

*International  
Virtual  
Observatory  
Alliance*



## Resource Metadata for the Virtual Observatory Version 1.01

IVOA Recommendation 2004 April 26

**This version:**

<http://www.ivoa.net/Documents/REC/ResMetadata/RM-20040426.html>

**Latest version:**

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

**Previous version(s):**

<http://www.ivoa.net/Documents/PR/ResMetadata/RM-20040323.html>  
<http://www.ivoa.net/Documents/PR/ResMetadata/RM-20040126.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RM-20031002.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RM-20030801.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RM-20030709.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RSM-20030509.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RSM-20030206.html>  
<http://www.ivoa.net/Documents/WD/ResMetadata/RSM-20021011.html>

**Editor(s):**

[Robert Hanisch](#)

**Author(s):**

IVOA Interoperability Working Group  
NVO Metadata Working Group

## Abstract

An essential capability of the Virtual Observatory is a means for describing what data and computational facilities are available where, and once identified, how to use them. The data themselves have associated metadata (e. g., FITS keywords), and similarly we require metadata about data collections and data services so that VO users can easily find information of interest. Furthermore, such metadata are needed in order to manage distributed queries efficiently; if a user is interested in finding x-ray images there is no point in querying the HST archive, for example. In this document we suggest an architecture for resource and service metadata and describe the relationship of this architecture to emerging Web Services standards. We also define an initial set of metadata concepts.

## Status of this document

This is a Recommendation. The [first release of this document](#) was 7 June 2002.

*This document has been produced by the [IVOA Resource Registry Working Group](#).*

*It has been reviewed by IVOA Members and other interested parties, and has been endorsed by the IVOA Executive Committee as an IVOA Recommendation. It is a stable document and may be used as reference material or cited as a normative reference from another document. IVOA's role in making the Recommendation is to draw attention to the specification and to promote its widespread deployment. This enhances the functionality and interoperability inside the Astronomical Community.*

*A list of [current IVOA Recommendations and other technical documents](#) can be found at <http://www.ivoa.net/Documents/>.*

## Acknowledgments

Many members of the IVOA Registry Working Group, IVOA Interoperability Working Group, AstroGrid project, and NVO Metadata Working Group have made significant contributions to this document. Contributors to this document have been partly or completely supported by the following projects and programs:

- The U.S. National Virtual Observatory project, which is funded by the National Science Foundation's Information Technology Research Program under Cooperative Agreement AST0122449 with The Johns Hopkins University.
- The UK AstroGrid project, which is funded by the Particle Physics and Astronomy Research Council.
- The Astrophysical Virtual Observatory, which is funded by the fifth framework program of the European Community for research, technological development, and demonstration activities (FP5).

## Contents

[Abstract](#)

[Status of this document](#)

[Acknowledgments](#)

[Contents](#)

[1 Introduction](#)

[2 Architecture](#)

[3 Resource metadata concepts](#)

[3.1 Identity metadata](#)

[3.2 Curation metadata](#)

[3.3 General content metadata](#)

- [3.4 Collection and service content metadata](#)
- [4 Data quality assessment](#)
- [5 Service metadata concepts](#)
  - [5.1 Interface metadata](#)
  - [5.2 Capabilities metadata](#)
- [6 Example](#)
- [7 Changes from previous versions](#)

## 1 Introduction

An essential capability of the Virtual Observatory is a means for describing what data and computational facilities are available where, and once identified, how to use them. The data themselves have associated metadata (e. g., FITS keywords), and similarly we require metadata about data collections and data services so that VO users can easily find information of interest. Furthermore, such metadata are needed in order to manage distributed queries efficiently; if a user is interested in finding x-ray images there is no point in querying the HST archive, for example. In this document we suggest an architecture for resource and service metadata and describe the relationship of this architecture to emerging Web Services standards. We also define an initial set of metadata concepts.

## 2 Architecture

In order to make it easy for astronomy information services to participate in the VO, we propose a hierarchical system for metadata management. At the top level we require a minimum amount of information, sufficient primarily to note the existence of a resource and to describe who is responsible for it. At lower levels, the metadata are more extensive and complex, allowing for the description of query syntax, access protocols, and usage policies.

A *resource* is a general term referring to a VO element that can be described in terms of who curates or maintains it and which can be given a name and a unique identifier. Just about anything can be a resource: it can be an abstract idea, such as sky coverage or an instrumental setup, or it can be fairly concrete, like an organization or a data collection. This definition is consistent with its use in the general Web community as “anything that has

an identity” (Berners-Lee 1998, IETF RFC2396). We expand on this definition by saying that it is also describable.

An *organization* is specific type of resource that brings people together to pursue participation in VO applications. Organizations can be hierarchical and range greatly in size and scope. At a high level, an organization could be a university, observatory, or government agency. At a finer level, it could be a specific scientific project, space mission, or individual researcher. A *provider* is an organization that makes data and/or services available to users over the network.

A *service* is any VO resource that can be invoked by the user to perform some action on their behalf. Associated with any service is descriptive *metadata* about the service. Metadata generally include information the user needs to determine if a service is of interest and how the service may be invoked. Specific types of metadata are described below. Note that the service itself need not be aware of the metadata that describe it.

A *query service* supports a query/response protocol. The user submits a query to the service that may define characteristics of interest, and the service returns a set of information to the user. The query may be null, e. g., a current-time service may only support a null query, and some services may respond to a null query with appropriate default actions. Non-query services may also exist, e.g., services to copy or delete files on remote files systems, to mail information to other users, to kill existing jobs, to authorize actions, etc.

