Virtual European Solar & Planetary Access (VESPA): A Planetary Science Virtual Observatory Cornerstone

S. Erard¹, B. Cecconi¹, P. Le Sidaner², C. Chauvin², A. P. Rossi³, M. Minin³, T. Capria⁴, S. Ivanovski⁵, B. Schmitt⁵, V. Génot⁶, N. André⁶, C. Marmo⁷, A. C. Vandaele⁸, L. Trompet⁹, M. Scherf⁹, R. Hueso¹⁰, A. Määttänen¹¹, B. Carry¹²,¹³, N. Achilleos¹⁴, J. Soucek¹⁵, D. Pisa¹⁵, K. Benson¹⁴, P. Fernique¹⁶ and E. Millour¹⁷

¹ LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, Meudon, FR
² DIO-VO/UMS2201 Observatoire de Paris/Université PSL/CNRS, FR
³ Jacobs University, Bremen, DE
⁴ INAF/IAPS, Rome, IT
⁵ Université Grenoble Alpes, CNRS, IPAG, FR
⁶ IRAP/UPS/CNRS, Toulouse, FR
⁷ GEOPS, Univ. Paris-Sud, CNRS, U. Paris-Saclay, FR
⁸ BIRA-IASB, Brussels, BE
⁹ OeAW, Graz, AT
¹⁰ UPV/EHU, Bilbao, ES
¹¹ LATMOS/IPSL, Sorbonne Université, UVSQ, U. Paris-Saclay, CNRS, Paris, FR
¹² Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS, Laboratoire Lagrange, FR
¹³ IMCCE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Lille, FR
¹⁴ University College London, UK
¹⁵ Inst. of Atmospheric Physics/CAS, Prague, CZ
¹⁶ Observatoire de Strasbourg/UMR 7550, FR
¹⁷ LMD/IPSL, CNRS, Sorbonne Université, École Normale supérieure, Université PSL, École polytechnique, Paris, FR

Corresponding author: S. Erard (stephane.erard@obspm.fr)

The Europlanet-2020 programme, which ended on Aug 31st, 2019, included an activity called VESPA (Virtual European Solar and Planetary Access), which focused on adapting Virtual Observatory (VO) techniques to handle Planetary Science data. This paper describes some aspects of VESPA at the end of this 4-years development phase and at the onset of the newly selected Europlanet-2024 programme starting in 2020. The main objectives of VESPA are to facilitate searches both in big archives and in small databases, to enable data analysis by providing simple data access and online visualization functions, and to allow research teams to publish derived data in an interoperable environment as easily as possible. VESPA encompasses a wide scope, including surfaces, atmospheres, magnetospheres and planetary plasmas, small bodies, heliophysics, exoplanets, and spectroscopy in solid phase. This system relies in particular on standards and tools developed for the Astronomy VO (IVOA) and extends them where required to handle specificities of Solar System studies. It also aims at making the VO compatible with tools and protocols developed in different contexts, for instance GIS for planetary surfaces, or time series tools for plasma-related measurements. An essential part of the activity is to publish a significant amount of high-quality data in this system, with a focus on derived products resulting from data analysis or simulations.

Keywords: Virtual Observatory; Solar System; GIS
Introduction
Modern space borne instruments often produce large and complex datasets, especially on long-lived missions. Detailed data analysis requires new ways to handle the data, in particular to locate specific observing conditions easily and efficiently. Virtual Observatory (VO) techniques have been developed in Astronomy during the past 15 years to address similar issues; they can be adapted to this context provided they are extended to take into account the specificities of Solar System studies. The VESPA (Virtual European Solar and Planetary Access) data access system focuses on applying VO techniques and tools to Planetary Science data, and supports all aspects of Solar System science (Erard et al. 2018). VESPA was developed in the framework of the EU-funded Europlanet-2020 programme, which started Sept 1st, 2015 and will be pursued in the recently selected Europlanet-2024 programme for another 4-year period. The objectives of VESPA are to facilitate searches both in big archives and in sparse databases, to enable data analysis by providing simple data access and on-line visualization functions, and to allow research teams to publish derived data in an interoperable environment as easily as possible. This system relies on studies and developments led in Astronomy (International Virtual Observatory Alliance, IVOA), Solar Physics (HELIO), Space Physics (SPASE), and space data archives (International Planetary Data Alliance, IPDA); it is responsive to FAIR principles and focuses on science-oriented description of data. We hereby provide a summary of achievements at the end of Europlanet-2020.

