



**CS³
MESH⁴
EOSC**

Connecting European Data



Rucio and ScienceMesh

Enabling Data Management for the CS3 Community

Rahul Chauhan (CERN-HSF GSoC), Giuseppe Lo Presti (CERN)

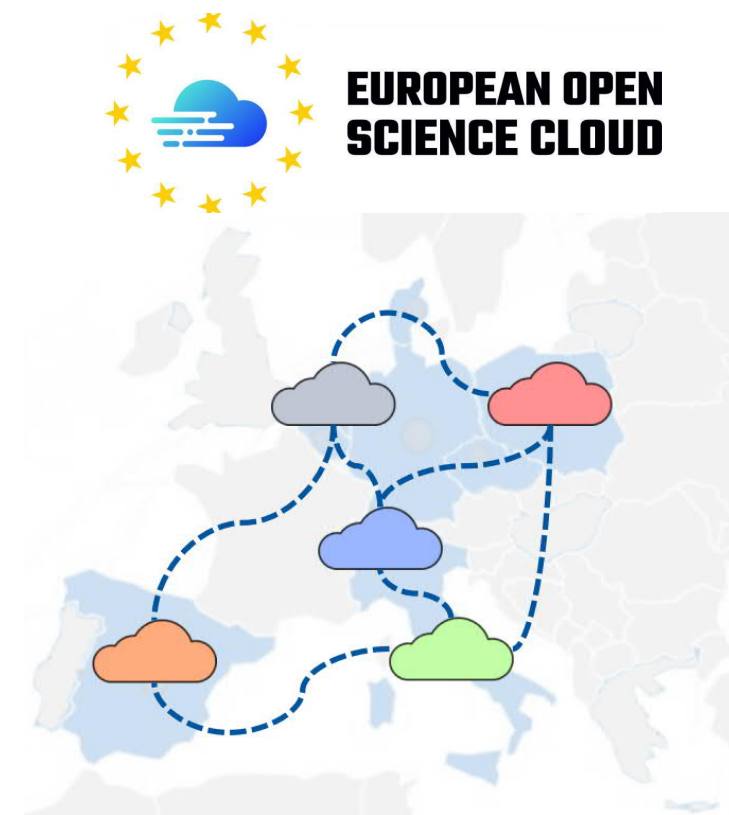


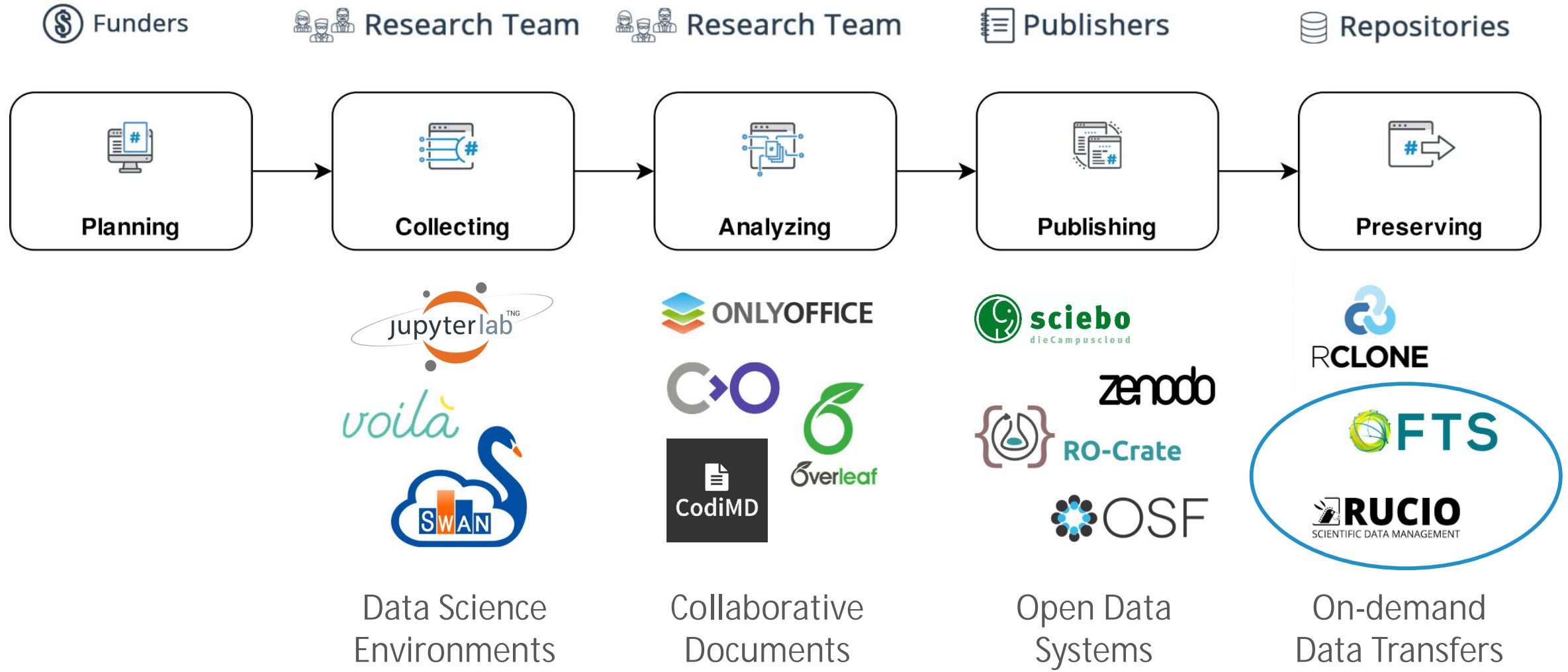
CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.

