



*International  
Virtual  
Observatory  
Alliance*

## Radio Astronomy in the VO: implementation note

### Version 1.0

#### IVOA Note 2021-05-21

Working group

Radio Astronomy interest group

This version

<https://www.ivoa.net/documents/RadioVOImplementation/20210521>

Latest version

<https://www.ivoa.net/documents/RadioVOImplementation>

Previous versions

This is the first public release

Author(s)

Mark Lacy, François Bonnarel, The RadioAstronomy Interest Group

Editor(s)

Mark Lacy, François Bonnarel

### Abstract

This document provides some hints to help Radio astronomy projects data provider to build VO services on top of their archives. It gathers experience gained by various groups around the world

### Status of this document

This is an IVOA Note expressing suggestions from and opinions of the authors. It is intended to share best practices, possible approaches, or other perspectives on interoperability with the Virtual Observatory. It should not be referenced or otherwise interpreted as a standard specification.

A list of current IVOA Recommendations and other technical documents can be found at <https://www.ivoa.net/documents/>.

## Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>3</b>  |
| <b>2</b> | <b>Science data</b>   | <b>4</b>  |
| 2.1      | Radio 2D images and cubes . . . . .                             | 4         |
| 2.1.1    | Data Access Layer (DAL) services . . . . .                      | 4         |
| 2.1.2    | HiPS . . . . .  | 5         |
| 2.1.3    | List of working services . . . . .                              | 5         |
| 2.2      | Polarization data . . . . .                                     | 5         |
| 2.3      | Radial velocity, Rotation and dispersion measure maps . . . . . | 6         |
| 2.4      | Source catalogues . . . . .                                     | 6         |
| 2.5      | TimeSeries, dynamical spectra, pulsar data . . . . .            | 6         |
| 2.6      | Using MOC for discovery . . . . .                               | 6         |
| 2.7      | Registration . . . . .  | 7         |
| <b>3</b> | <b>Raw data</b>   | <b>7</b>  |
| 3.1      | Interferometry . . . . .  | 7         |
| 3.2      | Single dish data . . . . .                                      | 8         |
| 3.3      | Beam forming data . . . . .                                     | 8         |
| <b>4</b> | <b>Provenance</b>   | <b>8</b>  |
|          | <b>References</b>   | <b>9</b>  |
| <b>A</b> | <b>Current Implementation stories</b>                           | <b>10</b> |
| A.1      | ALMA . . . . .  | 10        |
| A.2      | INAF Radio Data . . . . .                                       | 11        |
| A.3      | The Joint Institute for VLBI in Europe . . . . .                | 12        |
| A.4      | The Murchison Wide-Field Array . . . . .                        | 13        |
| A.5      | CADC . . . . .  | 14        |
| A.6      | Nançais . . . . .   | 14        |
| A.6.1    | NDA : . . . . .   | 14        |
| A.6.2    | NRH / ORFEES : . . . . .  | 15        |
| A.6.3    | NenuFAR : . . . . .   | 15        |
| A.7      | CSIRO Radio Archives . . . . .                                  | 16        |
| A.7.1    | CASDA . . . . .   | 16        |
| A.7.2    | PSRDA . . . . .   | 16        |
| A.7.3    | ATOA . . . . .  | 17        |
| A.8      | ASTRON VO services . . . . .                                    | 17        |
| A.9      | MERLIN . . . . .  | 17        |

## Acknowledgments

Authors gratefully thank all the radioastronomy projects who provided storytelling gathered in Appendix

## Conformance-related definitions

The words “MUST”, “SHALL”, “SHOULD”, “MAY”, “RECOMMENDED”, and “OPTIONAL” (in upper or lower case) used in this document are to be interpreted as described in IETF standard RFC2119 [? ].

The *Virtual Observatory (VO)* is a general term for a collection of federated resources that can be used to conduct astronomical research, education, and outreach. The *International Virtual Observatory Alliance (IVOA)* is a global collaboration of separately funded projects to develop standards and infrastructure that enable VO applications.

