

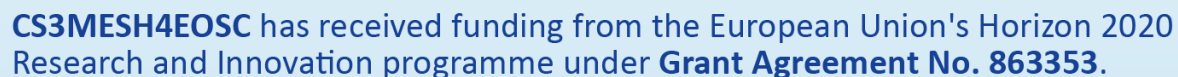


# Connecting European Data



# The Science Mesh and its core applications

# Holger Angenent



- Science Mesh - global collaboration platform for researchers
  - Integrating (open) data and (open-source) tools
  - Provided as building block for European Open Science Cloud (EOSC)
- Open, bottom-up approach
  - Applications: working closely with user communities we take existing best practices, services and technologies, improve them and open them up for other scientific communities



**Science  
Mesh**



**EUROPEAN OPEN  
SCIENCE CLOUD**

What are our users doing?

- # LHC produces unprecedented volumes of data
  - # Raw stream from detectors: **600TB per second**, or **50 000 PB per day**
    - # The raw data per event ~1MB, 600 million events per second
  - # Storing only fraction of this (total volume of **all stored data: 350PB**)
    - # Filtering of events/data need to be smart and fast!
  - # Large distributed infrastructure for transferring and large-scale processing
- # Constant innovation in tools and methods for analytics.
  - # Data streams from LHC increase with each upgrade
- # Distributed teams of scientists from institutes all over the world
  - # a variety of storage systems and processing tools.
- # **One of the challenges: providing tools in this distributed environment for effective collaboration in Data Science**



## MISLAND - NORTH AFRICA - MONITORING INTEGRATED SERVICES FOR LAND DEGRADATION

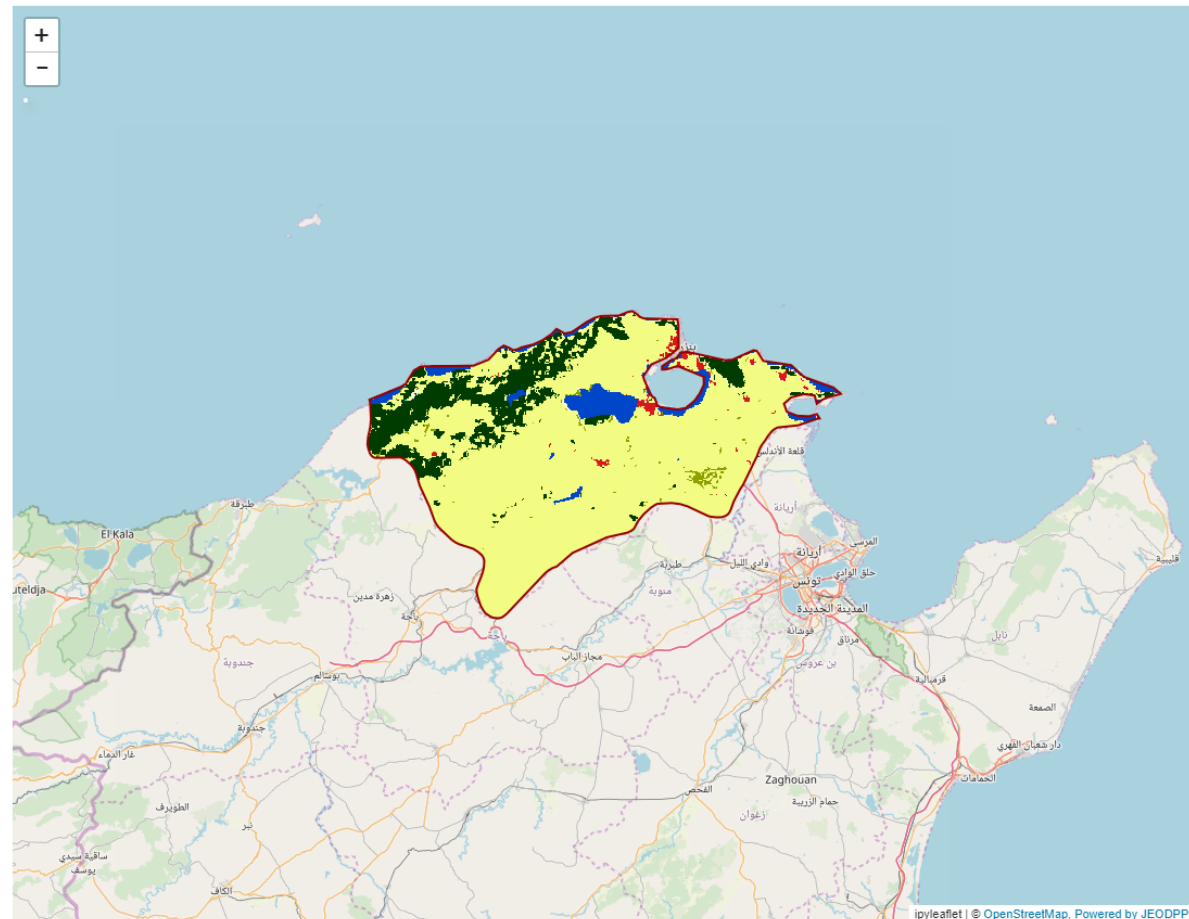
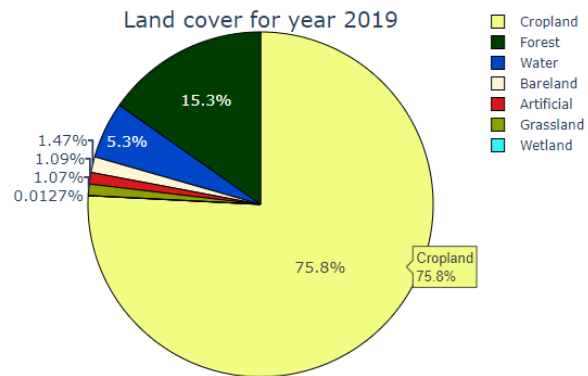


