

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

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

<sup>1</sup>Cellular and Organismic Networks, Faculty of Biology, Ludwig-Maximilians-Universität München

<sup>2</sup>Center for Computational and Theoretical Biology, Julius-Maximilians-Universität Würzburg

<sup>3</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig

<sup>4</sup>Flanders Marine Institute (VLIZ)

<sup>5</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson

<sup>6</sup>Department of Biology, Faculty of Science, Ferdowsi University of Mashhad

<sup>7</sup>Evolution & Ecology Research Centre, University of New South Wales Sydney

<sup>8</sup>Hawkesbury Institute for the Environment, Western Sydney University

<sup>9</sup>National Museum of Natural History, Smithsonian Institution

<sup>10</sup>Max Planck Institute for Biogeochemistry, Jena

<sup>11</sup>Plant-Insect Interactions, TUM School of Life Science Systems, Technical University of Munich

<sup>12</sup>Hawaii Institute of Marine Biology, University of Hawaii at Manoa

<sup>13</sup>University at Buffalo, Department of Geography

<sup>14</sup>Senckenberg Biodiversity and Climate Research Center (SBik-F)

<sup>15</sup>Department of Life Sciences, Imperial College London

<sup>16</sup>Ronin Institute for Independent Scholarship

<sup>17</sup>Department of Zoology, University of Oxford

<sup>18</sup>ISOE - Institute for Social-Ecological Research

<sup>19</sup>School of Biological Earth and Environmental Sciences, University of New South Wales

<sup>20</sup>Institute of Plant Sciences, University of Bern

July 23, 2022

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

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

9  
10 <sup>1</sup> Cellular and Organismic Networks, Faculty of Biology, Ludwig-Maximilians-Universität München,  
11 82152 Planegg-Martinsried, Germany

12 <sup>2</sup> Center for Computational and Theoretical Biology, Julius-Maximilians-Universität Würzburg,  
13 97070 Würzburg, Germany

14 <sup>3</sup> Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg,  
15 Am Kirchtor 1, Halle (Saale), 06108, Germany

16 <sup>4</sup> German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4,  
17 Leipzig, 04103, Germany

18 <sup>5</sup> Flanders Marine Institute (VLIZ), InnovOcean Site, Wandelaarkaai 7, 8400 Oostende, Belgium

19 <sup>6</sup> Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721,  
20 United States and The Santa Fe Institute, Santa Fe, NM 87501

21 <sup>7</sup> Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

22 <sup>8</sup> Evolution & Ecology Research Centre, University of New South Wales Sydney, Australia

23 <sup>9</sup> Hawkesbury Institute for the Environment, Western Sydney University, Richmond NSW, Australia

24 <sup>10</sup> National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

25 <sup>11</sup> Max Planck Institute for Biogeochemistry, Jena, Germany

26 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

27 <sup>12</sup> Plant-Insect Interactions, TUM School of Life Science Systems, Technical University of Munich,  
28 Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising

29 <sup>13</sup> Hawai‘i Institute of Marine Biology, University of Hawai‘i at Manoa, Kāne‘ohe, Hawai‘i, USA

30 <sup>14</sup> University at Buffalo, Department of Geography, 120 Wilkeson Quadrangle, 14261 Buffalo, NY,  
31 United States. Department of Environment and Sustainability. Department of Environment and  
32 Sustainability, 602 Clemens Hall, North Campus, 14260 Buffalo, NY United States

33 <sup>15</sup> Senckenberg Biodiversity and Climate Research Center (SBik-F), Frankfurt, Germany

34 <sup>16</sup> Department of Life Sciences, Imperial College London, Ascot SL5 7PY, United Kingdom

35 <sup>17</sup> Ronin Institute for Independent Scholarship, Montclair, New Jersey, United States

36 <sup>18</sup> Cheadle Center for Biodiversity and Ecological Restoration, UC Santa Barbara, Santa Barbara,  
37 California, United States

38 <sup>19</sup> Department of Zoology, University of Oxford, 11a Mansfield Road, OX1 3SZ, Oxford, UK

39 <sup>20</sup> ISOE - Institute for Social-Ecological Research, Frankfurt am Main, Germany

40 <sup>21</sup> Centre for Ecosystem Science, School of Biological Earth and Environmental Sciences, University  
41 of New South Wales, Sydney Australia

42 <sup>22</sup> Institute of Plant Sciences, University of Bern, Switzerland

43

44 <sup>§</sup> Corresponding author: Alexander Keller, Cellular and Organismic Networks, Faculty of Biology,  
45 Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany, keller@bio.lmu.de

46

47 <sup>&</sup> Corresponding author: Caterina Penone, Institute of Plant Sciences, University of Bern,  
48 Switzerland, caterina.penone@ips.unibe.ch

49

50

51 **Running head:** Ten rules to future-proof trait data

52

53

## 54 Abstract

55 1. Traits have become a crucial part of ecological and evolutionary sciences, helping  
56 researchers understand the function of an organism's morphology, physiology, growth and  
57 life-history, with effects on fitness, behaviour, interactions with the environment, and  
58 ecosystem processes. However, compiling and analysing trait data comes with data-scientific  
59 challenges due to the complex nature of trait data.

60 2. We offer 10 (mostly) simple rules, with some detailed extensions, as a guide in making  
61 critical decisions that consider the entire life cycle of trait data.

62 3. This article is particularly motivated by its last rule, i.e. to propagate the good practices,  
63 and has the intention of bringing awareness of the different facets of a trait's life cycle to the  
64 community.

65 4. When it comes to working with trait data, we gain particularly as an interdisciplinary  
66 community of field biologists, synthesis ecologists, computer scientists and database  
67 managers. We hope these basic guidelines can be useful as a starter for active communication  
68 in disseminating such integrative knowledge and how to make trait data future-proof.

69

70 Keywords: data life cycle, data science, FAIR principles, good practices, metadata, open  
71 science, phenotype, trait data

72

73

74

75

76

## 77 Introduction

78 As early as 300 BC, Greek philosophers such as Theophrastus forged the first formal systems  
79 defining and classifying organisms by their combination of morphological, physiological,  
80 behavioural, and phenological characteristics, *i.e.*, their traits (Weiher et al., 1999). Having  
81 knowledge about an organism's traits often allows deeper understanding of its life-history,  
82 behaviour, fitness, interactions and potential responses to and effects on ecosystem processes  
83 (Violle et al., 2007). Trait-based research questions in ecology and evolution often focus on  
84 the processes that drive trait patterns (Sutherland et al., 2013), and view the organism as a  
85 functional rather than a taxonomic unit. Traits offer a bridge between different dimensions:  
86 from organismal biology, *e.g.* population abundance (Webb et al., 2010), species distribution  
87 (Sporbert et al., 2021), and phylogeny (Junker et al., 2015; Tucker et al., 2018) to ecosystem  
88 functioning (Lavorel & Garnier, 2002; Wright et al., 2016).

89

90 Traits are commonly defined as a measure of an entity (Garnier et al., 2017), where the entity  
91 can be the whole individual, or a specific organ or tissue (*e.g.* a tail fin, a leaf or wood) and  
92 the quantity is an observable characteristic of that entity (*e.g.* the length of a fish, its colour,  
93 the length of a tail fin, or the density of wood tissue). Together, all traits of an individual  
94 organism represent its phenotype resulting from the historical evolution of the genotype and  
95 potential current interactions with the environment. Therefore, a trait record can inform not  
96 only about the entity that was observed (*e.g.* taxonomic classification or age) and the  
97 quantity/characteristic that was measured, but also about the environment in which the  
98 individual has developed that trait (*e.g.* where a fish was caught, where a tree lived or the soil  
99 depth where an invertebrate was observed)(Kattge et al., 2011).

