Mapping.bio: Piloting FAIR semantic mappings for biodiversity digital twins

Alexander Wolodkin[‡], Claus Weiland[‡,§], Jonas Grieb[‡]

‡ Senckenberg – Leibniz Institution for Biodiversity and Earth System Research, Frankfurt am Main, Germany
§ DiSSCo-D, Frankfurt am Main, Germany

Corresponding author: Alexander Wolodkin (alexander.wol@gmail.com), Claus Weiland (cweiland@senckenberg.de)

Abstract

Biodiversity research has a strong focus on the links between environment and functional traits, e.g., to assess how anthropogenic drivers of change impact ecological systems (Díaz et al. 2013). Interoperable exchange and integration of such data is enabled through the use of ontologies that provide "meaning" to data and enable downstream processing involving learning and inference over graph-structured models of these data (Kulmanov et al. 2020). However, the development of thematically similar semantic artifacts, e.g., the Environmental Ontology (ENVO, Buttigieg et al. 2016) and the Semantic Web for Earth and Environment Technology Ontology (SWEET, DiGiuseppe et al. 2014), in biodiversity-related disciplines (e.g., environmental genomics and earth observation) can introduce substantial conceptual overlaps, and highlights the need for bridging technologies to facilitate reuse of biodiversity data across those knowledge fields (Karam et al. 2020).

A recent design study, funded by the European Open Science Cloud (EOSC), proposes a framework to create, document and publish mappings and crosswalks linking different semantic artifacts within a particular scientific community and across scientific domains under the label of "Flexible Semantic Mapping Framework" (SEMAF, Broeder et al. 2021). SEMAF puts a strong emphasis on so-called pragmatic mappings, i.e., mappings that are driven by specific interoperability goals such as translations between specific observation measurements (e.g., sensor configurations) and metadata descriptions. Within the Horizon Europe Project "Biodiversity Digital Twin for Advanced Modelling, Simulation and Prediction Capabilities" (BioDT), a mapping tool leveraging SEMAF is currently under development: Mapping.bio provides a lightweight web service to read semantic artifacts, visualize them, add mappings as graphical connections and store the mappings as FAIR (Findable, Accessible, Interoperable Reusable) Digital Objects (FDOs, De Smedt et al. 2020) in a repository. To foster reusability, sustainably and long-term availability of digital objects, mapping.bio features mappings compliant with the Simple Standard for Sharing Ontological Mappings (SSSOM, Matentzoglu et al. 2022), a machine-interpretable and

extensible vocabulary enabling the self-contained exploration and processing of annotated mappings by machines (machine actionability, Jacobsen et al. 2020).

Keywords

FAIR Digital Object, BioDT, ontology mapping, ENVO, environment ontology, phenotype ontology, machine actionability, EOSC

Presenting author

Alexander Wolodkin

Presented at

TDWG 2023

Funding program

HORIZON-INFRA-2021-TECH-01-01 – Grant Agreement No. 101057437

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Broeder D, Budroni P, Degl'Innocenti E, Le Franc Y, Hugo W, Jeffery K, Weiland C, Wittenburg P, Zwolf CM (2021) SEMAF: A Proposal for a Flexible Semantic Mapping Framework. Zenodo https://doi.org/10.5281/zenodo.4651420
- Buttigieg PL, Pafilis E, Lewis S, Schildhauer M, Walls R, Mungall C (2016) The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. Journal of Biomedical Semantics 7 (1). https://doi.org/10.1186/s13326-016-0097-6
- De Smedt K, Koureas D, Wittenburg P (2020) FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. Publications 8 (2). https://doi.org/10.3390/publications8020021
- Díaz S, Purvis A, Cornelissen JC, Mace G, Donoghue M, Ewers R, Jordano P, Pearse W (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. Ecology and Evolution 3 (9): 2958-2975. https://doi.org/10.1002/ece3.601
- DiGiuseppe N, Pouchard L, Noy N (2014) SWEET ontology coverage for earth system sciences. Earth Science Informatics 7 (4): 249-264. https://doi.org/10.1007/s12145-013-0143-1

- Jacobsen A, de Miranda Azevedo R, Juty N, al. e, Guizzardi G, Hansen KK, Hasnain A, Hettne K, Heringa J, Hooft RW, Imming M, Jeffery K, Kaliyaperumal R, Kersloot M, Kirkpatrick C, Kuhn T, Labastida I, Magagna B, McQuilton P, Meyers N, Montesanti A, van Reisen M, Rocca-Serra P, Pergl R, Sansone S, da Silva Santos LOB, Schneider J, Strawn G, Thompson M, Waagmeester A, Weigel T, Wilkinson M, Willighagen E, Wittenburg P, Roos M, Mons B, Schultes E, et al. (2020) FAIR Principles: Interpretations and Implementation Considerations. Data Intelligence 2: 10-29. https://doi.org/10.1162/dint_r_00024
- Karam N, Khiat A, Algergawy A, Sattler M, Weiland C, Schmidt M (2020) Matching biodiversity and ecology ontologies: challenges and evaluation results. The Knowledge Engineering Review 35 https://doi.org/10.1017/s0269888920000132
- Kulmanov M, Smaili FZ, Gao X, Hoehndorf R (2020) Semantic similarity and machine learning with ontologies. Briefings in Bioinformatics 22 (4). https://doi.org/10.1093/bib/bbaa199
- Matentzoglu N, Balhoff JP, Bello SM, Bizon C, Brush M, Callahan TJ, Chute CG, Duncan WD, Evelo CT, Gabriel D, Graybeal J, Gray A, Gyori BM, Haendel M, Harmse H, Harris NL, Harrow I, Hegde HB, Hoyt AL, Hoyt CT, Jiao D, Jiménez-Ruiz E, Jupp S, Kim H, Koehler S, Liener T, Long Q, Malone J, McLaughlin JA, McMurry JA, Moxon S, Munoz-Torres MC, Osumi-Sutherland D, Overton JA, Peters B, Putman T, Queralt-Rosinach N, Shefchek K, Solbrig H, Thessen A, Tudorache T, Vasilevsky N, Wagner AH, Mungall CJ (2022) A Simple Standard for Sharing Ontological Mappings (SSSOM). Database 2022 https://doi.org/10.1093/database/baac035