

# Winter School Day 1: Introduction to AI in medical imaging

Ivan Bocharov, dr. ir. Stefan Klein

The aim of this session is to give a high-level introduction to using AI in medical imaging. We provide an example of a principled workflow starting from managing your dataset in XNAT (a specialized imaging informatics platform for imaging-based research) to image preprocessing and machine learning model application.

## Tutorial information

We offer this tutorial in the form of a Jupyter notebook. This tutorial heavily relies on using an XNAT server. We leave the choice of which XNAT to use to the participant. A local, self-hosted XNAT instance gives you more flexibility and control. The benefit of a central server is that you don't have to install anything locally. During the tutorial session we suggest that participants use the [central XNAT](#). Nevertheless, we encourage the participants to install XNAT locally afterwards and try all its functionality (e.g. creating projects, uploading data, etc.) since this functionality is not available to the guest users of XNAT.

This tutorial consists of three parts:

- High-level introduction to XNAT
- Two Jupyter notebooks:
  - XNAT and image preprocessing
  - Using XNAT in machine learning workflows

## License information

The code and the data used in this tutorial can only be used by the participants of AI in Medical Imaging Winter School. The data (including pretrained machine learning model) can only be used by the participants for the duration of the winter school.

## Acknowledgements

We would like to thank dr. Bo Li and dr. Esther Bron for their help with creating this tutorial and providing the pretrained segmentation model.

## Installing the tutorial dependencies

In order to run this tutorial, you have to use Python 3.7+. We recommend installing and running this tutorial locally using a Python virtual environment. We suggest using [virtualenv](#) for managing virtual environments. Installation instructions for virtualenv can be found [here](#). After installing virtualenv, create a virtual environment by running `virtualenv venv` and [activate it](#). Finally, run `pip install -r requirements.txt` to install all requirements needed for running the Jupyter notebooks. Confirm that the installation worked correctly by running the command `jupyter notebook` in the terminal.

This tutorial can also be installed when using conda distributions: `conda install --file requirements.txt`.

### (Optional) Installing XNAT

1. Install Docker and Docker Compose (usually bundled with Docker). You may find the installation instructions here: <https://docs.docker.com/get-docker/>
2. Run ``docker-compose pull && docker-compose``
3. After these commands finished running, run ``docker-compose up``
4. Navigate to <http://localhost:8080>. The XNAT login page should open. Input default username and password (admin:admin) into the login form and you should be greeted by the XNAT setup page. Change the email to something meaningless (e.g. a@b), disable SMTP and finish the setup. You should have full access to your local XNAT server!

Once you have all the dependencies installed and ready, it is time to start the tutorial.

## Introduction

XNAT is an open source imaging informatics platform developed by the Neuroinformatics Research Group at Washington University. XNAT was originally developed at Washington University in the Buckner Lab, which is now located at Harvard University. It facilitates common management, productivity, and quality assurance tasks for imaging and associated data (from <https://xnat.org/about>).