### SELECT AREA

Country: Tunisia Region: Bizerte

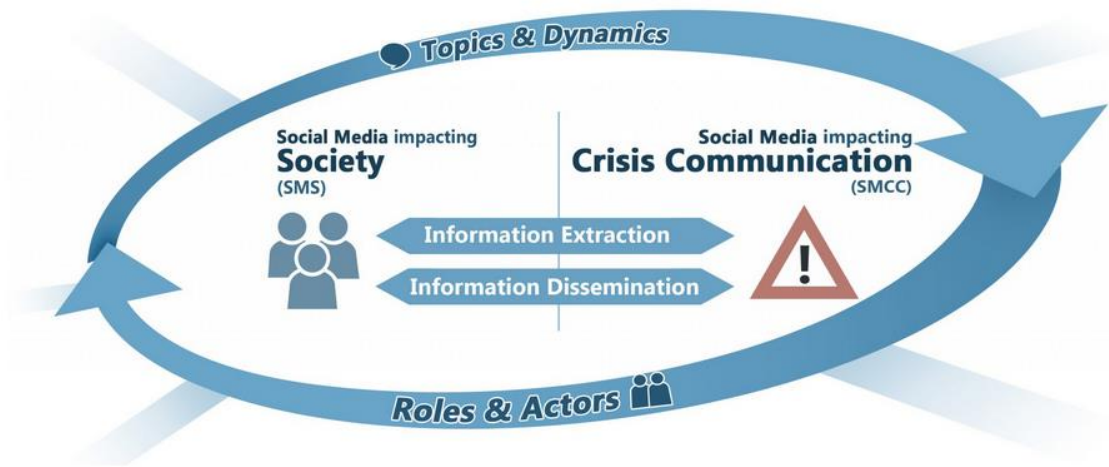
< **SDG 15.3.1** VEGETATION LOSS/GAIN FOREST CHANGE FOREST FIRE >

Select SDG 15.3.1: LULC Analysis: Land Cover Year: 2019





- Social Media Analytics for Society and Crisis Communication
- <https://social-media-analytics.org/>



