



Welcome!

<https://indi.to/cs3mesh-kickoff>

This Morning

- Introduction to the Project
- Round table
- Discussion

CS3 Science Mesh Introduction

Project Kickoff Meeting

30 Jan 2020

Naming mess



Science Mesh



CS3MESH4EOSC

Naming mess



community
conferences

Science Mesh



the resulting service

CS3MESH4EOSC

EC-funded project



CS3MESH4EOSC Project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 224732

CS3MESH4EOSC

- New EU-funded project to capitalise on CS3 Community
 - 2020, 3 years, 5.8M EUR



- Objectives

- Deliver a global collaboration service for European(*) research, education and public institutions
- Create an environment for collaboration on technology, applications, use-cases, software and operations

(*) Europe, Australia and beyond

Project Partners


Coordinator



EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

Address

Route De Meyrin Cern
1211 Geneva 23


 Switzerland

Activity type

Research Organisations


RESEARCH

DANMARKS TEKNISKE UNIVERSITET

 Denmark


UNI / NREN

SURFSARA BV

 Netherlands


HPC / NREN

INSTYTUT CHEMII BIOORGANICZNEJ POLSKIEJ AKADEMII
NAUK

 Poland

HPC

CESNET ZAJMOVE SDRUZENI PRAVNICKYCH OSOB

 Czechia


NREN

AARNET PTY LTD

 Australia


NREN

SWITCH

 Switzerland

NREN

WESTFAELISCHE WILHELMS-UNIVERSITAET MUENSTER

 Germany

UNI

AILLERON SA

 Poland


TECH

CUBBIT SRL

 Italy

TECH

JRC -JOINT RESEARCH CENTRE- EUROPEAN COMMISSION

 Belgium

EC RESEARCH

FUNDACION ESADE

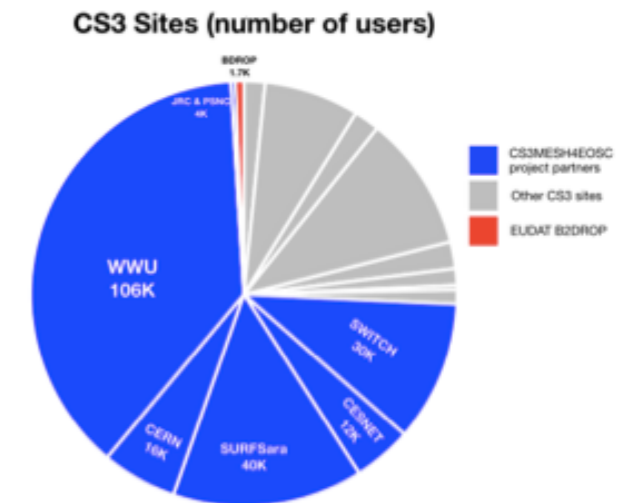
 Spain

BUSINESS SCHOOL

WP Number⁹	WP Title	Lead beneficiary¹⁰
WP1	Project Management	1 - CERN
WP2	Federated Infrastructure	3 - SURFSARA BV
WP3	Technology and Foundation	1 - CERN
WP4	Users and Applications	4 - PSNC
WP5	Dissemination, exploitation & outreach	2 - DTU

EFSS Starting Point

- 7 sites to **create initial infrastructure**
 - **PSNCBox (Seafile)**
 - **ScienceData (Nextcloud)**
 - **SwitchDrive, Surfdrive, CESNETBox, CERNBox (Owncloud)**
- **connect existing, sustainable services**
 - all major EFSS platforms included (multivendor)
 - 200K+ existing users, 10PB of sync&share data, >1billion files and objects
- The infrastructure will be **gradually expanded** and **integrating the entire community**, education and research in Europe and beyond.



