Updated 5/2020

**Description for using NUNDA for multisite neuroimaging data collection and analysis**

A secure web-based portal will be implemented that will allow contributing sites to upload data, quality test neuroimaging data, and test potential image processing pipelines. The archive will be searchable by authorized personnel to identify data from the entire database that match the search criteria.

**Description for Northwestern University Neuroimaging Data Archive – NUNDA** <https://nunda.northwestern.edu/>

**Short version to be included in the grant:**

**Northwestern University Neuroimaging Data Archive (NUNDA)**

We will leverage the existing Northwestern University Neuroimaging Data Archive (NUNDA, co-directed by Drs. Lei Wang and Todd Parrish, [you are encouraged to work with Drs. Wang and Parrish on your grant]). For this project, we will deploy an XNAT-based data repository for capturing imaging data on NUNDA. DICOM files collected on the MR and PET scanners will be transmitted over secure channels to NUNDA’s integrated DICOM C-STORE Service Class Provider (SCP) via SSL and open source dcm4che tools (<http://www.dcm4che.org/>). Using tools available on NUNDA, QA will be performed by validating imaging sequence details against a study specific protocol to ensure acquisition compliance. Noncompliant scans are flagged for immediate follow-up. Optional web-based radiological evaluations are also supported. In addition, automated routines will be executed to determine overall image quality; for BOLD data, for example, signal-to-noise and movement histograms will be generated to inform investigators.

NUNDA Description: NUNDA is built on the NIH Biomedical Informatics Research Network (BIRN) infrastructure, XNAT (Extensible Neuroimaging Archive Toolkit ([Marcus et al., 2007a](#_ENREF_2); [Marcus et al., 2007b](#_ENREF_3))), and is designed to securely archive, efficiently process, and conveniently retrieve multisite, multimodality imaging data: MRI, PET, and CT. NUNDA currently contains over 16,980 subjects with 26,500 imaging sessions (replace with most current numbers). NUNDA’s features include: 1) a range of image upload/download methods, including DICOM (as mentioned above), FTP, and web interface; 2) incorporation of new data types by automatically generating the necessary databasing components; 3) quality control modules and audit trails; 4) a secure web-based user interface; 5) a sophisticated search engine; 6) an online image viewer; and 7) automatic image processing pipeline support. In NUNDA, data security is addressed at several levels. Scan data will be transmitted to NUNDA over secure channels using SSL. NUNDA resides on enterprise-grade linux servers with restricted access; each server has its own firewall, which will be monitored by the NUNDA system administrator and Northwestern IT personnel. The entire computer system is housed in the NU Data Center facilities, which have additional security. NUNDA access will be restricted to authorized users with specific privileges (create, delete, read, edit) according to their role in the study. Only selected personnel will have data modification privileges. NUNDA storage is managed by NU IT with duplication. Daily data backup will include copies of the data in various stages of analysis to be used to reproduce results from the same data origins.

**Long version to be included in Facilities:**

**Northwestern University Neuroimaging Data Archive (NUNDA)**

The scans aggregated for this project will be archived using the Northwestern University Neuroimaging Data Archive (NUNDA) (<http://nunda.northwestern.edu/>). Directed by Dr. Wang, NUNDA is a tool for the integrated storage and pre-processing of structural and functional imaging data: Acquisition 🡪 Archive 🡪 Processing & Analysis 🡪 Integration.

NUNDA was developed jointly by the Departments of Psychiatry and Behavioral Sciences and Radiology to integrate neuroimaging research data (multi-modal imaging, demographic, and behavioral) to efficiently process and securely store data. NUNDA currently stores about 12,000 imaging sessions from 8,000 subjects, representing over 190 projects from various research groups at NU and NUFSM. Image files are directly transferred to the data archive from MRI scanners. Newly entered data are placed in a virtual quarantine until an authorized user validates the integrity of the data. Once the data have been validated, they are moved into a secure archive. Archived data are made available to data-specific automated processing and analysis pipelines. Automated image processing routines generate calculated images and quantitative ROI-based summary measures that are integrated with demographic, clinical and neurobiological measures. The summary data are stored in the database for use in the statistical modeling of treatment effects and natural recovery. Data are archived using the Extensible Neuroimaging Archive Toolkit (XNAT). XNAT follows a three-tiered architecture that includes a data archive, user interface, and middleware engine, designed to facilitate management and exploration of neuroimaging and related data. It includes a secure database backend and a rich web-based user interface. XNAT uses an XML data model from which a relational database is generated. Non-imaging data are either entered via web-based forms, spreadsheet uploads, or XML on NUNDA, or captured in REDCap databases, which can be linked with the imaging database on NUNDA. XNAT’s web-based user interface provides tools for monitoring workflow and for exploring the resulting archive. As XNAT-managed studies progress, the data are made available to broader groups of users, from collaborators to members of the scientific community. The databases from this project are then created and integrated into NUNDA. Description of NUNDA is published in: “Alpert, K., Kogan, A., Parrish, T., Marcus, D., Wang, L., 2016. The Northwestern University Neuroimaging Data Archive (NUNDA). Neuroimage 124, 1131-1136.”([Alpert et al., 2016](#_ENREF_1))

Security for imaging data: Data security is addressed in several ways. The database will reside on enterprise-grade Linux servers with restricted access; each server has its own firewall, which will be monitored by NUNDA system administrator and Northwestern University IT personnel. The entire computer system is housed in the Northwestern University Data Center which has additional security. Users must have password access to the API or browser interface for database access. NUNDA access will be restricted to authorized users with specific privileges (create, delete, read, edit) according to their role in the study. Only selected personnel will have data modification privileges. Selected personnel will be able to upload data to a quarantine space that will verify the integrity and completeness of the dataset. If the data passes QA and is verified to be valid, it will be moved into the permanent archive for all authorized researchers to access. The uploaded data will be checked for any personal identification data prior to moving to the permanent archive.

Infrastructure and Backup: NUNDA hardware are housed in the NU Data Center facilities (<http://www.it.northwestern.edu/data-centers/>), operated by the Northwestern University IT, which provides 24/7 HVAC, secure data networks, service monitoring and environmental management. Technical staff at the data center is available 24/7 for emergencies. NUNDA storage is managed by NU IT with duplication. Daily data backup will include copies of the data in various stages of analysis to be used to reproduce results from the same data origins.

In addition, all NUNDA quality assurance and image processing pipelines will be performed on the NU High Performance Computing (HPC) Cluster, known as Quest (<http://www.it.northwestern.edu/research/services/quest.html>). See below.

**Northwestern University High Performance Computing System**

Northwestern University's high performance computing (HPC) system is referred to as QUEST. Housed at the University's secure Data Center facilities, the 3.5 PB, >16,000 core computing cluster offers a large shared computational facility. NIACAL has basic access to run 32 jobs in queue at any one time. Intensive computing jobs are allocated to run at QUEST. A number of different software packages are available on the cluster, and nodes on the cluster are being transformed with success to run "cloud computing" virtual machines. Together with NUNDA, we have developed a high performance cluster to rapid analysis and a customizable searchable database.

**References:**

Alpert, K., Kogan, A., Parrish, T., Marcus, D., Wang, L., 2016. The Northwestern University Neuroimaging Data Archive (NUNDA). Neuroimage 124, 1131-1136.

Marcus, D.S., Olsen, T.R., Ramaratnam, M., Buckner, R.L., 2007a. The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. Neuroinformatics 5, 11-34.

Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L., 2007b. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J Cogn Neurosci 19, 1498-1507.