### A Global Network of Excellence



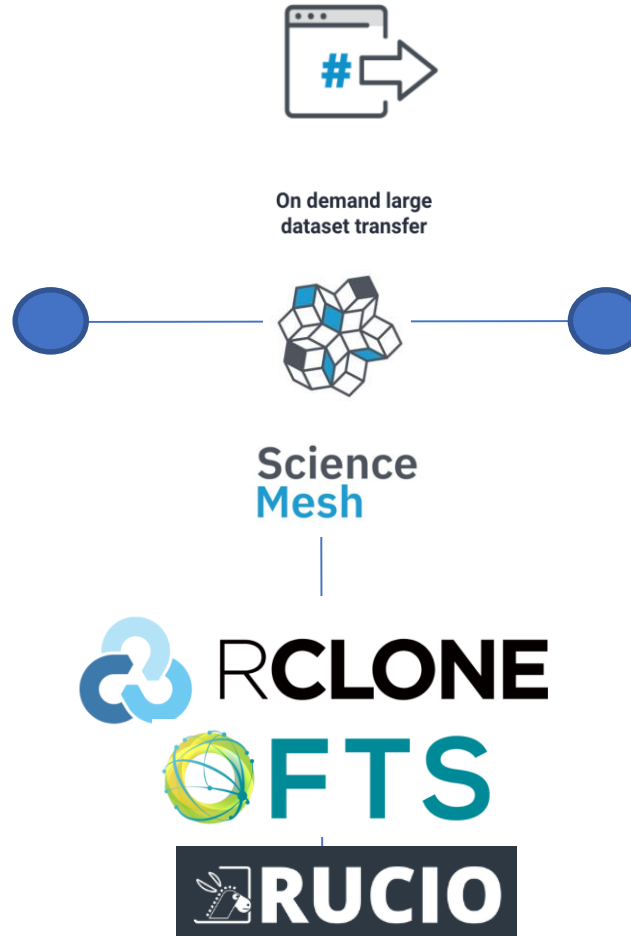


*Data stored at PSNC, SURF and FZJ.  
Initially processing (64x reduction).*



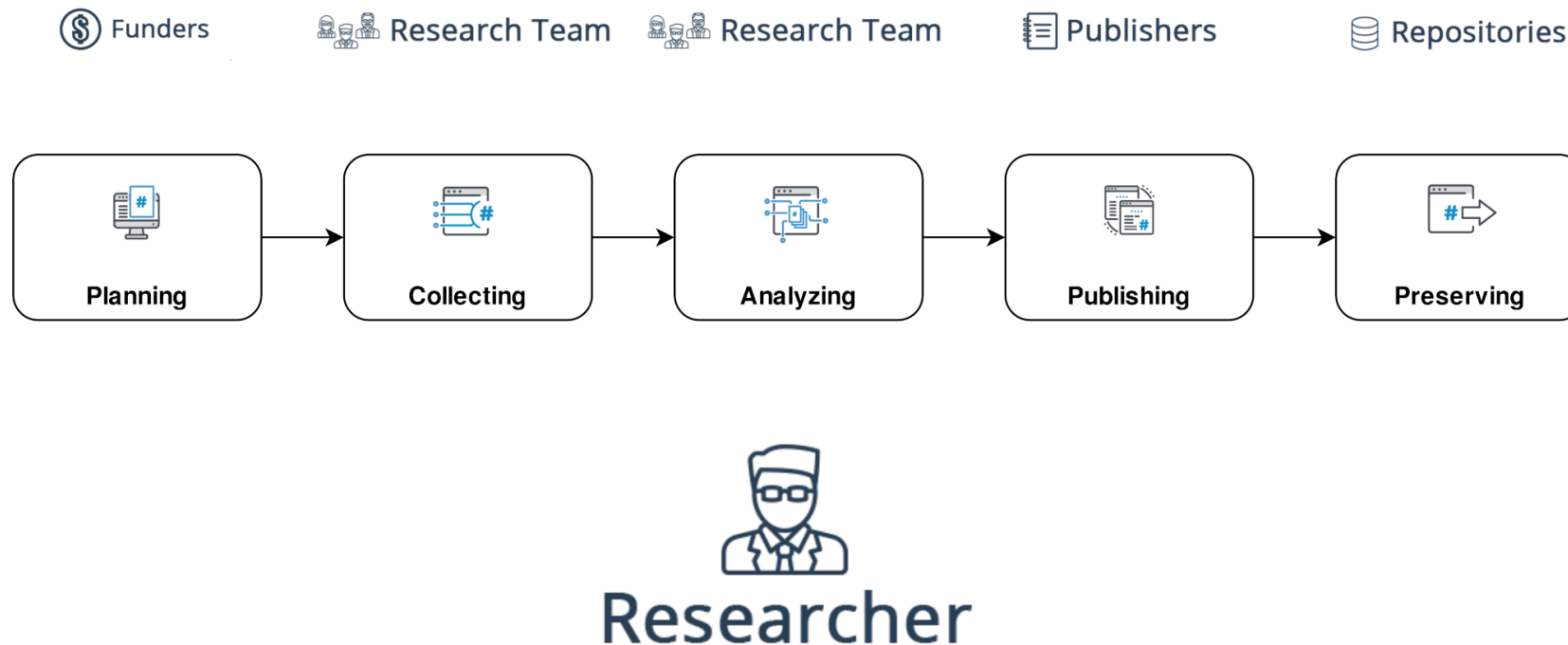
LOFAR Surveys Key Science Project  
Collaboration between researchers

