



HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2017 January 01.

Published in final edited form as:

Neuroimage. 2016 January 1; 124(Pt B): 1093–1096. doi:10.1016/j.neuroimage.2015.06.076.

XNAT Central: Open Sourcing Imaging Research Data

Rick Herrick^a, William Horton^a, Timothy Olsen^b, Michael McKay^a, Kevin A. Archie^a, and Daniel S. Marcus^a

^aDepartment of Radiology, Washington University School of Medicine, St. Louis, MO, USA

^bDeck5 Consulting, Normal, IL, USA

Abstract

XNAT Central is a publicly accessible medical imaging data repository based on the XNAT open-source imaging informatics platform. It hosts a wide variety of research imaging data sets. The primary motivation for creating XNAT Central was to provide a central repository to host and provide access to a wide variety of neuroimaging data. In this capacity, XNAT Central hosts a number of data sets from research labs and investigative efforts from around the world, including the OASIS Brains imaging studies, the NUSDAST study of schizophrenia, and more. Over time, XNAT Central has expanded to include imaging data from many different fields of research, including oncology, orthopedics, cardiology, and animal studies, but continues to emphasize neuroimaging data. Through the use of XNAT'S DICOM metadata extraction capabilities, XNAT Central provides a searchable repository of imaging data that can be referenced by groups, labs, or individuals working in many different areas of research. The future development of XNAT Central will be geared towards greater ease of use as a reference library of heterogeneous neuroimaging data and associated synthetic data. It will also become a tool for making data available supporting published research and academic articles.

Keywords

XNAT; XNAT Central; Neuroinformatics Databases; Open Access; Data Sharing; Open Source

Introduction

XNAT Central (<https://central.xnat.org>) is a publicly accessible medical imaging data repository based on the XNAT open-source imaging informatics platform (Marcus, Olsen, et al. 2007). In contrast with most other image repositories, XNAT Central is not moderated to control content or tailored to support any particular disease, modality, scientific approach, or data use terms. Furthermore, data sets can be continually edited and expanded to enable

Corresponding author: Rick Herrick, Neuroinformatics Research Group, Mallinckrodt Institute of Radiology, Washington University School of Medicine, Campus Box 8225, 4525 Scott Avenue, Saint Louis, Missouri 63119-2451, rick.herrick@wustl.edu, +1-314-740-5961.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

open-ended development. This open, laissez-faire approach complements more controlled repositories like ConnectomeDB (Hodge et al, this issue) and Open fMRI (Poldrack et al. 2013) by providing an avenue for rapid, self-styled data sharing and informatics software development.

XNAT Central was established in 2006 with the dual purposes of providing a place to share data and of serving as a sandbox environment for developing and assessing XNAT and related software. Thus a mix of data ranging from high quality open access published data sets to error-filled test data sets quickly emerged on the site. The ability to store structured data along with more loosely organized data has made XNAT Central a popular option for supporting software development challenges hosted by MICCAI and other organizations. Over time, research groups have used XNAT Central both to stage data for further processing and to host and aggregate clinical and other data associated with the initial raw imaging studies. This has expanded the function of XNAT Central from its original focus as a means to make data available to actually providing ways to enrich and further understand the meaning and significance of the data. XNAT Central is also typically the first public XNAT site upgraded when new XNAT versions are released, thus providing a resource for exploring new XNAT features.

Current Services

XNAT Central is an unmodified instantiation of the current release of the XNAT imaging informatics platform (Figure 1). Additional data types, including behavioral and clinical instruments, have been added via XNAT's standard XML Schema extensions mechanism to support specific data sets. The system includes a web-based data upload tool for uploading individual DICOM-formatted imaging session data. These tools automatically remove personally identifying metadata from the DICOM headers following the DICOM Supplement 142 anonymization profile (DICOM WG 18 n.d.). Upon upload, XNAT extracts relevant DICOM metadata into its database, enabling users to identify and search for data sets using specific acquisition parameters. Batches of DICOM data can be uploaded using XNAT's scriptable application programming interface (API) in combination with tools such as DICOM Browser (Archie and Marcus 2012) and Clinical Trial Processor (RSNA 2013). Non-DICOM data can also be uploaded to the system using XNAT's API as well as through upload links on the website. Data hosted on the site can be easily searched, reviewed, visualized, and downloaded using XNAT's standard tools as well as by external tools developed by the community, including pyxnat (Schwartz et al. 2012). Data identifiers (e.g. subject IDs, session IDs) are chosen by the contributors of the data and are then accessible via permanent URLs associated with those identifiers via the XNAT API (Figure 2). Similarly, search terms for projects and individual study resources are selected by contributors and can be modified by them at any time.

