An integrated software solution for improving neuroimaging data archival, management, and processing

- The experience from the Queen Square MS Centre

Dr Baris Kanber,
Translational Imaging Group,
Centre for Medical Image Computing,
Department of Medical Physics and Biomedical Engineering,
University College London

UNIVERSITÉ DE RENNES 1

UCL

Why was such a solution required?

- There were increasing numbers of methods available for processing neuroimaging data.
- Additionally, increasingly larger datasets were being acquired in clinical trials and research projects.
- It was difficult and cumbersome to manually manage and process such large datasets with different, and often linked, processing methods.

Figure: Irons, tele-aliens, and chairs.

TIG. TRANSLATIONAL IMAGING GROUP

cmic Centre for Medical Image Computing

Why was such a solution required?

- Often, data and processing efforts were duplicated by researchers and it was difficult for a new-starter to make sense of what data and results were already available to him/her.

- Furthermore, variations between software versions and pipeline parameters used by researchers could bias the results.
    - Controlling for inter-operator variability required time and effort.

Figure: Needle in a haystack. Copyright © Jolyon Trochianko.

Background

Tree structure for imaging data management:

- Adequate for small projects
  - Project root
    - Patient A
      - Visit 1
        - DICOM files
        - NIFTIs
        - Files resulting from processing
        - …
      - Visit 2
      - …
    - Patient B
    - Patient C
    - …

Figure: Needle in a haystack. Copyright © Jolyon Trochianko.

Background

- Becomes troublesome for larger projects or in multi-user environments

Figure: Needle in a haystack. Copyright © Jolyon Trochianko.

Example

0801_T1vol/

0901_DTI/

0902/

1001_b0/

1101/

1201/

1301_MT/

A1120130_B7033991-2010-00475_344776560_000601_000006-DC-Gm.roi

cortcomp_475/

DTI_analysis_FSL/

fill_t1_lesions/

FlairasDIR.nii

FlairasPSIR.nii

mtr-vbm/

outputResult.nii

PSIR-lobar-masks/

roi_varun/

T2_222_bet.nii

- The directories starting with numbers appear to be individual scans but it is pretty much anybody's guess what the rest of the files/directories are.

B0060473-2012-01536/0801_T1vol:

20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008.nii.gz

B0060473-2012-01536-T1vol_axial.roi

B0060473-2012-01536_T1vol_ROImask.hdr

B0060473-2012-01536_T1vol_ROImask.img

co20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_mixeltype.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_mixeltype.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_0.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_0.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_1.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_1.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_2.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pve_2.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pveseg.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_pveseg.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_seg_0.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_seg_1.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_seg_2.nii.gz

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_seg.nii

o20120305_101354MR001SAGT13DGRE1mmisoSENSEA9380733s801a1008_seg.nii.gz

- Inside a sub-directory, there appear to be many files that are either duplicated or look very similar.

B0060473-2012-01536/fill_t1_lesions:

Which scan is this?

B0060473-2012-01536_t1vol_axial.nii

B0060473-2012-01536_t1vol_bet_mask.nii

B0060473-2012-01536_t1vol_bet_mixeltype.nii.gz

B0060473-2012-01536_t1vol_bet.nii

B0060473-2012-01536_t1vol_bet_pve_0.nii.gz

B0060473-2012-01536_t1vol_bet_pve_1_c100.nii.gz

B0060473-2012-01536_t1vol_bet_pve_1.nii.gz

B0060473-2012-01536_t1vol_bet_pve_2.nii.gz

B0060473-2012-01536_t1vol_bet_pveseg.nii.gz

Also exist in the upper level directory

B0060473-2012-01536_t1vol_bet_seg_0.nii.gz

B0060473-2012-01536_t1vol_bet_seg_1.nii.gz

B0060473-2012-01536_t1vol_bet_seg_2.nii.gz

B0060473-2012-01536_t1vol_bet_seg.nii.gz

B0060473-2012-01536_T1vol_ROImask.hdr

B0060473-2012-01536_T1vol_ROImask.img

new_lfill_analyze/

PROBLEMS

- Data duplication
- Carrying out processing which has been done before
- Making sense of the data, and processing results that are available
- Processing/using wrong files

UNIVERSITÉ DE RENNES 1

UCL

In essence,

- Having a tree structure for imaging data and processing results is good practice but it is much better if it is actively managed…

TIG. TRANSLATIONAL IMAGING GROUP

cmic Centre for Medical Image Computing

I will,

- Present the Queen Square MS Centre XNAT Platform
- Highlight our pleasant experience with it

Figure: Needle in a haystack. Copyright © Jolyon Trochianko.

