The Open Traits Network: Using Open Science principles to accelerate trait-based science across the Tree of Life

Gallagher, R. V.[1], Falster, D. S.[2], Maitner, B. S.[3], Salguero-Gómez , R.[4,5,6], Vandvik, V.[7,8], Pearse, W. D.[9], Schneider, F. D.[10], Kattge, J.[11,12], Alroy, J.[1], Ankenbrand, M. J.[13,14], Andrew, S. C.[15], Balk, M.[16], Bland, L. M.[17], Boyle, B. L.[3], Bravo-Avila, C. H.[18,19], Brennan, I.[20], Carthey, A. J. R.[21], Catullo, R.[20], Cavazos, B. R.[22], Chown, S.[23], Fadrique, B.[24], Feng, X.[3], Gibb, H.[25], Halbritter, A. H.[7,8], Hammock, J.[26], Hogan, J. A.[27], Holewa, H.[15], Hope, M.[15], Iversen, C. M.[28], Jochum, M.[29,30,31], Kearney, M.[32], Keller, A.[13], Mabee, P.[33], Madin, J. S.[34], Manning, P.[35], McCormack, L.[36], Michaletz, S. T.[37], Park, D. S.[38], Penone, C.[29], Perez, T. M.[24,39], Pineda-Munoz, S.[40], Poelen, J.[41], Ray, C. A.[42], Rossetto, M.[43], Sauquet, H.[43,44], Sparrow, B.[45], Spasojevic, M. J.[46], Telford, R. J.[7,8], Tobias, J. A.[47], Violle, C.[48], Walls, R.[49], Weiss, K. C. B.[42], Westoby, M.[1], Wright, I. J.[1], Enquist, B. J.[3,50]

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia , [2]Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia, [3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA, [4]Department of Zoology, Oxford University, Oxford, UK, [5]Centre for Biodiversity and Conservation Science, University of Queensland, Queensland, Australia, [6]Evolutionary Demography Laboratory, Max Plank Institute for Demographic Research, Rostock, Germany, [7]Department of Biological Sciences, University of Bergen, Bergen, Norway, [8]Bjerknes Centre for Climate Research, University of Bergen, N-5020 Bergen, Norway, [9]Ecology Center and Department of Biology, Utah State University, Utah, USA, [10]Niederstr. 23 64285 Darmstadt, Germany, [11]Max Planck Institute for Biogeochemistry, Jena, Germany,[12]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany,[13]Department of Bioinformatics, Biocenter, University of Wuerzburg, Wuerzburg, Germany, [14]Center for Computational and Theoretical Biology, Biocenter, University of Wuerzburg, Wuerzburg, Germany,[15]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, Australia, [16]Bio5 Institute, University of Arizona, Tucson, Arizona, USA, [17]School of Life and Environmental Sciences, Centre for Integrative Ecology, Deakin University, Victoria, Australia, [18]Department of Biology, University of Miami, Coral Gables, Florida,

USA, [19]Fairchild Tropical Botanic Garden, Coral Gables, Florida, USA, [20]Research School of Biology, Australian National University, Canberra, Australia, [21]Department of Biological Sciences, Macquarie University, North Ryde, NSW, Australia , [22]Department of BioSciences, Program in Ecology and Evolutionary Biology, Rice University, Houston, Texas, USA, [23]School of Biological Sciences, Monash University, Victoria, Australia, [24]Department of Biology, University of Miami, Florida, USA, [25]Department of Ecology, Environment and Evolution and Centre for Future Landscapes, La Trobe University, Victoria,[26]National Museum of Natural History, Smithsonian Institution, Washington DC, USA, [27]International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Florida, USA,[28]Climate Change Science Institute and Environmental Sciences Division, Oak Ridge National Laboratory, Tennessee, USA, [29]Institute of Plant Sciences, University of Bern, Bern, Switzerland, [30]German Centre for Integrative Biodiversity Research (iDiv) Halle?Jena?Leipzig, Leipzig, Germany, [31]Leipzig University, Institute of Biology, Leipzig, Germany, [32]School of BioSciences, The University of Melbourne, Victoria, Australia,[33]Department of Biology, University of South Dakota, South Dakota, USA, [34]Hawai‘i Institute of Marine Biology, University of Hawai‘i at Manoa, Hawai‘i, USA, [35]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt,, [36]Center for Tree Science, The Morton Arboretum, Illinois, USA, [37]Department of Botany and Biodiversity Research Centre, University of British Columbia, British Columbia, Canada,[38]Department of Organismic and Evolutionary Biology and Harvard University Herbaria, Harvard University, Cambridge, Massachusetts, USA, [39]Fairchild Tropical Botanic Garden, Florida, USA, [40]School of Biological Sciences, School of Earth & Atmospheric Sciences, Georgia Institute of Technology, Georgia, USA, [41]400 Perkins Street Apt 104, Oakland, California, United States, [42]School of Life Sciences, Arizona State University, Tempe, Arizona, USA, [43]National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, NSW, Australia, [44]Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Universite? Paris-Saclay, Orsay, France, [45]TERN / School of Biological Sciences, Faculty of Science, The University of Adelaide, Adelaide, Australia, [46]Department of Evolution, Ecology, and Organismal Biology, University of California Riverside, California, USA, [47]Department of Life Sciences, Imperial College London, Silwood Park, UK, [48]CEFE, CNRS, Univ Montpellier, Université Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France, [49]CyVerse, University of Arizona, Tucson, AZ, USA, [50]Santa Fe Institute, Santa Fe, New Mexico, USA

ABSTRACT

Synthesising trait observations and knowledge across the Tree of Life remains a grand challenge for biodiversity science. Despite the well-recognised importance of traits for addressing ecological and evolutionary questions, trait-based approaches still struggle with several basic data requirements to deliver openly accessible, reproducible, and transparent science. Here, we introduce the Open Traits Network (OTN) – a decentralised alliance of international researchers and institutions focused on collaborative integration and standardisation of the exponentially increasing availability of trait data across all organisms. The OTN embraces the use of Open Science principles in trait research, particularly open data, open source, and open methodology protocols and workflows, to accelerate the synthesis of trait data across the Tree of Life. Increased efforts at all levels – from individual scientists, research networks, scientific societies, funding agencies, to publishers – are necessary to fully exploit the opportunities offered by Open Science in trait research. Democratising access to data, tools and resources will facilitate rapid advances in the biological sciences and our ability to address pressing environmental and societal demands.