Future Federated Analysis Platform

Advancing state of the art

One-click to create user groups, share projects and data



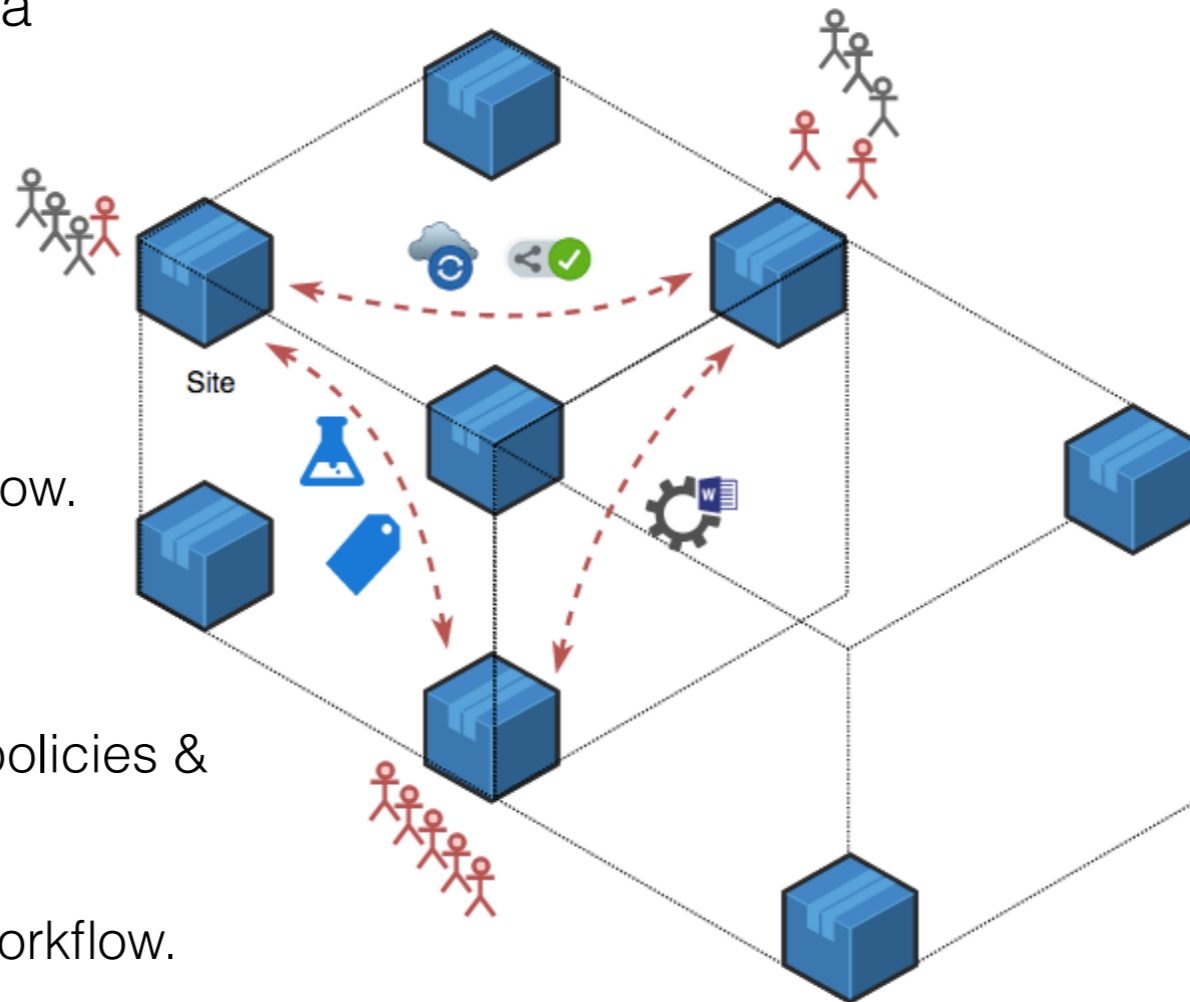
Domestic and remote users in the same collaborative workflow.

Collaborate on data across institutions, respecting local policies & GDPR

Application&data workflow.

