



**DATA SCIENCE
INSTITUTE®**
AMERICAN COLLEGE OF RADIOLOGY

ACR AI-Lab – MGH Pilot

Romane Gauriau
Sr Machine Learning Scientist, MGH & BWH Center for Clinical Data Science (CCDS)

Goal of the pilot

- The Primary Aim of this study is to determine whether the software can be successfully deployed at multiple institutions to allow them to download an AI model developed at another institution, add new data from their own institution, update the model, and share the results.
- Secondary Aim of this study is to compare the difference in performance of a Breast Density model before and after the model has been updated with images from the pilot sites.



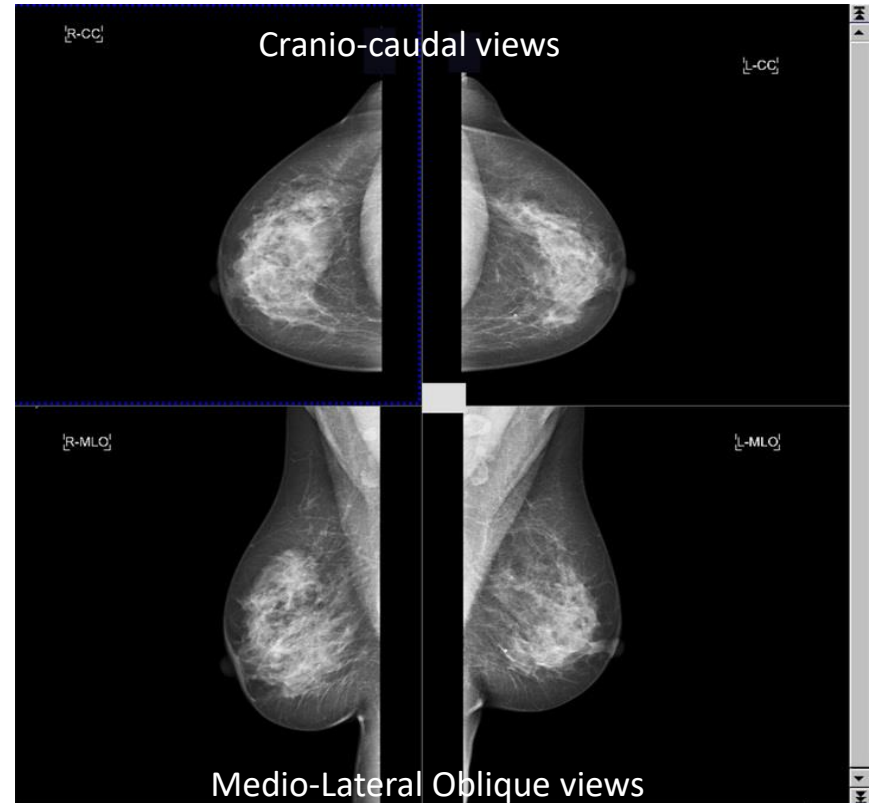
Program plan

| Task | Owner | Institution |
|--|--|-------------------|
| IRB submission | Ali Pourvaziri | CCDS |
| Online Lab notebook | Ali Pourvaziri | CCDS |
| Provision Hardware | Adam McCarthy | CCDS |
| Enable user access | Adam McCarthy | CCDS |
| Install SW | Stephanie Bossong, Nir Neumark | ACR, CCDS |
| Identify data set, 2D Mammography Studies | Nir Neumark, Jayashree Kalpathy-Cramer | CCDS, Martinis |
| Request data pull | Min Yun | CCDS |
| Pull Data | Sean Doyle | CCDS |
| Select series and annotate | Nir Neumark, Ken Chang, Jayashree Kalpathy-Cramer | CCDS Martinis |
| Run model | Romane Gauriau | CCDS |
| Collate results | Romane Gauriau | CCDS |



