Quarterly Report
October-December 2004

Building the Framework for the National Virtual Observatory

NSF Cooperative Agreement
AST0122449

NATIONAL VIRTUAL OBSERVATORY

INTERNATIONAL VIRTUAL OBSERVATORY ALLIANCE

Executive Summary 1
1 Management 2
1.1 General (planning, reporting, communications, team meetings, etc.) 2
1.2 Science 2
1.3 Technical (including standards, configuration management) 3
1.4 Financial 3
1.5 International coordination/collaboration 3
2 Science Requirements 3
2.1 Usage scenarios for all areas of astronomy research, including theoretical simulations 3
2.2 Requirements analysis 4
2.3 Demonstration definition and review 4
3 System Architecture 4
3.1 System design, components; relationships to Grid components 5
3.2 Computational facilities (processing, bulk data storage, network access, security, authentication) 5
3.3 Digital library integration 6
4 Registries 7
4.1 Resource metadata 7
4.2 Resource metadata schema 7
4.3 Publishing and harvesting protocols 7
4.4 Search protocols 8
4.5 Replication, synchronization, maintenance, revision control, and curation 8
5 Data Models 8
5.1 High-level (image, spectrum, time series, event lists, visibilities, catalogs, simulations, data quality) 8
5.2 Low-level (measurement, quantity, uncertainty, relationship) 9
5.3 Descriptors and ontologies (UCDs) 9
5.4 Space-Time and regions 9
5.5 Standard schema 9
6 Data Access Layer 9
6.1 Data access services (catalog, image, spectrum, time series, visibilities, …) 9
6.2 Data representation (VOTable, etc.) 11
6.3 Framework (mediators, components) 12
6.4 Data provider/consumer implementations and end-to-end testing 13
7 Query Language 14
7.1 Low-level: Astronomical Data Query Language 14
7.2 Mid-level: VOQL and OpenSkyQuery/OpenSkyNode 14
7.3 High-level: Complex queries 14
8 Web and Grid Services 15
8.1 Web Services (SOAP, WSDL, etc.) 15
8.2 Grid Services (OGSA) 15
8.3 Computational resource management 16
8.4 Virtual data 16
8.5 Application and service integration with Grid 16

9 Applications ... 17
9.1 Data location services ... 17
9.2 Cross-correlation services ... 18
9.3 Image combination, registration ... 18
9.4 Visualization tools and services ... 20
9.5 Theory ... 20
9.6 Statistical analysis ... 20
9.7 Datamining, outlier identification ... 20
9.8 Interfaces to/from legacy software systems ... 20
10 Community Engagement ... 21
10.1 Documentation ... 21
10.2 Web site ... 21
10.3 Technical training initiatives ... 21
10.4 Advocacy ... 21
11 Education and Public Outreach ... 22
11.1 Strategic partnerships ... 22
11.2 Formal education ... 22
11.3 Informal education ... 23
11.4 Outreach and press activities ... 23
11.5 Technical development ... 24
Activities by Organization ... 25
Caltech–Astronomy Department and Center for Advanced Computational Research (CACR) ... 25
Caltech–Infrared Processing and Analysis Center (IPAC) ... 25
Canadian Astronomy Data Centre/Canadian Virtual Observatory ... 25
Carnegie-Mellon University/University of Pittsburgh (CMU/UPitt) ... 25
Fermi National Accelerator Laboratory (FNAL) ... 26
High Energy Astrophysics Science Archive Research Center (HEASARC) ... 26
Johns Hopkins University ... 27
Microsoft Research ... 28
National Optical Astronomy Observatories (NOAO) ... 28
National Radio Astronomy Observatory (NRAO) ... 29
Raytheon/ADC (University of Maryland and George Mason University) ... 30
San Diego Supercomputer Center ... 30
Smithsonian Astrophysical Observatory ... 31
Space Telescope Science Institute ... 31
United States Naval Observatory ... 32
University of Illinois-Urbana/Champaign/National Center for Supercomputer Applications (UIUC/NCSA) ... 32
University of Pennsylvania ... 32
University of Southern California (USC/ISI) ... 32
University of Wisconsin ... 33
Publications ... 35
Virtual Observatory Articles in the Popular and Technical Press ... 37
Acronyms ... 38

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

Period covered by this report: 1 October—31 December 2004*
Submitted by: Dr. Robert Hanisch (STScI), Project Manager

Executive Summary

During the first quarter of FY2005 (Project Year 4) the NVO project team concentrated on preparing a suite of five science applications for release to the research community at the January 2005 meeting of the American Astronomical Society in San Diego. All five applications were successfully deployed, including provisions for mirror sites. Initial feedback from astronomers visiting the NVO demonstration booth at the AAS Meeting was extremely positive.

NVO registry services were upgraded to the current IVOA-approved resource metadata schema, and the OpenSkyNode/OpenSkyQuery system was made fully compliant with the Astronomy Data Query Language (ADQL) standard. A number of new catalogs were added as SkyNodes, including the NOAO Deep Wide-Field Survey. The Simple Spectral Access Protocol has matured to the point where prototype implementations have been possible, and these are showing that the protocol has the required capabilities.

NVO interactions with the Grid are increasing. A subset of the NVO team is actively using the TeraGrid. This year we will be exploring ways to simplify access to Grid-based computational resources and we will start to make greater use of the Grid for generating standard products.

A small number of education and outreach partnerships have been established and progress is starting to be seen. The domain name *virtualobservatory.org* has been secured as an EPO portal for NVO.

International collaboration remains strong. NVO personnel attended the fall IVOA Interoperability Workshop in Pune, India. NVO proposed that the IVOA create a new working group aimed at defining a common protocol for transient event notification. The new VOEvent WG was approved shortly thereafter. The IVOA Executive Committee submitted a proposal for an IAU Symposium on VO science.

Project spending is as expected. Carry-forward funds are being reduced at a rate consistent with the plan to maintain team activities through most of Year 5.

The External Advisory Committee met just after the January AAS meeting, and initial feedback was extremely positive.

---

* This report actually covers activities up through the AAS Meeting and External Advisory Committee meeting, i.e., 14 Jan 2005. Financial information is for the period closing on 31 Dec 2004.

Activities by WBS

1 Management

1.1 General (planning, reporting, communications, team meetings, etc.)

Regular weekly telecons of the Technical Working Group (TWG) continue, with this name replacing Metadata Working Group in recognition of the broader technical issues under discussion. Similarly, the Executive Committee also continues to meet weekly by telecon, and the Project Manager and WBS leaders have a biweekly telecon to discuss programmatic issues.

J. Good (IRSA/IPAC/Caltech) has assumed the role of chair of the NVO Technical Working Group, replacing Ray Plante (NCSA) who so ably led this group for three years. The group has begun to focus more on the requirements for supporting formal distributed development and system operations. This includes work in several areas:

- Operational Registry entry monitoring and validation
- Formal service testing, regression testbed, and change control
- NVO access toolkits
- Fast inventory services
- Distributed user subset collections
- Generic forms-based database search capabilities.

The primary emphasis is currently on setting-up and populating a formal NVO code repository.

The fall NVO team meeting was held in Socorro, NM in November, hosted by the National Radio Astronomy Observatory. The focus of this meeting was a final review of the applications to be released for community use in conjunction with the January meeting of the American Astronomical Society in San Diego. We also invited representatives from the LSST project to this meeting, and we are continuing to strengthen our collaboration. LSST offered to host the NVO spring team meeting in Tucson.

1.2 Science

The guidance of the Science Steering Committee, which met in July 2004, was essential in shaping the science applications released in January 2005. In Executive Committee and team discussions, we consciously agreed to emphasize general services and tools rather than end-to-end research, aiming for breadth rather than depth. This is in contrast to some of our IVOA collaborators, but overall we benefit from taking complementary approaches.

By releasing science applications to the community we are entering into a new phase for NVO. The availability of such applications calls for user support, documentation, call-tracking, etc. We are endeavoring to support such activities, within the current budget, and without seriously impacting work on the remaining development tasks.

1.3 Technical (including standards, configuration management)

Following the September Summer School, and with a mind toward providing a level of community support, we have deployed code management systems based on CVS at both Caltech and U. Maryland. There are differences in the user interfaces for these systems and differences in the level of institutional support, and an evaluation got under way in this quarter to see which system is a better match for the NVO team and the general community.

1.4 Financial

Overall project spending continues as planned. Our current carry-forward is slightly over $1M, down some $600k from one year ago. This is as planned, and is sufficient to keep the project team in place well into Year 5. Some reallocations of funds among subawards were made this year, accounting for an increase in the scope of work at SAO as well as the relocation of U Penn collaborators to SAO.

1.5 International coordination/collaboration

International coordination and collaboration remains strong. NVO team members attended the October 2004 IVOA Interoperability Workshop, which was hosted by VO India in Pune. NVO proposed that the IVOA create a new working group aimed at defining a common protocol for transient event notification. The new VOEvent WG was approved shortly thereafter. NVO hosted the IVOA Executive Committee meeting, just prior to the San Diego AAS Meeting (January 2005), at the San Diego Supercomputer Center in La Jolla. R. Hanisch planned to attend the AVO (Euro-VO) Science Working Group meeting near Madrid at the end of January 2005.

Members of the IVOA Executive Committee led a proposal to the IAU for having a Symposium on VO-enabled science in conjunction with the 2006 IAU General Assembly in Prague. Support for the proposal was obtained from numerous IAU Divisions and Working Groups.

2 Science Requirements

2.1 Usage scenarios for all areas of astronomy research, including theoretical simulations

An immediate priority for the US NVO is the creation and dissemination of science applications that will be readily accepted by the US astronomy community. The NVO Science Steering Committee (SSC), created in the spring of 2004, was assigned the task of creating a prioritized list of such applications without regard to issues of implementation. The SSC created such a list during their July 2004 meeting, and subsequent review by the team resulted in a list of five science applications that were felt to be highly relevant to the astronomy community and also to be practical in terms of their being deployable by the January 2005 AAS meeting. The five applications chosen were:

- Data Registry Service
- DataScope
- Spectrum Services

- Web Enabled Source Identification with Cross-Matching
- Open SkyQuery.

