Joining forces to define the European Health Data Sharing Landscape for Research

23 November 2022

In the next 10 minutes I'd like to introduce you:

* Science in transition - Open Science and EOSC are European priorities

* Challenges to systematically use health-related data for research

* What opportunities we have at the moment to address those challenges, what EOSC elements can be utilised.

Science in Transition

- The **Open Science** paradigm affects the whole research cycle and all its stakeholders

- It implies sharing knowledge and tools:

- "<u>as early as possible</u>" in the research process;

- "<u>as openly as possible</u>";

- "<u>as FAIR as possible</u>";

- not only within a discipline but also between disciplines and society at large.
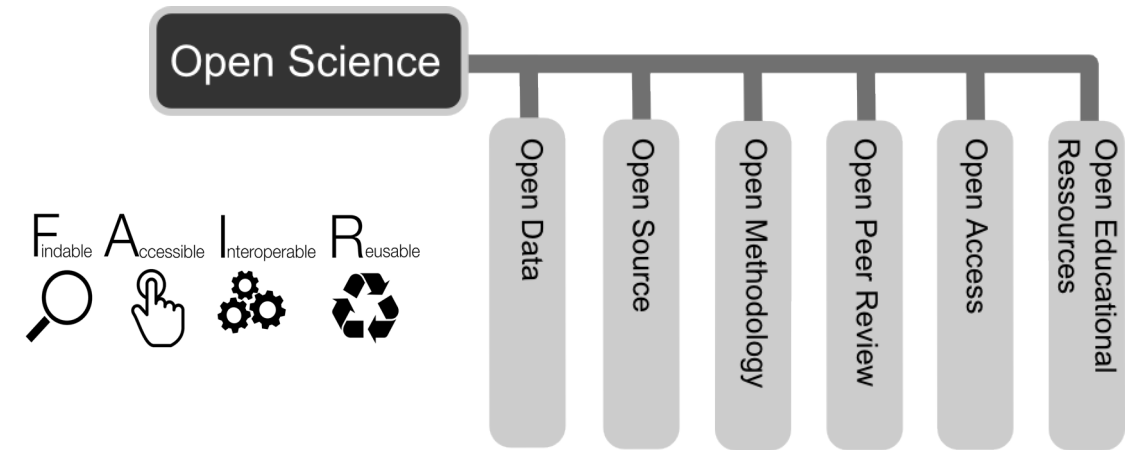
https://commons.wikimedia.org/wiki/File:Open_Science_-_Prinzipien.png

➢ Developing an enabling policy environment for open science
➢ Investing in open science infrastructures and services
➢ Investing in human resources, training, education, digital literacy and capacity building for open science
➢ Provide incentives

Open Science: a Political Priority of the Union

- 2016: Council Conclusions on the 'Transition Towards an Open Science System'
- 2018: EC Recommendation on 'Access to and Preservation of Scientific Information'
- 2020: EC Communication on the 'New ERA'
- 2021: Council Recommendation on a 'Pact for R&I in Europe'
- 2021: Council Conclusions on the 'Future Governance of the ERA' including the 'ERA Policy Agenda
- 2022: Council Conclusions on 'Research Assessment"

Member States' Commitment to 20 ERA Actions
Status: 13/07/2022

Priority area: 1 Open Science including EOSC
8 Research Infrastructures

Challenges to systematically use health-related data for research

Challenges to systematically use health-related data for research

- **Interoperability of health related data**: need of set of standards and guidelines to support interoperability

- **Application of the FAIR principles** balancing the protection of privacy and promoting the safe use of citizens' health data: a clear governance and rules for access and sharing of health data from medical records and genetic and genomic data for research purposes is needed.

- **Infrastructures**: need to develop technical computational infrastructures, certified and GDPR compatible

- **Security management:** Health data processed and stored in secure systems need specific (cyber) security management. Conditions for the secure access to health data for research purposes need to be defined: responsible access to health data is essential for maintaining citizens' trust.

- **Legal guidance** : harmonised legal basis for access and use of data and respect of data subjects' rights. GDPR implementation in Member States is still fragmented. A new legislation might lead to even more complexity. GDPR. It is necessary to complete the EU regulatory panorama to fill the gaps left in the GDPR concerning the use of health data.