100

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

111

112 In essence, trait data are a special kind of data: they are diverse, often not well defined,  
113 relatively simple or potentially complex, independent of one-another or correlated, and range  
114 between cheap and costly to collect. However, they are very informative as they represent the  
115 evolutionary adaptation or developmental acclimation of the individual organisms to their  
116 environment and allow for quantitative and predictive ecology and biodiversity research.  
117 Therefore - if collected, stored and published in a meaningful way - organismal trait data  
118 have an extraordinary value for reuse, which is *e.g.* indicated by the >20.000 data requests to  
119 the TRY Plant Trait Database since 2015 (Kattge et al., 2020).

120

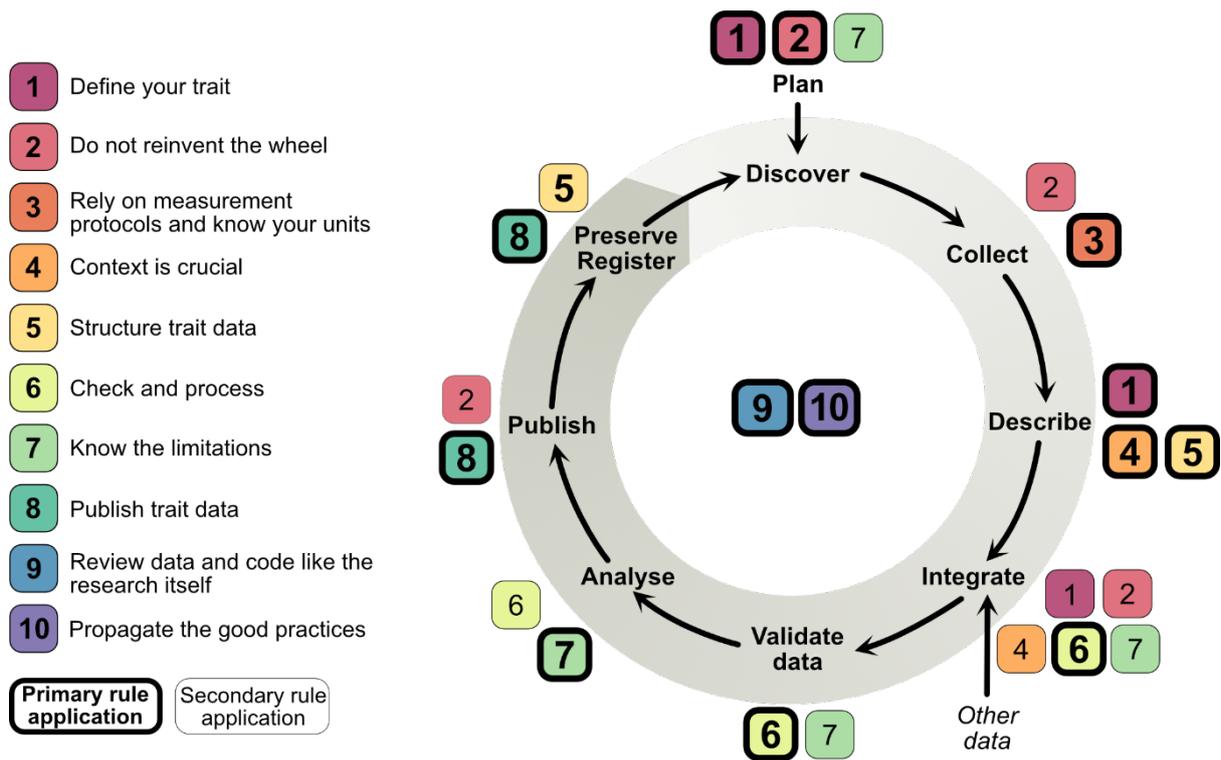
121 To enable the reuse of these trait data beyond their original research campaign, to make them  
122 meaningful in other contexts and to avoid data degradation, observation records must be  
123 clearly defined, where possible the environmental and meteorological context given, as well as  
124 provenance and protocols for collection documented (Michener, 2006). Recent efforts to  
125 expand trait knowledge across the Tree of Life (Gallagher et al., 2020) call for datasets which

126 are open and FAIR (Findable, Accessible, Interoperable, Reusable) (Wilkinson et al., 2016),  
127 key principles at the heart of the emergent Open Science movement (Nosek et al., 2015).  
128 Global and local datasets of organismal traits have rapidly grown in scope since the 1990's  
129 (e.g. Herberstein et al., 2022; Kattge et al., 2020; J. S. Madin et al., 2016). However, these  
130 datasets bear a variety of new challenges linked to harmonisation, biases, expertise and  
131 communication (Salguero-Gómez et al., 2021). These challenges result in a major trade-off  
132 between investing in collection of new trait data or reusing open trait data (Westoby et al.,  
133 2021). Indeed, many studies in trait-based research reuse available trait data or collect  
134 additional trait data and/or assemble new data (e.g. examples in Kattge et al., 2020). Thus,  
135 these studies also often involve linking to different types of data, which requires  
136 interoperability between datasets (Feng et al., 2022; Gallagher et al., 2020).

137

138 These key aspects are just a few of many dimensions illustrating how and why researchers  
139 not only have to make biological decisions, but also a wide range of data-science decisions  
140 when working with traits. Multiple complexities of trait data structure and manipulation are  
141 not obvious at a first glance (Michener, 2006). For instance, there is sometimes confusion and  
142 lack of awareness of trait standards, measurement units, and trait data are particularly prone  
143 to errors in recording, translation, and understanding. By offering a larger perspective, the  
144 trait data life cycle can help clarify these confusions and inform about good practices when  
145 working with trait data (Fig.1). In this article, we bring awareness to some common pitfalls in  
146 the usage of trait data and offer 10 rules for making critical decisions that consider the entire  
147 life cycle of trait data. We start each rule with a general and simple statement and develop the  
148 complexity of each rule within more detailed subsections.

149



150

151 **Figure 1:** Ten (mostly) simple rules and where they apply in the overall trait data life cycle. Each rule is  
 152 primarily applied to a specific element of the cycle (in bold) but can also be important to other elements  
 153 (secondary application). Rules 9 and 10 apply to the whole cycle.

154

## 155 Rule 1: Define your trait

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

160

161 **Follow your hypothesis:** Increasingly, trait data describing organisms of interest are publicly  
 162 available for reuse. Vast availability extends the potential scope of what is possible with  
 163 limited resources (e.g. Falster et al., 2021; Kattge et al., 2020). However, when reusing trait  
 164 data, we relinquish control of what variables are collected, which species are sampled, and

165 the methods used for collection (Koricheva et al., 2013). Undirected fishing expeditions for  
166 traits can yield large datasets, but these may not necessarily be appropriate to answer a given  
167 research question, for various reasons (*e.g.* coverage, distribution, meaningfulness and  
168 resolution, Violle et al., 2015). Furthermore, the wealth of available trait data may distract  
169 from initial hypotheses, risking random exploration of the available traits and fishing for  
170 significant relationships without a clear focus. Thus, trait selection and collection should in  
171 most cases be primarily tethered to a concrete hypothesis, not defined by availability of  
172 existing data. This rule does, however, not completely exclude extensive data exploration and  
173 data-driven discovery within a given range as relevant to the research question and  
174 subsequent streamlining (Violle et al., 2015).

175

176 ***Consider the scale:*** Research questions define the appropriate hierarchical level for sampling:  
177 a continental-scale study of thousands of species may treat the intra-specific variation as  
178 statistical noise. In contrast, this variation may be the study focus in locally scaled projects.  
179 There is no "correct" scale, either in terms of spatial grain (*e.g.*, km<sup>2</sup>, m<sup>2</sup>), temporal duration  
180 (*e.g.*, seconds, years), or taxonomic coverage (*e.g.*, clade, species, population or individual),  
181 but not every scale will be appropriate for every question. So, when defining the traits of  
182 interest, it is important to define the scale at which these need to be collected or aggregated to  
183 match the research question (Messier et al., 2017).