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



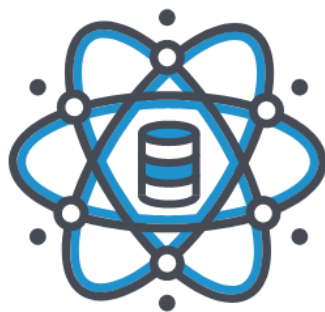
- Of course, there are many other use cases which can profit from Science Mesh:
  - Cross-border data sharing
  - In need of streamlined work flows
  - Intuitive interfaces for data transfer
  - Doing computations with data stored in your EFSS system
  - Pushing the same data to HPC cluster
  - Pushing data to other users (from same or other EFSS)
  - Direct editing of files from your EFSS (with non-MS editors...)

More general, what does the typical researcher's workflow look like?



What does Science Mesh offer for the users?

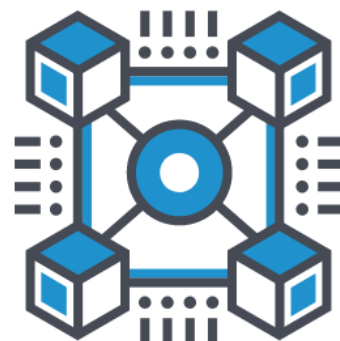
### THE SCIENCE MESH DATA SERVICES



#### Data Science Environments

To facilitate collaborative research and enable **cross-federation sharing** of **computational tools, algorithms** and **resources**.

[Read more](#)



#### Open Data Systems

To add **metadata** and publish datasets with **persistent identifiers** directly on the **Science Mesh** sites or to external data **repositories**.

[Read more](#)



#### On-demand Data Transfers

To allow efficient **data-based collaboration** on **on-demand** basis.

[Read more](#)



#### Collaborative Documents

To be able to **simultaneously edit documents** in **safe, EU-based**, cloud environments.

[Read more](#)



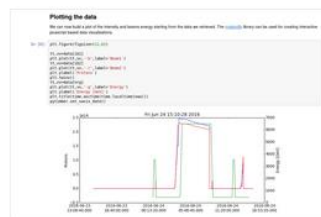
Basic Examples  
ROOT Primer  
**Accelerator Complex**  
Beam Dynamics  
Machine Learning  
Apache Spark  
Outreach  
AWAKE

## Accelerator Complex

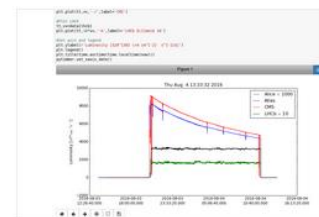
This gallery shows examples of machine studies relative to the CERN accelerators' complex.

