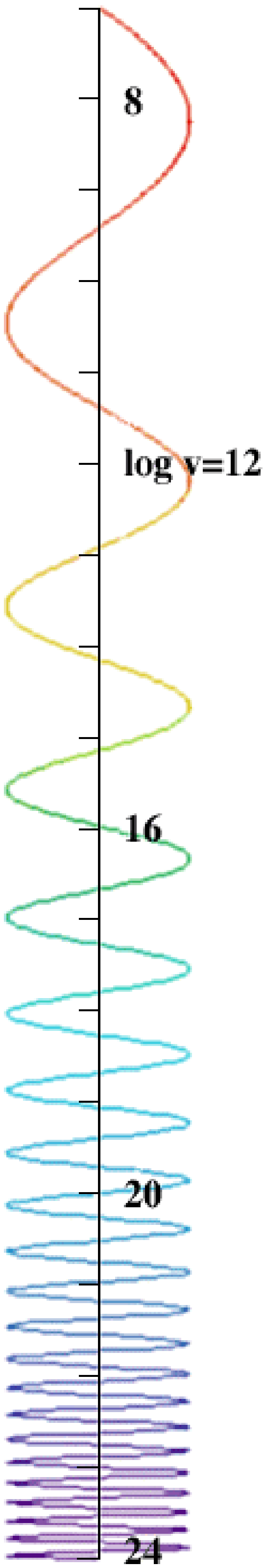

Quarterly Report
April–June 2002

Building the Framework for the National Virtual Observatory

NSF Cooperative Agreement
AST0122449

NATIONAL VIRTUAL OBSERVATORY

Building the Framework for the National Virtual Observatory
NSF Cooperative Agreement AST0122449
Quarterly Report

Period covered by this report: 1 April-30 June 2002
Submitted by: Dr. Robert Hanisch (STScI), Project Manager

Executive Summary

Highlights: In this quarter the project team made substantial progress on scientific, technical, and programmatic matters.

- *Scientific.* At the project team meeting (Tucson, 16-17 April) three scientific demonstration projects were selected from an extensive list of potential projects compiled by the Science Working Group. The demonstrations were chosen based on a number of criteria, including availability of necessary data, feasibility of completion by January 2003, and ability to show results in a matter of a few minutes (i.e., the time one can typically hold the attention of an astronomer passing by a display booth at an AAS meeting). The selected demonstrations are
    - Brown dwarf candidate search.
    - Gamma-ray burst follow-up service.
    - Galaxy morphology measurement and analysis.

  These are described in more detail in WBS 10.2 of this report.

- *Technical.* In collaboration with the European virtual observatory development projects, AstroGrid and AVO, we released V1.0 of the VOTable XML formatting standard for astronomical tables. Using VOTable as a standard output product, some 50 "cone search" services were implemented by 7 different groups within the team. The cone search services respond to a request for information based on a right ascension, declination, and radius about that position. Four software libraries for parsing VOTable documents were written and made available via the team web site. Also, a JHU-based team developed a catalog cross-correlation service for SDSS, 2MASS, and FIRST using Microsoft's .NET facilities and won second place in a nationwide software development contest.

- *Programmatic.* Closure was reached on all collaborators' subawards.

In addition, project team members gave 12 invited or contributed talks and 5 poster papers at the conference "Towards an International Virtual Observatory" (June 10-14, Garching, Germany). The project leadership worked closely with our European counterparts to establish the International Virtual Observatory Alliance, a coordinating body that will help assure interoperability among the various national VO initiatives. Through the IVOA we also expect to minimize redundant evaluation and development efforts.

The focal point of work going on now and for the coming quarter will be implementation of the science demonstrations and their underlying software components.

Issues and Concerns: It took as long as six months to finalize some of the subawards. As a result, several groups have not been able to commit as much effort to the project as originally planned, and spending for the first year funds is lagging behind expectations. In our second year renewal package we will present plans for application of these funds to ongoing and future work.

Activities by WBS

1 Management

1.1 Science Oversight (Executive Committee)

Highlights: The Science Working Group assembled an extensive list of possible science demonstration projects, and at the April team meeting three of these were selected for implementation:

- A brown dwarf candidate search
- A galaxy morphology analysis
- A gamma-ray burst follow-up service

Leads were selected and teams formed for implementation of each demonstration, and schedules were defined for each effort.

Status: Work is proceeding essentially as planned. At our July 30-31 team meeting we will review progress in detail and verify that all technology elements are in place, or will be in place in sufficient time to support the demonstrations next January.

1.2 Technical Oversight (Executive Committee)

Highlights: Substantial progress was made in metadata standards with 1) the release of the XML-based VOTable format specification, 2) a draft specification for the VO data model, 3) a draft specification for spatial-temporal metadata, 4) a draft specification for resource and service metadata, and 5) a general metadata definition framework. This work enabled a first, simple interoperability service to be prototyped: the Cone Search. Several groups both within and outside the collaboration developed software to write and parse VOTable documents.

Issues and Concerns: The rich toolkit of Web Services (UDDI, WSDL, RDF, SOAP, etc.) needs to be explored and understood more fully by the team.

Status: The science demonstration projects are being used to determine technical implementation priorities.

1.3 Project and Budget Oversight (Executive Committee)

Highlights: All subawards have been successfully placed. NSF was briefed on the overall project status on May 21, and was provided with updates to the project management plan and general project schedule. We also provided detailed schedules for the three initial science demonstrations.

Issues and Concerns: It took as much as 6 months to get several subawards in place, owing to lengthy negotiations over certain terms and conditions. Some organizations were not able to commit to hiring and others were somewhat slow to engage efforts in the project. Thus, most organizations in the collaboration are currently under-running their first-year spending plans.

Status: As we process our second-year renewal we will rebudget, using unspent funds from year one to carry forward and smooth out the overall funding profile.

2 Data Models

2.1 Data Models / Data Model Architecture (McDowell, SAO)

Highlights: Draft architecture document circulated.

Issues and Concerns: Need to elucidate the role of Uniform Content Descriptors (UCDs) or their successors in data models.

Status: A draft document on the data model architecture has been circulated among the team and to members of the international collaborations. Fruitful discussions at the April NVO team meeting in Tucson and at the VO conference in Garching have led to agreement on a basic approach, in which we will make small models of aspects of the data and agree on a mechanism for associating such models with datasets and representing them in formats such as VOTable.

2.2 Data Models / Data Types (McDowell, SAO)

Highlights: The image data model is discussed in the draft document.

Status: We are supporting the work of the metadata group in their definition of space-time metadata. From the data model point of view, it is important to ensure that the mechanisms used to associate the space-time metadata with a dataset are defined generically so that they can also be used with other kinds of metadata.

Identified several kinds of metadata as candidates for early modeling, including data quality and data subspace.