- # Focus is **users** and **collaborative workflows**
- # **The community**: Cloud Storage for Sync&Share sites
- # Objective: delivering **ScienceMesh**, a federated platform for researchers, educators, data curators, analysts...
 - # Users in the HEP community and beyond will be enabled to more easily collaborate on common projects

More info: <https://cs3mesh4eosc.eu/data-services>

- # Federation of existing CS3 sites
 - # 30+ sites (e.g. CERNBox, DesyBox, GarrBox, universities...)
 - # 400K+ users
- # Global collaborative environment for research
 - # Within EOSC
 - # Share documents, files, projects, data, ...
 - # Connected application hubs
 - # Data/metadata-aware workflows
 - # FAIR: Find, Access, Interoperate, Reuse
- # EU-funded project
 - # 6MEUR, 12 partners, 2020-2022







Data stored at SURF and FZJ.
Initially processing (64x reduction).



LOFAR Surveys Key Science Project
Collaboration between researchers

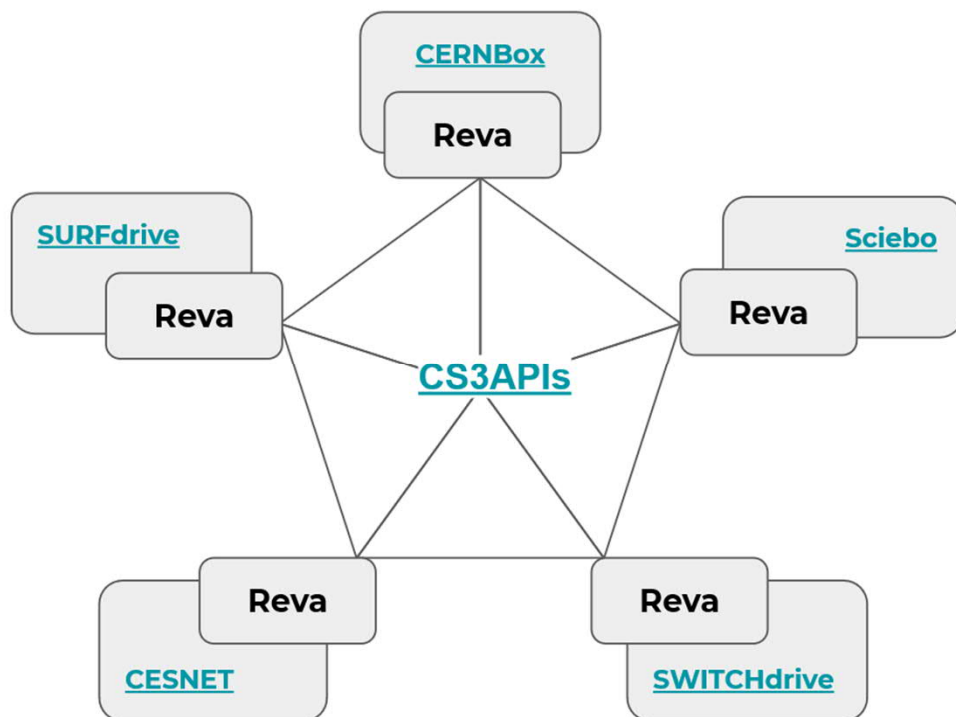
- Leiden University and ASTRON (NL)
- Jagiellonian University, Kraków (PL)



