Paleobiology

www.cambridge.org/pab

Data equity in paleobiology: progress, challenges, and future outlook

Emma M. Dunne[1], Devapriya Chattopadhyay[2], Christopher D. Dean[3], Erin M. Dillon[4], Elizabeth M. Dowding[1], Pedro L. Godoy[5], Jansen A. Smith[6] and Nussaïbah B. Raja[1]

[1]GeoZentrum Nordbayern, Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Erlangen, Germany
[2]IISER Pune, Department of Earth and Climate Science, Pashan, Pune, India
[3]Department of Earth Sciences, University College London, London, U.K.
[4]Smithsonian Tropical Research Institute, Balboa, Republic of Panama
[5]Department of Zoology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil
[6]Department of Earth and Environmental Sciences, University of Minnesota Duluth, Duluth, Minnesota, U.S.A.

Invited Article

Cite this article: Dunne, E. M., D. Chattopadhyay, C. D. Dean, E. M. Dillon, E. M. Dowding, P. L. Godoy, J. A. Smith, and N. B. Raja (2025). Data equity in paleobiology: progress, challenges, and future outlook. *Paleobiology* **51**, 237–249.
https://doi.org/10.1017/pab.2024.61

Received: 14 March 2024
Revised: 20 November 2024
Accepted: 20 November 2024

Handling Editor:
Mark Patzkowsky

Corresponding author:
Emma M. Dunne;
Email: dunne.emma.m@gmail.com

Abstract

In the last 50 years, the field of paleobiology has undergone a computational revolution that opened multiple new avenues for recording, storing, and analyzing vital data on the history of life on Earth. With these advances, the amount of data available for research has grown, but so too has our responsibility to ensure that our data tools and infrastructures continue to innovate in order to best serve our diverse community. This review focuses on data equity in paleobiology, an aspirational goal, wherein data in all forms are collected, stored, shared and analyzed in a responsible, equitable, and sustainable manner. While there have been many advancements across the last five decades, inequities persist. Our most significant challenges relate to several interconnected factors, including ethical data collection, sustainable infrastructure, socioeconomic biases, and global inequalities. We highlight the ways in which data equity is critical for paleobiology and stress the need for collaborative efforts across the paleobiological community to urgently address these data equity challenges. We also provide recommendations for actions from individuals, teams, academic publishers, and academic societies in order to continue enhancing data equity and ensuring an equitable and sustainable future for our field.

Non-technical Summary

The study of the history of life (paleobiology) relies heavily on data preserved in various forms. Over the past 50 years, there has been a big shift toward using computers and other technology to store and analyze these data, opening up new possibilities for research. As a result, the amount of data available has increased dramatically. However, with this growth comes the responsibility to make sure everyone in the paleobiological community can collect, store, share, analyze, and use data in a fair and sustainable way. This review looks at how well we have done in this regard over the past five decades and what challenges still lie ahead. While progress has been made in creating tools for sharing digital data, there are still many issues we must address. These include the process of how fossil data are collected and biases based on social and economic factors (e.g., wealth and access to resources). To address these challenges, everyone in the paleobiological community needs to work together. We provide suggestions for actions that individuals, their teams, academic journals, and societies can take to promote equity in the field now and into the future.

Introduction

The history of life on Earth is uniquely preserved in paleobiological data. These data take numerous forms, including taxonomic, anatomical, molecular, morphological, (paleo)ecological, geographic, and stratigraphic. They have been used for centuries to answer fundamental and diverse questions about biodiversity and evolutionary patterns throughout the Phanerozoic and beyond (Phillips 1860; Raup 1972; Sepkoski 1996; Betts et al. 2018; Cohen and Kodner 2022; Finnegan et al. 2024).

The journal *Paleobiology* was founded in 1975 just as the field of paleobiology, and paleontology more boradly, was undergoing a computational revolution. This revolution opened multiple new avenues for paleobiologists to record, store, and analyze paleobiological data (Davies et al. 2017; Pandolfi et al. 2020). The research community that *Paleobiology* now encompasses builds on this rich history to explore exciting and dynamic research rooted in data-driven and computational methods (Raup 1991; Cunningham et al. 2014; Davies et al. 2017; Dillon et al. 2023). Over the last 50 years, as the variety of research questions and approaches have increased in number (e.g., Seddon et al. 2014), so too have compilations of paleobiological data. With an ever-growing amount of data, it is critical that our tools, systems, and repositories keep up with growing research demand and continue to innovate in order to best serve our diverse community (Payne et al. 2012; Seddon et al. 2014; Kaufman and PAGES 2k Special-Issue Editorial Team 2018; Smith et al. 2023b).

Data Equity

Data equity is a growing movement for more responsible data work, from data collection and storage to analysis and data-driven decision making (Jagadish et al. 2023; data.org 2024). "Equity" is the term given to the pursuit of ensuring fair treatment, equality of opportunity, and fairness in access to information and resources for all. Achieving equity in science, including paleobiology, is an ongoing challenge that requires sustained action from across the scientific community (Sugimoto et al. 2017; Bernard and Cooperdock 2018; Dutt 2020; North et al. 2020; Posselt 2020; Ranganathan et al. 2021; Muralidhar and Ananthanarayanan 2024). Equity strengthens science in a multitude of interconnected ways, providing diversity of experiences, knowledge, and skills (Dutt 2018; A.-M. Núñez et al. 2020; Emery et al. 2021; M. A. Nuñez et al. 2024).

Data equity is an essential component of equity in science. We define data equity as the responsible, accessible, and sustainable collection, sharing, analysis, and use of scientific data. In this article, we present an overview of data equity in paleobiology, focusing on the kinds of research typically published by *Paleobiology*. We outline where significant progress has and is being made and point to current and future challenges that our field must systemically address in order to increase data equity over the next half century and beyond.

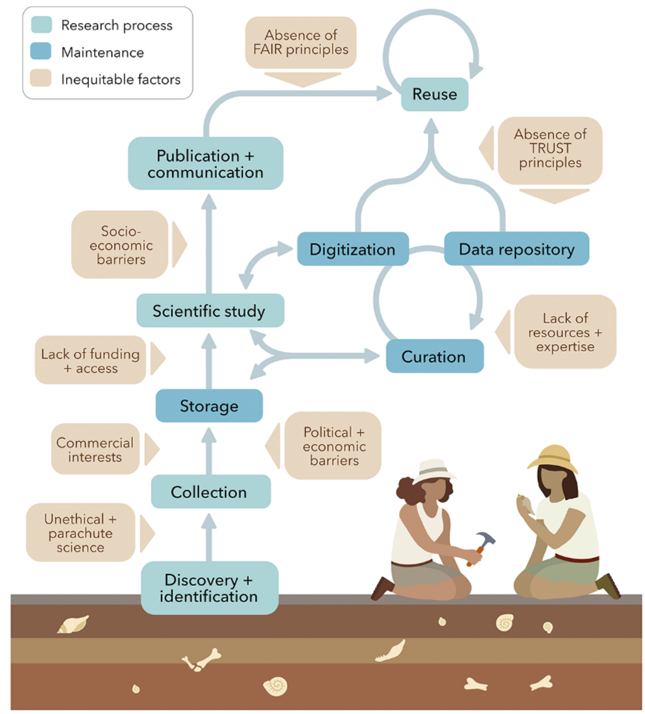

Figure 1. Simplified flowchart illustrating generalized steps in paleobiological research processes and the various factors that introduce inequity with regard to data collection, storage, study, analysis, publication, and reuse. Note: inequitable factors may be relevant at more steps than indicated, but are anticipated to be acute where included. FAIR, Findable, Accessible, Interoperable, and Reusable; TRUST, Transparency, Responsibility, User focus, Sustainability, and Technology.

Data Collection

Fossil Collection and Study

Fossils are the primary unit of paleobiological data, and their associated geographic, stratigraphic, ecological, and morphological data underpin a vast array of paleobiological studies (i.e., the "extended specimen"; Webster 2017; Lendemer et al. 2020). The collection and documentation of fossil specimens is therefore foundational to any paleobiological investigation on any scale (Fig. 1). Principles of data equity apply to paleobiological data collection in that all data collection should be carried out in a responsible and sustainable way that is widely accessible to all paleobiologists.

It is widely documented that the known fossil record provides an invaluable but imperfect view of how biodiversity has changed over Earth's history, due to various taphonomic, geological, and anthropogenically introduced sampling biases (Raup 1972; Behrensmeyer et al. 2000; Alroy et al. 2001; Smith and McGowan 2011; Vilhena and Smith 2013; Close et al. 2018; Whitaker and Kimmig 2020; Benson et al. 2021). While there is a large and growing body of research on how quantitative methods can alleviate some of these limitations imposed by fossil record biases (e.g., Warnock et al. 2020; Smith et al. 2022; Antell et al. 2023; Dillon et al. 2023; Reitan et al. 2024), considerably less attention has been paid to how we, as paleobiologists, impose additional biases through data collection and handling that are far less easy to solve through quantitative means. Inequities and global geographic biases in how paleobiological data are collected can influence downstream analysis of global/regional patterns.

There is a long history of fossils being collected by humans (e.g., Cortés-Sánchez et al. 2020), but the scale of collection increased dramatically as paleontology developed around an extractive process connected to mining, quarrying, and systematic mapping surveys (Schofer 2003; Manias 2015; Das and Lowe 2018; Monarrez et al. 2022; Stewens et al. 2022). Over centuries, this extraction and exploitation of fossil-rich and paleontologically significant regions of the world has continued, facilitated in particular by European colonialism in the nineteenth century (Aldrich 2009; Manias 2015; Zuroski 2017; Das and Lowe 2018; Yen 2024). Today, due to the fundamental nature of fossil specimens (as limited, sought-after physical specimens), their extraction continues to underpin paleobiological research. However, there is a growing awareness within the paleobiological community regarding the connection between fossil collecting and issues related to equity, ethics, socioeconomics, legality, environmental degradation, and distress for Indigenous communities (Bradley 2010; Cisneros et al. 2022; Dunne et al. 2022; Monarrez et al. 2022; Raja et al. 2022; Kempf et al. 2023).

