Mem. S.A.It. Suppl. Vol. 9, 423

Memorie della
Supplementi

The Virtual Observatory in Italy: status and prospect

R. Smareglia[1], F. Pasian[1], U. Becciani[2], G. Longo[3], A. Prete Martinez[4], A. Volpicelli[5]

[1] INAF/O.A.Trieste, Via Tiepolo 11, I-34131 Trieste, Italy
[2] INAF/O.A.Catania, Via S.Sofia 78, I-95123 Catania, Italy
[3] Univiversita' di Napoli Federico II, Dip. Scienze Fisiche, Via Cintia, I-80126 Napoli, Italy
[4] INAF/IASF Via del Fosso del Cavaliere100, I-00133 Roma –Italy
[5] INAF/O.A.Torino, Via Osservatorio 20, I-10025 Pino Torinese(TO), Italy

Abstract.
With the increase of the potentiality of the astronomical research system through the use of new and expensive tools on multi-wavelength, both in the space and ground structures and through the availability of simulated data too, scientists are now fully aware that this amount of investments require to maximize the exploitation of such tools and their scientific return. A typical theme to face in this attempt is the need to preserve the patrimony inherent in the acquired data. The data must be acquired, described, processed and preserved in order to make possible their reuse from the future generations of researchers and scientific programs whose purposes can be different from those that have originally pushed the acquisition of such data.
The Virtual Observatory (VO) is the middleware that will allow the astronomers of the XXI century to fully perceive the return of their investments through the setup of long-term data archives. The VO, therefore, becomes a key element of every new astronomical facility .
The first Italian step toward this philosophy is the creation of the "Italian Center for the Astronomical Archive" (IA2). IA2 is the point of contact for the Italian astronomic communities managing catalogs and astronomic archives, so that related data are federated and usable from the national and international astronomers in the VO philosophy. INAF is also involved in the Euro-VO project; it will make available to the astronomical community tools, developed in Italy, that guarantee the interoperability of data like visualization and data mining tools.

1. The Virtual Observatory: this stranger

When a new tool is developed, the purpose for which this has been created should not be forgot. The astronomer could like to get information on a particular set of sources: know if an object is a source nominee to be an extensive GRB; try to check if the X-flux of a source could be contaminated from the background; check if there is a radio flux of this source and check if there is near a SNR. One more exam-

Send offprint requests to: R. Smareglia ple could be getting information on a class of sources, know what of them are QSO and what is the average X-flux of the B stars in the sample.

Therefore, when we want to make correlations between data of different archives and catalogs (which are the bright objects in the SDSS that don't result in 2MASS and vice versa; which is the statistic excess of the IR-flux in the position of some X sources?), the problem is the interoperability of such different archives. This is the reason pushing astronomical data providers to ask for the creation of what it's currently known as Virtual Observatory (VO).

It is important to understand exactly what the VO is. The VO is a suite of international standards for the exploration, exchange, interoperability and analysis of astronomic data through the web. It also foresees the creation of a working environment for the access and the analysis of data exploiting the emergent technologies tied to the Grid; such data processing environment, in turn, make easy to undertake training activities and encourages the use and the development of new algorithms. The VO, furthermore, is suitable to become a tool normally used by every researcher for the scientific planning allowing the identification of gaps in particular spaces of the parameters. The astronomer, in this way, is able to identify and develop new missions, tools or experiments that could have great impact in the scientific research. Finally, the VO is able to become a good vehicle for the education and the scientific dissemination.

It is very important to stress and well understand that the Vo isn't a substitute to the construction of new telescopes and tools or a centralized data center.

The VO concentrates its attention on aspects like how data are seen by the international community both from the point of view of applications accessing them and answers got from the given requests. The implementation techniques of the underlying services and how data are saved and manipulated are not relevant.

