NATURAL HISTORY COLLECTIONS AND MUSEOMICS

Natural History Collections and Museomics 2: 1–20 (2025)
DOI: 10.3897/nhcm.2.148855

Research Article

Collectomics – towards a new framework to integrate museum collections to address global challenges

Julia D. Sigwart[1,2], Matthias Schleuning[3], Angelika Brandt[1,2], Markus Pfenninger[3], Hanieh Saeedi[1], Thomas Borsch[4], Eva Häffner[4], Robert Lücking[4], Anton Güntsch[4], Helmuth Trischler[5], Till Töpfer[6], Karsten Wesche[7,8], Collectomics Consortium[9]

1 Senckenberg Research Institute and Natural History Museum, Department of Marine Zoology, Frankfurt am Main, Germany
2 Institute for Ecology, Evolution and Diversity, Goethe-University of Frankfurt, Frankfurt am Main, Germany
3 Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, Germany
4 Botanischer Garten und Botanisches Museum Berlin, Freie Universität Berlin, Berlin, Germany
5 Deutsches Museum, Munich, Germany
6 Leibniz Institute for the Analysis of Biodiversity Change, Bonn and Hamburg, Germany
7 Senckenberg Museum of Natural History Görlitz, Görlitz, Germany
8 Dresden University of Technology, Zittau, Germany
9 Collectomics Consortium, Frankfurt, Germany

Corresponding author: Julia D. Sigwart (julia.sigwart@senckenberg.de)

Abstract

Collections' digitisation is a priority in many natural history collections, and publicly available datasets are expanding rapidly. The potential value of collections remains largely untapped even in modern research, because the vast scope of collections dwarfs current efforts at data mobilisation. Collections are continually expanding, and there are an estimated 3 billion undigitised specimen records worldwide. In this review, we use a simple model to illustrate that current efforts at global digitisation will not succeed until the late 21st century at the earliest, unless new technologies are harnessed and commitments by funding bodies and society are made. As we advance toward specimen digitisation, an equally important consideration is that the majority of these digital records only represent a fraction of the information potentially available from the collection objects. The term "collectomics" was coined in discussions within the Senckenberg institution as a phrase for digital frameworks that embrace all current and future data and knowledge derived from specimens. This expands on the concept of museomics, which was originally defined to focus on molecular data generated from museum specimens. Rooted in the concept of the extended specimen, collectomics encompasses metadata, images, traits, DNA, and further data extracted in the future with yet unknown applications, all of which are connected to environmental data and other historical contextual information. Thus, a view of digitisation under the collectomics concept is not limited to natural history collections but directly integrates evolutionary, ecosystem and social sciences, including the human contributions of collectors, donors, and researchers in the past and future. A "collectomics" view envisions seamless integration of multidimensional specimen-based data, with interoperability among historical, artistic, ethnographic, and natural history collections, to generate knowledge that is needed to tackle global challenges.

Key words: Collectiomics, digitisation, FAIR, museum collections

OPEN ACCESS

Academic editor: Franco Andreone
Received: 5 February 2025
Accepted: 18 March 2025
Published: 28 March 2025

ZooBank: https://zoobank.org/E0F93BE8-2B9F-42F6-9AD0-4E44E3512389

Citation: Sigwart JD, Schleuning M, Brandt A, Pfenninger M, Saeedi H, Borsch T, Häffner E, Lücking R, Güntsch A, Trischler H, Töpfer T, Wesche K, Collectomics Consortium (2025) Collectomics – towards a new framework to integrate museum collections to address global challenges. Natural History Collections and Museomics 2: 1–20. https://doi.org/10.3897/nhcm.2.148855

Introduction

Research collections represent key scientific infrastructure for biology and they constitute some of the oldest continuously maintained and expanded scientific resources on Earth. Climate change and other human-driven transformations to our planet are impacting global biodiversity at an unprecedented speed and extent, with cascading impacts on our livelihoods (Díaz et al. 2019). Solutions for modern social-ecological problems will require perspectives that bridge the traditional divide between natural sciences and humanities. This challenge requires integrative research across space, time, methods, disciplines, research fields, and lines of evidence (Sanders et al. 2023).

Scientific advice on safeguarding biological and cultural diversity should rely on all available information, to understand past and current developments of our natural and cultural heritage and to predict future responses of Earth's systems. To assess the impact of ever-accelerating global change, the new situation must be measured against some prior condition. Scientific museum collections with their manifold objects and associated data represent tangible evidence and an irreplaceable historical record of our changing planet. Natural history museums collectively house billions of biological records spanning more than three centuries (Johnson et al. 2023). Large collections in arts, cultures, and technologies also span multiple millennia and number in the millions of artefacts. The majority of these objects are in the purview of institutions in the Global North, but represent global knowledge (Asase et al. 2022; Park et al. 2023).

To realise the potential value of these collections for addressing issues of global change, there is increasing awareness that both the objects and their associated metadata must be made available through digitally accessible, integrated research platforms. Digital access to the authoritative datasets connected to museums empowers local researchers to study local biodiversity trends (Asase et al. 2022). Such platforms need to be designed as knowledge powerhouses of public accessibility. The scope of this effort must also be integrated in a historical and geographical context, which requires interlinking natural history collections with technical and cultural research collections. An integrated library of bio-techno-cultural knowledge would allow a more planetary perspective on diversity (Misof et al. 2023).

Current technological developments provide an unprecedented opportunity to unleash the full potential of collections by fully integrating the myriad data dimensions from collection objects. The term "collectomics" originated in digitisation discussions and internal documents in the Senckenberg institution, and has been used in the literature in the last few years (Sigwart et al. 2021; Misof et al. 2023; Peters et al. 2023; Boudinot et al. 2024), but an explicit and clear definition has not previously been published. Collectomics is not a new way to use collections, it is a succinct term to rally efforts around the many facets of research that could benefit from open, digitally accessible, interlinked collections. It aspires to increase the integration with other diverse types of relevant data, such as social and historical context, or those derived from remote-sensing or large citizen-science initiatives (Spear et al. 2017). Global data mobilization of the authoritative data source—scientific collections—is further important for equitable accessibility for science and society (Sigwart 2018).

Unleashing the potential of collections to address global challenges also depends on increasing accessibility of multidimensional collections data, via their integration across disciplines, and ensuring that FAIR criteria (Findability, Accessibility, Interoperability, and Reusability) for past, present and future collection objects and data are met (Miller et al. 2020). This is possible now, as tremendous developments in digitisation, data standards, and the build-up of global research networks offer unprecedented opportunities to unlock collections and make their information available in integrated research platforms (Hedrick et al. 2020). Rapidly developing applications based on artificial intelligence and machine learning finally promise major advances in mobilising specimen data (Blagoderov et al. 2012; Høye et al. 2021), as well as long-anticipated applications for computer-aided species identification and trait recognition (Cordier et al. 2018; Høye et al. 2021; Kothari et al. 2023; Hofmann et al. 2024). Advances in taxonomy and building global taxonomic backbones at species level are now taking place and will increasingly provide globally consistent species information to be used as references (Hobern et al. 2021). Mass digitisation efforts to unlock the temporal and spatial data attached to physical specimens will provide the main reference for large-scale and long-term studies. What is still needed is investment in scaling up digitisation efforts and in fully connected, interoperable data repositories that can bridge the two ends of the spectrum: from detailed insights derived from single objects to large-scale analyses of ecosystems (Kissling et al. 2018).

Here we briefly review two important factors in planning for a more integrated future for natural history collections with a broad view of their power for confronting global problems. The first issue is the current pace of digitisation of collection metadata, and the need for significant acceleration. The second part is to define collectomics and a holistic vision with specimen objects at the heart of powerful interdisciplinary data approaches.

Natural history collections as one global data bank

Natural history collections have already started to become integrated into globally coherent research platforms, and the "Global Museum" (Bakker et al. 2020) functions as decentralised infrastructure for evolutionary biology, ecosystem sciences and public understanding of science. These collections represent a truly global resource, and cover many centuries of collecting. Meanwhile, much biodiversity research that underpins current policy discussions, such as monitoring programs, does not adequately capture long term trends (Dornelas et al. 2018). Datasets from collections have revealed fundamental mechanisms of Earth-system dynamics, such as mass extinctions and tipping points, latitudinal diversity gradients, and evolution through escalation (Vermeij 1993; Roy et al. 1998; Barthlott et al. 2007). The centuries-long time spans provided by collections data – and millennia and even millions of years in prehistoric or geological collections – are increasingly recognised as the most important sources to provide comparative baselines on historical time scales and in deep time (Meineke et al. 2019; Dominici et al. 2020; Lopez et al. 2020). Since digitised data have much higher visibility than those in un-digitised physical collections, it is important to reiterate that specimens are data even before they are digitised (Bakker et al. 2020); physical voucher specimens represent important records that inform research as they have prior to the digital age.

