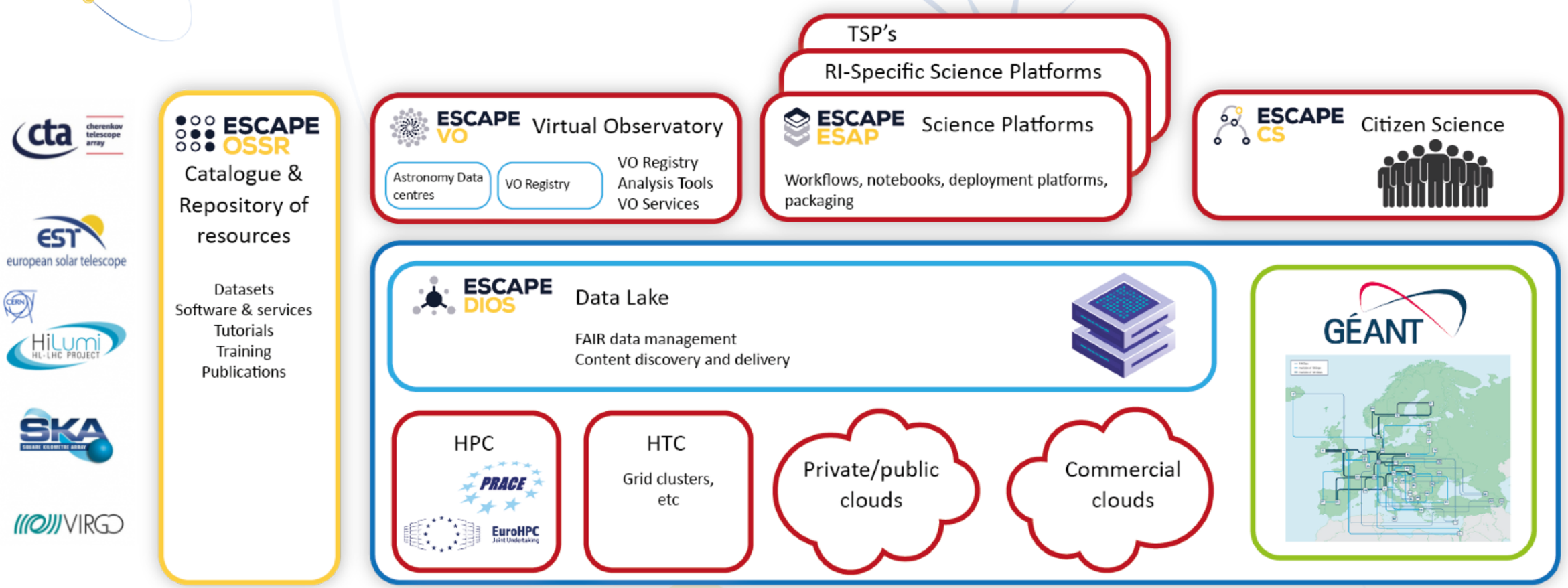
ESAP
The ESCAPE ESFRI Science Analysis Platform

John D. Swinbank — swinbank@astron.nl

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

ESCAPE

www.projectescape.eu

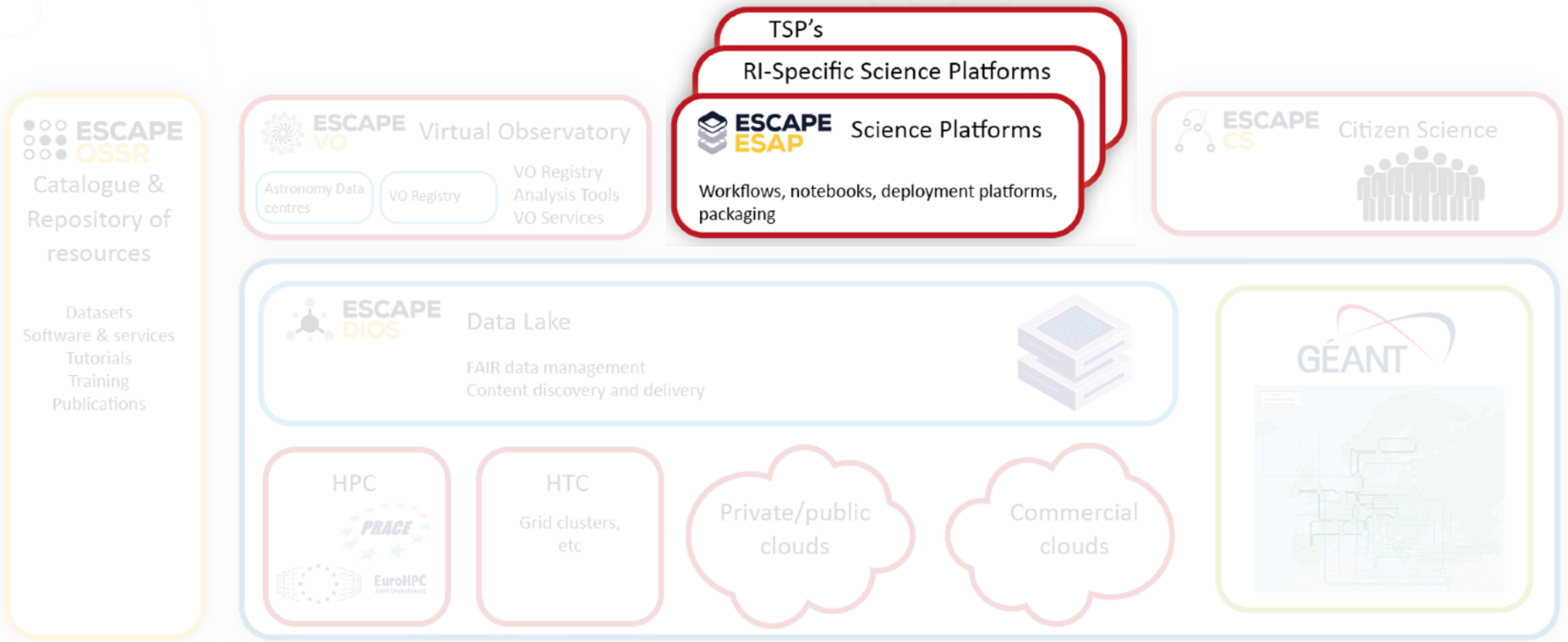


European Science Cluster of Astronomy & Particle Physics
ESFRI Research Infrastructures

