Astronomical Data Analysis Software and Systems XIII
ASP Conference Series, Vol. 314, 2004
F. Ochsenbein, M. Allen, and D. Egret, eds.

Organization of Data Sets in Virtual Observatories

Mark G. Allen, François Bonnarel, Thomas Boch, Pierre Fernique

Centre de Données astronomiques de Strasbourg, France

Mireille Louys

Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédetection, France

Abstract. Browsing and accessing local and distributed datasets is an important aspect enabling Virtual Observatories. We present an example implementation of a data tree in the Astrophysical Virtual Observatory prototype tool. This is a dynamically built data tree containing information on image datasets, based on the IDHA data model. We show how the meta-data representation of the GOODS dataset in the AVO demo allows efficient data browsing and selection, and how the tree may be used to access local and distributed data. This capability is being developed in the framework of the CDS Aladin image browser, and AVO prototypes.

1. Introduction

Browsing and visualization of image datasets will be an important part of Virtual Observatory operations. Such datasets may range from a small set of images stored on a local disk, to the tera-byte collections of modern surveys. Standardized and scalable descriptions of image metadata will be required to enable dataset browsing, selection and visualization. The data tree that was developed for the Astrophysical Virtual Observatory (AVO[1]) 1st year demonstration represents a prototype implementation of a scalable, hierarchical metadata description (in VOTable[2]/XML) for image datasets. The data tree mechanism allows *any* image data available via URL to be described and accessed via an hierarchical tree.

Here we provide a very brief outline of the image metadata description that forms the basis of this data access scheme, and show how the data tree is used in the AVO prototype to enable "smart browsing" and selection of data.

[1] `http://www.euro-vo.org`

[2] `http://www.ivoa.net/twiki/bin/view/IVOA/IvoaVOTable`

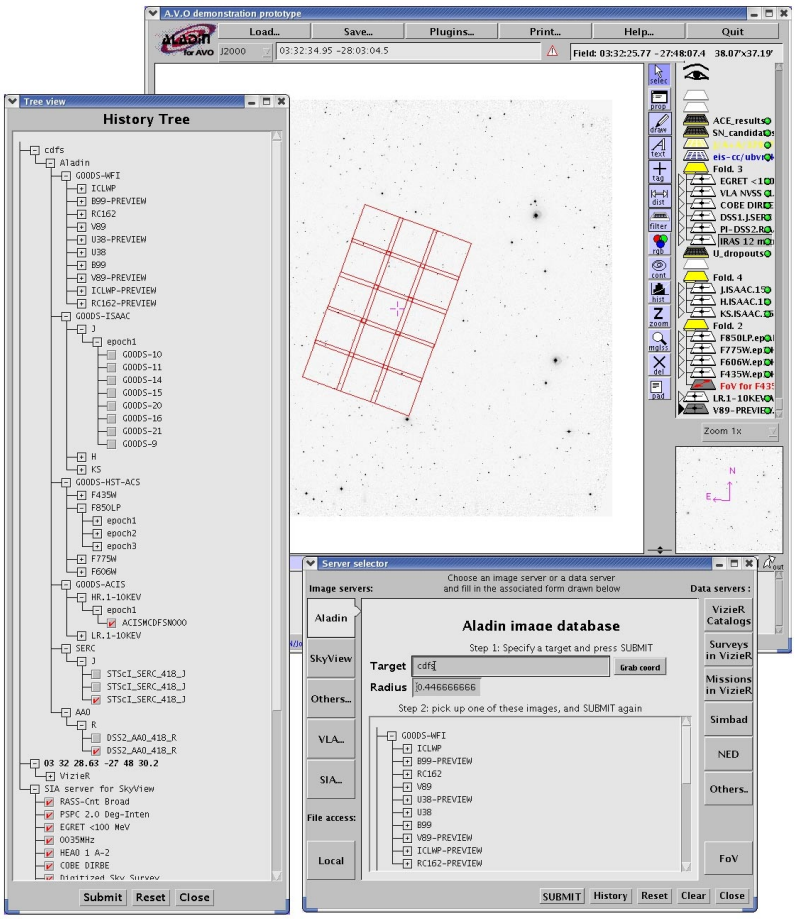

Figure 1. The Data Tree in the AVO Prototype. The lower right Server Selector window shows a request sent to the Aladin image server for data available in the CDFS, and the resulting list of images. This tree is presented in the Treeview window along with the results of subsequent requests. The fields of view of images in the tree are interactively shown in the top right visualization window, where the images can be selected and loaded into the image and catalog plane stack.

