



FAIR data Expert Group Call for Contributions

The remit and objectives of the FAIR Data Expert Group

The European Commission has established an Expert Group on FAIR data to support the Research and Innovation policy development on Open Science. The overall objective of the Expert Group is turning the [FAIR data principles](#) into an operational reality, to ensure that research data are Findable, Accessible, Interoperable and Reusable. The Group will address five specific objectives:

- I. To develop recommendations on what needs to be done to turn each component of the FAIR data principles into reality
- II. To propose indicators to measure progress on each of the FAIR components
- III. To provide input to the proposed European Open Science Cloud (EOSC) action plan on how to make data FAIR
- IV. To contribute to the evaluation of the Horizon 2020 Data Management Plan (DMP) template and development of associated sector / discipline-specific guidance
- V. To provide input on the issue of costing and financing data management activities

The Group will review existing initiatives and analyse the components of FAIR to recommend how they could be implemented and supported. Several topics will be addressed in the report, including research data culture, skills, incentives, service components, data management planning, metrics to evaluate FAIR, cost models and sustainability.

Further information is available on the [FAIR data expert group webpage](#).

What to contribute?

We invite contributions on each proposed section of the report as follows:

- Concepts - why FAIR?
- Research data culture
- Making FAIR data real
- Measuring change
- Facilitating change

In the **concepts and context** section of the report, we will introduce FAIR and other similar principles that promote effective data documentation and sharing to enable reuse. These will include the OECD principles, the Royal Society concept of 'intelligent openness' and the RDA Global Digital Object Cloud. This section will present clear definitions of what it means for data to be FAIR, and why this is important. We will analyse the relationship between FAIR

and the principle that data should be as open as possible and as closed as necessary. It will be important to understand - in terms of concepts and definitions - what are the component attributes which enable data to be Findable, Accessible, Interoperable and Reusable. Identifying such definitions will allow us to explore and assess proposals for degrees or scales of FAIRness, which will help researchers and institutions with implementation. The role of data selection and case studies on the benefits of FAIR will also be considered.

Specific questions to guide external contributions are:

- Why do we need FAIR data? What challenges does science face without it?
- What case studies can be shared of FAIR data in practice and the benefits this brings?
- What is a good definition of X, where X is a key concept related to FAIR data?
- What other principles or guidelines that parallel FAIR should be considered?
- What are the characteristics that allow data to be Findable, Accessible, Interoperable and Reusable?
- How are these attributes best deployed in a model that presents a scale of FAIRness (e.g. increasingly FAIR Digital Objects) to assist implementation?
- Are there any issues you have encountered, from a conceptual and definitional perspective, in interpreting the FAIR principles?
- What models or criteria exist for identifying 'valuable' data that define what should be retained?

In the **research data culture** section of the report, we will consider who are the critical players in the research community and what needs to be done by each of them to make FAIR data a reality. This will address research culture, workflows and skills requirements. We will also critically assess how the FAIR principles apply to different disciplines outside of the life sciences - where they originated - and whether differences in working practices necessitate amendments. This section will also address the European Commission template for FAIR Data Management Plans. Here, we will assess the appropriateness of the existing EC template and make recommendations for discipline specific and machine-actionable approaches.

Specific questions to guide external contributions are:

- How FAIR are current research practices in different parts of the research ecosystem (e.g., disciplines, sectors, geographic regions...)?
- What are good examples of aggregating large amounts of data of different origins and how this offers new scientific possibilities?
- What are the key barriers to FAIR practices (e.g. lack of metadata standards, domain repositories, data sharing norms, ...)?
- What could different players do to increase the FAIRness of data within their remit? How have research workflows actually been adapted in order to make data more FAIR? How much of this can be automated?
- What training opportunities and career paths are needed to support researchers and other players in the research ecosystem with data management and sharing?
- What improvements could be made to the current EC approach to DMPs?
- How can DMPs become more integrated and machine-actionable?

In the **making FAIR data real** section of the report, we will consider the operational challenges and components needed at a national and international level to provide a FAIR data ecosystem. This will address the role of repositories, registries, standards, identifiers, workflows, skills and legal interoperability. The report will explore how national and domain infrastructures interoperate and make recommendations for the European Open Science Cloud (EOSC) roadmap.

