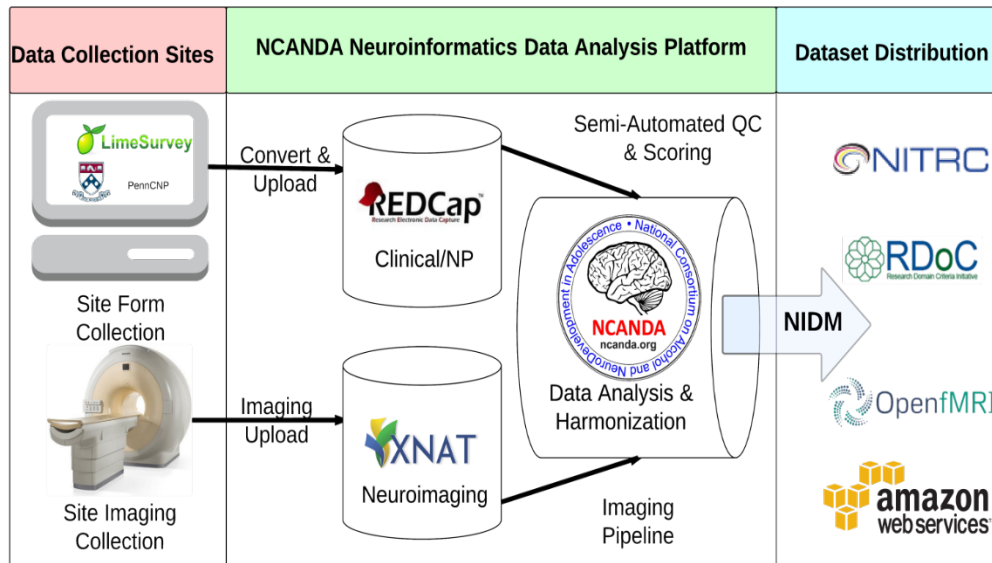# Curating, Releasing, and Access NCANDA Data

Kilian M. Pohl
Adolf Pfefferbaum

Ehsan Adeli
Dongjin Kwon
Michael Hasak
Qingyu Zhao
Sara Benito
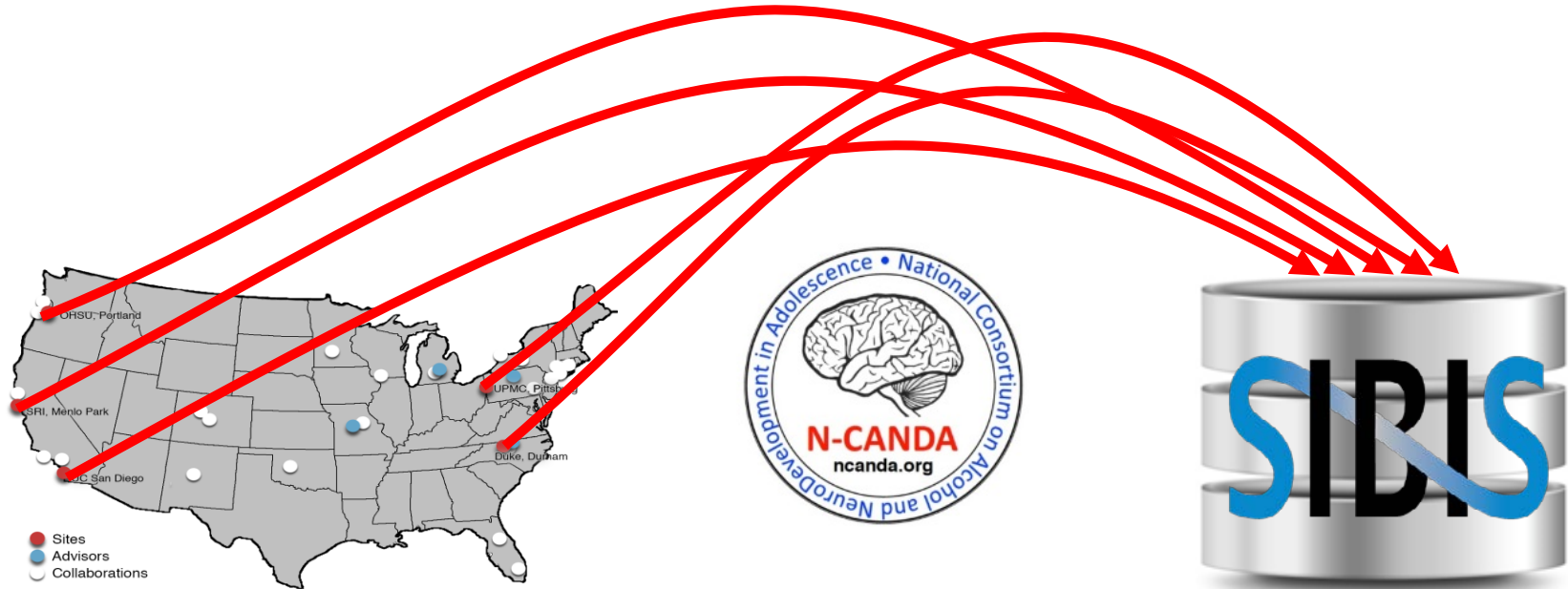Weiwei Chu
SRI International

