

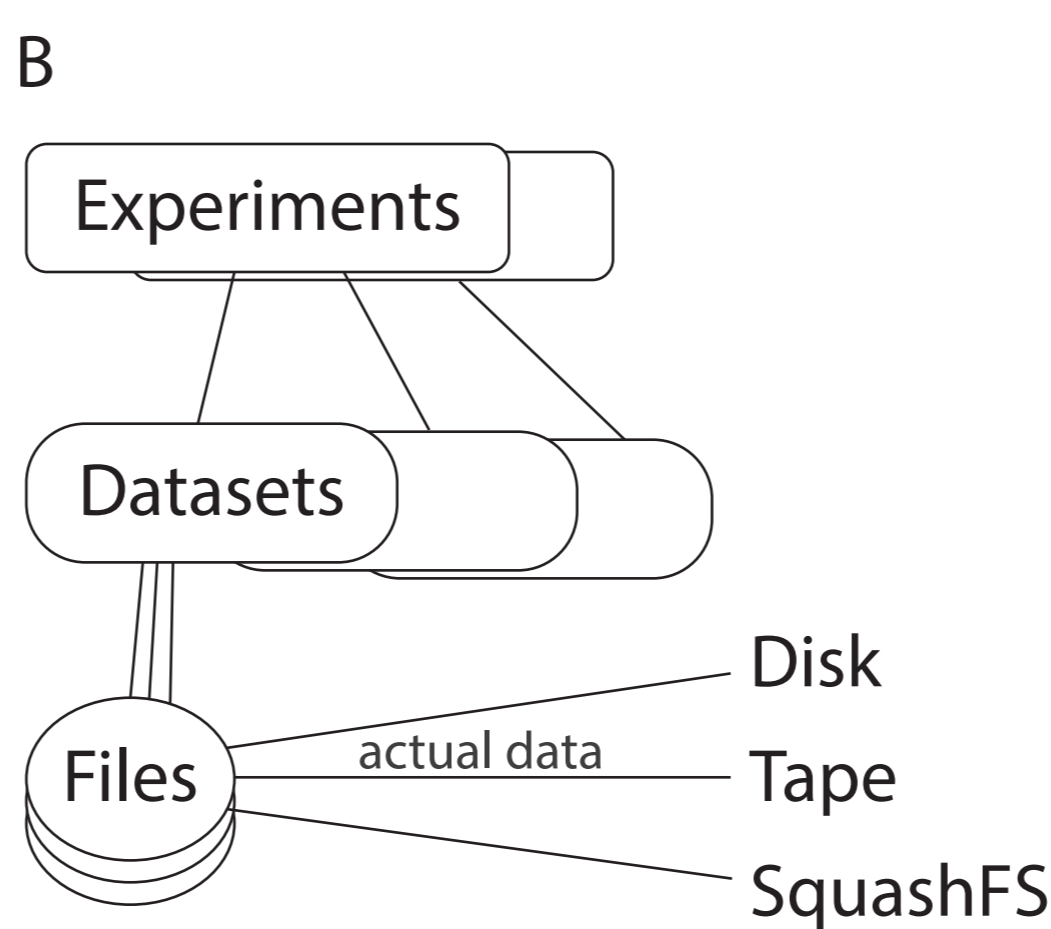
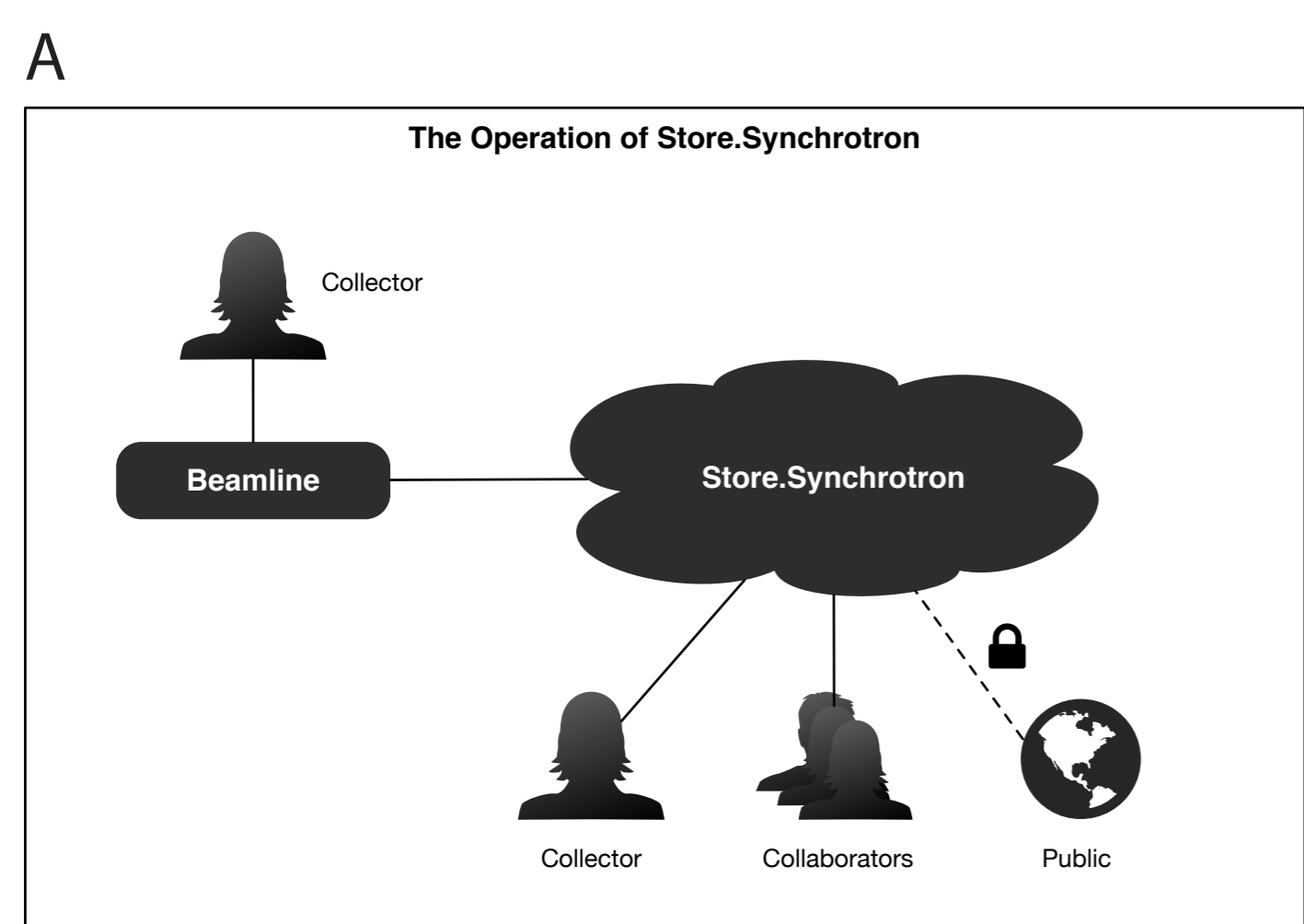
The Australian Store.Synchrotron Data Management Service for Macromolecular Crystallography

Grischa R Meyer, Steve Androulakis, Philip Bertling, Ashley Buckle, Wojtek James Goscinski, David Groenewegen, Chris Hines, Anitha Kannan, Sheena McGowan, Stevan Quenette, Jason Rigby, Patrick Splawa-Neyman, James M Wettenhall (Monash University, Clayton, Australia), David Aragao, Tom Caradoc-Davies, Nathan Mudie (Synchrotron Light Source Australia, Clayton), Charles Bond (University of Western Australia, Crawley, Australia)

