Check for updates

Received: 25 July 2022 | Accepted: 7 November 2022

DOI: 10.1111/2041-210X.14033

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

REVIEW

Ten (mostly) simple rules to future-proof trait data in ecological and evolutionary sciences

Alexander Keller[1] | Markus J. Ankenbrand[2] | Helge Bruelheide[3,4] | Stefanie Dekeyzer[5] | Brian J. Enquist[6,7] | Mohammad Bagher Erfanian[8] | Daniel S. Falster[9] | Rachael V. Gallagher[10] | Jennifer Hammock[11] | Jens Kattge[4,12] | Sara D. Leonhardt[13] | Joshua S. Madin[14] | Brian Maitner[15,16] | Margot Neyret[17] | Renske E. Onstein[4,18] | William D. Pearse[19] | Jorrit H. Poelen[20,21] | Roberto Salguero-Gomez[22] | Florian D. Schneider[17,23] | Anikó B. Tóth[24] | Caterina Penone[25]

[1]Cellular and Organismic Networks, Faculty of Biology, Ludwig-Maximilians-Universität München, Martinsried, Germany; [2]Center for Computational and Theoretical Biology, Julius-Maximilians-Universität Würzburg, Würzburg, Germany; [3]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; [4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany; [5]Flanders Marine Institute (VLIZ), InnovOcean Site, Oostende, Belgium; [6]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA; [7]The Santa Fe Institute, Santa Fe, New Mexico, USA; [8]Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran; [9]Evolution & Ecology Research Centre, University of New South Wales Sydney, Sydney, New South Wales, Australia; [10]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, New South Wales, Australia; [11]National Museum of Natural History, Smithsonian Institution, Washington, District of Columbia, USA; [12]Max Planck Institute for Biogeochemistry, Jena, Germany; [13]Plant-Insect Interactions, TUM School of Life Science Systems, Technical University of Munich, Freising, Germany; [14]Hawai'i Institute of Marine Biology, University of Hawai'i at Manoa, Kāne'ohe, Hawai'i, USA; [15]Department of Geography, University at Buffalo, Buffalo, New York, USA; [16]Department of Environment and Sustainability, University at Buffalo, Buffalo, New York, USA; [17]Senckenberg Biodiversity and Climate Research Center (SBik-F), Frankfurt, Germany; [18]Naturalis Biodiversity Center, Leiden, The Netherlands; [19]Department of Life Sciences, Imperial College London, Ascot, UK; [20]Ronin Institute for Independent Scholarship, Montclair, New Jersey, USA; [21]Cheadle Center for Biodiversity and Ecological Restoration, UC Santa Barbara, Santa Barbara, California, USA; [22]Department of Zoology, University of Oxford, Oxford, UK; [23]ISOE - Institute for Social-Ecological Research, Frankfurt am Main, Germany; [24]Centre for Ecosystem Science, School of Biological Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia and [25]Institute of Plant Sciences, University of Bern, Bern, Switzerland

Correspondence
Alexander Keller
Email: keller@bio.lmu.de

Caterina Penone
Email: caterina.penone@gmail.com

Funding information
Deutsche Forschungsgemeinschaft; Deutsches Zentrum für integrative Biodiversitätsforschung Halle-Jena-Leipzig, Grant/Award Number: W7.15; LifeWatch Belgium

Handling Editor: Karen Bacon

Abstract

1. Traits have become a crucial part of ecological and evolutionary sciences, helping researchers understand the function of an organism's morphology, physiology, growth and life history, with effects on fitness, behaviour, interactions with the environment and ecosystem processes. However, measuring, compiling and analysing trait data comes with data-scientific challenges.
2. We offer 10 (mostly) simple rules, with some detailed extensions, as a guide in making critical decisions that consider the entire life cycle of trait data.
3. This article is particularly motivated by its last rule, that is, to propagate good practice. It has the intention of bringing awareness of how data on the traits of organisms can be collected and managed for reuse by the research community.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4. Trait observations are relevant to a broad interdisciplinary community of field biologists, synthesis ecologists, evolutionary biologists, computer scientists and database managers. We hope these basic guidelines can be useful as a starter for active communication in disseminating such integrative knowledge and in how to make trait data future-proof. We invite the scientific community to participate in this effort at http://opentraits.org/best-practices.html.

KEYWORDS
data life cycle, data science, FAIR principles, good practices, metadata, open science, phenotype, trait data

1 | INTRODUCTION

As early as 300 BC, Greek philosophers such as Theophrastus forged the first formal systems defining and classifying organisms by their combination of morphological, physiological, behavioural and phenological characteristics, that is, their traits (Weiher et al., 1999). Knowing an organism's traits often allows deeper understanding of its life history, behaviour, fitness, biotic interactions and potential responses to and effects on ecosystem processes (Violle et al., 2007). Traits can also allow a better understanding of the processes behind ecological and evolutionary patterns (Sutherland et al., 2013), and offer a bridge between different dimensions: from organismal biology, for example, population abundance (Webb et al., 2010), species distribution (Sporbert et al., 2021) and phylogeny (Junker et al., 2015; Tucker et al., 2018) to ecosystem functioning (Lavorel & Garnier, 2002; Wright et al., 2016).

Traits are commonly defined as a measure of an entity (Garnier et al., 2017), where the entity can be the whole individual, or a specific organ or tissue (e.g. a fish, its tail fin or wood) and the quantity is an observable characteristic of that entity (e.g. the length or colour of a fish, the length of its tail fin or the density of wood tissue). Together, all traits of an individual organism represent its phenotype, which results from the historical evolution of the genotype and potential current interactions with the environment. Therefore, a trait record should inform not only about the entity that was observed (e.g. taxonomic classification or age) and the quantity/characteristic that was measured, but also about the environment in which the individual has developed that trait (de Bello, Carmona, et al., 2021; Kattge et al., 2011), for example, where a fish was caught, where a tree lived or the soil depth where an invertebrate was observed.

There are many ways to describe and measure the traits of organisms (Kearney et al., 2021; Walker et al., 2022). For example, a plant leaf can be described by several hundred measurable characteristics, or ‘traits’. These include surface area, sodium concentration, phenology and maximum photosynthetic rate (see e.g. Kattge et al., 2020). On the one hand, different traits of an individual are often correlated, for example, for a tree to grow tall it usually needs a thick stem. Recognising these correlations in how the data are collected (e.g. on the same tree) and stored is essential. In this case, for a trait record to be meaningful, it needs to be connected to a combination of multiple trait measurements. In contrast, a trait record can also be rather simple, if the given trait is well defined, if it depends ‘only’ on the genotype, or if it is not affected by current interaction with the environment.

In essence, trait data are a special kind of data: they are diverse (e.g. categorical or numeric, with a multitude of units), relatively simple (e.g. length) or potentially complex (e.g. behavioural traits), largely independent of one another (e.g. fish colour) or correlated with other traits (e.g. brain and body mass), and range between cheap and costly to measure (e.g. simple colour vs. metabolome data). However, they are very informative as they represent the evolutionary adaptation or developmental acclimation of the individual organisms to their environment and allow for quantitative and predictive ecology and biodiversity research. Therefore—if collected, stored and published in a meaningful way—organismal trait data have an extraordinary value for reuse, as indicated by, for example, the >20,000 data requests to the TRY Plant Trait Database since 2015 (Kattge et al., 2020).

To enable the reuse of trait data beyond their original research campaign, to make them meaningful in other contexts and to avoid data degradation, observation records must be clearly defined, where possible the environmental context given, as well as provenance and sampling and measuring protocols for collection documented (Michener, 2006). Recent efforts to expand trait knowledge across the Tree of Life (Gallagher et al., 2020) call for datasets that are open and FAIR (Findable, Accessible, Interoperable, Reusable; Wilkinson et al., 2016), fundamental principles at the heart of the emergent Open Science movement (Nosek et al., 2015). Global and local datasets of organismal traits have rapidly grown since the 1990s (e.g. Herberstein et al., 2022; Kattge et al., 2020; Madin et al., 2016, 2020). However, these datasets bear various new challenges linked to harmonisation, biases, expertise and communication (Salguero-Gómez et al., 2021). These challenges result in a significant trade-off between investing in collecting new trait data or reusing open trait data (Westoby et al., 2021). Indeed, many studies in trait-based research reuse available trait data or collect additional trait data and/or assemble new data (e.g. examples in Kattge et al., 2020). Thus, these studies also often involve linking different types of data, which requires interoperability between datasets (Feng et al., 2022; Gallagher et al., 2020).

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

These key aspects are just a few of many dimensions illustrating how and why researchers have to make biological decisions, and a wide range of data-science choices when collecting and working with trait data. Multiple complexities of trait data structure and manipulation are not obvious at first glance (Michener, 2006). For instance, there is sometimes confusion, and lack of awareness of trait standards, measurement units and trait data are particularly prone to errors in recording, language translation and understanding (Dawson et al., 2021; Kunz et al., 2022). By offering a larger perspective, a 'trait data life cycle' (i.e. a data life cycle specific for trait data, Rüegg et al., 2014) can help clarify these confusions and inform about good practices when working with trait data (Figure 1). In this article, we highlight some common pitfalls in the usage of trait data and offer 10 rules for making critical decisions that consider the entire life cycle of trait data. We start each rule with a general and simple statement and develop the complexity of each rule within more detailed subsections.

2 | RULE 1: SELECT THE RIGHT TRAIT

Let your study question or hypothesis determine both the trait(s) to be used and how those traits are collected and analysed. Clear, upfront definitions of traits will avoid errors through, for example, confusion of scales and definitions, data gaps or inclusion of inadequate traits (Dawson et al., 2021; González-Suárez et al., 2012; Hulme et al., 2013; Messier et al., 2017).

