Developing a federated XNAT portal for medical image datasets

Kranthi Thej Kandula

Molecular Imaging Center, Department of Molecular Biotechnology and Health Sciences, University of Torino
Project team members and affiliations

- **Dario Longo**, Institute of Biostructures and Bioimaging (IBB), Italian National Research Council (CNR)

- **Kranthi Thej Kandula**, Molecular Imaging Center (CIM), Department of Molecular Biotechnology and Health Sciences, University of Torino (UniTO)

- **Walter Dastrù**, Molecular Imaging Center (CIM), Department of Molecular Biotechnology and Health Sciences, University of Torino (UniTO)

- **Jean-Karim Hériche**, Cell Biology & Biophysics Unit, European Molecular Biology Laboratory (EMBL)

- **Yi Sun**, Cell Biology & Biophysics Unit, European Molecular Biology Laboratory (EMBL)
Project scientific and technical background

• Preclinical imaging facilities generate large volumes of biomedical image datasets.
• These centers usually lack of standard platforms for managing, storing and sharing image data.
• Datasets are usually stored in personal computers or local databases making it not findable and shareable.
• Local resources can be accessed only by authorized users.
• Several platform have been developed for clinical imaging, among these XNAT is an open source imaging informatics platform developed by the Neuroinformatics Research Group at Washington University.
• Thanks to its extensibility, XNAT can be used to support a wide range of imaging-based projects.

1 PID5815/Integration of EU-wide cardiovascular datasets, Adriano Barbosa da Silva - CORBEL project
2 https://www.health-ri.nl/services/xnat, BBMRI-NL and Lygature TraIT
Project goals

Enable data discovery and reuse of medical image datasets by:

- Deploying a federated XNAT to sync medical imaging datasets from local XNAT installations.

- Annotating image datasets by adding metadata/custom variables to allow findability of these images.
Federation means the action of forming states or organizations into a single group with centralized control.

For implementing a federated XNAT we used the following XNAT instances (HPC4AI with very high computational and storage capacity as “Central” and two other instances as “local”).

We used and adapted the Xsync plugin. The current development includes communication between XNAT instances and syncing the projects from any instance to the centralized XNAT instance. Since there are some limitations of this plugin, we customized the plugin in such a way that it can accept the connection between different versions (1.7.6, 1.8.1) and installation types (vagrant, docker) of XNAT.
Xsync plugin enables automatic synchronization of image datasets from a project in one XNAT instance to a project in a second instance. Xsync is configurable to ensure that only the desired data is delivered, and if required, data is properly de-identified, and that it is delivered on a pre-set schedule.

Deploy Xsync Plugin on Source XNAT.

Create a Project on the Destination/Central XNAT.

Login to Source XNAT as a Source Project owner
- Navigate to Manage Tab on the project report page.
- From the Manage Tab, navigate to Xsync Configuration
- Click on Begin Configuration

If Anonymize Images was selected, enter the desired XNAT Anonymization script. A sample script is here.
We can monitor the sync status.

Log file with sync details is saved as a Source Project resource file. These files are never synced.
Comparing Source and destination instances after Syncing

Central Instance (HPC4AI)
To keep track of which local XNAT instance is feeding the central XNAT instance, or to know which node/imaging center provided a specific image datasets, we added a new custom variables (imaging facility, EUBI node etc..) to keep track of this information.

We made some mandatory fields like Institution name, EUBI Node etc.. to make sure the owner of the project provides during creation of the project at the Central XNAT instance.

When we combine and use the two plugins that we developed it will be possible to sync the information from the local instances (source) to the Destination (Central XNAT) and also to have the precise information about each project that is being synced. Using our metadata which makes it easy to search and filter projects.
Findable and Reusable (FaiR) image datasets require annotation: definition of minimal metadata information to annotate image datasets.

As a proof of concept we defined a first list of 30 metadata describing several aspects of an image dataset, including study information, institution, linked publication, disease, location, species, number of subjects, imaging modality (MRI/PET/SPECT/CT/US/X-ray...), therapy regimen, etc..