Store.Synchrotron is a service for management and publication of diffraction data from the macromolecular crystallography (MX) beamlines of the Australian Synchrotron. Since the start of the development, in 2013, the service has handled over 79 TB of raw data (~ 5.8 million files). Raw data is made available securely via the web and SFTP so experimenters can sync it to their labs for further analysis. With the goal of becoming a large public repository of raw diffraction data, a guided publishing workflow which optionally captures discipline specific information was built. The MX-specific workflow links

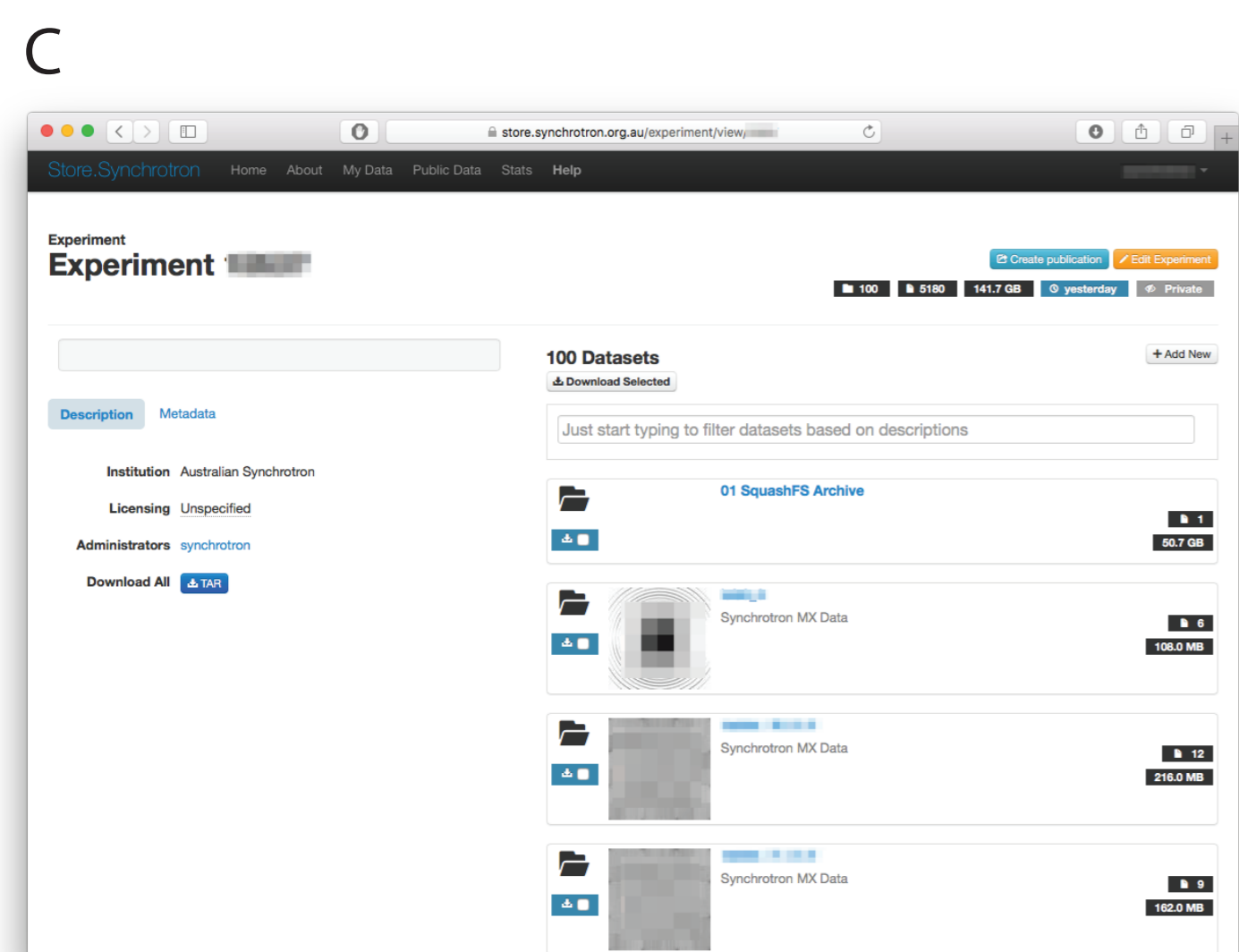
PDB coordinates from the PDB to raw data. An optionally embargoed DOI is created for convenient citation. This repository will be a valuable tool for crystallography software developers. To support complex projects, integration of other instruments such as microscopes is underway. We developed an application that captures data from instrument computers, enabling centralised data management without the need for custom ingestion workflows. The next step is to integrate the hosted data with automatic and interactive processing and analysis tools on virtual desktops.

