

# Advanced Querying for Multi-site and Multi-modal Data in Neuroscience Science



Jordan Stirling<sup>1,2</sup>, Samir Das<sup>1,2</sup>, Leigh MacIntyre<sup>1,2</sup>, Cecile Madjar<sup>3</sup>, Penelope Kostopoulos<sup>1,2</sup>, David MacFarlane<sup>1,2</sup>, Dahlia Sniderman<sup>1,2</sup>, Alan Evans<sup>1,2</sup>

<sup>1</sup>McGill Centre for Integrative Neuroscience, Montreal, Canada, <sup>2</sup>Montreal Neurological Institution, Montreal, Canada, <sup>3</sup>Douglas Mental Health University Hospital, Montreal, Canada

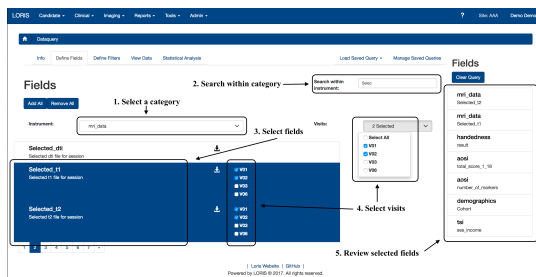
## Introduction

Neuroscience studies have come to depend heavily on neuroinformatics infrastructures that include databasing, data management and curation. The ability to efficiently query data from large-scale, longitudinal studies is a necessity for statistical analysis, versioning, and data mining. LORIS's Data Query Tool (DQT) is an easy-to-use tool integrated directly into the web browser that allows users to create complex queries without needing to know how to create queries for MySQL.

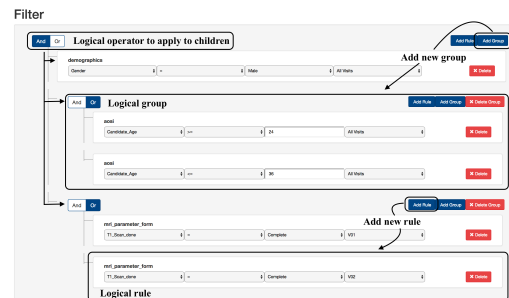
## Example Query

The following example shows how one would run a query for behavioural and imaging data for all males which had an AOSI administered when they were between 24 and 36 months, and who had a scan done at either V01 or V02.

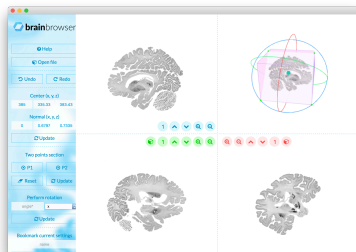
## Workflow to example



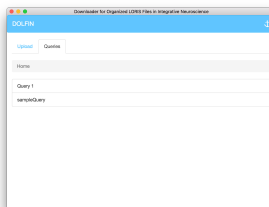
**Figure 1:** The field selector allows the user to select multiple fields from different categories. Users select a category (1), and within the selected categories, users can search for desired fields (2). They must then click on the fields (3) to add them to the query. For any field, users can specify the desired visits (4). A summary of selected fields is displayed on the right (5).



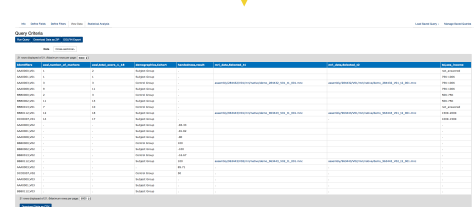
**Figure 2:** The filter builder allows users to apply query logic to get specific subsets of subjects. Users can switch between AND or OR logical operators, and apply additional conditions to compare one subject's resulting subset of data with its siblings' results. Adding groups allows users to utilize nested logic.



**Figure 5:** BrainBrowser is built in to DOLFIN. Users can view an image in oblique planes simply by clicking on the desired image.



**Figure 4:** Using the desktop application DOLFIN (see results), users can download the imaging data from their query. Since DOLFIN is a desktop application, it has the ability to access the user's file system to organize the imaging files. If the connection is interrupted during the download, DOLFIN can later pick up where it left off instead of downloading the entire dataset



**Figure 3:** Once both fields and filters are defined, the user can run their query. The result is a table with the requested data. Users can choose to view the data cross-sectionally or longitudinally, and can download the resulting table as a csv. For imaging data, users have the option to download as a zip file, or to download a JSON file which can be imported to DOLFIN. Users can also save and share their query by clicking on the Manage Saved Queries tab.

## Results

With the new features recently added to the DQT, such as defining visits and advanced query logic, users are able to extract more precise data subsets from the database. Additionally, the implementation of desktop application Downloader For Organized LORIS Files in Integrative Neuroscience (DOLFIN) has streamlined the process of downloading imaging data. Users no longer have to rely on constant connectivity or coordination with the Data Coordinating Centre to get the imaging data from the database.

## Conclusion

The newest iteration of the DQT has streamlined the process of filtering and downloading large sets of data for multi-modal studies. The most common issues in the system have been addressed, and the most time-consuming steps of the process have been simplified to address users' need for fast, easy and shareable access to data.

## References

- Das, S., et al., 2011. LORIS: A web-based data management system for multi-center studies. *Front. Neuroinform.* 5:37. Das, S., et al., 2016. The MNI data-sharing and processing ecosystem. *NeuroImage* 124:1188–1195. MacFarlane, D., et al., 2014. Enhanced Data Querying for Neuroinformatics Databases. OHBM 2014. Wolff, J.J., et al., 2012. Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *Am. J. Psychiatry.* 169(6), 589–600. <http://dx.doi.org/10.1176/appi.ajp.2011.11091447>.