Goals of the build-up phase of the EOSC Federation

Users

By **Q4 202...**

... access t...

... search a... enhance...

... use an I...

... access I... from div...

... be **cont**... of the E...

Federation

By **Q4 202...**

... consist...

... have es... clear de...

... have de... organisa...

... have es... Access a...

... have es...

... have de... monitor...

Communities and stakeholders

By **Q4 2025** communities/stakeholders interested to contribute to the EOSC Federation will be provided with ...

... detailed **practical information** about the key aspects and components of the EOSC Federation, such the governance, operations, policies, procedures, technical specifications and standards (included in the EOSC Federation Handbook)

... **guidance** on the procedures for enrolment of Nodes and onboarding of resources, as well as the estimated costs and effort

... demonstration of multi-node end-to-end **use cases** that showcase the added-value of joining the EOSC Federation

... continuous opportunities to be **consulted** in the stages of development of the build-up phase

Identification of potential candidate Nodes
Implementation of sequencing criteria

Interest to become an EOSC Node

✔ Respondent expresses intention to build an EOSC Node and can identify legal entity

Maturity and **diversity of resources** that can be offered through the EOSC Federation

✔ Respondent can offer a diverse set of resources/services proven

✔ Services / resources offered are in mature state, priority to those with track record of operating them in production grade environment

✔ Respondent is referred by other respondents as representative candidate Node to onboard resources

Inclusivity and **representation** of research communities and different types of organisations

✔ Inclusion of organisations from the thematic domains of the **five science clusters,** as well as of **national scope**

✔ Inclusion of **RPOs, service providers, research infrastructures, national organisations and institutions**

✔ Broad **geographic representation**

✔ Potential for **coordination** amongst groups of respondents

✔ Take account of participation in **EOSC EU Node consortium**

2-stage dialogues with potential candidate Nodes

1st stage – confirm commitment to join build-up phase

- Objectives
  - Present key **goals/deliverables** of the build-up phase & **timeline** for delivery
  - Discuss the **expectations of candidates** in joining the Federation
  - Clarify and confirm expected levels of **commitment** and required resources
  - Encourage community coordination to identify **representative organisations** to take part in the build-up phase
  - **Identify** the potential candidate Nodes to participate in the 2nd stage
- Participants
  - all potential candidate Nodes; the Tripartite Group
- Timing
  - 4 meetings; **December 2024 to January 2025**

2-stage dialogues with potential candidate Nodes

2nd stage – technical and organisational deep-dives (one-on-one)

- Objectives
  - Clarify the **organisational capacities** to achieve the goals of the build-up phase
  - Clarify the **technical requirements** to achieve the goals of the build-up phase
  - Discuss the **technical challenges the candidates aim to tackle** by joining the Federation
  - Identify a list of around **7-10 candidate Nodes** to participate in the first wave of enrolment
- Participants
  - potential candidate Nodes (reduced list); EOSC EU Node, EOSC Tripartite Group
- Timing
  - Around 4-8 meetings*, each w/ 2-5 organisations
  - **January 2025 to early February 2025**

* Will be determined by the number of organisations continuing to the 2nd stage.

2-stage dialogues with potential candidate Nodes
1st stage: 4 meetings during December - January

Dialogue #1
*
BBMRI ERIC
Elixir Hub
EMBL
Euro-BioImaging
ERIC
Instruct ERIC

Dialogue #2
**
ACTRIS ERIC
CNR (Blue-Cloud)
ESRF (PANOSC)
CNRS-LAPP
(ESCAPE)
CERN
ARIADNE RI
CLARIN ERIC

Dialogue #3
*
DE   **NFDI**
FR   **CNRS**
(Data Terra)
HU   **HUN**
REN
IT

Foundation ICSC
LU   **Lux**
NDS
PL   **NCN**
SK   **CVTI**
SR
UA   **BITP**

Dialogue #4
*
AT   **ACOnet / EOSC**
Sup. Of.
DK   **DeiC**
FI   **CSC – IT Center**
for Science
HR   **SRCE**
NL   **SURF**
NO   **NRIS**
SE   **SND**
SI   **ARNES**
              EUDAT

* Max 2 to continue to the 2nd stage
** Max 7 to continue to the 2nd stage

EOSC EU Node officially launched on 22 October

As the first building block in the EOSC Federation, the EOSC EU Node makes a clear contribution to the adoption of Open Science principles across Europe.