There are important fundamental differences between a specimen collection, which is an analogue storage system, and a database. A physical, spatial organization of knowledge has an inherently human-centric organisation—specimens are arranged in a way that makes sense to people (by systematics or stratigraphy), and although these relationships are scientific, they rely on human interpretation. Relationships in a database are formalised and rely on predefined rules. (This is a constant source of low-level friction in the experience of collections digitisation, for example where species identifications are ambiguous or uncertain.) A collection occupies physical space, which requires material investment and limits the feasibility of reorganisation. Databases can be dissected, cross-referenced, and reassembled dynamically. Collections depend on a user physically browsing to retrieve information; however, like a physical bookshelf in a library, the information retrieved usually exceeds the boundary of the original question, from observations of the specimen itself or serendipitous information in the context of nearby comparative material. Databases allow structured queries across categories. For example, extracting all records from one geographical location is extremely difficult when the physical collection is organised taxonomically, and close to impossible when the collection is not digitized.

Collections-based research has in part been stymied because data are stored in different formats and databases (Kissling et al. 2018; Gallagher et al. 2020), which limits their inter-operability. The definition and application of broadly defined biodiversity data standards (e.g., Darwin Core and ABCD), have contributed to address and reduce such challenges (Blum et al. 2019). However, the largest sources of biodiversity data are still contained in physical repositories that are not fully accessible due to the lack of electronic records. For more than 200 years, relevant information was most often recorded in the form of hand-written labels and inventories (Fig. 1), posing a challenge to digitisation (Owen et al. 2020). In most museums, mass data capture (metadata, high-throughput methods) and in-depth digitisation (high-end imaging and advanced analytics) is undertaken continuously. However, the focus may shift sporadically, based on available funds, leading to patchy data capture across collections (Borsch et al. 2020). Current approaches usually focus on "priority" collections as discrete projects, but the more urgent aim is the full range of global collections as an integrated whole. Moreover, natural history collections are historical as well as biological archives, and the interaction of biological and historical records represents untapped potential for insights to cultural change.

Digitisation forecasting

Digitisation of collections is chasing a moving target, because data standards and applications are evolving, but most importantly because collections are continuously expanding. Any consideration of digitisation must also include future provisioning; this is not a final solution, full digitisation is a step in improving data infrastructure that must also include commitments to collections maintenance and support of curators and technicians with specialist expertise (Andreone et al. 2024). New data records are usually rapidly available in digital form, but the vast majority of records pre-date digitisation. A recent study of major natural history museum holdings included 27 major European

Figure 1. Top: for more than 200 years, relevant object information was most often recorded in the form of hand-written labels and inventories. Bottom: Natural history museums directly intersect with social sciences, although the connections often go unrecognised. **Top left:** Jan-Peter Kasper/Universität Jena, **Top right:** Sigrid Hof / Senckenberg Research Institute and Museum Frankfurt, **Bottom left:** image of Dr Fritz Haas (seated) and unnamed companions (men and women), in the act of collecting a new species *Unio valentinus*, **Bottom right:** natural history objects also appear in the context of art objects, photo: Emőke Dénes.

institutions with 545.5 million objects among 73 world museums with 1.1 billion records globally (Johnson et al. 2023). The large EU-consortium DiSSCo has recently estimated that its 170 European members collectively host 1.5 billion specimens (https://www.dissco.eu/), three times larger than the major research museums alone; this suggests that globally, museums would hold around 3.2 billion objects. In contrast, the Global Biodiversity Information Facility (GBIF) currently contains ~255 million digital records tied to physical specimens (gbif.org, occurrences with the basis "preserved specimen" on 2 January 2025), that number is inflated by the inclusion of over 4 million records not connected to typical natural history collections (bacteria and virus collections, records marked "incertae cedis").

A review of major herbaria estimated that only 21% of preserved collections were available via GBIF (Paton et al. 2020), and the proportion of zoological coverage is substantially lower. Although the number of digitised records is large, it still represents less than 10% of the true global data bank in natural history museum collections, and the selective digitisation process introduces geographic and taxonomic bias in the currently available data.

To consider how the rate of global digital record capture compares with the ongoing growth of collections, we constructed a simple model to illustrate the intersection between the trajectories of these two trends (Fig. 2). For total

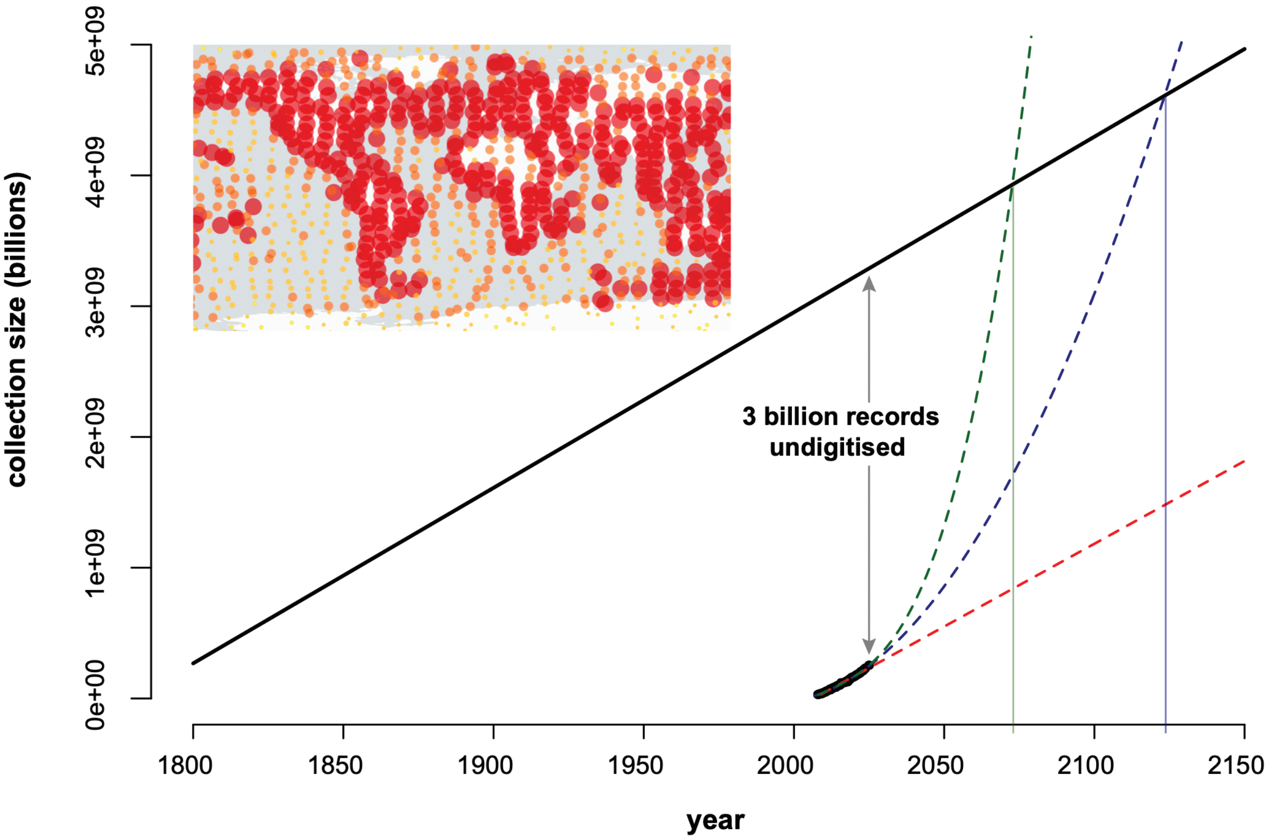

Figure 2. Although global coverage of biological specimen data are good (inset, specimen-based records from GBIF 2023), it is only a fraction of the data that we need to understand global change. Main figure: Simplified scheme of projected global collections growth. The black line represents a linear increase of the number of collection objects in time from the late 1700s to 2020 projected into the future. Dashed lines are model projections based on data records from GBIF that are specimen occurrence records combining all available data including incomplete records (GBIF 2023). In 2020 there are ~200 million digitised records, compared to ~ 3 billion records without digital footprints. The current pace of progress thus requires more than 100 years to achieve full digitisation (blue line). In the best-case model prediction, we could achieve complete digitisation at the earliest around the year 2071; if there is no acceleration (red line) and digitisation continues to increase linearly, the long-term pace of collection growth is higher than the pace of digitisation and the global digitisation gap will even increase.

global collections, we assume a linear increase starting from the late 1700s to 2020 and that this linear rate will continue into the future. The starting point is inferred to be 100,000 objects in 1780, which is likely an under-estimate, increasing to 3.2 billion in 2020 as noted above. Collections growth is much more likely to be accelerating because of the growth in museum resources in the Global South (Gallagher et al. 2020; Wu et al. 2024) and effective techniques for sampling remote habitats such as in the deep sea. A stronger growth curve means more challenge for digitisation, so it should be noted that a linear increase of collections is the best-case scenario for digitisation.

