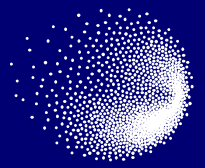


PSI Center for Scientific Computing, Theory and Data

ETH-ORD overview

SciCatCon 2024

Alun Ashton
PSI, 03 July 2024

What Is The ETH Domain

PSI

ETH BOARD ETH zürich EPFL PAUL SCHERRER INSTITUT PSI WSL Empa Materials Science and Technology eawag aquatic research

The ETH Domain comprises the two Federal Institutes of Technology in Zurich (ETH Zurich) and Lausanne (EPFL), as well as the four research institutes: the Paul Scherrer Institute (PSI), the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), the Swiss Federal Laboratories for Materials Science and Technology (Empa), and the Swiss Federal Institute of Aquatic Science and Technology (Eawag). With its numerous locations, the ETH Domain is firmly rooted in the Swiss regions.

swissuniversities

Open Research Data – Position of the ETH Domain

"Science often benefits from multiple evaluations of the same data sets by researcherswith different backgrounds, perceptions and ideas. ORD allows – and stimulates – new research and discoveries"

"A governance framework that defines the rights and duties of data producers, data owners and data users is needed. Indeed, a culture that gives proper credit to data producers needs to be promoted, along with the requisite tools; and research communities need to develop best practices for data citation."

"A strategy should be formulated for the development of ORD-related infrastructure and services that responds to future demands and to develop funding schemes in order to support these infrastructures and services in the long-term."

ETH Domain ORD Initiative Measures (swissuniversities) PSI

- An initiative to support Open Research Data (ORD) / Open Science, with about 15 million francs* until 2024**. The program includes five distinct measures, including a bottom-up, research-oriented projects fund and a measure to improve ORD services and infrastructure across ETH Domain institutions.

M1: Support researchers engaging in, or developing, ORD practices with and for their community. Assist these researchers in becoming Open Science leaders in their field.

M2: Improve the ecosystem of research data management (RDM) services and infrastructure that support ORD practices, through Domain-wide, coordinated action focused on interoperability and access.

M3: Raise availability and quality of training on RDM in application to ORD practices.

M4: Legal aspects of RDM

M5: Careers of staff supporting ORD practices

*Half the 15mCHF will be matched from the ETH domain institutions

**Until 2024 = all funds committed by end of 2024 and spent by end of 2026

ORD@PSI Steering Committee (representatives as in 2023)

Oversight of ORD activities in PSI.

Review and prioritization of PSI central funding* as matching for measures in the ETH ORD calls

Alex Amato	**DIRK Representative and Chair**
Oliver Bunk	**SLS**
Michel Kenzelmann	**SINQ**
Andreas Suter	**SμS**
Klaus Kirch	**CHRISP**
Markus Ammann	**ENE**
Gregor Cicchetti	**BIO**
Didier Gavillet	**NES**
Alun Ashton	**SCD and Science IT Infrastructure Department**
Michèle Erat	**Directors Staff, support and external liaison**

*1 MCHF

ETH ORD EG members join as required:

- EG-R: Uwe Filges, Markus Ammann
- EG-SI: Spencer Bliven, Leonardo Sala

Measure 1:
Call for Field-Specific Actions - handled by EG-R

Goals

– Support ETH Domain actors to become ORD leaders in their field by funding bottom-up proposals for ORD initiatives.
– Strengthen (in synergy with Measure 2) the connections between research communities and services and infrastructures providers.

Projects with different maturity level:
– Explore projects (circa 30k)
– Establish projects (circa 150k)
– Contribute projects (circa 1.4m)

Measure 2:
Coordination of Access to Research Data Management (RDM) Services & Infrastructures – handled by EG-SI

Goal

Standardization, federation, or mutualization of existing services across the ETH Domain and the improvement of interoperability of existing services and filling of gaps in services.

The full data lifecycle will be considered:
(1) Data management planning,
(2) Data acquisition,
(3) Data processing and analysis,
(4) Data publication and re-use,
(5) Long-term preservation,
(6) Data disposal after the end of the retention period.

Also envision to increase standardization, federation and mutualization of ETH Domain hardware infrastructures and services.

Measure 3: Development of Online Course Material for RDM Training – handled by EG-R

Goal

Develop a training concept and online teaching material for RDM building on the existing expertise in the ETH Domain. The course material should be structured into small, freely reusable and modifiable modules (using Creative Commons CC-BY license and existing platforms) and free of costs.

Measure 4: Information on Legal Questions related to ORD – handled by EG-SI

Goal

Develop a consolidated legal document, which can be used as a common reference and guidance in understanding the legal situation of ORD within the ETH Domain.

Objectives and tasks

Collection of a question catalogue for legal questions related to ORD. Questionnaire to be answered by experts in digital law (i.e. https://ccdigitallaw.ch, swissuniversities P5 program or EMPA legal team).
Legal document to be drafted (in accordance with ETH Regulation) and harmonized with each institution. Training material to be elaborated for the onboarding of new employees.

