Conference Abstract

Improving Search Efficiency in the Biodiversity Heritage Library Corpus

Carolyn A. Sheffield ‡

‡ Smithsonian Libraries/Biodiversity Heritage Library, Washington, DC, United States of America

Abstract

Biodiversity literature and archival collections are not only indispensable in taxonomic research, they provide crucial information for understanding of museums' natural history collections. Literature and archives document collecting events resulting in specimen collections, contain original descriptions based on those specimens, and provide a wealth of other contextual information for the study of life on earth. The Biodiversity Heritage Library is committed to improving research efficiency by providing open access to a growing body of biodiversity literature and archives. While descriptive metadata is widely available for both specimen collections (i.e., DarwinCore) and literature (i.e., MARCXML), connections between the two collection types cannot generally be found at these descriptive levels thus hindering efficient discovery of relevant materials. The integration of name finding services, powered by Global Names Architecture, provides a significant value-add through page-level access to mentions of a given taxon name. Yet how might one search based on a museum code, a common name, or a place name? This presentation will share how BHL's top technical priorities for 2018 will help facilitate more efficient searching and discovery of information in the pages of the BHL corpus. Specifically, updates on BHL's top two priorities – implementation of full text search and incorporation of available crowdsourced transcriptions—will be covered.

Keywords

Presenting author

Carolyn A. Sheffield