U. Penn expanded their strawman design of a new standard for the incorporation of time-series data into a federated database system. It appears that an extension of FITS (with NVO compliant metadata (see WBS 3)) will be the most readily accepted proposal. Alternative, non-FITS models may also be successful, but have been explored less extensively. Work has begun on a proposal to the FITS community.

2.3 Data Models / Data Associations (McDowell, SAO)

Status: Work scheduled to begin July 1, 2002.

3 Metadata Standards

3.1 Metadata Standards / Basic Profile Elements (Rots, SAO)

Highlights: Worked on DTD for space-time coordinates and regions.

Issues and Concerns:

Status: Published the DTD for Space-time coordinate metadata and a draft DTD for region definition. Hence, the former is ready to be tried out in resource description, query, and data return context in the cone search experiment. Working on schema representations of the two and examples.

A proposal was developed for reclassification of position-related UCDs and forwarded to the CDS (keepers of the UCD list).

R. Plante, J. McDowell, and A. Rots have a basic agreement for the X-ray side of the work on the galaxy classification demo.

U. Penn. concentrated on the issues of data provenance, especially in federated time series databases which may be assembled in semi-automated data pipelines. This is especially a concern with moving objects, which may be detectable only with analyses of multiple image datasets.

3.2 Specific Profile Implementations (McGlynn, USRA/HEASARC)

Highlights: A write-up describing the interactions and areas of usefulness for various Web technologies (WSDL, DAML, UDDI, SOAP, GLU, RDF) was circulated through the Metadata Working Group for discussion.

In the related NASA ITWG effort preliminary WSDL profiles were written for HEASARC services and similar efforts were collected for other archive sites.

Issues and Concerns: There is uncertainty about the scope of the metadata effort and boundaries with the API and data model issues. The responsibilities of these parties are not well defined. However since all parties are working together closely the consequences to the project may not be large.

The project team needs to acquire a better understanding of Web Services and their role in the NVO.

Status: The catalog profile described in the previous quarter has been updated in response to comments noting the similarity of certain elements to Dublin Core Metadata standards. Elements of the catalog metadata description may be broken off into a more generic profile describing a wide variety of VO services.

3.3 Metadata Representations and Encoding (Plante, UIUC/NCSA)

Highlights: Version 1.0 of the VOTable specification was released April 15 as part of an international collaboration led by R. Williams (SDSC) and F. Ochsenbein (CDS/AVO). Since then, A. Rots has defined (in XML) a schema for representing a Space-Time coordinates and coordinate systems. Work is now underway on a proposal to integrate this schema into VOTable. It is expected that this work will lead to a general way to describe reference frames/ancillary descriptions of for other types of data included in a VOTable.

Plante presented a white paper on metadata frameworks at the international VO meeting in Garching that attempts to unify the use of metadata across the different contexts of a VO application (http://monet.astro.uiuc.edu/~rplante/VO/metadata/fw-draft.pdf).

Issues and Concerns: The relationship between metadata and data models is a current topic of discussion.

Status: The main effort in this area is driven by the needs of our first-year science demonstrations. In particular, D. Tody (NOAO) is developing a specification for standard image access; this will include a specific specification for describing images within a VOTable. This is needed in some form for at least two of our three science demonstrations. Effort this summer will be focused on simple single-image access.

3.4 Profile Applications (Plante, UIUC/NCSA)

Highlights: There are now over 50 implementations of the Cone Search standard service. This has been used to experiment with the registration and searching for resources. R. Hanisch led the development of a draft specification for general resource metadata; this white paper document was presented at the Interoperability Working Group breakout session at the international VO meeting in Garching.

A design for a special service that computes X-ray fluxes from observations by the Chandra X-ray telescope was developed to support the Galaxy Morphology science demonstration. The service will be implemented this summer.

Issues and Concerns: Intelligent use of VO services requires access to descriptive metadata that go beyond the service interface; how to integrate this information with an WSDL description is under investigation.

Status: Much of the work toward the first year demonstrations are metadata related, and prototyping of the integrating services is underway. In addition, several investigations are underway on the use of Web Services for describing services like the Cone Search.

3.5 Metadata Standards / Relationships (Rots, SAO)

Status: No work scheduled at this time.

3.6 Metadata APIs (Plante, UIUC/NCSA)

Highlights: VOTable parsers are now available for three languages are now available, including two developed within the NVO project: a Perl module by Eric Winter (HEASARC) and Java library (JAVOT) by R. Williams (CACR). Williams consulted for the C++ version developed by the VO-India project.

Status: See WBS 3.4.

4 Systems Architecture

4.1 System Design (Moore, SDSC)

Highlights: The system design will be strongly influenced by the requirements that were expressed at the conference "Towards an International Virtual Observatory." From this meeting, the following central components were identified: portals for accessing images, catalogs, and procedures; interactive web services; batch-oriented survey processing pipelines; and grid services. While these components are oriented towards data and information management, a similar infrastructure is required for knowledge management that expresses the sets of operations that can be performed on a given data model, and defines the relationships between the UCDs that express exact semantics for physical quantities.

Issues and Concerns: An MOU has been established with IRSA governing use of the 2MASS collection. The MOU states that released material can be freely replicated, that UCSD will ensure that only public data will be released, and that UCSD will inform IRSA about the sites that receive copies. The remaining concern is to build an appropriate set of access controls for all data that is still proprietary.

Status:
4.1.1 – The system design for most of the NVO architecture will be driven by existing systems for managing records, images, and processing pipelines. These systems typically are oriented towards either processing a small amount of data (1000 records or 90 seconds access) or processing the images that are taken on that day. Larger scale accesses will need to be supported by the NVO testbed. The design of the testbed requires an engineering estimate of the computation capacity, I/O bandwidth, and caching capacity. To ascertain a reasonable scale, we are continuing the implementation of a background analysis of the 2MASS collection, in collaboration with the Montage Project team at IPAC. This will require a complete sweep through 10 TB of data, at an expected rate of 3 GB/sec. Good has created the initial analysis routine, which is being tested at SDSC.

We have 2 TB of the 2MASS collection replicated onto a disk cache at SDSC. We are continuing the replication of the 2MASS collection onto the HPSS archive at Caltech. This is important to improve reliability by a factor of 10. When the HPSS archive at

SDSC is off line, we will be able to retrieve images from the Caltech copy. To support the automated replica fail over, we are installing version 1.1.8 of the SRB (Storage Resource Broker) at Caltech.

We have done a test run of a re-analysis for the DPOSS collection, in collaboration with R. Williams of Caltech. This was done on a 64-processor Sun platform, accessing data from a disk cache. The computation was CPU-limited, taking 410 seconds to process a single 1-GB image on one processor. Using the entire platform, the re-analysis of the complete DPOSS collection could be done in 11 hours, at a sustained I/O rate of 135 MB/sec. This includes writing a new version of the entire collection back to disk, or moving 5.6 TB of data. The goal is to gain a factor of 10 in performance by moving to the Teraflops compute platform, and the large 30 TB disk cache.