Edith V. Sullivan
Stanford University

kilian.pohl@sri.com ♦ https://www.stanford.edu/~kpohl

Financial Interest Disclosure

Salary and Research Support
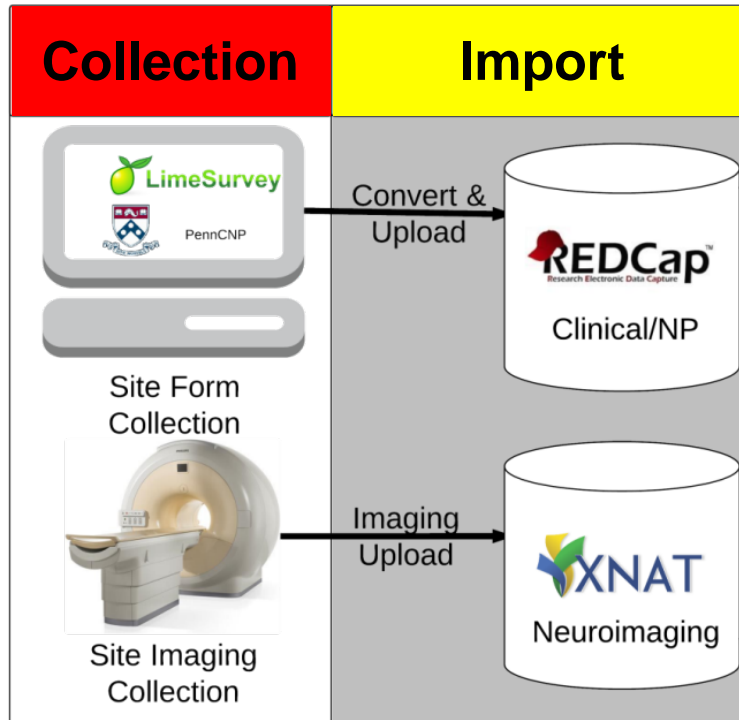of
Kilian M Pohl, Ph.D.