2. IVOA Architectural

To develop a real VO, all national VO infrastructures are federated in an alliance: International Virtual Observatory Alliance IVOA (http://www.ivoa.net), whose target is supporting and developing the worldwide Virtual Observatory. The whole set of information about the IVOA structure is well illustrated in the IVOA web site where the most relevant technical articles and the full list of involved people can be found. In any case, we want to focus our attention on some part of the IVOA architecture. First of all the Data Access Layer (DAL). It's the mediator between the applications of the VO consumers and the data provider. It uses some Web Technologies (Web Services and SOAP protocol) and VO Technologies (Astronomical Data Query Language, Cone Search, Single Image Access Protocol and Spectral Image Access Protocol) developed ad hoc from VO people.

After that we must remember some basic standards on top of which the VO is built:

Resource Metadata and Registry: they use standards developed inside the Open Archives Initiative (http://www.openarchives.org) for the metadata collection, and describe data and services that are gathered in the registry;

Uniform Content Descriptors (UCD) which provide a common "dictionary" able to describe the contents of catalogs;

VOTable: starting up from the XML standard, internal or external data can be included. It's a base protocol to access catalogs and images, and it is complementary with the FITS format.

All these standards already have a first well defined version; as the web technology is continuously evolving, however, they are evolving to improve and increase their performances too.

3. Italian Archives and VO

A list of Italian archives and catalogs, exhaustive enough, of astronomical data can be found on http://www.ccra.inaf.it/archivi/. Here we can perceive that a national awareness to make public the acquired, processed and simulated data, already exists. The integration of such data within the VO, however, is still not mature. In this last year, however, the italian community has become more present in this important international initiative. The already VO-compliant archives are the ”Italian Center for the Astronomical Archives” (IA2) and the SOLARNET. Some INAF personnel, moreover, is already involved in international VO initiatives like developing the Data Access Layer (DAL) for the Galex and the GSC2 archives. Least but not last INAF is present, as an active partner, within the European VO (Euro-VO) with Vo-Tech and VO-DCA projects.

3.1. IA2

Starting up from the experience done in the development on the prototype of the Long Term Archive of the National Telescope Galileo (TNG/LTA) in 2002, in the beginning of 2005 INAF has financed the creation of the IA2 (http://wwwas.oat.ts.astro.it/IA2/). The purpose of IA2 is the creation of an infrastructure able to archive the whole data of the TNG: raw and calibrated scientific data, as well as housekeeping data. It is foreseen to have the centre ready for the community in the autumn 2005 with a standard web interface; some work is still in progress to become VO-compliant in a very short time. The goal is making it an italian reference point for the VO technology reducing in this way the gap that INAF has with respect to other international data providers. The installation of some main VO services like the Italian VO registry is also foreseen.

Thanks to this acquired experience, the IA2 will store the raw data coming up from the Italian time of the Large Binocular Telescope (LBT) when these will become available in late 2005.

3.2. Solarnet

Another Italian experience based on the VO technology is the national project SOLARNET (SOLar ARchive NETwork) (http://solarnet.to.astro.it:8080/portal/). The aim of this initiative is to federate all the Italian solar archives as a distributed database, which is the first step toward an Italian Virtual Solar Observatory, interconnecting the distributed resources and the whole set of available solar data in a unified database by a web user interfaces. Different user interfaces allow searches of all participating data services using different input parameters. Currently there are 5 data providers in SOLARNET: SOLAR , SOLRA , PSPT , DISCO/VAMOS and CATANIA archives, plus two services that the Italian solar comunity has developed for the EGSO project: SEC (Solar Events Catalog) and DSO (Database for Solar Observatory).

3.3. Euro-VO

As stated above, INAF is already involved in the European Virtual Observatory (Euro-VO) initiative (http://euro-vo.org). It is an integrated and coordinated program designed to provide the European astronomical community with tools, systems, research support and data interoperability based on Vo philosophies. The project is based on the research, development experience and prototypes produced in the FP5 RTD entitled: ”Astrophysical Virtual Observatory (AVO) Project”. The Euro-VO program seeks to support and deploy VO capabilities to data centres and observatories across the entire electromagnetic spectrum. It will therefore be closely coupled with the two other major integrating and networking activities for astronomy in FP6: OPTICON and RADIONET. The Euro-VO will consist of a new organizational structures which will meet the goals of the total work program and which will provide a platform for a long term European VO research infrastructure and capability.

