EOSC & DG-SANTE Round Table on EHDS and EOSC

8 October 2024, 9:30-17:10

The objective of the event was to understand the complementarities, overlaps and gaps of EHDS and EOSC in the secondary use of health data.

Agenda

9:30 Welcome and introduction from partners

9:50 Setting the scene

- The EHDS. Brief presentation of the concept, principles and actors (Jerome Barros / Irini Kessissoglou - DG SANTE).
- The EOSC Federation. Concept of EOSC, concept of nodes, main actors in EOSC, role of existing RIs in Health (Karel Luyben /Ignacio Blanquer - EOSC-A BoD).

10:30 Discussion of four topics

10:30 Discoverability

11:45 Permission management

14:00 Accessing and processing health data

15:00 Reusability of data and outputs

16:00 Summary of discussions and actions to be taken

16:35 Brief summary per topic by the rapporteur

17:00 Conclusions from the chairs of the session

17:10 *End of meeting*

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

Welcome and introduction

The event was attended by 27 people (2 of them remotely), representing key actors in the area of EOSC, EHDS and related projects and initiatives. The attendees came from six different stakeholder groups:

EOSC Association membership

- Niels Bolding, HealthRI
- Lene Krøl Andersen, DTU, Co-Chair of the EOSC-A Health Data Task Force
- Petr Holub, BBMRI-ERIC, Co-Chair of the EOSC-A Health Data Task Force
- Maddalena Fratelli, EATRIS-ERIC
- Jonathan Tedds, ELIXIR
- Jacques Demotes ECRIN-ERIC
- Petra Ritter, EBRAINS

EOSC Association Board of Directors

- Karel Luyben, President EOSC-A
- Ignacio Blanquer, Director EOSC-A
- Marialuisa Lavitrano, Director EOSC-A
- Ute Gunsenheimer, EOSC-A Secretary General

EHDS-related Projects or initiatives of interest

- Manuel de la Guia Solaz, IMI Joint Undertaking
- Enrique Bernal-Delgado, IACS
- Juan González-García, IACS
- Luis Marti-Bonmatí, Hospital LA Fe
- Salvador Capella, BSC (online)
- Pascal Derycke, Sciencsano

HDABs-Community of Practice

- Tamara Buble Croatian Institute of Public Health, Chair of the Steering Board
- Fidelia Cascini, UCSC, Chair of the Stakeholders engagement subgroup

European Commission

- Licinio Kustra Mano, DG SANTÉ (online)
- Jerome de Barros, DG SANTÉ
- Irini Kessissoglou, DG SANTÉ
- Aleksandra WESOLOWSKA, DG CNECT
- Peter SZEGEDI, DG CNECT
- Angelo Solimi, DG RTD
- Bertil Egger-Beck, DG-RTD
- Christina Kyriakopoulou, DG RTD

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

Setting the scene

The presentations about the EHDS and the EOSC Federation helped to reach a common background from both the EHDS and EOSC. In general:

- EOSC and EHDS concur on activities related to the secondary use of health data for research.
- Both EOSC and EHDS focus on a federation of data and service providers, although EHDS is a common European regulatory framework that defines the provision and access to data and services (EHDS has a harmonised European legal framework agreed by the countries and EOSC does not). In both frameworks, Member States play an important role.
- EOSC federation model is structured into 3 tiers: 1) Services fully integrated in the EOSC EU node; 2) Onboarded services from third-party providers, operating at the same quality level; and 3) Discoverable services from third parties, with a lower entry barrier.
- EOSC focuses on the management of both open and restricted data, and EHDS focuses strongly on the access to sensitive personal data for secondary use.
- EOSC is multidisciplinary and heterogeneous by nature and EHDS is firmly based on health thematic. However, EOSC has a strong base of research performing institutions in health (EOSC-A members include 8 Health-related RIs, 12 health research institutes and more than 25 research performing institutions including research on health).
- The implementation force of EOSC is mainly through projects (and the procurement of EOSC managed services), with the strategy guidance supported by the EOSC-A Task Forces and other expert groups. On the other hand, as a European legislation, the EHDS construction is compulsory to all the Member States as well as EEA States and implementing acts will define specific components. Sustainability and cost analysis has been provided in the regulation proposal.
- EOSC is focusing on the implementation of the EOSC federation starting in 2025.
- EHDS has a hierarchical country-based design approach and EOSC is mainly bottom-up.
- There are several actors in the EHDS ecosystem that have similar roles to EOSC actors (especially important is the potential role of RIs as Authorised Participants, but also the role of Data Intermediation Entities and single Trusted Data Holders).
- The EOSC EU node is a multi-disciplinary node commissioned by the EC which will serve as a blueprint for other EOSC nodes, defining an open-source based software architecture that provides federating capabilities to other EOSC nodes, resources and services.
- The implementation path of EOSC and EHDS will overlap, although EOSC has already committed an important amount of resources on building the community and

