europlanet 2024 Research Infrastructure

H2020-INFRAIA-2019-1

Europlanet 2024 RI has received funding from the
European Union's Horizon 2020 Research and Innovation Programme under

Grant agreement no: 871149

Deliverable D6.11

VESPA documentation of standards

Due date of deliverable:	31/03/2024
Nature[1]:	R
Dissemination level[2]:	PU
Work package:	6
Lead beneficiary:	ObsParis
Contributing beneficiaries:	
Document status:	Final

Start date of project:	01 February 2020
Project Duration:	54 months
Co-ordinator:	Prof Nigel Mason, University of Kent

1. **Nature:** R = Report, P = Prototype, D = Demonstrator, O = Other
2. **Dissemination level:**

PU	PP	RE	CO
Public	Restricted to other programme participants (including the Commission Service)	Restricted to a group specified by the consortium (including the Commission Services)	Confidential, only for members of the consortium (excluding the Commission Services)

Executive Summary / Abstract:

This document summarises and links to standards defined or used by the VESPA data infrastructure.

Table of Contents

1 VESPA Elements

The goal of VESPA is to set up and finalize a global distribution and search system for Solar System data consistent and compliant with the Virtual Observatory (VO), the mature infrastructure used for Astronomy data.

The VESPA ecosystem consists of 4 major original elements:

1) A Data Access Protocol able to describe and query the main parameters in this field, including coverages, data type, observational and instrumental conditions, and physical quantities. This is the EPN-TAP protocol.

2) A dedicated EPN-TAP client demonstrating the capacities of this protocol and its interactions with visualisation and analysis tools existing in the VO. This is the VESPA portal, a web interface mostly oriented toward the discovery of data content. Alternative data access paths are also provided.

3) Data services responsive to the EPN-TAP protocol, provided either by VESPA partners, other work packages in Europlanet, or the broad Solar System community through collaborative work.

4) Tools to visualise and analyse metadata and data. To save unnecessary development time and efforts, VESPA asks VO developers to add specific support of Solar System data in existing VO tools (e.g., reflected light, coordinate systems, etc), and connects external tools of interest to the VO (e.g., GIS, time series, image processing, etc).

This deliverable focuses on standards used to make these 4 elements work together. At the end of the feasibility study conducted in the Europlanet RI programme (2009-2012), VESPA identified the astronomy Virtual Observatory (VO) as the most promising framework to develop a data infrastructure for the Solar System. The work has therefore consisted either of updating the existing VO standards to handle these data, or on some occasions of designing the entirely missing parts.

2 VESPA new standards for Solar System data

2.1 EPN-TAP

EPN-TAP (Europlanet-TAP) is an original standard to publish and access solar system data in the VO. It relies on the standard TAP mechanism (Table Access Protocol) from the International Virtual Observatory Alliance (IVOA), associated with a vocabulary describing these data (EPNCore) and a number of syntactic rules.

EPN-TAP v2.0 has been approved in Aug 2022 by the IVOA, and has become the international standard to publish Solar System Data in the VO: https://ivoa.net/documents/EPNTAP/

EPN-TAP is a restriction of the TAP standard and is therefore supported by all TAP clients: TOPCAT (tables), Aladin (images and cubes), CASSIS and SPLAT-VO (spectra), and web clients such as TAPhandle. In contrast, the VESPA (https://vespa.obspm.fr) portal implements the whole capacities of EPN-TAP, and specific EPN-TAP clients are included in the CASSIS and 3Dview tools.

EPN-TAP specificities have been implemented in the DaCHS framework used to implement data services (as a mixin of predefined parameters), and in the taplint service validator (associated to the TOPCAT tool). A specific EPN-TAP mode is implemented in the IVOA Services Validation Portal at PADC: https://voparis-validation-reports.obspm.fr/

EPNCore includes both mandatory and optional parameters. Mandatory parameters allow generic queries across all data services, and the ability to discover data collections unknown to the user – this functionality is implemented in the VESPA portal and is unique in the VO so far. Optional parameters describe the data in more details and are typically defined in groups, or thematic extensions, by agreement among experts in each field. To address the whole field of Planetary Science data, this approach proved much more rapid, efficient, and flexible than designing a complex data model.

2.2 TFCat

TFCat (Time-Frequency Catalogue) is a data interchange format based on JSON. It defines several types of JSON objects and the manner in which they are combined to represent time-frequency features in a time spectrogram (a.k.a. dynamic spectrum), their properties, and their temporal and spectral extents. The standard is available on the MASER web site: https://doi.org/10.25935/6068-8528

TFCat is associated to an original python module and a labelling tool: https://zenodo.org/records/7821364

See also :

Cecconi, B., Louis, C.K., Bonnin, X., Loh, A., and Taylor, M.B. (2023), Time-frequency catalogue: JSON implementation and python library. Front. Astron. Space Sci. 9:1049677. https://doi.org/10.3389/fspas.2022.1049677

Louis et al (2022). The “SPectrogram Analysis and Cataloguing Environment” (SPACE) labelling tool. Front. Astron. Space Sci., 9:1001166. https://doi.org/10.3389/fspas.2022.1001166

D6.5- New radio services: MASER

2.3 Solid Spectroscopy Data Model (SSDM)

The "Solid Spectroscopy Data Model" (SSDM) is a relational data model allowing a complete description of spectral data of solid materials. It includes a detailed description of the solid samples through their layers, materials, constituents and species. Natural matters have their specific description. The instruments and techniques used for the measurements are also described.

SSDM is the basis of the SSHADE database infrastructure of spectroscopy of solids, and has been first published in Europlanet 2020 RI (July 2018). It has been deeply upgraded during Europlanet 2024 RI to support more materials and descriptions.

