nature ecology & evolution

PERSPECTIVE

https://doi.org/10.1038/s41559-018-0667-3

OPEN

Towards global data products of Essential Biodiversity Variables on species traits

W. Daniel Kissling[1*], Ramona Walls[2], Anne Bowser[3], Matthew O. Jones[4], Jens Kattge[5,6], Donat Agosti[7], Josep Amengual[8], Alberto Basset[9], Peter M. van Bodegom[10], Johannes H. C. Cornelissen[11], Ellen G. Denny[12], Salud Deudero[13], Willi Egloff[7], Sarah C. Elmendorf[14,15], Enrique Alonso García[16], Katherine D. Jones[14], Owen R. Jones[17], Sandra Lavorel[18], Dan Lear[19], Laetitia M. Navarro[6,20], Samraat Pawar[21], Rebecca Pirzl[22], Nadja Rüger[6,23], Sofia Sal[21], Roberto Salguero-Gómez[24,25,26,27], Dmitry Schigel[28], Katja-Sabine Schulz[29], Andrew Skidmore[30,31] and Robert P. Guralnick[32]

Essential Biodiversity Variables (EBVs) allow observation and reporting of global biodiversity change, but a detailed framework for the empirical derivation of specific EBVs has yet to be developed. Here, we re-examine and refine the previous candidate set of species traits EBVs and show how traits related to phenology, morphology, reproduction, physiology and movement can contribute to EBV operationalization. The selected EBVs express intra-specific trait variation and allow monitoring of how organisms respond to global change. We evaluate the societal relevance of species traits EBVs for policy targets and demonstrate how open, interoperable and machine-readable trait data enable the building of EBV data products. We outline collection methods, meta(data) standardization, reproducible workflows, semantic tools and licence requirements for producing species traits EBVs. An operationalization is critical for assessing progress towards biodiversity conservation and sustainable development goals and has wide implications for data-intensive science in ecology, biogeography, conservation and Earth observation.

In 2013, the Group on Earth Observations Biodiversity Observation Network (GEO BON) introduced the framework of Essential Biodiversity Variables (EBVs) to derive coordinated measurements critical for detecting and reporting biodiversity change[1]. Through this process, 22 candidate EBVs were proposed and organized within six classes ('genetic composition', 'species populations', 'species traits', 'community composition', 'ecosystem structure' and 'ecosystem function')[1]. These EBVs provide a foundation for assessing progress towards national and international policy goals, including the 20 Aichi Biodiversity Targets developed by the Parties to the United Nations (UN) Convention on Biological Diversity (CBD) and the 17 Sustainable Development Goals (SDGs) identified by the UN 2030 Agenda for Sustainable Development[2]. EBVs are conceptually located on a continuum between primary data observations ('raw data') and synthetic or derived metrics ('indicators'), and can be represented as 'data cubes' with several basic dimensions (for example, time, space, taxonomy or Earth observation data types)[3–5]. Hence, EBVs allow derivation of biodiversity indicators (for example, trends of biodiversity change) such as those developed for the Aichi Biodiversity Targets, with several EBVs (for example, species population abundance) informing multiple targets[1,6]. Specific EBVs in the classes species populations, ecosystem structure and

[1]Department of Theoretical and Computational Ecology, Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands. [2]CyVerse, University of Arizona, Tucson, AZ, USA. [3]Woodrow Wilson International Center for Scholars, Washington DC, USA. [4]University of Montana, W. A. Franke Department of Forestry and Conservation, Missoula, MT, USA. [5]Max Planck Institute for Biogeochemistry, Jena, Germany. [6]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [7]Plazi, Bern, Switzerland. [8]Area de Conservacion, Seguimiento y Programas de la Red, Organismo Autonomo Parques Nacionales, Ministerio de Agricultura y Pesca, Madrid, Spain. [9]Department of Biological and Environmental Sciences and Technologies, University of Salento, Lecce, Italy. [10]Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands. [11]Systems Ecology, Department of Ecological Science, Vrije Universiteit, Amsterdam, The Netherlands. [12]USA National Phenology Network, University of Arizona, Tucson, AZ, USA. [13]Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Palma de Mallorca, Spain. [14]National Ecological Observatory Network, Battelle Ecology, Boulder, CO, USA. [15]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA. [16]Franklin Institute, University of Alcala, Madrid, Spain. [17]Department of Biology, University of Southern Denmark, Odense M, Denmark. [18]Laboratoire d'Ecologie Alpine, CNRS - Université Grenoble Alpes, Grenoble, France. [19]Marine Biological Association of the United Kingdom, Plymouth, Devon, UK. [20]Institute of Biology, Martin Luther University Halle Wittenberg, Halle (Saale), Germany. [21]Department of Life Sciences, Imperial College London, Ascot, Berkshire, UK. [22]CSIRO and Atlas of Living Australia, Canberra, Australian Capital Territory, Australia. [23]Smithsonian Tropical Research Institute, Ancon, Panama. [24]Department of Zoology, Oxford University, Oxford, UK. [25]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK. [26]Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, Queensland, Australia. [27]Evolutionary Demography Laboratory, Max Plank Institute for Demographic Research, Rostock, Germany. [28]Global Biodiversity Information Facility (GBIF), Secretariat, Copenhagen, Denmark. [29]Smithsonian Institution, National Museum of Natural History, Washington DC, USA. [30]Department of Natural Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands. [31]Department of Environmental Science, Macquarie University, New South Wales, Australia. [32]Florida Museum of Natural History, University of Florida, Gainesville, FL, USA.
*e-mail: wdkissling@gmail.com

Box 1 | Definition and societal relevance of species traits EBVs

A species trait can be defined as any phenological, morphological, physiological, reproductive or behavioural characteristic of a species that can be measured at an individual level[11,91]. Hence, species traits can be quantified by measuring characteristics of individuals (for example, timing of flowering, body lengths of fish individuals, stem heights and diameters of tree individuals, leaf nitrogen and chlorophyll content) or parts of individuals (for example, area of an individual leaf).

Individual variation in trait measurements can be summarized at different hierarchical levels, for instance at the population level (for example, mean body length of a fish species population), at the species level (for example, intra-specific variability of body lengths of a fish species across its entire geographic range), or across multiple species (for example, as community-weighted means[91] or as spectral trait variation when using airborne or spaceborne remote sensing[43,92]). Quantifying trait variation across multiple species (that is, within a community, ecosystem or landscape) is highly relevant for mapping and monitoring ecosystem processes and functional diversity[43,51]. However, such community- and ecosystem-level trait variation is mainly relevant for the EBV classes 'community composition', 'ecosystem structure' and 'ecosystem function'[1], but not for 'species traits' because it does not allow attribution of trait variation to the species level[1].