Data available on **all devices**: mobile, laptops, desktops

Full metadata awareness in the research workflow.



Global Computing Data Fabric & Digital Repositories

Integrated with data sharing capabilities

- For example, a medical researcher at the University of Southern Denmark, on his home service (**sciencedata.dk**) can share an anonymised, but classified patient dataset with a collaborating, trusted, professor at the University of Amsterdam just like that: The professor appears in the sharing search field of the Dane and the folder immediately appears in the left-menu item “Shared with me” of the Dutch professor on his home service (**surfdrive.nl**). All trusted and vetted behind the scenes via eduGAIN and the deployed trust model.
- As an example scenario of what will be possible with the Science Mesh: consider a lecturer at Charles University in Prague teaching statistical programming in R for data scientists: From the web interface on his home data service (**owncloud.cesnet.cz**), the lecturer will be able to share example notebooks with his students who will be able to execute the notebooks via the Jupyter notebook service operated by CERN and integrated with CERNBox (**cernbox.cern.ch**). The results will be automatically synchronized to the **cloudstor.aarnet.edu.au** collaboration space in Australia, where local collaborators in Melbourne will simply run these notebooks to process data stored in their system. With the results in hand, they will add relevant metadata and publish the curated datasets to an Open Data repository such as Zenodo. All modifications to the notebooks will automatically be synchronized and available in the same collaboration space for both the students and the lecturer in Prague. Each of the actors of this workflow will interact with the others without leaving their home’s institution service interface.