Global Inequalities in Knowledge Generation

Compilations of paleontological data exhibit a strong association between the production of published knowledge and wealthier, more politically stable countries, especially in North America and western Europe (Raja et al. 2022). This same asymmetrical pattern is also seen in disciplines allied to paleobiology, such as in modern biodiversity data compilations (Boakes et al. 2010; Amano and Sutherland 2013; Hughes et al. 2021; Trisos et al. 2021). Many socioeconomic factors related to wealth, access to education, security, and working conditions determine who can participate in

Table 1. Recommended actions for improving and enhancing data equity in paleobiology

Stakeholder(s)	Actions
Individuals and research teams	• Regularly engage with the literature and other media on the topics of data equity, accessibility, research ethics, and open data in paleobiology and beyond. • Develop data equity protocols for your research team, based on established principles; (e.g., FAIR [Findable, Accessible, Interoperable, and Reusable] Data Principles). • Develop protocols for data collection that center equity and sustainability, and that are in accordance with data sovereignty guidelines; (e.g., CARE [Collective Benefit, Authority to Control, Responsibility, and Ethics] Principles for Indigenous Data Governance. • Provide educational resources on data equity and associated topics (e.g., as part of doctoral training programs). • Consider data equity at each step in the research process, from applying for funding to conducting research and sharing results, including in funding applications. • Conduct literature and data searches in multiple languages. • Increase the visibility of data equity challenges through conversations with colleagues, leaders, and students, as well as in online engagement, conference presentations, and publications.
Academic journals	• Develop strict and enforceable data repository protocols based on FAIR, CARE, and TRUST (Transparency, Responsibility, User focus, Sustainability, and Technology) principles. • Introduce specialist data editors to assist authors and peer reviewers in engaging with best practices in data equity (e.g., FAIR principles). • Permit and encourage the use of non–English language summaries or abstracts. • Remove limits on numbers of references to facilitate citation of data sources, especially in online-only journals. • Adopt strict editorial policies for the handling of parachute science and scientific colonialism that are accompanied by comprehensive guidance for peer reviewers and authors on how to avoid such practices. • Work towards implementing the principles of diamond open access (OA). • Introduce fee waivers for OA, especially for authors in countries that do not provide funding for OA.
Academic societies	• Work with membership to develop community guidelines to increase understanding of data equity and data sharing. • Engage with society journals to develop guidelines for data sharing focused on equity and accessibility. • Require statements on actions towards data equity as part of applications for research funding. • Highlight the diversity and equity issues faced by the membership and wider community. • Provide and/or highlight training opportunities for membership to educate themselves on topics related to data equity. • Advocate to relevant government bodies for legal protection and safeguarding of geoheritage and museums. • Publicly comment on local and international events related to data equity.

scientific research (Bernard and Cooperdock 2018; Nuñez et al. 2021; Valenzuela-Toro and Viglino 2021). If we are to work toward a world with equitable knowledge production in paleobiology, it is important to not only be aware of these entrenched biases when working with paleobiological data, but also to actively work to mitigate and counteract them in our own research (e.g., more equitable sharing of funding, tools, and training; Table 1).

Parachute science (also scientific colonialism or expropriation) refers to when researchers, typically from higher-income countries, "drop in" on lower-income countries to extract scientific material and do so without acknowledging local expertise or connecting with local communities (Stefanoudis et al. 2021; Asase et al. 2022; Cisneros et al. 2022). This practice is prevalent in paleontology and can threaten both the ethical and legal integrity of paleobiological data (Cisneros et al. 2022; Dunne et al. 2022; Raja and Dunne 2022; Raja et al. 2022). One such recent example is the widely publicized case of "*Ubirajara jubatus*", a dinosaur fossil removed from Brazil contrary to long-established national laws and housed in Germany while being studied by German and British researchers until its eventual repatriation to Brazil in 2023 (Pérez Ortega 2022). Parachute science can also threaten the scientific integrity of the data; for instance, important contextual stratigraphic or geographic information might be overlooked or missing, a situation that could be greatly improved through collaboration with local experts. In the case of "*Ubirajara jubatus*", the fossil specimen was left without a valid taxonomic name following the retraction of the original publication at the onset of the legal investigation (Cisneros et al. 2022). While these issues appear at first to be restricted to fossil specimens and not their associated data, there is a growing informal discussion around the most ethical way to handle these data when they are part of larger compilations (e.g., see Dunne et al. 2022). As awareness grows, more actions, such as repatriation of fossil specimens, are being undertaken to remedy unethical and illegal actions (Harris 2015; Cisneros et al. 2021; Stewens et al. 2022). However, extraction goes beyond material objects. For example, parachute science may not only plunder geological, paleontological, or biological specimens, but also local knowledge about prospective study sites and how to navigate to those areas. (Cisneros et al. 2022; Nóbrega et al. 2023; Raposo et al. 2023; Coningham et al. 2024).

(Un)ethical Collection Practices

Fossils have long been used not only for scientific reasons, but also for cultural and commercial purposes, notably by Indigenous peoples across the world (Mayor 2005; Cortés-Sánchez et al. 2020). Extractive practices have led to a worldwide paucity of input from Indigenous peoples and local communities in the generation of scientific and paleobiological data (Jennings et al. 2023; Carvalho et al. 2023; Kempf et al. 2023), highlighting a global lack of Indigenous data Sovereignty. For example, the United States has a long history of fossil dispossession from Indigenous peoples in North America (Bradley 2010; Kempf et al. 2023). Data sovereignty, in the most general sense, ensures that data are subject to the laws and governance structures of the country or nation where they are collected. Indigenous Data Sovereignty is the right of Indigenous peoples to own and govern data about their communities, resources, and lands, meaning they are in control of how these data are accessed and used (Kukutai and Taylor 2016; Smith 2016; Rainie et al. 2019; McCartney et al. 2022; Diviacchi 2023). Indigenous Data Sovereignty can be implemented through Indigenous

https://doi.org/10.1017/pab.2024.61 Published online by Cambridge University Press

data governance, which respects and leverages the values, traditions, and roles that communities have for the care and use of their data (Carroll et al. 2019; Jennings et al. 2023). The CARE Principles for Indigenous Data Governance (Collective Benefit, Authority to Control, Responsibility, and Ethics) were developed to ensure that data collected on Indigenous lands will ultimately benefit the peoples of those lands and that this will be conducted in a manner that is not harmful to their communities (Carroll et al. 2020; www.gida-global.org/care). The CARE Principles guide researchers to include Indigenous peoples in data governance while increasing their access to, use of, and benefit from these data (Carroll et al. 2021). Adopting the CARE Principles in paleobiological research is critical for establishing more mutually beneficial research projects involving Indigenous lands across the world (Jennings et al. 2023; Kempf et al. 2023). First, nonlocal paleobiologists must be highly proactive in their outreach to and engagement with local Indigenous peoples, tribes and tribal-serving organizations, as it is the right of these groups to decide what data can be collected and shared and how that may be undertaken (Table 1). Visiting paleobiologists need to then work to understand and respect the wishes of Indigenous communities, especially when restrictions are placed on collection and sharing (publishing) of certain data (Jennings et al. 2023; Kempf et al. 2023). Implementation of the CARE Principles, in collaboration with local tailored resources (e.g., the AIATSIS Code for Australian territories) is already reframing research partnerships and data stewardship in ecology, conservation science, and the geosciences beyond paleobiology (Taitingfong and Carroll 2023; O'Brien et al. 2024). In paleobiology, there is enormous potential to apply these principles to build more ethical data stewardship practices, infrastructures, and technologies.

Fossils are essential for all kinds of paleobiological research, yet not all fossils are made available for scientific use. Some argue that commercial fossil collecting (i.e., collecting fossils to sell for profit) ultimately benefits both science and the seller (Larson and Russell 2014), but many others are concerned about the loss of these specimens to science and the public, as it poses a threat to data equity and accessibility (Shimada et al. 2014). Numerous high-profile auctions of exceptionally complete vertebrate fossils have drawn criticism from the paleontological community, particularly with regard to their enormous price tags (Reynolds 2018; Greshko 2020). At millions of U.S. dollars, these specimens are often far outside the budget of natural history museums, which leads to the loss of paleobiological data to commercial ventures (Lukiv 2024). Even if some scientists were to have the financial means to buy certain fossils (e.g., specimens with lower price tags), this would eventually create a hierarchy among scientists: wealthier scientists would be able to gather more data than those without such resources. While the impact of these activities is primarily felt by vertebrate paleontologists, unethical practices related to commercial interests can lead to other critical issues, such as irreparable damage and loss of access to fossil sites (Raja and Dunne 2022; Swallow et al. 2023), which disproportionately affects countries with fewer resources to designate and protect important fossil sites and perpetuates global data inequity (Kumar 2018; Gutiérrez-Marco and García-Bellido 2022).

Building awareness of these issues is a necessary first step in improving data-collection practices in paleobiology and moving toward equity. Several paleontological societies have committed to developing and providing guidelines to their memberships. For example, the Society for Vertebrate Paleontology provides guidance documents for working with and publishing on amber from Myanmar, as well as the commercial sale of vertebrate fossils (see www.vertpaleo.org/governance-documents). These guidelines were developed through specially formed working groups composed of researchers with experience in these areas and have since been applied to the society's journal, the *Journal of Vertebrate Paleontology* (Barrett and Johanson 2020). However, these actions can come with a significant time lag between the catalyst and implementation. In the case of Myanmar amber, the human rights abuses associated with amber mining in the north of the country were reported on by the United Nations several years before the paleontological community began developing guidance, and the commercial nature of fossils in Myanmar amber had been widely known for decades (Zin-Maung-Maung-Thein and Khin Zaw 2021; Dunne et al. 2022). Other academic publishers are becoming increasingly aware of unethical and inequitable paleobiological data collection and are establishing stricter editorial standards, including requesting official documents outlining fossil provenance and ethical declarations. For example, *Nature Ecology and Evolution* and *Palaeontology* have strict editorial policies on the publication of research surrounding Myanmar amber. The majority of other journals catering to paleobiology currently lag further behind in their implementation of such policies, highlighting how much work is yet to be done to preserve ethical and equitable data collection and sharing (Table 1).

Data Storage and Curation

Paleobiological Databases

Paleobiological data come in various forms and are stored in a variety of different places: physically (as specimens in museums and research institutions) and digitally (in large databases, various dedicated online repositories, and supplementary files). The way in which paleobiological data are stored, maintained, and managed has far-reaching implications for data equity. The issue of data equity is an integral component of the FAIR Data Principles, which are aimed at making data Findable, Accessible, Interoperable, and Reusable (Wilkinson et al. 2016). These principles, which have rapidly come to define best practices in the management of research data, are relevant to all stakeholders involved with paleobiological data, including data collectors, curators, managers, publishers, repositories, and users. However, they have not yet been widely applied across all facets of our field, including as a required framework for data management and sharing in funding applications and as standard principles in research institutions (Table 1).