A key aspect of EBV development is to standardize, aggregate and harmonize data across time (for example, temporal resolution), space (for example, spatial resolution and geographic extent) and biological organization (for example, taxonomy or Earth observation data type)[3–5]. Species traits EBVs can therefore be defined as standardized and harmonized data of phenological, morphological, physiological, reproductive or behavioural trait measurements that can be quantified at the level of individual organisms. To distinguish species traits EBVs from other EBV classes, we constrain them to trait measurements that allow quantification of trait changes within species populations (that is, intra-specific variation). Hence, trait measurements of individuals or populations must be attributable to the taxonomic level of a species (rather than to communities, landscapes or ecosystems). Alternatively (as in the case of micro-organisms), individuals might be identified at the level of operational taxonomic units (OTUs), that is, grouped by DNA sequence similarity rather than by a classical Linnaean taxonomy. Hence, taxonomic information, as well as time and location of trait data collection, is key for monitoring intra-specific trait changes.

The societal relevance of EBVs becomes crucial when assessing progress towards biodiversity targets and policy goals[1,2]. Species traits EBVs can be important for such targets, including the 20 Aichi Biodiversity Targets developed by Parties to the UN CBD and the 17 SDGs identified by the UN 2030 Agenda for Sustainable Development. For instance, the impact of harvesting large fish individuals for commercial fisheries could be monitored by trait measurements that quantify changes in mean or maximum body size (for example, body length at first maturity) in economically important fish populations[15,79]. This would allow deriving size-based indicators (for example, trends of maximal fish body lengths over time) and hence measuring overexploitation and unsustainable harvesting as specified in Aichi Target 6 (sustainable harvesting of fish and invertebrate stocks and aquatic plants) or SDG 2 (sustainable food production).

Species traits are also important for understanding the response of organisms to their environment ('response traits')[8]. For instance, phenological trait information (for example, related to changes in timing of bird egg laying, phytoplankton population peaks, or plant leafing, flowering and fruiting) can be an early indicator of climate change impacts[21] and has relevance for SDG 13 (combating climate change and its impacts). Other examples include trait measurements related to movement behaviour (for example, dispersal distances and pathways, animal home range size) and reproduction (for example, fruit and seed size). These trait measurements can be of societal relevance, for instance if they determine the success of alien invasive species[16], describe how organisms respond to habitat fragmentation[14], or indicate how species adapt to global change drivers[93]. This information is directly related to Aichi Target 5 (habitat loss and forest fragmentation) and Aichi Target 9 (invasive species control), but has yet to be developed into indicators.

Species traits EBVs can therefore provide critical information for monitoring biodiversity change, which cannot be captured by measuring changes in species distributions alone or ecosystem structure and functioning. Moreover, different species traits differ in their importance across policy targets and each species traits EBV contains important information with societal and policy relevance that cannot be substituted by other species traits EBVs (Supplementary Note 2).

ecosystem function are now being developed by GEO BON working groups[7]. However, other EBV classes have received less attention, and the research community has yet to fully coalesce efforts to develop the conceptual and empirical frameworks for those variables and their associated data products.

Species traits are a key component of biodiversity because they determine how organisms respond to disturbances and changing environmental conditions, with impacts at a population level and beyond[8–10]. Within the EBV framework, the EBV class 'species traits' has yet to be formally conceptualized in detail and therefore cannot yet be made operational. In line with previous work[8,11,12], we here define a species trait as any phenological, morphological, physiological, reproductive or behavioural characteristic of an individual that can be assigned to a species (Box 1). Because the building of EBV data products requires standardization and harmonization of raw measurements[1,3,5], we further define species traits EBVs as standardized and harmonized measurements of species' characteristics that allow monitoring of intra-specific trait changes within species populations across space and time (Box 1). Specific species traits selected for EBVs (for example, body mass, plant height and specific leaf area as examples of morphological traits) allow quantification of how species respond to global change including climate change, biological invasions, overexploitation and habitat fragmentation[8,13–16] (Box 1). The time frame of species traits responses should be policy relevant, that is, intra-specific trait changes should be detectable within a decade rather than only seasonally, annually or over evolutionary time scales[6]. This is needed because EBVs will feed into biodiversity change indicators (Box 1) that allow the assessment of progress towards policy goals including the SDGs and Aichi Biodiversity Targets as well as National Biodiversity Strategies and Action Plans (NBSAPs). They can also help to inform global and regional assessments of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES)[1,17]. Other aspects of species traits that reflect traits expressions at the community or ecosystem level are not considered here as they belong to other EBV classes (Box 1). To our knowledge there are currently no global data products available that allow direct measurement and monitoring of trait changes within species populations across time[17].

Here, we develop the conceptual and empirical basis for species traits EBVs to help to operationalize the development of global EBV

data products. We start by critically re-examining the current set of candidate species traits EBVs (phenology, body mass, natal dispersal distance, migratory behaviour, demographic traits and physiological traits). We then explore how trait data are collected, how they can be standardized and harmonized and what bottlenecks currently prevent them from becoming findable, accessible, interoperable and reusable (FAIR guiding principles)[18]. We further outline workflow steps to produce EBV data products of species traits, using an example of plant phenology. Our perspective provides a conceptual framework with practical guidelines for building global, integrated and reusable EBV data products of species traits. This will promote the use of species trait information in national and international policy assessments and requires significant advancements and new tools in ecology, biogeography, conservation and environmental science. Beyond the direct relevance to species traits EBVs, our perspective further explores cross-cutting issues related to data-intensive science, interoperability, and legal and policy aspects of biodiversity monitoring and Earth observation that will help to advance the EBV framework.

A critical re-examination

GEO BON has proposed six candidate EBVs in the EBV class species traits (Supplementary Table 1): phenology, body mass, natal dispersal distance, migratory behaviour, demographic traits and physiological traits. These candidate EBVs were discussed in detail during a three-day experts' workshop in Amsterdam (March 2017) organized by the GLOBIS-B project (http://www.globis-b.eu/)[19]. We suggest several key improvements of that initial list of candidate species traits EBVs.

Identified inconsistencies. We identified several inconsistencies in the proposed candidate list of species traits EBVs (summarized in Supplementary Table 2). First, some previously listed measurements — such as ocean and river flows, extent of wetlands and net primary productivity — do not occur at the species level (Box 1) and should therefore be placed within community or ecosystem-scale EBV classes such as community composition, ecosystem function or ecosystem structure. Second, several candidate EBVs (for example, body mass and natal dispersal distance) are narrowly defined compared to other candidate EBVs (for example, phenology, demographic traits, physiological traits), resulting in an inconsistent scope across EBVs. Third, a few candidate EBVs represent a similar category but are split into different EBVs (for example, both natal dispersal distance and migratory behaviour are aspects of movement behaviour), and should therefore be represented together. Fourth, the candidate EBV 'demographic traits' reflects population-level quantities that cannot be measured on individual organisms (for example, population growth rate, generation time, survival rate). These population-level metrics are derived from data that are captured by the EBV population structure by age/size/stage class belonging to another EBV class (species populations). It is therefore inconsistent to capture the same set of underlying measurements in two different EBV classes.