## 1 Introduction

Radio astronomy data were considered in the first decade of IVOA, within the *Char*/Observation datamodel and the first version Simple Image Access (SIA1, [12]). When multidimensional data science became a CSP priority (2013), this drove the emergence of new protocols that included treatment of polarization and image cubes, both important for radio astronomy: ObsCore [9], SIA2 [5], DataLink [4] and SODA [2]. Since that time, radio astronomy projects have further increased in size and variety, necessitating a review of how the VO can better work with radio data. In this document, we begin this process as follows:

- In Sections 2 and 3 below, we give a list of existing radio astronomy VO services with protocol used for ALMA (section A.1), LOFAR (section A.8), ASKAP, ATCA, (section A.7) NRAO and CGPS (section A.5) and MERLIN (section A.9). The content of these Sections is organized by category of data and then type of services. Section 2 presents science data, Section 3 concentrates on raw (visibility) data. We try to extract some generic lessons on how services can proceed to build services.
- We asked projects to provide us some “story-telling” of what they have done (and how they did it) : this can be found in the appendix

The scope of this document is restricted to data from heterodyne receivers working at frequencies of 0-1 THz. Far-infrared data has a lot of data from space missions, and often makes use of bolometer technology, which has a lot more in common with optical/near-infrared data than most radio data.

## 2 Science data

In this section we only consider high level dataproducts that can be produced from any of interferometers, single dishes etc. We defer discussion of observational data to Section 3.

### 2.1 Radio 2D images and cubes

#### 2.1.1 Data Access Layer (DAL) services

The combination of the ObsCore data model and the VO's Table Access Protocol (TAP; [6]), known as ObsTAP, and the SIA are classical "discovery, then access" services adapted to images and cubes, and can be used to serve radio image products. Metadata characterizing the observation needed for ObsCore can be extracted from FITS headers or observation logs. The main issues for metadata extraction are to choose the granularity and quality of datasets to expose, and to find solutions for access in cases where the datasets are huge. ObsCore tables can be served by two different interfaces, TAP (queryable via Astronomy Data Query Language (ADQL) [10] ), or SIA2 (a parameter based interface). The SIA1 interface remains a VO standard. It is a parameter based interface with a different set of parameters than SIA2. The standardised SIA1 response is not based on ObsCore and has archaic features. The SIA1 interface should be considered as a legacy standard for the first-generation of VO services. We do not recommend its implementation for new services.



For large datasets, where retrieval of entire images/cubes may be cumbersome, DataLink and SODA are appropriate solutions that have been experimented with at CADC (section A.5) and by CASDA. SODA allows the extraction of subcubes by selecting a smaller spatial coverage or spectral range. DataLink enables the linkage of a SODA interface to a specific dataset discovered with ObsTAP or SIA. Access to raw data and provenance information can also be provided via DataLink, as it is the case for INAF radio archive and ALMA science archive.

The Common Archive Observation Model (CAOM) may be a good intermediary between the set of metadata in image headers metadata and VO standards. Some CAOM standard metadata map easily to Obscore attributes while others can be used to help develop SODA or DataLink services. This is the way CADC implements DAL services. DACHS <sup>1</sup> is a complete python toolkit allowing to create all DAL services and register them. Data centers such as CDS, CSIRO and CADC have developed their own open java libraries (vollt, CASDA VO Tools, openCADC). These libraries can be used to develop new services.

---

<sup>1</sup><http://docs.g-vo.org/DaCHS/>

### 2.1.2 HiPS

A completely different way to make image data available through the VO is by using HiPS [7]. HiPS provides a hierarchical progressive access to pixels. It can be used for 2D images and 3D cubes. In the radio domain, it is often used for 2D images produced from spectral cubes. The “tile progenitor” functionality may be used to help discover original images or cubes. Astron has published HiPS for Apertif and LOFAR data, and CADC for CGPS(section A.5). Several software packages are able to produce HiPS images, and when that is done no more software is needed to install a HiPS server - a simple http server allows access to the hierarchical data. There are several VO applications able to visualize HiPS (AladinLite, Aladin Desktop, Stellarium, etc...).

### 2.1.3 List of working services

- ObsTAP services: ALMA(section A.1), CADC (VLA Sky Survey, Canadian Galactic Plane Survey -DRAO-, other DRAO data, VLA Galactic Plane Survey, Section A.5), CSIRO (CASDA, ATOA - Section A.7), Astron (section A.8).
- SIA1: CADC, Astron (APERTIF, LOTSS), Skyview/GSFC (FIRST, NVSS)
- SIA2: ALMA, CADC, CASDA, ASTRON (APERTIF, LOTSS)
- DataLink: ALMA, CADC, CASDA
- SODA: CADC, CASDA
- HiPS: CASDA, Apertif, LOTSS, CGPS at CADC
- Applications: Aladin, DS9, TOPCAT allow discovery and simple visualisation of cubes. Neither the current CASA viewer, nor the new CARTA viewer support VO searches or SAMP [1], though CARTA has long-term plans to do so.