A *registry* is a query service for which the response is a structured description of resources. The resources described by a registry may be of any type. The registry may support a query that allows the user to indicate which resources might be of interest.

In our model, the hierarchy of resources is one in terms of management and curation. For example, an organization may manage a collection of one or more services and even smaller organizations or projects. For example, MAST, HEASARC, IRSA, NED et al. are all resources. Each of these manages other resources, e. g., the HST archive in MAST. They also support specific services (which are also resources) such as an HST observation log query service or a cone search service. One could in principle describe all of NASA astrophysics data holdings as a resource, or all of NVO as a resource, but aggregates of this scale circumvent the goal of being able to locate the specific resources and services of interest for a particular application.

All resources are described by metadata. *Resource metadata* are generic, high-level, and independent of any specific service. Resource metadata include

- *Identity metadata*, which gives the resource a name and an identifier,

- *Curation metadata*, which describe who supports the resource and its availability (i.e., version, release date), and
- *Content metadata*, which describe what kind of information is available (types of data, sky coverage, spectral coverage, etc.). Content metadata can be either general, applying to all resources, or associated more specifically with data collections and the services that deliver data from them.

Resource metadata are typically not queryable parameters in the underlying services, but rather they encompass information that now is simply “known” to users, or must be discovered through other means. Astronomers know that the HST archive includes optical images and spectra, for example, or that VizieR provides access to catalogs and tables. Resource metadata constitute a “yellow pages” of astronomical information. Resource metadata are analogous to the UDDI (Universal Description, Discovery and Integration) Web Service, and are analogous to the high-level descriptions included in the CDS GLU.

Organizations, data collections, and services can be considered as classes of resources that may each require additional metadata to fully describe it, but which are not shared by other classes. For example, a service description would need to include its inputs, outputs, and how it can be accessed. *Service metadata*, therefore, can be thought of as an extension of the general resource metadata: where as the resource metadata, through its content metadata, describes *what* is available, the service metadata describes *how* to access it.

Resource metadata will be collected through resource registration services, e.g., web forms that present a resource curator with the requisite fields and enumerated lists, and construct a resource descriptor in a standard format (such as VOTable). The resource registration service should not allow fields to be left unspecified. Some metadata elements may be irrelevant, unknown, or not provided by the publisher of a resource. Since “irrelevant” conveys different information than “not provided”, we will adopt standard representations of these conditions:

|                  |   |
|------------------|---|
| “Not Applicable” | irrelevant or not applicable to this resource         |
| “Unknown”        | unknown, cannot be defined                            |
| “Not Provided”   | no information was provided by the resource publisher |

Various applications based on the registry may choose to include or exclude certain resources based on these attributes. If a metadata element is “Not Provided” the application should make no assumption regarding applicability or relevance.

Similarly, some resources may provide quite large aggregations or collections, covering many bandpasses, types, or formats. It may be prohibitive to list all such options. In such cases acceptable representations for the metadata entries would be:

|       |  |
|-------|--|
| “Any” | resource will respond to requests for any of the available types (though some may not actually be available)               |
| “All” | resource will respond to requests for all of the available types, and all are actually available in some non-zero quantity |

The most general resource metadata is similar in concept to the Dublin Core metadata definitions (<http://dublincore.org/documents/dces/>), and where possible DC metadata elements have been used. VO metadata elements that correspond directly to DC counterparts are noted. The Dublin Core elements Language and Relation are not currently used in the VO metadata.

### 3 Resource metadata concepts

Below we describe the *concepts* we believe are needed in the resource metadata. These concepts may be instantiated in a variety of standard forms, e.g. XML, UCD tags, or FITS keywords, and with a variety of mechanisms, such as Topic Maps, OWL, or RDBMSs. Consequently, the exact names and rendering of the values may depend on the particular form in which they are represented. For example, when *Coverage.Spatial* is rendered as a FITS keyword record, the name will need to be limited to 8 characters and the value rendered in a pure ASCII form; in contrast, when rendered in XML, it might be better to tag the different components of the value separately. It will be necessary to define standard renderings for each of these common forms.

A limited number of keywords are considered essential for a basic understanding of the resource, and are thus denoted as *required*. All others are optional, or may be applied to certain classes of resources only.

#### 3.1 Identity metadata

*Title* (string) [Dublin Core] [Required]

Definition: A name given to the resource.

Comment: Typically, a Title will be a name by which the resource is formally known.

*ShortName* (string)

Definition: A short abbreviation for the name given to the resource.

Comment: The ShortName will be used where brief annotations for the resource name are required.

ShortName strings are limited to a maximum of sixteen characters.

*Identifier* (URI) [Dublin Core] [Required]

Definition: An unambiguous reference to the resource within a given context. The syntax for Identifiers is described in *IVOA Identifiers* in the IVOA document collection (<http://www.ivoa.net/Documents/>).

Comment: The URI corresponding to the resource.

## 3.2 Curation metadata

*Publisher* (string) [Dublin Core] [Required]

Definition: An entity responsible for making the resource available

Comment: Examples of a Publisher include a person or an organization. Users of the resource should include Publisher in subsequent credits and acknowledgments.

*PublisherID* (URI)

Definition: The identifier for the entity responsible for making the resource available. The syntax for Identifiers is described in *IVOA Identifiers* in the IVOA document collection (<http://www.ivoa.net/Documents/>).

Comment: This item is optional; an ID for the publisher may not yet be established (e.g., if the publisher has not yet been registered).

*Creator* (string) [Dublin Core]

Definition: An entity primarily responsible for making the content of the resource.

