CIP Survey of Biomedical Imaging Archives

Note: When contributing to this page please use Firefox as your browser to allow use of the "Rich Text" WYSIWYG interface for editing this page. Internet Explorer, Chrome, and Safari only allow editing the wiki markup code.

Definition of project

The goal of this project is to create a survey of Publicly Available InVivo Medical Imaging Archives and the underlying software capabilities. It is generally agreed that there is a need for public medical imaging archives to provide the biomedical research community, industry, and academia with access to images that support:

- Lesion detection and classification
- Accelerated diagnostic imaging decision
- Quantitative imaging assessment of drug response

The purpose of this project is to provide a practical guide for the community which allows them to:

1. to assess existing software and instantiations that are appropriate to their research or clinical needs.
2. to locate relevant publicly available data for research

We encourage any feedback from the wider community that may help improve this information or correct any misconceptions stated below. The survey is divided into two sections:

1. Publicly hosted biomedical imaging archives which are populated with actual data which researchers, teachers, industry, etc may wish to utilize
2. Image archive software solutions which one could download and use to host their own DICOM image data sets

Please contact Justin Kirby (kirbyju@mail.nih.gov) or John Freymann (freymanj@mail.nih.gov) with any questions, error reports, updates, additions, etc.

Acknowledgements

We would like to thank the following people for volunteering their time and effort in helping us populate this survey.

- Dan Marcus (WUSTL)
- Brian Hughes (Terpsys)
- Dan Hall (NIH)
- Patrick Reynolds (Kitware)
- Julien Jomier (Kitware)
- Ivo Dinov (UCLA)
- Matthew McAuliffe (NIH)
- David Keator (UCI)
- Christian Haselgrove (NITRC)

Publicly Hosted Biomedical Imaging Archives

The following table attempts to summarize publicly accessible biomedical image archives. This survey originally initiated in August of 2010. Information in the tables are being updated periodically.

NOTE: Due to the large size of this table you may need to use the horizontal scroll bar at the bottom of the table to view some of the archives listed furthest to the right.

	The Cancer Imaging Archive (TCIA)	NBIA	NIAMS	XNAT Central	Image Data Archive	Function BIRN Data Repository	Give A Scan	Optical Society of America (OSA)	Insight Journal (MIDAS)	National Database for Autism Research (NDAR)	Pediatric MRI Data Repository	FITBIR	NITRC-IR
Supporting Institution(s)	NCI Cancer Imaging Program	NCI Center for Bioinformatics and Information Technology	NIAMS, NCI Center for Bioinformatics and Information Technology	WUSTL, BIRN	Lab of NeuroImaging UCLA (LONI)	FBIRN Institutions	Lung Cancer Alliance, Kitware	Optical Society of America, Kitware	Kitware, Insight Software Consortium	NIH, NIMH, NINDS, NICHD, NIEHS	NIH, NIMH, NICHD, NIDA, NINDS	NINDS, DoD	NIH Blueprint, NIBIB, NINDS, NIMH, NIDA
Content Type	In Vivo Cancer Imaging, phantom imaging and related metada (see full Collection list)	Demo instance for showing the latest features of NBIA. All data are re-used from The Cancer Imaging Archive.	Osteoarthritis	Biomedical images, meta data, other phenotypic data (behavioral, clinical, etc)	ADNI (Alzheimers), CRYO (histology), ICBM (Brain mapping), AIBL (Autralian Aging)	FMRI/MRI images, behavioral data, and clinical data from schizophrenics and healthy volunteers. Willing to accept data on other neurological disorders.	Patient-contributed Lung Cancer Medical scans	Optical, digital holography, 2D/3D modalities, etc	Biomedical images, meta data, and journal articles	Autism - standard phenotypic data, imaging and genomic /pedigree data related to human subjects	Normal brain development	TBI related data: imaging, phenotypic and some genomics; human but expanding to preclinical models	Neuroimages