184

185 ***Be aware of existing trait definitions and homologies:*** Much effort has already gone into  
186 creating definitions and protocols for traits collection. Yet, trait naming and corresponding  
187 definitions may differ between studies and trait databases (Ankenbrand et al., 2018; Dawson  
188 et al., 2021; Kunz et al., 2022). For example, specific leaf area (SLA) and leaf mass per area  
189 (LMA) are essentially the same trait, one being the inverse of the other. Similarly, for discrete

190 traits, values may differ between resources (*e.g.* “therophyte” and “annual” are synonyms).  
191 Furthermore, when comparing traits and trait states across organisms, it is important to be  
192 aware of the ‘homology’ of the character. Homologous traits share similarity of structure,  
193 physiology, or development (often by common evolutionary ancestry), whereas non-  
194 homologous (or analogous) characters may perform a similar function, but differ in structure,  
195 physiology, or development.

196

197 ***Be pragmatic:*** In a perfect trait research world, we could measure or retrieve the exact traits  
198 for the exact scale and organisms needed to answer our specific question. This vision is rarely  
199 real. Instead, we often need to work with proxies of traits that are difficult to measure (*e.g.*  
200 reproductive output is often used as a proxy for fitness (McGraw & Caswell, 1996)) or are  
201 incomplete in a database (*e.g.* diet or behavioural traits are less complete than morphological  
202 traits, Oliveira et al., 2017). There is a common understanding of these technical or financial  
203 limitations in the scientific community; ultimately, we need to be pragmatic to advance  
204 research questions. However, it is crucial to explain and justify the choice of traits, especially  
205 when these are used as proxies or “best available data”.

## 206 Rule 2: Do not reinvent the wheel

207 Build on existing trait resources to reduce the likelihood of redundancy and ensure  
208 compatibility with current data. The decision when to collect new trait data is generally based  
209 on the research question, the scope of the analysis (*e.g.* local, global) and the availability of  
210 the existing data. Financial and geographic constraints may also influence the decision to use  
211 current trait data instead of embarking on a measurement campaign. However, the existing  
212 trait data must be ‘fit for purpose’ to avoid compromising the capacity to answer the research  
213 question.

214

215 **Check public data sources:** Most data probably exist decentralised as individual trait datasets  
216 in the form of raw data attachments to publications, data papers, or data uploads to unspecific  
217 public databases (e.g. Zenodo <https://zenodo.org>, DataDryad <https://datadryad.org>). However,  
218 these datasets can be challenging if not registered at central hubs (e.g. <https://opentraits.org>).  
219 To counter this challenge, dedicated centralised trait databases have been and continue to be  
220 developed (e.g., TRY (Kattge et al., 2020), Encyclopedia of Life (EOL) TraitBank (Parr et  
221 al., 2015), Marine Traits Portal of the World Register of Marine Species (WoRMS) (Marine  
222 Species Traits editorial board), AusTraits (Falster et al., 2021). Common to these efforts is  
223 the fact that they contain already harmonised, error-checked, and standardised values. These  
224 resources usually provide user-friendly interfaces for searches and dynamic, up-to-date  
225 aggregations of data. Particularly for studies of larger scale (e.g. many taxa, many  
226 bioregions), it often makes sense to consult these existing big databases and data registries.  
227

228 **Identify data origins:** Trait data are not always raw or first-hand: they can be created and  
229 perhaps aggregated from original observations and measurements (e.g. Kattge et al., 2020),  
230 but also mobilised from literature or undigitised legacy trait data (e.g. Parr et al., 2015),  
231 synthesised as imputed trait data (e.g. Penone et al., 2014), reused from data publications  
232 (e.g. Kattge et al., 2020), or mined from texts with automated algorithms or other contexts  
233 (Thessen et al., 2018). Thus, when reusing trait data, it is important to check and report  
234 information about the source to downstream analyses and subsequent publications.  
235 Importantly, providing this information also gives credit to the original trait data collectors.  
236

237 **Fill the gaps:** Existing databases are taxonomically and biogeographically biased, ‘gappy’,  
238 and traits assigned to the same species are rarely collected in the same locations or conditions

239 (Etard et al., 2020; Penone et al., 2014). Despite the presence of large trait databases, new  
240 trait collections continue to remain valuable. When collecting new data, we encourage  
241 researchers to first check available trait databases, identify such gaps, and contribute to the  
242 wider trait community by filling these gaps even if this collection goes beyond the current  
243 project. Additional traits may be easily collected with little extra effort yet provide the  
244 possibility to close gaps in trait coverage. Filling gaps may be especially valuable in  
245 biodiverse but hard to access regions (Etard et al., 2020), for rare but functionally important  
246 species which may be less likely to have traits documented (Leitão et al., 2016), or for  
247 threatened species which will benefit from functional approaches to their conservation  
248 (Gallagher et al., 2021).

### 249 **Rule 3: Rely on measurement protocols and know your units**

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

257 ***Beware of ambiguities:*** In most cases, researchers of a domain (*e.g.* plants) have adopted  
258 sufficiently specific trait definitions to allow comparison of widely used measurements and  
259 enable synthesis within the field. However, some difficulties in measurement remain. To  
260 illustrate, specific leaf area (SLA) is the ratio of the surface area to leaf biomass of an  
261 individual leaf. However, the application of the concept of SLA may differ between research

262 contexts, because the value reported may relate to measurements of individual leaves or an  
263 average of all leaves on the shoot, for one or both sides of the leaf, including or excluding the  
264 petiole, and focus on the leaf or leaflet (example taken from Garnier et al., 2017). While fully  
265 justified in the specific research setting, identifying and dealing with semantic  
266 disambiguation is a major challenge in trait-based synthesis.

267 ***Adhere to existing standards:*** Methodological handbooks for trait measurements have been  
268 proposed, *e.g.* for plant (Cornelissen et al., 2003), macrofungi (Dawson et al., 2019) or for  
269 terrestrial invertebrate functional traits (Moretti et al., 2017). These handbooks provide  
270 precise, taxon-specific definitions and recommended methods for trait measurement,  
271 measurement precision and replication. They also provide considerations and warnings of  
272 misconception and error, and point to the key literature debating the methodology. Taking  
273 formalisation of trait concepts even one step further are thesauri of trait concepts (Garnier et  
274 al., 2016, 2017), *e.g.* TOP. The bottom line is: research that provides original trait  
275 measurements should consider existing measurement protocols, make an explicit choice, and  
276 describe any deviations from or additions to protocols. When such handbooks do not exist it  
277 is good practice to also report accompanied the specific measurement protocol in the  
278 metadata (see Rule 4). For instance, how the length of a fish has been measured and if  
279 potential extensions of the tail fin were taken into account.

280

281 ***Understand your units:*** Trait data are necessarily ‘rich in dimensions’. That is, trait data may  
282 require multiple SI base units and may also be measured and reported in multiple alternative  
283 configurations of units. For example, photosynthetic rate involves three SI base units *e.g.*  
284 mass per area per time, and is often reported in units of  $\mu\text{mol CO}_2\text{m}^{-2}\text{s}^{-1}$  or an amount per unit  
285 area per unit time. Simple measures of size, area, and time are often reported in differing  
286 units, though all can be related to more fundamental base units. All metric trait data can be

287 reduced to the seven base units as defined by the SI standard (m, g, s, K, A, cd, mol).  
288 Significant data management effort is needed to record units accurately, preserve them  
289 through metadata, and convert them correctly to avoid propagating errors (Calder, 1982).