Comment: Examples of a Creator include a person or an organization. Users of the resource should include Creator in subsequent credits and acknowledgments.

*Creator.Logo* (URL)

Definition: A URL pointing to a graphical logo, which may be used to help identify the information resource.

*Contributor* (string) [Dublin Core]

Definition: An entity responsible for making contributions to the content of the resource.



Comment: Examples of a Contributor include a person or an organization. Users of the resource should include Contributor in subsequent credits and acknowledgments.

*Date* (string) [Dublin Core]

Definition: A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability (i.e., most recent release or version) of the resource. ISO8601 is the preferred format (YYYY-MM-DD).

*Version* (string)

Definition: A label associated with the creation or availability (i.e., most recent release or version) of the resource.

*Contact* (string, e-mail address)

Definition: The e-mail address for contacting the persons responsible for the resource.

Comment: Contact is split into two components for clarity.

*Contact.Name* (string)

Definition: The name of the contact.

Comment: A person's name, "John P. Jones", or a group, "Archive Support Team".

*Contact.Email* (e-mail address)

Definition: The e-mail address of the contact.

Comment: For example, "John.P.Jones@navy.gov", or "archive@datacenter.org".

### 3.3 General content metadata

*Subject* (string, list) [Dublin Core] [Required]

Definition: A list of the topics, object types, or other descriptive keywords about the resource.

Comment: Subject is intended to provide additional information about the nature of the information provided by the resource. Is this a catalog of quasars? Of planetary nebulae? Is this a tool for computing ephemerides?

Terms for Subject should be drawn from the IAU Astronomy Thesaurus (<http://msowwww.anu.edu.au/library/thesaurus/>), though in the absence of suitable terms (the IAU Thesaurus is not complete in all areas

of astronomical research) the following alternate collections of astronomical research terms may be used:

Vizier keywords (CDS): <http://vizier.u-strasbg.fr/doc/ADCkwsd.htm>

Astronomy journal keywords: [http://www.edpsciences.org/journal/statique/doc/aa\\_keywords.html](http://www.edpsciences.org/journal/statique/doc/aa_keywords.html)

*Description* (string, free text) [Dublin Core] [Required]

Definition: An account of the content of the resource.

Comment: Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

*Source* (string) [Dublin Core]

Definition: A bibliographic reference from which the present resource is derived or extracted.

Comment: The present resource may be derived from the Source in whole or in part. Recommended best practice is to use the standard *bibcode* (see <http://cdsweb.u-strasbg.fr/simbad/refcode.html>), where available. If no *bibcode* is available, Source should use a string or number conforming to a formal identification or citation system.

*ReferenceURL* (URL) [Required]

Definition: A URL pointing to additional information about the resource. In general, this information should be human-readable.

*Type* (string, list) [Dublin Core] [Required]

Definition: The nature or genre of the content of the resource.

Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. VO Types include:

| <u>Type</u>  | <u>Description</u>  |
|--------------|---|
| Archive      | Collection of pointed observations  |
| Bibliography | Collection of bibliographic references, abstracts, and publications             |
| Catalog      | Collection of derived data, primarily in tabular form                           |
| Journal      | Collection of scholarly publications under common editorial policy              |
| Library      | Collection of published materials (journals, books, etc.)                       |
| Simulation   | Theoretical simulation or model   |
| Survey       | Collection of observations covering substantial and contiguous areas of the sky |

|              |   |
|--------------|---|
| Education    | Collection of materials appropriate for educational use, such as teaching resources, curricula, etc.  |
| Outreach     | Collection of materials appropriate for public outreach, such as press releases and photo galleries   |
| EPOResource  | Collection of materials that may be suitable for EPO products but which are not in final product form, as in Type Outreach or Type Education. EPOResource would apply, e.g., to archives with easily accessed preview images or to surveys with easy-to-use images. |
| Animation    | Animation clips of astronomical phenomena   |
| Artwork      | Artists' renderings of astronomical phenomena or objects  |
| Background   | Background information on astronomical phenomena or objects   |
| BasicData    | Compilations of basic astronomical facts about objects, such as approximate distance or membership in constellation.  |
| Historical   | Historical information about astronomical objects.  |
| Photographic | Publication-quality photographs of astronomical objects.  |
| Press        | Press releases about astronomical objects.  |
| Organization | An organization that is a publisher or curator of other resources.  |
| Project      | A project that is a publisher or curator of other resources.  |
| Registry     | A query service for which the response is a structured description of resources.  |
| Other        | A resource not described by any of the above types.   |

This list is extensible. Resources providing more than one type of content should list all relevant types.

*ContentLevel* (string, list)

Definition: A description of the content level, or intended audience.

Comment: VO resources will be available to professional astronomers, amateur astronomers, educators, and the general public. These different audiences need a way to find material appropriate for their needs.

ContentLevel

General

Definition

Resource provides information appropriate for all users

|                         |  |
|-------------------------|--|
| Elementary Education    | Resource provides information appropriate for grades K-4 education   |
| Middle School Education | Resource provides information appropriate for grades 5-8 education   |
| Secondary Education     | Resource provides information appropriate for grades 9-12 education  |
| Community College       | Resource provides information appropriate for education at community colleges  |
| University              | Resource provides information appropriate for university-level education   |
| Research                | Resource provides information appropriate for professional-level research and graduate school education                  |
| Amateur                 | Resource provides information of interest to amateur astronomers   |
| Informal Education      | Resource provides information appropriate for education at museums, planetariums, and other centers of informal learning |

#### *Relationship* (string)

Definition: A resource may be related to another resource in a way that is important to document, so that associated services or duplicate copies may easily be located.

|              |   |
|--------------|---|
| mirror-of    | The resource is a mirror of another resource. Information gathered from the resources is indistinguishable.                                       |
| service-for  | The resource is a service associated with a data collection.  |
| derived-from | The resource is a derivative of another resource, e.g., a subset selected for a particular scientific interest, or a reprocessed data collection. |

#### *RelationshipID* (URI)

Definition: The identifier of an associated resource. The relationship is described in the Relationship metadata element. The syntax for Identifiers is described in *IVOA Identifiers* in the IVOA document collection (<http://www.ivoa.net/Documents/>).

