

Improving data findability and accessibility / D1.2

Introduction

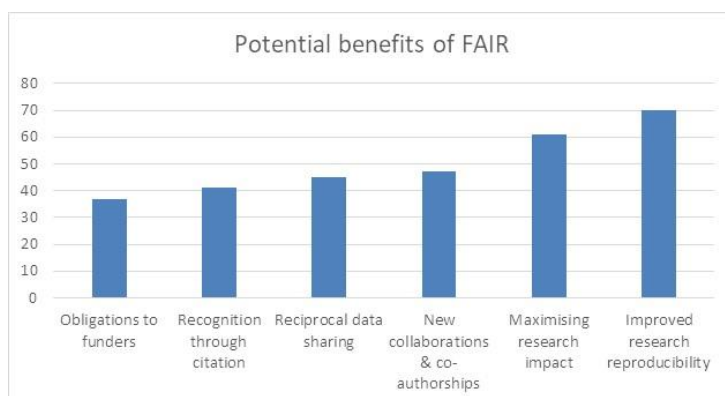
The aim of this work package is to improve the implementation of FAIR (Findable, Accessible, Interoperable and Reusable) principles in nano-EHS databases. This report focuses in particular on ways to improve data Findability and Accessibility.

Description of Work

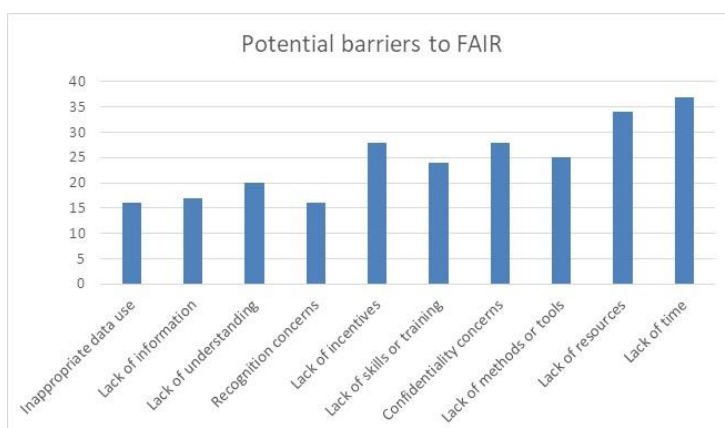
Incentives for scientists to make their data findable and accessible, and related barriers to implementation in practice, were investigated through a review of the published literature, supplemented by an online survey of stakeholders working in a range of scientific areas. The findings of these activities have been summarised, and supplemented by additional investigations of feasibility and methods of citing and storing data sets, Intellectual Property and data privacy issues, and the development of machine-readable metadata. Based on the results of the review work, training resources to guide researchers in maximising the findability and accessibility of their data, were identified, with links provided to key sources.

Main Results

There were 128 responses to the online survey and results were consistent with the themes identified in the literature review, although it was notable that 40% of respondents had low awareness of the FAIR principles.



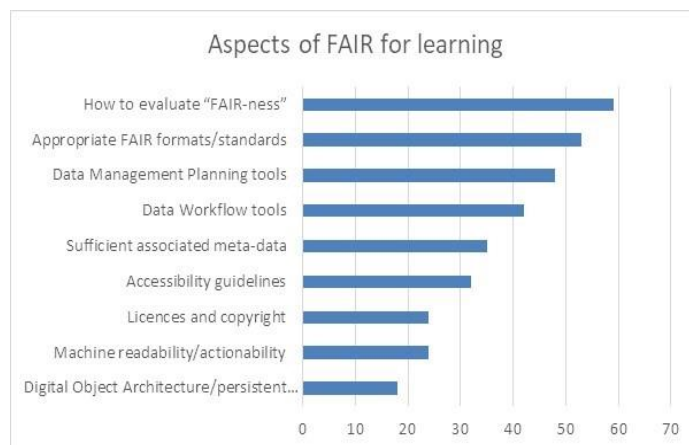
Findings from the work showed that while there is a general enthusiasm for data sharing among researchers, many fewer report actually doing so in practice.



The main benefits of and barriers to FAIRification of data highlight opportunities for data sharing, reproducibility and collaborative research contrasted with concerns about confidentiality and inappropriate data use, and a lack of time, resources, incentives, skills and methods/tools.

Increased awareness and provision of guidance regarding the FAIR principles could help to encourage more researchers

to make their data FAIR and available, and respondents to the survey expressed a preference for this to be in the form of practical data management tools, introductory information and materials and web-based guidance, with a focus on how to evaluate 'FAIRness' and appropriate formats and standards. Provision of detailed metadata and contextual data providing information on, for example, how and for what purpose the data was collected would also help to address researchers' concerns about data quality.



Concerns that efforts in collecting and processing the data would not be properly acknowledged or recognised once data was in the public domain, can be addressed through the use of standard citation formats for data sets, and by more stringent implementation of data sharing requirements by journals. While many journals provide detailed guides on data citation and on the use of recommended data repositories, most journals only 'encourage' or 'expect' rather than 'mandate' sharing of data associated with the papers that they publish. Lower fees for open access publications would also encourage wider implementation of data sharing by authors.

Perceived shortage of time and resources, and lack of skills and tools can be addressed through the provision of training resources and simple-to-use FAIRification tools and guidance, with completion of relevant training addressing the lack of skills and leading to more efficient implementation of FAIR thus saving researchers time and resource. A key training source is the eNanomapper interface. Measures to address concerns regarding availability of resources could also include a stronger recognition of the need for, and benefits of, FAIRification by funders and sponsors, and the inclusion of this topic within research calls, and invitations to tender. This would allow provision for FAIRification to be integrally included in study proposals and protocols and for resource time and funding to be explicitly costed into the bids.

Summary

There is a general enthusiasm for application of the FAIR principles to datasets, with key advantages in collaborative working and publications, but the major concerns of researchers are the time and resources required for implementation, and the lack of incentives, skills and tools. Training resources, in particular eNanomapper, should be used to enhance skills and knowledge of FAIR, with resulting efficiencies minimising the time and resource required. Standard citation formats, and further encouragement from journal publishers to share data, would contribute to improved recognition and implementation of sharing of FAIR data.

For more details about the Gov4Nano project please visit the Gov4Nano website. Public deliverables will be made available in due time via this website.