



# US National Virtual Observatory

[Home](#) | [Get Started with NVO](#) | [About Registries](#) | [About Publishing](#) | [Registry@STScI](#) | [Carnivore@Caltech](#)

[Registration Portal](#) | [Documents](#) | [Contact Us](#)

## How to Publish to the NVO

**Version 2.0 (3 Jan 2008)**

**Editor: [Raymond Plante](#) (NCSA)**

The [National Virtual Observatory \(NVO\)](#)<sup>1</sup> is a powerful environment for locating and integrating a wide variety of data originating from many different instruments and exploring many different research questions in astronomy. But how does data get into that environment in the first place? Data is exposed to the NVO environment through a process called *publishing*.

This "how-to" document is intended for anyone who has data and would like to share it with the astronomy community through the NVO. Remember, though: data is not the only thing you can publish—you can also publish *services*. That is, if you have a piece of software that might be useful to others and would like to make it accessible over the network, publishing it as a service makes it possible for other NVO applications to make use of it.

There are two things to keep in mind as you join the NVO community as a data or service provider. First, we've tried to make the publishing process an incremental one. You can decide the level of exposure that you want for your data and the amount of effort you want to put in. You can gradually build more visibility as you have time and expertise available. Second, the NVO is an evolving environment. As it matures, standards will improve, and there will be a new ways to expose your data.

## Contents

- [1. Introduction](#)
  - [1.1. An Overview of Data Discovery in NVO](#)
  - [1.2. Who are you?](#)
    - ["I'm an individual with a small data collection."](#)
    - ["I run a web-based archive of data."](#)
    - ["I have a cool service."](#)
- [2. Contributing Data to VO-ready Repositories](#)

- [3. Federating an Archive with the NVO](#)
- [4. Deploying a service to the NVO](#)

## 1. Introduction

### 1.1. An Overview of Data Discovery in NVO

In general, data or services are considered **published** if one can use NVO facilities to find them. The first step in data discovery is the NVO Registry system. A **registry** is a web accessible database that contains descriptions of data and services. A registry can contain descriptions of other things, too--namely, organizations or software. We refer to all things that can be described in a registry as **resources**.

You can [search the main NVO registry](#)<sup>2</sup> interactively with a web browser. However, the registry is more often used by higher-level applications like the [DataScope](#)<sup>3</sup> that help users find things in the NVO.

NVO registries tend to have descriptions of "coarse-grained" things, like data collections. Locating more "fine-grained" things, like images and spectra, is done through specialized services like the [Simple Image Access Protocol \(SIAP\)](#)<sup>4</sup>. Normally, users don't usually access such services directly; instead, they use higher-level services. For example, the [DataScope](#)<sup>3</sup> will use SIAP services behind the scenes to find images for the user.

This document describes how the publishing process exposes your data through these various services so that users can find your data through applications like the [DataScope](#)<sup>3</sup>.

### 1.2. Who are you?

How you publish your data generally depends on who you are and what you want to publish. If you fall into one of the following categories, you just skip to the appropriate part of this document to learn the process.

#### **"I'm an individual with a small data collection."**

You might have a modest number of datasets--images, spectra, or catalogs--that you wish to share, but you don't have permanent place to store them. In this case, skip to [section 2](#) to find how you can take advantage of repositories that accept data deposits.

#### **"I run a web-based archive of data."**

You might already maintain an archive, perhaps associated with a telescope or a large research project, that serves data to a community. You might also support a number of specialized services that operate on that data. If so, skip to [section 3](#) to learn how to expose your data and services to the NVO.

**"I have a cool service."**

You might have a piece of useful software you would like to wrap up and deploy on your web server. Or, you may already maintain services on a web site. If so, skip to [section 4](#) to learn how to let the NVO know that they exist.

## 2. Contributing Data to VO-Ready Repositories

If you have a collection of data you would like to share with the community, you might consider depositing it to an NVO-federated repository. You may not have access to a web site where you can host the data yourself, or if you do, you may not be able to guarantee that the data will be supported for the long-term. (If you can take on that responsibility, see [section 3](#).)

Open repositories are more than just a web site that will host your data. They take responsibility for long-term curation of the data. They often provide a number of value-added services that can operate on your data. The most important advantage of an NVO-federated repository is that it already supports many of the NVO publishing standards. By simply depositing your data into the repository, it is automatically exposed to the NVO.

In the following subsections, we give an overview of the VO-ready repositories that we are aware of at this time. We expect more such repositories to emerge in the coming years; check the [NVO web site](#)<sup>1</sup> for the latest pointers.