Suggestions for improvement. Based on our assessment, we suggest reducing the initial candidate list to five species traits EBVs (Fig. 1): phenology (timing of periodic biological events), morphology (dimensions, shape and other physical attributes of organisms), reproduction (sexual or asexual production of new individual organisms), physiology (chemical or physiological functions promoting organism fitness) and movement (spatial mobility of organisms) (see overview in Fig. 1 and detailed description in Supplementary Note 1). This improves the previous classification of species traits EBVs by standardizing the breadth and scope of EBVs, better recognizing the importance and relevance of reproductive traits and excluding ecosystem variables that cannot be measured at the scale of the individual and are thus not species-specific traits (Supplementary Note 1). These five species traits EBVs provide a conceptual framework for the EBV class species traits and are relevant to the Aichi Biodiversity Targets and SDGs (Fig. 1, Supplementary Table 3). Because GEO BON has the main responsibility for developing EBVs, we suggest that the new GEO BON working group on species traits (as recommended in the GEO BON implementation plan 2017–2020[7]) should take our suggestions into consideration when updating the EBV class species traits.

Collecting trait data

Many trait databases have recently emerged that support assembling trait measurements from published literature, specimen collections, in situ collections and close-range, airborne or spaceborne remote sensing (for examples see Supplementary Table 4). Nevertheless, the total demand for species traits in the EBV context is still unmet for the following reasons.

Aggregated species-level trait values are not sufficient. Many ongoing trait data collections assemble species trait information from published literature (Fig. 2). When aggregated to the species-level without location and time information (for example, mean species body length for morphology, or typical month of flowering or fruiting for phenology), this information does not allow measurement of trait changes within species populations over space or time, and hence lacks the ability to yield species traits EBVs (Fig. 2, Box 1). However, if the variation in the aggregated trait (that is, variance) can be calculated from a sufficiently large sample, then changes in species populations over time (or space) can be statistically estimated[15,20–22]. Nevertheless, many projects aggregate trait data at the species level from multiple sources such as published and unpublished trait datasets, natural history collections, citizen science projects and text mining[23–28]. These trait data remain limited in their application for species traits EBVs if they do not keep the resolution of the original data in terms of space, time and individual measurement information. The lack of individual or population measures therefore makes it difficult to assess intra-specific trait changes and the drivers and scales at which they operate.

Natural history collections offer historical data that remain underutilized. Museum and herbarium specimens allow study of individuals' traits in species populations of the recent past[29]. Specimen collections can therefore be an important source for individual-level trait measurements through time (Fig. 2). For example, specimens have been used to document temporal changes in morphology (for example, bird and beetle body size[30,31]) and phenology (for example, timing of flowering[32,33]) during the past century. Billions of specimens are available for study, but efforts to digitize and store trait data associated with specimens are still in their infancy[29]. Hence, trait data from digitized specimen collections remain underutilized and are currently too often constrained and biased in space, time and number of individuals[25]. New ways to digitize biocollections and to automate trait data extraction from specimens are needed[25], and analyses must take into account the constraints and biases inherent in these data[34].

In situ monitoring of traits is promising but labour intensive. A promising approach for developing species traits EBVs is to collect in situ trait data through monitoring schemes (Fig. 2). These include repeated trait measurements (for example, of animal body size, plant size, lichen length, flower and fruit phenology, leaf morphology and chemistry) with standardized protocols using long-term ecological research sites[35,36] or national and international monitoring programmes and citizen science networks[20,37,38]. Such sites and networks can monitor a comprehensive set of trait measurements for targeted species or sites through time and at continental

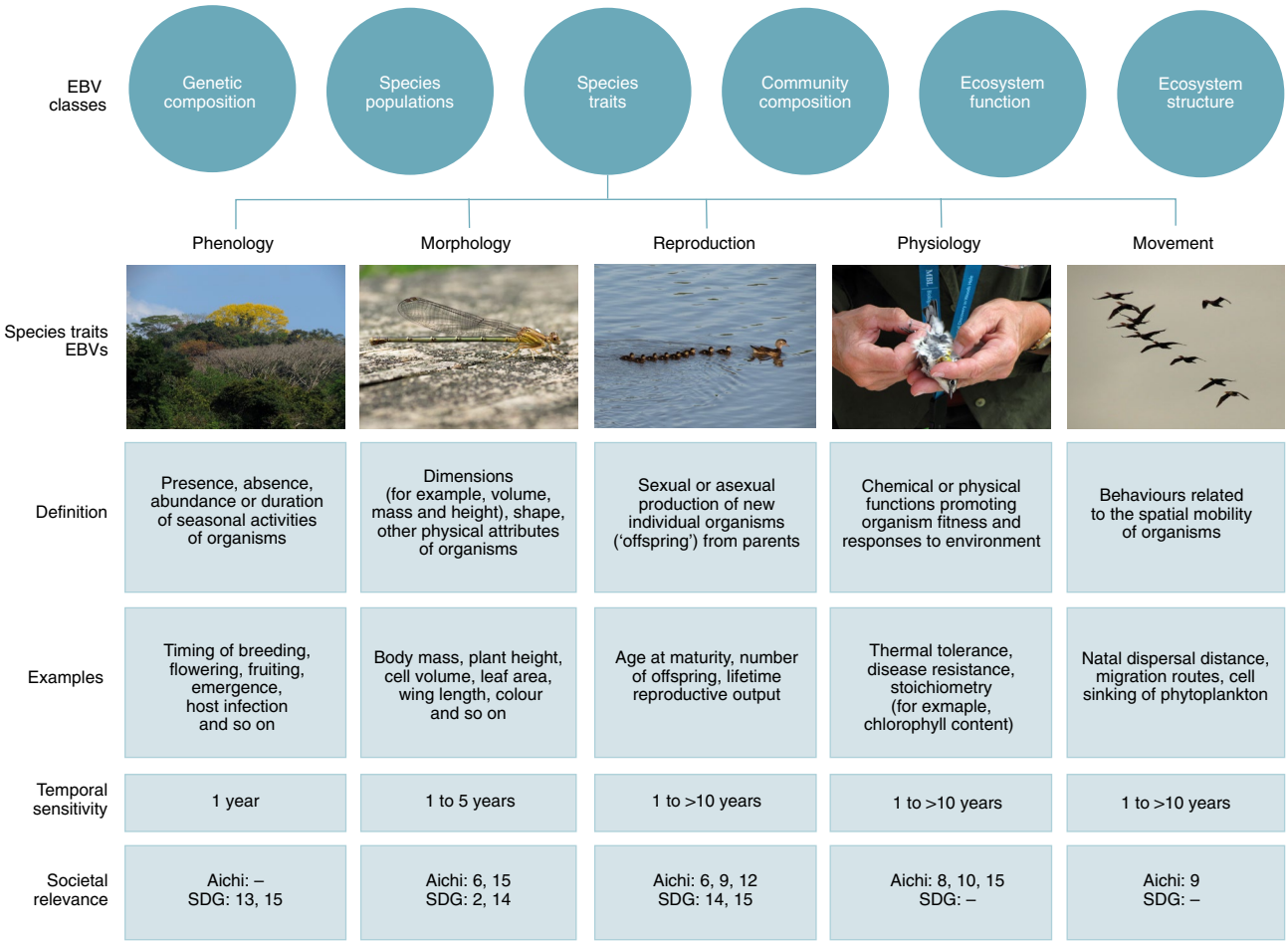


Fig. 1 | A framework for EBVs on species traits. We suggest five EBVs within the EBV class 'species traits', comprising (1) phenology, (2) morphology, (3) reproduction, (4) physiology and (5) movement. For each EBV, a definition, examples of species trait measurements, temporal sensitivity and societal relevance are given. Societal relevance refers to those Aichi Biodiversity Targets and SDGs to which the specific EBV is of highest relevance (for details on societal relevance see Supplementary Note 2 and Supplementary Table 2). Photo credits: Katja-Sabine Schulz.

extents[38,39], but remain costly and labour intensive. The future collection of trait data time series through in situ monitoring therefore requires prioritization according to global and regional biodiversity and sustainability goals, and a robust temporal replication and spatial/environmental stratification of the sampling design[40].