We are also working on engineering estimates for the manipulation of large catalogs. J. Gray has shipped us a copy of the SDSS metadata (80 GB). We are attempting to dedicate disk space and compute resources to the analysis of this catalog.

4.1.2 and 4.1.3 – E-mail exchanges are being conducted on OGSA interfaces to the grid environment, metadata management, data model specification, and knowledge management with E. Deelman, D. Tody, R. Williams, and R. Plante.

4.1.4 – We are upgrading the SRB server at Caltech to version 1.1.8, to support automatic fail over to an alternate replica. This will improve reliability of the system for image access by a factor of 10. This requires testing the new version with the existing IPAC 2MASS portal.

4.2 Interface Definition (Williams, CACR)

Status: See WBS 3.4.

4.3 Network Requirements (Williams, CACR)

Status: Work not scheduled until late in CY2002.

4.4 Computational Requirements (Williams, CACR)

Status: Work not scheduled until late in CY2002.

4.5 Security Requirements (Kesselman, USC)

Status: See WBS 6.2.2.

5 Data Access/Resource Layer

5.1 Resource and Information Discovery (Szalay, JHU)

Status: Work has begun on developing efficient mechanisms for comparing sky coverage between large-scale surveys and archives of pointed observations. Current development utilizes the Hierarchical Triangular Mesh (HTM) spatial indexing scheme. Consideration is also being given to performance issues for positional cross-correlation methods with large catalogs, and it looks feasible to speed up nearest neighbor identification by as much as a factor of 100.

5.2 Data Access Mechanisms (Kesselman, USC)

5.2.1 Data Replication (Deelman, USC)

Highlights: In a wide area computing system, it may be desirable to create remote read-only copies (replicas) of data elements (files)—for example, to reduce access latency, increase robustness, or increase the probability that a file can be found associated with idle computing capacity. A system that includes such replicas requires a mechanism for locating them. The Replica Location Service (RLS), in development at USC/ISI maps a given unique logical identifier for desired data content to the physical locations of one or more copies of this content. The RLS is extensible in that the users and applications can extend the information contained within it to other application specific attributes.

Issues and Concerns: The Replica Location Service has now entered the phase of alpha testing. During this period we are testing the functionality of the service as well as its performance. So far the results are encouraging in both areas however, further testing still need to be conducted.

Status: USC/ISI is developing a Replica Location Service, the next generation of the Globus Replica Catalog (RC). RC permitted a mapping from logical file names to the physical locations of the particular file. Although the functionality of RC in terms of the mapping was adequate the performance and the reliability of the system (a centralized server) was low. The new generation, the Replica Location Service, allows for the system to be distributed, and replicated. The testing on the alpha prototype of the service is underway. As we progress in the development cycle, we will look forward to setting up a testing environment within the NVO framework.

5.3 Data Access Protocols (Williams, CACR)

Highlights: An XML-based standard for encoding of astronomical tables, VOTable V1.0, was released on April 15 following several months of active development with the European VO projects.

Status: The VOTable XML format has advanced from the official April 15 release, with several parsing packages now available, including the JAVOT Java package. This is built with an automatic creation of an API from the XML, using a product called Breeze Factor. The software parses the XML to a memory-efficient, column-wise structure that can be easily manipulated for the purposes of transformation, filtering, and presentation. JAVOT is freely downloadable (http://www.us-vo.org/VOTable/JAVOT)

The JAVOT package has been used for several purposes in connection with the Cone Search registry. There is a primitive cross-match engine, and a verifier service to determine is a purported VOTable is properly formatted. From the published cone searches and JAVOT, we have build a service that makes a density chart of any con service, by repeated sampling and counting the number of records returned.

5.4 Data Access Portals (Tody, NOAO)

Highlights: We identified the data access/analysis portal as a fourth primary service required to support researcher data analysis with NVO. Further investigation of data grid, Teragrid, and Web Services technology was conducted to determine the best role for these technologies in distributed data access and analysis. We conducted multiple exchanges with participants regarding the requirements and approach for distributed data access and analysis, principally Reagan Moore (SDSC), Jonathan McDowell (CXC), and Tim Cornwell (NRAO).

We implemented several cone search prototypes and reviewed VOTable in the context of data access. In collaboration with SDSC, we installed SRB on a NOAO computer and established a demonstration data link to SDSC. A survey was conducted of current Web Services technology and toolkits and a review of popular Web Services toolkits was written and posted to the metadata discussion group.

We initiated a discussion of a set of four web services for querying image databases and retrieving image cutouts or mosaics. A white paper specifying prototype versions of these services was in production at the end of the quarter.

Issues and Concerns: There is considerable overlap of functionality between the four major NVO services thus far identified: Web portals, a request management framework, the data access/analysis portal, and the Teragrid large-scale testbed. Further study and experimentation is needed to determine how these services should be integrated. The first-year science prototypes will provide an opportunity to conduct experiments in all four areas, and should provide the experience needed to develop an overall system architecture linking all service classes.

Status: Our main emphasis over the next quarter will be to specify the image query and retrieval services, and to participate in implementation of some initial services for image access and retrieval. These services are expected to support image service registration, querying of image databases, and single-file and bulk retrieval of image data, including both simple synchronous access and general asynchronous or bulk image access using messaging. Image data will range from archival image files to atlas images to image cutouts or mosaics. These prototype image access services are needed to support the first-year science prototypes, and will provide an opportunity to experiment with the use of Web Services technology for data access.

6 NVO Services

6.1 Computational Services (Berriman, IRSA)

Status: IPAC has secured in-kind funding through the NASA ESTO-CT program to develop *Montage*, an image mosaic service for the NVO. Achievements to June 30:

- Development of a Software Engineering Plan, Software Requirements and High-Level design.
- Development of a working, portable prototype that has been used on the IBM Blue Horizon (8 processors) to generate demonstration mosaics. The code consists of stand-alone modules that perform functions such as image reprojection; these modules are of value in their own right and will be made available to other NVO image services.
- Development of an initial design for deploying Montage on a computing grid within the NVO framework; distributed to the NVO architecture team for review.

IPAC is also beginning work on a request management service called *ROME* Request Object Management Environment). They prepared and distributed requirements and design documents and developed a prototype request management system for handling file transfers and downloads; this is now in the IRSA testbed as part of the OASIS v2.0 release.