To estimate the potential forecast for digital record growth, we extracted historical data from GBIF for the number of "preserved specimen" records in the database each quarter from 2008 to the start of 2025. These include all records, so including the largest possible set of records, including nonstandard records, and all without validation and without geographical data, and not the preferred scenario of high-quality records data. We fit a polynomial regression to these datapoints using R (R Core Team 2021) (Suppl. material 2) and tested the fit of curves with orders up to 20. The best fit models (based on AIC and RSS) were order 1, 2, and 5 (linear, parabolic, and polynominal). Projections of these fitted curves were then used to calculate the potential intersection of the digitisation trajectory and the inferred linear growth of the museums being digitised.

The current pace of progress could require 100 years to achieve full digitisation (Fig. 2). The fastest rate that we considered would achieve global digitisation in 2071, whereas the second line intercepts in 2118, whereas constant linear rates of digitisation (which we consider implausible) would never catch up with ongoing collections growth and the global digitisation gap would even increase. There are signs that digitisation is already accelerating, which is likely connected to rapid technological developments: the same quick analysis, but using data only up to 2023, would find intercepts at 2096 and 2131. The two most recent years of digitisation progress have skewed the acquisition curve upward so that the projections are already significantly more optimistic. However, the caveats remain that these data represent a simplified model of a best-case scenario with absolutely minimal data quality.

Even if all collections were eventually databased, their digital records, as they appear in collections databases or global aggregators, such as GBIF, are not equivalent to specimens. Occurrence records from metadata are one very important and relevant dimension of collections data, but as noted above, there are other dimensions, attributes, and data types associated with every physical specimen. In order to capture these additional features of the extended specimen, we must first dramatically scale up mass data retrieval and provisioning to keep up with ongoing growth of collections and to capture the backlog. Representative digital coverage of these objects at global scale is therefore imperative.

The current revolutionary advances of artificial intelligence (AI) promise to accelerate the current pace of digitalisation substantially. While application of AI in automated cataloguing and metadata creation is already widespread, there are limitations and human expertise is still a core requirement (Wu et al. 2023). The FAIR use of the collected data can strongly profit from ongoing developments as long as broad user communities are kept in view (Liu et al. 2023; Tiribelli et al. 2024).

Collections as integrated resources for research

Once digitised, data must be linked to other massive species-based information hubs (Kattge et al. 2020), creating an environment for integrated research across temporal and spatial scales. The questions of where species are, and are not, is critical to environmental management. Occurrence data, such as those captured in GBIF and OBIS, are also fundamentally collections-based data. Beyond metadata, which allow one to identify and localize a collection on the global map, collections-based researchers have long sought to champion the many applications and uses of natural history collections, from evolutionary research, to potential uses, to conservation (Committee on Biological Collections 2020).

Collectomics describes the multiple ways that voucher specimens can be harnessed for additional datasets, as well as the many ways these data interact synergistically with other research fields including science and humanities in order to integrate information across scales and generate new knowledge. Collections form the core data for biological and geological sciences, and digital information derived from research on collection objects feeds into interconnected research platforms that are readily accessible to science and society across the globe. What is currently missing is ensuring that these data projects connect directly back to specimen identifiers. Collectomics implies that large scale data platforms interact with other big-data approaches (e.g., remote sensing, citizen science) to form interconnected data networks. These are expected to inspire new developments such as in analyses of biodiversity trends across scales or process-based models of ecosystem functions, but also interdisciplinary research including sciences and humanities.

The point that natural history museum collection objects are useful to myriad subjects has been raised repeatedly (Andreone et al. 2024), yet the major scientific users of collections remain in the realm of taxonomy and systematics. These disciplines shape the physical organisation of natural history collections in zoology and botany and historically drove digitisation efforts. Taxonomic work depends on and produces high quality identification tools and applies integrative, multi-faceted study and description of specimens, but is time and labour intensive (Sigwart et al. 2023).

Is it more useful to provide publicly accessible, large scale, but incomplete and partially inaccurate data (the actual status quo of any museum), or to provide a small fraction of available data but ensuring high quality (expert validated records)? This is a conflict that is not unique to natural history collections, but all “big data” approaches more generally. The benefits of public accessibility are clear, to make hidden data available to a larger pool of potential experts for further validation. But this does not entirely sit comfortably with the cultural values of museums as sources of expertise and trust.

Observation efforts including large-scale citizen-science platforms have provided an explosion of biological records data; however, records that are not connected to voucher specimens have lower geographic, temporal, and species coverage (Daru and Rodriguez 2023). Specimens also have the important advantage of the additional potential data dimensions that cannot be extracted from observations alone. New frontiers in collections’ applications are continuously expanding to meet modern challenges, including historical ecology of parasites (Wood et al. 2023) or identifying potential novel zoonotic pathogens (Colella et al. 2021).

The preservation of physical voucher specimens or objects underpins replicability of all collections-based research (Colella et al. 2020; Miller et al. 2020; Paton et al. 2020). Collections, through their objects can ensure that derived data which are permanently connected to specimens are contributing to an additive knowledge generation process (Kilian et al. 2015), thus allowing to deepen knowledge in an effective and structured way. Collection-based research includes the discovery of hitherto completely unknown entities, which add to biodiversity knowledge as well as theoretical advances (e.g. Borsch et al. 2015).

Mobilising data on species identity and occurrence is often the focus of studies considering the utility of natural history collections (Miller et al. 2020), but recent work has highlighted the potential of specimen collections for drug discovery or emergent diseases (DiEuliis et al. 2016), and trait-based approaches and models to capture processes and functions of whole ecosystems (Funk et al. 2017; Gallagher et al. 2020; Schleuning et al. 2020). The real added value of collection-based data is therefore not only large-scale occurrence data over time and space (Pfenninger and Schwenk 2007), but the fact that the information is tied to physical objects that validate the records but also provide different dimensions for further study (Sumner-Rooney and Sigwart 2017; Webster 2017).

The potential of the “extended specimen” (Webster 2017) is increasingly unlocked with rapidly advancing technologies, such as biogeochemical methods, genome sequencing, hyperspectral imaging, or X-ray tomography, as well as future advances that cannot be imagined now (Sumner-Rooney and Sigwart 2017; Miller et al. 2020; Høye et al. 2021; Hardisty et al. 2022). The term “museomics” was originally coined to highlight the possibilities to obtain ancient DNA from museum artifacts (Wolinsky 2010); today, this signals the potential to extract molecular data from historical specimens (Lalueza-Fox 2022; Fong et al. 2023), and this remains an important breakthrough for biodiversity research. Integrating multiple data dimensions of the genome and phenome of the digital extended specimen enables research centred on different levels of organisation (e.g., genes, traits, taxa, functions, ecosystems, Earth system) and bridge geo- and biosciences, including genomics, and social-ecological research (Ryan et al. 2018). Ultimately, such in-depth digitisation might enable multi-scale approaches to understand biodiversity dynamics and the consequences of biodiversity loss, from genetic diversity over trait and species diversity to changes in ecosystems and the Earth system (Sakschewski et al. 2016; Sanders et al. 2023). The next generation of collection-based approaches share the focus on specimens as a physical data source, but with the promise of accelerating and expanding access (Miralles et al. 2020).

The integration of specimen- and object-based data across collections, object categories and research disciplines is not trivial (Gallagher et al. 2020). There are established global data standards for collections, with unique, permanent identifiers for essential objects such as physical specimens, digital specimens, and people, and for the different types of data derived from specimens (Güntsch et al. 2021; Hardisty et al. 2021). Established repositories exist for some types of data derived from specimens (e.g., genetic sequences, biological traits), but rarely ensure seamless interoperability by directly interlinking different types of databases with the underlying specimens. Our envisioned integrative research agenda requires global initiatives of data sharing and integration, aiming at decentralised, but highly connected and interoperable networks of scientific collection data. This conceptual framework provided by collectomics is thus intended to support this integration and emphasise the centrality of specimen- and object-based data.