Services offered by EOSC EU Node

- **Bulk Data Transfer:** Move data effortlessly to data-intensive execution environments
- **Large File Transfer:** Streamline large file transfers online with added security and integrity
- **Virtual Machines:** Design and conduct experiments with flexibility while ensuring reproducibility
- **Cloud Container Platform:** Deploy cloud-native containerised applications that can easily scale
- **Interactive Notebooks:** Create and share documents with real-time code execution
- **File Sync & Share:** Enable automatic file syncing and secure sharing across locations and teams

EOSC Federation Handbook

## Progress update by the EOSC Association Board

Why a Handbook?

The "How To" manual for building the EOSC Federation

1 Most Deliverables from EOSC projects are not "How To's"

2 EOSC Federation and Node concepts (2023) not well defined

3 Goal to build the 1st Federation in 2025 is very ambitious

4 Need a clear manual (+ guides) on how to build the Federation for Nodes to commit to joining the Federation

Outline of EOSC Federation Handbook

- Chapter 1 – Purpose
- Chapter 2 – Governance
- Chapter 3 – Operations

- Chapter 4 – Architecture
- Chapter 5 – Scientific Resources
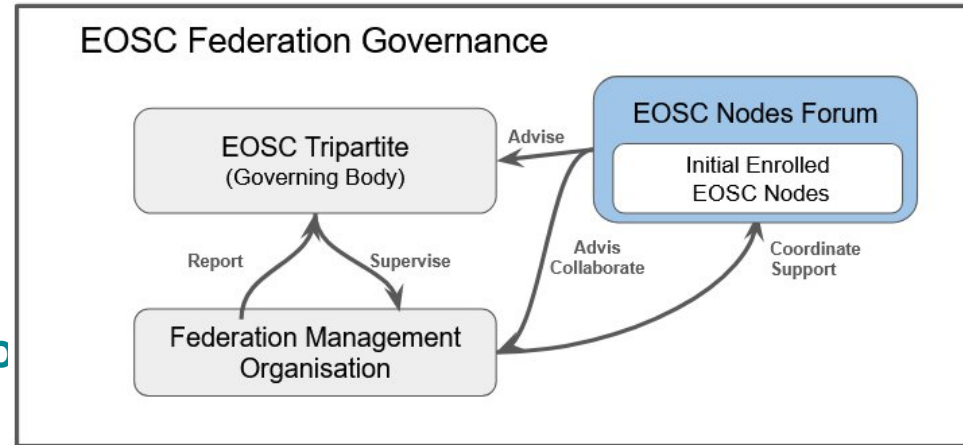- Chapter 6 – Policies

Purpose of 1st version of Handbook
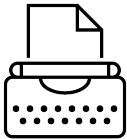
1. How to build the EOSC Federation during the period 2025 to 2027

2. What will the EOSC Federation enable and provide for researchers

3. How will the success of the EOSC Federation be measured

Chapter 2 - Governance

- Governance has been the trickiest chapter to write because:
  - Long-term view depends on newly formed Commission e.g. FP10 or not
  - Additional funding is coming from the European Commission
  - Resources are coming from the Nodes

- Clearly the **Tripartite Governance** is in charge of governance today and determines the roadmap and timeline

- However, the Federation needs a dedicated organisation to operate and manage it on a daily basis – **Federation Management Organisatio**

EOSC Federation Governance

EOSC Tripartite (Governing Body)

EOSC Nodes Forum

Initial Enrolled EOSC Nodes

Advise

Report   Supervise

Advis Collaborate

Coordinate Support

Federation Management Organisation

Chapter 3 - Operations

"Operations management (OM) is the administration of business practices to create the highest level of efficiency possible within an organisation." [Investopedia]

This chapter describes what is needed to operate the Federation :