Scalable Informatics for Biomedical Imaging Studies

- NCANDA collection sites upload data acquired in 831 study participants to the SIBIS portal
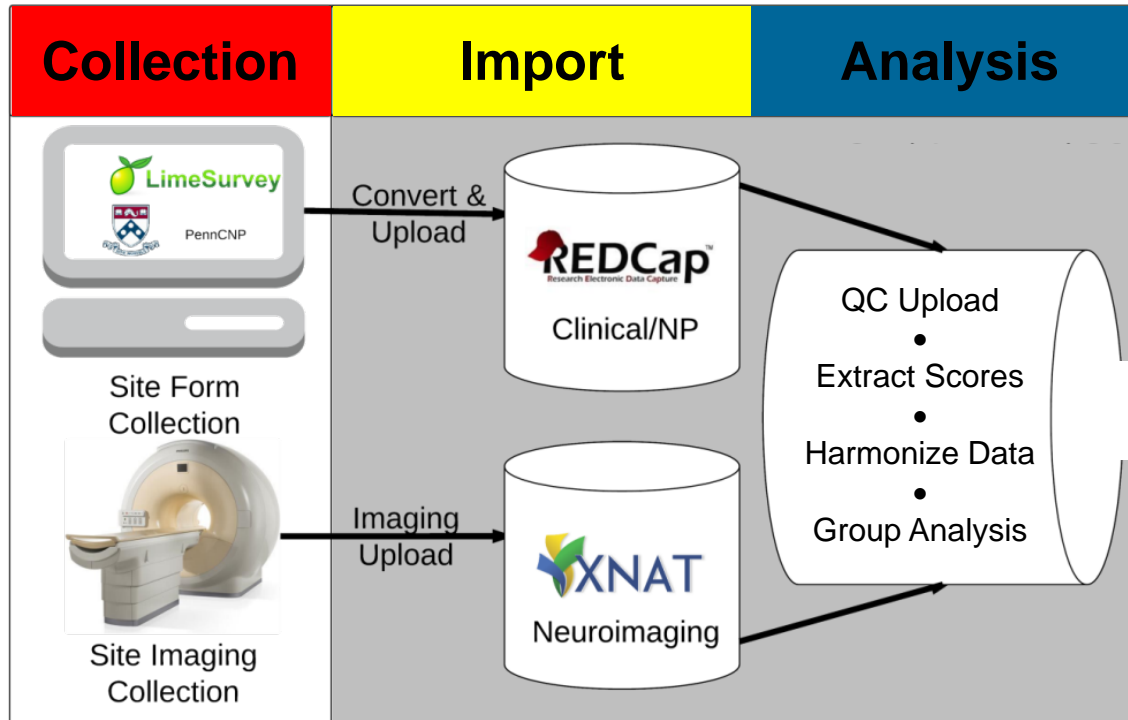- SIBIS combines, curates, analyzes, and distributes the data using open source software.

Scalable Informatics for Biomedical Imaging Studies
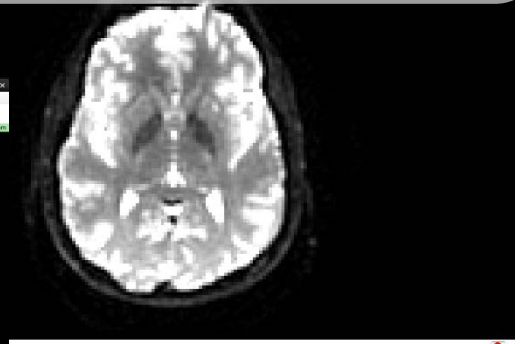

