ESA/HST and ESO archive access and visualisation in the VO era

D. Sforza [ESA/Space Telescope European Coordinating Facility (ST-ECF)]
Karl Schwarzschild strasse 2 - 85748 Garching bei München (Germany)
dsforza@eso.org

A. Micol [ESA/Space Telescope European Coordinating Facility (ST-ECF)]
Karl Schwarzschild strasse 2 - 85748 Garching bei München (Germany)
amicol@eso.org

T. Boch [Centre de Données astronomiques de Strasbourg (CDS)]
Observatoire de Strasbourg, UMR 7550, 11 rue de l'Université, 67000 Strasbourg (France)
boch@astro.u-strasbg.fr

D. Durand [National Research Council Canada, Canadian Astronomy Data Center]
National Research Council Canada, 5071 W. Saanich Road Victoria, B.C. Canada V9E 2E7
Daniel.Durand@nrc.ca

Turning the ESA Hubble Space Telescope Archive along with ESO Archive Facility into a VO-compliant data centre is a complex process: new technologies and standards have to be adopted and a more detailed description (metadata) of the data holdings is required. While Astrovirtel helped us understanding and exploring ways to achieve the latter, the current world-wide effort of developing VO-compliant software can allow us to provide with limited in-house effort the aforementioned metadata through a modern VO interface. This technology has proved to be mature enough to be used in the presented example, where HST-ACS & HST-WFPC2 associations (a ST-ECF, CADC, STScI project) metadata and field of view are visualised with Aladin. This is not only a useful application for the end user but also a first step of the on-going effort toward a VO-compliant Archive. In spite of its novelty and being still under development, VO technology is mature to cope with real life scenarios, allowing the implementation of powerful concepts like the metadata tree presented in our example

ESA, ESO, ST-ECF, HST, Hubble Space Telescope, ACS, WFPC2, Astrovirtel, IVOA, AVO, VO, Virtual Observatory, Interoperability, Data Mining, Metadata, Metadata Tree, IDHA

1. INTRODUCTION

VO paradigm stands on two main pillars: on one side standards and tools used to exchange and analyse data, and on the other the existence of valuable data provided to the community. The two sides are equally important, because for the end users the best tool's worth is measured against its ability to provide data, and the best data are fully exploited if they can be retrieved and integrated together. While developing standards and related technologies is a distributed effort, embraced by the VO community [L1],[L2] as a whole [R1], offering good quality data is a task specific to the data provider that manages the data. Under the VO view, "good quality" is not only something referred to the sensitivity or precision of a detector, but also takes the meaning of "well-characterised" data: i.e. metadata associated with data. Detailed and standardised metadata are useful in a number of different situations, as:

1. They offer a homogeneous description for data from different archive.
2. They allow interoperability between software tools.
3. They help translating the "engineering" view in which data are stored into the "scientific" view in which they are actually used.
4. They allow better and more scientific archive search engines, as it is shown by the experience of Querator (a data mining software tool) [L3].

Of course, all of the above is actually possible only in presence of standards to define metadata, as for instance the Unified Content Descriptor (UCD, see [L4] and [R2]), and to exchange them, as the VOTable format [L5]. Another issue to be considered is also which metadata are actually useful, and this is a question that can be answered only with the involvement of the scientific community, i.e. the final user of the data. Experiences like Astrovirtel [L6], [R3] a project founded by the European Commission and aimed at improving access to the ESO/ST-ECF archive, helped us in this respect.

2. EXPLOITING VO

There is a great deal of expectation on what the VO will be able to do for the astronomy community in the future, but what is currently possible using the already existing VO frame? We present here an example. The interested reader

can also download, for a detailed presentation of current VO status, the AVO DEMO tool from the link in the references [L7].