Open in SWAN

LHC Page1



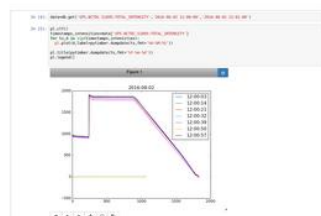
Experiments' Luminosities



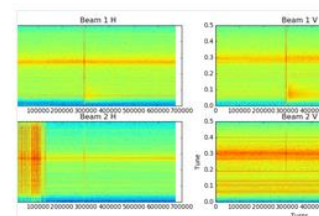
PyTimber Tutorial



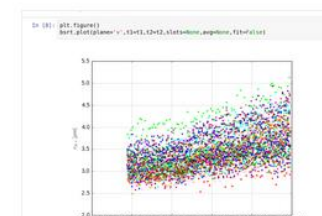
SPS Intensity



LHC BBQ Example



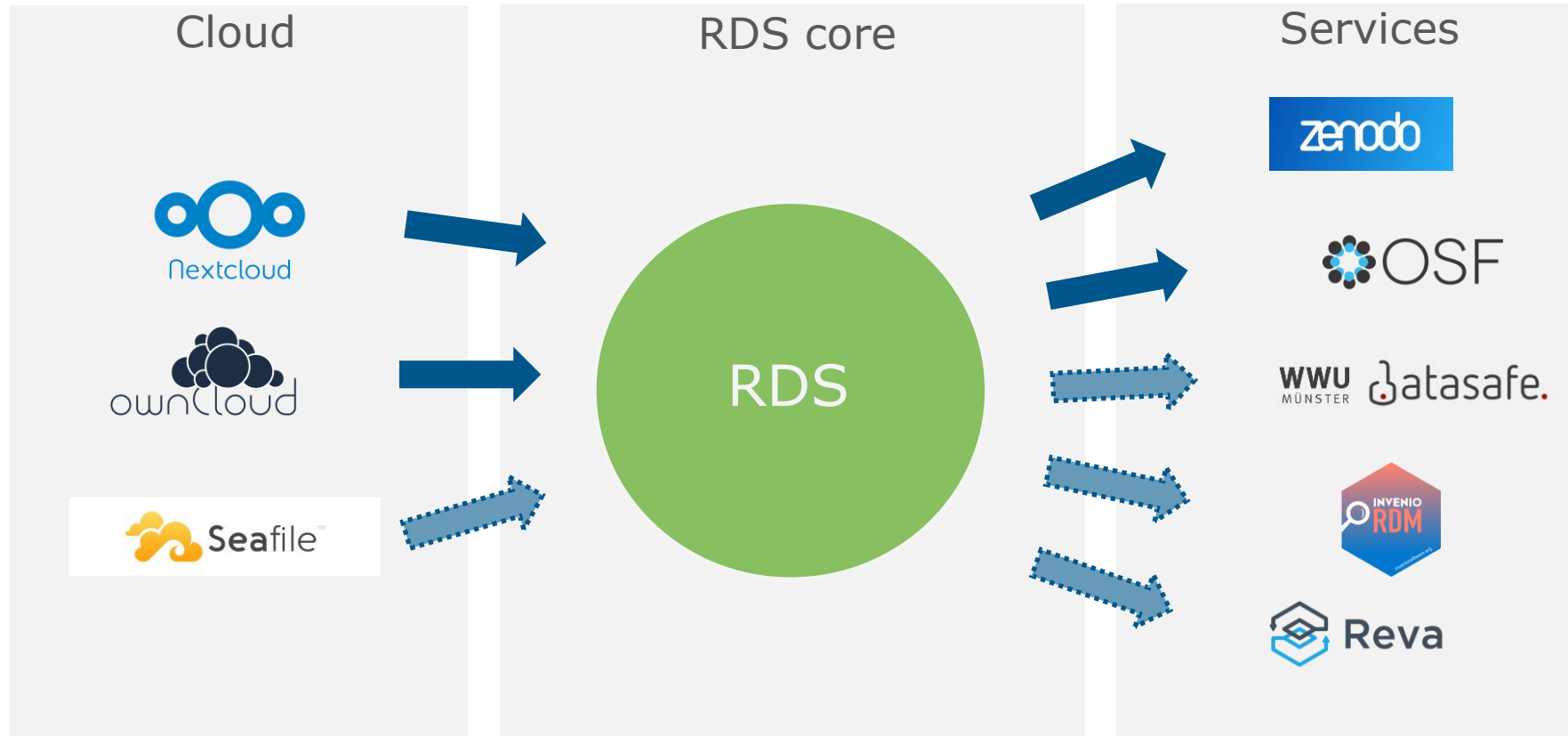
BSRT Example



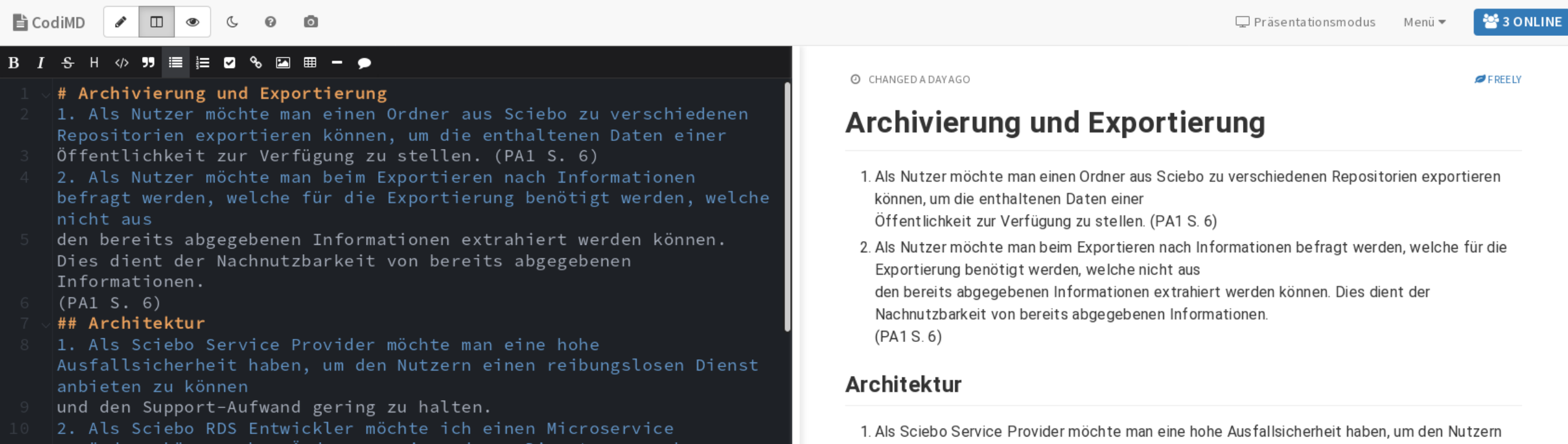
← Previous  
ROOT Primer

Next  
Beam Dynamics →

## Integration of different research data management tools



- Lightweight markdown editor
- No big overhead or complicated menus, just type your content
- Ideal as note taking tool, grasp some ideas while talking about it, easily write simple structured documentation
- Data still resides in EFSS