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE

www.projectescape.eu

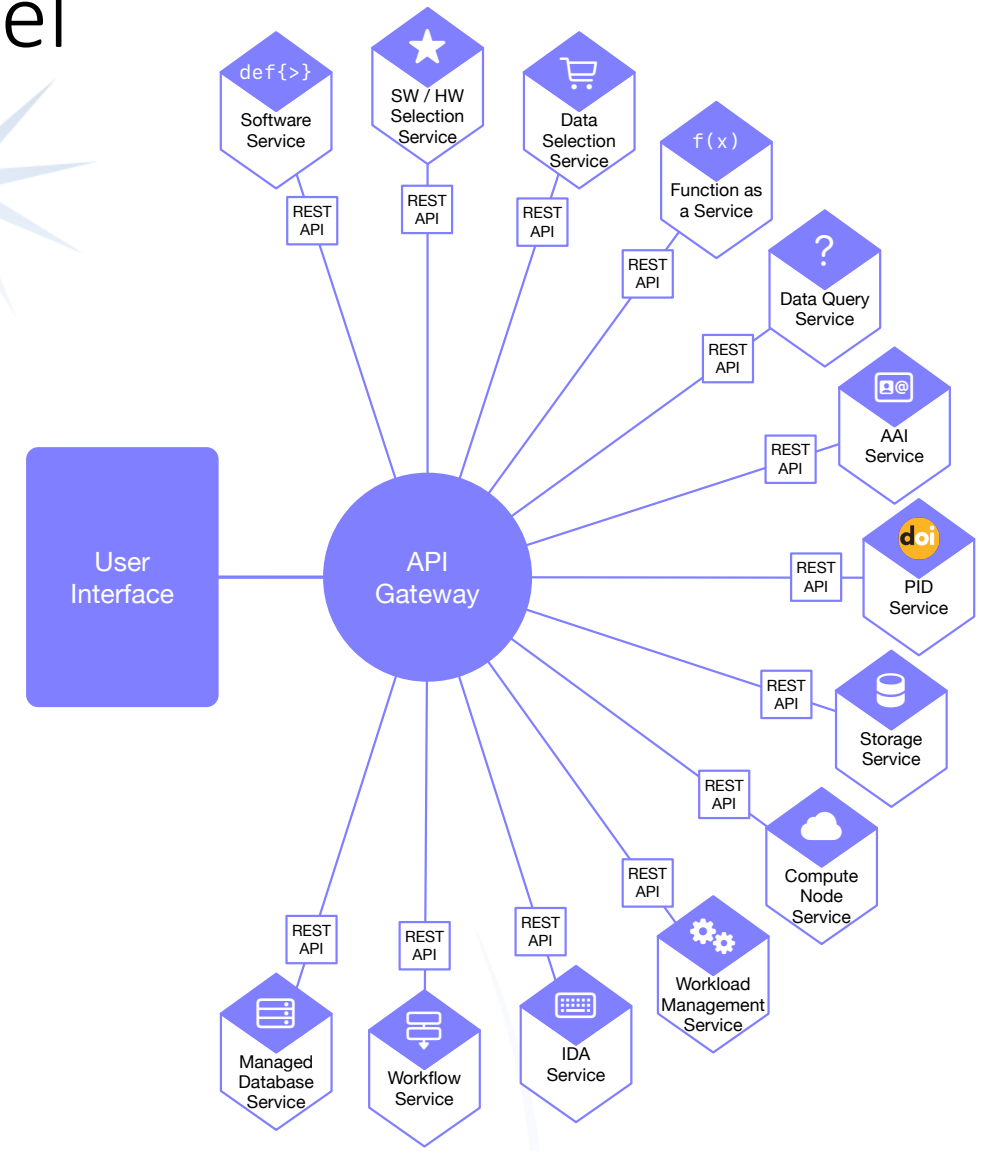


European Science Cluster of Astronomy & Particle Physics
ESFRI Research Infrastructures

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESAP: The hub in ESCAPE's wheel

- A *focal point* for integrating diverse services which are drawn from other providers.
- Two part structure: Web UI, API Gateway.
- Focal point of a range of pluggable, independent services (some of which are ESCAPE deliverables).
- Designed to be *robust* & *extensible*.

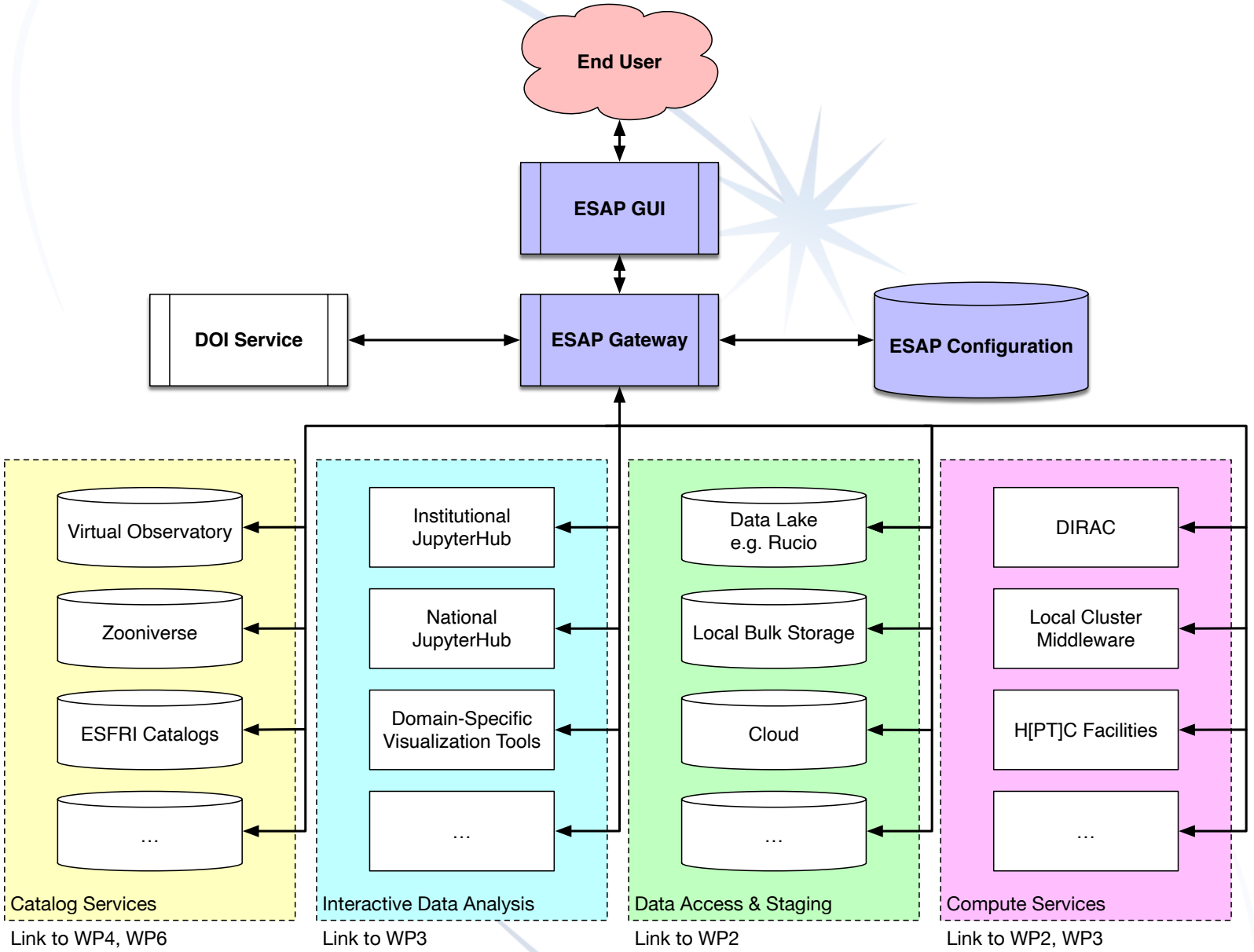


Funded by the European Union's Horizon 2020 - Grant N° 824064

A platform... or a toolkit?