- ○ **Day-to-day** operations
- ○ Regular **management** activities
- ○ How to **Apply to become a Node**
- ○ **EOSC Node roles** identified
  - ○ **Coordinator**
  - ○ **Operation Manager**
  - ○ **Technical Coordinator**
  - ○ **Security Officer**
  - ○ **Scientific Officer** (for Thematic Nodes)
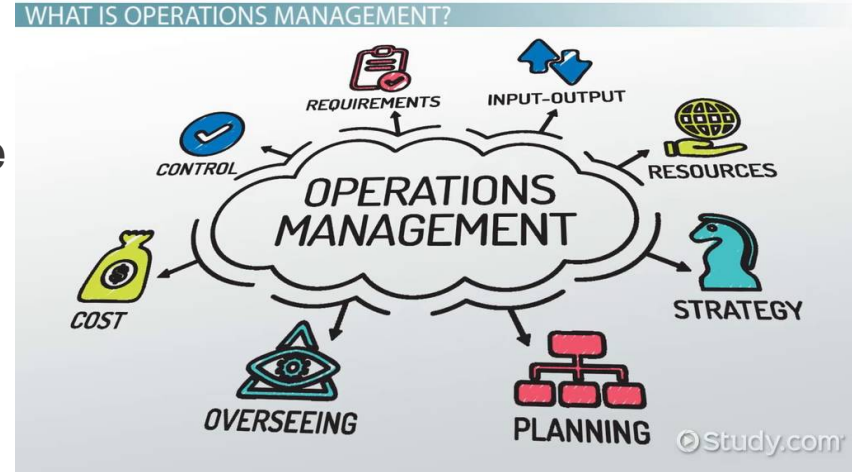
Chapter 4 - EOSC Federation Platform / Architecture

Feedback = all sections are relevant or highly relevant

Describing the EOSC Node architecture will cover:

- EOSC **Node Architecture**
- EOSC **Federation Interoperability Framework**
- EOSC **Federating Capabilities**
- EOSC **EU Node** Federating Capabilities and Exchange Services
- **Technical Operations**
- **Compute** and **Storage** Resources
- **Cybersecurity**
- Links to **External Entities**
- **Training**

18 Pages so far !

Interoperability

Chapter 5 - Scientific resources

Feedback = all sections are relevant or highly relevant

- EOSC is supposed to be the Web of Scientific FAIR Data

    → it should therefore federate [**eventually all**] Scientific FAIR Data in Europe

- The chapter will describe what Scientific Resources can be found in EOSC Nodes:
    o Data Repositories
    o Metadata standards
    o Training material for FAIR
    o Scientific productivity tools
    o Fostering domain specific communities

- It will describe how to connect them and make them visible in the EOSC Federation

