SchizConnect: Flexible, Dynamic Platform for Mediating Multiple Schizophrenia Neuroimaging Databases

Lei Wang[1], Kathryn I. Alpert[1], Jessica Berger[1], Vince Calhoun[2,3], David Keator[4], Margaret King[2], Alex Kogan[1], Drew Landis[2], Steven G. Potkin[4], Jessica A. Turner[5,2], Jose Luis Ambite[6]

[1]Northwestern University Feinberg School of Medicine, [2]The Mind Research Network, [3]The University of New Mexico, [4]University of California, Irvine, [5]Georgia State University, [6]University of Southern California

Background

- Schizophrenia is highly heterogeneous
  - Clinical, behavioral, cognitive, neurobiological & genetic variability
  - Neuroimaging studies require increasingly large datasets
- Data sharing for integrative analysis
  - Aggregation into larger datasets through repository and/or federation
  - Access individual level data & identify specific subsets
  - Replication & new discovery
- Data Sharing Challenges
  - Different interfaces & download mechanisms
 - Different data organization & variable ontology

SchizConnect

- Federation through mediation (virtual data integration)
  - Data resides at each source, unchanged
  - Data integration by mediation software (mediator)
- SchizConnect (www.schizconnect.org) links consortia data to form compatible mega-datasets
- SchizConnect has 3 components
  - Data sources: individual databases with varying platforms, schemas, and interfaces
  - Mediator: data integration engine, translating compatible data terms across the sources to a common model\
  - Web portal: a user-friendly interface for query and download