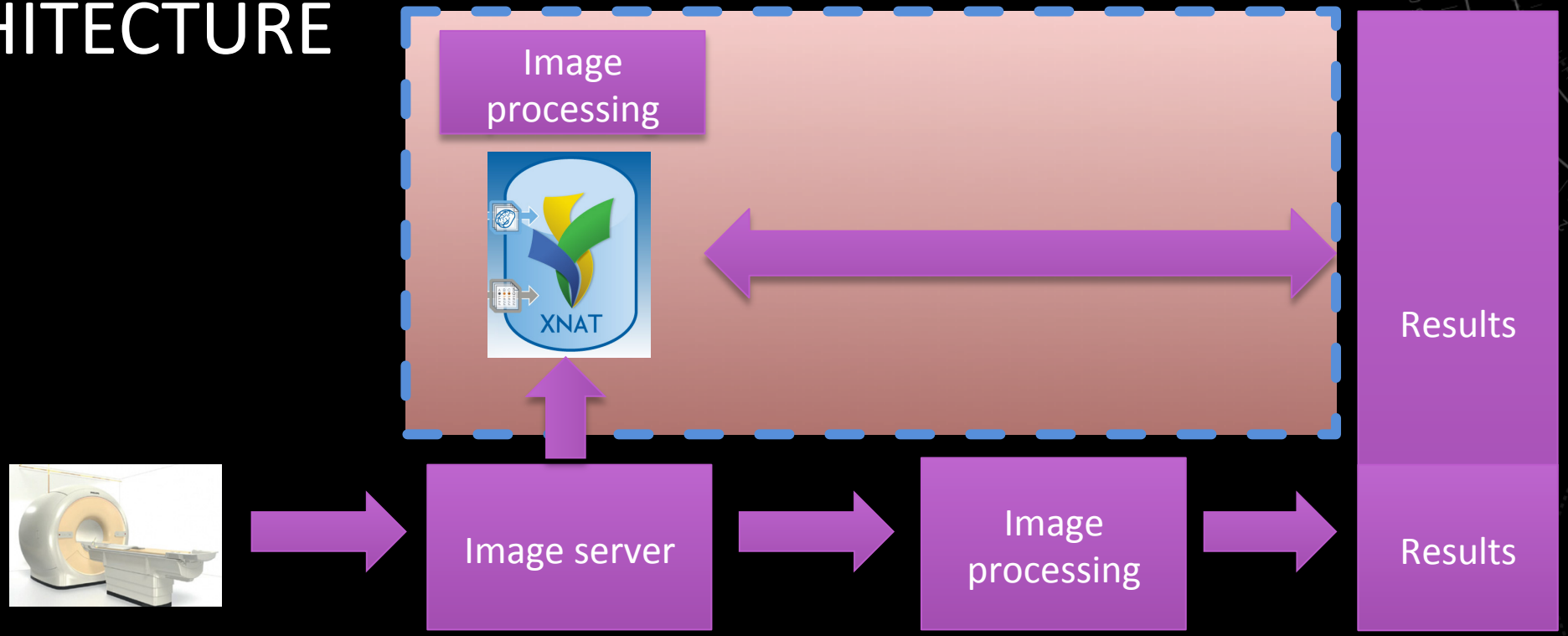
Methods

- Details to be presented by Dr Marc Modat later this afternoon.