6.2 Computational Resource Management (Moore/SDSC)

6.2.1 Computational Request and Planning

Status: We are tracking technology that is being developed through the NSF GriPhyN project, the DOE Particle Physics Data Grid, the DOE SciDAC projects, and the NASA Information Power Grid, for the management of computational resources. The two central components are management of the computational resources, and management of the processes that are being run on the computational resources. The former is handled by the Globus toolkit, version 2. The latter is still a research activity. There are multiple versions of work flow management under development, including the Condor DAGman and associated data scheduling mechanisms, the survey pipeline processing systems used in astronomy, and an advanced knowledge-based processing system under development at SDSC for a DOE SciDAC project. We would expect to start with the current survey pipeline systems, switch processing to a grid managed environment under Condor when computer resources are exceeded, and then switch to the knowledge-based processing systems for complex queries. The advantage of the knowledge-based systems is their ability to dynamically adjust the workflow based upon results of complex queries to information collections. The conditional relationships between processing steps can be quite complex, as opposed to the simple semantic mapping of output files to input files for the DAGman system.

6.2.2 Authentication

Highlights: USC/ISI has evaluated Spitfire, a database access service, which allows access to a variety of databases. Spitfire, developed as part of the European Data Grid project, consist of a server as well as client tools. The Spitfire server connects through JDBC to a database using predefined roles. The client can connect directly to the server through HTTP, and perform database operations. Even though Spitfire sounds like an interesting technology, it has many drawbacks in terms of security and support for transactions that span multiple database tables.

Issues and Concerns: Although Spitfire is based on the Globus Grid Security Infrastructure for Authentication, it exemplifies security problems in terms of authorization. In tests performed at USC/ISI, we are able to modify the database using a new version of the Spitfire server with an unauthorized client (a client from an earlier version of the code, which did not implement any security). Spitfire also does not currently support transactions that span multiple DB tables. The documentation was also inadequate, as it showed only examples of query operations and not example templates for create, update or delete operations.

Status: USC/ISI has communicated the authentication concerns to the Spitfire developers and is currently studying the possibility of adding transactional support to Spitfire. We are also following the developments within the UK e-Science program for any development in the area of grid-enabled interfaced to databases.

6.2.4 Virtual Data

Highlights: USC/ISI is working with University of Chicago on a Virtual Data System, Chimera, which allows users to specify virtual data in terms of transformations and input data. The system is composed of a language and a database for storing the information needed to derive virtual data products. USC/ISI focused on designing and implementing a planner which enables the translation between an abstract representation of the workflow necessary to produce the virtual data and the concrete steps needed to schedule the computation and data movement.

Issues and Concerns: USC/ISI is currently working on the first version of the planner, which is part of the Chimera. Currently the planner is rudimental and further research is needed to increase the level of sophistication of the planning algorithm as well as increase the level of the planner's fault tolerance.

Status: The Virtual Data System language (VDL), developed at University of Chicago is specified in both a textual and XML format. The textual version is intended for use in the manual creation of VDL definitions, for use in tutorial, discussion, and publication contexts. The XML version is intended for use in all machine-to-machine communication contexts, such as when VDL definitions will be automatically generated by application components for inclusion into a VDL definition database.

The VDS-1 system, also known as Chimera, is implemented in Java, and currently uses a very simple XML text file format for the persistent storage of VDL definitions. Its virtual data language provides a simple and consistent mechanism for the specification of formal and actual parameters, and a convenient paradigm for the specification of input parameter files. VDS-1 will be packaged and released during Summer 2002.

This planner takes an abstract Directed Acyclic Graph (DAG) specified by Chimera and builds a concrete DAG that can then be executed by Condor-G. In the abstract DAG neither the locations of where the computation is to take place nor the location of the data are specified. The planner consults the replica catalog to determine which data specified in the abstract DAG already exists and reduces the DAG to only the minimum number of required computations and data movements. Finally, the planner transforms the abstract DAG into a concrete DAG where the execution locations and the sources of the input data are specified. This DAG is then sent to Condor-G for execution.

7 Service/Data Provider Implementation and Integration

7.1 Service/Data Provider Implementation (Hanisch/STScI)

Status: No work scheduled this quarter. Service and data providers are participating in the metadata standards discussions and are investigating prototypes, e.g., via the Cone Search servers described in WBS 3.4.

7.2 Service/Data Provider Integration (Hanisch/STScI)

Status: No work scheduled this quarter. Service and data providers are participating in the metadata standards discussions and are investigating prototypes.

8 Portals and Workbenches

8.1 Data Location Services (McGlynn, USRA/HEASARC)

Status: As part of the activities associated with the GRB demonstration project, simple scripts for evaluating where the 'best' survey datasets were available were developed. These simple scripts search a hierarchy of surveys for each wavelength until a survey with coverage in the requested region is found.

8.2 Cross-Correlation Services (Djorgovski, Caltech)

Status: No scheduled activities prior to CY2002 Q3.

8.3 Visualization Services (Williams, CACR)

Status: No scheduled activities prior to CY2003 Q3.

8.4 Theoretical Models (De Young, NOAO)

Status: No scheduled activities prior to CY2002 Q4.

9 Test-Bed

Highlights: USC/ISI is actively working with the iVDGL project to understand the infrastructure needed to support testbed build up as well as Virtual Organization management. We are evaluating the software available from the European Data Grid that allows for the setup of a Virtual Organization Server. This server holds information about the people belonging to a given virtual organization.

Issues and Concerns: During our evaluation we have found the tools provided with the virtual organization server to be difficult to use. We are also evaluating the tools built for the Virtual Organization server at Caltech to see if they are more robust.

Status: We are evaluating tools for testbed set up and management.

10 Science Prototypes

10.1 Definition of Essential Astronomical Services (Szalay, JHU)

Status: Discussion has been active about Web Services, with general agreement among the NVO, AVO, and AstroGrid projects that VO functions should be built with a hierarchy of Web Services. Within the NVO project several prototypes have been implemented (such as the ConeSearch). The focus of current work is on high-level content descriptors and the relationship between Web Services and facilities such as GLU (CDS) and metadata services (UCDs).

10.2 Definition of Representative Query Cases (De Young, NOAO)

Highlights: Three scientific demonstration projects were selected from an extensive list of potential projects compiled by the Science Working Group. The demonstrations were chosen based on a number of criteria, including availability of necessary data, feasibility of completion by January 2003, and ability to show results in a matter of a few minutes (i.e., the time one can typically hold the attention of an astronomer passing by a display booth at an AAS meeting). The selected demonstrations are

- Brown dwarf candidate search.
- Gamma-ray burst follow-up service.
- Galaxy morphology measurement and analysis.

Status: The scientific goals of each demonstration are described below.

- *Brown dwarf candidate search:* We will deploy web-based cross-matching applications and demonstrate their science value by identifying T-type brown dwarfs through cross-matching the 2MASS (infrared) and several optical catalogs that have been published in an NVO-compliant manner. A major target will be the Sloan (optical) source catalogs to identify objects with extreme colors. We will also investigate the use of the newly-released USNOB catalog in determining the completeness of such searches for brown dwarfs having less extreme colors: while the USNOB catalog is not as deep as the Sloan survey and will likely not contain any of the extreme T-dwarfs seen there, it does contain proper motion information and has much greater sky coverage. We will also consider a few other selection criteria, like high redshift QSOs, or high-redshift galaxies (dropouts) as part of our demo suite.

