The All Sky Virtual Observatory

Yeshe Fenner 1, Darren Croton 2, Jon Smillie 3

1 Astronomy Australia Ltd, Melbourne, Australia, yeshe.fenner@astronomyaustralia.org.au
2 Swinburne University of Technology, Hawthorn, Australia, dcroton@astro.swin.edu.au
3 National Computational Infrastructure, Australian National University, Canberra, Australia, Jon.Smillie@anu.edu.au

ABSTRACT

New astronomical telescopes are producing data in volumes never previously experienced in Australian astronomy. To gain maximum scientific benefit from this flood of information, datasets must become part of an astronomical data fabric that uses international standards and services to connect data to infrastructure and users. The All-Sky Virtual Observatory (ASVO) [1] is a coordinated effort to build data infrastructure and services to link significant Australian astronomy datasets to the global astronomical data fabric, beginning with the following "Nodes" that were completed in early 2014:

- *ASVO-SkyMapper Node*: This Node is deployed at NCI/ANU and provides access to data from ANU's SkyMapper telescope, which is being used to produce the most detailed and sensitive map of the southern sky at optical wavelengths, and will be a fundamental reference for astronomers [2].
- *ASVO-TAO Node*: The Theoretical Astrophysical Observatory (TAO), deployed at Swinburne University of Technology, houses a growing ensemble of theory data sets and galaxy formation models, with value-add tools that allow users to easily create virtual universes to compare with real observations.

This talk will describe the framework developed for ASVO, early outcomes, lessons learned, and plans for the future of ASVO, which includes creating the first radio Node of ASVO and supporting new optical datasets from the Anglo-Australian Telescope.

ASVO v1.0 was funded by National eResearch Collaboration Tools and Resources (NeCTAR) and the Department of Education, and built through a partnership involving Astronomy Australia Ltd (AAL), Swinburne University, NCI/ANU and Intersect Australia Ltd.

DEVELOPMENT & DESIGN

A guiding principle in designing and developing ASVO is to adopt, wherever possible, the data interoperability standards and protocols developed by the International Virtual Observatory Alliance (IVOA). IVOA standards and services are powering a new era in astronomical research and discovery, by enabling astronomers to more easily share their data and access and combine datasets spanning radio, optical, ultraviolet and X-ray wavelengths, and theoretical data.

ASVO currently comprises two primary Nodes that support very different types of data:

- *ASVO-SkyMapper Node*: A high-performance database was created at NCI to support the expected 1.5 petabytes of SkyMapper image data and 200 terabytes of catalogue metadata. The image data are in the very common astronomical FITS format, while the catalogue data are in PostgreSQL databases, optimized to improve performance for common types of queries. The data and the co-located services are compliant with IVOA standards and tools. Two modes for data access have been developed: 1) a simple user-friendly web portal [3] for performing basic searches of the databases; and 2) IVOA services that can be accessed using very popular astronomical VO tools like TOPCAT [4] that support powerful and complex queries and allow the SkyMapper data to be cross-matched and compared with other VO-compliant datasets.
- *ASVO-TAO Node*: The Theoretical Astrophysical Observatory (TAO) provides a user-friendly web portal to access mock extragalactic survey data generated using sophisticated semi-analytic galaxy formation models that are coupled to large N-body cosmological simulations and the output stored in an SQL database [5]. A major feature of TAO is its ability to post-process the core data for different scientific applications. This is achieved through a modular design that makes TAO readily expandable with new functionality in the future. A particularly powerful and unique feature of TAO is the telescope simulator that allows users to predict the observations expected from specific telescopes or instruments, including SkyMapper. A growing number of telescopes and surveys are supported in TAO, and we expect that this feature will be developed in parallel with the creation of new observational data Nodes within ASVO. This provides a direct and vital link between the theory and the observations.

EARLY OUTCOMES

In the first two months of operations of ASVO-TAO, there were 150 registered users who executed ~120 jobs/queries. Another positive indicator of the value of ASVO has been the strong interest from astronomers to contribute their datasets and/or contribute to the development of ASVO. The ASVO-SkyMapper infrastructure is complete, but we are awaiting the first public release of SkyMapper data, expected in early 2015.

In this talk, I will describe our efforts and strategies to improve awareness, uptake and usage. I will also present examples of early science outcomes and use cases, including the use of TAO to predict the deep night sky expected to be seen by the Australian Square Kilometer Array (ASKAP) telescope [6, 7] and the use of SkyMapper data to discover the oldest known star, which was published in the journal Nature [8].