2.1 | Follow your hypothesis

Increasingly, trait data describing organisms of interest are publicly available for reuse. However, primary trait collection is necessary for a large number of research questions, for instance those involving rare species, understudied regions or small spatial scales. Vast public availability extends the potential scope of what is possible with limited resources (e.g. Falster et al., 2021; Kattge et al., 2020). However, when reusing trait data, we relinquish control of what variables are collected, which species are sampled, and the methods used for collection (Koricheva et al., 2013). Undirected fishing expeditions for traits can yield large datasets. Still, these may not be appropriate to answer a given research question, for various reasons (e.g. coverage, geographical origin, distribution, meaningfulness, and resolution, Violle et al., 2015). Furthermore, the wealth of available trait data may distract from initial hypotheses, risking random exploration of the available traits and fishing for significant relationships without a clear focus. Thus, trait selection and collection should in most cases be primarily tethered to a concrete hypothesis, not defined by the availability of existing data. This rule does, however, not completely exclude extensive data exploration and data-driven discovery within

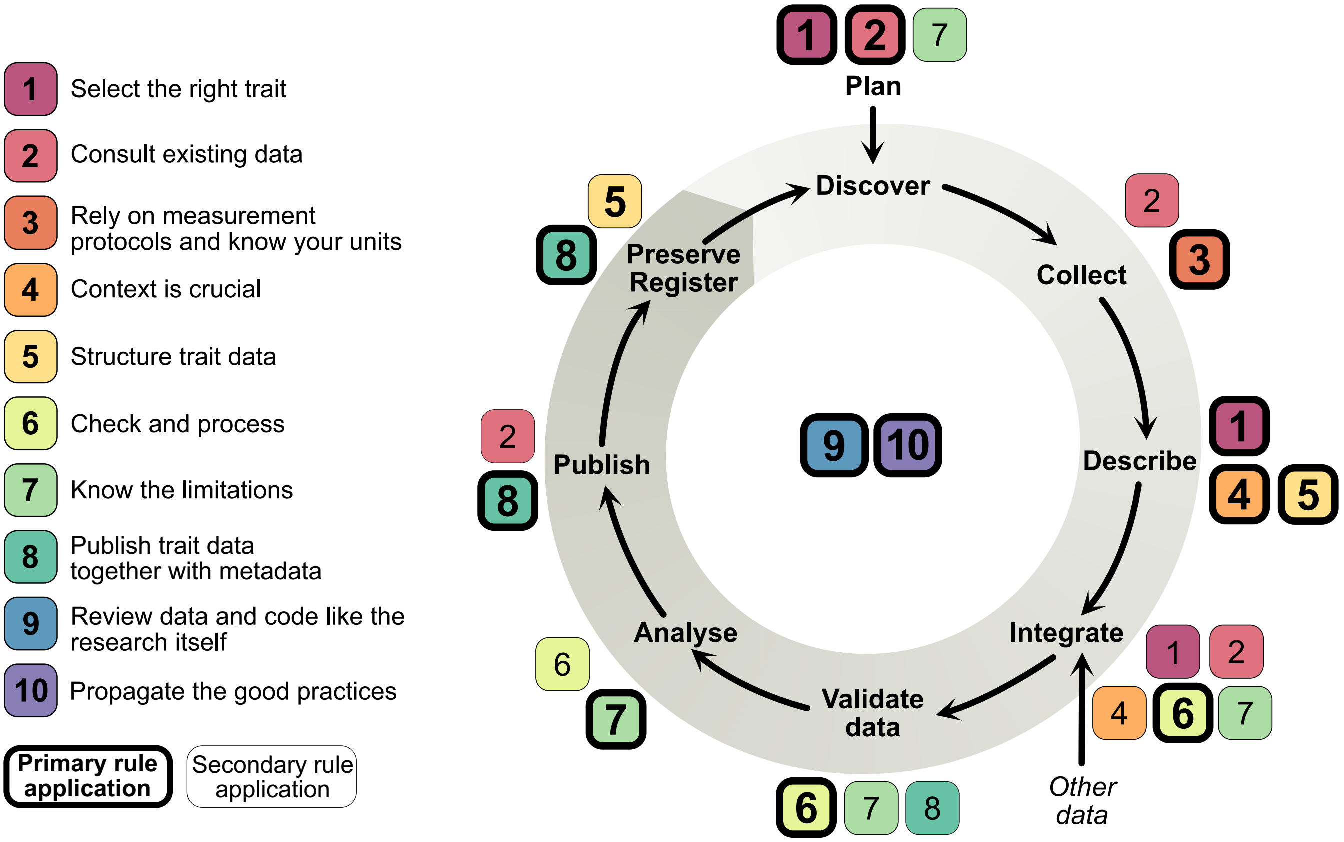

FIGURE 1 Ten (mostly) simply rules and where they apply in the overall trait data life cycle. Each rule is primarily applied to a specific element of the cycle (in bold) but can also be necessary to other elements (secondary application). Rules 9 and 10 apply to the whole cycle.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a given range as relevant to the research question and subsequent streamlining (Violle et al., 2015).

2.2 | Consider the scale

Research questions define the appropriate hierarchical level for sampling: a continental-scale study of thousands of species may treat the intraspecific variation as statistical noise. In contrast, this variation may be the study focus on locally scaled projects. There is no 'correct' scale, either in terms of spatial grain (e.g. km^2, m^2), temporal duration (e.g. seconds, years) or taxonomic coverage (e.g. clade, species, population or individual), but not every scale will be appropriate for every question. So, when defining the traits of interest, it is important to determine the scale at which these need to be collected or aggregated to match the research question (Messier et al., 2017).

2.3 | Be aware of existing trait definitions and homologies

Much effort has already gone into creating definitions and protocols for trait collection (Pérez-Harguindeguy et al., 2013). Yet, trait naming and corresponding descriptions may differ between studies and trait databases (Ankenbrand et al., 2018; Dawson et al., 2021; Kunz et al., 2022). For example, the activity cycle of animals is sometimes reported as a discrete value (e.g. Jones et al., 2009), or sometimes split into multiple binary traits such as 'nocturnal', 'crepuscular', 'diurnal' (e.g. Wilman et al., 2014). Similarly, values may differ between resources (e.g. 'therophyte' and 'annual' are synonyms). Furthermore, when comparing traits and trait states across organisms, it is important to be aware of the 'homology' of the character. Homologous traits share similarity of structure, physiology or development (often by common evolutionary ancestry), whereas non-homologous (or analogous) characters may perform a similar function, but differ in structure, physiology or development.

2.4 | Be pragmatic and transparent

In a perfect trait research world, we could measure or retrieve the exact traits for the precise scale and organisms needed to answer our specific question. This vision is rarely applicable in practice. Instead, we often need to work with proxies for traits that are difficult to measure (e.g. hairiness of pollinators as a proxy for pollination effectiveness, Stavert et al., 2016), for inference of fitness (e.g. reproductive output as a performance trait, McGraw & Caswell, 1996, Violle et al., 2007) or for traits that are incomplete in a database (e.g. diet or behavioural traits are less complete than morphological traits, Oliveira et al., 2017). There is a common understanding of these technical or financial limitations in the scientific community; ultimately, we must be pragmatic to advance research questions. However, it is crucial to explain and justify the choice of traits, especially when these are used as proxies or 'best available data' to allow fellow researchers to understand and evaluate whether such choices were valid for the specific research question at hand.

3 | RULE 2: CONSULT EXISTING DATA

Build on existing trait resources to reduce the likelihood of redundancy and ensure compatibility with current data. The decision when to collect new trait data is generally based on the research question, the scope of the analysis (e.g. local, global), and the availability of the existing data. Financial and geographic constraints may also influence the decision to use current trait data instead of embarking on a measurement campaign. However, the existing trait data must be 'fit for purpose' to avoid compromising the capacity to answer the research question and in many cases, new trait measurements will still be needed.

3.1 | Check public data sources

Most data probably exist decentralised as individual trait datasets in the form of raw data attachments to publications, data papers or data uploads to unspecific public databases (e.g. Zenodo https://zenodo.org, Dryad https://datadryad.org). However, these datasets can be challenging to find if not registered at central hubs (e.g. https://opentraits.org). To counter this challenge, dedicated centralised trait databases have been and continue to be developed (e.g. TRY, Kattge et al., 2020), Encyclopedia of Life (EOL) TraitBank (Parr et al., 2015), Marine Traits Portal of the World Register of Marine Species (WoRMS, Marine Species Traits Editorial Board, 2022), AusTraits (Falster et al., 2021). Common to these efforts is the fact that they contain already harmonised, error-checked and standardised values. These resources usually provide user-friendly interfaces for searches and dynamic, up-to-date aggregations of data. Particularly for studies of larger scale (e.g. many taxa, many bioregions), it often makes sense to consult these existing big databases and data registries.

3.2 | Identify and cite data origins

Trait data are not always raw or first-hand: they can be created and perhaps aggregated from original observations and measurements (e.g. Kattge et al., 2020) but also mobilised from literature or undigitised legacy trait data (e.g. Parr et al., 2015), synthesised as imputed trait data (e.g. Penone et al., 2014), reused from data publications (e.g. Kattge et al., 2020) or mined from texts with automated algorithms or other contexts (Thessen et al., 2018). Thus, when reusing trait data, it is essential to check and report information about the source to downstream analyses and

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

subsequent publications (i.e. data provenance). Importantly, providing this information also gives credit to the original trait data collectors.

