

A distributed workflow for managing and processing neuroimaging data





In every large-scale neuroscience project, storing, managing, and processing a large amount of neuroimaging data can be a challenge especially when the data is heterogeneous, and sourced from multiple sites.

Besides getting all the acquired data together in one convenient place and having people from different centers distributed worldwide working on the processing and analysis of the data, one has to deal with the regulations for access control and de-identification for the different locations of the contributing sites. In this whitepaper, we describe one such neuroscience project and we present our solution at QMENTA implemented together with the UCSF School of Medicine, Department of Neurology. In the end, we summarize the benefits of our approach and inform you on how you can also use our platform with minimum effort.



For a large multi-site study on patients with Multiple Sclerosis (MS), a large amount of datasets will need to be uploaded over an extensive period of time (5 years). At the end of the project, it is estimated for the project to contain a total of 3000-4000 MRI datasets, uploaded from 16 sites around the world, including sites in USA and Germany.

The data will be heterogeneous: there is retrospective data, which consists mostly of T1W and in some cases T2W or FLAIR MRI scans, and there will be prospective data, which is mostly T1W and FLAIR. All this data needs to be stored and managed in a unified way.

Because the sites providing the data are located in various countries, we have to adhere to the laws and restrictions of the countries involved. To avoid issues, we always apply rules that are at least as strict as the most strict laws involved. After the data that has been uploaded, it is preprocessed automatically. This pre-processing may include bias-field correction, registration, defacing, and conversion to different file formats such as NIFTI. We use several publicly-available tools for the pre-processing, which are listed in the next section. When the pre-processing is done, a specialist will review the images and approve or disapprove the dataset for inclusion in the study. This manual quality control (QC) step is needed to avoid introducing incorrect data in the study that was acquired with, for example, wrong scanning parameters or if there was any problem during file format conversion. For all approved datasets, the MS lesions need to be segmented, as well as different brain structures so that they can be later used to derive measurements of cortical atrophy, brain degeneration, lesion volume, white matter integrity, etc. The goal is to automate this process as much as possible. Each step however, will still need a QC step where a specialist evaluates and potentially edits the results, and can approve or disapprove the results.



The QMENTA platform allows users from the different sites to collaborate on the project. The project owner can add users, and can set the permissions per user and per site.

Datasets can be easily and conveniently uploaded through the browser, or using our specialized QMENTA app which will automatically anonymize the data before it is actually sent to the platform. The platform stores all data in a secure way by applying proper de-identification and encryption, and complies with the various privacy and security laws in different countries. This is elaborated upon in more detail in our security whitepaper. The datasets will be processed using several open source neuroimaging tools:

- FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL) for registration and defacing in the pre-processing step
- MRtrix3 (http://www.mrtrix.org/) to convert DICOM to NIFTI file format in the pre-processing step
- LST: Lesion Segmentation Tool (http://www.statistical-modelling.de/lst.html) for SPM to automatically segmentation of MS lesions
- Lesion filling is by default our own algorithm, but can optionally use FSL or ANTS
- Freesurfer image analysis suite (http://surfer.nmr.mgh.harvard.edu/) for skull stripping and brain parcellation,
- ANTS: Advanced Normalization Tools (http://stnava.github.io/ANTs/) for bias-field correction (in pre-processing), skull stripping and brain parcellation.

Each of these tools is specialized in one or more neuroimaging tasks that are performed automatically. By making these tools available in our system, we remove the overhead for the user to download, install and configure the tools, and to apply them manually or scripted to many datasets. The users can set the parameters of each of the tools using our platform, but proper default values are chosen when the user does not set them.

All the tools and QC steps are combined together in a single MultipleMS Workflow which is depicted in the diagram below:



Pre-processing will automatically start after a dataset has been uploaded, and after preprocessing is done, the data will be added to the queue for manual quality control (QC). Users that have been assigned a QC role will see a list of datasets to verify when they log onto the platform. Several of the other advanced analyses are followed by a QC step as well. For the QC after lesion segmentation, we use MindControl¹, built on top of the Papaya² medical image viewer for the user to review and edit the detected MS lesions. Using this automatic detection followed by manual QC saves a lot of time compared to full manual detection of lesions. The QC steps after the other analyses consist of a simple approval step where the specialist has to choose PASS or FAIL to confirm or reject the results of an analysis.

1. https://www.ncbi.nlm.nih.gov/pubmed/28365419, https://github.com/akeshavan/mindcontrol

2. https://github.com/rii-mango/Papaya



QMENTA offers a cloud-based neuroimaging platform for hospitals, research centers, and pharmaceutical companies to help accelerate the discovery and development of new treatments for brain diseases. It provides a unique infrastructure with scalable computing power to help save valuable time, effort, and resources on intensive R&D activities.

The QMENTA cloud platform supports the entire R&D workflow, including data collection, data management, image processing, 3D visualization, and sharing. Images and clinical scores can be easily uploaded to the cloud, automatically anonymized, and managed from the browser without the need to install or maintain any software. All data is securely stored in a HIPAA compliant environment with fine-grained access settings. Proprietary, open-source, and licensed tools such as FSL, MRtrix, LST, ANTs and FreeSurfer are offered together. Thus, the standardized computing environment with these validated tools enables reproducibility in research.

By using the QMENTA platform, the data is stored securely in the cloud, so no need to worry about missing CDs or broken hard drives. The data is secure, accessible, and easy to share with other users that have the right permissions. For the MultipleMS project, we have created a new workflow in the QMENTA platform that combines various open source neuroimaging tools and manual quality control steps. Using this workflow on our platform enables researchers and specialists worldwide to effectively and efficiently work together. It helps them to easily view the data uploaded from different sites, automatically starts pre-processing of the data, and ties together a variety of available neuroimaging tools to acquire the desired results. Individual processing steps are followed by manual QC steps by researchers and specialists to guard the overall guality of the project. The combined automatic processing combined with manual QC saves a lot of time when compared to other approaches.



Learn more about all the AI analysis and workflows included in our dedicated neuro disease package for Multiple Sclerosis

LEARN MORE

For more information about MultipleMS go to www.multiplems.eu

US Headquarters

75 State Street, Suite 100 Boston, MA 02109 +1 339 368 8040

EU Offices

C/ Roger de Llúria 46, Pral. 1ª 08009 Barcelona, Spain +34 933 282 007