- Challenge: how to attract trustworthy scientific resources of each domain

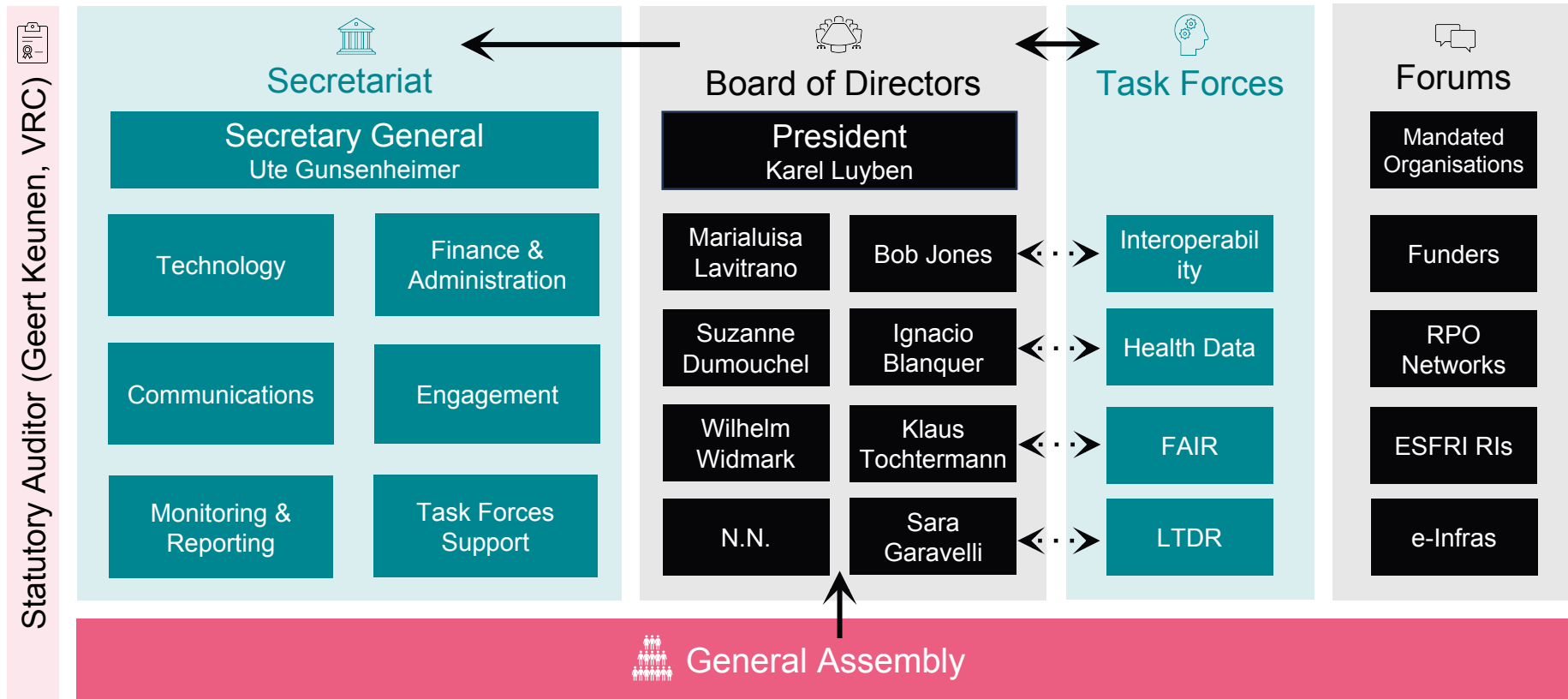2.5 Pages so far

Chapter 6 - Policies

Feedback = all sections are relevant or highly relevant

- Rules of Participation for Nodes
  - Must be Open, FAIR, follow standards e.g. cybersecurity
  - Includes requirements to help sustain EOSC Federation

- Access Policy
  - Defines who can access EOSC Resources

- Acceptable Usage Policy

  - Ethical, Legal and Social Implications (ELSI) Policy
  - Cybersecurity, Privacy and Data Protection Policy
  - Intellectual Property (IP) Policy
  - Code of Conduct

11 Pages so far …

La communauté impliquée
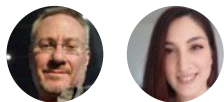
EOSC Association: Organisational chart

Statutory Auditor (Geert Keunen, VRC)

Secretariat

Secretary General
Ute Gunsenheimer

Technology	Finance & Administration
Communications	Engagement
Monitoring & Reporting	Task Forces Support

Board of Directors

President
Karel Luyben

Marialuisa Lavitrano	Bob Jones
Suzanne Dumouchel	Ignacio Blanquer
Wilhelm Widmark	Klaus Tochtermann
N.N.	Sara Garavelli

Task Forces

- Interoperability
- Health Data
- FAIR
- LTDR

Forums

- Mandated Organisations
- Funders
- RPO Networks
- ESFRI RIs
- e-Infras

General Assembly

EOSC Association Task Forces

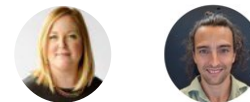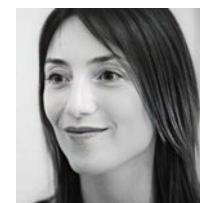|	FAIR Metrics and Digital Objects	Health Data	EOSC Technical and Semantic Interoperability	Long-Term Data Retention
Co-chairs	Elli Papadopoulou, Athena RC; and Mark Wilkinson, UPM	Lene Krøl Andersen, Computerome, DTU; and Petr Holub, BBMRI-ERIC & Masaryk University	Diego Scardaci, EGI Foundation; Jiří Marek, Masaryk University; Christos Kanellopoulos, GÉANT	Jenny O'Neill, HEAnet; Jacques Flores, Utrecht University
Members	51	44	60	37
Board liason	Klaus Tochtermann	Ignacio Blanquer	Bob Jones	Sara Garavelli
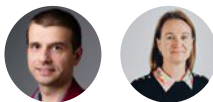
EOSC Opportunity Area Expert Groups

