

Euro BioImaging

Preparatory Phase II Project

D6.1 Report presenting the first comprehensive catalogue of current and future image data repositories in EU institutions and states, covering the domains of biological microscopy, preclinical imaging and medical imaging, and all imaging modalities included in Euro-BioImaging

Project N.	688945
Project Title	Euro-BioImaging Preparatory Phase II
Project Acronym	EuBI PPII
Associated Work Package	WP6
Associated Task	Task 6.6
Lead Beneficiary (short name)	FISABIO, COIMBRA, UNIVDUN
Nature	Report
Dissemination Level	Public
Estimated Delivery Date (Grant Agreement, Annex I)	30/06/2016
Actual Delivery Date	03/08/2016
Task leader	Jason Swedlow ¹
Contributors	Gabriella Rustici



Funded by the
Horizon 2020
Framework
Program of the
European Union

¹ Compared to the Grant Agreement the Task Leads have been changed from FISABIO and COIMBRA to UNIVDUN.

Abstract

Despite significant advances in biological and biomedical imaging and analysis, resources for publishing image datasets are limited. In order to provide an objective measure of the current state of the art of publicly available image repositories, we have surveyed the Euro-BioImaging and global communities to identify the location, capabilities domain coverage of public image data resources. The survey was run during July 2016. The records in the resulting catalogue have been cross-checked, and if complete information and approval by owner is provided, they are published in a specially constructed Euro-BioImaging collection in Bio-Sharing, the definitive, ELIXIR-associated community resources for standards, databases and collections. The Euro-BioImaging Bio-Sharing collection is the first public collection of image data repositories and is available at <https://www.biosharing.org/collection/EuroBioImaging>.

Table of Contents

1. Introduction	Page 2
2. A catalogue of Image Data Repositories	Page 3
3. Publishing and maintaining the Euro-BioImaging Image Data Repository Catalogue	Page 3
4. The Breadth and Scope of a Euro-BioImaging Image Data Repository Catalogue	Page 4
5. Table 1. Summary of Image Data Repositories reported in the Community Survey	Page 5

1. Introduction

Access to primary research data is vital for the advancement of the scientific enterprise. It facilitates the validation of existing observations and provides the raw materials to build on those observations. In the life sciences, research communities have repeatedly collaborated to build resources that allow public submission and access to particular types of datasets. These include gene sequences, protein structural data, and gene and protein expression profiles. In these cases the community united to standardize the structure of the data and its associated metadata, and to create centralized repositories to facilitate deposition, promote discoverability, and ensure the longevity of the data.

Much of the published research in the life sciences carries with it detailed image data. These images are routinely used for quantitative measures of biological processes and structures that form the foundation of many of the results published in peer-reviewed life sciences journals. In almost all cases, however, images are presented in published articles in processed, compressed formats that do not accurately convey the quality and complexity of the original image data.

2. A Catalogue of Image Data Repositories

Currently, there are several public image data repositories each that publish data from one or more specific experiments. These are independent resources, and to date, there is no connectivity between them. As a first step to building a fully functional Euro-BioImaging Image Data Repository, we have attempted to build a catalogue of these resources by surveying the global biological and biomedical communities and collecting details of known public image repositories. The survey was announced on several mailing lists, distributed to contacts at Euro-Imaging Nodes, and distributed by the members of Euro-BioImaging Prep Phase II WP6 to their colleagues in several institutions.

The survey was designed by members of WP6 to be easy to fill out within ten minutes and aimed at capturing, for each image data repository, basic information including: full name, URL, year of creation, free text description, contact details of the person responsible for its maintenance, its status, data access policy, the imaging domains covered, organisms represented and the data standards implemented.

Additionally we wanted to acquire details of the organizations involved with the resource (e.g. maintainers, funders, collaborators, etc), grants supporting it, publications, tools developed to work with the resources, web service access point, license and means of support available (e.g. forum, helpdesk, etc.)

The survey was opened in June 2016 and closed at the end of July 2016. The survey results can be found at: <https://www.surveymonkey.net/results/SM-RCVQG8GT/>

As of 15/07/2016, details of 22 image data repositories were collected in the survey. 17 were based in Europe, three are based in the USA, and one is based in Japan (See Table 1).

3. Publishing and maintaining the Euro-Bioimaging Image Data Repository Catalogue

In order to make our findings as accessible as possible, and to encourage the deposition of information on more repositories, we have created, for each reported repository, a record in the BioSharing repository². BioSharing is a public, curated, searchable portal of inter-related data standards, databases, and policies in the life, environmental and biomedical sciences. BioSharing harnesses community curation to collate and cross-reference resources across the life sciences from around the world, making them findable and accessible. Every record is interlinked, providing a detailed description not only on the resource itself, but also on its relations with other life science infrastructures. BioSharing is in the process of being recognized as an ELIXIR³ resource and a collaboration between Euro-BioImaging and Biosharing is in line with the Image Data Strategy jointly published by ELIXIR and Euro-BioImaging⁴.

² <https://biosharing.org/>

³ <https://www.elixir-europe.org/>

⁴ https://www.elixir-europe.org/sites/default/files/documents/euro-bioimaging_elixir_image_data_strategy.pdf

All BioSharing records are grouped in a “Euro-BioImaging” collection, which can be expanded in the future to include additional imaging resource (standards, databases or policies) as they are developed and released. Records will also be assigned to the maintainers of each image data repository allowing them to update the corresponding record when necessary.