### 2.1. Image Repositories

The [NCSA Astronomy Digital Image Library](#)<sup>5</sup> (ADIL) allows astronomers to upload their research-quality FITS images, making them available to the astronomy community and the general public. Each deposit is a collection of one or more images all related to a single scientific study. The collection can also include other data files, such as spectra, visibility data, or figures; however, there must be at least one image. Finally, the ADIL requires the images be associated with a published paper describing the images.

The ADIL provides searching and browsing services for images in the library. Furthermore, all the data is automatically cross-linked with the [NASA Astronomical Data System \(ADS\) Abstract Service](#)<sup>6</sup>: in particular, ADS users that find the published paper through the abstract service will see a link to the data in the ADIL. The ADIL is not only [registered](#) with the NVO registry, it implements several of the NVO standard services.

### 2.2. Spectral Repositories

You can publish spectra via the [Spectral Services for the VO](#)<sup>7</sup> site. This repository allows users to search for and plot spectra in a variety of ways. This repository also contains an archive of filter passbands that can be used for analyzing spectra taken from different telescopes.

### 2.3. Catalog Repositories

As of this writing, there are no known catalog repositories that allow users to simply upload tabular data and expose them to the NVO, although this is expected to change soon. However, the [NASA Extragalactic Database \(NED\)](#)<sup>8</sup> will often work with scientists on a one-on-one basis to host catalogs via the specialized NED catalog services.

Note that if you have the facilities to host your catalogue yourself, consult [Section 3](#) about how you integrate it with the VO.

## 3. Federating an Archive with the NVO

If you already maintain an archive, you can increase its visibility by federating it with NVO. Publishing starts, then, when you register your data and services with the NVO registry system. How much you register is up to you. Afterwards, you may wish to implement one or more standard NVO services. The more information you provide to the NVO and the more standards you support, the easier it will be for others to make use of your data.

### 3.1. Basic Registration

The simplest thing you can do to publish your data to the NVO is to let the NVO know that your archive exists by [registering](#) with the NVO registry system. This is done by visiting the [NVO Registration Portal](#)<sup>9</sup>.

When you visit the portal for the first time, you will first have to register your *organization* as a publisher of VO [resources](#). Afterward, you can register any number of additional resources, each of which will refer back to your organization as the publisher. You should register your archive next as a *data collection*.

**Note:**

If you have a large number of [resources](#) that you wish to register (e.g. more than twenty resources), you may prefer to run your own publishing registry. There are two packages that allow you to deploy your own registry as part of the NVO registry system: [VORegistry-in-a-Box](#)<sup>10</sup> and [Carnivore](#)<sup>11</sup>. If you manage a very large number of collections (e.g. hundreds), you may wish to implement your own publishing registry that connects directly into your metadata management system. For more information on how to do this, consult the NVO registry team via [usvo-feedback@us-vo.org](mailto:usvo-feedback@us-vo.org).

Registering a [resource](#) via the Registration Portal is done by filling out a form where you describe your organization, data collection, or service. Some of the information you provide includes:

- a title for the resource
- a short name for the resource
- a [global identifier](#)
- a short description of the resource
- information about the people behind the resource

- various kinds of keywords
- information on how to access the resource

With some resources, such as standard services, you will be asked for some specialized information that describes how the service behaves. The registration portal will guide you through the registration form. When you register your archive as a *data collection*, one of the most important bits of information will be the URL for accessing the archive.

#### **A word about Identifiers:**

As part of the registration process, you may be asked to make up a globally unique identifier for your [resource](#). These identifiers follow a standard format that looks something like this: `ivo://adil.ncsa/targeted/SIA`. Every identifier has two major parts.

The first is the **authority ID**, e.g. `adil.ncsa`. This names a potential set or *space* of identifiers that is under the control of a single publisher. As a publisher, you get to choose your own authority IDs as long as they are not already "owned" by another publisher. NVO uses the convention that is hierarchical based like a DNS name; however, we don't include fields like "edu", "org", and "www". The publishing registry should help you with choosing an authority ID.

The second part of the ID is called the **resource key**, e.g. `targeted/SIA`. This, too, can be anything you want as long as you have not already used it with any of your other resources. Since you are the only one allowed to create identifiers with your authority IDs, the two parts together constitute a unique, global identifier.

The registration portal form will assist you with choosing an appropriate identifier for your resource.

Included in the information you provide is a short description of your archive and its home page. Thus, this minimal registration lets VO users find your home page via a search of the [registry](#).

