

FAIR but not Necessarily Open: Sensitive data in the domain of biodiversity

Patricia Mergen^{‡,§}, Sofie Meeus[§], Frederik Leljaert[§]

[‡] Royal Museum for Central Africa, Tervuren, Belgium

[§] Meise Botanic Garden, Meise, Belgium

Corresponding author: Patricia Mergen (patricia.mergen@plantentuinmeise.be)

Abstract

In the framework of implementing the European Open Science Cloud ([EOSC](#)), there is still confusion between the concept of data FAIRness (Findable, Accessible, Interoperable and Re-usable, Wilkinson et al. 2016) and the idea of open and freely accessible data, which are not necessarily the same. Data can indeed comply with the requirements of FAIRness even if their access is moderated or behind a paywall. Therefore the [motto](#) of EOSC is actually “As open as possible, as closed as necessary”. This [confusion or misinterpretation](#) of definitions has raised concerns among potential data providers who fear being obligated to make sensitive data openly accessible and freely available, even if there are valid reasons for restrictions, or to forfeit any charges or hamper profit making if the data generate revenue. As a result, there has been some reluctance to fully engage in the activities related to FAIR data and the EOSC.

When addressing sensitive data, what comes to mind are personal data governed by the General Data Protection Regulation ([GDPR](#)), as well as clinical, security, military, or commercially valuable data protected by patents. In the domain of biodiversity or natural history collections, it is often reported that these issues surrounding sensitive data regulations have less impact, especially when contributors are properly cited and embargo periods are respected. However, there are cases in this domain where sensitive data must be considered for legal or ethical purposes. Examples include protected or endangered species, where the exact geographic coordinates might not be shared openly to avoid poaching; cases of Access and Benefit sharing ([ABS](#)), depending on the country of origin of the species; the respect of traditional knowledge; and a desire to limit the commercial exploitation of the data. The requirements of the [Nagoya Protocol](#) on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity, as well as the upcoming Digital Sequence Information regulations ([DSI](#)), play an important role here. The Digital Services Act ([DSA](#)) was recently adopted with the aim of the protection of the digital space against the spread of illegal content, which sets the interoperability requirements for operators of data spaces. This raises questions on the actual definition of data spaces and how they would be

affected by this new European legislation but with a worldwide impact on widely used social media and content platforms such as Google or YouTube.

During the implementation and updating activities in projects and initiatives like Biodiversity Community Integrated Knowledge Library ([BiCIKL](#)), it became clear that there is a need to offer a secure data repository and management system that can deal with both open and non-open data in order to effectively include all potential data providers and mobilise their content while adhering to FAIR requirements.

In this talk, after a general introduction about sensitive data, we will provide several examples in the biodiversity and natural sciences domains on how to deal with sensitive data and their management, such as recommended by [GBIF](#). Last, but not least, we will highlight how important it is to use internationally accepted standards such as those from Biodiversity Information Standards ([TDWG](#)) to achieve such developments in the context of the Biodiversity Knowledge Hub ([BKH](#)) implemented by BiCIKL. Notably, by providing clear metadata about the terms of use, citation requirements and licensing, actual re-use of the data is made possible both legally and efficiently.

Keywords

FAIR data, EOSC, BiCIKL, legislation, ethics, metadata, re-use

Presenting author

Patricia Mergen

Presented at

TDWG 2023

Acknowledgements

European Commission, Meise Botanic Garden, Flemish Research Foundation Flanders (FWO), TDWG (Biodiversity Information Standards)

BiCIKL (Biodiversity Community Integrated Knowledge Library)

FOSB (Flemish Open Science Board)

Funding program

BiCIKL (Biodiversity Community Integrated Knowledge Library)

FOSB (Flemish Open Science Board)

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Wilkinson M, Dumontier M, Aalbersberg I, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. <https://doi.org/10.1038/sdata.2016.18>