INTRODUCTION

Traits, broadly speaking, are attributes or characteristics of organisms. Traits related to function – such as leaf size, body mass, or growth form – are often used to understand how organisms interact with the environment and with other species via key vital rates such as survival, development, and reproduction[1-6]. Trait-based approaches have long been used in systematics and macroevolution to delineate taxa and reconstruct ancestral morphology and function[7-9] and to link candidate genes to phentoypes[10-12]. The broad appeal of the trait concept is its ability to facilitate quantitative comparisons of biological form and function and to mechanistically link organismal responses to abiotic and biotic factors using measurements that are, in principle, relatively easy to capture across large numbers of individuals. For example, appropriately chosen and defined traits can help identify different lineages that share similar life-history strategies for a given environmental regime[13,14]. Thus, documenting and understanding the diversity and composition of traits in ecosystems directly contributes to our understanding of organismal and ecosystem processes, functionality, productivity, and resilience in the face of environmental change[15-20].

Functional traits are important indicators for the socio-economic value of ecosystems and their services. They are increasingly used to model food and energy security, and to inform conservation decision-making[21-25]. Traits are also key to improving modelling of earth systems and their responses to climate change by linking energy and resource fluxes between organisms and the environment[26,27]. In short, trait data can help bridge disciplines within biology, and link biology to the physical sciences and human systems.

In recent decades there has been an acceleration in the collection, compilation, and availability of trait data for a variety of organisms. Substantial trait databases now exist for plants[28-31], reptiles[32,33], invertebrates[31,34-37], fish[38,39], corals[40], birds[31,41,42], amphibians[43], mammals[31,42,44,45], and fungi[31] (see also https://github.com/traitecoevo/fungaltraits) and parallel efforts are no doubt underway for other taxa. Though considerable effort has been made to quantify traits for some key groups (*e.g.*, Fig. 1), substantial work remains. In order to advance knowledge by developing and testing theory in biodiversity science much greater effort is needed to combine and integrate data[46]. There are two central questions that together define the scope and nature of the challenge to trait science. First, how can we most effectively advance the synthesis of trait data within and across disciplines to address questions of global significance (Box 1)? Second, how can we best deliver accurate and understandable biodiversity knowledge to non-academic audiences who should have equitable access to quality data on the traits of Earths' species?

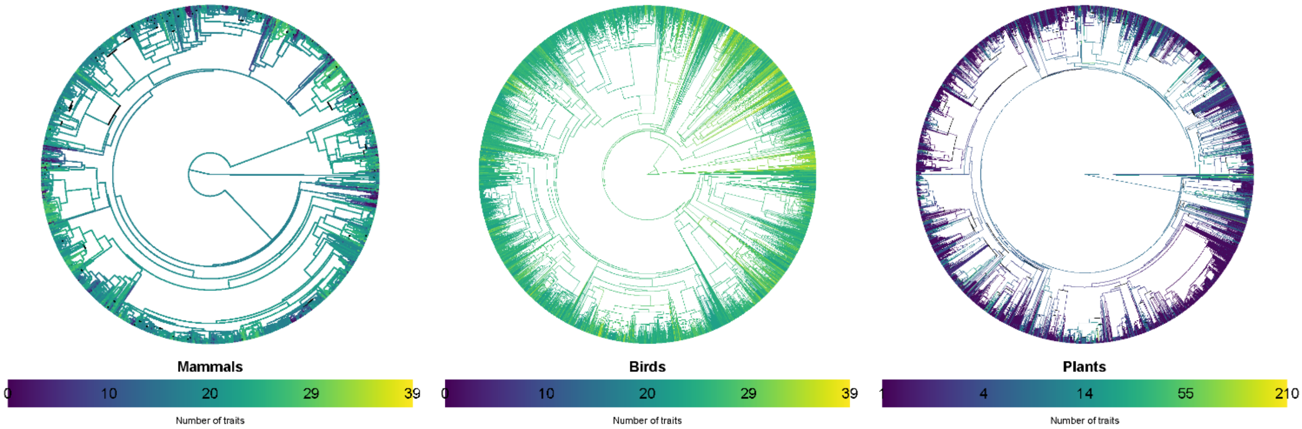

Figure 1. Mammal, bird, and plant phylogenies coloured according to the number of traits for which we have data for each species and lineage. Trait data were downloaded from [33,42,47], the number of traits present across these datasets for each species counted, and then mapped onto single phylogenies from the posteriors of [45,48], and a random subset of plant species within a single phylogeny from[49]. Terminal branches (representing species) and ancestral lineages (using ancestral state reconstruction[50]) were then coloured according to the number of (reconstructed) traits. Note that this is an exploratory analysis, conducted purely to show the variation across taxonomic groups in the amount of available trait data.

Using Open Science principles to accelerate trait-based science

This is an exciting time to advocate for greater coordination and transparency in trait-based research. Indeed, across the sciences, Open Science principles are rapidly emerging and being adopted. Open Science principles (Fig. 2) outline a movement towards making all aspects of the scientific process transparent and accessible to a wide audience[51,52].