Essential to the success of these applications is their ease of use by the general member of the astronomy community. Progress in achieving this was greatly enhanced by the participation of D. Norman, an NOAO post-doctoral fellow at CTIO, who agreed to act as a "test particle" in trying out the applications as they were developed. Her advice was very useful in removing jargon from the applications descriptions and in streamlining the interface between the applications software and the potential user. Norman also assisted in the further development of the interface between the NVO project and the astronomy community by creating detailed documentation that describes how to use the NVO applications in solving a few specific sample scientific problems.

In addition, documentation was prepared on "How to Publish to the NVO" together with a one-page leaflet on the NVO, designed for appeal to the general astronomical community, for distribution at the AAS meeting.

2.2 Requirements analysis

During the period covered by this report, the development of the applications for the January 2005 release was continually reviewed and modified by the team members responsible for each application. As described above, continuing input about the ease of use of the applications by astronomers was incorporated in addition to the technical issues arising during the development.

2.3 Demonstration definition and review

One-page documentation sheets for each application were developed during this period for circulation at the January AAS meeting. At the meeting itself, demonstrations of the applications were presented daily to the astronomers attending the AAS meeting. After the meeting the NVO web site will contain links to each of the applications together with their documentation and the step-by-step "use case" demonstrations.

3 System Architecture

The goal of the architecture group is to define a coherent infrastructure that supports the services required by the astronomy community. The fundamental components based on web, grid, and digital library technologies have provided a viable infrastructure for building NVO services. However, consensus has not been reached on:

- Security. The issue is whether to rely on authentication by a user's home institution, or whether the NVO should run its own certificate authority.
- Grid services standard. The release of the Web Services Resource Framework is expected this year. The NVO will need to evaluate the robustness and performance of services based on WSRF.
- Workflow processing. Many implementations exist both as image processing pipelines within the Astrophysics community, and as generic workflow systems within the grid community. At the moment, the best approach is to seek interoperability between the systems.

- Preservation. Multiple communities are developing technologies for long-term preservation. The leading approach is based on use of data grids to provide infrastructure independence and manage technology evolution. Collaborations between the NVO and NSF, Library of Congress, and National Archives and Record Administration projects on preservation will lead to a standard.

A collaboration is being started with the Large Synoptic Survey Telescope project. The LSST is planning to use NVO services for standard access to their data. The LSST is also evaluating the capabilities of the NVO testbed to manage the massive amount of data that will be generated (on the order of 15 terabytes per day and a billion records per day). The LSST project will put pressure on the development of standard bulk data movement and bulk data management services in the NVO.

3.1 System design, components; relationships to Grid components

NVO continues to participate in the Global Grid Forum to evaluate grid technology standards for use in the NVO testbed. This is a joint effort with the IVOA.

3.2 Computational facilities (processing, bulk data storage, network access, security, authentication)

A major demonstration of the bulk processing capabilities of the NVO testbed is planned through use of the NSF TeraGrid. The NVO plans to process the entire 2MASS and DPOSS image archives to generate a HyperAtlas (led by R. Williams (CACR), J. Good (IPAC), and L. Brieger (SDSC)). Each page in the HyperAtlas will be assembled as a mosaic of images using the Montage mosaicing service. The steps required to analyze the 2MASS 10-TB collection include:

- Porting of the Montage environment onto the TeraGrid (J. Good, L. Brieger)
- Testing workflow environments for controlling the generation of the Mosaics (E. Deelman, L. Brieger)
- Porting the final releases of the 2MASS and DPOSS image archives onto TeraGrid storage systems (R. Williams, J. Good, G. Kremenek)
- Validating selected HyperAtlas pages for accuracy (R. Williams, J. Good)
- Monitoring the production workflow (L. Brieger)

This project exercises all of the computational facilities on the NVO testbed, including the distribution of images between TeraGrid nodes (SDSC and NCSA), authentication using grid certificates, bulk access to image collections stored in the SRB data grid, network access over the TeraGrid network, and sustained execution for approximately 30,000 CPU hours. The demonstration is being supported by the San Diego Supercomputer Center as an SDSC Strategic Application Collaboration. The goal is to complete the computation in 1st Quarter 2005 for the 2MASS collection, and then process the DPOSS collection using the same services.

Processing. A new release of the GridFTP transport technology has been made, version 3.9. To better support interactions between grid tools and data grids, this version of GridFTP has been ported on top of the SDSC Storage Resource Broker data grid, version

3.2.1. The integrated system will be available in the first quarter, 2005. The Montage/Chimera workflow system for generating Mosaics assumes the use of GridFTP to access images. Thus the GridFTP port was essential for interoperability between collections housed in the SRB data grid and Chimera workflow.

Bulk Data Storage. The final release of the 2MASS sky survey is being replicated onto TeraGrid storage. This required copying a terabyte at a time onto a terabyte disk, carrying the disk from IPAC to CACR, copying the data onto a TeraGrid node at CACR, and then moving the images over the TeraGrid network to a storage system at SDSC (R. Williams, G. Kremenek). Since the data retrieved from CACR is stored in the order seen by the telescope, the images will be sorted to co-locate images for the same area of the sky in a single container. This will make generation of mosaics much more efficient. A copy of the 2MASS collection will be kept at Caltech, sorted in temporal order, and a copy of the 2MASS collection will be kept at SDSC sorted in spatial order.

The final release of the DPOSS sky survey is also being replicated onto TeraGrid storage. This has required replacing selected images and adding missing images. However the bulk of the images did not have to be changed.

The terabyte disk used to support the 2MASS image replication is now being used to support replication of the USNO-B image collection onto the TeraGrid. This project is expected to be completed by June 2005 (collaboration with S. Levine).

The replication of the MACHO collection onto TeraGrid storage is being restarted in support of the LSST project (collaboration with T. Axelrod). The collection will be used to test processing pipelines.

Workflow Systems. The Chimera/Pegasus workflow system (E. Deelman) has been tested on the TeraGrid in support of Montage mosaic generation. Timing tests demonstrate that a substantial amount of I/O can be overlapped with computation. The amount of I/O that is not overlapped is about 10% of the total execution time. The total amount of I/O (including I/O overlapped with computation) is about 35% of the total execution time. The creation of mosaics remains a CPU intensive task.

3.3 Digital library integration

The integration of digital library services with the NVO testbed is being pursued at JHU (A. Szalay) and SDSC (R. Moore). The two leading candidates for digital library front ends are DSpace (from MIT) and Fedora (from Cornell). An integration of DSpace version 1 with the SRB has been done. A collaboration with Carl Lagoze at Cornell is being initiated to integrate Fedora with the SRB data grid.

In a second effort, JHU (W. O'Mullane) is developing a specification for a “VOSpace” personal digital library that can be used to organize image collections. Once the specification is complete, an implementation will be demonstrated on top of the SRB data grid.

Preservation. Multiple projects within the NSF are leading to the implementation of preservation environments for scientific data. The approach that is emerging with the strongest support differentiates between the community expertise needed to select, organize, and describe important digital holdings, and the facility that actually manages the preservation. From this perspective, the NVO is a preferred collaboration for providing the astrophysics community expertise. The SDSC NSF supercomputer center has submitted preservation proposals to NSF to provide the preservation facility. Integration of a preservation proposal from the NVO with the preservation facility proposal from SDSC promises to provide the quickest path to a viable preservation environment.

4 Registries

In the last quarter, the effort in this area has focused on readying our registries for the January product release. Most of this activity is reported in WBS 4.5.

4.1 Resource metadata
A. Rots (Harvard-CfA) released V1.1 of the Space-Time Coordinates metadata Working Draft; we continue to work on issues of integration of this metadata in to the resource metadata used by registries.

An important push for the coming year will be to submit standard documents for some of the resource metadata extensions now in use with the registry.

4.2 Resource metadata schema
In this last quarter, we exercised a previously unused subset of service metadata that allows one to register HTTP Get services, the most common form of non-standard web service. These metadata allows data providers to register their existing services. NED services represent an important example of this class of services and served as our driver. Not only is it possible for users and applications (like DataScope) to locate these existing services, it's possible to generate browser-based interfaces to them on-the-fly from the registry description.

Full support for VOResource V0.10 (zero-point-ten) was completed for the searchable registry at STScI in this last quarter in preparation for the January release.

4.3 Publishing and harvesting protocols
R. Plante (NCSA) continues to make contributions to the refinement of the Registry Interface Working Draft standard. NVO registry upgrades in preparation for the January release included supporting the standard harvesting protocol. In particular in this last quarter, full harvesting compliance was completed for the searchable archive at STScI.

M. Graham (Caltech) has released an exportable version of Carnivore (http://mercury.cacr.caltech.edu:8080/carnivore) that allows data providers with many resources to run their own registries. This package is currently being used to set up a publishing registry at the Harvard-Smithsonian Center for Astrophysics (CfA). With this latest Carnivore release, Graham and Plante plan now to combine the capabilities of Carnivore and NCSA's VORegistry-in-a-Box into a single product, using Carnivore as a base.

4.4 Search protocols
NVO registry upgrades also included support for the standard search protocol defined by the Registry Interface standard. Basic support for the ADQL based standard search interface has been implemented for both Carnivore and the STScI registry based on the working draft. These implementations will likely require additional refining in the coming year as the standard evolves to a proposed recommendation.

4.5 Replication, synchronization, maintenance, revision control, and curation
Significant effort has gone into reviewing the user interface to the STScI registry in preparation for the January release. The project enlisted the help of D. Norman (NOAO) as an "external" reviewer of the user interface. R. Hanisch, G. Greene, and W. O'Mullane used her feedback to improve the user experience. These improvements included better organization of registry query results to make it easier to zero in on resources of interest. For example, the top of the results page features a summary of the number of matches of each type of resource (e.g. Cone Search, SIA, etc.), with each linking to a display of just resources of that type. In addition, result summaries for Cone Search and SIA resources include a "Try Me!" button; when pressed, a simple interface for querying that service directly is displayed. A significant number of improvements came in the form of subtle changes to the look and feel that makes the registry easier to understand for first time users.