3.3 | Fill the gaps

Existing databases are taxonomically and biogeographically biased, 'gappy', and traits assigned to the same species are rarely collected in the exact locations or conditions (Etard et al., 2020). Despite the presence of large trait databases, new trait collections continue to remain valuable. When collecting new data, we encourage researchers first to check available trait databases, identify such gaps and contribute to the broader trait community by filling these gaps even if this collection goes beyond the current project. Additional traits may be easily collected with little extra effort, yet provide the possibility to close gaps in trait coverage. Filling gaps may be especially valuable in biodiverse but hard-to-access regions (Etard et al., 2020), for rare but functionally important species which may be less likely to have traits documented (Leitão et al., 2016), or for threatened species which will benefit from functional approaches to their conservation (Gallagher et al., 2021).

4 | RULE 3: RELY ON MEASUREMENT PROTOCOLS AND KNOW YOUR UNITS

To ensure comparability, future data reuse and synthesis, relate primary measurements of your traits to the wider body of published trait data. Conform your measurement procedures to existing trait measurement protocols, or—if no such standard protocols exist—document with precision and build upon unambiguous concepts. Also, confusion and errors in terms of recording and reporting of units can be propagated through large trait compilations. Thus, define your units clearly; they are essential for harmonising different trait datasets, approximations and uncertainties.

4.1 | Beware of ambiguities

In most cases, researchers of a domain (e.g. plants) have adopted sufficiently specific trait definitions to allow comparison of widely used measurements and enable synthesis within the field. However, some difficulties in measurement remain. To illustrate, specific leaf area (SLA) is the ratio of the surface area to leaf biomass of an individual leaf. However, the application of the concept of SLA may differ between research contexts, because the value reported may relate to measurements of individual leaves or an average of all leaves on the shoot, for one or both sides of the leaf, including or excluding the petiole, and focus on the leaf or leaflet (example taken from Garnier et al., 2017). While fully justified in the specific research setting, identifying and dealing with semantic disambiguation is a major challenge in trait-based synthesis.

4.2 | Adhere to existing standards

Methodological handbooks for trait measurements have been proposed, for example, for plant (Cornelissen et al., 2003; Pérez-Harguindeguy et al., 2013), macrofungi (Dawson et al., 2019) or terrestrial invertebrate functional traits (Moretti et al., 2017). These handbooks provide precise, domain-specific definitions and recommended methods for trait measurement, measurement precision and replication. They also provide considerations and warnings of misconception and error, and point to the key literature debating the methodology. Taking formalisation of trait concepts even one step further are thesauri of trait concepts (Garnier et al., 2016, 2017), for example, the TOP thesaurus of plant characteristics (https://top-thesaurus.org/). The bottom line is: research that provides original trait measurements should consider existing measurement protocols, make an explicit choice and describe any deviations from or additions to protocols. When such handbooks do not exist, it is good practice to report specific measurement protocols in the metadata (see Rule 4). For instance, how the length of a fish has been measured and if potential extensions of the tail fin were taken into account.

4.3 | Understand your units

Trait data are necessarily 'rich in dimensions'. That is, trait data may require multiple SI (International System of Units) base units and may also be measured and reported in various alternative configurations of units. For example, photosynthetic rate involves three SI base units, mass per area per time, and is often reported in units of μmol $CO_2 m^{-2} s^{-1}$ an amount per unit area per unit time. Measures of size, area and time are often reported in different units, although all can be related to more fundamental base units. All metric trait data can be reduced to the seven base units as defined by the SI standard (m, kg, s, K, A, cd, mole). Significant data management effort is needed to record units accurately, preserve them through metadata and convert them correctly to avoid propagating errors (Calder, 1982).

5 | RULE 4: CONTEXT IS CRUCIAL

Always pair your data points with metadata. Sampling protocols ideally also define metadata that can be considered as covariates of the measurement procedure or inform the user about the provenance of the trait data. Together with the trait measurements, metadata defines an observation and its context (Madin et al., 2008). While such metadata may already be necessary for the proximate research question, it further helps future users to understand better and reproduce the methods and correctly interpret the trait values. The reuse value of existing datasets increases with the quantity and quality of metadata, so, datasets with sufficient context information are more likely to be reused in future synthesis analyses or included in more extensive databases.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.1 | Define at least the minimum context

Some metadata are considered essential and universal between all domains, such as an unique ID for observations to cross-reference to other measures, geolocation, time and date, life stage (e.g. juvenile), health status, scale (e.g. leaf), habitat type (e.g. semi-natural grassland or botanical garden) and measurement details (e.g. following standards, devices used, etc., Schneider et al., 2019). Further metadata must include the source and authorship of the trait measurement. To permit effective reuse, authorship attributes should consist of the original data collectors and the databases where these data were gathered, as they may have undergone processing therein (Rule 2).

5.2 | Cover the domain-specific standard, if possible

Deciding which further metadata to collect often involves a trade-off between which data are commonly collected in a specific domain (e.g. plants) and the time and expense involved in collecting or processing such data. Metadata preferably includes detailed documentation and code of how traits were measured (e.g. manufacturer and version of devices used) and processed (e.g. standardizations or species means). We recommend checking existing well-used datasets and databases of the specific domains before collecting new trait data to determine which common metadata should be covered.

5.3 | Link to other data by metadata

A good practice is to link the data with publications directly (e.g. by DOI) for the scientific context and further information in the materials and methods sections, as well as identification of trait data providers (e.g. by ORCID) to provide opportunities for feedback and requests for additional information. Traits are often measured also to collect other data, such as ecosystem function (e.g. Bongers et al., 2021) or species composition or interactions (e.g. Breitschwerdt et al., 2018). In these cases, functions measured, and species composition recorded, would be part of the metadata or links to those data in other repositories.

6 | RULE 5: STRUCTURE TRAIT DATA

Do not underestimate the importance of the structure of your dataset. It might sound trivial at first glance to think about how to structure the data, but poorly structured data may become a nightmare to work with in downstream analyses, or to reformat for publication, deposit in a public database, or synthesise in meta-analyses. It thus makes sense to consider structural aspects even in the early stages of a project using traits.

6.1 | Minimum trait data standards

The minimal, essential information for a trait record includes taxon name, trait name, observation ID, trait value, unit (if applicable) and source. Several standards are available to help structure this minimal information set (Fegraus et al., 2005; Kattge et al., 2011; Madin et al., 2007; Parr et al., 2015; Schneider et al., 2019; Wieczorek et al., 2012). A good start for data structuring is to adopt one of these well-established schemes.

6.2 | Preserve metadata and further observations recorded together

A complex aspect of structuring trait data is how to keep metadata and data links. This linkage is critical for various questions that address intraspecific variation, derived traits or multivariate modelling on individuals. A good practice is to have unique identifiers for every entity that requires relations. Using such IDs allows linking to other traits and different data types, like community records of the plot in which the individual was found, DNA barcodes or information on the experiment that involved the individual. One approach is the Extensible Observation Ontology (OBOE, Madin et al., 2007), which various large databases have also adopted (e.g. TRY, Kattge et al., 2020).

6.3 | Apply version control

The process from gathering to analysing trait data is long, or trait data may change as measurement technologies improve (e.g. genomic traits, Madin et al., 2020), which may result in many different versions of a trait dataset. Thus, it is important to keep track of these different versions. It is recommended to keep the raw trait dataset and the processed trait dataset used for analysis as separate copies. Version names should be structured in a logical way, for example, combining project acronyms, researchers' initials, short names of the trait dataset, version numbers, file status and/or dates. Use of versioning or change control systems like Git (Spinellis, 2012) is highly recommended to keep track of changes to data. Also, conventions that include a current (e.g. my_data_current.tsv) and versioned copy (e.g. my_data_1.0.tsv) help automated systems by providing a stable interface for indexing.

7 | RULE 6: CHECK AND PROCESS

Rigorously check your data quality, integrity and compatibility during each step of data processing. Trait-based analyses, mainly when data are consolidated from different sources, can harbour various inherent incompatibilities that may cause biases and severe scientific misinterpretations. For trait compilations, data usually need to be harmonised, subset, transformed, derived and/or aggregated into

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

comparable formats to fit the research question. Wherever possible, steps must be scripted and directly reproducible, and where not, manual steps should be well documented.

7.1 | Harmonise trait data

If trait data originate from multiple sources, each source may identify the same entities or concepts differently (Kunz et al., 2022). Harmonisation is crucial to reconcile equivalent entities and explicitly connect related entities by 'similar' or subclass relationships. Ideally, these entities or concepts should be identified by standard identifiers (see Rule 5). Manual harmonisation may be necessary to detect and reconcile spelling variations before text strings are mapped to identifiers. But for common classes of data, there are a variety of services available that allow automated and reproducible harmonisation, for example, for taxonomic names (Boyle et al., 2013; Chamberlain & Szöcs, 2013; Global Names Architecture, reviewed by Grenié et al., 2022), units (Gama, 2016) or geographic names (Boyle et al., 2022). Other covariates and categorical trait values may be semantically reconciled where appropriate ontologies exist (Kunz et al., 2022; Violle et al., 2015).