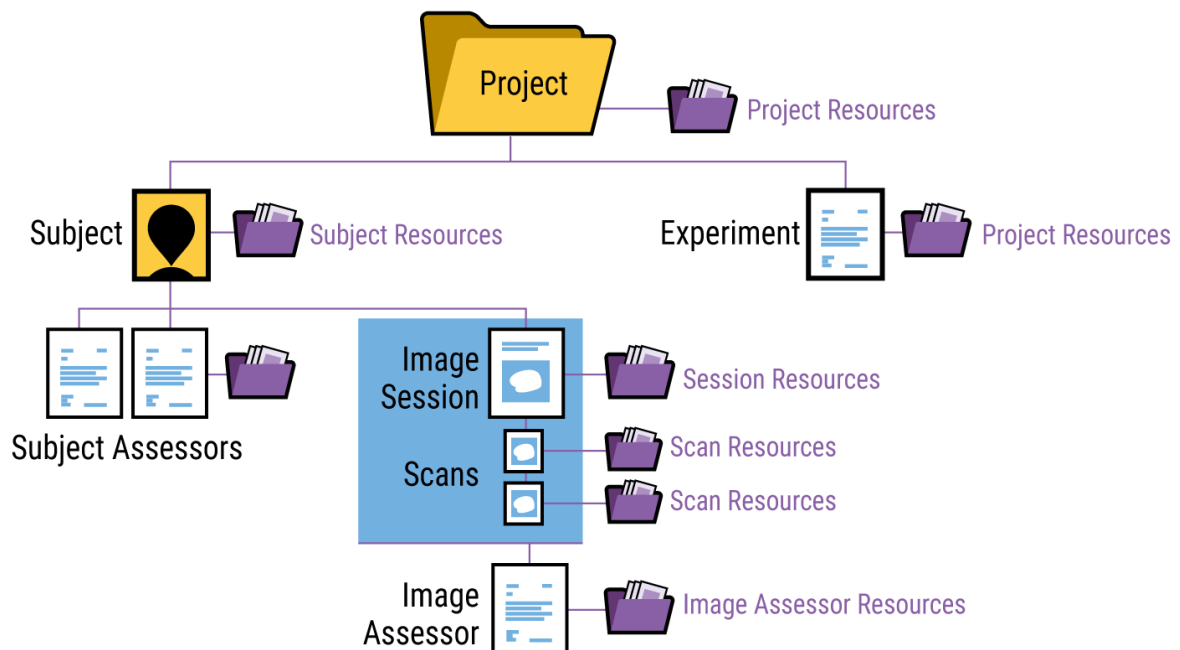
In this session we will discuss basics of XNAT, including data archival and programmatic access to XNAT using XNAT API.

# Data model

XNAT has a hierarchical data model:

1. Projects
2. Subjects
3. Sessions (a specific visit to a scanner)
4. Scans (a specific type of scan during that visit)

You can also store any kind of resources attached to an instance (e.g. metadata about a study at project level) on any level of this hierarchy.



*XNAT Data model<sup>1</sup>*

---

<sup>1</sup> image from: <https://wiki.xnat.org/documentation/how-to-use-xnat/understanding-the-xnat-data-model>

# XNAT overview

Go to <https://xnat.bmia.nl>.

The screenshot shows the XNAT overview page. At the top, there is a navigation bar with 'Browse' and 'Tools' menus. Below this, the XNAT logo is displayed, followed by the text 'a Health-RI platform'. A status bar indicates 'BMIA XNAT currently contains 229 projects, 30184 subjects, and 83884 imaging sessions.' Below this, there are tabs for 'Projects', 'Subjects', 'MR', 'PET', and 'CT'. The 'Projects' tab is active, showing a list of projects. The 'WORC' project is highlighted in the 'Projects' pane. The 'Recent Data Activity' pane shows a table with columns 'worc', 'MR', and 'Lipo-039\_MR'.

Navigate to the **WORC** project at the projects pane in the bottom left corner of the page.