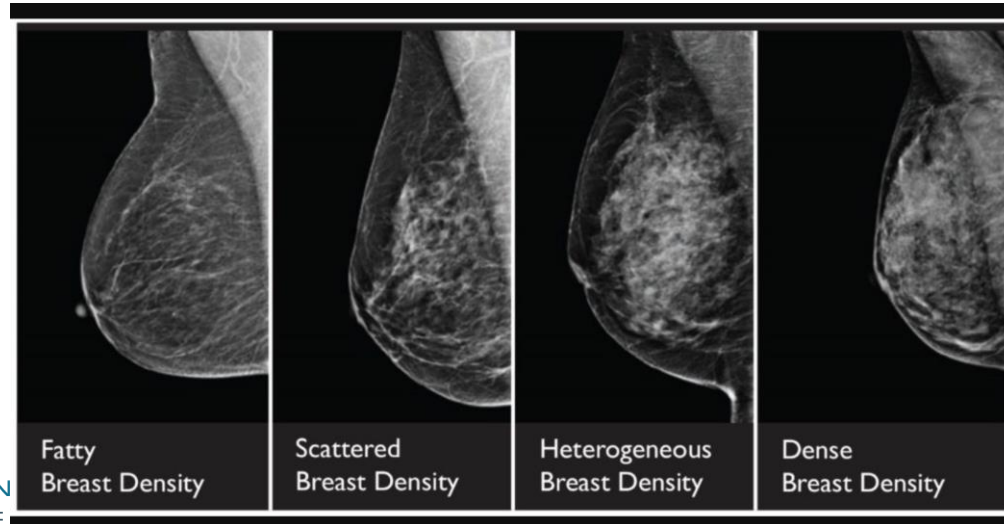
Use case – Breast Density in 2D mammography

- Mammogram:
low-dose X-ray of the breasts
- Used as a diagnostic
and screening tool
- Early tumors can be detected
- Different views with different
angles are performed



Use case – Breast Density in 2D mammography

- Each examination was interpreted by a radiologist
- The mammograms were annotated using ACR BI-RADS breast density lexicon (fatty, scattered, heterogeneously dense, dense)



Pilot script

Off the platform

- Data pull
- Cohort creation
- Create excel file with list of studies and annotations
- Copy data on the local machine where the platform is installed

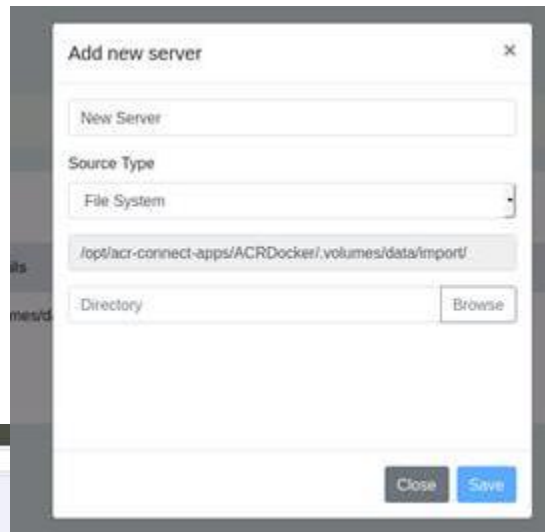
On the platform (Graphical Use Interface)

- **Import DICOM studies in the platform**
- **Prepare dataset**
- Evaluate ACR model
- Train new model on MGH data
- Evaluate new model on MGH data