7.2 | Filter where needed and double-check data contexts

Not all trait data are equally suitable for all purposes. Erroneous or duplicate data points need to be identified and removed before analyses, for example, by validation of data origins and metadata to make sure identical values were indeed from individual measurements. As with other kinds of data, outlier detection and data visualisation provide valuable methods for the detection of such data errors (de Bello et al., 2013). For trait data that are primarily compiled from different sources, other reasons may also render data points inappropriate. For example, if metadata suggests that the observation is from a cultivated occurrence such as a botanical garden, greenhouse, zoo or farm, values might not be representative of wild specimens (Gering et al., 2019). Observations stemming from introduced or experimental populations may violate assumptions as well. Observations can be collected from different subsets of the population (e.g. adult vs. juvenile, healthy vs. diseased), at different times of year (e.g. breeding season vs. overwintering), in different contexts (e.g. experimental temperature treatments), and using other protocols. It is essential to exclude unsuitable observations, usually by making use of the associated metadata.

7.3 | Derive traits from raw data

Research questions may concern composite or derived traits, such as the 'hand-wing index' (a wing's aspect ratio in birds). It is advisable to calculate derived traits directly from the raw data where possible to avoid bias and allow for new calculations. This procedure may not always be possible because of data gaps; in this case the calculation can be done at a higher level (e.g. at the taxonomic level of interest).

7.4 | Aggregate trait data

Trait data may come at different levels of resolution. A dataset may include multiple measurements per individual, per population, species or even higher taxonomic levels. Such structures may imply aggregating (e.g. to calculate average trait values) within individuals, then populations, then species derived from a particular data source, and then across data sources if the species is represented in several of these (Schneider et al., 2019). The way trait values were aggregated has to be precisely described, in particular when data transformation is involved. For example, when it is desired to express leaf area on a log scale, it makes a difference to take the log before or after aggregating the data. Importantly, suppose multiple successive steps of aggregation are necessary. In that case, there is the need to properly measure the uncertainty of the final trait values and assess the effect of aggregation on the results and conclusions, for example, by sensitivity analyses with different aggregated datasets (Kunz et al., 2022).

7.5 | Transform and standardise where applicable

Likewise for other types of data, transformations such as the natural logarithm or square root may be essential to conform to the requirements of analytical models. Beyond these, data challenges include how to combine binary, categorical and continuous traits into the same analysis (de Bello, Carmona, et al., 2021). It is thus very useful to explore transformation and standardisation options applied in current trait scientific literature. For example, to compare the effects of several explanatory traits on a specific response in a linear model approach, values can be standardised for each trait to range between 0 and 1, or by scaling their mean to 0 and their standard deviation to 1 or 0.5 (in case of making continuous traits comparable with categorical traits, Gelman, 2008).

7.6 | Work with relative errors

Units are essential when we deal with approximations, uncertainties and errors (Langtangen & Pedersen, 2016). An example is, a trait measurement where the length scale is typically measured in mm and approximates 12.5 m to the exact value of 12.52 m with an error of 0.02 m. Switching units to mm leads to an error of 200 mm. A study working in mm would report 2 x10^2 as the error, while a study working in m would report 0.02 as the error. As a result, knowing the original measurement units is essential, and the downstream use of the unitless relative error is recommended (Langtangen & Pedersen, 2016).

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

8 | RULE 7: KNOW THE LIMITATIONS

Follow the latest developments for best practices in trait data analyses. As the downstream part of data analysis is directly linked to the research question, generalisation of analytical methods is rarely possible. Given the diversity of research questions, the analytical steps can thus broadly diverge. However, the following notions can help identifying some common mistakes made with trait data due to their nature. Beyond this, we recommend referring to closely domain-specific and topic-related literature that can provide appropriate solutions.

8.1 | Mind the level

Traits encompass different levels: organ, individual, population, species and community (Violle et al., 2007), and this structure determines the tools used for data analyses. For instance, trait–environment relationships investigated at the species or community level require different analysis types (e.g. comparative models vs. simple linear models, see below). It is important to choose the appropriate level early in the research program to fit the target scientific question and to be able to analyse the data correctly.

8.2 | Be careful with categorical traits

By having fewer possible values, categorical traits might influence the outcome of statistical analyses. In particular, categorical traits might have disproportionate effects when aggregated with continuous traits in a standard metric (e.g. functional diversity). For instance, when computing Gower distances of traits between species pairs, a categorical trait treated as a binary trait (0/1) for each possible categorical value will only result in distances of 0 or 1. In contrast, only the species pair with the highest/lowest trait values for a continuous trait will have a value of 1. When averaging Gower distances of these two traits, the categorical one will have more influence than the continuous one (example from de Bello, Carmona, et al., 2021). The joint use of continuous and categorical traits thus needs particular scaling (Gelman, 2008). Generally, it is important to be aware of this issue and to account for it with existing proposed methods (e.g. de Bello, Botta-Dukát, et al., 2021).

8.3 | Do not confuse richness and abundance signals in trait metrics

Metrics aggregating traits at the community level (e.g. functional diversity or community-weighted means—CWMs), are influenced by the richness, the abundance of species and the overall species composition of the community. Choosing metrics unrelated to abundance (e.g. unweighted means) or null models (Hawkins et al., 2017) is necessary to separate species abundance, composition or richness signals from trait information.

8.4 | Handle correlations with care

Traits are often correlated, causing issues with statistical analyses (e.g. collinearity in linear models when traits are explanatory variables). Often, these correlations are due to biological constraints (e.g. allometries), or 'strategies' (Díaz et al., 2016). In some cases, it is possible to use multivariate analyses (e.g. principal component analysis) to reduce trait space dimensionality by use of axes in further analyses. Keeping the original traits is advised when possible, as the biological meaning can be lost when using multivariate principal component axes. On another level, since the start of trait analyses at the community level, numerous metrics have been proposed to characterise functional diversity analogous to species diversity. When choosing a metric for a specific analysis, it is essential to be aware that several of these functional diversity metrics are highly correlated and with species richness (de Bello, Carmona, et al., 2021).

8.5 | Consider correction for phylogenetic relatedness

When analysing data from multiple species in trait–trait correlations, or when using traits as responses, and depending on whether the focus of the question is ecological or evolutionary, it may become necessary to account for the fact that species are not independent units (Pillar et al., 2021). The whole field of comparative analyses tackles this issue. It proposes tools to account for phylogenetic relatedness in trait analyses (e.g. see Garamszegi, 2014), although care should be taken to justify the use of such analytical corrections relative to the aims of the research question (Freckleton, 2000; Westoby et al., 1995).

8.6 | Account for variability and uncertainty

Very often, intraspecific data are aggregated at the species level to obtain one trait value per species. All information on variability and measurement uncertainty is then lost. When information on variability is available and reasonable in the scope of the study, it is possible to include it, for example, by weighting species-level measures in functional diversity metrics (de Bello, Carmona, et al., 2021) or by explicitly including it when inferring trait evolution across lineages (Kostikova et al., 2016; Purschke et al., 2017). This can be an issue, especially if variability is phylogenetically structured (Garamszegi, 2014; Paterno et al., 2018).

9 | RULE 8: PUBLISH TRAIT DATA TOGETHER WITH METADATA

Openly publish trait data to facilitate answering yet unknown questions beyond their original study, lay the groundwork for understanding ecological processes beyond clearcut niches (Elton, 1927;

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schneider et al., 2019), and democratise access to valuable trait datasets (Soranno et al., 2015). Each data point of trait measurements has a considerable value for the scientific community and future generations working on trait-related research questions.

9.1 | Consider the stakeholders

As our scholarly processes evolve to better find, access, integrate and reuse scientific data, we face the communal task of treating trait datasets as first-class research citizens. However, doing so is not easy as it involves different stakeholders: publishers have to make their publications open and FAIR (Wilkinson et al., 2016), scientists have to improve their skills to publish, reuse and correctly cite datasets, and funding agencies have to find ways to reward exemplary projects. A welcome development is that many publishers now consider trait data papers (e.g. Falster et al., 2021; Guerrero-Ramírez et al., 2021; Tobias et al., 2022; Vandvik et al., 2020), which allow for a detailed methodological and context description, open access, and at the same time, accreditation of trait data collectors by citations.

9.2 | Accept the additional responsibility

Erroneous data might bias a current project, but also the future works of others. Currently, no common established practices exist on how peer review is also extended to trait data. A way to ensure that a dataset conforms to community standards is to submit it to an established curated database (e.g. TRY (Kattge et al., 2020) for plant traits; Coral Traits (Madin et al., 2016) for anthozoans). Furthermore, consider publicly depositing raw and processed data and clearly differentiating between the two types. This allows tracing errors generated during processing and grants future users access to the original values.

9.3 | Aim for redundancy

Public trait data suffer from the same generic issues as other data, for example, hardware failures, linkrot (URLs not entirely reliable) or content-drift (content changes, but URLs do not, Koehler, 1999). To mitigate such issues and reliably preserve data in the long term, data can be submitted to multiple repositories, for example, beside trait databases, also in general storage platforms such as FigShare (https://figshare.com) or Zenodo (https://zenodo.org). This procedure, however, requires systematic methods to track changes and separately citable versions, for example, by unique DOIs.

9.4 | Make data accessible for machines and humans

In order to facilitate trait data reuse in general, machine-readable and nonproprietary data formats should be preferred (i.e. plain csv over excel or pdf). In this context, the licence under which data are released should also be correctly chosen (e.g. CC0, Creative Commons, 2009). When reporting already published data, future studies might also run the risk of using the same trait from independent sources, thus resulting in pseudo-replication of measurements. This makes it important to render data traceable throughout the life cycle; especially because trait data collections often carry large numbers of references and republished original data. Data and reference tracing thus particularly calls for systematic, reproducible and automated methods (Elliott et al., 2020) that rely on machine-readable data.

9.5 | Register trait data

Independent of the choice of actual data deposition, it is important that datasets are registered in a trait data registry (e.g. https://opentraits.org) to allow fellow scientists to find the data quickly.