Data shipped to Kraków
for creating science quality images



All participating sites run *Reva*, an Inter-Operability Platform, and interact between each other via generic *CS3APIs*



The [Reva](#) project aims to make cloud storage and application providers inter-operable through a common platform.

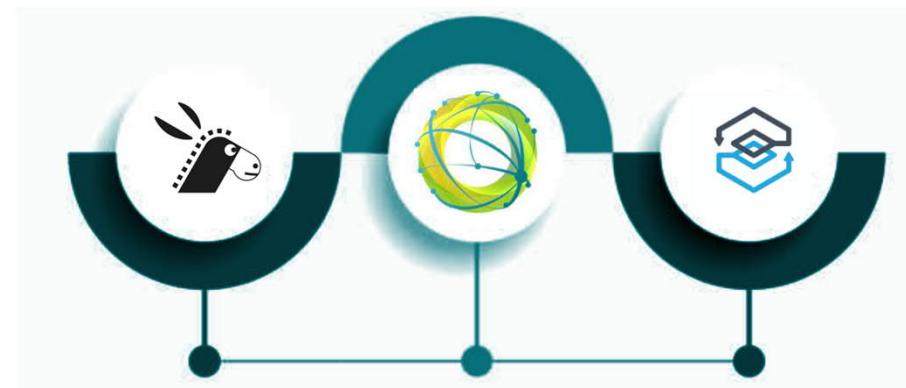
The [CS3APIs](#) are a set of *protobuf* interfaces that enable portability of integrations across different platforms, application providers and data providers.

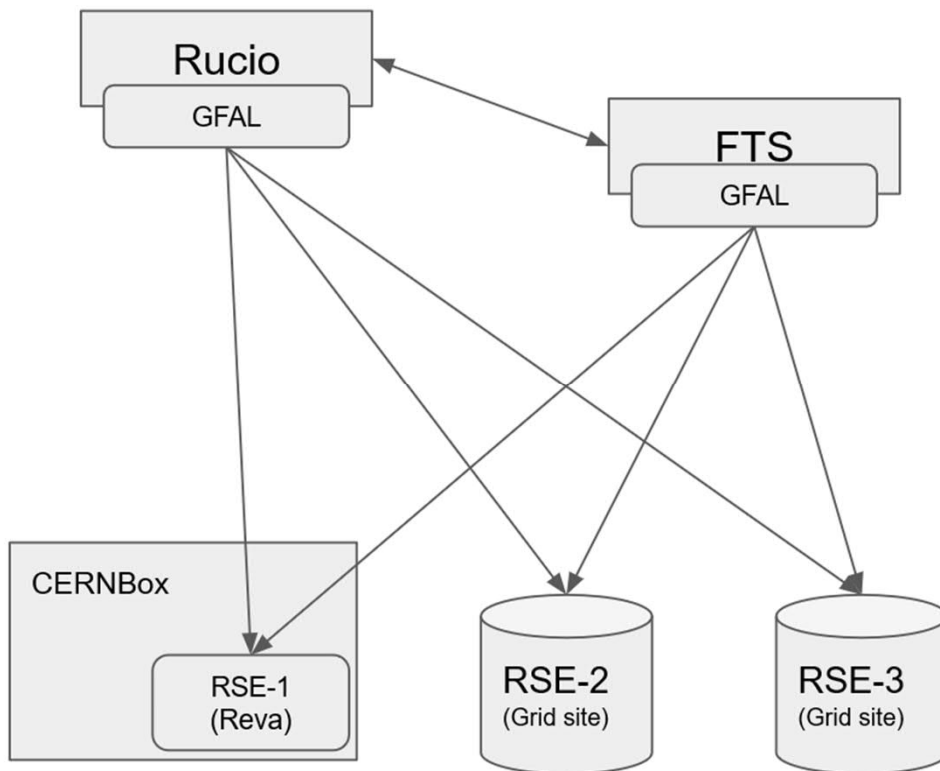


Rucio and CS3APIs To enable Data Management for the ScienceMesh Cloud

Mentee: [Rahul Chauhan](#)