The establishment of online data repositories, such as the Paleobiology Database (PBDB; paleobiodb.org; Uhen et al. 2023), Neotoma (neotoma.org), Triton (Fenton et al. 2021), the Geobiodiversity Database (www.geobiodiversity.com; Fan et al. 2013), DigiMorph (https://digimorph.org), and MorphoSource (Boyer et al. 2016; https://www.morphosource.org), has undoubtedly increased the accessibility and usability of paleobiological data over the last half century. The majority of these databases were instigated by, are physically stored in, and are managed by teams based in countries of the Global North, particularly in North America and western Europe (Fig. 2). Many paleobiological databases are established with a particular goal in mind (e.g., BioDeepTime; biodeeptime.github.io: cross-scale time-series analysis; Smith et al. 2023a) or were developed to answer specific paleobiological questions (e.g., PBDB, magnitude of Phanerozoic diversification; Alroy et al. 2001). Over time, paleobiological databases can morph and expand in various directions, often surpassing their original purposes, and new databases are created for new research purposes. This dynamism poses challenges for

Figure 2. Locations of non-governmental/community-developed digital databases that store paleobiological data or are regularly associated with studies in paleobiology. A tile grid map was used to avoid distorting the representation of the data that is typical of standard map projections.

the future of these repositories, such as increased need for better data integration, infrastructure updates, and long-term financial support, all of which have been highlighted extensively in biodiversity, ecology, and conservation science (Kamp et al. 2016; Kindsvater et al. 2018; Peterson and Soberón 2018; Isaac et al. 2020). This is further compounded by the current international funding landscape, which often does not cater to expenses related to digital infrastructure, producing disproportionate effects on the development of data infrastructure in regions of the Global South (Fig. 2). A notable exception in domestic funding is the NSF Geoinformatics funding scheme in the United States, which supports the development of community cyberinfrastructure to advance research and education in Earth science (https://new.nsf.gov/funding/opportunities/gi-geoinformatics). These challenges also point to the need for a multifaceted approach to paleobiological data storage and curation (e.g., systems that can integrate multiple forms of data input). Although several initiatives have attempted to do this for paleobiological data (e.g., ePANDDA and iDigBio; www.idigbio.org), an all-inclusive "data lake" for integration of data from across paleobiology and allied sciences is currently lacking and could facilitate even greater synthesis research in paleobiology and beyond. Importantly, this would also greatly enhance data equity through increased accessibility and data sharing (Drew et al. 2017).

Museum Collection Data

Large fossil occurrence databases, such as the PBDB and Neotoma, are invaluable resources for studies of past biodiversity. However, they rely primarily on published literature, which represents only a small proportion of the paleobiological data housed in natural history collections. One study conducted several decades ago demonstrated that the published record underestimates diversity within a specific time or geographic interval by three to five times, depending on the taxonomic group (Koch 1978). A more recent survey found that 9 museum collections in the United States held up to 23 times more unique localities for marine invertebrates than were contained in the equivalent geographic region in the PBDB (Marshall et al. 2018), highlighting how much data could be mobilized from museum collections, particularly invertebrate data. Indeed, some biodiversity databases, such as GBIF (gbif.org) and iDigBio (idigbio.org), do integrate specimen information alongside occurrence data, which could provide feasible examples for paleobiological databases.

Mobilizing "dark data" has the potential to shine a light on underutilized fossil material that, due to a lack of resources, could be missing from the published literature, especially from institutions in the Global South (Kaiser et al. 2023). Yet mobilization of this data is restricted by several factors. Substantial time, money, and effort are required to move through what can be a tedious digitization process. This often includes verifying information about how, when, and where a specimen was collected (e.g., upholding Darwin Core standards); holding requisite taxonomic expertise to check and update taxonomic assignments; detailed photography with specialized equipment (e.g., StackShot photography); and entry into databasing software (e.g., Specify). This information then may or may not be integrated into broader tools such as iDigBio, GBIF, or the field-specific databases mentioned earlier (Nelson et al. 2012; Paterson et al. 2016; Allmon et al. 2018; Marshall et al. 2018). This extensive work can only be carried out by larger, resource-rich institutions (Booth et al. 2021), further highlighting the interconnectedness of data equity challenges in paleobiology.

Just as data from online data repositories are not an accurate reflection of the fossil record, museum collections are not complete reflections of fossil-bearing outcrops (Lieberman and Kaesler 2000). Studies have found that both invertebrate (Whitaker and Kimmig 2020; Nanglu and Cullen 2023) and vertebrate (Davis and Pyenson 2007) collections show anthropogenically introduced biases, such as those based on gender, sex, and race, which can result in greater uncertainty in diversity and abundance estimates and have important implications for data equity (Das and Lowe 2018; Cooper et al. 2019). Specific collection criteria (e.g., set by institutions) and methodological choices by collectors (e.g., what pieces of information to record) are likely to be invisible to data users, meaning that these issues are challenging to identify and mitigate, especially for researchers who are more isolated due to geography, economics, and politics. Future work should endeavor to report the differences between collection protocols and initial sampling of material to provide transparency and allow these issues to be mitigated (e.g., Nanglu and Cullen 2023).

Natural history collections are bastions in which paleobiological research is rooted, yet they are also facing significant challenges across the globe. In the Global North, reduced government funding due to political and socioeconomic changes (Dalton 2003; Kemp 2015; Zamudio et al. 2018) has led to museums needing to reorient themselves as market actors (DesRoches 2015). One such example is related to "new museology," a necessary and relevant discourse around the roles of museums in society and politics (Vergo 1989), following which the museum landscape has diversified to reorient museums to include a greater focus on societal and political issues (McCall and Gray 2014). While this is a necessary step to building more equitable and inclusive museum spaces, it can also be argued that this has hastened a shift in museum priorities toward entertainment and education (DesRoches 2015). This overall "marketization" of museums has therefore resulted in a change of organizational approach from an internal focus on scholarship and curation to an externally oriented corporatist model of growth (McCall and Gray 2014; DesRoches 2015). The prioritization of short-term economic goals over long-term values results in staff cuts, the commodification of labor, and reduced financial support for collection maintenance (Suarez and Tsutsui 2004; DesRoches 2015; Miller et al. 2020). With improvements to collections being reduced to triage and more critical data (specimens, archive materials, etc.) being added year upon year, stockpiled information will be increasingly difficult to access. Further exacerbating this situation is the "widening role" of museum curators, who are increasingly tasked with expanded managerial, administrative, educational, outreach, conservation, and digitization responsibilities (discussed in the next section) alongside traditional curation (McCall and Gray 2014). Inadequate staffing of museum collections can also lead to the loss and degradation of both physical material and data, creating a vicious cycle wherein the workload is ever increasing. In natural history museums of the Global South, these same challenges exist but are greatly exacerbated through the legacy of colonial practices and socioeconomic inequalities (Booth et al. 2021). Many fossil specimens from the Global South have been, and continue to be, transported to repositories in the Global North, making them largely inaccessible to local communities and curators. Repatriation of fossils is one important step in rebalancing global inequalities in paleobiology, but it requires careful considerations about infrastructure, as well as financial resources and available expertise (Cisneros et al. 2021; Sebuliba et al. 2021; Zin-Maung-Maung-Thein and Khin Zaw 2021; Stewens et al. 2022). This highlights the need for a multifaceted approach to improving data equity in paleobiology that not only encompasses technological advancements, but also historical, socioeconomic, and political factors (Table 1).

Digitization

Advances in technology and computing over the last 50 years have allowed paleobiologists to develop an increasing diversity of tools to collect, collate, store, share, analyze, and visualize paleobiological data. In particular, the process of digitization has democratized and greatly improved the accessibility of data in paleobiology and allied sciences (Maschner and Schou 2013; Drew et al. 2017; Science Europe 2018; Nagaraj et al. 2020; Nagendra et al. 2024). Digitization transforms data from physical material, such as fossil specimens into a digital format (e.g., digital images, occurrence data, and morphological measurements), typically stored in digital repositories and databases. However, access to these digital data, as well as the resources to manage, store, and analyze these data, are not equitable across the whole paleobiological community.

Importantly for paleobiology, it is not just the digital versions of the data that require long-term preservation, the physical materials (e.g., fossil specimens) that are foundational to these data also require long-term preservation, which is often as, or even more, challenging. Natural history museums are critical for studies of past, present, and future life on Earth, particularly as they provide data that have enormous potential for tackling the current biodiversity crisis (Suarez and Tsutsui 2004; Plotnick et al. 2016; Meineke et al. 2018). They not only serve research and educational needs, but also provide a high monetary and social return for communities (Booth et al. 2021; Popov et al. 2021). Many natural history museums across the world are therefore committed to digitizing their collections, thus providing wider access to their fossil data (Nelson and Ellis 2019; Bakker et al. 2020; Hedrick et al. 2020; Sandramo et al. 2021). Digital representations of physical fossil specimens (e.g., photographs, 3D scans, or genetic sequences) have dramatically expanded the impact of natural history collections, making them more accessible and transforming research and researchers (Drew et al. 2017; Blackburn et al. 2024).

Several museums have already made subsets of their data publicly accessible on their websites, for example, the Natural History Museum in London, UK, the Smithsonian Institution in Washington DC, USA, and the Paleontological Research Institution in Ithaca, New York, USA (Hendricks et al. 2015). However, digitization (and the associated data storage) requires a significant amount of resources (e.g., financial) that are not available equally to all museums (Vollmar et al. 2010; Allmon et al. 2018). It also requires a diverse array of expertise, not only in terms of technological and museum expertise, but also regarding taxonomic expertise to ensure that preexisting errors and biases are not exaggerated. Above all else, digitization must fit the requirements and needs of those who use the data. Some digitization processes result in images being made available online, which is both attractive to the general public wanting to explore the content of various collections and useful for researchers and students wishing to access anatomical, morphological, and taxonomic information for specimens without the need to access the specimens in person. One example of a project to increase the accessibility of fossil specimens through digital 2D and 3D representations is the University of Michigan Museum of Paleontology's UMORF project (University of Michigan Online Repository of Fossils; https://umorf.ummp.lsa.umich.edu/wp/). The goal of the project is to serve a range of different user communities, from researchers and students to the general public, as well as to highlight the type and figured collection (specimens that are the basis for descriptions of new species and new interpretations of known species) and the parts of the museum's collection that could be used for comparative study.