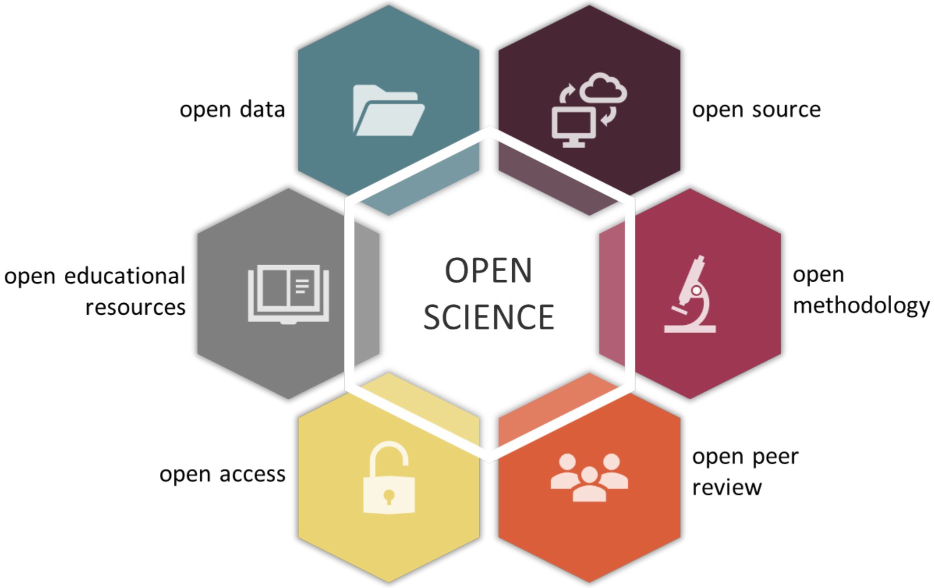

Figure 2. The six core principles of Open Science which guide the Open Traits Network.

In this context, knowledge is considered open if anyone can freely access, use, modify, and share it – subject, at most, to measures that preserve provenance and openness (http://opendefinition.org/). Several pronouncements about Open Science principles have already been made, including the Berlin (https://openaccess.mpg.de/Berlin-Declaration), Bouchout (http://www.bouchoutdeclaration.org/declaration/) and Denton Declarations (https://openaccess.unt.edu/denton-declaration) on open access to science data. Other initiatives champion some open practices such as the Bari Manifesto on interoperability[53] and the FORCE 11 network, which developed the 'Joint Declaration of Data Citation Principles' (https://www.force11.org/datacitationprinciples) and 'FAIR' principles (https://www.force11.org/group/fairgroup/fairprinciples). The FAIR principles address several of the major challenges facing trait-based research, namely making data Findable, Accessible, Interoperable, and Reusable.

While the adoption of Open Science Principles has the potential to rapidly advance global trait synthesis by democratising access to data, tools, and resources, numerous hurdles remain. Trait-based science continues to struggle with realising these principles because of several issues including: 1) a lack of readily available, machine-accessible primary data released under a clear license arrangement; 2) the need for standardised protocols, handbooks or metadata formats for data collection, documentation and management, but see[54-56]; and 3) the complexity of integrating existing legacy data from disparate sources (*e.g.*, specimens, published literature, citizen science initiatives[57,58], large scale digitisation efforts (*e.g.*, Biodiversity Heritage Library) which will have systematic differences in the error rates, validation, context, reproducibility, and objectivity relative to field-collected trait observations.

For many researchers and institutions, Open Science principles are increasingly becoming part of their everyday approach to trait research. Connecting those wishing to transition to an Open Science model – whether individuals, research groups, or institutions – to those already actively adopting these principles will facilitate transfer of skills and knowledge. These connections could be made via model examples, standards, and networks. Trait science has clear potential to rapidly increase its taxonomic, phylogenetic, and spatial scope, by stronger advocacy of Open Science and greater connection between researchers, institutions, publishers, and funding bodies.

Introducing the Open Traits Network

The Open Traits Network (OTN) is a decentralised network accessible to any international researcher, institution, or research infrastructure provider focused on collaboration and standardization in the collection of trait data based on Open Science principles (Fig. 2). The primary goal of the OTN is to increase cross-community exchange of ideas, tools, resources, and data.

Network diagrams can represent the way researchers interact when building interlinked trait sources and databases; they show how individual nodes (researchers, groups, institutions) are connected (Fig. 3).