## 2.2 Polarization data

ObsCore attributes `pol_states` and `pol_xel` allow to describe the polarization content of datasets. Polarization add a 4th dimension to spectral cubes, but this does not prevent ObsCore from describing them. Several of the ObsTAP and SIA services listed above provide Polarization data (eg: Astron, CADC...)

## 2.3 Radial velocity, Rotation and dispersion measure maps

These are actually classical 2D maps derived from spectral cubes, as the continuum, line or "moment zero" maps are. Except that the Observable quantity is not a flux density. These are thus manageable with an ObsCore/SIA service. The provider only has to choose the best `o_ucd` value, for example:

- rotation measure: `phys.polarization.rotMeasure;instr.rmsf`
- radial velocities : `spect.dopplerVeloc`
- dispersion measure : `phys.veloc.dispersion`

## 2.4 Source catalogues

Extraction of radio sources from survey images generates catalogues of sources. Their provenance is specific, but otherwise they are no different from other type of catalogues that can be served by many VO services, following either SCS [13] or TAP standards. A few examples are:

- Simple Cone Search (SCS) services: CASDA, ATOA, PSRDA (Parkes Pulsar data archive) - see section A.7 - , CDS/VizieR
- TAP services: CDS/VizieR (NVSS, FIRST, etc...), ASTRON (WSRT, TGSS, LoFAR), CASDA, ATOA, PSRDA

## 2.5 TimeSeries, dynamical spectra, pulsar data

Data sampled on rapid (millisecond) timescales are common in pulsar astronomy and the study of Fast Radio Bursts (FRBs). Solar system objects (the Sun, and some planets, such as Jupiter), also can show rapid time variability in their radio emission. Data streams typically capture flux variations along with frequency and polarization information, in the form of "dynamical spectra". These products may be described with ObsCore and exposed through TAP or, in some cases, SIA. Dataproduct\_type "timeseries" ~~may be~~ ambiguous for dynamical spectra, so a `dataproduct_subtype` ~~can~~ be used. Such data are made available through JIVE (section A.3) and Nançay (section A.6) services and also by the Parkes Pulsar Data Archive (PSRDA).

## 2.6 Using MOC for discovery

The Multi-Order Coverage (MOC; [8], spatial MOC) standard enables the description of the coverage of surveys or data collections on the sky using the Healpix tessellation. The IVOA registry integrates these survey MOCs. This is very useful for all sky discovery of the radio data collections. CADC, CASDA, AStron, .. provide MOCs for their data embedded in VO services.

An extension of MOC to Time (combined or separated from space) will be standardized by next version of the MOC specification currently in discussion.



## 2.7 Registration

To be discoverable in the VO, a service must be registered. When a registered VO service is specific to some radio data collection it's easy to discover it. When a generic service contains radio data as well as other types, it may be a good idea to register the service with a specific auxiliary capability, as explained in the IVOA endorsed note discovering data collections [3].

## 3 Raw data

The nature and format of raw radio astronomical data depends on the observational technique. Interferometric telescopes typically store data in the form of complex visibilities (post correlation), while single dish or beam-forming observations will provide on-the-fly maps or other kind of mapping data, but also uncalibrated flux variation as a function of position, time and frequency.

The IVOA does not promote specific formats. Visibilities may be provided as Measurement Sets, FITS-IDI, RPFITS or UVFITS. Pulsar data are provided in PSRFITS and single dish data may use SDFITS or HDF5. An important feature of IVOA is to be able to describe the media type (format) of the data before retrieval. Extensions of the mime type vocabulary may be standardized in a near future.



### 3.1 Interferometry

Complex visibilities can be discovered using three different services according to their relationship with science data and to their characterization : SCS, ObsTAP and DataLink. SCS is adapted to datasets with small field of view, where a pointed observation can be represented by an entry in a catalog. ObsTAP is well adapted for visibility data matching the science spectral cube contours derived from them. Special attention may be needed to describe interferometry data for observations containing several targets and various spectral resolutions. They have to be split into several datasets for the discovery (JIVE has experimented with this). DataLink exposition allows the linking of visibility data to the science data they produce but makes direct discovery more difficult.