OA Expert Groups are an **important mechanism for collaboration** on technical and related matters within the EOSC Partnership

All OA Expert Groups have developed work plans and **meet roughly every four weeks**

EOSC Symposium 2024 featured a **panel discussion** dedicated to the collaboration within OA Expert Groups

Winter School 2024, Thessaloniki

OA Expert Group: Open Scholarly Communication

OA Expert Group: User and Resource Environments

OA Expert Group: Research Software

OA Expert Group: Persistent Identifiers

OA Expert Group: Metadata, Ontologies and Interoperability

OA Expert Group: FAIR Assessment and Alignment

OA Expert Group: Skills, Training, Rewards, Recognition and Upscaling

Le rôle des projets EOSC dans la fédération

- Développer des services qui seront intégrés dans de futurs noeuds
- Préparer l'engagement des communautés dans EOSC
- Contribuer à l'EOSC federation handbook
- Contribuer aux OAs pour favoriser la transversalité
- Répondre aux enquêtes de l'Asso EOSC

Quelques exemples de projets qui façonnent le paysage EOSC

Skills4EOSC ou comment *mettre en pratique la science ouverte*

Skills4EOSC

- Compétences pour EOSC : créer un **écosystème de formation homogène** pour la **science ouverte et FAIR** à l'échelle européenne

- **36 mois** : du 01/09/2022 au 31/08/2025 (8 WP)

44 **participants**, 18 pays

"**Acteurs clés**" en science ouverte dans leur pays/région/domaine

2 infrastructures de recherche ESFRI

7 millions €, co-financé par l'Union Européenne et le UK Research and Innovation

CNRS DDOR & INIST
IPSL
OPERAS
AMU

Objectifs du projet

Faire progresser les compétences en Science ouverte en **unifiant** le paysage actuel de la formation en un **écosystème paneuropéen commun et fiable**.

Combler **3 grandes lacunes** :

Manque d'expertise en science ouverte et en science des données	**Créer une expertise reconnue**
Manque d'une définition claire de profils de métiers liés aux données + de parcours professionnels correspondants	**Création de profils métiers (MVS) et de parcours professionnels correspondants**
Fragmentation dans les ressources de formation	**Capitalisation et mutualisation des ressources (matériaux FAIR)**

Harmoniser les formations (WP2)

- Respecter une **méthodologie FAIR-by-Design** (T2.3) dans la création des formations

La méthodologie a été testée par **DoRANum** sur **deux ressources**

- Création de **profils métiers** (T2.1)

Publication du catalogue en juin 2023 ✓

- **Certification des compétences** par des **Open Badges** (T2.5)

Publication de la proposition des Open Badges comme certifications aux formations de formateurs (décembre 2023)

Proposer des formations (WP4)

- **Des parcours de formation** (T4.2)
  → basés sur les MVS (profils métier) correspondants.

Data Stewards : formation pilote dispensée (module sur les licences, *data and code*) le 19 juin 2024

Data Librarians : formation pilote prévue le 17 janvier 2025 (module sur la reproductibilité)

ELSI (professionnels éthiques, légaux, sociaux) : lancement de la création du module en février 2024

Juin 2025 : **Publication de l'ensemble des parcours de formation**

- **Plan de formation pour les Data Stewards** (T4.1)

Basé sur le MVS (profil métier) correspondant

Développement de **9 modules :**

1. Foundation	6. Ethics
2. Research Data Management	7. Personal Data and GDPR
3. Research Software Management	8. Teaching and Education
4. Policy and Governance	9. Management and Soft Skills
5. Usage Rights and Licences	

WORK IN PROGRESS

Les Centres de Compétences

Réseau de coordination
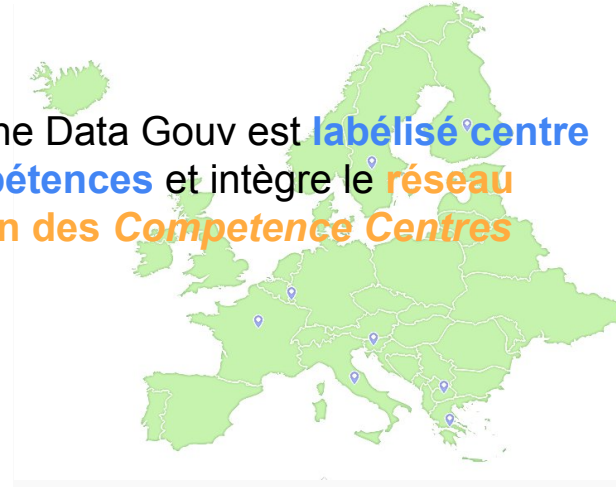
Harmonisation

Centre de compétence

Diversité du niveau local

- **Former les formateurs** et gagner en expertise

- Bénéficier d'un **large panel de services, de formations et de ressources** de formation