- *Gamma-ray burst follow-up service:* We will develop a quick response service that provides a summary of all known astronomical information regarding a specified region in the sky (typical a few arcminutes to a degree in size). This should include at least:
    - optical, IR, radio and X-ray images
    - known stars and galaxies
    - special objects
    - observations by major telescopes including links to appropriate archival data
    - bibliography links

  This service shall be integrable with notification services for rapid response events (e.g., BACODYNE GRB notifications).

- *Galaxy morphology measurement and analysis:* This demonstration looks for relationships in galaxy clusters between morphology and the intergalactic environment, specifically the X-ray emitting gas. Given a user-specified list of clusters, this application will create a sample of candidate member galaxies. For each galaxy that has been observed in the X-ray, we will calculate the X-ray flux emitted from the surrounding environment. In addition, we calculate three morphological parameters (mean surface brightness, concentration index, asymmetry index) from image cutouts of the galaxies from a chosen image survey. These calculated values will be combined with other relevant data mined from existing catalogs to create correlation plots.

10.3 Design, Definition, and Demonstration of Science Capabilities (De Young, NOAO)

Highlights: Implementation plans and schedules were agreed upon for all three science demonstrations. All lead to completion in November, allowing contingency time for adjustments and improvements prior to the January 2003 AAS meeting.

Issues and Concerns: As noted below, once detailed design work began we were able to identify some holes in technical services and the need to refine some of the science requirements. Thus far these problems appear to be solvable within the current schedule and resource allocations.

Status: A simple mockup of data assembly for the gamma-ray burst follow-up service was created using SkyView and VizieR services only. This enabled the creation of complex VOTable files which describe multiple images and illustrated that there remain unresolved issues in the structure of complex VOTables regarding how various elements (e.g., separate resources) interact.

Discussions were held with F. Ochsenbein and P. Fernique (for Aladin) and J. Good (for OASIS) to discuss the user interface requirements for the GRB demo. Both indicated that they felt that the interfaces would meet the minimal requirements of scriptability needed.

It is unclear if the 2MASS dataset will be usable in the context of the GRB demo, since the anticipated field of view (~1 degree) may encompass hundreds of individual images. The Montage server, which could address this issue, will not be running at the time of the demos. However, a composite image may be generated from the IRSA catalogs.

A description of the information sets that are to be included in the GRB demo was circulated among the collaboration. This list is currently being translated into explicit URLs. Discussions have been had with D. Tody regarding the desirability that any image services used by the GRB demo be included in the image cutout service. Similarly all catalog services should be accessible through the standardized cone search protocol.

IPAC has assumed the role of technical lead for the brown dwarf demonstration. As part of this effort we are leveraging in-kind funding provided by the NPACI Digital Sky project to deploy a distributed cross-matching service. Progress to date includes:

- Preparing a technical description and work plan for the Brown Dwarf demonstration project.
- Developing a design for a general, distributed cross-comparison service for calculating cross-matches between two distributed catalogs. Began development work on this service.

Plante and Annis are completing the design of the Galaxy Morphology demo. They have identified the target input datasets, user inputs, and the form of the output results. Minimum requirements for grid-based calculation of morphology parameters have been worked out and a development plan outlined. Plante and Tody have specified the requirements on the image cutout service. Plante, Rots, and McDowell have outlined the design of the x-ray flux service using the Chandra archive.

11 Outreach and Education

11.1 Strategic Partnerships (Voit, STScI)

Highlights: We have planned and will hold an Education and Public Outreach workshop in Baltimore on July 11-12. The focus of the meeting will be on identifying key capabilities that developers of E/PO products and services would like to see integrated into the NVO infrastructure. Representatives will be attending from the following institutions:

- Space Telescope Science Institute
- Space Sciences Laboratory / Berkeley
- Hands-on Universe
- Starry Night / Space.com
- Spitz Planetaria
- Hayden Planetarium / American Museum of Natural History
- Chandra Science Center
- National Optical Astronomy Observatories
- National Center for Supercomputing Applications
- American Association of Variable Star Astronomers
- Loch Ness Productions (planetarium shows)
- Sloan Digital Sky Survey
- Contemporary Exercises in Laboratory Astronomy / Gettysburg U.

We also presented our outreach plans at the international meeting on virtual observatories in Garching.

Status: Our list of outreach partners continues to grow, and through the outreach workshop, they will assist us in drafting an outreach requirements document during the month of August.

11.2 Education Initiatives (Voit, STScI)

Status: No activities planned this quarter.

11.3 Outreach and Press Activities (Voit, STScI)

Status: No activities planned this quarter.

Activities by Organization

Astronomical Data Center (Raytheon)
Raytheon staff at the Astronomical Data Center supported work in four WBS areas.

WBS 2 Data Models. Raytheon staff at the ADC wrote a separate white paper on a Data Model for space systems data and telemetry. Staff also provided analysis and commentary on various NVO data model documents.

WBS 3 Metadata Standards. Raytheon/ADC staff participated most significantly in the Metadata Standards (WBS 3) activities, which included: (a) participation in weekly telecons of the NVO Metadata Working Group; (b) development of science use cases and the science demos; (c) analysis of various NVO metadata framework documents and resource description documents; (d) research into Resource Description standards (DAML+OIL) to see what applications are useful for NVO; (e) worked on improvements of VOTable and its API; (f) made improvements to the proposed NVO Space-Time schema and validation of its DTD; and (g) investigated applications of Web Services (UDDI, WSDL, SOAP) to NVO prototypes.

WBS 7 Service/Data Provider Implementation and Integration. Raytheon/ADC staff began work to define the ADC metadata and to set up a prototype WSDL service demo.

WBS 10 Science Prototypes. Raytheon/ADC staff members K. Borne and E. Shaya attended the second Project Team meeting in Tucson in April 2002. At that meeting and subsequently, Raytheon/ADC staff contributed to the development of two of the science demos: the Gamma-Ray Burst Follow-up and the Galaxy Morphology Demo.

Caltech–Astronomy Department
We have experimented with Topic Maps as a way to represent knowledge (Mahabal et al.,http://www.astro.caltech.edu/~aam/science/topicmaps/). The knowledgebase considers galaxies, papers, and authors, with relationships “paper mentions object” and “is an author of this paper,” etc.

Caltech–Center for Advanced Computational Research
The VOTable XML format has advanced from the official April 15 release, with several parsing packages now available, including the JAVOT Java package developed at CACR. This is built with an automatic creation of an API from the XML, using a product called Breeze Factor. The software parses the XML to a memory-efficient, column-wise structure that can be easily manipulated for the purposes of transformation, filtering, and presentation. JAVOT is freely downloadable (http://www.us-vo.org/VOTable/JAVOT).