To achieve this, we

• first created and deployed a new running plugin for the XNAT instance
• implemented a new UI to add such metadata information when creating a new project
• modified the UI by adding a new query box for looking into the metadata table

We first developed a prototype and then implemented these functionalities into the dedicated XNAT plugin.
Metadata list

- **Study Information**
  1. **Institution**: University of Torino
  2. **Imaging Facility**: Molecular Imaging Center
  3. **EuBI node**: Multi-sited Multi Modal Molecular Imaging MMMI
  4. **Pl**: Dario Longo
  5. **Study name**: Dichloroacetate tumor acidosis MRI-CEST
  7. **doi**: 10.3892/ijo.2017.4029

- **Demographics**
  1. **Species**: mouse
  2. **Strain**: Balb/C
  3. **Genotype**: homozygote
  4. **Gender**: male
  5. **Groups**: control, treated
  6. **Number of mice**: 10, 8

- **Medical**
  1. **Disease model**: cancer
  2. **Organ**: breast
  3. **Tumor model**: xenograft
  4. **Site of engraftment**: subcutaneous

- **Biological**
  1. **Tumor cell line**: TS/A
  2. **Protein/Enzyme/Receptor**: Pyruvate dehydrogenase kinase (PDK)

- **Chemical**
  1. **Intervention substance**: metabolism inhibitor
  2. **Drug**: Dichloroacetate
  3. **Dose**: 100 mg/kg/day
  4. **Administration route**: oral
  5. **Contrast agent**: iopamidol

- **Imaging**
  1. **Imaging technique**: Magnetic Resonance Imaging
  2. **Imaging method**: CEST
  3. **Scanner**: 7 Tesla
  4. **Scanner vendor**: Bruker
  5. **Time points**: 2 (pre, post 15 days)
New UI for adding metadata in new projects

<table>
<thead>
<tr>
<th>Institution Name</th>
<th>Imaging facility</th>
<th>EuBi node</th>
<th>Study Name</th>
<th>Publication</th>
<th>DOI</th>
</tr>
</thead>
</table>

**Study Information**

**Demographics**

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
</table>
New Search projects button based on Metadata information
XNAT currently contains 2 projects, 1 subjects, and 1 imaging sessions.

<table>
<thead>
<tr>
<th>Projects</th>
<th>Subjects</th>
<th>MR</th>
<th>PET</th>
<th>CT</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Keywords</td>
<td>Title</td>
<td>Investigator</td>
<td>Keywords</td>
<td>Metadata</td>
</tr>
</tbody>
</table>

- **Projects:**
  - **Project Title:**
    - ProjectID: ProjectID
    - Project Description
    - You are an owner for this project.
  - **Project_Title2**
    - ProjectID: ProjectID
    - Project Description
    - You are an owner for this project.
  - **Project_Title3**
    - ProjectID: ProjectID
    - Project Description
    - You are an owner for this project.

- **Recent Data Activity:**

**Powered by XNAT**

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824087*
CHALLENGES

• We faced some challenges during API development that accepts the metadata and making it searchable. The reason is because of the existing XNAT algorithm creates tables and API according to the defined classes. Till now there is no datatype which let us add the custom fields at the project level.

• We tried to fix it or to look for an alternative solution, at the end we contacted Rick Herrick, Lead Developer of XNAT, who confirmed that adding metadata at project level is extremely difficult right now but that dedicated APIs will be implemented in the next XNAT release end of September 2022)

• Below is the response from Rick:

```java
XnatProjectData project = new XnatProjectData();
project.setCustomFields("\"institutionName\": \"xxx\", \"imagingFacility\": \"yyy\", \"studyName\": \"zzz\")
project.set(true, false, false, true, false);
```

I think the second alternative would be the best way to do this: it’ll be very easy to query the database for projects matching some desired set of attributes, the values are stored directly with the project entry, and soon the XNAT search engine will know how to work with these JSON attributes and it’ll be easy to locate data just from the values stored in there.
OVERALL Achievements

• Implementation of a federated XNAT across several local nodes and XNAT versions/instances.

• Implementation and deployment of a XNAT plugin to store metadata information.

• Development of a new UI for adding metadata information at project level and of a new query box for searching such information at metadata level to populate the information of specific projects which has the metadata/keywords that matched to the search keys.
NEXT STEPS

• To evaluate whether deploy the XNAT instance in the Embassy Cloud or keep the central XNAT inside the High Computing Center (HPC4AI) hosted by the EuroBioImaging med-hub (for long-term hosting/service)

• To finalize the plugin once the new XNAT version will be released (Q3-2022)

• To share the plugin to various institutions that are interested to collaborate and ask them to sync their image datasets, to test the efficiency of the system for the large scale synchronization.
A1209
Developing a federated XNAT portal for medical image datasets

Thank you!

kranthithej.kandula@unito.it