### 3.4 Collection and service content metadata

*Facility* (string, list)

Definition: The observatory or facility where the data was obtained.

Comments: Some resources are likely to hold data from multiple observatories. If just a few, this could be a list; if very many, just say “many”. Theoretical data will not originate with an observatory, but rather might be characterized by the computational facility used to create them (NCSA, SDSC, etc.).

*Instrument* (string, list)

Definition: The instrument used to collect the data.

Comments: Can be a specific instrument name (Wide Field/Planetary Camera 2) or generic instrument type (CCD camera). Theoretical data is produced by a computer code, and the name of the code could be specified.

*Coverage* (string) [Dublin Core, with modifications]

Definition: The extent of scope of the content of the resource.

Comment: The Dublin Core notion of coverage is too generic to be of much use in the VO, where we need more specific information. We propose to subset this element as follows:

*Coverage.Spatial* (string)

Definition: The sky coverage of the resource.

Comment: The complete syntax for the spatial coverage specification is described in the [Space-Time Metadata definition document](#) Region definition. Resource metadata may be somewhat simplified (i.e., do not give detailed spatial coverage of a large archive), but should be expressed in a syntax which adheres to the STM specification. All positions should be given in degrees unless specified otherwise. The list below suggests a basic representation of regions, but implementations should rely on the Region schema definition.

| <u>Region Name</u> | <u>Specification</u>   |
|--------------------|--|
| Circle             | circle (cframe, $\xi_{\text{cen}}$ , $\eta_{\text{cen}}$ , radius)   |
| Ellipse            | ellipse (cframe, $\xi_{\text{cen}}$ , $\eta_{\text{cen}}$ , semi-major axis, semi-minor axis, position angle)  |
| Polygon            | polygon (cframe, $\xi_1$ , $\eta_1$ , $\xi_2$ , $\eta_2$ , $\xi_3$ , $\eta_3$ , ...) where $\xi, \eta$ pairs give the vertices of a polygon in counterclockwise order. In spherical coordinates, the segments between vertices are assumed to be great circles |

|            |   |
|------------|---|
| Sector     | unless small circles are explicitly noted.<br>sector (cframe, $\xi$ , $\eta$ , $\theta_1$ , $\theta_2$ ) where $\theta_1$ and $\theta_2$ are position angles bounding the sector                    |
| Constraint | constraint (cframe, x, y, z, d) where x, y, and z are the components of a normal vector on the unit sphere and d is the distance from the center of the sphere. Constraint defines a spherical cap. |

$\xi$  and  $\eta$  represent coordinates in the appropriate frame ( $\alpha$ ,  $\delta$ ; l, b; ...). Compound regions may be constructed with unions (logical *or*), intersections (logical *and*), and negation (logical *not*). The coordinate system reference frame is specified as follows (<http://aladin.u-strasbg.fr/java/doctech/cds/astro/Astroframe.html> for additional details):

| <u>cframe</u> | <u>Description</u>                         |
|---------------|--|
| ICRS          | International Celestial Reference System   |
| FK5           | Equatorial coordinates, FK5 system (J2000) |
| FK4           | Equatorial coordinates, FK4 system (B1950) |
| ECL           | Ecliptic coordinates (J2000)               |
| GAL           | Galactic coordinates (J2000)               |
| SGAL          | Supergalactic coordinates (J2000)          |

*Coverage.RegionOfRegard* (float, decimal degrees)

Definition: Both data archives and catalogs have an intrinsic scale size. For example, a source catalog created from an instrument with one degree angular resolution would have a RegionOfRegard of 0.5 degree, meaning that if one is searching for information pertinent to a given position, objects in this catalog within 0.5 degree of the position of interest would need to be included. For an image archive the RegionOfRegard corresponds to the image field of view.

Comment: RegionOfRegard corresponds to CoordArea in the Space-Time Metadata definition document.

*Coverage.Spectral* (string, list)

Definition: The spectral coverage of the resource.

Comment: Spectral coverage at the resource level will be in terms of general spectral regions (gamma-ray, x-ray, extreme UV, UV, optical, infrared, radio). The general spectral regions are defined specifically as follows:

| <u>Coverage.Spectral</u> | <u>Represents</u> |
|--------------------------|-------------------|
|--------------------------|-------------------|

|             |   |
|-------------|---|
| Radio       | $\lambda \geq 10 \text{ mm}$<br>$\nu \leq 30 \text{ GHz}$   |
| Millimeter  | $0.1 \text{ mm} \leq \lambda \leq 10 \text{ mm}$<br>$3000 \text{ GHz} \geq \nu \geq 30 \text{ GHz}$   |
| Infrared    | $1 \text{ } \mu \leq \lambda \leq 100 \text{ } \mu$   |
| Optical     | $0.3 \text{ } \mu \leq \lambda \leq 1 \text{ } \mu$<br>$300 \text{ nm} \leq \lambda \leq 1000 \text{ nm}$<br>$3000 \text{ } \text{\AA} \leq \lambda \leq 10000 \text{ } \text{\AA}$ |
| Ultraviolet | $0.01 \text{ } \mu \leq \lambda \leq 0.3 \text{ } \mu$<br>$100 \text{ } \text{\AA} \leq \lambda \leq 3000 \text{ } \text{\AA}$<br>$1.2 \text{ eV} \leq E \leq 120 \text{ eV}$       |
| X-ray       | $0.1 \text{ } \text{\AA} \leq \lambda \leq 100 \text{ } \text{\AA}$<br>$0.12 \text{ keV} \leq E \leq 120 \text{ keV}$   |
| Gamma-ray   | $E \geq 120 \text{ keV}$  |

Resources containing data in multiple spectral regions may give a list (e.g., “Radio, Infrared”).

