



HAL
open science

FAIR solutions for a science platform to analyse Cherenkov data online

Mathieu Servillat, Paula Kornecki, Catherine Boisson

► **To cite this version:**

Mathieu Servillat, Paula Kornecki, Catherine Boisson. FAIR solutions for a science platform to analyse Cherenkov data online. ADASS XXXII, Oct 2022, Victoria, Canada. obspm-03929759

HAL Id: obspm-03929759

<https://hal-obspm.ccsd.cnrs.fr/obspm-03929759>

Submitted on 9 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FAIR solutions for a science platform to analyse Cherenkov data online

Mathieu Servillat,¹ Paula Kornecki,¹ Catherine Boisson¹

¹*LUTH, Observatoire de Paris, Université PSL, Université de Paris, CNRS, F-92190 Meudon, France; mathieu.servillat@obspm.fr*

Abstract.

We developed a system to run quick analyses of Cherenkov data in compliance with the FAIR Guiding Principles for scientific data management (FAIR: Findable, Accessible, Interoperable and Reusable), through the use of interoperability standards and technologies, particularly those provided by the International Virtual Observatory Alliance (IVOA) to build the Virtual Observatory (VO).

We therefore provide a controlled and stable environment on a computing cluster, in order to execute and re-execute well defined jobs. User-specific input parameters can be specified to configure the execution of an analysis job. Provenance information is automatically captured by the system and accessible to the user. To avoid long transfers, the data can be placed close to the computing nodes. This system is primarily used to analyse Cherenkov astronomy data, though it may be used for other purposes.

1. Introduction and context

In the context of Open Science, data providers include more and more in their requirements the distribution of FAIR data (FAIR: Findable, Accessible, Interoperable and Reusable), as described by the FAIR Guiding Principles for scientific data management (Wilkinson et al. 2016). However, transforming the FAIR Principles into working implementations of a data access system is not always straightforward.

We propose an implementation where data access, job definitions, and computing resources are separated modules, and use standards to implement the interfaces. Data access is based on the IVOA Table Access Protocole (TAP, Dowler et al. 2019), job definitions are compatible with the IVOA Provenance Data Model (Servillat et al. 2020), and the execution of jobs on computing resources follow the IVOA Universal Worker Service pattern (UWS, Harrison & Rixon 2016). Authentication is possible through the OpenID Connect standard, and user management follows the SCIM standard (System for Cross-domain Identity Management).

2. From principles to implementation

2.1. Findable

- Data should be described with *standard metadata*: we provide metadata based on the IVOA Observation Core Metadata model (Louys et al. 2017) adapted to Cherenkov event data, and suitable for high energy astronomy.

- Metadata is served with a *standard protocol*: we use an IVOA TAP server based on GAVO DaCHS (Demleitner et al. 2014) that answers to ADQL queries.
- This metadata service is *publicly declared* in the VO Registry (Dower et al. 2018).

2.2. Accessible

- Data can be retrieved directly (through link or service): we indicate in the metadata the file type and its *access URL*.
- Authentication may be necessary: we use *federated and token based authentication* (OpenID Connect¹) to connect to services like KeyCloak (at Observatoire de Paris), eduTeams or ESCAPE IAM.
- Data is adequately identified to be found by the computing resources: data is placed in a storage mounted to computing nodes, in order to avoid long file transfers.

2.3. Interoperable

- Interfaces of the data access system are *based on standards*: we use IVOA TAP and UWS as standard interfaces with OPUS (Servillat et al. 2022).
- Data is stored in a well defined and *standard data format*: a common data format is proposed and discussed for Very-high-energy Observatories (GADF² and now VODF³).
- Related software follow the *FAIR principles for Research Software* (Chue Hong et al. 2021): Cherenkov data analyses are performed with the open source software GammaPy (Nigro et al. 2019).

2.4. Reusable

- Data is presented as complete datasets that contain all relevant pieces of information for further analysis (including instrument response functions).
- Data is connected to relevant calibration files and other datasets. We propose to use IVOA DataLink (Dowler et al. 2015) to interconnect data.
- Full and *detailed provenance* is available for each dataset: Provenance is automatically captured and exposed following the IVOA Provenance Data Model (Servillat et al. 2020).

3. Test implementations

Based on technologies and choices listed in Section §2, we thus developed a data access prototype in line with the FAIR Principles described in Figure 1, built on the following test implementations:

¹<https://openid.net/connect>

²<https://gamma-astro-data-formats.readthedocs.io>

³<https://vodf.readthedocs.io>

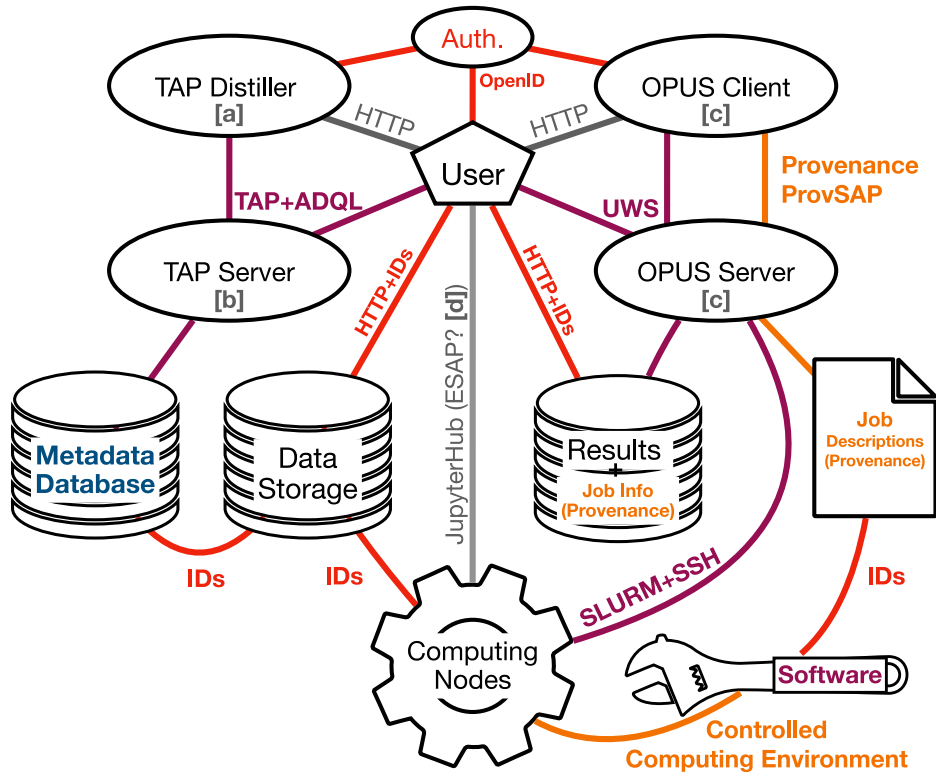


Figure 1. Functional diagram of a Science Platform based on the FAIR Principles. Test implementations are further described in §3. Colors indicate a compliance with the FAIR principles: Findable (blue), Accessible (red), Interoperable (purple) and Reusable (orange).

- [a] **TAP Distiller**⁴: a web based interface to query a TAP server.
- [b] **TAP Server**⁵: direct access to the TAP server implemented with DaCHS.
- [c] **OPUS**⁶ client/server: UWS interface to run jobs on computing nodes using OPUS (Servillat et al. 2022)
- [d] **ESAP**⁷: ESFRI Science Analysis Platform developed as a toolkit during the ESCAPE European project.

Figure 1 exposes the central position of the user that can authenticate to several web based modules to perform data searches (TAP Distiller) or data analyses online (OPUS CLient). The user can also directly access the TAP server or the OPUS server

⁴<https://voparis-cta-test.obspm.fr>

⁵http://voparis-tap-he.obspm.fr/browse/hess_dr/q

⁶<https://voparis-uws-test.obspm.fr>

⁷<https://git.astron.nl/astron-sdc/esap-api-gateway>

that implement IVOA standards (TAP+ADQL and UWS). The Data is described by standard metadata distributed by the TAP server. The same data can be accessed by the computing nodes. The OPUS server describes a set of jobs deployed in controlled environments and run on the computing nodes. We currently explore a parallel access to computing nodes and to the data based on a JupyterHub cluster (or a BinderHub cluster), possibly deployed via the ESAP interface.

4. Working use case

The FAIR solutions for a science platform were further tested to run a quick gammapy analysis of Cherenkov data online, as presented in Kornecki et al. (2022).

References

- Chue Hong, N. P., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., Honeyman, T., & et al. 2021, FAIR Principles for Research Software (FAIR4RS Principles), Research Data Alliance. URL <https://doi.org/10.15497/RDA00065>
- Demleitner, M., Neves, M. C., Rothmaier, F., & Wambsganss, J. 2014, *Astronomy and Computing*, 7, 27. 1408. 5733
- Dower, T., Demleitner, M., Benson, K., Plante, R., Auden, E., Graham, M., Greene, G., Hill, M., Linde, T., Morris, D., O'Mullane, W., Rixon, G., Stébé, A., & Andrews, K. 2018, Registry Interfaces Version 1.1, IVOA Recommendation 23 July 2018
- Dowler, P., Bonnarel, F., Michel, L., & Demleitner, M. 2015, IVOA DataLink Version 1.0, IVOA Recommendation 17 June 2015
- Dowler, P., Rixon, G., Tody, D., & Demleitner, M. 2019, Table Access Protocol Version 1.1, IVOA Recommendation 27 September 2019
- Harrison, P. A., & Rixon, G. 2016, Universal Worker Service Pattern Version 1.1, IVOA Recommendation 24 October 2016
- Kornecki, P., Servillat, M., Boisson, C., & Fuessling, M. 2022, in Proceedings of Gamma 2022 Symposium, PoS(Gamma2022), vol. 417, TBD. <https://pos.sissa.it/417/>
- Louys, M., Tody, D., Dowler, P., Durand, D., Michel, L., Bonnarel, F., Micol, A., & IVOA DataModel Working Group 2017, Observation Data Model Core Components, its Implementation in the Table Access Protocol Version 1.1, IVOA Recommendation 09 May 2017
- Nigro, C., Deil, C., Zanin, R., Hassan, T., King, J., Ruiz, J. E., Saha, L., Terrier, R., Brügge, K., Nöthe, M., Bird, R., Lin, T. T. Y., Aleksić, J., Boisson, C., Contreras, J. L., Donath, A., Jouvin, L., Kelley-Hoskins, N., Khelifi, B., Kosack, K., Rico, J., & Sinha, A. 2019, *A&A*, 625, A10. [arXiv:1903.06621](https://arxiv.org/abs/1903.06621)
- Servillat, M., Aicardi, S., Cecconi, B., & Mancini, M. 2022, in *Astronomical Society of the Pacific Conference Series*, edited by J. E. Ruiz, F. Pierfederici, & P. Teuben, vol. 532 of *Astronomical Society of the Pacific Conference Series*, 451. [arXiv:2101.08683](https://arxiv.org/abs/2101.08683)
- Servillat, M., Riebe, K., Boisson, C., Bonnarel, F., Galkin, A., Louys, M., Nullmeier, M., Renault-Tinacci, N., Sanguillon, M., & Streicher, O. 2020, IVOA Provenance Data Model Version 1.0, IVOA Recommendation 11 April 2020
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. 2016, *Scientific Data*, 3