- **Training:** for actual and the new generations of data managers, researchers and clinicians

- **Artificial Intelligence:** raises ethical concerns: discrimination of the sexes, socio-economic groups or minorities by being underrepresented in the data sets. Algorithms are trained and, if not scrutinized appropriately, they open the doors to new forms of discrimination and marginalization.

What opportunities we have at the moment to address those challenges

What EOSC elements can be utilised

What is EOSC?

Image: Martin Reisch https://unsplash.com/photos/6DivtP_WRYs

# "A web of scientific insight"

- **Web of FAIR Data and related Services**

- **Federation of relevant existing and future data sources**

- **Federation of e-Infra and Research Infrastructures**

- **Virtual space where science producers and consumers come together**

- **Environment in which data can be brought together with services to perform analyses and address societal challenges**
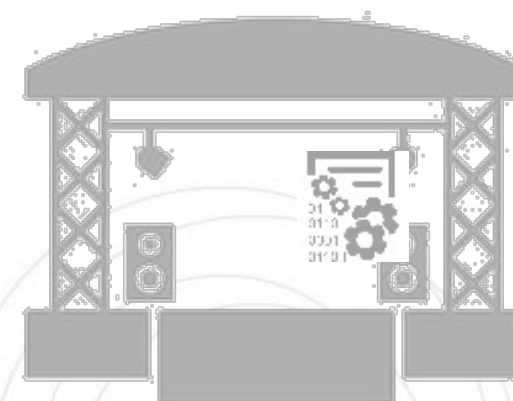
Align with European Partnerships & data spaces

The nine European Common Data Spaces: Hybrid Public and Private Data
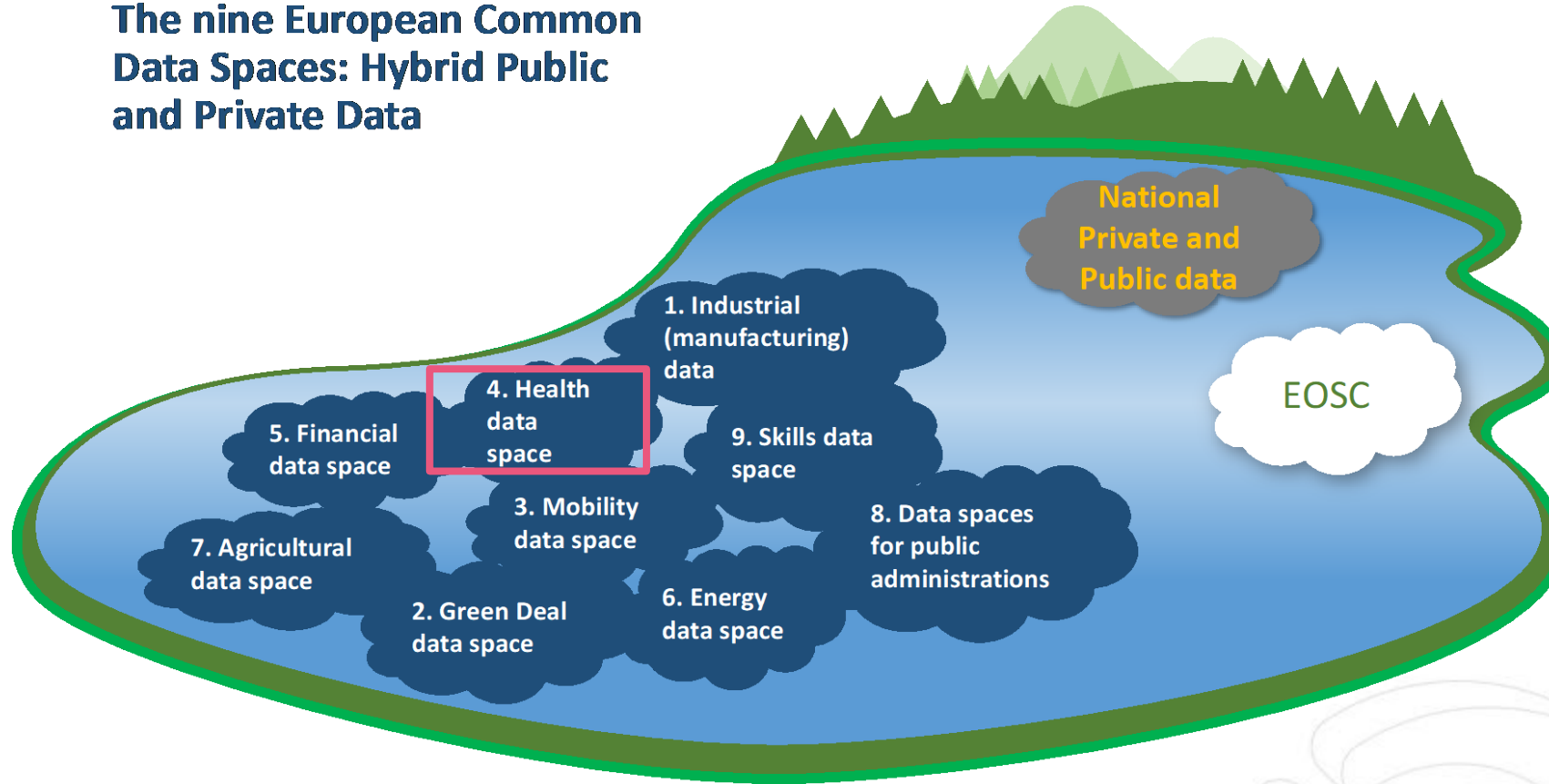
EOSC can be seen as a crosscutting data infrastructure focused on research and innovation that intersects all sectorial FAIR Data Spaces