- Science Platforms: all things to all people.
- We are *not resourced* to build and maintain Jupyter, batch computing services, or similar for common/EOSC access.
  - Many ESCAPE partner institutions do make available systems for testing, developing, and experimenting on.
- **ESAP is a *toolkit* for building "science platforms" which are customized to particular applications.**
- At a variety of scales:
  - "Centralized ESAP", providing flexible and convenient access to a wide spectrum of ESCAPE services.
  - "ESFRI ESAP", providing a way for individual infrastructures, projects, etc to quickly integrate diverse capabilities into a unified service offering.

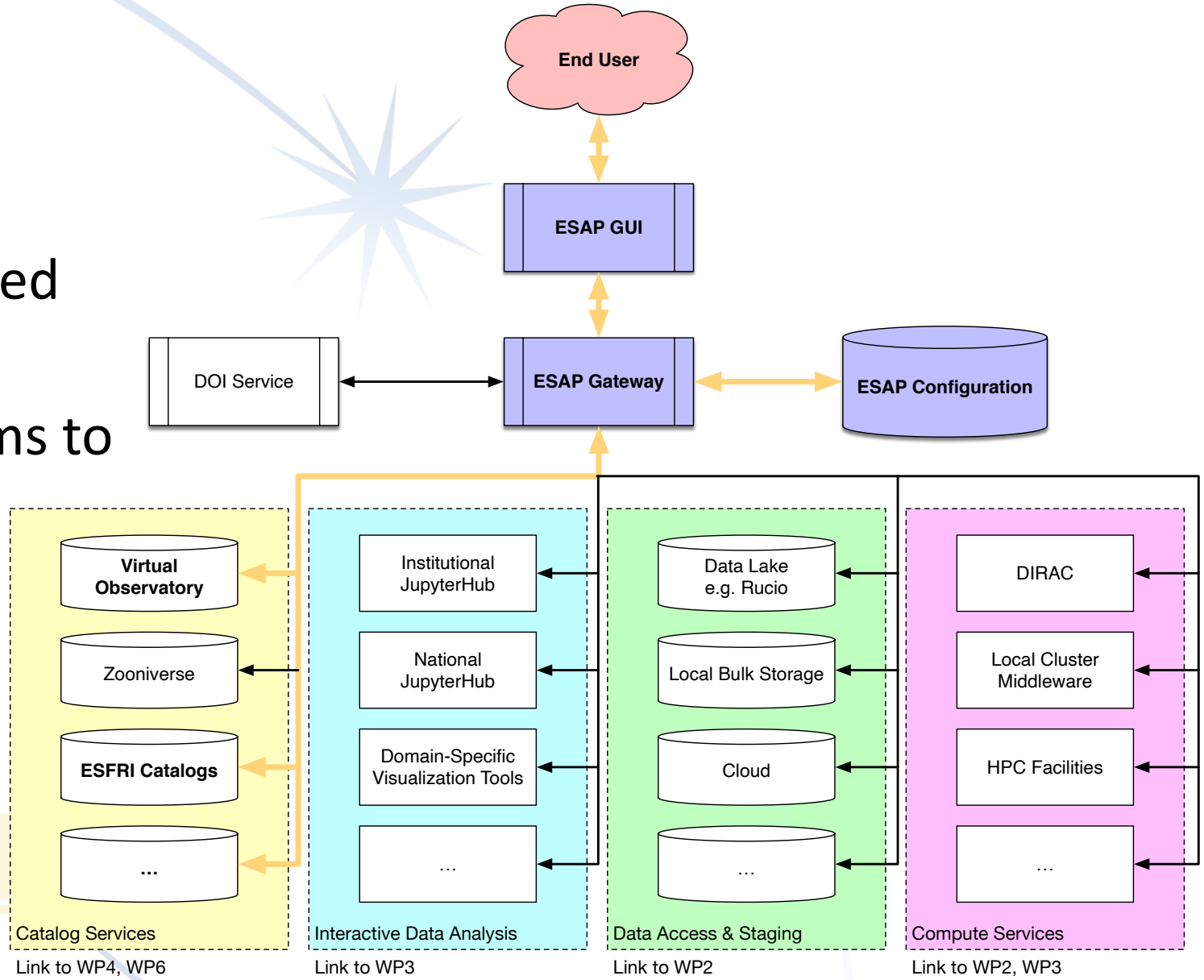
Funded by the European Union's Horizon 2020 - Grant N° 824064



Funded by the European Union's Horizon 2020 - Grant N° 824064

Example Workflow
1. Query

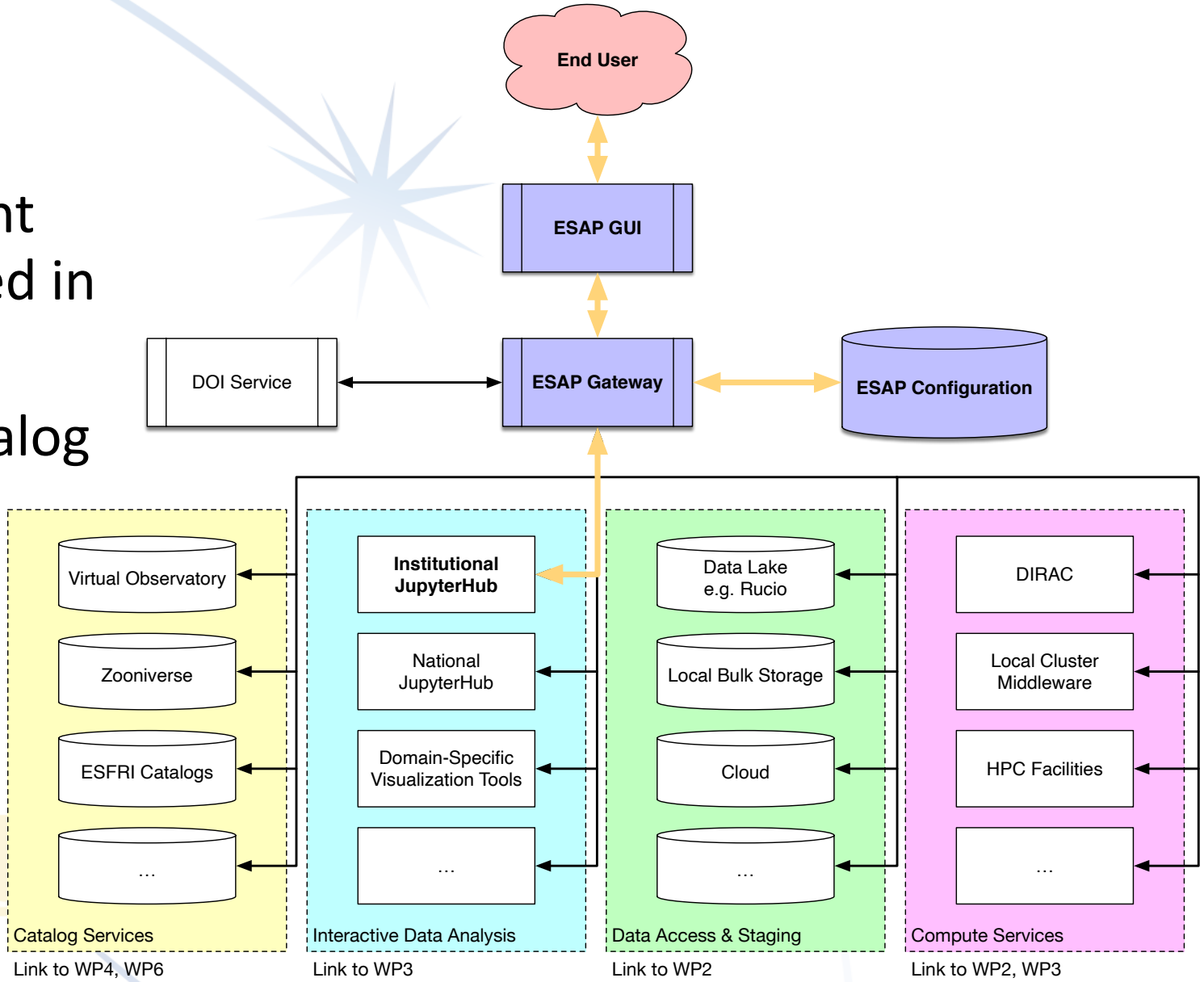
- User identifies relevant catalog services configured in this instance of ESAP.
- User submits search terms to multiple catalogs using consistent ESAP UI.
- Search results returned to user and displayed in unified form.