Accelerating mass digitisation has a vast potential to make collections data available where they are actually needed. In the context of Anthropocene change, resources and contextual data are most needed in tropical countries of the Global South (Lira-Noriega and Soberón 2015); however, collections resources that underpin the understanding of global social-ecological systems, are mainly located in Europe and North America (Asase et al. 2022). For herbaria there is a documented inverse relationship for plants between the countries where species occur and where collections occur, in a trend that persists into modern herbaria (Park et al. 2023). The number of herbaria and specimens in these has been considerably increasing in recent decades in the Global South, which sometimes have a better staff-specimen ratio as compared to Europe or North America (Thiers 2024) although this is obviously directly connected to collection size. Nevertheless, resources limitations hinder digitisation, and for the Global South in particular this results in limited visibility and accessibility of information from the country of origin; where collections are undigitised, they are invisible, but older collections in the Global North have higher visibility while collections in the country of origin remain unknown to global audiences. The botanical community established a global registry of herbaria in the 1930s, which is credited as a prerequisite for the successful and globally comprehensive digitisation of plant type specimens (http://plants.jstor.org). A lack of visibility can hinder effective participation in scientific workflows despite any commitment to the FAIR principles, which underscores the need for digitisation and data sharing at a global scale, not only from European and North American institutions. New collections often focus on specific projects, with specific collecting goals constrained by interest and/or available permits. There is a high value in broad-spectrum collecting, to expand local natural history collections in species-rich regions (Wu et al. 2024). The goal for global science must be a network of knowledge hubs—physical collections and digital access—including the Global South in order to avoid allowing the gap between the Global North and South to continue expanding.

Specimens and objects in research collections were collected by historical actors who frequently intersect multiple disciplines. Contributions from a certain person are often distributed across multiple holding institutions. Collections credited to a certain person also often connect to uncredited local knowledge holders – the local collector who passed on knowledge or materials that ultimately end up in museums. All of these intersecting agents are an important legacy in tracing and confronting the colonial history of cultural and scientific museum collections in the Global North.

Sensitivity to provenance issues is potentially better developed in art and cultural collections (Grigo and Laely 2022) but there is increasing awareness of the importance of these issues in contemporary collection practices for natural history collections (Van Wyhe and Drawhorn 2015). In natural history collections, much attention has been focussed on the implications of the Nagoya protocol for current and future collections (e.g. Sigwart 2018; Bakker et al. 2020). Consideration of the CARE principles (Collective benefit, Authority to control, Responsibility, Ethics) for indigenous data governance helps to further increase data transparency (https://www.gida-global.org/care). The time has come to learn from each other and to integrate information across traditional disciplines to arrive at a much more comprehensive picture of Anthropocene Earth as a social-ecological system.

The vision of collectomics includes all types of specimen data, linking an object to descriptive metadata, images, sequences, and also its history. The historical and cultural context of the people who originally collected it, and later researchers who brought new approaches to include that specimen in another analysis, connected through geographical, biological, and biographical information. Making collections fully accessible, fully integrated, and visible to the broadest group of global users, is crucial to protecting collections into the future. A fully digitised collections object is not just a species name, an image and some coordinates, but it is a rich, complex, and ongoing history. Our perspective focusses primarily on natural history collections, but this is only one facet we bring into focus, whereas other users will be more driven by other facets among historical, ethnographic, and artistic endeavours that create a rich unified tapestry.

Acknowledgements

We are very grateful to this new journal, Natural History Collections and Museomics, for providing a platform for the global efforts in collectomics. This paper benefited from input of three reviewers and discussions with many colleagues who are too numerous to thank here but all have our gratitude, and we especially acknowledge input from Steffen Pauls, Dieter Uhl, and Georg Zizka (Senckenberg Research Institute and Museum Frankfurt), Klaus Klass and Ulf Linnemann (Senckenberg Natural History Collections Dresden), and Johannes-Geert Hagmann (Deutsches Museum).

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

No funding was reported.

Author contributions

Conceptualization: JDS, AB, MP, MS, KW, CC. Data curation: JDS. Formal analysis: JDS. Visualization: JDS. Writing – original draft: JDS. Writing – review and editing: TB, HT, AG, RL, MS, EH, KW, TT, MP, AB, HS, JDS.

Author ORCIDs

Julia D. Sigwart https://orcid.org/0000-0002-3005-6246
Angelika Brandt https://orcid.org/0000-0002-5807-1632
Markus Pfenninger https://orcid.org/0000-0002-1547-7245
Hanieh Saeedi https://orcid.org/0000-0002-4845-0241
Eva Häffner https://orcid.org/0000-0001-6448-5826
Helmuth Trischler https://orcid.org/0000-0001-6923-2465
Karsten Wesche https://orcid.org/0000-0002-0088-6492

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.

References

Andreone F, James SA, Li S, Paul DL, Stoev P, Penev L (2024) Enlighten the Marvels: A new journal dedicated to natural history museums, collections and their role in interpreting a changing world. Natural History Collections and Museomics 1: 1–7. https://doi.org/10.3897/nhcm.1.142018

Asase A, Mzumara-Gawa TI, Owino JO, Peterson AT, Saupe E (2022) Replacing "parachute science" with "global science" in ecology and conservation biology. Conservation Science and Practice 4: e517. https://doi.org/10.1111/csp2.517

Bakker FT, Antonelli A, Clarke JA, Cook JA, Edwards SV, Ericson PGP, Faurby S, Ferrand N, Gelang M, Gillespie RG, Irestedt M, Lundin K, Larsson E, Matos-Maraví P, Müller J, von Proschwitz T, Roderick GK, Schliep A, Wahlberg N, Wiedenhoeft J, Källersjö M (2020) The Global Museum: natural history collections and the future of evolutionary science and public education. PeerJ 8: e8225. https://doi.org/10.7717/peerj.8225

Barthlott W, Hostert A, Kier G, Kueper W, Kreft H, Mutke J, Rafiqpoor MD, Sommer JH (2007) Geographic patterns of vascular plant diversity at continental to global scales. Erdkunde 61: 305–315. https://doi.org/10.3112/erdkunde.2007.04.01

Blagoderov V, Kitching IJ, Livermore L, Simonsen TJ, Smith VS (2012) No specimen left behind: industrial scale digitization of natural history collections. ZooKeys 209: 133–146. https://doi.org/10.3897/zookeys.209.3178

Blum S, Barker K, Baskauf SJ, Berendsohn WG, Buttigieg PL, Döring M, Droege G, Fichtmueller D, Glöckler F, Güntsch A, Guralnick R, Hoffmann J, Klazenga N, Macklin J, Morris PJ, Paul DL, Petersen M, Robertson T, Sachs J, Shorthouse D, Walls R, Wieczorek J, Zermoglio PF (2019) Integrating ABCD and DarwinCore: Toward a better foundation for biodiversity information standards. Biodiversity Information Science and Standards 3: e37491. https://doi.org/10.3897/biss.3.37491

Borsch T, Hernández-Ledesma P, Berendsohn WG, Flores-Olvera H, Ochoterena H, Zuloaga FO, Von Mering S, Kilian N (2015) An integrative and dynamic approach for monographing species-rich plant groups – Building the global synthesis of the angiosperm order Caryophyllales. Perspectives in Plant Ecology, Evolution and Systematics 17: 284–300. https://doi.org/10.1016/j.ppees.2015.05.003

Borsch T, Stevens A-D, Häffner E, Güntsch A, Berendsohn WG, Appelhans M, Barilaro C, Beszteri B, Blattner F, Bossdorf O, Dalitz H, Dressler S, Duque-Thüs R, Esser H-J, Franzke A, Goetze D, Grein M, Grünert U, Hellwig F, Hentschel J, Hörandl E, Janßen T, Jürgens N, Kadereit G, Karisch T, Koch M, Müller F, Müller J, Ober D, Porembski S, Poschlod P, Printzen C, Röser M, Sack P, Schlüter P, Schmidt M, Schnittler M, Scholler M, Schultz M, Seeber E, Simmel J, Stiller M, Thiv M, Thüs H, Tkach N, Triebel D, Warnke U, Weibulat T, Wesche K, Yurkov A, Zizka G (2020) A complete digitization of German herbaria is possible, sensible and should be started now. Research Ideas and Outcomes 6: e50675. https://doi.org/10.3897/rio.6.e50675

Boudinot BE, Van De Kamp T, Peters P, Knöllinger K (2024) Male genitalia, hierarchical homology, and the anatomy of the bullet ant (*Paraponera clavata*; Hymenoptera, Formicidae). Journal of Morphology 285: e21757. https://doi.org/10.1002/jmor.21757

Colella JP, Stephens RB, Campbell ML, Kohli BA, Parsons DJ, Mclean BS (2020) The Open-Specimen Movement. BioScience 71: 405–414. https://doi.org/10.1093/biosci/biaa146

Colella JP, Bates J, Burneo SF, Camacho MA, Carrion Bonilla C, Constable I, D'Elía G, Dunnum JL, Greiman S, Hoberg EP, Lessa E, Liphardt SW, Londoño-Gaviria M, Losos E, Lutz HL, Ordóñez Garza N, Peterson AT, Martin ML, Ribas CC, Struminger B, Torres-Pérez F, Thompson CW, Weksler M, Cook JA (2021) Leveraging natural history biorepositories as a global, decentralized, pathogen surveillance network. Morrison TE (Ed.). PLOS Pathogens 17: e1009583. https://doi.org/10.1371/journal.ppat.1009583

