WP5 - ESAP
ESFRI Science Analysis Platform

Zheng Meyer-Zhao, Michiel van Haarlem

ASTRON, The Netherlands

ESCAPE Progress Meeting, 26-27 February 2020

ESCAPE

Connect science platform with existing astronomical data archives and VO-enabled data collections

Access to software & services in ESCAPE-EOSC catalogue

EOSC and ESFRI Science Communities

ESCAPE

WP5: ESFRI Science Analysis Platform

| Task 5.1 - Discovery and Staging | Task 5.2 - Software deployment and virtualization | Task 5.3 - Analysis interfaces, work flows, and reproducibility | Task 5.4 - Integration with HPC infrastructures |

Portals | Data Staging | Workflows | User Tools

ESCAPE

WP4: CEVO

Data Discovery

VO Services and Tools

ESFRI Data

User Work Spaces

User Compute Resources

ESFRI Software

EOSC Services

ESCAPE

WP3: OSSR

Data Lake Storage and Computing Infrastructure

ESCAPE WP2: DIOS
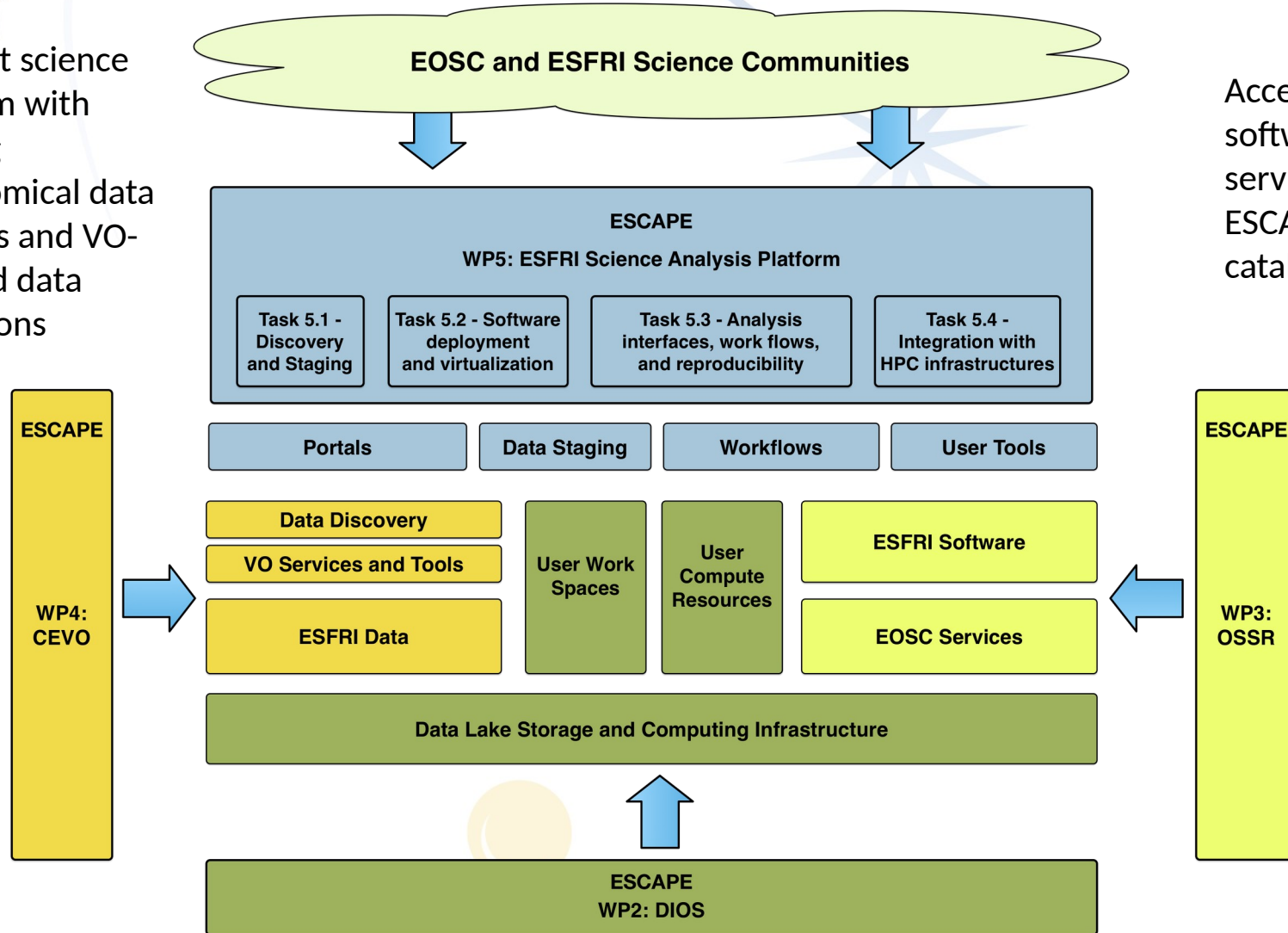
Integration with Data Lake - distributed computing and storage