Measure 5:
Career Paths for ORD Professionals - handled by EG-R

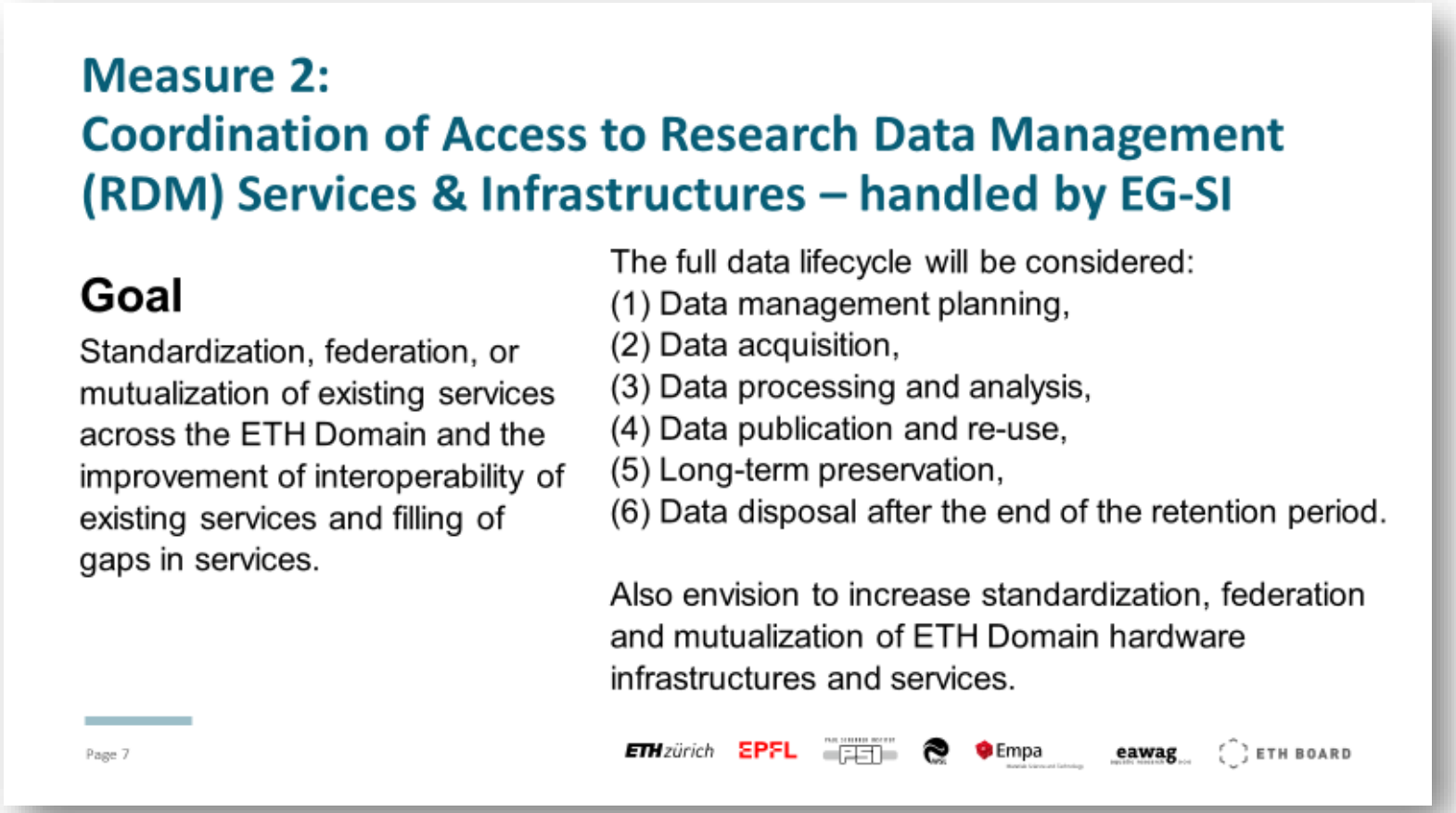
Goal

Develop proposals for future career paths of ORD professionals in the ETH Domain.
In order to offer top scientific IT services to ETH Domain researchers, it is crucial to recruit skilled professionals.

Of biggest interest to PSI and SciCatCon

- M1 – OpenEM, Spencer Bliven

- Landscape analysis
- ELN interoperability call
- API call: Caterina Barillari
- Focus topic for the central point of information
- Focus topic on ID/access management
- Additional activities: additional activities proposals out for review.



Future horizon beyond 2024 (and expenditure to 2026)

SDSC and SDSCHub@PSI

PSI

SDSC

THE SDSC | RESEARCH | INNOVATION | EDUCATION | EVENTS | CONTACT

In 2017, a national Data Science initiative from the ETH Board resulted in the creation of a unique joint venture: the Swiss Data Science Center, bringing together the ETH Zürich, the EPFL, and the PSI . The Center's mission is to accelerate the use of data science and machine learning techniques within academic disciplines of the ETH Domain, the Swiss academic community at large, the public institutions and the industrial sector.

A team of senior data scientists and experts in domains such as personalized health & medicine, earth & environmental science, social sciences, digital humanities and economics collaborate on academic and industrial projects. This unique positioning, at the crossroad of academic excellence and a fast-paced business environment, is key to simplifying a complex data science journey.

SDSC+ First National Calls for Projects

https://www.datascience.ch/call-for-projects

PSI

SDSC

First National Calls for Projects

1 June 2024

Focus Areas

Submission of 15th July 2024
Final Decision 31st October 2024
Start Q12025

- **Collaborative Research Data Science Projects**
- **End-User Innovation Projects (Track suggestions)**
- **Collaborative Open Research Data (ORD) Infrastructure Projects**

Tracks

- **Environment, Climate and Energy (ECE)**
- **Health and Biomedical Sciences (HBMS)**
- **Data Science for Large-Scale Infrastructures (LSI)**
- **General Track (G)**
- **Energy and Sustainability**
- **Health and Biomedical**
- **Societal Transformation**
- Sensitive Research Data Management (SRDM)
- Research Data Governance and FAIRification (RDGF)
- Standardized and Interoperable Components (SIC)
- Metadata Enhancement and Semantic Integration (MESI)
- Production Environments for Research Code Integration (PER)
- Enhancing, improving, or extending the Renku platform through infrastructure integration (RPI)
- Improving the sustainability of data science projects by leveraging the Renku platform (RPDS)