EUROPEAN OPEN SCIENCE CLOUD



share application services

data analysis & notebooks

collaborate interactively

tag files & add metadata

synchronize

On-premise infrastructure

high volume data transfer

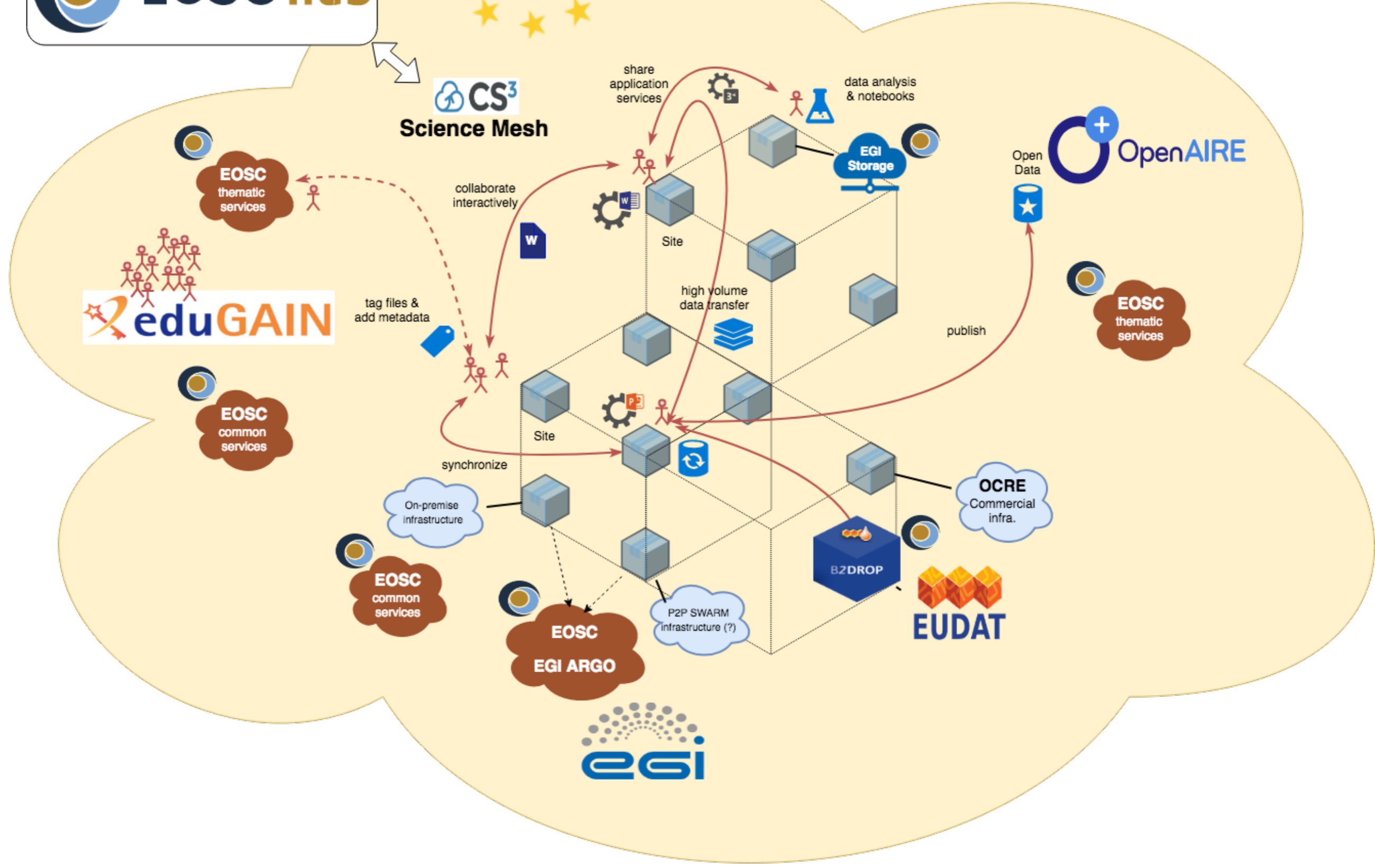
publish

OCRE Commercial infra.

P2P SWARM infrastructure (?)


B2DROP

EUDAT












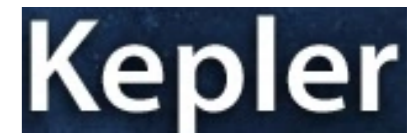
Pilot users

Students, educators and researchers at large

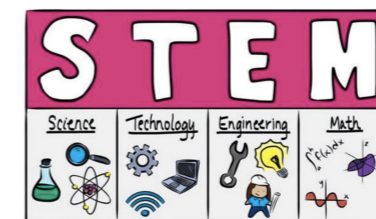
-  • connecting large university campuses

Target specific research and application areas

-  • Earth Observation (Copernicus Sentinel)
-  • High Energy Physics (LHC)
-  • Astroseismology (NASA Kepler telescope)
-  • Cultural Heritage and Archival Collections
-  • Material Science
-  • Astrophysics (LOFAR)
-  • Plasma Physics
-  • Video processing technology development
-  • Diabetes Research



Pacific and Regional Archive for Digital Sources in Endangered Cultures



Edu & Outreach

Collaborative Workflows

Integrate existing experience and technology



Share, access, synchronize



Metadata&tagging, Open Data (OpenAIRE, Zenodo,...)



Data Science: Jupyter Notebooks (SWAN,...)



Collaborative editing, Latex, Markdown, Indico, ...



On-demand data transfers (Rucio, FTS, FileSender,...)

Technology Integration

- **Connect EFSS platforms with research services and digital repositories**
 - Promote vendor-neutral APIs and protocols
 - Open-source software development and service delivery
- **Sustainable technology**
 - Collaboration on technology with **all** EFSS, storage and application vendors

Interoperability

- Add thin layer on top of existing services



- Use existing fabric



- authentication, monitoring,...



- close collaboration with GEANT, EGI, EOSC-hub,...



OPENCLOUDMESH

EFSS Native

A vendor neutral standard under the GÉANT umbrella

- Use existing standards

- Introduce new APIs only if needed



Project Goals

Interoperable Service Foundation

SO1: Create the software foundation for the Science Mesh service: extend and standardise vendor-neutral APIs and protocols (based on existing work on OCM and CS3APIs), provide a reference implementation based on community software and integrated with a choice of commercial implementations and plugins.