Archive Software	NBIA, AIM Data Service (XML image metadata), and a Clinical Data relational database	NBIA	NBIA	XNAT	Image Data Archive	Human Imaging Database (HID)	MIDAS	MIDAS	MIDAS	custom	Same as NDAR (custom)	Biomedical Research Informatics Computing System (BRICS) NIH developed – custom	XNAT
Login account required	No. Public data is accessible without logging in. Some special restricted-access data sets do require free registration. Register on the TCIA website.	For advanced site features or limited access data sets, but is not required for accessing public data. Click here to register.	Yes. Register on the NIAMS website.	For accessing limited access data sets, but not for public data	Yes, via web https://ida.loni.ucla.edu/login.jsp	No (email requested)	No	For accessing limited access data sets, but not for public data	Only for submitting data.	Yes	Yes	Yes	For accessing limited access data sets, but not for public data
Explicit data sharing policy	Yes, with options for uploading fully open or limited access data sets	Yes, with options for uploading fully open or limited access data sets	Yes, found here	No, data is made public or restricted as specified by the user who uploads it.	Yes. IDA User Manual.	All data is made publicly accessible.	All data is made publicly accessible.	Yes, found here	All data is made publicly accessible (varying licenses)	Yes, found here	Similar to NDAR but there is no explicit policy	Yes, https://fitbir.nih.gov/jsp/about/policy.jsp	No
Number of Registered Users (or NA)	6,117	2,712	46	~1,000	>1,000	N/A	N/A		2,657	60 for data access 100-200 for data submission	~30	15 – just starting	1976 users as of April, 2015
Accepting new data	Yes, proposals are accepted via email and reviewed monthly by the TCIA Advisory Committee. Acceptance criteria is summarized here: Requesting Permission to Upload your Data	No	More data is being added as part of the official initiative, but external proposals are not being accepted.	Yes, users can register accounts and upload data	Yes, see section 9 in the Appendix of the LONI Policies & Procedures	Yes, https://www.birncommunity.org/about/contact/	Yes, through Lung Cancer Alliance. Learn more here	Yes, new data may be added as part of future Optical Society of America publications are released.	Yes, users can register accounts and upload data.	Yes, learn more here	No	Yes, FITBIR has established a two-tiered submission strategy to ensure high quality and to provide maximum benefit to investigators. See the Data Submission Procedures for more information.	Yes. Community-generated data sets may be suggested for inclusion. Contact moderator@nitrc.org.
Central curation /review	Yes, a multiple tiered de-identification and QC process is utilized involving both human review and systematic analysis. The process is summarized in detail on the TCIA De-Identification Knowledge Base and What to Expect as an Image Provider	No	Yes, performed by NIAMS staff.	No	Collaborators strip all personal info from data prior to submission to LONI. Then LONI auto filters again, to ensure that there are no PHI in the files (especially if the data is binary) and stores the data in quarantine, until it's approved for posting to the web interface.	PHI must be removed by the submitting institution prior to giving the data to the FBIRN. FBIRN also performs a review to make sure there aren't any de-identification problems.	Yes, performed by Lung Cancer Alliance	Yes, performed by the Optical Society of America	Yes, some QC performed by Kitware staff and peer reviews. Most data is de-identified by the submitter prior to upload.	Sites do their own de-identification any way they prefer, so long as it meets their IRB's approval. Pre-validation is performed to ensure all data conforms /harmonizes to the autism data dictionary. QA is also performed by NDAR staff to check for identifiable information.	Archived project, no longer receiving new data.	Yes, pre-validation is performed to ensure all data conforms to the NINDS CDEs. QA is also performed by staff to check for personal identifiable information.	Yes, performed by NITRC.
Availability /Uptime	~99%, hosted on a redundant production system at WUSTL	~99%, hosted on a redundant production system at NCI CBIIT	~99%, hosted on a redundant production system at NCI CBIIT	~99%	Continuous (no exact % specified)	Continuous (no exact % specified)	~99.9%, hosted on a production server at Kitware	~99.9%, hosted on a production server at OSA	~99%, hosted on a production server at Kitware	~99%, imaging data hosted on Amazon and metadata hosted by NIH.	~99%, hosted by NIH.	~99%, hosted by NIH.	~99.9%, hosted on production servers at UCSD.
Project- or Collection-based groupings?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Size of Current Volume	TCIA: 2.26 TB NLST-LSS: 6.5 TB NLST-ACRIN: 4.8 TB **Total: 13.56 TB**	~21GB	~7.5TB	1 TB	0.7 PB	~2TB	33GB	~50GB	~60GB	~2TB	~2TB	0.5TB	1.5 TB
Number of patients /subjects with imaging	TCIA: 5,997 NLST-LSS: 17,043 NLST-ACRIN: 9,211 **Total: 32,251**	32	4,796	3,494	> 120,000	~300	37	N/A	> 200, plus some non-patient data	2500 NDAR	550 (migrated into NDAR)	200	6854 subjects