Funded by the European Union's Horizon 2020 - Grant N° 824064

Example Workflow
2. Winnow

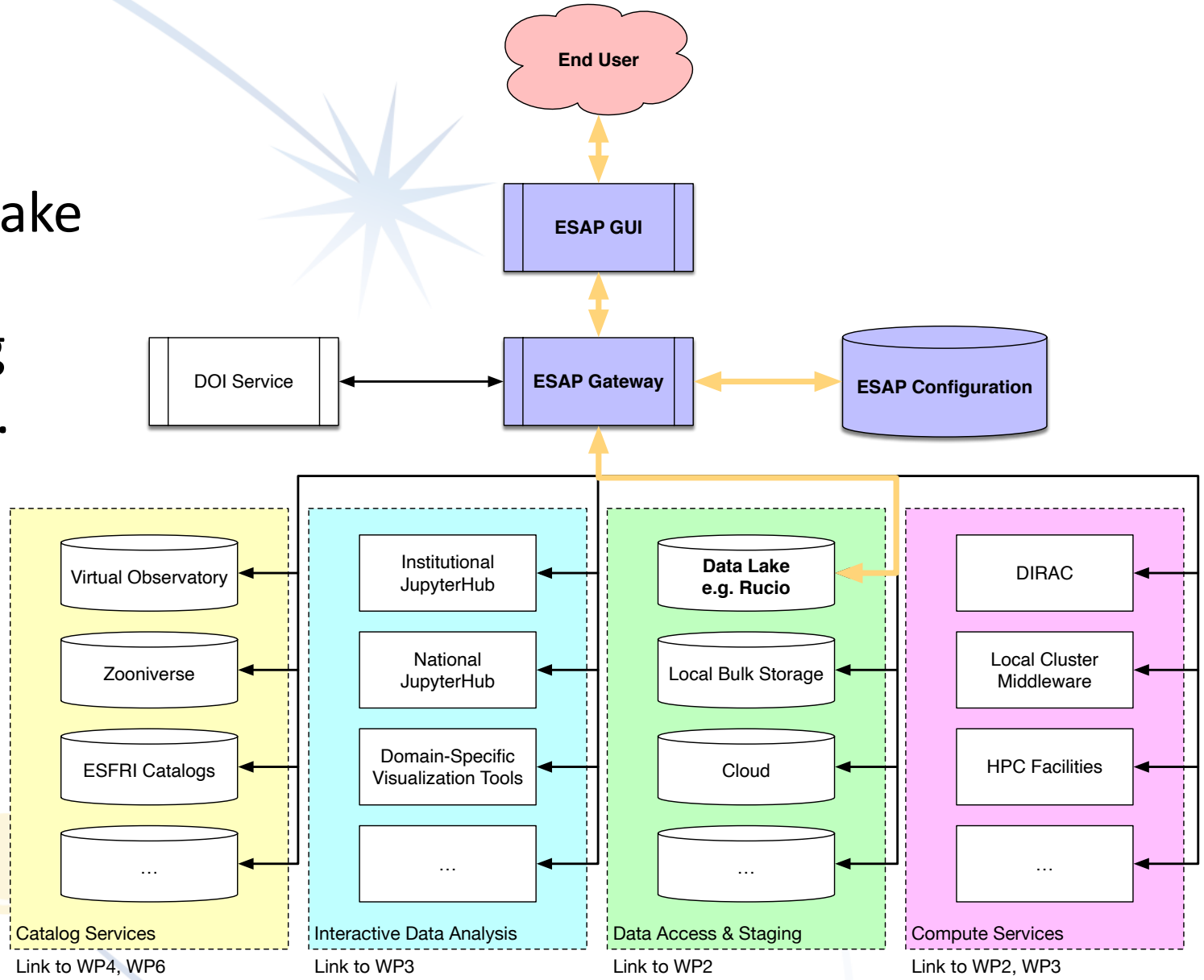
- User identifies convenient Jupyter system configured in this instance of ESAP.
- User sends retrieved catalog data to notebook.
- Interactive analysis session in notebook identifies most promising candidates for bulk processing.



Funded by the European Union's Horizon 2020 - Grant N° 824064

Example Workflow
3. Stage Data

- User instructs the Data Lake to “stage” archived data corresponding to catalog entries to online storage.