Remote sensing observations are promising but often not species specific. Airborne, spaceborne and close-range remote sensing techniques are promising tools (Fig. 2) because they can extend the geographic and temporal dimensions of trait measurements considerably[9,41–43]. Increasingly, ground-based light detection and ranging (that is, terrestrial LiDAR) is automating in situ data collection and allows retrieval of species trait information for individual plants (for example, height[44] and leaf water content[45]). Moreover, sensor-derived trait data can provide individual- or population-level trait measurements from close-range instruments such as camera traps, phenology cameras[46,47], field spectrometers[48], wireless sensor networks, unmanned aerial vehicle (UAV) and aircraft mounted instruments such as airborne LiDAR and hyperspectral sensors[49,50]. Combining airborne LiDAR and imaging spectroscopy also allows mapping of individual-level variation in morphological and physiological traits (for example, canopy height, leaf chlorophyll and water content) at regional scales[43]. For species traits EBVs, the remotely sensed trait measurements require fine enough spatial resolution to attribute them to an individual or population of a particular species (Box 1). A synergy of hyperspectral and LiDAR remote sensing with airborne sensors has great potential for developing species traits EBVs, but is not available at a global extent. Spaceborne remote sensing systems can provide global coverage, but they still show a large deficit for providing an operational combination of data at high spatial and spectral resolution[9,42,51]. In other words, spaceborne instruments are in their infancy for monitoring species traits due to limitations with very high spatial resolution (pixel area) and spectral resolution (high number and small width of spectral bands), though new spaceborne imaging spectrometers and LiDAR are planned which will go some way towards closing this gap[42,52,53]. Further developments in instrumentation and data[52], planned satellite sensor missions[53], species-level spectral library databases (for example, EcoSIS; https://ecosis.org) and spectranomics[54,55] — the coupling of spectroscopy with plant phylogeny and canopy chemistry — will further enhance the ability to retrieve species-specific trait data.

Standardizing trait data

A current bottleneck for integrating trait datasets from multiple sources is that measurements, data and metadata are not sufficiently standardized. We highlight three focal areas to improve this.

Standardizing protocols for measuring traits. The use of standardized measurement protocols during the phase of trait data collection is foundational for integrating data into EBV data

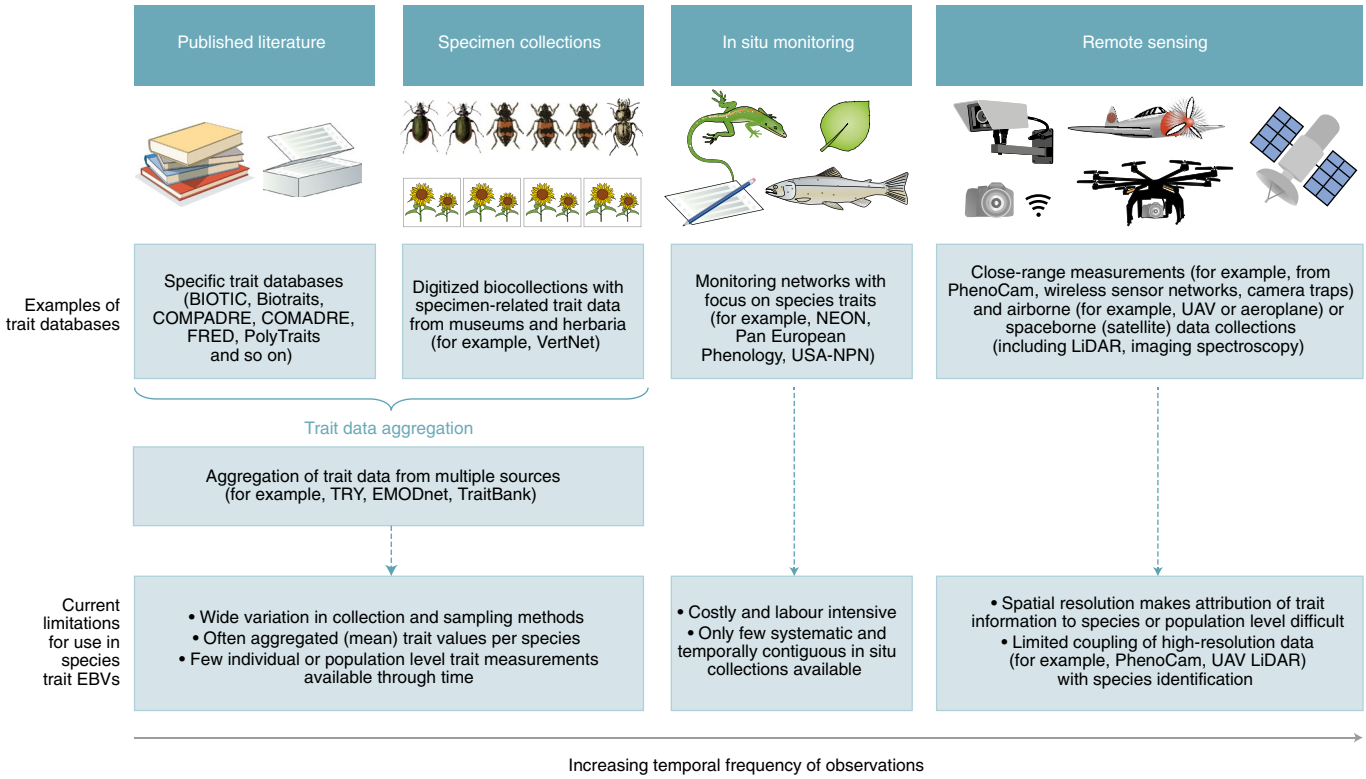


Fig. 2 | Methods for trait data collection with examples of trait databases and limitations for developing EBVs. Several methods are used to assemble comprehensive trait databases, for example, from published literature, specimen collections, in situ monitoring and remote sensing (close-range, airborne and spaceborne). These methods can be ordered along a gradient of increasing temporal frequency of observations. Aggregation of trait data from multiple sources often does not provide measurements repeated in time and hence typically does not allow monitoring of trait changes within species populations. More information about trait databases (abbreviations) is provided in Supplementary Table 3.

products. Good examples of comprehensive protocols for standardized measurements of morphological, reproductive, physiological and behavioural traits exist for vascular plants[56,57] and terrestrial invertebrates[58]. However, such comprehensive definitions of measurement protocols are still missing for most traits and taxa, and some remain little-known and difficult to access[59]. This is particularly true for remote sensing measurements of species traits (for example, leaf chlorophyll concentration and canopy chlorophyll content) where the instrumentation and required pre-processing of data to derive information on species-specific traits may vary considerably even within the same class of sensors (for example, within different types of spectrometers, phenology cameras or LiDAR instruments). A coordinated effort is therefore needed to develop and harmonize standardized measurement protocols for various taxa and across data types, sensors and regions, and to support consistent monitoring across political boundaries.