10 | RULE 9: REVIEW DATA AND CODE LIKE THE RESEARCH ITSELF

Best practices in peer review have already been discussed in detail (Roberts, 2004; Spigt & Arts, 2010), but can perhaps be summarised with this statement: 'Be polite, fair, specific, and constructive'. A reviewer should provide information for the editorial team to decide; this process also applies to the data. Specifically for trait-based papers, it includes considering the entire life cycle of the trait data:

1. First, are the traits themselves appropriate for the question being asked? It should be considered how these traits have been used in the past and how they fit into biological theory. Are they being contextualised appropriately, and are they fit for the purpose for which they are being used?
2. How were the data collected? Does the protocol conform to current standards, bearing in mind that the purpose of many papers is to improve standards and so they may not? Is the collection of new data well justified? Are units and metadata properly provided?
3. How were the data processed? Consider not just quality assurance and quality control but also how the traits were generally processed into a format that can be analysed. Ensure that relevant code (ranging from simple cleaning scripts to full-featured analysis pipelines and models) is openly available, functional and conforms to community standards. Guidelines for source code review like those by the rOpenSci community (https://ropensci.org/software-review) or The Journal of Open Source Software (https://joss.readthedocs.io/en/latest/reviewer_guidelines.html) can be helpful.
4. Ensure that a distinction is made between the 'raw' data collected and the 'clean' final product used for analysis, and that both forms of data are released.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5. Ensure that all data sources are appropriately cited and the provenance of data is explained.

While these standards apply to all manuscripts, a distinction should be made between data papers and research papers. Standards for data release are the same for both types of publication. However, the expectations for demonstrating data cleaning and conceptual novelty likely differ between the two (but are, ultimately, determined by the journal's guidelines).

11 | RULE 10: PROPAGATE THE GOOD PRACTICES

Inform the community and the next generation of ecologists about the issues discussed here and in other resources (e.g. de Bello, Carmona, et al., 2021). As outlined above, there are several issues to be aware of when measuring, collecting, handling, analysing and publishing trait data, that is, the life cycle of trait data. Some may be straightforward; others require more technical knowledge or extensive reading of existing resources. In many cases, good procedures are not applied simply because ecological or evolutionary scientists are unaware that they exist, for example, in trait quality control or using a standard structure. Educating can be done from the small scale of an informal conversation with a colleague, to teaching a large undergraduate class, up to participating in the collective creation of open access materials in several languages, accessible from any part of the world. Integrating trait-data-specific sections into ecology textbooks and modules in ecology courses could become a standard practice, which will undoubtedly be made more accessible by disseminating open access material by the trait scientific community. You are welcome to use these 10 rules as a starter when teaching your students, colleagues and friends. For diving deeper into every single topic, we encourage you also to check more comprehensive resources such as the Handbook of Trait-Based Ecology (de Bello, Carmona, et al., 2021) or the activities of the Open Traits Network (https://opentraits.org; Gallagher et al., 2020).

11.1 | Train students

Courses specific to trait-based research are often lacking at undergraduate and graduate levels. Where courses or modules are taught, the focus may be limited to a subset of the trait data life cycle (e.g. Collection and Analysis; Figure 1), leaving students lacking critical skills (Feng et al., 2020). Open Educational Resources, including those built using incubators (Ryder et al., 2020), are one promising method for implementing such courses and modules more easily. In particular, authentic teaching experiences provide several benefits over traditional lectures or 'cook-book' experiments (Brownell et al., 2012). They seem well suited to trait-based ecology given that many traits can be collected quickly and inexpensively and that many tools are available (see, e.g. de Bello, Carmona, et al., 2021). One example of such authentic teaching experiences, the TraitTrain plant functional trait courses (https://plantfunctionaltraitscourses.w.uib.no/), has provided training across the entire trait data life cycle to hundreds of participants and has created scientific (Henn et al., 2018), data (Vandvik et al., 2020), methodological (Maitner et al., 2021) and pedagogical (Geange et al., 2021) publications.

11.2 | Train colleagues

Making colleagues aware of important developments in trait-based research via either formal (e.g. publishing protocols, giving talks) or informal means (e.g. conversations, social media, email) is a critical way of helping to advance the field. Furthermore, trait-based research is an integrative field. It provides many opportunities for collaboration and idea-sharing across branches of life science, so discussing traits with a wide variety of colleagues is useful.

11.3 | Train the world

There is an urgent need for more comprehensive trait data across the globe and the tree of life (Feng et al., 2022), thus, increasing global access to training. Open access publications, tools, data and educational resources help lower the barriers to participation (Evans & Reimer, 2009). Furthermore, due to the relative ease, low cost and tangible nature of many functional traits, they are well suited for inclusion in elementary education and citizen science (e.g. Isaac & Martin, 2019; Schiller et al., 2021).

12 | CONCLUSIONS

This 10 rules document is mainly motivated by its last rule and intends to of bring awareness of the different facets of a trait's life cycle to the community. Most scientists working with traits are usually experts on only one or few of numerous aspects of the trait data life cycle, ranging from biological theory and field research to computer sciences and publishing and reviewing ethics. Furthermore, current data infrastructures, definitions and methods, may pose limits mainly when working between domains. When working with trait data, we gain particularly as an interdisciplinary community of field biologists, synthesis ecologists, evolutionary biologists, computer scientists and database managers from a broad taxonomic range. This allows for the development of tools, methods and infrastructures that connect the entirety of trait science in an interoperable fashion. We hope that these basic guidelines can be useful as a starter for active communication in disseminating such integrative knowledge and how to make trait data future-proof. We encourage the scientific community to contribute to these rules when new tools and practices emerge in the 'living document' version of this article at http://opentraits.org/best-practices.html.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

AUTHOR CONTRIBUTIONS

Alexander Keller and Markus J. Ankenbrand conceived the idea. Caterina Penone, Jorrit H. Poelen, Markus Ankenbrand and Alexander Keller led the writing of the manuscript. Alexander Keller, Markus J. Ankenbrand, Brian J. Enquist, Helge Bruelheide, Stefanie Dekeyzer, Daniel S. Falster, Renske E. Onstein, Rachael V. Gallagher, Jennifer Hammock, Jens Kattge, Joshua S. Madin, Brian Maitner, William D. Pearse, Caterina Penone, Florian D. Schneider and Jorrit H. Poelen contributed in writing the first draft during a workshop session and thereafter. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

The manuscript was drafted during a workshop by participants of the sDiv working group sDevTrait hosted by the German Centre for Integrative Biodiversity Research (iDiv; DFG FZT 118). We appreciate funding (grant number W7.15 to AK and MA) and support before, during and after the workshop by the iDiv, particularly staff members Marten Winter, Carolin Kögler and Doreen Brückner. SD acknowledges additional support from LifeWatch Belgium. Funding to keep the World Register of Marine Species (WoRMS) is currently provided through the LifeWatch Belgium project. WoRMS constitutes a major contribution to the LifeWatch Species Information Backbone. CP acknowledges support from the DFG Priority Program 1374 'Biodiversity- Exploratories' (DFG-193921238). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

All authors declare to have no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.14033.

DATA AVAILABILITY STATEMENT

This paper does not use data.

ORCID

Alexander Keller https://orcid.org/0000-0001-5716-3634
Markus J. Ankenbrand https://orcid.org/0000-0002-6620-807X
Helge Bruelheide https://orcid.org/0000-0003-3135-0356
Stefanie Dekeyzer https://orcid.org/0000-0001-9525-2742
Brian J. Enquist https://orcid.org/0000-0002-6124-7096
Mohammad Bagher Erfanian https://orcid.org/0000-0002-8671-9037
Daniel S. Falster https://orcid.org/0000-0002-9814-092X
Rachael V. Gallagher https://orcid.org/0000-0002-4680-8115
Jennifer Hammock https://orcid.org/0000-0002-9943-2342
Jens Kattge https://orcid.org/0000-0002-1022-8469
Sara D. Leonhardt https://orcid.org/0000-0002-8154-9569
Joshua S. Madin https://orcid.org/0000-0002-5005-6227
Brian Maitner https://orcid.org/0000-0002-2118-9880
Margot Neyret https://orcid.org/0000-0001-9435-1634
Renske E. Onstein https://orcid.org/0000-0002-2295-3510
William D. Pearse https://orcid.org/0000-0002-6241-3164
Jorrit H. Poelen https://orcid.org/0000-0003-3138-4118
Roberto Salguero-Gomez https://orcid.org/0000-0002-6085-4433
Florian D. Schneider https://orcid.org/0000-0002-1494-5684
Anikó B. Tóth https://orcid.org/0000-0002-3063-1917
Caterina Penone https://orcid.org/0000-0002-8170-6659

REFERENCES

Ankenbrand, M. J., Hohlfeld, S. C. Y., Weber, L., Förster, F., & Keller, A. (2018). Functional exploration of natural networks and ecological communities. *Methods in Ecology and Evolution*, *9*, 2028–2033. https://doi.org/10.1111/2041-210X.13060

Bongers, F. J., Schmid, B., Bruelheide, H., Bongers, F., Li, S., von Oheimb, G., Li, Y., Cheng, A., Ma, K., & Liu, X. (2021). Functional diversity effects on productivity increase with age in a forest biodiversity experiment. *Nature Ecology & Evolution*, *5*(12), Article 12. https://doi.org/10.1038/s41559-021-01564-3

Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J. A., Mozzherin, D., Rees, T., Matasci, N., Narro, M. L., Piel, W. H., Mckay, S. J., Lowry, S., Freeland, C., Peet, R. K., & Enquist, B. J. (2013). The taxonomic name resolution service: An online tool for automated standardization of plant names. *BMC Bioinformatics*, *14*(1), 16. https://doi.org/10.1186/1471-2105-14-16

Boyle, B. L., Maitner, B. S., Barbosa, G. G. C., Sajja, R. K., Feng, X., Merow, C., Newman, E. A., Park, D. S., Roehrdanz, P. R., & Enquist, B. J. (2022). *Geographic name resolution service: A tool for the standardization and indexing of world political division names, with applications to species distribution modeling* (p. 2022.04.25.489424). *bioRxiv*. https://doi.org/10.1101/2022.04.25.489424

Breitschwerdt, E., Jandt, U., & Bruelheide, H. (2018). Using co-occurrence information and trait composition to understand individual plant performance in grassland communities. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-27017-9

Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). Undergraduate biology lab courses: Comparing the impact of traditionally based "cookbook" and authentic research-based courses on student lab experiences. *Journal of College Science Teaching*, *41*(4), 36–45.