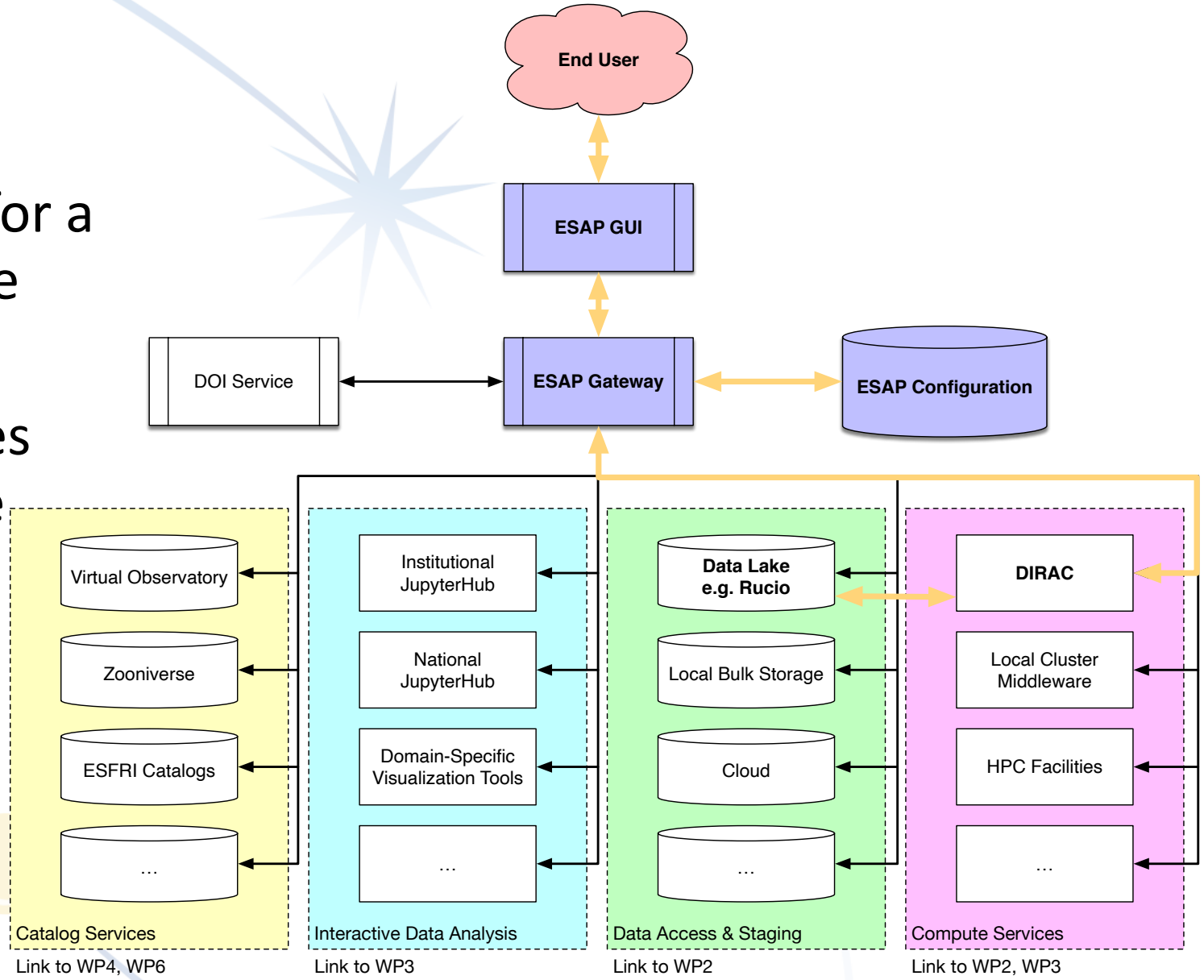


Funded by the European Union’s Horizon 2020 - Grant N° 824064

Example Workflow

4. Compute

- User sends instructions for a batch compute job to the DIRAC cluster.
- Compute cluster retrieves the staged data from the Data Lake.
- Batch processing happens.
- Results are stored to Data Lake, and the user notified of completion.