Historical and Global Inequalities

Digitization of museum specimens should also be carried out with an awareness of historic injustices and inequalities, otherwise it has the capacity to perpetuate these issues (Kaiser et al. 2023). Many natural history collections have colonial or exploitative roots, and digitization of these data integrates assumptions about communities, capacities, and values that can reinforce inequalities in paleobiology (Heumann et al. 2018; Cisneros et al. 2022; Kaiser et al. 2023). In 2021, the Biodiversity Heritage Library (BHL), the largest open access digital library for biodiversity literature and archives, adopted an Acknowledgement of Harmful Content. This initiative acknowledges the existence of harmful content in many of the BHL's collections, which reflects centuries of historical decisions, practices, and colonial processes. The BHL website provides users with resources to critically evaluate content alongside an opportunity to report instances of harmful content through a feedback form (www.about.biodiversitylibrary.org/about/harmful-content). At the Museum für Naturkunde in Berlin, Germany, interdisciplinary researchers and museum staff are leading an ambitious digitization project that aims to increase the accessibility and equity of the Tendaguru dinosaur collection. Between 1909 and 1913, during German colonial rule, countless dinosaur fossils were taken from the Tendaguru Formation in southeastern Tanzania, then German East Africa (Schwarz and Heumann 2023). The project, funded by the German research foundation (Deutsche Forschungsgemeinschaft; DFG), is working to digitize a vast amount of data from the Tendaguru dinosaur collection, including the fossils themselves, through photographs, 3D models, and archival material associated with the colonial expeditions. The digitized material will be stored on a single data platform, which will enhance both the accessibility and transparency of the collection and will enable more equitable research development with Tanzanian colleagues. Furthermore, the project is specifically being conducted within the FAIR framework (Wilkinson et al. 2016), highlighting the potential of this framework for similar future projects in paleobiology.

Data collection in paleobiology often involves visiting museums or gathering digital data from museum specimens, yet this can be logistically, financially, or politically infeasible—or even impossible—for many researchers across the globe (Bezuidenhout and Chakauya 2018). Most major natural history museums are located in high-income countries of the Global North, such as in the United States or Europe, and traveling to these museums from lower-income countries can carry a heavy financial and administrative burden. Some scientists simply will not be granted entry to certain countries solely on the basis of their citizenship (Talavera-Soza 2023; Chugh and Joseph 2024). Further compounding this issue is the "digital divide" (Lythreatis et al. 2022), whereby different demographic regions have varying degrees of access to the tools and resources necessary for processing and analyzing digital data. In paleobiology, the digital divide often shapes the kind of work that researchers in particular regions can engage in (Abungu 2002; Mogajane 2022; Sánchez Membrilla 2024). For example, licensed computer software essential for the processing of 3D images and scans can be prohibitively expensive, especially when combined with the need for powerful devices to run such software. The urgent need for greater resources to be pooled into increasing the accessibility of museum collections was highlighted during the COVID-19 pandemic. At the height of the pandemic, researchers were not able to travel to collections for study, but with the urgency of the pandemic noticeably reduced and the changing landscape in the museum sector, the realities of making museum collections more accessible are becoming increasingly complex. Equitable research collaborations that are built on mutual trust and the sharing of resources are critical to overcoming many accessibility barriers. Researchers based in proximity to specimens or with expertise in certain data repositories can be invaluable connections for those who are more marginalized.

Data Sharing

Data Accessibility in Paleobiology

Since the 1970s, paleobiology has continued to embrace new and emerging digital technologies, and a knock-on effect of this is that data sharing has become increasingly simple. The sharing of data is an integral part of the research process, as it permits increased access to knowledge for all and improves transparency and reproducibility in science. Sharing research data can also have individual benefits, such as increased citation rates (Piwowar et al. 2007). However, uptake of open access (OA) and open data practices is not uniform across the world (see Fig. 2). For academics within low- and middle-income countries, current academic rewards (citations, altmetrics, institutional visibility, etc.) are severely compromised by structural limitations on data generation and aforementioned failures in data-collection practices, such as parachute science. This results in a disincentivization of participation in Open Data (Bezuidenhout and Chakauya 2018). Data sharing, consequently, requires an environment where there is sufficient academic security and benefit for the contributing scientists (Bezuidenhout and Chakauya 2018; Smith et al. 2023b).

Efforts in increasing social and scientific equity within paleontology has meant that there has been an overall positive trend toward open science and open data in recent years, fueled in part by new infrastructure and policies introduced by funding agencies and scientific journals (e.g., means-tested fee waivers for OA publishing). However, there is not yet a consistent requirement or protocol for the digital storage or sharing of paleobiological data (Rowe and Frank 2011; Davies et al. 2017; Dillon et al. 2023, Smith et al. 2023b). In this regard, paleobiology is lagging somewhat behind other allied disciplines such as the geospatial research and environmental sciences that are embracing community standards for data repositories, access, and reuse (Seltmann et al. 2018; Kinkade and Shepherd 2021; Crystal-Ornelas et al. 2022).

As paleontology data resources proliferate, the broad application of the TRUST principles (Transparency, Responsibility, User focus, Sustainability, and Technology) may serve to ensure good governance of shared resources and proper attribution for all contributions. The TRUST principles were developed to demonstrate the trustworthiness of digital repositories (Lin et al. 2020). This includes communicating a clear mission statement and repository policies, which would facilitate paleobiological data discovery and provide governance for necessary long-term preservation of data. These principles also provide a common framework to facilitate discussion and implementation of best practices in digital preservation by all stakeholders, including researchers, their institutions, academic publishers, and the digital repositories. For these principles to serve their intention, data equity must be centered in all research processes from data access to data entry, use, and attribution.

Principles in practice require guidelines and facilitation. Acknowledging the increase in quantitative studies in paleobiology due to data sharing, some journals have introduced data editors who ensure that all data related to a publication are available, together with associated materials (e.g., coding scripts) (Table 1). These editors serve to enhance the reproducibility of studies and are currently installed at journals such as *Proceedings of the Royal Society B*, *Journal of Evolutionary Biology*, and the *American Naturalist*. Some journals, such as those published by the British Ecological Society, have produced guides for authors on the topics of data management and reproducible code (www.britishecologicalsociety.org/publications/better-science). To further promote better data-sharing practices, paleobiological journals could enforce requirements for authors to make their data freely accessible in data repositories such as MorphoBank, the Paleobiology Database, and Phenome10K (phenome10k.org), instead of in supplementary data files. Indeed, some paleobiological journals (e.g., *Paleobiology* and *Journal for Vertebrate Paleontology*) already do this through associations with Figshare, Dryad, Zenodo, and MorphoBank. Although this entails an additional step in the publication process, authors are also likely to benefit through greater citations of their work when data are shared (Colavizza et al. 2020; Dorta-González et al. 2021).

Open Access Publishing

Open access (OA), the movement that seeks to grant free and open online access to academic information, is becoming increasingly more widespread, including within the field of paleobiology (Fig. 3). Despite the promise to make science more inclusive, capacities to engage with OA vary considerably across regions, institutions, and demographics (Bezuidenhout and Chakauya 2018; Ross-Hellauer et al. 2022). For example, green and gold OA options do not charge fees to readers, but charge fees to authors, making them unattainable for many researchers and institutions who may lack the means to pay the large fees required. For example, many funding agencies in India do not support publication charges, making it impossible for researchers to choose OA for their publications. Without a free OA or generous waiver policy, these researchers are excluded from circulating their published articles to a wider audience.

In 2014, the Alliance of German Science Organisations initiated Projekt DEAL (now known as DEAL Konsortium; www.deal-konsortium.de), which sought to negotiate transformative OA agreements with the largest commercial publishers of academic journals (Elsevier, Springer Nature, and Wiley). The consortium was successful in negotiations and continues to work toward fair pricing structures for academic publishing with the aim of increasing accessibility and visibility of researchers' work (Vogel 2023). In 2018, Plan S was launched by a consortium of national funding agencies in Europe, which requires researchers who benefit from state funding to publish their work in open repositories or journals. Although this mandate will only apply to authors who produced approximately 6% of the world's papers (based on an estimate from 2017), it is still expected to make a sizable impact in the long term, as the mandate applies to about one-third of papers published in *Nature* and *Science* (Brainard 2020).

The diamond OA publishing model, in which fees are not charged to authors or readers, addresses many issues related to financial constraints. Some journals catering to paleobiological research already adhere to the diamond OA model, including *Palaeontologia Electronica* and *Lethaia*. Despite the obvious advantages for increasing access and data equity, there continues to be a noticeable lack of transparent open science publications in paleobiology (Drage and Wong Hearing 2023). Preprints are an increasingly popular option, not only for researchers to showcase their work to a public audience, but also to gain additional feedback beyond the closed peer-review system (Sarabipour et al. 2019). Preprint services, such as bioRxiv, EcoEvoRxiv, and EarthArXiv, that cater to paleontological research papers, are rapidly becoming more popular.

Paleobiology recently announced that it will be transitioning entirely to an OA model, with a four-tiered approached based on agreements with institutions and funding bodies, locations of authors, and a waiver request form. Encouragingly, when announced, Cambridge University Press and the Paleontological Society stated a commitment to approving all waivers not covered by other funding sources. This move appears to bring *Paleobiology*'s publishing model to meet the criteria for diamond OA. Bibliometric data indexed by Web of Science show that the total number of articles published in *Paleobiology* has remained steady over the last five decades, while the number of OA articles has generally increased since the mid-1990s (Fig. 3). Fifteen percent of all *Paleobiology* articles indexed by Web of Science in the last 47 years are available under green, gold, or hybrid OA agreements (Web of Science 2024). In 2023, OA articles accounted for more than three-quarters (76%) of publications in *Paleobiology* (Fig. 3). Looking at the trends for *Paleobiology*, the field of paleobiology is comparatively more open than it was a decade ago (Fig. 3), yet there is much room for improvement. The majority of paleobiological journals still use a publishing model that is unfair, inequitable, and unsustainable for global science. Models such as diamond OA promote greater flexibility, accessibility, and data equity, and there is an increasing appetite within the research community for transformative changes to the status quo (Drage and Wong Hearing 2023).