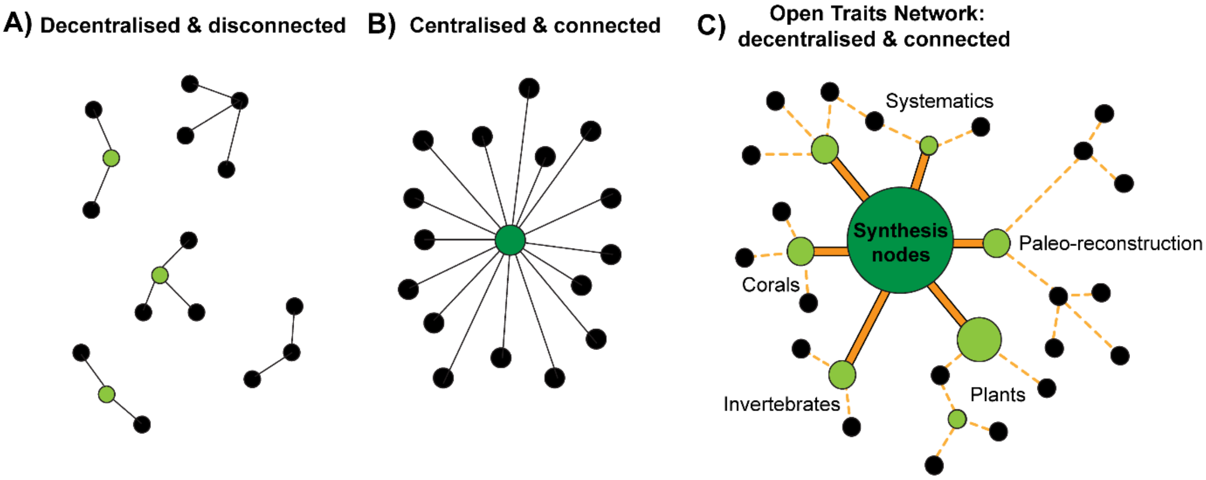

Figure 3. Architectures of three alternative networks in which research groups (nodes) interact in collecting and organising trait data. Black nodes are individuals, groups or institutions pursuing trait projects. Light green nodes are those synthesising data and protocols, where node size is proportional to available resources. Dark green nodes are synthesising nodes, which benefit from the flow of standardised trait data and knowledge. (A) Groups are disconnected and decentralised, risking duplication of effort (often the status quo); (B) Groups are inflexibly linked to a centralised repository, potentially limiting innovation; (C) The Open Traits Network, which is represented by orange connecting lines. Nodes are linked together within biological domains (*e.g.,* plants, corals) and/or disciplines (*e.g.,* systematics, palaeobiology), allowing for more efficient and specialised decisions about trait collection. However, data synthesis across domains or disciplines is then facilitated by joining nodes based on common workflows and data sharing protocols which adhere to the guiding principles of the Open Traits Network. Central synthesising nodes may be, for example, the Smithsonian Encyclopedia of Life https://eol.org/.

In many instances, groups building smaller-scale databases do so in isolation, using their own tools and workflows tailored to their research question (Fig. 3A). These decentralised researchers are often best placed to determine which traits are most useful in their study system and how they should be compiled. However, there is little formal support or interaction across this style of network, so researchers often collect redundant data, use different data structures, and may develop similar tools for data collection, cleaning, and integration in isolation, which can lead to duplication of effort. There are many small, isolated, and heterogeneous data sources of this sort, increasing the disconnect between pools of trait data[59].

For some organisms, centralised hubs already exist which aggregate and standardise trait data from across disparate sources (*e.g.*,[29,40,60-63]) (Fig. 3B). These trait repository hubs have increasingly become the main access point for trait data, but they remain mostly isolated from each other, limiting the sharing of expertise. Other large-scale informatic challenges have also followed the centralised and connected network model (*e.g.,* the Global Biodiversity Information Facility https://www.gbif.org/; GenBank https://www.ncbi.nlm.nih.gov/genbank/) and mandate strict data exchange protocols to facilitate synthesis. Standardized, controlled vocabularies – such as Darwin Core and Humboldt Core – have been essential to the explosive growth of biodiversity data because they facilitate free exchange of information using common data formats[46,58,64]. Further, ontologies necessary to represent traits have been developed (*e.g.,* Uberon – the multispecies anatomy ontology for animals[65], TOP – the Thesaurus of Plant characteristics[66]) that enable the integration with genetic and environmental data through corresponding ontologies (*e.g.,* Gene Ontology[67]; Environmental Ontology[68]). Although the centralised and connected model (Fig. 3B) has clear advantages, it can lack the flexibility to dynamically connect trait data where ontologies and exchange formats do not exist. The result is that established trait networks will remain isolated and disconnected.

The OTN (represented by the orange connections in Fig. 3C) maintains the key advantages of a decentralised network (*e.g.,* taxon/discipline specific decision-making) while enhancing the level of connectivity among groups, allowing for easier aggregation and sharing of expertise, tools, and data. These network characteristics may also buffer against node loss (*e.g.,* due to lack of funding). Decentralised and connected networks are also known to be characterised by socially-mediated improvements in learning[69]. The OTN is, in part, about capitalising on existing robust connections within disciplines to more efficiently disseminate granular knowledge about traits.

Principles of the Open Traits Network

The OTN can strengthen and bridge between current initiatives in trait science and encourage efficient new enterprises by serving as a platform for sharing principles, methods, tools, examples, and approaches to support the wider traits community in developing its scientific practice. The OTN is based on core principles of Open Science (Fig. 2), namely:

1. Openly sharing data, methods, protocols, code, and workflows;
2. Appropriate citation of original data collectors, providing scholarly credit;
3. Provision of appropriate metadata together with trait observations;
4. Collection of trait data following reproducible, standardised methods and protocols (when available) or commitment to their development;
5. Providing training resources in trait collection and database construction using Open Science principles.

Below, we highlight key activities for the OTN designed to empower researchers to gather and make better use of trait data.

Key activities for the Open Traits Network

Activity 1: Maintaining a global registry of trait-based initiatives

The OTN maintains a global registry of trait-based initiatives (https://protectau.mimecast.com/s/Fu2DCyoj8PuyX879TZG_y5?domain=docs.google.com; Table 1) to help (i) connect the research community, (ii) identify data and knowledge gaps, (iii) prioritise trait collection, and (iv) allow researchers to avoid inadvertent duplication of efforts when collating trait data. The heterogeneous ways in which trait data have been collected to date have resulted in a patchy and unrepresentative trait landscape across trait types, taxa, regions, and times of the year[70]. These gaps impede synthetic analyses across taxa, space, and ontogeny.

The OTN Registry contains information on existing datasets so that gaps can be easily identified, and ultimately filled, through collective effort. Core information for the registry includes trait name, geographic extent, taxonomic coverage, and temporal period (Table 1), and existing knowledge from Ecological Metadata Language (http://rd-alliance.github.io/metadata-directory/standards/eml-ecological-metadata-language.html) and Darwin Core is adopted. Critically, the OTN Registry provides the opportunity for contributors to identify where code to process and manipulate their raw data is located (see *Activity 2* below). The OTN Registry will also link to the ontology resource OBO Foundry (http://www.obofoundry.org). Thus, the OTN registry maps to several Open Science principles (Fig. 2; *e.g.*, Open Source, Open Data, Open Access) and is designed, from the ground up, to support resolving the issue of data integration.

