Mem. S.A.It. Vol. 80, 352

**Memorie** della

# Science in the Virtual Observatory

M.G. Allen

Observatoire de Strasbourg UMR 7550 11 Rue de l'Universite, F-67000 Strasbourg, France
e-mail: allen@astro.u-strasbg.fr

**Abstract.** The Virtual Observatory opens up new possibilities for doing science with large databases and on-line resources. Demonstrations and early science use of the VO have been important for the development of VO interoperability standards, and early use and prototypes are equally important for integrating theory into the VO. Here I describe some of the science drivers, early science demonstrations, and selected VO-enabled science papers along with some of the issues related to the comparison of observations with models and simulations using the VO.

**Key words.** Virtual Observatory - Science - Software - Simulations - Observations

## 1. Introduction

The Virtual Observatory (VO) addresses the imminent data volume and computational challenges facing astronomy, and aims to provide a new kind of research environment for scientific research based on data discovery, efficient access to data, and interoperability. The VO is ultimately science driven, with the development of science requirements, and early usage of the systems, occurring in step with the technological developments. Science requirements are guiding development of VO technologies within the various VO projects around the world, and science demonstrations of the VO tools and services are important for informing the astronomy community of new scientific capabilities. These new capabilities have been used to generate the first published science papers using VO technologies.

In this proceedings, I describe some of the science drivers being used to guide VO developments in the European EuroVO project, and the requirements they place on the VO. Then I highlight some early VO science demonstrations and selected published examples of VO enabled science, including the use of theoretical models and simulations, and some of the challenges involved for more detailed use of theory data and services.

## 2. Example use cases and science drivers

There are a number of generic 'use cases' for the VO. For example, the ability to collect 'all' the information on a given astronomical object via a relatively simple query. The results would include entries from telescope observation logs, entries in catalogues and publications, derived measurements, and science ready data products. In addition to such positional (or name based) queries, it should also be possible to perform queries for objects with other constraints, such as the brightness of objects, or their colours, morphology, observation date, uncertainty limits and so on. Also highly sought in the VO is the ability to combine data from across many different multi-wavelength archives, to build up a spectral energy distribution (SED), taking into account aperture affects, instrumental sensitivities and resolution. The vision of the VO is also that it provides the capability for managing large $10^7$-$10^{10+}$ size samples so that data mining techniques can be used for finding rare or unexpected objects, often referred to as out-lier science. Other high volume type capabilities include being able to re-analyse the SDSS or MACHO data bases. 'Virtual observations' are also a key idea, whereby theoretical models are 'observed' in a way the mimics what a real telescope measures as a way of comparing models and observation.

In Europe the national and agency VO projects are coordinated by the EuroVO project. This followed the earlier phase-A study called the Astrophysical Virtual Observatory (AVO) project. In 2005 AVO generated a set of formal science goals for the VO, called the Science Reference Mission (SRM). This outlined a set of key scientific results that a European VO should be able to achieve when fully implemented. It consists of science cases over broad range of astronomy and the related requirements. The science cases, listed below range over different areas of research with Galactic, Extragalactic and Cosmological science cases (www.euro-vo.org/internal/Avo/AvoSRM/srm.pdf).

- Circumstellar disks: from pre-Main Sequence stars to stars harboring planets
- Intermediate Velocity Clouds
- Which star will go Supernova next?
- Initial Mass Function within 1 kpc: from planetary to stellar masses
- Initial Mass Function for massive stars
- The contribution from low and intermediate mass stars to the interstellar medium
- Galaxy Formation and Evolution from z = 10 to z = 0:1
- Build-up of supermassive black holes
- The Formation and Evolution of Galaxy Clusters
- Correlation of Cosmic Microwave Background, radio/mm and optical/NIR Galaxy Surveys

While the scientific goals of the SRM programs are diverse, there are many common aspects to these programs and these provide important requirements for VO developments in order to enable a wide range of science. The most common aspect is the requirement to browse and search for data and distributed information. Other common aspects are

- The management of large amounts of distributed heterogeneous data.
- Cross-Matching of catalogues - This includes catalogues that may be large or small, sparse or dense. Cross matches would need to take into account the positional uncertainties, resolution, completeness as well as extra constraints such as colour or object type. Some of the science cases also require taking into account the environment, i.e. such as crowded fields for stars.
- Combining multi-wavelength data - taking into account different units, photometric systems, spatial, wavelength and time coverage resolutions observing technique.
- Compare observations with models - including making 'virtual observations of models. A common requirement across many areas of galactic and extragalactic astronomy is spectral fitting and classification using (stellar and stellar population synthesis) models.