Committee on Biological Collections (2020) Biological Collections: Ensuring Critical Research and Education for the 21st Century. National Academies Press, Washington, D.C. https://doi.org/10.17226/25592

Cordier T, Forster D, Dufresne Y, Martins CIM, Stoeck T, Pawlowski J (2018) Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Molecular Ecology Resources 18: 1381–1391. https://doi.org/10.1111/1755-0998.12926

Daru BH, Rodriguez J (2023) Mass production of unvouchered records fails to represent global biodiversity patterns. Nature Ecology & Evolution 7: 816–831. https://doi.org/10.1038/s41559-023-02047-3

Díaz S, Settele J, Brondízio ES, Ngo HT, Agard J, Arneth A, Balvanera P, Brauman KA, Butchart SHM, Chan KMA, Garibaldi LA, Ichii K, Liu J, Subramanian SM, Midgley GF, Miloslavich P, Molnár Z, Obura D, Pfaff A, Polasky S, Purvis A, Razzaque J, Reyers B, Chowdhury RR, Shin Y-J, Visseren-Hamakers I, Willis KJ, Zayas CN (2019) Pervasive human-driven decline of life on Earth points to the need for transformative change. Science 366: eaax3100. https://doi.org/10.1126/science.aax3100

DiEuliis D, Johnson KR, Morse SS, Schindel DE (2016) Specimen collections should have a much bigger role in infectious disease research and response. Proceedings of the National Academy of Sciences 113: 4–7. https://doi.org/10.1073/pnas.1522680112

Dominici S, Forli M, Bogi C, Guerrini A, Benvenuti M (2020) Paleobiology from museum collections: comparing historical and novel data on upper Miocene molluscs of the Livorno hills. Rivista Italiana di Paleontologia e Stratigrafia 126: 65–109. https://doi.org/10.13130/2039-4942/13006

Dornelas M, Antão LH, Moyes F, Bates AE, Magurran AE, Adam D, Akhmetzhanova AA, Appeltans W, Arcos JM, Arnold H, Ayyappan N, Badihi G, Baird AH, Barbosa M, Barreto TE, Bässler C, Bellgrove A, Belmaker J, Benedetti-Cecchi L, Bett BJ, Bjorkman AD, Błażewicz M, Blowes SA, Bloch CP, Bonebrake TC, Boyd S, Bradford M, Brooks AJ, Brown JH, Bruelheide H, Budy P, Carvalho F, Castañeda-Moya E, Chen CA, Chamblee JF, Chase TJ, Siegwart Collier L, Collinge SK, Condit R, Cooper EJ, Cornelissen JHC, Cotano U, Kyle Crow S, Damasceno G, Davies CH, Davis RA, Day FP, Degraer S, Doherty TS, Dunn TE, Durigan G, Duffy JE, Edelist D, Edgar GJ, Elahi R, Elmendorf SC, Enemar A, Ernest SKM, Escribano R, Estiarte M, Evans BS, Fan T-Y, Turini Farah F, Loureiro Fernandes L, Farneda FZ, Fidelis A, Fitt R, Fosaa AM, Daher Correa Franco GA, Frank GE, Fraser WR, García H, Cazzolla Gatti R, Givan O, Gorgone-Barbosa E, Gould WA, Gries C, Grossman GD, Gutierréz JR, Hale S, Harmon ME, Harte J, Haskins G, Henshaw DL, Hermanutz L, Hidalgo P, Higuchi P, Hoey A, Van Hoey G, Hofgaard A, Holeck K, Hollister RD, Holmes R, Hoogenboom M, Hsieh C, Hubbell SP, Huettmann F, Huffard CL, Hurlbert AH, Macedo Ivanauskas N, Janík D, Jandt U, Jażdżewska A, Johannessen T, Johnstone J, Jones J, Jones FAM, Kang J, Kartawijaya T, Keeley EC, Kelt DA, Kinnear R, Klanderud K, Knutsen H, Koenig CC, Kortz AR, Král K, Kuhnz LA, Kuo C-Y, Kushner DJ, Laguionie-Marchais C, Lancaster LT, Min Lee C, Lefcheck JS, Lévesque E, Lightfoot D, Lloret F, Lloyd JD, López-Baucells A, Louzao M, Madin JS, Magnússon B, Malamud S, Matthews I, McFarland KP, McGill B, McKnight D, McLarney WO, Meador J, Meserve PL, Metcalfe DJ, Meyer CFJ, Michelsen A, Milchakova N, Moens T, Moland E, Moore J, Mathias Moreira C, Müller J, Murphy G, Myers-Smith IH, Myster RW, Naumov A, Neat F, Nelson JA, Paul Nelson M, Newton SF, Norden N, Oliver JC, Olsen EM, Onipchenko VG, Pabis K, Pabst RJ, Paquette A, Pardede S, Paterson DM, Pélissier R, Peñuelas J, Pérez-Matus A, Pizarro O, Pomati F, Post E, Prins HHT, Priscu JC, Provoost P, Prudic KL, Pulliainen E, Ramesh BR, Mendivil Ramos O, Rassweiler A, Rebelo JE, Reed DC, Reich PB, Remillard SM, Richardson AJ, Richardson JP, van Rijn I, Rocha R, Rivera-Monroy VH, Rixen C, Robinson KP, Ribeiro Rodrigues R, de Cerqueira Rossa-Feres D, Rudstam L, Ruhl H, Ruz CS, Sampaio EM, Rybicki N, Rypel A, Sal S, Salgado B, Santos FAM, Savassi-Coutinho AP, Scanga S, Schmidt J, Schooley R, Setiawan F, Shao K-T, Shaver GR, Sherman S, Sherry TW, Siciński J, Sievers C, da Silva AC, Rodrigues da Silva F, Silveira FL, Slingsby J, Smart T, Snell SJ, Soudzilovskaia NA, Souza GBG, Maluf Souza F, Castro Souza V, Stallings CD, Stanforth R, Stanley EH, Mauro Sterza J, Stevens M, Stuart-Smith R, Rondon Suarez Y, Supp S, Yoshio Tamashiro J, Tarigan S, Thiede GP, Thorn S, Tolvanen A, Teresa Zugliani Toniato M, Totland Ø, Twilley RR, Vaitkus G, Valdivia N, Vallejo MI, Valone TJ, Van Colen C, Vanaverbeke J, Venturoli F, Verheye HM, Vianna M, Vieira RP, Vrška T, Quang Vu C, Van Vu L, Waide RB, Waldock C, Watts D, Webb S, Wesołowski T, White EP, Widdicombe CE, Wilgers D, Williams R, Williams SB, Williamson M, Willig MR, Willis TJ, Wipf S, Woods KD, Woehler EJ, Zawada K, Zettler ML (2018) BioTIME: A database of biodiversity time series for the Anthropocene. Global Ecology and Biogeography 27: 760–786. https://doi.org/10.1111/geb.12729

Fong JJ, Blom MPK, Aowphol A, McGuire JA, Sutcharit C, Soltis PS (2023) Editorial: Recent advances in museomics: revolutionizing biodiversity research. Frontiers in Ecology and Evolution 11: 1188172. https://doi.org/10.3389/fevo.2023.1188172

Funk JL, Larson JE, Ames GM, Butterfield BJ, Cavender-Bares J, Firn J, Laughlin DC, Sutton-Grier AE, Williams L, Wright J (2017) Revisiting the Holy Grail: Using plant functional traits to understand ecological processes. Biological Reviews 92: 1156–1173. https://doi.org/10.1111/brv.12275

Gallagher RV, Falster DS, Maitner BS, Salguero-Gómez R, Vandvik V, Pearse WD, Schneider FD, Kattge J, Poelen JH, Madin JS, Ankenbrand MJ, Penone C, Feng X, Adams VM, Alroy J, Andrew SC, Balk MA, Bland LM, Boyle BL, Bravo-Avila CH, Brennan I, Carthey AJR, Catullo R, Cavazos BR, Conde DA, Chown SL, Fadrique B, Gibb H, Halbritter AH, Hammock J, Hogan JA, Holewa H, Hope M, Iversen CM, Jochum M, Kearney M, Keller A, Mabee P, Manning P, McCormack L, Michaletz ST, Park DS, Perez TM, Pineda-Munoz S, Ray CA, Rossetto M, Sauquet H, Sparrow B, Spasojevic MJ, Telford RJ, Tobias JA, Violle C, Walls R, Weiss KCB, Westoby M, Wright IJ, Enquist BJ (2020) Open Science principles for accelerating trait-based science across the Tree of Life. Nature Ecology & Evolution 4: 294–303. https://doi.org/10.1038/s41559-020-1109-6