The development of the EHDS should avoid that data guidelines and management of Health FAIR Data may grow apart with the risk of lessening the impact on research and overall European data strategy

eosc science

"Create better health outcomes for citizens/patients tomorrow than we have today"

The existence of the **EOSC** initiative should be pointed out in all 9 sectorial Data Space initiatives, in their regulations and communications, as a fundamental **source of synergies** with the overall European efforts towards a coherent overarching strategy, therefore also **avoiding parallel and unconnected development efforts**.

eosc

It is in the interest of the EU member states that investments at national and European level, in terms of direct financial effort, in-kind contributions and especially deployment of human resources, are well coordinated in order to maximize effectiveness and impact, as well as to avoid overspending.

European Health Data Space should develop synergiostically with the EOSC initiative, exploiting and strengthening the relevant existing public infrastructures and services.

EHDS should in general build on and make use of already existing infrastructures and FAIRdata Services
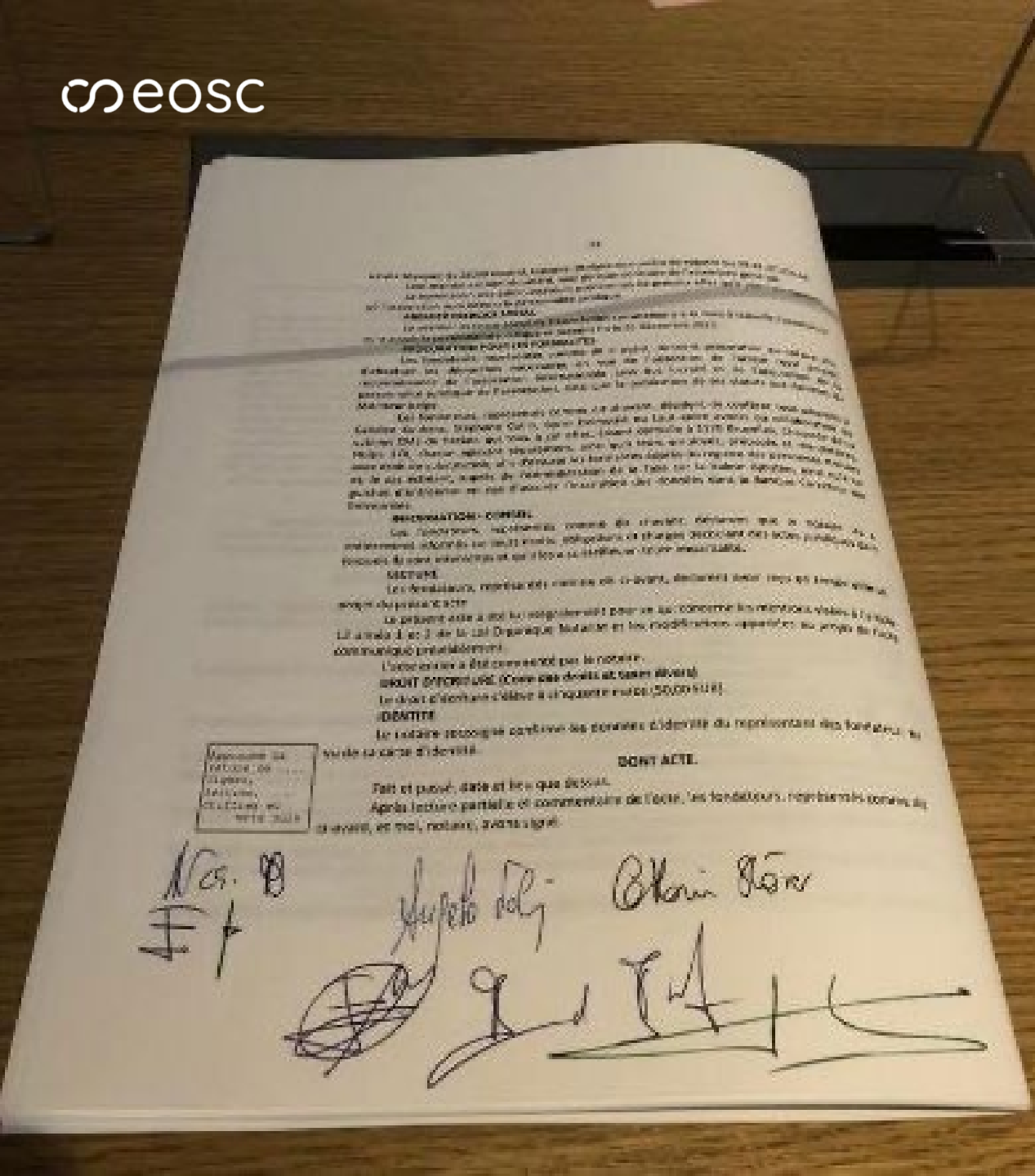
