

Collections do not have to Remain Ambiguous

Forever: Seven steps to getting the correct people into your data

Quentin Groom[‡], Christian Bräuchler[§], Robert W. N. Cubey[‡], Mathias Dillen[‡], Pieter Huybrechts[‡], Nicole Kearney[¶], Siobhan Leachman[#], Deborah L Paul^{‡,«}, Heather Rogers[»], Joaquim Santos[^], David Peter Shorthouse[∨], Alison Vaughan[‡], Sabine von Mering[?], Elspeth M Haston[‡]

[‡] Meise Botanic Garden, Meise, Belgium

[§] Naturhistorisches Museum Wien, Wien, Austria

[|] Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom

[¶] Biodiversity Heritage Library (BHL) Australia, Melbourne, Australia

[#] Independent Researcher, Wellington, New Zealand

[‡] University of Illinois Urbana-Champaign, Urbana, United States of America

[«] Florida State University, Tallahassee, United States of America

[»] McGill University, Montreal, Canada

[^] Centre for Functional Ecology, Coimbra, Portugal

[∨] Agriculture & Agri-Food Canada, Ottawa, Canada

[|] Royal Botanic Gardens Victoria, Melbourne, Australia

[?] Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

Corresponding author: Quentin Groom (quentin.groom@plantentuinmeise.be)

Abstract

People are involved with the collection and curation of all biodiversity data, whether they are researchers, members of the public, taxonomists, conservationists, collection managers or wildlife managers. Knowing who those people are and connecting their biographical information to the biodiversity data they collect helps us contextualise their scientific work. We are particularly concerned with those people and communities involved in the collection and identification of biological specimens. People from herbaria and natural science museums have been collecting and preserving specimens from all over the world for more than 200 years. The problem is that many of these people are only known by unstandardized names written on specimen labels, often with only initials and without any biographical information. The process of identifying and linking individuals to their biographies enables us to improve the quality of the data held by collections while also quantifying the contributions of the often underappreciated people who collected and identified these specimens. This process improves our understanding of the history of collecting, and addresses current and future needs for maintaining the provenance of specimens so as to comply with national and international practices and regulations.

In this talk we will outline the steps that collection managers, data scientists, curators, software engineers, and collectors can take to work towards fully disambiguated

collections. With examples, we can show how they can use these data to help them in their work, in the evaluation of their collections, and in measuring the impact of individuals and organisations, local to global.

Keywords

disambiguate, herbarium, museum, collection management system, ORCID

Presenting author

Quentin Groom

Presented at

TDWG 2022

Acknowledgements

This paper is a product of the People in Biodiversity Data task group of the [Biodiversity Information Standards \(TDWG\)](#) organisation.

Funding program

This work was supported by European Cooperation in Science and Technology (COST) as part of the Mobilise Action CA17106 on Mobilising Data, Experts and Policies in Scientific Collections; SYNTHESYS+ a Research and Innovation Action (Grant agreement 823827) and DiSSCo Prepare a Coordination and Support Action (Grant Agreement 871043), both funded by the Horizon 2020 Framework Programme of the European Union. This work was also facilitated by the Research Foundation – Flanders (FWO) research infrastructure under grant number I001721N. Additional support provided by the National Science Foundation (NSF) grant number #2033973.

Conflicts of interest