Language Barriers

English remains the lingua franca of scientific research, including paleobiology. Today, 98% of all scientific publications are estimated to be in English (Gordin 2015). In the last 30 years, 92% of publications recorded in the PBDB were written in English, with Chinese, German, French, and Spanish making up the majority of the remainder (Raja et al. 2022). This dominance of English disadvantages paleobiologists for whom English is a secondary language or who are based in countries with low English proficiency (Ramírez-Castañeda 2020). In paleobiological research, the dominance of English could lead to biases through the exclusion of non-English publications, for example, in literature searches, as has been demonstrated in ecology and biodiversity research (Amano et al. 2016; Konno et al. 2020; Nuñez and Amano 2021). Within publishing, journals should adopt linguistically inclusive policies to overcome language barriers, including providing translation tools on their web pages, promoting the use of non–English language references, and providing author guidelines in multiple languages (Arenas-Castro et al. 2024). *Paleobiology* is among a number of journals, including *Palaeontologia Electronica*, *Integrative Organismal Biology*, and *Geodiversitas*, that permit authors to submit manuscript abstracts in several different languages. In many cases, these journals are based in countries where English is not the first language, such as in the case of *Revista Brasileira de Paleontologia*, and multiple language abstracts allow them to be accessed by a wider audience and more discoverable in more online searches (Amano et al. 2021a; Arenas-Castro et al. 2024). Out of 55 journals catering to paleobiology, a quarter (n = 13, 24%) are indexed by Web of Science as being multilingual in some way

https://doi.org/10.1017/pab.2024.61 Published online by Cambridge University Press

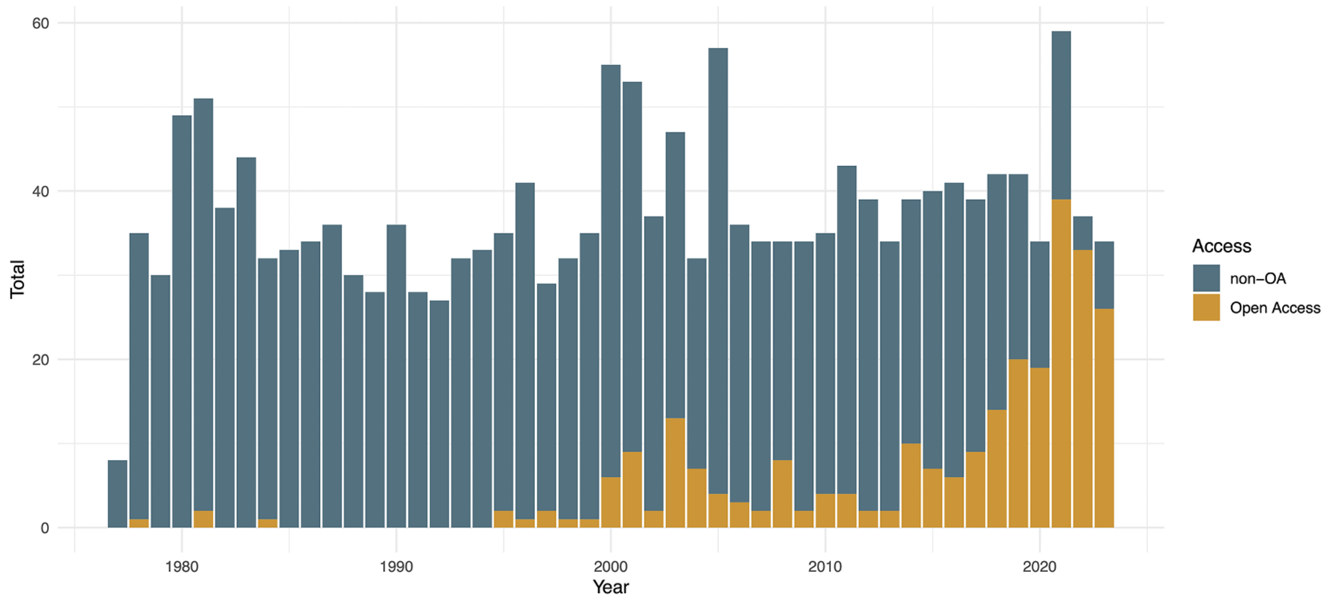

Figure 3. The total number of articles published in the journal *Paleobiology* from 1977 to 2023 according to the data indexed by Web of Science (2024; see reference for access details). The number of open access (OA) articles has increased steadily over time since the 1990s.

(Smith et al. 2023b). Journals and venues catering to paleobiological research should continue to embrace multilingualism, as it is a simple step toward increasing data equity for both paleobiologists and the general public (Table 1).

Future Outlook and Actions

Over the past half century, paleobiology has undergone a computational revolution that has given rise to a multitude of new avenues for recording, storing, and analyzing data on the history of life on Earth. With these advances, the amount of data available for research has grown considerably, accompanied by an expansion in the definition of paleontological data. Paleobiology once almost exclusively comprised data as counts of taxa at localities from different geological times. Now, paleontological data consists of terabytes in single images and high-resolution 3D models, as well as databases of millions of fossil occurrences, stratigraphic units, paleoenvironmental variables, and even molecular signatures. This translates to an increasing array of exciting opportunities for new research questions, but also to a critical responsibility to ensure that our data tools and infrastructures continue to innovate in order to best serve our diverse community.

In this review, we have highlighted how individual and systemic action is required to continue increasing data equity in paleobiology and to tackle ongoing challenges related to inequality, accessibility, and sustainability.

Paleobiologists can engage with data equity in a multitude of ways. The actions recommended here are by no means exhaustive, but instead should be considered a minimum requirement for all those working within and governing the paleobiology research community. There is undoubtedly a need for greater governmental and institutional support for fundamental resources that increase data equity, such as for digitization and OA publishing. However, there are still numerous actions that can be taken by individuals, teams, institutions, funders, journals, and societies to center and enhance data equity in paleobiological research.

Individuals should be proactive in ensuring that their data are collected in an ethical, equitable, and sustainable manner and can work toward promoting data equity within their collaborative networks and teams (Table 1). This can include regularly engaging with the topic of data equity through literature and other media and conducting multilingual data and literature searches. Institutions can support paleobiologists, curators, and students by offering (or even mandating) training on data equity protocols and principles, as well as providing robust infrastructure for managing and sharing data (Table 1). Academic journals that are not already working toward increasing data equity and openness should make commitments to do so. Increasing data equity requires strict data repository protocols, such as those based on the FAIR, CARE, and TRUST principles (Table 1). Journals should strive to have linguistically inclusive policies, including translation services and multi-language abstracts (Amano et al. 2021b; Arenas-Castro et al. 2024). Academic societies can also work to promote data equity, for example, through advocating to relevant government bodies for legal protection and safeguarding of geoheritage and museums (Table 1).

We cannot know with certainty what data advances will follow in the next 50 years or further into the future. Regardless, we paleobiologists have a responsibility to safeguard the potentiality of those future data for everyone in our community by developing and committing to ethical, sustainable, and equitable data practices today.

Acknowledgments. We are grateful to colleagues for invaluable wider discussions on the topics featured in this article, particularly, Á. Kocsis, L. Mulvey, and T. Clements. We also thank the editors of the 50th Anniversary Issue for inviting us to contribute this article and promoting such a critical topic in our field. We are enormously grateful for the constructive and positive reviews of B. Allen, H. L. Kempf, and an anonymous reviewer. E. M. Dunne was supported by an Emerging Talents Initiative grant at FAU Erlangen-Nürnberg. D.C. was supported by a SERB Core Research Grant (CRG/2022/001658). C.D.D. was supported by a Royal Society grant (RF_ERE_210013). E. M. Dillon was supported by an Earl S. Tupper Postdoctoral Fellowship from the Smithsonian Tropical Research Institute. E. M. Dowding was supported by the Paleosynthesis Project and the Volkswagen Stiftung (Az 96 796). P.L.G was supported by the University of São Paulo (grant no. 22.1.09345.01.2).

Competing Interests. The authors declare no competing interests.

Data Availability Statement. All data are available on GitHub: https://github.com/emmadunne/data_equity.

Code Availability Statement. All code used to manipulate and visualize data for this manuscript are available on GitHub: https://github.com/emmadunne/data_equity.

Literature Cited

Abungu, L. 2002. Access to Digital Heritage in Africa: bridging the digital divide. *Museum International* **54**:29–34.

Aldrich, R. 2009. Colonial museums in a postcolonial Europe. *African and Black Diaspora: An International Journal* **2**:137–156.

Allmon, W. D., G. P. Dietl, J. R. Hendricks, and R. M. Ross. 2018. Bridging the two fossils records: paleontology's "big data" future resides in museum collections. *In* Gary D. Rosenberg and Renee M. Clary, eds. *Museums at the forefront of the history and philosophy of geology: history made, history in the making*. Geological Society of America Special Paper 535. https://doi.org/10.1130/2018.2535(03).

Alroy, J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fürsich, T. A. Hansen, et al. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences USA* **98**:6261–6266.

Amano, T., and W. J. Sutherland. 2013. Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. *Proceedings of the Royal Society B* **280**:20122649.

Amano, T., J. P. González-Varo, and W. J. Sutherland. 2016. Languages are still a major barrier to global science. *PLoS Biology* **14**:e2000933.

Amano, T., C. Rios Rojas, Y. Boum II, M. Calvo, and B. B. Misra. 2021a. Ten tips for overcoming language barriers in science. *Nature Human Behaviour* **5**: 1119–1122.

Amano, T., V. Berdejo-Espinola, A. P. Christie, K. Willott, M. Akasaka, A. Báldi, A. Berthinussen, et al. 2021b. Tapping into non-English-language science for the conservation of global biodiversity. *PLoS Biology* **19**:e3001296.

Antell, G. T., R. B. J. Benson, and E. E. Saupe. 2023. Spatial standardization of taxon occurrence data—a call to action. *Paleobiology* **50**:177–193.

Arenas-Castro, H., V. Berdejo-Espinola, S. Chowdhury, A. Rodríguez-Contreras, A. R. M. James, N. B. Raja, E. M. Dunne, et al. 2024. Academic publishing requires linguistically inclusive policies. *Proceedings of the Royal Society B* **291**:20232840.

Asase, A., T. I. Mzumara-Gawa, J. O. Owino, A. T. Peterson, and E. Saupe. 2022. Replacing "parachute science" with "global science" in ecology and conservation biology. *Conservation Science and Practice* **4**:e517.

Bakker, F. T., A. Antonelli, J. A. Clarke, J. A. Cook, S. V. Edwards, P. G. P. Ericson, S. Faurby, et al. 2020. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ* **8**: e8225.

Barrett, P. M., and Z. Johanson. 2020. Myanmar amber fossils: a legal as well as ethical quagmire. *Nature* **586**:674–674.

Behrensmeyer, A. K., S. M. Kidwell, and R. A. Gastaldo. 2000. Taphonomy and paleobiology. *Paleobiology* **26**:103–147.

Benson, R. B. J., R. Butler, R. A. Close, E. Saupe, and D. L. Rabosky. 2021. Biodiversity across space and time in the fossil record. *Current Biology* **31**: R1225–R1236.

Bernard, R. E., and E. H. G. Cooperdock. 2018. No progress on diversity in 40 years. *Nature Geoscience* **11**:292–295.