SO2: Integrate the Science Mesh service into the EOSC-hub catalog to enable integration of Science Mesh with the existing EOSC-hub services and to complement the service catalog with new, interactive collaboration sharing capabilities.

SO3: Create an open software environment and process for technology providers and users (technical experts in the research community) of Science Mesh to provide contributions. This is to allow them to develop new applications and solutions/components to be easily integrated in the service, and at the individual sites. Project will achieve this by designing, developing and promoting interoperability protocols for sharing storage, data and applications. This environment can be used to validate and evolve industry standards for cloud services.

Infrastructure: Sustainability and Federated Sites

- **SO4: Establish a trusted sharing and collaboration model and policies for the Science Mesh service** (authentication, authorization, resource access, delegation, responsibilities, etc.) based on the state-of-the-art AAI technologies and deployed trust model based on EduGAIN and additional services (such as EduTEAMs, AARC or ScienceCollaborationZone,...). KPI: % of European researchers who may be reached by the service via a known identity mechanism; ability to form groups of researchers from different institutions. Impact: key enabler of trust and reachability of the users.
- **SO5: Define criteria and establish the measures for achieving service excellence of federated sites via Quality of Service**, monitoring, testing, reporting and accounting, define criteria for sites joining the federation, as well as continuous assessment of all relevant technologies and components at TRL-6 or higher. KPI: measures effectively achieved. Impact: service fully operational.
- **SO6: Assure low and predictable costs for sites and technology providers to join the service/infrastructure in a sustainable way.** This will be done by supporting exchange of knowledge, develop common operational skills for the service and define requirements as well as support tools to extend the reach of the service to integrate new actors at low and predictable cost. KPI: number of connected sites. Impact: service fully supported within the community and sustained by the participating sites.

Applications and Users

SO7: Integrate high-level application services into the sharing mesh platform and create additional service offering for:

- a) **Interactive Data Science Environments** based on Jupyter Notebooks and the SWAN service,
- b) **Collaborative editing of content** for Office documents based on tools such as Only Office and Collabora Online,
- c) **Open Data Applications** such as FAIR Metadata Collections or Open Data Publishing systems based on OpenAIRE,
- d) **Distributed Processing Workflows and Large Dataset Sharing** based on state-of-the-art protocols (such as S3) and existing file transfer infrastructures (such as GridFTP).

SO8: Validate the system in close connection with research groups and users. This will be done for the selected number of benchmark use-cases from different scientific disciplines as well as for new use-cases identified during the execution of the project

Applications and Users

SO7: Disseminate project results during the project execution to provide training and networking opportunities for the CS3 community (and beyond: attract new user communities).

SO10: Deliver the impact assessment of the service and technology in relation to new business cases and market evolution. This will be done by identifying and analysing the financial impacts and potential business cases with relevant business partners, and offering guidelines towards a longer term preparation for market changes based on studies of the Open Source communities as well as established bodies of literature.

WP2 : federated infrastructure

- Establish a lightweight, minimal add-on infrastructure based on existing self-sustained CS3 sites and existing sustainable operational infrastructure in the EOSC-hub project.
- Establish operational and deployment processes, procedures, policies and tools for the federated community service at the site level.
- Implementing accounting, monitoring, incident- and change management and security procedures and policies.
- Establish trust relationships, the authentication and authorization infrastructure, and the defining operational procedures and policies and provide documentation.
- Support for sites that want to join this federated infrastructure.

WP2: Federated Infrastructure



WP5
sciencemesh.io
website

dashboards
 packages
 guidelines
 policies



Integration
 Task 4

Identity, AAI
 EDUGAIN,
 EGI-CheckIn,...

Task 1

Grouping
 EDUTEAMS,...