Data services
The VESPA architecture (Figure 1) consists in a new data access protocol, a specific user interface querying the available data services, and intensive usage of standards and tools developed for the Astronomy VO (Erard et al. 2018 and references therein). The Europlanet data access protocol, EPN-TAP, is based on the general Table Access Protocol (TAP) associated to a set of parameters describing the content of a data service (Erard et al. 2014). These parameters describe in particular the observational and instrumental conditions for each data element, which can thus be queried by the scientific user. Data services are required to return the metadata of matching results in VOtable format, which is supported by all standard VO tools.

Data services are installed in the provider institutes and are declared in the standard IVOA registries; they are thus always visible and searchable by query interfaces. A standard procedure has been identified for design and publication, based on the DaCHS server (Demleitner et al. 2014), although other solutions are possible. At
the time of writing, 54 data services are publicly open, and about 15 more are being finalized.\textsuperscript{1} They encompass a wide scope, including surfaces, atmospheres, magnetospheres and planetary plasmas, small bodies, heliophysics, exoplanets, and experimental data such as spectroscopy in solid phase. VESPA mostly focuses on derived data, typically associated to publications. Some of these derive from a significant preparation phase, e.g.: DynAstVO which computes orbital parameters for Near Earth Asteroids; APIS which provides data on planetary aurorae derived from several archives; SSHADE which currently gathers 17 experimental databases of spectroscopy of minerals, organics, ices, and cosmo-materials in a consistent and very detailed format; solar feature catalogues from the HELIO EU programme; a refined and corrected version of the Robbins & Hynek Mars craters database (being published) will also be accessible, together with the original database.

Some large existing data archives also received an additional EPN-TAP interface during the programme, in particular: ESA’s Planetary Science Archive (PSA), which gathers data from all European space missions (Besse et al. 2018); planets and satellites data collected from the Hubble Space Telescope archive in Canada; planetary plasma data at CDPP, Toulouse, which includes many unique derived products; the IAU Minor Planet Center which distributes the properties of most minor bodies; the Encyclopaedia of Extra Solar Planets, which is the world-wide reference in the field; several ground-based solar archives (currently BASS2000 and CLIMSO) also distribute their data through VESPA.

To favour the emergence of this kind of material, VESPA has organized a yearly call to the community to select projects of interest, typically associated to a publication; 4 or 5 selected teams were invited to a one-week workshop to design and install the service in their institute.\textsuperscript{2} Several services providing amateur astronomy data were also selected at the onset of the programme for implementation in research institutes, including PVOL (planetary images) in EHU/Bilbao (Hueso et al. 2018) and RadioJove (Jupiter radio measurements) at Paris Observatory. In addition, a special type of services will gather tables of VOEvents produced by alert systems in various fields (Cecconi et al. 2018b).

According to the contributory nature of the VO, any team can publish EPN-TAP services in the IVOA registry. A validator is available to certify TAP compliance, but manual checks are also performed to ensure consistency of EPN-TAP services prior to publication. By default the VESPA portal uses a local registry listing only services certified by the core VESPA team. The status of published services is monitored automatically through Nagios to inform data providers in case of problems.

Data access

EPN-TAP data services are best queried from the VESPA portal (Figure 2), an optimized search interface providing specific user support (http://vespa.obspm.fr). Alternate access modes are discussed below.