T5.1 - Data aggregation and staging

- Stage data in the Data Lake (WP2)
- Data discovery, VO (WP4) - to be expanded
- Dynamically allocate user workspace across distributed infrastructure
- Tools to estimate availability & latency
- Demonstrate for a range of data collections (CTA, ESO, EST, FAIR, JIVE, LOFAR,...)

T5.2 - Software deployment and virtualisation

- Integrate software and service repository (WP3), allow access to software components developed by ESFRIs
- Provide access to software repository metadata
- Support containerisation of additional tools
- Demonstrate with variety of examples (ESO, FAIR, JIVE, LOFAR)

T5.3 – Analysis interface, work flows and reproducibility

- Interactive analysis interface which Integrates data access & staging (T5.1)
- Provides access to EOSC software repository (T5.2)
- Simplify porting workflows to science platform environment
  - support common deployment language (e.g. CWL)
  - deploy across EOSC infrastructure
  - promote preservation & sharing of workflows
- Start with small number of representative workflows
- Evaluate performance, monitor compliance w/ FAIR principles

T5.4 - Integration with HPC and HTC infrastructures

- Deploy user-initiated workflows on HPC and HTC infrastructure
- but… maintain interactivity and responsiveness
- Obviously close links with WP2 – integrate Science Platform with Data Lake
- Expand number of ESFRIs supported

Links with other ESCAPE WPs

- WP1 - General EOSC policy for services & infrastructure access

- WP2 - Integration with Data Lake - distributed computing and storage, AAI

- WP3 - Access to software & services in ESCAPE-EOSC catalogue

- WP4 - connect science platform with existing astronomical data archives and VO-enabled data collections

Links with other ESCAPE WPs

Deliverables

D5.1	Preliminary report on requirements for ESFRI science analysis use cases	WP5	ASTRON	R	PU	M6
D5.2	Detailed project plan for WP5	WP5	ASTRON	R	PU	M9
D5.3	Performance assessment of initial Science Platform prototype	WP5	ASTRON	R	PU	M24
D5.4	Final assessment of the performance of the Science Platform prototype and plan for deployment of production version within the EOSC.	WP5	ASTRON	R	PU	M42

Milestones

M5.1	First WP5 workshop on Science Platform design and requirements	WP5	M4	Workshop summary report
M5.2	Review of preliminary report on requirements for ESFRI science analysis use cases by WP5 task leader and ESFRI representatives	WP5	M5	Review report
M5.3	Initial science platform prototype with discovery and data staging	WP5	M18	Prototype science platform online and available to partners (online service)
M5.4	Deployment of initial set of ESFRI software on prototype platform	WP5	M20	Initial set of ESFRI software available for use in platform prototype (online service)
M5.5	Second WP5 workshop to analyse prototype performance	WP5	M22	Workshop summary report
M5.6	Integration of Science Platform with OSSR repository	WP5	M28	OSSR software repository available for use in platform prototype (online service)
M5.7	Integration of Science Platform with Data Lake expanded prototype	WP5	M30	External compute resources accessible through the platform prototype (online service)
M5.8	Delivery and integration of new ESFRI visualization and analysis tools	WP5	M36	New ESFRI tools available in OSSR repository and accessible through the science platform (online service)
M5.9	Final WP5 ESFRI user training workshop on the Science Platform	WP5	M38	Workshop summary report

1st year ESCAPE WP5 Activities

Organised Workshops and F2F meeting:

- 16- 17 April 2019, Groningen
  ESFRI Use Case Requirements workshop

- 01-03 July 2019, Amsterdam
  Joint ESCAPE WP2/WP5 workshop

- 03-04 July 2019, Amsterdam
  WP5 1st Deliverable discussion

- Monthly Video conference

- Bi-weekly Tech meeting

1st year ESCAPE WP5 Activities contd.