- ObsTAP services: ATOA, MWRA (section A.4), ASKAP, JIVE (section A.3)

- DataLink : ALMA, ASTRON (apertif)
- SCS : ASTRON (apertif)

### 3.2 Single dish data

Before being used to produce science data in a more classical format (nD cubes and images, dynamical spectra or spectra), single dish raw or calibrated data are generally stored or transported as collections of spectra for a given position, time, polarization state, as it is the case eg in the SDFITS standard. Such datasets could be accessed from science data using DataLink but can also be discovered in an ObsCore service. In the latter case `dataproduct_type`, `dataproduct_subtype`, `calibration_level` and `access_format` should be adapted to single dish specificities.

- ATOA
- INAF (section A.2)
- Other candidates for VO integration
  - FAST
  - Arecibo

### 3.3 Beam forming data

Some telescopes arrays (LOFAR, Nenufar, VLBI) may provide "beam forming data" by coherently summing up signals from the different antennae, instead of registering the visibilities of the interference patterns. In this way they obtain time/frequency data. At the time of writing these are the projects providing such kind of data:

- JIVE
- Nançay

## 4 Provenance

IVOA has specified the Provenance DataModel [11] which allows the history of the data to be traced via structured metadata. The Provenance model enables high-level data products to be linked the raw data, and the processing and QA steps to be identified. At the time of writing there are no radioastronomy VO service providing IVOA provenance metadata, though a service is planned by the Spanish VO.

## References

- [1] Boch, T., Fitzpatrick, M., Taylor, M., Allan, A., Paioro, L., Taylor, J. and Tody, D. [2009], ‘SAMP — Simple Application Messaging Protocol Version 1.11’, IVOA Recommendation 21 April 2009.  
<http://doi.org/10.5479/ADS/bib/2009ivoa.spec.0421B>
- [2] Bonnarel, F., Dowler, P., Demleitner, M., Tody, D. and Dempsey, J. [2017], ‘IVOA Server-side Operations for Data Access Version 1.0’, IVOA Recommendation 17 May 2017, arXiv:1710.08791.  
<http://adsabs.harvard.edu/abs/2017ivoa.spec.0517B>
- [3] Demleitner, M. and Taylor, M. [2019], ‘Discovering Data Collections Within Services Version 1.1’, IVOA Note 20 May 2019.  
<http://doi.org/10.5479/ADS/bib/2019ivoa.rept.0520D>
- [4] Dowler, P., Bonnarel, F., Michel, L. and Demleitner, M. [2015], ‘IVOA DataLink Version 1.0’, IVOA Recommendation 17 June 2015, arXiv:1509.06152.  
<http://doi.org/10.5479/ADS/bib/2015ivoa.spec.0617D>
- [5] Dowler, P., Bonnarel, F. and Tody, D. [2015], ‘IVOA Simple Image Access Version 2.0’, IVOA Recommendation 23 December 2015.  
<http://adsabs.harvard.edu/abs/2015ivoa.spec.1223D>
- [6] Dowler, P., Rixon, G. and Tody, D. [2010], ‘Table access protocol version 1.0’, IVOA Recommendation.  
<http://www.ivoa.net/documents/TAP>
- [7] Fernique, P., Allen, M., Boch, T., Donaldson, T., Durand, D., Ebisawa, K., Michel, L., Salgado, J. and Stoehr, F. [2017], ‘HiPS - Hierarchical Progressive Survey Version 1.0’, IVOA Recommendation 19 May 2017, arXiv:1708.09704.  
<http://doi.org/10.5479/ADS/bib/2017ivoa.spec.0519F>
- [8] Fernique, P., Boch, T., Donaldson, T., Durand, D., O’Mullane, W., Reinecke, M. and Taylor, M. [2019], ‘MOC - HEALPix Multi-Order Coverage map Version 1.1’, IVOA Recommendation 07 October 2019.  
<https://ui.adsabs.harvard.edu/abs/2019ivoa.spec.1007F>
- [9] Louys, M., Bonnarel, F., Schade, D., Dowler, P., Micol, A., Durand, D., Tody, D., Michel, L., Salgado, J., Chilingarian, I., Rino, B., de Dios Santander, J. and Skoda, P. [2011], ‘Observation data model core components and its implementation in the Table Access Protocol, version 1.0’, IVOA Recommendation.  
<http://www.ivoa.net/documents/ObsCore/20111028/REC-ObsCore-v1.0-20111028.pdf>