XNAT Central uses XNAT's standard user management, project and data organization, and data access controls. Individuals navigating to the site without logging in are treated as guests with highly restricted access to site functionality. To create projects, upload data, and access non-public data, an individual must create a password-protected user account and login each time they visit the site. XNAT's standard project management mechanism

provides project owners with flexible control over access to their data. Projects can be set to public, protected, or private. Public projects are accessible to all individuals navigating to the XNAT Central site, including those who do not explicitly login to the system. Protected project data is only accessible to users who have been explicitly granted access to the project but descriptive project metadata (title, investigator name, keywords) are accessible to all users. A link on protected projects is provided to request access to the data. Private projects are entirely hidden from users who have not been granted explicit access. By default, projects are set to “private” mode. Users added explicitly to projects can be enabled with read-only or fully-edit permissions. XNAT Central’s operations staff currently maintains no control or oversight over the data sharing status of projects created by external users. Furthermore, data use agreements must be maintained by individual projects through mechanisms external to XNAT Central. The NUSDAST project (Wang et al. 2013), for example, maintains their data in a protected project and hosts an independent website for individuals to review and agree to data use terms (http://niacal.northwestern.edu/nusdast_accessors/new). These individuals separately request access through XNAT Central’s access request form, and NUSDAST staff grants access to those who have agreed to their data use terms.

Hosted Data

XNAT Central hosts data sets from a number of different research labs and investigative efforts. The majority of data are maintained in private projects, many of which may eventually be made publicly accessible (see *Contributing Data* below). A majority of data hosted in XNAT Central is neuroimaging though data sets in oncology and other domains are also present. Example open access data sets include the OASIS Brains imaging studies, which focus on cross-sectional (Marcus, Wang, et al. 2007) and longitudinal (Marcus et al. 2009) imaging across the human lifespan, including individuals with Alzheimer’s disease; the Northwestern University Schizophrenia Data and Software Tool (NUSDAST), which provides clinical, cognitive, genetic, and imaging data for 450 subjects, including many with diagnosed schizophrenia (Wang et al. 2013); and animal studies on cerebral amyloidosis (Grandjean et al. 2014) and functional networks measured with diffuse optical tomography (Eggebrecht et al. 2014). Various calibration, tutorial, and reference data sets are also present for validating calculations and processing scripts, and as real-world data samples for learning and teaching purposes. These include phantom DICOM data, reference data for image registration validation, sample data sets for programming or analytical challenges, and data to support tutorials and teaching tools.

Contributing Data

Any registered user can contribute data to XNAT Central. Thus far private projects are allowed for an unlimited duration, though in order to promote open sharing, policies will likely be introduced to delete private data after a reasonable period. While XNAT Central is open to all data contributions, we are particularly interested in data sets that are less likely to be supported by complementary data repositories. Such data include reference data that can be used to validate image processing and analysis software. Similarly, reference data that represent advances in scanner technology, acquisition sequences, and DICOM and other

imaging formats. These advances pose particular challenges for informaticians designing and implementing the systems used to store, manipulate, and publish these data. The XNAT Central team plans to recruit research groups and laboratories throughout the existing XNAT user community to contribute data from real-world investigative efforts to help provide working samples of as many of these new modalities and formats as possible. The goal is to provide thorough coverage of newer or leading edge modalities such as PET-MR and DICOM-RT, proprietary DICOM formatting and data handling such as Siemens-, Philips-, and GE-specific DICOM headers and sequences that may hide or obscure PHI or other sensitive data, and methods for processing and managing these different types of data. The worldwide XNAT community is a valuable resource that not only has access to these technologies in the neuroimaging field but also actively advances those technologies. XNAT Central provides a critical tool to leverage and share that resource. XNAT Central's open support for unstructured data is also well-suited data sets that include ancillary data such as raw k-space files from MR acquisitions, region of interest files, and task files for fMRI experiments. Permanent URIs to XNAT Central hosted data are generated at the individual project, subject, and session/experiment levels. In addition, specific file resources (e.g. an individual DICOM series, a single NIFTI file) can be referenced by URI.

Future Development

Going forward, there are a number of steps planned to increase the available data sets hosted on XNAT Central, as well as to increase the accessibility and utility of the hosted data.

Reducing clutter

XNAT Central's open submission model has enabled a broad range of collaborative research and facilitated publication of a number of open access data sets. However, this unmoderated approach has also worked to the detriment of XNAT Central as a usable resource. Many users have used XNAT Central as a sandbox to test out XNAT itself, store data for a short period of time, or run into issues when trying to maintain their own project. As a result, the signal-to-noise ratio—in the form of useful and interesting projects compared to one-off test projects or abandoned half-completed efforts at a data collection—has gotten low enough that it detracts from XNAT Central's usability and search capabilities. As a first pass at reducing the clutter, we intend to review all existing projects and delete those that are of little or no value. Second, a dedicated sandbox XNAT environment is being instantiated for users primarily interested in “kicking the tires” of the XNAT platform. Finally, there are a number of newer features in the base XNAT platform that should help prevent casual or inexpert use from obscuring the repository with noise, including event automation, automated access control, and security on project creation and image uploads.

