Herbaria Heritage:

Visualising Colonial Bias in Natural History Collections

Sakura Morales Furuta*[1], Ana Reviejo Salamanca*[2], Michaela Todd*[1], Lise Stork[2], Andreas Weber[3]

[1] Leiden University, Postbus 9500, 2300 RA Leiden, The Netherlands
sakuramoralesf@gmail.com, mickeyirenetodd@gmail.com

[2] Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
a.reviejosalamanca@student.vu.nl, l.stork@vu.nl

[3] University of Twente, BMS-KiTeS, 7500 AE Enschede, The Netherlands
a.weber@utwente.nl

* Sakura Morales Furuta, Ana Reviejo Salamanca and Michaela Todd share the first authorship of this paper.

Abstract. As a result of the colonial entanglements of many natural history collections, much of the world's biodiversity heritage is housed in Europe. Increasingly, natural history institutions have started to address this history. However, computational methods for analysing large collections often consist in static visualisations of collection provenance. We argue that interactive visualisations allow users to understand collections better: their content gaps as well as interesting patterns and trends. Using a dataset containing metadata of five million entries from the botanical collection of *Naturalis Biodiversity Center*, we created an interactive visualisation with Microsoft PowerBI. The visualisation depicts colonial origins and the botany collection's movement to the Netherlands over time on an interactive map and timeline. This project thus addresses a gap in historical research on the colonial legacy of Dutch botanical collections, and also a gap in computer science research regarding computational techniques for distant reading of natural heritage data. Our interactive visualisation increases the accessibility of the available scientific data. It also contributes to understanding the cultural history of natural history collections and ultimately, highlights the importance of accurate and accessible visual representations of biodiversity collection histories. This project suggests a way forward for natural history museums grappling with their colonial past.

Introduction

Natural History museums and herbaria in the Global North house much of the world's biodiversity heritage (NatSCA, 2020). However, these institutions often overlook their colonial entanglements, distancing themselves from cultural histories (Rijksmuseum & Slavery, 2024). Acknowledging this entanglement is crucial to unveil the complex political context, and ensure it is not obscured by scientific rationality and digitisation (Keiser, et al. 2023). This recognition becomes vital for understanding the impact on global biodiversity and conservation efforts (Park et al., 2023).

When addressing colonial histories of natural history collections in the Dutch context, researchers have predominantly used textual or object-focused approaches, lacking computational methods and visualisations (e.g. Drieënhuizen & Sysling, 2021; Van Wingerden, 2023). Some studies by Mohammed et al. (2022) and Park et al. (2023) utilised heat maps, world maps, etc. to illustrate specimen movements and categorise trends. However, we argue that existing visualisations, like herbaria snapshots, lack interactivity.

Inspired by these studies, we describe a simple method to create an *interactive* visualisation through a digital database, focusing on the provenance of botanical specimens, specifically those stored in Naturalis Biodiversity Center (Leiden, Netherlands). The interactive and multi-layered visualisation provides a starting point to reveal hidden histories behind the botanical collection, bridging cultural history and digital representations of Dutch biodiversity heritage. Our paper comprises three parts: we first explain the dataset, then our interactive visualisation and lastly, the obtained results and evaluation. In a final section we reflect upon our methodology, suggesting follow-up work.

Dataset and Preprocessing

We use an openly available dataset that comprises metadata of more than five million plant specimens belonging to 1,122 distinct families, collected in 259 different countries by more than 101,100 naturalists from 1550 to 2023. The dataset can be downloaded from the Global Biodiversity Information Facility (GBIF) (Creuwels 2023). The physical plant specimens that this dataset represents are stored in the depots of Museum Naturalis in Leiden.

Out of these specimens, 88,740 are type specimens, collected in 231 countries by 7,239 naturalists. Generally, a type specimen is a name-bearing specimen that taxonomists use as a reference for new species, therefore of scientifically high value. Metadata for each collected specimen is organised in terms of the Darwin Core vocabulary (Figure 1), consisting of terms such as *dwc:scientificName*, which refers to a specimen's full scientific name, or *dwc:eventDate*, the date at which the specimen was recorded at Naturalis. We preprocessed the dataset by splitting the *dwc:eventDate* values when two dates were given (a start and end-dates, likely of the month in which the specimen was brought in), and kept only start-dates. As a final preprocessing step, we filtered out rows with empty, null and "?" values.
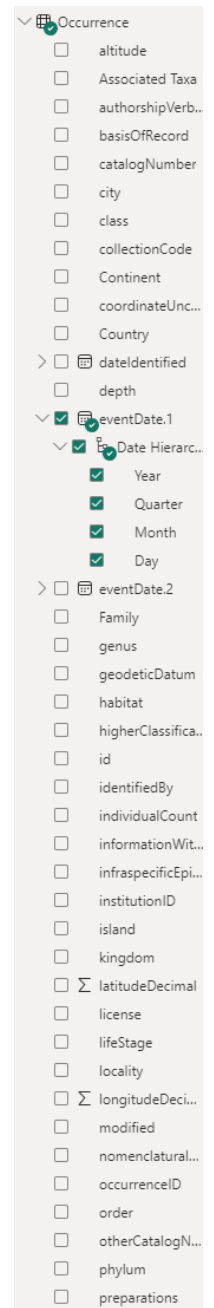

Figure 1: Dataset's columns loaded into Power BI, focusing on the *eventDate1* column.