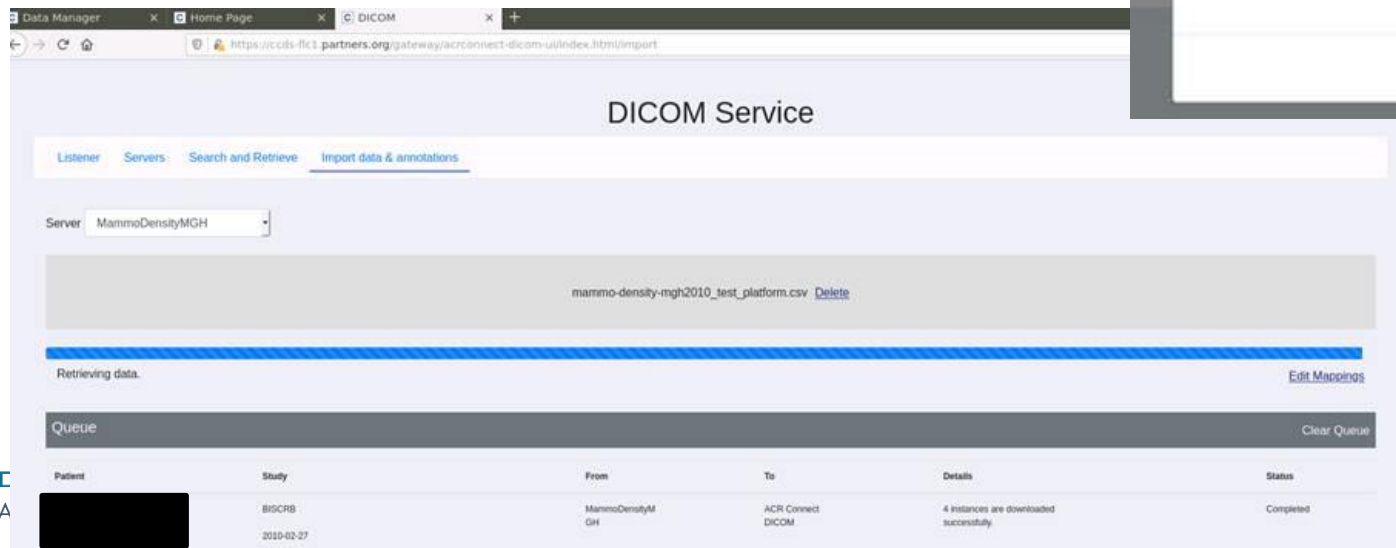
DICOM import

- Copy the data on the local machine
- Configure the location on the platform
- Load excel file with the list of studies (+annotations) to create the cohort



The 'Add new server' dialog box contains the following fields and controls:

- New Server:** A text input field.
- Source Type:** A dropdown menu currently set to 'File System'.
- Directory:** A text input field containing the path '/opt/acr-connect-apps/ACRDOCKER/volumes/data/import/'.
- Browse:** A button next to the directory field.
- Close:** A grey button at the bottom right.
- Save:** A blue button at the bottom right.



The screenshot shows the 'DICOM Service' web interface. The browser address bar displays 'https://ccds-fic1.partners.org/gateway/acrconnect-dicom-ui/index.html#import'. The page title is 'DICOM Service'. The navigation menu includes 'Listener', 'Servers', 'Search and Retrieve', and 'Import data & annotations'. The 'Server' dropdown is set to 'MammoDensityMGH'. A file 'mammo-density-mgh2010_test_platform.csv' is listed with a 'Delete' link. A blue progress bar indicates 'Retrieving data.' with an 'Edit Mappings' link. A 'Queue' section shows a 'Clear Queue' link. At the bottom, a table displays the import progress:

| Patient | Study | From | To | Details | Status |
|------------|----------------------|---------------------|----------------------|---|-----------|
| [REDACTED] | BISCRB 2010-02-27 | MammoDensityM GH | ACR Connect DICOM | 4 instances are downloaded successfully. | Completed |



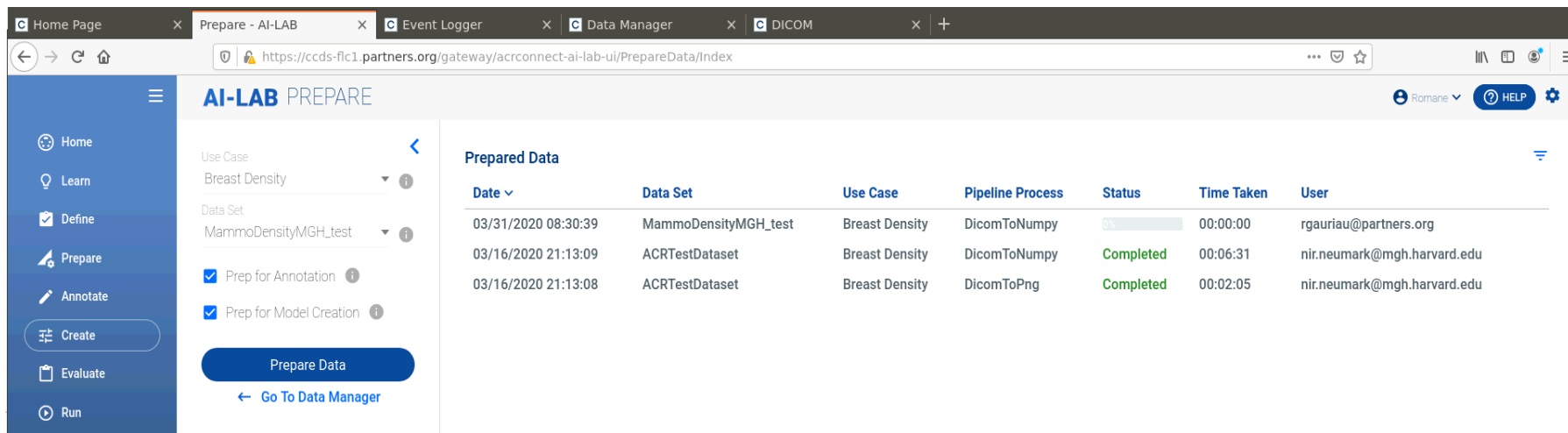
DICOM import

- Easy to setup
- Relatively fast (240 studies loaded in X min)
- Limitations found:
 - Requires enough disk space on the local machine
 - Multithreading was causing issues when loading many studies (>1000 studies)



