Planning the Migration to a New Database: Implications for the Collections of Meise Botanic Garden

Henry Riley Engledow‡, Ann Bogaerts‡, Sofie De Smedt‡, Wesley Tack‡

‡ Meise Botanic Garden, Meise, Belgium

Corresponding author: Henry Riley Engledow (henry.engledow@plantentuinmeise.be), Ann Bogaerts (ann.bogaerts@plantentuinmeise.be), Sofie De Smedt (sofie.desmedt@plantentuinmeise.be), Wesley Tack (wesley.tack@plantentuinmeise.be)

Abstract

Established in 1796, the Meise Botanic Garden has undergone multiple changes in nationality, name, focus, and locations. These changes have impacted its herbarium, living collections, and associated data. The way specimens are filed in the herbarium varies according to specimen type, taxonomic group and historical importance. Pragmatic solutions were used in the past for dealing with the divergent collections and their users' specific needs. There are at least 15 filing systems used at the Botanic Garden at present. In 2015, funding facilitated the digitization of 1.2 million herbarium specimens. A follow-on project in 2018 digitzed another 1.4 million specimens. Together with other initiatives, like the Global Plants Initiative, we have more than 2.7 million imaged specimens, representing about 77% of our herbarium collections. Of the imaged specimens, 89% have minimal data (barcode, species name, collector(s), collector number, collection date, country and locality), while the remaining 11% still need the metadata captured from the scanned images (mostly being done through the crowdsourcing project DoeDat).

In 2023, a call to tender was started for a new Collection Management System (CMS) as part of the Distributed System of Scientific Collections (DiSSCo) Flanders*[1] project. This call for tender was compiled in consultation with various natural history institutions, botanic gardens and museums in Belgium. At present, we are in the process of evaluating the various tenders. As part of the preparation to migrate the data from our present database to the new database, data gaps are being filled and data is being cleaned where possible. This will facilitate the migration from one system to another. This is particularly important for fields relating to taxonomy, countries and collectors, as these have implications for the physical location of specimens in the herbarium. In the past months, much time has been dedicated to linking specimens to the taxon table. At present fewer than 0.6% of specimens with a filing name (the scientific name under which the specimen is filed in the herbarium cupboards) remain to be linked to the taxon table.

The use of the Taxonomic Name Resolution Service helped to resolve some of the issues associated with unlinked names. Many of the remaining unlinked names are names of infraspecific taxa (which are not yet in the taxon table), orthographic variations, or names of commercial cultivars or hybrids. However, this does not include the 11% of the specimens without data mentioned above. The number of specimens lacking a country code is 7%. The African Herbarium, which represents about a third of the specimens at the Botanic Garden, is sorted according to species, country and collector. Resolving missing country codes would facilitate the curatorial process for this collection. Likewise, 6% of specimen data only have collectors as free-text that still needs to be linked to the standardised collectors table. Atomizing and linking these collector strings would also ease the migration to a new CMS.

In 2015, a master plan was also drawn up for the physical renovation of the Garden, including the herbarium's infrastructure. Work on the herbarium is planned for 2025, but is contingent on funding. The CMS will form a crucial part of this move as specimens will be reclassified. As is mentioned above, there are several classification systems used at the Botanic Garden. The African and Belgian collections currently follow Mabberley (1997) for angiosperms. This system is no longer up-to-date, and we are considering moving to APG IV (Angiosperm Phylogeny Group) system for flowering plants. For example, the 213 genera present at the Botanic Garden for the family Liliaceae *sensu lato* (Mabberley 1997) would resolve to 24 families in APG IV. As the specimens are arranged alphabetically by family and genus, this would mean a significant physical reorganization of specimens. The CMS will also be used to track the movement of the various collections. Once the new CMS has been chosen, we will have a better idea of what needs to be done in order to facilitate data migration. For example, linking our taxonomic names to external sources using their permanent identifiers (e.g., World Flora Online, International Plant Names Index, Global Biodiversity Information Facility, Index Fungorum, MycoBank, World Register of Marine Species) would allow us to track and maintain taxonomic changes more easily in the future. Importing the various permanent identifiers into our current database would be prone to inconsistencies as the existing fields are not fit for the purpose. A possible solution would be to manage this data externally and import them into the new CMS. A similar approach could be followed with collectors and Wikidata permanent identifiers. We hope to develop a pragmatic approach with practical solutions.

Keywords

mass digitisation, data cleaning, curatorial tool

Presenting author

Henry Riley Engledow

Presented at

TDWG 2023

Funding program

DiSSCo Flanders

Hosting institution

Meise Botanic Garden

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Mabberley DJ (1997) The Plant-book: A Portable Dictionary of the Vascular Plants. 2nd. Cambridge University Press, New York, 858 pp. [ISBN 0-521-41421-0]

Endnotes

*1 https://dissco-flanders.be/