

## Locating Virtual Observatory Resources With the Astrogrid Registry

E. Auden

*Mullard Space Science Laboratory, University College London,  
Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*

T. Linde, K. Noddle

*University of Leicester, Dept of Physics and Astronomy, University  
Road, Leicester, LE1 7RH, UK*

A. Richards

*Jodrell Bank Observatory, Dept of Physics and Astronomy, The  
University of Manchester, Macclesfield, Cheshire SK11 9DL, UK*

N. Walton

*Nic Walton, Institute of Astronomy, University of Cambridge,  
Madingley Road, Cambridge, CB3 0HA, UK*

**Abstract.** The Astrogrid registry stores metadata for every resource available to the Astrogrid virtual observatory. A resource can be defined as a data archive, a data storage area, data modelling and astronomical algorithms, communities, and other registries to name but a few. The Astrogrid registry uses the registry and resource schemas developed in conjunction with the IVOA registry workgroup. Each registry entry describes a resource in terms of identity, curation, content, and service access metadata. An Astrogrid user or component can conduct nested searches on the registry using XQuery, an implementation of the XML query language. In addition, the registry administration service allows privileged users to add, edit, and delete resource entries in a registry.

### 1. Overview

The Astrogrid registry stores metadata about astronomical data, storage resources, and data processing services for the Astrogrid project, a PPARC funded virtual observatory (VO). Astrogrid is a member of the International Virtual Observatory Alliance (IVOA), so the Astrogrid registry has been implemented with a primary goal of interoperating with other VOs. The registry component consists of a browser, an administration service, and a storage system holding registered grid resource metadata. All resources registered with Astrogrid are described in terms of metadata established by the IVOA.

Users query the registry through the browser web service. This web service can be accessed through either a web browser GUI or through a web service

call written inside another program. A user can query the registry through equality, wildcard, boolean and arithmetic conditions; multiple conditions may be nested in a single query. The registry browser receives an XML-formatted query, converts the query to XQL, and executes it against the Astrogrid registry. Any matching resource entries are returned to the user as a series of name / value key pairs. The key pairs can be displayed as HTML on the GUI or interpreted by other grid components.

Privileged users can access the registry administration service through a web browser GUI or service call. The administration service allows a user to add a new resource entry, edit an existing entry, or delete an existing entry. At the time of writing, all resource entries in the current registry XML file conform to the document, "Resource and Service Metadata for the Virtual Observatory: Version 7" (RSM v0.7) (Hanish, 2003). However, the Astrogrid registry is currently converting its resource descriptions to the IVOA's VOResource 0.9 schema (Plante et al., 2003).

The resources described in the Astrogrid registry are currently limited to astronomical datasets such as WFC-SUR, 1XMM, SURF, MERLIN, and USNO-B. However, the registry can also describe a variety of resources that model data (such as galactic simulations), process data (such as astrometric alignment algorithms), and store data. Other VO registries will complement the resource registry; the Community and MySpace registries will register information about datasets, storage, and user permissions.

## 2. Registry Browser

The registry browser communicates with users and grid components through a web service. This web service receives an XML query from a web browser GUI or another grid component's service call. The query, formatted to the Astrogrid registry query schema, is passed to a query processing class and executed against the registry XML file. The query response of matching resource entries is returned to the user through the web service. The registry browser web service is implemented as an Axis SOAP service and deployed through the Tomcat engine. This service is an integrated component of the Astrogrid Iteration 3 deployed at both the Mullard Space Science Laboratory and the University of Leicester.

The query processing class converts an XML query to an XQL query. The Astrogrid registry uses the GMD-IPSI XQL implementation of XQuery to perform registry queries. A registry query can be a combination of boolean operators like AND, OR, and NOT and the arithmetic operators EQ, NE, GT, LT, GE, and LE, and these conditions may be nested inside a single query. Wild card characters can be used in queries to search for all metadata element names or all element values. The XML query response is returned by the registry formatted as one or more resource entries. The query parser reformats the response as a series of name / value key pairs which are returned to the user.

The Astrogrid registry is stored as an XML file. Each resource is described as a "service" in accordance with the RSM v 0.7 schema. A copy of the registry XML file may be stored in any location accessible by a web service; registry XML files do not need to be exact mirrors of each other. In fact, a registry file may be populated only with specialized datasets or data processing services

tailored to a specific discipline, institution, or target audience. Other copies of the registry may choose to continuously harvest updated metadata from other VO resources to obtain maximum coverage of available grid resources. After Astrogrid Iteration 3, the XML files will be migrated to an XML database.