Web access to live data SFTP access Data publishing Link to PDB DOI minting

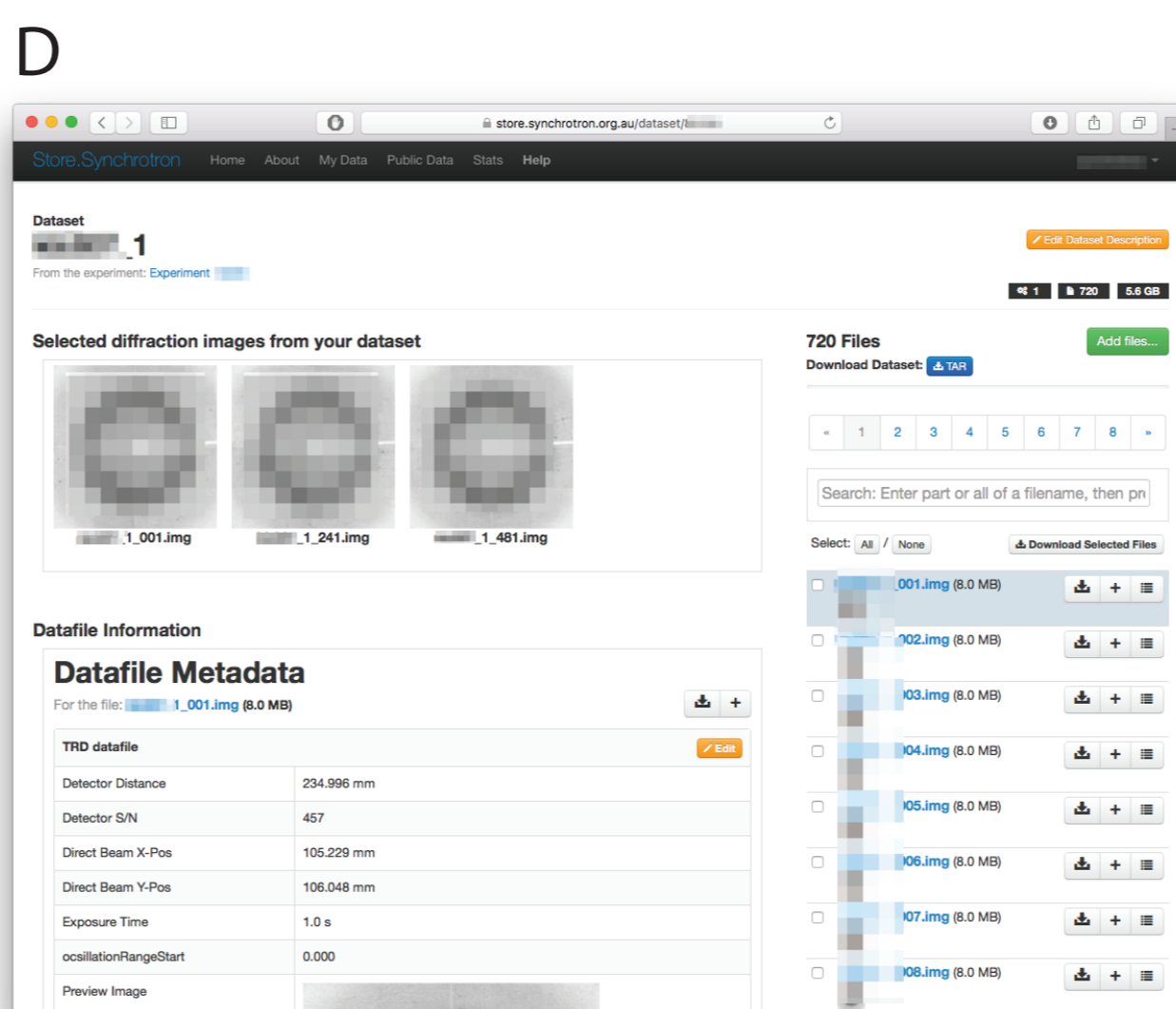


A: Schematic overview of the Store.Synchrotron service

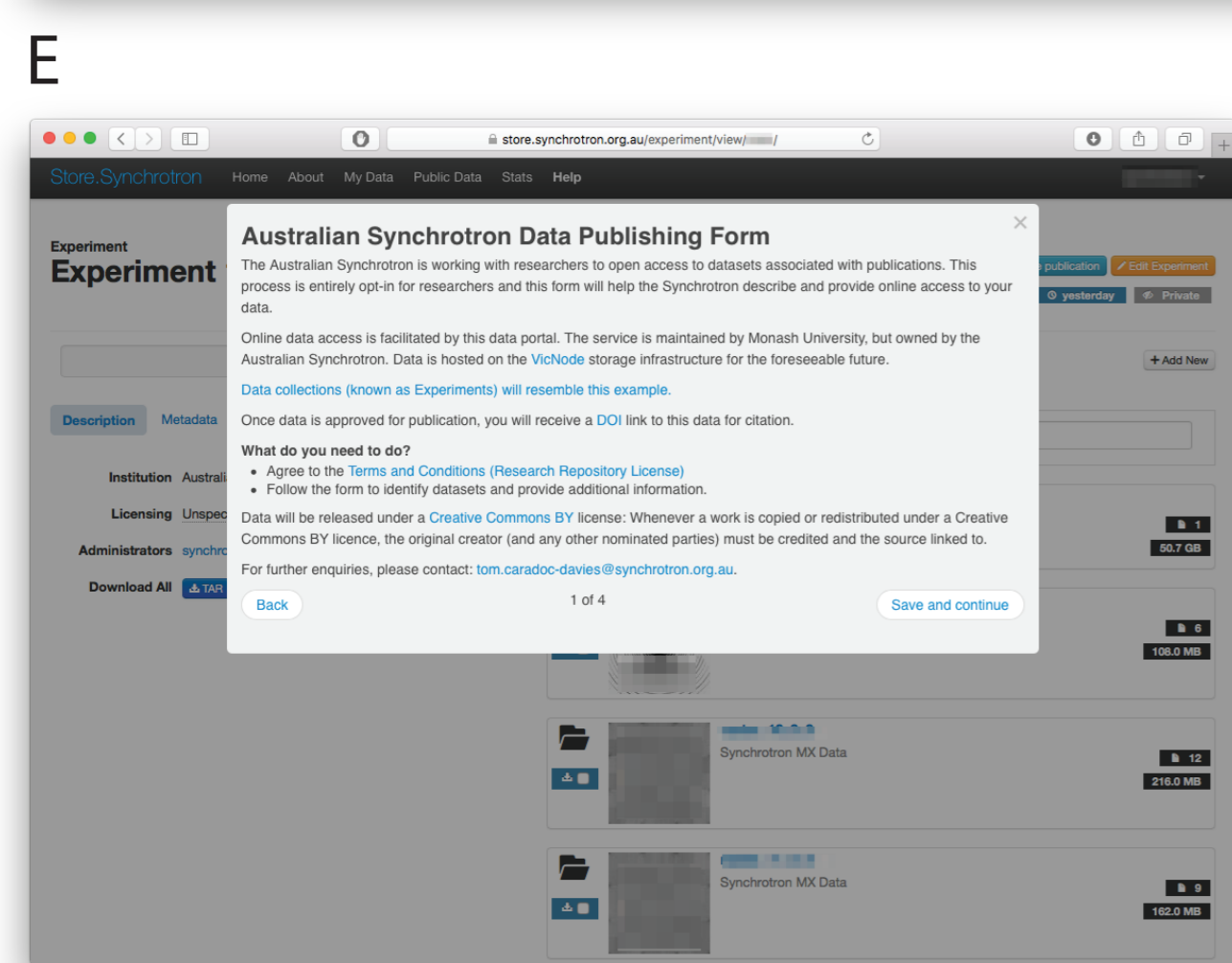