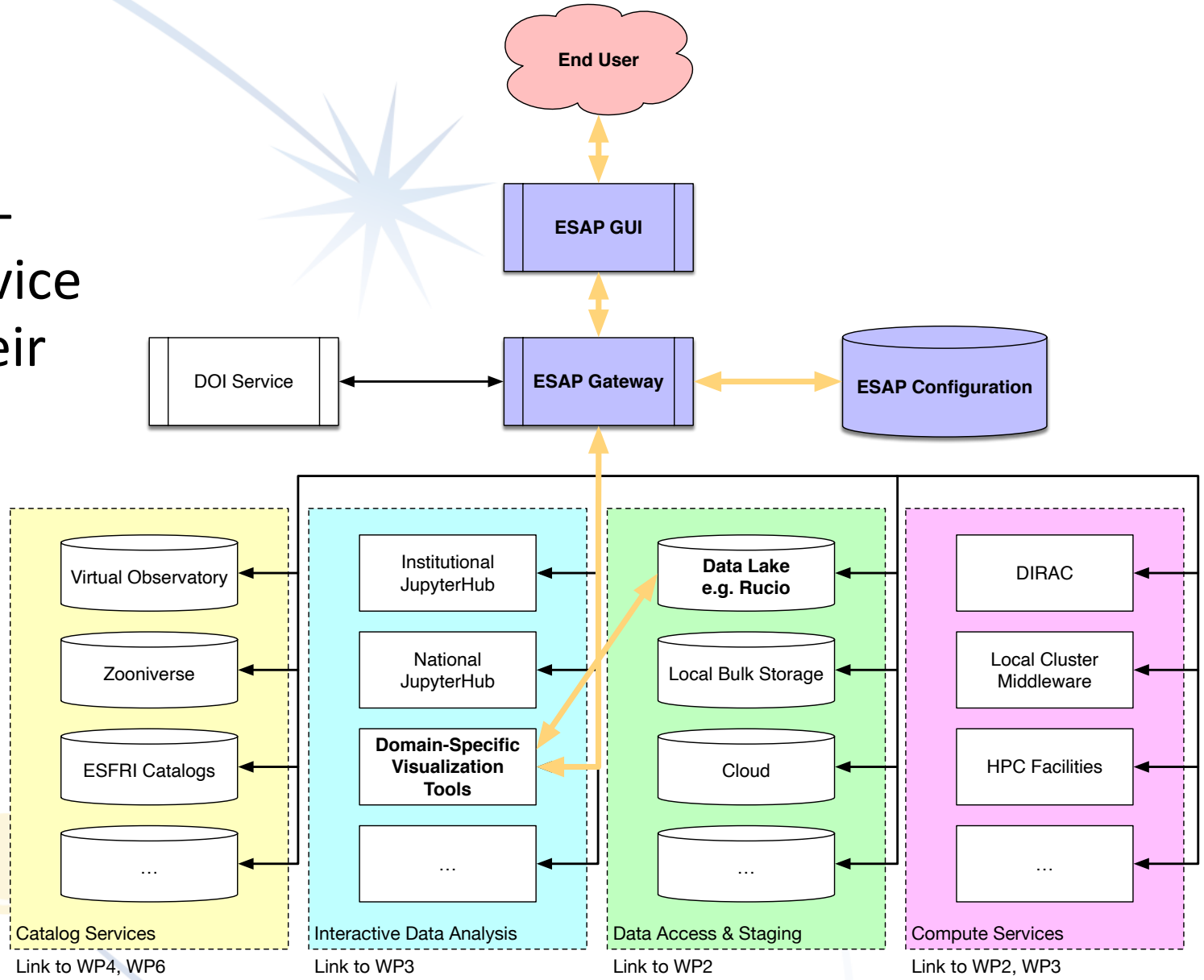


Funded by the European Union's Horizon 2020 - Grant N° 824064

Example Workflow

5. Visualize Results

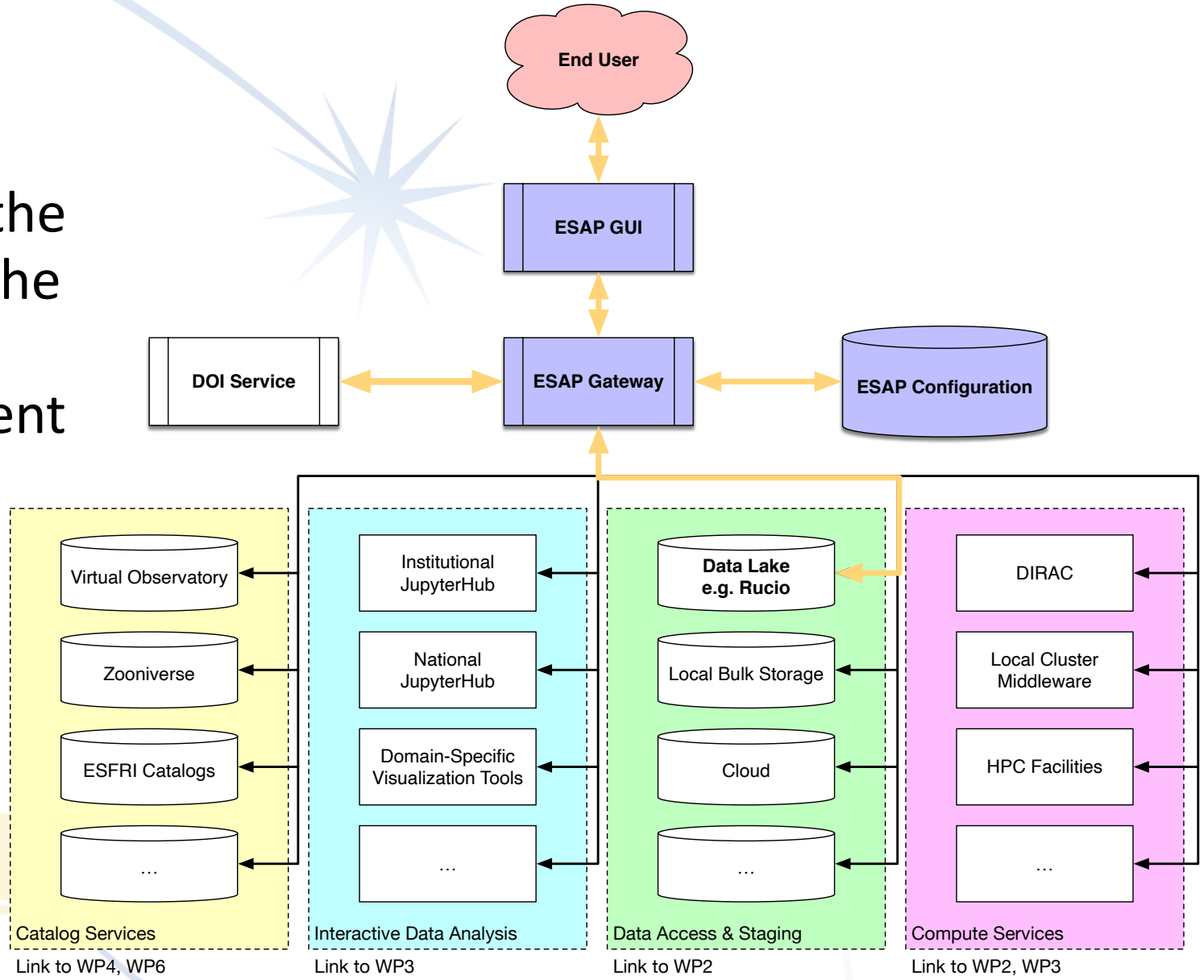
- User identifies a domain-specific visualization service that can help analyze their data.
- User initiates a visualization session, passing location of compute results.



Funded by the European Union's Horizon 2020 - Grant N° 824064

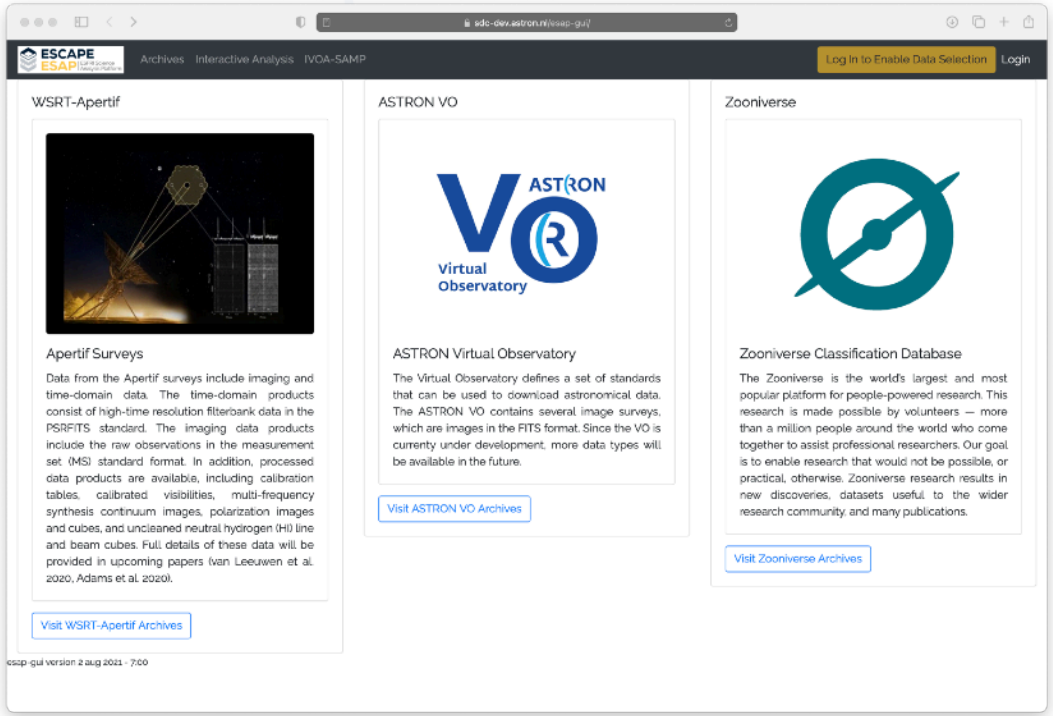
Example Workflow
6. Publish

- Having established that the results are noteworthy, the user instructs the DOI Service to mint a persistent identifier for them.
- The results are made available to the wider community.