For all entries, Euro-BioImaging requests relevant background information, including

- Full name,
 - URL,
 - Year of creation,
 - Free text description,
 - Contact details of the person responsible for its maintenance,
 - Its status,
 - Data access policy,
 - Imaging domains covered,
 - Organisms represented
 - Data standards implemented
-
- Approval by owner/publisher of image data repository that this resources is linked in the Euro-BioImaging collection on www.biosharing.org.

We will continue to submit records to the Euro-BioImaging Collection during the term of the Prep Phase II project, e.g. via the interim Web Access Portal (<http://www.eurobioimaging-interim.eu/>), and this resource can continue to be developed once Euro-BioImaging moves into construction.

All image data repositories with complete background information and approval by the owner, are included in the Euro-BioImaging collection available at: <https://www.biosharing.org/collection/EuroBioImaging>

4. The Breadth and Scope of a Euro-BioImaging Image Data Repository Catalogue

The diversity of data resources, domains and applications in the Euro-BioImaging collection in BioSharing matches the range of application domains served by Euro-BioImaging. The listed resources have been built over the last 10 years and span different levels of scale, functionality and technical capability and suggests that, today, the creation of a single, central repository that holds all image data is not yet possible or practical. Regardless, the breadth of resources highlights the importance of public image resources across a broad cross-section of the technologies and communities served by Euro-BioImaging. Moreover, along with these full submissions, the announcement of our survey brought several less formal submissions, usually via email, naming public resources that make image datasets available for viewing and download. We have included these resources as suggestions in Table 1, and as soon as the related background data is complete and approval by owner is provided, records for them appear in BioSharing.

In all these submissions, we have not yet identified a single comprehensive resource that catalogues public image data resources. Instead it appears there are lists known by different fields, and by different members of each field. This highlights the need for continued updates to the Euro-BioImaging Image Data Catalogue and for the future commitment of resources to expanding and maintaining this resource, so that members of Euro-BioImaging's scientific community can find, use and where appropriate integrate these resources into their own work.

5. Table 1. Summary of Image Data Repositories reported in the Community Survey

Full name of image data repository	Homepage/URL
A Image Storage Platform for Analysis Management and Mining (ISPAMM)	https://gbw-s-omero01.luna.kuleuven.be
Allen Brain Atlas	http://www.brain-map.org/
Atlases - Pathology images	http://atlases.muni.cz
Berkeley Drosophila Transcription Network Project	http://bdtnp.lbl.gov/Fly-Net/
BioEmergences	http://www.bioemergences.eu
Bioimaging	http://bioimg.weizmann.ac.il
BioStudies	http://www.ebi.ac.uk/biostudies/
Broad Bioimage Benchmark Collection (BBBC)	http://www.broadinstitute.org/bbbc/
The Cancer Digital Slide Archive	http://cancer.digitalslidearchive.net/
CELL Image Library*	http://www.cellimagelibrary.org
Curie Image Database (CID)	https://cid.curie.fr
Educational and Research PACS	http://www.medimed.cz
eMouseAtlas*	https://emouseatlas.org
Haeckaliens	http://www.gabygmartins.info/research/haeckaliens
Image Data Repository (IDR)	http://idr-demo.openmicroscopy.org
Liverpool CCI OMERO	http://cci02.liv.ac.uk/gallery/
Liverpool CCI OMERO Gallery	http://cci02.liv.ac.uk/gallery/
Mammography Image Analysis Society Database	https://www.repository.cam.ac.uk/handle/1810/250394
Microenvironment Perturbagen (MEP) LINCS Center image server	https://meplincs.ohsu.edu
MitoCheck	http://www.mitocheck.org
Neurovault*	http://neurovault.org
OpenfMRI*	http://www.openfmri.org
Pathology images	http://atlases.muni.cz
SSBD	http://ssbd.qbic.riken.jp
Structural and functional MRI data	https://www.ceitec.eu/mafil
SYSGRO: A resource of fission yeast phenotypic data & analysis	http://smc.sysgro.org/
The Human Protein Atlas*	http://www.proteinatlas.org
The JCB DataViewer*	http://jcb-dataviewer.rupress.org
ViBE-Z: The Virtual Brain Explorer for Zebrafish	http://vibez.informatik.uni-freiburg.de
A Image Storage Platform for Analysis Management and Mining (ISPAMM)	https://gbw-s-omero01.luna.kuleuven.be
Medical Imaging Databank Valencia Region (BIMCV)	https://ceib.cipf.es/bimcv
Brain Imaging Data Bank of the National Brain Imaging Network of Portugal	www.brainimaging.pt
BIRN / XNat	http://www.birncommunity.org/resources/data/

National Alliance for Medical Image Computing (NAMIC)	http://www.insight-journal.org/midas/community/view/17
NLM: Imaging Methods Assessment and Reporting	http://www.insight-journal.org/midas/community/view/15
100 Healthy Brain MRI: 18-90 years old	http://www.insight-journal.org/midas/community/view/21
University of South Florida Digital Mammography	http://marathon.csee.usf.edu/Mammography/Database.html
Mini-MIAS (Mammographic Image Analysis Society)	http://peipa.essex.ac.uk/info/mias.html
ELCAP public database of whole lung CT images	http://www.via.cornell.edu/databases/lungdb.html
Public Lung Database To Address Drug Response	http://www.via.cornell.edu/databases/crpf.html
The Osteoarthritis Initiative	https://oai.epi-ucsf.org/datarelease/
Open Access Series of Imaging Studies - brain images	http://www.oasis-brains.org/
Alzheimer's Disease Neuroimaging Initiative	http://adni.loni.usc.edu/

Table 1. List of image data repositories for which details have been acquired. *For these repositories, a record already existed in Biosharing and was added to the Euro-Biolmaging collection.