2.1 VO In Action

Astronomers may need multi-wavelength observations of the same astronomical object. These observations come likely from different data providers, are taken by different instruments, with different telescopes and different configurations. Visualising them in a homogeneous way involves transforming between different units and coordinate systems, taking into consideration instrument characteristics and filters, matching size, orientation and scale of each observations. These helpful tasks are only preliminary to the real scientific job, but require time-consuming efforts together with a specialised knowledge on each archive/instrument. In our example we use the Aladin software tool [L8] to superimpose both the Eta Carinae (Figure 1 to 4) and the NGC6397 globular cluster (Figure 5, 6) images, taken from the DSS digitised sky survey, with the Field of View (FOV) of the matching observations taken with the HST-WFPC2 instrument [L9]. The ST-ECF archive provides the observations with all the needed metadata so that Aladin can automatically take care of most of these preliminary tasks. Metadata of the HST-WFPC2 observations (as used filter, size, orientation, and so on) can be easily browsed using a metadata tree, where they are logically grouped in hierarchical view that helps the user to choose the subset of observations suitable to their scientific work. The example clearly shows how the VO paradigm helps astronomers: The data provider, ST-ECF in our case, can focus on providing better data characterisation while the existence of a powerful VO tool for visualisation like Aladin helps to make the best use of these data.

2.2 The Metadata Tree

The data access in the example is based on the IDHA protocol, used to implement the hierarchical access of a Metadata Tree [L10]. The Metadata Tree that was developed for the AVO demo and represents an implementation of a scalable, hierarchical metadata description of an image dataset. Using a VOTable description, it displays a tree view (i.e. a logical organisation) of the available metadata. The three contexts where the metadata tree concept has been successfully used are:

1. Associations of data, like the WFPC2 associations.
2. ESO/WFI (Wide Field Imager) multi-chip images (8 CCDs).
3. Datacube access like the Canadian Galactic Plane Survey.

In Figure 3 the Eta Carinae data tree shows in the Aladin branch the available images, while the HST/WFPC2 Associations branch presents in a hierarchical view all the WFPC2 associations available in that field, organised per filter band pass. More details and capabilities on individual observations are also available in pop-up panels, like the button to overlay the field of view onto the current selected image. Going one level up on the metadata tree hierarchy, the metadata belonging to a selected filter are presented, as shown in Figure 4.

FIGURE 1: DSS Eta Carinae image visualised in Aladin.

Figure 2: Eta Carinae image with all the WFPC2 FOV associations superimposed with the proper scale and orientation.

Figure 3: Metadata tree and detailed information for WFPC2 associations.

Figure 4: A filter is selected on the metadata tree and its metadata are shown in a popup panel.

Figure 5: Selecting a WFPC2 association for NGC6397 globular cluster (FOV in green)

Figure 6: Preview of the selected WFPC2 association.

3. CURRENT VO ACTIVITIES AT ST-ECF

VO development is at the technology cutting edge. This alone is an issue to cope with, but the real challenge for an archive group (i.e. a data provider) is embracing the VO technologies while at the same time keeping up with all the current activities and maintenance of legacy applications, required by the daily operations. While a massive effort is on going to generate well characterised enhanced data products, here we concentrate on the infrastructure required to offer such products in a VO compliant way to the community, and on the standards and technologies that a data provider moving into the VO frame is required to adopt. In this respect our main areas of development are:

1. Use of VOTable, i.e. a standardised way to offer metadata together with data for input/output operations.
2. Linked with the above is the actual ability to access metadata from software applications, in order to offer them as part of the output VOTable. This implies building a (database) system that is able to store and retrieve all the needed metadata.
3. On the Web Technology side, we need the ability to offer and use web services. We believe that using Apache Tomcat (servlet container) [L11] and Apache Axis (SOAP engine for web services) [L12] is a feasible option. In particular, Axis offers useful and easy to use tools to handle the WSDL files [L13] needed for web services. They are both open source software packages.
4. Publishing services and data products to the VO Registry. That means implementing a standard query interface for the user, i.e. compliance with the IVOA SkyNode interface [L14].

The graphical representation of the above activities given in figure 7 helps visualising how they integrate together and how they stand in respect to our current status.