B: Storage schema and abstraction. Experiments are beamline allocations, datasets combine files for each crystal. The files can be stored in any number of locations and accessed transparently via Web or SFTP



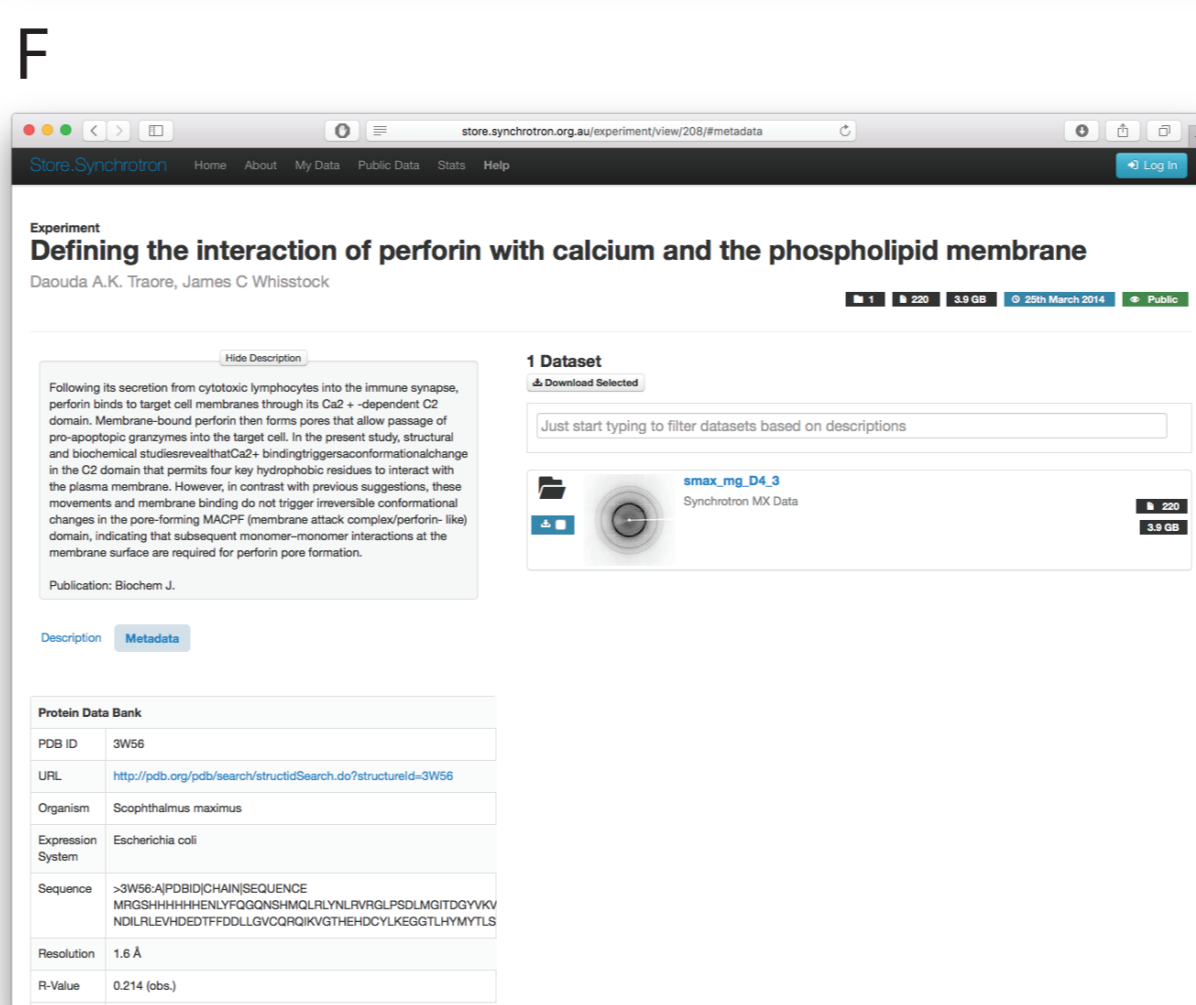
C: Experiment page right after collection. SquashFS archive added after Experiment completion, auto-processing coming soon.