The **implementation of the new infrastructures for secondary uses of electronic health data**, will be done in coherence with the overall IT governance framework of the European Commission, specifically **building upon the expertise and technical solutions developed in the field of research and innovation**, therefore contributing to the strengthening of the EU effort for a **coherent FAIR data policy**

The EHDS Board shall cooperate with all other relevant bodies, entities and experts, including EOSC, in the effort of reaching **advanced solutions for the FAIR data usage** in research and innovation

Health data access bodies shall cooperate with the governance of EOSC to build upon the experience in building a European FAIR data space and to elaborate with the research efforts towards the **interoperability solutions in the framework of the EOSC**

The Purpose of EOSC-A

(1) to provide a single voice for the advocay and representation of the boarder EOSC stakeholder community,

(2) to promote the alignment of European Union research policy and priorities with activities coordinated by the Association;

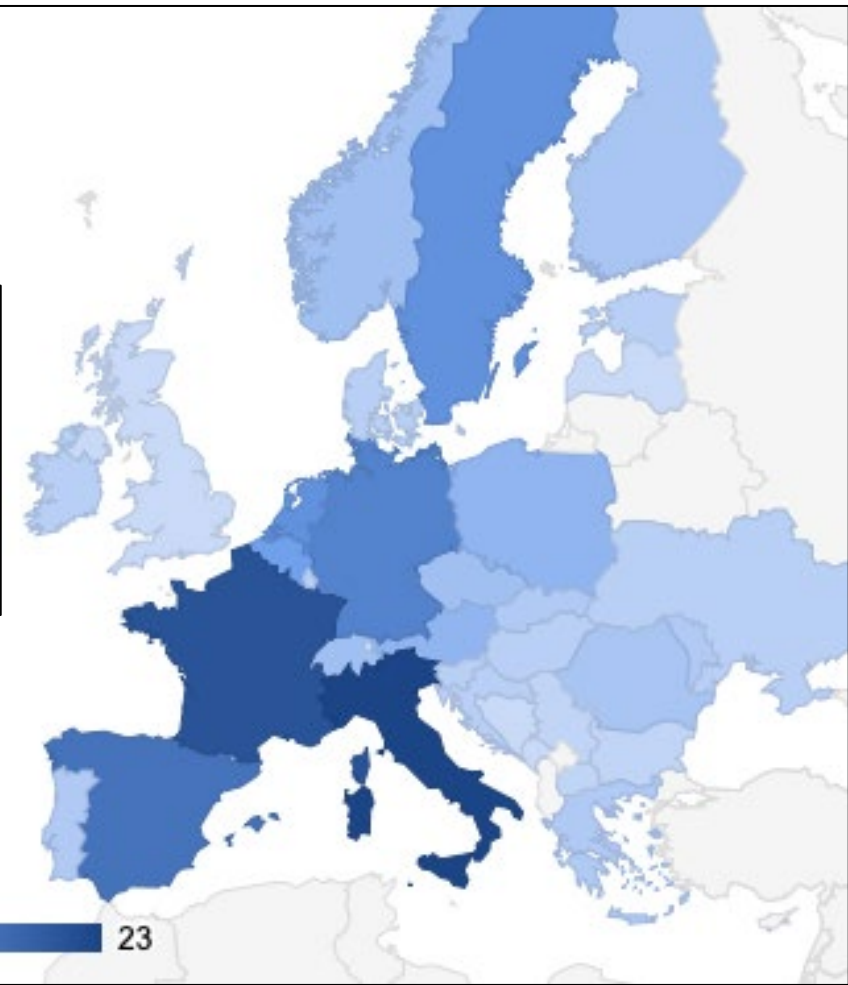(3) to enable seamless access to data through interoperable services that address the entire research data life cycle.

