VESPA: PROGRESS AND PROSPECTS. S. Erard[1], B. Cecconi[1], P. Le Sidaner[2], C. Chauvin[2], A. P. Rossi[3], M. Minin[3], T. Capria[4], S. Ivanovski[4], B. Schmitt[5], V. Génot[6], N. André[6], C. Marmo[7], A. C. Vandaele[8], L. Trompet[8], M. Scherf[9], R. Hueso[10], A. Määttänen[11], B. Carry[12,13], N. Achilleos[14], J. Soucek[15], D. Pisa[15], K. Benson[14], P. Fernique[16], E. Millour[17]. [1]LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité, 5 place Jules Janssen, 92195 Meudon, France [2]DIO-VO/UMS2201 Observatoire de Paris/Université PSL/CNRS, Fr, [3]Jacobs University, Bremen, Ge [4]INAF/IAPS, Rome, It [5]IPAG UGA/CNRS, Grenoble, Fr [6]IRAP/UPS/CNRS, Toulouse, Fr [7]GEOPS/CNRS/U. Paris-Sud, Fr [8]BIRA-IASB, Brussels, Be [9]OeAW, Graz, Aut [10]UPV/EHU, Bilbao, Sp [11]LATMOS/IPSL, UVSQ, U. Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, 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, ENS, Université PSL, École polytechnique, Paris, Fr

Introduction: The VESPA data access system focuses on applying Virtual Observatory (VO) techniques and tools to Planetary Science data, and supports all aspects of Solar System science [1]. VESPA (Virtual European Solar and Planetary Access) is developed in the framework of the EU-funded Europlanet-2020 program started Sept 1st, 2015 for 4 years. The objective of this activity is to facilitate searches in big archives as well as in sparse databases, to provide simple data access and on-line visualization tools to users, and to allow small data providers to make their data available in an interoperable environment with minimum effort.

As we approach the end of the contract, we provide a summary of achievements of the current program. A new proposal has been recently submitted for a 4th Europlanet program, in which VESPA will evolve from the current "data access" concept to a new "enabling data analysis" concept, while securing existing and future data resources. If selected, this program will start in February 2020.

Data access system: VESPA adapts the concepts of the Astronomical Virtual Observatory to Planetary Science and Solar System studies in general, including planetary plasmas and heliophysics. The paradigm is to connect distributed data services with a light infrastructure: data resources are delocalized and declared in a central registry; clients formalize user queries and send them to the data services; results can be forwarded to visualization and analysis tools, which can also exchange data.

The infrastructure relies on standards and tools developed in Astronomy (International Virtual Observatory Alliance, IVOA [2]), Solar Physics (HELIO), and space data archives (International Planetary Data Alliance, IPDA [3]), and connects other developments. It is responsive to FAIR principles, focuses on science-oriented description of data, and matches the definition of a Planetary Spatial Data Infrastructure [4] but extends its scope, beyond planetary surfaces, to all areas of Solar System studies.

VESPA uses the same IVOA registry of services as the Virtual Observatory, therefore following its standards. Data services are provided and hosted by the community, and may have very different sizes and contents; the largest one is currently ESA's Planetary Science Archive, but many research teams contribute directly with services of derived data, typically related to a publication or an experiment. All these data resources have a similar interface. The main VESPA Data Access Protocol is based on a global IVOA mechanism called TAP (Table Access Protocol), which formalizes queries and answers from data servers based on a specific SQL-type language. The novelty of VESPA is to associate this mechanism to a set of uniform parameters (EPN-TAP) describing the data content: position along several axes (spatial, time, spectral, but also illumination angles), target properties, observing and acquisition modes, origin and credits, etc. This set of parameters is extensible to accommodate new services and new fields, but the core of mandatory parameters allows the user to query many services all at once, favoring the discovery of new datasets.

Clients and services: the main VESPA portal is optimized to query all connected services together on common parameters; results are accessed for each service independently. Special modes to access individual or non-public services are also implemented. An all-VO mode mixing results from many services is being devised. Alternatively, all TAP clients (e.g., TapHandle TAPsh, and many VO tools) can access EPN-TAP services directly, although with lesser support. In all cases, the EPN-TAP protocol allows to locate data files selected from conditions which make sense for the scientific user.