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

implementing services to spark the implementation process of EOSC. EOSC nodes are expected by 2025. Implementation of the EHDS will be through a staged approach. Assuming that the EHDS Regulation is published in the Official Journal of the European Union in early 2025, regarding the secondary use of electronic health data, Member States shall comply with the EHDS by early 2029 (access to some health data type such as genomic will be further delayed). For complying with EHDS, Member states have started establishing or nominating their national Health Data Access Bodies for the HealthData@EU infrastructure of the EHDS for secondary use and EOSC has started building the EOSC federation through the EOSC EU node and a call for EOSC nodes.

- Under the EHDS Regulation, preparation of data and services such as SPE resources can require a cost in which the user may be expected to contribute. Services in EOSC (such as computing and storage) may require agreements and quotas, although there is no defined path for direct monetary support from the users.
- The EHDS Data User Journey resonates quite well with the equivalent in EOSC when dealing with sensitive personal data.

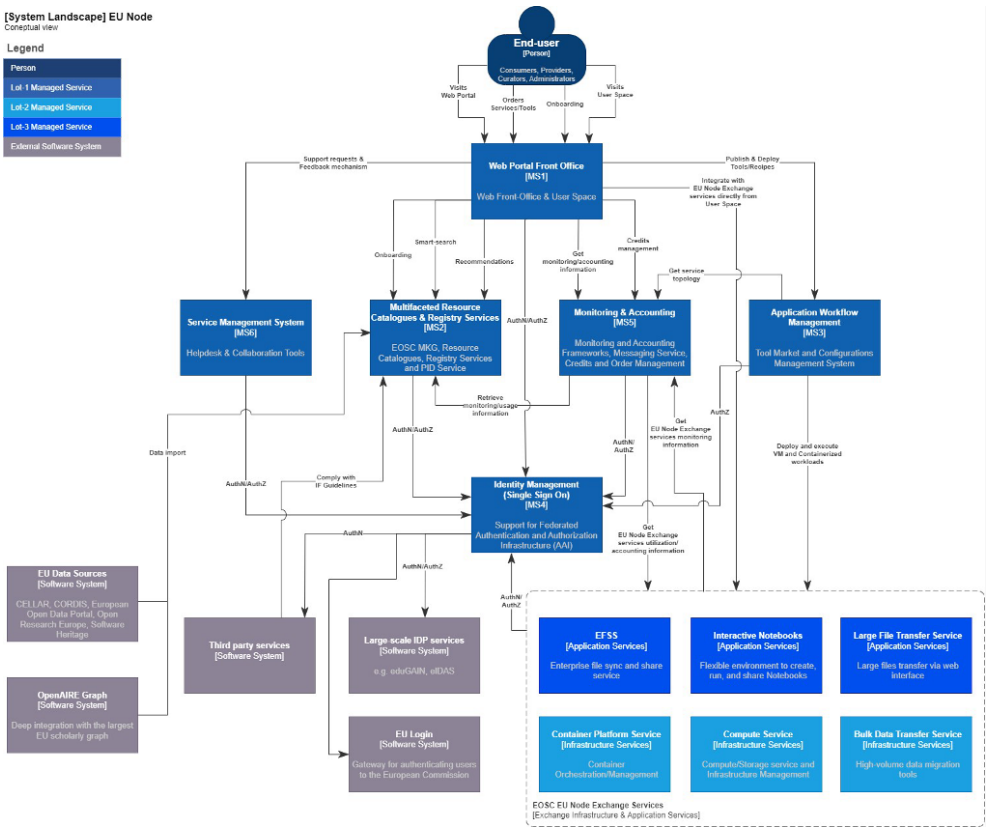

Figure 1: EOSC EU Node architecture

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

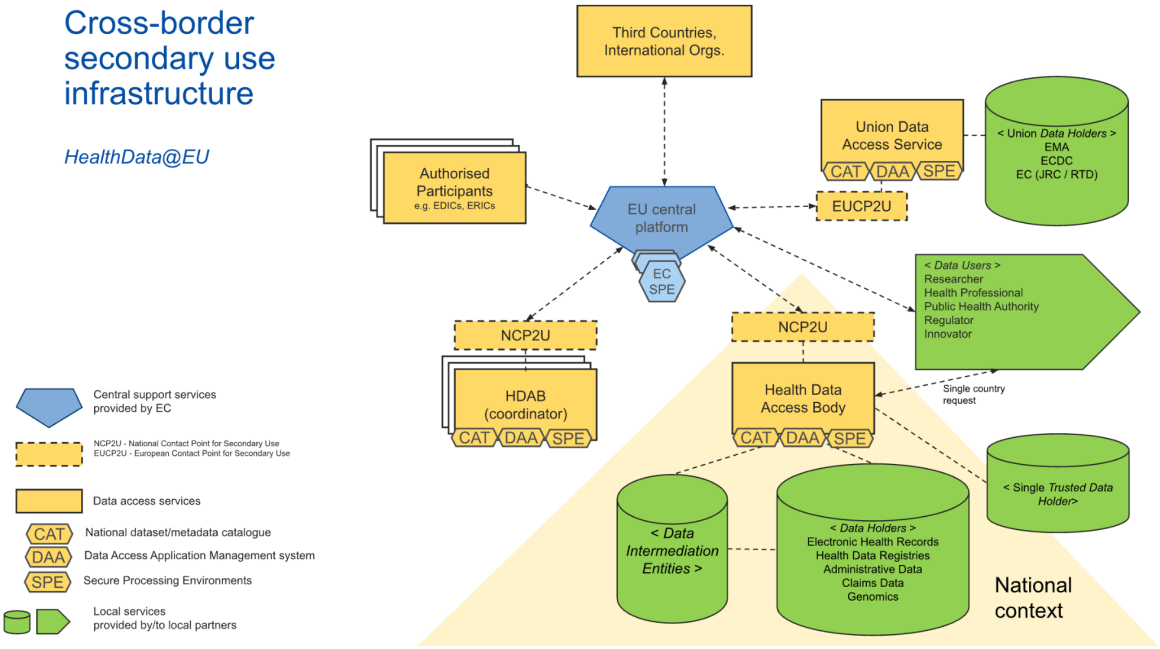

Figure 2: Schema of the HealthData@EU EHDS infrastructure for secondary use of health data.

Discussion

Four topics were proposed to the group for discussion. Each one of the topics was introduced by EOSC and the EHDS and the discussion was conducted by a moderator and documented by a rapporteur. The discussions on the topics are summarised in the following subsections.

Discoverability

This topic focused on How EOSC and EHDS are considering the findability of the data through catalogues, registries, identifiers, metadata schemas, requests for observational studies, networks of data holders, etc. Several key points in the discussion were:

- Both EOSC and EHDS deal with the findability and accessibility to health data, including the cataloguing of metadata records of datasets. EOSC data catalogues mainly focus on data that is already collected and curated for research and EHDS, in addition to existing curated data, it will deal with data extracted and prepared on request from the data holders.
- EHDS focuses on describing and publishing the data, creating the datasets which will be findable through the European Datasets Catalogue and national datasets catalogues, access to which can be requested. RIs in EOSC (e.g., BBMRI-ERIC and ELIXIR) are aiming

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

to enlarge the discoverability beyond the catalogued datasets by providing search engines to extract “virtual” datasets according to the requirements of the users..