Number of DICOM Tags query-able	2 via the Simple Search, ~90 via the Advanced Search, and all DICOM metadata can be queried using the Text Search option.	2 via the Simple Search, ~90 via the Advanced Search, and all DICOM metadata can be queried using the Text Search option.	~90	~50	50	0	22 (imaging parameters in query interface)	N/A	22 (imaging parameters in query interface)	9, with full listing from data dictionary. (4 NDAR, 5 Pediatric MRI)	9, with full listing from data dictionary. (4 NDAR, 5 Pediatric MRI)	10	N/A
Metadata Availability	Wide variety of clinical, genetic, and image segmentation /annotation available is available for various data sets. Full summary can be viewed here in the "Supporting Data Available" column.	None	None	Various clinical and other metadata	Biospecimen, clinical, pathological, neuropsychiatric, and demographic.	Yes, extensive behavioral data and clinical data.	Some unstructured clinical data such as patient age, cancer stage, recurrence, and treatment information.	Associated articles, figures, publication-specific metadata, etc	Unstructured clinical data as well as publication-specific metadata	NDAR contains all human subjects data related to autism research funded by the NIH and others. Outside of the NDAR data dictionary, Metadata supporting project definition and research results are provided (see data from papers).		Yes, all metadata is collected using NINDS CDEs	Gender, age, handedness, diagnosis, and MR acquisition parameters. Various cognitive assessments and other rich metadata available for some data.
Data submission /download methods	Submission via DICOM or HTTPS protocols using CTP. Download via Java Webstart client. A REST API is also available with documentation.	Submission via DICOM or HTTPS protocols using CTP. Download via Web (zip), FTP, Java Webstart client	Submission via DICOM or HTTPS protocols using CTP. Download via Web (zip), FTP, Java Webstart client	Submission via Web UI or DICOM protocol. Download via Web (zip) or Java applet.	Secure web upload	Downloads via Web UI, Submissions via https://www.birncommunity.org/about/contact/	Submission via Web UI, Download via Web (zip).	Submission via Web UI, Download via Web (zip).	Submission via Web UI, DICOM push, MIDASDesktop. Download via Web (zip), MIDASDesktop	A custom Java Webstart application allows SFTP /Amazon S3 transfers. MIPAV is offered as an optional method for de-identification. Submission is harmonized to the autism data standard using custom data validation software. Download methods include multithreaded download from the Amazon Cloud or push to cloud computational pipeline.	Not applicable.	They can use MIPAV for submitting images. Submissions must conform to FITBIR Data Dictionary (NINDS CDEs). A custom Java Webstart application allows SFTP transfers.	Upload by arrangement. Download via Web (zip), Java applet, or REST API.
Helpdesk Support	Yes, the TCIA Helpdesk supports both end users and submitters. They provide phone and email support during regular business hours Mon-Fri.	Yes, CBIIT Application Support	Yes, CBIIT Application Support	Via XNAT discussion group	dba@loni.ucla.edu	Yes, via https://www.birncommunity.org/about/contact/	Technical issues can be sent to midas@public.kitware.com or click here for Administrative support and other questions	Contact infobase@osa.org.	Contact midas@public.kitware.com	Yes, available at ndarhelp@mail.nih.gov.	Yes, pedsmri@mail.nih.gov.	Yes, FITBIR-help@mail.nih.gov	Contact moderator@nitrc.org.
Affiliation with Journal	Not directly, but Digital Object Identifiers can be provided for integration with publications. TCIA is also a recommended repository for Nature's Scientific Data journal.	No	No	No	Yes, NeuroImage	No	No	Yes, Optics Info Base	Yes, Insight Journal	No	No	No	No
Intended Audience(s)	Cancer researchers, engineers and developers, professors	Anyone interested in testing the functionality of the NBIA software.	Osteoarthritis researchers	All imaging research	Neuroimaging and genetics research	Neuroimaging research	Lung cancer researchers	Optical Society of America subscribers	All imaging research	Autism researchers (clinical /phenotype /genomic), both those receiving autism related NIH grants and other investigators sponsored by an NIH recognized institution with a current federal-wide assurance.	Neuroscientists interested in normative brain study of child development.	TBI researchers	Neuroimaging research

