Multidimensional Data in the Virtual Observatory

Alumno: José Enrique Ruiz del Mazo

Tutora: Lourdes Verdes-Montenegro

Instituto de Astrofísica de Andalucía - CSIC

Diciembre 2010

MASTER FISYMAT

TRABAJO DE INVESTIGACIÓN TUTELADA

UNIVERSIDAD DE GRANADA

Contents

1 Introduction

The concept of *"data deluge"* has been introduced in Astronomy to reflect the growing amount of data placed at the disposal of the astronomical community. An *"exaflood"* of observational data threatens to overwhelm scientists with the advent of a new generation of facilities and automatized surveys (e.g. LSST, panSTARRS, LOFAR, SKA, etc.) As pointed out in *The Fourh Paradigm* [HTT09], the only way to cope with it is a new kind of scientific computing tools to manage, visualize and analyze the data flood. These transformations should be made within the frame of the **Virtual Observatory** (VO). The Astronet Infrastructure Roadmap [BCM08] states clearly: *"In the longer term (ten years) the development of the VO is expected to merge into the standard practices for delivery of astronomy data. The scientific development is expected to be rich in innovations as VO leverages on data mining and semantic technologies."*

Since complex **multidimensional datasets** are becoming more and more frequent in astronomy, the Virtual Observatory has to be prepared in order to provide models and metadata for a complete description of these observations and standard protocols for their discovery and access. These observed data are multidimensional in the sense that they may cover several axes in the spatial, time, electromagnetic field frequency and polarization, and flux domains. E.g. radiointerferometry cubes, long-slit spectra, integral field spectroscopic observations, time series images datasets, etc. But because they may be seen as associated dicrete data products of different nature, they also conform complex generic datasets.

Present IVOA[1] (International Virtual Observatory Alliance) data models and Data Access Layer (DAL) recommendations provide the possibility to describe, discover and access most of the astronomical catalogs of numerical data, 1D spectra and 2D images. Nevertheless a **second generation of DAL services** [DTDS08] is preparing their way to reach the status of final IVOA recommendations. These services define a family of access protocols for all possible types of astronomical data, taking into account the addition of GRID capabilities for authentication, asynchronous data generation and distributed storage [TBDS08].

New DAL services need to provide standardized access methods for the generation of derived data products from complex server-side operations on the original datasets. In particular standard parameters are needed (redshift, velocity, spectral lines, etc.) for discovering and filtering science-ready datasets, together with a model for the complete description of the data [RB09]. Furthermore, more technical issues need to be addressed, since the large volume of these datasets may raise new challenges for improving transfer rates and latency issues.

Because of the size and complexity of the next generation of observational datasets, data providers should supply on-line processing and analysis services that will deliver on-the-fly generated data. These services, if they are properly self-described and characterized, can be reused as components for **internet-based workflows** that capture and preserve the

[1] http://www.ivoa.net/

scientific methodology. Since their execution is independent of the investigator's platform and given their modularity and their universal availability, they ensure the reproducibility of the results and their dissemination.

The work presented here has been developed within the AMIGA[2] (Analysis of the interstellar Medium of Isolated Galaxies) project, an international collaboration led from the Instituto de Astrofísica de Andalucía – CSIC. The scientific project focuses on a multi-wavelength analysis of a statistically significant sample of isolated galaxies, in order to provide baseline to compare with the behavior of galaxies in denser environments. Since intensive and complex analysis of 3D data at several wavelengths is needed for achieving the scientific goals of the project, the group is actively involved in the migration of analysis tasks on multidimensional data to VO services deployed in a storage and computing distributed environment (GRID), as well as in developments of astronomical workflows and their preservation.

The here presented work is a natural evolution of the **AMIGA group software activities** within the VO context [RdG+10], which started with the development of the first VO-compliant data model for single-dish radio observations RADAMS [San09]. In this document we present the next step, consisting in a detailed search and compilation of different solutions for the complete description of and access to multidimensional data within the VO framework, with the final aim of implementing standard VO access to 3D data. We hope this work foster and advance the discussions inside the community in order to achieve a final recommendation for IVOA DAL protocols that will take into account complex multidimensional datasets.

This document starts with a general description of the two objects of study: what we have called *"multidimensional data"* in Astronomy in §2 and the state of the art of the VO in §3. We will show in §4 how uniformly and non-uniformly sampled multidimensional datasets may be accessed and described in a very different way in the VO, and we will present the *"generic dataset"* concept, since multidimensional datasets may very often be considered as complex collections of associated data products of different nature. At this point we will propose in §5 a set of input parameters for the discovery of generic datasets and a structured response for their complete description, fully compliant with the present VO data models.

The main contribution of the undertaken study is a proposed data discovery operation for a *"generic dataset service"* (GDS) which serves as a blueprint for the development of subsequents standards for multidimensional data access methods on second generation DAL protocols. The here proposed discovery GDS operation has been implemented on one of the AMIGA group archives (B0DEGA) containing datacubes of galaxies, as it is described in §6. The B0DEGA archive will host future developments and implementations of eventual data access operations for a next generation of DAL services, where more simple virtual data products will be generated on-the-fly at user demand from complex multidimensional datasets physically stored on the archive. Data access operations will be discussed in §7, followed by a list of representative scientific uses cases.

[2] http://amiga.iaa.es

2 Multidimensional Data in Astronomy

Most of the astronomical data can be seen as projections of a five dimensional parameter space: two celestial coordinates, photon energy, light polarization and time. In this scenario the measured flux is a function of one or several of these parameters and the photon energy axis is generally interpreted as the frequency/wavelength of the received light in the electromagnetic field spectrum, quite often converted to velocity units by means of the Doppler law.

Although the majority of the twentieth century astronomical data have been two-dimensional images or one-dimensional spectra and time series, multidimensional data are becoming increasingly prevalent in the XXI century. The idealized picture of an astronomical data cube (Fig. 1) is an orthogonal cube with two spatial axes and a third for the frequency/wavelength/velocity coordinate.

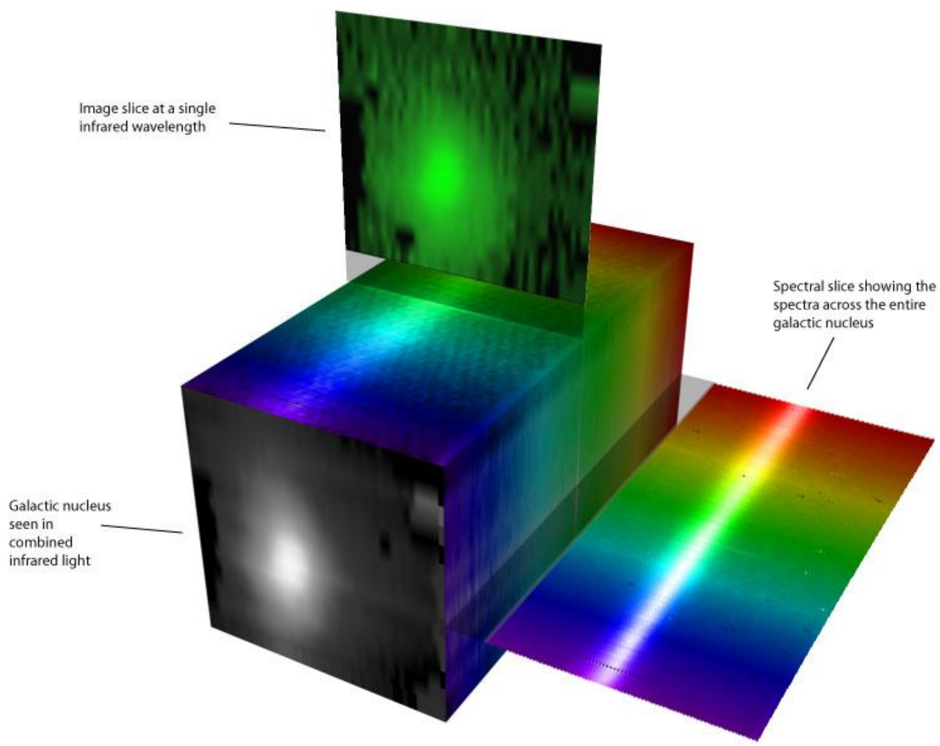


Figure 1. The concept of a data cube. Credit: Stephen Todd and Douglas Pierce-Price

Astronomical data cubes are ideal for exploring spatially extended sources in a frequency extended range. Any astronomical object that requires spatially resolved spectroscopic information is ideally suited for study with 3D data; some of the astrophysical fields where they may be useful are given in the following list.

- Pre-main sequence objects: proto-stellar discs, jets, Herbig Haro objects, collimation, mass loss, mass accretion, pre-solar nebula.
- Resolved stellar populations: local group and nearby galaxies, crowded fields, nuclear and bulge regions, most luminous stars, B and A supergiants, supernovae, Luminous Blue Variables, planetary nebulae, novae and cataclysmic variables, HII regions.
- Field normal galaxies: stellar and gas dynamics, supermassive black holes, galaxy nuclei, elliptical galaxies, spirals/bars.
- Active galaxies: gas dynamics, stellar populations, nuclear activity/starburst, fuelling of the active nucleus, extended narrow emission-line regions.
- Groups and cluster of galaxies: galaxy formation, galaxy evolution, tidal interaction and merging, star formation history.
- High redshift galaxies: galaxy evolution, galaxy dynamics, galaxy formation, star formation history, cosmology.
- Gravitational lensing: gravitational lens models, QSO unresolved structure, dark matter in lens galaxies, extinction laws in galaxies

Upcoming instruments (EVLA, ALMA, LOFAR, ASKAP, MeerKAT, etc.) will routinely produce data products resolved in the spatial, time, spectral, and polarization axes simultaneously. Moreover, the associated advances in sensitivity, field-of-view, frequency range and spectral resolution will provide multidimensional data cubes hundreds of GB larger than the existing ones, leading to transfer latency issues which will have to be considered by the data providers. The complications of dealing with these data as both imaging and spectroscopy make it a special challenge for the Virtual Observatory. In the next pages we describe the different kind of multidimensional data in Astronomy focusing on the techniques that are used to produce them.

2.1 Radiointerferometric

Radio interferometers are used in radio astronomy for aperture synthesis imaging. This technique allows radio telescopes to have a resolution equivalent to very large effective apertures by using an array of smaller widely spaced antennas. Many variations on the technique are possible, but all rely on collecting samples in the Fourier transform plane *(u-v plane)* of the image, taking advantage of the fact that the ideal interference pattern follows the complex Fourier transform of the emitted signal. Each point in the *u-v plane* corresponds to a particular orientation and physical separation of the antennas *(baselines)* in the interferometer. Sampling of the *u-v plane* is required, usually provided by the rotation of the earth or, in the case of Space Very Long Baseline Interferometry (VLBI), by the motion of an orbiting antenna.

The results of such measurements can yield an image of the intensity of the radio emitting source at each frequencies in the configured observed band. Radiointerferometric datacubes are the most common of 3D astronomical data. Radio telescopes feeds normally observe in two orthogonal polarizations and the cross products can also be performed, giving four polarization products.

2.2 Integral Field Spectroscopy

In integral field spectrographs (IFS) the focal plane is sampled into spaxels ('spatial pixels') and then the light from each is dispersed by a standard grating spectrograph, forming a collection of spectra. In a similar way to radio interferometers, all of the spectra are obtained simultaneously. The instrument inserted in the focal plane of a telescope in order to provide 3D astronomical data is usually referred as integral-field-unit (IFU).

The simplest IFS are derivatives of fiber multiobject spectrographs (MOS) where input ends of optical fibers are placed on pinholes masks defining the positions of the objects for which the spectra will be gathered. In IFS the fibers are packed as closely together as possible in the focal plane, the output of the fibers reformats the 2D field into a pseudo slit at the entrance of the spectrograph.

This technique has been implemented with different variations (Fig.2) in several instruments. Among the most known by the community are OASIS and SAURON IFUs on the William Herschel Telescope, OSIRIS on Keck Observatory and the PMAS spectrophotomer at the Calar Alto Observatory.

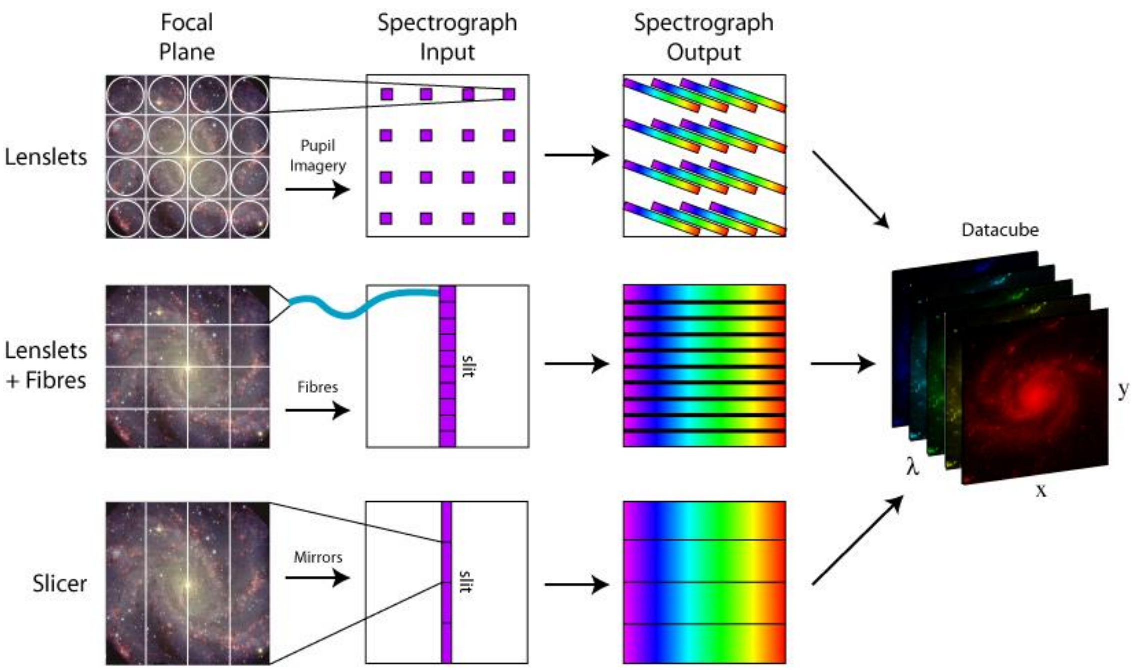


Figure 2. The main techniques in integral field spectroscopy. Credit: M. Westmoquette, adapted from [AC98]

Each of these intruments provide a unique native format. In some cases the reduced dataset is a set of 1D spectra, one for each spaxel, and a table of the relative positions of each spaxel on the sky. This format is called raw-stacked-spectra (RSS). Often the result is not a rectangularly gridded data cube. The spaxel coordinate may or may not be orthonormal; it is only through the process of interpolation in the spatial axes that arbitrary IFU geometries are converted to orthonormal.

2.3 Fabry-Perot

In the Fabry-Perot imaging technique, the interference between two closely spaced, adjustable, partially-transmissive plates creates a narrow bandpass. The main advantage of the tunable filter is that it can provide a very wide field-of-view, however the spectral resolution is relatively low. The data format provided is a regularly gridded datacube. Each spatial plane is a typical image that can be manipulated by most image processing software. This data is similar to a radio datacube and is relatively easy to process and visualize with current software.

2.4 Other cases

Although less frequently used, other multidimensional data can be identified. If we consider the time axis, time series for spectra and images arise in a first approach. Some use cases for images time series range from solar studies[3] to distant masers. A next generation of telescopes (PanSTARRS, LSST) will perform multi-epoch surveys delivering huge datasets of time-image video cubes. Another example of multidimensional time series is spectral variability of the broad components of emission lines in Seyfert galaxies. There is a large international project, AGN Watch[4], dealing with photometrical and spectral variability of Active Galactic Nuclei (AGN). Another example is image stacks produced with small exposures for infrared instruments, to avoid fringing, or those produced by lucky imaging instruments, such as Astrolux, for dynamically selecting conditions with best seeing, but also valid for short term source variability.

On-The-Fly (OTF) mapping observations also provide spectral datacubes. These observations are usually carried on single-dish radio antennas where the telescope beam moves across the source at a constant velocity while spectral data are taken on a selected band. More cases for multidimensional data could be added if we also take into account the Stokes polarization values, obtaining multiple combinations of images or spectra in different polarized planes and for different epochs of time.

[3] http://wiki.astrogrid.org/bin/view/Astrogrid/MovieMaker

[4] http://www.astronomy.ohio-state.edu/~agnwatch/

3 The Virtual Observatory

3.1 General Scope

In the last two decades observational projects have started to make available to the whole community the data resulting from space or ground based missions. With the advent of new technologies, interoperability [Gen02] between data and services has gained relevance because it allows an easy access and sharing of data in order to exploit the wealth of multi-instrument panchromatic datasets. The International Virtual Observatory Alliance (IVOA) aims to guarantee the interoperability of applications and data, developing and providing common standards for data access and description.

IVOA activities are structured in Working Groups and Interest Groups, each responsible of a given area where key interoperability standards have to be defined and agreed upon. In particular, the task of the Data Access Layer (DAL) working group is to define and formulate VO standards for remote data access protocols, and it is in this working group where the bulk of the here presented study and results can find the most suitable audience. Other groups are related to applications, data models, GRID and web services, semantics or registry issues.

Astronomical data are usually managed by large data centers. These *data providers* supply data and quite often also computing services to the *consumers*: astronomers, research teams or even computing systems. The VO is the necessary *middle layer* framework connecting the *resource layer* to the *user layer.* The VO provides technical standards for providers to share their data and services, and for developers of applications used by consumers to discover and find the resources and to use the data.

The *registry* can be described as a *yellow pages* service that collects into databases metadata about all data resources and services in order to record and also provide the information through specific discovery services. The registry, as most of the VO resources and services, is geographically distributed where several replicas exist in different parts of the world. All the registries harvest each other to know the new added data and services added to other VO registries. In the registry data of interest can be searched for where it is often possible to make requests based on position in the sky data.

A *resource* is a general term referring to a VO element describing who curates or maintains it and that can have a name and a unique identifier. Anything could be a resource as long as it has a VO identitifier, e.g. sky coverage, instrumental set up, data access services or a data collection. A VO resource identifier starts always with `ivo://`, and may contain links to related resources, as well as external links to the literature. An overall description of the VO architecture can be seen in Fig. 3.