- EOSC is multidisciplinary and can provide integration with data from different disciplines, some of them already developing and using DCAT.-AP specifications. The wide adoption of DCAT-AP in EOSC will be key.
- The interoperability in EHDS is intrinsic to the management of distributed, multicountry health data. EOSC focuses on interoperability and it is developing a cross-discipline interoperability framework. The consideration of other disciplines could help in expanding the specification of EHDS metadata for the integration of other relevant data (e.g. climate, socio-economic, geographic, etc.).
- EHDS metadata records will be structured using the HealthDCAT-AP metadata standard. Therefore, if EOSC metadata records also use this standard then the metadata records can be findable through the EU datasets catalogue of the EHDS and viceversa the EHDS metadata records findable through the EOSC metadata catalogue.

Permission Management

This topic focused on the access to health data. Health data has been collected under specific conditions, which must be considered when reusing the data. RIs also have health data collections. Interactions between the EHDS legal framework and the EOSC access models. Several key points in the discussion were:

- Data in RIs has been already collected under specific consent. EHDS regulation will provide a legal basis for reusing health data without the need to receive consent from natural persons. Only a right to opt-out for the reuse of their health data is provided within the EHDS Regulation to the natural persons. If RIs connect to the HealthData@EU infrastructure of the EHDS as Authorised Participants then the data they control can be accessed in the context of EHDS, based on the legal framework of the RI but also any article in the EHDS regulation mentioning the Authorised Participants. Outside of the use cases of the EHDS, RIs can continue using their data under the legal ground provided by the consent used for their collection.
- EHDS defines an authorisation process based on the national HDABs (Health Data Access Bodies) for managing the data access applications and requests. If access is granted, the HDAB is responsible for ensuring that the data is provided into a Secure Processing Environment (SPE) for the user to do the analysis.
- Some alignment at the level of the AAI may be needed. EOSC uses the AARC framework common standards for AAI which has been developed over 10+ years and is being

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

implemented in different domains. EHDS uses eIDAS, which will also be supported by the EOSC EU node.

- EOSC actors involved in the EHDS should analyse how the roles in EHDS could facilitate them in pursuing their objectives as Data Holders, Data Users, Health Data Intermediation Entities, Trusted Data Holders, Cross-Border registries, or Authorised Participants within the HealthData@EU infrastructure. It will be important that EOSC actors analyse the EHDS framework and discuss the challenges, as it may be difficult for some providers to meet the requirements for some roles such as Secure Processing Environment providers or Authorised participants.
- Permission management in EHDS is happening at the national HDAB-level, but there can be more than one HDAB per member state.. In EOSC it is much more distributed, as it can involve different service providers with different heterogeneous resources.

Accessing and processing health data

The focus of the discussion of this topic covered the accessibility through Secure Processing Environments, federated models, data infrastructures supporting the processing, catalogues and registries of certified software, interoperability of data, benchmarking. Several key points in the discussion were:

- HDABs decide on the SPEs to be used, and researchers and Trusted Data Holders can make suggestions, potentially accepting the use of transnational SPEs. EOSC is developing Trusted Research Environments for research on sensitive data (ENTRUST, TITAN, SIESTA, etc.). A possible action would be to align the concept of TREs in the context of the definition of the SPEs of the Data Governance Act and collect feedback for the requirements and even the specification of the implementation of SPEs.
- There is a need to ensure that output results do not contain sensitive information either directly or indirectly through generative models. This could be a common problem in EOSC and EHDS when dealing with sensitive data .
- EOSC can support HDAB in understanding the requirements and conditions for research use cases beyond the scope of the health data, as the needs and scenarios of the researchers are wider and more diverse in the EOSC context.
- SPE requirements will define the conditions and restrictions of the software running on pseudonymised data to avoid data privacy leakage risks and other unallowed data uses. Best practices on Software certification would be needed, especially on trustworthiness including privacy preserving aspects. Projects in the EOSC ecosystem could contribute. Registries of certified tools would also be important.