Data Publication Capabilities

As an unmodified instance of XNAT, XNAT Central's user interface feels less like a data repository and more like the data management system that is more typical of XNAT's use. Its primary features that focus largely enabling the acquisition, processing, and aggregating of neuroimaging research data have been well defined and established for a while. This has led to widespread acceptance of XNAT in the neuroimaging research community as a

flexible and performant tool for these purposes. Other aspects of the research lifecycle are significantly less developed within the core XNAT system. However, emerging uses of XNAT as a publication framework, particularly by the Human Connectome Project (Hodge et al., this issue), have led to the development of a number of new capabilities to curate and disseminate neuroimaging and related data. We anticipate that XNAT Central will be the primary proving ground for the continued development of XNAT as a publishing and sharing platform will be the primary proving ground for this work. Key capabilities will include streamlined data submission and review procedures; support for project-specific data use terms; improved project metadata management and search functionality (Herrick et al. 2014); and support for emerging data sharing formats and interfaces (Poline et al. 2012). In addition to these core features, we anticipate creating a custom user interface to build navigation around disease- and domain-specific portfolios.

XNAT Federation through XNAT Central

In many cases, high value data sets are already available in independently operated XNAT instances around the globe, including many described in this issue. We anticipate evolving XNAT Central into a system that federates data hosted in remote XNAT instances into a unified, searchable “repository of repositories”. This could theoretically give a researcher working with XNAT Central worldwide reach across vast data sets, without the burden of actually moving the bulk of binary imaging data into a single collection point. By combining this federation capability with planned advances in components for sharing and publishing research data, XNAT Central can serve as a primary resource for sharing the wealth of openly available neuroimaging data. An important balance will need to emerge between the general interface provided by XNAT Central and focused interfaces made available by ConnectomeDB and other systems described in this issue.

Acknowledgments

This project was supported by NIH grants 5 R01 EB009352, U24 RR026057, U54 EB005149, and U24 RR025736. We wish to thank all contributors to XNAT Central, especially the NUSDAST and OASIS projects.

References

- Archie, Kevin A.; Marcus, Daniel S. DicomBrowser: Software for Viewing and Modifying DICOM Metadata. *Journal of Digital Imaging: The Official Journal of the Society for Computer Applications in Radiology*. 2012; 25(5):635–645.
- DICOM WG 18. [accessed January 28, 2015] Digital Imaging and Communications in Medicine (DICOM) Supplement 142: Clinical Trial De-Identification Profiles. N.d.[ftp://medical.nema.org/medical/dicom/final/sup142_ft.pdf](http://medical.nema.org/medical/dicom/final/sup142_ft.pdf)
- Eggebrecht, Adam T.; Ferradal, Silvina L.; Robichaux-Viehoever, Amy, et al. Mapping Distributed Brain Function and Networks with Diffuse Optical Tomography. *Nature Photonics*. 2014; 8(6):448–454. [PubMed: 25083161]
- Grandjean, Joanes; Schroeter, Aileen; He, Pan, et al. Early Alterations in Functional Connectivity and White Matter Structure in a Transgenic Mouse Model of Cerebral Amyloidosis. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2014; 34(41):13780–13789. [PubMed: 25297104]
- Herrick, Rick; McKay, Michael; Olsen, Timothy, et al. Data Dictionary Services in XNAT and the Human Connectome Project. *Frontiers in Neuroinformatics*. 2014 In press.

- Marcus, Daniel S.; Olsen, Timothy R.; Ramaratnam, Mohana; Buckner, Randy L. The Extensible Neuroimaging Archive Toolkit: An Informatics Platform for Managing, Exploring, and Sharing Neuroimaging Data. *Neuroinformatics*. 2007; 5(1):11–34. [PubMed: 17426351]
- Marcus, Daniel S.; Wang, Tracy H.; Parker, Jamie, et al. Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience*. 2007; 19(9):1498–1507. [PubMed: 17714011]
- Marcus, Daniel S.; Fotenos, Anthony F.; Csernansky, John G.; Morris, John C.; Buckner, Randy L. Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults. *Journal of Cognitive Neuroscience*. 2009; 22(12):2677–2684. [PubMed: 19929323]
- Poldrack, Russell A.; Barch, Deanna M.; Mitchell, Jason P., et al. Toward Open Sharing of Task-Based fMRI Data: The OpenfMRI Project. *Frontiers in Neuroinformatics*. 2013; 7:12. [PubMed: 23847528]
- Poline, Jean-Baptiste; Breeze, Janis L.; Ghosh, Satrajit, et al. Data Sharing in Neuroimaging Research. *Frontiers in Neuroinformatics*. 2012; 6:9. [PubMed: 22493576]
- RSNA. Clinical Trial Processor. 2013. http://mircwiki.rsna.org/index.php?title=CTP_Articles
- Schwartz, Yannick; Barbot, Alexis; Thyreau, Benjamin, et al. PyXNAT: XNAT in Python. *Frontiers in Neuroinformatics*. 2012; 6:12. [PubMed: 22654752]
- Wang, Lei; Kogan, Alex; Cobia, Derin, et al. Northwestern University Schizophrenia Data and Software Tool (NUSDAST). *Frontiers in Neuroinformatics*. 2013; 7:25. [PubMed: 24223551]