The JAVOT package has been used for several purposes in connection with the Cone Search registry. There is a primitive cross-match engine, and a verifier service to determine is a purported VOTable is properly formatted. From the published cone searches and JAVOT, we have built a service that makes a density chart of any cone service, by repeated sampling and counting the number of records returned.

Working with G. Kremenek of SDSC, we have been working with bulk (terabyte) image datasets from DPOSS and 2MASS. For 2MASS, we have begun the mosaicing of the entire dataset for ingestion into the Virtual Sky presentation system (http://virtualsky.org). This involves background subtraction and image resampling for over a trillion pixels. To extend the usefulness of Virtual Sky, we have also built a cutout service, so that users can drill further down from Virtual Sky to the original, astrometric pixels. This service will be available as long as there is 3 TB of disk storage at SDSC to hold the data.

A summary paper has appeared on applications of Grid: "Grids and the Virtual Observatory", R. D. Williams, for publication as part of a collection on "The State of the Grid 2002" (http://www.cacr.caltech.edu/~roy/papers/vogrid.pdf).

Caltech–Infrared Processing and Analysis Center
During this period IPAC

- Assumed technical lead for Brown Dwarf demonstration project; prepared technical description and work plan
- Deployed demonstration cone searches that are NVO compliant
- Developed requirements, design and working prototype code for Montage, the image mosaic service
- Began development of cross-comparison service
- Delivered requirements and design of the Request Object Management Environment (ROME).

Canadian Astronomy Data Centre/Canadian Virtual Observatory Project
In the area of science prototypes we have completed the initial implementation and internal release of the CADC VO prototype system. The system allows queries on a VO-level representation of pixels and source catalogues. We have generated content suitable for testing the system from HST WFPC2 archive images and extracted source catalogues with 18.6 million sources.

A primary concern is database performance. It is unclear whether queries can be done effectively on large source catalogues. Funding for a major increase in database capability is in hand and the acquisition process is underway. Overall, we have achieved all of the major goals including interface development and data content generation. There have been delays due to unforeseen systems/database administration hurdles.

Carnegie-Mellon University/University of Pittsburgh
We have used the NVO support to partially finance a professional programmer, Paul Hsuing. Hsuing joined us only recently but has already worked closely with A. Szalay and T. Budavari at JHU. He is focusing on Web Services to help export the advanced statistical tools developed in Pittsburgh to the NVO. In particular, Hsuing is working on providing a Miscrosoft .NET interface to our Mixture Model clustering algorithm to be part of the first NVO science tutorials. We plan to continue this work, and Hsuing will help the Pittsburgh group understand the .NET architecture and other Web Services.

Fermi National Accelerator Laboratory
Fermilab continued to participate in the weekly Metadata working group phone conferences (WBS 3). J. Annis attended the NVO project meeting in April. He worked on developing the galaxy morphology science demo (WBS 10.3.1) and continues to work on its implementation. Four to five datasets with the necessary catalog and atlas image cutouts have been identified for use in the demo. A person has been hired to work on NVO/ iVDGL.

High Energy Astrophysics Science Archive Research Center (USRA)
HEASARC personnel attended the "Towards an International Virtual Observatory" conference in Garching including an Opticon Interoperability Working Group meeting as well the VO team meeting in Tucson.

The HEASARC has taken a lead role in developing a demonstration project for the January AAS meeting to show how the VO could be used to effectively present users with information from disparate VO resources in a timely fashion. The particular scenario being addressed is to provide immediate information regarding the vicinity of a poorly localized gamma ray burst.

The HEASARC has participated in the weekly metadata telecons and bi-weekly management telecons. Several memos have been sent out for comment by the collaboration. The HEASARC continues to work on developing a clear understanding of the metadata representations of catalog information and other resources. Work has begun – largely as a byproduct of the GRB demo – to integrate HEASARC resources, notably SkyView, Astrobrowse and Browse, into the VO.

Related Activities

The ClassX project to develop automated classifiers for X-ray data continues. A VOTable pipeline has been developed and a generic VOTable library has been released. A Web page comparing the various existing VOTable libraries including their native languages and capabilities has been developed.

The NASA Astronomy Data Centers Executive Committee Interoperability Technical Working Group (ADEC ITWG) has continued to work at building interlinked NASA archive systems. Development of common linking profiles and WSDL descriptions continues.

Johns Hopkins University
The SQL parser and function extensions developed quickly for the SkyQuery project are being generalized and made robust. JHU team members assisted A. Moore (CMU) in making his expectation maximization (EM) clustering algorithm available as a Web Service. The template developed for this was also used to improve the SDSS image cut-out service, and work is in progress on adding support for FITS files (in addition to JPEG format) to the cut-out server. A visualization Web Service that displays query results as a density map was developed. Also, a Web Service that computes various cosmological functions was deployed.

Microsoft Research
In this period J. Gray worked closely with the SDSS consortium on making their online data more accessible. This includes improvement to the SkyServer (http://skyserver.sdss.org) and a cutout service (http://skyservice.pha.jhu.edu/SdssCutout) that gives the SDSS images at requested sizes and scales optionally annotated with a grid and with object markers. This is part of the larger JHU SkyQuery.net effort that integrates several web-service archives. Gray also coauthored quite a few papers (and gave many talks) on the computer science challenges of the VO (see Publications).

National Optical Astronomy Observatories
D. De Young continued working with NVO SWG to refine the science goals of the three science demonstrations for NVO. He also revised and edited the final draft of the NVO Science Definition Team Report, and participated in the presentation of this report to NASA and NSF at NASA Headquarters.

De Young presented an invited paper, "Theoretical Astrophysics and the US-NVO Initiative," at the international VO meeting in Garching.

De Young continued to revise and assist in the refined definition of the work scope for the NOAO participation in the NVO, participated in weekly Executive Committee telecons and biweekly WBS Level 2 telecons, and hosted the NVO Team Meeting in Tucson, 16-17 April.

D. Tody led the data access portals design effort (WBS 5.4), and participated in system architecture design, data model architecture development, and metadata standards development. Activities focused on support for the first year science prototypes, and technology studies and discussions relating to the development of an architecture for data model-based data access. Cone search services were implemented for several catalogs. Data grid and Web services technologies were investigated to better understand the role they might play in data access. This included installing and reviewing the Storage Resource Broker (SRB) software from SDSC, and reviewing the computational architecture of the Teragrid to see how distributed analysis software might scale. A review of interesting Web Services toolkits was prepared and distributed. A concept for a set of Web services for querying image databases and retrieving image data (primarily image cutouts or mosaics) was developed and circulated for discussion. Discussions were held to further develop the concept of an extensible data model architecture.