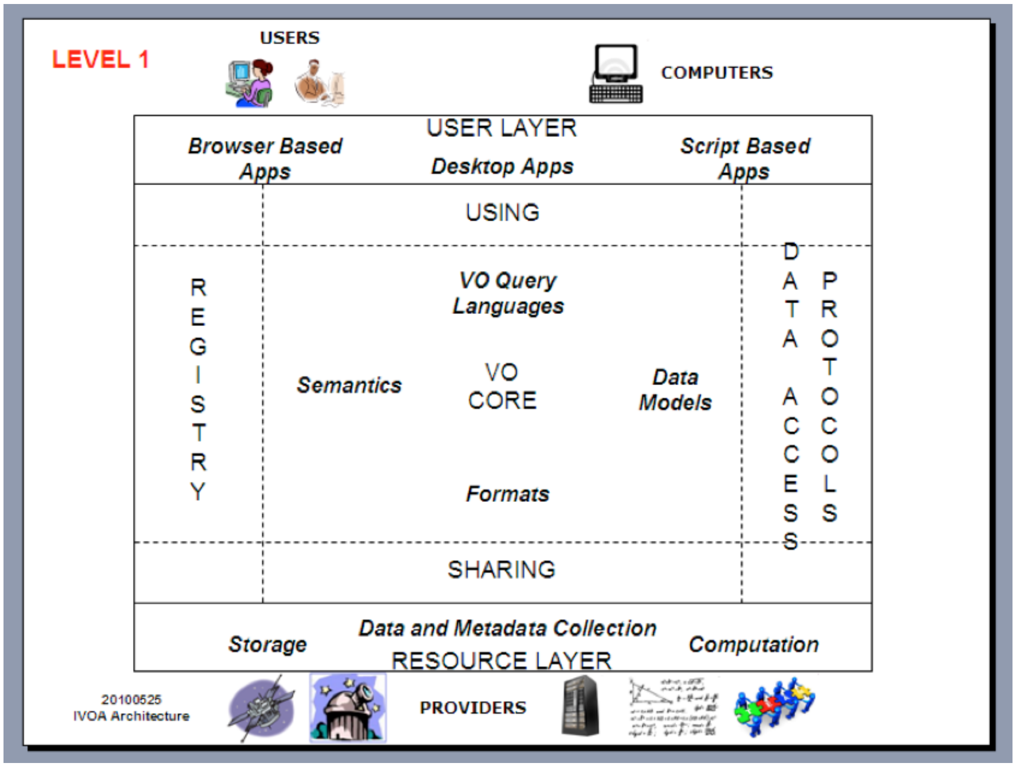


Figure 3. IVOA general architecture as seen in 'IVOA 2010 Technical assessment and roadmap' [Arv10]

GRID middleware is used in the VO for high-performance computing, data storage and transfer and authentication. This technology implements VOSpace, a service which enable users to store data within the VO. VOSpace stores files and intermediate database tables issued from linked operations on services. It avoids copying results to the user desktop and also establishes rights on access and allows the user to remotely manage data storage. Development of standards for more sophisticated services is expected, e.g. image processing and source detection, spectral analysis, visualization of complex datasets and data mining for catalogues.

3.2 Basic Concepts

Below we summarize basic concepts associated to the Virtual Observatory, relevant for the here presented work.

VOTables

The VOTable [OW09] format is an XML standard for the interchange of data represented as a set of tables. It is the standard response to any VO query that provides astronomical data of any kind and/or its complete description and characterization based on standard VO data models. Some examples of VOTables are provided in Appendix A and Appendix B.

Data Models

A data model is a logical model detailing the decomposition of a complex dataset into simpler elements, specifying the semantics for each element and the relationships among them, the metadata used for their description, and the concepts upon which the data model is based. Specifying a data model helps greatly to document the structure and the meaning of the data, allowing software to work correctly, since data is understood at a fundamental level.

UTypes

UTypes tags are used to provide uniform means to identify the elements of a data model in any language or environment. UTypes come from the need of flattening the hierarchy of a data model into a VOTable, so that all fields are represented by fixed strings in a flat name-space allowing, a wide variety of software to manipulate or use the model.

UCDs

The UCDs (Unified Content Descriptors) [PDD+07] are a standardized vocabulary used to describe astronomical quantities and related concepts. A UCD identifies the semantic type of a data value or data model element, saying what type of quantity, in a physical sense, is stored in the value. UCDs may be used independently of any data model. Multiple data models may define fields which share the same UCD, or multiple fields of a single data model may share the same UCD. For this reason UCDs cannot be used to uniquely identify data model fields, but they provide a unique capability to identify or associate similar types of fields.

Services, Interfaces, Operations and Protocols

A service operates at a defined service endpoint (e.g., an Internet URL; an URL for a service endpoint is often called a baseURL), and implements one or more service operations, also known as requests, or methods. All these operations define the interface of the service. Because interfaces may change with time they can be versioned. Each operation accepts as input zero or more request parameters. The details of how a client talks to a service interface defines the protocol used to interact with the service.

Synchronous stateless vs. asynchronous operations

A service operation executed synchronously does not usually return a response to the client until the operation is complete. An asynchronous operation returns an immediate response to the client indicating whether or not the request was accepted, with the operation continuing to execute as a background job on the server. A service operation that executes asynchronously performs the same action as a synchronous request, but may take an arbitrarily long time to execute.

Simple Cone Search Protocol

This protocol defines a service for retrieving records from a catalog of astronomical sources. The query describes a sky position and an angular distance, defining a cone on the sky. The response returns a list of astronomical sources from the catalog whose positions lie within the cone, formatted as a VOTable.

Simple Line Access Protocol

This protocol defines a service for retrieving spectral lines from different spectral line databases through a uniform interface, mainly based on querys over a spectral range although extra search parameters can also be implemented to restrict the search.

Simple Spectral Access Protocol

This protocol defines a service to remotely discover and access one-dimensional spectra, including simple aggregations of 1-D spectra.

Simple Image Access Protocol

This protocol defines a service for retrieving image data from a variety of astronomical image repositories through a uniform interface. A typical use consists in defining a rectangular region on the sky in order to query for candidate images. The service returns a list of candidate images formatted as a VOTable. For each candidate image an access reference URL may be used to retrieve the image. Images may be returned in a variety of formats including the most common digital file format in Astronomy (FITS) and various graphics formats. Referenced images are often computed on the fly, e.g., as cutouts from larger images.

Table Access Protocol

This protocol defines a service for accessing tabular data, including astronomical catalogs as well as general database tables. Access is provided for both database and table metadata as well as for actual table data. It also includes support for both synchronous and asynchronous queries. A multi-position query capability permits queries against an arbitrarily large list of astronomical targets, providing a simple spatial cross-matching capability.

Generic and typed interfaces

Generic interfaces can describe any type of astronomical data whereas the typed interfaces can describe only a single type of data, but can do so in greater detail, with a more refined data model specific to the data. While generic interefaces can only describe static files stored in some archive, typed interfaces can describe, generate, and provide access to both static archival datasets as well as virtual data. We call *virtual data* those datasets which may not actually exist when described, but which can be computed on demand if accessed. Since the generic interface can describe any type of data it can also describe relationships between different types of data, for example to associate multiple discrete data products as elements of complex data of some sort.

3.3 VO Software

The VO provides a centralised and uniform access to data without having the users worrying about different formats and access mechanisms, providing all the necessary information, in the form of metadata, for a scientific usage of data collections. There is a variety of VO tools and services allowing for imaging, spectra and catalogue discovery, such as Aladin[5] developed by the Centre de Données Astronomiques de Strasbourg in France, Datascope[6]

[5] http://aladin.u-strasbg.fr/

[6] http://heasarc.gsfc.nasa.gov/cgi-bin/vo/datascope/init.pl

provided by the US-VO, or VODesktop[7] developed by AstroGrid, to name a few. They allow the user to look for all the reduced data available in all the VO-compliant archives and data repositories (or registries) for a source, browse through them, select those of interest and visualise in various forms. Most of the tools allow either for single-object multi-resource query or for a multi-object single-resource query.

One of the most advanced VO tools to date for imaging analysis is Aladin. It is an interactive sky atlas that allows the user to search, browse through and visualise images retrieved from VO resources, perform astrometric and photometric calibration on them, superpose catalogues and instrument's fields-of-view, create colour-composite images and many more operations. It also includes a scripting capability, which enables multi-object (by target name or coordinates) multi-resource search. Concerning spectral analysis, the VOSpec[8], developed by ESA-VO, allows the user to directly query all the available registries that contain ground-based and space spectral data and/or load local files. Apart from display capabilities, it also provides various high level analysis module,s such as line measurement and fitting, or line identification.

For tabular data handling TOPCAT[9] is a powerful tool that provides a wide variety of data visualisation functionalities, such as histogrammes, scatter plots, 3D and density plots, table and column metadata visualisation and editing (e.g. operation among columns, coordinate transformations), table format conversions, and cross-correlation capabilities. It makes use of VO standards in order to query VO-compliant resources, therefore allowing the user to combine data residing on a local disk with data available through the VO. The development of SAMP [TBF+09], the Simple Application Messaging Protocol, allows applications to communicate with each other, building thus a complete VO environment in the investigator' desktop for efficient multiwavelength datasets analysis, crossmatching and comparisons.

3.4 AMIGA VO Activities

The AMIGA group has widely contributed to the VO with technical implementations of VO-compliant radio archives, as well as with developments of standards for radioastronomical data models (RADAMS) [San09]. The RADAMS data model has been implemented in the developments of several radio archives: the DSS-63 VO Archive[10] and TAPAS (Telescope Archive for Public Access System) IRAM 30m VO Archive[11]. Two screenshots of different VO software displaying data from VOTables provided by TAPAS VO services can be seen in Fig. 4 and Fig. 5. The developments of these two VO-compliant radio archives and the query interfaces have been performed in the context of a collaboration between the AMIGA group and both the Institut de Radio Astronomie Millimétrique (IRAM) in Sierra Nevada and the Spanish Virtual Observatory (SVO).

[7] http://www.astrogrid.org/wiki/Help/IntroVODesktop
[8] http://www.sciops.esa.int/index.php?project=ESAVO&page=vospec
[9] http://www.star.bris.ac.uk/~mbt/topcat/
[10] http://sdc.laeff.inta.es/robledo/index.jsp
[11] https://mrt-lx3.iram.es/tapas/

Since intensive and complex analysis of 3D data is needed inside the AMIGA group and given the experience acquired by the group in radio-VO developments, a collaboration has started with the Kapteyn Institute for upgrading the GIPSY software in order to produce a friendly VO-compliant package for high-level analysis of datacubes, which will fill the gap found at present where no 3D analisys and modeling software exist in the VO framework. It is in the context of this development that the here presented study has been performed.

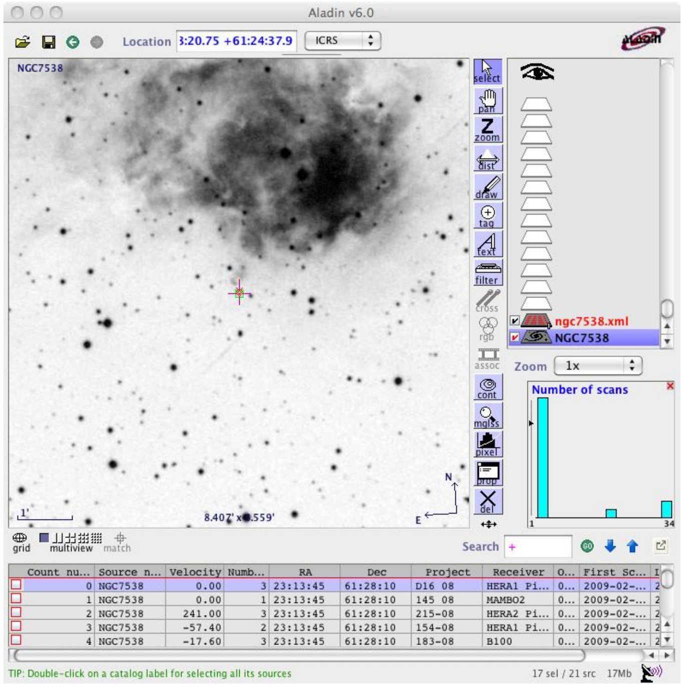


Figure 4. TAPAS VOTable data superimposed over an optical image of NGC7538 in an Aladin window.

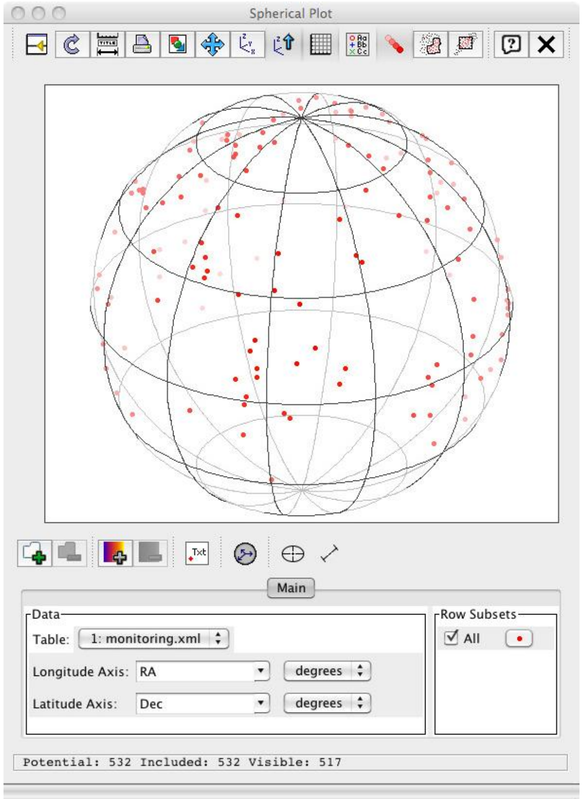


Figure 5. Topcat window displaying all the observations related to one IRAM-30m monitoring project (from a TAPAS VOTable)

4 Multidimensional Data Discovery

In this section we show how the sampling in the datasets axes determines the process to follow for discovery and access different kinds of multidimensional datasets. Some succesful attempts to deal with multidimensional data have been made for small collection of fiber-bundled spectra and SSA services. Looking at the developments and discussions that are taking place inside the DAL Working Group, we think that a new strategy has to be defined for the discovery and access of most of the multidimensional datasets, mainly based on second generation DAL interfaces and generic dataset discovery.

4.1 Uniformly Sampled Multidimensional Datasets

At this moment some efforts inside the VO community are being steered toward the achievement of standards for a second generation of Data Access Layer (DAL) protocols. These protocols will form a family of service interfaces sharing most of their query parameters, response metadata and methods [DTDS08], providing uniformity for both clients and servers in the access and deployment of VO data repositories. This **unification** of DAL interfaces requires a unified generic data model that describes not only the list of available data products fulfilling the search criteria but also each specific dataset to be accessed.

It is in this context that general multidimensional complex datasets have found support in a second version of the Simple Image Access protocol SIAv2 [BDM+09], where the specifications are currently being defined inside the DAL Working Group. It is important to note that the datasets supported by SIAv2 must be sampled in a uniform consistent way in all the axes, so that they can only be irregularly sampled as long as they can be described by a world coordinate system (**WCS**). Any other dataset not properly described with an associated WCS will be considered out of the scope of what SIAv2 can deal with. For example, continuum images at different irregularly spaced frequencies or time series of 2D images taken at different time intervals can be seen as uniformly sampled datasets or as simple associations of data depending on the existence of a related WCS which describes them.

There exist standard mechanisms that allow the inclusion of eventual metadata describing very specific aspects of complex datasets not foreseen in the data model, as well as the method and parameters that will be used for accessing each specific dataset. The DAL query response may use the metadata extension RESOURCE mechanism [BDM+09] in order to give any additional information. Likewise, the REF-ID association mechanisms provided by VOTable format definition [OW09] allow the description of relationships among datasets.

The strategy for managing datasets and archives proposed by DAL architecture breaks down into two different parts: **discovery** of generic datasets and an ensuing **access** to a specific kind of dataset. These two different operations need one generic and one or more typed DAL interfaces. The first one provides global data discovery and description, including general metadata to associate related data products and additional detailed descriptive metadata, giving information about the available specific access methods and the parameters needed for retrieval of each data product. The second one is provided by typed access interfaces, used to retrieve a single given dataset, e.g. 1D spectrum, 2D image, supporting much more specific query parameters than the generic dataset discovery interface.

The multidimensional nature of these datasets enables different ways to get information from them, e.g. a radiointerferometric data cube can be viewed as a collection of images/channels or as a series of spectra uniformly sampled in the spatial plane. This fact takes us into a generic dataset approach where the response to a discovery request may be composed by multiple associated elements, described by generic dataset metadata, many of them generated in the very moment of a subsequent access request.

At present, most of the current first generation interfaces handle specific access to datasets by providing links to static data products in their native format. But several issues arise for complex multidimensional datasets access:

- Original data products can be very large, worsening transfer rates and latency.
- Client applications do not support all native observatory-dependent formats.
- Users are often not interested in the whole product but in a smaller derived portion.

These three points can be solved if **virtual data** are generated on-the-fly for any given specific access method providing data extractions, flagging, filtering, cut-outs and other transformations on the original data product. Because of the size and complexity of 3D data, processing and analysis on the server should be considered, although users should also have the option to download the whole data for local analysis. Depending on the access request method, the generation of virtual data can be quite trivial or be really complex and computationally expensive. Most of these operations need precise WCS metadata and other additional information, which must be found in the generic dataset query response and based on the generic characterization data model.

Virtual data access in DAL architecture can be seen as a command to the service to generate a derived product, rather than a query to get a static file as it is now the case for most of the present data access interfaces. The second generation of DAL interfaces for spectra and images must provide standard methods to deliver derived virtual data from a larger multidimensional dataset. For example, a radio interferometric data cube can be accessed with a SSA (Simple Spectral Access) interface in order to extract a single 1D spectrum from a given pair of spatial coordinates and aperture size, or with a SIA (Simple Image Access) interface in order to extract a 2D plane for a given frequency or a 2D projection along the spectral axis.

The DAL architecture also gives the possibility to provide asynchronous operations so to compute one or more virtual datasets off line and stage them for later retrieval via sign-on authentication. These operations are usually called GRID capabilities [TBDS08] since most of them refer to services deployed by means of a GRID infrastructure, although many of the virtual data generation server-side mechanisms may also make use of a computing and storage distributed infrastructure. Access to derived 2D image data as velocity maps or velocity dispersion maps can be provided with SIAv2 interfaces as long as the needed query parameters are given and the operations are supported by the data provider.

4.2 Associated Data Collections

Multi-fiber spectra datasets such as IFU and MOS, as well as OTF single-dish observations can be treated as collections of associated calibrated 1D spectra and therefore described and accessed by first generation SSA [TD08] interfaces. The presence of different measuring apertures, possibly irregularly spaced, instead of a regular grid of pixels favors this choice where the related spectra can be grouped and described in the VOTable response as a single 3D observation by the means of the REF-ID association mechanism. A successful attempt for the implementation of 3D spectroscopic VO archives based on SSA interfaces and the standard Characterization data model has been carried out by Chilingarian et al [CBL+08].

Unlike uniformly sampled multidimensional datasets, 3D spectroscopic data cubes may be presented in the VOTable response as collections of their basic constituting elements. In this case a typed SSA interface can be used to query a VO compliant spectroscopic data repository in order to retrieve a list of observations fulfilling some given criteria. The VOTable response will describe the list of available spectra where many of them may be deeply related forming a single 3D spectroscopic observation.

The present Characterization [LRB+08] and Spectral [McD07] data models enable a complete description of every one of these spectra including their precise spatial coordinates. The association of the related elements comprising a single observation is provided by means of the REF-ID association mechanisms. Present data models allow for more detailed information such as the 3D observation footprint or instrument specific information and provenance, which may be added making use of the metadata extension RESOURCE mechanism.

Access to each one of these spectra is provided by links to the static files in their original native format, which can be very different depending on the instrument or the used reduction package. This is partially due to the fact that most commonly used software do not share a common data structure for 3D data, but also because different instruments produce different types of data, e.g. one spectrum per spaxel or a series of 2D long-slit spectra. One of the most commonly used formats for IFU datasets is the Euro3D format, where the data do not necessarily fit into a regularly spaced grid.