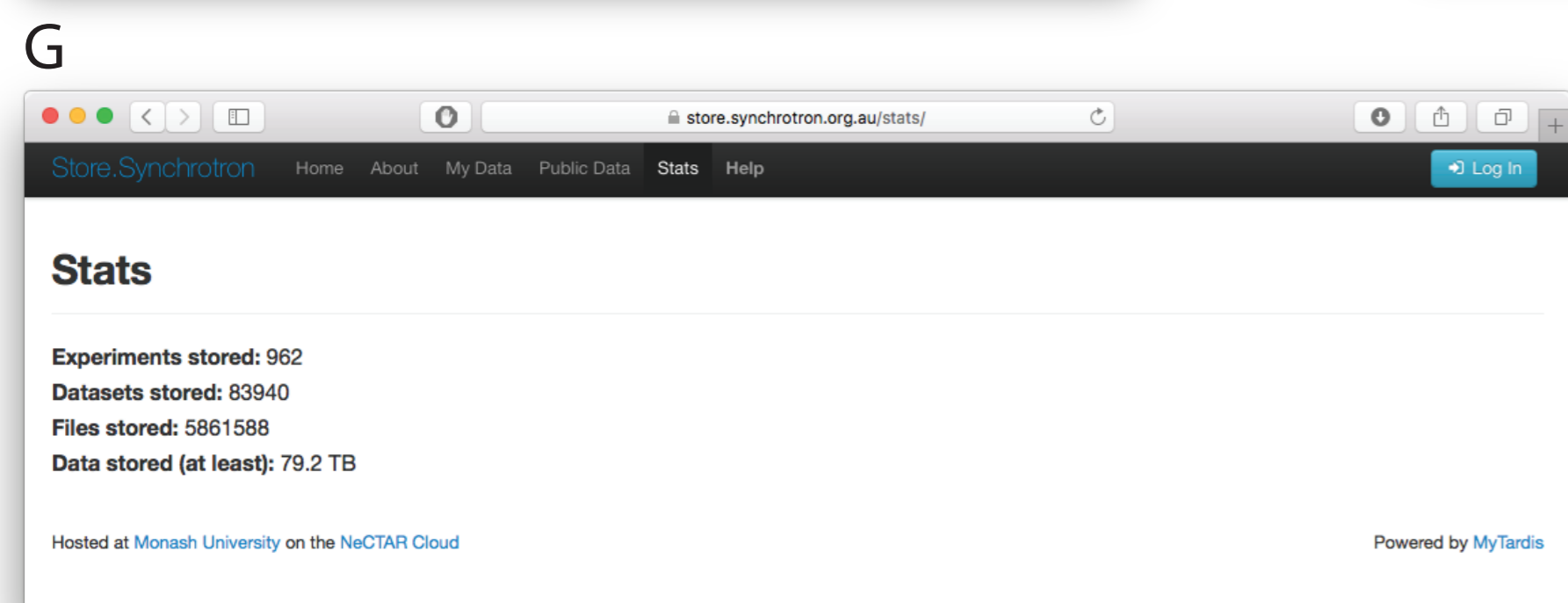
D: Dataset view, shows previews for image files, metadata for dataset (hidden here), a list of files and rich metadata for each file.