National Radio Astronomy Observatory
NRAO participated in the metadata working group.

NRAO has been advertising for a scientist or software developer for our NVO funded work. D. Tody will be joining NRAO in September 2002 to fill this position and to lead NRAO's NVO work.

Even without a dedicated NVO person, we have been able to make progress on making connections between AIPS++ and web resources (including services). We now have a connection between Glish (the scripting language used in AIPS++) and any web page.

This is being deployed in our prototype archive server. We are also working to enable connections from Glish to and from Java, with the goal of eventually providing a simple way of accessing and providing web services.

We have developed and deployed a prototype radio astronomy archive server using the above Glish-Web page connection. Doing this required development of a first pass at meta-data for radio astronomy data sets: specifically meta-data for any data set stored in an AIPS++ MeasurementSet.

San Diego Supercomputer Center
SDSC is supporting formation of an initial NVO testbed that includes replication of collections onto environments where large-scale re-analysis operations can be performed. This includes working with collections to demonstrate replication, collaborating on re-analysis tasks to understand computation requirements, and collaborating on the NVO system architecture design.

Smithsonian Astrophysical Observatory
SAO continued to lead the Data Model design (WBS 2.1, 2.2) and the Metadata design (WBS 3.1) efforts. G. Fabbiano, J. McDowell, and A. Rots attended the Tucson team meeting in April. Fabbiano also participated in the international VO meeting in Garching in June and the AVO Science Working Group meeting. Fabbiano, McDowell, and Rots presented talks at the June VO meeting in Garching; I. Evans et al. presented a poster paper.

The SAO team monitored various design and implementation proposals related to data models, metadata, and the Grid. Sent comments on VOTable version 0.99; reviewing Web Services, SOAP, and UDDI.

Space Telescope Science Institute
In the period from April through June, activities at STScI (R. Hanisch, project manager) included the following:

- Organized and chaired the NVO project team meeting (April 16-17, Tucson). Achieved meeting goals of selecting and defining science demonstration projects for January 2003.
- Participated in presentation of NASA-NSF Science Definition Team final report to the agencies (April 11).
- Gave presentation to the new Science Archives Working Group at NASA HQ (April 30) on status of NVO project, focusing on our approach to project management and international collaboration.
- Co-organized the VO session for the forthcoming SPIE meeting.
- Gave project review presentation at NSF Headquarters (May 21) (also attended by J. Bredekamp, NASA HQ). In preparation for this meeting, updated NVO project management plan, project schedule, and science demonstration schedules.
- Wrote up material on NVO for Origins Roadmap.
- Wrote draft document describing resource and service metadata.
- Discussed NVO EPO activities with staff at the American Museum of Natural History (Hayden Planetarium).

- Met with C. Page and R. Mann of the UK AstroGrid project.
- Attended the conference "Toward an International Virtual Observatory," Garching, Germany, June 10-14 and presented an invited talk on the NVO Project and a report on resource and service metadata at the Interoperability Working Group discussion.
- Gave a talk on the NVO Project at the Oort Memorial Workshop on Astronomical Surveys, Leiden, The Netherlands, on June 17.
- Developed agenda for NVO Team Meeting (30-31 July).
- Participated in regular meetings of the NVO Executive Committee and NVO Metadata Working Group, and conducted biweekly Project Status Review telecons.

United States Naval Observatory

University of Illinois Urbana-Champaign/National Center for Supercomputer Applications

R. Plante continues to chair the weekly telecons of the Metadata Working Group. As part of general work for the MWG, he maintains the website of documents, participated in the development of the resource metadata, image access standard (on-going), general discussion of the first year science prototypes.

Plante is the lead for the Galaxy Morphology science prototype. He and J. Annis (FNAL) have developed the goals of the demo, the general recipe, and an implementation plan. Plante has worked out a plan with Rots and McDowell (CXC) to develop a special service for obtaining x-ray fluxes for given regions.

Plante developed a white paper discussing a Metadata Framework for the NVO, which he presented at the international VO meeting in Garching, June 10-14, 2002.

University of Pennsylvania

Penn participated in the design discussions regarding metadata concepts, initiated the design of a plausible new standard for time series data, and participated in the exploration of new science test-beds for the NVO program. A major, related activity at Penn is the design of a modular data analysis pipeline for NVO compliant data.

University of Southern California (ISI)

USC/ISC has been active in several areas of technology evaluation:

- Virtual Data Systems (Chimera)
- Data replica location services
- Authenticated remote database access

Also, in collaboration with the iVSDL project, we are evaluating tools for testbed set up and management.

University of Wisconsin

We participated in the April NVO meeting in Tucson, primarily in discussions about systems architecture and the use of Condor. We expect that as the demos are developed further we will contribute more substantially.

Publications and Presentations

Berriman, G., Kong, M, & Good, J. C. "OASIS: A Data Fusion System Optimized for Access to Distributed Archives," AAS Meeting 200, 60.01, 2002

Brunner, R. J., Djorgovski, S. G., Prince, T. A., & Szalay, A. S., "The National Virtual Observatory," ASP Conf. Proc. 254, 383, 2002

Davis, L. E., "Tools for Automated Astrometric and Photometric Calibration of Survey Data," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

De Young, D., "Theoretical Astrophysics and the US-NVO Initiative," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Djorgovski, G., "Data Mining Challenges in a Virtual Observatory," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Djorgovski, S. G., "The National Virtual Observatory Science Definition Team: Report and Status," AAS Meeting 200, 87.01, 2002

Evans, I., DePonte Evans, J., Cresitello-Dittmar, M., Germain, G., Glotfelty, K., Plummer, D., & Zografou, P., "Chandra Pipelines—Meeting the Science and Engineering Challenges," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Fabbiano, G., "CfA VO Activities and Data Model Testbed," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Genova, F., Benvenuti, P., De Young, D. S., Hanisch, R. J., Lawrence, A., Linde, T., Quinn, P. J., Szalay, A. S., Walton, N. A., & Williams, R. D., "International Collaboration for the Virtual Observatory," AAS Meeting 200, 87.03, 2002

Gray, J., "Computer Science Aspects of the Virtual Observatory—How Can We Federate All These Archives?" Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Gray, J., & Szalay, A., "The World Wide Telescope: An Archetype for Online Science," MSR TR 2002-75, pp. 4, June 2002

Gray, J., Slutz, D., Szalay, A., Thakar, A., Kuntz, P., & Stoughton, C., "Data Mining the SDSS SkyServer Database," MSR TR 2002-1, pp.1-40, 2002

Hanisch, R. J., "Building the Infrastructure for the National Virtual Observatory: An Information Technology Research Initiative of the National Science Foundation," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Hawkins, I., & Mattei, J., "Education and Public Outreach Summary from the NVO Science Definition Team Report," AAS Meeting 200, 60.08, 2002