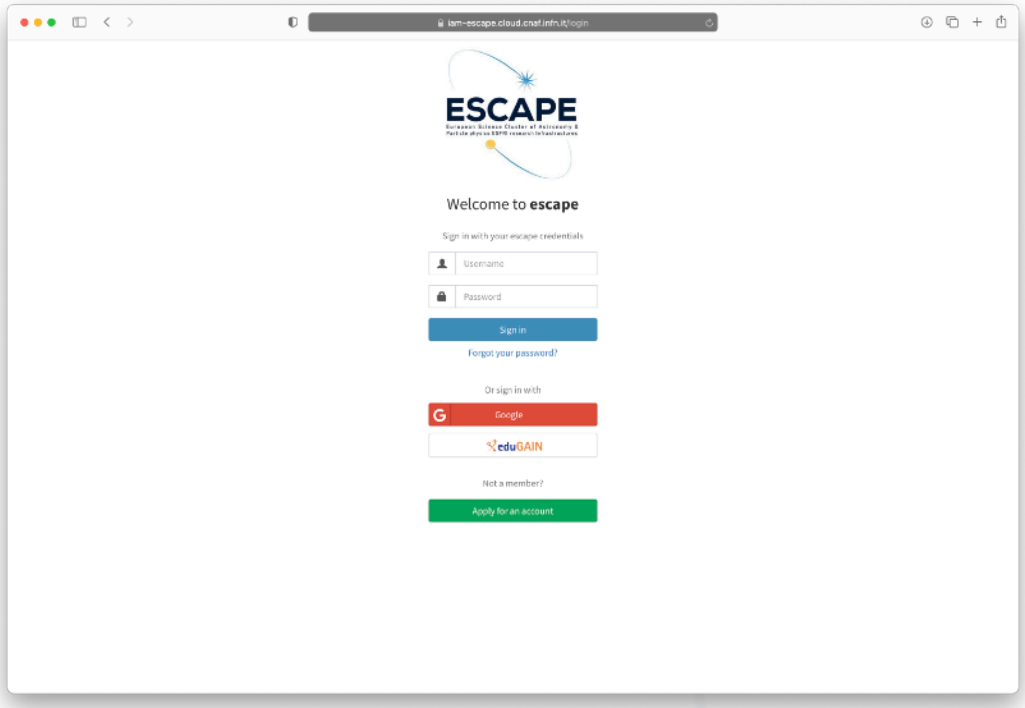


Funded by the European Union's Horizon 2020 - Grant N° 824064

Current Status: Prototype Deployment



- Refined & attractive user interface.
- Integrated with ESCAPE IAM.
- Containerized deployment; moving towards production environment.

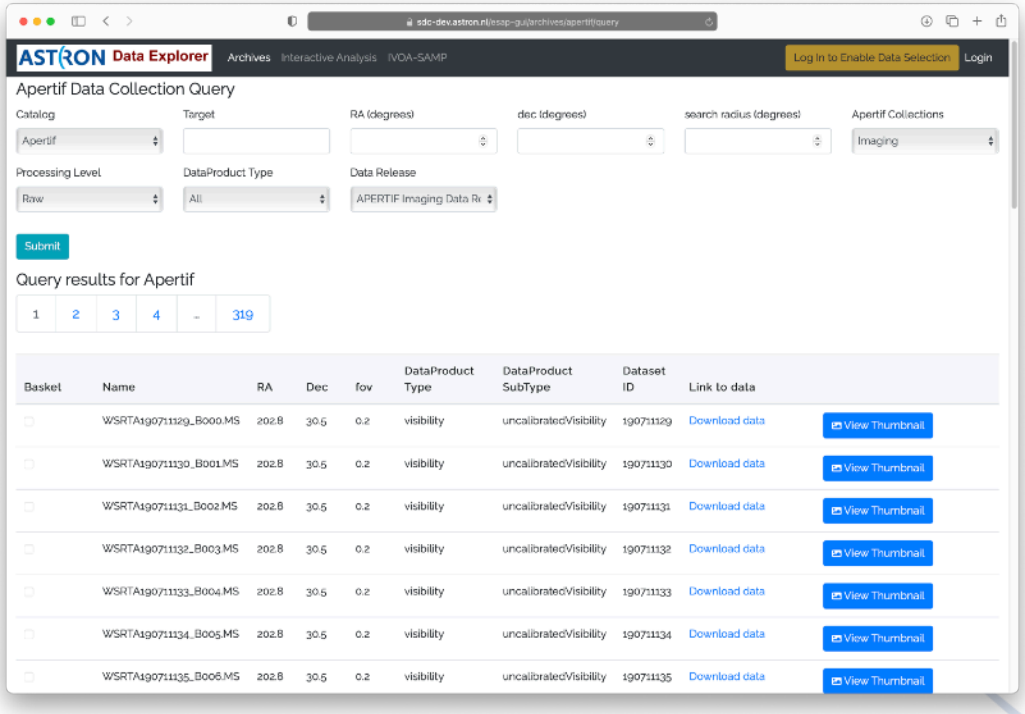


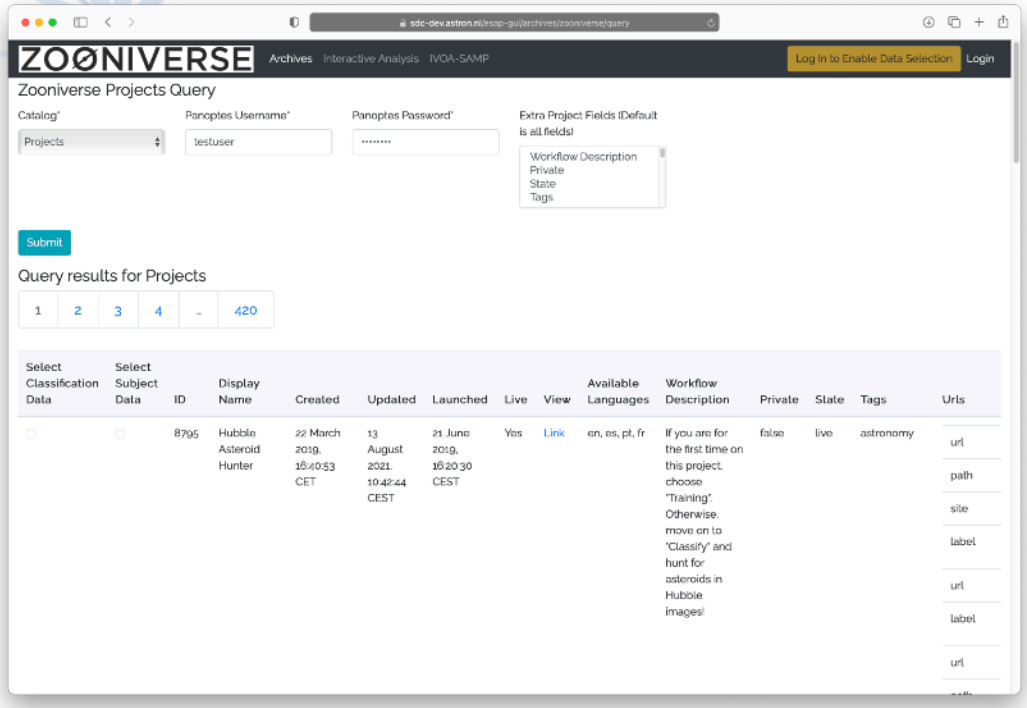
- https://sdc-dev.astron.nl/esap-gui
- *Unsupported prototype*: try at your own risk! 💥

Funded by the European Union's Horizon 2020 - Grant N° 824064

Current Status: Data Discovery

- Data discovery through multiple archive types
- VO (WP4) and non-VO enabled catalogues (Zooniverse, Apertif)