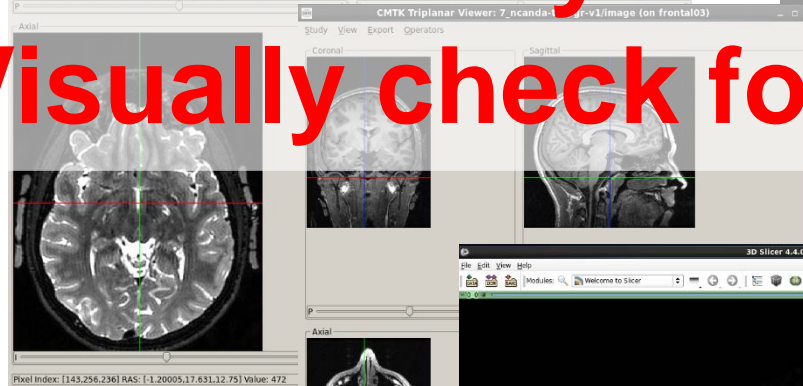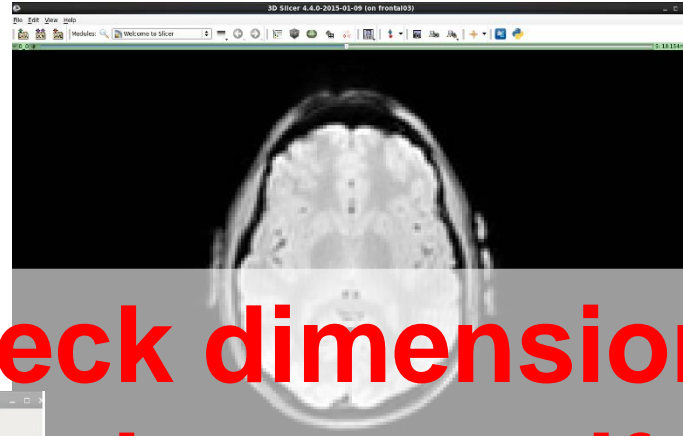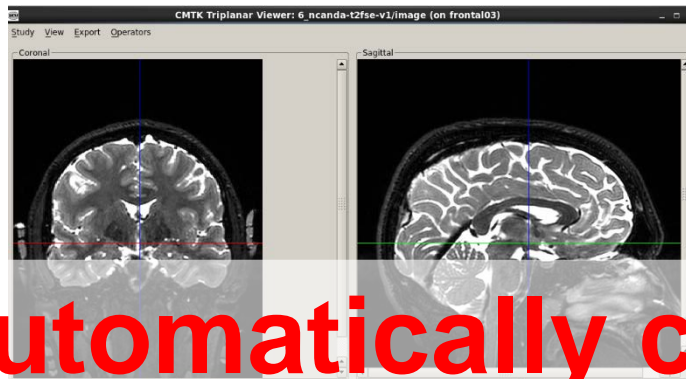**Sites collect:**
- Non Imaging Data
  - Demographic Information
  - Clinical Data
  - Neuropsychological Test Scores
- MRI
  - Structural
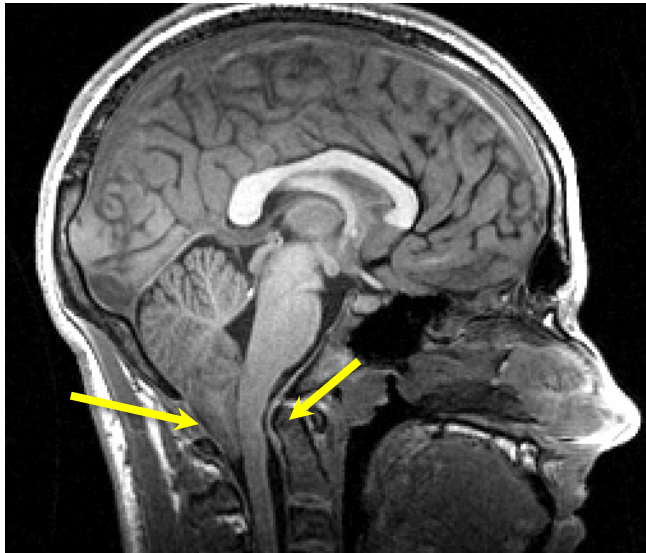 - Diffusion
  - Functional

Data Analysis

Semi-Automatic Quality Control

Automatically check dimension, ….
Visually check for image artifacts
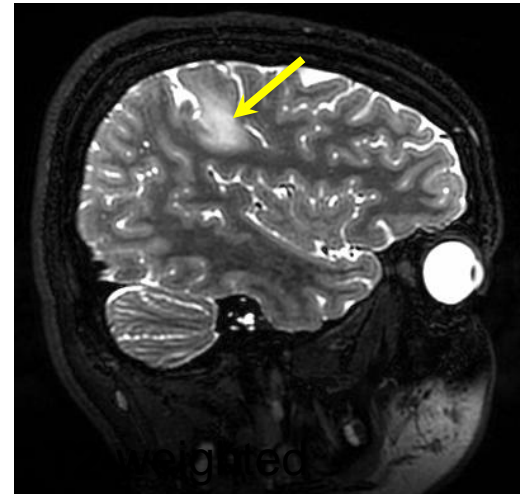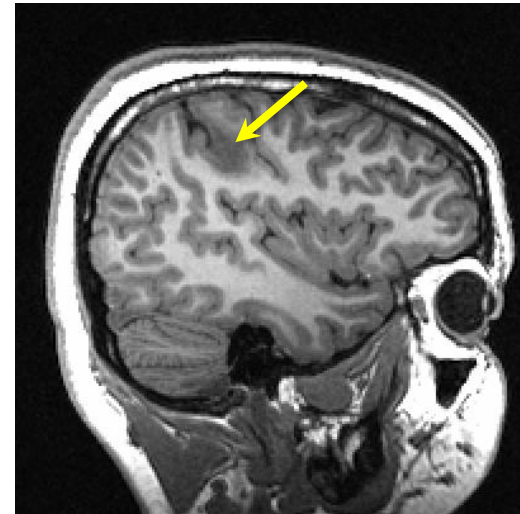
Neuroradiologist Reading