Work continues on improving the registries' user interfaces. Currently in development is an intermediate-level search form—i.e., between the simple keyword search and the advanced SQL-based interface in terms of ease of use—which can be adapted to any existing searchable registry. Such a form will make it easier to find specific types of resources—such as a specific service type, like SIA—with constraints on specific resource metadata (e.g. the publisher) without having to know SQL. In addition, we plan to propagate the improvements made to the STScI registry to our other registries.

5 Data Models

5.1 High-level (image, spectrum, time series, event lists, visibilities, catalogs, simulations, data quality)
The SED/Spectrum model has been the main development during this reporting period. After extensive discussions among J. McDowell, T. Budavari, D. Tody, and M. Dolensky, we froze version 0.92 of the model for implementation. The current document includes the model and FITS, VOTABLE, and XML-schema serializations. The JHU team has generated a spectral service using the XML-schema serialization of this model. The SAO team is working on a full software implementation of the model and will carry out interoperability experiments with the JHU group.

Recent changes to the model have included a flattening of the elements for each point (for compactness), an improvement of the sub-model for redshifts, addition of specified units, and clarification of when items are or are not optional, as well as overall clarification of the document.

5.2 Low-level (measurement, quantity, uncertainty, relationship)

J. McDowell has carried out discussions on the Characterization model with the Strasbourg group. A further technical meeting was held at the January AAS.

5.3 Descriptors and ontologies (UCDs)

J. McDowell proposed additional UCDs for use with the SED model. Feedback is pending.

5.4 Space-Time and regions

The Space-Time Coordinate metadata design was revised by A. Rots following comments and an extensive discussion with D. Berry (AstroGrid). The modified design was published as STC Version 1.10. The XML schemata were also revised to reflect the changes in the design and improve its generality through increasing the robustness of its inheritance and polymorphism aspects. These schemata were also published. In addition, Rots created a project page for STC on the us-vo.org web site and added to the documentation an explicit description of the XML schema implementation design.

The most important changes are:

- Pixel Space specification
- Allow UNKNOWN for certain elements
- Allow ranges for Coordinate components other than Name and Value
- Provide for the specification of rest frequencies
- Improved generalization, inheritance, polymorphism
- Added AllSky to Region shapes

Finally, Rots prepared a poster on STC for the AAS meeting.

Version 1.10 is expected to be a stable version and to mark the end of the Version 1 development. The effort will now concentrate on coding certain region functionalities.

5.5 Standard schema

J. McDowell proposed an interim schema for Characterization, which was discussed in Pune (and later at San Diego).

6 Data Access Layer

6.1 Data access services (catalog, image, spectrum, time series, visibilities, …)

Catalog access. The primary DAL interface for catalog access is the basic SkyNode, developed by W. O'Mullane and others within the VOQL group. SkyNode development has been mostly frozen over the past quarter to provide a stable environment for the January applications rollout at the AAS.

The older cone search interface continues to be heavily used despite its simplicity. The simple REST-based cone search interface will continue to be supported as an alternate interface for catalog access services, with both REST and web service interfaces accessing the same underlying service and returning the same VOTable response.

Image access. The simple image access interface (SIA) is heavily used in the current NVO applications, hence is frozen until summer 2005 when the next version is scheduled to appear. In the meanwhile the technology required for the next version is being developed and tested in the new spectral access interface (SSA) and in various prototypes.

The intention is that DAL interfaces such as SIA and SSA share much the same query interface. Generic dataset metadata for dataset identification, coverage, characterization, etc., will be identical for all classes of data. The main difference will be in the object model and associated metadata (2D sky image, 1D spectrum, etc.), and in the service functionality required to generate such a dataset. The functionality required to generate an actual image, spectrum, or time series dataset will in general be very different.

The data modeling effort for the next version of SIA is thus part of SSA development, described below. A group led by F. Bonnarel (CDS) and A. Micol (ST-ECF) is developing a prototype version of SIA to test both data characterization and the new SIA metadata extension mechanism. This prototype will use data from HST WFPC2 associations as a use-case to test the capability of the next version of SIA to model complex datasets. Y. Shirasaki (NAOJ) is developing a prototype extension of SIA to use ADQL and virtual table technology to add general query language capabilities to SIA (we will extend this later to all the DAL interfaces). The Grid and Web services group is developing the asynchronous job execution technology needed to add asynchronous data staging capabilities to SIA.

Spectrum access. The Simple Spectral Access (SSA) interface is the newest DAL interface, and the prototype for the next generation of DAL interfaces. SSA includes support for all tabular spectrophotometric data, including SEDs, 1D spectra, and time series. The SSA interface consists of the SSA data model, mappings of the data model to dataset representations in various formats (VOTable, FITS, XML, etc.), and a query interface.

The SSA data model (led by J. McDowell and D. Tody) reached V0.92 in late November, suitable for initial implementations, several of which have already occurred (see below). Refinement of the spectral data model and the more general characterization of, e.g., dataset models, continues. A version of flux scaling based on dimensional analysis contributed by P. Osuna was integrated.

DAL data access deals with both simple, precomputed archive data files and with virtual data, which is generated on the fly. In the process of developing SSA we have found that virtual data can be further broken down into several subcategories. For spectral data these are the following:

- Observed. A calibrated spectrum as observed by some instrument. This case applies to most archived spectra.
- Composite. A combination two or more observational spectra. Similar to “observed” but the processing is more involved since the spectrum is composite.
- Synthetic. A spectrum computed from some model (including a composite of more than one model, which is just another model).
- Simulated. A spectrum computed OTF from more fundamental data, e.g., image data, event data, pulsar data, etc.

Observed, composite, and synthetic data are precomputed but data access may require on the fly calibration, data model mediation, or filtering. Simulated data in effect implements a simulated instrument, e.g., producing a 1D spectrum from a spectral data cube, or an image from event or visibility data (for example the SED generator identified as a priority by the NVO Science Advisory Committee).

With completion of the V0.9 SSA data model, work has gone forward on the query interface, which is based upon the SSA data model. A partial draft of the interface document was completed by M. Dolensky and D. Tody around the end of the year. Completion of the V0.9 interface is expected in early 2005, to be followed by the first full end-to-end implementations. The goal is to complete an initial round of implementations by the time of the May interoperability meeting.

Time series. Basic support for access to spectrophotometric time series data is included in SSA. No further work on time series data is scheduled until trial implementations go forward later in 2005.

Event and visibility data. At this stage event and visibility data are being handled by accessing such data in the VO through a higher level interface (as an image, spectrum, etc.), while working to generate standard science data models (SDM) for the native data. In the case of interferometric visibility data most of the effort recently has been on a new SDM for ALMA, which will operate in the millimeter wavelength regime. This is in the process of being extended to EVLA data (centimeter wavelengths) and to VLBI, which has unique problems of very widely spaced antennas requiring precise timing. Work has also begun to define a new SDM for single dish radio data. These data models will build upon the work done on the VO data models for dataset identification, characterization, etc.

6.2 Data representation (VOTable, etc.)

Further revisions were made to the SSA VOTable, FITS, and XML data formats for V0.9 by J. McDowell, based on feedback from T. Budavari and others. The FITS representation was modified to conform to the proposed new FITS spectral WCS specification for coordinate tables (-TAB).

An issue is whether the spectral coordinate should be fully sampled or should permit use of a dispersion relation as for FITS WCS (especially for linearized spectra). In the initial version we decided to use a fully sampled coordinate vector since it works for all cases and avoids us having to develop a standard way to specify WCS for all data representations.

6.3 Framework (mediators, components)

Much of the work in the DAL area in this past quarter dealt with framework issues and trying to understand what is needed to provide a service framework for constructing DAL services and for interfacing the DAL services to client data analysis software.

The primary focus of NVO up until now has been on data mining of fully processed data stored in archives. While this is central to VO, a large portion of astronomy still deals with hands-on data processing and analysis of data from PI observing programs. In this scenario users process or reprocess data from specific observing programs, often writing their own software or processing scripts to customize processing, and then perform multiwavelength analysis comparing "their" data with data from public archives. Data processing and analysis is driven from the desktop but (in the VO era) data, computation, and collaborators may be either local or remote. The amounts of data and computation involved range in size from modest to very large.

Both data processing and analysis and the "data access" part of a DAL service (to construct the virtual data product returned to the client) are similar in nature, involving complex and numerically intensive scientific computation that works directly on astronomical data. For performance reasons, the most numerically intensive computations always take place on a localized, high performance computer (workstation or cluster).

Hence the framework needed emphasizes complex, numerically intensive scientific computation on a workstation or cluster. Most of the numerical software (science software) should be in the form of reusable components which do not know how they are being used, hence may be used either on the client or the server, stand-alone or connected to the VO/Grid. The VO/Grid interface occurs at the boundaries of this science data processing engine. A "VO-Client" layer integrates with the local data analysis system and implements the client side of the VO services required for data analysis (primarily registry and DAL). A conventional Web/Grid services layer is used to interface services built upon the processing engine to the VO, e.g., to build computational components to participate in large-scale Grid workflows.

D. Tody and R. Williams participated in a number of meetings over the past quarter to develop these concepts and help build a community consensus on what is required. A special session was held at the ADASS conference in October to discuss these issues with the broader community. D. Tody has participated over the past year as the VO representative in an OPTICON working group on future astronomical software environments that includes representatives from all the major data analysis systems. A meeting of the OPTICON working group was held at ESO in December to consider both the requirements for such an environment, and the architecture required. A white paper on the architecture required was prepared for this meeting and has since been endorsed by the working group. Further reviews within the NVO and IVOA communities are in progress. A special session to discuss the issue of data analysis and VO is planned for the May interoperability meeting in Kyoto.

6.4 Data provider/consumer implementations and end-to-end testing

Several sites including JHU (T. Budavari and L. Dobos) have implemented spectral access using the SSA V0.9 data model and spectral data representations. We will need to add a SSA query interface later before these are fully standard services.