Participated in

- International Symposium on Grids & Clouds 2019 (ISGC 2019)

- The 29th annual international Astronomical Data Analysis Software & Systems (ADASS 2019)

- Workshop on Cloud Services for Synchronisation and Sharing (CS3 2020)

- CS3MESH4EOSC project kick-off meeting (2020)

- ESCAPE WP4 Tech Forum (2020)

- Other ESCAPE work packages meetings (including telecons)

1st year ESCAPE WP5 Achievements

Two deliverables:

- D5.1 Preliminary report on requirements for ESFRI science analysis use cases

- D5.2 - Detailed Project Plan

Two milestones:

- M5.1 First WP5 workshop on Science Platform design and requirements

- M5.2 Review of preliminary report on requirements for ESFRI science analysis use cases by WP5 task leader and ESFRI representatives

1ˢᵗ year ESCAPE WP5 Achievements contd.

- Identified ESAP service components

- Finished ESAP Architectural Design

- Established ESAP Minimum Viable Product (MVP) Core team

Looking Forward

- 31 March – 3 April, ASTRON (Dwingeloo, The Netherlands)
  WP5 Busy Week

- WP5 Hackathon, date & location TBD

- 26-30 October, EGO (Pisa, Italy)
 ESAP prototype performance workshop

ESAP UI Service Components
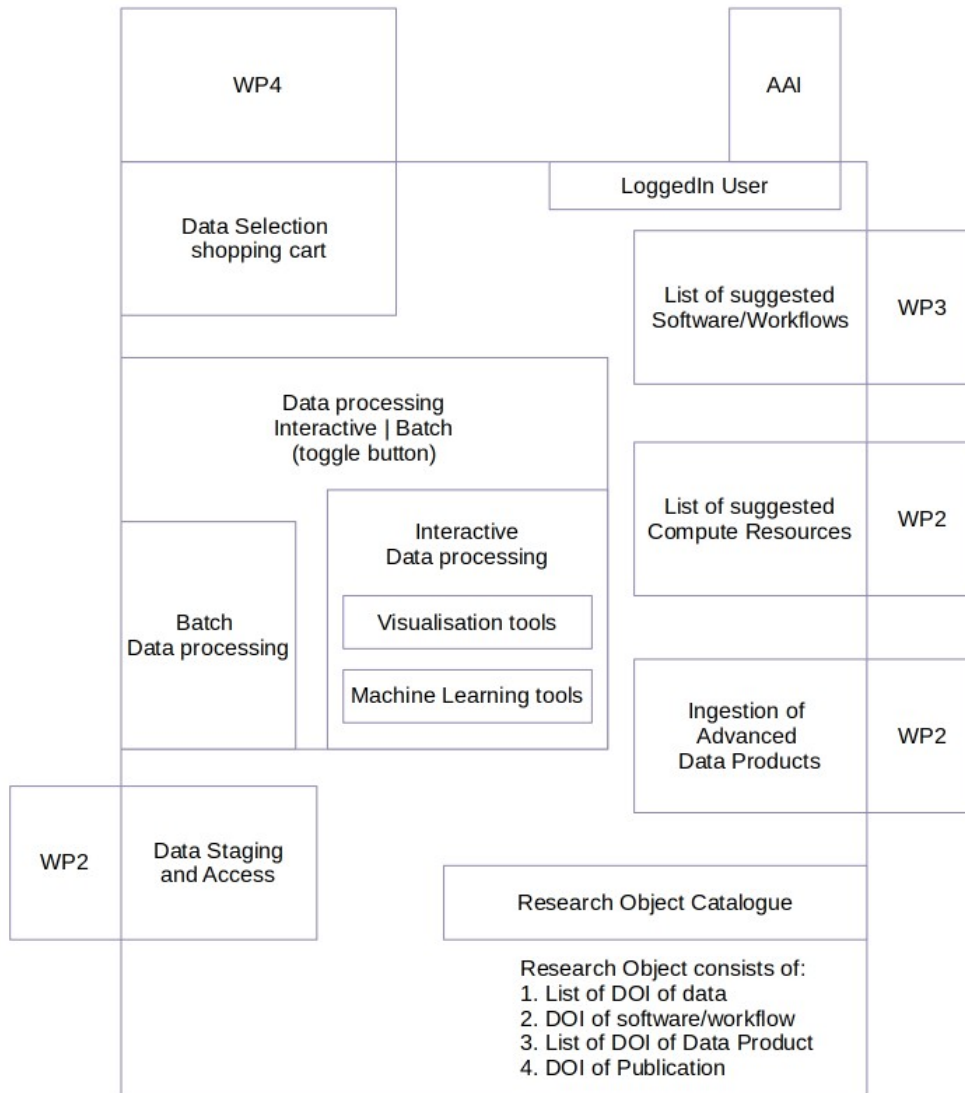
- AAI
- Data selection shopping cart
- Data Staging and Access
- List of suggested Sofware/Workflows
- List of suggested Compute Resources
- Batch data processing
- Interactive Data analysis
- Data analysis with visualisation tools
- Data analysis with machine learning tools
- Research object catalogue
- Ingestion of advanced data products