Task 2

Acceptance

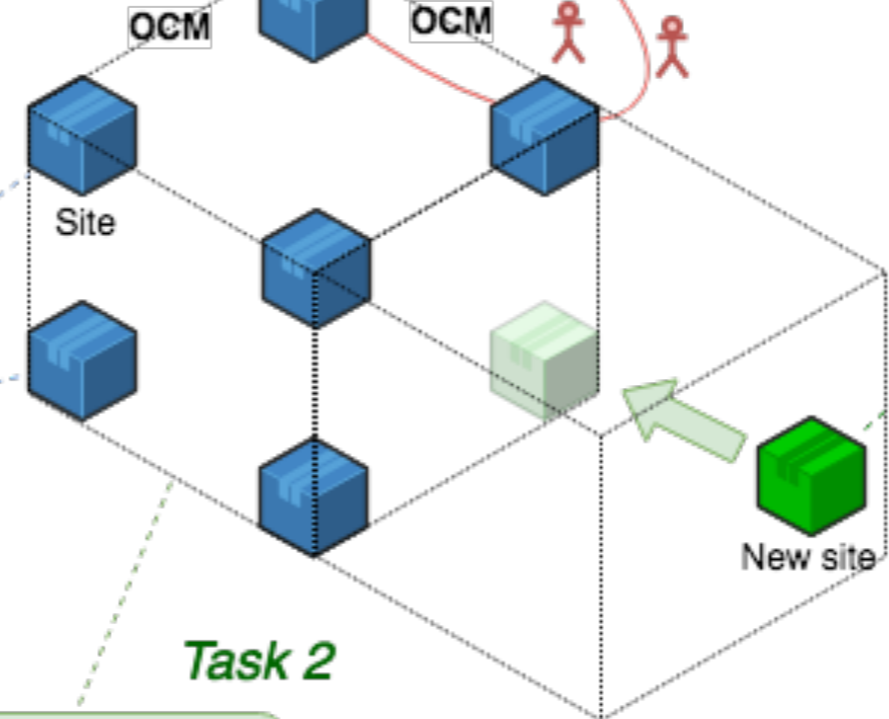
criteria,
 interoperability,
 tests,
 automation,
 documentaiton