## 290 Rule 4: Context is crucial

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

300

301 ***Define at least the minimum context:*** Some metadata are considered essential and universal  
302 between all domains, such as unique ID for observations to cross-references to other  
303 measures, geolocation, time and date, life stage (*e.g.* juvenile), health status, scale (*e.g.* leaf),  
304 habitat type (*e.g.* natural habitat) and measurement details (*e.g.* following standards, devices  
305 used, etc.)(Schneider et al., 2019). Further metadata must include the source and authorship  
306 of the trait measurement. To permit effective reuse, authorship attributes should include the  
307 original data collectors and the databases where these data were gathered, as they may have  
308 undergone processing therein (Rule 2).

309

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

317

318 ***Link to other data by metadata:*** A good practice is to directly link the data with publications  
319 (*e.g.* by DOI) for the scientific context and further information in the materials and methods  
320 sections, as well as identification of trait data providers (*e.g.* by ORCID) to provide  
321 opportunities for feedback and requests for additional information. Often traits are measured  
322 with the objective to also collect other data, such as ecosystem function (*e.g.* Bongers et al.,  
323 2021) or species composition (*e.g.* Breitschwerdt et al., 2018). In these cases, functions  
324 measured, and species composition recorded would be part of the metadata, or links to that  
325 data in other repositories.

## 326 **Rule 5: Structure trait data**

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

332

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

339

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

349

350 **Apply version control:** The process from gathering to analysing trait data is long, which  
351 might result in many different versions of a trait dataset. Thus, it is important to keep track of  
352 these different versions. It is recommended to keep the raw trait dataset and the processed  
353 trait dataset used for analysis as separate copies. Version names should be structured in a  
354 logical way, *e.g.* combining project acronyms, researchers' initials, short names of the trait  
355 dataset, version numbers, file status and/or dates. Use of versioning or change control  
356 systems like Git (Spinellis, 2012) is highly recommended to keep track of changes to data.

357 Also, conventions that include a current (*e.g.* my\_data\_current.tsv), as well as versioned copy  
358 (*e.g.* my\_data\_1.0.tsv) help automated systems by providing a stable interface for indexing.

## 359 Rule 6: Check and process

360 Rigorously check your data quality, integrity and compatibility during each step of data  
361 processing. Trait-based analyses, mainly when data are consolidated from different sources,  
362 can harbour various inherent incompatibilities that may cause biases and severe scientific  
363 misinterpretations. For trait compilations, data usually need to be harmonised, subset,  
364 transformed, derived and/or aggregated into comparable formats to fit the research question.  
365 It is crucial that steps are wherever possible scripted, and by that directly reproducible, and  
366 where not that manual steps are well documented.

367

368 ***Harmonise trait data:*** If trait data originate from multiple sources, each source may identify  
369 the same entities or concepts differently (Kunz et al., 2022). Harmonisation is crucial to  
370 reconcile equivalent entities and explicitly connect related entities by “similar” or subclass  
371 relationships. Ideally, these entities or concepts should be identified by standard identifiers  
372 (see Rule 5). Manual harmonisation may be necessary to detect and reconcile spelling  
373 variations, before text strings are mapped to identifiers. But for common classes of data there  
374 are a variety of services available that allow automated and reproducible harmonisation, *e.g.*  
375 for taxonomic names (B. Boyle et al., 2013; Chamberlain & Szöcs, 2013; Global Names  
376 Architecture.) (reviewed by Grenié et al., 2022), units (Gama, 2014) or geographic names  
377 (Boyle et al., 2022). Other covariates and categorical trait values may be semantically  
378 reconciled where appropriate ontologies exist (Kunz et al., 2022; Violle et al., 2015).

379

380 ***Filter where needed and double-check data contexts:*** Not all trait data are equally suitable  
381 for all purposes. Erroneous or duplicate data points need to be identified and removed before  
382 analyses (Kattge et al., 2020). As with other kinds of data, outlier detection and data  
383 visualisation provide valuable methods for the detection of such data errors (de Bello et al.,  
384 2013). For trait data mainly compiled from different sources, other reasons may also render  
385 data points inappropriate. For example, if metadata suggest that the observation is from a  
386 cultivated occurrence such as a botanical garden, greenhouse, zoo or farm, values might not  
387 be representative for wild specimens (Gering et al., 2019). Observations stemming from  
388 introduced or experimental populations may violate assumptions as well. Observations can be  
389 collected from different subsets of the population (*e.g.* adult vs juvenile, healthy vs diseased),  
390 at different times of year (*e.g.* breeding season vs overwintering), in different contexts (*e.g.*,  
391 experimental temperature treatments), and using other protocols. It is essential to exclude  
392 unsuitable observations, usually by making use of the associated metadata.

393

394 ***Derive traits from raw data:*** Research questions may concern composite or derived traits,  
395 such as ‘hand-wing index’ (the ratio between wing length and body length in birds). Where  
396 possible it is advisable to calculate derived traits directly from the raw data to avoid bias and  
397 allow for new calculations. This procedure may not always be possible because of data gaps;  
398 in this case the calculation can be done at a higher level (*e.g.* at the taxonomic level of  
399 interest).

400

401 ***Aggregate trait data:*** Trait data may come at different levels of resolution. A dataset may  
402 include multiple measurements per individual, per population, or species, or even higher  
403 taxonomic levels. Such structures may imply first aggregating (*e.g.* to calculate average trait  
404 values) within individuals, then populations, then species derived from a particular data

405 source, and then across data sources if the species is represented in several of these  
406 (Schneider et al., 2019). The way trait values were aggregated has to be precisely described,  
407 in particular when data transformation is involved. For example, when it is desired to express  
408 leaf area on a log scale, it might make a big difference to take the log before or after  
409 aggregating the data. Importantly, if multiple successive steps of aggregation are necessary,  
410 there is the need to properly measure the uncertainty of the final trait values and assess the  
411 effect of aggregation on the results and conclusions, *e.g.*, by sensitivity analyses with  
412 different aggregated datasets (Kunz et al., 2022).

413

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

423

424 ***Work with relative errors:*** Units are important when we deal with approximations,  
425 uncertainties, and errors (Langtangen & Pedersen, 2016). As an example, a trait measurement  
426 where the length scale is typically measured in mm and has an approximation of 12.5 m to  
427 the exact value 12.52 m with an error of 0.02 m. Switching units to mm leads to an error of  
428 200 mm. A study working in mm would report  $2 \times 10^2$  as the error, while a study working in  
429 m would report 0.02 as the error. As a result, knowing the original measurement units is

430 essential and the downstream use of the unitless relative error is recommended (Langtangen  
431 & Pedersen, 2016).

## 432 Rule 7: Know the limitations

433 Follow the latest developments for best practices in trait data analyses. As the downstream  
434 part of data analysis is directly linked to the research question, generalisation of analytical  
435 methods is rarely possible. Given the diversity of research questions, the analytical steps can  
436 thus broadly diverge. However, some best practices can help to avoid common mistakes  
437 made with trait data due to their nature.

438

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

445

446 ***Be careful with categorical traits:*** By having fewer possible values, categorical traits might  
447 influence the outcome of statistical analyses. In particular, categorical traits might have  
448 disproportionate effects when aggregated with continuous traits in a common metric (*e.g.*  
449 functional diversity). For instance, when computing Gower distances of traits between  
450 species pairs, a categorical trait treated as a binary trait (0/1) for each possible categorical value  
451 will only result in distances of 0 or 1. In contrast, for a continuous trait, only the species pair  
452 with the highest/lowest trait values will have a value of 1. When averaging Gower distances  
453 of these two traits, the categorical one will have more influence than the continuous one

454 (example from Gelman, 2008). The joint use of continuous and categorical traits thus needs  
455 particular scaling (Gelman, 2008). In general, it is important to be aware of this issue and to  
456 account for it with existing proposed methods (see de Bello et al., 2021).

457

458 ***Do not confuse trait signals for richness and abundance signals:*** Metrics aggregating traits  
459 at the community level (e.g. functional diversity or community-weighted means - CWM), are  
460 influenced by the richness, the abundance of species and the overall species composition of  
461 the community. In general, choosing metrics unrelated to abundance (e.g. unweighted means)  
462 or null models (Hawkins et al., 2017) is necessary to separate species abundance,  
463 composition or richness signals from trait information.