- [10] Osuna, P., Ortiz, I., Lusted, J., Dowler, P., Szalay, A., Shirasaki, Y., Nieto-Santisteban, M. A., Ohishi, M., O’Mullane, W., VOQL-TEG Group and VOQL Working Group. [2008], ‘IVOA Astronomical Data Query Language Version 2.00’, IVOA Recommendation 30 October 2008, arXiv:1110.0503.  
<http://doi.org/10.5479/ADS/bib/2008ivoa.spec.10300>
- [11] Servillat, M., Riebe, K., Boisson, C., Bonnarel, F., Galkin, A., Louys, M., Nullmeier, M., Renault-Tinacci, N., Sanguillon, M. and Streicher, O. [2020], ‘IVOA Provenance Data Model Version 1.0’, IVOA Recommendation 11 April 2020.  
<https://ui.adsabs.harvard.edu/abs/2020ivoa.spec.0411S>
- [12] Tody, D. and Plante, R. [2009], ‘Simple image access specification’, IVOA Recommendation.  
<http://www.ivoa.net/documents/latest/SIA.html>
- [13] Williams, R., Hanisch, R., Szalay, A. and Plante, R. [2008], ‘Simple cone search’, IVOA Recommendation.  
<http://www.ivoa.net/documents/latest/ConeSearch.html>

## A Current Implementation stories

### A.1 ALMA

The ALMA archive has implemented ObsTAP (<https://almascience.org/tap>), SIAv2 (<https://almascience.org/siav2>) and DataLink (<https://almascience.org/datalink>) services. The datalink service offers the same files that are also offered through the normal download, including visibility data. The searches, however, happen on a higher-level of the data where metadata from several sets of visibilities is combined.

The decision of which granularity of the very hierarchical ALMA data structure to offer took quite some time. In the end we opted for serving one row corresponding to a `project_code+memberObsUnitSet+SpectralWindow`. This roughly relates to one row per primary FITS product that can be created.

A particular difficulty for ALMA which is not yet entirely resolved is that three identical services are running at the three ALMA regional centres where the URLs like <https://almascience.org/tap> get redirected to the closest regional centre e.g. <https://almascience.eso.org/tap>. Whether it is best to register the three TAP, three SIAv2 and three DataLink services or only registering the original link is not clear and depends also on how the VO tools handle such cases.

For the downloads through DataLink there are at least four hierarchy levels of the data that people could be interested in accessing (Project, ObsUnitSet, Source or Spectral Window). We hope that through astroquery those can be made available to the users. The download of proprietary data through astroquery and DataLink is in the works.

## A.2 INAF Radio Data

INAF manages three fully steerable reflector antennas for radio astronomy: the 32m dishes at Medicina and Noto and the 64m Sardinia Radio Telescope (SRT). These observing facilities can be used separately as single-dish instruments (SD) or in a coordinated national network, sometimes joined by other EVN facilities, that delivers VLBI interferometric data. Pulsar observations are supported in both single-dish and interferometric mode. The various raw data types (single dish, interferometric and pulsar) are characterized by different output data formats: FITS, FITS-IDI and PSRFITS. Aiming at persistence and future scientific exploitation of data, ancillary information contained in the observing schedules and telescopes/correlator logs is archived in separate files for each dataset as well. The radio archive database stores the metadata from the currently adopted formats and is capable to include further instruments provided their data can be described by means of the general database (internal) radio datamodel.

Currently the archive stores continuum and spectropolarimetric raw data from single dish, pulsar and VLBI observations (in this last case the correlated visibilities are archived as raw data). Storage of processed data is planned once the Archive Science Gateway and the User Space will be finalized. At that point, processing pipelines, calibration and processing information as well as some level of quality metrics will be provided.

The Radio Archive is accessible through a dedicated web interface. According to INAF policies, to guarantee the data proprietary period a Single-Sign-On login authentication mechanism is foreseen while public data are available for download without user registration. In both public or private cases, users perform dedicated queries in the Radio Archive by means of web forms. Query parameters can be common to the various data types, like for instance the celestial object coordinates, or specific to the observing mode, like the scan geometry for single dish or the subset of antennas for interferometric data. In order to increase both the open data re-usage and the accessibility of scientific data to the astronomical community and the general public, we are planning to adopt IVOA standards or, at least, solutions that take into account the current and under-development VO architecture for radio astronomy.