Activity 2: Sharing reproducible workflows and tools for aggregating trait data

The OTN leverages collaborative software development via platforms like GitHub (https://github.com/) to create a toolbox of modular open source software for access and harmonisation and re-use of trait data, with seamless piping of data from one tool to the next. OTN contributors have already developed several open source tools. For instance, the *traitdataform* package assists R users to format their data and harmonise units (http://ecologicaltraitdata.github.io/traitdataform); the code for the Coral Traits database[40] (https://github.com/jmadin/traits) could be easily modified to guide the creation of databases on other organisms; and the FENNEC project provides a tool for accessing and viewing community trait data as a self-hosted website service[71] (https://github.com/molbiodiv/fennec). The OTN acts as a connector for interactions between developers and the broader community seeking to synthesise trait data, facilitating the training of scientists in all aspects of reproducible data management.

Activity 3: Advocating for a free flow of data and appropriate recognition of efforts

A goal of the OTN is to improve how researchers receive credit, via citations, for the effort they have made to collect or synthesise primary data on species traits. Without effective reward or motivation for collecting new trait observations or liberating legacy data (*e.g.*, observations from field guides, specimens, publications without data supplements) a broad trait synthesis across the Tree of Life will remain unattainable. Currently, however, motivation for collecting and sharing new primary data is not strong.

The OTN can strengthen the attribution of credit to data providers via two paths. Firstly, by encouraging citation back to primary source via CC-BY licensing. There is an important distinction between sharing data within a network and making data publicly available under an open license. Clear license arrangements increase visibility and promote fair attribution/citation (*e.g.*, using creative commons licenses such as CC-BY or CC0). CC-BY requires attribution (*i.e.,* citation) to the original creator, whereas CC0 doesn't legally require users of the data to cite the source, but it does not affect the ethical norms for attribution in scientific and research communities (https://creativecommons.org/share-your-work/public-domain/cc0/). However, identifying where credit for prior work should be directed for legacy data is complicated, particularly where data involve a chain of expertise (*e.g.*, when trait data are extracted from taxonomic treatments which involve specimen collectors, digitisers, taxonomists, and curators).

Secondly, the OTN Trait Registry (*Activity 1*) can be used to identify high-value data gaps, helping would-be collectors of primary data on traits to have studies pre-registered via the Centre for Open Science (https://cos.io/prereg/). Already as of March 2019, 168 journals are willing to give in-principle acceptance before field or experimental work is conducted but following pre-review of the study design. Approximately ten of these participating journals regularly feature papers on trait-based science, including *BMC Ecology* and *Ecology and Evolution.*

Activity 4: Advocating a common metadata standard across nodes

Given the highly contextual nature of trait data, metadata are as important as the measurements themselves. The OTN provides a platform for the development of metadata standards, controlled vocabularies, and a suite of trait ontologies which can be recorded in the OTN Registry (*Activity 1*). Several initiatives have developed metadata standards (*e.g.*, Darwin Core[64]; Humboldt Core[58]; Ecological Metadata Language[72]). However, these metadata standards are yet to be commonly applied in trait-based data publications and syntheses. Using referencing terms from anatomy or phenotype ontologies (*e.g.,* the Plant Ontology[73]; the Vertebrate Trait Ontology[74]) relates traits semantically to publicly-defined terms and allows data thus annotated to be processed computationally[75-77]. Over time, the further development and implementation of metadata standards in the OTN will help to avoid downstream issues in data re-use and synthesis.

1 **Table 1. Structure of the OTN global registry of trait-based initiatives**

Field name	**Definition**	**Format/Values**	**Condition**
datasetID	A unique identifier for each dataset	Integer	Required (automatically generated)
datasetDOI_URL	Location of the dataset or database on the internet, e.g. the database website or the address of the dataset on a file hosting service	Character	Required
contactName	Corresponding author or maintainer for the dataset or database (name)	Characters separated by vertical bar space | (if multiple names)	Required
contactEmail	Corresponding author or maintainer for the dataset or database (email address)	Characters by vertical bar space | (if multiple names)	Recommended
License	A legal document giving official permission to do something with the resource (http://purl.org/dc/terms/license); see https://creativecommons.org/licenses/	Character	Required
traitList	List of traits. Names should correspond to existing controlled vocabulary (when applicable)	Characters separated by vertical bar space | (if multiple traits)	Recommended
higherGeography	A list (concatenated and separated) of geographic names less specific than the information captured in the locality term. (from: http://rs.tdwg.org/dwc/terms/higherGeography)	Characters separated by vertical bar space | (if multiple traits)	Recommended
decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic centre of a Location. Positive values are north of the Equator, negative values are south of it. Legal	Decimal degrees	Recommended

	values lie between -90 and 90, inclusive. (http://rs.tdwg.org/dwc/terms/decimalLatitude)		
decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.	Decimal degrees	Recommended
Taxon	A group of organisms (*sensu* http://purl.obolibrary.org/obo/OBI_0100026) considered by taxonomists to form a homogeneous unit. (see http://rs.tdwg.org/dwc/terms/Taxon)	Characters separated by vertical bar space | (if multiple groups)	Required
eventDate	The date-time or interval during which an Event occurred. For occurrences, this is the date-time when the event was recorded. Not suitable for a time in a geological context. (from http://rs.tdwg.org/dwc/terms/eventDate)	Use a date that conforms to ISO 8601:2004E (see http://rs.tdwg.org/dwc/terms/eventDate)	Recommended
paperDOIcitation	Immutable document object identifier of the dataset or database; or actual citation (if collection of multiple datasets, DOIs or citations must appear in metadata)	Character	Required
Description	Description of the dataset or any other useful information	Character	Recommended
taxaList	List of all taxa with some trait information	Characters separated by vertical bar space | (if multiple taxa)	Recommended
usefulClasses	Any other useful information for improved searching (e.g., life stage, body part, inferred traits	Character	Recommended

	etc.)		
dataStandard	Automated data consolidation (e.g., Darwin Core or Ecological Trait-Data Standard)	Character	Recommended
standardizationScripts	Links to scripts for data standardization	Character (using a standard that can be parsed)	Recommended