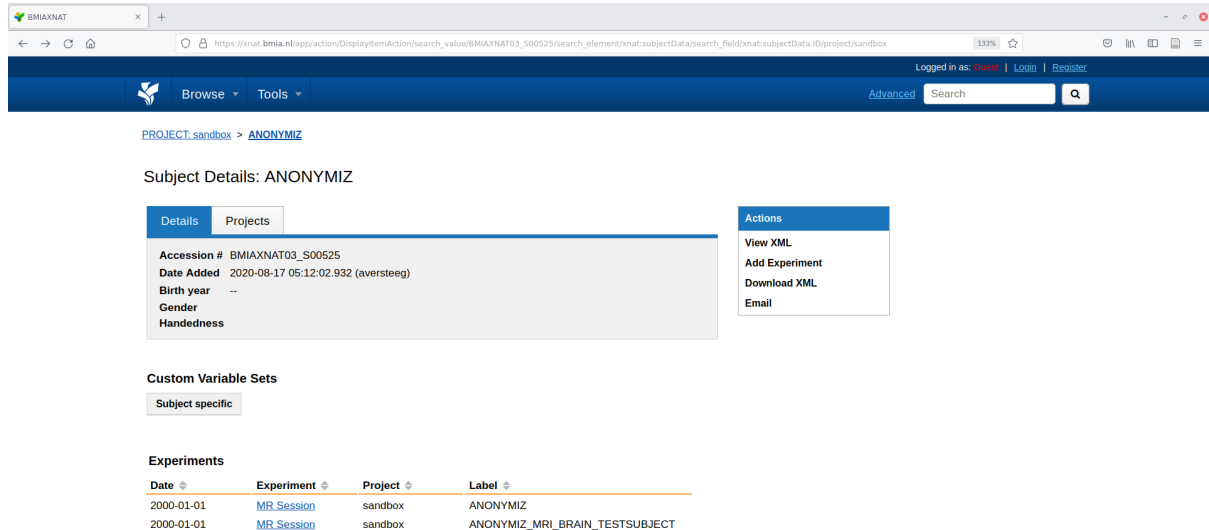
The screenshot shows the XNAT project overview page for the WORC project. The page is divided into three main sections: 'Details', 'Subjects', and 'Actions'. The 'Details' section contains information about the project, including its ID, description, keywords, PI, and investigators. The 'Subjects' section displays a table of subjects with columns for Subject, M/F, Hand, YOB, Index\_Knee, Notes, Age, Sex, Dataset, Diagnosis, Diagnosis\_binary, Phenotype, and MR Sessions. The 'Actions' pane on the right contains buttons for 'Download XML' and 'Download Images'.

| Subject  | M/F | Hand | YOB | Index_Knee | Notes | Age | Sex | Dataset | Diagnosis | Diagnosis_binary | Phenotype | MR Sessions |
|----------|-----|------|-----|------------|-------|-----|-----|---------|-----------|------------------|-----------|-------------|
| CRLM-001 | U   |      |     |            |       | 85  | M   | CRLM    | rHGP      | 0                | nan       | 1           |
| CRLM-002 | U   |      |     |            |       | 49  | F   | CRLM    | rHGP      | 0                | nan       | 1           |

Open the project overview page by clicking on the project's name on the XNAT main page.

This page contains metainformation about the project (PI, Description), the **Subjects** table and **Actions** pane. For each subject, some non-imaging data (e.g. sex and age) might already be shown, and you can see which experiments (i.e., scan sessions) are available for each patient. You can export this data by clicking on the **Options** button, e.g. export to a spreadsheet.

Now go back to the main page of the XNAT, and open the **sandbox** project. This is a project that we use for most of our demos and development. After getting to the project overview, open a subject page by clicking on a subject ID (**ANONYMIZ**) in the **Subjects** table.



The screenshot shows the XNAT web interface. The browser address bar displays the URL: `https://xnat.bmia.nl/app/action/DisplayItemAction/search_value/BMIAxNAT03_500525/search_element/xnat:subjectData/search_field/xnat:subjectData.ID/project/sandbox`. The page title is "Subject Details: ANONYMIZ".

On the left, there are two tabs: "Details" (selected) and "Projects". The "Details" tab shows the following information:

- Accession # BMIAxNAT03\_S00525
- Date Added 2020-08-17 05:12:02.932 (aversteeg)
- Birth year --
- Gender --
- Handedness --