September 2022 by EOSC-A

The EOSC Association Today

27 Mandated Organisation - 133 Members - 78 Observers

A strong member base

EOSC-A Board

Karel Luyben President, 2021-22	Marialuisa Lavitrano Vice-President, 2021-23	Suzanne Dumouchel 2021-23	Klaus Tochtermann Treasurer, 2021-23	Sarah Jones 2021-22
Ignacio Blanquer 2021-22	Sara Garavelli 2022-24	Bob Jones 2022-24	Wilhelm Widmark 2022/24	

Other organisation 7,3%
Research funding org… 7,6%
Service providing orga… 34,3%
Research performing… 50,9%

EOSC Steering Board

23

17

EOSC-A Brain-Pool: 13 Task Forces

Over 400 volunteers

Implementation of EOSC

- Rules of Participation compliance monitoring
- PID policy and implementation
- Researcher engagement and adoption

Technical challenges on EOSC

- Technical interoperability of data and services
- Infrastructure for quality research software
- AAI Architecture

Metadata and data quality

- Semantic interoperability
- FAIR metrics and data quality

Research careers and curricula

- Data stewardship curricula and career paths
- Research careers, recognition and credit
- Upskilling countries to engage in EOSC

Sustaining EOSC

- Financial sustainability
- Long-term data preservation

Metadata and data quality

Semantic interoperability TF

- Foster **Knowledge Exchange** around Semantic Interoperability and **Provide Recommendations** to Produce a **refined version** of the **Interoperability Framework**.
- Explorations into **Semantic Interoperability**
  - Select/Recommend **common metadata standards** for a broad range of 'data'.
  - Suggest **next steps** in the development of **catalogues** for **metadata standards**.
  - Evolve **syntactic interoperability** for data formats, metadata schemas and services.
  - Support the **alignment** of **semantic artefacts**, both at the domain level and at the top level.
  - Evolve the **descriptions** of the **technical components** for **semantic interoperability** from the original EOSC IF.
  - Propose **strategies** for **long-term preservation** of semantic artefacts.

FAIR metrics and data quality TF

- **Implement** the proposed **FAIR metrics** for EOSC by assessing their **applicability** across research communities and **testing** a range of tools to enable uptake.
- Give **recommendations** will be made to **update metrics** and adopt tools as appropriate.
- Undertake a **state of the art** to generate mutual understanding about data **quality** and conduct several case studies to identify **common features** and dimensions to define a data quality approach for EOSC.

Research careers and curricula

Data stewardship curricula and career paths TF

- **Focus** on the **Data Steward** role and their **core activities**.
- Work on data stewardship job roles and curricula to ensure these are **recognised** and **aligned** across Europe.
- Develop **recommendations** for potential **career paths** taking into account appropriate **recognition** and rewards for data management activities.
- Actively **engage** with stakeholders and **build on previous work**.
- Ensure a **co-creation process** between **theoretical** development and **implementation** examples.