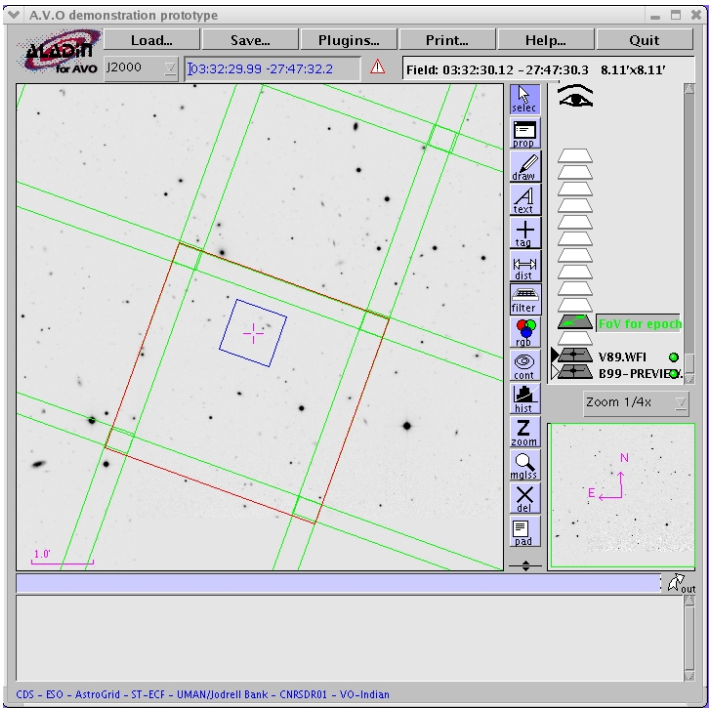

Figure 2. Field of view display of the GOODS ACS image tiles. The original image tiles are shown, along with the currently selected tile. The small square shows the outline of the image cut-out centred on the cursor position, that can be generated by the Aladin image server on request.

2. Image Metadata Description in VOTable

The information requirements for the VOTable description of the data tree have been designed to be flexible enough to allow for a simple listing of a set of images, or a rich description based on the IDHA[3] data model. Details on the required, and optional fields, including an annotated example are described on the Euro-VO web page. The IDHA data model seeks to provide a generic description of astronomical data. Using the model as a basis for image metadata description is beneficial because data described in a tree that conforms to the IDHA model schema can be organized into nodes corresponding to objects of the model. Also uniform description of the image metadata (such as 'coordinates' and 'observed wavelength') allows software interfaces to make use of this information for data browsing and selection. Such capabilities have been built into the AVO prototype.

[3] `http://cdsweb.u-strasbg.fr/idha.html`

3. The Data Tree in the AVO Prototype

Figure 1 shows the data tree interface in the AVO prototype. The hierarchical tree displays information on the image data available in a given region of the sky. In this example the tree was dynamically generated by the Aladin image server as a result of a request for data within a $\sim 0.5'$ radius of the Chandra Deep Field South (CDFS). The results show the GOODS data (Giavalisco et al. 2004) in the Aladin image server, including WFI, ISAAC, HST-ACS and Chandra images. The images are organized by instrument, bandpass and epoch and the interface provides a mechanism to interactively re-sort the nodes based on the metadata.

The metadata stored in the tree is utilised to provide an efficient means for "smart browsing" and selection of data. Browsing the cursor over the tree causes the field of view outlines for each tree node to be displayed. An example is shown in figure 1 where image tiles of a single epoch of the CDFS ACS images are overlaid on a WFI image. Conversely, browsing the cursor over the image highlights nodes in the tree when data are available at the cursor coordinate. This interaction between the tree and the image display provides an efficient means to simultaneously select all the data available at a given point in the sky.

Image servers that provide image cut-outs are also supported by the full IDHA implementation of the data tree. Figure 2 shows the fields of view of the original HST ACS image tiles, plus the outline of the image section that would be generated if the data request was submitted to the image server.

4. Access to Image Data via Treeview

In addition to data stored in the Aladin image server, images available via the Simple Image Access (SIA) protocol, and indeed any images available via a direct URL including images on a local disk, may be accessed and viewed in the tree with their field of view outlines shown in the display. Interfaces to the SIA servers for SkyView, and the NOAO Science Archive have been implemented in the prototype. For example, the result of a query to SkyView for the CDFS is shown in the tree in figure 1. Data trees for images available via direct URLs may be constructed and loaded into the prototype. Such trees have been demonstrated for WFPC2 association images served from ESO, and for VLA data served from Jodrell Bank. Data trees for local images may be automatically generated by specifying the top level directory at the load interface.

The data tree functions developed for the AVO prototype are also fully enabled in the Aladin 2.0 release (Bonnarel, Fernique & Boch 2004).

References

Bonnarel, F., Fernique, P., Boch. T. 2004, this volume, 221

Giavalisco, M. & the GOODS team, 2003, ApJ in press