The VESPA portal currently connects 48 data services, and 10 more will be open within the end of the

program. They include large archives (ESA's Planetary Science Archive, containing datasets from 70+ space borne instruments [5], Hubble planetary data, IAU's Minor Planets Center catalogue, etc.), world-reference services (the Encyclopaedia of Exoplanets, CDPP plasma data in AMDA, the Mars Climate Database simulations, USGS maps, crater catalogues, etc.) and data derived for VESPA from science analysis or computation (APIS for aurorae, DynAstVO for Near Earth Asteroids, event predictions, or the large solid spectroscopy service SSHADE). Databases derived from past EU programs are accessible from VESPA (e.g. HELIO, IMPEx, PlanetServer and SBNAF databases, etc), as well as derived products from several space instruments and large telescopes not available otherwise.

Tools: Data provided in convenient formats can be forwarded from the portal to VO tools for visualization and analysis. Several standard IVOA tools are connected via the SAMP protocol: TOPCAT (tabular data), Aladin (images and cubes), CASSIS and SPLAT-VO (spectra), Miriade (ephemerides computation), SsODNet (name resolver), etc. Some new or external tools were also connected with new plugins: 3DView (plots along a spacecraft trajectory), MATISSE, Mizar and Planetary Cesium Viewer (3D plots, including on shape models), AMDA and Autoplot (time series), iPECMAN (analysis of electromagnetic fields), ImageJ (image processing and conversions), etc.

In addition, a VO-GIS bridge was developed in the course of the current program. EPN-TAP services can provide links to WMS or WCS services, and QGIS plugins were developed to exchange data with VO tools via SAMP. An extension of the *fits* format supporting georeferentiation on planets has been devised and a driver was merged into the official distribution of GDAL [6]. This provides QGIS and other software with the ability to import such images.

All developments are open source and available on VESPA github.

Prospects: The next program will focus on two directions: enlarging and securing the data content; enabling modern data analysis. The first point will include the setup of regional hubs; those will mirror the existing data services, and will take over in case of difficulty from the original data providers. This "stewardship program" will ensure service sustainability in the long term. Further solutions will also be studied, including data storage on the European Open Science Cloud, sustainable data packaging (e.g., Docker), and emerging RDA standards. DOI will systematically be affected to datasets in VESPA, and possibly to query results. Finally, all data produced in Europlanet-2024 (from experimental activities, field studies, or telescopic campaigns) will be distributed in VESPA data services. Interfaces with existing data distribution systems will be enlarged: the portal currently queries the PDAP services at ESA and JAXA; this will be complemented with additional accesses to PDS and ESA services such as the PDS keywordsearch service [7], the OPUS or ODE platforms. This interface will rely on a dictionary of metadata from various standards (OGC/PDS/VO).

The second main activity will be to develop new tools to exploit the data made available in VESPA, and to produce more derived results. First, a run-on-demand platform will be installed on local clusters to run specific simulations in several fields (radio observations, exoplanets atmospheres, etc). Second, specific data analysis tools will be made available to exploit the observations, such as spectral signatures matches or photometric inversions; this will also include further developments of the VO-GIS bridge, e.g. related to geofits formatting of imaging spectrometry data [8]. Finally, the VESPA data services will be used as inputs by mapping and machine learning (ML) activities developed in Europlanet-2024. The former will establish a solid infrastructure to support the production of planetary geological maps and related products following standard procedures; the latter will develop ML-powered data analysis and exploitation tools targeting selected representative scientific cases for specific applications of ML in Planetary Science.

References:

[1] Erard et al 2018, *Planet Space Sc.* 10.1016/j.pss.2017.05.013 [2] http://www.ivoa.net [3] https://planetarydata.org [4] Laura et al 2017 *Int. J. Geo-Inf. 6*, 181 [5] Besse et al. 2018, *Planet Space Sc.* 150, 131–140, 10.1016/j.pss.2017.07.013 [6] Marmo et al 2018 *Earth and Space Science* 5, 640-651 10.1029/2018EA000388 [7] Le Sidaner et al 2019, this conference [8] Marmo et al 2019, this conference.

Acknowledgements:

The Europlanet 2020 Research Infrastructure project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.

VESPA web site: http://www.europlanet-vespa.eu