4.3 The Generic Dataset

Unlike the typed interfaces for image, spectrum, etc., the generic dataset describes any type of dataset found in global data discovery search and can be also used to describe complex data associations by logically associating multiple individual typed datasets. A complex observation may consist of several related data products that can be described via the generic dataset query mechanism. For example, a generic dataset might include one or more spectral data cubes in different spectral bands, some 2-D projections or continuum images of the same field, extracted spectra or object catalogs, and so forth.

We might have a survey field consisting of a spectral data cube, some 2-D projections of the cube, a source catalog for the field computed from the 2-D continuum, and possibly some extracted spectra of objects in the field. Client applications that do not understand the complex data association could still be used to access and analyze the individual primary datasets. Data objects such as extracted spectra or projections could be either precomputed or computed on demand as virtual data. If the client application is sufficiently knowledgeable of the data it could use the instrumental or collection-specific metadata provided to perform a more detailed analysis of the data.

5 Generic Dataset Service

Since we are considering complex multidimensional datasets as those covering spatial, time, frequency, polarization and flux axes, we are approaching the goal of the fully general observation. In this context the DAL interfaces class hierarchy [DTDS08], where a generic dataset is at the root and more specific and typed datasets are represented as subclasses, is the starting point for the definition of a strategy to discover, describe and access multidimensional data.

The current DAL interfaces use an HTTP GET-based interface to submit parameterized requests, with responses being returned as structured documents in VOTables. The discovery method defined in this section can be accessed through a *queryData* value in the REQUEST input parameter. We can define another operation which describes the capabilities of the service through a *getCapabilities* value in the REQUEST input parameter.

Although a Generic Dataset Protocol has not yet been defined in the DAL working group, the query parameters for the discovery operation should not be very different from those proposed in the present work. They are provided in §5.1 where essentially query parameters for existing SSA and SIA protocols have been merged, and others added where needed, in order to retrieve all kind of data in a single query.

The VOTable response describes the list of available static archival datasets matching the generic search criteria, taking into account any complex data associations. The available data, as well the as the single datasets and their associations may be described implementing a generic data model, which has been built considering the existing Characterization, Spectra and Observational data models and classes.

In the following, mandatory parameters are indicated by MAN, recommended parameters by REC, and optional parameters by OPT. A table schema with UTypes and UCD is given §5.2 and an example of a VOTable serialization of the response of a generic dataset service can be found in Appendix A for a discovery queryData method and in Appendix B for a self-descriptive getCapabilities method.

5.1 Input Parameters

Name	Req	Value	Type	Unit	Description
REQUEST	**MAN**	**queryData/getData/getCapabilities**	**string**		**The request or operation name**
VERSION	OPT	1.0	string		The version number of the interface
POS	**MAN**	**52,-27.8;GALACTIC**	**string**	**degrees**	**POS defaults to right-ascension and declination in decimal degrees in the ICRS coordinate system. A coordinate system reference frame may optionally be specified to specify a coordinate system other than ICRS.**
SIZE	**MAN**	**0.05[,0.03]**	**double**	**degrees**	**The width and height of the rectangular region of interest. If only a single value is given it applies to both the width and height of the search region, otherwise the two values may be specified separately.**
BAND	**MAN**	**2.7E-7/0.13;source**	**string**	**meters**	**The spectral band pass is specified in range-list format either numerically as a wavelength value or range, or textually as a spectral band pass identifier. If a band pass is specified as a string identifier it is assumed to be a band pass identifier such as a standard VO band pass name ("radio", "U", "V", "B", "R", "I", "J", etc..)**
TIME	**MAN**	**1998-05-21/1999**	**string**	**ISO 8601 UTC**	**The temporal coverage (epoch of observation) is specified as a single value or range in ISO8601 format.**
POL	**MAN**	**any/none/circular/stokes/circular**	**string**		**Specifies whether or not data is desired which measures polarization, and if so the type of polarization desired.**
FORMAT	**MAN**		**string**		**The FORMAT parameter specifies the allowable data formats for a retrieved file.**
REGION	OPT	Circle ICRS 148.9 69.1 2.0	string		In the case of services which return archival images (whole images) the REGION parameter may be used to specify the spatial region to be searched more precisely than can be done with POS, SIZE
INTERSECT	OPT	COVERS/ENCLOSED/CENTER/OVERLAPS	string		A parameter that indicates how matched images should intersect the region of interest
APERTURE	OPT	0.00028	double	degrees	The aperture parameter is only used for spectral extraction; a spectral extraction SSA service must support this parameter
SPECRES/RP	REC	2.500	double		The minimum spectral resolution, specified as the spectral resolving power /d in dimensionless units
SPATRES	REC	0.05	double	degrees	The minimum spatial resolution (corresponding to the PSF of the observed signal) specified in decimal degrees
TIMERES	OPT	31536000	double	seconds	The minimum time resolution, time coverage of the exposure
FLUXLIMIT	OPT	0.005	double		The minimum measurable flux in the image
SNR	OPT	5.0	double		The minimum signal-to-noise ratio of a candidate dataset, e.g. specified as the ratio of the mean signal to the RMS noise of the background
REDSHIFT	OPT	1.3 / 3.0	string		A photometric (observed) redshift range specified as a single element open or closed range-list
VARAMPL	OPT	0.77	double		The acceptable range of variability amplitude, specified as a single element open or closed range-list, with values in the range 0.0 to 1.0
TARGETNAME	OPT	NGC 4589	string		The target name, suitable for input to a name resolver.
TARGETCLASS	OPT	star, galaxy, pulsar,..	string		Types of astronomical objects to be searched for.
FLUXCALIB	OPT	absolute/relative/any	string		Specifies the minimum level of flux calibration for acceptable data
WAVECALIB	OPT	absolute/relative/any	string		Specifies the minimum level of spectral coordinate calibration for acceptable data

ASTCALIB	OPT	absolute/relative/any	string		Specifies the minimum level of astrometric (spatial) calibration for acceptable data
PUBID	REC	ADS/col#R5983	string		The IVOA publisher's dataset identifier
CREATORID	REC	ivo://auth/col#R124	string		The IVOA dataset identifier assigned at creation time
COLLECTION	REC	SDSS-R5	string		The IVOA resource identifier for a collection
TYPE	REC	Spectra/Cube/Image/Longslit	string		Type of dataset
TOP	REC	9	int		Number of returned top ranked records
MAXREC	REC	5 000	int		Maximum number of records to be returned
MTIME	REC	2005-01-01/2006-01-01	string	ISO 8601 UTC	Only datasets modified, created or deleted at a given range
COMPRESS	REC	true	boolean		If flag is present datasets returned via getData method may be returned in compressed form
RUNID	REC		string		Hidden variable to link several processes and services

FORMAT PARAM VALUES

application/x-votable+xml
application/fits
application/xml
text/csv
image/jpeg
text/html

5.2 Query Response

Name	Req	UType	UCD	Type	Size	Unit	Description
		Query					**Query Metadata**
Score	REC	**gds:**Query.Score		float			Degree of match to the query parameters
		Association					**Association Metadata**
AssocType	OPT	**gds:**Association.Type		char	*		Type of association
AssocID	OPT	**gds:**Association.ID		char	*		Association identifier
AssocKey	OPT	**gds:**Association.Key		char	*		Key used to distinguish association elements
		Access					**Access Metadata**
AcRef	**MAN**	**gds:Access.Reference**	**meta.ref.url**	**char**	*		**URL used to access dataset**
Format	**MAN**	**gds:Access.Format**		**char**	*		**Content or MIME type of dataset**
DatasetSize	REC	**gds:**Access.Size		long		kB	Estimated dataset size
Parameters	OPT	**gds:**Access.Parameters		string	*		Parameters for getData operation
		DataSet					**General Dataset Metadata**
DataModel	**MAN**	**gds:DataSet.DataModel**		**char**	*		**Data model name and version**
DatasetType	**MAN**	**gds:DataSet.Type**		**char**	*		**Type of dataset**
DataLength	**MAN**	**gds:DataSet.Length**	**meta.number**	**long**			**Number of pixels in image/cube**
TimeSI	OPT	gds:DataSet.TimeSI	time;arith.zp	char	*		SI factor and dimensions
SpectralSI	OPT	gds:DataSet.SpectralSI		char	*		SI factor and dimensions
FluxSI	OPT	gds:DataSet.FluxSI		char	*		SI factor and dimensions
Dataset.SpatialAxis		gds:DataSet.SpatialAxis		char	*		Spatial axes column names (native data)
Dataset.SpectralAxis	OPT	gds:DataSet.SpectralAxis		char	*		Spectral axis column name (native data)
Dataset.FluxAxis	OPT	gds:DataSet.FluxAxis		char	*		Flux axis column name (native data)
Dataset.TimeAxis		gds:DataSet.TimeAxis		char	*		Time axis column name (native data)
Dataset.PolarizationAxis		gds:DataSet.PolarizationAxis		char	*		Polarization axis column name (native data)
Deleted	OPT	gds:DataSet. Deleted		char	*		Set to deletion time, if dataset is deleted
		DataID					**Dataset Identification Metadata**
Title	**MAN**	**gds:DataID.Title**	**meta.title;meta.dataset**	**char**	*		**Dataset Title**
Creator	REC	gds:DataID.Creator		char	*		Dataset creator

Collection	REC	gds:DataID.Collection		char	*		Data collection to which dataset belongs
DatasetID	OPT	gds:DataID.DatasetID	meta.id;meta.dataset	char	*		IVOA Dataset ID
CreatorDID	REC	gds:DataID.CreatorDID	meta.id	char	*		Creator's ID for the dataset
CreatorDate	OPT	gds:DataID.Date	time;meta.dataset	char	*		Data processing/creation date
CreatorVersion	OPT	gds:DataID.Version	meta.version;meta.dataset	char	*		Version of dataset
CreationType	REC	gds:DataID.CreationType		char	*		Dataset creation type
		Provenance					**Provenance Metadata**
Instrument	OPT	gds:Provenance.ObsConfig.Instrument	meta.id;instr	char	*		Instrument name
Bandpass	OPT	gds:Provenance.ObsConfig.Bandpass	instr.bandpass	char	*		Band as in RSM Coverage.Spectral
DataSource	REC	gds:Provenance.ObsConfig.DataSource		char	*		Original source of the data
BeamMajorAxix		gds:Provenance.ObsConfig.BeamMajorAxis	phys.angSize;instr.beam	double		deg	Beam Major Axis
BeamMinorAxis		gds:Provenance.ObsConfig.BeamMinorAxis	phys.angSize;instr.beam	double		deg	Beam Mminor Axis
BeamPositionAngle		gds:Provenance.ObsConfig.BeamPositionAngle	pos.posAng;instr.beam	double		deg	Beam Position Angle
		Curation					**Curation Metadata**
Publisher	**MAN**	**gds:Curation.Publisher**	**meta.curation**	**char**	*		**Dataset publisher**
Reference	REC	gds:Curation.Reference	meta.bib.bibcode	char	*		URL or Bibcode for documentation
PublisherID	OPT	gds:Curation.PublisherID	meta.ref.url;meta.curation	char	*		URI for VO Publisher
PublisherDID	REC	gds:Curation.PublisherDID	meta.ref.url;meta.curation	char	*		Publisher's ID for the dataset ID
PublisherDate	OPT	gds:Curation.Date	time; meta.curation	char	*		Date curated dataset last modified
PublisherVersion	OPT	gds:Curation.Version	meta.version;meta.curation	char	*		Publisher's version of the dataset
Rights	OPT	gds:Curation.Rights		char	*		Restrictions on data access
		Target					**Target Metadata**
TargetName	OPT	gds:Target.Name	meta.id;src	char	*		Target name
TargetDescription	OPT	gds:Target.Description	meta.note;src	char	*		Target description
TargetClass	OPT	gds:Target.Class	src.class	char	*		Object class of observed target
TargetPos	OPT	gds:Target.Pos	pos.eq;src	double	2	deg	Target RA and Dec
SpectralClass	OPT	gds:Target.SpectralClass	src.spType	char	*		Object spectral class
Velocity		gds:Target.Velocity	phys.veloc	double			Target velocity
Redshift	OPT	gds:Target.Redshift	src.redshift	double			Target redshift
VarAmpl	OPT	gds:Target.VarAmpl	src.var.amplitude	float			Target variability amplitude (typical)
		Derived					**Derived Metadata**
DerivedSNR	OPT	gds:Derived.SNR	stat.snr	float			Signal-to-noise for spectrum
DerivedVelocity		gds:Derived.Velocity.Value	phys.veloc	double			Derived mean velocity from spectrum

VelocityStatError		gds:Derived.Velocity.StatError	stat.error;phys.veloc	float			Error on derived velocity
VelocityConfidence		gds:Derived.Velocity.Confidence		float			Confidence value on velocity
DerivedRedshift	OPT	gds:Derived.Redshift.Value	src.redshift	double			Measured redshift for spectrum
RedshiftStatError	OPT	gds:Derived.Redshift.StatError	stat.error;src.redshift	float			Error on measured redshift
RedshiftConfidence	OPT	gds:Derived.Redshift.Confidence		float			Confidence value on redshift
DerivedVarAmpl	OPT	gds:Derived.VarAmpl	src.var.amplitude;arith.ratio	float			Variability amplitude as fraction of mean
VarAmplStatError		gds:Derived.VarAmpl.StatError	stat.error;src.var.amplitude	float			Error on variability amplitude
VarAmplConfidence		gds:Derived.VarAmpl.Confidence		float			Confidence value on variability amplitude
		CoordSys					**Coordinate System Metadata**
CoordSysID	OPT	gds:CoordSys.ID		char	*		ID string for coordinate system
SpaceFrameName	REC	gds:CoordSys.SpaceFrame.Name		char	*		Spatial coordinate frame name
SpaceFrameUcd	OPT	gds:CoordSys.SpaceFrame.Ucd		char	*		Space frame UCD
SpaceFrameRefPos	OPT	gds:CoordSys.SpaceFrame.RefPos		char	*		Origin of SpaceFrame
SpaceFrameEquinox	OPT	gds:CoordSys.SpaceFrame.Equinox	time.equinox;pos.frame	double		y	Equinox
TimeFrameName	OPT	gds:CoordSys.TimeFrame.Name	time.scale	char	*		Timescale
TimeFrameUcd	OPT	gds:CoordSys.TimeFrame.Ucd		char	*		Time frame UCD
TimeFrameRefPos	OPT	gds:CoordSys.TimeFrame.RefPos	time.scale	char	*		Location for times of photon arrival
TimeFrameZero	OPT	gds:CoordSys.TimeFrame.Zero	time;arith.zp	double		d	Zero point of timescale in MJD
SpectralFrameName	OPT	gds:CoordSys.SpectralFrame.Name		char	*		Spectral frame name
SpectralFrameUcd	OPT	gds:CoordSys.SpectralFrame.Ucd		char	*		Spectral frame UCD
SpectralFrameRefPos	OPT	gds:CoordSys.SpectralFrame.RefPos		char	*		Spectral frame origin
SpectralFrameRedshift	OPT	gds:CoordSys.SpectralFrame.Redshift		double			Redshift value used if restframe corrected
RedshiftFrameName	OPT	gds:CoordSys.RedshiftFrame.Name		char	*		Redshift frame name
RedshiftFrameUcd		gds:CoordSys.RedshiftFrame.Ucd		char	*		Redshit frame UCD
RedshiftFrameRefPos	OPT	gds:CoordSys.RedshiftFrame.RefPos		char	*		Redshift frame origin
DopplerDefinition	OPT	gds:CoordSys.RedshiftFrame.DopplerDefinition		char	*		Type of redshift
		Char.SpatialAxis					**Spatial Axis Characterization**
SpatialAxisName	OPT	gds:Char.SpatialAxis.AxisName	meta.id	char	*		Name for spatial axis
SpatialAxisUcd	REC	gds:Char.SpatialAxis.Ucd	meta.ucd	char	*		UCD for spatial coordinate
SpatialAxisUnit	OPT	gds:Char.SpatialAxis.Unit	meta.unit	char	*		Unit for spatial coordinate

SpatialLocation	**MAN**	**gds:Char.SpatialAxis.Coverage.Location.Coord**	**pos.eq**	**double**	**2**	**deg**	**Spatial Position**
SpatialExtent	**MAN**	**gds:Char.SpatialAxis.Coverage.Bounds.Extent**	**instr.fov**	**double**	**2**	**deg**	**Angular area in square degrees**
SpatialStart	OPT	gds:Char.SpatialAxis.Coverage.Bounds.Limits.Start	pos.eq	double	2		Minimum value for coordinates
SpatialStop	OPT	gds:Char.SpatialAxis.Coverage.Bounds.Limits.Stop	pos.eq	double	2		Maxmum value for coordinates
SpatialArea	OPT	gds:Char.SpatialAxis.Coverage.Support.Area		double		deg	Filed of view or aperture region for spectra
SpatialAreaType	OPT	gds:Char.SpatialAxis.Coverage.Support.AreaType		char	*		Geometry of region
SpatialSampleExtent	OPT	gds:Char.SpatialAxis.SamplingPrecision.SamplingPrecisionRefVal	instr.pixel	float			Spatial bin size
SpatialFillFactor	OPT	gds:Char.SpatialAxis.SamplingPrecision.SamplingPrecisionRefVal.FillFactor	stat.fill;pos.eq	float			Spatial sampling filling factor
SpatialStatError	REC	gds:Char.SpatialAxis.Accuracy.StatError	stat.error;pos.eq	double		deg	Astrometric statistical error
SpatialSysError	OPT	gds:Char.SpatialAxis.Accuracy.SysError	stat.error.sys;pos.eq	double		deg	Astrometric systematic error
SpatialCalibration	REC	gds:Char.SpatialAxis.CalibrationStatus	meta.code.qual	char	*		Type of spatial coord calibration
SpatialResolution	REC	gds:Char.SpatialAxis.Resolution	pos.angResolution	double		deg	Spatial resolution of data
		Char.SpectralAxis					**Spectral Axis Characterization**
SpectralAxisName	OPT	gds:Char.SpectralAxis.AxisName	meta.id	char	*		Name for spectral axis
SpectralAxisUcd	REC	gds:Char.SpectralAxis.Ucd	meta.ucd	char	*		UCD for spectral coordinate
SpectralAxisUnit	OPT	gds:Char.SpectralAxis.Unit	meta.unit	char	*		Unit for spectral coordinate
SpectralLocation	**MAN**	**gds:Char.SpectralAxis.Coverage.Location.Coord**	**instr.bandpass**	**double**		**m**	**Spectral coord value**
SpectralExtent	**MAN**	**gds:Char.SpectralAxis.Coverage.Bounds.Extent**	**instr.bandwidth**	**double**		**m**	**Width of spectrum**
SpectralStart	REC	gds:Char.SpectralAxis.Coverage.Bounds.Limits.Start	em;stat.min	double		m	Start in spectral coordinate
SpectralStop	REC	gds:Char.SpectralAxis.Coverage.Bounds.Limits.Stop	em;stat.max	double		m	Stop in spectral coordinate
SpectralSampleExtent	OPT	gds:Char.SpectralAxis.SamplingPrecision.SamplingPrecisionRefVal	em;spect.binSize	double		m	Wavelength bin size
SpectralFillFactor	OPT	gds:Char.SpectralAxis.SamplingPrecision.SamplingPrecisionRefVal.FillFactor	stat.fill;em	float			Spectral sampling filling factor
SpectralStatError	OPT	gds:Char.SpectralAxis.Accuracy.StatError	stat.error;em	double		m	Spectral coord statistical error
SpectralSysError	OPT	gds:Char.SpectralAxis.Accuracy.SysError	stat.error.sys;em	double		m	Spectral coord systematic error
SpectralCalibration	REC	gds:Char.SpectralAxis.CalibrationStatus	meta.code.qual	char	*		Type of spectral coord calibration
SpectralResolution	REC	gds:Char.SpectralAxis.Resolution	spect.resolution;em	double		m	Spectral resolution FWHM
SpectralResPower	OPT	gds:Char.SpectralAxis.ResPower	spect.resolution	float			Spectral resolving power
		Char.TimeAxis					**Time Axis Characterization**
TimeAxisName	OPT	gds:Char.TimeAxis.AxisName	meta.id	char	*		Name for time axis
TimeAxisUcd	REC	gds:Char.TimeAxis.Ucd	meta.ucd	char	*		UCD for time coordinate