Data preparation

- Conversion to PNG and to numpy
- For the annotations and for the models

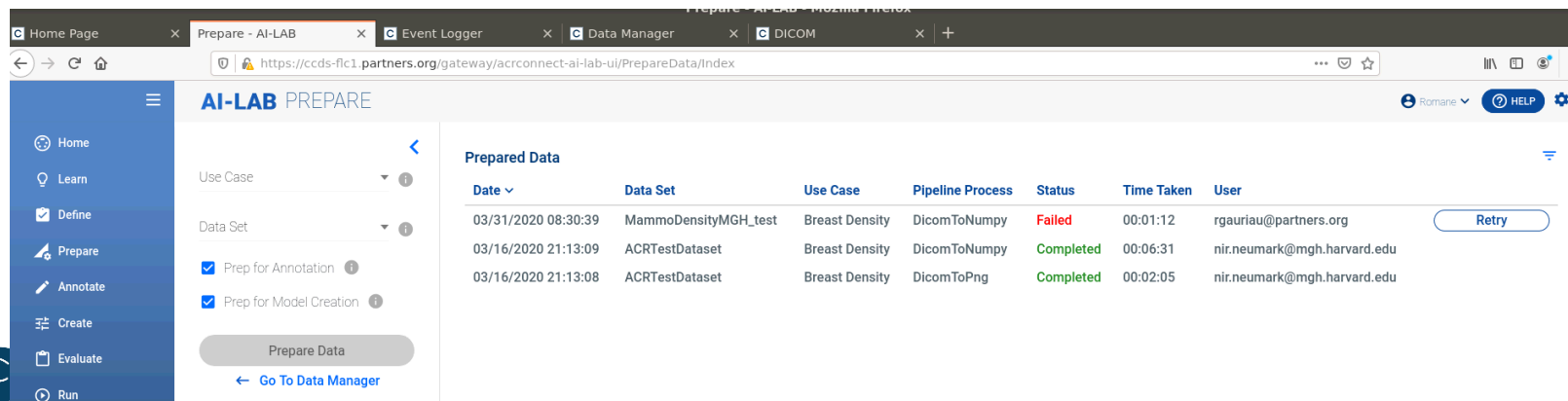


The screenshot displays the AI-LAB PREPARE web application interface. The browser address bar shows the URL: <https://ccds-flc1.partners.org/gateway/acrconnect-ai-lab-ui/PrepareData/Index>. The application header includes the title "AI-LAB PREPARE" and a user profile for "Romane" with a "HELP" button. A left-hand navigation menu contains options: Home, Learn, Define, Prepare, Annotate, Create, Evaluate, and Run. The main content area is titled "Prepared Data" and features a table with the following columns: Date, Data Set, Use Case, Pipeline Process, Status, Time Taken, and User. The table contains three rows of data, with the first row showing a 0% completion status and the others as "Completed".

| Date | Data Set | Use Case | Pipeline Process | Status | Time Taken | User |
|---------------------|----------------------|----------------|------------------|-----------|------------|-----------------------------|
| 03/31/2020 08:30:39 | MammoDensityMGH_test | Breast Density | DicomToNumpy | 0% | 00:00:00 | rgauriau@partners.org |
| 03/16/2020 21:13:09 | ACRTestDataset | Breast Density | DicomToNumpy | Completed | 00:06:31 | nir.neumark@mgh.harvard.edu |
| 03/16/2020 21:13:08 | ACRTestDataset | Breast Density | DicomToPng | Completed | 00:02:05 | nir.neumark@mgh.harvard.edu |

Data preparation

- User-friendly
- Easy to do
- Ran into problems because of some DICOM files that were not images (DICOM SR files had not been filtered from the raw DICOM studies)