Standardizing trait terminology. Aggregating trait data from multiple sources requires standardized lists of trait terms or controlled vocabularies (that is, carefully selected lists of words and phrases)[11,27,60,61]. For instance, in the marine domain the formalization of a standardized list of trait terms and definitions has been achieved across a wide range of taxa[26,60]. Similar examples exist for other taxa and realms, for example, the thesaurus of plant characteristics[11]. Nevertheless, comprehensive trait vocabularies that provide standardized terms, definitions, units and synonyms for trait data and their metadata remain scarce. The further development and linking of such trait vocabularies is therefore needed to achieve semantic interoperability and facilitate integration of trait datasets[11,23,27,62].

Ontologies. Integrating trait data from disparate sources requires mapping trait data to ontologies[23,25,61,63–66], that is, to semantic models that allow formal descriptions of the relationships among trait concepts and vocabulary terms (Box 2). For trait data in particular, not only information about the occurrence of a species and the identification process needs to be reported, but also information about the entity (that is, whether specific parts of organisms, individual organisms, populations or species are measured), the measurement focus (for example, mass, length or area), the measurement units (for example, plant height in m, leaf nitrogen content in mg g^{-1}, photosynthetic rate in µmol m^2 s^{-1}) and the protocols used. Because many traits exhibit phenotypic plasticity, information about the individuals' living conditions before trait measurements (for example, if a plant was exposed to direct sunlight or shaded in the understory) is also essential to understand and interpret trait measurements[67]. Such reporting can be standardized by connecting two types of ontology: (1) observation and measurement ontologies for traits and environmental conditions and (2) ontologies for entities and qualities (Box 2). Various examples of both types of ontology already exist (Box 2), but their wider integration for developing comprehensive species traits data products has not yet been achieved.

Making trait data open and machine-readable

A workflow-oriented production of EBVs requires trait datasets and their metadata to be openly accessible and machine-readable[3,18]. Although openness and sharing of biodiversity data are improving[68–70] and trait databases increasingly develop data management policies around open access principles (see Supplementary Note 3 for an assessment of openness of individual species traits datasets), the actual levels of open and FAIR[18] access to trait data are still

Box 2 | Semantic tools for reporting trait measurements

Reporting trait data is best accomplished using two types of ontologies (that is, semantic models): those that describe the processes, inputs and outputs around data collection, and those that systematically describe the traits themselves. The first type of ontology standardizes observation and measurement data that is important for capturing how trait measurements were performed (for example, protocols), metadata on taxon, sampling location, sampling time and so on, and tracking data provenance. A key example is the Extensible Observation Ontology (OBOE), which captures the semantics of observational datasets, including field, experimental, simulation and monitoring data[94]. Similarly, the Biological Collections Ontology (BCO) allows sampling, specimen collection and observations to be reported in a standardized way[95]. For geospatial data, the Observations and Measurements (O&M) ontology allows interoperability with sensor data and could be valuable to report information such as optical traits related to plant function[51]. Further progress is still needed to create interoperability across different observation ontologies and develop easy-to-use implementations. Moreover, comprehensive definitions of measurement protocols and methods are lacking.

The second type of ontology (that is, semantic models for describing traits) is most commonly based on the Entity–Quality (E–Q) model[63]. The E–Q model provides a framework for adequately describing the entity (for example, a leaf of a plant, of individual organisms, populations or species) and the quality of that entity being measured, such as mass, length or area. Standardized trait data must also include information on how they are measured (for example, protocols), and the units used for coding the trait value[96]. While the E–Q model was originally developed for the description of phenotypes in the field of biomedicine[63], there are now many applications to ecological trait data. Examples for plant traits include the Thesaurus of Plant Characteristics (TOP)[11], the Flora Phenotype Ontology (FLOPO)[64], the Plant Trait Ontology (TO)[65] and the PPO[73]. Similar examples can be found for animal traits[61,66,97]. In addition, trait measurements should also be linked to descriptions of the environment in which the individuals have been living[67], for example, using the Environment Ontology (ENVO)[98]. The combination of trait ontologies with observation process ontologies provides a strong basis for standardizing how traits are measured, compiled, shared and made semantically interoperable (see Box 3).

lagging behind the ideal, although remote-sensing data are increasingly freely available, especially through space agencies (for example, NASA and the European Space Agency). Here, we highlight two key steps for enhancing openness and machine-driven integration of trait datasets.

Use standardized copyright waivers and licences. Waivers and licences support legal interoperability by clearly defining the conditions for both creation and use of combined or derivative data products, and allow users to legally access and use data without seeking additional authorization from the rights holders[71]. Many trait datasets do not yet use standardized copyright waiver or licence information such as those published through the Creative Commons (CC) framework[72]. In the context of EBVs, the formal designation with a CC0 copyright waiver or an open CC BY licence have been recommended because they minimize constraints on legal interoperability that emerge from restrictions on data use, modification and sharing[3]. Although a waiver of copyright through CC0 makes sharing and reuse much easier, the appropriate 'attribution' and maintenance of data provenance is important in a scientific context[18], and the CC BY licence provides the opportunity for acknowledgement and citation.

Provide standardized and machine-readable metadata. Many trait datasets are already available through web portals and other developed infrastructures (Supplementary Table 4), but access to standardized and machine-readable trait data and metadata remains a key bottleneck for technical and legal interoperability. For instance, licence and citation information is often not available in standardized and machine-readable form (for example, by using hyperlinks or embedded code, Supplementary Note 3) and many research projects publish their trait data on file hosting services (for example, Figshare, Dryad, Zenodo and so on) where no data and metadata standards are forced upon the uploaded material[27]. Moreover, metadata on the level of individual trait records is usually missing and data provenance is rarely documented (Supplementary Note 3). Hence, sufficient, consistent and well-documented metadata in a standardized form should be provided to successfully integrate trait measurements into workflows for building EBV data products of species traits.

A workflow for integrating EBV-relevant trait data

The production of species traits EBVs can only be achieved if multiple trait datasets are harmonized and combined into open, accessible and reusable products[3]. However, most trait data are currently stored in siloed resources and not available in an interoperable and machine-readable format. We therefore outline a generalized workflow for integrating EBV-relevant trait data (Fig. 3) and show how this workflow is currently applied to produce a new integrated plant phenology dataset (Box 3).

Collecting and provisioning trait data. The first part of the workflow represents the collection and initial processing of raw measurements of traits (for example, on flower and leaf phenology) following standardized sampling protocols, for example, by people (specimen collection and in situ observations) or close-range, airborne and spaceborne remote sensing (Fig. 3, top). After collection, raw data are validated through data quality assurance (QA, for example, by following standard protocols for trait data cleaning) and quality control (QC, for example, normalizing trait distributions, checking for outliers) (Fig. 3, top). Metadata about trait data collection and validation processes (for example, description of protocols) and about the dataset itself (for example, specimen IDs, ownership and licensing) need to be associated with the data when bundling the trait datasets (Box 3). Most currently existing trait datasets are only published in repositories with little metadata documentation and data standardization, but efforts to integrate them into more comprehensive data products are beginning to emerge.