TimeAxisUnit	OPT	gds:Char.TimeAxis.Unit	meta.unit	char	*		Unit for time coordinate
TimeLocation	**MAN**	**gds:Char.TimeAxis.Coverage.Location.Coord**	**time.epoch**	**double**		**d**	**Midpoint of exposure on MJD scale**
TimeExtent	REC	gds:Char.TimeAxis.Coverage.Bounds.Extent	time.duration	double		s	Total exposure time
TimeStart	OPT	gds:Char.TimeAxis.Coverage.Bounds.Start	time.start;obs.exposure	double			Start time
TimeStop	OPT	gds:Char.TimeAxis.Coverage.Bounds.Stop	time.stop;obs.exposure	double			Stop time
TimeSupportExtent	OPT	gds:Char.TimeAxis.Coverage.Support.Extent	time.duration;obs.exposure	double		s	Effective exposure time
TimeSampleExtent	OPT	gds:Char.TimeAxis.SamplingPrecision.SamplingPrecisionRefVal	time.interval	double		s	Time bin size
TimeFillFactor	OPT	gds:Char.TimeAxis.SamplingPrecision.SamplingPrecisionRefVal.FillFactor	time;stat.fill;time	float			Time sampling filling factor
TimeStatError	OPT	gds:Char.TimeAxis.Accuracy.StatError	stat.error;time	double		s	Time coord statistical error
TimeSysError	OPT	gds:Char.TimeAxis.Accuracy.SysError	stat.error.sys;time	double		s	Time coord systematic error
TimeCalibration	OPT	gds:Char.TimeAxis.CalibrationStatus	meta.code.qual	char	*		Type of coord calibration
TimeResolution	OPT	gds:Char.TimeAxis.Resolution	time.resolution	double		s	Time resolution
		Char.FluxAxis					**Flux Axis Characterization**
FluxAxisName	OPT	gds:Char.FluxAxis.AxisName	meta.id	char	*		Name for flux axis
FluxAxisUcd	REC	gds:Char.FluxAxis.Ucd	meta.ucd	char	*		UCD for flux coordinate
FluxAxisUnit	OPT	gds:Char.FluxAxis.Unit	meta.unit	char	*		Unit for flux coordinate
FluxAverage		gds:Char.FluxAxis.Coverage.Bounds.Location.Coord	phot.flux.density	double			Average flux density
FluxMin		gds:Char.FluxAxis.Coverage.Bounds.Start	phot.flux.density;stat.min	double			Limiting value for flux density
FluxSaturation		gds:Char.FluxAxis.Coverage.Bounds.Stop	phot.flux.density;stat.max	double			Saturation value for flux density
FluxSupportExtent		gds:Char.FluxAxis.Coverage.Support.Extent	phot.flux.density	double	2		Minimum, maximun value for flux density
FluxStatError	OPT	gds:Char.FluxAxis.Accuracy.StatError	stat.error;phot.flux.density;em	double			Flux statistical error
FluxSysError	OPT	gds:Char.FluxAxis.Accuracy.SysError	stat.error.sys;phot.flux.density;em	double			Flux systematic error
FluxCalibration	REC	gds:Char.FluxAxis.CalibrationStatus		char	*		Type of flux calibration
FluxResolution		gds:Char.FluxAxis.Resolution	phot.flux.density;stat.stdev	double			Flux density standard deviation
		Char.PolarizationAxis					**Polarization Axis Characterization**
PolarizationAxisName	OPT	gds:Char.PolarizationAxis.AxisName	meta.id	char	*		Name for polarization axis
PolarizationAxisUcd	REC	gds:Char.PolarizationAxis.Ucd	meta.ucd	char	*		UCD for polarization
PolarizationValues	OPT	gds:Char.PolarizationAxis.Enumeration		char	*		List of available polarization parameters

6 Generic Dataset Service Implementation

The discovery method for a Generic Dataset Service proposed in §5 has been implemented on a VO archive of datacubes of galaxies (B0DEGA). The AMIGA software group has developed a multidimensional data VO archive for observations of a sample of galaxies made with SMA radio interferometer. The archive has been built based on a slightly modified Characterization data model [LRB+08] that accounts for the metadata needed for the implementation of a Generic Dataset Service.

6.1 The B0DEGA Project

B0DEGA[12] (Below 0 Degree Galaxies) is a legacy project of the Submillimeter Array (SMA), a radio interferometer telescope. This project aims to address the molecular gas properties in the circumnuclear regions (1 arcmin) of the largest survey of nearby galaxies been performed to date. It is a collaboration with members of the IAA-CSIC, CfA (Harvard-Smithsonian Center for Astrophysics, USA) and ASIAA (Institute of Academia Sinica Astronomy and Astrophysics, Taiwan), and is led by Daniel Espada (IAA-CSIC, CfA), member of the AMIGA group. The sample of B0DEGA is composed of spiral galaxies that have undergone recent interaction, many of them characterized by central starbursts.

6.2 The B0DEGA Generic Dataset Service

The B0DEGA datasets are processed datacubes of observations of galaxies for various molecular lines. Because of their multidimensionality the B0DEGA datacubes represent an excellent base for the development of a VO archive that can be used to implement the proposed Generic Dataset Service discovery method. Fig. 6 shows the relational database serialization schema for the proposed generic datamodel, which has been implemented on a PostgreSQL DB engine. The B0DEGA archive has been built using the Django Python framework and the PgSphere libraries for PostgreSQL in order to provide celestial coordinates transformations and efficient indexing for positional queries. This development has been performed inside by the AMIGA group technical team (Jose Enrique Ruiz, Victor Espigares and Susana Sánchez).

Fig. 7 shows a web form interface to query the archive with the most representative input parameters proposed for the generic dataset service in §5. The response provided depends evidently on the search criteria; a list of datacubes delivered by the service is shown in Fig. 8. Since the generic dataset model allows the description of both a list of datasets and also independent entities, a more detailed description of a single dataset can be offered as seen in Fig. 9. Access to some associated data products is also provided as links to static files (channel map of a spectral datacube, moments as integrated flux and velocity maps, spectral profiles and radial distribution of the intensity). These derived data products should be

[12] http://b0dega.iaa.es

accessed by means of standardized data access methods in future implementations of a second generation DAL services.

Te B0DEGA archive web interface offers the possibility to access the whole datacube with a link to the reduced FITS file of the observation, but it also provides a way to visualize the cube without the need of installing any new software in the desktop user. Fig 10 shows the Aladin applet movie window displaying a B0DEGA archive datacube. This functionality can be accessed by means of a single mouse click in the web interface

B0DEGA GDS discovery method is not constrained to a web interface and can be also accessed via any other VO software. Fig. 11 shows a Topcat window displaying data values issued from a B0DEGA GDS VOTable response. In Fig 12 and Fig 13 we can see a Topcat window displaying columns metadata and general descriptive parameters of the service for a B0DEGA GDS VOTable response.

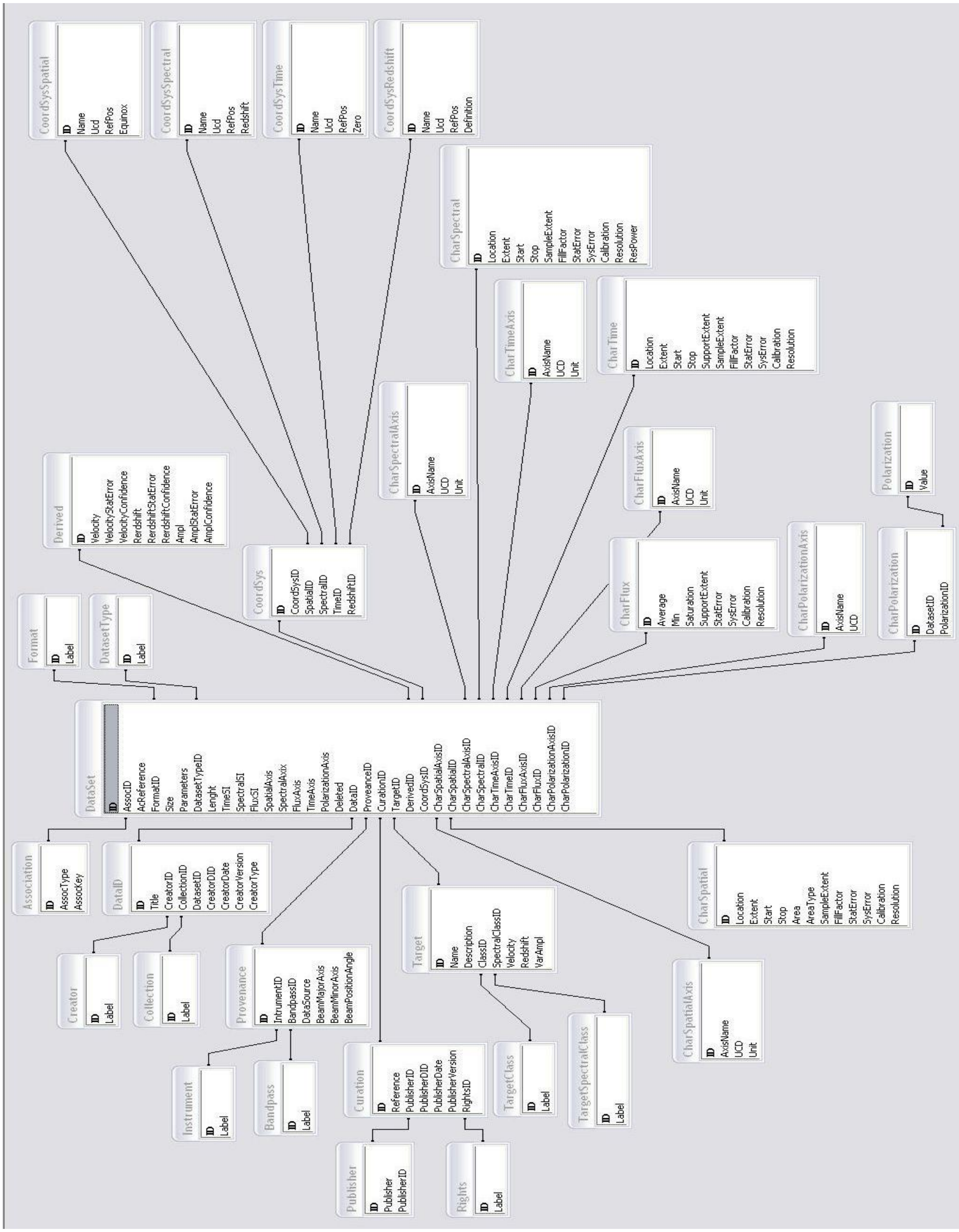


Figure 6. Relational database serialization schema for a GDS datamodel

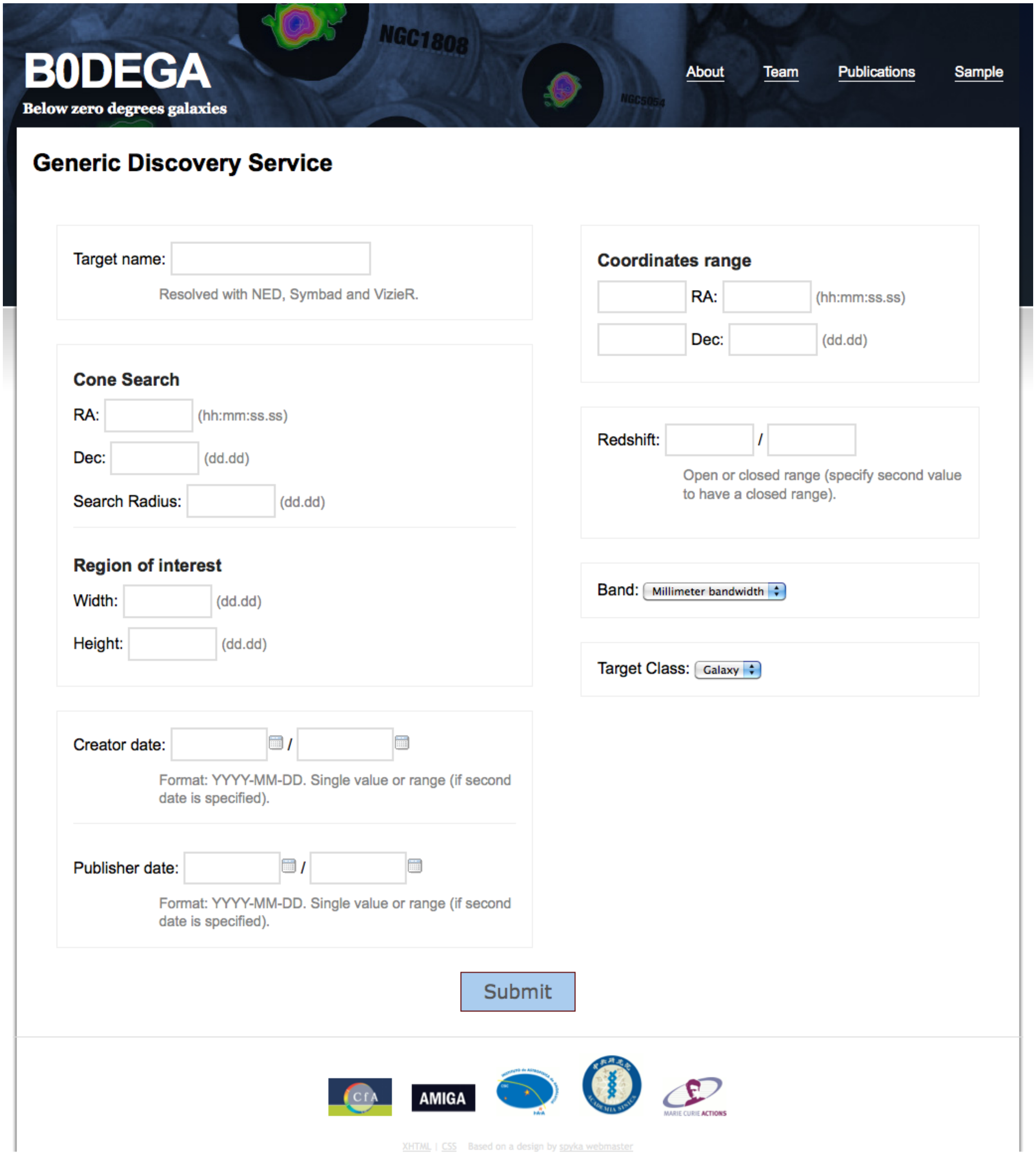
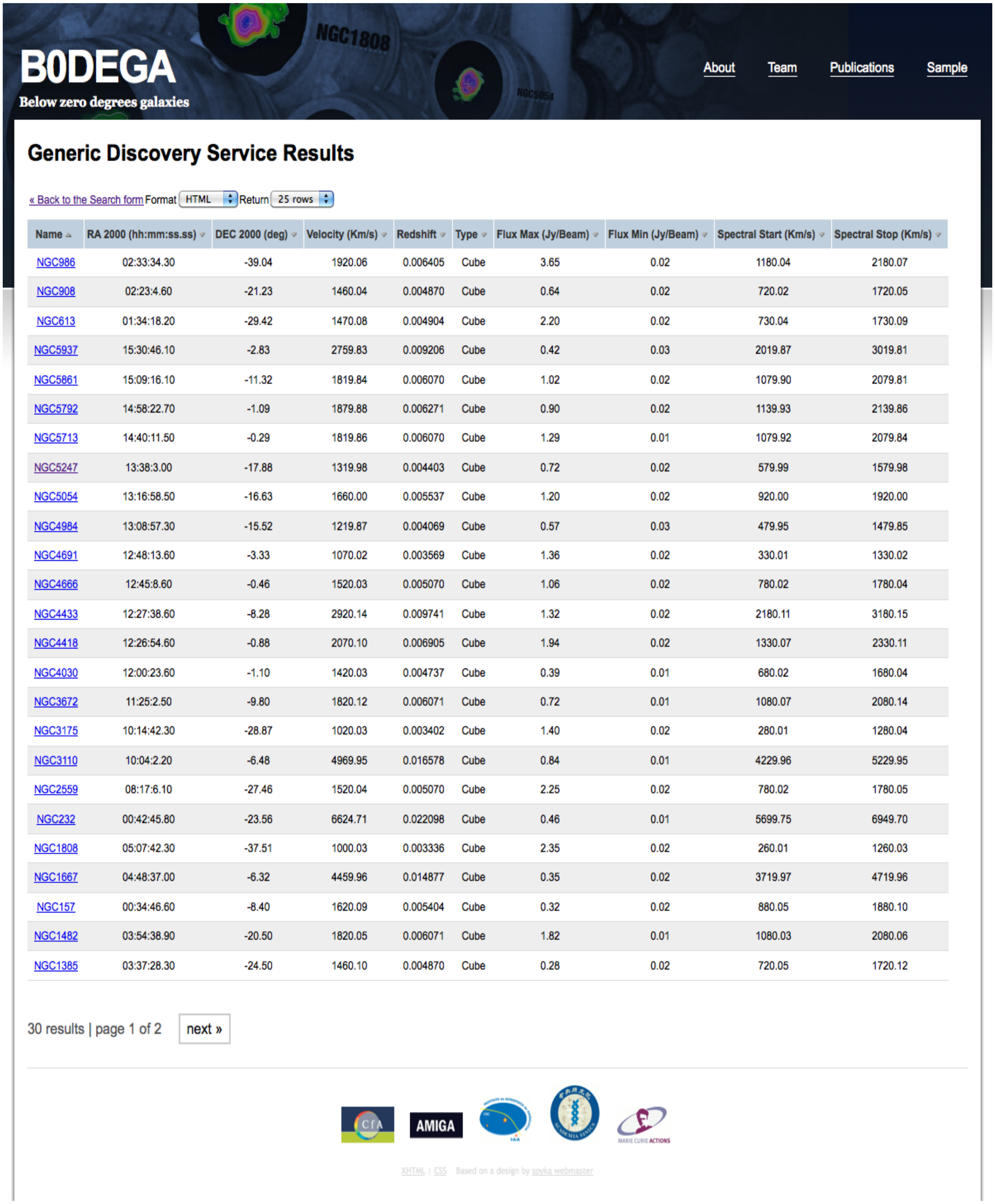


