

# Phase II of the Imaging Data Service: Expanding XNAT

Ryan Sullivan<sup>1</sup>

Haofei Feng<sup>1</sup>, Vipul Patel<sup>1</sup>, Xu Fang<sup>1</sup>, Stephen Kolmann<sup>1</sup>, Chris Albone<sup>1</sup>, Andrew Janke<sup>1</sup>

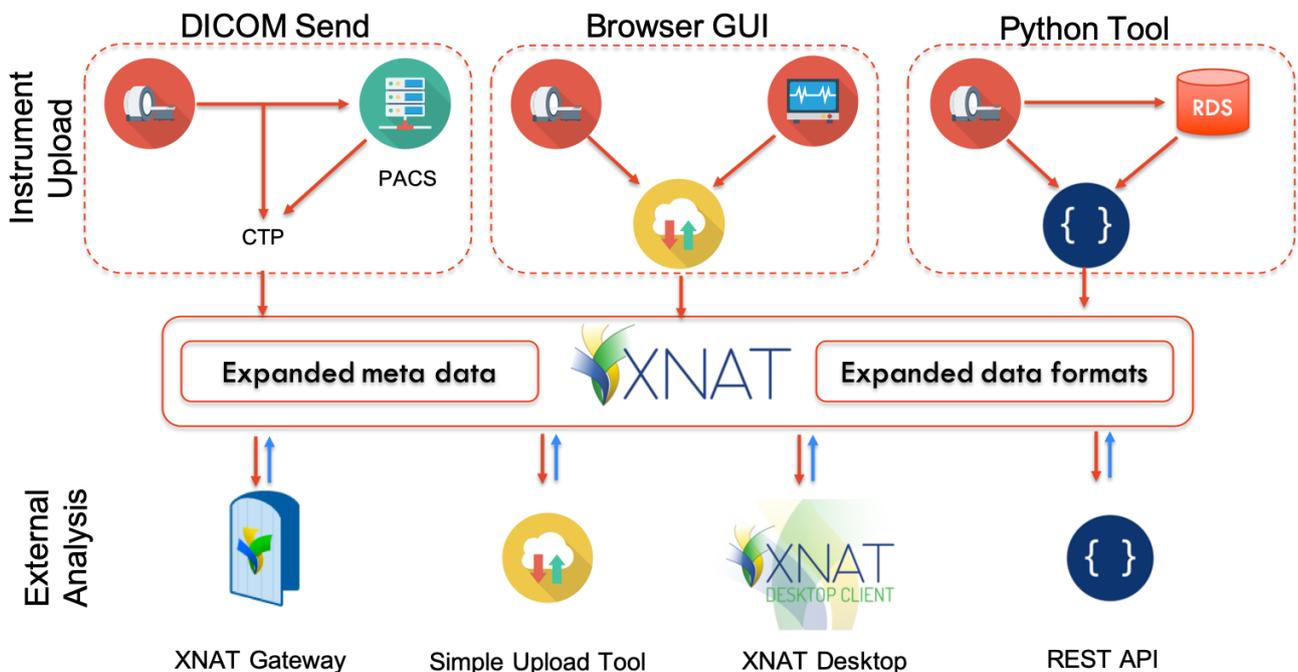
<sup>1</sup>The University of Sydney, Sydney, AU [ryan.sullivan@sydney.edu.au](mailto:ryan.sullivan@sydney.edu.au)

## INTRODUCTION

In 2018, ICT and Sydney Imaging released Phase I of our Imaging Data Service, a university wide data management and informatics platform targeted at medical imaging, using XNAT as the core technology. This was well received by researchers in clinical imaging, where data formats are highly standardized, but met some limitations in preclinical imaging, where standards are often the exception rather than the rule. We present our continued work in Phase II, refining and operationalizing the platform, integrating with other university services, and setting us up to expand past our initial ~85 end users to the broader 500+ imaging community.

## EXPANDED DATA AND METADATA

We have expanded our data ingestion toolkit to cover instruments with non-DICOM data and have additionally expanded the parsable metadata so that it can be reported on by the system. The automated tools now capture NIF TDR metadata such as Instrument ID (IID) and Quality Control Project ID (QCPID) and operational metadata useful for facility and ethical compliance reporting, with the aim of being able to link specific instrument usage to grants and research outputs in a robust manner. As the toolkit consists of a general upload library developed with Radiologics and vendor specific configurations and scripts to harmonize data structure, it is extendable and translatable to instruments at other sites.



We have also more thoroughly integrated open source tools to better facilitate user workflows for highly interactive analysis on such platforms such as the Characterization Virtual Lab (CVL) and for real time clinical viewing.

### **REFINED HYBRID ARCHITECTURE**

In order to increase the reproducibility of research analysis, we have implemented dual containerized pipeline options to replace the original XML based pipelines of XNAT. The first runs on a Docker Swarm on our AWS cloud while the second translates to Singularity containers that run on the University's HPC. This gives us extra versatility to balance load and cost, running pipelines where most appropriate. Extensive performance and security testing have helped us understand the limitations of the software and our deployment from Phase I, and lead to security patches upstream, benefit the wider community. The changes we are implementing to address these will help solidify IDS and XNAT as a scalable platform that can serve the research base at an institutional level along with affiliated partners.