- Travailler sur des problématiques partagées et **identifier des solutions**

- Recherche Data Gouv est **labélisé centre de compétences** et intègre le **réseau européen des *Competence Centres***

Le projet FAIR-IMPACT

Jeudi 16 mai 2024 - 15h30 / 17h
Salomé Landel (CNRS-DDOR)

CNRS

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

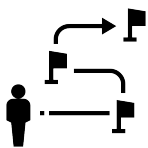Funded by
the European Union

Fil de la présentation

1. Présentation du projet FAIR-IMPACT
2. Participation des équipes françaises dans le projet
3. Les effets du projet : standards et outils
4. Contacts et liens utiles

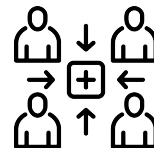# 1. FAIR-IMPACT : programme de soutien pour le FAIR dans l'EOSC

1. **Engager** les parties-prenantes de l'EOSC

4. **Définir** la gouvernance et la coordination pour garantir des données FAIR dans l'EOSC

eosc

2. **Rassembler** les solutions pour le FAIR : bonnes pratiques, règlements, outils, etc.

3. **Adapter** les solutions déjà développées et encourager leur mise en oeuvre (guide, cadre de travail)
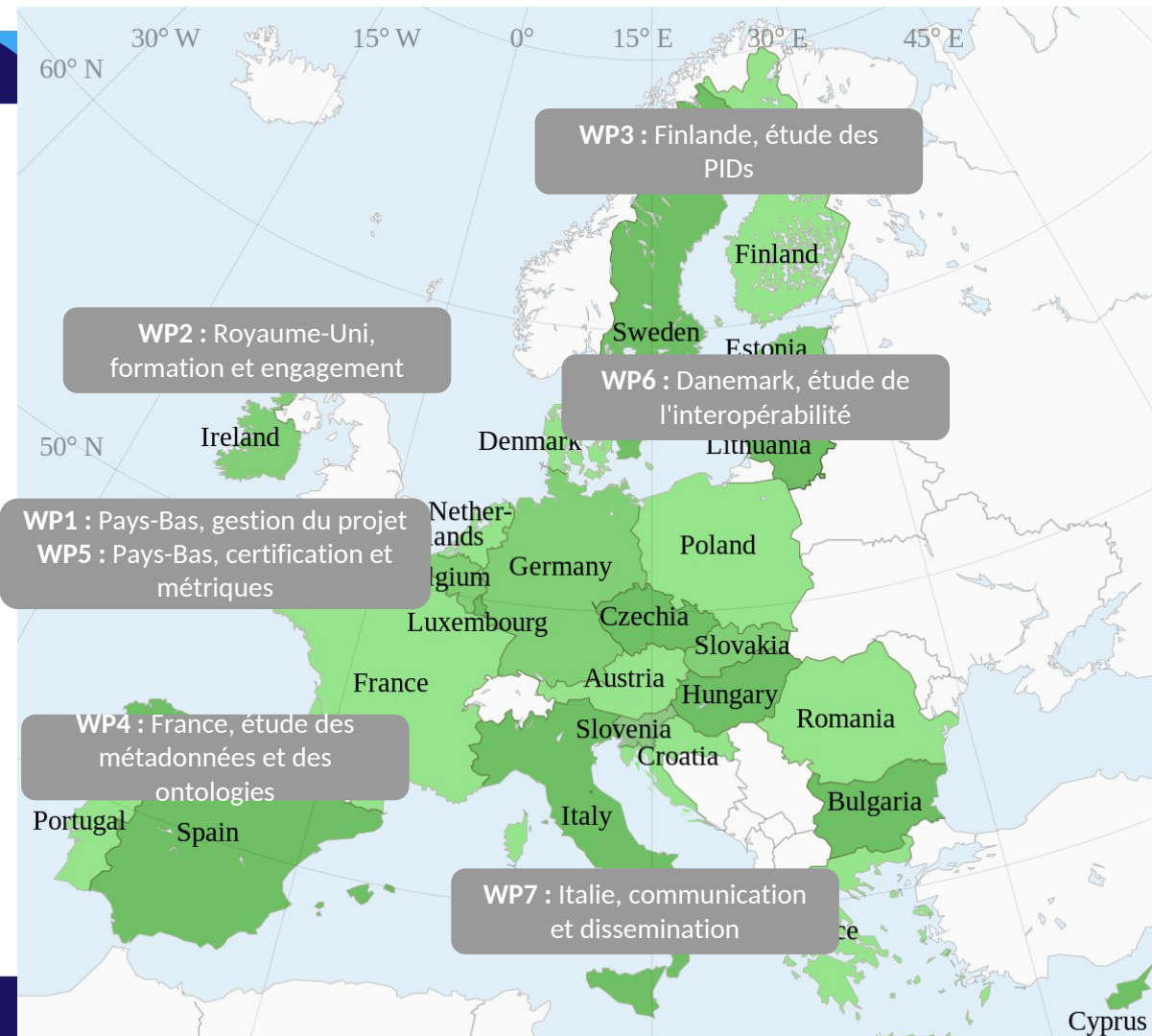
2. FAIR-IMPACT : un projet européen