Y. Shirasaki (NAOJ) has implemented SIA and SkyNode services. While there are numerous such services already available, this is the first implementation we are aware of within Japan, and will be useful for prototyping the ADQL extensions to SIA.

The DPOSS (Digital Palomar Observatory Sky Survey) image data (3 TB) has been rebuilt and cleaned, and is available on disk from the TeraGrid resource at Caltech. The images are available through the Simple Image Access Protocol, in several derivations: original images, a cleaned set (cropped, flattened, re-moded), and a tiled version with simplified headers. The 2MASS atlas images (9 TB) have been copied from IPAC to the Caltech TeraGrid machine, and services are being built for both point and bulk access to these. The DPOSS catalog is now being served as a SkyNode from Caltech, so that dynamic cross-matches can be done between SkyNodes in Baltimore (Sloan, 2MASS, etc) and Pasadena.

The NASA Extragalactic Database project has been funded by NVO to make all of its holdings available as NVO services using standard VO data formats for requests and responses. This is an ongoing process, in part because the NED holdings are much richer in detail and nuance than the more basic datasets NVO has dealt with to date. However, this has proven to be very valuable as a “test particle” and driver in several NVO areas. Specifically in the last quarter the NED services have been used as a first case of parameter-based services in the NVO Registry. Also, NED is an early user of the UCD+ development for dealing with its spectral holdings.

R. Plante has implemented Web clients for detailed testing of cone search and SIA service instances.

M. Fitzpatrick has implemented a web services front-end to the IRAF system. Essentially any IRAF functionality encapsulated in a CL script can be exposed as a Web service. This work to interface IRAF to NVO is in accord with the framework architecture discussed in the last section; in particular the VO/Grid interface is the same.

A group within ALMA and NRAO are working on "componentization" of AIPS++ code in accord with the DAL framework architecture. This includes implementation of a prototype parameter mechanism as well as a C++ container. Related prototypes are starting up within ESO (the SAMPO project) and RadioNet.

7 Query Language

7.1 Low-level: Astronomical Data Query Language

With the publication of ADQL V0.7.4 consensus was reached at the fall IVOA Interoperability Workshop in Pune, India that it was time to go forward with a Proposed Recommendation (PR). M. Ohishi (Chair of IVOA VOQL Working Group) requested publication of ADQL V0.9 to bring it closer to 1.0. The only addition from V0.7.4 was to add a bracket notation [ ] to allow reserved words for column names. A request from the AstroGrid project for further testing of the full V0.9 specification led to a delay in the call for promotion to Proposed Recommendation. We will probably now delay ADQL V1.0 until the May 2005 Interop Workshop in Kyoto, when we should have three implementations (OpenSkyQuery, AstroGrid, JVO).

7.2 Mid-level: VOQL and OpenSkyQuery/OpenSkyNode

Work continued on the OpenSkyQuery (OSQ) portal at JHU. A good deal of effort was put into making this ready for the AAS meeting in January 2005. A major deficiency mentioned by several parties was the ability to build a query. OSQ now contains a highly original and quite sophisticated ADQL Query builder; see www.openskyquery.net and click on Tutorial. VOPlot was also integrated to allow plotting of results directly. This was presented at ADASS in Pasadena. A good deal of effort also went into documentation of the site.

Assistance was provided to aid the integration of the web services SExtractor (WESIX); the integration of this system improved and tested the portal interface. It also went fairly smoothly.

Turning the Java node into a Full SkyNode with the Xmatch cross-match function continues to be a problem. A meeting was held at ESAC (European Space Astronomy Center, Villafranca, Spain) that concentrated on this and SSAP. We ran into a technical problem with the AXIS SOAP library. NCSA believes they have this working, we hope to resolve this issue post AAS. We would then soon have a full implementation in Java.

7.3 High-level: Complex queries

E. Shaya described current work on ontologies to team members during the ADASS conference. Discussion continues on the feasibility of basing a high-level query language on an astronomy ontology.

Y. Shirasaki (Japan) continues to develop his ideas on extensions for ADQL, which will be looked into in more detail after ADQL1.0 is a Proposed Recommendation. We continue to work with the DAL and the Data Model groups on this front.

8 Web and Grid Services

8.1 Web Services (SOAP, WSDL, etc.)

A prototype framework was successfully implemented for interfacing any IRAF task to a SOAP web service. The framework required the development of a web-service adapter acting as a front-end for an unmodified IRAF installation running persistently in the background, the web service itself is implemented using standard Java tools (Tomcat/ AXIS) and custom servlets. New functionality can be deployed using IRAF scripts tasks or interfacing directly to existing IRAF tasks. Additionally, work was begun on a tool to automate the generation of the web service interface using an XML configuration file that describes the service and generates the Java implementation. The tool itself will be available as a web service for community use, allowing non-programmers to create web services more easily from IRAF tasks.

WESIX (Web Enabled Source Extraction and Cross-match) does source extraction and cross-matching for any astrometric FITS image. The client uses the web page (or SOAP API) to upload the FITS image, and the remote service runs the SExtractor software for source extraction. The resulting catalog can be cross-matched with any of several major surveys, and the results returned as VOTable. The web page also allows use of Aladin or VOPlot to visualize results

In the future, the coaddition, source detection and cross matching will be interconnected through the data grid services to provide an application that enables an arbitrary stack of images to be constructed (selected based on frequency or quality of data), sources detected and cross-matched. This, we believe, will be an example of one of the core products of the NVO (the ability for a user to select, analyze and return arbitrary sets of data). The interplay between data selected by a user, the resulting object catalogs and existing databases of catalogs will provide a detailed test of the grid operations and services.

8.2 Grid Services (OGSA)

USC/ISI has created a web service interface for submitting montage workflows for execution over the Grid resources. This web service interface can be invoked using a command line client. This client, called mGridExec, takes as argument the name of a zip file. The zip file contains all the information and the abstract workflow required for processing the request. This web interface decouples the development of the front-end, where the request is submitted, from the processing on the Grid. The client returns a job identifier that can be used to track the status of the request. Upon completion, an email is send to the user indicating the result of the processing. This web service also creates a HTML page containing links to all the data products created by the workflow. The URL of this html page is sent to the user in the notification email in case the request completes successfully.

SDSC has developed a new command-line HTTP data subsetting service for us with scripts. The service provides a subset of a FITS image and recomputes the FITS header. The service has been used for extracting sub-images from the DPOSS image archive that is stored in the SRB data grid. The service uses SRB server side proxies to extract the sub-image at the remote storage system, and is based on the latest release of the Montage code (J. Good, G. Kremenek).

For example, the following http command extracts a sub-image:

```
"http://users.sdsc.edu/~kremenek/SRB/get_DPOSS_file_by_name_in_pixels.c
gi?FN=f176.fits&X1=555&X2=500&Y1=21&Y2=600" > /tmp/f176.4.fits
```

SDSC also developed a statistical image co-add function to composite images taken from multiple locations. The goal is to see whether a sufficiently large number of co-added images will be able to detect a signal in the 2MASS survey for selected objects that are visible in the SDSS survey but not visible in the 2MASS survey. (A. Szalay, in collaboration with A. Jagatheesan (SDSC) and J. Jacobs (IPAC)).

8.3 Computational resource management

ROME, the Request Object Management Environment, has been developed at IPAC to provide an EJB-based toolset for keeping track of jobs and messages for NVO users. It allows service and data providers to control job queues on a application-by-application basis, thereby ensuring that resources do not become overloaded and ensures that jobs get done (even if underlying systems fail) and that users are notified of job completion, failure, and even running status.

In the last quarter, ROME has been deployed to underpin the Montage web service described in WBS 9.3. The deployment included two major efforts: large-scale load testing was completed (tens of thousands of jobs submitted and successfully tracked, even with errors in the underlying applications), and the reference DBMS implementation was successfully migrated to mySQL, completing evolution to a completely open-source architecture (the other components being JAVA and the JBOSS EJB container).

8.4 Virtual data

SDSC plans to release the next version of the SRB data grid in 1sst quarter 2005. The updated version will support bulk data movement between federated data grids, queries across federated data grids, and improved synchronization scripts for federating data grids. This is work primarily supported by other projects.

8.5 Application and service integration with Grid

Prototypes have been developed at Caltech for a Science Gateway, meaning an application-specific task wrapped as a web service, and these have been implemented on the NSF TeraGrid cyber infrastructure. The Gateway is for mosaicing astronomical images, and features a graduated security model, "HotGrid", which is a means of providing simple, immediate access to the Grid. The secondary purpose of HotGrid is to acclimate a science community to the concepts of certificate use. The system provides these weakly authenticated users with immediate power to use the Grid resources for science, but without the dangerous power of running arbitrary code. The implementation of these Science Gateways is done with the Clarens secure web server.

9 Applications

During this quarter a major focus of activities for the NVO was the development of a suite of applications for release at the January AAS meeting. Five software applications had been selected for development during the summer of 2004. The applications were built, tested, refined and documented during the final quarter of 2004. The selected applications, the VO registry portal, DataScope, SkyQuery, Spectrum Services, and WESIX source extractor, provided a sampling of the broad range of services that the VO will ultimately provide.

A coordinated effort by many members of the NVO team worked to ensure that the software was effective, robust and well-documented. Many developers tested services and provided bug reports. An external review of the tools was performed by D. Norman (NOAO) and simple user guides were prepared by her. In several cases an animated help movie was prepared. Mirrors of some services were provided to provide backups in case of failure of network connections. This also ensured that the code base of the service was sufficiently robust to work in different locations. Code for the services was deposited in the NVO code repository or made available at the service web site.

The successful preparation of these services involved the participation of a large fraction of the NVO team. The NCSA, U. Pittsburg, IRSA, JHU, STScI and HEASARC were particularly involved. Individual services are described below in the appropriate subsections. The development of these services often involved efforts in other areas of the WBS. All of the services were linked to documentation at the NVO web site at http://us-vo.org/apps.

All of these services were successfully demonstrated at the AAS meeting with many positive comments from viewers. For example, during one demonstration a user found that an important feature visible in GALEX data was also easily visible in UIT imagery, likely changing a paper in the process of being submitted.