### **3. Registry Administration**

The registry administration service allows privileged users to add, edit, and delete resource entries stored in the registry. Much like the registry browser web service, the administration web service is an Axis SOAP service deployed through the Tomcat engine; it is also part of the registry component integrated with the Astrogrid Iteration 3 implementation. The administration service accepts an XML query formatted to the Astrogrid registry admin query schema. Administrative actions can only be performed on one resource entry at a time. After the administrative action has been performed, the web service returns a success or failure message to the user.

The registry administration service utilizes the functionality of the registry browser to ensure that entries for resources to be added do not already exist, and to gather existing metadata for entries to be edited or deleted. A user first selects a specific registry to administer and then selects an add, edit, or delete function. If a new resource is to be added, the user fills in an empty resource entry template with mandatory identity metadata and all other applicable curation, content, and service accessibility information. If the resource is to be edited, the user can update any of the fields displayed in a template populated with metadata previously stored in the registry. For services to be deleted, the user can simply view the existing metadata without edit facilities before executing the delete action. Once the add, edit, or delete action has been performed, the administered registry will immediately reflect the changes.

### **4. Metadata and Resources**

The Astrogrid registry is stored as an XML file. This file is described by the Astrogrid registry schema, which in turn references a further four schemas. The identity and curation schemas describe uniquely identifying metadata about the resource, as well as information about the people and institutions responsible for the resource's maintenance. The content schema contains the bulk of a resource entry's metadata, in particular UCDs, resolution, data quality, observation facilities and instruments, data rights, and finally spatial, spectral, and temporal coverage. The service accessibility metadata defines the location and protocol for accessing a resource through a web service, GLU service, HTML form, or other interface. For further information about Astrogrid registry schemas, please visit <http://wiki.astrogrid.org/bin/view/Astrogrid/AgCd06Schema>.

In the RSM v0.7, the IVOA defines a resource as any virtual observatory element that can be described in terms of a name, a unique identifier, and curation / maintenance metadata. The resources described in the Astrogrid registry are primary astronomical datasets stored at universities in the United Kingdom. Registered datasets include the USNO-B Catalogue (USNO-B), the Wide Field Survey on the Isaac Newton Telescope (WFC-SUR), the XMM-Newton

Serendipitous Source Catalogue (1XMM), the Multi-Element Radio Linked Interferometer Network archive (MERLIN), and the Solar UK Research Facility archive (SURF). In addition, three other Astrogrid resources are registered: the MSSL MySpace area, the MSSL community, and the Astrogrid Data Warehouse. More resources, services, and external registries will be added as the grid of virtual observatories expands.

## 5. Future Developments

The Astrogrid project anticipates an interoperability demonstration with other VOs participating in the IVOA in January of 2004. Registry interoperability will be a key issue of this demonstration. One of the major goals of the IVOA registry demonstration in January will be the implementation of metadata harvesters using the httpGET interface with OAI protocols. The web service interface for each registry will be defined in WSDL, and both SQL and XQL queries will be enabled.

The Astrogrid registry is also undergoing internal developments. The Astrogrid registry schemas will be upgraded from RSM v0.7 to the series of schemas released by the IVOA in November of 2003, including the VOResource 0.9 schema and its extensions. In addition, the registry backend will be migrated from its current format as an XML file to storage in an XML database. A third development will associate community permissions with registry administration privileges for specific users.

## 6. Conclusions

The Astrogrid registry stores metadata about datasets, data processing services and data storage services so that both scientific users and other grid software components can locate these resources on the grid. The registry consists of java-based browser, administration service, and XML file backend, and all three components have been integrated with the Astrogrid Iteration 3 release. The registry is designed to be a standalone element, to work in cooperation with other grid components, and to interoperate with other VO registries.

Through the registry browser, users can search available grid resources through queries composed of boolean and arithmetic conditions, exact matches, and wildcards. Queryable resource entries can describe data archives, models, software routines, educational materials, and a host of other resources. The registry administration service allows privileged users to maintain up to date metadata that describes available resources; resource entries can be added, edited, and deleted from a registry. The registry components will enable VO users to locate globally distributed resource of public, educational, and research interest.

## References

- Hanisch, R. 2003, "Resource and Service Metadata for the Virtual Observatory: Version 7", IVOA Working Draft.
- Plante, R. et al. 2003, "VOResource v0.9", IVOA schema release.