



Conference Abstract

# Data ownership and data publishing

Lyubomir Penev ‡

‡ Pensoft Publishers & Bulgarian Academy of Sciences, Sofia, Bulgaria

Corresponding author: Lyubomir Penev ([penev@pensoft.net](mailto:penev@pensoft.net))

Received: 19 Aug 2019 | Published: 20 Aug 2019

Citation: Penev L (2019) Data ownership and data publishing. ARPHA Conference Abstracts 2: e39250.

<https://doi.org/10.3897/aca.2.e39250>

## Abstract

"Data ownership" is actually an oxymoron, because there could not be a copyright (ownership) on facts or ideas, hence no data ownership rights and law exist. The term refers to various kinds of data protection instruments: Intellectual Property Rights (IPR) (mostly copyright) asserted to indicate some kind of data ownership, confidentiality clauses/rules, database right protection (in the European Union only), or personal data protection (GDPR) (Scassa 2018). Data protection is often realised via different mechanisms of "data hoarding", that is withholding access to data for various reasons (Sieber 1989). Data hoarding, however, does not put the data into someone's ownership. Nonetheless, the access to and the re-use of data, and biodiversity data in particular, is hampered by technical, economic, sociological, legal and other factors, although there should be no formal legal provisions related to copyright that may prevent anyone who needs to use them (Egloff et al. 2014, Egloff et al. 2017, see also the [Bouchout Declaration](#)).

One of the best ways to provide access to data is to publish these so that the data creators and holders are credited for their efforts. As one of the pioneers in biodiversity data publishing, Pensoft has adopted a multiple-approach data publishing model, resulting in the ARPHA-BioDiv toolbox and in extensive [Strategies and Guidelines for Publishing of](#)

[Biodiversity Data](#) (Penev et al. 2017a, Penev et al. 2017b). ARPHA-BioDiv consists of several data publishing workflows:

1. Deposition of underlying data in an external repository and/or its publication as supplementary file(s) to the related article which are then linked and/or cited in-text. Supplementary files are published under their own DOIs to increase citability).
2. Description of data in data papers after they have been deposited in trusted repositories and/or as supplementary files; the system allows for data papers to be submitted both as plain text or converted into manuscripts from Ecological Metadata Language (EML) metadata.
3. Import of structured data into the article text from tables or via web services and their subsequent download/distribution from the published article as part of the integrated narrative and data publishing workflow realised by the [Biodiversity Data Journal](#).
4. Publication of data in structured, semantically enriched, full-text XMLs where data elements are machine-readable and easy-to-harvest.
5. Extraction of Linked Open Data (LOD) from literature, which is then converted into interoperable RDF triples (in accordance with the OpenBiodiv-O ontology) (Senderov et al. 2018) and stored in the [OpenBiodiv](#) Biodiversity Knowledge Graph

In combination with text and data mining (TDM) technologies for legacy literature (PDF) developed by Plazi, these approaches show different angles to the future of biodiversity data publishing and, lay the foundations of an entire data publishing ecosystem in the field, while also supplying FAIR (Findable, Accessible, Interoperable and Reusable) data to several interoperable overarching infrastructures, such as [Global Biodiversity Information Facility \(GBIF\)](#), [Biodiversity Literature Repository \(BLR\)](#), [Plazi TreatmentBank](#), [OpenBiodiv](#), as well as to various end users.

## Presenting author

Lyubomir Penev

## References

- Egloff W, Patterson D, Agosti D, Hagedorn G (2014) Open exchange of scientific knowledge and European copyright: The case of biodiversity information. *ZooKeys* 414: 109-135. <https://doi.org/10.3897/zookeys.414.7717>
- Egloff W, Agosti D, Kishor P, Patterson D, Miller J (2017) Copyright and the use of images as biodiversity data. *Research Ideas and Outcomes* 3 <https://doi.org/10.3897/rio.3.e12502>
- Penev L, Georgiev T, Geshev P, Demirov S, Senderov V, Kuzmova I, Kostadinova I, Peneva S, Stoev P (2017a) ARPHA-BioDiv: A toolbox for scholarly publication and dissemination of biodiversity data based on the ARPHA Publishing Platform. *Research Ideas and Outcomes* 3 <https://doi.org/10.3897/rio.3.e13088>

- Penev L, Mietchen D, Chavan V, Hagedorn G, Smith V, Shotton D, Ó Tuama É, Senderov V, Georgiev T, Stoev P, Groom Q, Remsen D, Edmunds S (2017b) Strategies and guidelines for scholarly publishing of biodiversity data. Research Ideas and Outcomes 3 <https://doi.org/10.3897/rio.3.e12431>
- Scassa T (2018) Data Ownership. CIGI papers 187 <https://doi.org/10.2139/ssrn.3251542>
- Senderov V, Simov K, Franz N, Stoev P, Catapano T, Agosti D, Sautter G, Morris R, Penev L (2018) OpenBiodiv-O: ontology of the OpenBiodiv knowledge management system. Journal of Biomedical Semantics 9 (1). <https://doi.org/10.1186/s13326-017-0174-5>
- Sieber J (1989) Sharing scientific data I: New problems for IRBs. IRB: Ethics and Human Research 11 (6). <https://doi.org/10.2307/3564184>