

Making nano data FAIR enough

The FAIR principles provide compelling guidelines on how to achieve reusability of nanotechnology data.

The widespread use of nanomaterials in real-world applications can only be possible if potential risks to health and the environment are comprehensively regulated. Research efforts aimed at assessing the environmental, health and safety (EHS) implications of nanomaterials have received attention from scientists and support from funding agencies for a long time. Analysis of this large data supply is paramount to maximize the potential offered by the acquired knowledge and lead to widely applicable policies and regulations.

Data sharing is a first essential step, of course. Indeed, making data, in every scientific discipline, publicly available is among the main objectives of open science advocates. Data availability is not enough, however, unless proper curation can aid its reuse and analysis.

A remarkable effort towards the efficient reuse of scientific data was the publication, in 2016, of the FAIR principles (M. D. Wilkinson et al. *Sci. Data* 3, 160018; 2016) by a large consortium of authors. The acronym FAIR stands for ‘findable, accessible, interoperable and reusable’. Briefly, this means that each set of data must have an identifier, that enough information is provided for the data and the metadata to be accessed, that a set of data can be integrated with other sets and can be analysed by a range of software, and that permissions for reuse are clearly specified. More details can be found in Fig. 1 and in the 2016 paper.

A noticeable feature of the FAIR principles is that they aim at “enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals”. This is becoming more urgent, especially when considering the potential of machine-learning procedures to analyse large datasets and provide guidance for further scientific developments, but also for policy and governance.

Despite being only five years old, the FAIR principles have already been recognized by several international institutions. A first endorsement emerged from the G20 summit in Hangzhou (https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_16_2967), and the FAIR principles are also strongly encouraged and officially recommended by a large number of institutes and funding agencies around the globe.

Now, writing in *Nature Nanotechnology*, Penny Nymark and co-authors report an

The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- 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. (meta)data 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. (meta)data use a formal, accessible, shared and broadly applicable language for knowledge representation
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. (meta)data are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage licence
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

The FAIR Guiding Principles. Adapted from M. D. Wilkinson et al. (*Sci. Data* 3, 160018; 2016).

extensive analysis of how nano EHS data align with the FAIR principles. The analysis focuses specifically on data included in the eNanoMapper database (Beilstein *J. Nanotechnol.* 6, 1609–1634; 2015), which was created to collect data from a number of nanosafety projects to guide risk and safety policies on nanomaterials.

The outcome of the analysis by Nymark et al. is the identification of precise challenges that prevent reusability of nanosafety data, and consequently a set of recommendations to improve the current situation. We refer our readers to that paper for the list of recommendations; but to illustrate how rooted the problem is and how much work still needs to be done, we emphasize that simply in terms of findability, the authors report a general lack of persistent identifiers for data, which also leads to low motivation to share rich metadata for data interpretation and accessibility.

The reuse of data is also an essential element of sustainable nanotechnology design, as discussed by Desiree Plata and Nina Z. Janković in their Comment article. The essence of sustainable design is to provide guidance for material synthesis aimed at optimizing both performance and environmental friendliness. As the authors emphasize, the importance of considering the EHS aspects of nanomaterials was recognized very soon after the development of nanotechnology-enabled applications. But the absence of coordinated efforts to

report data on both EHS and performance aspects in a way that aids its reuse by human scholars and machines prevents progress in sustainable design. Plata and Janković provide a series of aspirations to guide progress in sustainable-by-design nanotechnology and emphasize the importance of interoperability, which is at the heart of the FAIR principles.

As stressed in both articles, the successful implementation of data deposition frameworks that allow appropriate reuse of scientific data needs the stewardship and continuous support of a number of stakeholders, including funding agencies and publishers. As a journal, *Nature Nanotechnology* fully supports data sharing inspired by the FAIR principles. We recommend our authors to follow suggestions from available resources on FAIRification of data, for example those of the GO FAIR initiative (www.go-fair.org/) and more specifically those of the FAIR implementation network for nano- and advanced materials (www.go-fair.org/implementation-networks/overview/advancednano/).

In the meantime, we invite you to join us on 24 June 2021 for an online panel discussion on data sharing and sustainable nanotechnology (www.nature.com/nnano/video). □

Published online: 11 June 2021
<https://doi.org/10.1038/s41565-021-00935-y>