Secure-DICOM-Uploader: A platform for anonymising and transferring imaging data from hospital sites to remote repositories

D.G. Beasley[1†], B. Eiben[2], S.J. Doran[4], J. D'Arcy[4], J.A. Petts[4], R. Patel[3], M. White[5], E. Miles[3], J.R. McClelland[2], M. Modat[1]

[1]*School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK*

[2]*Centre for Medical Image Computing, University College London, London, UK*

[3]*National Radiotherapy Trials QA Group (RTTQA), Mount Vernon Cancer Centre, Rickmansworth Road, Northwood, Middlesex, HA6 2RN, UK*

[4]*Institute of Cancer Research, 123 Old Brompton Road, London, UK*

[5]*UCLH NHS Foundation Trust, London, UK*

Introduction

For a large multi-site project, imaging (DICOM and other formats) and non-imaging data from different hospitals are anonymised and sent to a central repository for analysis. To provide secure data transfer from different hospital sites to a central repository, we have developed a user-friendly solution based on software components designed for and established in large-scale clinical trials.

Motivation and Requirements

A major challenge for large multi-site projects is to ensure that staff at local sites are able to efficiently and consistently transfer medical data to remote central repositories while ensuring the data is anonymised. Transfer using storage media (DVDs/USB drives) or sFTP can be very labour-intensive, unreliable and insecure. For the purpose of Quality Assurance for radiotherapy clinical trials, a new solution is required, with the essential components to include: adequate security measures, ease of installation and management so that local sites are willing and able to install the software; minimal steps required to transfer data, with the maximum automation possible; data must be pseudonymised, and checks in place to ensure that this is the case; checks are required that data being transferred fit the protocols for a particular clinical trial; a link must exist between anonymised and identifiable data to ensure all data can be identified by the clinical trial management; finally, each site must be in control of its own data, and data can only be uploaded off-site after the local user is satisfied that the data is properly anonymised and correct.

Software Description and Workflow

XNAT is a platform that provides for easy data management, image viewing and synchronisation of data [1]. It is growing in popularity and user-friendly, providing a web front-end for fully controlling and reviewing data. In the developed framework, a centralised XNAT on a remote server is used to collect the anonymised data from many sites. Local hospital sites host preconfigured XNAT servers running within Docker containers. XNAT allows for manual checks and validation scripts. This ensures only fully anonymised data is uploaded to a remote repository.

We have developed a customised XNAT-based workflow and placed this within a Docker service for easy distribution, reliability and control. Docker is a system to package software and its dependencies into lightweight 'containers' – these are effectively stripped-down virtual machines. As it is self-contained and

independent of the host operating system, it installs and runs the same on any Windows or Linux system. In addition, any updates can be easily installed with a simple command.

The Uploader consists of two separate XNAT servers running within the same Docker service, one hosting non-anonymised data, the other anonymised. DICOM images are pushed from PACS or placed in a network folder which is then automatically imported into the non-anonymised XNAT server. The local user can log into the XNAT server and view session details and visualise the images via the OHIF (Open Health Imaging Foundation) viewer [2] available via a plugin for XNAT. The data can then be anonymised either automatically or manually. For our purposes, we have designed two anonymisation routes; automatic, where subject and session IDs are generated using hashing and synchronised to a remote research database, or manual assignment of IDs for clinical trials. A clinical trial is selected from a drop-down list, and the clinical trial protocol loaded into XNAT and compared with the session data. For radiotherapy, ROI labels are checked to ensure that they conform to standard nomenclature and that all expected structures are present. Labels can be changed within XNAT if they do not. When the proposed clinical trial information is successfully validated against the protocol, the data can then be anonymised.

Anonymisation is performed in three parts. A configurable DicomEdit script is used that can be edited locally. A separate Python script is used for more complex manipulation of DICOM files; private tags are removed, and clinical trial information is also entered into the DICOM header at this stage. In addition, every file is checked to ensure that anonymisation is complete. The anonymised data residing on the non-anonymised XNAT is sent to a separate local anonymised XNAT server. From here, additional checks can be made before the user, when happy with the anonymisation, can upload the data to a remote repository.

Each local group has access and control of a corresponding project within the remote XNAT repository. For clinical trials, the central repository has a project for each site and, in addition, projects for each clinical trial. When an imaging session, and associated non-imaging data, is uploaded to a corresponding site project, it is automatically shared to the corresponding clinical trial.

Discussion & Conclusions

The Secure DICOM Uploader improves the experience and reliability of data transfer for research. The software requires minimal training while providing a number of checks to prevent identifiable information being transferred. It is planned to distribute the software for upcoming clinical trials and research.

References

[1] The Extensible Neuroimaging Archive Toolkit (XNAT): An informatics platform for managing, exploring, and sharing neuroimaging data, Marcus, D.S. et al., Neuroinformatics 5(1) (2007) 11-34

[1] OHIF Viewer (http://ohif.org/)

Acknowledgements

The RTTQA group is funded by the NIHR. This work was supported by the Wellcome/EPSRC Centre for Medical Engineering [WT 203148/Z/16/Z], the CRUK ART-NET Network Accelerator Award (A219932) and the STFC Global Challenge Network+ in Advanced Radiotherapy (ST/N002423/1).

†Corresponding Author: daniel.beasley@kcl.ac.uk