Betts, H. C., M. N. Puttick, J. W. Clark, T. A. Williams, P. C. J. Donoghue, and D. Pisani. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nature Ecology and Evolution* **2**: 1556–1562.

Bezuidenhout, L., and E. Chakauya. 2018. Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics* **29**:39–54.

Blackburn, D. C., D. M. Boyer, J. A. Gray, J. Winchester, J. M. Bates, S. L. Baumgart, E. Braker, et al. 2024. Increasing the impact of vertebrate scientific collections through 3D imaging: the openVertebrate (oVert) Thematic Collections Network. *BioScience* **74**:169–186.

Boakes, E. H., P. J. K. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O'Connor, and G. M. Mace. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology* **8**:e1000385.

Booth, P., T. Navarrete, and A. Ogundipe. 2021. Museum open data ecosystems: a comparative study. *Journal of Documentation* **78**:761–779.

Boyer, D. M., G. F. Gunnell, S. Kaufman, and T. M. McGeary. 2016. Morphosource: Archiving and sharing 3-D digital specimen data. *Paleontological Society Papers* **22**:157–181.

Bradley, L. W. 2010. *Dinosaurs and Indians: paleontology resource dispossession from Sioux lands*. (First edition) ISBN 978-1478737063. ETD Collection for University of Nebraska–Lincoln. https://digitalcommons.unl.edu/dissertations/AAI3397860.

Brainard, J. 2020. A new mandate highlights costs, benefits of making all scientific articles free to read. *Science*. https://doi.org/10.1126/science.abg3557.

Carroll, S. R., D. Rodriguez-Lonebear, and A. Martinez. 2019. Indigenous Data Governance: strategies from United States Native Nations. *Data Science Journal* **18**. https://doi.org/10.5334/dsj-2019-031.

Carroll, S. R., I. Garba, O. L. Figueroa-Rodríguez, J. Holbrook, R. Lovett, S. Materechera, M. Parsons, et al. 2020. The CARE Principles for Indigenous Data Governance. *Data Science Journal* **19**. https://doi.org/10.5334/dsj-2020-043.

Carroll, S. R., E. Herczog, M. Hudson, K. Russell, and S. Stall. 2021. Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Scientific Data* **8**:108.

Carvalho, R. L., A. F. Resende, J. Barlow, F. M. França, M. R. Moura, R. Maciel, F. Alves-Martins, et al. 2023. Pervasive gaps in Amazonian ecological research. *Current Biology* **33**:3495–3504.e4.

Chugh, M., and T. Joseph. 2024. Citizenship privilege harms science. *Nature* **628**:499–501.

Cisneros, J. C., A. M. Ghilardi, N. B. Raja, and P. P. Stewens. 2021. The moral and legal imperative to return illegally exported fossils. *Nature Ecology and Evolution* **6**:2–3

Cisneros, J. C., N. B. Raja, A. M. Ghilardi, E. M. Dunne, F. L. Pinheiro, O. R. Regalado Fernández, M. A. F. Sales, et al. 2022. Digging deeper into colonial palaeontological practices in modern day Mexico and Brazil. *Royal Society Open Science* **9**:210898.

Close, R. A., S. W. Evers, J. Alroy, and R. J. Butler. 2018. How should we estimate diversity in the fossil record? Testing richness estimators using sampling-standardised discovery curves. *Methods in Ecology and Evolution* **9**:1386–1400.

Cohen, P. A., and R. B. Kodner. 2022. The earliest history of eukaryotic life: uncovering an evolutionary story through the integration of biological and geological data. *Trends in Ecology and Evolution* **37**:246–256.

Colavizza, G., I. Hrynaszkiewicz, I. Staden, K. Whitaker, and B. McGillivray. 2020. The citation advantage of linking publications to research data. *PLoS ONE* **15**:e0230416.

Coningham, R., N. Lewer, K. P. Acharya, K. Weise, R. B. Kunwar, A. Joshi, and S. Parajuli Khanal. 2024. Enabling equitable and ethical research partnerships in crisis situations: lessons learned from post-disaster heritage protection interventions following Nepal's 2015 earthquake. *Research Ethics* **20**:835–846.

Cooper, N., A. L. Bond, J. L. Davis, R. Portela Miguez, L. Tomsett, and K. M. Helgen. 2019. Sex biases in bird and mammal natural history collections. *Proceedings of the Royal Society B* **286**:20192025.

Cortés-Sánchez, M., M. D. Simón-Vallejo, J.-C. Corral, M. del C. Lozano-Francisco, J. L. Vera-Peláez, F. J. Jiménez-Espejo, A. García-Alix, et al. 2020. Fossils in Iberian prehistory: A review of the palaeozoological evidence. *Quaternary Science Reviews* **250**:106676.

Crystal-Ornelas, R., C. Varadharajan, D. O'Ryan, K. Beilsmith, B. Bond-Lamberty, K. Boye, M. Burrus, et al. 2022. Enabling FAIR data in Earth and environmental science with community-centric (meta)data reporting formats. *Scientific Data* **9**:700.

Cunningham, J. A., I. A. Rahman, S. Lautenschlager, E. J. Rayfield, and P. C. J. Donoghue. 2014. A virtual world of paleontology. *Trends in Ecology and Evolution* **29**:347–357.

Dalton, R. 2003. Natural history collections in crisis as funding is slashed. *Nature* **423**:575–575.

Das, S., and M. Lowe. 2018. Nature Read in Black and White: decolonial approaches to interpreting natural history collections. *Journal of Natural Science Collections*:4–14.

data.org. 2024. What is data equity and why does it matter? https://data.org/resources/what-is-data-equity-and-why-does-itmatter/.

Davies, T. G., I. A. Rahman, S. Lautenschlager, J. A. Cunningham, R. J. Asher, P. M. Barrett, K. T. Bates, et al. 2017. Open data and digital morphology. *Proceedings of the Royal Society B* **284**:20170194.

Davis, E. B., and N. D. Pyenson. 2007. Diversity biases in terrestrial mammalian assemblages and quantifying the differences between museum collections and published accounts: a case study from the Miocene of Nevada. *Palaeogeography, Palaeoclimatology, Palaeoecology* **250**:139–149.

DesRoches, D. 2015. The marketized museum: new museology in a corporatized world. *Political Economy of Communication* **3**:2–24.

Dillon, E. M., E. M. Dunne, T. M. Womack, M. Kouvari, E. Larina, J. R. Claytor, A. Ivkić, et al. 2023. Challenges and directions in analytical paleobiology. *Paleobiology* **49**:377–393.

Diviacchi, T. 2023. Indigenous data sovereignty and open data. *EveryONE*. Accessed August 15th 2024.

Dorta-González, P., S. M. González-Betancor, and M. I. Dorta-González. 2021. To what extent is researchers' data-sharing motivated by formal mechanisms of recognition and credit? *Scientometrics* **126**:2209–2225.

Drage, H. B., and T. W. Wong Hearing. 2023. Diamond open access with preregistration: a new publishing model for palaeontology. EarthArXiv. https://doi.org/10.31223/X5KH30.

Drew, J. A., C. S. Moreau, and M. L. J. Stiassny. 2017. Digitization of museum collections holds the potential to enhance researcher diversity. *Nature Ecology and Evolution* **1**:1789–1790.

Dunne, E. M., N. B. Raja, P. P. Stewens, Zin-Maung-Maung-Thein, and K. Zaw. 2022. Ethics, law, and politics in palaeontological research: the case of Myanmar amber. *Communications Biology* **5**:1–10.

Dutt, K. 2018. *How implicit bias and lack of diversity undermine science.* Scientific American. https://www.scientificamerican.com/blog/voices/how-implicit-bias-and-lack-of-diversity-undermine-science. Accessed 18th August 2024.

Dutt, K. 2020. Race and racism in the geosciences. *Nature Geoscience* **13**:2–3.

Emery, N. C., E. K. Bledsoe, A. O. Hasley, and C. D. Eaton. 2021. Cultivating inclusive instructional and research environments in ecology and evolutionary science. *Ecology and Evolution* **11**:1480–1491.

Fan, J., Q. Chen, X. Hou, A. I. Miller, M. J. Melchin, S. Shen, S. Wu, et al. 2013. Geobiodiversity Database: a comprehensive section-based integration of stratigraphic and paleontological data. *Newsletters on Stratigraphy* **46**: 111–136.

Fenton, I. S., A. Woodhouse, T. Aze, D. Lazarus, J. Renaudie, A. M. Dunhill, J. R. Young, and E. E. Saupe. 2021. Triton, a new species-level database of Cenozoic planktonic foraminiferal occurrences. *Scientific Data* **8**:160.

Finnegan, S., P. G. Harnik, R. Lockwood, H. K. Lotze, L. McClenachan, and S. S. Kahanamoku. 2024. Using the fossil record to understand extinction risk and inform marine conservation in a changing world. *Annual Review of Marine Science* **16**:307–333.

Gordin, M. D. 2015. *Scientific Babel*. University of Chicago Press, Chicago.

Greshko, M. 2020. "Stan" the *T. rex* just sold for $31.8 million—and scientists are furious. *Science*. https://www.nationalgeographic.com/science/article/stan-tyrannosaurus-rex-sold-at-auction-paleontologists-are-furious, accessed 13 March 2024.

Gutiérrez-Marco, J. C., and D. C. García-Bellido. 2022. The international fossil trade from the Paleozoic of the Anti-Atlas, Morocco. *Geological Society of London Special Publication* **485**:69–96.

Harris, F. 2015. Understanding human remains repatriation: practice procedures at the British Museum and the Natural History Museum. *Museum Management and Curatorship* **30**:138–153.

Hedrick, B. P., J. M. Heberling, E. K. Meineke, K. G. Turner, C. J. Grassa, D. S. Park, J. Kennedy, et al. 2020. Digitization and the future of natural history collections. *BioScience* **70**:243–251.

Hendricks, J. R., A. L. Stigall, and B. S. Lieberman. 2015. The *Digital Atlas of Ancient Life:* delivering information on paleontology and biogeography via the web. *Palaeontologia Electronica*. https://palaeo-electronica.org/content/2015/1269-commentary-digital-atlas-of-ancient-life, accessed 11 March 2024.

Heumann, I., H. Stoecker, M. Tamborini, and M. Vennen. 2018. *Dinosaurierfragmente: zur Geschichte der Tendaguru-Expedition und ihrer Objekte, 1906–2018*. Wallstein, Göttingen.