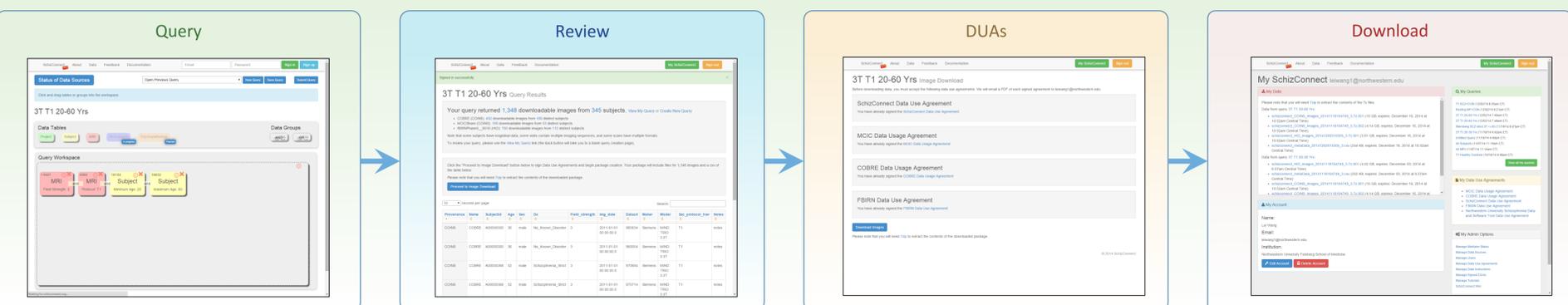## SchizConnect Domain Models
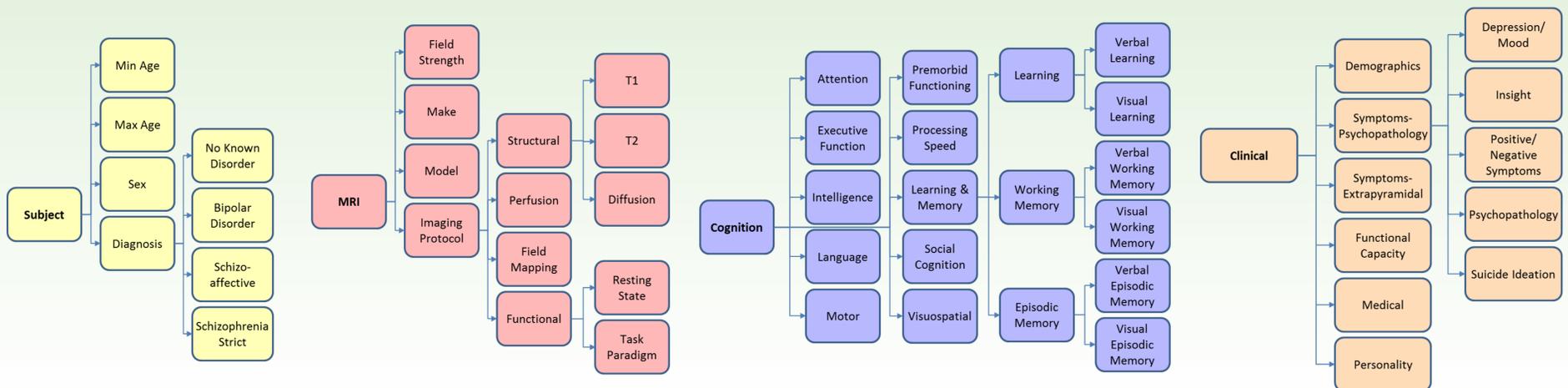
SchizConnect Architecture
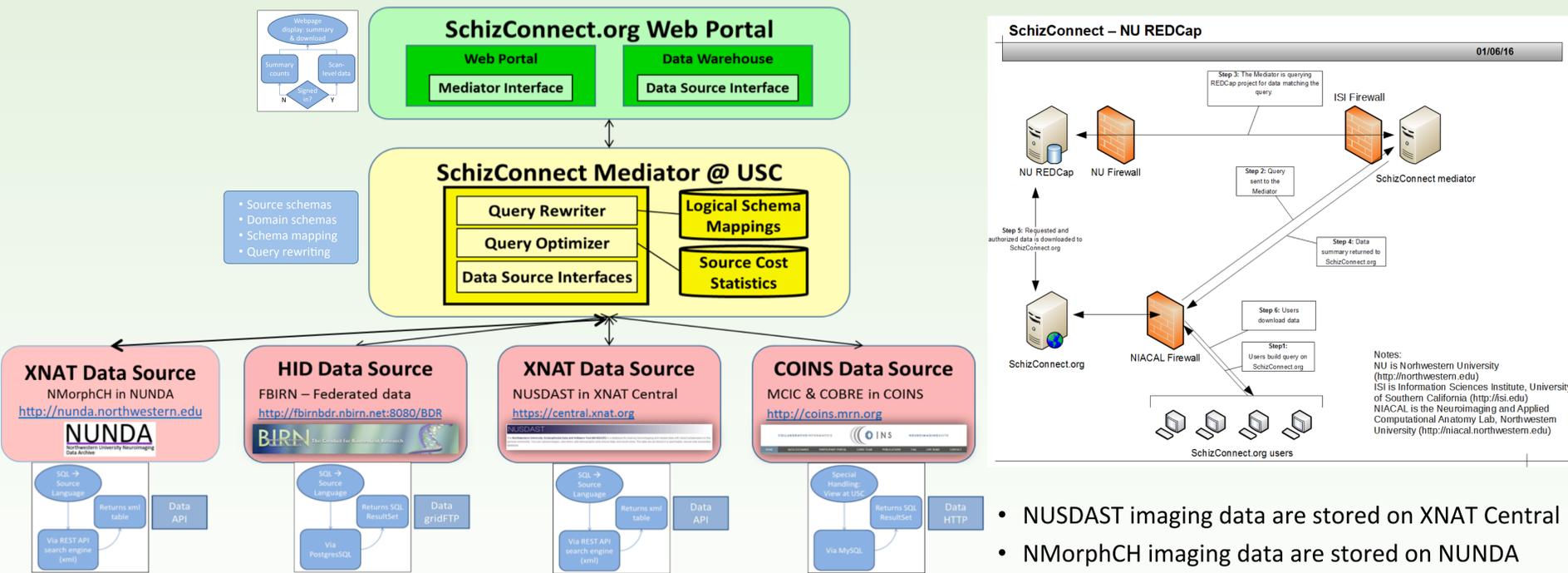
- NUSDAST imaging data are stored on XNAT Central
- NMorphCH imaging data are stored on NUNDA
- Non-imaging data for these projects are stored in a protected REDCap clinical database located behind a firewall

SchizConnect.org Web Portal

Query → Review → DUAs → Download

Available Subjects and Data

	FBIRN	NU	MRN	NU
Database	HID	XNAT Central	COINS	NUNDA
Number of Subjects	236 (Phase II)	451	423 (213 MCIC/210 COBRE)	89

To Be Released

FBIRN	MRN	WUSTL
HID	COINS	NUNDA
400 (Phase 3)	200 (BrainGluSchi)	135 (Conte)

BIRN Biomedical Informatics Research Network
The Mind RESEARCH NETWORK FOR NEURODIAGNOSTIC DISCOVERY
USC University of Southern California
NORTHWESTERN UNIVERSITY FEINBERG SCHOOL OF MEDICINE