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

Reusability of data and outputs

Focus of the discussion: Usage licensing conditions, conditions for derived data, publication and sharing of results, referencing and provenance, compliance to standards. Several key points in the discussion were:

- Researchers may only retrieve anonymised data from the SPE via EHDS. Peer reviewers of journals may automatically also get access to the study protocol used in the research in order to validate the results . This could also include the study process, code repository and tools.
- Data processing tracing. The provenance information is key for the reusability of the output results. EOSC promotes the use of Persistent Identifiers and rich metadata to find and describe the objects used to build an output result (e.g. the RO-Crate and ISO 23494 communities). For data objects in EHDS built on demand, information about the filtering and preprocessing will be key. However, this information may become itself sensitive, and hence distributed provenance chains are being developed (ISO 23494-2). Provenance links will be likely removed during the anonymisation of the results.
- EHDS focuses on Reusability, and EOSC is also deeply interested in Reproducibility, which requires the same conditions (much more challenging in EHDS as datasets in SPEs are destroyed after being used, and reproducibility should be possible through new data sets, using the same extracting, preprocessing, timestamps and new DOIs).
- All data access application forms, permits, results and publications from the EHDS will be publically available.
- EOSC could act as the "Open Science" window of EHDS outputs, providing catalogues, persistence and reusability services for such open results.
- The experience on the evaluation and development of data according to FAIR principles in EOSC will be useful in the EHDS context, where EHR data will not necessarily be FAIR - it needs to be mediated and prepared for usage. EOSC has experience in evaluating the FAIRness of a system, which is at least as important as the FAIRness of a dataset.
- The EHDS Transparency requirements could be supported by EOSC.

Final Conclusions

EHDS and EOSC concur on the use and exploitation of research outputs. Data sources in EHDS are defined within the Regulation (for example: Real World Health Data from clinical data holders and derived data warehouses) and differ from the data sources on health data in EOSC (mainly comprising anonymised datasets from research and innovation activities). EOSC is focused

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

more on the research outputs and important synergies can be found in this part. Considering the Data User's Journey of the EHDS, we identify the following potential synergies:

- **Data discovery**: The EU Datasets Catalogue and the national and institutional catalogues should be compatible and complement each other. EHDS and EOSC federation should pursue interoperability at the level of services and metadata specifications, exploiting the efforts of EHDS in Health DCAT-AP, the EOSC Interoperability framework, the EOSC PID infrastructure and the technical specifications of services. EOSC EU node and SIMPL will be interoperable by their contracts. The adoption of Health DCAT-AP for health related data in EOSC will make these data visible at the EHDS level.
- **Permit application**: EOSC AAI in the EOSC EU Node supports EU Login (eIDAS). Therefore, EOSC and EHDS services are interoperable at the level of the authentication. Interoperability at the level of attributes necessary for authorisation (e.g., institutional affiliation) and results of authorisation process could be explored. The eduGAIN model supported in EOSC AAI includes information about the user affiliation and role which are relevant for the definition of user roles and their liability (permissions are much more complex, though). A joint model will facilitate the use of resources from EOSC nodes related to health.
- **Data preparation**: EOSC services related to knowledge graphs in EOSC that link different related research objects will also facilitate the discovery of additional information relevant for conducting the research. The link to cross-discipline data in EOSC will be relevant for research studies that require socioeconomic, environmental mobility or any other relevant data.
- **Data provision**: Conceptually SPEs, can leverage technology and resources from EOSC TREs (if they were EHDS compliant).The EHDS technical specifications of SPEs can be influenced by the development of EOSC TREs and actual services of e-Infrastructures in Health RIs can act as SPEs. A common definition of TRE will be needed.
- **Data use**: On one hand, EOSC (certified) software registries and best practices can facilitate SPE software availability, control, software provenance. On the other hand, EHDS actors can leverage experience on the management of sensitive data for research, by revising architectures, expanding requirements and reusing existing developments in EOSC, beyond the services and measures implemented in SPEs.
- **Results output**: The results obtained in activities related in the frame of EHDS should be preserved in EOSC research object registries. EOSC can describe in detail the standards, formats, policies and services for the registration and reuse of research output and EHDS define the restrictions related to the reusability policies.

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

Figure 3 shows a diagram of the potential interactions among HealthData@EU infrastructure and the EOSC Federation.