Accurate characterisation of the dataset is mandatory given the variety of observing projects and the heterogeneity of the data hosted in the Radio Archive. A thorough analysis of the necessary information to be stored in

the header has been conducted prior to the definition of the FITS metadata content for single dish data. Additional provenance information for a specific dataset can be retrieved also in the telescope log files and schedules, which are stored in the Archive as ancillary data. Other information to describe e.g. the UV coverage of a specific interferometric dataset can be added in graphical form but currently is yet to be implemented. Provenance information already stored as metadata in the raw data files is, for instance, the unique project ID of the scientific program and the actor performing the observation (Observer name). Other information on weather parameters and receiver performance (e.g. wind conditions and system temperature) are saved in the ancillary associated files. A more accurate analysis of the Provenance information needed to document all the phases of data reduction and assess the data products quality is planned in view of archiving also processed data. Provenance information for processed data, once stored in the Archive, will be added in different ways. Additional keywords will be used to describe the data reduction process with particular focus on the calibration steps (including RFI or atmospheric opacity removal). Also, we plan to use accompanying graphical plots to describe some specific characteristics of the data processing that benefit from a more visual approach, for instance the stability in the counts-to-Jansky transformation during prolonged observing sessions.

Global data discovery and access requires to expose a uniform, standard data model. To this aim, we have verified that the radio data model maps to the ObsCore data model, at least for what concerns the mandatory elements.

A TAP service, exposing an ObsCore table, is under preparation for the holdings of the Radio Archive.

A comparative analysis between the (internal) radio archive model and the CAOM one has been recently performed and can potentially address some specific use cases for provenance, metadata annotation and provide the basis for a more general discovery and access solution.

We are evaluating the usage of the DataLink protocol to access complex datasets like those hosted in the radio archive. For instance, to access single dish raw datasets, typically composed by a number of data files organized in a hierarchical structure plus some additional files containing ancillary information like the association with calibration files. Another DataLink use case could be relating additional metadata (for instance provenance and data quality) at different stages of the processing chain.

### **A.3 The Joint Institute for VLBI in Europe**

The Joint Institute for VLBI in Europe (JIVE) hosts the archive of all radio-astronomical VLBI observations of the European VLBI Network (EVN). The datasets are archived as raw interferometric visibilities in FITS-IDI format

(a registered FITS convention<sup>2</sup>) or in the deprecated UVFITS format for the oldest datasets. The latter format is deprecated because it is not an officially registered FITS convention. The raw visibilities are the result of processing ("correlation") and should be regarded as Data Level (DL) 1. The DL0 data, the raw digitized, sampled, voltages from the antennas, is too big to archive and are deleted after correlation.

Connected to the raw visibilities, the EVN archive stores initial calibration information, flagging information, diagnostic plots and the output products of running a generic pipeline. All these artefacts are either necessary to create science data products or are a valuable resource in assessing data quality and/or scientific potential.

JIVE is working towards making these raw visibilities and support information findable and downloadable through ObsTAP and DataLink services; none of the other IVOA services apply naturally to unprocessed interferometric data. The pipeline outputs crude images of a subset of the observed sources (the calibrator(s)). These images will be added as thumbnail/preview images to ObsCore entries through DataLink. There is no point in offering these through e.g. SIAP since the images are not at all science quality. An important envisaged use case of making the raw visibilities available through VO protocols is that this will allow users to find and download historic observations and reprocess them with modern software.

The FITS-IDI files contain sufficient meta-data to parse almost all the necessary ObsCore parameters from there. The raw visibility data sets tend to be large (2020 standards). Typical sizes vary from several GB to a few TB per dataset. Radio-interferometric datasets can represent a rather large theoretical FoV on the sky. Combined, these properties lead to a large probability of a VO query returning GBs of data which may turn out to be unuseable. Two parameters to characterize the useability of visibility datasets are proposed to be added to the ObsCore data model such that these parameters can be filtered on. The parameters eccentricity and filling factor of the u,v-plane are currently being experimented with in ASTRON and JIVE to assess feasibility in computing and usefulness in practice.

It might be possible to create a union `s_region` of all the FoVs over all EVN observations in the archive and offer that as a MOC, however this is not a primary goal. If time and resources allow it might be considered.