*Coverage.Spectral.Bandpass* (string, list)

Definition: A specific bandpass specification.

Comment: Some resources and services may choose to give spectral coverage in more specific terms than the general spectral regions. The list of possible bandpass names is too lengthy to enumerate here, but would include optical bandpasses (U, V, B, R, I), narrow line filters (H-alpha, [OIII]), or other specific bandpass names.

*Coverage.Spectral.CentralWavelength* (float)

Definition: The central wavelength of the spectral bandpass, in meters.

Comment: This should be the most representative wavelength of the bandpass.

*Coverage.Spectral.MinimumWavelength* (float)

Definition: The minimum wavelength of the spectral bandpass, in meters.

Comment: Implementers are encouraged to set the minimum wavelength to be as inclusive as possible, allowing all relevant resources to be discovered.

*Coverage.Spectral.MaximumWavelength* (float)

Definition: The maximum wavelength of the spectral bandpass, in meters.

Comment: Implementers are encouraged to set the maximum wavelength to be as inclusive as possible, allowing

all relevant resources to be discovered.

*Coverage.Temporal.StartTime* (string)

Definition: The earliest temporal coverage of the resource.

Comment: Temporal coverage specifications will be given in ISO8601 format. An empty value field implies that there is no known earliest temporal coverage.

*Coverage.Temporal.StopTime* (string)

Definition: The latest temporal coverage of the resource.

Comment: Temporal coverage specifications will be given in ISO8601 format. An empty value field implies that there is no known latest temporal coverage, i.e., that information continues to be added to the resource.

*Coverage.Depth* (float)

Definition: The (typical) depth coverage, or sensitivity, of the resource. Coverage.Depth is specified in flux density (Jy).

*Coverage.ObjectDensity* (float)

Definition: The (typical) density of objects, catalog entries, telescope pointings, etc., on the sky, in number per square degree.

*Coverage.ObjectCount* (int)

Definition: The total number of objects, catalog entries, telescope pointings, etc., in the resource.

*Coverage.SkyFraction* (float)

Definition: The fraction of the sky represented in the observations, ranging from 0 to 1.

*Resolution* (float)

Definition: The resolution of the resource contents.

Comment: Resolution is divided into the following sub-elements:

*Resolution.Spatial* (float)

Definition: The spatial (angular) resolution that is typical of the observations, in decimal degrees.

*Resolution.Spectral* (float)

Definition: The spectral resolution that is typical of the observations, given as the ratio  $\lambda/\Delta\lambda$  (so that higher spectral resolution has a larger number).



*Resolution.Temporal* (float)

Definition: The temporal resolution that is typical of the observations, given in seconds.

*UCD* (string, list)

Definition: A list of the UCDs (Unified Content Descriptors, <http://cdsweb.u-strasbg.fr/doc/UCD.htx>) represented in the resource.

Comment: Some large or complex resources will have hundreds of associated UCDs and are unlikely to be specified in the resource metadata. Users of the resource metadata should not assume that an empty specification implies that the resource has no associated UCDs.

*Format* (string, list) [Dublin Core]

Definition: The physical or digital manifestation of the information provided by the resource.

Comments: Typical values would be “image/fits”, “image/gif”, “text/plain”, “text/html”, “text/xml” (for VOTable), etc. MIME types should be used where available to specify digital information formats in order to utilize existing standards.

Other format values will be used to describe the physical medium of the information: CDROM, Digital Planetarium, Online, Presentation, Print, Slides, Video. Format specifications may be combined, as in “Video, video/mpeg” (both hardcopy video cassettes and on-line MPEG files) or “CDROM, image/fits, image/gif” (FITS and GIF images are available on-line and on CDROM hardcopy).

*Rights* (string) [Dublin Core]

Definition: Information about rights held in and over the resource.

Comment: Dublin Core uses Rights to describe copyright and other intellectual property rights issues. In the VO context Rights would describe access privileges, using the following values: public, proprietary, mixed.

## 4 Data quality assessment

Users of virtual observatory resources need some way to assess the quality of the data. Data quality is both subjective and quantitative, and data collections may have no single data quality metric. While the completeness and consistency of the resource metadata itself may be a reasonable indicator of the quality of the associated resource, this is at best a qualitative measure. The following metadata elements are intended to capture the most basic measures of data quality, and may well require extensions as VO usage practices

evolve and become more sophisticated.

*DataQuality* (char)

Definition: An overall assessment of the integrity, consistency, and level of documentation concerning uncertainty estimates and calibration procedures, of the data provided by the resource. We suggest 3 general grade levels, plus codes for unknown or undocumented cases:

- A Data are fully calibrated, fully documented, and suitable for professional research.
- B Data are calibrated and documented, but calibration quality is inconsistent. Users are advised to check data carefully and recalibrate.
- C Data are uncalibrated.
- U Data quality is unknown. If a resource does not provide a data quality assessment, class U should be assumed.