9.1 Data location services

JHU and STScI enhanced the existing interfaces to the VO registry to provide an easy to use portal to the VO registry for the January AAS. The basic portal provides a simple keyword search capability to look for appropriate resources in the registry, a registry browser, and a simple tool for accessing standard service types. Both primary and backup sites for the registry portal were provided. More sophisticated queries of the registry can be performed using advanced interfaces, but the basic portal service makes the VO registry accessible to the novice user of the VO.

The Data Inventory Service, rechristened NVO DataScope, was enhanced and upgraded by the HEASARC for release at the January AAS meeting. Features were added to the service to address issues with the underlying cone and SIAP services and to make the transition to analysis tool more straightforward. The NCSA brought up a mirror of the DataScope at their site, addressing issues with different versions of the underlying Perl libraries. The documentation for the DataScope was greatly expanded and enhanced.

9.2 Cross-correlation services

The OpenSkyQuery service was enhanced and upgraded for release at the January 2005 AAS meeting. Several new SkyNodes, including the GALEX UV survey catalogs, were added to OpenSkyQuery and the documentation and appearance of the interface were brought in line with NVO standards.

A meeting was held at JHU with R. Grossman's group from the Laboratory for Advanced Computing (University of Illinois at Chicago). We discussed building a cross-match for entire distributed large catalogues. This type of query cannot be handled by OpenSkyQuery at the moment. The LAC UIC group agreed to see if they could reproduce the brown dwarf search for the entire SDSS-DR2 and 2MASS as a demonstration. Help on XMATCH and the search were pledged by JHU (M. Nieto-Santisteban). D. Hanley has reported some success, with it taking only 15 minutes to do the search on a local setup. We still need to assess this with distributed databases and validate the results.

A proposal to build a *Parallel Cross-Match Engine for Astronomy* was submitted to Microsoft Research (M. Nieto-Santisteban). The proposal was highly ranked and has been accepted. The project aims to build an engine able to perform massive cross-match jobs between big catalogs (about a billion rows each) in parallel using HPC Beowulf clusters hosting SQL-Server. The goal is to create a scalable and flexible system where users can specify their own XMATCH procedures.

9.3 Image combination, registration

The University of Pittsburg released WESIX (Web Enabled Source Identification with Cross-Matching). WESIX enables users to upload an image to an encapsulated SExtractor engine, extract sources from the image and then cross-match the resulting source list with major catalogs using VO OpenSkyQuery protocols. The user is returned a list of identified and matched objects. The WESIX approach of encapsulating legacy software (SExtractor in this case) is essential if the VO is to be successful.

Many of the high-quality sky surveys concentrate on the catalog output as the primary product. However, this science product comes from the images; it is a set of large mosaics with careful reprojection to minimize astrometric and photometric distortion. All of the following surveys are being projected to the same pixel grids (the "HyperAtlas" set), so that these surveys can be federated by stacking and deeply mined. The 2MASS image dataset is a high-quality infrared sky survey of about 10 TB. It has been replicated at Caltech on disk, and also at SDSC on the SRB. SDSC is using the TeraGrid Datastar machines and the NASA Montage software to reproject this dataset. This work will enter full production in 2005.

The Digital Palomar Observatory Sky Survey (DPOSS) is a 3 TB optical sky survey that has been rebuilt and cleaned in 2004 using TeraGrid resources. A derived version has been created and stored on the Caltech PVFS, which has registration marks cropped, with flattening and re-moding, to allow for mosaicing. Some 20 HyperAtlas pages have been built with the very slow Montage 1.7, using some 4000 node-hours at Caltech to produce 30 GB. The algorithm is being rebuilt, replacing the complex polynomial plate solution with a much simpler piecewise TAN projection, and we expect a 20-fold speedup. The reprojection is more complex than 2MASS, but the survey covers only the northern hemisphere, and we estimate that 50000 CPU-hours will be needed. The full DPOSS HyperAtlas will be produced in 2005. The service is accessible from

http://pissarro.ipac.caltech.edu:8000/applications/Montage/

and the password to run it is available on request. Figure 1 below shows a 0.2 degree x 0.2 degree J-band mosaic of M42 computed through this portal. The mosaic took 69 seconds, and was returned along with a list of if input files and associated metadata and a link to the OASIS for interactive study of the mosaic. Work is currently underway to integrate the Montage image mosaicing software, NVO image archive access functionality, the Pegasus Grid execution toolkit, and ROME (for request management; see WBS 8.3) into a fully operational mosaic-on-demand service for NVO. IPAC is collaborating with ISI on the latter.

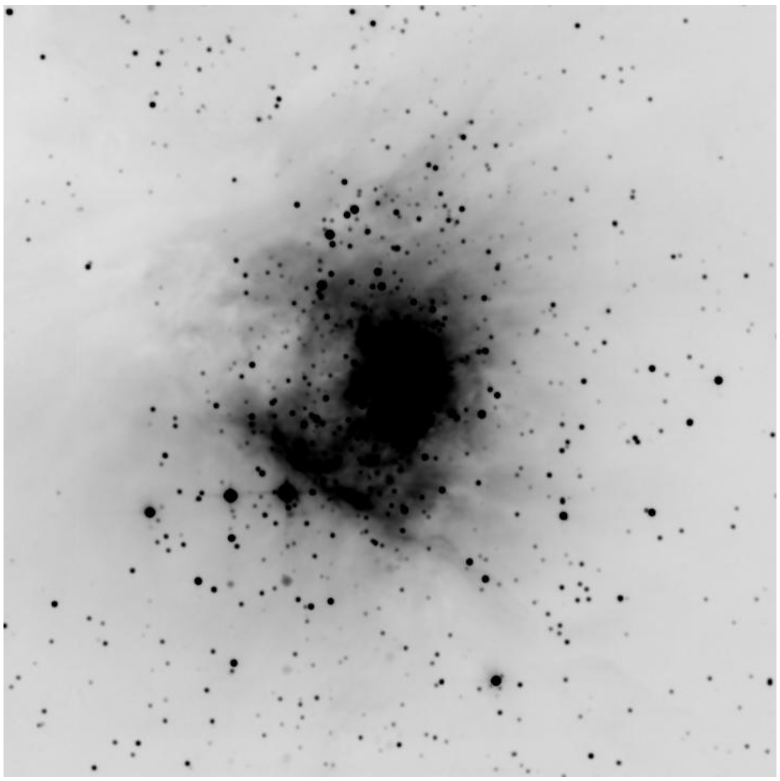

Figure 1. An image mosaic of M42 produced with the image mosaic service.

A new algorithm has been developed at Caltech for coadding overlapping images and removing edge effects. This "edge-removal" algorithm uses a weight function for each contributing image that tapers to zero at the edges, and maximizes at the center of the image.

The Palomar-Quest survey is running a catalog pipeline at NCSA and an image pipeline at Caltech, to reproject images to HyperAtlas so they can be stacked. This will allow searching for very faint quasars, as well as image subtraction to find transient sky sources.

9.4 Visualization tools and services

JHU enhanced and upgraded the spectral and filter services for release at the January AAS meeting. These allow users to easily view and combine spectra from the SDSS and 2dF surveys. They also make it easy for users to publish their own spectra within the VO. Documentation was improved and the appearance of the service was made consistent with the other NVO services.

9.5 Theory

Discussions continued in the IVOA Theory Interest Group during the IVOA Interop Workshop in Pune, India. Three types of theory applications were identified, and sub-groups were identified to study the issues associated with making each type available in the VO framework:

- Large-scale simulations (hydro codes, n-body simulations)
- Models (spectral synthesis, galaxy evolution)
- Atomic and molecular databases

9.6 Statistical analysis

Development continues on the VOStat software package, which was shown at the September 2004 NVO Summer School and at the October 2004 ADASS Conference.

9.7 Datamining, outlier identification

Nothing to report this quarter.

9.8 Interfaces to/from legacy software systems

A prototype framework was successfully implemented for interfacing any IRAF task to a SOAP web service. The goal is to provide a capability of fielding IRAF applications as web services, and to package that capability such that any user with an IRAF installation could field one or more IRAF applications for community use. The framework required the development of a web-service adapter acting as a front-end for an unmodified IRAF installation running persistently in the background, the web service itself is implemented using standard Java tools (Tomcat/AXIS) and custom servlets. New functionality can be deployed using IRAF scripts tasks or interfacing directly to existing IRAF tasks. Additionally, work was begun on a tool to automate the generation of the web service interface using an XML configuration file that describes the service and generates the Java implementation. The tool itself will be available as a web service for community use, allowing non-programmers to create web services more easily from IRAF tasks.

10 Community Engagement

10.1 Documentation

A handout was prepared for distribution to the astronomical community at the AAS meeting in San Diego in January 2005 by D. Norman (NOAO), R. Hanisch (STScI), and R. Williams (Caltech). This handout provides an introduction to five of the NVO core applications: the Registry Portal, DataScope, Open SkyQuery, Spectrum Services, and the SExtractor Web Service. In a semi-tutorial manner, the use of each application is explained through a series of screenshots with associated text boxes. The use cases demonstrate how one might solve a realistic problem using each tool. There were also six posters, one for each application, and a single tri-fold "handout" explaining the NVO and the applications very concisely to astronomers. A rack of NVO and IVOA publications was prepared.

10.2 Web site

The NVO web site (http://us-vo.org) is the main point-of-entry for the research community. We continue to improve the site to make it easy to navigate and easy for users to find the tools and applications of interest. The software that we released for general use in January 2005 is featured directly on the home page.

The web site has been running with dynamic content management throughout 2004. In this quarter, the front page has been kept current with news announcements and upcoming events, such as AAS, ADASS, NVO team meetings, and IVOA meetings. Specific changes include addition of the modify-file capability in publication upload, creation of a web page highlighting the NVO Core applications for AAS, and modification of the search function to search publications as well as news items and projects. The projects under the content-management can now have mirror sites.

A new machine has been ordered for the NVO web server. This is in part for speed, and also so that the old machine can act as a hot-spare backup in case of failure. The web server operating system has also been completely reinstalled as Red Hat Linux in place of Debian Linux.