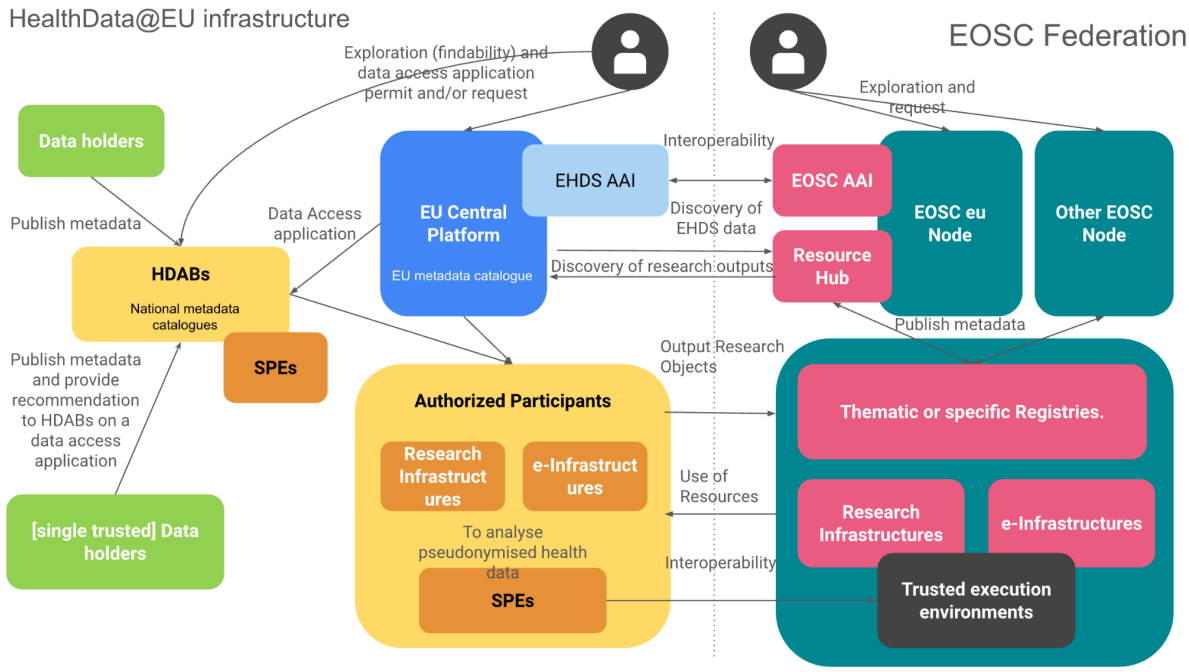

Figure 3: Illustrative diagram with some interactions between EHDS and EOSC discussed at the meeting. A partner can play different roles.

Potential actions

- Identify gaps, redundancies, and synergies among the work implemented in the EOSC projects that are relevant for the above points in the EHDS by means of a session in the EOSC Winter school in January 2025. This will be within the remit of EOSC Health Data Task Force.
- Analyse the interoperability between PIDs in EHDS and EOSC. Activity that can be proposed to the Opportunity Area Expert Group 1 (Persistent Identifiers).
- Promotion of the use of Health DCAT-AP and the development of reference implementations. This activity can be proposed to the EOSC Health Data TF.
- Inclusion in the EU Central platform resources form the Resource Hub and vice-versa. This activity can be proposed as part of the EOSC Federation validation to be started in 2025.
- Analyse the suitability of the EOSC AAI framework for EHDS. This is a service being tested in the EOSC Federation and the output of this analysis will be relevant for EHDS.

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

- Make an inventory of services and best practices for the secure processing of data from the EOSC related projects (e.g. EOSC ENTRUST Blueprint v1.0 for SPEs and EOSC EVERSE SW catalogues and best practices, EOSC TITAN and SIESTA privacy preserving services). This activity can be performed by the EOSC Health Data TF.
- EOSC can detail the paths for the preservation and reuse of research output data to the EHDS so protocols are adjusted and EOSC services for this purpose are listed in EHDS Health Data Infrastructure.

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