Mentors: [Martin Barisits](#), [Hugo Gonzalez Labrador](#), [Mario Lassnig](#), [Giuseppe Lo Presti](#)





- # [GFAL](#) (Grid File Access Library) is the abstraction layer to manipulate files supporting multiple protocols (HTTP/WebDAV, GridFTP, ...)
- # REVA natively speaks HTTP/WebDAV and GRPC
- # The work:
 - # Enable Reva to understand the control protocol and perform a data transfer according to the data protocol of HTTP TPC.
 - # Extend the HTTP-Plugin of the GFAL library and allow the multiplexer to use the custom flow for ScienceMesh sites based on a new URL prefix (**cs3**).

- # Coming up with a solution
 - # (Deliberately) weakly specified requirements and constraints
 - # Rethink solutions as we became aware of new dependencies
- # Multiple software stacks and technologies
 - # FTS, GFAL, Davix, Reva, Rucio
 - # C++, Golang, Python
- # Lot of Documentation :)
 - # Rucio, Reva, ScienceMesh, XrootD, CS3APIs, HTTP-TPC, WebDav, Gfal2, Davix...

The HTTP TPC in action with Reva

Protocol specificities within GFAL are implemented using plugins.
Our example triggers the use of the HTTP Plugin which uses [libdavix](#) as HTTP client.

```
# gfal-copy -vf \  
  --copy-mode=pull \  
  cs3://reva:19001/remote.php/webdav/home/srcFile \  
  cs3://reva2:17001/remote.php/webdav/home/dstFile
```

```
[root@1480da8e24fc build]# gfal-copy -vf --copy-mode=pull cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile  
Copying 10737418240 bytes cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile  
event: [1629654517438] BOTH GFAL2:CORE:COPY LIST:ENTER  
event: [1629654517438] BOTH GFAL2:CORE:COPY LIST:ITEM cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile  
event: [1629654517438] BOTH GFAL2:CORE:COPY LIST:EXIT  
event: [1629654517438] BOTH http_plugin PREPARE:ENTER cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile  
event: [1629654517449] BOTH http_plugin PREPARE:EXIT cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile  
event: [1629654517449] BOTH http_plugin TRANSFER:ENTER cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile  
event: [1629654517449] BOTH http_plugin TRANSFER:TYPE 3rd pull  
monitor: cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile 635102822 274255052 3175514112 5  
monitor: cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile 467435520 299768217 4674355200 10  
monitor: cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile 407378329 287263948 6110674944 15  
monitor: cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile 379887616 297415475 7597752320 20  
monitor: cs3://reva2:19001/remote.php/webdav/home/large-file cs3://reva2:17001/remote.php/webdav/home/largeFile 364234670 301622886 9105866752 25  
event: [1629654552213] BOTH http_plugin TRANSFER:EXIT cs3://reva2:19001/remote.php/webdav/home/large-file => cs3://reva2:17001/remote.php/webdav/home/largeFile
```

- # GSoC report: [Google Summer of Code Work Product Submission](#)
- # Authentication and Authorization still under study
 - # Going to adopt IAM as deployed for ESCAPE (cf. yesterday's talk)
 - # The real added value is enabling data transfers between WLCG/ESCAPE sites and ScienceMesh sites
- # Explore UI integrations
 - # Possible synergies with the *Data-Lake-as-a-Service* work successfully deployed within ESCAPE!



**CS³
MESH⁴
EOSC**

Connecting European Data

Thank you!
Discover more on...

 [cs3mesh4eosc.eu](https://www.cs3mesh4eosc.eu)

 [company/cs3mesh4eosc](https://www.linkedin.com/company/cs3mesh4eosc)

 [CS3org](https://twitter.com/CS3org)

 [CS3MESH4EOSC Project](https://www.youtube.com/channel/UCHKcZEKMaXjCvc3MLFjFxbw)

<https://www.youtube.com/channel/UCHKcZEKMaXjCvc3MLFjFxbw>

Credits to H. Angenent and J. Moscicki for the CS3MESH4EOSC project material



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.