In the frame of TAP, data services consist in a list of “granules”, or data elements, described by a series of parameters. EPN-TAP defines a set of mandatory parameters that introduce metadata for individual granules; this scheme is inspired by the IVOA ObstTAP protocol for observational Astronomy datasets. EPN-TAP parameters provide the observational and instrumental conditions of acquisition, accounting for the specific diversity and complexity of Planetary Science: ranges along several axes (spatial, temporal, spectral, photometric), measurement type, origin of data, and several references. Localisation can be provided in various coordinate systems (celestial or planetary coordinates, either body-fixed or rotating). Time is provided in Julian days, but also as local time or season if relevant. The benefit of EPN-TAP is that all data services are described uniformly with parameters that are relevant for the science user. The VESPA portal queries all registered data services at once based on the mandatory EPN-TAP parameters, and returns individual granules matching the query; this allows in particular the user to discover unknown data content in the field of interest. In addition, thematic extensions are defined to describe new fields uniformly (e.g., lab spectroscopy). Specific parameters may also be made-up to describe individual services with more details; when querying a single service, they can be used to identify granules more precisely. Independently from VESPA and the data distribution context, the EPN-TAP Data Model (Erard et al., 2019 submitted to IVOA) can also be used or adapted to describe private databases in a consistent way, e.g. to share proprietary data inside an experiment team.

The VESPA portal also supports other query systems. User queries are converted and forwarded to space agencies’ PDS (Planetary Data System) archives: ESA and JAXA are queried via the PDAP protocol from IPDA, and NASA through its PDS keyword-search interface. Owing to limitations in these protocols, such queries are performed only at dataset level.

\textsuperscript{1} A roadmap for service publication is maintained here: https://voparis-wiki.obspm.fr/display/VES/EPN-TAP+Services.

\textsuperscript{2} Implementation tutorials are available: https://voparis-wiki.obspm.fr/display/VES/Implementing+a+VESPA+service.
EPN-TAP services can be accessed via other search interfaces. An EPN-TAP library was developed and included in several tools (3DView, CASSIS, and AMDA) to issue direct queries from these environments. Since EPN-TAP relies on the more general TAP mechanism, EPN-TAP data services can also be accessed individually via standard TAP clients; these include general query interfaces (e.g., TAPHandle) as well as standard VO tools (e.g., TOPCAT, Aladin, etc). Programmatic access is possible using existing libraries in python (pyvo), IDL (SSW), or shell scripts (TAPsh) – making such data services handy for pipeline processing. Finally, a mapping app prototype was also developed to explore new types of access from mobile and computer uses.

EPN-TAP tables either provide links to data files, or include the data itself when they consist in a small set of scalar quantities. A variety of tools are available to handle the data in the VO. Adequate VO tools are identified through data description parameters, which not only indicate a file format but also specify dimensions, units, and physical quantities, based on IVOA Data Models adapted for VESPA. For instance, images and spectra will open in different tools, and the spectral tools will recognize spectra in radiance or in reflectance, and handle them differently.

**Tools**

Selected metadata can be transferred from the VESPA portal to VO tools via the SAMP protocol from IVOA. Standard VO tools are connected to the VESPA portal so that they readily display metadata, e.g., spatial footprints are plotted on a 3D sphere in Aladin or Mizar; other metadata such as local time or instrument modes, can be plotted in 2D or 3D with TOPCAT (Figure 3).

The data themselves can be transferred in a similar way for display and standard analyses. Data description is used to select appropriate tools, e.g., TOPCAT handles all types of tabular data (Taylor 2017), Aladin most images and spectral cubes, CASSIS and SPLAT-VO spectra in general, 3DView displays observations of various types in 3D planetary contexts (Génot et al. 2018), Autoplot is dedicated to extracting data from long time series.

Most of these tools have been updated in collaboration with their developers to support Planetary Science and specificities of Solar System data, e.g., coordinate systems on surfaces and in magnetospheres, or measurements in reflected light (Figure 4). In some instances new tools have been developed for VESPA,

---

3 See Erard et al. 2018 and references therein for tools; IVOA standards are documented here: [http://www.ivoa.net/documents/](http://www.ivoa.net/documents/).

Figure 3: Location of Mars in the sky from historical telescopic images (1905 to 1976) displayed in TOPCAT, evidencing retrograde motion around each opposition (the colour scale provides time in Julian Days).

