Implementing IVOA services and tools for Australian optical astronomy datasets

Jon Smillie, Rod Harris, Ben Evans
ANU Supercomputer Facility, Australian National University, Canberra, Ben.Evans@anu.edu.au

ABSTRACT
In the last decade it has become clear that high-level database facilities are essential for not just data distribution, but also the basic scientific analysis of large datasets. For optical astronomy in Australia, the ANU, UQ, and AAO have begun to build a central facility for the storage and analysis of four significant optical astronomy datasets: SkyMapper Southern Sky Survey; WiggleZ Dark Energy Survey; GAMA Galaxy Survey, and MACHO. This has achieved efficient quality control on the underlying data and compliance with the IVOA standards for data accessibility. In this talk we will describe the standards, services and tools that have brought these to bear on these nationally significant datasets to begin to provide the scientific value from the investments in these projects.

INTRODUCTION
The national optical astronomy facility has emerged with initial support by an ARC LIEF grant – LE100100111. Based on the ANU using infrastructure at the ANU Supercomputer Facility, the astronomy facility has been targeted at the four of the largest Australian optical astronomy projects. Each of these major surveys has been recognised by substantial ARC funding, in addition to extremely large amounts of telescope time. Together they will address some of the most important questions in modern astronomy ranging from the nature of dark energy to the formation of stars and galaxies. In the case of SkyMapper we have set up a procedure to continuously store observations as they occur. The WiggleZ and GAMA data are periodically ingesting their public data releases. The MACHO dataset is complete, and has the most complete of the datasets for which services have been deployed.

The SkyMapper Southern Survey will measure the most distant objects in our Solar System to help understand how it formed; find the oldest stars in our Milky Way galaxy to determine how the first stars in the Universe formed; and find the oldest stars in our Milky Way galaxy to determine how the first stars in the Universe formed. The WiggleZ Dark Energy Survey will provide a test the currently favoured “cosmological constant” model for dark energy by measuring a characteristic scale in the distribution of 240,000 distant galaxies. It will also be the first experiment to measure the epoch at which dark energy starts to overpower gravity in the Universe. Importantly, it will also provide a measurement that is completely independent of existing evidence for dark energy. The GAMA Galaxy Survey will test our understanding of how structure forms in the Universe. Specifically, it will:
identify the processes (such as star formation and magneto-hydrodynamics) that begin to dominate over simple gravitational interactions on scales smaller than galaxy clusters, by measuring the relationship between visible galaxy mass and the masses of the larger dark matter concentrations they are found in;
determine the efficiency of star formation in galaxies by measuring the distribution of galaxy masses associated with their stellar components; and directly measure the recent rates of collisions (or mergers) between galaxies, to determine the significance of mergers in building up the mass of stars in galaxies, compared to in situ star formation.

The focus for the resultant datasets has been to develop a general and extensible framework for IVOA services that can be deployed on top of new datasets with a minimum of work. This involves an abstraction layer between the underlying project database, which may be any variant of SQL, and the web service itself, implemented over HTTP. Another motivation for the abstraction layer was to allow the astronomers to modify the service without having to touch any code themselves. Our work has also included developing the Conesearch and Simple Image Access VO services for these datasets and a corresponding visualisation tool that can communicate with these (or other VO compliant) services and render the results onto a 3 dimensional sky-sphere.
ARCHITECTURE AND WORKFLOW

The basic architecture for a query is shown below in Figure 1.

![Figure 1: query architecture](image)

We have defined an XML map for each service and dataset. This file is parsed by the base VOMapper class and defines how to map between parameters passed to the service via a HTTP get and the final SQL statement that is executed and the table or view that is searched. The map definition includes optional type and range attributes for parameter sanitisation to help prevent SQL injection attacks. Finally the XML map file also includes instructions on how to transform the rows returned from the SQL select statement to a list of key, value dictionaries. These dictionaries are then passed to a renderer (such as VOTable or HtmlTable) to be converted into textual data and returned to the caller.

The VOMapper class can be extended and certain methods overridden to allow any extended functionality (such as complex joins in the database or programatically generated 'synthetic' fields) not provided by the base service to be implemented.

As an example of a tool, we have developed the VO Search Visualisation tool to validate the results returned from our VO services. It is an OpenGL powered desktop application capable of querying Conesearch and SIA VO compliant services, parsing and rendering the results. It is capable of simultaneously connecting to several services thus enabling the user to query several projects for the region of sky they are interested in. The tool can also save the VOTables returned from the services and render them to the browser as HTML tables. We have included the ability to display astronomical markers (such as certain well known stars/galaxies) to aid the user in navigating the virtual sky sphere.

![Figure 1: VO Search Visualisation Tool](image)

REFERENCES

4. IVOA VO services standards definitions: http://www.ivoa.net/Documents/

Melbourne, Australia
6-10 Nov 2011
5th eResearch Australasia Conference
ABOUT THE AUTHOR(S)

Jon Smillie is a programming consultant in the ANU Supercomputer Facility with a focus in Astronomy and Astrophysics. Rod Harris is a programmer in the ANU Supercomputer Facility. Ben Evans is the Head of the ANU Supercomputer Facility.