Towards cross-domain interoperability frameworks for scientific data

EOSC Symposium 2022

Interoperability Session

Prague, 16 November 2022

Dr. Michel Schouppe
Senior Expert, Open Science
DG R&I, European Commission

"*The value of data lies in its use and re-use*"

European strategy for data (COM(2020)66 final)

- Research data need to **integrate with other data** and **interoperate with applications or workflows** for analysis, storage and processing

- **Key for EOSC to generate added value** from the resources made available through the EOSC federation

- The overall research domain is a very **dynamic and heterogeneous environment** with many different stakeholder communities

  - **Disciplinary interoperability frameworks** exist in several communities
  - = f(culture, data types and infrastructures used by that community, maturity level and dynamics, etc)

- EOSC role to break down silos and broker across borders and disciplines

- ➢ To provide an **Interoperability Framework** [framework of frameworks?] **across domains**, communities and infrastructures

Some milestones (amongst others)

2014-2016: Birth of the FAIR principles
The FAIR principles were conceived at the Lorentz conference in 2014 and published by Wilkinson et al. in 2016 following consultation via FORCE11

March 2018: Study "Cost of not having FAIR research data"
by PwC EU Services. Overall cost estimated at €10.2bn/year for the EU alone

November 2018: "Turning FAIR into reality"
Final report and action plan from the EC expert group on FAIR data

2019-2020: RDA Working Group on "FAIR data maturity model"
Core assessment criteria for FAIRness

Feb 2021: EOSC Interoperability Framework (EOSC IF)
Report from the EOSC Executive Board Working Groups FAIR and Architecture identifies the general principles that should drive the creation of the EOSC IF

Some milestones (amongst others)

7 June 2021: Declaration 'Riding the Next Wave of Research Data: Leveraging the COVID-19 response towards advancing data interoperability (workshop under the Portuguese Presidency of the EU)

- Coronavirus crisis:
  - Pressure on the research community to speed up scientific discovery
  - Response: global collaboration and significant increase in early and open sharing of knowledge, data and tools.

- Hurdles:
  - Ability to quickly find and integrate different data types, such as virus and host genomics, clinical, epidemiological and social science data, originating from laboratory, hospital or industrial settings across countries, organisations and communities.

- Pre-requisite for impactful science:
  - Be able to ride the next wave of data and to increase the volumes of relevant FAIR and open data and their reuse in multidisciplinary science.
  - Call for a wide mobilisation across the European Union towards developing – in the context of the European Open Science Cloud - coherent interoperability frameworks that build on research community practices to enable interdisciplinary research and innovation.

Where are we now ?

Where will we be in three years from now?

Thank you

European Commission