

I. Title

Linking Biodiversity Literature and Data: Persistent Identifier Exchange Between BHL and Wikidata

II. Abstract

The Biodiversity Heritage Library (BHL) is a global consortium of natural history and botanical institutions that cooperate to digitize and share biodiversity literature. Wikidata, a free, collaborative, multilingual knowledge base, has become a central hub for persistent identifiers (PIDs) across diverse domains, many of which link directly to BHL resources.

BHL is now leveraging the structured data curated by Wikidata contributors to enhance its own metadata. This presentation will describe an automated pipeline developed to extract and align PIDs from Wikidata with existing BHL records. The process involves tailored queries, automated scripts, and data formats designed for reliable integration. These tools are open, adaptable, and intended to support similar efforts by other organizations seeking to enrich their data with PIDs extracted from Wikidata.

Equally important is the bi-directional flow of information: BHL also contributes structured metadata back to Wikidata volunteers, closing the loop and enabling a collaborative exchange of biodiversity knowledge. This round-tripping of data strengthens the overall biodiversity knowledge graph by connecting literature and associated identifiers across platforms.

Sharing identifiers between BHL and Wikidata offers multiple benefits. It enhances the discoverability of biodiversity literature, improves entity disambiguation, and builds stronger links between BHL and other infrastructures that rely on shared identifiers. It also reinforces Wikidata's role as a PID hub within the global biodiversity informatics landscape.

This talk will highlight how the BHL-Wikidata collaboration demonstrates the utility of Wikidata as a source of persistent identifiers, and we hope to encourage use of Wikidata as a tool for connecting biodiversity infrastructures and making knowledge more open, structured, and accessible.

III. Session

6787919 Wikimedia and Biodiversity Data: A Mutualistic Relationship in the Open Knowledge Ecosystem

Authors and affiliations

Lichtenberg, Mike

Biodiversity Heritage Library, St. Louis, United States

ORCID ID: <https://orcid.org/0009-0001-0220-5275>