Activity 5: Facilitating consistent approaches to measuring traits within major groups

The OTN will advocate for the development of protocols and handbooks for major clades that standardise approaches to capture traits. Protocols are necessary because downstream activities such as developing metadata standards (*Activity 4*) will be impossible to create if trait measurement protocols do not exist. Some trait-research communities have already adopted standardised lists of terms[66] and trait data collection protocols (*e.g.*, plants[28,56,78-80], invertebrates[37,55,81,82], mammals[44], aquatic life[38,40] https://www.sealifebase.org/home/index.php). New protocols and handbooks may not emerge rapidly and should have the flexibility to be open to innovation through a commitment to creating versions and updates as techniques evolve (*e.g.*, from[83] to[56], or as in http://prometheuswiki.org/tiki-custom_home.php).

Standardising approaches to trait measurement *a priori* across communities of researchers will reduce ambiguity when aggregating data and improve the quality of resulting datasets. Also, integrating trait standardization and databasing in to taxonomic workflows represents both a considerable challenge and opportunity[8], but also holds the promise of bridging a long disconnect between structural and functional traits. The presence of a range of biodiversity collections personnel in the OTN, and an open invitation for more to join, has the potential to catalyse the adoption of trait-based thinking into taxonomic practices.

Concluding remarks

The creation of an open, trait-based view of global diversity is now possible given the myriad tools and approaches to data mobilisation and aggregation, harmonisation, and processing. These advances in technology are also accompanied by numerous emerging opportunities to work with institutions seeking to deliver biodiversity information to the public, and with citizen scientists[84] keen to gather trait data through platforms like (*e.g.,* Zooniverse https://www.zooniverse.org/; iNaturalist https://www.inaturalist.org/). The OTN aims at supporting a reciprocal exchange of expertise and outputs between researchers, institutions, and citizen scientists based on Open Science principles to accelerate a cross-taxa, worldwide, trait-based data resource to examine, understand, and predict nature's responses.

ACKNOWLEDGEMENTS

Ideas presented stem from initial discussions at three separate international meetings – the Australian National Climate Change Adaptation Research Facility Roundtable on Species Traits, the iDigBio ALA Traits workshop, and a preliminary Open Traits workshop held at the Ecological Society of America. RVG is supported by an Australian Research Council DECRA Felllowship (DE170100208). DSF is supported by an Australian Research Council Future Felllowship (FT160100113). WDP is supported by NSF ABI-1759965, NSF EF-1802605, and USDA Forest Service agreement 18-CS-11046000-041. CMI was supported by the Biological and Environmental Research program in the United States Department of Energy's Office of Science. MJ was supported by the German Research Foundation within the framework of the Jena Experiment (FOR 1451) and by the Swiss National Science Foundation. AK received financial support for MJA by the German Research Foundation (DFG KE1743/7-1). STM is supported by SERDP project RC18-1346. CP is supported by the DFG Priority Program 1374.

DATA AVAILABILITY STATEMENT

There are no original data associated with this manuscript.

BOXES

Box 1: Potential research programs that could be carried out with comprehensive trait data across the Tree of Life

- A multi-kingdom analysis of adult size exploring mechanistic constraints and ecological correlates.
- Whole-ecosystem or multi-trophic analyses of common traits which influence function (*e.g.*, relating traits to ecosystem processes, such as how differences in traits and migration patterns in birds can influence soil nutrient cycling).
- Prediction of community assembly processes across time and space (*e.g.*, from deep-time via the fossil record to present day human-dominated systems).
- Reserve selection optimised for protecting function (*e.g.*, reserve design in both marine and terrestrial realms based functional attributes of the broad community of organisms).
- Developing better Earth System models, biophysical niche models, and process-based mortality models through the integration of trait data.

REFERENCES

1 Chapin III, F. S., Autumn, K. & Pugnaire, F. Evolution of suites of traits in response to environmental stress. *The American Naturalist* **142**, S78-S92 (1993).

2 Diaz, S., Cabido, M. & Casanoves, F. Plant functional traits and environmental filters at a regional scale. *Journal of Vegetation Science* **9**, 113-122 (1998).

3 Violle, C. *et al.* Let the concept of trait be functional! *Oikos* **116**, 882-892 (2007).

4 Adler, P. B. *et al.* Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences* **111**, 740-745 (2014).

5 Kooijman, S. A. L. M. *Dynamic energy and mass budgets in biological systems*. (Cambridge University Press, 2000).

6 Chown, S. L. & Gaston, K. J. Macrophysiology–progress and prospects. *Functional Ecology* **30**, 330-344 (2016).

7 Harmon, L. J. *et al.* Early bursts of body size and shape evolution are rare in comparative data. *Evolution: International Journal of Organic Evolution* **64**, 2385-2396 (2010).

8 Sauquet, H. & Magallón, S. Key questions and challenges in angiosperm macroevolution. *New Phytologist* **219**, 1170-1187 (2018).