Calder, W. A. (1982). A proposal for the standardization of units and symbols in ecology. *Bulletin of the Ecological Society of America*, *63*(1), 7–10.

Chamberlain, S. A., & Szöcs, E. (2013). Taxize: Taxonomic search and retrieval in R. *F1000Research*, *2*, 191. https://doi.org/10.12688/f1000research.2-191.v2

Cornelissen, J. H. C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D. E., Reich, P. B., Ter Steege, H., Morgan, H. D., Van Der Heijden, M. G. A., Pausas, J. G., & Poorter, H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, *51*(4), 335–380. https://doi.org/10.1071/BT02124

Creative Commons. (2009). CC0 1.0 universal. https://Creativecommons.Org/Publicdomain/Zero/1.0/

Dawson, S. K., Boddy, L., Halbwachs, H., Bässler, C., Andrew, C., Crowther, T. W., Heilmann-Clausen, J., Nordén, J., Ovaskainen, O., & Jönsson, M. (2019). Handbook for the measurement of macrofungal functional traits: A start with basidiomycete wood fungi. *Functional Ecology*, *33*(3), 372–387. https://doi.org/10.1111/1365-2435.13239

Dawson, S. K., Carmona, C. P., González-Suárez, M., Jönsson, M., Chichorro, F., Mallen-Cooper, M., Melero, Y., Moor, H., Simaika, J. P., & Duthie, A. B. (2021). The traits of "trait ecologists": An analysis

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of the use of trait and functional trait terminology. *Ecology and Evolution*, *11*(23), 16434–16445.

de Bello, F., Botta-Dukát, Z., Lepš, J., & Fibich, P. (2021). Towards a more balanced combination of multiple traits when computing functional differences between species. *Methods in Ecology and Evolution*, *12*, 443–448. https://doi.org/10.1111/2041-210X.13537

de Bello, F., Carmona, C. P., Dias, A. T. C., Götzenberger, L., Moretti, M., & Berg, M. P. (2021). *Handbook of trait-based ecology: From theory to R tools*. Cambridge University Press. https://doi.org/10.1017/9781108628426

de Bello, F., Carmona, C. P., Mason, N. W. H., Sebastià, M.-T., & Lepš, J. (2013). Which trait dissimilarity for functional diversity: Trait means or trait overlap? *Journal of Vegetation Science*, *24*(5), 807–819. https://doi.org/10.1111/jvs.12008

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P. B., Moles, A. T., Dickie, J., Gillison, A. N., Zanne, A. E., … Gorné, L. D. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171.

Elliott, M. J., Poelen, J. H., & Fortes, J. A. B. (2020). Toward reliable biodiversity dataset references. *Ecological Informatics*, *59*, 101132. https://doi.org/10.1016/j.ecoinf.2020.101132

Elton, C. S. (1927). *Animal ecology* (pp. 1–256). Macmillan Co. https://www.biodiversitylibrary.org/item/31642

Etard, A., Morrill, S., & Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. *Global Ecology and Biogeography*, *29*(12), 2143–2158. https://doi.org/10.1111/geb.13184

Evans, J. A., & Reimer, J. (2009). Open access and global participation in science. *Science*, *323*(5917), 1025. https://doi.org/10.1126/science.1154562

Falster, D., Gallagher, R., Wenk, E. H., Wright, I. J., Indiarto, D., Andrew, S. C., Baxter, C., Lawson, J., Allen, S., Fuchs, A., Monro, A., Kar, F., Adams, M. A., Ahrens, C. W., Alfonzetti, M., Angevin, T., Apgaua, D. M. G., Arndt, S., Atkin, O. K., … Ziemińska, K. (2021). AusTraits, a curated plant trait database for the Australian flora. *Scientific Data*, *8*(1), Article 1. https://doi.org/10.1038/s41597-021-01006-6

Fegraus, E. H., Andelman, S., Jones, M. B., & Schildhauer, M. (2005). Maximizing the value of ecological data with structured metadata: An introduction to ecological metadata language (EML) and principles for metadata creation. *The Bulletin of the Ecological Society of America*, *86*(3), 158–168. https://doi.org/10.1890/0012-9623(2005)86[158:MTVOED]2.0.CO;2

Feng, X., Enquist, B. J., Park, D. S., Boyle, B., Breshears, D. D., Gallagher, R. V., Lien, A., Newman, E. A., Burger, J. R., Maitner, B. S., Merow, C., Li, Y., Huynh, K. M., Ernst, K., Baldwin, E., Foden, W., Hannah, L., Jørgensen, P. M., Kraft, N. J. B., … López-Hoffman, L. (2022). A review of the heterogeneous landscape of biodiversity databases: Opportunities and challenges for a synthesized biodiversity knowledge base. *Global Ecology and Biogeography*, *31*(7), 1242–1260. https://doi.org/10.1111/geb.13497

Feng, X., Qiao, H., & Enquist, B. J. (2020). Doubling demands in programming skills call for ecoinformatics education. *Frontiers in Ecology and the Environment*, *18*(3), 123–124. https://doi.org/10.1002/fee.2179

Freckleton, R. P. (2000). Phylogenetic tests of ecological and evolutionary hypotheses: Checking for phylogenetic Independence. *Functional Ecology*, *14*(1), 129–134.

Gallagher, R. V., Butt, N., Carthey, A. J. R., Tulloch, A., Bland, L., Clulow, S., Newsome, T., Dudaniec, R. Y., & Adams, V. M. (2021). A guide to using species trait data in conservation. *One Earth*, *4*(7), 927–936. https://doi.org/10.1016/j.oneear.2021.06.013

Gallagher, R. V., Falster, D. S., Maitner, B. S., Salguero-Gómez, R., Vandvik, V., Pearse, W. D., Schneider, F. D., Kattge, J., Poelen, J. H., Madin, J. S., Ankenbrand, M. J., Penone, C., Feng, X., Adams, V. M., Alroy, J., Andrew, S. C., Balk, M. A., Bland, L. M., Boyle, B. L., … Enquist, B. J. (2020). Open Science principles for accelerating trait-based science across the tree of life. *Nature Ecology & Evolution*, *4*(3), Article 3. https://doi.org/10.1038/s41559-020-1109-6

Gama, J. (2016). *NISTunits: Fundamental physical constants and unit conversions from NIST*. R package version 1.0.1. https://CRAN.R-project.org/package=NISTunits

Garamszegi, L. Z. (Ed.). (2014). *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice*. Springer. https://doi.org/10.1007/978-3-662-43550-2

Garnier, E., Navas, M.-L., & Grigulis, K. (2016). *Plant functional diversity: Organism traits, community structure, and ecosystem properties*. Oxford University Press.

Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougenot, I., Kühn, I., Laporte, B., Amiaud, B., Ahrestani, F. S., Bönisch, G., Bunker, D. E., Cornelissen, J. H. C., Díaz, S., Enquist, B. J., Gachet, S., Jaureguiberry, P., Kleyer, M., Lavorel, S., Maicher, L., … Klotz, S. (2017). Towards a thesaurus of plant characteristics: An ecological contribution. *Journal of Ecology*, *105*(2), 298–309. https://doi.org/10.1111/1365-2745.12698

Geange, S. R., von Oppen, J., Strydom, T., Boakye, M., Gauthier, T.-L. J., Gya, R., Halbritter, A. H., Jessup, L. H., Middleton, S. L., Navarro, J., Pierfederici, M. E., Chacón-Labella, J., Cotner, S., Farfan-Rios, W., Maitner, B. S., Michaletz, S. T., Telford, R. J., Enquist, B. J., & Vandvik, V. (2021). Next-generation field courses: Integrating Open Science and online learning. *Ecology and Evolution*, *11*(8), 3577–3587. https://doi.org/10.1002/ece3.7009

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, *27*(15), 2865–2873. https://doi.org/10.1002/sim.3107

Gering, E., Incorvaia, D., Henriksen, R., Wright, D., & Getty, T. (2019). Maladaptation in feral and domesticated animals. *Evolutionary Applications*, *12*(7), 1274–1286. https://doi.org/10.1111/eva.12784

Global Names Architecture. (n.d.). *Globalnames.org*. https://globalnames.org/

González-Suárez, M., Lucas, P. M., & Revilla, E. (2012). Biases in comparative analyses of extinction risk: Mind the gap. *Journal of Animal Ecology*, *81*(6), 1211–1222. https://doi.org/10.1111/j.1365-2656.2012.01999.x

