

# Facilitating the use of imaging data repository for hypothesis testing: a tool for creating and querying imaging-data

Samuel Danso<sup>1,2</sup>, Dominic E. Job<sup>1,2</sup>, Rudolph Pienaar, Ellen P. Grant<sup>3</sup> and Joanna M. Wardlaw<sup>1,2</sup>

1. Centre for Clinical Brain Sciences (CCBS), The University of Edinburgh; 2. Scottish Imaging Network, SINAPSE, Glasgow.  
3. Fetal-Neonatal Neuroimaging and Developmental Science Centre, Boston Childrens Hospital, Harvard Medical School

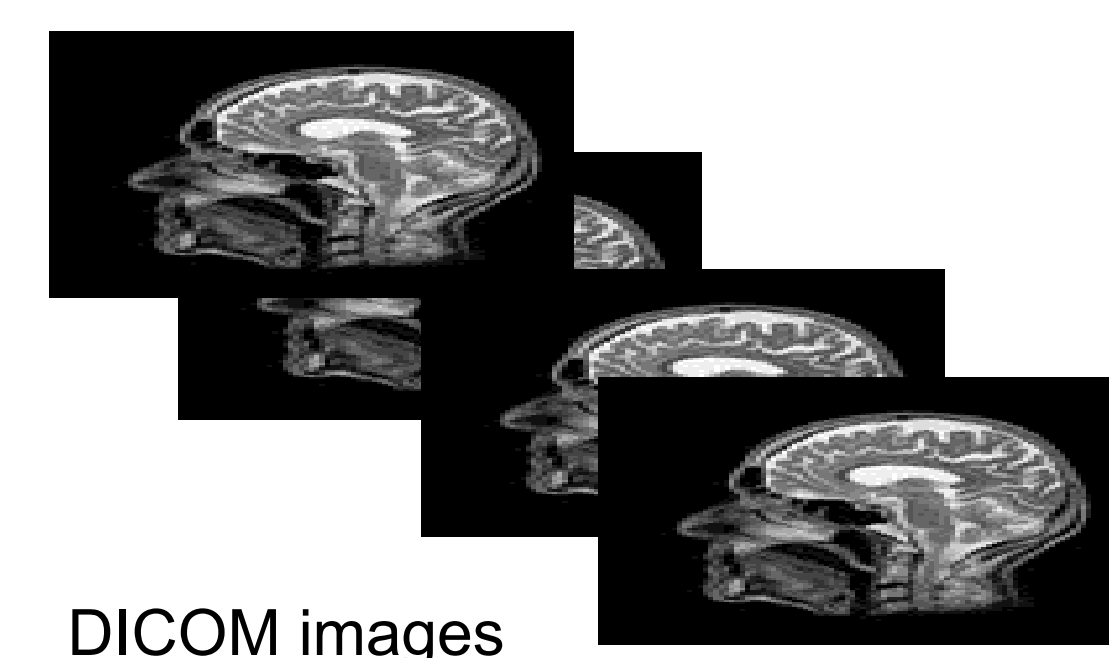
## Background & rationale

- DICOM standards allow image meta-data to be embedded in the header of the image file during image acquisition.
- Researchers with relatively little computing expertise have difficulty querying an imaging data repository particularly when associated data is not readily available.
- We present a framework and a set of python scripts which automatically extracts meta-data contained in the DICOM image files to a CSV and subsequently into MySQL database.
- The work was carried out during a Postdoctoral and Early Career Researcher Exchange funded visit to Harvard Medical School.