464

465 ***Handle correlations with care:*** Traits are often correlated causing issues with statistical  
466 analyses (e.g. collinearity in linear models when traits are explanatory variables). Often, these  
467 correlations are due to biological constraints (e.g. allometries), or “strategies” (Díaz et al.,  
468 2016). In some cases, it is possible to use multivariate analyses (e.g. principal component  
469 analysis) to reduce trait space dimensionality by use of axes in further analyses. Keeping the  
470 original traits is advised when possible, as the biological meaning can be lost when using  
471 multivariate principal component axes. On another level, since the start of trait analyses at the  
472 community level, numerous metrics have been proposed to characterise functional diversity,  
473 analogous to species diversity. When choosing a metric for a specific analysis, it is essential  
474 to be aware that several of these functional diversity metrics are highly correlated among  
475 themselves and with species richness (see de Bello et al., 2021).

476

477 ***Consider correction for phylogenetic relatedness:*** When analysing data from multiple  
478 species in trait-trait correlations or using traits as responses and depending on whether the

479 focus of the question is ecological or evolutionary, it may become necessary to account for  
480 the fact that species are not independent units (Pillar et al., 2021). The whole field of  
481 comparative analyses tackles this issue and proposes tools to account for phylogenetic  
482 relatedness in trait analyses (*e.g.* see Garamszegi, 2014), though care should be taken to  
483 justify the use of such analytical corrections relative to the aims of the research question  
484 (Freckleton, 2000; Westoby et al., 1995).

485

486 ***Account for variability and uncertainty:*** Very often, intra-specific data are aggregated at the  
487 species level to obtain one trait value per species. All information on variability and  
488 measurement uncertainty is then lost. When information on variability is available and  
489 reasonable in the scope of the study, it is possible to include it, *e.g.*, by weighting species-  
490 level measures in functional diversity metrics (de Bello et al., 2021) or by explicitly including  
491 it when inferring trait evolution across lineages (Kostikova et al., 2016; Purschke et al.,  
492 2017). This can be an issue, especially if variability is phylogenetically structured  
493 (Garamszegi, 2014; Paterno et al., 2018).

## 494 Rule 8: Publish trait data

495 Openly publish trait data to facilitate answering yet unknown questions beyond their original  
496 study, lay the groundwork for understanding ecological processes beyond clearcut niches  
497 (Elton, 1927; Schneider et al., 2019) and democratise access to valuable trait datasets  
498 (Soranno et al., 2015). Each data point of trait measurements has a considerable value for the  
499 scientific community and future generations working on trait-related research questions.

500

501 ***Consider the stakeholders:*** As our scholarly processes evolve to find better, access, integrate,  
502 and reuse scientific data, we face the communal task of treating trait datasets as first-class

503 research citizens. However, doing so is not easy as it involves different stakeholders:  
504 publishers have to make their publications open and FAIR (Wilkinson et al., 2016), scientists  
505 have to improve their skills to publish and reuse datasets, and funding agencies have to find  
506 ways to reward exemplary projects. A welcome development is that many publishers now  
507 consider trait data papers (e.g. Falster et al., 2021; Guerrero-Ramírez et al., 2021; Tobias et  
508 al., 2022; Vandvik et al., 2020), which allow for a detailed methodological and context  
509 description, open access, and at the same time, accreditation of trait data collectors by  
510 citations.

511

512 ***Accept the additional responsibility:*** Erroneous data might not only bias a current project but  
513 also the future works of others. There are currently no common established practices on how  
514 peer review is also extended to trait data. A way to ensure that a dataset conforms to  
515 community standards is to submit it to an established curated database (e.g. TRY (Kattge et  
516 al., 2020) for plant traits; Coral Traits (Madin et al., 2016)).

517

518 ***Aim for redundancy:*** Public trait data suffer from the same generic issues as other data, e.g.  
519 hardware failures, linkrot (URLs not fully reliable) or content-drift (content changes, but  
520 URLs do not) (Koehler, 1999). To mitigate such issues and reliably preserve data in the long  
521 term, data can be submitted to multiple repositories, e.g. beside trait databases, also in general  
522 storage platforms such as FigShare (<https://figshare.com>) or Zenodo (<https://zenodo.org>).

523 This procedure however requires systematic methods to track changes and separately citable  
524 versions e.g. by unique DOIs.

525

526 ***Make data accessible for machines and humans:*** Trait data often carry large numbers of  
527 references and republished original data. Tracing both call for systematic and automated

528 methods (Elliott et al., 2020) that rely on machine-readable data. This issue becomes  
529 particularly relevant when reporting already published data, as future studies might run the  
530 risk of using the same trait from independent sources, thus resulting in pseudo-replication of  
531 measurements. Also, in order to facilitate reuse, machine-readable, non-proprietary data  
532 formats should be preferred (*i.e.* plain csv over excel or pdf).

533

534 **Register trait data:** Independent of the choice of actual data deposition, it is of great  
535 importance that datasets are also registered in a trait data registry (*e.g.* <https://opentraits.org>)  
536 to allow fellow scientists to easily find the data.

## 537 Rule 9: Review data and code like the research itself

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

- 543 1. First, are the traits themselves appropriate for the question being asked? It should be  
544 considered how these traits have been used in the past and how they fit into biological  
545 theory. Are they being contextualised appropriately, and are they fit for the purpose to  
546 which they are being used?
- 547 2. How were the data collected? Does the protocol conform to current standards, bearing  
548 in mind that the purpose of many papers is to improve standards and so they may not?  
549 Is the collection of new data well justified? Are units and metadata properly  
550 provided?

- 551 3. How were the data processed? Consider not just quality assurance and quality control  
552 but also how the traits were generally processed into a format that can be analysed.  
553 Ensure that relevant code (ranging from simple cleaning scripts to full-featured  
554 analysis pipelines and models) is openly available, functional, and conforms to  
555 community standards. Guidelines for source code review like those by the rOpenSci  
556 community (<https://ropensci.org/software-review>) or The Journal of Open Source  
557 Software ([https://joss.readthedocs.io/en/latest/reviewer\\_guidelines.html](https://joss.readthedocs.io/en/latest/reviewer_guidelines.html)) can be  
558 helpful.
- 559 4. Ensure that a distinction is made between the ‘raw’ data collected, and the ‘clean’  
560 final product used for analysis, and that both forms of data are released.
- 561 5. Ensure that all data sources are appropriately cited, and the provenance of data is  
562 explained.

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

## 568 Rule 10: Propagate the good practices

569 Inform the community and the next generation of ecologists about the issues discussed here  
570 and in other resources (*e.g.* de Bello et al., 2021). As we saw above, there are several issues  
571 to be aware of when collecting, handling, analysing, and publishing trait data, *i.e.* the life  
572 cycle of trait data. Some of them might be straightforward; others require more technical  
573 knowledge or extensive reading of existing resources. In many cases, good procedures are not  
574 applied simply because ecological or evolutionary scientists are unaware that they exist, *e.g.*

575 when it comes to trait quality control or using a standard structure. Educating can be done  
576 from the small scale of an informal conversation with a colleague, to teaching a large  
577 undergraduate class, up to participating in the collective creation of open access materials in  
578 several languages, accessible from any part of the world. Integrating trait-data-specific  
579 sections into ecology textbooks and modules in ecology courses could become a standard  
580 practice, which will undoubtedly be made easier in the future by disseminating open-access  
581 material by the trait scientific community. You are welcome to use these ten rules as a starter  
582 when teaching your students, colleagues and friends. For diving deeper into each single topic,  
583 we encourage you to also check more comprehensive resources such as the Handbook of  
584 Trait-Based Ecology (de Bello et al., 2021) or the activities of the Open Traits Network  
585 (<https://opentraits.org>) (Gallagher et al., 2020).