Upskilling countries to engage in EOSC TF

Research careers, recognition & credit TF

Expected results and workplan:

[MS] Identify Stakeholders and related initiatives (June)
[MS] Recommendations for Data Steward Curricula v1 (October)
[MS][D] Data Stewards roles mapping (Nov)
[MS] Use Cases Identification (Nov)

Long-term data preservation TF

Provide **recommendations** on the **vision** and **sustainable implementation** of **long-term data preservation** policies and practices, as well as suggestions to later strategy execution:

- Define the **functions necessary** to create, store, curate, preserve, and engage with data (re) users.
- **Map** and **promote** the roles, responsibilities and accountability required for preservation of data
- **Mapping** of the **financial aspects** of long-term data preservation, including business model aspects.
- Recommendations on the creation of **a European network** of **trustworthy FAIR digital repositories** to mature preservation and FAIR-enabling practices across disciplines and geographies.

Financial sustainability TF

Expected results and workplan:

[MS] Shared vision and understanding on LTDP agreed upon by LTDP TF (Oct)

[D] Vision and policy recommendations on LTDP v1 (Oct)

[D] Vision and recommendations on roles, responsibilities and accountability in relation to LTDP (Oct)

[MS] Roles, responsibilities and accountability agreed upon by LTDP TF & related task forces (Oct)

[D] Vision and recommendations on financial aspects in relation to LTDP (Oct)

[MS] Recommendations concerning financial aspects of LTDP agreed upon by LTDP TF & related TFs (Oct)

[D] Vision and recommendations on a TDR network in relation to LTDP (Oct)

[MS] Recommendations for EOSC-A concerning strategy and support actions for a TDR network agreed upon by LTDP TF & related TFs (Oct).