## **3.2. Registering Your Existing Services**

If you want VO users to find more than just your home page, you should consider registering your existing services. These services can come in three forms:

- Web browser-based services  
These are Web pages containing a form interface to an interactive service. These might include your archive's search page.
- HTTP-CGI services  
These are HTTP-based services with a well-defined interface (using GET or POST) which you allow access to outside of a Web browser. An example might be a name resolving service.
- SOAP-based Web Services  
These are your own custom services intended for non-interactive access.

You can register these kinds of services using the same [registration portal](#) you used to [register your archive](#).

### 3.3. Deploying Standard Services

The real power of the NVO comes in the form of integrating services like the [DataScope](#)<sup>3</sup> and the [OpenSkyQuery](#)<sup>12</sup> Portal. These services are able to bring data together from multiple archives by taking advantage of *standard services*. Currently, there are four types of standard services that you may want to expose your data through:

- [Simple Cone Search Protocol \(SCS\)](#): an interface for searching catalogs by sky position.
- [Simple Image Access Protocol \(SIA\)](#): an interface for searching and retrieving images.
- [Simple Spectral Access Protocol \(SSA\)](#): an interface for searching and retrieving spectra.
- [OpenSkyNode](#): a general purpose database search interface.

These are services that make up what we call the VO's *data access layer*: they run on your web site and provide access to your data. Implementing these requires some programming skills and familiarity with XML; however, there exist a number software packages that can make deployment easier. The above list is ordered by complexity, with the Cone Search Interface being the simplest to create. The first three "simple" protocols use HTTP's traditional CGI model for a service, so it is likely not much more complex than the typical services you already run for your repository. OpenSkyNode is a SOAP Web Service and designed to allow you to make use of any of the common "off-the-shelf" Web Service toolkits available. (For a nice overview of all of the data access layer services, consult the presentation on [DAL Servers](#)<sup>16</sup>.)

To get started, consult the section below for the service you want to implement. There you will find links to the specification and useful libraries. It's worth reviewing the specification document to help ensure your service is compliant with the standard. For some of the standards, there exist verifiers that will help you test your service once deployed on your web site.

#### A word about VOTable, a standard table interchange format:

All four of the services listed here make use of [VOTable](#)<sup>13</sup>, an XML standard for exchanging tabular data. A number of [libraries for manipulating VOTables](#)<sup>14</sup> exist in several programming languages. More importantly, a number [visualization tools](#)<sup>15</sup> support the VOTable format as well. Thus, you may find it useful to incorporate VOTable as an export format for

**Helpful Software from the NVO Summer School:**

A software bundle was developed for the [NVO Summer School](#)<sup>17</sup> which includes sample implementations of all three of the standard services discussed in this document. It is available from the [Summer School Software Page](#)<sup>18</sup>. You should also consult the summer school course presentations available from the [proceedings page](#)<sup>19</sup> for "how-to" tutorials. You can also get the software and a wealth of supporting tutorial information from the forth-coming [ASP](#) publication, *The National Virtual Observatory: Tools and Techniques for Astronomical Research*<sup>20</sup>.

your own custom services.

Once you have deployed one of these standard services, you are almost done. You have one more thing to do--that's right: [register the service](#) through the registration portal (see [section 3.1](#)).

### 3.3.1. Simple Cone Search (SCS)

This service is intended for searching catalogs by sky position. The interface is very simple: your service will get a query encoded in the URL containing a right ascension, a declination, and a search radius--all in degrees (J2000). Your service returns a table containing all the rows that fall within that circle or *cone*.

Simple Cone Search is typically implemented for two types of catalogs:

- *Object catalogs*, where each row describes a different astronomical object.
- *Observation catalogs*, where each row describes an observation by a telescope toward some direction.

The standard is not necessarily restricted to these two types. An implementation must merely be able to respond sensibly to a position query with a list of matching items.

To get started, consult the following links:

**Specification:**

<http://www.ivoa.net/Documents/latest/ConeSearch.html>

**Verifier Service:**

<http://nvo.ncsa.uiuc.edu/VO/services/csvalidate.html>

**Software Libraries:**

<http://us-vo.org/summer-school/2006/proceedings/software/NVOSS2006-Software.html>

**How-tos/FAQ:**

<http://us-vo.org/summer-school/2006/proceedings/presentations/conesearchservice.html>

### 3.3.2. Simple Image Access Protocol (SIA)

The Simple Image Access Protocol (referred to in short as either SIAP or just SIA) provides a way to search for and retrieve images. This interface can be used to access either static images or "cutout" images that are generated on-the-fly. Use of an SIA service has two distinct steps. First, the user searches for images of interest by providing a sky position and an image size. The service responds with a table where rows describe the images that match the query. (For many cutout services, the table will contain only one row describing a custom-made image.) Each row includes a URL for the image, so that in the second step, the user can download the image.