Chiari 1 malformation

right parietal cortical mass

bilateral tonsillar herniation
with medullary distortion

- 95/833 adolescents = 11.4%
- 2 excluded from the study

Issues Reported To Github

Issues are resolved with collection sites

Process Data

XNAT

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
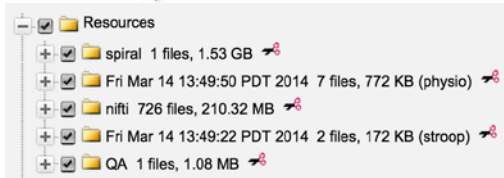
Import to

Pipeline

Light Weight Data Pipeline

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

Resting State

Diffusion
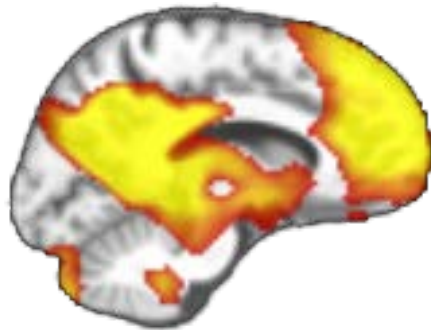
Anatomical

Nipype-based Preprocessing
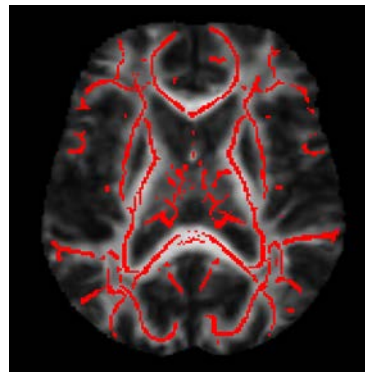
FSL Tract-Based Spatial Statistics

Volume Scores

Kilian M. Pohl
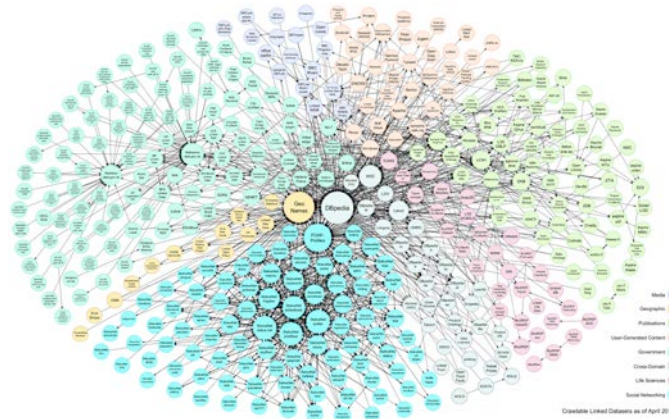
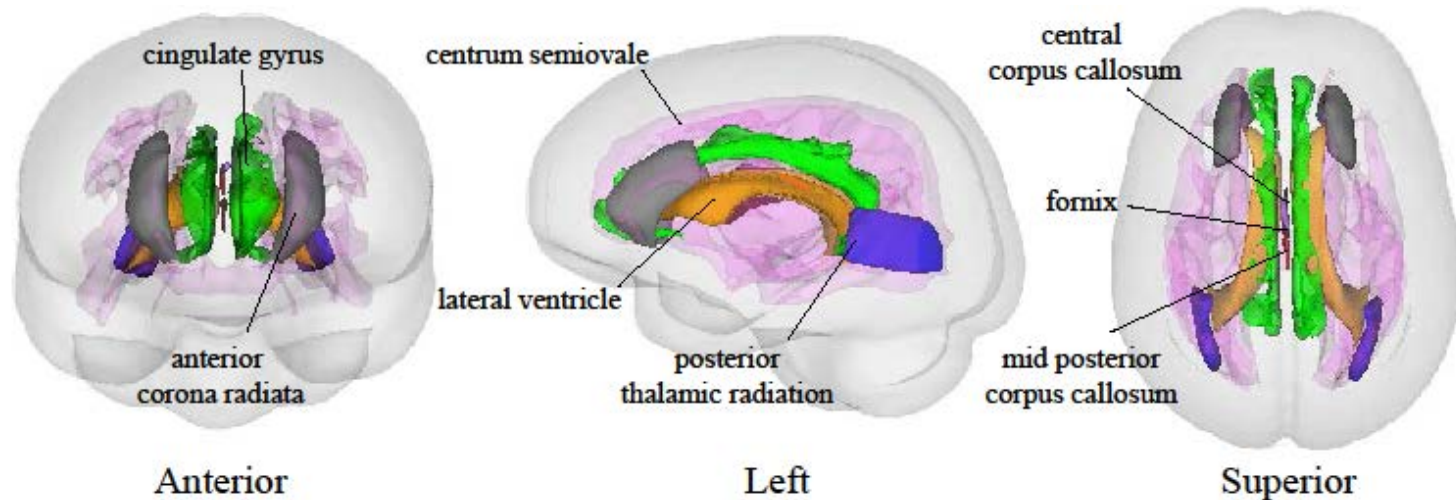- 9 -
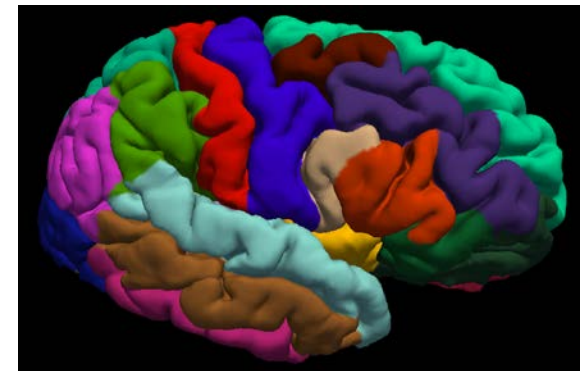
Create Data Releases