Technical challenges on EOSC

Technical interoperability of data and services TF

- A **first principles** discussion on the guiding principles for Interoperability.
- A **landscape overview** and analysis of the existing systems and **interoperability standards** for data and services.
- Promote **alignment between EOSC standards** and other **major activities** (such as RDA, EuroHPC and GAIA-X).
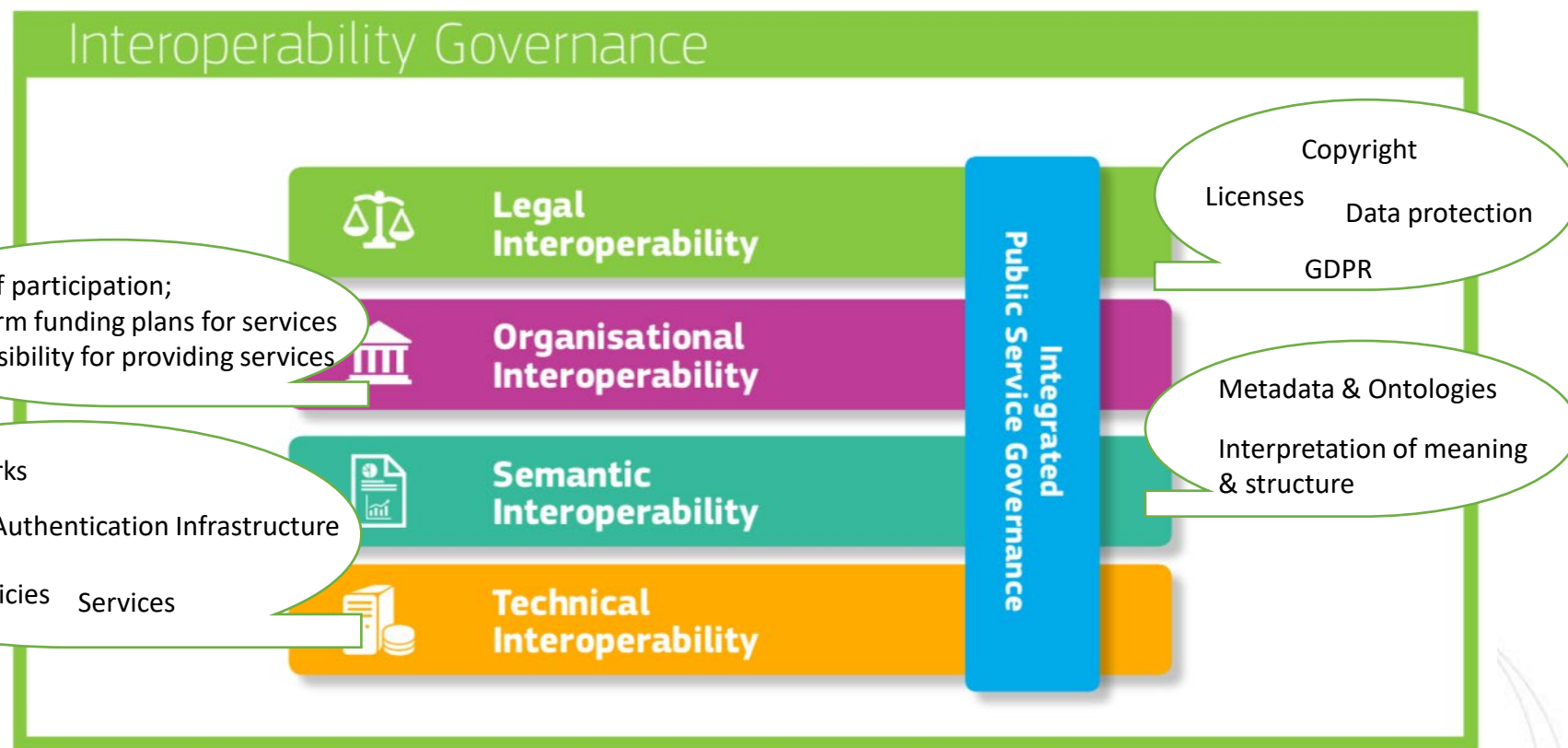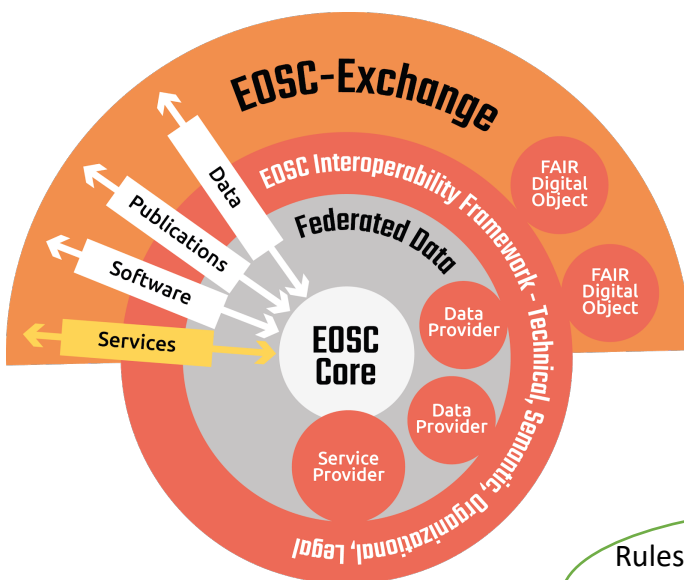- A **technical architecture** discussion for the **EIF**.

Infrastructure for quality research software TF

AAI Architecture TF

Expected results/workplan:
- [D] First Principles Document (In progress)
- [D] Landscape overview of the EIF v2022 (Oct)
- [D] Draft Technical Architecture (Dec)

EOSC Interoperability Framework

The European Interoperability Framework four levels of interoperability

Oscar Corcho - EOSC Interoperability Framework

eosc

Research is at the basis of an efficient healthcare, and, in turn, healthcare can offer challenging paths to research.

Division between primary use and secondary use of health data cannot be strictly upheld in practice

Healthcare systems, should be set up, from the start, in a compatible way with the needs of research.

Indeed, health data and digital workflows are at the intersection of primary care and research and need to be conceptualized from the beginning as complementary to reduce errors, avoid duplication of procedures and allow for innovative digital health solutions.

AlmaHealthDB
FAIR-by design

- The AlmaHealthDB (AHDB) infrastructure will allow researchers from UNIBO and the three research hospitals in Bologna to collect and access large volumes of data in compliance with legal, organizational, and regulatory requirements.

- Input data sources undergo an assessment process to establish their origin, permitted uses, presence of valid consents and authorizations, and conditions for processing. Outcomes of this assessment process are coded into a set of standardized metadata.