*Uncertainty.Photometric* (float)

Definition: The uncertainty of the photometric measurements provided by the resource, given in Jy.

*Uncertainty.Spatial* (float)

Definition: The uncertainty of the astrometric, or positional measurements, provided by the resource, given in degrees.

*Uncertainty.Spectral* (float)

Definition: The uncertainty of the wavelengths provided by the resource, given in meters.

*Uncertainty.Temporal* (float)

Definition: The uncertainty of the temporal measurements provided by the resource, given in seconds.

## 5 Service metadata concepts

The metadata necessary for describing a service will vary quite a bit depending on the type of service it is. We propose two general categories of service metadata:

*Interface metadata*, which describe how to access the service—the inputs and the outputs. There will be

standard types of interfaces that could include a web-browser-based interface (i.e., HTML Forms), a Web Service interface (describable by a WSDL document), a general HTTP Get interface (e.g., using *key=value* arguments), and a GLU-described interface.

*Capability metadata*, which describe what the service does, its limitations, and other behavioral characteristics.

Note that these categories are reasonably orthogonal. We can imagine the same basic service—in terms of its capabilities—accessible through multiple interfaces.

We expect that for each standard service recognized by the VO there will be a specification document that defines all the specific metadata necessary to describe a particular implementation of that service; thus, we do not include them all here. However, we can identify a few metadata concepts that might be employed to describe a particular service. Described below, these concepts should be employed by standard service specifications wherever they are applicable. We note also that metadata associated with the VOTable schema can also be reused to describe the inputs and outputs of a service that returns a VOTable.

## 5.1 Interface metadata

*Service.InterfaceURL* (URL)

Definition: A URL pointing to a document that presents or describes the service interface.

Comment: For a Web Service, this would point to the WSDL document, for a GLU-described service, it would point to the GLU record, and for a browser-based service, this would be the Web page that actually contains the Web Form.

*Service.BaseURL* (URL)

Definition: The base portion of a URL used to invoke a service with the expectation that an additional string must be appended for the service to execute properly. The syntax of the appended string is defined by the specific service.

*Service.HTTPResultsMIMEType* (MIME type)

Definition: The MIME type that is returned by a service.

## 5.2 Capabilities metadata

*Service.StandardURI* (URI)

Definition: An identifier for a standard service. The syntax for Identifiers is described in *IVOA Identifiers* in the

IVOA document collection (<http://www.ivoa.net/Documents/>).

Comment: This provides a unique way to refer to a service specification standard, such as a Simple Image Access service. It assumes that such standard is registered and accessible.

*Service.StandardURL* (URL)

Definition: A URL that points to a human-readable document that describes the standard upon which a service is based.

*Service.MaxSearchRadius* (float, decimal degrees)

Definition: Service providers may choose to restrict the scope of searches done against their services, lest they be swamped with requests for millions or billions of results records. *Service.MaxSearchRadius* restricts searches to some maximum radius (in decimal degrees) about a celestial coordinate.

Comment: A value of 180.0 or greater denotes that there is no restriction.

*Service.MaxReturnRecords* (int)

Definition: Service providers may choose to restrict the number of records returned in order to avoid swamping the user with responses to an overly general query. If no value is provided, it is assumed that there is no restriction on the number of records returned.

## 6 Example

Example: The Sloan Digital Sky Survey data as hosted by MAST at STScI (with no assertion that the metadata element values are actually correct, though they are not unreasonable).

### Identity metadata

|            |                           |
|------------|---------------------------|
| Title      | Sloan Digital Sky Survey  |
| ShortName  | SDSS                      |
| Identifier | ivo://stsci.edu/mast/sdss |

### Curation metadata

|              |   |
|--------------|---|
| Publisher    | Space Telescope Science Institute/MAST  |
| PublisherID  | ivo://stsci.edu/mast  |
| Creator      | Sloan Digital Sky Survey Consortium   |
| Creator.Logo | <a href="http://archive.stsci.edu/images/sdss_logo.gif">http://archive.stsci.edu/images/sdss_logo.gif</a> |
| Contributor  | Sloan Digital Sky Survey Consortium   |

|               |   |
|---------------|---|
| Date          | 2003-02-01  |
| Version       | SDSS EDR  |
| ReferenceURL  | <a href="http://archive.stsci.edu/sdss/index.html">http://archive.stsci.edu/sdss/index.html</a> |
| Contact.Name  | Archive Branch, Space Telescope Science Institute   |
| Contact.Email | <a href="mailto:archive@stsci.edu">archive@stsci.edu</a>  |

### General content metadata

|                |  |
|----------------|--|
| Subject        | galaxies, quasars, stars, CCD photometry,<br>spectroscopy, redshift, sky surveys   |
| Description    | The Sloan Digital Sky Survey is using a dedicated 2.5 m telescope and a large format CCD camera to obtain images of over 10,000 square degrees of high Galactic latitude sky in five broad bands (u', g', r', i' and z', centered at 3540, 4770, 6230, 7630, and 9130 Å, respectively). Medium resolution spectra will be obtained for approximately 10 <sup>6</sup> galaxies and 100,000 quasars. The early data release (EDR), on June 2001, includes searchable catalogs of images and spectra, images for display and scientific purpose in both 2-D FITS and JPEG formats, and spectra in both 1-D FITS and GIF formats. The EDR covers about 460 square degrees of sky. The next data releases will occur every 18 months or so. |
| Source         | 2002AJ....123..485S  |
| Type           | Survey, Catalog, EPOResource   |
| ContentLevel   | Research   |
| Relationship   | mirror-of  |
| RelationshipID | <a href="http://ivo://sdss.org/sdss/edr">ivo://sdss.org/sdss/edr</a>   |