9 Sneath, P. H. & Sokal, R. R. *Numerical taxonomy. The principles and practice of numerical classification*. (W.H. Freeman, 1973).

10 Edmunds, R. C. *et al.* Phenoscape: identifying candidate genes for evolutionary phenotypes. *Molecular Biology and Evolution* **33**, 13-24 (2015).

11 Mungall, C. J. *et al.* The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Research* **45**, D712-D722 (2016).

12 Gkoutos, G. V., Schofield, P. N. & Hoehndorf, R. The anatomy of phenotype ontologies: principles, properties and applications. *Briefings in Bioinformatics* **19**, 1008-1021 (2017).

13 Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A. & Wright, I. J. Plant ecological strategies: some leading dimensions of variation between species. *Annual Review of Ecology and Systematics* **33**, 125-159 (2002).

14 Kiørboe, T., Visser, A. & Andersen, K. H. A trait-based approach to ocean ecology. *ICES Journal of Marine Science* **75**, 1849-1863 (2018).

15 Kunstler, G. *et al.* Plant functional traits have globally consistent effects on competition. *Nature* **529**, 204 (2016).

16 Laughlin, D. C. Nitrification is linked to dominant leaf traits rather than functional diversity. *Journal of Ecology* **99**, 1091-1099 (2011).

17 Finegan, B. *et al.* Does functional trait diversity predict above-ground biomass and productivity of tropical forests? Testing three alternative hypotheses. *Journal of Ecology* **103**, 191-201 (2015).

18 Laigle, I. *et al.* Species traits as drivers of food web structure. *Oikos* **127**, 316-326 (2018).

19 Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771-1789 (2004).

20 West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122-126 (1997).

21 Butt, N. & Gallagher, R. Using species traits to guide conservation actions under climate change. *Climatic Change* **151**, 317-332 (2018).

22 Devictor, V. *et al.* Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: the need for integrative conservation strategies in a changing world. *Ecology letters* **13**, 1030-1040 (2010).

23 Martin, A. R. & Isaac, M. E. Plant functional traits in agroecosystems: a blueprint for research. *Journal of Applied Ecology* **52**, 1425-1435 (2015).

24 Burkle, L. A., Delphia, C. M. & O'Neill, K. M. A dual role for farmlands: food security and pollinator conservation. *Journal of Ecology* **105**, 890-899 (2017).

25 Foden, W. B. *et al.* Identifying the world's most climate change vulnerable species: a systematic trait-based assessment of all birds, amphibians and corals. *PloS one* **8**, e65427 (2013).

26 Fyllas, N. *et al.* Analysing Amazonian forest productivity using a new individual and trait-based model (TFS v. 1). *Geoscientific Model Development* **7**, 1251-1269 (2014).

27 van Bodegom, P. M., Douma, J. C. & Verheijen, L. M. A fully traits-based approach to modeling global vegetation distribution. *Proceedings of the National Academy of Sciences* **111**, 13733-13738 (2014).

28 Iversen, C. M. *et al.* A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. *New Phytologist* **215**, 15-26 (2017).

29 Kattge, J. *et al.* TRY–a global database of plant traits. *Global Change Biology* **17**, 2905-2935 (2011).

30 Bernhardt-Römermann, M., Poschlod, P. & Hentschel, J. BryForTrait–A life-history trait database of forest bryophytes. *Journal of Vegetation Science* **29**, 798-800 (2018).

31 Bennett, J. M. *et al.* GlobTherm, a global database on thermal tolerances for aquatic and terrestrial organisms. *Scientific Data* **5**, 180022 (2018).

32 Meiri, S. Traits of lizards of the world: Variation around a successful evolutionary design. *Global Ecology and Biogeography* **27**, 1168-1172 (2018).

33 Myhrvold, N. P. *et al.* An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles. *Ecology* **96**, 3109-3109 (2015).

34 Schäfer, R. B. *et al.* A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in South-East Australia. *Science of the Total Environment* **409**, 2055-2063 (2011).

35 Bland, L. Global correlates of extinction risk in freshwater crayfish. *Animal Conservation* **20**, 532-542 (2017).

36 Brun, P., Payne, M. R. & Kiørboe, T. A trait database for marine copepods. *Earth System Science Data* **9**, 99-113, doi:10.5194/essd-9-99-2017 (2017).

37 Parr, C. L. *et al.* GlobalAnts: a new database on the geography of ant traits (Hymenoptera: Formicidae). *Insect Conservation and Diversity* **10**, 5-20 (2017).

38 Froese, R. & Pauly, D. *Fishbase* (Fisheries Centre, University of British Columbia, 2010).

39 Frimpong, E. A. & Angermeier, P. L. Fish traits: a database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries* **34**, 487-495 (2009).

40 Madin, J. S. *et al.* The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Scientific Data* **3**, 160017 (2016).

41 Garnett, S. T. *et al.* Biological, ecological, conservation and legal information for all species and subspecies of Australian bird. *Scientific Data* **2**, 150061 (2015).

42 Wilman, H. *et al.* EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals: Ecological Archives E095-178. *Ecology* **95**, 2027-2027 (2014).

43 Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C. & Costa, G. C. AmphiBIO, a global database for amphibian ecological traits. *Scientific Data* **4**, 170123 (2017).

44 Jones, K. E. *et al.* PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648-2648 (2009).

45 Faurby, S. *et al.* PHYLACINE 1.2: The phylogenetic atlas of mammal macroecology. *Ecology* **99**, 2626-2626 (2018).

46 Deans, A. R. *et al.* Finding our way through phenotypes. *PLoS Biology* **13**, e1002033 (2015).

47 Maitner, B. S. *et al.* The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution* **9**, 373-379 (2018).

48 Jetz, W., Thomas, G., Joy, J., Hartmann, K. & Mooers, A. The global diversity of birds in space and time. *Nature* **491**, 444 (2012).