586

587 ***Train students:*** Courses specific to trait-based research are often lacking at both the  
588 undergraduate and graduate levels. Where courses or modules are taught, the focus may be  
589 limited to a subset of the trait data life cycle (*e.g.* Collection and Analysis; Fig. 1), leaving  
590 students lacking critical skills (Feng et al., 2020). The use of Open Educational Resources,  
591 including those built using incubators (Ryder et al., 2020), is one promising method for  
592 making the implementation of such courses and modules easier. In particular, authentic  
593 teaching experiences provide a number of benefits over traditional lectures or “cook-book”  
594 experiments (Brownell et al., 2012), and seem well suited to trait-based ecology given that  
595 many traits can be collected easily and inexpensively, and that there are a host of tools  
596 available (see *e.g.* de Bello et al., 2021). One example of such authentic teaching experiences,  
597 the TraitTrain plant functional trait courses (<https://plantfunctionaltraitcourses.w.uib.no/>),  
598 has provided training across the entire trait data life cycle to hundreds of participants and has

599 created scientific (Henn et al., 2018), data (Vandvik et al., 2020), methodological (Maitner et  
600 al., 2021), and pedagogical (Geange et al., 2021) publications.

601

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

608

609 ***Train the world:*** There is an urgent need for more comprehensive trait data across the globe  
610 and the tree of life (Feng et al., 2022) and thus increasing access to training globally is  
611 critically important. Open access publications, tools, data, and educational resources provide  
612 ways to help lower the barriers to participation (Evans & Reimer, 2009). Further, due to the  
613 relative ease, low cost, and tangible nature of many functional traits, they are well-suited to  
614 inclusion in elementary education and citizen science (*e.g.* Isaac & Martin, 2019; Schiller et  
615 al., 2021).

## 616 Conclusions

617 This ten rules document is particularly motivated by its last rule and has the intention of  
618 bringing awareness of the different facets of a trait's life cycle to the community. Most  
619 scientists working with traits usually are experts on only one or few aspects of the life cycle  
620 of trait data, which are numerous, ranging from biological theory and field research to  
621 computer sciences as well as publishing and review ethics. When it comes to working with  
622 trait data, we gain particularly as an interdisciplinary community of field biologists, synthesis

623 ecologists, computer scientists and database managers. We hope these basic guidelines can be  
624 useful as a starter for active communication in disseminating such integrative knowledge and  
625 how to make trait data future-proof.

## 626 Conflict of Interest statement

627 All authors declare to have no conflicts of interest.

## 628 Acknowledgements

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

## 638 Author contributions

639 AK and MA conceived the idea. CP, JP, MA and AK led the writing of the manuscript. AK,  
640 MA, BE, HB, SD, DF, REO, RG, JH, JK, JM, BM, WP, CP, FS and JP contributed in writing  
641 the first draft during a workshop session and thereafter. All authors contributed critically to  
642 the drafts and gave final approval for publication.

## 643 Data availability

644 Not applicable.

## 645 References

- 646 Ankenbrand, M. J., Hohlfeld, S. C. Y., Weber, L., Förster, F., & Keller, A. (2018). Functional  
647 exploration of natural networks and ecological communities. *Methods in Ecology and*  
648 *Evolution*, 0(0). <https://doi.org/10.1111/2041-210X.13060>
- 649 Bongers, F. J., Schmid, B., Bruelheide, H., Bongers, F., Li, S., von Oheimb, G., Li, Y.,  
650 Cheng, A., Ma, K., & Liu, X. (2021). Functional diversity effects on productivity  
651 increase with age in a forest biodiversity experiment. *Nature Ecology & Evolution*,  
652 5(12), 1594–1603. <https://doi.org/10.1038/s41559-021-01564-3>
- 653 Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J. A., Mozzherin, D., Rees, T., Matasci, N.,  
654 Narro, M. L., Piel, W. H., Mckay, S. J., Lowry, S., Freeland, C., Peet, R. K., &  
655 Enquist, B. J. (2013). The taxonomic name resolution service: An online tool for  
656 automated standardization of plant names. *BMC Bioinformatics*, 14(1), 16.  
657 <https://doi.org/10.1186/1471-2105-14-16>
- 658 Boyle, B. L., Maitner, B. S., Barbosa, G. G. C., Sajja, R. K., Feng, X., Merow, C., Newman,  
659 E. A., Park, D. S., Roehrdanz, P. R., & Enquist, B. J. (2022). *Geographic Name*  
660 *Resolution Service: A tool for the standardization and indexing of world political*  
661 *division names, with applications to species distribution modeling* (p.  
662 2022.04.25.489424). bioRxiv. <https://doi.org/10.1101/2022.04.25.489424>
- 663 Breitschwerdt, E., Jandt, U., & Bruelheide, H. (2018). Using co-occurrence information and  
664 trait composition to understand individual plant performance in grassland  
665 communities. *Scientific Reports*, 8(1), 9076. [https://doi.org/10.1038/s41598-018-](https://doi.org/10.1038/s41598-018-27017-9)  
666 27017-9
- 667 Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). *Undergraduate Biology*  
668 *Lab Courses: Comparing the Impact of Traditionally Based “Cookbook” and*  
669 *Authentic Research-Based Courses on Student Lab Experiences*. 10.

670 Calder, W. A. (1982). A Proposal for the Standardization of Units and Symbols in Ecology.  
671 *Bulletin of the Ecological Society of America*, 63(1), 7–10.

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

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

679 Dawson, S. K., Boddy, L., Halbwachs, H., Bässler, C., Andrew, C., Crowther, T. W.,  
680 Heilmann-Clausen, J., Nordén, J., Ovaskainen, O., & Jönsson, M. (2019). Handbook  
681 for the measurement of macrofungal functional traits: A start with basidiomycete  
682 wood fungi. *Functional Ecology*, 33(3), 372–387. [https://doi.org/10.1111/1365-](https://doi.org/10.1111/1365-2435.13239)  
683 [2435.13239](https://doi.org/10.1111/1365-2435.13239)

684 Dawson, S. K., Carmona, C. P., González-Suárez, M., Jönsson, M., Chichorro, F., Mallen-  
685 Cooper, M., Melero, Y., Moor, H., Simaika, J. P., & Duthie, A. B. (2021). The traits  
686 of “trait ecologists”: An analysis of the use of trait and functional trait terminology.  
687 *Ecology and Evolution*, 11(23), 16434–16445.

688 de Bello, F., Carmona, C. P., Dias, A. T. C., Götzenberger, L., Moretti, M., & Berg, M. P.  
689 (2021). *Handbook of Trait-Based Ecology: From Theory to R Tools*. Cambridge  
690 University Press. <https://doi.org/10.1017/9781108628426>

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

694 Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer,  
695 M., Wirth, C., Colin Prentice, I., & others. (2016). The global spectrum of plant form  
696 and function. *Nature*, 529(7585), 167–171.