Grigo J, Laely T (2022) Attempts to decolonize knowledge production in museum practice. Critical reflections on a collaboration between Uganda and Switzerland. Recherches sociologiques et anthropologiques 53–2: 119–151. https://doi.org/10.4000/rsa.5709

Güntsch A, Groom Q, Ernst M, Holetschek J, Plank A, Röpert D, Fichtmüller D, Shorthouse DP, Hyam R, Dillen M, Trekels M, Haston E, Rainer H (2021) A botanical demonstration of the potential of linking data using unique identifiers for people. Corpas M (Ed.). PLOS ONE 16: e0261130. https://doi.org/10.1371/journal.pone.0261130

Hardisty A, Addink W, Glöckler F, Güntsch A, Islam S, Weiland C (2021) A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes 7: e67379. https://doi.org/10.3897/rio.7.e67379

Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, Rabeler RK, Bates J, Bentley A, Fortes JAB, Hansen S, Macklin JA, Mast AR, Miller JT, Monfils AK, Paul DL, Wallis E, Webster M (2022) Digital extended specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the Internet. BioScience 72: 978–987. https://doi.org/10.1093/biosci/biac060

Hedrick BP, Heberling JM, Meineke EK, Turner KG, Grassa CJ, Park DS, Kennedy J, Clarke JA, Cook JA, Blackburn DC, Edwards SV, Davis CC (2020) Digitization and the future of natural history collections. BioScience 70: 243–251. https://doi.org/10.1093/biosci/biz163

Hobern D, Barik SK, Christidis L, T.Garnett S, Kirk P, Orrell TM, Pape T, Pyle RL, Thiele KR, Zachos FE, Bánki O (2021) Towards a global list of accepted species VI: The Catalogue of Life checklist. Organisms Diversity & Evolution 21: 677–690. https://doi.org/10.1007/s13127-021-00516-w

Hofmann M, Kiel S, Kösters LM, Wäldchen J, Mäder P (2024) Inferring taxonomic affinities and genetic distances using morphological features extracted from specimen images: a case study with a bivalve data set. Systematic Biology 73: 920–940. https://doi.org/10.1093/sysbio/syae042

Høye TT, Ärje J, Bjerge K, Hansen OLP, Iosifidis A, Leese F, Mann HMR, Meissner K, Melvad C, Raitoharju J (2021) Deep learning and computer vision will transform entomology. Proceedings of the National Academy of Sciences 118: e2002545117. https://doi.org/10.1073/pnas.2002545117

Johnson KR, Owens IF, The Global Collection Group (2023) A global approach for natural history museum collections. Science 379: 1192–1194. https://doi.org/10.1126/science.adf6434

Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P, Tautenhahn S, Werner GDA, Aakala T, Abedi M, Acosta ATR, Adamidis GC, Adamson K, Aiba M, Albert CH, Alcántara JM, Alcázar CC, Aleixo I, Ali H, Amiaud B, Ammer C, Amoroso MM, Anand M, Anderson C, Anten N, Antos J, Apgaua DMG, Ashman T, Asmara DH, Asner GP, Aspinwall M, Atkin O, Aubin I, Baastrup-Spohr L, Bahalkeh K, Bahn M, Baker T, Baker WJ, Bakker JP, Baldocchi D, Baltzer J, Banerjee A, Baranger A, Barlow J, Barneche DR, Baruch Z, Bastianelli D, Battles J, Bauerle W, Bauters M, Bazzato E, Beckmann M, Beeckman H, Beierkuhnlein C, Bekker R, Belfry G, Belluau M, Beloiu M, Benavides R, Benomar L, Berdugo-Lattke ML, Berenguer E, Bergamin R, Bergmann J, Bergmann Carlucci M, Berner L, Bernhardt-Römermann M, Bigler C, Bjorkman AD, Blackman C, Blanco C, Blonder B, Blumenthal D, Bocanegra-González KT, Boeckx P, Bohlman S, Böhning-Gaese K, Boisvert-Marsh L, Bond W, Bond-Lamberty B, Boom A, Boonman CCF, Bordin K, Boughton EH, Boukili V, Bowman DMJS, Bravo S, Brendel MR, Broadley MR, Brown KA, Bruelheide H, Brumnich F, Bruun HH, Bruy D, Buchanan SW, Bucher SF, Buchmann N, Buitenwerf R, Bunker DE, Bürger J, Burrascano S, Burslem DFRP, Butterfield BJ, Byun C, Marques M, Scalon MC, Caccianiga M, Cadotte M, Cailleret M, Camac J, Camarero JJ, Campany C, Campetella G, Campos JA, Cano-Arboleda L, Canullo R, Carbognani M, Carvalho F, Casanoves F, Castagneyrol B, Catford JA, Cavender-Bares J, Cerabolini BEL, Cervellini M, Chacón-Madrigal E, Chapin K, Chapin FS, Chelli S, Chen S, Chen A, Cherubini P, Chianucci F, Choat B, Chung K, Chytrý M, Ciccarelli D, Coll L, Collins CG, Conti L, Coomes D, Cornelissen JHC, Cornwell WK, Corona P, Coyea M, Craine J, Craven D, Cromsigt JPGM, Csecserits A, Cufar K, Cuntz M, Da Silva AC, Dahlin KM, Dainese M, Dalke I, Dalle Fratte M, Dang-Le AT, Danihelka J, Dannoura M, Dawson S, De Beer AJ, De Frutos A, De Long JR, Dechant B, Delagrange S, Delpierre N, Derroire G, Dias AS, Diaz-Toribio MH, Dimitrakopoulos PG, Dobrowolski M, Doktor D, Dřevojan P, Dong N, Dransfield J, Dressler S, Duarte L, Ducouret E, Dullinger S, Durka W, Duursma R, Dymova O, E-Vojtkó A, Eckstein RL, Ejtehadi H, Elser J, Emilio T, Engemann K, Erfanian MB, Erfmeier A, Esquivel-Muelbert A, Esser G, Estiarte M, Domingues TF, Fagan WF, Fagúndez J, Falster DS, Fan Y, Fang J, Farris E, Fazlioglu F, Feng Y, Fernandez-Mendez F, Ferrara C, Ferreira J, Fidelis A, Finegan B, Firn J, Flowers TJ, Flynn DFB, Fontana V, Forey E, Forgiarini C, François L, Frangipani M, Frank D, Frenette-Dussault C, Freschet GT, Fry EL, Fyllas NM, Mazzochini GG, Gachet S, Gallagher R, Ganade G, Ganga F, García-Palacios P, Gargaglione V, Garnier E, Garrido JL, De Gasper AL, Gea-Izquierdo G, Gibson D, Gillison AN, Giroldo A, Glasenhardt M, Gleason S, Gliesch M, Goldberg E, Göldel B, Gonzalez-Akre E, Gonzalez-Andujar JL, González-Melo A, González-Robles A, Graae BJ, Granda E, Graves S, Green WA, Gregor T, Gross N, Guerin GR, Günther A, Gutiérrez AG, Haddock L, Haines A, Hall J, Hambuckers A, Han W, Harrison SP, Hattingh W, Hawes JE, He T, He P, Heberling JM, Helm A, Hempel S, Hentschel J, Hérault B, Hereş A, Herz K, Heuertz M, Hickler T, Hietz P, Higuchi P, Hipp AL, Hirons A, Hock M, Hogan JA, Holl K, Honnay O, Hornstein D, Hou E, Hough-Snee N, Hovstad KA, Ichie T, Igić B, Illa E, Isaac M, Ishihara M, Ivanov L, Ivanova L, Iversen CM, Izquierdo J, Jackson RB, Jackson B, Jactel H, Jagodzinski AM, Jandt U, Jansen S, Jenkins T, Jentsch A, Jespersen JRP, Jiang G, Johansen JL, Johnson D, Jokela EJ, Joly CA, Jordan GJ, Joseph GS, Junaedi D, Junker RR, Justes E, Kabzems R, Kane J, Kaplan Z, Kattenborn T, Kavelenova L, Kearsley E, Kempel A, Kenzo T, Kerkhoff A, Khalil MI, Kinlock NL, Kissling WD, Kitajima K, Kitzberger T, Kjøller R, Klein T, Kleyer M, Klimešová J, Klipel J, Kloeppel B, Klotz S, Knops JMH, Kohyama T, Koike F, Kollmann J, Komac B, Komatsu K, König C, Kraft NJB, Kramer K, Kreft H, Kühn I, Kumarathunge D, Kuppler J, Kurokawa H, Kurosawa Y, Kuyah S, Laclau J, Lafleur B, Lallai E, Lamb E, Lamprecht A, Larkin DJ, Laughlin D, Le Bagousse-Pinguet Y, Le Maire G, Le Roux PC, Le Roux E, Lee T, Lens F, Lewis SL, Lhotsky B, Li Y, Li X, Lichstein JW, Liebergesell M, Lim JY, Lin Y, Linares JC, Liu C, Liu D, Liu U, Livingstone S, Llusià J, Lohbeck M, López-García Á, Lopez-Gonzalez G, Lososová Z, Louault F, Lukács BA, Lukeš P, Luo Y, Lussu M, Ma S, Maciel Rabelo Pereira C, Mack M, Maire V, Mäkelä A, Mäkinen H, Malhado ACM, Mallik A, Manning P, Manzoni S, Marchetti Z, Marchino L, Marcilio-Silva V, Marcon E, Marignani M, Markesteijn L, Martin A, Martínez-Garza C, Martínez-Vilalta J, Mašková T, Mason K, Mason N, Massad TJ, Masse J, Mayrose I, McCarthy J, McCormack ML, McCulloh K, McFadden IR, McGill BJ, McPartland MY, Medeiros JS, Medlyn B, Meerts P, Mehrabi Z, Meir P, Melo FPL, Mencuccini M, Meredieu C, Messier J, Mészáros I, Metsaranta J, Michaletz ST, Michelaki C, Migalina S, Milla R, Miller JED, Minden V, Ming R, Mokany K, Moles AT, Molnár A, Molofsky J, Molz M, Montgomery RA, Monty A, Moravcová L, Moreno-Martínez A, Moretti M, Mori AS, Mori S, Morris D, Morrison J, Mucina L, Mueller S, Muir CD, Müller SC, Munoz F, Myers-Smith IH, Myster RW, Nagano M, Naidu S, Narayanan A, Natesan B, Negoita L, Nelson AS, Neuschulz EL, Ni J, Niedrist G, Nieto J, Niinemets Ü, Nolan R, Nottebrock H, Nouvellon Y, Novakovskiy A, The Nutrient Network, Nystuen KO, O'Grady A, O'Hara K, O'Reilly-Nugent A, Oakley S, Oberhuber W, Ohtsuka T, Oliveira R, Öllerer K, Olson ME, Onipchenko V, Onoda Y, Onstein RE, Ordonez JC, Osada N, Ostonen I, Ottaviani G, Otto S, Overbeck GE, Ozinga WA, Pahl AT, Paine CET, Pakeman RJ, Papageorgiou AC, Parfionova E, Pärtel M, Patacca M, Paula S, Paule J, Pauli H, Pausas JG, Peco B, Penuelas J, Perea A, Peri PL, Petisco-Souza AC, Petraglia A, Petritan AM, Phillips OL, Pierce S, Pillar VD, Pisek J, Pomogaybin A, Poorter H, Portsmuth A, Poschlod P, Potvin C, Pounds D, Powell AS, Power SA, Prinzing A, Puglielli G, Pyšek P, Raevel V, Rammig A, Ransijn J, Ray CA, Reich PB, Reichstein M, Reid DEB, Réjou-Méchain M, De Dios VR, Ribeiro S, Richardson S, Riibak K, Rillig MC, Riviera F, Robert EMR, Roberts S, Robroek B, Roddy A, Rodrigues AV, Rogers A, Rollinson E, Rolo V, Römermann C, Ronzhina D, Roscher C, Rosell JA, Rosenfield MF, Rossi C, Roy DB, Royer-Tardif S, Rüger N, Ruiz-Peinado R, Rumpf SB, Rusch GM, Ryo M, Sack L, Saldaña A, Salgado-Negret B, Salguero-Gomez R, Santa-Regina I, Santacruz-García AC, Santos J, Sardans J, Schamp B, Scherer-Lorenzen M, Schleuning M, Schmid B, Schmidt M, Schmitt S, Schneider JV, Schowanek SD, Schrader J, Schrodt F, Schuldt B, Schurr F, Selaya Garvizu G, Semchenko M, Seymour C, Sfair JC, Sharpe JM, Sheppard CS, Sheremetiev S, Shiodera S, Shipley B, Shovon TA, Siebenkäs A, Sierra C, Silva V, Silva M, Sitzia T, Sjöman H, Slot M, Smith NG, Sodhi D, Soltis P, Soltis D, Somers B, Sonnier