49 Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany* **105**, 302-314 (2018).

50 Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217-223 (2012).

51 Munafò, M. R. *et al.* A manifesto for reproducible science. *Nature Human Behaviour* **1**, 0021, doi:10.1038/s41562-016-0021 (2017).

52 Nosek, B. A. *et al.* Promoting an open research culture. *Science* **348**, 1422-1425 (2015).

53 Hardisty, A. R. *et al.* The Bari Manifesto: An interoperability framework for essential biodiversity variables. *Ecological Informatics* **49**, 22-31 (2019).

54 Dawson, S. K. *et al.* Handbook for the measurement of macrofungal functional traits: A start with basidiomycete wood fungi. *Functional Ecology* **33**, 372-387 (2019).

55 Moretti, M. *et al.* Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology* **31**, 558-567 (2017).

56 Perez-Harguindeguy, N. *et al.* A new handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany* **64**, 715-716 (2013).

57 Schilthuizen, M., Vairappan, C. S., Slade, E. M., Mann, D. J. & Miller, J. A. Specimens as primary data: Museums and 'open science'. *Trends in Ecology and Evolution* **30**, 237-238, doi:10.1016/j.tree.2015.03.002 (2015).

58 Guralnick, R., Walls, R. & Jetz, W. Humboldt Core–toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* **41**, 713-725 (2018).

59 Farley, S. S., Dawson, A., Goring, S. J. & Williams, J. W. Situating ecology as a big-data science: Current advances, challenges, and solutions. *BioScience* **68**, 563-576 (2018).

60 Falster, D. S. *et al.* BAAD: a Biomass And Allometry Database for woody plants. *Ecology* **96**, 1445-1445 (2015).

61 Salguero-Gómez, R. *et al.* COMADRE: a global data base of animal demography. *Journal of Animal Ecology* **85**, 371-384 (2016).

62 Salguero-Gómez, R. *et al.* The COMPADRE Plant Matrix Database: an open online repository for plant demography. *Journal of Ecology* **103**, 202-218 (2015).

63 Marques, G. M. *et al.* The AmP project: comparing species on the basis of dynamic energy budget parameters. *PLoS Computational Biology* **14**, e1006100 (2018).

64 Wieczorek, J. *et al.* Darwin Core: an evolving community-developed biodiversity data standard. *PloS one* **7**, e29715 (2012).

65 Haendel, M. A. *et al.* Unification of multi-species vertebrate anatomy ontologies for comparative biology in Uberon. *Journal of Biomedical Semantics* **5**, 21 (2014).

66 Garnier, E. *et al.* Towards a thesaurus of plant characteristics: an ecological contribution. *Journal of Ecology* **105**, 298-309 (2017).

67 Gene Ontology Consortium. The Gene Ontology resource: 20 years and still GOing strong. *Nucleic Acids Research* **47**, D330-D338 (2018).

68 Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J. & Lewis, S. E. The environment ontology: contextualising biological and biomedical entities. *Journal of Biomedical Semantics* **4**, 43 (2013).

69 Becker, J., Brackbill, D. & Centola, D. Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 201615978 (2017).

70 Kissling, W. D. *et al.* Towards global data products of Essential Biodiversity Variables on species traits. *Nature Ecology & Evolution*, **1** (2018).

71 Ankenbrand, M. J., Hohlfeld, S. C., Weber, L., Förster, F. & Keller, A. Functional exploration of natural networks and ecological communities. *Methods in Ecology and Evolution* **9**, 2028-2033 (2018).

72 Fegraus, E. H., Andelman, S., Jones, M. B. & Schildhauer, M. 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**, 158-168 (2005).

73 Jaiswal, P. *et al.* Plant Ontology (PO): a controlled vocabulary of plant structures and growth stages. *International Journal of Genomics* **6**, 388-397 (2005).

74 Park, C. A. *et al.* The Vertebrate Trait Ontology: a controlled vocabulary for the annotation of trait data across species. *Journal of Biomedical Semantics* **4**, 13 (2013).

75 Dececchi, T. A., Balhoff, J. P., Lapp, H. & Mabee, P. M. Toward synthesizing our knowledge of morphology: Using ontologies and machine reasoning to extract presence/absence evolutionary phenotypes across studies. *Systematic Biology* **64**, 936-952 (2015).

76 Walls, R. L. *et al.* Semantics in support of biodiversity knowledge discovery: an introduction to the biological collections ontology and related ontologies. *PLoS One* **9**, e89606 (2014).

77 Schneider, F. D. *et al.* Towards an Ecological Trait-data Standard. *bioRxiv*, doi:10.1101/328302 (2018).

78 Fang, J. *et al.* Methods and protocols for plant community inventory. *Biodiversity Science* **17**, 533-548 (2009).

79 Sack, L. *et al.* A unique web resource for physiology, ecology and the environmental sciences: PrometheusWiki. *Functional Plant Biology* **37**, 687-693 (2010).

80 Bjorkman, A. D. & Myers-Smith, I. H. Tundra Trait Team: A database of plant traits spanning the tundra biome. *Global Ecology and Biogeography* **27**, 1402-1411 (2018).

81 Ferris, H. NEMAPLEX: The Nematode-Plant Expert Information System. Online at: *http://nemaplex.ucdavis.edu* (2001).

82 Tennessen, J. M., Barry, W. E., Cox, J. & Thummel, C. S. Methods for studying metabolism in Drosophila. *Methods* **68**, 105-115 (2014).

83 Cornelissen, J. *et al.* A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany* **51**, 335-380 (2003).

84 Ballard, H. L. *et al.* Contributions to conservation outcomes by natural history museum-led citizen science: Examining evidence and next steps. *Biological Conservation* **208**, 87-97 (2017).