## Methods & Results

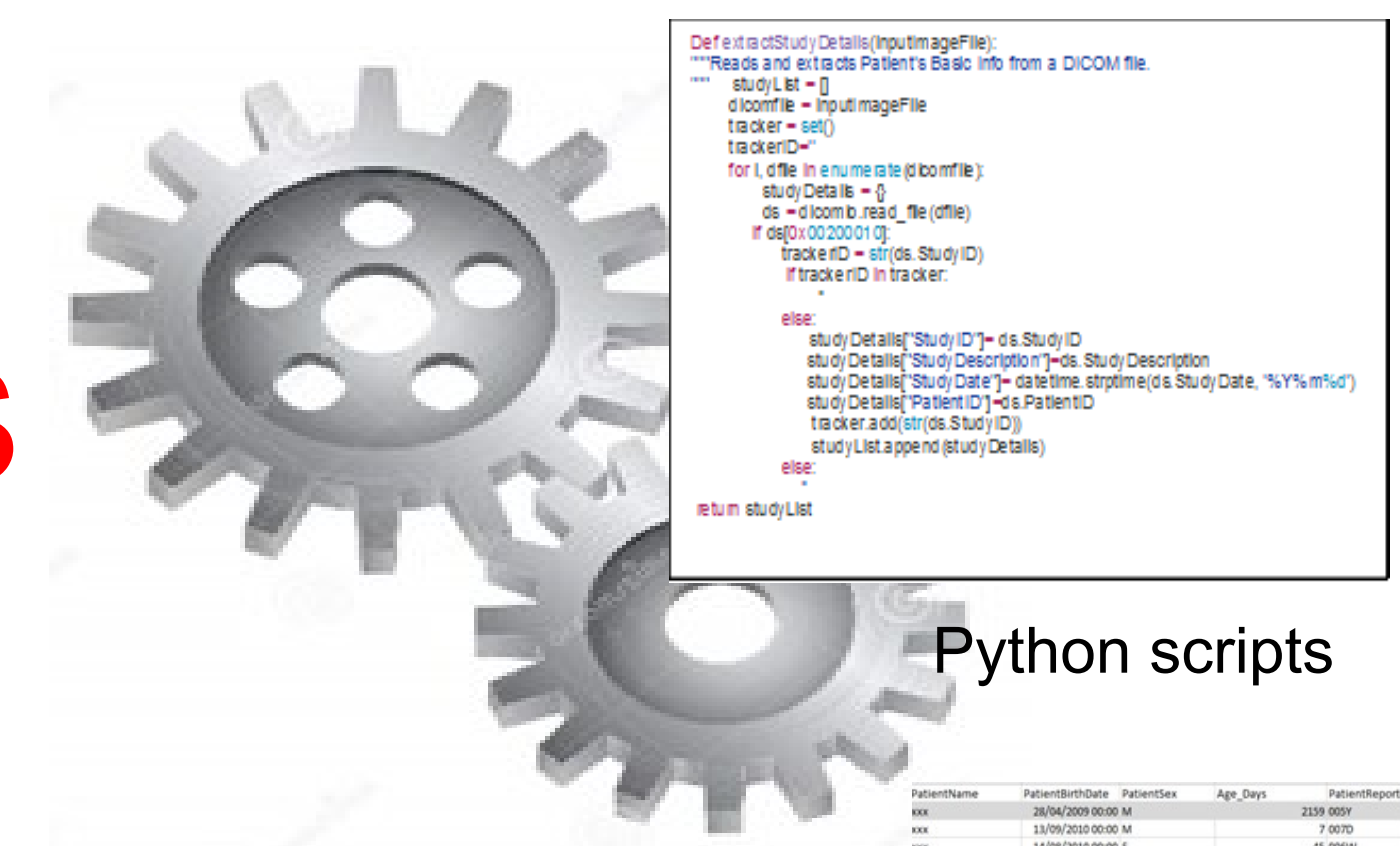
- The concepts of this work arose from the BRAINS project.  
<http://www.brainsimagebank.ac.uk/>
- Our python based scripts use a publicly available python-based library:  
<https://github.com/darcymason/pydicom>
- The scripts traverses through an organised tree of directories containing DICOM images
- Specify DICOM images top location and output file path
- Import DicomInfoExtract.py
- Run `extractDetails(retrieveDicomFiles(),'filename')`
- Automatically generates a set of organised csv files with details which include: anonymised patientID, modalities, and series descriptions for all unique imaging data sets in the directories.
- Scripts and instructions to use are freely available from:  
<https://github.com/FNNDSC/dmd2b>