Mahabal, A. A., McDermott, J. P., Babu, G. J., Feigelson, E. D., Djorgovski, S. G., Nichol, R., & Wasserman, L., "Implementing Astrostatistics in the Virtual Observatory," AAS Meeting 200, 60.05, 2002

McDowell, J., "Towards an Image Data Model for the VO," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

McGlynn, T. A., Angellini, L., Corcoran, M., Drake, S., Pence, W., Winter, E., Hanisch, R., Suchkov, A., White, R., Postman, M., Donahue, M., Genova, F., Ochsenbein, F., Fernique, P., & Wenger, M., "Classifying the High Energy Universe with ClassX," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

McGlynn, T. A., Suchkov, A., Angellini, L., Corcoran, M., Drake, S., Pence, W., Winter, E., Hanisch, R., White, R., Postman, M., Donahue, M., Genova, F., Ochsenbein, F., Fernique, P., & Wenger, M., "Classifying X-Ray Sources Using Multi-Wavelength Data," AAS Meeting 200, 60.03, 2002

McGlynn, T., Borne, K., Thomas, B., Eichhorn, G., Accomazzi, A., Rots, A., Berriman, B., Good, J., Kimball, T., & Mazzarella, J., "Linking NASA's Astronomy Data Archives," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Moore, R. W., "Data Grids for Collection Federation," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Moore, R., & Baru, C., "Virtualization Services for Data Grids," for publication as part of a collection on "The State of the Grid 2002"

Ochsenbein, F., Williams, R., Davenhall, C., Durand, D., Fernique, P., Giaretta, D., Hanisch, R., McGlynn, T., Szalay, A., & Wicenec, A., "VOTable—A Proposed XML Format for Astronomical Tables," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Plante, R., "Scalable Metadata Definition Frameworks," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Rots, A., "Space-Time Metadata for the VO," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Schade, D., "The Canadian Virtual Observatory Online: A VO Prototype System at the Canadian Astronomy Data Centre," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Szalay, A., "Analyzing Large Data Sets in Astrophysics (Distributed Computing and the Virtual Observatory)", Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Szalay, A., & Williams, R., "The NST ITR Project: Framework for the National Virtual Observatory," AAS Meeting 200, 87.02, 2002

Szalay, A., Gray, J., Thakar, A., Kuntz, P., Malik, T., Raddick, J., Stoughton, C., Vandenberg, J., "The SDSS SkyServer—Public Access to the Sloan Digital Sky Server Data," ACM SIGMOD 2002, MSR TR 2001 104

Tody, D., Davis, L., Fitzpatrick, M., & Valdes, F., "A Prototype VO Data Access Portal," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Voit, G. M., "Enabling Outreach with Virtual Observatories," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Williams, R., "The Uphill Battle of Semantic Interoperability," Towards an International Virtual Observatory, Garching, Germany, June 10-14, 2002

Williams, R. D., "Grids and the Virtual Observatory", for publication as part of a collection on "The State of the Grid 2002"

Acronyms

AAS	American Astronomical Society
ADC	Astronomical Data Center
ADEC	Astrophysics Data Centers Executive Committee (NASA)
AIPS++	Astronomical Image Processing System++ (NRAO)
API	Applications Programming Interface
AVO	Astrophysical Virtual Observatory
CACR	Center for Advanced Computational Research (Caltech)
CADC	Canadian Astronomy Data Centre
CDS	Centre de Données astronomiques de Strasbourg
CMU	Carnegie Mellon University
CXC	Chandra X-Ray Center
CY	calendar year
DAGMan	Directed Acyclic Graph Manager (Condor)
DAML	DARPA Agent Markup Language
DARPA	Defense Advanced Research Projects Agency
DM	Data Model
DOE	Department of Energy
DPOSS	Digitized Palomar Observatory Sky Survey
DTD	Document Type Description
EPO	Education and Public Outreach
ESTO	Earth Science Technology Office (NASA)
ESTO-CT	ESTO Computational Technologies (NASA)
FIRST	Faint Images of the Radio Sky at Twenty Centimeters
FITS	Flexible Image Transport System
FNAL	Fermi National Accelerator Laboratory
FTP	File Transport Protocol
FY	fiscal year
GB	gigabyte
GLU	Générateur de Liens Uniformes (uniform link generator)
GRB	Gamma Ray Burst
GriPhyN	Grid Physics Network
HEASARC	High Energy Astrophysics Science Archive Center
HTM	Hierarchical Triangular Mesh
HTTP	HyperText Transport Protocol
IPAC	Infrared Processing and Analysis Center (Caltech)
IRAF	Image Reduction and Analysis Facility (NOAO)
IRSA	Infrared Science Archive (IPAC)
ISI	Information Sciences Institute (USC)
ITWG	Information Technology Working Group (NASA data centers)
iVDGL	International Virtual Data Grid Laboratory
IVOA	International Virtual Observatory Alliance
JDBC	Java Data Base Connectivity (Sun, Inc., trademark)
JHU	The Johns Hopkins University

MAST	Multimission Archive at Space Telescope (STScI)
MB	megabyte
MOU	Memorandum of Understanding
MWG	Metadata Working Group
NASA	National Aeronautics and Space Administration
NCSA	National Center for Supercomputer Applications
NOAO	National Optical Astronomy Observatories
NPACI	National Partnership for Advanced Computational Infrastructure
NRAO	National Radio Astronomy Observatory
NSF	National Science Foundation
NVO	National Virtual Observatory
OASIS	On-line Archive Science Information Services (IRSA)
OGSA	Open Grid Services Architecture
OIL	Ontology Inference Layer
PB	petabyte
Q	quarter
QSO	Quasi-Stellar Object
RC	Replica Catalog
RDF	Resource Description Framework
RLS	Replica Location Service
ROME	Request Object Management Environment
SAO	Smithsonian Astrophysical Observatory
SAWG	Science Archives Working Group (NASA)
SAWG	System Architecture Working Group (this project)
SciDAC	Scientific Discovery through Advanced Computing (DOE)
SDSC	San Diego Supercomputer Center
SDSS	Sloan Digital Sky Survey
SDT	Science Definition Team
SOAP	Simple Object Access Protocol
SRB	Storage Resource Broker
STScI	Space Telescope Science Institute
SWG	Science Working Group
TB	terabyte
UCD	Uniform Content Descriptor
USC	University of Southern California
UDDI	Universal Description, Discovery, and Integration
UIUC	University of Illinois Champaign-Urbana
USNO	United States Naval Observatory
USRA	Universities Space Research Association
VDL	Virtual Data System Language
VDS	Virtual Data System
VO	Virtual Observatory
VO	Virtual Organization
WBS	Work Breakdown Structure
WSDL	Web Services Description Language
XML	Extensible Mark-up Language

2MASS Two-Micron All Sky Survey