Figure 7. B0DEGA archive GDS discovery web form interface



Generic Discovery Service Results

« Back to the Search form Format HTML Return 25 rows

Name	RA 2000 (hh:mm:ss.ss)	DEC 2000 (deg)	Velocity (Km/s)	Redshift	Type	Flux Max (Jy/Beam)	Flux Min (Jy/Beam)	Spectral Start (Km/s)	Spectral Stop (Km/s)
NGC986	02:33:34.30	-39.04	1920.06	0.006405	Cube	3.65	0.02	1180.04	2180.07
NGC908	02:23:4.60	-21.23	1460.04	0.004870	Cube	0.64	0.02	720.02	1720.05
NGC613	01:34:18.20	-29.42	1470.08	0.004904	Cube	2.20	0.02	730.04	1730.09
NGC5937	15:30:46.10	-2.83	2759.83	0.009206	Cube	0.42	0.03	2019.87	3019.81
NGC5861	15:09:16.10	-11.32	1819.84	0.006070	Cube	1.02	0.02	1079.90	2079.81
NGC5792	14:58:22.70	-1.09	1879.88	0.006271	Cube	0.90	0.02	1139.93	2139.86
NGC5713	14:40:11.50	-0.29	1819.86	0.006070	Cube	1.29	0.01	1079.92	2079.84
NGC5247	13:38:3.00	-17.88	1319.98	0.004403	Cube	0.72	0.02	579.99	1579.98
NGC5054	13:16:58.50	-16.63	1660.00	0.005537	Cube	1.20	0.02	920.00	1920.00
NGC4984	13:08:57.30	-15.52	1219.87	0.004069	Cube	0.57	0.03	479.95	1479.85
NGC4691	12:48:13.60	-3.33	1070.02	0.003569	Cube	1.36	0.02	330.01	1330.02
NGC4666	12:45:8.60	-0.46	1520.03	0.005070	Cube	1.06	0.02	780.02	1780.04
NGC4433	12:27:38.60	-8.28	2920.14	0.009741	Cube	1.32	0.02	2180.11	3180.15
NGC4418	12:26:54.60	-0.88	2070.10	0.006905	Cube	1.94	0.02	1330.07	2330.11
NGC4030	12:00:23.60	-1.10	1420.03	0.004737	Cube	0.39	0.01	680.02	1680.04
NGC3672	11:25:2.50	-9.80	1820.12	0.006071	Cube	0.72	0.01	1080.07	2080.14
NGC3175	10:14:42.30	-28.87	1020.03	0.003402	Cube	1.40	0.02	280.01	1280.04
NGC3110	10:04:2.20	-6.48	4969.95	0.016578	Cube	0.84	0.01	4229.96	5229.95
NGC2559	08:17:6.10	-27.46	1520.04	0.005070	Cube	2.25	0.02	780.02	1780.05
NGC232	00:42:45.80	-23.56	6624.71	0.022098	Cube	0.46	0.01	5699.75	6949.70
NGC1808	05:07:42.30	-37.51	1000.03	0.003336	Cube	2.35	0.02	260.01	1260.03
NGC1667	04:48:37.00	-6.32	4459.96	0.014877	Cube	0.35	0.02	3719.97	4719.96
NGC157	00:34:46.60	-8.40	1620.09	0.005404	Cube	0.32	0.02	880.05	1880.10
NGC1482	03:54:38.90	-20.50	1820.05	0.006071	Cube	1.82	0.01	1080.03	2080.06
NGC1385	03:37:28.30	-24.50	1460.10	0.004870	Cube	0.28	0.02	720.05	1720.12

30 results | page 1 of 2 next »

AMIGA

XHTML | CSS Based on a design by spyka webmaster

Figure 8. B0DEGA archive GDS discovery web interface response

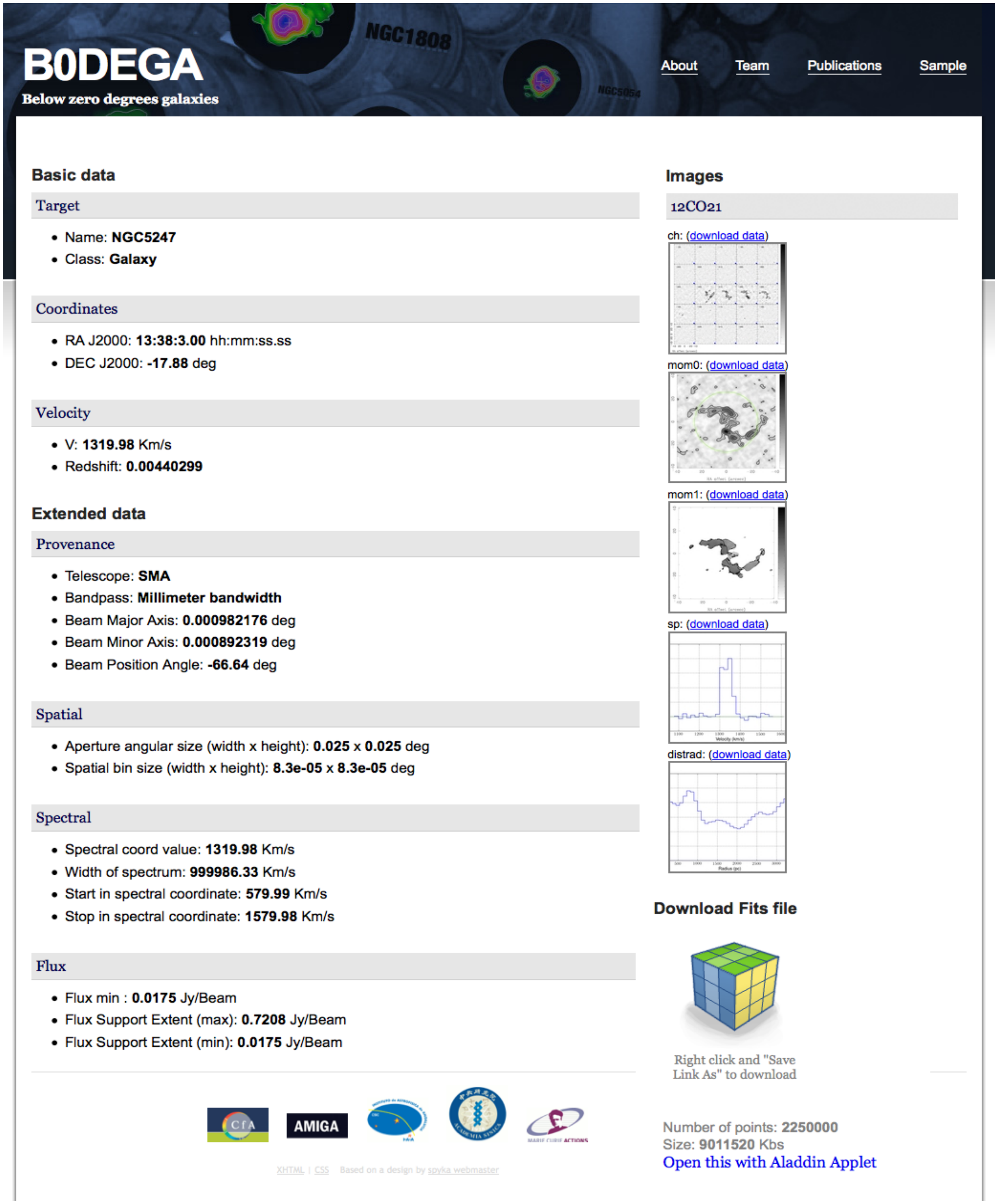


Figure 9. Metadata provided by the GDS discovery web interface for a single dataset in the B0DEGA archive

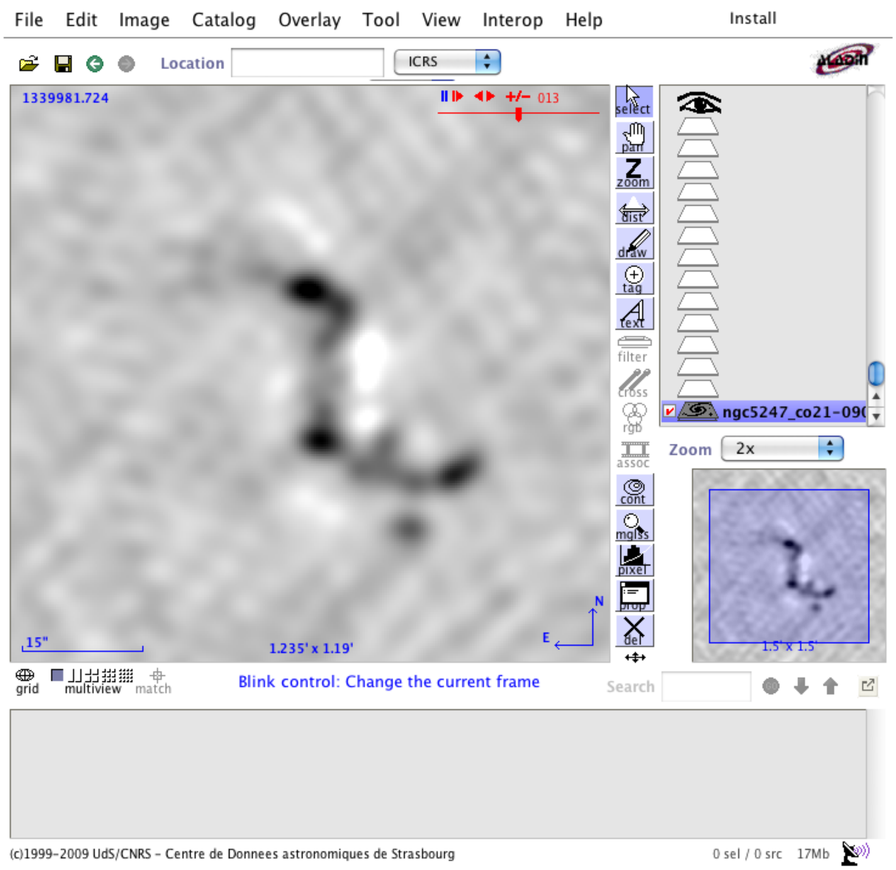
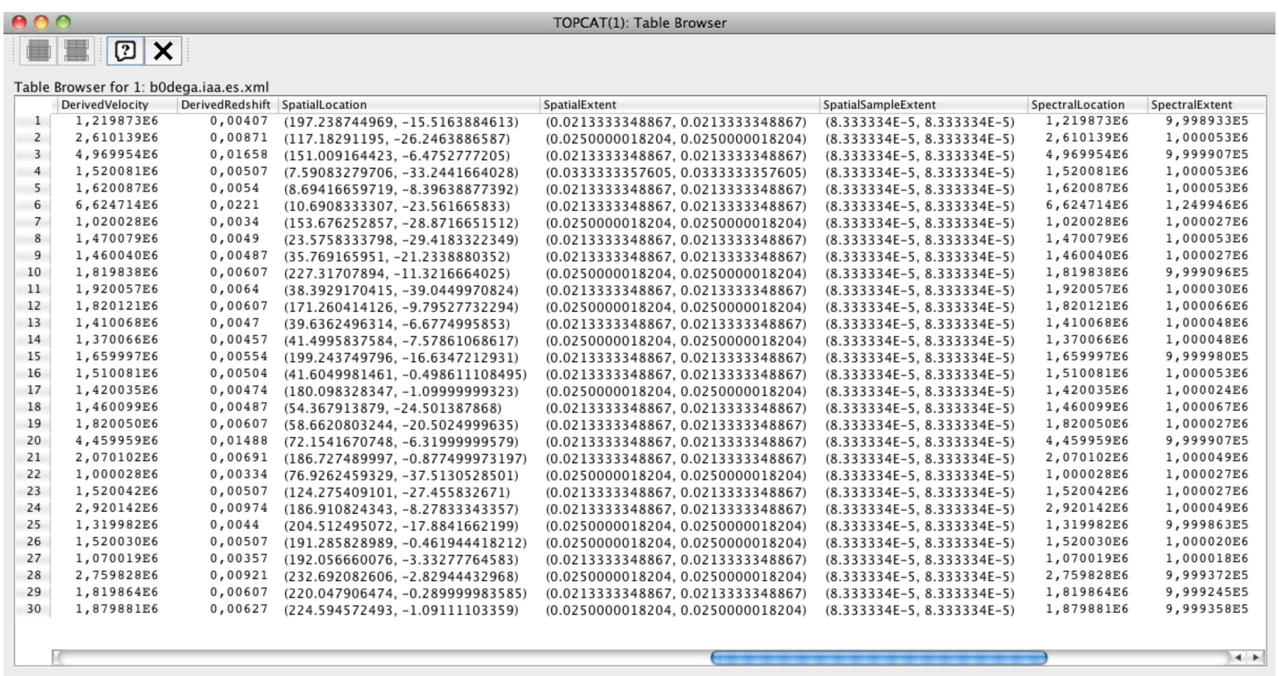


Figure 10. Aladin applet movie window displaying a datacube in the B0DEGA archive