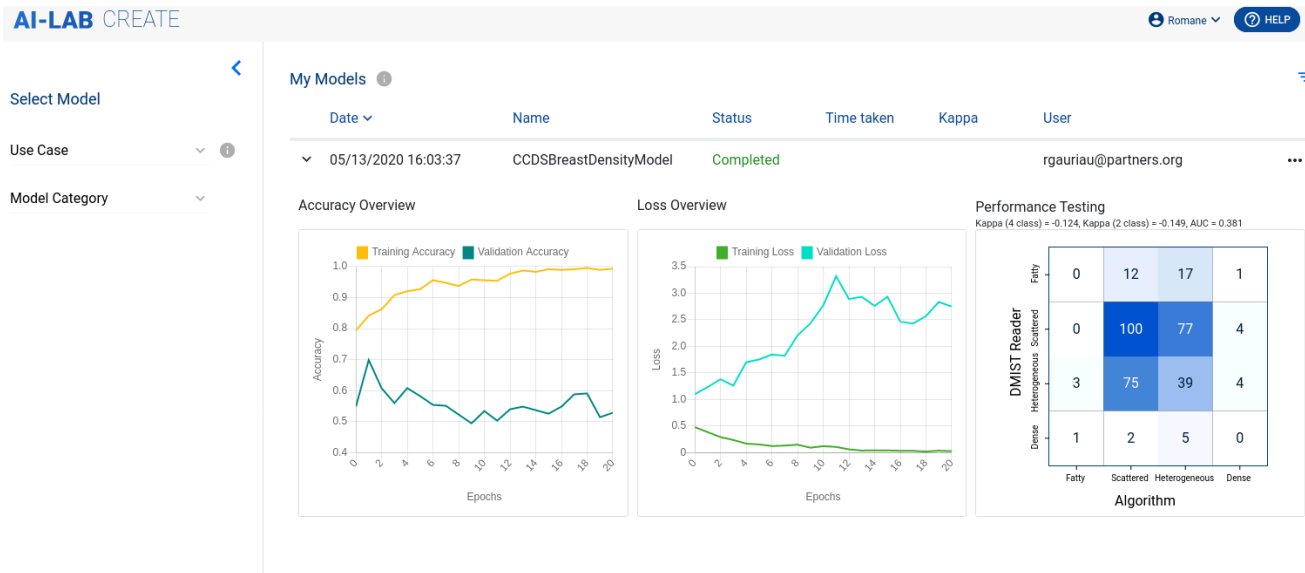
The screenshot displays the AI-LAB PREPARE web application interface. The browser tabs include 'Home Page', 'Prepare - AI-LAB', 'Event Logger', 'Data Manager', and 'DICOM'. The URL is <https://ccds-flc1.partners.org/gateway/acrconnect-ai-lab-ui/PrepareData/Index>. The interface features a left sidebar with navigation options: Home, Learn, Define, Prepare, Annotate, Create, Evaluate, and Run. The main content area is titled 'AI-LAB PREPARE' and includes a 'Use Case' dropdown set to 'Breast Density' and a 'Data Set' dropdown set to 'ACRTestDataset'. There are two checked options: 'Prep for Annotation' and 'Prep for Model Creation'. A 'Prepare Data' button is visible, along with a 'Go To Data Manager' link. Below this, a table titled 'Prepared Data' shows the following entries:

| Date | Data Set | Use Case | Pipeline Process | Status | Time Taken | User | |
|---------------------|----------------------|----------------|------------------|-----------|------------|-----------------------------|-----------------------|
| 03/31/2020 08:30:39 | MammoDensityMGH_test | Breast Density | DicomToNumpy | Failed | 00:01:12 | rgauriau@partners.org | Retry |
| 03/16/2020 21:13:09 | ACRTestDataset | Breast Density | DicomToNumpy | Completed | 00:06:31 | nir.neumark@mgh.harvard.edu | |
| 03/16/2020 21:13:08 | ACRTestDataset | Breast Density | DicomToPng | Completed | 00:02:05 | nir.neumark@mgh.harvard.edu | |



Model training and evaluation

- Straightforward
- Fast
- User-friendly



Conclusion

- User-friendly
- Well adapted for
 - well defined use cases
 - retrain existing models on new data
 - machine learning beginners
- Some limitations:
 - Requires precise data cleaning/curation of the raw DICOM studies
 - ... but problem easy to fix
- Very promising tool for the democratization of AI in clinical environment



Acknowledgements

- **CCDS:** Nir Neumark, Shaimaa Sharaf, Ali Pourvaziri, Adam McCarthy, Min Yun, Sean Doyle, Ram Naidu
- **MGH:** Tarik Alkasab
- **Martinos Center:** Jayashree Kalpathy-Cramer, Ken Chang
- **ACR:** Stephanie Bossong, Laura Coombs, Kris Glassmir, Deepak Kattil Veettil