Hughes, A. C., M. C. Orr, K. Ma, M. J. Costello, J. Waller, P. Provoost, Q. Yang, C. Zhu, and H. Qiao. 2021. Sampling biases shape our view of the natural world. *Ecography* **44**:1259–1269.

Isaac, N. J. B., M. A. Jarzyna, P. Keil, L. I. Dambly, P. H. Boersch-Supan, E. Browning, S. N. Freeman, et al. 2020. Data integration for large-scale models of species distributions. *Trends in Ecology and Evolution* **35**:56–67.

Jagadish, H., J. Stoyanovich, and B. Howe. 2023. The many facets of data equity. J. *Data and Information Quality* **14**:27:1–27:21.

Jennings, L., T. Anderson, A. Martinez, R. Sterling, D. D. Chavez, I. Garba, M. Hudson, N. A. Garrison, and S. R. Carroll. 2023. Applying the "CARE Principles for Indigenous Data Governance" to ecology and biodiversity research. *Nature Ecology and Evolution* **7**:1547–1551.

Kaiser, K., I. Heumann, T. Nadim, H. Keysar, M. Petersen, M. Korun, and F. Berger. 2023. Promises of mass digitisation and the colonial realities of natural history collections. *Journal of Natural Science Collections* **11**:13–25.

Kamp, J., S. Oppel, H. Heldbjerg, T. Nyegaard, and P. F. Donald. 2016. Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. *Diversity and Distributions* **22**:1024–1035.

Kaufman, D. S., and **PAGES 2k Special-Issue Editorial Team**. 2018. Technical note: open-paleo-data implementation pilot—the PAGES 2k special issue. *Climate of the Past* **14**:593–600.

Kemp, C. 2015. Museums: the endangered dead. *Nature* **518**:292–294.

Kempf, H. L., H. C. Olson, P. M. Monarrez, L. Bradley, C. Keane, and S. J. Carlson. 2023. History of Native American land and natural resource policy in the United States: impacts on the field of paleontology. *Paleobiology* **49**: 191–203.

Kindsvater, H. K., N. K. Dulvy, C. Horswill, M.-J. Juan-Jordá, M. Mangel, and J. Matthiopoulos. 2018. Overcoming the data crisis in biodiversity conservation. *Trends in Ecology and Evolution* **33**:676–688.

Kinkade, D., and A. Shepherd. 2021. Geoscience data publication: practices and perspectives on enabling the FAIR guiding principles. *Geoscience Data Journal* **9**:177–186.

Koch, C. F. 1978. Bias in the published fossil record. *Paleobiology* **4**:367–372.

Konno, K., M. Akasaka, C. Koshida, N. Katayama, N. Osada, R. Spake, and T. Amano. 2020. Ignoring non-English-language studies may bias ecological meta-analyses. *Ecology and Evolution* **10**:6373–6384.

Kukutai, T., and J. Taylor. 2016. Data sovereignty for indigenous peoples: current practice and future needs. Pp. 1–22 *in* T. Kukutai and J. Taylor, eds. *Indigenous data sovereignty: toward an agenda*. CAEPR Research Monograph No. **38**. ANU Press, Canberra, Australia.

Kumar, S. 2018. Edge of extinction. *Science* **360**:22–25.

Larson, P., and D. Russell. 2014. The benefits of commercial fossil sales to 21st-century paleontology. *Palaeontologia Electronica* **17**(1):2E.

Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. LeVan, et al. 2020. The Extended Specimen Network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience* **70**: 23–30.

Lieberman, B. S., and R. L. Kaesler. 2000. The scientific value of natural history museum collections: the concept of completeness. *Paleontological Society Special Publication* **10**:109–118.

Lin, D., J. Crabtree, I. Dillo, R. R. Downs, R. Edmunds, D. Giaretta, M. De Giusti, et al. 2020. The TRUST Principles for digital repositories. *Scientific Data* **7**:144.

Lukiv, J. 2024. Dinosaur skeleton fetches record $44.6m at New York auction. *BBC News*, July 17.

Lythreatis, S., S. K. Singh, and A.-N. El-Kassar. 2022. The digital divide: a review and future research agenda. *Technological Forecasting and Social Change* **175**:121359.

Manias, C. 2015. Building *Baluchitherium* and *Indricotherium*: imperial and international networks in early-twentieth century paleontology. *Journal of the History of Biology* **48**:237–278.

Marshall, C. R., S. Finnegan, E. C. Clites, P. A. Holroyd, N. Bonuso, C. Cortez, E. Davis, et al. 2018. Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital revolution. *Biology Letters* **14**: 20180431.

Maschner, H. D. G., and C. D. Schou. 2013. Virtualization and the democratization of science: 3D technologies revolutionize museum research and access. *2013 Digital Heritage International Congress (DigitalHeritage)*:265–271.

Mayor, A. 2005. *Fossil legends of the First Americans*. Princeton University Press, Princeton, N.J.

McCall, V., and C. Gray. 2014. Museums and the "new museology": theory, practice and organisational change. *Museum Management and Curatorship* **29**:19–35.

McCartney, A. M., J. Anderson, L. Liggins, M. L. Hudson, M. Z. Anderson, B. TeAika, J. Geary, R. Cook-Deegan, H. R. Patel, and A. M. Phillippy. 2022. Balancing openness with Indigenous data sovereignty: An opportunity to leave no one behind in the journey to sequence all of life. *Proceedings of the National Academy of Sciences USA* **119**:e2115860119.

Meineke, E. K., T. J. Davies, B. H. Daru, and C. C. Davis. 2018. Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society B* **374**:20170386.

Miller, S. E., L. N. Barrow, S. M. Ehlman, J. A. Goodheart, S. E. Greiman, H. L. Lutz, T. M. Misiewicz, et al. 2020. Building natural history collections for the twenty-first century and beyond. *BioScience* **70**:674–687.

Mogajane, N. M. 2022. *Digital technology impact in the access, preservation and curation of indigenous artefacts at Mafikeng Museum*. Master's thesis, North-West University, South Africa.

Monarrez, P. M., J. B. Zimmt, A. M. Clement, W. Gearty, J. J. Jacisin, K. M. Jenkins, K. M. Kusnerik, et al. 2022. Our past creates our present: a brief overview of racism and colonialism in Western paleontology. *Paleobiology* **48**:173–185.

Muralidhar, S., and V. Ananthanarayanan. 2024. Women's representation in Indian academia and conferences. *Communications Biology* **7**:1–6.

Nagaraj, A., E. Shears, and M. de Vaan. 2020. Improving data access democratizes and diversifies science. *Proceedings of the National Academy of Sciences USA* **117**:23490–23498.

Nagendra, H., S. Goswami, and S. Mundoli. 2024. Climate data democracy is essential for climate action. *PLoS Climate* **3**:e0000447.

Nanglu, K., and T. M. Cullen. 2023. Across space and time: a review of sampling, preservational, analytical, and anthropogenic biases in fossil data across macroecological scales. *Earth-Science Reviews* **244**:104537.

Nelson, G., and S. Ellis. 2019. The history and impact of digitization and digital data mobilization on biodiversity research. *Philosophical Transactions of the Royal Society B* **374**:20170391.

Nelson, G., D. Paul, G. Riccardi, and A. Mast. 2012. Five task clusters that enable efficient and effective digitization of biological collections. *ZooKeys* **209**:19–45.

Nóbrega, R. L. B., P. H. L. Alencar, B. Baniwa, M.-C. Buell, P. L. B. Chaffe, D. Munduruku Pinto Correa, D. Munduruku do Santos Correa, et al. 2023. Co-developing pathways to protect nature, land, territory, and well-being in Amazonia. *Communications Earth and Environment* **4**:1–5.

North, M. A., W. W. Hastie, and L. Hoyer. 2020. Out of Africa: the underrepresentation of African authors in high-impact geoscience literature. *Earth-Science Reviews* **208**:103262.

Núñez, A.-M., J. Rivera, and T. Hallmark. 2020. Applying an intersectionality lens to expand equity in the geosciences. *Journal of Geoscience Education* **68**:97–114.

Nuñez, M. A., and T. Amano. 2021. Monolingual searches can limit and bias results in global literature reviews. *Nature Ecology and Evolution* **5**:264–264.

Nuñez, M. A., M. C. Chiuffo, A. Pauchard, and R. D. Zenni. 2021. Making ecology really global. *Trends in Ecology and Evolution* **36**:766–769.

Nuñez, M. A., T. August, S. Bacher, B. S. Galil, P. E. Hulme, T. Ikeda, M. A. McGeoch, et al. 2024. Including a diverse set of voices to address biological invasions. *Trends in Ecology and Evolution* **39**:409–412.

O'Brien, M., R. Duerr, R. Taitingfong, A. Martinez, L. Vera, L. L. Jennings, R. R. Downs, et al. 2024. Earth science data repositories: implementing the CARE Principles. *Data Science Journal* **23**. https://doi.org/10.5334/dsj-2024-037.

Pandolfi, L., P. Raia, J. Fortuny, and L. Rook. 2020. Editorial: evolving virtual and computational paleontology. *Frontiers in Earth Science* **8**:479.

Paterson, G., S. Albuquerque, V. Blagoderov, S. Brooks, S. Cafferty, E. Cane, V. Carter, et al. 2016. iCollections—digitising the British and Irish butterflies in the Natural History Museum, London. *Biodiversity Data Journal* **4**:e9559.

Payne, J. L., F. A. Smith, M. Kowalewski, R. A. Krause Jr., A. G. Boyer, C. R. McClain, S. Finnegan, P. M. Novack-Gottshall, and L. Sheble. 2012. A lack of attribution: closing the citation gap through a reform of citation and indexing practices. *Taxon* **61**:1349–1351.

Pérez Ortega, R. 2022. Rare dinosaur heads home as Germany agrees to return Brazilian fossil. *Science*. https://doi.org/10.1126/science.ade0219.

Peterson, A. T., and J. Soberón. 2018. Essential biodiversity variables are not global. *Biodiversity and Conservation* **27**:1277–1288.

Phillips, J. 1860. *Life on the Earth. Its origins and succession*. Cambridge, U.K.: Macmillan.

Piwowar, H. A., R. S. Day, and D. B. Fridsma. 2007. Sharing detailed research data is associated with increased citation rate. *PLoS ONE* **2**:e308.

Plotnick, R. E., F. A. Smith, and S. K. Lyons. 2016. The fossil record of the sixth extinction. *Ecology Letters* **19**:546–553.

Popov, D., P. Roychoudhury, H. Hardy, L. Livermore, and K. Norris. 2021. The value of digitising natural history collections. *Research Ideas and Outcomes* **7**:e78844.