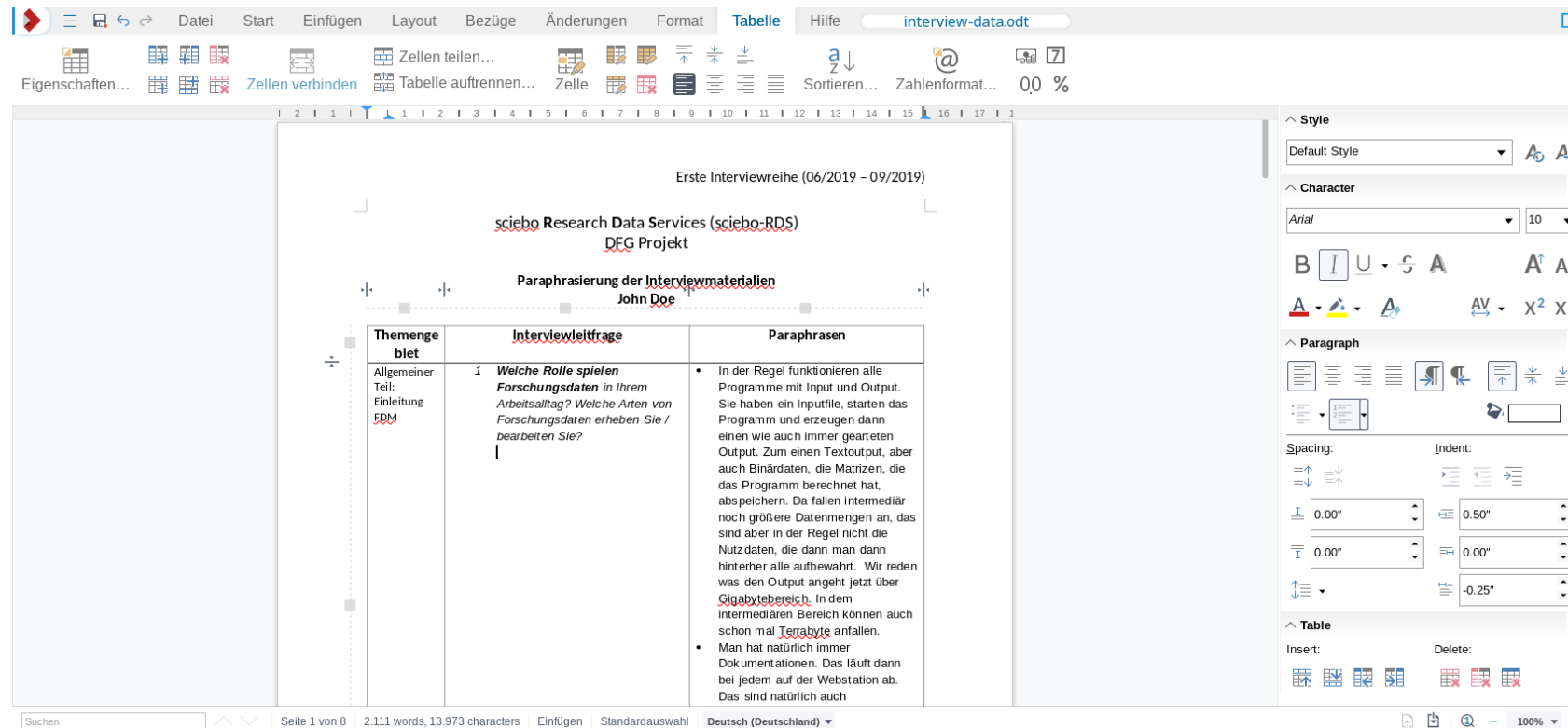


The screenshot displays the CodiMD web interface. The top bar includes the CodiMD logo, a toolbar with icons for editing, viewing, and settings, and a status bar showing 'Präsentationsmodus', a 'Menü' dropdown, and '3 ONLINE' users. The main area is split into two panes. The left pane is a dark-themed code editor showing a markdown document with the following content:

```
1 # Archivierung und Exportierung
2 1. Als Nutzer möchte man einen Ordner aus Sciebo zu verschiedenen
3 Repositorien exportieren können, um die enthaltenen Daten einer
4 Öffentlichkeit zur Verfügung zu stellen. (PA1 S. 6)
5 2. Als Nutzer möchte man beim Exportieren nach Informationen
6 befragt werden, welche für die Exportierung benötigt werden, welche
7 nicht aus
8 den bereits abgegebenen Informationen extrahiert werden können.
9 Dies dient der Nachnutzbarkeit von bereits abgegebenen
10 Informationen.
11 (PA1 S. 6)
12 ## Architektur
13 1. Als Sciebo Service Provider möchte man eine hohe
14 Ausfallsicherheit haben, um den Nutzern einen reibungslosen Dienst
15 anbieten zu können
16 und den Support-Aufwand gering zu halten.
17 2. Als Sciebo RDS Entwickler möchte ich einen Microservice
18 entwickeln, um die Daten in einer Datenbank zu speichern und sie
19 für die Nutzer verfügbar zu machen.
```

The right pane is a light-themed preview of the rendered markdown document. It shows the title 'Archivierung und Exportierung' and the list of requirements. The status bar at the bottom of the preview indicates 'CHANGED A DAY AGO' and 'FREELY'.

- Powerful editor especially for, but not limited to Open Document format
- Texts documents, spread sheets and presentations can be edited