LESSONS LEARNED

This presentation will outline key lessons learned during the project, including the value of:

- bringing in project management skills and implementing professional software development methodologies
- senior scientists being closely involved and invested in the project, and acting as key champions and promoters
- developing robust automated test suites
- utilising established and mature International Virtual Observatory Alliance (IVOA) standards
- reassessing and refining user requirements at key stages during the project lifecycle
- allocating plenty of time to the verification of the science output from complex software, user testing, and developing a fixed set of automated verification and statistical tests as early as possible

THE FUTURE

This presentation will describe our intentions to expand the capabilities of ASVO, including creating the first radioastronomy ASVO Node to support the ~10 petabytes of data from Murchison Widefield Array telescope in Western Australia, which will be stored at Pawsey/iVEC. We are also taking steps towards the development of an ASVO-AAT Node, which will serve up key optical datasets from the Anglo-Australian Telescope (AAT). The current status and planned expansion of ASVO is outlined in Figure 1.

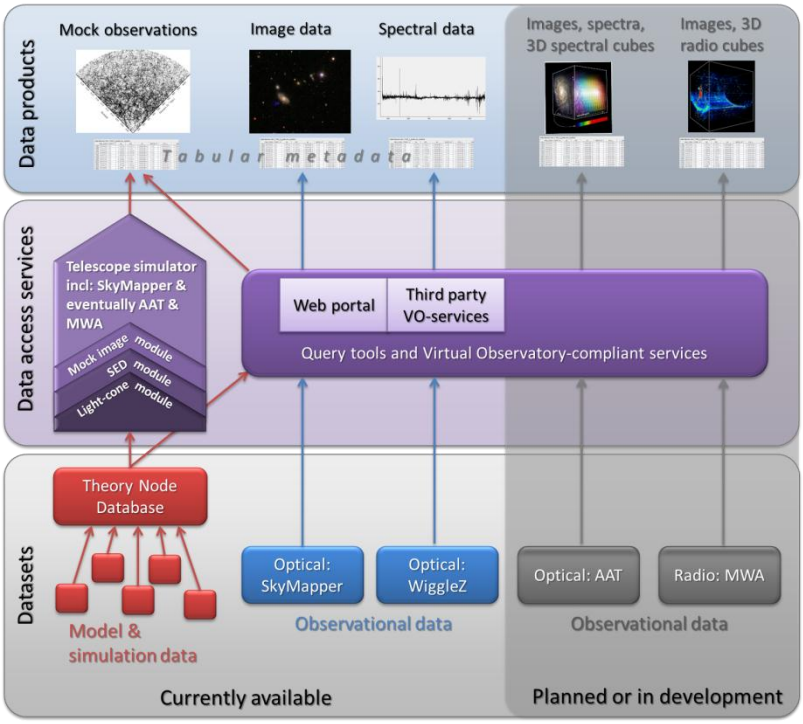

Figure 1: ASVO design, current status, and future plans

REFERENCES

1. http://www.asvo.org.au
2. Keller S. C. et al., 2007, PASA, 24, 1
3. http://skymapper.asvo.org.au
4. http://www.star.bris.ac.uk/~mbt/topcat/
5. Bernyk et al. 2014, pre-print: http://arxiv.org/abs/1403.5270
6. Duffy, Meyer, Staveley-Smith et al 2012 MNRAS 426 3385D
7. Visualisations of the simulations at: http://vimeo.com/53066502
8. Keller et al. Nature 506, 463–466 (2014) doi:10.1038/nature12990

ABOUT THE AUTHOR(S)

Dr Yeshe Fenner is the Executive Officer at Astronomy Australia Ltd, having obtained her PhD in astrophysics before moving into project management roles in other areas of science and biomedicine. Previous roles include: Project Manager of a women's health study at Murdoch Childrens Research Institute, Operations Coordinator/Project Manager at the National Resource for Imaging Mass Spectrometry, Harvard Medical School, ITC Postdoctoral Fellow at Harvard-Smithsonian Center for Astrophysics, and astrophysics PhD at Swinburne University.

A/Prof Darren Croton is an Associate Professor and QEII Research Fellow in the Centre for Astrophysics and Supercomputing at Swinburne University of Technology in Melbourne, Australia. He primarily works on theoretical aspects of the formation and evolution of galaxies in the nearby and distant Universe, using both supercomputer simulations and large observational data sets. A recent focus has been the "big data" challenges facing astronomy, brought about by the ever increasing sensitivity and size of the telescopes and supercomputers astronomers routinely use. Dr. Croton is the scientist leading the development and operations of ASVO-TAO.

Jon Smillie is a Senior Data Specialist Programmer at the National Computational Infrastructure. He has considerable experience with data analysis, database design, and software development for astronomy and other research projects involving large and complex datasets. Jon led the technical development of ASVO-SkyMapper at NCI.