10.3 Technical training initiatives

The Proceedings of the 2004 Summer School has been created, and is available on the NVO web site. A large number of CDs have been made with all the Summer School software for distribution at AAS and beyond.

A sign-up form has been added to the NVO web page so that astronomers can be on the mailing list when the 2005 Summer School is announced. Planning is underway for the 2005 Summer School, and a return to Aspen, Colorado is the most likely venue.

10.4 Advocacy

R. Hanisch gave a presentation on NVO to the Optical/IR Long Range Planning Committee at their January meeting in San Diego. A. Szalay is scheduled to discuss NVO with the AAAC in February.

The NVO team had a display booth at the fall ADASS Conference. Fifty copies of the Summer School software set were distributed and numerous demonstrations were shown.

T. Boroson gave a presentation at a Local Area Workshop on the Future of Ground-Based O/IR Astronomy, held at Yale University in October 2004. The talk, "Archiving, Data Mining, and the NVO Concept" provided an introduction for young researchers to the DataScope and Open SkyQuery tools.

At the request of the CfA director, C. Alcock, P. Fabbiano gave a presentation on the VO initiative to a congressional delegation visiting SAO.

11 Education and Public Outreach

11.1 Strategic partnerships

The NVO EPO program is crafted to optimize interaction with a variety of EPO professionals with established experience and programs reaching a wide variety of audiences. The program spans news and public information, formal and informal science education and undergraduate education. Outreach to the science and technical communities are covered in WBS 10. The program is coordinated by C. Christian.

The strategy we have taken is that the NVO project will coordinate the activities to insure that technical support for all projects is robust and that interactions between the EPO project personnel and the NVO project personnel facilitates deployment of appropriate resources. The second facet of the program is to identify inter-relationships between programs that capitalize on use of NVO data. Finally, identification of new partnerships is a strategy for growth.

11.2 Formal education

The first education project (Adopt a Galaxy) was produced and presented as a poster at ADASS in Pasadena (http://voservices.net/nvoedu) by J. Raddick, leveraging our experience with the Sloan Digital Sky Survey education program. Animated tutorials were prepared for each of the Applications to be shown at AAS including:

- Registry
- DataScope
- OpenSkyQuery
- Spectrum Services.

Further tutorials were made for Aladin and how to plot data in general especially using Excel. Further pre-college education projects are in development (student and teacher pages). We also purchased a domain name to create a portal into the NVO EPO area, http://www.virtualobservatory.org.

A trimmed down version of the Summer School Course was presented at STScI for 20 participants. This was given by Gretchen Greene and William O'Mullane and concentrated mainly on the web services aspects of the Summer School. The feedback was very positive.

Project LITE, under PI K. Brecher (Boston U.), continues to develop modules in the Spectrum Explorer for undergraduate student exercises. The resources are being developed to allow students to discover the temperature and compositions of stars through comparison with templates. Project LITE technical personnel attended the Summer School. The project is in a good position to being accessing NVO spectra and, when available, press release images for instructional purposes.

Interaction with the CLEA (PI L. Marshall, Gettysburg College) project has been re-initiated and technical details will be addressed to explore the feasibility of bringing NVO data into that environment

11.3 Informal education

Our partners at the Adler Planetarium in Chicago are using access to NVO data for interactive exhibits within the museum hall. Plans are underway to take advantage of NVO for SkyRider planetarium shows. M. SubbaRao at Adler attended the Summer School and has incorporated his new expertise into workshops and classes hosted by Adler.

The NVO group at Caltech has been working with the Griffith Observatory, a planetarium and education center that is a Los Angeles landmark. The Observatory is building a new facility that will include a sky mural, “The Big Picture,” that will be 150 feet wide by 20 feet high. The mural will be a picture of about 15 degrees of sky, including the Virgo cluster, and is being built with NVO-supported software and TeraGrid hardware, to be delivered in March 2005.

11.4 Outreach and press activities

The Office of Public Outreach (OPO) at STScI is the major party responsible for producing news releases related to breakthrough science conducted with the NVO and for technical advances worthy of news attention. In this period no releases were produced.

OPO is also creating a structure for ingesting press release images into the NVO. This work is done in collaboration with the University of California Berkeley under an NASA Applied Information Systems Research Program (AISRP) grant. Meta-tags will be defined and a schema for describing the images adopted. OPO will take advantage of the considerable work already accomplished in this area for the NASA Education Resource Directory. The other half of the AISRP grand involves creation of a public interface for the press images to be constructed by Berkeley.

All of the major players in the EPO programs are coordinating along with European colleagues at the European Space Agency to define a rational approach for offering the public access to a variety of materials of interest to the public such as the press release images, animations, simulations, graphics and etc.

11.5 Technical development

The NVO Registry accommodates metadata geared to EPO consumers, but until now the completeness and relevance of that metadata has not been tested. Now that partnerships are growing in strength and number, feedback will allow us to adapt the NVO infrastructure as needed to accommodate this new clientele.

Activities by Organization

Caltech–Astronomy Department and Center for Advanced Computational Research (CACR)

The Carnivore registry, developed at Caltech, has been released in open source form as a "Registry in a Box". Several institutions are actively engaged in deploying VO-standard registries using Carnivore, including CfA Harvard, European Southern Observatory, and CTIO/NOAO Tucson. The UK AstroGrid project is also using the Carnivore infrastructure for publishing to their registry. Carnivore is a fully featured open source VO registry service developed at Caltech. VO registry services are meant to facilitate publication and discovery of services. Carnivore supports publishing, searching and harvesting, and is fully compliant with the IVOA Registry Interface specification. Carnivore is primarily intended for use by data providers who want to set up their own registry. The installation at Caltech can be used by data and service providers who just have a few resources to publish.

Caltech staff have also set up a CVS-based software repository for testing and evaluation by the NVO team, and made available DPOSS and 2MASS data collections through SIAP and OpenSkyQuery interfaces. Prototypes for providing easy access to TeraGrid computational resources are being explored. CACR staff provided logistical support for the NVO demonstrations at the January AAS meeting, and contributed to the extensive documentation effort.

Caltech–Infrared Processing and Analysis Center (IPAC)

Staff at IPAC:

- Assumed responsibility for leadership of the Technical Working Group
- Deployed the Montage image mosaic service
- Continued to exploit NED as a test bed for NVO protocols and registries

Canadian Astronomy Data Centre/Canadian Virtual Observatory

No report received.

Carnegie-Mellon University/University of Pittsburgh (CMU/UPitt)

Over the last quarter the focus of the Pitt/CMU effort has been twofold: the development of web services that enable users to interface their data into the NVO; the implementation of grid services that are applicable to NVO research programs. Under the web services research and development an initial application, WESIX (Web Enabled Source Extraction and Cross-match; http://nvo.phyast.pitt.edu), has been developed by S. Krughoff, A. Connolly, and J. Colberg together with R. Williams (Caltech), A. Szalay, and W. O'Mullane (JHU), and R. Hanisch (STScI). WESIX provides a facility for users to upload images and detect sources within these images using SExtractor, a source extraction program from E. Bertin. The resulting catalog, returned as a VOTable, is cross-matched with the datasets available through openskyquery.net. A user selects both the catalog with which to cross-match and the variables within that catalog to return (e.g. fluxes, colors, positions). The identification of available catalogs and parameters is determined dynamically based on the resources available through openskyquery.net. Visualization tools, VOPlot and Aladin, are interfaced into WESIX. VOPlot provides a 2D plotting capability with which to compare output parameters and Aladin enables the output catalogs to be overlaid on the original image. Catalogs are returned for all sources that match between the two catalogs together with catalogs for those sources in either the image or database that do not match. WESIX is provided through a web service and web client interface and was demonstrated at the January AAS in San Diego.

A. Connolly and J. Gardner were CoIs with R, Williams (Caltech) on the renewal proposal for the NVO TeraGrid proposal. Pitt has undertaken an analysis of the Cosmic Microwave Background (R. Scranton and A. Connolly) and the 3-point function of galaxies (J. Gardner, C. McBride, and A. Connolly) using the initial years allocation of resources. These works are currently in preparation for publication.

Fermi National Accelerator Laboratory (FNAL)
S. Kent and N. Sharma worked with W. O'Mullane on testing and documenting the OpenSkyQuery web service portal. Small problems with web browser compatibility and XML code were identified and fixed. They also began exploring ways for end users to automate queries without requiring user intervention and to import returned data into existing data analysis applications. OpenSkyQuery provides both SOAP and HTTP forms interfaces, each with advantages and disadvantages. This work will continue to next quarter.

High Energy Astrophysics Science Archive Research Center (HEASARC)
HEASARC activities focused in two areas: the release of DataScope as a VO tool in the January 2005 AAS, and the upgrading of the metadata and VO services provided by the HEASARC. Many small changes were made to the DataScope although the basic architecture was not affected. Changes include adding a category of major services, supporting non-JavaScript enabled browsers, ensuring section 508 compliance in the web site, working with a number of data providers to ensure that the services were correct (e.g., fixing HTTP header errors and errors with unescaped ampersands in XML text), developing a test facility, installation scripts, and versions, addressing some deficiencies in OASIS processing of compressed FITS files and VOTables, as well as fixing many small (and occasionally large) bugs. The HEASARC worked with the NCSA to help them in their successful effort to bring up a mirror of the DataScope. The internal documentation for the DataScope was substantially upgraded.

The HEASARC updated all HEASARC metadata records to use the latest version of the XML schemas. The data was also updated to ensure accuracy of the textual descriptions. A number of records were deleted from the registry to get rid of redundant links, or to ensure that novice users would not be confused by links to relatively obscure data holdings. The HEASARC cone search services were moved from the *SkyView* servers to the main HEASARC machines. The *SkyView* SIAP services were completely reorganized with distinct SIAP services for each of the major *SkyView* surveys as well as an overall *SkyView* service that provides links to all of the surveys. This model matches other SIAP services and works more effectively in the registry.

HEASARC staff supported weekly and bi-weekly telecons and attended the fall team meeting by phone. HEASARC staff helped demonstrate the new VO application suite at the January AAS.

