

# Applications of Jupyter Notebook to VO-GIS interoperability

Mikhail Minin (1), Angelo Pio Rossi (1), Baptiste Cecconi (2), Chiara Marmo (3) and Stéphane Erard (2)

(1) Jacobs University Bremen, Bremen, Germany, (2) LESIA, Observatoire de Paris/CNRS/UPMC/Univ. Paris-Diderot, Meudon, France, (3) GEOPS, Univ. Paris-Sud, CNRS, Univ. Paris-Saclay, Orsay, France

## 1. Introduction

Jupyter Notebook [1] is an interactive environment inside a web browser capable of running a variety of scripting languages, primarily python. The notebook allows annotating code with markdown text and multimedia, thus allowing presenting reproducible workflows accompanied by a rich narrative.

Jupyter Notebook can be connected to Virtual Observatory (VO) tools [2], VO registries (GAVO DaCHS servers) [3], QGIS [2], as well as geospatial web servers such as PlanetServer [4, 5] by implementing appropriate protocols (Figure 1).

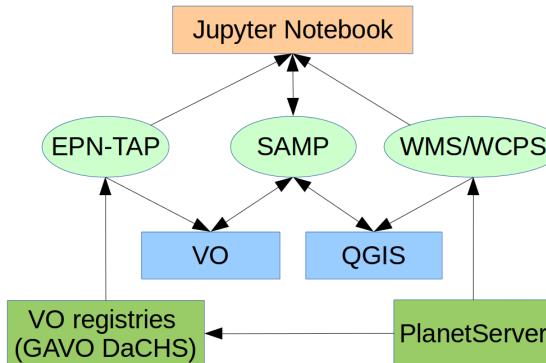


Figure 1: Enabling VO/GIS interoperability in Jupyter Notebook with EPN-TAP, SAMP, and WMS / WCPS protocol interfaces.

## 2. VO Registries

Access to Planetary Science data in Virtual Observatory is provided by VO Registries (GAVO DaCHS servers) through EPN-TAP data access protocol [1]. The requests to the Table Access Protocol (TAP) endpoint are formatted in Astronomical Data Query Language (ADQL). Astropy affiliated package Pyvo can be used to interface with TAP Service from Jupyter Notebook, thus removing a need to use intermediate

data mining software, while at the same time making it easy to share TAP queries.

Functionality similar to TOPCAT/STILTS [6] is available through STILTS python wrapper, allowing complex workflows involving tables from different sources, and data visualization. Furthermore, capabilities are extended through data analysis libraries available for Python and R in Jupyter.

## 3. SAMP interface

The communication between different VO tools is done by RPC/XML messages sent through Simple Application Messaging Protocol (SAMP) interfaces [2]. SAMP clients connect to a SAMP hub, which routes the messages between the clients. Files are shared between SAMP clients by sending an access link. A SAMP interface for QGIS exists [2], and can be easily added to Jupyter Notebook by instantiating it from Astropy library [7].

The default standard for tabular data in VO is VOTable, support for which in python is provided in Astropy library. Geospatial datasets can be converted to VOTable in Jupyter and forwarded via SAMP to other VO tools. Alternatively, Jupyter Notebook can be used to simplify the process of Resource Descriptor creation for ingesting new data into GAVO DaCHS. This has the potential to streamline the process of data publication to VO registries.

Jupyter Notebook connection to SAMP allows for rapid prototyping of work-flows to geospatial problems. SAMP interface between QGIS and Jupyter Notebook would allow seamless interoperability and data sharing.

## 4. PlanetServer

PlanetServer is a web GIS providing access to hyperspectral raster data cubes for Moon and Mars served by Rasdaman Array DBMS [4, 5]. Hyperspectral coverages stored on PlanetServer can be accessed via HTTP using Web Coverage Processing Service

(WCPS), additionally a python API exists for generating common derived data products [8].

## 5. Summary and Conclusions

Smoothen interoperability between geospatial and astronomical software is an ongoing effort. Jupyter Notebook is a platform which can accommodate WCPS, SAMP and TAP interfaces, while presenting reproducible workflows with rich annotations. Capacity to interface simultaneously with different systems enables development of high degree of automation, assisting machine learning. For instance, a workflow may involve discovering hyperspectral data with EPN-TAP, accessing data with WCPS, analyzing it with numpy library, then forwarding the results via SAMP interface to QGIS or Aladin for plotting. Within a Jupyter Notebook such a workflow could be fully automated, easily modifiable, shareable and reproducible. Derived data can even be published back to the Virtual Observatory.

An example of using VO interfaces in Jupyter Notebook can be found at [9].

## Acknowledgements

This work benefits from support of VESPA/Europlanet. Europlanet 2020 RI has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.

## References

- [1] Ragan-Kelley, M., Perez, F., Granger, B., et al., "The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication." AGU Fall Meeting Abstracts. 2014.
- [2] Erard, S., Cecconi, B., Le Sidaner, P., et al. "Vespa: a community-driven virtual observatory in planetary science", Planetary and Space Science, Vol. 150, pp. 65–85, 2017.
- [3] Erard, S., Cecconi, B., Le Sidaner, P., et al. "The epn-tap protocol for the planetary science virtual observatory", Astronomy and Computing, Vol. 7, pp. 52–61, 2014.
- [4] Figuera, R. M., Huu, B. P., Rossi, et al. "Online characterization of planetary surfaces: Planetserver, an open-source analysis and visualization tool", Planetary and Space Science, Vol. 150, pp. 141–156, 2018.
- [5] Marco Figuera, R., Pham Huu, B., Rossi, A. P., et al., "PlanetServer – A web GIS and Python API for planetary hyperspectral images analysis", this meeting.
- [6] Taylor, M. B., "TOPCAT & STIL: Starlink Table/VOTable Processing Software", in Astronomical Data Analysis Software and Systems XIV, Astronomical Society of the Pacific Conference Series, Vol. 347, p. 29, 2005.
- [7] Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al., "Astropy: A community python package for astronomy", Astronomy & Astrophysics, Vol. 558, A33, 2013.
- [8] Halder, A., Marco Figuera, R., Rossi, A. P., et al., "PlanetServer Python API–Visualization and Analysis of CRISM images." Lunar and Planetary Science Conference, Vol. 48. 2017.
- [9] Applications of Jupyter Notebook to VO-GIS interoperability, GitHub gist, <https://gist.github.com/mminin/01a4ae4af245efa96eb9692306cb18a9>, accessed May 2018.