Converting trait data into interoperable formats. To achieve integrated trait data products, data and metadata from different sources have to be standardized (Fig. 3, middle). This involves converting all data to comparable units and formats, the mapping of trait data to ontologies and automated reasoning over mapped data to discover new facts (Fig. 3, middle). The use of ontologies, for example, the Plant Phenology Ontology (PPO)[73] for flower and leaf phenology traits (Box 3), provides a formal, generalized, logical structure that helps to automate integration across different datasets. Ontologies can also be used to further improve quality of trait data integration through inferring new facts through machine reasoning (see Box 3 for examples). This process converts trait datasets into fully interoperable formats and enables future researchers as well as machines to interpret the data.

Providing integrated and reusable trait data products via web services. To make an integrated trait data product FAIR[18] (see above), a public domain designation (for example, CC0) or an open access licence (for example, CC BY) should be applied and provided together with other metadata in a machine-readable format (Fig. 3,

Box 3 | Example of a workflow integrating plant phenology data

The USA National Phenology Network (USA-NPN)[20] and the Pan-European Phenology Network (PEP725)[75] are two separate networks with differing protocols for capturing plant phenology traits (for example, timing of leafing, flowering and fruiting) at continental scales. The networks mobilize scientists and volunteers to collect data according to phenology trait or phase definitions. In addition, the National Ecological Observatory Network (NEON)[99] gathers trait measurements of many taxa (including leaf and flower phenology) across multiple field sites in the US. All three networks use data assurance and QC mechanisms, for example, constraining trait data entry to specific formats and including a set of consistency and completeness checks to ensure trait data quality. Their online portals provide bundled data and metadata on plant phenology, and the networks therefore follow typical workflow steps for collecting and provisioning species traits datasets (Fig. 3 top). However, the integration of plant phenology data products from these three sources is challenging because these networks use different frameworks.

As a response to the challenge of multiple frameworks, the PPO[73] was newly developed to standardize reporting from any in situ phenology resource, including professional and citizen science efforts such as USA-NPN and PEP725, more standardized surveys from NEON, and phenology data scored from herbarium records. The PPO defines a set of hierarchically organized 'phenological traits', that is, observable features of a plant that provide phenologically relevant information such as whether a plant has flowers, how many ripe fruits are on a plant, or whether a plant's leaves are senescing. Definitions of phenological traits therefore depend on classes for particular plant structures taken from the Plant Ontology[100]. Phenology terms from USA-NPN, PEP725, NEON and herbarium datasets have been mapped to the PPO, and plant phenology data can therefore be converted into a fully interoperable format through standardizing data and metadata (Fig. 3 middle). An added benefit of using ontologies is that automated procedures can produce new information from standardized data. For example, automated reasoning tools can use the PPO to infer that any plant that has open flower buds present must also have flowers and reproductive structures present.

To make integrated phenology trait data products accessible, a new web platform has been created (the Global Plant Phenology Data Portal, https://www.plantphenology.org/). Each individual phenology record is annotated to its source (for example, USA-NPN, PEP725 or NEON) and the licence of the source applied to the records. To allow efficient queries, harmonized data are processed using virtual machines run on CyVerse (formerly iPlant Collaborative)[90] and then loaded into Elasticsearch, a distributed, RESTful search and analytics engine (https://www.elastic.co/). This allows scalable searching of billions of trait data points that deliver outputs from standard queries very quickly. The backend is connected to an API which provides simple mechanisms for building front-end queries. Such a web platform allows open access to fine-resolution, population-level plant phenology data from different regions and continents (Fig. 3 bottom).

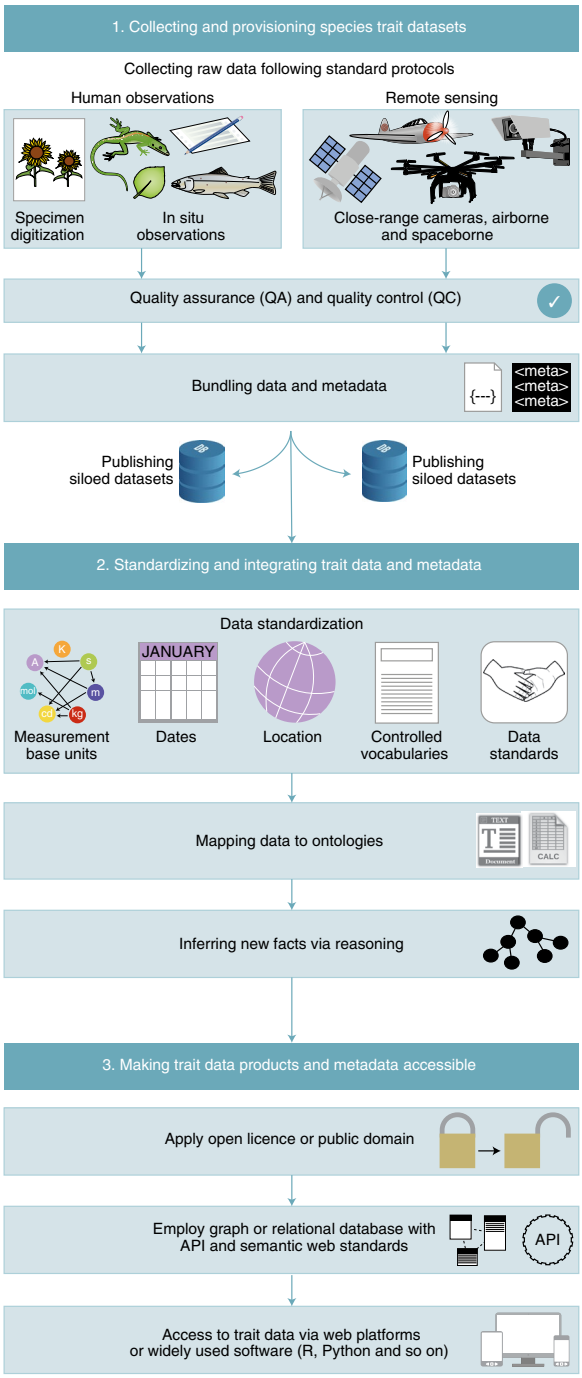


Fig. 3 | A generalized workflow for integrating species trait measurements into harmonized, open, accessible and reusable data products for EBVs. Initial species trait measurements are collected through human observations and remote sensing and subsequently quality checked and bundled into datasets (1). Because such datasets often have different sampling protocols, reporting processes and metadata descriptions, they commonly end up as siloed datasets in file hosting services with little metadata documentation and data standardization. To achieve integration of different measurements and data collections, datasets must be harmonized through standardization of data and metadata and mapped to community-developed standards, including metadata standards, controlled vocabularies and ontologies (2). Standardization often includes a second QA and QC process to assure data quality across datasets (not shown). Such harmonized data products can then be made accessible through open licences, databases that employ semantic web standards and APIs, and web platforms or widely used software (3).

bottom). In the best case, licence information should be available for each trait record and original source (Box 3). Further, it is important that data structures of trait data products align with semantic web standards (for example, multi-layered, relational databases rather than two-dimensional data tables). Hence, trait data products

should be housed with a graph database that allows on-the-fly reasoning via semantic queries, or with relational database if on-the-fly reasoning is not needed (Fig. 3, bottom). In both cases, an application programming interface (API) should allow communication and access to the trait data product via a web platform (Box 3) or via widely used software such as R or Python (Fig. 3, bottom).

