BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e176413
doi: 10.3897/biss.9.176413

OPEN ACCESS

Conference Abstract

Taxonomy OWLizer: A Tool for Converting Taxonomic Data into OWL

Filipi Miranda Soares[‡,§], Antonio Mauro Saraiva[|], Luís Ferreira Pires[§], Luiz Olavo Bonino da Silva Santos[§], Debora Pignatari Drucker[¶]

‡ MISTEA INRAE, Montpellier, France
§ University of Twente, Enschede, Netherlands
| Universidade de São Paulo, São Paulo, Brazil
¶ Embrapa, Campinas, Brazil

Corresponding author: Filipi Miranda Soares (filipisoares@usp.br)

Received: 29 Oct 2025 | Published: 20 Nov 2025

Citation: Soares F, Saraiva A, Ferreira Pires L, Santos LBdaS, Drucker D (2025) Taxonomy OWLizer: A Tool for Converting Taxonomic Data into OWL. Biodiversity Information Science and Standards 9: e176413.
https://doi.org/10.3897/biss.9.176413

Abstract

The growing demand for high-quality, interoperable biodiversity data exposes the current challenges of standardizing taxonomic information across platforms. While the Global Biodiversity Information Facility Backbone Taxonomy (GBIF Secretariat 2023) offers a valuable reference for species names, translating this information into formal ontology structures for semantic integration remains a complex task. A key challenge lies in the dynamic nature of taxonomic knowledge: as GBIF-BT is continuously updated, ontologies built manually from its data risk becoming outdated or misaligned. Without automated mechanisms for synchronizing with these updates, such ontologies may fail to reflect current taxonomic consensus, limiting their interoperability and long-term usefulness in biodiversity informatics.

To address this problem, we developed the Taxonomy OWLizer (TOWLizer), a lightweight web application that allows users to convert species names into Web Ontology Language (OWL) ontologies based on taxonomic data retrieved directly from the GBIF Species API, as detailed in Fig. 1. Users input one or multiple scientific names, and the application fetches, organizes, and outputs a structured OWL file, facilitating taxonomic integration into semantic web projects. OWLizer automates synonym handling (see Fig. 2) and

taxonomic hierarchy generation, and reuses GBIF URIs in the code, following linked data principles (Berners-Lee 2006). It aims to lower the technical barriers for researchers working in biodiversity informatics who need machine-actionable taxonomies but may lack programming expertise.

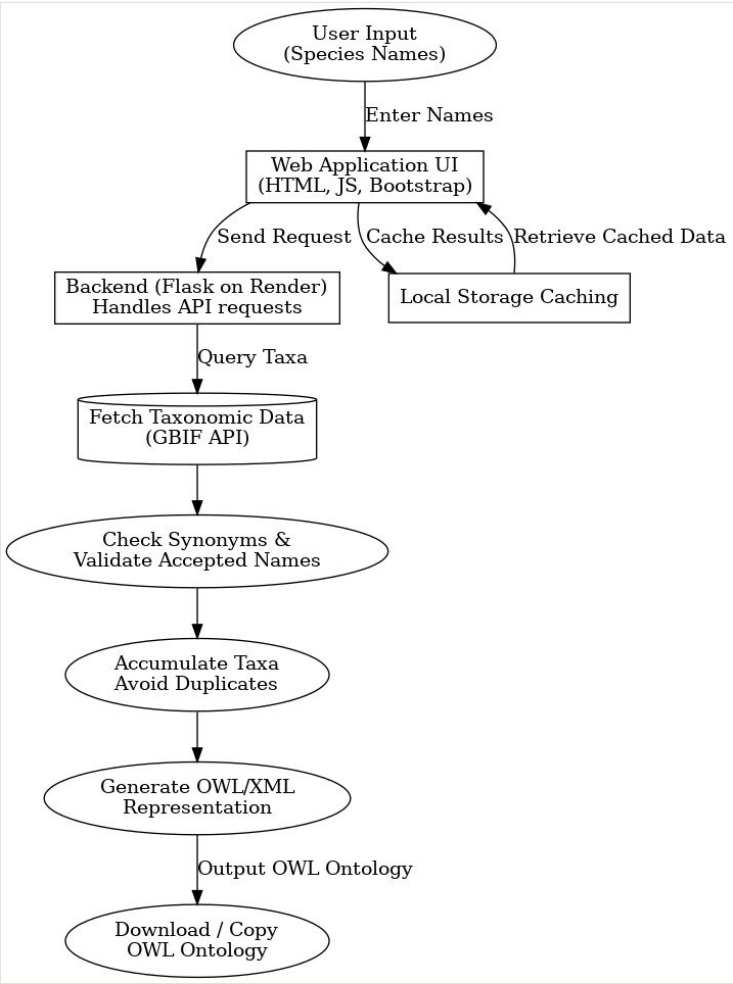

Figure 1.

Taxonomy OWLizer workflow: the user interface submits one or more species names to a Flask backend on Render, which queries the GBIF API for taxonomy. The backend validates accepted names/synonyms, deduplicates taxa, and outputs OWL/XML. The browser caches prior queries locally for faster reuse on the same device. Source: Filipi Soares, Antonio Saraiva, Luís Ferreira Pires, Luiz Santos, and Debora Drucker (2025). License: CC-By 4.0.

The application was implemented using HTML, JavaScript, Bootstrap, and a Flask backend, with ChatGPT-4 providing support for code drafting and debugging, which accelerated prototyping and streamlined development. The tool is freely available

through a GitHub-hosted interface with the backend deployed on Render, and a backup archived in Zenodo (Soares 2025).

