



# eur PLANET 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 D7.6

<b>Deliverable Title:</b>	Run on demand platform (OPUS) with applications to Planetary Science context & Study for installation on EOSC
<b>Due date of deliverable:</b>	30/09/2022
<b>Nature<sup>1</sup>:</b>	R
<b>Dissemination level<sup>2</sup>:</b>	P
<b>Work package:</b>	7
<b>Lead beneficiary:</b>	OBSPARIS
<b>Contributing beneficiaries:</b>	
<b>Document status:</b>	Final

<b>Start date of project:</b>	01 February 2020
<b>Project Duration:</b>	48 months
<b>Co-ordinator:</b>	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 deliverable concerns the adaptation of the OPUS workflow platform to planetary science data in VESPA, and an assessment study of its use in the EOSC environment. The deployment on EOSC servers has been successfully achieved and provides a strong basis for future applications in the programme. Specific configurations adapted to particular applications will be studied in a future step.

### 1. Introduction

**OPUS** (Observatoire de Paris UWS System) is a job control system that aims at facilitating the access to analysis and simulation codes through an interoperable interface. OPUS was initially developed in support of the large CTA (Cherenkov Telescope Array) observatory and in the frame of the ASTERICS H2020 programme around 2016.

The Universal Worker System pattern v1.1 (UWS), as defined by the International Virtual Observatory Alliance (IVOA), is implemented as a REST service to control the asynchronous execution of a job on a work cluster. OPUS also follows the recent IVOA Provenance Data Model recommendation to capture and expose the provenance information of jobs and results. By following well-defined standards, the tool is interoperable and jobs can be run either through a web interface or a script, and can be integrated to existing web platforms.

The code is publicly available on Github:

<https://github.com/ParisAstronomicalDataCentre/OPUS/>

The documentation can be accessed at <http://opus-job-manager.readthedocs.io/>

A first assessment of OPUS in the VESPA context was performed during the Europlanet 2020 RI programme (D11.11, [http://www.europlanet-2020-ri.eu/sites/open.ac.uk.europlanet/files/ePN2020\\_d11.11.pdf](http://www.europlanet-2020-ri.eu/sites/open.ac.uk.europlanet/files/ePN2020_d11.11.pdf)), and is used by ESA to prepare the JUICE space mission. The current study is about porting OPUS on a local cloud, then studying possible uses in the EOSC (European Open Science Cloud) environment, and identifying possible hard points requiring further developments.

Future implementations can be performed at other participants of the VESPA activity; in particular OPUS will be used by the SPIDER Work Package.

### 2. Use Case & Overview of Progress

#### a) Objective

OPUS was already installed at Paris Observatory and used by the CTA and ComPOSE projects in astronomy. A first application to perform mass processing on VESPA EPN-TAP data services was also implemented in Europlanet 2020 RI.

The objective is:

- a) to adapt OPUS on the local computational cluster in Observatory of Paris,
- b) to test it with EOSC-related functionalities and to identify possible blocking points

### **b) Use cases**

Several tests were performed:

1) Previous use cases focused on requesting runs of codes developed within the MASER project, most notably the Exoplanetary and Planetary Radio Emission Simulator (ExpRES) simulator.

Since it uses computational resources at the Observatory of Paris, this instance of OPUS is not public but access-protected, see <https://voparis-uws-maser.obspm.fr/client/>. Past runs are accessible from the OPUS interface. They are also included and searchable in an EPN-TAP service (ExpRES service).

This workflow runs routinely in Observatory of Paris and can be installed in other places benefiting from significant computational resources. In particular, OPUS will be installed in CNRS/IRAP for use in the SPIDER Work Package (WP5).

2) A simple example workflow

This queries the M4ast EPN-TAP service for spectra of asteroid 4 Vesta.

The result is a set of table rows, each containing a URL to a spectrum file, plus associated metadata describing the observational configuration and provenance data. The files are in raw ascii format and are not self-described. They are converted to VOTable (a standard xml VO format) and the associated metadata are extracted to fill the VOTable header. The result is a self-described file, which is a standard way to publish spectra in the Virtual Observatory (VO); a thumbnail image (plot) is also created for each spectrum. This workflow uses the target name as an input parameter. It can be used to pre-process a data collection while setting up an interoperable service, e.g., when ingesting new spectral collections in the M4ast service.

This simple but realistic workflow concatenates an EPN-TAP query (as a shell script) and a call to a python routine to re-arrange the data, write the VOTables and plot the spectra retrieved from the query. The workflow can be modified to extract spectra of Vesta from all available EPN-TAP services, which may require a different python routine adapted to services that provide non-standard files. This modified workflow would be adapted to a study of Vesta's spectral variability from all existing data sources — a realistic use case.

### **c) Installations**

Use case (2) has been first tested on the local cluster in the Observatory of Paris, with restricted access (local account required). In this configuration, a federated Authentication and Authorisation Infrastructure (AAI) can be installed to grant access to external persons with ID provided by their institutes, or with their ORCID – this solves the problem of opening local accounts for external users, which is often forbidden by local IT policies.

Such an AAI has been studied earlier in this programme with the help of GÉANT / EduTeams and a VESPA / Europlanet community has been declared in EduTeams to identify authorized persons. It is routinely used to access the VOParis GitLab where external data services from all VESPA contributors are stored (<https://voparis-wiki.obspm.fr/display/VES/VESPA-Cloud+2021+Activity+Report>).

Second, the workflow was installed on EOSC servers, again based on this earlier study to deploy OpenStack Virtual Machines either locally or on EOSC. A virtual cluster is defined during this procedure, which can be retired after use. The procedure and details are provided here: <https://gitlab.obspm.fr/aicardi/tutorial-opus-eosc>

An instance is available here at the time of writing (by principle, this can be retired anytime by EOSC managers if not used): [http://78.128.251.190/opus\\_client/](http://78.128.251.190/opus_client/)  
This is a basic configuration on a minimal cluster of two nodes, which is not intended as a production service – that would require a real domain name and specific configuration, significantly increasing the housekeeping work and not central to this study.  
The EduTeams AAI is not connected here, although OPUS can handle it, as it makes little sense with a non-permanent server with no DNS name.

### 3. Results & Prospects

Altogether, the outcome of this study is positive: we can deploy the OPUS workflow platform on a local cloud as well as on EOSC servers and run it with no particular issue. A community AAI can be used to manage the user community's authentication and authorisation in simple cases. This provides a strong basis for future applications in the programme.

Possible improvements include:

- those concerning the OPUS interface, which were identified during this study. Some small improvements to the OPUS platform were reported to the main developer and fixed in the course of this study. Other were transmitted for a later update (e.g., add a function to send images and products to VO tools through SAMP from the OPUS page).
- more specifically, those related to cluster configuration and management on EOSC, e.g., shutting off unused nodes to save power, or dynamically adding new nodes to face peak activity periods. Such improvements require familiarity with EOSC interfaces and will be studied in a next step.
- an Authentication and Authorisation Infrastructure developed earlier in the programme has been installed and tested to grant access to external users through the VESPA / Europlanet community. Although usage with short-lived Virtual Machines does not appear relevant, the infrastructure is working as expected. More advanced use cases will be studied later in the programme, based on actual user requirements.