Johns Hopkins University
T. Budavari continued to work on the Simple Spectral Access Protocol (SSAP) and data model with the team. The Spectrum Services web site was tidied up for the AAS meeting. Budavari also assisted with the STScI web services course and attended the November NVO team meeting in Socorro.

W. O'Mullane has continued to work on the OpenSkyQuery portal with N. Li. This was demonstrated at ADASS meeting in Pasadena. O'Mullane spent a week at ESAC working on SSAP for SDSS and making sure it worked in VOSpec for the AVO demos in January, and on implementing Basic and Full SkyNodes in Java. VOPlot support was added to OpenSkyQuery. Work continued with G. Greene (STScI) on the searchable registry. Enhancements to the user interface were made, including grouping results by service types.

N. Li worked on OpenSkyQuery portal, most notably implementing an innovative dynamic query builder.

M. Nieto-Santisteban continued working on effective ways to utilize grid-based computing in large databases. She submitted a paper to the Conference in Innovative Data Research (CIDR05) that was accepted for publication. Nieto-Santisteban spent one week at ESAC working with the ESAC and Spanish VO teams to set up SkyNodes and SSAP services. She carefully tested several of the AAS applications (DataScope, OpenSkyQuery, Registry, and Spectrum Services). Bugs and suggestions for changes were reported. Maria presented several posters at ADASS and participated in the Sky Indexing BoF talking about how to index the sky by mapping the sphere into Zones. She worked with NOAO on ImgCutout and SkyNode issues for the NOAO Deep Wide-Field Survey and submitted a successful proposal to Microsoft Research to build a *Parallel Cross-Match Engine for Astronomy.*

A. Thakar continues to work on logging.

A. Szalay worked with NOAO on ImgCutout and SkyNode issues for NDWFS.

G. Fekete, A. Szalay, and J. Gray worked on the HTM and region support. Region intersect is now working.

J. Raddick has produced the virtualobservatory.org website. This included some excellent animated tutorials for the AAS applications

S. Carliles made minor modifications to the Mirage package, and started looking into making a FITS support package for C# which does not just wrap CFITSIO.

A meeting was held with R. Grossman's group from the Laboratory for Advanced Computing (University of Illinois at Chicago). We discussed building a cross-match for entire geographically distributed large catalogues. This is the type of query that OpenSkyQuery cannot perform at the moment. The LAC group has highly efficient methods for moving large datasets that may form the basis for such services. A. Thakar, M. Nieto-Santisteban, and A. Szalay continue working on this.

Microsoft Research
J. Gray worked with JHU on a spatial data library, the SDSS SkyServer architecture, data loading services, database design issues, and database performance issues. Also worked on GBps (yes Giga Byte Per Second) file transfers across Internet2 from disk to disk (this work is primarily with Caltech and CERN). Worked with JHU on the BCG cluster-finding algorithm and wrote several papers; one was accepted at the CIDR conference. Also did some early database sizing and performance work for LSST.

National Optical Astronomy Observatories (NOAO)
The NVO development staff at NOAO was augmented during this quarter by the addition of K. Holdaway, who is finishing a graduate degree at the University of Arizona, and by Post-Doc D. Norman, who is based at CTIO in Chile. M. Fitzpatrick and D. De Young attended the NVO Team Meeting in Socorro, NM.

NVO Project Scientist D. De Young continues to provide guidance to the NVO project in terms of setting project priorities from the perspective of scientific relevance to the astronomy and astrophysics community. With the critical goal of engagement of that community in mind, additional guidance is provided in terms of science demonstration projects and in terms of NVO functionality required to meet the desired NVO science capability. De Young also provided two use case scenarios for a possible Science Reference Mission for the European Astrophysical Virtual Observatory (AVO), and he continues to serve on the AVO Science Working Group and on the Executive Committee of the International VO Alliance (IVOA). In addition, De Young is the US liaison member of the Theoretical Astrophysics Interest Group of the IVOA, and in that context he has established a preliminary network of theory dataset links to be incorporated into an NVO and IVOA theory website.

D. Norman developed documentation for distribution to the community at the San Diego AAS meeting. This handout detailed the use of the NVO core applications, including DataScope, the Registry Portal, Open SkyQuery, Spectrum Services, and the SExtractor Web Service.

NOAO/NVO technical lead M. Fitzpatrick and student intern K. Holdaway successfully implemented a prototype framework for interfacing any IRAF task to a SOAP web service. New functionality can be deployed using IRAF scripts tasks or interfacing directly to existing IRAF tasks. Additionally, work was begun on a tool to automate the generation of the web service interface using an XML configuration file that describes the service and generates the Java implementation. The tool itself will be available as a web service for community use, allowing non-programmers to create web services more easily from IRAF tasks. The initial IRAF web services were demonstrated at the NVO team meeting in Socorro, NM, and at the ADASS meeting in Pasadena. The initial set of services is planned for release in the next quarter. NOAO proposed further development of IRAF web services, participation in the VOEvent specification, and the investigation of efficient VOTable data transport in its proposal to the LSST Corp. for LSST Data Management design & development work.

Software engineers P. Warner, I. Barg, and S. Yao have placed data release 3 from the NOAO Deep-Wide Survey (by project PIs B. Jannuzi and A. Dey) in the NOAO Science Archive (NSA) for public access. NDWFS is a multi-band survey of a 10 square-degree area of the sky, with an optical limiting magnitude fainter than 24. These data are accessible though NVO services such as SIA and cone search. As a part of the NDWFS release-3, NOAO has published and registered the source catalog, and built an Open SkyNode through which catalog access is provided. These efforts, funded by NOAO and with generous help from J. Gray (Microsoft) and A. Szalay and his JHU/NVO staff, place valuable resources on-line to support community research. In the process, NOAO staff members were able to contribute to the OpenSkyNode development effort by identifying and characterizing certain performance weaknesses in the software. It also revealed the importance of database tuning and performance monitoring in deploying a successful OpenSkyNode. S. Yao installed an open-source web statistics module called "AWStats" on the NOAO node to monitor the detailed web server workload to provide insight on the characteristics of web visits. It may be viewed at: *http://opensky-web.tuc.noao.edu/cgi-bin/awstats.pl?config=opensky-web.tuc.noao.edu*. Utilities such as this also have great potential for illustrating the actual use of the NVO and the resources to which it connects.

National Radio Astronomy Observatory (NRAO)

Highlights of NRAO participation in the NVO over the past quarter included the following:

- D. Tody chaired the DAL working group sessions at the fall IVOA workshop in Pune.
- NRAO sponsored the fall NVO team meeting held at the AOC in Socorro Nov 17-18.
- J. Ulvestad prepared a proposal for a VLA archive imaging pilot project that was submitted to the NSF (Dec 9). This would produce images for continuum data for the VLA B configuration at 5 and 8.4 GHz, using data from a single semester (late 1999 to early 2000, consisting of about 300 separate observing programs). The resultant images would be published to the VO. If successful, a more ambitious effort to data mine the VLA archive would be possible. At this point, this is just a proposal and additional resources are needed to actually proceed with the project.
- The SSA data model was advanced to V0.9 at the end of November. An updated draft of the SSA interface document is in preparation. The JHU spectral service can now return data conformant to the SSA V0.9 data model.
- In coordination with EU-OPTICON and ESO, a birds-of-a-feather (BOF) session was held at the fall ADASS conference (Oct 25) to discuss requirements and plans for

future astronomical software environments. This included side meetings with STScI, ESO (SAMPO), Euro-VO, NOAO, and others to discuss plans for a Python-based CLI, scalable framework, and for integration of general data processing and analysis software with VO.

- The first face-to-face meeting of the OPTICON working group on future astronomical software environments was held at ESO Dec 2-3. A white paper prepared by D. Tody for this meeting presented the scope of the project and proposed a scalable component-framework architecture for the software, including a strategy for how to integrate with VO/Grid.

Work continues to replicate the NRAO archive to NCSA. At the end of the quarter a broader effort was getting started to see what can be done in the short term to publish image data to the VO for radio data observations. Except for the large radio surveys, this is going to be a hard problem to address until the capability exists within NRAO to routinely pipeline process radio data observations.

Raytheon/ADC (University of Maryland and George Mason University)
In a poster at the ADASS meeting, E. Shaya presented first drafts for general astronomical schema and OWL (Web Ontology Language). These will eventually be used to support a high level VOQuery in a distributed astronomical data system.

E. Shaya continues to attend the weekly Technical Working Group telecons and help to inform the NVO about the latest developments in XML applications. E. Shaya contributed discussions on upper and lower limits for the spectral and SED data models.

Brian Thomas worked on the Catalog data model.

The Maryland group continues to host an NVO CVS repository.

George Mason University (GMU) staff attended the ADASS XIV conference in Pasadena, California, and supported the NVO display booth, providing several demonstrations of NVO science applications. GMU staff continued to seek additional funding to support NVO-enabled distributed data mining activities. GMU staff held discussions with the LSST Project Data Management team in order to identify potential NVO-LSST synergies in the areas of distributed data mining, science use case development, scientific database end-user requirements, and education/public outreach.

San Diego Supercomputer Center
SDSC is establishing multiple collaborations with the LSST project, both through the NVO and through collaborations with the Lawrence Livermore National Laboratory. The collaborations focus on the management of massive amounts of data and support for extremely large databases. Multiple meetings have been held with the LSST team to identify relevant data management technologies.

Management activities include participation in Global Grid Forum working groups on Preservation, AstroGrid, Grid File System, and Grid Reliability and Robustness (R. Moore). R. Moore also participates on the IVOA and NVO executive committees.

Smithsonian Astrophysical Observatory
A. Rots and J. McDowell attended the team meeting in Socorro, NM, 17-18 November 2004.

Rots and P. Fabbiano attended the ADEC meeting in Pasadena, where discussions of NASA archive and ADS interoperability were held. SAO is a strong proponent and pathfinder for this effort.

C. Alcock and P. Protopapas are now at CfA and we will be coordinating efforts. P. Fabbiano met with A. Goodman, M. Elvis, and C. Alcock to address how best could the Harvard Initiative for Innovative Computing help VO initiatives. Meetings with people from the Harvard Medical School were also held, to discuss approaches to visualization.