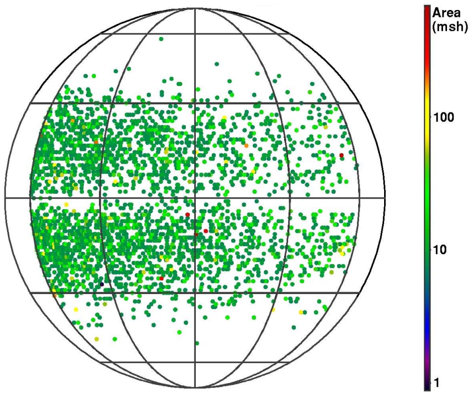


**Fig. 1.** Centre-to-limb variation of sunspot visibility - Dalla et al. (2008)

## 3. Science demonstrations and early VO science

Science demonstrations are an important way of showing the capabilities of the VO to the astronomy community, and also help guide future development. Here I briefly outline some of the VO demos and early science use.

An early VO project, AstroVirtel, had the goal to exploit archives as virtual telescopes.

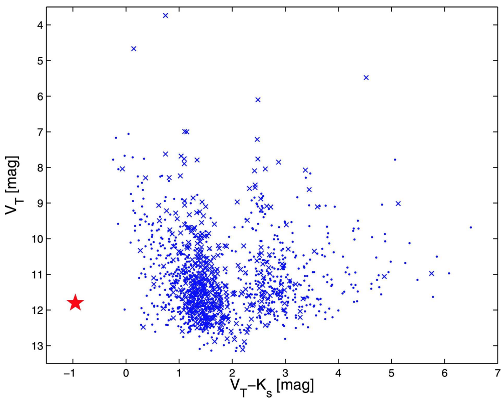


**Fig. 2.** Out-lier 'candidate white dwarf/ He subdwarf found in - Caballero & Solano (2007)

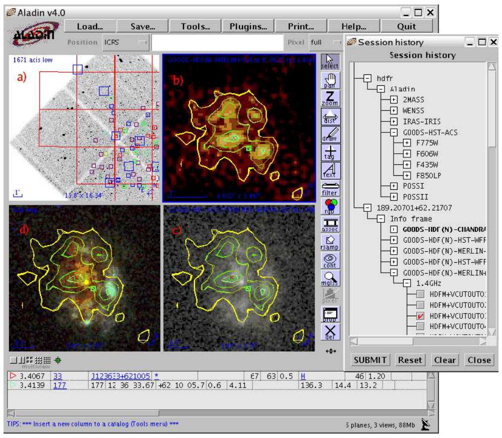

**Fig. 3.** The Aladin tool being used to investigate the starburst candidate J123633+621005, see - Richards et al. (2007)

The project involved 3 proposal cycles 2000-2003 and performed archival projects using HST, ESO: VLT, NTT, WFI, and ISO archives. This involved advanced searching of these archives including constraints on exposure, seeing, field density and location, e.g., for pre-discovery images for asteroids (www.euro-vo.org/astrovirtel/index.html). Some 20+ publications related to this project are listed listed on the AstroVirtel pages.

The US National Virtual Observatory (NVO), using prototype VO systems in a demonstration project identified a new L-type Brown Dwarf. This rare object was found via a cross-correlation of the SDSS and 2MASS surveys in a demonstration project that was designed to reproduce known (∼ 200) Brown Dwarfs. These surveys involved some 15M SDSS sources, and 160M 2MASS sources (see – www.us-vo.org).