#### A.4 The Murchison Wide-Field Array

The MWA supports the implementation of various VO standards, which helps to ensure that all data follows FAIR access principals. Our usage of some of these protocols are described below:

---

<sup>2</sup><http://fits.gsfc.nasa.gov/registry/fitsidi.html>

TAP - The MWA ASVO provides an IVOA compliant Table Access Protocol (TAP) service, which can be used by any VO/TAP compliant software such as TOPCAT to retrieve WMA observation metadata. Our TAP service supports the standard IVOA "ObsCore" schema, in addition to some new MWA-specific schema for richer observation metadata. Our implementation is courtesy of CASDA VOTools [https://github.com/csiro-rds/casda\\_vo\\_tools](https://github.com/csiro-rds/casda_vo_tools).

VOTable - Users of the MWA ASVO are able to download observation metadata in a VOTable XML format which is compatible with widely used astronomical software.

Cone Search - The MWA ASVO implements the simple cone search protocol allowing users to quickly and easily find astronomical sources of interest.

ADQL - In addition to our custom-developed search interface, users can also query the MWA TAP service using Astronomical Data Query Language (ADQL).

SSO - While not specific to the IVOA, the MWA ASVO implements a number of different authentication methods which are approved by the IVOA. These include logging in with a large number of federated identity providers with SAML, as well as oAuth integration with AAO Data Central.

## A.5 CADC

because we use the CAOM data model for all collections, implementation has been at 3 levels: CAOM workshops have made a point of involving people working with radio and millimetre data so their input on supporting radio data use cases drove some of the evolution of the CAOM data model (currently at version 2.4) CADC has always been involved in the IVOA and has taken on the effort to support VO protocols and models within the core of the CADC software architecture. In that sense, all CADC users use "the VO". An example: if a data centre is building a query page, build it with a TAP backend. The data engineering work for a new collection is third type of implementation. This is the effort to map a collection's metadata into CAOM2 concepts. That involves working with science domain experts both at the source of the data and at CADC, and with developers that translate the mapping into code (pipelines) built from common libraries and customized for the collection. These pipelines create and persist CAOM instances from metadata databased and/or from files.

## A.6 Nançais

### A.6.1 NDA :

An EPN-TAP service, distributed through the Virtual European Solar and Planetary Access (VESPA, <http://vespa.obspm.fr/planetary/data/>), en-

ables accessibility to the Nançay Decameter Array (NDA, <https://www.obs-nancay.fr/reseau-decametrique/>) products from its Routine receiver. The NDA is a phased array of 144 helicoidal antennae divided in to sub-arrays of 72 antennae, sensitive to either Right-Handed or Left-Handed circular polarizations. This low-frequency radio instrument is observing the Sun and Jupiter on a daily basis and record Routine dynamic spectrum data within the 10 – 40 MHz bandwidth at a time sampling of 1 sec and a frequency sampling of 75 kHz. Three types of products may be queried from its associated EPN-TAP service: - Raw data files in binary format; - Raw data converted to Common Data Format (CDF) files, including richer metadata compliant with the International Solar Terrestrial Program (ISTP) specifications; - Observation quicklooks in PDF format for both polarizations.

### A.6.2 NRH / ORFEES :

Two EPN-TAP services, accessible through the Virtual European Solar and Planetary Access (VESPA, <http://vespa.obspm.fr>), are prepared to share space weather and solar observations acquired in Nançay, with the Nançay Radio Heliograph (NRH, <https://www.obs-nancay.fr/radioheliographe/>) and ORFEES (Observations Radio pour FEDOME et l'Etude des Eruptions Solaires, <https://www.obs-nancay.fr/orfees/>). The NRH is a T-shaped interferometer composed of 47 antennas, and observing between 150 MHz and 450 MHz, at a spatial resolution of about 2.5 arc-minutes. High temporal resolution data is available, as well as low resolution 2D images. The ORFEES instrument is a spectrograph coupled with a 5m diameter antenna, observing the Sun from 130 MHz to 1 GHz, with a 100 ms temporal resolution. The data are accessible from the Radio Solar Data Base portal (<https://rsdb.obs-nancay.fr>) with an interactive interface. The NRH and ORFEES data files available through EPN-TAP are: - ORFEES dynamic spectra (FITS format) - NRH image data files (raw binary format) - NRH image data files (animated GIF files)