**Accounting,
 Monitoring,
 QoS levels**
 ARGO

Task 3



monitoring



Registry

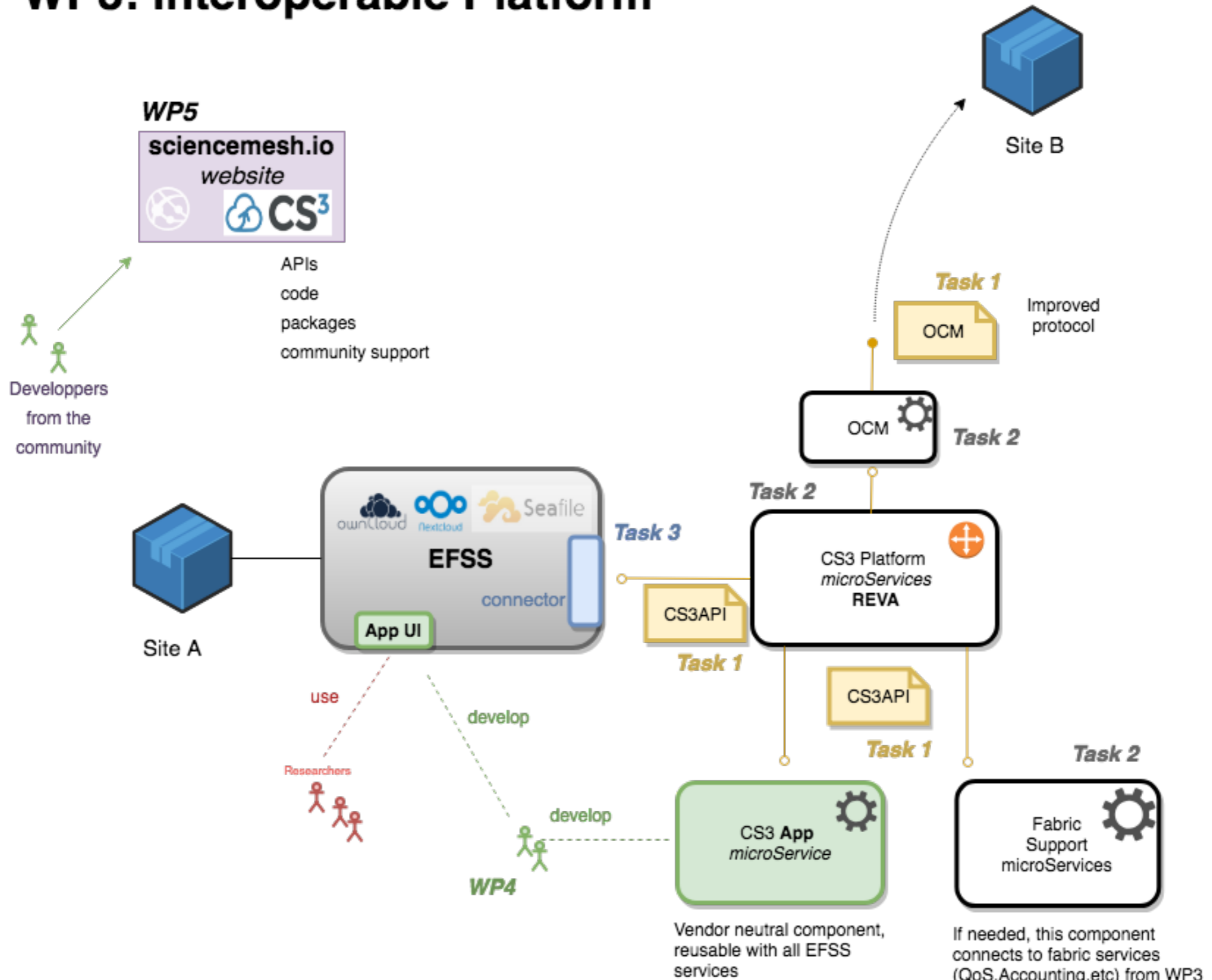
Task 2

Registry of trusted sites with associated metadata (e.g. authorization decisions)

WP3: Interoperability platform

Provide the technical foundation for the site and application interoperability (protocols, APIs, implementation and runtime architecture) and a blueprint for evolution of the shared community service based on common protocols and interfaces.

WP3: Interoperable Platform



WP4: Users & applications

To integrate core applications into the federated service for inclusion into the production workflows and daily practice of all considered research groups.

The results of the work performed by the leading partners in each application Task (1-4) will be tested with benchmark use-cases and respective user community. This will be performed as a part of each task. Early adopter validation will be performed before applications are made available to the entire community. Application tasks will provide feedback to WP3 to extend OCM API and CS3APIs via Task 5.5 in order to enable exploiting benefits from the federation layer.