Figure 4: NIR telescopic spectra of 4 Vesta (from the m4ast service) compared to basaltic meteorites from the PDS spectral library in CASSIS, on a reflectance scale.
such as Planetary Cesium Viewer to visualise catalogues of features on planetary surfaces; APERICubes to slice PDS3 spectral cubes (Savalle et al. 2016); iPECMA to analyse multi-dimensional measurements of planetary electromagnetic fields (Pisa et al. 2017). Some existing non-VO tools have been provided with a SAMP interface to exchange data in a VO context and can be integrated in workflows, e.g.: Image] which now provides conversions for many formats, as well as image processing functions to the VO; QGIS, an Open Source GIS application; Autoplot (see below); MATISSE to plot data on 3D shape models (Longobardo et al. 2018). Finally, specific web tools developed as part of larger data services have been made accessible for use with external data, e.g.: AMDA for time series at CDPP (Génot et al. 2014), or the new SSHADE service for lab spectroscopy (Schmitt et al. 2018).

TOPCAT can easily integrate sparse surface observations (e.g. from a point spectrometer) using the healpix tessellation system, while Aladin can produce multi-resolution maps (HiPS) from large datasets, which allow for smooth and fast change of scale (Fernique et al. 2015). Currently, 60 planetary maps from USGS have been converted to HiPS and are available from the Aladin data tree. The same technique applied to large panoramas from planetary landers provides a very exciting way to navigate within such images, by smoothly changing from the global picture to the highest local details.

A significant activity is the development of a connection between the VO world and Geographic Information Systems (GIS). In a first step, EPN-TAP services were used to provide links as queries to WMS or similar services, i.e. using different, non-VO, access protocols. Traditionally, such links are only handled in GIS applications such as the open source QGIS. While the intermediate VO layer allows for powerful search functions in the databases, cross-examinations with other datasets remains difficult because of the variety of query systems and image formats (e.g., Hare et al. 2018). In a second step, the goal was to provide bridges between these two worlds, so that VO (e.g., fits) and GIS (e.g., geotiff) images can be displayed in all applications. This is done on one hand by providing improved georeferentiation support in fits headers and conversion routines in the GDAL library, which is widely used to import data in applications such as QGIS (Rossi et al. 2016, Marmo et al. 2018, Figure 5), on the other hand with new QGIS plug-ins to add SAMP connectivity (Minin et al. 2019).

A similar situation applies to time series depicting radio emission of the planets. A protocol of choice in this case is das2, which allows the distribution of data with adaptive temporal resolution. Data services are

![Figure 5: A MESSENGER multispectral image of Mercury converted to geofits format is directly opened in QGIS and plotted on a coordinate grid through the updated GDAL library.](image-url)
responsive to EPN-TAP but provide data as queries to such servers, the results of which can be fetched via SAMP to the Autoplot tool for display (Cecconi et al. 2018a, Figure 6).

As far as 2D data are concerned, VESPA makes use of two IVOA protocols to handle footprints. The first one is the STC-S standard (used in particular by the ObsTAP protocol) which provides oriented contours; the second one is the Multi-Order Coverage (MOC, healpix based) used e.g. in Aladin, TOPCAT, and Mizar. Both standards can be used to issue powerful searches on intersections or inclusions, and to select objects within arbitrary footprints (Figure 7).

Figure 6: Cassini RPWS/HFR Jupiter Flyby Calibrated Dataset displayed in Autoplot, using the das2server interface.

Figure 7: Selection of Mars Reconnaissance Orbiter CRISM spectral cubes located in large Martian craters, based on MOC in Aladin (over THEMIS daytime HiPS).
Simulation services
Another important goal for VESPA is to connect online computation services with an interface similar to that of data services, so as to compare observations and simulations more routinely. This activity has obvious applications, e.g., for radiative transfer in planetary atmospheres or for magnetospheres, but also to connect ephemeris systems (e.g. Miriade) with data services. The datalink protocol of IVOA is used to call web services with parameters retrieved from existing data services, e.g.: Mars-Express SPICAM vertical profiles are linked to simulations from the Mars Climate Database; HST data to physical ephemerides from Miriade. More direct ways to launch simulations on demand are being set up (Trompet et al. 2017, Cecconi et al. 2018a). Independently, a new function called ViSiON has been developed on Miriade to help plan observations of planetary objects from arbitrary locations (Carry & Berthier 2018).