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

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

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

705 Evans, J. A., & Reimer, J. (2009). Open Access and Global Participation in Science. *Science*,  
706 323(5917), 1025–1025. <https://doi.org/10.1126/science.1154562>

707 Falster, D., Gallagher, R., Wenk, E. H., Wright, I. J., Indiarto, D., Andrew, S. C., Baxter, C.,  
708 Lawson, J., Allen, S., Fuchs, A., Monro, A., Kar, F., Adams, M. A., Ahrens, C. W.,  
709 Alfonzetti, M., Angevin, T., Apgaua, D. M. G., Arndt, S., Atkin, O. K., ... Ziemińska,  
710 K. (2021). AusTraits, a curated plant trait database for the Australian flora. *Scientific  
711 Data*, 8(1), 254. <https://doi.org/10.1038/s41597-021-01006-6>

712 Fegraus, E. H., Andelman, S., Jones, M. B., & Schildhauer, M. (2005). Maximizing the Value  
713 of Ecological Data with Structured Metadata: An Introduction to Ecological Metadata  
714 Language (EML) and Principles for Metadata Creation. *The Bulletin of the Ecological  
715 Society of America*, 86(3), 158–168. [https://doi.org/10.1890/0012-  
716 9623\(2005\)86\[158:MTVOED\]2.0.CO;2](https://doi.org/10.1890/0012-9623(2005)86[158:MTVOED]2.0.CO;2)

717 Feng, X., Enquist, B. J., Park, D. S., Boyle, B., Breshears, D. D., Gallagher, R. V., Lien, A.,  
718 Newman, E. A., Burger, J. R., Maitner, B. S., Merow, C., Li, Y., Huynh, K. M., Ernst,

719 K., Baldwin, E., Foden, W., Hannah, L., Jørgensen, P. M., Kraft, N. J. B., ... López-  
720 Hoffman, L. (2022). A review of the heterogeneous landscape of biodiversity  
721 databases: Opportunities and challenges for a synthesized biodiversity knowledge  
722 base. *Global Ecology and Biogeography*, 31(7), 1242–1260.  
723 <https://doi.org/10.1111/geb.13497>

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

727 Freckleton, R. P. (2000). Phylogenetic Tests of Ecological and Evolutionary Hypotheses:  
728 Checking for Phylogenetic Independence. *Functional Ecology*, 14(1), 129–134.

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

732 Gallagher, R. V., Falster, D. S., Maitner, B. S., Salguero-Gómez, R., Vandvik, V., Pearse, W.  
733 D., Schneider, F. D., Kattge, J., Poelen, J. H., Madin, J. S., Ankenbrand, M. J.,  
734 Penone, C., Feng, X., Adams, V. M., Alroy, J., Andrew, S. C., Balk, M. A., Bland, L.  
735 M., Boyle, B. L., ... Enquist, B. J. (2020). Open Science principles for accelerating  
736 trait-based science across the Tree of Life. *Nature Ecology & Evolution*, 4(3), 294–  
737 303. <https://doi.org/10.1038/s41559-020-1109-6>

738 Gama, J. (2014). *NISTunits: Fundamental physical constants and unit conversions from*  
739 *NIST*.

740 Garamszegi, L. Z. (Ed.). (2014). *Modern Phylogenetic Comparative Methods and Their*  
741 *Application in Evolutionary Biology: Concepts and Practice*. Springer Berlin  
742 Heidelberg. <https://doi.org/10.1007/978-3-662-43550-2>

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

745 Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougenot, I., Kühn, I., Laporte, B.,  
746 Amiaud, B., Ahrestani, F. S., Bönisch, G., Bunker, D. E., Cornelissen, J. H. C., Díaz,  
747 S., Enquist, B. J., Gachet, S., Jaureguiberry, P., Kleyer, M., Lavorel, S., Maicher, L.,  
748 ... Klotz, S. (2017). Towards a thesaurus of plant characteristics: An ecological  
749 contribution. *Journal of Ecology*, *105*(2), 298–309. [https://doi.org/10.1111/1365-](https://doi.org/10.1111/1365-2745.12698)  
750 [2745.12698](https://doi.org/10.1111/1365-2745.12698)

751 Geange, S. R., von Oppen, J., Strydom, T., Boakye, M., Gauthier, T.-L. J., Gya, R.,  
752 Halbritter, A. H., Jessup, L. H., Middleton, S. L., Navarro, J., Pierfederici, M. E.,  
753 Chacón-Labela, J., Cotner, S., Farfan-Rios, W., Maitner, B. S., Michaletz, S. T.,  
754 Telford, R. J., Enquist, B. J., & Vandvik, V. (2021). Next-generation field courses:  
755 Integrating Open Science and online learning. *Ecology and Evolution*, *11*(8), 3577–  
756 [3587](https://doi.org/10.1002/ece3.7009). <https://doi.org/10.1002/ece3.7009>

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

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

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

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

766 Grenié, M., Berti, E., Carvajal-Quintero, J., Dädlow, G. M. L., Sagouis, A., & Winter, M.  
767 (2022). Harmonizing taxon names in biodiversity data: A review of tools, databases

768 and best practices. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041->  
769 210X.13802

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

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

780 Henn, J. J., Buzzard, V., Enquist, B. J., Halbritter, A. H., Klanderud, K., Maitner, B. S.,  
781 Michaletz, S. T., Pötsch, C., Seltzer, L., Telford, R. J., Yang, Y., Zhang, L., &  
782 Vandvik, V. (2018). Intraspecific Trait Variation and Phenotypic Plasticity Mediate  
783 Alpine Plant Species Response to Climate Change. *Frontiers in Plant Science*, 9.  
784 <https://www.frontiersin.org/article/10.3389/fpls.2018.01548>

785 Herberstein, M. E., McLean, D. J., Lowe, E., Wolff, J. O., Khan, M. K., Smith, K., Allen, A.  
786 P., Bulbert, M., Buzatto, B. A., Eldridge, M. D. B., Falster, D., Fernandez Winzer, L.,  
787 Griffith, S. C., Madin, J. S., Narendra, A., Westoby, M., Whiting, M. J., Wright, I. J.,  
788 & Carthey, A. J. R. (2022). AnimalTraits—A curated animal trait database for body  
789 mass, metabolic rate and brain size. *Scientific Data*, 9(1), 265.  
790 <https://doi.org/10.1038/s41597-022-01364-9>

791 Hulme, P. E., Pyšek, P., Jarošík, V., Pergl, J., Schaffner, U., & Vilà, M. (2013). Bias and  
792 error in understanding plant invasion impacts. *Trends in Ecology & Evolution*, 28(4),  
793 212–218. <https://doi.org/10.1016/j.tree.2012.10.010>

794 Isaac, M. E., & Martin, A. R. (2019). Accumulating crop functional trait data with citizen  
795 science. *Scientific Reports*, 9(1), 15715. <https://doi.org/10.1038/s41598-019-51927-x>

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

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

804 Kattge, J., Ogle, K., Bönisch, G., Díaz, S., Lavorel, S., Madin, J., Nadrowski, K., Nöllert, S.,  
805 Sartor, K., & Wirth, C. (2011). A generic structure for plant trait databases. *Methods*  
806 *in Ecology and Evolution*, 2(2), 202–213. <https://doi.org/10.1111/j.2041->  
807 [210X.2010.00067.x](https://doi.org/10.1111/j.2041-210X.2010.00067.x)

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

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

814 Koricheva, J., Gurevitch, J., & Mengersen, K. (2013). *Handbook of Meta-analysis in Ecology*  
815 *and Evolution*. Princeton University Press.

816 Kostikova, A., Silvestro, D., Pearman, P. B., & Salamin, N. (2016). Bridging Inter- and  
817 Intraspecific Trait Evolution with a Hierarchical Bayesian Approach. *Systematic*  
818 *Biology*, 65(3), 417–431. <https://doi.org/10.1093/sysbio/syw010>

819 Kunz, S., Kefford, B. J., Schmidt-Kloiber, A., Matthaei, C. D., Usseglio-Polatera, P., Graf,  
820 W., Poff, N. L., Metzeling, L., Twardochleb, L., Hawkins, C. P., & Schäfer, R. B.  
821 (2022). Tackling inconsistencies among freshwater invertebrate trait databases:  
822 Harmonising across continents and aggregating taxonomic resolution. *Freshwater*  
823 *Biology*, 67(2), 275–291. <https://doi.org/10.1111/fwb.13840>

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

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

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

833 Madin, J., Bowers, S., Schildhauer, M., & Jones, M. (2008). Advancing Ecological Research  
834 with Ontologies. *Trends in Ecology & Evolution*, 23, 159–168.  
835 <https://doi.org/10.1016/j.tree.2007.11.007>

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

839 Madin, J. S., Anderson, K. D., Andreasen, M. H., Bridge, T. C. L., Cairns, S. D., Connolly, S.  
840 R., Darling, E. S., Diaz, M., Falster, D. S., Franklin, E. C., Gates, R. D., Harmer, A.