G, Sørensen MV, Sosinski EE, Soudzilovskaia NA, Souza AF, Spasojevic M, Sperandii MG, Stan AB, Stegen J, Steinbauer K, Stephan JG, Sterck F, Stojanovic DB, Strydom T, Suarez ML, Svenning J, Svitková I, Svitok M, Svoboda M, Swaine E, Swenson N, Tabarelli M, Takagi K, Tappeiner U, Tarifa R, Tauugourdeau S, Tavsanoglu C, Te Beest M, Tedersoo L, Thiffault N, Thom D, Thomas E, Thompson K, Thornton PE, Thuiller W, Tichý L, Tissue D, Tjoelker MG, Tng DYP, Tobias J, Török P, Tarin T, Torres-Ruiz JM, Tóthmérész B, Treurnicht M, Trivellone V, Trolliet F, Trotsiuk V, Tsakalos JL, Tsiripidis I, Tysklind N, Umehara T, Usoltsev V, Vadeboncoeur M, Vaezi J, Valladares F, Vamosi J, Van Bodegom PM, Van Breugel M, Van Cleemput E, Van De Weg M, Van Der Merwe S, Van Der Plas F, Van Der Sande MT, Van Kleunen M, Van Meerbeek K, Vanderwel M, Vanselow KA, Vårhammar A, Varone L, Vasquez Valderrama MY, Vassilev K, Vellend M, Veneklaas EJ, Verbeeck H, Verheyen K, Vibrans A, Vieira I, Villacís J, Violle C, Vivek P, Wagner K, Waldram M, Waldron A, Walker AP, Waller M, Walther G, Wang H, Wang F, Wang W, Watkins H, Watkins J, Weber U, Weedon JT, Wei L, Weigelt P, Weiher E, Wells AW, Wellstein C, Wenk E, Westoby M, Westwood A, White PJ, Whitten M, Williams M, Winkler DE, Winter K, Womack C, Wright IJ, Wright SJ, Wright J, Pinho BX, Ximenes F, Yamada T, Yamaji K, Yanai R, Yankov N, Yguel B, Zanini KJ, Zanne AE, Zelený D, Zhao Y, Zheng J, Zheng J, Ziemińska K, Zirbel CR, Zizka G, Zo-Bi IC, Zotz G, Wirth C (2020) TRY plant trait database – enhanced coverage and open access. Global Change Biology 26: 119–188. https://doi.org/10.1111/gcb.14904

Kilian N, Henning T, Plitzner P, Müller A, Güntsch A, Stöver BC, Müller KF, Berendsohn WG, Borsch T (2015) Sample data processing in an additive and reproducible taxonomic workflow by using character data persistently linked to preserved individual specimens. Database 2015: bav094. https://doi.org/10.1093/database/bav094

Kissling WD, Walls R, Bowser A, Jones MO, Kattge J, Agosti D, Amengual J, Basset A, van Bodegom PM, Cornelissen JHC, Denny EG, Deudero S, Egloff W, Elmendorf SC, Alonso García E, Jones KD, Jones OR, Lavorel S, Lear D, Navarro LM, Pawar S, Pirzl R, Rüger N, Sal S, Salguero-Gómez R, Schigel D, Schulz K-S, Skidmore A, Guralnick RP (2018) Towards global data products of Essential Biodiversity Variables on species traits. Nature Ecology & Evolution 2: 1531–1540. https://doi.org/10.1038/s41559-018-0667-3

Kothari S, Beauchamp-Rioux R, Laliberté E, Cavender-Bares J (2023) Reflectance spectroscopy allows rapid, accurate and non-destructive estimates of functional traits from pressed leaves. Methods in Ecology and Evolution 14: 385–401. https://doi.org/10.1111/2041-210X.13958

Lalueza-Fox C (2022) Museomics. Current Biology 32: R1214–R1215. https://doi.org/10.1016/j.cub.2022.09.019

Lira-Noriega A, Soberón J (2015) The relationship among biodiversity, governance, wealth, and scientific capacity at a country level: Disaggregation and prioritization. Ambio 44: 391–400. https://doi.org/10.1007/s13280-014-0581-0

Liu F, Hindmarch J, Hess M (2023) A review of the cultural heritage linked open data ontologies and models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-M-2–2023: 943–950. https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-943-2023

Lopez L, Turner KG, Bellis ES, Lasky JR (2020) Genomics of natural history collections for understanding evolution in the wild. Molecular Ecology Resources 20: 1153–1160. https://doi.org/10.1111/1755-0998.13245

Meineke EK, Davies TJ, Daru BH, Davis CC (2019) Biological collections for understanding biodiversity in the Anthropocene. Philosophical Transactions of the Royal Society B: Biological Sciences 374: 20170386. https://doi.org/10.1098/rstb.2017.0386