- Contains text files with de-identified demographic and neuropsychological test scores, raw and derived imaging data, and composite scores

- Data dictionary describing variables
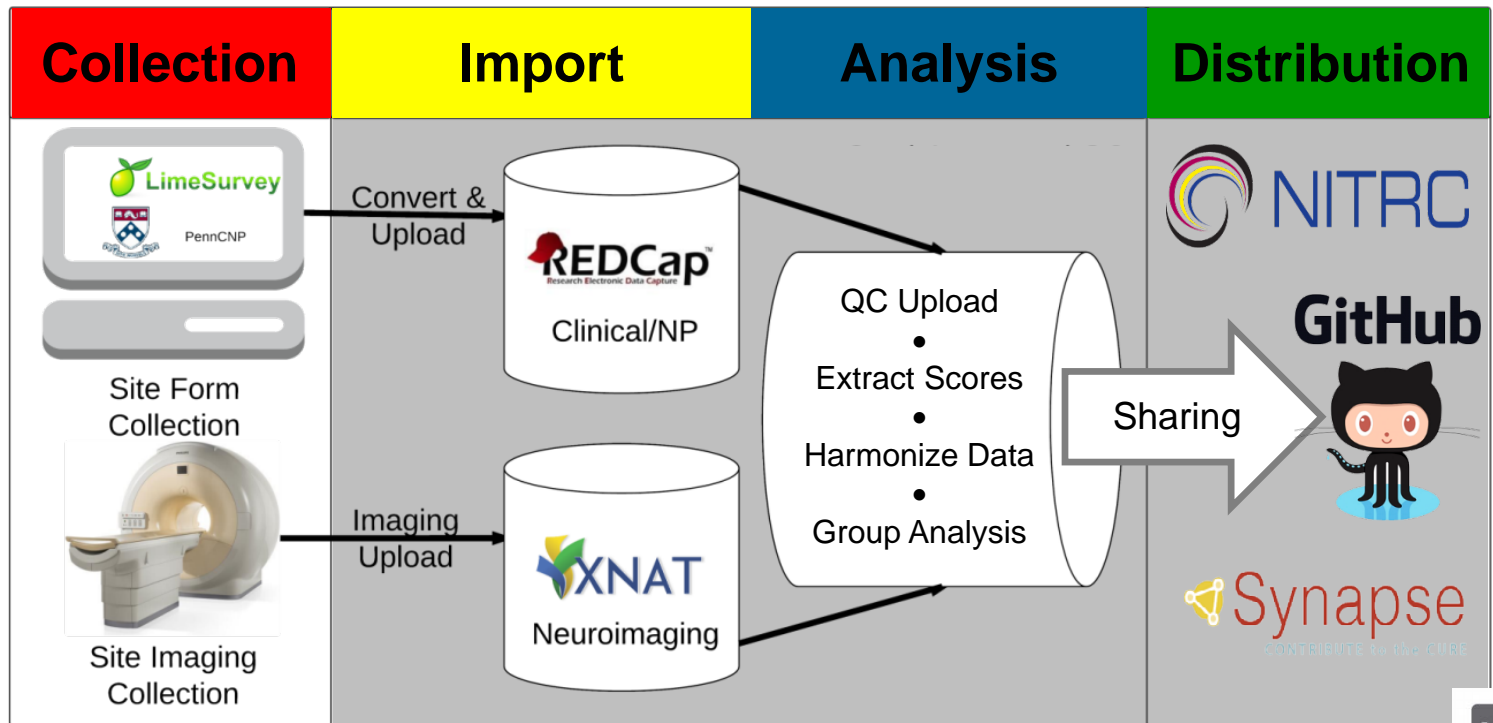
- Software used to generate the data