UNIVERSITÉ DE RENNES 1

UCL

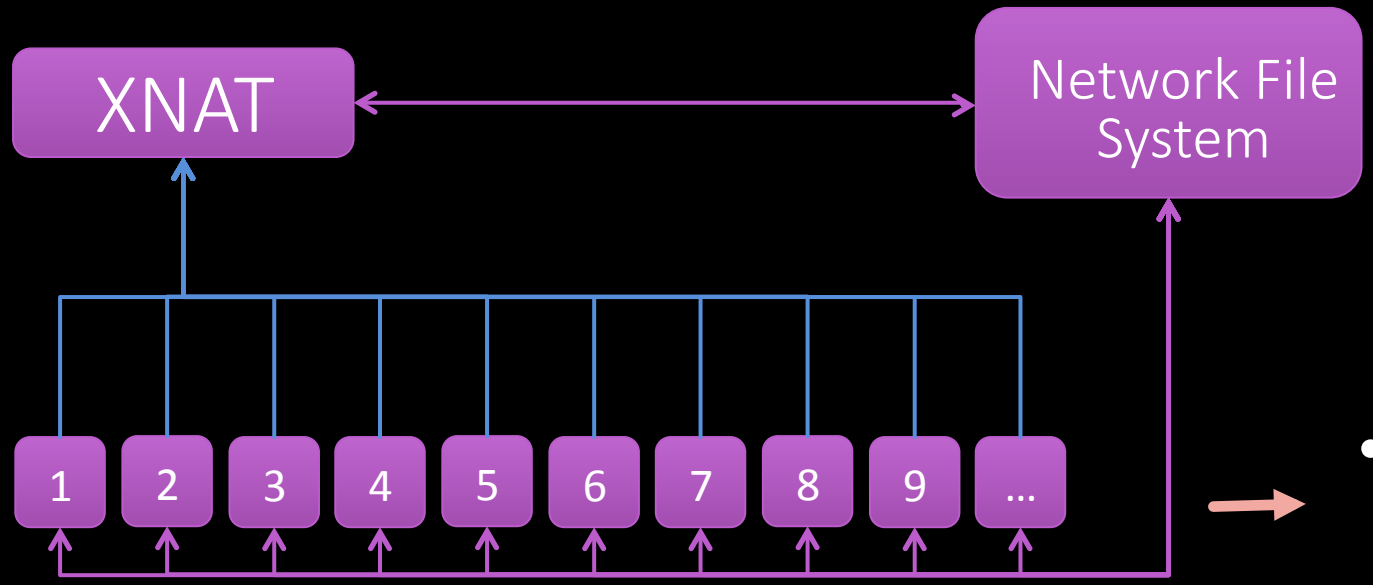
ARCHITECTURE



TRANSLATIONAL IMAGING GROUP

cmic
Centre for Medical Image Computing

AUTOMATED PROCESSING SYSTEM



Processing Nodes (minimal Linux, 8GB Memory, 40 GB disk)

- We can use:
  - A dedicated cluster, or
  - Existing workstations

Communications link

Data link

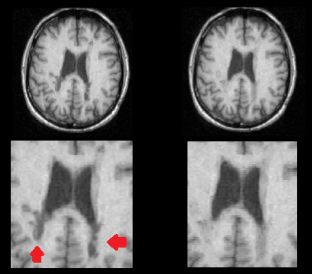
TRANSLATIONAL IMAGING GROUP

Centre for Medical Image Computing

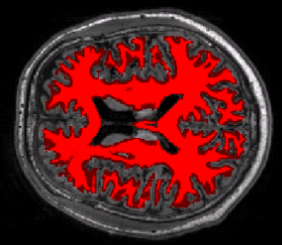
Queen Square MS Centre XNAT

- We currently have 20 projects, 1123 subjects, and 3953 MRI sessions.
- ~10 users.

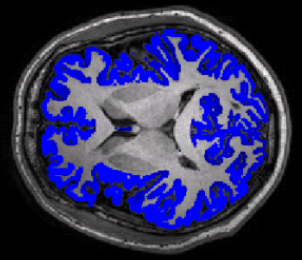
- Examples of automated processes we carry out:



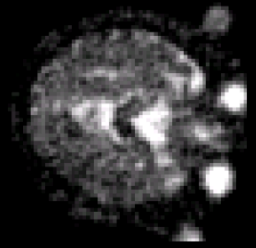
lesion filling



brain parcellations and tissue segmentation



cortical thickness mapping



brain sodium concentration maps

segmentations of hippocampi

TRANSLATIONAL IMAGING GROUP

Centre for Medical Image Computing

The experience

- We no longer experience the problems commonly encountered in manual processing methods such as the processing of the wrong data, the use of wrong flags or parameters.

- We are saving a lot of time: during the automated processing, we use our time to work on other things.

Other benefits

- A central repository for everyone
- Avoids duplication of data and efforts
- Users can be assigned appropriate access rights to individual projects
- Ability to also store non-imaging data (e.g. clinical data, spreadsheets)

Other benefits

- Setup and run pipelines depending on project and requirements
- View data online, download or upload
- Can be searched easily

Overall,

- The automated system for cataloguing, archiving, and processing neuroimaging data has been more time- and resource-efficient than the manual methods that were previously in use.

Finally,

- Introduction of this service has helped to standardize results across studies, as the same processing pipelines, software versions and parameters are used for all the processing.
- We now have a centralised location for all the data and processing results for new studies.
- New starters can easily get access to the existing results, which has increased the speed and efficiency of this process.
- This has been very cost-effective: we used the existing computational resources in the MS Centre.