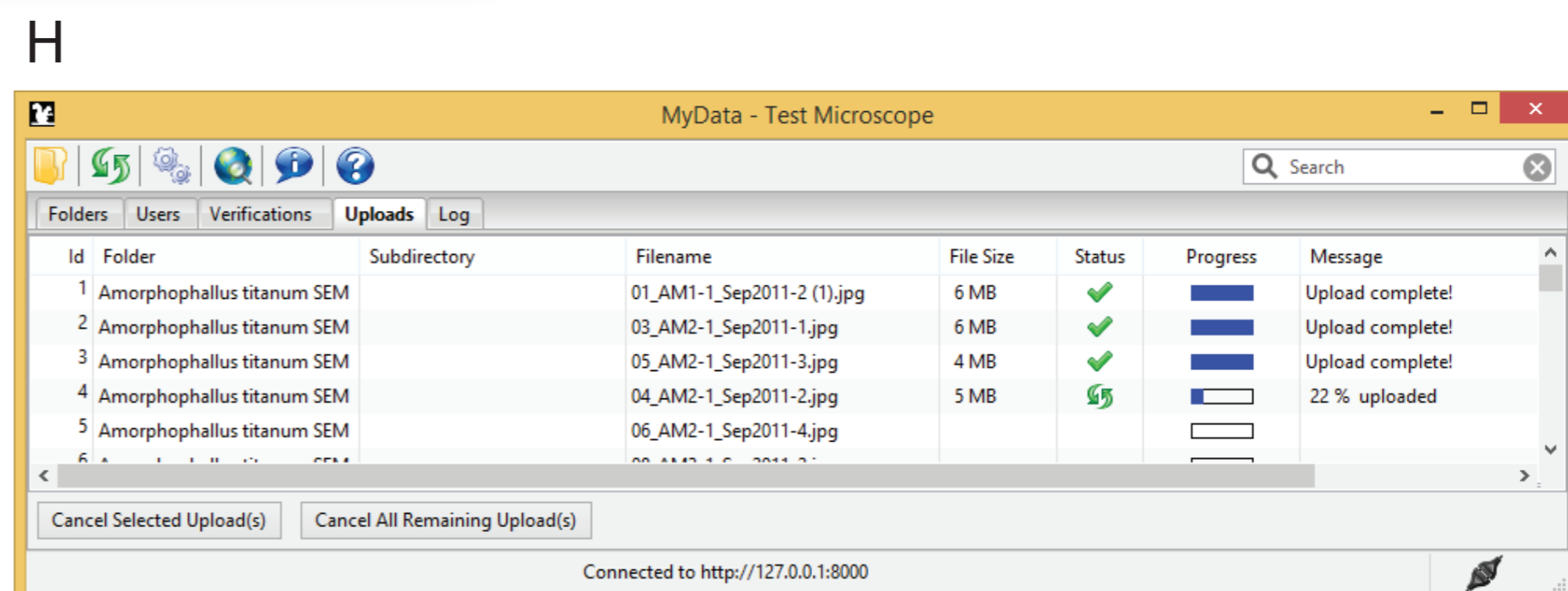
E: Publication form start page. Three dialogs present streamlined query for information on data, PDB ID, authors, license, embargo.



F: Public experiment view, shows title, detailed description, metadata and a list of datasets. Some metadata and related publications (hidden here) auto-harvested from PDB website.



G: Latest statistics of Store.Synchrotron site.



H: MyData desktop application. Used for data upload from microscopes and many other instruments.