One of this structures is the Euro-VO Technology Centre (VO-Tech http://wiki.eurovotech.org): a distributed organization that coordinates a set of research and development projects on VO technology, systems and tools. The Vo-Tech is FP6 funded. The project is divided up into a six Design Study areas. INAF people are involved in three of this:

DS4: UserTools (OATrieste, OACatania) This task will produce designs and prototypes for new VO-compliant end-user tools. We will produce a list of suggested priority tools. This will partially be based on the experience and feedback of precursor projects and will require inputs from the scientific user community;

DS5: ResourceDiscovery (IASF Roma) The first goal of DS5 is to undertake a feasibility study for developing components based on emergent technologies in the areas of the Semantic Web and Ontologies. This study will be used as a reference for actual component designs and trial implementations, and ultimately developments of standards at the IVOA level;

DS6:DataExploration (Data Mining and Visualization)(Univ. Napoli Federico II, OACatania) The aim of DS6 is to complete all technical preparatory work necessary to enable effective data exploration within the EuroVO. This will involve assessment of a range of data mining and visualization algorithms and packages, with a view to determine how they can be run as distributed services, how they can be made VObs-compliant and how they can be extended to extremely large datasets.

One additional structure in which INAF is involved and that will be submitted on FP6/EU tu be funded will be the The Euro-VO Data Centre Alliance (VO-DCA): a collaborative and operational network of European data centres publishing data and metadata to the Euro-VO and providing a research infrastructure of GRID-enabled processing and storage facilities.

3.4. Grid and VO

When we are talking about the VO it is easy to make notice of its possible interactions with the Grid: the first as a data services provider, the second as a computational services provider. The computational grid, indeed, is a distributed system, complementary of the High Performance Computing system, where it is ”easy” to run a lot of independent jobs. It is possible to setup an easy way to connect the Grid and the VO so that their interaction and integration results are immediately clear: these jobs can use the data coming from the distributed archives and the results can, in turn, be saved in the archives; such archives may be inside or outside the grid but in any case they must be VO compliant.

Also in this field INAF is present with financed projects: from the FIRB (Grid.it, WP10), from the PRIN03 (DRACO, third more evaluation project in the field of the physics) and in EGEE (financed by EU/FP6 where INAF is an unfounded member). Starting up from our experience handling issues related to archives, one of the main targets of these projects is the Gird-VO integration.

4. Conclusions

The VO is the main infrastructure that will allow the astronomers of the XXI century to fully perceive the return of the investments through a correct long-term maintenance of the data. Therefore, with this view the VO becomes a key element for every new Astronomical facility so that the maximum scientific return of the investments is assured and the scientific knowledge is distributes to the whole community.

Acknowledgements. I am grateful to all the researchers that are working in the VO and Grid area for the excellent work done and for allowing me to present this view at the italian comunity.

References

Becciani, U., Gheller, C. et al. 2003 AstroMD, a tool for Stereographic Visualization and data analysis for astrophysical data; Mem Sait Suppl, Vol.1, 80

Pasian, F., Smareglia, R. et al. 2003 INAF Archives in the Framework of the Astronomical data grid; Mem Sait Suppl, v.3, p. 368 2003

Longo, G. et al. 2002 Data mining of large astronomical databases with neural tools; Spie, Vol 4847, pp 265,276

Smareglia, R. et al. 2003 The pilot project for the TNG Long-Term Archive; Mem Sait Suppl, Vol.74, 514

Williams, R., Hanischet, B. et al. 2004 Virtual Observatory Architecture Overview; Version 1.0 IVOA Note 15 June 2004