- AHDB will support data FAIRness with dedicated services for data curation, data quality management, analysis of legal requirements, standardization of data and metadata.
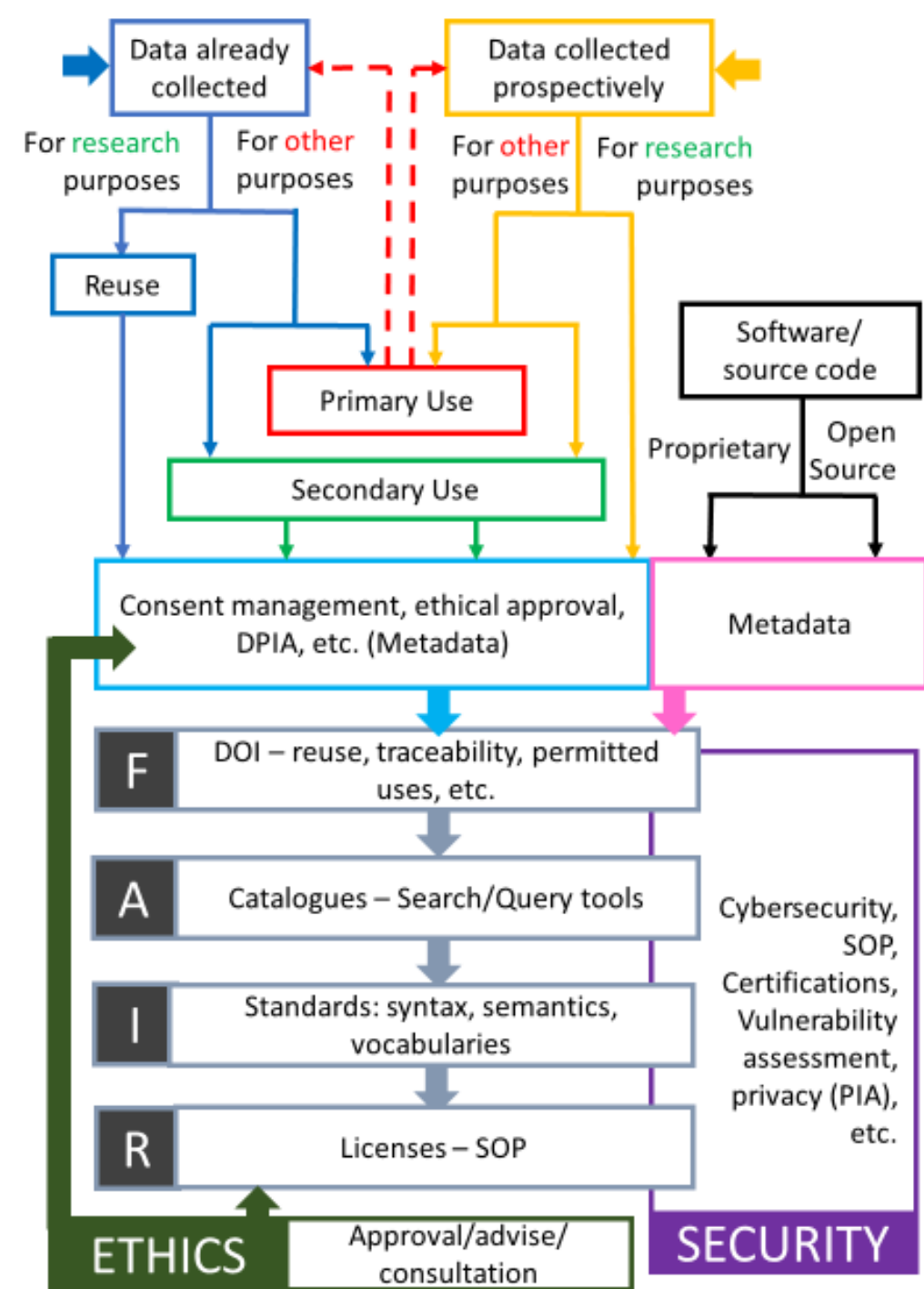
EOSC Association AISBL

Rue du Luxembourg 3
BE-1000 Brussels, Belgium
+32 2 537 73 18
info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931
VAT number: BE0755 723 931

Slides made by EOSC-A Board and Secretariat
Marialuisa Lavitrano, EOSC-A Director and vice-President
marialuisa.lavitrano@unimib.it