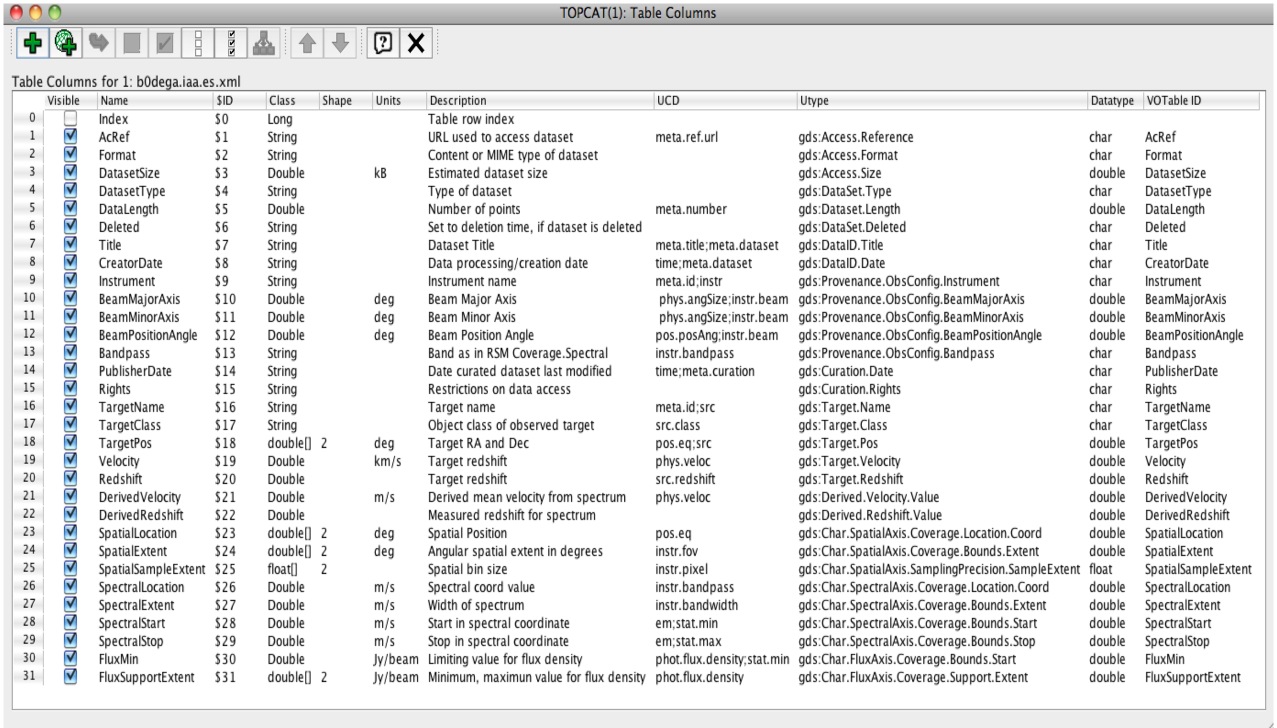


TOPCAT(1): Table Browser

Table Browser for 1: b0dega.iaa.es.xml

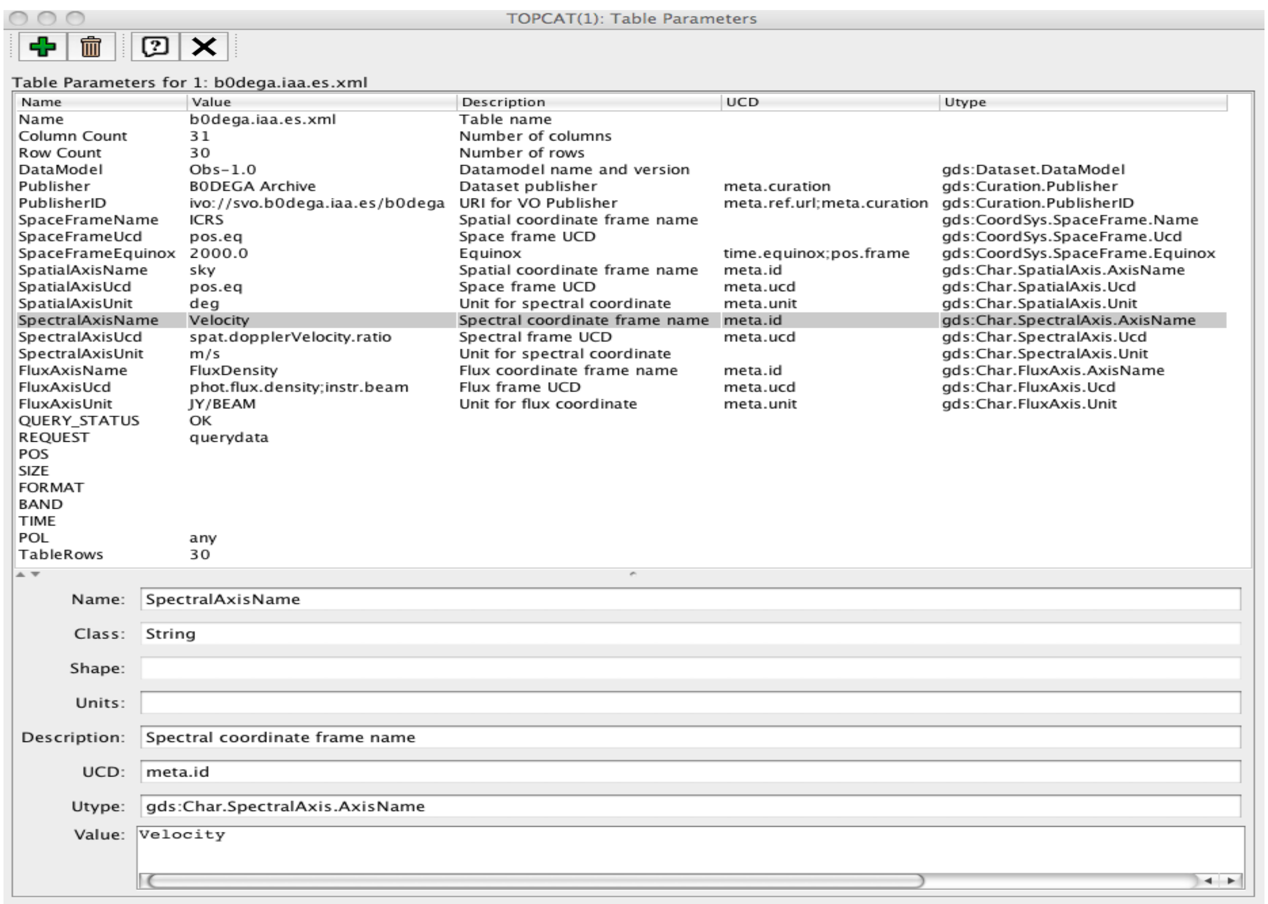
	DerivedVelocity	DerivedRedshift	SpatialLocation	SpatialExtent	SpatialSampleExtent	SpectralLocation	SpectralExtent
1	1,219873E6	0,00407	(197.238744969, -15.5163884613)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,219873E6	9,998933E5
2	2,610139E6	0,00871	(117.18291195, -26.2463886587)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	2,610139E6	1,000053E6
3	4,969954E6	0,01658	(151.009164423, -6.4752777205)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	4,969954E6	9,999907E5
4	1,520081E6	0,00507	(7.59083279706, -33.2441664028)	(0.0333333357605, 0.0333333357605)	(8.333334E-5, 8.333334E-5)	1,520081E6	1,000053E6
5	1,620087E6	0,0054	(8.69416659719, -8.39638877392)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,620087E6	1,000053E6
6	6,624714E6	0,0221	(10.6908333307, -23.561665833)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	6,624714E6	1,249946E6
7	1,020028E6	0,0034	(153.676252857, -28.8716651512)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,020028E6	1,000027E6
8	1,470079E6	0,0049	(23.5758333798, -29.4183322349)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,470079E6	1,000053E6
9	1,460040E6	0,00487	(35.769165951, -21.2338880352)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,460040E6	1,000027E6
10	1,819838E6	0,00607	(227.31707894, -11.3216664025)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,819838E6	9,999096E5
11	1,920057E6	0,0064	(38.3929170415, -39.0449970824)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,920057E6	1,000030E6
12	1,820121E6	0,00607	(171.260414126, -9.79527732294)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,820121E6	1,000066E6
13	1,410068E6	0,0047	(39.6362496314, -6.6774995853)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,410068E6	1,000048E6
14	1,370066E6	0,00457	(41.4995837584, -7.57861068617)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,370066E6	1,000048E6
15	1,659997E6	0,00554	(199.243749796, -16.6347212931)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,659997E6	9,999980E5
16	1,510081E6	0,00504	(41.6049981461, -0.498611108495)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,510081E6	1,000053E6
17	1,420035E6	0,00474	(180.098328347, -1.09999999323)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,420035E6	1,000024E6
18	1,460099E6	0,00487	(54.367913879, -24.501387868)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,460099E6	1,000067E6
19	1,820050E6	0,00607	(58.6620803244, -20.5024999635)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,820050E6	1,000027E6
20	4,459959E6	0,01488	(72.1541670748, -6.31999999579)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	4,459959E6	9,999907E5
21	2,070102E6	0,00691	(186.727489997, -0.877499973197)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	2,070102E6	1,000049E6
22	1,000028E6	0,00334	(76.9262459329, -37.5130528501)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,000028E6	1,000027E6
23	1,520042E6	0,00507	(124.275409101, -27.455832671)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,520042E6	1,000027E6
24	2,920142E6	0,00974	(186.910824343, -8.27833343357)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	2,920142E6	1,000049E6
25	1,319982E6	0,0044	(204.512495072, -17.8841662199)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,319982E6	9,999863E5
26	1,520030E6	0,00507	(191.285828989, -0.461944418212)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,520030E6	1,000020E6
27	1,070019E6	0,00357	(192.056660076, -3.33277764583)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,070019E6	1,000018E6
28	2,759828E6	0,00921	(232.692082606, -2.82944432968)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	2,759828E6	9,999372E5
29	1,819864E6	0,00607	(220.047906474, -0.289999983585)	(0.0213333348867, 0.0213333348867)	(8.333334E-5, 8.333334E-5)	1,819864E6	9,999245E5
30	1,879881E6	0,00627	(224.594572493, -1.09111103359)	(0.0250000018204, 0.0250000018204)	(8.333334E-5, 8.333334E-5)	1,879881E6	9,999358E5

Figure 11. Topcat window displaying data values for a B0DEGA GDS VOTable response.



	Visible	Name	$ID	Class	Shape	Units	Description	UCD	Utype	Datatype	VOTable ID
0		Index	$0	Long			Table row index				
1	☑	AcRef	$1	String			URL used to access dataset	meta.ref.url	gds:Access.Reference	char	AcRef
2	☑	Format	$2	String			Content or MIME type of dataset		gds:Access.Format	char	Format
3	☑	DatasetSize	$3	Double		kB	Estimated dataset size		gds:Access.Size	double	DatasetSize
4	☑	DatasetType	$4	String			Type of dataset		gds:DataSet.Type	char	DatasetType
5	☑	DataLength	$5	Double			Number of points	meta.number	gds:Dataset.Length	double	DataLength
6	☑	Deleted	$6	String			Set to deletion time, if dataset is deleted		gds:DataSet.Deleted	char	Deleted
7	☑	Title	$7	String			Dataset Title	meta.title;meta.dataset	gds:DataID.Title	char	Title
8	☑	CreatorDate	$8	String			Data processing/creation date	time;meta.dataset	gds:DataID.Date	char	CreatorDate
9	☑	Instrument	$9	String			Instrument name	meta.id;instr	gds:Provenance.ObsConfig.Instrument	char	Instrument
10	☑	BeamMajorAxis	$10	Double		deg	Beam Major Axis	phys.angSize;instr.beam	gds:Provenance.ObsConfig.BeamMajorAxis	double	BeamMajorAxis
11	☑	BeamMinorAxis	$11	Double		deg	Beam Minor Axis	phys.angSize;instr.beam	gds:Provenance.ObsConfig.BeamMinorAxis	double	BeamMinorAxis
12	☑	BeamPositionAngle	$12	Double		deg	Beam Position Angle	pos.posAng;instr.beam	gds:Provenance.ObsConfig.BeamPositionAngle	double	BeamPositionAngle
13	☑	Bandpass	$13	String			Band as in RSM Coverage.Spectral	instr.bandpass	gds:Provenance.ObsConfig.Bandpass	char	Bandpass
14	☑	PublisherDate	$14	String			Date curated dataset last modified	time;meta.curation	gds:Curation.Date	char	PublisherDate
15	☑	Rights	$15	String			Restrictions on data access		gds:Curation.Rights	char	Rights
16	☑	TargetName	$16	String			Target name	meta.id;src	gds:Target.Name	char	TargetName
17	☑	TargetClass	$17	String			Object class of observed target	src.class	gds:Target.Class	char	TargetClass
18	☑	TargetPos	$18	double[]	2	deg	Target RA and Dec	pos.eq;src	gds:Target.Pos	double	TargetPos
19	☑	Velocity	$19	Double		km/s	Target redshift	phys.veloc	gds:Target.Velocity	double	Velocity
20	☑	Redshift	$20	Double			Target redshift	src.redshift	gds:Target.Redshift	double	Redshift
21	☑	DerivedVelocity	$21	Double		m/s	Derived mean velocity from spectrum	phys.veloc	gds:Derived.Velocity.Value	double	DerivedVelocity
22	☑	DerivedRedshift	$22	Double			Measured redshift for spectrum		gds:Derived.Redshift.Value	double	DerivedRedshift
23	☑	SpatialLocation	$23	double[]	2	deg	Spatial Position	pos.eq	gds:Char.SpatialAxis.Coverage.Location.Coord	double	SpatialLocation
24	☑	SpatialExtent	$24	double[]	2	deg	Angular spatial extent in degrees	instr.fov	gds:Char.SpatialAxis.Coverage.Bounds.Extent	double	SpatialExtent
25	☑	SpatialSampleExtent	$25	float[]	2		Spatial bin size	instr.pixel	gds:Char.SpatialAxis.SamplingPrecision.SampleExtent	float	SpatialSampleExtent
26	☑	SpectralLocation	$26	Double		m/s	Spectral coord value	instr.bandpass	gds:Char.SpectralAxis.Coverage.Location.Coord	double	SpectralLocation
27	☑	SpectralExtent	$27	Double		m/s	Width of spectrum	instr.bandwidth	gds:Char.SpectralAxis.Coverage.Bounds.Extent	double	SpectralExtent
28	☑	SpectralStart	$28	Double		m/s	Start in spectral coordinate	em;stat.min	gds:Char.SpectralAxis.Coverage.Bounds.Start	double	SpectralStart
29	☑	SpectralStop	$29	Double		m/s	Stop in spectral coordinate	em;stat.max	gds:Char.SpectralAxis.Coverage.Bounds.Stop	double	SpectralStop
30	☑	FluxMin	$30	Double		Jy/beam	Limiting value for flux density	phot.flux.density;stat.min	gds:Char.FluxAxis.Coverage.Bounds.Start	double	FluxMin
31	☑	FluxSupportExtent	$31	double[]	2	Jy/beam	Minimum, maximun value for flux density	phot.flux.density	gds:Char.FluxAxis.Coverage.Support.Extent	double	FluxSupportExtent

Figure 12. Topcat window displaying columns metadata for a B0DEGA GDS VOTable response



Name	Value	Description	UCD	Utype
Name	b0dega.iaa.es.xml	Table name		
Column Count	31	Number of columns		
Row Count	30	Number of rows		
DataModel	Obs-1.0	Datamodel name and version		gds:Dataset.DataModel
Publisher	B0DEGA Archive	Dataset publisher	meta.curation	gds:Curation.Publisher
PublisherID	ivo://svo.b0dega.iaa.es/b0dega	URI for VO Publisher	meta.ref.url;meta.curation	gds:Curation.PublisherID
SpaceFrameName	ICRS	Spatial coordinate frame name		gds:CoordSys.SpaceFrame.Name
SpaceFrameUcd	pos.eq	Space frame UCD		gds:CoordSys.SpaceFrame.Ucd
SpaceFrameEquinox	2000.0	Equinox	time.equinox;pos.frame	gds:CoordSys.SpaceFrame.Equinox
SpatialAxisName	sky	Spatial coordinate frame name	meta.id	gds:Char.SpatialAxis.AxisName
SpatialAxisUcd	pos.eq	Space frame UCD	meta.ucd	gds:Char.SpatialAxis.Ucd
SpatialAxisUnit	deg	Unit for spectral coordinate	meta.unit	gds:Char.SpatialAxis.Unit
SpectralAxisName	Velocity	Spectral coordinate frame name	meta.id	gds:Char.SpectralAxis.AxisName
SpectralAxisUcd	spat.dopplerVelocity.ratio	Spectral frame UCD	meta.ucd	gds:Char.SpectralAxis.Ucd
SpectralAxisUnit	m/s	Unit for spectral coordinate		gds:Char.SpectralAxis.Unit
FluxAxisName	FluxDensity	Flux coordinate frame name	meta.id	gds:Char.FluxAxis.AxisName
FluxAxisUcd	phot.flux.density;instr.beam	Flux frame UCD	meta.ucd	gds:Char.FluxAxis.Ucd
FluxAxisUnit	JY/BEAM	Unit for flux coordinate	meta.unit	gds:Char.FluxAxis.Unit
QUERY_STATUS	OK			
REQUEST	querydata			
POS				
SIZE				
FORMAT				
BAND				
TIME				
POL	any			
TableRows	30			

Figure 13. Topcat window displaying descriptive parameters for a B0DEGA GDS VOTable response

7 Multidimensional Data Access

The main capabilities of the existing DAL protocols are focused in data discovery, and this is also the method that has been defined and implemented in the previous sections for a Generic Dataset Service. While one might imagine separating discovery from access methods in the case of access to static archival data, this would be difficult or impossible in the case of virtual data. For virtual data, the discovery includes negotiating with the service the details of what the service can actually deliver, given the parameters of the ideal dataset as requested by the client.

As it was discussed in §4 a second generation DAL service, and the SIAv2 in particular, must support access methods that fully deal with the description and generation of virtual data products. These methods should be defined considering the needs of the user through a careful study of common scientific use cases. In the following some eventual access method are presented and use cases are discussed as a guideline for a future definition of generic dataset access methods for a second generation DAL service.

7.1 Data Access Methods

In the following we provide different data access methods that may be supported by a SIAv2 service dealing with virtual datasets generation in order to provide the user with specific views of the same spectral datacube.

Whole dataset
This is the simplest access method but it can be impractical for very large datasets. It is the method to implement to access smaller datasets, and for applications that need to track actual physical datasets in archives, e.g., to index data or to track replicas. The whole dataset access will normally return instrument native format data.

Spectrum extraction
In the case of a spectral data cube, a 1D spectrum is extracted along the spectral axis through a synthetic aperture whose location and size is defined by the client. The details of how the extracted spectrum is computed can be data-dependent and are determined by the service.

Cutout 2D planes
This method allows the extraction of 2D planes from a dataset. The entire 2D plane does not have to be retrieved; a subset of a 2D plane can be cutout if desired, as for a 2D image. A cutout does not resample the data; the original data samples are returned. The image axes are not changed. For many types of analysis resampling degrades the data more than is acceptable and a simple cutout is what is desired. In the case of 3D data, a 2D cutout can be made along any 2 of the 3 axes. Spatial planes cutouts are often needed in order to select the spectral range to perform a cutout 3D sub-cube.

Cutout 3D sub-cube

This is similar to a 2D cutout except that a 3D sub-cube is extracted. This is a common case for true 3D analysis of datacubes where the individual cubes are too large to be practical to retrieve without subsetting, but also when the extreme pixels in all axes are known to have a worse signal-to-noise ration than the rest of the cube. Most of the times a general inspection of the whole datacube is needed to determine the region where a signal has been detected. If several sources are found in distant parts of the datacube it is usually less time-consuming to work with different 3D sub-cubes.

2D projection

In this case a 3D cube is collapsed along one axis to produce a 2D image. This is a form of projection since the pixels are resampled along one axis: collapsing an axis in effect changes the sampling on that axis, causing the sampling bin-width to encompass the full region being collapsed. The effect is similar to looking through the cube face-on along the selected axis. Different functions can be used for the reprojection (averaging, median value, standard deviation, and covariance), and may correspond to the definition of datacubes moments in astronomy.

3D projection

This is the generalization of the 3D cutout. This case produces a 3D sub-cube like the 3D cutout, but the pixels are resampled in 3D. The axes of the extracted sub-cube do not have to be aligned with the original cube, axes can be transposed or flipped, the sampling can be changed along any axis, any axis projections can be changed (such as a 2D sky projection or a velocity axis scaling), and any coordinate systems (e.g., equatorial vs galactic) can be modified as desired. The client is allowed to specify the WCS of the 3D sub-cube to be generated. It is directly analogous to reprojection of a 2D image.

General 2D slices through a 3D cube

This is the generalization of the 2D cutout, and is equivalent to the 3D projection except that a 2D image is produced. The reference point of the 2D image can be anywhere in the 3D space of the cube, and the orientation of the 2D slice can be arbitrary; hence this case represents a general 2D slice through the cube. For example, one might slice the cube along the major axis of a galaxy as determined from a 2D continuum image, with the Y axis of the extracted 2D slice representing velocity. The analogous case for a 2D image is an arbitrary 2D slice to generate a plot of flux versus 2D position. This access mode occurs commonly in 3D cube visualization.

7.2 Scientific Use Cases

In the following we provide two scientific use cases detailing the way to proceed in the discovery and access of multidimensional generic datasets, the services and methods involved and the parameters needed.

Integrated flux

In this case we want to collapse a spectral datacube in the optical range along the spectral axis to produce a 2D image. The cube is not sky-subtracted, and we wish to exclude wavelength regions containing night sky emission or absorption lines.

- Perform a generic dataset discovery query to locate the data and get a description of all the data elements. A simple positional query or a time range can be specified to select a single observation if multiple observations of the same field are available.
- The response is a list of all the individual data elements (datasets) comprising the observation. These include the main spectral datacube plus a 2D continuum image of the same field. The cube provides a 3D SIA cube access method plus a 1D SSA spectrum extraction access method.
- The 2D continuum image is retrieved and displayed or otherwise analyzed by the client to determine one or more sky regions.
- One or more 1D spectra are extracted from the cube via the SSA access method provided. These are analyzed to locate all the night sky lines.
- The 3D SIA cube access method is called to produce the 2D projection, collapsing the cube along the spectral axis. The BANDPASS parameter is passed a range list specifying the wavelength regions (those which do not contain night sky lines) to be used to compute the projection.

Sub-cube extraction

In this case we want to extract a sub-cube from a larger cube, using cutout techniques. The axes of the cube are RA, DEC, and velocity.

- The generic dataset query may or may not be required depending upon the complexity of the data and the analysis to be performed. Either the generic dataset query, or an SIA query, is performed to locate the cube and determine the characteristics and geometry of the cube.
- A query to the 3D SIA cutout access method for the cube is then performed to get an access reference for the sub-cube. The query parameters identify the cube in some fashion, e.g., by its dataset ID, and specify the coverage of the sub-cube in equatorial coordinates for X, Y, and in velocity for Z.
- The query response verifies that the operation is correct and gives the client standard metadata for the image to be returned. The access reference is used to retrieve the cube, which comes back as a FITS file.
- Some pre-existing cube visualization tool is used to view and analyze the sub-cube returned.

7.3 Other considerations

The main needs for spectral analysis on datacubes are single and multiple line fitting, radial velocity and velocity dispersion measurements using observed templates or models, and the measurement of line fluxes, equivalent widths, and line indices. Radial velocity measurements, and stellar template fitting[13], are applications well fit for the VO since the template observations or models could be accessed via the VO.

Until now we have been treating the spectral and spatial domains independently. However, certain operations on 3D data require that they be treated simultaneously. One example is the need to subtract a spectral continuum from a region using an annulus around it. This would involve fitting in both wavelength and position.

Once a 2D map is produced it should be possible to perform all standard imaging analysis on it. This includes calculating statistics, aperture photometry, ellipse or isophote fitting, surface fitting, and background subtraction. In addition, many maps may not represent intensity data but the results of the spectral analysis such as velocity fields or line-ratio maps. These maps may also need quantitative analysis.

More advanced capabilities can also be envisioned. For example, it is likely that an investigator would want to apply new and perhaps proprietary techniques or models to data available through the VO. For example, someone could want to fit kinematical models to velocity fields. This person could download the data to fit to the models, but when datasets are very large it may be more efficient to upload the model or analysis technique to the data on the server and let the powerful server handle the processing. In cases like this the intellectual property of the user would have to be respected and the use of expensive processing resources might need to be negotiated.

[13] In order to spearate the emission line and absorption line spectra

8 Conclusions and Future Work

In the presented work we have studied the state of the art of both multimensional datasets in Astronomy and Data Access Layer Protocols in the VO in order to determine the best strategy for the discovery and access of complex datasets. The solution proposed has been deliberately conceived to solve most of the issues identified in the less possible intrusive way, reusing existing data models and protocols with minor modifications and considering the present efforts in the DAL working group in the design of a family of second generation DAL protocols.