Specific questions to guide external contributions are:

- To what extent are the FAIR principles alone sufficient to reduce fragmentation and increase interoperability? The principles have a great potential to influence the minds of stakeholders towards more efficient data sharing and reuse, but perhaps additional measures and more specifics are needed to guide implementation?
- What are the necessary components of a FAIR data ecosystem in terms of technologies, standards, legal framework, skills etc?
- What existing components can be built on, and are there promising examples of joined-up architectures and interoperability around research data such as those based on Digital Objects?
- Do we need a layered approach to tackle the complexity of building a global data infrastructure ecosystem, and if so, what are the layers?
- Which global initiatives are working on relevant architectural frameworks to put FAIR into practice?
- A large proportion of data-driven research has been shown to not be reproducible. Do we need to turn to automated processing guided by documented workflows, and if so how should this be organised?
- What kind of roles and professions are required to put the FAIR principles into place?

In the **measuring change** section of the report, we will focus on metrics, incentives and methods to assess the FAIRness of data and repositories. Current career progression for academic researchers is deeply dependent on author-centric metrics linked to publications. Researchers who devote time and expertise to data curation and sharing are not sufficiently rewarded for this. We welcome examples of incentives that are changing practice towards FAIR, such as data being recognised in research assessment exercises, data sharing and curation forming part of promotion criteria, or data being required in reference lists in grant proposals. Emerging models for assessing the FAIRness of data and repositories, such as those from DANS and 4TU, will also be reviewed. Feedback on these methods or examples of how organisations plan to assess FAIR are welcomed.

Specific questions to guide external contributions are:

- What metrics are you using to track the impact of research data?
- How are you incentivising FAIR data practices, and which methods are working best?
- How can FAIR data practice be incorporated into career progression criteria?
- Do repositories plan to include a FAIR rating alongside dataset metadata?
- Are the FAIR criteria a useful way to assess the trustworthiness of data repositories?
- What factors should be monitored to evidence a change in practice towards FAIR data?
- Given variations in disciplinary practices, are different factors more/less relevant to note a shift in practice in certain domains?
- How long will it take to see measurable change across the broad research community or specific subsets thereof?

In **the facilitating change** section of the report, we will consider what enablers are needed in terms of data policy, investment and skills. Recognising that many of the issues faced are social rather than technical, we will make recommendations for creating data stewardship career paths and recognising these roles. Cost models for investing in the FAIR data ecosystem at a national and international level, and recognising additional resources required in proposals via DMPs will also be covered.

Specific questions to guide external contributions are:

- What sustainable business models exist for data infrastructure investments?
- How can policymakers on an EU and member state level ensure there is sufficient, sustainable funding to support a FAIR data ecosystem?
- Can lessons be learned from the existing practice of costing RDM activities into proposals?
- How do we build the community of 500,000 data stewards suggested for EOSC?
- What data stewardship roles and career trajectories need to be supported?
- What incentives are needed for individuals to drive the transition to FAIR practices?
- How can we promote cross-disciplinary and cross-border re-use of data?

How and when to contribute?

A [GitHub repository](#) has been established to enable the community to contribute suggestions and resources in the open. Please see the [guidelines and examples](#) on how to contribute. We suggest that you review the existing tickets first and add to these if your suggestion is related to previously shared ideas and resources. When you have something different to share, proceed to [raise a new issue](#).

We will also host two open community sessions to allow people to contribute ideas to the FAIR data expert group in real-time and in a moderated forum to allow discussion with the group. These will take place on Monday 3rd July from 14:00-16:00 CEST and Wednesday 26th July from 10:00-12:00 CEST. Further details on how to join will be circulated nearer the time.

Your contributions will be shared with the group members responsible for the associated topic. We may request further information or invite external experts to address the group at future telecons or face-to-face meetings.

The call for contributions is open between 12 June 2017 and 31 July 2017. Early input is encouraged to allow time for follow-up discussion.

Members of the expert group will be present at the European Open Science Cloud summit on 12th June and the Estonian Presidency event in Tallinn on 12-13th October 2017. Further information on our progress will be shared then.