ESAP Architectural Design
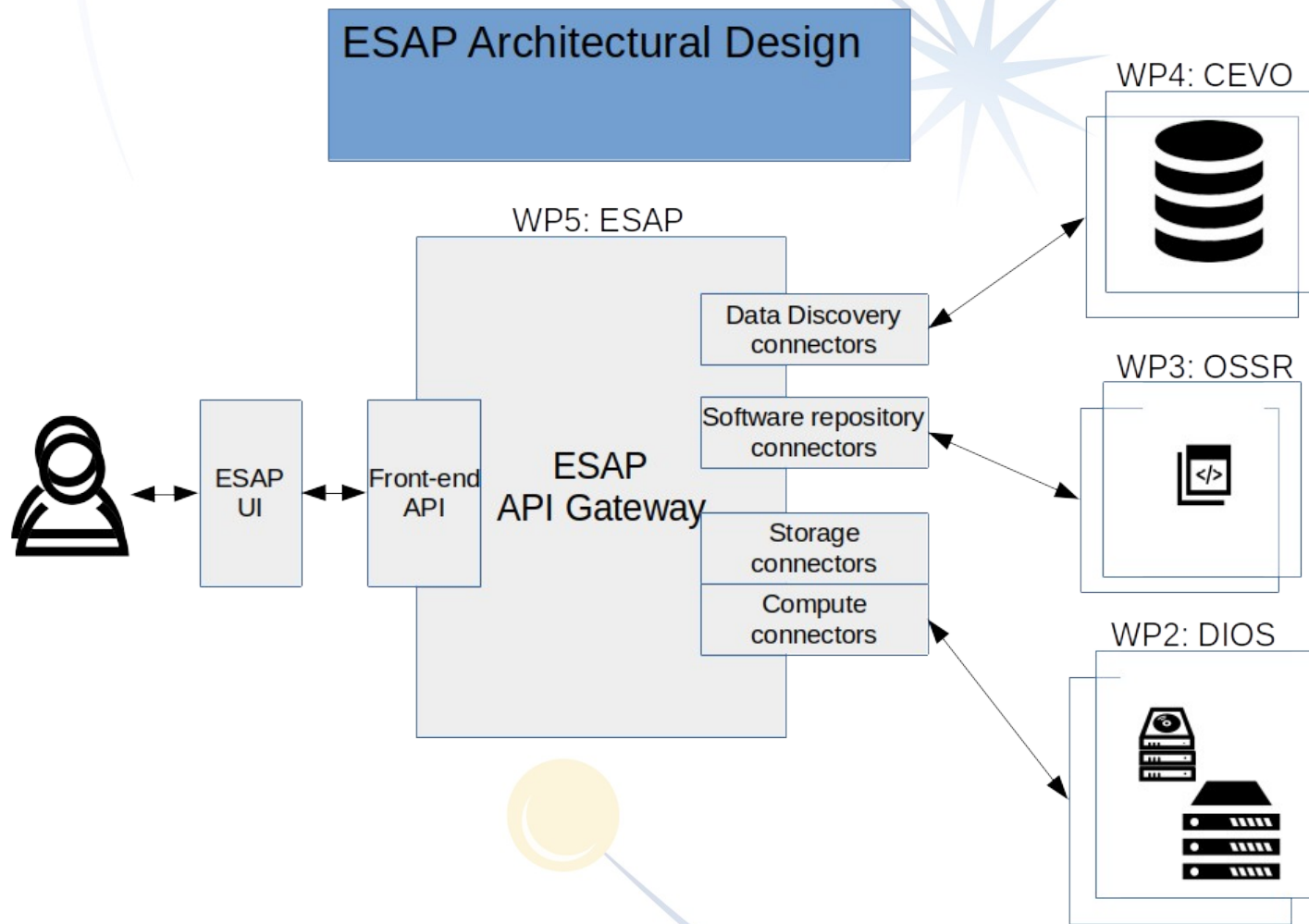
ESAP Minimum Viable Product Core Team

Each core team member is leading one or more service categories:

- ASTRON is leading the development of the ESAP UI and API Gateway.