Future work will be oriented towards the conception and development of scientific workflows. The AMIGA group participates in a recently approved STREP project "WF4ever: Advanced Workflow Preservation Technologies for Enhanced Science" under the FP7ICT-2009-6 Program, where a workflow for kinematical modelling of cubes for galaxies will be developed in the existing B0DEGA V0 archive framework, taking advantage of the incipient infrastructure of data and services developed and building new ones based on both upcoming DAL and Workflow Preservation standards.

The quantitative leap in volume and complexity of the next generation of archives will need analysis and data mining tasks to live closer to the data, in computing and distributed storage environments (GRID), but they should also be modular enough to allow customization from scientists and be easily accessible to foster their dissemination among the community.

Appendix A: B0DEGA GDS QueryData VOTable Response

```
<VOTABLE xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="xmlns:http://www.ivoa.net/xml/VOTable/VOTable-1.2.xsd"
xmlns:sia2="http:www.ivoa.net/xml/gds/v0.1" version="1.2">
<RESOURCE type="Results">
<DESCRIPTION>B0DEGA Generic Data Service</DESCRIPTION>
<INFO name="QUERY_STATUS" value="OK"/>
<INFO name="REQUEST" value="querydata"/>
<INFO name="POS" value=""/>
<INFO name="SIZE" value=""/>
<INFO name="FORMAT" value=""/>
<INFO name="BAND" value=""/>
<INFO name="TIME" value="" />
<INFO name="POL" value="any"/>
<INFO name="TableRows" value="30"/>
<TABLE>
    <FIELD ID="AcRef" name="AcRef" datatype="char" utype="gds:Access.Reference" arraysize="*"
ucd="meta.ref.url">
       <DESCRIPTION>URL used to access dataset</DESCRIPTION>
    </FIELD>
    <FIELD ID="Format" name="Format" datatype="char" utype="gds:Access.Format" arraysize="*">
       <DESCRIPTION>Content or MIME type of dataset</DESCRIPTION>
    </FIELD>
    <FIELD ID="DatasetSize" name="DatasetSize" datatype="double" utype="gds:Access.Size"
unit="kB">
       <DESCRIPTION>Estimated dataset size</DESCRIPTION>
    </FIELD>
    <FIELD ID="DatasetType" name="DatasetType" datatype="char" utype="gds:DataSet.Type"
arraysize="*">
       <DESCRIPTION>Type of dataset</DESCRIPTION>
    </FIELD>
    <FIELD ID="DataLength" name="DataLength" datatype="double" utype="gds:Dataset.Length"
ucd="meta.number">
       <DESCRIPTION>Number of points</DESCRIPTION>
    </FIELD>
    <FIELD ID="Deleted" name="Deleted" datatype="char" utype="gds:DataSet.Deleted" arraysize="*">
       <DESCRIPTION>Set to deletion time, if dataset is deleted </DESCRIPTION>
    </FIELD>
    <FIELD ID="Title" name="Title" datatype="char" utype="gds:DataID.Title" arraysize="*"
ucd="meta.title;meta.dataset">
       <DESCRIPTION>Dataset Title</DESCRIPTION>
    </FIELD>
    <FIELD ID="CreatorDate" name="CreatorDate" datatype="char" utype="gds:DataID.Date"
arraysize="*" ucd="time;meta.dataset">
       <DESCRIPTION>Data processing/creation date</DESCRIPTION>
    </FIELD>
    <FIELD ID="Instrument" name="Instrument" datatype="char"
utype="gds:Provenance.ObsConfig.Instrument" arraysize="*" ucd="meta.id;instr">
       <DESCRIPTION>Instrument name</DESCRIPTION>
    </FIELD>
    <FIELD ID="BeamMajorAxis" name="BeamMajorAxis" datatype="double"
utype="gds:Provenance.ObsConfig.BeamMajorAxis " ucd=" phys.angSize;instr.beam" unit="deg">
       <DESCRIPTION>Beam Major Axis</DESCRIPTION>
    </FIELD>
    <FIELD ID="BeamMinorAxis" name="BeamMinorAxis" datatype="double"
utype="gds:Provenance.ObsConfig.BeamMinorAxis " ucd=" phys.angSize;instr.beam" unit="deg">
       <DESCRIPTION>Beam Minor Axis</DESCRIPTION>
    </FIELD>
    <FIELD ID="BeamPositionAngle" name="BeamPositionAngle" datatype="double"
utype="gds:Provenance.ObsConfig.BeamPositionAngle" ucd="pos.posAng;instr.beam" unit="deg">
       <DESCRIPTION>Beam Position Angle</DESCRIPTION>
    </FIELD>
    <FIELD ID="Bandpass" name="Bandpass" datatype="char" utype="gds:Provenance.ObsConfig.Bandpass"
arraysize="*" ucd="instr.bandpass">
       <DESCRIPTION>Band as in RSM Coverage.Spectral</DESCRIPTION>
    </FIELD>
```

```
    <FIELD ID="PublisherDate" name="PublisherDate" datatype="char" utype="gds:Curation.Date"
arraysize="*" ucd="time;meta.curation">
        <DESCRIPTION>Date curated dataset last modified</DESCRIPTION>
    </FIELD>
    <FIELD ID="Rights" name="Rights" datatype="char" utype="gds:Curation.Rights" arraysize="*">
        <DESCRIPTION>Restrictions on data access</DESCRIPTION>
    </FIELD>
    <FIELD ID="TargetName" name="TargetName" datatype="char" utype="gds:Target.Name" arraysize="*"
ucd="meta.id;src">
        <DESCRIPTION>Target name</DESCRIPTION>
    </FIELD>
    <FIELD ID="TargetClass" name="TargetClass" datatype="char" utype="gds:Target.Class"
arraysize="*" ucd="src.class">
        <DESCRIPTION>Object class of observed target</DESCRIPTION>
    </FIELD>
    <FIELD ID="TargetPos" name="TargetPos" datatype="double" utype="gds:Target.Pos" arraysize="2"
ucd="pos.eq;src" unit="deg">
        <DESCRIPTION>Target RA and Dec</DESCRIPTION>
    </FIELD>
    <FIELD ID="Velocity" name="Velocity" datatype="double" utype="gds:Target.Velocity"
ucd="phys.veloc" unit="km/s">
        <DESCRIPTION>Target redshift</DESCRIPTION>
    </FIELD>
    <FIELD ID="Redshift" name="Redshift" datatype="double" utype="gds:Target.Redshift"
ucd="src.redshift">
        <DESCRIPTION>Target redshift</DESCRIPTION>
    </FIELD>
    <FIELD ID="DerivedVelocity" name="DerivedVelocity" datatype="double"
utype="gds:Derived.Velocity.Value" ucd="phys.veloc" unit="m/s">
        <DESCRIPTION>Derived mean velocity from spectrum</DESCRIPTION>
    </FIELD>
    <FIELD ID="DerivedRedshift" name="DerivedRedshift" datatype="double"
utype="gds:Derived.Redshift.Value">
        <DESCRIPTION>Measured redshift for spectrum</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpatialLocation" name="SpatialLocation" datatype="double"
utype="gds:Char.SpatialAxis.Coverage.Location.Coord" arraysize="2" ucd="pos.eq" unit="deg">
        <DESCRIPTION>Spatial Position</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpatialExtent" name="SpatialExtent" datatype="double"
utype="gds:Char.SpatialAxis.Coverage.Bounds.Extent" arraysize="2" ucd="instr.fov" unit="deg">
        <DESCRIPTION>Angular spatial extent in degrees</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpatialSampleExtent" name="SpatialSampleExtent" datatype="float"
utype="gds:Char.SpatialAxis.SamplingPrecision.SampleExtent" arraysize="2" ucd="instr.pixel">
        <DESCRIPTION>Spatial bin size</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpectralLocation" name="SpectralLocation" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Location.Coord" ucd="instr.bandpass" unit="m/s">
        <DESCRIPTION>Spectral coord value</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpectralExtent" name="SpectralExtent" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Extent" ucd="instr.bandwidth" unit="m/s">
        <DESCRIPTION>Width of spectrum</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpectralStart" name="SpectralStart" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Start" ucd="em;stat.min" unit="m/s">
        <DESCRIPTION>Start in spectral coordinate</DESCRIPTION>
    </FIELD>
    <FIELD ID="SpectralStop" name="SpectralStop" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Stop" ucd="em;stat.max" unit="m/s">
        <DESCRIPTION>Stop in spectral coordinate</DESCRIPTION>
    </FIELD>
    <FIELD ID="FluxMin" name="FluxMin" datatype="double"
utype="gds:Char.FluxAxis.Coverage.Bounds.Start" ucd="phot.flux.density;stat.min" unit="Jy/beam">
        <DESCRIPTION>Limiting value for flux density</DESCRIPTION>
    </FIELD>
    <FIELD ID="FluxSupportExtent" name="FluxSupportExtent" datatype="double"
utype="gds:Char.FluxAxis.Coverage.Support.Extent" arraysize="2" ucd="phot.flux.density"
unit="Jy/beam">
        <DESCRIPTION>Minimum, maximun value for flux density</DESCRIPTION>
```

```
    </FIELD>

    <GROUP ID="Access" name="Access" utype="gds:Access">
       <DESCRIPTION>Access Metadata</DESCRIPTION>
       <FIELDref ref="AcRef"/>
       <FIELDref ref="Format"/>
       <FIELDref ref="DatasetSize"/>
    </GROUP>
    <GROUP ID="Dataset" name="Dataset" utype="gds:Dataset">
       <DESCRIPTION>General Dataset Metadata</DESCRIPTION>
       <FIELDref ref="DataLength"/>
       <FIELDref ref="Deleted"/>
       <PARAM ID="DataModel" datatype="char" name="DataModel" utype="gds:Dataset.DataModel"
value="Obs-1.0" arraysize="*">
           <DESCRIPTION>Datamodel name and version</DESCRIPTION>
       </PARAM>
    </GROUP>
    <GROUP ID="DataID" name="DataID" utype="gds:DataID">
       <DESCRIPTION>Dataset Identification Metadata</DESCRIPTION>
       <FIELDref ref="Title"/>
       <FIELDref ref="CreatorDate"/>
    </GROUP>
    <GROUP ID="Provenance" name="Provenance" utype="gds:Proveanance">
       <DESCRIPTION>Provenance Metadata</DESCRIPTION>
       <FIELDref ref="Instrument"/>
       <FIELDref ref="Bandpass"/>
       <FIELDref ref="BeamMajorAxis"/>
       <FIELDref ref="BeamMinorAxis"/>
       <FIELDref ref="BeamPositionAngle"/>
    </GROUP>
    <GROUP ID="Curation" name="Curation" utype="gds:Curation">
       <DESCRIPTION>Curation Metadata</DESCRIPTION>
       <PARAM ID="Publisher" datatype="char" name="Publisher" utype="gds:Curation.Publisher"
value="B0DEGA Archive" arraysize="*" ucd="meta.curation">
           <DESCRIPTION>Dataset publisher</DESCRIPTION>
       </PARAM>
       <PARAM ID="PublisherID" datatype="char" name="PublisherID" utype="gds:Curation.PublisherID"
value="ivo://svo.b0dega.iaa.es/b0dega" arraysize="*" ucd="meta.ref.url;meta.curation">
           <DESCRIPTION>URI for VO Publisher</DESCRIPTION>
       </PARAM>
       <FIELDref ref="PublisherDate"/>
       <FIELDref ref="Rights"/>
    </GROUP>
    <GROUP ID="Target" name="Target" utype="gds:Target">
       <DESCRIPTION>Target Metadata</DESCRIPTION>
       <FIELDref ref="TargetName"/>
       <FIELDref ref="TargetClass"/>
       <FIELDref ref="TargetPos"/>
       <FIELDref ref="Velocity"/>
       <FIELDref ref="Redshift"/>
    </GROUP>
    <GROUP ID="Derived" name="Derived" utype="gds:Derived">
       <DESCRIPTION>Derived Metadata</DESCRIPTION>
       <FIELDref ref="DerivedVelocity"/>
       <FIELDref ref="DerivedRedshift"/>
    </GROUP>
    <GROUP ID="CoordSys" name="CoordSys" utype="gds:CoordSys">
       <DESCRIPTION>Coordinate System Metadata</DESCRIPTION>
       <PARAM ID="SpaceFrameName" datatype="char" name="SpaceFrameName"
utype="gds:CoordSys.SpaceFrame.Name" value="ICRS" arraysize="*">
           <DESCRIPTION>Spatial coordinate frame name</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpaceFrameUcd" datatype="char" name="SpaceFrameUcd"
utype="gds:CoordSys.SpaceFrame.Ucd" value="pos.eq" arraysize="*">
           <DESCRIPTION>Space frame UCD</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpaceFrameEquinox" datatype="double" name="SpaceFrameEquinox"
utype="gds:CoordSys.SpaceFrame.Equinox" value="2000.0" ucd="time.equinox;pos.frame">
           <DESCRIPTION>Equinox</DESCRIPTION>
       </PARAM>
    </GROUP>
```

```
    <GROUP ID="Char.SpatialAxis" name="Char.SpatialAxis" utype="gds:Char.SpatialAxis">
       <DESCRIPTION>Spatial Axis Characterization</DESCRIPTION>
       <PARAM ID="SpatialAxisName" datatype="char" name="SpatialAxisName"
utype="gds:Char.SpatialAxis.AxisName" value="sky" arraysize="*" ucd="meta.id">
           <DESCRIPTION>Spatial coordinate frame name</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpatialAxisUcd" datatype="char" name="SpatialAxisUcd"
utype="gds:Char.SpatialAxis.Ucd" value="pos.eq" arraysize="*" ucd="meta.ucd">
           <DESCRIPTION>Space frame UCD</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpatialAxisUnit" datatype="char" name="SpatialAxisUnit"
utype="gds:Char.SpatialAxis.Unit" value="deg" arraysize="*" ucd="meta.unit">
           <DESCRIPTION>Unit for spectral coordinate</DESCRIPTION>
       </PARAM>
       <FIELDref ref="SpatialLocation"/>
       <FIELDref ref="SpatialExtent"/>
       <FIELDref ref="SpatialSampleExtent"/>
    </GROUP>
    <GROUP ID="Char.SpectralAxis" name="Char.SpectralAxis" utype="gds:Char.SpectralAxis">
       <DESCRIPTION>Spectral Axis Characterization</DESCRIPTION>
       <PARAM ID="SpectralAxisName" datatype="char" name="SpectralAxisName"
utype="gds:Char.SpectralAxis.AxisName" value="Velocity" arraysize="*" ucd="meta.id">
           <DESCRIPTION>Spectral coordinate frame name</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpectralAxisUcd" datatype="char" name="SpectralAxisUcd"
utype="gds:Char.SpectralAxis.Ucd" value="spat.dopplerVelocity.ratio" arraysize="*" ucd="meta.ucd">
           <DESCRIPTION>Spectral frame UCD</DESCRIPTION>
       </PARAM>
       <PARAM ID="SpectralAxisUnit" datatype="char" name="SpectralAxisUnit"
utype="gds:Char.SpectralAxis.Unit" value="m/s" arraysize="*">
           <DESCRIPTION>Unit for spectral coordinate</DESCRIPTION>
       </PARAM>
       <FIELDref ref="SpectralLocation"/>
       <FIELDref ref="SpectralExtent"/>
       <FIELDref ref="SpectralStart"/>
       <FIELDref ref="SpectralStop"/>
    </GROUP>
    <GROUP ID="Char.FluxAxis" name="Char.FluxAxis" utype="gds:Char.FluxAxis">
       <DESCRIPTION>Flux Axis Characterization</DESCRIPTION>
       <PARAM ID="FluxAxisName" datatype="char" name="FluxAxisName"
utype="gds:Char.FluxAxis.AxisName" value="FluxDensity" arraysize="*" ucd="meta.id">
           <DESCRIPTION>Flux coordinate frame name</DESCRIPTION>
       </PARAM>
       <PARAM ID="FluxAxisUcd" datatype="char" name="FluxAxisUcd" utype="gds:Char.FluxAxis.Ucd"
value="phot.flux.density;instr.beam" arraysize="*" ucd="meta.ucd">
           <DESCRIPTION>Flux frame UCD</DESCRIPTION>
       </PARAM>
       <PARAM ID="FluxAxisUnit" datatype="char" name="FluxAxisUnit" utype="gds:Char.FluxAxis.Unit"
value="JY/BEAM" arraysize="*" ucd="meta.unit">
           <DESCRIPTION>Unit for flux coordinate</DESCRIPTION>
       </PARAM>
       <FIELDref ref="FluxMin"/>
       <FIELDref ref="FluxSupportExtent"/>
    </GROUP>
   <DATA>
    <TABLEDATA>
    </TABLEDATA>
   </DATA>
  </TABLE>
 </RESOURCE>
</VOTABLE>
```