### Collection and service content metadata

|                                     |   |
|-------------------------------------|---|
| Facility                            | Apache Point Observatory, Sloan 2.5-m Telescope   |
| Instrument                          | Five-band clocked CCD camera  |
| Coverage.Spatial                    | polygon (FK5, 145.17, 1.25, 235.9, 1.25, 235.9, -1.25, 145.17, 1.25) or polygon (FK5, 250.71, 66.29, 267.0, 66.29, 267.0, 52.15, 250.71, 66.29) or polygon (FK5, 350.43, 1.17, 360.0, 1.17, 360.0, -1.25, 350.43, -1.25) or polygon (FK5, 0.0, 1.17, 56.37, 1.17, 56.37, -1.25, 0.0, -1.25) |
| Coverage.RegionOfRegard             | 0.0001  |
| Coverage.Spectral                   | Optical   |
| Coverage.Spectral.Bandpass          | u', g', r', i', z'  |
| Coverage.Spectral.MinimumWavelength | 400.e-9   |
| Coverage.Spectral.MaximumWavelength | 850.e-9   |
| Coverage.Temporal.StartTime         | 1999-12-25  |
| Coverage.Temporal.StopTime          | 2001-07-15  |
| Coverage.Depth                      | 3.e-6   |

|                        |              |
|------------------------|--------------|
| Coverage.ObjectDensity | 6.e4         |
| Coverage.ObjectCount   | 2.e7         |
| Coverage.SkyFraction   | 0.01         |
| Resolution.Spatial     | 0.00028      |
| Resolution.Spectral    | 5000         |
| Resolution.Temporal    | 120          |
| UCD                    | Not Provided |
| Format                 | text/xml     |
| Rights                 | Public       |

#### Data quality metadata

|                         |         |
|-------------------------|---------|
| DataQuality             | A       |
| Uncertainty.Photometric | 3.e-7   |
| Uncertainty.Spatial     | 0.00003 |
| Uncertainty.Spectral    | 1.e-11  |
| Uncertainty.Temporal    | 0.1     |

#### Service metadata

|                             |  |
|-----------------------------|--|
| Service.InterfaceURL        | <a href="http://archive.stsci.edu/sdss/catalog.html">http://archive.stsci.edu/sdss/catalog.html</a>              |
| Service.BaseURL             | <a href="http://archive.stsci.edu/cgi-bin/sdss/catalog">http://archive.stsci.edu/cgi-bin/sdss/catalog</a>        |
| Service.HTTPResultsMIMEType | text/xml   |
| Service.StandardID          | ivo://ivoa.net/Services/ConeSearch   |
| Service.StandardURL         | <a href="http://www.ivoa.net/Documents/REC/ConeSearch.html">ivo://www.ivoa.net/Documents/REC/ConeSearch.html</a> |
| Service.MaxSearchRadius     | 0.2  |
| Service.MaxReturnRecords    | 5000   |

## 7 Changes from previous versions

### From V1.0 to V1.01

- Revised wording in Section 2 to clarify phrase concerning “independence” of resource metadata.
- Added text to indicate that PublisherID must have the form of a valid IVOA Identifier (Section 3.2).
- Clarified definition of ReferenceURL (Section 3.3).
- Added text to indicate that RelationshipID must have the form of a valid IVOA Identifier (Section 3.3).
- Deleted Convex from list of possible values for Spatial.Coverage, as it is redundant with extant ability to specify region intersections (Section 3.4).
- Corrected URL to coordinate system descriptions at CDS (<http://aladin.u-strasbg.fr/java/doctech/cds/astro/>)

[Astroframe.html](#), Section 3.4).

- Renamed Coverage.Spectral value of UV to Ultraviolet, and deleted the value EUV. EUV spectral range is merged into UV. (Section 3.4.)
- Renamed Service.HTTPResults to Service.HTTPResultsMIMEType (Section 5.1 and in example).
- Renamed Service.StandardURI to Service.StandardID, consistent with example and with other elements that have the form of URIs/Identifiers. This was an inconsistency not noticed by reviewers previously. Also added text to indicate that Service.StandardID must have the form of a valid IVOA Identifier. (Section 5.2)
- Corrected URL in the example for element Service.InterfaceURL.

From V0.82 to V1.0

- Removed Creator from list of required elements (Section 3.2).
- Removed Contact.Address and Contact.Telephone, and modified description of Contact to reflect the removal of these elements (Section 3.2).
- Updated the example in Section 6 to reflect the above changes.

From V0.81 to V0.82:

- Included full change history from previous versions.
- Identified a subset of metadata elements as *required* for all resources. These are Title, Identifier, Publisher, Creator, Subject, Description, ReferenceURL, and Type.
- Added reference to IVOA Identifiers specification under metadata element Identifier (Section 3.1).
- Added Contact.Address and Contact.Telephone, and modified description of Contact to reflect these new elements (Section 3.2).
- Added alternate sources of Subject keywords in addition to the IAU Thesaurus (Section 3.3).
- Clarified definition of Source to avoid confusion with Dublin Core sense of “resource” (Section 3.3).
- Added Registry to list of recognized resource Types (Section 3.3).
- Added Other to list of recognized resource Types (Section 3.3).
- Added Relationship and RelationshipID to provide means to show relationships (mirror of, service for, derived from) between resources (Section 3.3).
- Restructured Section 3.3 to distinguish the most generic content metadata (which remains in Section 3.3) from content metadata that applies more specifically to data collections and their corresponding data access services (new Section 3.4).
- Refined definition of Coverage.Spatial (Section 3.4).
- Inserted Millimeter bandpass definition into Coverage.Spectral (Section 3.4).
- Added UCD to Section 3.4 to provide a way to show the contents of resources via their associated Unified Content Descriptors.