The document is available from this page:
https://wiki.sshade.eu/sshade/documentation/ssdm

2.4 Bandlist Data Model (SSDM-BL)

Bandlists are a parameterised approach to describing spectra in solid phase, with promising applications to the interpretation of observational data.

The SSDM-BL data model, partly inherited from SSHADE-Spectra, describes the various parameters of a band list and its bands (position, width, intensities, errors, mode attribution, environment, references, …). It is implemented in the SSHADE bandlist database and interface, both released during Europlanet 2024 RI.

The document is available from this page:
https://wiki.sshade.eu/sshade/documentation/ssdm

See also:
MS83- final SSHADE bandlist database

2.5 ObsFacility

ObsFacility is an on-going standard to help match possible names or identifiers of observing facilities, starting with space missions and ground observatories. It derives from cross-analysis of various institutional sources and the Wikidata database (maintained through public contributions) and will eventually provide a standard vocabulary to refer to observing facilities, plus an API to resolve their various designations.

ObsFacility is a shared activity in the frame of the IVOA, with a major contribution from Europlanet VESPA and inputs from other communities (SPASE, IHDEA, etc). It is currently available as an IVOA note:
https://ivoa.net/documents/Notes/ObsFacilityWikidata/20231115/NOTE-ObsFacilityWikidata-1.0-20231115.pdf

A resolver prototype is available at:
https://voparis-elasticsearch.obspm.fr/obsfacility/resolve/aliases?q=ad
https://voparis-elasticsearch.obspm.fr/obsfacility/resolve?q=ad

2.6 Cartography standards and Coordinate systems

Cartography standards are new in the VO and specifically related to Solar System studies but need to integrate smoothly with existing ways to provide celestial coordinates. They also need to be interoperable with standards used for Earth Observation, in particular those from the OGC consortium.

An extension of the fits WCS (Word System Coordinate) was proposed by Marmo et al (2018) during Europlanet 2020 RI to support surface coordinates on planetary bodies. It has been adopted and included in WCSlib 8.0 in July 2023, and support for bodycentric latitudes and longitudes was added to the astropy library by the VESPA team at this moment.

One difficulty with planetary coordinate reference systems (CRS) is their diversity, although their parameters are identified in successive publications by an IAU WG. A global naming scheme has been proposed by Hare and Malapert (2021), intended to be compliant with OGC usages – it is currently discussed in the OGC Planetary WG. If accepted, it will be used as the baseline in EPNCore at least for solid bodies, facilitating access of planetary scientists to tools from the Earth Observation community. A registry provides information on these CRS and supports associated projections in OGC-compliant format (provided by CNES): http://voparis-vespa-crs.obspm.fr:8080/web/

Concerning heliophysics and plasma-related coordinate frames, the SPASE consortium has already defined a consistent system which is being studied for implementation as an IVOA vocabulary and in future versions of EPN-TAP.

See also:
Marmo et al 2018 (EPN 2020): https://doi.org:10.1029/2018EA000388
WCSlib: https://www.atnf.csiro.au/people/mcalabre/WCS/wcslib/overview.html
astropy: https://www.astropy.org/
Hare and Malapert (2021) 5th PSIDA conference:
https://www.hou.usra.edu/meetings/planetdata2021/pdf/7012.pdf
MS40- Cartographic standards & dictionary of metadata OGC/VO/PDS

3 VESPA contributions to existing standards

VESPA is a big user of IVOA standards, and has brought a number of unexpected use cases in this context — VESPA is therefore a noticeable contributor to IVOA standards, in particular:

UCD (Uniform Content Descriptors) are used to provide a quick identification of physical quantities in astronomy. VESPA has defined many quantities that weren't considered by astronomers, in particular related to reflected light, coordinate systems, dynamics, spectroscopy of solid samples, etc.

Datalink is a protocol allowing to link datasets with various resources such as other related datasets, metadata or other services. Specific usages related to VESPA have been incorporated in recent versions.

HiPS (Hierarchical Progressive Survey) are multi-resolution maps based on the healpix tessellation system. Although originally designed to visualise celestial surveys, they proved to be a natural way to publish global planetary maps, provided some extensions of the associated metadata — nearly 70 planetary HiPS are now available, and are used in the VESPA geospatial portal prototype not only as contextual maps, but also as a source of data.

MOC (Multi-Order Coverages) provide spatial footprints in the form of healpix cell lists. They have been recently extended to space-time coverages and can potentially also cover the frequency axis. MOCs prove to be extremely efficient in support of geospatial information on planetary surfaces — such MOCs have been computed for all EPN-TAP services to support the prototypes of VESPA discovery portal and geospatial portal. Again, specific use cases from VESPA have been included.

VESPA uses many other IVOA standards (VOTable, SAMP, TAP, VOunits, VOevents, DALI, UWS, registry declarations and usage, etc) and some non-VO ones (das2, HAPI, etc). Regular feedback is provided to developers and WG in charge of these standards, and contributions to the new formal vocabularies of the VO are provided.

IVOA standards are available from this page:
https://ivoa.net/documents/

4 Prospects

Future versions of EPN-TAP will include updates and more parameter extensions to support new fields of activity. Work has started on the next version (v2.1), in particular related to coordinate systems (with OGC and SPASE standards), dynamical classes of objects (discussed with the PDS Small Bodies Node), events, bandlists, and other topics approaching maturity.

Interoperability with related fields will be improved further by relating terms of similar use in various contexts, e.g., heliophysics, PDS4 (space archives), OGC (Earth Observation), etc. Such ontologies are developed in the frame of new EOSC-related projects, in particular FAIR-IMPACT and OSTRAILS.