- \* On-demand data transfers in two flavours

- \* Ad-hoc data transfers

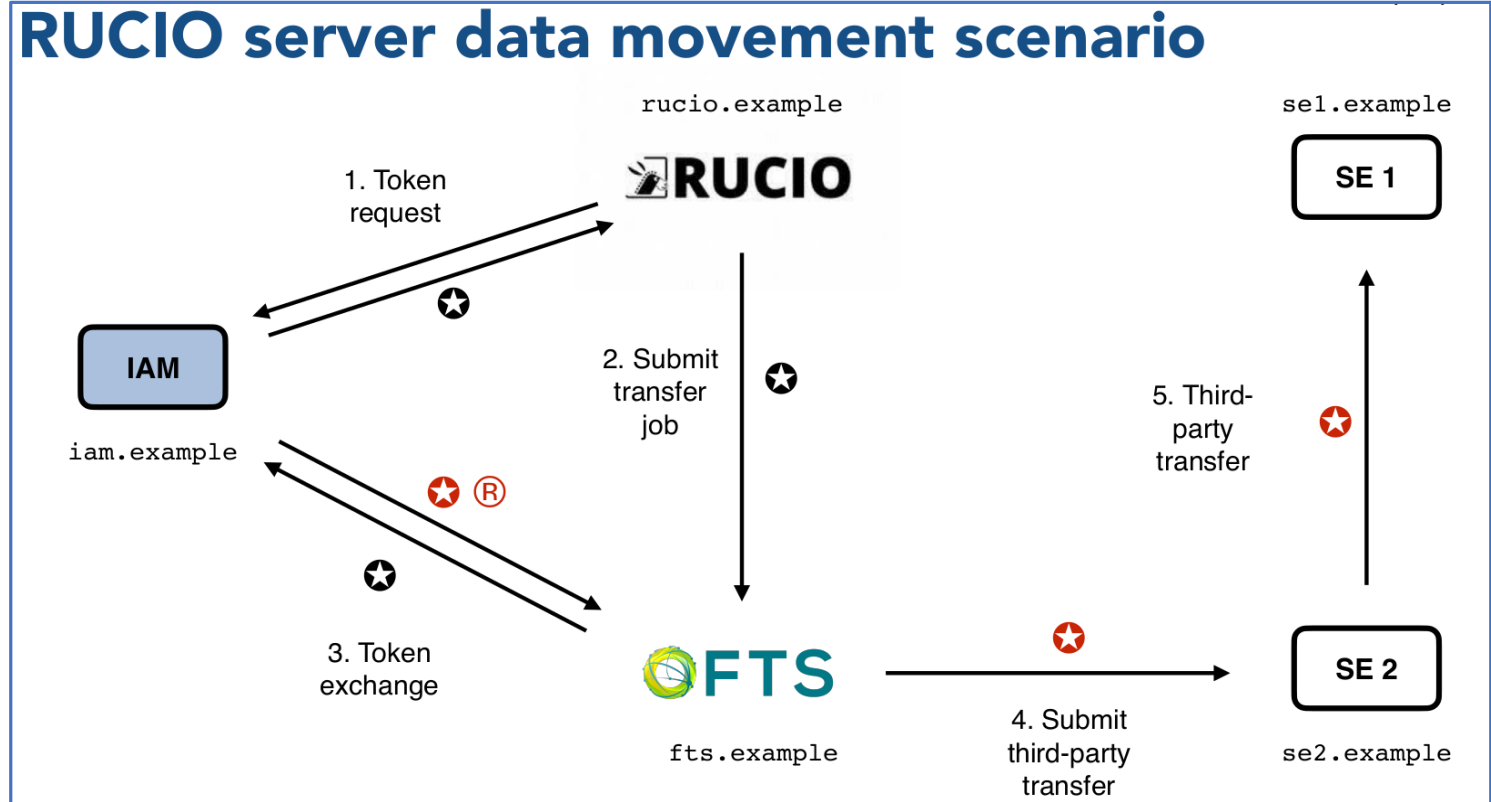
- \* For small/medium sized data transfer needs
    - \* Individual researchers or individual research groups having access to their own sync-and-share service
    - \* Typically local identity management system
    - \* Based on [Rclone](#)
      - \* Rclone is a tool to manage files on cloud storages

- \* Managed data transfers

- \* Connecting Big(ger) Science with sync-and-share (EFSS)
    - \* Scientific communities having access to multiple storage systems and EFSS systems
    - \* Federated AAI using [Indigo IAM](#)
    - \* Based on [Rucio](#) & [FTS](#)
      - \* Rucio is a data management tool
      - \* FTS is a data transfer scheduler



- \* EFSS as Rucio/FTS storage endpoint
- \* From EFSS issue
  - \* Uploads
  - \* Downloads
  - \* Replications
  - \* Metadata queries
  - \* .....



(Slide stolen from A. Ceccanti)



**Thank you!**  
Discover more on...

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

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

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

 [CS3MESH4EOSC Project](https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw)

<https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw>



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