Posselt, J. R. 2020. *Equity in science: representation, culture, and the dynamics of change in graduate education*. Stanford University Press, Stanford, Calif.

Rainie, S.C., Kukutai, T., Walter, M., Figueroa-Rodríguez, O.L., Walker, J., & Axelsson, P. (2019). Issues in open data: Indigenous data sovereignty. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The state of open data: Histories and horizons*. (pp. 300–319). Cape Town and Ottawa: African Minds and International Development Research Centre. http://stateofopendata.od4d.net

Raja, N. B., and E. Dunne. 2022. Fossil trafficking, fraud, and fakery. Pp. 61–79 *in* D. Yates and N. Oosterman, eds. *Art crime in context*. Springer, Cham, Switzerland.

Raja, N. B., E. M. Dunne, A. Matiwane, T. M. Khan, P. S. Nätscher, A. M. Ghilardi, and D. Chattopadhyay. 2022. Colonial history and global economics distort our understanding of deep-time biodiversity. *Nature Ecology and Evolution* **6**:145–154.

Ramírez-Castañeda, V. 2020. Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: the case of Colombian researchers in biological sciences. *PLoS ONE* **15**:e0238372.

Ranganathan, M., E. Lalk, L. M. Freese, M. A. Freilich, J. Wilcots, M. L. Duffy, and R. Shivamoggi. 2021. Trends in the representation of women among US geoscience faculty from 1999 to 2020: the long road toward gender parity. *AGU Advances* **2**:e2021AV000436.

Raposo, M. A., H. R. da Silva, B. C. S. Francisco, O. Vieira, O. V. da Fonseca, C. P. de Assis, G. M. Kirwan, et al. 2023. Is stability too revered in zoological nomenclature? *Zoological Journal of the Linnean Society* **199**:7–9.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. *Science* **177**:1065–1071.

Raup, D. M. 1991. The future of analytical paleobiology. *Short Courses in Paleontology* **4**:207–216.

Reitan, T., E. D. Martino, and L. H. Liow. 2024. Estimating relative species abundance using fossil data identified to different taxonomic levels. *Ecography* **2024**:e06866.

Reynolds, M. 2018. The dinosaur trade: how celebrity collectors and glitzy auctions could be damaging science. *Wired UK*, June 21.

Ross-Hellauer, T., S. Reichmann, N. L. Cole, A. Fessl, T. Klebel, and N. Pontika. 2022. Dynamics of cumulative advantage and threats to equity in open science: a scoping review. *Royal Society Open Science* **9**:211032.

Rowe, T., and L. R. Frank. 2011. The disappearing third dimension. *Science* **331**:712–714.

Sánchez Membrilla, S. 2024. *Indigenous collections at the Museum of World Culture: digitisation, decolonisation and other stories*. Master's thesis. Linnaeus University, Sweden.

Sandramo, D., E. Nicosia, S. Cianciullo, B. Muatinte, and A. Guissamulo. 2021. Unlocking the entomological collection of the Natural History Museum of Maputo, Mozambique. *Biodiversity Data Journal* **9**:e64461.

Sarabipour, S., H. J. Debat, E. Emmott, S. J. Burgess, B. Schwessinger, and Z. Hensel. 2019. On the value of preprints: an early career researcher perspective. *PLoS Biology* **17**:e3000151.

Schofer, E. 2003. The global institutionalization of geological science, 1800 to 1990. *American* Sociological Review **68**:730–759.

Schwarz, D., and I. Heumann. 2023. Virtual access to fossil and archival material from the German Tendaguru Expedition (1909-1913). Museum für Naturkunde. https://www.museumfuernaturkunde.berlin/en/science/virtual-access-fossil-and-archival-material-german-tendaguru-expedition-1909-1913, accessed 14 March 2024.

Science Europe. 2018. The rationales of open science: digitalisation and democratisation in research. Zenodo. https://doi.org/10.5281/zenodo.4925890.

Sebuliba, S., K. Wesche, and W. E. R. Xylander. 2021. Ready for restitution? Meeting challenges of colonial legacies in Africa's collections. *Bioscience* **71**:322–324.

Seddon, A. W. R., A. W. Mackay, A. G. Baker, H. J. B. Birks, E. Breman, C. E. Buck, E. C. Ellis, et al. 2014. Looking forward through the past: identification of 50 priority research questions in palaeoecology. *Journal of Ecology* **102**:256–267.

Seltmann, K., S. Lafia, D. Paul, S. James, D. Bloom, N. Rios, S. Ellis, et al. 2018. Georeferencing for Research Use (GRU): an integrated geospatial training paradigm for biocollections researchers and data providers. *Research Ideas and Outcomes* **4**:e32449.

Sepkoski, J. J. 1996. Patterns of Phanerozoic extinction: a perspective from global data bases. Pp. 35–51 *in* O. H. Walliser, ed. *Global events and event stratigraphy in the Phanerozoic: results of the international interdisciplinary cooperation in the IGCP-Project 216 "Global Biological Events in Earth History."* Springer, Berlin.

Shimada, K., P. J. Currie, E. Scott, and S. S. Sumida. 2014. The greatest challenge to 21st century paleontology: when commercialization of fossils threatens the science. *Palaeontologia Electronica* **17**:1–4.

Smith, A. B., and A. J. McGowan. 2011. The ties linking rock and fossil records and why they are important for palaeobiodiversity studies. *Geological Society of London Special Publication* **358**:1–7.

Smith, D. E. 2016. Governing data and data for governance: the everyday practice of Indigenous sovereignty. Pp. 117–136 *in* T. Kukutai and J. Taylor, eds. *Indigenous data sovereignty: toward an agenda*. CAEPR Research Monograph No. **38**. ANU Press, Canberra, Australia.

Smith, J., M. C. Rillo, Á. T. Kocsis, M. Dornelas, D. Fastovich, H.-H. M. Huang, L. Jonkers, et al. 2023a. BioDeepTime: a database of biodiversity time series for modern and fossil assemblages. *Global Ecology and Biogeography* **32**:1680–1689.

Smith, J. A., J. C. Handley, and G. P. Dietl. 2022. Accounting for uncertainty from zero inflation and overdispersion in paleoecological studies of predation using a hierarchical Bayesian framework. *Paleobiology* **48**:65–82.

Smith, J. A., N. B. Raja, T. Clements, D. Dimitrijević, E. M. Dowding, E. M. Dunne, B. M. Gee, et al. 2023b. Increasing the equitability of data citation in paleontology: capacity building for the big data future. *Paleobiology* **50**: 165–176.

Stefanoudis, P. V., W. Y. Licuanan, T. H. Morrison, S. Talma, J. Veitayaki, and L. C. Woodall. 2021. Turning the tide of parachute science. *Current Biology* **31**:R184–R185.

Stewens, P. P., N. B. Raja, and E. M. Dunne. 2022. The return of fossils removed under colonial rule. *Santander Art and Culture Law Review (SAACLR)* **2022**:89–114.

Suarez, A. V., and N. D. Tsutsui. 2004. The value of museum collections for research and society. *BioScience* **54**:66–74.

Sugimoto, C. R., N. Robinson-Garcia, D. S. Murray, A. Yegros-Yegros, R. Costas, and V. Larivière. 2017. Scientists have most impact when they're free to move. *Nature* **550**:29–31.

Swallow, B., M. Faulkner, and C. Dennis. 2023. Somerset beach damaged by "irresponsible" fossil hunters. *BBC News*, December 28.

Taitingfong, R., and S. Carroll. 2023. Implementing the CARE Principles for Indigenous data governance in biodiversity data management. *Biodiversity Information Science and Standards* **7**:e112615.

Talavera-Soza, S. 2023. Citizenship a determining factor in a geoscientist's career. *Nature Geoscience* **16**:550–551.

Trisos, C. H., J. Auerbach, and M. Katti. 2021. Decoloniality and anti-oppressive practices for a more ethical ecology. *Nature Ecology and Evolution* **5**:1205–1212.

Uhen, M. D., B. Allen, N. Behboudi, M. E. Clapham, E. Dunne, A. Hendy, P. A. Holroyd, et al. 2023. Paleobiology Database User Guide, Version 1.0. *PaleoBios* **40**(11):1–56.

Valenzuela-Toro, A. M., and M. Viglino. 2021. How Latin American researchers suffer in science. *Nature* **598**:374–375.

Vergo, P., ed. 1989. *The new museology*. Critical Views. Reaktion Books, London.

Vilhena, D. A., and A. B. Smith. 2013. Spatial Bias in the Marine Fossil Record. *PLoS ONE* **8**:e74470.

Vogel, G. 2023. *German science organizations strike open-access deal with Elsevier*. Science.

Vollmar, A., J. A. Macklin, and L. Ford. 2010. Natural history specimen digitization: challenges and concerns. *Biodiversity Informatics* **7**. https://doi.org/10.17161/bi.v7i2.3992.

Warnock, R. C. M., T. A. Heath, and T. Stadler. 2020. Assessing the impact of incomplete species sampling on estimates of speciation and extinction rates. *Paleobiology* **46**:137–157.

Web of Science. 2024. Results for *Paleobiology*. Clarivate. https://www.webofscience.com/wos/woscc/summary/2f42a9d1-b935-4eb7-a104-99e59719e90b-d4f9e23d/relevance/1, accessed 17 August 2024.

Webster, M. S. 2017. The extended specimen: emerging frontiers in collections-based ornithological research. *Studies in Avian Biology*. CRC Press, Taylor & Francis Group, Boca Raton, Fla.

Whitaker, A. F., and J. Kimmig. 2020. Anthropologically introduced biases in natural history collections, with a case study on the invertebrate paleontology collections from the middle Cambrian Spence Shale Lagerstätte. *Palaeontologia Electronica* **23**:1–26.

Wilkinson, M. D., M. Dumontier, Ij. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**:160018.

Yen, H. 2024. Fossils and sovereignty: science diplomacy and the politics of deep time in the Sino-American fossil dispute of the 1920s. *Isis* **115**:1–22.

Zamudio, K. R., A. Kellner, C. Serejo, M. R. de Britto, C. B. Castro, P. A. Buckup, D. O. Pires, et al. 2018. Lack of science support fails Brazil. *Science* **361**:1322–1323.

Zin-Maung-Maung-Thein and Khin Zaw. 2021. Parachute research is another ethical problem for Myanmar amber. *Nature Ecology and Evolution* **5**:707–707.

Zuroski, E. 2017. Situating the local in a global expedition: HMS *Challenger* expedition in New Zealand, 1874. *Journal of the Royal Society of New Zealand* **47**:107–111.