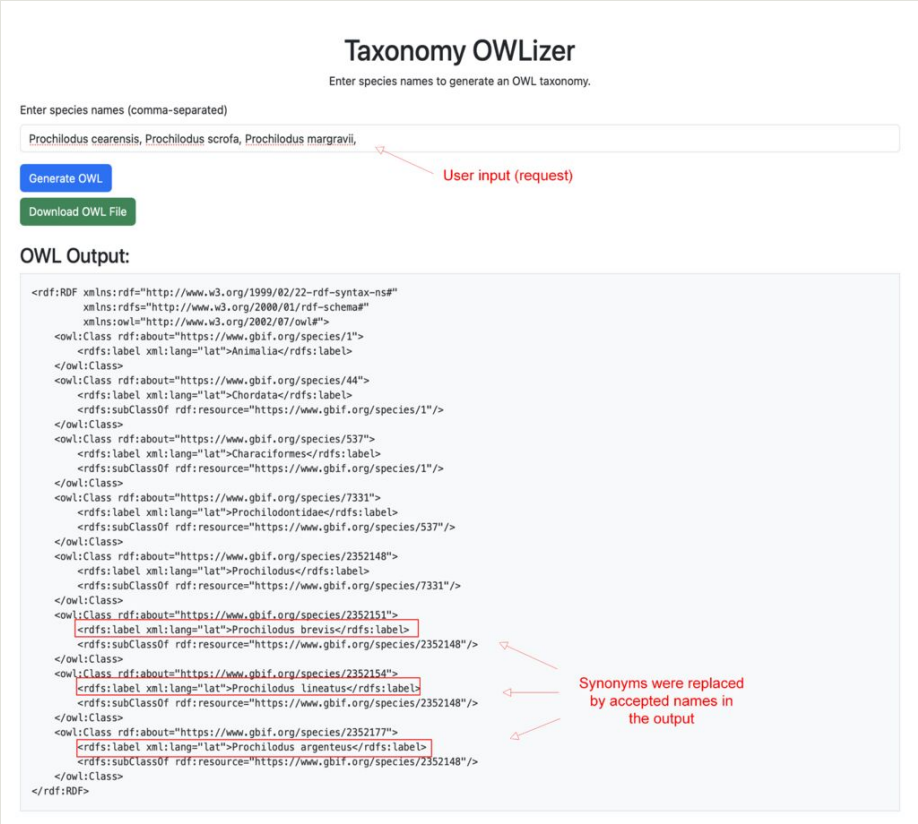

Figure 2.

TOWLizer usage example. In this example, a request was submitted with the names *Prochilodus cearensis*, *Prochilodus scrofa*, and *Prochilodus margravii*, which are known synonyms. The figure shows that the tool correctly identified the valid names and included *Prochilodus brevis*, *Prochilodus lineatus*, and *Prochilodus argenteus* in the generated ontology file. Source: Filipi Soares, Antonio Saraiva, Luís Ferreira Pires, Luiz Santos, and Debora Drucker (2025). License: CC-By 4.0.

TOWLizer has some limitations though. Its performance depends on free-tier hosting services: the backend may enter a sleep state after periods of inactivity, causing occasional delays, while caching is limited to local storage on a single device and browser. The algorithm is also sensitive to typographical errors in scientific names; a recommended workflow is to combine ChatGPT for spelling verification with TOWLizer for synonym management. Finally, GBIF URLs used in the application do not yet support content negotiation, which constrains their reusability in some semantic web contexts. We tested TOWLizer in the development of a real-world ontology, namely the Agricultural Product Types Ontology (APTO), which was designed to represent agricultural

commodities in Brazil. In this context, we addressed the content negotiation issue by replacing the GBIF URLs with URIs from the APTO namespace. These URIs are served via AgroPortal (Jonquet et al. 2018), enabling proper content negotiation.

Despite these constraints, the Taxonomy OWLizer has proven to be a practical and accessible way to generate OWL representations of taxonomic data. It provides an initial step toward automating taxonomic ontology construction, highlights the potential of AI-assisted development, and contributes to ongoing discussions in biodiversity informatics about sustainable, interoperable, and machine-actionable taxonomies. More details on the tool development are provided in Soares et al. 2025.

Keywords

ontology, semantic web, RDF, ChatGPT

Presenting author

Filipi Miranda Soares

Presented at

Living Data 2025

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Berners-Lee T (2006) Linked Data. www.w3.org/DesignIssues/LinkedData.html. Accessed on: 2025-10-02.
- GBIF Secretariat (2023) GBIF Backbone Taxonomy. GBIF. Release date: 2023-8-28. URL: https://doi.org/10.15468/39omei
- Jonquet C, Toulet A, Arnaud E, et al. (2018) AgroPortal: A vocabulary and ontology repository for agronomy. Computers and Electronics in Agriculture 144: 126-143. https://doi.org/10.1016/j.compag.2017.10.012
- Soares FM (2025) Taxonomy OWLizer. v.2. Zenodo. Release date: 2025-3-06. URL: https://doi.org/10.5281/zenodo.14984136
- Soares FM, Saraiva A, Ferreira Pires L, et al. (2025) Exploring a Large Language Model for Transforming Taxonomic Data into OWL: Lessons Learned and Implications for Ontology Development. Data Intelligence 7 (2): 265-302. https://doi.org/10.3724/2096-7004.di.2025.0020