Like the Cone Search interface, SIA uses HTTP GET queries to encode query arguments in the request URL. However, the SIA interface is much more flexible than the Cone Search interface. It specifies a number of optional parameters that can control what images are returned, how cutouts are to be made, and what formats to provide. It also allows you, as the data provider, to support non-standard arguments. The registry plays an important role in the use of SIA implementations: registry descriptions of SIA services indicate not only what type of SIA service it is, but also what arguments it supports.

To get started on implementing an SIA service, consult the following links:

**Specification:**

<http://www.ivoa.net/Documents/latest/SIA.html>

**Verifier Service:**

<http://nvo.ncsa.uiuc.edu/VO/services/siavalidate.html>

**Software Libraries:**

<http://us-vo.org/summer-school/2006/proceedings/software/NVOSS2006-Software.html>

**How-tos/FAQ:**

SIA Service Tutorial:

[http://us-vo.org/summer-school/2005/proceedings/presentations/server\\_side.html](http://us-vo.org/summer-school/2005/proceedings/presentations/server_side.html)

For info on connecting the server to a database<sup>†</sup>, see also:

<http://us-vo.org/summer-school/2006/proceedings/presentations/conesearchservice.html>

<sup>†</sup>The 2006 Summer School does not have a tutorial specifically on creating an SIA service; however, the `siapserver` toolkit



is very similar to the `coneserver` toolkit; thus, you will also find the ConeSearch service tutorial from 2006 helpful as well. In particular, consult this tutorial to learn about connecting your service to a database.

### 3.3.3. Simple Spectral Access Protocol (SSA)

The Simple Spectral Access Protocol does for spectra what [SIA](#) does for images: it provides a way to search and retrieve spectra through a simple URL-based interface. Also like SIA, the spectra it returns can either be static or calculated on-the-fly according to the user's specifications. The use of an SSA service has the familiar two-step process of the SIA: a search for available spectra for a given region of the sky followed by a request for individual matched spectra. There are some differences in the interface with SIA, particularly in the choice of standard formats that the service can return spectra in.

SSA is the newest of the so-called "simple" DAL protocols, and as of this writing, helpful tools and documentation for creating an SSA service are limited; however, these will begin appearing soon. Until then, you can consider consulting the specification and adapting the tools available for [deploying an SIA service](#).

<b>Specification:</b> <a href="http://www.ivoa.net/Documents/latest/SSA.html">http://www.ivoa.net/Documents/latest/SSA.html</a>
<b>Verifier Service:</b> <i>n/a</i>
<b>Software Libraries:</b> <i>see previous section on <a href="#">SIA</a>.</i>
<b>How-tos/FAQ:</b> Data Access Layer Overview (with discussion of SSA): <a href="http://us-vo.org/summer-school/2006/proceedings/presentations/dalserver.ppt">http://us-vo.org/summer-school/2006/proceedings/presentations/dalserver.ppt</a>

### 3.3.4. OpenSkyNode

The OpenSkyNode standard provides web-access to tabular data that is a bit more sophisticated than the Cone Search interface. When you implement the OpenSkyNode interface, applications can connect directly to your site and execute complex SQL-based searches of any of the tables you wish to expose. However, an important capability of the OpenSkyNode interface is in its ability to support efficient joint queries--in particular, object cross-matching--across multiple OpenSkyNode sites.

OpenSkyNode is defined as a SOAP-based service using a WSDL (Web Service Description Language) document. Two levels of compliance are defined: Basic and Full. The Basic SkyNode allows clients to send a query using the Astronomical Data Query Language (ADQL), a form of SQL that is specialized for the VO, and receive back a table of matching records as

a VOTable. The Basic SkyNode interface also provides a few methods for discovering information about the service, like what can be queried. The Full SkyNode interface adds the support for participating in cross site joins.

As with all standard services, you are free to implement the interface using whatever language and backend database that you wish; however, toolkits are available for some commonly used technologies. In particular, a plug-in implementation using SQLServer and .Net is available for Windows platforms. There is also a toolkit based on Java and Apache Axis (a Web Service toolkit); this toolkit supports several common databases, including MySQL, PostgreSQL, and Sybase. It is easily extended to support other databases as well as archive-specific customizations.