Task 4.1: Data Science Environments

Task 4.2: Open Data Systems

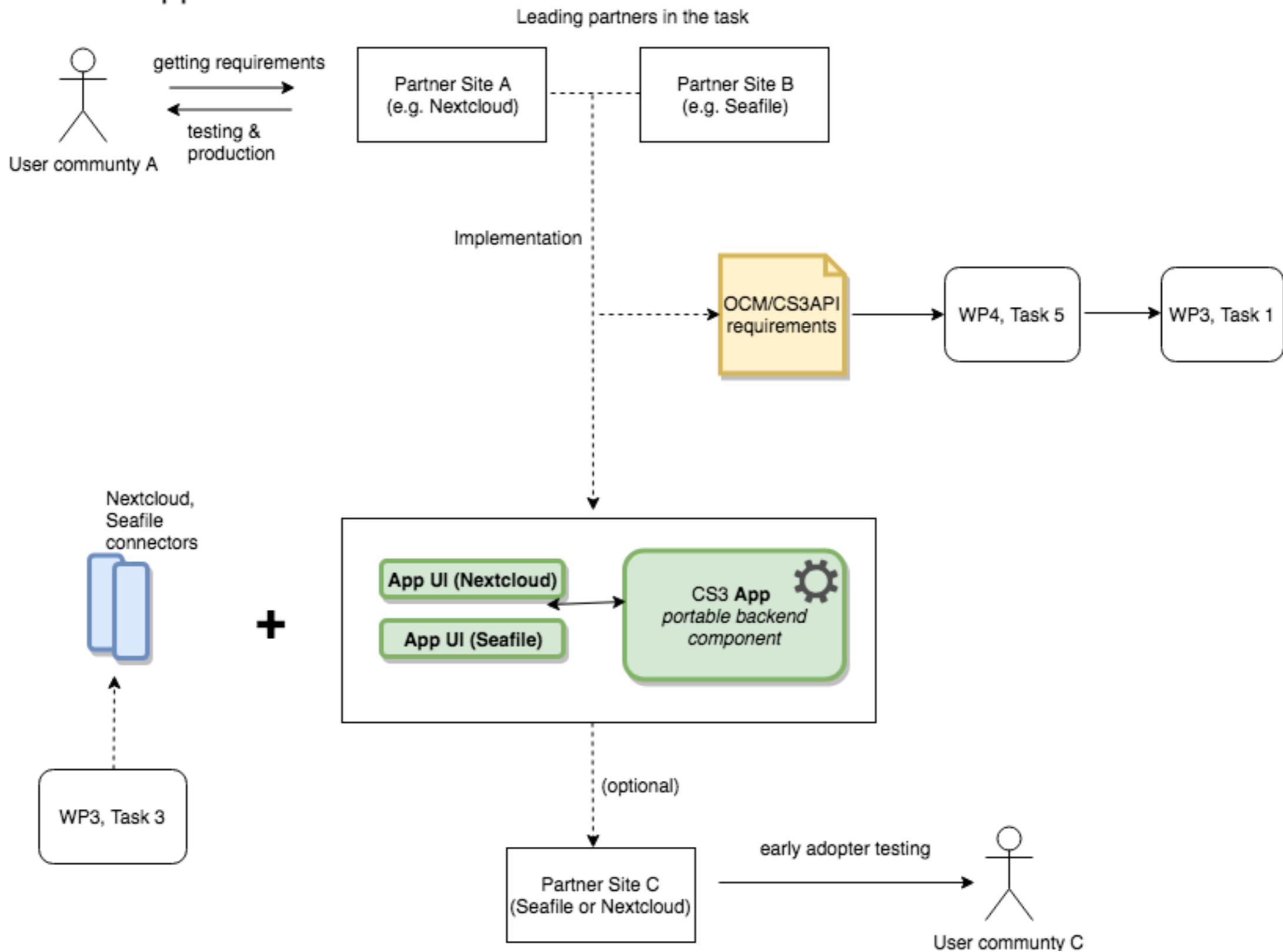
Task 4.3: Collaborative Documents

Task 4.4: On-demand data transfers

Task 4.5: Common application requirements and deployment mechanisms

Task 4.6: Evaluation of the dynamics of collaborative users communities and governance of the CS3 mesh consortium

WP4 Application Tasks 1-4



Bootstrapping the Community

Early Adopters

ARCHIVER

COS

EGI

Elettra Synchrotron

ESCAPE

GO FAIR

HIFIS

RENATER

Early Adopters

What we offer:

- Possibility to make your voice heard (needs)
- Privileged access to prototypes
- Possibility to provide early feedback

Early Adopters

What we ask from you:

- Your use cases, especially “game changers”?
 - * Or those of the communities you support;
- Your experience - challenges you can foresee?
- Testing and Feedback

Early Adopters

Today let's:

- Discuss obvious use cases and target communities;
- Identify possible synergies;
- Brainstorm ideas and challenges;
- Answer your questions;