Fabbiano joined the SOC of the proposed Prague VO Symposium, and provided feedback on the proposal to SOC chair A. Lawrence (AstroGrid).

K. McCusker joined the SAO NVO team on Nov 1 as a programmer, replacing M. Harris.

SAO submitted a revision to the NVO contract, with a revised statement of work, addressing additional funding needed to complete the Data Model work. The effort of the Alcock-Protopapas team was also included in this proposal.

Space Telescope Science Institute
NVO work at STScI focused on the NVO Registry. The STScI implementation was brought into compliance with the VOResource schema V0.10, and harvesting interfaces to NCSA, Caltech, CDS, VO Japan, and AstroGrid were validated. A “Try It” feature was added to the Registry interface, allowing users to directly invoke a Cone Search or SIAP request. We worked with JHU to set up a mirror site for the Registry on a JHU server. (G. Greene working with W. O’Mullane, JHU.)

A GSC2 OpenSkyNode was built for the All-Sky GSC 2.3 catalog, and a mirror site was set up at STScI for the primary OpenSkyQuery portal at JHU (G. Greene, B. McLean).

R. Hanisch was active in testing the NVO applications and in providing documentation. Hanisch worked with D. Norman (NOAO) and R. Williams and S. Emery Bunn (CACR) in preparing handouts for the January AAS Meeting.

R. Hanisch coordinated the agenda for the fall NVO team meeting, IVOA Executive Committee meeting, and NVO External Advisory Committee meeting.

C. Christian coordinated NVO EPO activities and presented an overview of the NVO EPO program to the Advisory Committee.

United States Naval Observatory

S. Levine worked on constructing a VO conesearch compliant server for several of the large USNO constructed catalogues (in particular USNO-A2, USNO-B1 and NOMAD). The conesearch server has been installed and tested, and will be made available to the community in early 2005. Work was begun on making a VO compliant SIA server for the USNO image archive.

Levine attended the NVO quarterly meeting at NRAO in Socorro, NM.

University of Illinois-Urbana/Champaign/National Center for Supercomputer Applications (UIUC/NCSA)

R. Plante, as lead for registries, continued to coordinate the upgrade of NVO registries to VOResource V0.10 and moderated discussion of future changes to the XML schemata. He continues to contribute to the refinement of the IVOA Registry Interface in collaboration with the IVOA Registry Working Group. He has provided consulting for Arnold Rots (CfA) in the development of the Space-Time Coordinates schema. He continues to develop validation services for VO standard services; greater focus will be put on validation as the project takes on registry curation issues in the coming year. Plante has also been working on an intermediate level registry search interface (see WBS 4.5). In October, he presented an invited talk on VO registries at the Astronomical Data Analysis Software and Systems conference in Pasadena, CA.

R. Williamson released an updated version of VORegistry-in-a-Box to incorporate the latest changes to VOResource V0.10. He continues to support the Java SkyNode server toolkit; in this quarter, he has been working with W. O'Mullane (JHU) and C. Stoughton (Fermilab) to integrate support for the full SkyNode. (This collaboration has its genesis in the NVO Summer School where Stoughton, as a student, started the coding as a personal project.) Williamson and Plante installed a mirror of the DataScope application at NCSA. This collaboration with T. McGlynn (HEASARC) resulted in a number of bug fixes and packaging improvements to this application.

In this quarter, Plante stepped down as moderator of the technical working group, passing this role to John Good (NASA IPAC). Originally referred to as the Metadata Working Group, the weekly telecon evolved over the life of the NVO project into a general-purpose working group addressing a wide range of technical issues. This quarter, we officially changed the name (and mailing list) to reflect the broader scope.

University of Pennsylvania

U. Penn participants C. Alcock and P. Protopapas have relocated to SAO, and their activities are described in the SAO report.

University of Southern California (USC/ISI)

This report highlights the work done in the last 3 months within the NVO project. Although most of the activities were related to the Montage processing on the Grid, the techniques developed in this work are applicable to many grid applications. ISI focused on three major changes areas:

1. The mAdd was tiled—to provide better application scalability.
2. A mGridExec web service was created to support generic application mapping onto the grid resources with Pegasus.
3. The Montage workflow was modified to add nodes for visualization, to enable a demonstration of the existing capabilities.

The final step in the Montage workflow is the co-addition of re-projected, background-corrected images into a final mosaic. This works fine for smaller montage workflows (less than 1 degree x 1 degree) but for larger workflows, the final mosaic becomes too large. It becomes difficult to transfer larger mosaics and also creates difficulties while visualizing them. In order to overcome these difficulties, the final step was distributed into multiple mAdds depending on the size of the desired mosaic. Thus a 6 degree x 6 degree Montage workflow now contains 36 mAdds. Each of these mAdds creates a smaller mosaic. All of these smaller mosaics can be combined into a bigger mosaic if required. It is now possible for the user to select which portion of the mosaic he/she wants to download instead of downloading the whole mosaic. All these mAdds can be executed in parallel, thus decreasing the workflow completion time.

ISI created a web service interface for submitting Montage workflows for execution over the Grid resources. This web service interface can be invoked using a command line client. This client, called mGridExec, takes as argument the name of a zip file. The zip file contains all the information and the abstract workflow required for processing the request. This web interface decouples the development of the front-end, where the request is submitted, from the processing on the Grid. The client returns a job identifier that can be used to track the status of the request. Upon completion, an email is send to the user indicating the result of the processing. This web service also creates a HTML page containing links to all the data products created by the workflow. The URL of this HTML page is sent to the user in the notification email in case the request completes successfully.

Nodes were added to the Montage workflow for creating a JPEG image of the final mosaic. As indicated earlier, multiple mosaics are created based on the size of the input request. Then these mosaics are shrunk using mShrink and co-added with mAdd. Then mJPEG is used to create an image of the final mosaic. The image is accessible on the URL send to the user in the notification email. Currently each request is specific to a particular band only (J, H or K).

University of Wisconsin
No activity reported this quarter.

Publications

"NVO LITE—Harnessing the VO for Education." K. Brecher, E. Weeks, & P. Carr, 2004. AAS 205, #130.10.

"Environments of Starburst Galaxies Diagnosed with the NVO." D. de Mello & M. Sosey, 2004. AAS 205, #130.07.

"The Simple Spectral Access Protocol." M. Dolensky & D. Tody, 2004. SPIE 5493, 262.

"Virtual Observatory Science Applications." T. McGlynn, 2004. AAS 205, #130.02.

"Automated Classification of ROSAT Sources Using Heterogeneous Multiwavelength Source Catalogs." T. A. McGlynn, et al., 2004. ApJ 616, 1284.

"A Dust Extinction Web Service in the VO Framework." C. Miller, K. S. Krughoff, T. K. Ho, 2004. AAS 205, #130.05.

"When Database Systems Meet The Grid," <http://research.microsoft.com/research/pubs/view.aspx?type=Technical%20Report&id=786> María A. Nieto-Santisteban; Alexander S. Szalay; Aniruddha R. Thakar; William J. O'Mullane; Jim Gray; James Annis, revised and accepted for publication by ACM CIDR, Jan 5, 2005, Pacific Grove, CA.

"Environments of Starburst Galaxies Diagnosed with the NVO." M. A. Nieto-Santisteban, M. Sosey, & D. de Mello, 2004. AAS 205, #113.02.

"New NED XML /VOTable Interfaces and Client Applications". O. Pevunova, B. Madore, J. Mazzarella, J. Good, G. B. Berriman, 2004. Submitted to Proceedings of ADASS XIV.

"The International Virtual Observatory Alliance: Recent technical developments and the road ahead." P. J. Quinn, et al., 2004. SPIE 5493, 137.

"Education and Outreach with the National Virtual Observatory." J. Raddick, 2004. AAS 205, #95.02.

"COMPLETE as a Virtual Observatory Testbed." N. A. Ridge & A. A. Goodman, 2004. AAS 205, #130.09.

"Space-Time Coordinate Metadata for the Virtual Observatory." A. Rots, 2004. AAS 205, #113.07.

"The NVO Comes of Age." A. S. Szalay, R. Cutri, D. De Young, R. Hanisch, R. Moore, E. Schreier, & R. Williams, 2004. AAS 205, #130.01.

"LSST and Astronomy in 2020." A. S. Szalay, et al., 2004. AAS 205, #108.13.

"Brave New World: Data Intensive Science with SDSS and the VO." A. R. Thakar, et al., 2004. AAS 205, #113.01.

Virtual Observatory Articles in the Popular and Technical Press

Nothing this quarter.

Acronyms

AAS	American Astronomical Society
ADC	Astronomical Data Center
ADEC	Astrophysics Data Centers Executive Committee (NASA)
ADQL	Astronomical Data Query Language
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
DAG	Directed Acyclic Graph
DAGMan	Directed Acyclic Graph Manager (Condor)
DAML	DARPA Agent Markup Language
DARPA	Defense Advanced Research Projects Agency
DIS	Data Inventory Service
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
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
OAI	Open Archives Initiative
OASIS	On-line Archive Science Information Services (IRSA)
OGSA	Open Grid Services Architecture
OIL	Ontology Inference Layer
OWL	Web Ontology Language
PB	petabyte
PMH	Protocol for Metadata Harvesting (of OAI)
Q	quarter
QSO	Quasi-Stellar Object
RC	Replica Catalog
RDF	Resource Description Framework
REST	Representational State Transfer
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
SIAP	Simple Image Access Protocol
SOAP	Simple Object Access Protocol
SRB	Storage Resource Broker
SSAP	Simple Spectral Access Protocol
STScI	Space Telescope Science Institute
SWG	Science Working Group
TB	terabyte
UCD	Unified 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 Language
VDS	Virtual Data System
VO	Virtual Observatory
VO	Virtual Organization
VOQL	Virtual Observatory Query Language
WBS	Work Breakdown Structure
WSDL	Web Services Description Language
XML	Extensible Mark-up Language
2MASS	Two-Micron All Sky Survey