Highlights

XNAT Central is a publicly accessible medical imaging data repository.

It runs on the widely used XNAT open-source imaging informatics platform.

It emphasizes neuroimaging data but includes imaging data from many different fields.

Research groups from around the world provide data on XNAT Central.

Future development will emphasize publishing and data management functions.

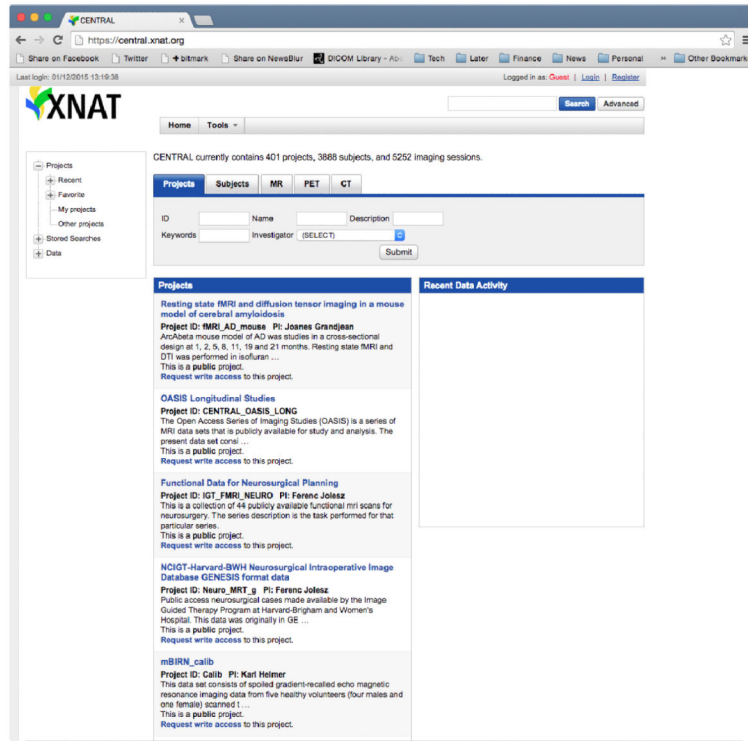


Figure 1.
XNAT Central landing page

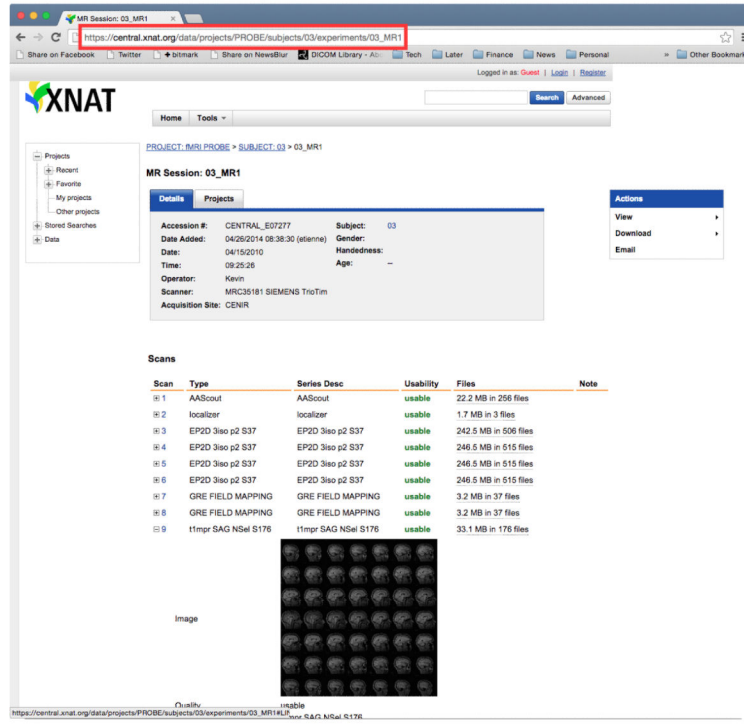


Figure 2.
Data on XNAT Central accessible via permanent URL