Grenié, M., Berti, E., Carvajal-Quintero, J., Dädlow, G. M. L., Sagouis, A., & Winter, M. (2022). Harmonizing taxon names in biodiversity data: A review of tools, databases and best practices. *Methods in Ecology and Evolution*, 1–14. https://doi.org/10.1111/2041-210X.13802

Guerrero-Ramírez, N. R., Mommer, L., Freschet, G. T., Iversen, C. M., McCormack, M. L., Kattge, J., Poorter, H., van der Plas, F., Bergmann, J., Kuyper, T. W., York, L. M., Bruelheide, H., Laughlin, D. C., Meier, I. C., Roumet, C., Semchenko, M., Sweeney, C. J., van Ruijven, J., Valverde-Barrantes, O. J., … Weigelt, A. (2021). Global root traits (GRooT) database. *Global Ecology and Biogeography*, *30*(1), 25–37. https://doi.org/10.1111/geb.13179

Hawkins, B. A., Leroy, B., Rodríguez, M. Á., Singer, A., Vilela, B., Villalobos, F., Wang, X., & Zelený, D. (2017). Structural bias in aggregated species-level variables driven by repeated species co-occurrences: A pervasive problem in community and assemblage data. *Journal of Biogeography*, *44*(6), 1199–1211. https://doi.org/10.1111/jbi.12953

Henn, J. J., Buzzard, V., Enquist, B. J., Halbritter, A. H., Klanderud, K., Maitner, B. S., Michaletz, S. T., Pötsch, C., Seltzer, L., Telford, R. J., Yang, Y., Zhang, L., & Vandvik, V. (2018). Intraspecific trait variation and phenotypic plasticity mediate alpine plant species response to climate change. *Frontiers in Plant Science*, *9*, 1548. https://doi.org/10.3389/fpls.2018.01548

Herberstein, M. E., McLean, D. J., Lowe, E., Wolff, J. O., Khan, M. K., Smith, K., Allen, A. P., Bulbert, M., Buzatto, B. A., Eldridge, M. D. B., Falster, D., Fernandez Winzer, L., Griffith, S. C., Madin, J. S., Narendra, A., Westoby, M., Whiting, M. J., Wright, I. J., & Carthey,

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

A. J. R. (2022). AnimalTraits—A curated animal trait database for body mass, metabolic rate and brain size. *Scientific Data*, *9*(1), Article 1. https://doi.org/10.1038/s41597-022-01364-9

Hulme, P. E., Pyšek, P., Jarošík, V., Pergl, J., Schaffner, U., & Vilà, M. (2013). Bias and error in understanding plant invasion impacts. *Trends in Ecology & Evolution*, *28*(4), 212–218. https://doi.org/10.1016/j.tree.2012.10.010

Isaac, M. E., & Martin, A. R. (2019). Accumulating crop functional trait data with citizen science. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-51927-x

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., … Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, *90*(9), 2648. https://doi.org/10.1890/08-1494.1

Junker, R. R., Blüthgen, N., & Keller, A. (2015). Functional and phylogenetic diversities of plant communities differently affect the structure of flower-visitor interactions and reveal convergences in floral traits. *Evolutionary Ecology*, *29*, 437–450.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar, C., Aleixo, I., Ali, H., … Wirth, C. (2020). TRY plant trait database – Enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. https://doi.org/10.1111/gcb.14904

Kattge, J., Ogle, K., Bönisch, G., Díaz, S., Lavorel, S., Madin, J., Nadrowski, K., Nöllert, S., Sartor, K., & Wirth, C. (2011). A generic structure for plant trait databases. *Methods in Ecology and Evolution*, *2*(2), 202–213. https://doi.org/10.1111/j.2041-210X.2010.00067.x

Kearney, M. R., Jusup, M., McGeoch, M. A., Kooijman, S. A. L. M., & Chown, S. L. (2021). Where do functional traits come from? The role of theory and models. *Functional Ecology*, *35*(7), 1385–1396. https://doi.org/10.1111/1365-2435.13829

Koehler, W. (1999). An analysis of web page and web site constancy and permanence. *Journal of the American Society for Information Science*, *50*(2), 162–180. https://doi.org/10.1002/(SICI)1097-4571(1999)50:2<162::AID-ASI7>3.0.CO;2-B

Koricheva, J., Gurevitch, J., & Mengersen, K. (2013). *Handbook of meta-analysis in ecology and evolution*. Princeton University Press.

Kostikova, A., Silvestro, D., Pearman, P. B., & Salamin, N. (2016). Bridging inter- and intraspecific trait evolution with a hierarchical Bayesian approach. *Systematic Biology*, *65*(3), 417–431. https://doi.org/10.1093/sysbio/syw010

Kunz, S., Kefford, B. J., Schmidt-Kloiber, A., Matthaei, C. D., Usseglio-Polatera, P., Graf, W., Poff, N. L., Metzeling, L., Twardochleb, L., Hawkins, C. P., & Schäfer, R. B. (2022). Tackling inconsistencies among freshwater invertebrate trait databases: Harmonising across continents and aggregating taxonomic resolution. *Freshwater Biology*, *67*(2), 275–291. https://doi.org/10.1111/fwb.13840

Langtangen, H. P., & Pedersen, G. K. (2016). *Scaling of differential equations*. Springer Nature.

Lavorel, S., & Garnier, E. (2002). Predicting changes in community composition and ecosystem functioning from plant traits: Revisiting the holy grail. *Functional Ecology*, *16*(5), 545–556. https://doi.org/10.1046/j.1365-2435.2002.00664.x

Leitão, R. P., Zuanon, J., Villéger, S., Williams, S. E., Baraloto, C., Fortunel, C., Mendonça, F. P., & Mouillot, D. (2016). Rare species contribute disproportionately to the functional structure of species assemblages. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1828), 20160084. https://doi.org/10.1098/rspb.2016.0084

Madin, J. S., Anderson, K. D., Andreasen, M. H., Bridge, T. C. L., Cairns, S. D., Connolly, S. R., Darling, E. S., Diaz, M., Falster, D. S., Franklin, E. C., Gates, R. D., Harmer, A. M. T., Hoogenboom, M. O., Huang, D., Keith, S. A., Kosnik, M. A., Kuo, C.-Y., Lough, J. M., Lovelock, C. E., … Baird, A. H. (2016). The coral trait database, a curated database of trait information for coral species from the global oceans. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.17

Madin, J. S., Bowers, S., Schildhauer, M., & Jones, M. (2008). Advancing ecological research with ontologies. *Trends in Ecology & Evolution*, *23*, 159–168. https://doi.org/10.1016/j.tree.2007.11.007

Madin, J. S., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., & Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, *2*(3), 279–296. https://doi.org/10.1016/j.ecoinf.2007.05.004

Madin, J. S., Nielsen, D. A., Brbic, M., Corkrey, R., Danko, D., Edwards, K., Engqvist, M. K. M., Fierer, N., Geoghegan, J. L., Gillings, M., Kyrpides, N. C., Litchman, E., Mason, C. E., Moore, L., Nielsen, S. L., Paulsen, I. T., Price, N. D., Reddy, T. B. K., Richards, M. A., … Westoby, M. (2020). A synthesis of bacterial and archaeal phenotypic trait data. *Scientific Data*, *7*(1), 170. https://doi.org/10.1038/s41597-020-0497-4

Maitner, B., Halbritter, A., Telford, R., Strydom, T., Chacon-Labella, J., Lamanna, C., Sloat, L., Kerkhoff, A., Messier, J., Rasmussen, N., Pomati, F., Merz, E., Vandvik, V., & Enquist, B. (2021). On estimating the shape and dynamics of phenotypic distributions in ecology and evolution [preprint]. https://doi.org/10.22541/au.162196147.76797968/v1

Marine Species Traits Editorial Board. (2022). *Marine species traits*. http://www.marinespecies.org/traits

McGraw, J. B., & Caswell, H. (1996). Estimation of individual fitness from life-history data. *The American Naturalist*, *147*(1), 47–64. https://doi.org/10.1086/285839

Messier, J., McGill, B. J., Enquist, B. J., & Lechowicz, M. J. (2017). Trait variation and integration across scales: Is the leaf economic spectrum present at local scales? *Ecography*, *40*(6), 685–697. https://doi.org/10.1111/ecog.02006

Michener, W. K. (2006). Meta-information concepts for ecological data management. *Ecological Informatics*, *1*(1), 3–7. https://doi.org/10.1016/j.ecoinf.2005.08.004

Moretti, M., Dias, A. T. C., de Bello, F., Altermatt, F., Chown, S. L., Azcárate, F. M., Bell, J. R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J. P., Ellers, J., & Berg, M. P. (2017). Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, *31*(3), 558–567. https://doi.org/10.1111/1365-2435.12776

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017). AmphiBIO, a global database for amphibian ecological traits. *Scientific Data*, *4*(1), Article 1. https://doi.org/10.1038/sdata.2017.123

Parr, C., Schulz, K., Hammock, J., Wilson, N., Leary, P., Rice, J., & Corrigan, R., Jr. (2015). TraitBank: Practical semantics for organism attribute data. *Semantic Web*, *7*, 577–588. https://doi.org/10.3233/SW-150190

Paterno, G. B., Penone, C., & Werner, G. D. A. (2018). sensiPhy: An r-package for sensitivity analysis in phylogenetic comparative methods. *Methods in Ecology and Evolution*, *9*(6), 1461–1467. https://doi.org/10.1111/2041-210X.12990

Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., & Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

approach performs the best? *Methods in Ecology and Evolution*, *5*(9), 961–970. https://doi.org/10.1111/2041-210X.12232

Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M. S., Cornwell, W. K., Craine, J. M., Gurvich, D. E., Urcelay, C., Veneklaas, E. J., Reich, P. B., Poorter, L., Wright, I. J., Ray, P., Enrico, L., Pausas, J. G., de Vos, A. C., … Cornelissen, J. H. C. (2013). New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, *61*(3), 167. https://doi.org/10.1071/BT12225

Pillar, V. D., Sabatini, F. M., Jandt, U., Camiz, S., & Bruelheide, H. (2021). Revealing the functional traits linked to hidden environmental factors in community assembly. *Journal of Vegetation Science*, *32*(1), e12976. https://doi.org/10.1111/jvs.12976

Purschke, O., Michalski, S. G., Bruelheide, H., & Durka, W. (2017). Phylogenetic turnover during subtropical forest succession across environmental and phylogenetic scales. *Ecology and Evolution*, *7*(24), 11079–11091. https://doi.org/10.1002/ece3.3564

Roberts, L. W. (2004). How to review a manuscript: A "Down-to-earth" approach. *Academic Psychiatry*, *28*(2), 81–87. https://doi.org/10.1176/appi.ap.28.2.81

Rüegg, J., Gries, C., Bond-Lamberty, B., Bowen, G. J., Felzer, B. S., McIntyre, N. E., Soranno, P. A., Vanderbilt, K. L., & Weathers, K. C. (2014). Completing the data life cycle: Using information management in macrosystems ecology research. *Frontiers in Ecology and the Environment*, *12*, 24–30. https://doi.org/10.1890/120375

Ryder, E. F., Morgan, W. R., Sierk, M., Donovan, S. S., Robertson, S. D., Orndorf, H. C., Rosenwald, A. G., Triplett, E. W., Dinsdale, E., Pauley, M. A., & Tapprich, W. E. (2020). Incubators: Building community networks and developing open educational resources to integrate bioinformatics into life science education. *Biochemistry and Molecular Biology Education*, *48*(4), 381–390. https://doi.org/10.1002/bmb.21387

Salguero-Gómez, R., Jackson, J., & Gascoigne, S. J. L. (2021). Four key challenges in the open-data revolution. *Journal of Animal Ecology*, *90*(9), 2000–2004. https://doi.org/10.1111/1365-2656.13567

Schiller, C., Schmidtlein, S., Boonman, C., Moreno-Martínez, A., & Kattenborn, T. (2021). Deep learning and citizen science enable automated plant trait predictions from photographs. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-95616-0

Schneider, F. D., Fichtmueller, D., Gossner, M. M., Güntsch, A., Jochum, M., König-Ries, B., Le Provost, G., Manning, P., Ostrowski, A., Penone, C., & Simons, N. K. (2019). Towards an ecological trait-data standard. *Methods in Ecology and Evolution*, *10*(12), 2006–2019. https://doi.org/10.1111/2041-210X.13288

Soranno, P. A., Cheruvelil, K. S., Elliott, K. C., & Montgomery, G. M. (2015). It's good to share: Why environmental Scientists' ethics are out of date. *Bioscience*, *65*(1), 69–73. https://doi.org/10.1093/biosci/biu169

Spigt, M., & Arts, I. C. W. (2010). How to review a manuscript. *Journal of Clinical Epidemiology*, *63*(12), 1385–1390. https://doi.org/10.1016/j.jclinepi.2010.09.001

Spinellis, D. (2012). Git. *IEEE Software*, *29*(3), 100–101. https://doi.org/10.1109/MS.2012.61

Sporbert, M., Welk, E., Seidler, G., Jandt, U., Aćić, S., Biurrun, I., Campos, J. A., Čarni, A., Cerabolini, B. E. L., Chytrý, M., Ćušterevska, R., Dengler, J., De Sanctis, M., Dziuba, T., Fagúndez, J., Field, R., Golub, V., He, T., Jansen, F., … Bruelheide, H. (2021). Different sets of traits explain abundance and distribution patterns of European plants at different spatial scales. *Journal of Vegetation Science*, *32*(2), e13016. https://doi.org/10.1111/jvs.13016

Stavert, J. R., Liñán-Cembrano, G., Beggs, J. R., Howlett, B. G., Pattemore, D. E., & Bartomeus, I. (2016). Hairiness: The missing link between pollinators and pollination. *PeerJ*, *4*, e2779. https://doi.org/10.7717/peerj.2779

Sutherland, W. J., Freckleton, R. P., Godfray, H. C. J., Beissinger, S. R., Benton, T., Cameron, D. D., Carmel, Y., Coomes, D. A., Coulson, T., Emmerson, M. C., Hails, R. S., Hays, G. C., Hodgson, D. J., Hutchings, M. J., Johnson, D., Jones, J. P. G., Keeling, M. J., Kokko, H., Kunin, W. E., … Wiegand, T. (2013). Identification of 100 fundamental ecological questions. *Journal of Ecology*, *101*(1), 58–67. https://doi.org/10.1111/1365-2745.12025

Thessen, A., Preciado, J., Jain, P., Martin, J., Palmer, M., & Bhat, R. (2018). Automated trait extraction using ClearEarth, a natural language processing system for text Mining in Natural Sciences. *Biodiversity Information Science and Standards*, *2*, e26080. https://doi.org/10.3897/biss.2.26080

Tobias, J. A., Sheard, C., Pigot, A. L., Devenish, A. J. M., Yang, J., Sayol, F., Neate-Clegg, M. H. C., Alioravainen, N., Weeks, T. L., Barber, R. A., Walkden, P. A., MacGregor, H. E. A., Jones, S. E. I., Vincent, C., Phillips, A. G., Marples, N. M., Montaño-Centellas, F. A., Leandro-Silva, V., Claramunt, S., … Schleuning, M. (2022). AVONET: Morphological, ecological and geographical data for all birds. *Ecology Letters*, *25*(3), 581–597. https://doi.org/10.1111/ele.13898

Tucker, C. M., Davies, T. J., Cadotte, M. W., & Pearse, W. D. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology*, *99*(6), 1473–1479. https://doi.org/10.1002/ecy.2349

Vandvik, V., Halbritter, A. H., Yang, Y., He, H., Zhang, L., Brummer, A. B., Klanderud, K., Maitner, B. S., Michaletz, S. T., Sun, X., Telford, R. J., Wang, G., Althuizen, I. H. J., Henn, J. J., Garcia, W. F. E., Gya, R., Jaroszynska, F., Joyce, B. L., Lehman, R., … Enquist, B. J. (2020). Plant traits and vegetation data from climate warming experiments along an 1100 m elevation gradient in Gongga Mountains, China. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-0529-0

Violle, C., Borgy, B., & Choler, P. (2015). Trait databases: Misuses and precautions. *Journal of Vegetation Science*, *26*(5), 826–827. https://doi.org/10.1111/jvs.12325

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, *116*(5), 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Walker, T. W. N., Alexander, J. M., Allard, P.-M., Baines, O., Baldy, V., Bardgett, R. D., Capdevila, P., Coley, P. D., David, B., Defossez, E., Endara, M.-J., Ernst, M., Fernandez, C., Forrister, D., Gargallo-Garriga, A., Jassey, V. E. J., Marr, S., Neumann, S., Pellissier, L., … Salguero-Gómez, R. (2022). Functional traits 2.0: The power of the metabolome for ecology. *Journal of Ecology*, *110*(1), 4–20. https://doi.org/10.1111/1365-2745.13826

Webb, C. T., Hoeting, J. A., Ames, G. M., Pyne, M. I., & LeRoy Poff, N. (2010). A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology Letters*, *13*(3), 267–283. https://doi.org/10.1111/j.1461-0248.2010.01444.x

Weiher, E., van der Werf, A., Thompson, K., Roderick, M., Garnier, E., & Eriksson, O. (1999). Challenging theophrastus: A common core list of plant traits for functional ecology. *Journal of Vegetation Science*, *10*(5), 609–620. https://doi.org/10.2307/3237076

Westoby, M., Falster, D. S., & Schrader, J. (2021). Motivating data contributions via a distinct career currency. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1946), 20202830. https://doi.org/10.1098/rspb.2020.2830

Westoby, M., Leishman, M. R., & Lord, J. M. (1995). On misinterpreting the 'phylogenetic correction'. *Journal of Ecology*, *83*(3), 531–534. https://doi.org/10.2307/2261605

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin core: An evolving community-developed biodiversity data standard. *PLoS One*, *7*(1), e29715. https://doi.org/10.1371/journal.pone.0029715

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B.,

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.18

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology*, *95*(7), 2027. https://doi.org/10.1890/13-1917.1

Wright, J. P., Ames, G. M., & Mitchell, R. M. (2016). The more things change, the more they stay the same? When is trait variability important for stability of ecosystem function in a changing environment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1694), 20150272. https://doi.org/10.1098/rstb.2015.0272

How to cite this article: Keller, A., Ankenbrand, M. J., Bruelheide, H., Dekeyzer, S., Enquist, B. J., Erfanian, M. B., Falster, D. S., Gallagher, R. V., Hammock, J., Kattge, J., Leonhardt, S. D., Madin, J. S., Maitner, B., Neyret, M., Onstein, R. E., Pearse, W. D., Poelen, J. H., Salguero-Gomez, R., Schneider, F. D. … Penone, C. (2022). Ten (mostly) simple rules to future-proof trait data in ecological and evolutionary sciences. *Methods in Ecology and Evolution*, *00*, 1–15. https://doi.org/10.1111/2041-210X.14033

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14033 by MPI 322 Chemical Ecology, Wiley Online Library on [23/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License