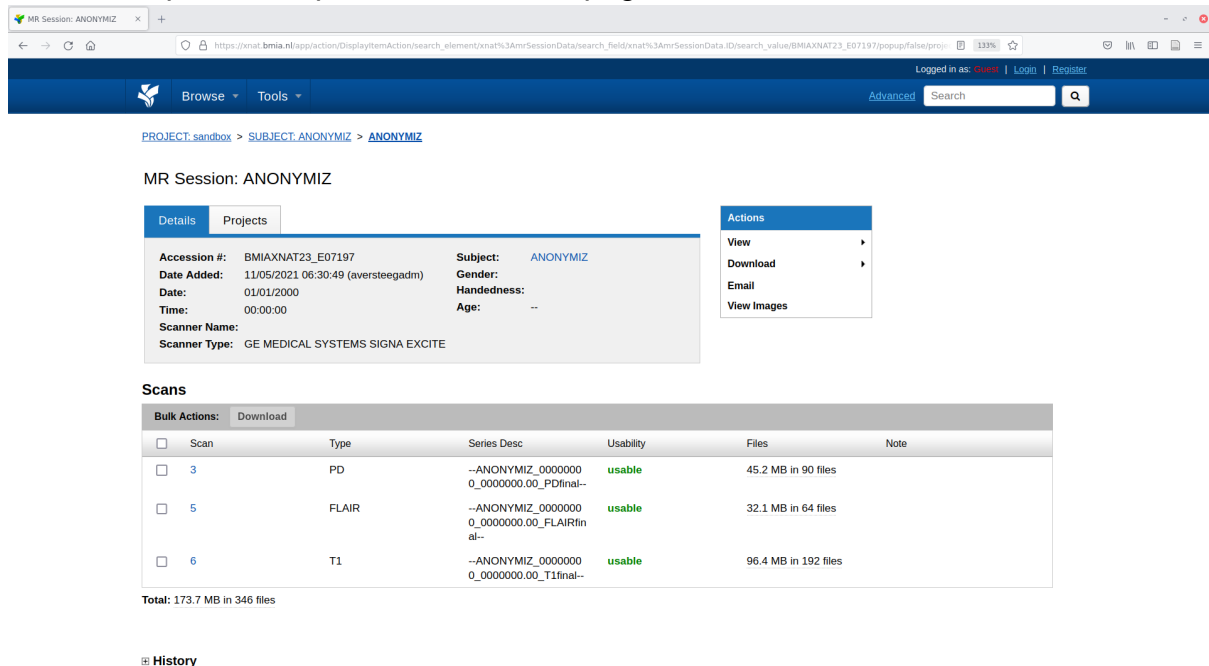
On the right, there is an "Actions" menu with the following options: View XML, Add Experiment, Download XML, and Email.

Below the details, there is a section for "Custom Variable Sets" with a tab for "Subject specific".

At the bottom, there is a table for "Experiments":

| Date       | Experiment                 | Project | Label                          |
|------------|----------------------------|---------|--------------------------------|
| 2000-01-01 | <a href="#">MR Session</a> | sandbox | ANONYMIZ                       |
| 2000-01-01 | <a href="#">MR Session</a> | sandbox | ANONYMIZ_MRI_BRAIN_TESTSUBJECT |

At the top of the page you can see the subject's details (e.g. internal accession number and (anonymized) year of birth). At the bottom of the page you can find all experiments (i.e. imaging sessions) associated with the subject. Click **MR Session**, which will open the experiment overview page.



The screenshot shows the XNAT web interface for the "MR Session: ANONYMIZ" page. The browser address bar displays the URL: `https://xnat.bmia.nl/app/action/DisplayItemAction/search_element/xnat%3AmrSessionData/search_field/xnat%3AmrSessionData.ID/search_value/BMIAxNAT23_E07197/popup/false/proj...`. The page title is "MR Session: ANONYMIZ".

On the left, there are two tabs: "Details" (selected) and "Projects". The "Details" tab shows the following information:

- Accession #: BMIAxNAT23\_E07197
- Date Added: 11/05/2021 06:30:49 (aversteegadm)
- Date: 01/01/2000
- Time: 00:00:00
- Scanner Name: GE MEDICAL SYSTEMS SIGNA EXCITE
- Scanner Type: GE MEDICAL SYSTEMS SIGNA EXCITE

On the right, there is an "Actions" menu with the following options: View, Download, Email, and View Images.