The European AVO project held three major demonstrations of VO capabilities between 2003 and 2005 (www.euro-vo.org/avo/). The 2003 *'First Light'* demonstration showed how multi-wavelength data from remote archives could be combined for identifying Supernovae candidates and high redshift galaxies, and how data processing could be performed on remote computing facilities. The 2004 *'First Science'* demonstration the AVO prototype was used to identify high redshift type 2 QSOs in the two Great Observatories Origins Deep Survey (GOODS) fields, and provided one of the first scientific results from VO science, see Padovani et al. (2004). The third *'Toward the Euro-VO'* demonstration showed the use of workflows to combine multiple VO services and tools to generate Spectral Energy Distributions (SEDs) of galaxies from multiple data archives, and to compare these with model spectra.

## 4. Published results of VO enabled science

As VO tools and services mature, and as the community becomes more aware of the capabilities of the VO, more VO-enabled science is appearing in the literature. A list of VO enabled science papers is maintained on the EuroVO web pages (www.euro-vo.org/pub/fc/papers.html). These papers are identified from journal and Astro-ph lists as using VO tools and services to some extent. While this list cannot be complete, it is a best effort to highlight some of the VO enabled science. Here I summarise a few selected examples to highlight the diversity of science being done with VO tools.

Papers on the subject of solar astrophysics have been published by Dalla et al. (2008) and Dalla et al. (2007). The first of these used AstroGrid tools to analyse a sample of 6862 sunspot regions, using databases of solar data to locate where sunspots appear and disappear on the solar disk. They are able to demonstrate that the visibility of small sunspots has a strong centre-to-limb variation (Fig. 1), and they find that the duration of the growth phase of solar regions has been previously underestimated. The second paper uses the AstroGrid workflow capabilities to analyse 2115 new sunspot regions obtained from the USAF/Mount Wilson catalogue of sunspot regions. They find that paired regions are more flare productive than isolated ones.

Caballero & Solano (2007) reports a bright white dwarf candidate found serendipitously in an optical-NIR study of young stars using the Tycho-2 and 2MASS catalogues. In their work they used the VO tool Aladin (http://aladin.u-strasbg.fr/aladin.gml) to analyse images, spectra and catalogues available from VO services. Figure 2 shows the out-lier which is a candidate white dwarf or He sub-dwarf.

Richards et al. (2007) present a detailed study of distant radio emitting starburst galaxies that also host AGN. They used VO tools (Vizier - http://vizier.u-strasbg.fr), Aladin and Topcat (www.star.bristol.ac.uk/~mbt/topcat) to select sources from catalogues of radio, X-ray and optical data, and make use of VO interoperability features that allow separate software tools to work together (Fig. 3). Their results show that star formation at $z > 1.5$ is closely linked to AGN activity and that it is also triggered differently compared with star formation at lower redshifts.

Santos et al. (2007) have used VO tools to search for 'fossil groups' in the SDSS. Fossil groups are systems with mass and X-ray luminosity comparable to galaxy clusters, but whose light is dominated by a single isolated elliptical galaxy. Their study involves a cross-match of all SDSS luminous red galaxies with X-ray source in the ROSAT All-sky survey, and a search for neighboring galaxies within a specified radius, leading to 34 candidate fossil groups. Their searches were performed using the NVO OpenSkyQuery capabilities.

## 5. Comparison with theory

Some of the early VO demonstrations showed comparison of observed data to theoretical models, for example the comparison of SEDs of extreme red and blue galaxies with theoretical spectra from stellar population models in the AVO demos (see Fig. 4). These demos showed how the distributed data, and models could be brought together technically, but the demos were relatively immature in the application of the models to the data. Since then there has been much progress in the use of theory 'data products' in the VO, both in terms of description of the models, and the tools with which to use them. This is of course the subject of this workshop, where many new advances have been reported.