- Budget : 10 millions d'euros
- Durée : 3 ans
        1er Juin 2022 - 31 Mai 2025
- 28 partenaires
- Coordinateur: KNAW-DANS (NL)

CSA : "*Coordination and Support Action***"**

Objectifs: coordonner et soutenir la mise en oeuvre de bonnes pratiques

⇒Capitalisation

WP3 : Finlande, étude des PIDs

WP2 : Royaume-Uni, formation et engagement

WP6 : Danemark, étude de l'interopérabilité

WP1 : Pays-Bas, gestion du projet
WP5 : Pays-Bas, certification et métriques

WP4 : France, étude des métadonnées et des ontologies

WP7 : Italie, communication et dissemination

3. FAIR-IMPACT : notre participation dans le projet

WP2

CNRS et INRAE engagés pour **promouvoir et former au FAIR**

- Ateliers de travail et mentorat (7 mois) pour des entrepôts de données, des organisations de recherche et des initiatives nationales

- Subventions après évaluation (4000 - 10 000 EUR)

WP4

INRAE, CNRS, AMU, INRIA et l'Observatoire de Paris participent à la **"FAIRisation"des métadonnées des jeux de données**

- Recommandations pour les métadonnées des logiciels
- Permettre l'interopérabilité sémantique entre les disciplines
- Méthodologie de cartographie des vocabulaires
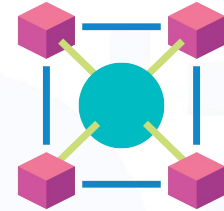
WP6

CNRS impliqué dans la **recommandation de standards pour l'interopérabilité**

- Modèles de SLAs et de MoUs intégrant l'interopérabilité
- Traduction du cadre de l'interopérabilité de l'EOSC en actions concrètes
- Sélection d'un schéma de métadonnées qui permet l'interopérabilité (DCAT)

4. FAIR-IMPACT, quels effets ?
Recommandation de standards, formations et subventions

Soutien pour l'utilisation de PIDs (Persistent identifiers) = ORCID, DOI, etc.