Below the details, there is a section for "Scans". It includes a table with columns: Scan, Type, Series Desc, Usability, Files, and Note.

| Scan                       | Type  | Series Desc                                  | Usability | Files                | Note |
|----------------------------|-------|--|-----------|----------------------|------|
| <input type="checkbox"/> 3 | PD    | --ANONYMIZ_0000000_0_0000000.00_PDfinal--    | usable    | 45.2 MB in 90 files  |      |
| <input type="checkbox"/> 5 | FLAIR | --ANONYMIZ_0000000_0_0000000.00_FLAIRfinal-- | usable    | 32.1 MB in 64 files  |      |
| <input type="checkbox"/> 6 | T1    | --ANONYMIZ_0000000_0_0000000.00_T1final--    | usable    | 96.4 MB in 192 files |      |

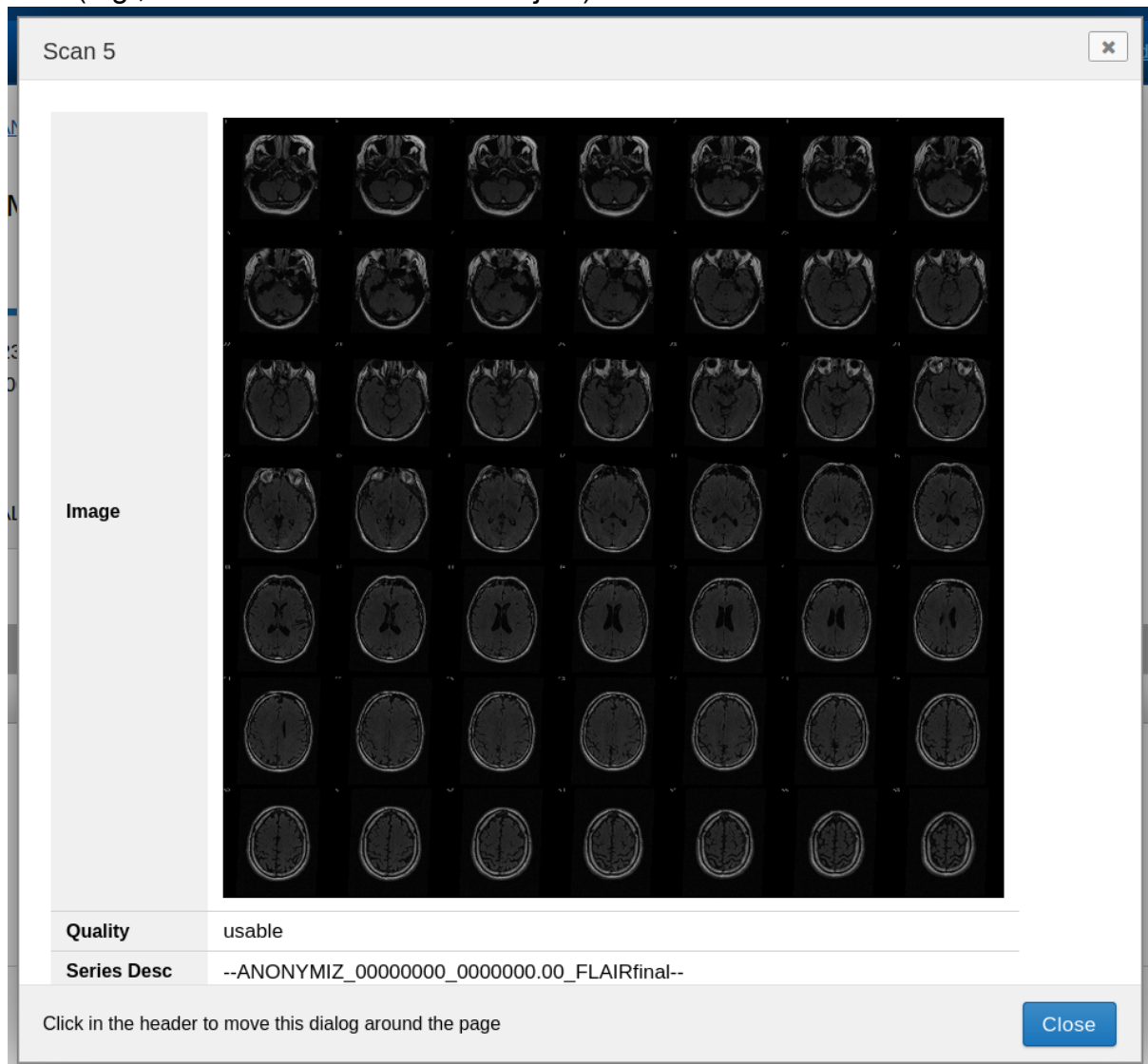
Total: 173.7 MB in 346 files

At the bottom, there is a link for "History".