Miller SE, Barrow LN, Ehlman SM, Goodheart JA, Greiman SE, Lutz HL, Misiewicz TM, Smith SM, Tan M, Thawley CJ, Cook JA, Light JE (2020) Building natural history collections for the Twenty-First Century and beyond. BioScience 70: 674–687. https://doi.org/10.1093/biosci/biaa069

Miralles A, Bruy T, Wolcott K, Scherz MD, Begerow D, Beszteri B, Bonkowski M, Felden J, Gemeinholzer B, Glaw F, Glöckner FO, Hawlitschek O, Kostadinov I, Nattkemper TW, Printzen C, Renz J, Rybalka N, Stadler M, Weibulat T, Wilke T, Renner SS, Vences M (2020) Repositories for taxonomic data: Where we are and what is missing. Friedman M (Ed.). Systematic Biology 69: 1231–1253. https://doi.org/10.1093/sysbio/syaa026

Misof B, Tockner K, Trischler H (2023) Die Bibliothek des biologischen, technischen und kulturellen Wissens – Warum brauchen wir eine integrierte Sammlungsinfrastruktur? Bibliothek Forschung und Praxis 47: 495–502. https://doi.org/10.1515/bfp-2023-0054

Owen D, Groom Q, Hardisty A, Leegwater T, Livermore L, Van Walsum M, Wijkamp N, Spasić I (2020) Towards a scientific workflow featuring Natural Language Processing for the digitisation of natural history collections. Research Ideas and Outcomes 6: e58030. https://doi.org/10.3897/rio.6.e58030

Park DS, Feng X, Akiyama S, Ardiyani M, Avendaño N, Barina Z, Bärtschi B, Belgrano M, Betancur J, Bijmoer R, Bogaerts A, Cano A, Danihelka J, Garg A, Giblin DE, Gogoi R, Guggisberg A, Hyvärinen M, James SA, Sebola RJ, Katagiri T, Kennedy JA, Komil TSh, Lee B, Lee SML, Magri D, Marcucci R, Masinde S, Melnikov D, Mráz P, Mulenko W, Musili P, Mwachala G, Nelson BE, Niezgoda C, Novoa Sepúlveda C, Orli S, Paton A, Payette S, Perkins KD, Ponce MJ, Rainer H, Rasingam L, Rustiami H, Shiyan NM, Bjorå CS, Solomon J, Stauffer F, Sumadijaya A, Thiébaut M, Thiers BM, Tsubota H, Vaughan A, Virtanen R, Whitfeld TJS, Zhang D, Zuloaga FO, Davis CC (2023) The colonial legacy of herbaria. Nature Human Behaviour 7: 1059–1068. https://doi.org/10.1038/s41562-023-01616-7

Paton A, Antonelli A, Carine M, Forzza RC, Davies N, Demissew S, Dröge G, Fulcher T, Grall A, Holstein N, Jones M, Liu U, Miller J, Moat J, Nicolson N, Ryan M, Sharrock S, Smith D, Thiers B, Victor J, Wilkinson T, Dickie J (2020) Plant and fungal collections: Current status, future perspectives. PLANTS, PEOPLE, PLANET 2: 499–514. https://doi.org/10.1002/ppp3.10141

Peters K, Blatt-Janmaat KL, Tkach N, Van Dam NM, Neumann S (2023) Untargeted Metabolomics for Integrative Taxonomy: Metabolomics, DNA Marker-Based Sequencing, and Phenotype Bioimaging. Plants 12: 881. https://doi.org/10.3390/plants12040881

Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evolutionary Biology 7: 121. https://doi.org/10.1186/1471-2148-7-121

R Core Team (2021) R: A language and environment for statistical computing. [Available from:] https://www.R-project.org/

Roy K, Jablonski D, Valentine JW, Rosenberg G (1998) Marine latitudinal diversity gradients: Tests of causal hypotheses. Proceedings of the National Academy of Sciences 95: 3699–3702. https://doi.org/10.1073/pnas.95.7.3699

Ryan SF, Deines JM, Scriber JM, Pfrender ME, Jones SE, Emrich SJ, Hellmann JJ (2018) Climate-mediated hybrid zone movement revealed with genomics, museum collection, and simulation modeling. Proceedings of the National Academy of Sciences 115: E2284–E2291. https://doi.org/10.1073/pnas.1714950115

Sakschewski B, von Bloh W, Boit A, Poorter L, Peña-Claros M, Heinke J, Joshi J, Thonicke K (2016) Resilience of Amazon forests emerges from plant trait diversity. Nature Climate Change 6: 1032–1036. https://doi.org/10.1038/nclimate3109

Sanders NJ, Cooper N, Davis Rabosky AR, Gibson DJ (2023) Leveraging natural history collections to understand the impacts of global change. Journal of Animal Ecology 92: 232–236. https://doi.org/10.1111/1365-2656.13882

Schleuning M, Neuschulz EL, Albrecht J, Bender IMA, Bowler DE, Dehling DM, Fritz SA, Hof C, Mueller T, Nowak L, Sorensen MC, Böhning-Gaese K, Kissling WD (2020) Trait-based assessments of climate-change impacts on interacting species. Trends in Ecology & Evolution 35: 319–328. https://doi.org/10.1016/j.tree.2019.12.010

Sigwart JD (2018) What Species Mean: A User's Guide to the Units of Biodiversity. CRC Press, 259 pp. https://doi.org/10.1201/9780429458972

Sigwart JD, Schleuning M, Brandt A, Pfenninger M, Senckenberg Collectomics Consortium (2021) Integrated use of enhanced natural history collections is key to solve the biodiversity crisis. Senckenberg Metadata Portal 2021.04.1. [preprint] https://doi.org/10.12761/sgn.2021.04.1

Sigwart JD, Chen C, Tilic E, Vences M, Riehl T (2023) Why is there no service to support taxonomy? BioEssays. 45: 2300070. https://doi.org/10.1002/bies.202300070

Spear DM, Pauly GB, Kaiser K (2017) Citizen science as a tool for augmenting museum collection data from urban areas. Frontiers in Ecology and Evolution 5: 86. https://doi.org/10.3389/fevo.2017.00086

Sumner-Rooney L, Sigwart JD (2017) Lazarus in the museum: resurrecting historic specimens through new technology. Invertebrate Zoology 14: 73–84. https://doi.org/10.15298/invertzool.14.1.11

Thiers BM (2024) Strengthening Partnerships to Safeguard the Future of Herbaria. Diversity 16: 36. https://doi.org/10.3390/d16010036

Tiribelli S, Pansoni S, Frontoni E, Giovanola B (2024) Ethics of Artificial Intelligence for Cultural Heritage: Opportunities and Challenges. IEEE Transactions on Technology and Society 5: 293–305. https://doi.org/10.1109/TTS.2024.3432407

Van Wyhe J, Drawhorn GM (2015) 'I am Ali Wallace' The Malay assistant of Alfred Russel Wallace. Journal of the Malaysian Branch of the Royal Asiatic Society 88: 3–31. https://doi.org/10.1353/ras.2015.0012

Vermeij GJ (1993) Evolution and Escalation: An Ecological History of Life. Princeton University Press, 547 pp.

Webster MS (2017) The Extended Specimen: Emerging Frontiers in Collections-Based Ornithological Research. CRC Press, 253 pp.

Wolinsky H (2010) History in a single hair: Museums are treasure troves of ancient DNA. Modern extraction and sequencing technology is key to putting the secrets of genetic history to use. EMBO reports 11: 427–430. https://doi.org/10.1038/embor.2010.70

Wood CL, Leslie KL, Claar D, Mastick N, Preisser W, Vanhove MPM, Welicky R (2023) How to use natural history collections to resurrect information on historical parasite abundances. Journal of Helminthology 97: e6. https://doi.org/10.1017/S0022149X2200075X

Wu M, Brandhorst H, Marinescu M-C, Lopez JM, Hlava M, Busch J (2023) Automated metadata annotation: What is and is not possible with machine learning. Data Intelligence 5: 122–138. https://doi.org/10.1162/dint_a_00162

Wu R, Zou Y, Liao S, Shi K, Nan X, Yan H, Luo J, Xiang Z, Bao Z (2024) Shall we promote natural history collection today?—Answered by reviewing Ernest Henry Wilson's plant collection process in China. Science of The Total Environment 915: 170179. https://doi.org/10.1016/j.scitotenv.2024.170179

Supplementary material 1

Additional contributors in the Collectomics Consortium, Germany

Authors: Collectomics Consortium

Data type: docx

Link: https://doi.org/10.3897/nhcm.2.148855.suppl1

Supplementary material 2

R commands and data used to generate Fig. 2

Authors: Julia D. Sigwart

Data type: R

Link: https://doi.org/10.3897/nhcm.2.148855.suppl2