Rare diseases GO FAIR!

A networking initiative to foster the implementation of principles to make rare disease data **F**indable, **A**ccessible, **I**nteroperable and **R**eusable for humans and computers (**FAIR**)

Our target: speeding up progress towards early diagnosis and new treatments of rare diseases.

Challenge: clinicians, patients and researchers should be able to perform analysis across data systems efficiently and without error. Even though rare disease data are collected everywhere, data are often not compatible, difficult to find and usage restrictions are often unclear.

Solution: Global Open FAIR implementation networks: **community networks** to help each other choose and adopt standards to implement FAIR. We introduce the initiative to form a GO FAIR implementation network for the rare disease community: **RDs GO FAIR!**

Lightweight organisation

A seed group of volunteers interacts with liaisons in the rare disease community

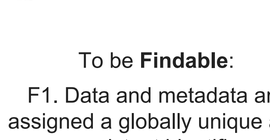
Represented are patients, EURORDIS, Orphanet, medical & FAIR information specialists, international infrastructures (ELIXIR, BBMRI-ERIC, ERNs, NIH), contacts for small enterprises and pharma

Patients

FAIR enables data custodians to take responsibility for making rare disease data as reusable as possible. Patients drive the culture change that this requires. The network helps organise workshops and webinars with patients.

IRDiRC RECOGNIZED RESOURCES

Metadata are data *about* other data, such as the name and purpose of a registry, and when and by whom data were collected. FAIR implies that metadata are also machine-readable.



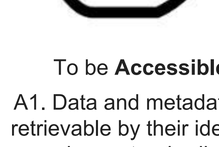
To be **Findable**:

F1. Data and metadata are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (see R1)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource



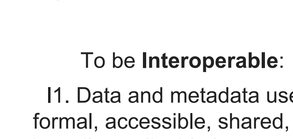
To be **Accessible**:

A1. Data and metadata are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

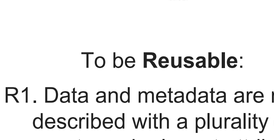


To be **Interoperable**:

I1. Data and metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. Data and metadata use vocabularies that follow FAIR principles

I3. Data and metadata include qualified references to other data and metadata



To be **Reusable**:

R1. Data and metadata are richly described with a plurality of accurate and relevant attributes

R1.1. Data and metadata are released with a clear and accessible data usage license

R1.2. Data and metadata are associated with detailed provenance

R1.3. Data and metadata meet domain-relevant community standards

Machine readable: values, value types, and relations between them are re-coded by globally understood computer codes.
Example: "Marco *has phenotype* blurred vision" becomes <oid:0...772X> <*obo:RO_0002200*> <hpo:0000622>
Non-personal codes can be looked up online (e.g. bioportal.bioontology.org). Similarly encoded resources are compatible and instantly analysable by computer.
Access permissions can also be encoded.

Standardisation

The network fosters activities to increase the level of standardisation when striving for conformity with the FAIR principles. This makes retrieving information and analysis over many data sources more efficient and insightful, and reduces error.

Data analytics

FAIR has a purpose: enable better (re-)use of data. This includes **advanced data analysis.** The network will expose the community to the possibilities of cross-resource analysis and stimulate innovation for the benefit of rare disease patients.

Adoption

RDs GO FAIR will contribute to efforts that organise the community towards adopting FAIR data principles, such as by FAIR data steward networks and business models for FAIR services that can serve all rare diseases.

Background In 2014, global leaders in data science conceived the FAIR principles to reduce the cost of cross-resource data analysis[1]. Preliminary estimates for the Research Data Alliance surpass 10 billion € per year for not having FAIR principles in place in Europe[2]. The FAIR principles are endorsed by EU infrastructures, the NIH, the G20 and G7, and IRDiRC for rare diseases. The GO FAIR concept was born from discussions about how to implement the European Open Science Cloud (with a global scope). RDs GO FAIR is one of several initiatives that receive organisational support from GO FAIR international support & coordination offices.
For a quick introduction to the concept of a FAIR infrastructure for health, search for 'Personal Health Train' on Vimeo or YouTube[3]

A '**seed group**' volunteered to organise the RDs GO FAIR network: Virginie Bros-Facer (EURORDIS), Claudio Carta (FAIR training), Ronald Cornet (clinical information specialist), David van Enckevort (biobanking and software development expert), Ian Harrow (contact for industry including pharmaceuticals R&D), Victoria Hedley (RD policy co& liaison with ERNs), Kristina Hettne (FAIR data analytics expert), Dipak Kalra (contact for profit and non-profit software tool developers), Veronica Popa (patient representative), Ana Rath & Marc Hanauer (Orphanet), Marco Roos (FAIR infrastructure for rare diseases), Yaffa Rubinstein (contact for the NIH Data Science special interest group, patient registries, biospecimens collections), Rachel Thompson (rare disease data specialist), Mark Wilkinson (information specialist, FAIR guiding principles). Additional contributors: Annika Jacobsen & Mark Thompson (FAIR data stewardship), Mascha Jansen & Luiz Bonino (International GO FAIR office, the Netherlands).

Help us create a powerful network of FAIR and secure rare disease data sources!
Express your interest in the RDs GO FAIR initiative by contacting the members of the seed group or visit
https://www.go-fair.org/implementation-networks/overview/go-fair-rare-disease/

[1] Wilkinson *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship", https://www.nature.com/articles/sdata201618 (doi:10.1038/sdata.2016.18)
[2] https://twitter.com/micmenn/status/976838188013703174, https://rd-alliance.org/how-expensive-fair-compliance-and-how-expensive-it-not-be-fair-compliant-rda-11th-plenary-bof
[3] https://www.youtube.com/watch?v=mktAtHmy-FM&t=27s