To get started on implementing an OpenSkyNode, consult the following links:

**Specification:**

*Working Draft:*

<http://www.ivoa.net/internal/IVOA/IvoaVOQL/SkyNodeInterface-0.7.4.pdf>

**Verifier Service:**

*n/a*

**Software Libraries:**

Summer School Package:

<http://us-vo.org/summer-school/2006/proceedings/software/>

**How-tos/FAQ:**

Building a SkyNode Server:

<http://us-vo.org/summer-school/2006/proceedings/presentations/Skynode2006.html>

Server Software Tutorial:

[http://us-vo.org/summer-school/2006/proceedings/presentations/Skynode\\_Exercise.htm](http://us-vo.org/summer-school/2006/proceedings/presentations/Skynode_Exercise.htm)

**The Future of Database Access in VO:**

As of this writing, the functionality of the SkyNode is being redesigned for a new service standard called *Table Access Protocol*. Deprecating the SkyNode working draft, this new standard will make it easier to both implement and use database services, while still supporting the ability to cross-correlate databases through the VO.

## 4. Deploying a service to the NVO

A service does not need to be a VO standard in order to be accessible from the VO. Some services are very specialized--say,

a service that does coordinate transformations--and do not benefit from having a large number of implementations available to users; nevertheless, they can be valuable tools for connecting data together. As you might guess, publishing begins with registering the service; however, there is more that you can do that will make your service more useful. The following are three things you could do, each taking a little more effort:

### 1. Register your service.

Simply follow the same instructions for registering as described in [section 3](#). Currently, there are three types of non-standard services you can register. When you visit the registry, you will have to choose which type of service it is.

- **Browser-based service:** an interactive web page, usually featuring a form interface.
- **CGI Get Service:** This service type is similar to a [Cone Search](#) or [SIA](#) in that parameters are provided as keyword-value pairs appended to a base URL.
- **SOAP-based Web Service:** You typically provide the endpoint URL for the Web Service when you register; the registry will assume that a WSDL can be retrieved by appending "?WSDL" to it.

### 2. Integrate support for standard formats and schemas.

Perhaps the most important standard formats you might incorporate are FITS and [VOTable](#). When you use a standard format, you make it easier to connect your service with existing data viewers. Making use of standard XML schemas, on either input or output, will make it easier to connect your service into a chain of services from the VO grid.

### 3. Implement the IVOA Standard Support Interface.

*Actually, you can't do this yet, because the standard is still in development.* The IVOA Standard Support Interface is intended to make it easier for higher level applications to interact with your service in a robust way. In particular, it has three basic functions:

- it allows a service to be more self-describing
- it provides the VO infrastructure with information on its health and availability.
- it provides information on its usage history.

## References

1. National Virtual Observatory Home Page: <http://www.us-vo.org>
2. The Searchable Registry at STScI: <http://nvo.stsci.edu/voregistry>
3. DataScope: <http://heasarc.gsfc.nasa.gov/vo/>
4. Simple Image Access Protocol: <http://www.ivoa.net/Documents/latest/SIA.html>
5. The NCSA Astronomy Digital Image Library: <http://adil.ncsa.uiuc.edu/>

6. The NASA Astronomical Data System Abstract Service: [http://adsabs.harvard.edu/abstract\\_service.html](http://adsabs.harvard.edu/abstract_service.html)
7. Spectral Services for the VO: <http://voservices.net/spectrum/>
8. NASA Extragalactic Database: <http://ned.ipac.caltech.edu/>
9. NVO Registration Portal: <http://nvo.ncsa.uiuc.edu/nvoregistration.html/>
10. VORegistry-in-a-Box: <http://trac.us-vo.org/nvo/wiki/VORegistryInABox>
11. Carnivore XML Registry: <http://nvo.caltech.edu:8080/carnivore/>
12. The OpenSkyQuery Portal: <http://openskyquery.net/Sky/skysite/>
13. About VOTable: <http://www.us-vo.org/VOTable/>
14. VOTable software: <http://www.ivoa.net/twiki/bin/view/IVOA/VOTableSoftware>
15. VOTable visualizers: <http://www.us-vo.org/projects/tools.cfm>
16. Tody, D. 2006, *Data Access Layer Servers*, NVO Summer School Presentation (MS Powerpoint format): <http://us-vo.org/summer-school/2006/proceedings/dalserver.ppt>
17. NVO Summer School: <http://us-vo.org/summer-school/>
18. NVO Summer School 2006 Software Distribution: <http://us-vo.org/summer-school/2006/proceedings/software/NVOSS2006-Software.html>
19. NVO Summer School 2006 Proceedings: <http://us-vo.org/summer-school/2006/proceedings/>
20. Graham, M., Fitzpatrick, M., and McGlynn, T. 2008, *The National Virtual Observatory: Tools and Techniques for Astronomical Research*, ASP Conference Series, Vol. 382 (ASP: San Francisco), *in press*.

---

Ray Plante

Last modified: Fri Jan 4 09:45:19 2008