Towards operationalizing species traits EBVs

Species traits are a key component of biodiversity, but species trait information is currently not well represented in indicators of biodiversity change used for national and international policy assessments[2,17,74]. The increasing willingness to share trait data in an open and machine-readable way (see Supplementary Note 3), coupled with emerging semantic tools (for example, new plant trait vocabularies[11], ontologies[64,73] and preliminary suggestions for trait data standards[27]) and a massive collection of trait data through in situ monitoring schemes and close-range sensors (for example, for phenology[20,39,47,75]) as well as on-going and forthcoming airborne and spaceborne missions (including radar, optical sensors, radiometers and spectrometers[42,43,50,53,76]), suggest that comprehensive data products on species traits are within reach in the near future. However, a cultural shift towards more openness, interoperability and reproducibility is needed within the broader science community[18,19,77] — including ecologists, biogeographers, global change biologists, biodiversity informaticians and Earth scientists — and with support from global coordinating institutions such as GEO BON, IPBES and the CBD.

Our refined list of species traits EBVs (Fig. 1) provides an improved conceptual framework for how phenological, morphological, reproductive, physiological and movement-related trait measurements can represent biodiversity in the EBV context and hence support international policies for biodiversity conservation and sustainable development. The specific species traits EBVs contain essential information with ecological, societal and policy relevance for biodiversity that cannot be substituted by other species traits EBVs (Supplementary Note 2). For instance, morphological and physiological measurements of leaves (for example, leaf area, nitrogen and chlorophyll content), stems (for example, height and stem density) and diaspores (for example, seed mass) allow quantification of fundamental dimensions of plant ecological strategies and how these organisms respond to competition, stress, environmental change and disturbances[8,12,43,50]. Phenological trait information of amphibians (spawning), birds (egg laying), plankton (population peaks), fish (spawning), insects (flight periods), mammals (birth dates) and plants (flowering, fruiting, leafing) is highly relevant for tracking changes in species' ecology in response to climate change[21] and other global changes (for example, nitrogen deposition inducing delayed foliar senescence). Morphological measurements (body sizes) of commercially relevant fish species[78–80] can allow assessments of sustainable food production and harvesting (Box 1). Similarly, morphological, reproductive and physiological traits of microbial species (for example, cell size, lifetime pattern of growth and microbial resistance to viruses) are essential for predicting their responses to environmental change[81]. A key aspect for the future operationalization of species traits EBVs is that they should be measurable with available technologies and have a proven track record of feasibility[6]. We suggest that a focus on trait measurements representing plant phenology, morphology and physiology (for example, from both in situ monitoring[20,39,47,75] and remote sensing[9,12,42,43,49,50,82]) as well as animal morphology[15,79] and movement[83] could provide a realistic prioritization for operationalizing species traits EBVs.

Compiling the necessary data for EBVs globally remains a major challenge, especially for species traits[7,17]. A key bottleneck is that the repeated and systematic collection of in situ trait data is not only costly and difficult but also spatially discontinuous. The global, spatially contiguous and periodic nature of spaceborne remote sensing observations therefore offers potential for building EBVs[82]. To date, spaceborne remote sensing products (for example, related to land surface phenology, canopy biochemistry and vegetation height) allow the mapping of ecosystem structure and processes as well as functional diversity[9,43,51,84], but not the quantification of species-level traits[1,82] because the spatial resolution is not fine enough to allow attribution of trait measurements to an individual or a population of a single species (Box 1). With airborne remote sensing it is possible to continuously map individual-level trait variation in morphological and physiological traits at fine (metre) resolution across regional scales (for example, forest trees[43]), often allowing assignment of trait measurements to the species level[85,86]. Since species-level resolution is required for many policy targets[76], assigning trait measurements to taxonomic information is key for monitoring intra-specific trait changes. A deeper integration of in situ and various close-range remote sensing trait measurements as well as a synergy of hyperspectral and LiDAR airborne remote sensing might help to achieve this. An avenue for building contiguous species traits EBVs could be to use information from Earth observation data for interpolating in situ trait point samples for building continuous landscape maps of trait distributions[76]. This would require the development of statistical and mechanistic models that allow mapping and prediction of trait distributions across space and time[87]. In this context, specimens from natural history collections could become useful for obtaining baseline trait data for regions that have been poorly studied[88].

Moving forward. Many dimensions of biodiversity still remain invisible when measuring and monitoring global biodiversity change[2,17,76]. Species traits EBVs will provide a deeper understanding of the species-level responses to global change and the benefits and services that individual species provide to humanity. For operationalizing species traits EBVs, we recommend the biodiversity research community to support trait data harmonization, reproducible workflows, interoperability and 'big data' biodiversity informatics for species traits[19,23,27,89,90]. Specifically, we suggest the following concrete steps to facilitate the building of EBV data products of species traits:

- Support the recording of species traits across time through repeated and periodic collection of in situ measurements of traits, through digitization of trait information from literature and biocollections and through developing species traits data products from close-range, airborne and spaceborne remote sensing observations.
- Develop and apply standardized protocols, controlled trait vocabularies and trait data standards when measuring, harmonizing and combining trait data and metadata.
- Support the semantic integration of trait data by mapping trait datasets to ontologies, facilitate training courses about semantic standards of the World Wide Web Consortium (W3C) and promote training tools for trait data integration within research institutions and educational programmes of universities.
- Publish trait databases with standardized licence information in machine-readable form and designate data as open access (for example, through CC BY) or in the public domain (for example, CC0). Encourage others to share trait data.
- Develop and apply reproducible statistical and mechanistic models for integrating in situ trait data with remote sensing observations to allow mapping and prediction of trait distributions across space and time.
- Establish consortia and interest groups on species traits. Contribute to the GEO BON working group on species traits and raise awareness of the need for semantic, technical and legal interoperability of trait data.
- Foster the integration of species traits EBVs into biodiversity indicators and biodiversity and sustainability goals.

These activities — which require substantial financial and in kind investments from universities, research infrastructures, governments, space agencies and other bodies — will facilitate the building of global EBV data products of species traits and allow significant steps towards incorporating intra-specific trait variability into global, regional and national biodiversity and policy assessments.

