

SchizConnect: Large-Scale Schizophrenia Neuroimaging Data Integration and Sharing

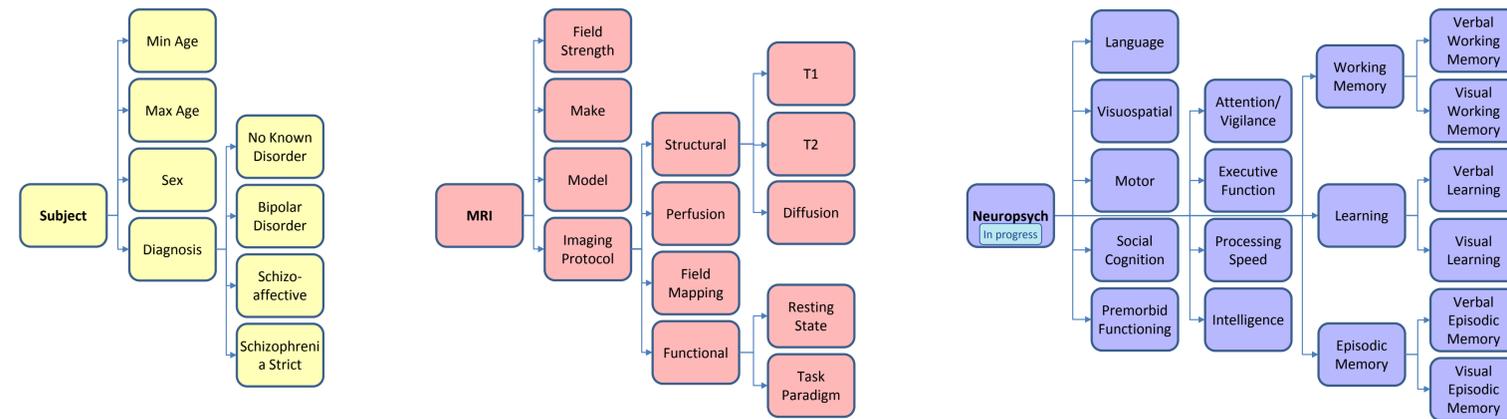
BETA

Lei Wang¹, Kathryn I. Alpert¹, Vince Calhoun^{2,3}, David Keator⁴, Margaret King², Alex Kogan¹, Drew Landis², Marcelo Tallis⁵, Steven G. Potkin⁴, Jessica A. Turner^{6,2}, Jose Luis Ambite⁵
¹Northwestern University Feinberg School of Medicine, ²The Mind Research Network, ³The University of New Mexico, ⁴University of California, Irvine, ⁵University of Southern California, ⁶Georgia State University
 Support: 1U01 MH097435, 1R01 MH084803, P50 MH071616, R01 MH056584, U24 RR025736-01, U24-RR021992, U24GM10420, P20 GM103472

Background

- Schizophrenia is highly heterogeneous
 - Clinical, behavioral, cognitive, neurobiological & genetic variability
- Data sharing for integrative analysis
 - Aggregation into larger datasets through repository and/or federation
 - Access individual level data & identify specific subsets
 - Replication & new discovery

SchizConnect Schema & Domain Models



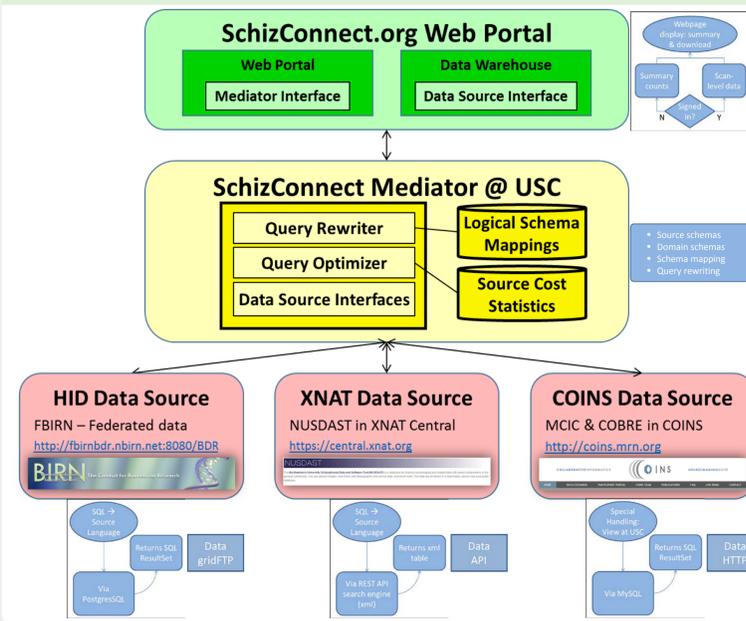
Summary

- One-click Interface
- Without having to register, the user can obtain a total count on the number of available datasets
- User registration and compliance with data use agreement with each site is facilitated
- User can download specific datasets without having to go to each data site
- Queries each data source in parallel
 - HID@UCI, NUSDAST@XNATCentral, COINS@MRN

Data Sharing Challenges & SchizConnect Solution

- Different interfaces & download mechanisms
- Different data organization & variable ontology
- Federation through mediation (virtual data integration)
 - Data resides at each source, unchanged
 - Data integration by mediation software (mediator)

SchizConnect Architecture



Example Queries

Search Terms	Return Number of Images & Distinct Subjects				
	Total	NUSDAST @ XNAT	COBRE @ COINS	MCICShare @ COINS	fbIRN @ HID
T1 SCZ+CON	3,192 738	1,382 284	484 173	264 95	1,054 185
3T T1 20-60 Yrs	1,348 345	-	492 180	106 53	750 112
T1 Healthy Controls	1,846 439	684 137	24 94	264 95	655 113
Sternberg SCZ 3T <=30	210 18	-	-	-	210 18
Resting BP+CON	103 101	-	103 101	-	-
All MRI	21,500 1,030	1,814 368	1,597 198	4,483 212	13,554 251

Available Subjects and Data

Currently Available			
	FBIRN	NU	MRN
Database	HID	XNAT Central	COINS
Number of Subjects	237 (Phase II)	570 (451 with 1.5T)	432 (213 MCIC 219 COBRE)
Image Type	Structural MRI Resting fMRI Task fMRI DTI	Structural MRI	Structural MRI Resting fMRI Task fMRI DTI
To be Released			
Number of New Subjects	400 (Phase 3)	130 (3T)	200 (R01)
Timeline	2016	2015	2017

Next Steps

- Enhanced interface for more complex queries
- Additional variables
 - Neuropsychological domains
 - Symptom domains
- Additional data sources
 - Prodromal, treatment-naive

Conclusions

- SchizConnect allows mediation and combining of neuroimaging data from different databases to form compatible mega-datasets with accuracy and fidelity
- SchizConnect shows potential for overcoming current barriers for creating large-scale datasets to increase statistical power, accelerating the testing of new hypotheses and methods, and creating a resource for developing advanced techniques to better integrate disparate data.

SchizConnect.org Web Portal