- SKAO and CTAO lead the aspects of data lake and workload management integration.
  - CERN, FAU

- Nikhef leads on the federated-AAI aspects.

- UEDIN leads the IVOA integration and interactive data processing services.
  - JIVE, EGO, FAIR, UCM

- INAF is leading the develpment of batch processing.
  - CSIC-IAA

ESAP Architectural Design
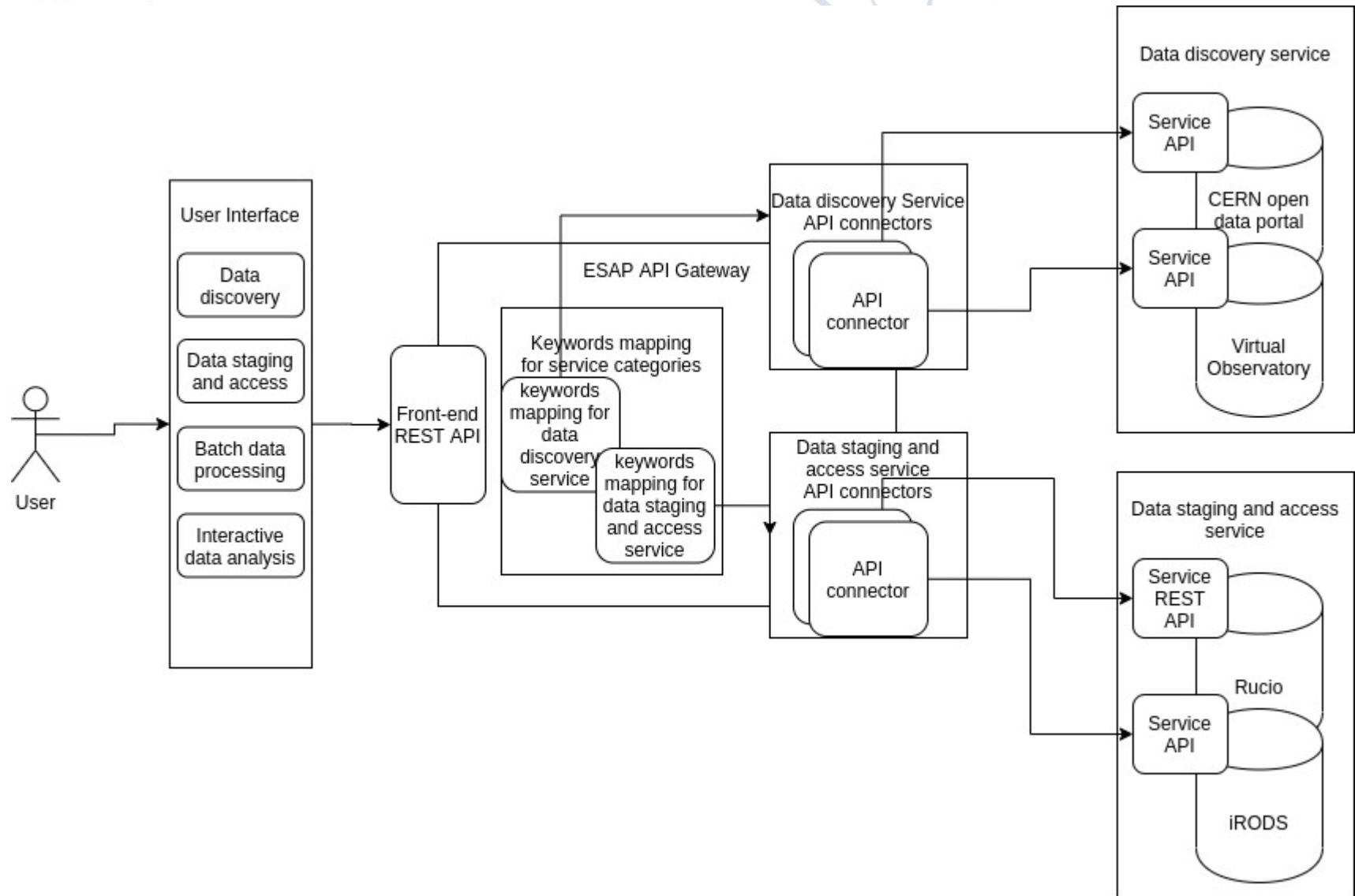
ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Thank you !