From a Dataset to a Visualisation of Colonial Bias

For our visualisation we used PowerBI Desktop, which allows interactive maps, timelines and filters based on database column headers. As filters we used the Darwin Core terms *dwc:country, dwc:recordedBy, dwc:continent, dwc:family and dwc:typeStatus,* and plotted specimen occurrences on a map, visualising the number of specimens originating from specific countries. Size was used to show quantitative differences: the larger the circle, the more specimens originate from that country. The interface also shows the total number of specimens, countries of origin, and a pie chart with the percentage of specimens per country. Importantly, there is a timeline plotting the number of specimens registered over time (see Figure 2). Or put differently, we represented the data in a way that shows how the collection's colonial bias evolved over time and space.
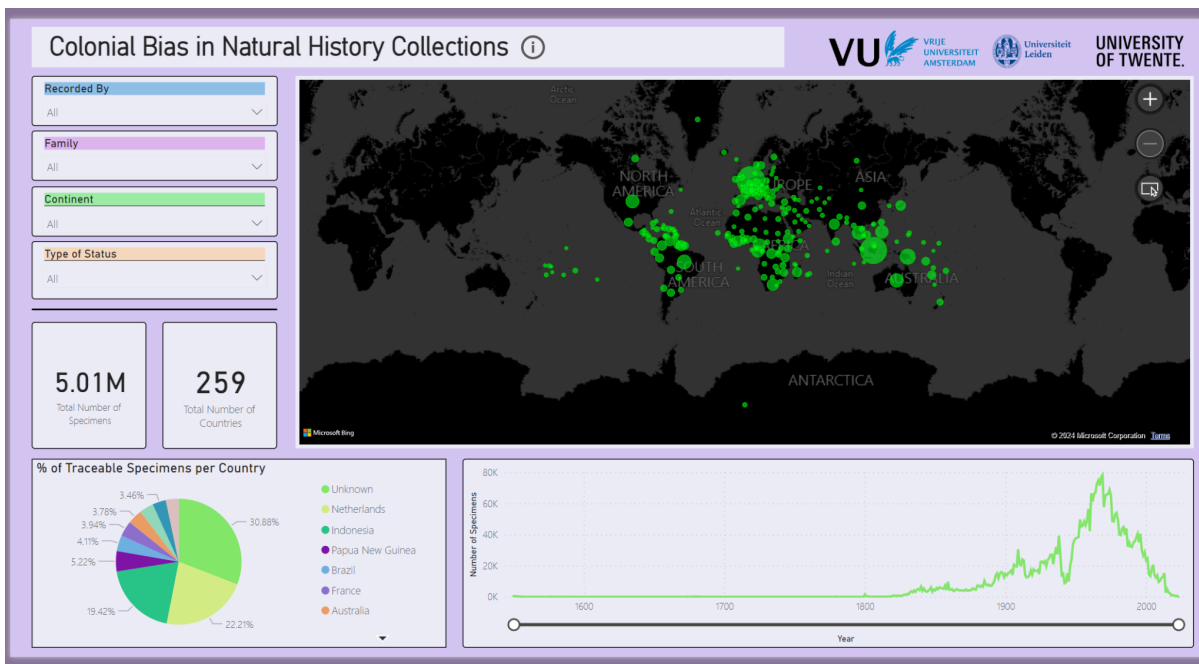

Figure 2: Main interface of the visualisation, featuring the filters, pie chart, map and timeline, among others.

The interface is aimed at historians, scientists, and the general public. We prioritised readability and simplicity, but ensured that a detailed investigation can be conducted through filters (top left on Figure 2). A filter example can be seen in Figure 3, with the *Type of Status* filter. Users are able to filter by each separate type to carry out in-depth quantitative analysis of the specimen of interest. The *recordedBy* filter lets the user filter by the botanist that obtained and recorded the specimen for the collection. This directly contributes to questions about the colonial links of collections, as many collectors and botanists (e.g. Carl Ludwig Blume, Pieter Willem Korthals, Jaheri, Kees van Steenis) were closely affiliated with colonial institutions such as the botanical garden in Bogor, Indonesia (Goss 2018 and 2023; Manse, 2013; Wille 2019; Weber 2018).
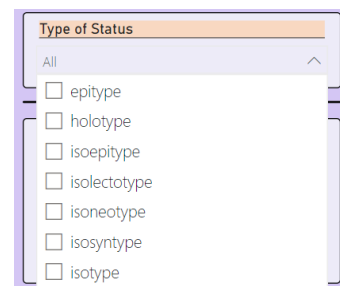

Figure 3: *Type of Status* filter.