- Updated the example in Section 6 to include new metadata elements and use Identifiers that are consistent with the most recent drafts of the Identifier specification.
- Added acknowledgements to funding agencies.

From V0.8 to V0.81:

- Restricted units of Coverage.Depth to Jy and eliminated Coverage.Depth.Units. It is assumed that a registry data entry interface could handle unit conversions.
- Added section on data quality metadata, including elements DataQuality, Uncertainty.Photometric, Uncertainty.Spatial, Uncertainty.Spectral, and Uncertainty.Temporal.
- Changed names of Service-related metadata elements to separate Service root from the rest of the name (e.g., ServiceInterfaceURL → Service.InterfaceURL).
- Changed name of ServiceMSR to Service.MaxSearchRadius.
- Added Service.MaxReturnRecords.
- Updated example to reflect above.

From V0.7 to V0.8:

- Title: Document renamed to “Resource Metadata for the Virtual Observatory” recognizing that services are a subset of resources.
- Section 2: Made several minor wording changes to reflect generic use of the term “resource”, which encompasses “service”.
- Section 2: Made definition of “Curation Metadata” more specific.
- Section 2: Added definitions for metadata elements that are not provided, unknown, or not applicable. Replaces previous discussion at end of Section 3.3.
- Section 2: Added definitions for aggregate resources whose contents might include “any” or “all” of a given metadata element’s types.
- Sections 2 and 3: Added note in Section 2 describing relationship of VO metadata to Dublin Core. In Section 3, changed references to Dublin Core to cite those elements that are the same as DC rather than those elements that are different from DC.
- Section 3.1: Changed Ticker to ShortName and extended string length from 8 to 16 characters.
- Section 3.2: Changed “based on” to “drawn from” the IAU Astronomy Thesaurus, concerning the selection of Subject metadata.
- Section 3.2: Moved elements Title, Subject, and ReferenceURL from 3.2 to 3.3.
- Section 3.3: Added metadata element Source (from Dublin Core), allowing for a cross reference to the bibcode (or other form of citation) for an associated bibliographic reference.
- Section 3.3: Extended Type to include Organization and Project.



- Section 3.3: Modified Coverage.Spatial to conform to Region specification of Space-Time Metadata definitions.
- Section 3.3: Clarified definition of RegionOfRegard (which is NOT the angular resolution).
- Section 3.3: Added Coverage.Depth.Units to provide separate field for units string.
- Section 3.3: Added Coverage.SkyFraction.
- Section 3.3: Added Resolution.Spatial, Resolution.Spectral, and Resolution.Temporal elements.
- Section 3.3: Allowed Facility and Instrument elements to be lists.
- Section 3.3: Changed examples for Format to use standard MIME types.
- Section 5: Updated example to include above changes and additions.

From V0.6 to V0.7:

- Reformatted document to be compatible with new IVOA documentation standards.
- Modified much of the architecture description to reflect current concepts about the function and structure of registries.
- Distinguished *identity metadata* from other curation metadata.
- Added Ticker element to provide shorthand abbreviation for resource Title.
- Added PublisherID element to provide unique identifier for the resource publisher.
- Replaced *ResourceURL* with *ReferenceURL*.
- Deleted *ServiceURL* and incorporated more general service metadata (Section 4).
- Incorporated more complete metadata definitions for Education and Public Outreach resources. Affects: Type, Format, ContentLevel.
- Added Type = Simulation for theoretical models.
- Coverage.Spatial refers to Space-Time Metadata document for spatial coverage specifications.
- Added Coverage.Spectral.CentralWavelength, Coverage.Spectral.MinimumWavelength, and Coverage.Spectral.MaximumWavelength to more fully describe the spectral bandpass.
- Replaced Coverage.Temporal with Coverage.Temporal.StartTime and Coverage.Temporal.StopTime to be consistent with Date and use of ISO 8601 time format.
- Added Coverage.RegionOfRegard to aid users of the registry in assessing the utility of a resource.
- Added Coverage. Depth, Coverage.Object Density, and Coverage.ObjectCount to provide information on resource depth of coverage (sensitivity) and richness.
- Broadened definition of Format, consistent with Dublin Core, to include specifications of both physical and digital formats.
- Added new section on service metadata, introducing metadata elements ServiceInterfaceURL, ServiceBaseURL, ServiceHTTPResults, ServiceStandardID, ServiceStandardURL, and ServiceMSR.
- Updated example to use new metadata elements.

From V0.5 to V0.6:

- Added *Date* element, following Dublin Core and in response to October 2002 discussion.
- Added *Version* element, in response to October 2002 discussion. Not in Dublin Core.
- Added *Creator.Logo* element, in response to October 2002 discussion. Not Dublin Core.
- Added *Subject* element, to replace *Coverage.Topics*, as the former is Dublin Core, the latter is not, and the intent is basically the same.
- Added *ContentLevel* element, to replace *Coverage.Level*. Upon reflection I thought that *ContentLevel* was both a better label and was of a different nature than the other *Coverage* elements. M. Voit will provide further inputs on appropriate element values.
- Updated example to use new metadata elements.