### A.6.3 NenuFAR :

NenuFAR (New Extension in Nançay Upgrading LOFAR, <https://nenufar.obs-nancay.fr/en/astromer/>) is a new radioastronomy instrument, operating in the 10 – 85 MHz frequency band. Data can be acquired through four distinct observing modes, which rely on specific receivers and instrument configurations. An Obs-TAP service will be provided for the 'Standalone Beamformer' mode. In this operating configuration, NenuFAR acts as a low-frequency phased array instrument, beamforming data from 96 sub-arrays called 'Mini-Arrays'. The latter are consisting of 19 analog-phased crossed-dipoles gathered in hexagonal tiles, each of them rotated with respect to the others. Dynamic Spectrum data computation is achieved thanks to the

UnDySPuTeD backend, with time and frequency resolutions ranging from 0.30 to 84.00 ms and from 0.10 to 12.20 kHz respectively. Reduced data (cleaned of radio-frequency interferences, corrected from instrument systematics, averaged in time/frequency and/or selected over given time periods or frequency bands) will be distributed, either in FITS or HDF5 format. Quicklooks will also be delivered in PDF format.

## A.7 CSIRO Radio Archives

### A.7.1 CASDA

The CSIRO ASKAP Science Data Archive (CASDA) aims to provide science ready data products from the Australian Square Kilometre Array Pathfinder (ASKAP) telescope. We use FITS format for images, cubes, spectra and moment maps, and VOTable format for Catalogue data. In addition we provide validation data-products as tar files and calibrated visibilities for some observations in tarred CASA measurement set format.

We provide three means of discovering and accessing data, a) via the website <sup>3</sup>, b) using VO protocols and c) using the astroquery.casda module. Discovery for the web site uses an ElasticSearch index while the TAP, SIAP and SSAP services query our PostgreSQL metadata database. Our Implementation of these protocols is available publicly as CASDA VO Tools <sup>4</sup>. All science data products are listed in the ObsCore view, however metadata for visibilities is limited. Data access for all three access paths goes through the same module implementing the Datalink and SODA protocols. Currently all data are held on tape and, if not already online, are recalled on demand to disc as part of the data access process. We are working to move to an object store and have all data online.

During deposit of data products we extract metadata from FITS and VOTable files including position and observation information. These are then stored in a PostgreSQL database which is directly queried by the VO services and is used to populate the ElasticSearch index. As part of the processing, validation reports are produced for each observation including quality metrics. These can be viewed on the website and queried via TAP. The science teams are also validating the data products and providing quality ratings and notes. These can be accessed via the website and the quality ratings can be found in all data-products in their ObsCore listing.

### A.7.2 PSRDA

The Pulsar Data Archive (PSRDA) provides pulsar observations taken predominantly at the Parkes 64m radio telescope. All data are in PSRFITS

---

<sup>3</sup><https://data.csiro.au>

<sup>4</sup><https://doi.org/10.25919/zy6w-h884>

format and data include both search and fold mode observations. The index of observations may be queried via TAP. Data access is only provided via the website <sup>5</sup>.

During deposit, metadata such as target, central frequency, the frontend and backend used, and the duration of the observation are extracted from the FITS headers and stored in a PostgreSQL database. This is indexed by ElasticSearch for use in queries on the website. Currently all data are held on tape and, if not already online, are recalled on demand to disc as part of the data access process. Future work will bring all data online via an object store.

### **A.7.3 ATOA**

The Australia Telescope Online Archive (ATOA) serves spectral line observations from both the Australia Telescope Compact Array (ATCA) and Parkes. Most ATCA data is in the form of visibilities in RPFITS format, with files split by time and thus containing multiple sources and spectral windows. Older spectral line Parkes data are in SDFITS format while data from the latest Parkes receivers are in HDF5 format with metadata based on SDFITS.

Metadata from the RPFITS, SDFITS and HDF5 files are retrieved as part of deposit processing and stored in a MySQL database. Web searches on the ATOA site <sup>6</sup> use this database. A copy of the metadata is made to a PostgreSQL database where it is made available for TAP queries against an ObsCore schema. All data are online and available to both web and VO users for download.

## **A.8 ASTRON VO services**

### **A.9 MERLIN**

---

<sup>5</sup><https://data.csiro.au>

<sup>6</sup><https://atoa.atnf.csiro.au/>