Image archive software solutions

Below is a list of image archive solutions that can be deployed by interested parties wishing to build their own DICOM based biomedical image archive. This list omits some of the archives above in cases where we could not find any information about how one might download and deploy their own instance of the software.

Software Name and Web Site	NBIA GForge project	XNAT	MIDAS
Interface /GUI	Web	Web	Web/Desktop Application
Query types /flexibility	Simple (9 parameters), Advanced (10 more parameters), Dynamic (boolean query of up to 90 DICOM tags)	Extensible set of DICOM tags as well as linked quantitative biomarkers, linked clinical data, and other non-imaging data.	Customizable, search by any tags registered in the system
Role Based Security	Yes	Yes	Yes
Public access option (no login req)	Yes	Yes	Yes
Active Development	Yes, NCI CBIIT	Yes, WUSTL Neuroinformatics Research Group	Yes, Kitware
License	Open source - NBIA License Agreement Details	Non-restrictive (BSD) open-source license - XNAT License Agreement Details	non-restrictive (BSD) open-source license
API available	Yes, REST	Yes, REST	Yes, REST, OAI-PMH
Supported image formats	DICOM	Automated import of DICOM and ECAT. Custom importers can be implemented for other formats. Any file type can be uploaded through the API and web interface.	DICOM and other ITK-based format
Supported metadata formats	XML, Zip	XML, CSV. Custom data import logic can be implemented via pluggable Groovy and Python scripts.	XML
Transfer protocols (import /export)	DICOM, HTTPS	DICOM, HTTPS	DICOM, HTTPS
Controlled Vocabulary	Follows caBIG standards (caDSR/EVS)	XNAT Schema	NIH Mesh and Dublin Core
Deployment Support	Yes, CBIIT Application Support or via NBIA User Listserv	XNAT Google Discussion group, monthly developer tcons, biannual user conference. Commercial technical support provided by Radiologics.	Yes, MIDAS mailing list
Support Operating Systems	Linux, Windows, Mac	Linux, Windows, Mac	Linux, Windows, Mac
Data submission options	Submission to NBIA is performed by a java tool called CTP developed by John Perry at the RSNA. CTP has options to import data from a hard drive or directly from a PACS or DICOM Workstation.	Direct upload is available through the web UI, direct DICOM transfer, scripts using REST API, optimized CTP workflow	Direct upload via web UI, direct DICOM transfer via push, MIDASDesktop transfer (includes command line tools), WebDAV support.
Standard of De-Identification	Incorporates DICOM de-identification standards from The Attribute Confidentiality Profile (DICOM PS 3.15: Appendix E) via CTP.	Built-in de-identification language based on DICOM Browser can be configured to comply with DICOM PS 3.15: Appendix E and other standards.	No, but pre-storage filters can be run automatically
Support for multi-site submissions	Yes	Yes	Yes

Help Downloading Files

For help accessing PDF, audio, video, and compressed files on this wiki, go to Help Downloading Files.