The **Details** give you the accession number of the scan, the upload date and time, the date and time of scanning, the scanner, and, if available, subject information

such as gender and age. Another useful feature here is the **Quality**, which can be set to usable, questionable, and unusable. This will not affect how the image is stored in the XNAT, but it gives a good visual cue in the scans overview table.

Under **Scans** you can inspect each of the uploaded scans, which, in our case contains one scan and an annotation of that scan. Click on the name of an actual scan (e.g., **5** in case of ANONYMIZ subject).



A scan preview popup should open after clicking on the scan name. This window shows a preview of all slices of the image as well as more detailed information of the scan such as field strength and others. Scroll down and click on “View **DICOM Headers**”.

| The DICOM tags for the first dicom file in this scan. |                                 |   |
|---|---------------------------------|---|
| Tag   | Description                     | Value   |
| (0002,0001)   | File Meta Information Version   | 00\01   |
| (0002,0002)   | Media Storage SOP Class UID     | 1.2.840.10008.5.1.4.1.1.4                                   |
| (0002,0003)   | Media Storage SOP Instance UID  | 1.3.6.1.4.1.40744.99.96774141932336415037960163561501719467 |
| (0002,0010)   | Transfer Syntax UID             | 1.2.840.10008.1.2   |
| (0002,0012)   | Implementation Class UID        | 1.2.40.0.13.1.1   |
| (0002,0013)   | Implementation Version Name     | dcm4che-2.0   |
| (0002,0016)   | Source Application Entity Title | CTP   |
| (0008,0005)   | Specific Character Set          | ISO_IR 100  |
| (0008,0008)   | Image Type                      | ORIGINAL\PRIMARY\OTHER                                      |
| (0008,0016)   | SOP Class UID                   | 1.2.840.10008.5.1.4.1.1.4                                   |
| (0008,0018)   | SOP Instance UID                | 1.3.6.1.4.1.40744.99.96774141932336415037960163561501719467 |
| (0008,0020)   | Study Date                      | 20000101  |
| (0008,0021)   | Series Date                     | 20000101  |
| (0008,0022)   | Acquisition Date                | 20000101  |
| (0008,0023)   | Content Date                    | 20000101  |
| (0008,0030)   | Study Time                      | 000000  |
| (0008,0031)   | Series Time                     | 000000  |
| (0008,0032)   | Acquisition Time                | 000000  |
| (0008,0033)   | Content Time                    | 000000  |
| (0008,0060)   | Modality                        | MR  |
| (0008,0070)   | Manufacturer                    | GE MEDICAL SYSTEMS  |
| (0008,0090)   | Referring Physician's Name      |   |
| (0008,1030)   | Study Description               | MRI BRAIN TESTSUBJECT                                       |
| (0008,103E)   | Series Description              | --ANONYMIZ_00000000_00000000.00_FLAIRfinal--                |
| (0008,1090)   | Manufacturer's Model Name       | SIGNA EXCITE  |

After pressing **View DICOM Headers** XNAT will display all the information contained in the DICOM headers of the first DICOM file in the scan. This might be useful for quick verification and visual inspection.

Now click on **View Images** in the Actions pane. An image viewer application should open.



Click on the scan (number 5) on the left side of the viewer. The image preview should open on the right side of the viewer.

You can visualize multiple images at the same time by clicking one of the small triangles and dragging an image in the new panel. You can scroll through the images with the slider on the bottom of the panel. If the images are not well visible, try to

increase the laptop's or computer's screen brightness.

We will not discuss all viewing features, but the three most important are:

1. The brightness/thresholds can be adjusted by hovering with the mouse over the chart on the right.
2. The layout and 2D/3D perspective can be changed by clicking on the top button of the center menu, between the center and left panel.
3. The image can be rotated with the left mouse button when in 3D mode. 3D mode is selectable via the previous option.

When you're done, close the viewer, which will bring you back to the MR session.

Switch to the **Notebook 1 (XNAT and Image preprocessing)**.