841 M. T., Hoogenboom, M. O., Huang, D., Keith, S. A., Kosnik, M. A., Kuo, C.-Y.,  
842 Lough, J. M., Lovelock, C. E., ... Baird, A. H. (2016). The Coral Trait Database, a  
843 curated database of trait information for coral species from the global oceans.  
844 *Scientific Data*, 3(1), 160017. <https://doi.org/10.1038/sdata.2016.17>

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

850 Marine Species Traits editorial board. (n.d.). *Marine Species Traits* (Accessed at 20.05.2022).  
851 Retrieved May 20, 2022, from <http://www.marinespecies.org/traits>

852 McGraw, J. B., & Caswell, H. (1996). Estimation of Individual Fitness from Life-History  
853 Data. *The American Naturalist*, 147(1), 47–64. <https://doi.org/10.1086/285839>

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

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

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

864 Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck,  
865 S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E.,

866 Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ...  
867 Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–  
868 1425. <https://doi.org/10.1126/science.aab2374>

869 Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017).  
870 AmphibiO, a global database for amphibian ecological traits. *Scientific Data*, 4(1),  
871 170123. <https://doi.org/10.1038/sdata.2017.123>

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

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

878 Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M.,  
879 Young, B. E., Graham, C. H., & Costa, G. C. (2014). Imputation of missing data in  
880 life-history trait datasets: Which approach performs the best? *Methods in Ecology and*  
881 *Evolution*, 5(9), 961–970. <https://doi.org/10.1111/2041-210X.12232>

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

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

888 Roberts, L. W. (2004). How to Review a Manuscript: A “Down-to-Earth” Approach.  
889 *Academic Psychiatry*, 28(2), 81–87. <https://doi.org/10.1176/appi.ap.28.2.81>

890 Ryder, E. F., Morgan, W. R., Sierk, M., Donovan, S. S., Robertson, S. D., Orndorf, H. C.,  
891 Rosenwald, A. G., Triplett, E. W., Dinsdale, E., Pauley, M. A., & Tapprich, W. E.  
892 (2020). Incubators: Building community networks and developing open educational  
893 resources to integrate bioinformatics into life science education. *Biochemistry and*  
894 *Molecular Biology Education*, 48(4), 381–390. <https://doi.org/10.1002/bmb.21387>

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

898 Schiller, C., Schmidlein, S., Boonman, C., Moreno-Martínez, A., & Kattenborn, T. (2021).  
899 Deep learning and citizen science enable automated plant trait predictions from  
900 photographs. *Scientific Reports*, 11(1), 16395. [https://doi.org/10.1038/s41598-021-](https://doi.org/10.1038/s41598-021-95616-0)  
901 [95616-0](https://doi.org/10.1038/s41598-021-95616-0)

902 Schneider, F. D., Fichtmueller, D., Gossner, M. M., Güntsch, A., Jochum, M., König-Ries,  
903 B., Le Provost, G., Manning, P., Ostrowski, A., Penone, C., & Simons, N. K. (2019).  
904 Towards an ecological trait-data standard. *Methods in Ecology and Evolution*, 10(12),  
905 2006–2019. <https://doi.org/10.1111/2041-210X.13288>

906 Soranno, P. A., Cheruvilil, K. S., Elliott, K. C., & Montgomery, G. M. (2015). It's Good to  
907 Share: Why Environmental Scientists' Ethics Are Out of Date. *BioScience*, 65(1), 69–  
908 73. <https://doi.org/10.1093/biosci/biu169>

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

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

913 Sporbert, M., Welk, E., Seidler, G., Jandt, U., Ačić, S., Biurrun, I., Campos, J. A., Čarni, A.,  
914 Cerabolini, B. E. L., Chytrý, M., Čušterevska, R., Dengler, J., De Sanctis, M., Dziuba,

915 T., Fagúndez, J., Field, R., Golub, V., He, T., Jansen, F., ... Bruelheide, H. (2021).  
916 Different sets of traits explain abundance and distribution patterns of European plants  
917 at different spatial scales. *Journal of Vegetation Science*, 32(2), e13016.  
918 <https://doi.org/10.1111/jvs.13016>

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

925 Thessen, A., Preciado, J., Jain, P., Martin, J., Palmer, M., & Bhat, R. (2018). Automated Trait  
926 Extraction using ClearEarth, a Natural Language Processing System for Text Mining  
927 in Natural Sciences. *Biodiversity Information Science and Standards*, 2, e26080.  
928 <https://doi.org/10.3897/biss.2.26080>

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

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

938 Vandvik, V., Halbritter, A. H., Yang, Y., He, H., Zhang, L., Brummer, A. B., Klanderud, K.,  
939 Maitner, B. S., Michaletz, S. T., Sun, X., Telford, R. J., Wang, G., Althuizen, I. H. J.,

940 Henn, J. J., Garcia, W. F. E., Gya, R., Jaroszynska, F., Joyce, B. L., Lehman, R., ...  
941 Enquist, B. J. (2020). Plant traits and vegetation data from climate warming  
942 experiments along an 1100 m elevation gradient in Gongga Mountains, China.  
943 *Scientific Data*, 7(1), 189. <https://doi.org/10.1038/s41597-020-0529-0>

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

946 Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E.  
947 (2007). Let the concept of trait be functional! *Oikos*, 116(5), 882–892.  
948 <https://doi.org/10.1111/j.0030-1299.2007.15559.x>

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

955 Webb, C. T., Hoeting, J. A., Ames, G. M., Pyne, M. I., & LeRoy Poff, N. (2010). A  
956 structured and dynamic framework to advance traits-based theory and prediction in  
957 ecology. *Ecology Letters*, 13(3), 267–283. [https://doi.org/10.1111/j.1461-](https://doi.org/10.1111/j.1461-0248.2010.01444.x)  
958 [0248.2010.01444.x](https://doi.org/10.1111/j.1461-0248.2010.01444.x)

959 Weiher, E., van der Werf, A., Thompson, K., Roderick, M., Garnier, E., & Eriksson, O.  
960 (1999). Challenging Theophrastus: A common core list of plant traits for functional  
961 ecology. *Journal of Vegetation Science*, 10(5), 609–620.  
962 <https://doi.org/10.2307/3237076>

963 Westoby, M., Falster, D. S., & Schrader, J. (2021). Motivating data contributions via a  
964 distinct career currency. *Proceedings of the Royal Society B: Biological Sciences*,  
965 288(1946), 20202830. <https://doi.org/10.1098/rspb.2020.2830>

966 Westoby, M., Leishman, M. R., & Lord, J. M. (1995). On Misinterpreting the 'Phylogenetic  
967 Correction'. *Journal of Ecology*, 83(3), 531–534. <https://doi.org/10.2307/2261605>

968 Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T.,  
969 & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed  
970 Biodiversity Data Standard. *PLOS ONE*, 7(1), e29715.  
971 <https://doi.org/10.1371/journal.pone.0029715>

972 Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A.,  
973 Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J.,  
974 Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T.,  
975 Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data  
976 management and stewardship. *Scientific Data*, 3(1), 160018.  
977 <https://doi.org/10.1038/sdata.2016.18>

978 Wright, J. P., Ames, G. M., & Mitchell, R. M. (2016). The more things change, the more they  
979 stay the same? When is trait variability important for stability of ecosystem function  
980 in a changing environment. *Philosophical Transactions of the Royal Society B:*  
981 *Biological Sciences*, 371(1694), 20150272. <https://doi.org/10.1098/rstb.2015.0272>  
982