Sélection de vocabulaires contrôlés pour décrire les métadonnées

Solutions pour l'interopérabilité prêtes à l'emploi

Université de Lille

Campus Condorcet

Université Paris 8

Université de Reims Champagne-Ardenne

Normandie Université

Université de Strasbourg

Universités de Rennes

Université Paris Nanterre

Université Gustave Eiffel

Université de Lorraine

Nantes Université

Université Paris-Saclay

Centre-Val de Loire

Université Bourgogne Franche-Comté

Université Clermont Auvergne

Université Grenoble Alpes

Aix-Marseille Université

Université de Toulouse

Université de Montpellier

Méthode pour rendre les métadonnées FAIR

Métriques d'évaluation sur la qualité d'un entrepôt de données

Programmes de soutien à destination des entrepôts de données et organisations de recherche

Source de la carte : site de Recherche Data Gouv

Contacts et liens utiles

Contacts DDOR:

Suzanne DUMOUCHEL :

suzanne.dumouchel@cnrs.fr

Olivier ROUCHON : olivier.rouchon@cnrs.fr

Salomé LANDEL : salome.landel@cnrs.fr

Site de FAIR IMPACT: https://fair-impact.eu/

⇒ **Informations sur les programmes de soutien** (financements et formations)

⇒ Accès aux livrables avec les **recommandations du projet**

Le projet LUMEN

- Janvier 2025-Dec 2027
- Coordonné par le CNRS-DDOR
- 19 partenaires
- Objectifs: construire des plateformes de découvertes des ressources pour différentes communautés scientifiques et intéropérables entre elles.
- S'appuie sur le projet TRIPLE (2019-2023), coordonnée par le CNRS, et qui a délivré la palteforme GoTriple.

Objectifs du projet LUMEN

SO1 - Create a shared and collaborative cross-disciplinary knowledge and data exchange space for the next generation EOSC

SO2- Enhance the qualitative discovery of research outputs by establishing an ecosystem of innovative features

SO3- Support the production of FAIR-by-design research outputs

SO4- Facilitate early stage of the research lifecycle by enhancing specific tools and services

LUMEN ARCHITECTURE OVERVIEW

Impacts principaux

The LUMEN project aims to revolutionise how researchers, public, and private sectors create, share, and utilise research outputs across four scientific domains: Maths, SSH, ES, and MD. By providing interoperable platforms tailored to researchers' needs, LUMEN will enhance scientific discoveries, streamline research phases, and lower barriers to accessing advanced tools, including those powered by AI. It will foster interdisciplinary collaboration and ensure seamless access and management of research data following FAIR principles, thereby boosting innovation and trust in European scientific research.

SI1: EU Scientists have easy access to high quality and FAIR multi-disciplinary data fostering cross-disciplinary collaboration.

SI2: Enrich diversity of services for EOSC-Exchange

SI3: Enhancing research efficiency, accessibility to expertise, scalability, and cost-effectiveness, while also providing valuable insights and educational opportunities for users.

SI4: Revolutionise the research process by enhancing user experience, improving reproducibility and analytical capabilities, accelerating discovery, facilitating collaboration, and promoting openness and inclusivity in scientific endeavours.

SI5: Transform how organisations manage and leverage data, fostering collaboration, innovation, and agility while ensuring data quality, interoperability, and alignment with business objectives across a connected ecosystem of communities.

LUMEN Outcome 1: A white-label data platform to offer discovery services and interdisciplinary features to the communities. Utilising a ready-to-use data platform, communities can reduce the costs and time required to develop their own data management solutions, freeing up resources for other activities.

LUMEN Outcome 2: federation of platforms in 4 different scientific domains. By setting up dedicated platforms for scientific data that can be used either in isolation or together, this outcome fosters interdisciplinarity practices and supports interoperability on service level.

LUMEN Outcome 3: FAIR Semantic Management Space

LUMEN Outcome 4: AI-driven discovery assistant (AIDA). A chatbot assisting researchers to perform discovery, moving from keyword-based searches to detailed interactive conversations in natural language.

LUMEN Outcome 5: Multiplatform innovative tools & services for the discovery platforms. The multiplatform innovative features are innovative tools and services for research development and interdisciplinary collaboration with the objective of maximising the discovery of contents between communities: the Interdisciplinary Metasearch Service, the LUMEN Chatbot AIDA, the Metrics services, advanced visualisation tools, and an automatically built software tools catalogue delivered in white label discovery platform

LUMEN Outcome 6: Data mesh architectural framework. Builds a connected ecosystem of communities for data-driven insights and links communities and scientific fields by empowering them to have high autonomy and ownership of their data domain, designing new usages and improving data access, interoperability, scalability and integration with EOSC.

LUMEN Outcome 7: Federated data sharing governance. Defines global policies regarding how the domain teams would have to share their data and create a data domain ecosystem based on community rules.