Input



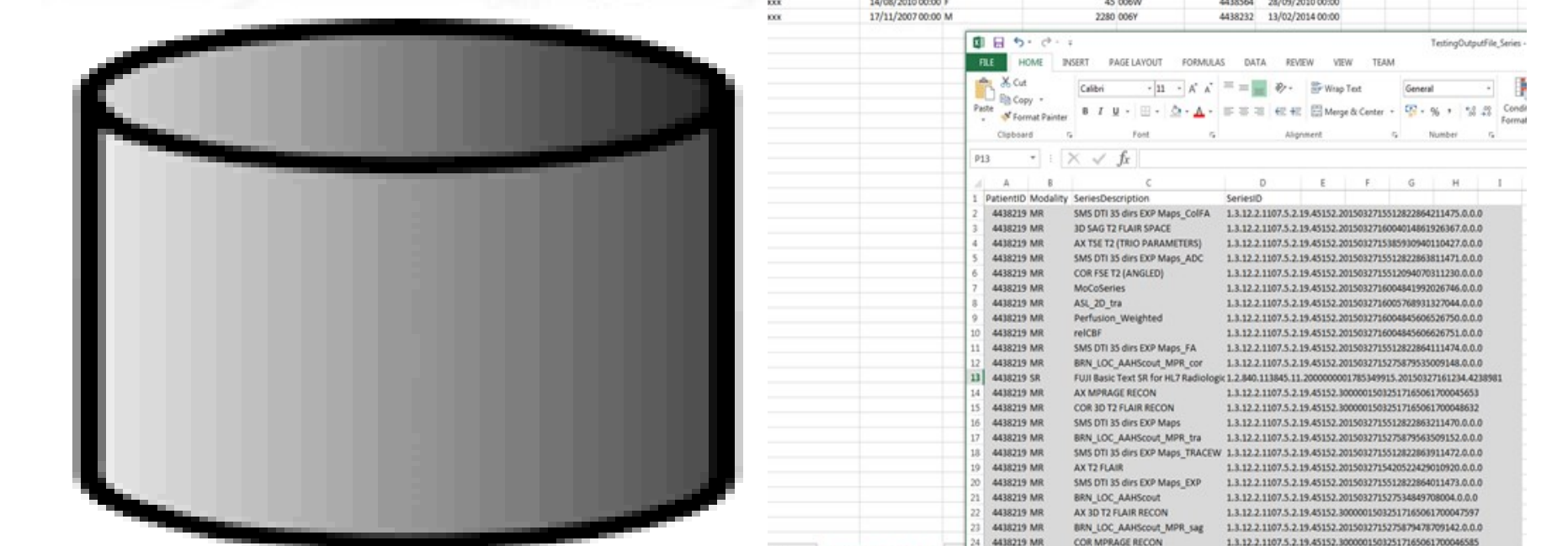
DICOM images

Process



Python scripts

Output



Database

csv spreadsheets

## Conclusion

- This work has provided a valuable resource for researchers with little computing expertise to automatically sort out and query images according to particular imaging types.
- This resource can also be useful at the pre-research stage when the imaging data is for example being sorted into a databank.
- The work was supported by the Scottish Funding Council through SINAPSE Postdoctoral and Early Career Researcher Exchange training.

## Acknowledgment

- Scottish Imaging Network - SINAPSE for an Early Career Researcher Exchange award.
- The BRAINS project team.
- The Computational Health Informatics Program, Harvard Medical School.

## References

- Mildenerger P, Eichelberg M, Martin E. Introduction to the DICOM standard. European radiology. 2002;12(4):920-7.
- Graham R, Perriss R, Scarsbrook A. DICOM demystified: a review of digital file formats and their use in radiological practice. Clinical radiology. 2005;60(11):1133-40.
- Job, D. E.; Dickie, D. A.; Rodriguez, D.; Robson, A.; Danso, S.; Pernet, C.; Bastin, M. E.; Boardman, J. P.; Murray, A. D.; Ahearn, T.; Waiter, G. D.; Staff, R. T.; Deary, I. J.; Shenkin S. D. and Wardlaw J. M. A brain imaging repository of normal structural MRI across the life course: Brain Images of Normal Subjects (BRAINS). NeuroImage. 2016.