3.1 Development Layer

There are a number of technical issues that we are evaluating, as briefly reported in the following. The preferred development platform would be Linux, but we need to retain sun solaris capabilities for the legacy applications. This alone is a hint for using a cross-platform software language like Java. Moreover VO development is nearly all java-based: a huge, increasing number of java packages developed by the VO community are freely available and we are aiming at the maximum level of reusability we can achieve. Another point in favour of java is the existence of an SQL-level API like JDBC, that greatly help database management from java applications. Last but not least, a development environment with the same look and feel on both our platforms is the natural choice, and among the possible options we are currently oriented towards the Eclipse IDE [L15], that appears easy to learn and use, and is also a free available software.

Figure 7: Present and Future of our Archive.

4. CONCLUSIONS

VO Infrastructure and standards are under development, meanwhile VO is already mature enough to provide useful features to the astronomer community. Becoming a VO-enabled data provider is a challenge that requires careful planning, but it is feasible and the road map is well established.

USEFUL LINKS.

- [L1] VO International: <http://www.ivoa.net>
- [L2] Euro VO: <http://www.euro-vo.org>
- [L3] Querator data mining tool: <http://archive.eso.org/querator>
- [L4] UCDs - Unified Content Descriptors: <http://cdsweb.u-strasbg.fr/UCD>
- [L5] VOTable Format definition: <http://www.ivoa.net/Documents/latest/VOT.html>
- [L6] Astrovirtel project: <http://www.stecf.org/astrovirtel/>
- [L7] AVO Demo: <http://www.euro-vo.org/twiki/bin/view/Avo/SwgDownload>
- [L8] CDS Aladin interactive sky atlas: <http://aladin.u-strasbg.fr>
- [L9] HST-WFPC2 associations:
 - at ST-ECF: http://archive.eso.org/archive/hst/wfpc2_asn/3sites
 - at CADC: <http://cadwww.hia.nrc.ca/wfpc2>
 - at STScI: <http://archive.stsci.edu/hst/wfpc2/index.html>
- [L10] IDHA (Images Distribuees Heterogenes pour l'Astronomie) Metadata Tree:
 - http://www.euro-vo.org/internal/Avo/MetadataTree/MetadataTree_v1.1.txt
- [L11] Apache Tomcat: <http://jakarta.apache.org/tomcat/index.html>
- [L12] Apache Axis: <http://ws.apache.org/axis/ref.html>
- [L13] WSDL files: <http://www.w3.org/TR/wsdl>
- [L14] SkyNode Interface: <http://www.ivoa.net/internal/IVOA/IvoaVOQL/SkyNodeInterface-0.7.4.pdf>
- [L15] Eclipse IDE: <http://www.eclipse.org/platform/>

REFERENCES.

- [R1] Allen, M. G., Genova, F., Ochsenbein, F., Derriere, S., Arviset, C., Didelon, P., Dolensky, M., Garrington, S. T., Mann, R. G., Micol, A., Richards, A. M. S., Rixon, G. T., & Wicenec, A. 2003, in ASP Conf. Ser., Vol. 295 Astronomical Data Analysis Software and Systems XII, eds. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook (San Francisco: ASP), 55 "Toward an AVO Interoperability Prototype".
- [R3] Derriere, S., Gray, N., McDowell, J. C., Mann, R., Ochsenbein, F., Osuna, P., Preite Martinez, A., Rixon, G., Williams, R., 2004 in ASP Conf. Ser. vol. 314 Astronomical Data Analysis Software and Systems XIII, eds. F. Ochsenbein, M. G. Allen & D. Egret, 315 "UCD in the IVOA context"
- [R3] A. Micol, M. Dolensky, F. Pierfederici, B. Pirenne 2002, in ASP Conf. Ser., Vol. 281 Astronomical Data Analysis Software and Systems XI, eds. D.A. Bohlender, D. Durand, T.H. Handley, 68 "Next Generation Science Archives: Lessons Learned for Archive Interoperability in the Virtual Observatory Era".