

## An EUDAT-based FAIR Data Approach for Data Interoperability



To achieve data publication in a FAIR manner and foster their findability, accessibility, interoperability and reusability, a set of (FAIR) data tools are being developed, including the FAIR Data Point (FDP), which is a software layer on top of datasets to expose them as FAIR (inter-linkable) data. The FDP provides information about the available datasets in terms of their metadata as well as the actual access to the data in an interoperable format.

### The Scientific Challenge

Although the FAIR Data Point service will be useful to any research community facing massive data management and interoperability issues, the proposed pilot will specifically target the life science community. Due to the complex nature of biology, life science data arguably represent one of the most heterogeneous, diverse and challenging types of research data. Exposing new and existing datasets following the FAIR data principles will facilitate the improvement of our ability to interpret and combine these data. The aim of this pilot is to implement and deploy a FDP using a combination of existing Semantic Web standards and frameworks for the front-end, and (existing or new) EUDAT services for the back-end. A FDP provides access to the data and metadata using REST-APIs conforming to the W3C Linked Data Platform specification.

### Who benefits and how?

The FAIR Data Point service will be useful to any community facing massive data management and interoperability issues, targetting specifically the life science community.

With the emergence of high-throughput methods (next-generation sequencing, microarrays, etc.), the life sciences have become increasingly data-intensive and novel insights and scientific breakthroughs will be dependent on our ability to interpret and combine existing as well as newly generated data. Exposing new and existing datasets following the FAIR data principles will facilitate these goals.

A FDP service built on the EUDAT infrastructure will offer advantages to individual researchers, as well as research groups and consortia. Existing, small-scale semantic data repositories are frequently managed by the researchers themselves and are notoriously difficult to maintain, resulting in frequent unavailability and short repository life spans. Therefore, one of the benefits of the service is to be able to completely remove this burden from the researcher. The FDP service would emphasize the novelty of such a service since to date; no Semantic Web-enabled repository services are available to the general research community. Moreover, FDPs are designed to enable data citation and maintain statistics about data accesses, which means that impact will be measurable for any FDP deployment.

### Technical Implementation

The FAIR Principles provide an implementation-independent, precise and measurable set of qualities for the publication of scientific research data. Following the FAIR Principles, data is easily discovered, easily evaluated,

EUDAT receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 654065.



and maximally reusable for both human agents and automatic (software) agents. The development of the FAIR Data infrastructure is supported by DTL and a number of national and European projects, including Elixir, BBMRI, FAIRdict and RDConnect.

The DTL FAIR Data infrastructure is aligned with the European Open Science Cloud, and it is the first general-purpose implementation of the FAIR principles. The infrastructure is composed of several applications that are being developed as stand-alone tools that will later be integrated to form the Data FAIRport.

## Preliminary Results

The first phase of the pilot has focused on knowledge exchange and exploring/evaluating EUDAT services. In the second phase, this data pilot will take the next steps towards the EUDAT infrastructure adopting the FDP concept (as described in B2SHARE/B2SAFE to FDP). It will finalize the nearly-completed integration of FDP in B2FIND (FDP to B2FIND). The adoption of PIDs (B2HANDLE for FDPs) will be taken up by the DTL FAIR development team.

## Contacts

- Luiz Olavo Bonino, DTL, [luiz.bonino@dtls.nl](mailto:luiz.bonino@dtls.nl)
- Mark Thompson, LUMC, [m.thompson@lumc.nl](mailto:m.thompson@lumc.nl)
- Christine Staiger, SURFsara, [christine.staiger@surfsara.nl](mailto:christine.staiger@surfsara.nl)

## Further Information

[Learn more about the EUDAT/FAIR Data Collaboration here](#)  
[Scientific Challenges behind the Pilot](#)

Category:

[Data Pilot Communities](#)

[Read more](#)