A national network of trusted data repositories for the Australian National Imaging Facility

Mehnert, A.[1], Janke, A.[2], Gruwel, M.[3], Goscinski, W.[4], Close, T.[4], Taylor, D.[1], Narayanan, A.[2], Vidalis, G.[4] and Galloway, G.[2]

[1] Centre for Microscopy, Characterisation and Analysis, The University of Western Australia, Perth, Australia, [2] National Imaging Facility, The University of Queensland, Brisbane, Australia, [3] Mark Wainwright Analytical Centre, The University of New South Wales, Sydney, Australia, [4] Monash eResearch Centre, Monash University, Melbourne, Australia

A Trusted Data Repository (TDR) service is essential for sharing data and ensures that data created and used by researchers is "managed, curated, and archived in such a way to preserve the initial investment in collecting them" and that the data "remain useful and meaningful into the future" (www.coretrustseal.org). In this talk we present a recently completed project to define the requirements and best practices for establishing TDR services for the Australian National Imaging Facility (NIF) and to establish exemplar services for both preclinical MRI data and clinical ataxia MRI data. The project "Delivering durable, reliable, high-quality image data" was jointly funded by the Australian National Data Service (ANDS) and Australian Research Data Services (RDS). It was motivated both by NIF's desire to enhance the quality of the data associated with the use of its facilities, and the desire of funders for TDRs that enable access to data for at least 10 years and include metadata that documents both the quality of the data and its provenance. The project involved four NIF nodes across four states: University of Western Australia (lead), University of Queensland, University of New South Wales and Monash University. Whilst the scope of the project was limited to MRI data, the outcomes can be and have been adapted to other instruments and modalities.

The key project outcomes include:

1. The "NIF agreed process for acquiring trusted data" that details the requirements that must be satisfied and the process that must be followed to obtain high-quality data suitable for ingestion in a NIF TDR service. They include provisioning of a unique instrument identifier, instrument registration with Research Data Australia (researchdata.ands.org.au), quality control (QC) and requisite metadata (including cross-reference to the QC data).
2. The "NIF requirements for a trusted data repository service", a platform-agnostic checklist of requirements that a basic NIF trusted data repository service should satisfy, including identification of data by a unique project identifier, ingestion of data from NIF-compliant instruments, authentication via the Australian Access Federation (aaf.edu.au), interoperability and easy deployment across NIF nodes.
3. Exemplar implementations of TDR services for preclinical MRI data and clinical ataxia MRI data among the participating NIF nodes.
4. Assessments of the resulting TDR services against a relevant international metric, the CoreTrustSeal "Core Trustworthy Data Repositories Requirements" (www.coretrustseal.org).

We also describe the benefits and outcomes for NIF users, the broader imaging community, the universities and NIF itself. These include improved reliability of research outputs, the provenance associated with it, making NIF data more FAIR (Findable, Accessible, Interoperable, Reusable) and facilitating compliance with the Australian Code for the Responsible Conduct of Research. Finally we present the transition plan post-funding including maintenance of the existing services for 10 years, the integration of additional instruments, planned new national and international service deployments and support for the newly co-funded ANDS, RDS and NeCTAR (Australian National eResearch Collaboration Tools and Resources) project entitled "The Characterisation Data Enhanced Virtual Laboratory".