Create machine learning technology to identify brain regions impacted by regular alcohol consumption during adolescence

Share Data and Software

Upload results to the data repository Synapse
Store software via
- NITRC: https://www.nitrc.org/projects/ncanda-datacore
- Github: https://github.com/sibis-platform

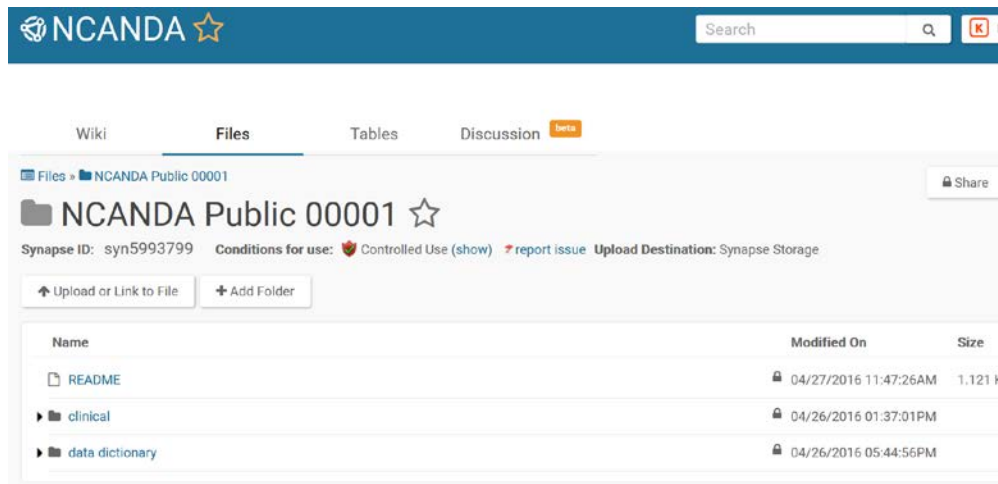# Gaining Access to Data

Apply to NIAAA:

- A **cover letter** on the letterhead of the sponsoring institution at which the research study will be conducted.

- **Curriculum vitae** of the principal investigator and all co-investigators

- A 1-2 page **description of the proposed research**

- A completed **Data Distribution Agreement** with signatures of the principal investigator and an authorized representative of the sponsoring institution.

Information available at https://www.niaaa.nih.gov/research/major-initiatives/national-consortium-alcohol-and-neurodevelopment-adolescence

Gaining Access to Data

After approval from NIAAA

- SRI will request from PI their Synapse user name
- PI will be able to access data via https://www.synapse.org/ncanda

Synapse allows users to query and download data releases, which are linked to publications.

Data Sample
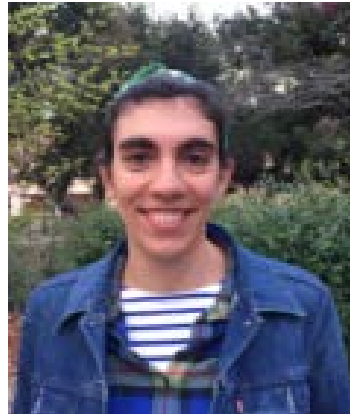

Data Dictionary

Score

Thank You

Dongjin Kwon,PhD

Ehsan Adeli,PhD

Qingyu Zhao,PhD

Michael Hasak

Sara Benito

Weiwei Chu

Kilian M. Pohl