First Results and Evaluation

Looking at the map and pie chart in Figure 3, one can see a significant clustering of specimens obtained from South-East Asia; most notably Indonesia. It comprises 19.42% of the total dataset (not accounting for specimens for which location data was unknown). More importantly, according to the dataset, 70,23% of all holotype specimens in the Naturalis plant collection (=14.580) stem from Indonesia, Papua New Guinea and Surinam. A holotype is a physical plant specimen that was used for the formal description of a plant species. The unknown cases also remain relevant: (i) they provide museums with a starting point for improving their datasets by highlighting data gaps, and (ii) could be a signal: such specimens could have come from former Dutch colonies, or perhaps locations that museums prefer not to publicise.

Regarding reusability, the PowerBI file is easily shareable, and can be used by researchers on their own data, specifically when published as a Darwin Core Archive. Furthermore, the visualisation is modular: chosen filters can be modified at any time. Through the visualisations' malleability, the dashboard can be adapted to fit specific users, and their historical or scientific questions.

Conclusion

Our interactive tool links the Netherlands' colonial past with Naturalis' scientific plant collection through multi-layered visualisations. The tool facilitates a comprehensive exploration of natural history collections through an interactive map, timeline and metadata filters, enabling users to investigate questions about the provenance of specific objects. The visualisation tool enhances the collections' accessibility through visual reading, with the potential to reveal insights otherwise hidden and uncover data-related concerns. The tool thus aids biologists and those interested in tracing the provenance, geographical spread of natural objects in collections and the ethics of obtaining these specimens in the context of colonial power asymmetries.

Ashby, J. (2020). *How do you do decolonial research in natural history museums?* https://natsca.blog/2023/10/26/how-do-you-do-decolonial-research-in-natural-history-museums/

Creuwels, J. (2015). Naturalis Biodiversity Center (NL) - Botany Leiden. Naturalis Biodiversity Center. Occurrence dataset https://doi.org/10.15468/pal1lb.

Drieënhuizen, C., and Sysling, F. Java Man and the politics of natural history: An object biography. *Bijdragen Tot de Taal-, Land- En Volkenkunde* 177, 290-311 (2021). https://doi.org/10.1163/22134379-bja10012

Goss, A., Reinventing the Kebun Raya in the New Republic: Scientific research at the Kebun Bogor botanical gardens in the age of decolonization. *Studium. Tijdschrift voor Wetenschaps- en Universiteitsgeschiedenis* 11, 206-219 (2018).

Goss, A., Decolonizing botany: Indonesia, UNESCO, and the Making of Global Science. *Journal of the history of biology* 56, 495-523 (2023).

Kaiser, K., Heumann, I., Nadim, T., et al. Promises of mass digitisation and the colonial realities of natural history collections. *Journal of Natural Science Collections* 11, 13-25 (2023). http://www.natsca.org/article/2796

Manse, M. Kennis is macht: de veelzijdige expedities van botanicus Pieter Willem Korthals (1807-1892). *Studium. Tijdschrift voor Wetenschaps- en Universiteitsgeschiedenis* 6, 42-50 (2013).

Mohammed, et al. (2022). Colonial Legacies Influence Biodiversity lessons: How past trade routes and power dynamics shape present-day scientific research and professional opportunities for Caribbean scientists. *The American Naturalist*, 200(1), pp. 140–155. https://doi.org/10.1086/720154

Park, D.S., Feng, X., Akiyama, S. et al. The colonial legacy of herbaria. *Nature Human Behaviour* **7**, 1059–1068 (2023). https://doi.org/10.1038/s41562-023-01616-7

Weber, A. A garden as a niche: Botany and imperial politics in the early nineteenth century Dutch empire. *Studium. Tijdschrift voor Wetenschaps- en Universiteitsgeschiedenis* 11, 178-190 (2018).

Wille, R.-J. *Mannen van de mikroskoop. De laboratoriumbiologie op veldtocht in Nederland en Indië, 1840-1910* (2019).

Wingerden, P. van, Layered loyalties: the Natuurkundige Commissie in the Netherlands Indies (1820-1850) PhD thesis, Leiden University (2023), https://hdl.handle.net/1887/3643457

Link to the utilised Dataset: Naturalis Biodiversity Center (NL) - Botany

Demo video: 📄 Demo Video Visualising Colonial Bias.mp4