One particular aspect of comparing models to data which has application in many of the science reference mission projects, and other VO science drivers, is the need to compare stellar or galaxy models to observed spectra and SEDs. Recently

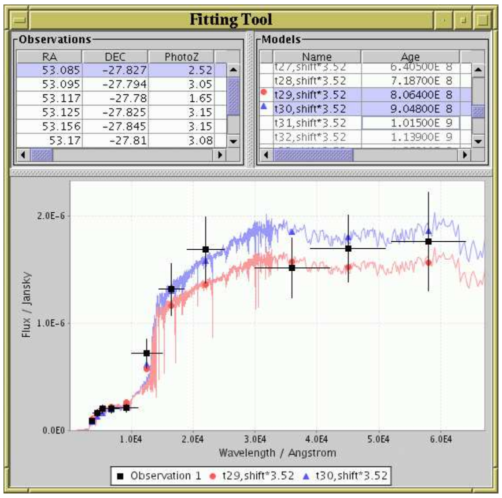


**Fig. 4.** Fitting theoretical spectra models to observed photometric data.

Bayo et al. (2008) have developed a workflow procedure for fitting synthetic spectra or photometry extracted from theoretical models to multi-wavelength observational data. They have applied this to Collinder 69, a stellar association, to derive physical parameters such as $T_e$, gravity, luminosity, etc. for the 170 candidate members to Collinder 69, providing an upper-limit for the age of this stellar association. The key step of their work-flow is performed by a new VO-tool, VOSA (http://svo.laeff.inta.es/theory/vosa2/), which performs all the steps in a VO environment, benefitting from the interoperability and data access that this provides.

In addition to being a useful tool, the VOSA effort, and similar projects are helping to develop the requirements for expressing the quantities and parameters needed when making detailed comparison of theoretical spectra to observed spectra and photometry. Many of the requirements for comparison of theoretical models with observed data requires a higher level of detailed scientific interoperability than is currently available, for both the observed and theory data products. Some of the issues are currently being tackled by the data model efforts in the IVOA, defining the essential elements of the description or 'metadata' required to characterize an observation in terms of parameters such as exposure time, sampling, resolution and details of the photometry or spectral measurements. Likewise, the Theory Interest Group of the IVOA is making progress toward standards that may be used to de-

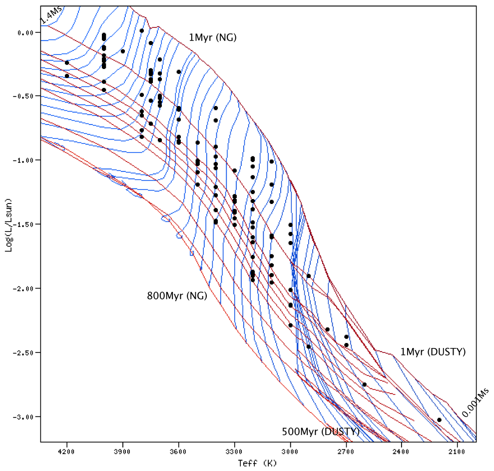


**Fig. 5.** Comparison of observation with stellar models, see - Bayo et al. (2008)

scribe theory resources, and ways to make models and simulations more generally interoperable. It is very important that the development of the interoperability standards for publishing theory models in the VO be closely linked to the real services that will provide them to the astronomy community. This meeting, with VO and Theory specialists has helped to build links within the community so that theory may be more closely integrated into the VO.

## References

Bayo, A., Rodrigo, C., Barrado Y Navascués, D., Solano, E., Gutiérrez, R., Morales-Calderón, M., & Allard, F. 2008, A&A, 492, 277
Caballero, J. A., & Solano, E. 2007, ApJ, 665, L151
Dalla, S., Fletcher, L., & Walton, N. A. 2008, A&A, 479, L1
Dalla, S., Fletcher, L., & Walton, N. A. 2007, A&A, 468, 1103
Padovani, P., Allen, M. G., Rosati, P., & Walton, N. A. 2004, A&A, 424, 545
Richards, A. M. S., et al. 2007, A&A, 472, 805
Santos, W. A., Mendes de Oliveira, C., & Sodré, L. J. 2007, AJ, 134, 1551