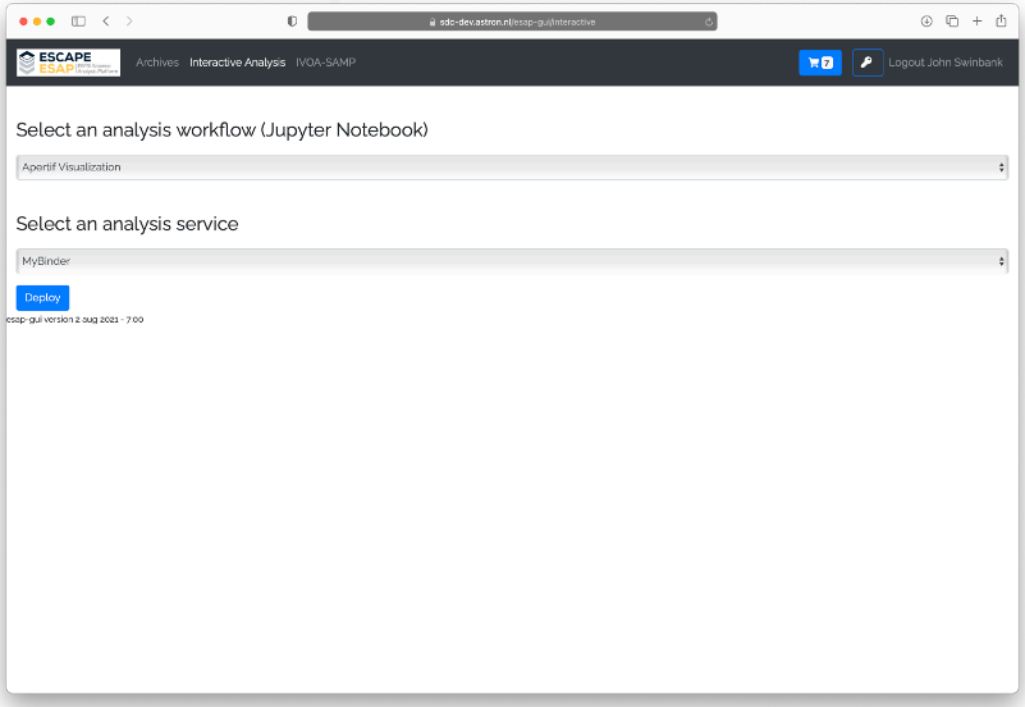




- Query the Rucio system which underlies the WP2 Data Lake
- Note interface is customized to reflect the particular archive and type of data being queried

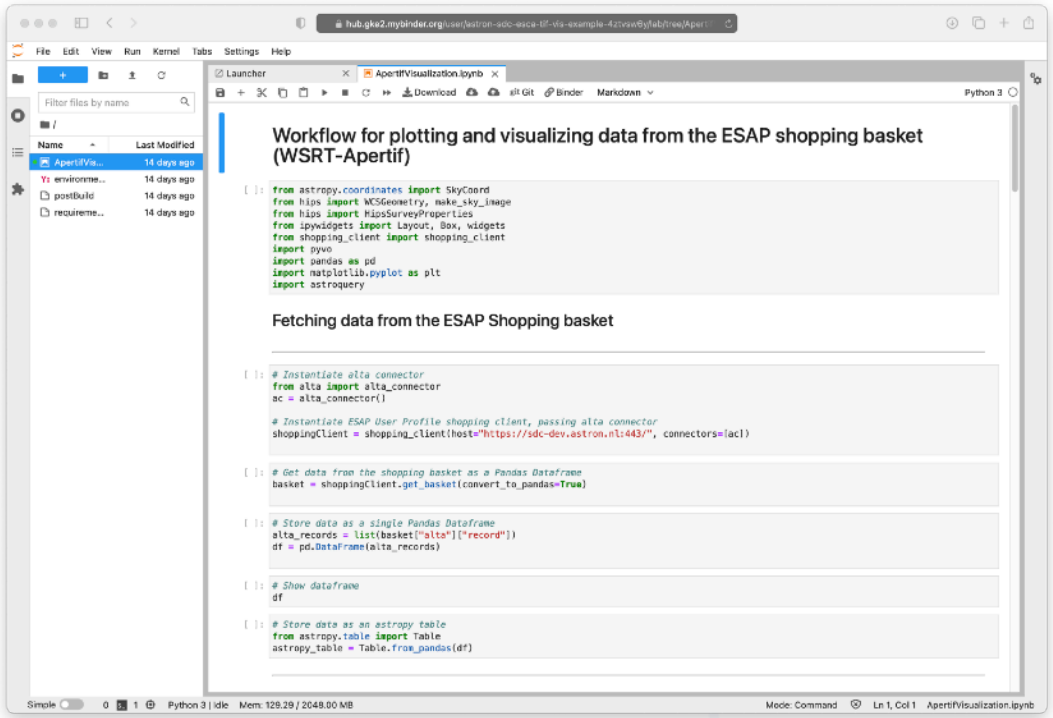
Funded by the European Union's Horizon 2020 - Grant N° 824064

Current Status: Interactive Data Analysis



- Select notebooks and analysis services from drop-downs in the ESAP system
- Software selection from the ESCAPE Open Source Scientific Software Repository (OSSR)
- BinderHub-based workflows most mature; alternative surveys/ technologies coming

- Python library integrates notebooks with ESAP shopping basket.
- Provides notebook user with full access to data discovered in ESAP.



Funded by the European Union's Horizon 2020 - Grant N° 824064

Current Status: Other Features

- Reinvigorated batch compute effort started recently.
  - Focusing on integration with the DIRAC system.
  - Aim to build generic interfaces which are usable for other systems.
- Managed Database service provides user-owned databases directly within the ESAP system.
  - Databases can integrate directly with archives discoverable through ESAP.
  - Enables complex query types, e.g. cross matching.
  - Currently an early prototype.
- Support for uploading data using IVOA Simple Application Messaging Protocol (SAMP).

Funded by the European Union's Horizon 2020 - Grant N° 824064

Future Plans

- Closer interaction with ESCAPE *Data Infrastructure for Open Science* (DIOS; data lake), building on “Data Lake as a Service” technology developed in ESCAPE WP2.
- Tighter links with the ESCAPE *Open Source Scientific Software Repository* (OSSR), helping the user find and select workflows which are most relevant to their interests and the data they have selected.
- Advanced matching of users, data, and workflows, taking account of data locality, resource needs, etc.
- Upgraded ESAP internal architecture, including support for async queries.

Funded by the European Union’s Horizon 2020 - Grant N° 824064

Downloads & Further Information

- ESAP is developed in Python, Django, and React.
- It is available under the Apache license, version 2.0.
- Get it here:
  - https://git.astron.nl/astron-sdc/esap-api-gateway
  - https://git.astron.nl/astron-sdc/esap-gui
- Try it here (no warranties; be tolerant):
  - https://sdc-dev.astron.nl/esap-gui
- Your feedback is always welcome!

Funded by the European Union's Horizon 2020 - Grant N° 824064