An additional aspect is to provide online low-level computation functions, e.g. averaging, resampling, deconvolution of actual data. This is currently supported only to some extent by standard VO tools and ImageJ; in addition, higher level processing such as retrieval of Hapke parameters from surface spectra, multivariate analyses, etc, would also be beneficial and are being investigated.

Building a community
Hands-on sessions have been organized twice a year at EGU and EPSC conferences in Europe to support new users, as well as contributions to similar workshops in Astronomy. Besides, a procedure has been developed to set up data services easily and with limited resources: a complete service can be installed within a week with no prior experience; the most challenging part is to provide adequate data description, which of course may be demanding for complex archives. This is expected to foster the installation of data services in research institutes, in particular to distribute derived data related to a publication. In parallel, discussions are held with big data providers, in particular space agencies in the frame of the IPDA. Finally, a Solar System Interest Group has been initiated in the IVOA in 2017, to which several VESPA partners contribute.

Prospects
At the end of Europlanet-2020, VESPA provides a consistent data distribution infrastructure open to contributions from the community. The data description system makes it possible to identify specific instrumental or observational conditions in datasets encompassing many areas of Solar System studies. The search system connects with powerful display and analysing tools. Collaborations started in several areas (in particular exoplanets, heliophysics, experimental spectroscopy, surface studies, small bodies, and radio observations of magnetized planets) will help coordinate these fields and make future data services interoperable.

During the coming Europlanet-2024 programme, VESPA will continue to set up new data services and to connect existing ones, e.g. by setting up bridges with PDS4 archives at NASA and ESA, and serving other space agencies as well. Connections between PDS4 and EPN-TAP dictionaries would make all PDS metadata searchable from the VESPA portal and vice versa. Another goal is to connect Solar System data present in astronomical VO catalogues, starting with the VizieR database (Ochsenbein et al. 2000). Finally, experimental work performed by other Europlanet activities will also be distributed in VESPA.

Using a light and distributed data system (as opposed to heavy and expensive data centres) to publish contributions from individual research teams is efficient, in particular in the context of Open Science and incitation to make research data widely available. However, there are drawbacks related to sustainability. Services may disappear or become obsolete whenever local interest vanishes, which often relies on a single person. A possible workaround, in addition to getting space and telescope agencies involved, is to rely on a series of regional centres taking care of existing services that are no longer maintained by the original publishers; in Europlanet-2024, Paris and Trieste observatories and the Heidelberg University, three major repositories of VO services, will act as regional VESPA hubs securing published resources. This is made easier if the data services are actually located on a cloud, rather than in the institutes; this may also provide a workaround to strict IT policy in some institutes. An obvious evolution will therefore be to consider the new EU-funded European Open Science Cloud (EOSC) being developed in the H2020 framework. Finally, other activities in Europlanet-2024 will develop state-of-the-art processing techniques in several fields, in particular related to planetary mapping and Machine Learning. These activities will make use of VESPA data services, and will distribute their results in VESPA. Such developments are retained necessary in view of future space missions such as BepiColombo and JUICE.
Acknowledgements
The Europlanet-2020 and Europlanet-2024 Research Infrastructure projects have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreements No 654208 and No 871149.
Support from Paris Astronomical Data Centre (PADC) is acknowledged.
VESPA web site: http://www.europlanet-vespa.eu.

Competing Interests
The authors have no competing interests to declare.

References
Cecconi, B, Le Sidaner, P, André, N, Tomasik, L, Gangloff, M and Marmo, C. 2018b. VOevent for Sun-Earth and planetary space weather. 42nd scientific assembly of the COSPAR, Pasadena, USA. Abst. PSW.4-0008.18


Submitted: 29 January 2019 Accepted: 11 January 2020 Published: 11 May 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.