Appendix B: B0DEGA GDS GetCapabilities VOTable Response


```
<VOTABLE xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="xmlns:http://www.ivoa.net/xml/VOTable/VOTable-1.2.xsd"
xmlns:sia2="http:www.ivoa.net/xml/gds/v0.1" version="1.2">
<RESOURCE type="Results">
        <DESCRIPTION>getCapabilities query response on B0DEGA Generic Data Service</DESCRIPTION>
        <INFO name="QUERY_STATUS" value="OK">Successful metadata query</INFO>
        <INFO name="SERVICE_PROTOCOL">GDS</INFO>
        <!--mandatory input parameters -->
        <PARAM name="INPUT:POS" datatype="char" arraysize="*">
                <DESCRIPTION>POS defaults to right-ascension and declination in decimal degrees in
the ICRS coordinate system. A coordinate system reference frame may optionally be specified to
specify a coordinate system such as 'GALACTIC', 'GALACTIC_I', 'GALACTIC_II', 'ECLIPTIC' but no
results will be provided; default is ICRS/FK7.</DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:SIZE" value="0.1" datatype="double" unit="deg">
                <DESCRIPTION>The width and height of the rectangular region of interest. If only a
single value is given it applies to both the width and height of the search region, otherwise the
two values may be specified separately.</DESCRIPTION>
                <VALUES>
                        <MIN value="0"/>
                        <MAX value="5.0"/>
                </VALUES>
        </PARAM>
        <PARAM name="INPUT:BAND" value="ALL" datatype="char" arraysize="*">
                <DESCRIPTION>The spectral band pass is specified in range-list format either
numerically as a wavelength value or range, or textually as a spectral band pass identifier. If a
band pass is specified as a string identifier it is assumed to be a band pass identifier such as a
standard VO band pass name.</DESCRIPTION>
                <VALUES>
                        <OPTION value="ALL"/>
                        <OPTION value="radio"/>
                        <OPTION value="millimeter"/>
                        <OPTION value="infrared"/>
                        <OPTION value="optical"/>
                        <OPTION value="ultraviolet"/>
                        <OPTION value="x-ray"/>
                        <OPTION value="gamma-ray"/>
                        <OPTION value="U"/>
                        <OPTION value="V"/>
                        <OPTION value="B"/>
                        <OPTION value="R"/>
                        <OPTION value="I"/>
                        <OPTION value="J"/>
                </VALUES>
        </PARAM>
        <PARAM name="INPUT:TIME" datatype="char" arraysize="*">
                <DESCRIPTION>The temporal coverage (epoch of observation) is specified as a single
value or range in ISO8601 format.
        </DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:POL" value="any" datatype="char" arraysize="*">
                <DESCRIPTION>Specifies whether or not data is desired which measures polarization,
and if so the type of polarization desired.
                <VALUES>
                        <OPTION value="any"/>
                        <OPTION value="none"/>
                        <OPTION value="circular"/>
                        <OPTION value="stokes"/>
                </VALUES>
        </DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:FORMAT" value="all" datatype="char" arraysize="*">
                <DESCRIPTION>The FORMAT parameter specifies the allowable data formats for a
retrieved file.</DESCRIPTION>
                <VALUES>
```

```
                        <OPTION>all</OPTION>
                        <OPTION>application/fits</OPTION>
                        <OPTION>application/x-votable+xml</OPTION>
                        <OPTION>application/xml</OPTION>
                        <OPTION>text/html</OPTION>
                        <OPTION>text/csv</OPTION>
                        <OPTION>image/jpeg</OPTION>
                </VALUES>
        </PARAM>
        <PARAM name="INPUT:REQUEST" datatype="char" arraysize="*">
                <DESCRIPTION>SIA2 protocol versions supported by this service. Reserved words for
future extensions are: getData, stageData, getCapabilities, getAvailability. Values are treated
case-insensitive.</DESCRIPTION>
                <VALUES>
                        <OPTION value="queryData"/>
                        <OPTION value="getCapabilities"/>
                </VALUES>
        </PARAM>
        <!--optional/recommended parameters and service defined input parameters -->
        <PARAM name="INPUT:REDSHIFT" datatype="char" arraysize="*">
                <DESCRIPTION>A photometric (observed) redshift range specified as a single element
open or closed range-list.</DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:TARGETNAME" datatype="char" arraysize="*">
                <DESCRIPTION>The target name, suitable for input to a name resolver.</DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:TARGETCLASS" value="galaxy" datatype="char" arraysize="*">
                <DESCRIPTION>Types of astronomical objects to be searched for.</DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:MTIME" datatype="char" arraysize="*">
                <DESCRIPTION>Only datasets modified, created or deleted at a given
range.</DESCRIPTION>
        </PARAM>
        <PARAM name="INPUT:MAXREC" datatype="int">
                <DESCRIPTION>Maximum number of records to be returned.</DESCRIPTION>
        </PARAM>

        <!--query response parameters (name="OUTPUT:param-name") -->
        <!-- service metadata: Access.* -->
        <PARAM name="OUTPUT:AcRef" datatype="char" arraysize="*" utype="gds:Access.Reference"
ucd="meta.ref.url">
                <DESCRIPTION>URL used to access dataset</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Format" datatype="char" arraysize="*" utype="gds:Access.Format">
                <VALUES>
                        <OPTION>application/fits</OPTION>
                        <OPTION>application/x-votable+xml</OPTION>
                        <OPTION>application/xml</OPTION>
                        <OPTION>text/html</OPTION>
                        <OPTION>text/csv</OPTION>
                        <OPTION>image/jpeg</OPTION>
                </VALUES>
                <DESCRIPTION>Content of MIME type of dataset</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:DatasetSize" datatype="long" arraysize="*" utype="gds:Access.Size"
unit="kB">
                <DESCRIPTION>Estimated dataset size</DESCRIPTION>
        </PARAM>

        <!-- data model metadata -->
        <!--data model metadata: Dataset.* -->
        <PARAM name="OUTPUT:DataModel" value="Obs-1.0" datatype="char" arraysize="*"
utype="gds:Dataset.DataModel" >
                <DESCRIPTION>Datamodel name and version</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:DatasetType" value="Cube" datatype="char" arraysize="*"
utype="gds:Dataset.Type">
                <DESCRIPTION>Type of dataset</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:DataLenght" datatype="long" utype="gds:Dataset.Lenght"
ucd="meta.number">
```

```
                <DESCRIPTION>Number of points</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Deleted" datatype="char" arraysize="*" utype="gds:Dataset.Deleted">
                <DESCRIPTION>Set to deletion time, if dataset is deleted</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: DataID.* -->
        <PARAM name="OUTPUT:Title" datatype="char" arraysize="*" utype="gds:DataID.Title"
ucd="meta.title;meta.dataset">
                <DESCRIPTION>Dataset Title</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:CreatorDate" datatype="char" arraysize="*" utype="gds:DataID.Date"
ucd="time;meta.dataset">
                <DESCRIPTION>Dataset creator</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: Provenance.* -->
        <PARAM name="OUTPUT:Instrument" datatype="char" arraysize="*"
utype="gds:Provenance.ObsConfig.Instrument" ucd="meta.id;instr">
                <DESCRIPTION>Instrument name</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:BeamMajorAxis" datatype="double"
utype="gds:Char.Provenance.ObsConfig.BeamMajorAxis " ucd="phys.angSize;instr.beam" unit="deg">
                <DESCRIPTION>Beam Major Axis</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:BeamMinorAxis" datatype="double"
utype="gds:Char.Provenance.ObsConfig.BeamMinorAxis " ucd="phys.angSize;instr.beam" unit="deg">
                <DESCRIPTION>Beam Minor Axis</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:BeamPositionAngle" datatype="double"
utype="gds:Char.Provenance.ObsConfig.BeamPositionAngle" ucd="pos.posAng;instr.beam" unit="deg">
                <DESCRIPTION>Beam Position Angle</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Bandpass" datatype="char"  arraysize="*"
utype="gds:Provenance.ObsConfig.Bandpass" ucd="instr.bandpass">
                <DESCRIPTION>Band as in RSM Coverage.Spectral</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: Curation.* -->
        <PARAM name="OUTPUT:PublisherDate" datatype="char" arraysize="*" utype="gds:Curation.Date"
ucd="time;meta.curation">
                <DESCRIPTION>Date curated dataset last modified</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Rights" datatype="char" arraysize="*" utype="gds:Curation.Rights" >
                <DESCRIPTION>Restrictions on data access</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Publisher" value="B0DEGA Archive" datatype="char" arraysize="*"
utype="gds:Curation.Publisher" ucd="meta.curation">
                <DESCRIPTION>Dataset publisher</DESCRIPTION>
        </PARAM>
        <PARAM name="UTPUT:PublisherID" value="ivo://svo.b0dega.iaa.es/b0dega" datatype="char"
arraysize="*" utype="gds:Curation.PublisherID" ucd="meta.ref.url;meta.curation">
                <DESCRIPTION>URI for VO Publisher</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: Target.* -->
        <PARAM name="OUTPUT:TargetName" datatype="char" arraysize="*" utype="gds:Target.Name"
ucd="meta.id;src">
                <DESCRIPTION>Target name</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:TargetClass" datatype="char" arraysize="*" utype="gds:Target.Class"
ucd="src.class">
                <DESCRIPTION>Object class of observed target</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:TargetPos" datatype="double" arraysize="2" utype="gds:Target.Pos"
ucd="pos.eq;src" unit="deg">
                <DESCRIPTION>Target RA and Dec</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:Velocity" datatype="double" utype="gds:Target.Velocity"
ucd="phys.veloc" unit="km/s">
                <DESCRIPTION>Target redshift</DESCRIPTION>
```

```
        </PARAM>
        <PARAM name="OUTPUT:Redshift" datatype="double" utype="gds:Target.Redshift"
ucd="src.redshift">
                <DESCRIPTION>Target redshift</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: Derived.* -->
        <PARAM name="OUTPUT:DerivedVelocity" datatype="double" utype="gds:Derived.Velocity.Value"
ucd="phys.veloc" unit="m/s">
                <DESCRIPTION>Derived mean velocity from spectrum</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:DerivedRedshift" datatype="double" utype="gds:Derived.Redshift.Value">
                <DESCRIPTION>Measured redshift for spectrum</DESCRIPTION>
        </PARAM>

        <!-- data model metadata: CoordSys.* -->
        <PARAM name="OUTPUT:SpaceFrameName" value="ICRS" arraysize="*" datatype="char"
utype="gds:CoordSys.SpaceFrame.Name">
                <DESCRIPTION>Spatial coordinate frame name</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpaceFrameUcd" value="pos.eq" arraysize="*" datatype="char"
utype="gds:CoordSys.SpaceFrame.Ucd">
                <DESCRIPTION>Space frame UCD</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpaceFrameEquinox" value="2000.0" datatype="double"
utype="gds:CoordSys.SpaceFrame.Equinox"  ucd="time.equinox;pos.frame">
                <DESCRIPTION>Equinox</DESCRIPTION>
        </PARAM>

        <!-- characterization metadata -->
        <!-- characterization metadata: Char.SpatialAxis -->
        <PARAM name="OUTPUT:SpatialAxisName" value="sky" datatype="char" arraysize="*"
utype="gds:Char.SpatialAxis.AxisName" ucd="meta.id">
                <DESCRIPTION>Spatial coordinate frame name
        </DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpatialAxisUcd" value="pos.eq" datatype="char" arraysize="*"
utype="gds:Char.SpatialAxis.Ucd" ucd="meta.ucd">
                <DESCRIPTION>Space frame UCD</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpatialAxisUnit" value="deg" datatype="char" arraysize="*"
utype="gds:Char.SpatialAxis.Unit" ucd="meta.unit">
                <DESCRIPTION>Unit for spectral coordinate</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpatialLocation" datatype="double" arraysize="2"
utype="gds:Char.SpatialAxis.Coverage.Location.Coord" ucd="pos.eq" unit="deg">
                <DESCRIPTION>Spatial Position</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpatialExtent" datatype="double" arraysize="2"
utype="gds:Char.SpatialAxis.Coverage.Bounds.Extent" ucd="instr.fov" unit="deg">
                <DESCRIPTION>Angular spatial extent in degrees</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:SpatialSampleExtent" datatype="float" arraysize="2"
utype="gds:Char.SpatialAxis.SamplingPrecision.SampleExtent" ucd="instr.pixel">
                <DESCRIPTION>Spatial bin size</DESCRIPTION>
        </PARAM>

        <!-- characterization metadata: Char.FluxAxis -->
        <PARAM name="OUTPUT:FluxAxisName" value="FluxDensity" datatype="char" arraysize="*"
utype="gds:Char.FluxAxis.AxisName" ucd="meta.id">
                <DESCRIPTION>Flux coordinate frame name</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:FluxAxisUcd" value="phot.flux.density;instr.beam" datatype="char"
arraysize="*" utype="gds:Char.FluxAxis.Ucd" ucd="meta.ucd">
                <DESCRIPTION>Flux frame UCD</DESCRIPTION>
        </PARAM>
        <PARAM name="OUTPUT:FluxAxisUnit" value="JY/BEAM" datatype="char" arraysize="*"
utype="gds:Char.FluxAxis.Unit" ucd="meta.unit">
                <DESCRIPTION>Unit for flux coordinate</DESCRIPTION>
        </PARAM>
        <PARAM name="FluxMin" datatype="double" utype="gds:Char.FluxAxis.Coverage.Bounds.Start"
```

```
ucd="phot.flux.density;stat.min" unit="Jy/beam">
               <DESCRIPTION>Limiting value for flux density</DESCRIPTION>
       </PARAM>
       <PARAM name="FluxSupportExtent" datatype="double" arraysize="2"
utype="gds:Char.FluxAxis.Coverage.Support.Extent" ucd="phot.flux.density" unit="Jy/beam">
               <DESCRIPTION>Minimum, maximun value for flux density</DESCRIPTION>
       </PARAM>

       <!-- characterization metadata: SpectralAxis -->
       <PARAM name="OUTPUT:SpectralAxisName" value="Velocity" datatype="char" arraysize="*"
utype="gds:Char.SpectralAxis.AxisName" ucd="meta.id">
               <DESCRIPTION>Spectral coordinate frame name</DESCRIPTION>
       </PARAM>
       <PARAM name="OUTPUT:SpectralAxisUcd" value="spat.dopplerVelocity.ratio" datatype="char"
arraysize="*" utype="gds:Char.SpectralAxis.Ucd" ucd="meta.ucd">
               <DESCRIPTION>Spectral frame UCD</DESCRIPTION>
       </PARAM>
       <PARAM name="OUTPUT:SpectralAxisUnit" value="m/s" datatype="char" arraysize="*"
utype="gds:Char.SpectralAxis.Unit">
               <DESCRIPTION>Unit for spectral coordinate</DESCRIPTION>
       </PARAM>
       <PARAM name="SpectralLocation" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Location.Coord" ucd="instr.bandpass" unit="m/s">
               <DESCRIPTION>Spectral coord value</DESCRIPTION>
       </PARAM>
       <PARAM name="SpectralExtent" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Extent" ucd="instr.bandwidth" unit="m/s">
               <DESCRIPTION>Width of spectrum</DESCRIPTION>
       </PARAM>
       <PARAM name="SpectralStart" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Start" ucd="em;stat.min" unit="m/s">
               <DESCRIPTION>Start in spectral coordinate</DESCRIPTION>
       </PARAM>
       <PARAM name="SpectralStop" datatype="double"
utype="gds:Char.SpectralAxis.Coverage.Bounds.Stop" ucd="em;stat.max" unit="m/s">
               <DESCRIPTION>Stop in spectral coordinate</DESCRIPTION>
       </PARAM>

</RESOURCE>
</VOTABLE>
```

Bibliography

[AC98] J. Allington-Smith and R. Content. *Sampling and Background Subtraction in Fiber-Lenslet Integral Field Spectrographs*. PASP, 110:1216–1234, 1998.

[Arv10] C. Arviset. *The IVOA in 2010: Technical assessment and roadmap*. IVOA Note Version 1.0, International Virtual Observatory Alliance, October 2010.

[BCM08] M. F. Bode, M. J. Cruz, and F. J. Molster, editors. *The ASTRONET Infrastructure Roadmap: A Strategic Plan for European Astronomy*. ASTRONET, 2008.

[BDM+09] F. Bonnarel, D. Durand, A. Micol, A. Richards, J. Salgado, and D. Tody. *Simple Image Access Protocol*. IVOA Working Draft Version 2.0, International Virtual Observatory Alliance, November 2009.

[CBLM07] I. Chilingarian, F. Bonnarel, M. Louys, and J. McDowell. *Handling IFU Datasets in the Virtual Observatory*. In M. Kissler-Patig, J. R. Walsh, & M. M. Roth, editor, Science Perspectives for 3D Spectroscopy, pages 29–+, 2007.

[DTDS08] F. Bonnarel, D. Tody, M. Dolensky, and J. Salgado. *IVOA Data Access Layer Service Architecture and Standard Profile*. IVOA Note Version 0.1, International Virtual Observatory Alliance, October 2008.

[Gen02] F. Genova. *Interoperability*. In D. A. Bohlender, D. Durand, & T. H. Handley, editor, Astronomical Data Analysis Software and Systems XI, volume 281 of Astronomical Society of the Pacific Conference Series, pages 41–+, 2002.

[HTT09] T. Hey, S. Tansley, and K. Toll, editors. *The Fourth Paradigm*. Microsoft Corporation, 2009.

[Hat10] E. Hatziminaoglou. *Virtual Observatory: Science Capabilities and Scientific Results.* In K. Tsinganos, D. Hatzidimitriou, & T. Matsakos, editor, Astronomical Society of the Pacific Conference Series, volume 424 of Astronomical Society of the Pacific Conference Series, pages 411–+, July 2010.

[LRB+08] M. Louys, A. Richards, F. Bonnarel, A. Micol, I. Chilingarian, and J. McDowell. *Data Model for Astronomical DataSet Characterisation*. IVOA Recommendation Version 1.13, International Virtual Observatory Al- liance, March 2008.

[MAR+10] E. Mediavilla, S. Arribas, M. Roth, J. Cepa-Nogué, and F. Sánchez, editors. *3D Spectroscopy in Astronomy*, volume XVII of Canary Islands Winter School of Astrophysics. Cambridge University Press, 2010.

[McD07] J. McDowell. *IVOA Spectral Data Model*. IVOA Recommendation Version 1.03, International Virtual Observatory Alliance, 2007.

[Mil08] B. W. Miller. *3D Spectroscopy and the Virtual Observatory*. In M. Guainazzi & P. Osuna, editor, Astronomical Spectroscopy and Virtual Observatory, pages 117–+, 2008.

[OW09] F. Ochsenbein and R. Williams. *VOTable Format Definition*. IVOA Recommendation Version 1.2, International Virtual Observatory Alliance, November 2009.

[PDD+07] A. Preite Martinez, S. Derriere, N. Delmotte, N. Gray, R. Mann, J. McDowell, T. Mc Glynn, F. Ochsenbein, P. Osuna, G. Rixon, and R. Williams. *The UCD1+ controlled vocabulary*. IVOA Recommendation Version 1.23, International Virtual Observatory Alliance, April 2007.

[RB09] A. Richards and F. Bonnarel. *Note on the description of polarization data*. IVOA Note Version 0.01, International Virtual Observatory Alliance, May 2009.

[RdG+10] J. E. Ruiz, J. de Dios Santander-Vela, E. Garcia, V. Espigares, S. Leon, and L. Verdes-Montenegro. *Virtual Observatory Activities in the AMIGA Group*. In J. M. Diego, L. J. Goicoechea, J. I. González-Serrano, & J. Gorgas, editor, Highlights of Spanish Astrophysics V, pages 533–+, 2010.

[RSE+09b] J. E. Ruiz, J. D. Santander-Vela, V. Espigares, L. Verdes-Montenegro, and J. M. van der Hulst. *GIPSY 3D: Analysis, Visualization and VO Tools for Datacubes*. In D. A. Bohlender, D. Durand, & P. Dowler, editor, Astronomical Society of the Pacific Conference Series, volume 411 of Astronomical Society of the Pacific Conference Series, pages 406–+, 2009.

[San09] J. D. Santander-Vela. *Integración de archivos y herramientas radioastronómicas en la arquitectura del Observatorio Virtual*. PhD thesis, Universidad de Granada, May 2009.

[Sch07] A. Schaaff. *Data in astronomy: From the pipeline to the virtual observatory*. In WISE'07 Proceedings of the 2007 international conference on Web information systems engineering, pages 52–62. Springer-Verlag Berlin, Heidelberg, 2007.

[TBDS08] D. Tody, F. Bonnarel, M. Dolensky, and J. Salgado. *Simple Image Access Protocol V2 Analysis, Scope and Concepts*. IVOA Note Version 0.2, Inter- national Virtual Observatory Alliance, October 2008.

[TBF+09] M. B. Taylor, T. Boch, M. Fitzpatrick, A. Allan, L. Paioro, J. Taylor, and D. Tody. *Simple Application Messaging Protocol*. IVOA Proposed Recommendation Version 1.11, International Virtual Observatory Alliance, April 2009.

[TD08] D. Tody and M. Dolensky. *Simple Spectral Access Protocol*. IVOA Recommendation Version 1.04, International Virtual Observatory Alliance, February 2008.

[TP09] D. Tody and R. Plante. *Simple Image Access Specification*. IVOA Recommendation Version 1.0, International Virtual Observatory Alliance, November, 2009.

[vTB+92] J. M. van der Hulst, J. P. Terlouw, K. G. Begeman, W. Zwitser, and P. R. Roelfsema. *The Groningen Image Processing SYstem, GIPSY*. In D. M. Worrall, C. Biemesderfer, & J. Barnes, editor, Astronomical Data Analysis Software and Systems I, volume 25 of Astronomical Society of the Pacific Conference Series, pages 131–+, 1992.