Received: 25 February 2018; Accepted: 16 July 2018;
Published online: 17 September 2018

References

1. Pereira, H. M. et al. Essential Biodiversity Variables. *Science* **339**, 277–278 (2013).
2. Geijzendorffer, I. R. et al. Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *J. Appl. Ecol.* **53**, 1341–1350 (2016).
3. Kissling, W. D. et al. Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev.* **93**, 600–625 (2018).
4. Giuliani, G. et al. Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data* **1**, 100–117 (2017).
5. Schmeller, D. S. et al. An operational definition of essential biodiversity variables. *Biodivers. Conserv.* **26**, 2967–2972 (2017).
6. Walters, M. et al. *Essential Biodiversity Variables* (Convention on Biological Diversity, 2013).
7. *GEO BON Implementation Plan 2017–2020 Version 1.3* (GEO BON Secretariat, 2017).
8. Lavorel, S. & Garnier, E. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Funct. Ecol.* **16**, 545–556 (2002).
9. Lausch, A. et al. Linking Earth observation and taxonomic, structural and functional biodiversity: local to ecosystem perspectives. *Ecol. Indic.* **70**, 317–339 (2016).
10. Gibert, J. P., Dell, A. I., DeLong, J. P. & Pawar, S. in *Advances in Ecological Research* Vol. 52 (eds Woodward, G., Pawar, S. & Dell, A. I.) 1–17 (Academic, Waltham, 2015).
11. Garnier, E. et al. Towards a thesaurus of plant characteristics: an ecological contribution. *J. Ecol.* **105**, 298–309 (2017).
12. Díaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
13. Eskildsen, A. et al. Ecological specialization matters: long-term trends in butterfly species richness and assemblage composition depend on multiple functional traits. *Divers. Distrib.* **21**, 792–802 (2015).
14. Hagen, M. et al. Biodiversity, species interactions and ecological networks in a fragmented world. *Adv. Ecol. Res.* **46**, 89–210 (2012).
15. Genner, M. J. et al. Body size-dependent responses of a marine fish assemblage to climate change and fishing over a century-long scale. *Glob. Change Biol.* **16**, 517–527 (2010).
16. Wilson, J. R. U., Dormontt, E. E., Prentis, P. J., Lowe, A. J. & Richardson, D. M. Something in the way you move: dispersal pathways affect invasion success. *Trends Ecol. Evol.* **24**, 136–144 (2009).
17. Navarro, L. M. et al. Monitoring biodiversity change through effective global coordination. *Curr. Opin. Environ. Sustain.* **29**, 158–169 (2017).
18. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
19. Kissling, W. D. et al. Towards global interoperability for supporting biodiversity research on essential biodiversity variables (EBVs). *Biodiversity* **16**, 99–107 (2015).
20. Schwartz, M. D., Betancourt, J. L. & Weltzin, J. F. From Caprio's lilacs to the USA National Phenology Network. *Front. Ecol. Environ.* **10**, 324–327 (2012).
21. Thackeray, S. J. et al. Phenological sensitivity to climate across taxa and trophic levels. *Nature* **535**, 241–245 (2016).
22. Kattge, J. et al. TRY – a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935 (2011).
23. Parr, C. S. et al. TraitBank: practical semantics for organism attribute data. *Semant. Web* **7**, 577–588 (2016).
24. Calewaert, J.-B., Weaver, P., Gunn, V., Gorringe, P. & Novellino, A. in *Quantitative Monitoring of the Underwater Environment* (eds Zerr, B. et al.) 31–46 (Springer, Cham, 2016).
25. Guralnick, R. P. et al. The importance of digitized biocollections as a source of trait data and a new VertNet resource. *Database* **2016**, baw158 (2016).
26. MarLIN *BIOTIC: Biological Traits Information Catalogue* (Marine Biological Association of the United Kingdom, 2006).
27. Schneider, F. D. et al. Towards an ecological trait-data standard. Preprint at https://www.biorxiv.org/content/early/2018/05/31/328302 (2018).
28. Kissling, W. D. et al. Establishing macroecological trait datasets: digitalization, extrapolation, and validation of diet preferences in terrestrial mammals worldwide. *Ecol. Evol.* **4**, 2913–2930 (2014).
29. Holmes, M. W. et al. Natural history collections as windows on evolutionary processes. *Mol. Ecol.* **25**, 864–881 (2016).
30. Gardner, J. L. et al. Temporal patterns of avian body size reflect linear size responses to broadscale environmental change over the last 50 years. *J. Avian Biol.* **45**, 529–535 (2014).
31. Tseng, M. et al. Decreases in beetle body size linked to climate change and warming temperatures. *J. Anim. Ecol.* **87**, 647–659 (2018).
32. Miller-Rushing, A. J., Primack, R. B., Primack, D. & Mukunda, S. Photographs and herbarium specimens as tools to document phenological changes in response to global warming. *Am. J. Bot.* **93**, 1667–1674 (2006).
33. Robbirt, K. M., Davy, A. J., Hutchings, M. J. & Roberts, D. L. Validation of biological collections as a source of phenological data for use in climate change studies: a case study with the orchid *Ophrys sphegodes*. *J. Ecol.* **99**, 235–241 (2011).
34. Willis, C. G. et al. Old plants, new tricks: phenological research using herbarium specimens. *Trends Ecol. Evol.* **32**, 531–546 (2017).
35. Kim, E.-S. Development, potentials, and challenges of the International Long-Term Ecological Research (ILTER) Network. *Ecol. Res.* **21**, 788–793 (2006).
36. Gosz, J. R., Waide, R. B. & Magnuson, J. J. in *Long-Term Ecological Research: Between Theory and Application* (eds Müller, F., Baessler, C., Schubert, H. & Klotz, S.) 59–74 (Springer, Dordrecht, 2010).
37. Silvertown, J. A new dawn for citizen science. *Trends Ecol. Evol.* **24**, 467–471 (2009).
38. Jackson, M. C. et al. in *Advances in Ecological Research* Vol. 55 (eds Dumbrell, A. J., Kordas, R. L. & Woodward, G.) 615–636 (Academic, Waltham, 2016).
39. Elmendorf, S. C. et al. The plant phenology monitoring design for the National Ecological Observatory Network. *Ecosphere* **7**, e01303 (2016).
40. Metzger, M. J. et al. Environmental stratifications as the basis for national, European and global ecological monitoring. *Ecol. Indic.* **33**, 26–35 (2013).
41. Pettorelli, N. et al. Framing the concept of satellite remote sensing essential biodiversity variables: challenges and future directions. *Remote Sens. Ecol. Conserv.* **2**, 122–131 (2016).
42. Skidmore, A. K. et al. Agree on biodiversity metrics to track from space. *Nature* **523**, 403–405 (2015).
43. Schneider, F. D. et al. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nat. Commun.* **8**, 1441 (2017).
44. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
45. Zhu, X. et al. Canopy leaf water content estimated using terrestrial LiDAR. *Agr. Forest Meteorol.* **232**, 152–162 (2017).
46. Richardson, A. D., Klosterman, S. & Toomey, M. in *Phenology: An Integrative Environmental Science* (ed Schwartz, M. D.) 413–430 (Springer, Dordrecht, 2013).
47. Nasahara, K. N. & Nagai, S. Development of an in situ observation network for terrestrial ecological remote sensing: the Phenological Eyes Network (PEN). *Ecol. Res.* **30**, 211–223 (2015).
48. Schweiger, A. K. et al. Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. *Nat. Ecol. Evol.* **2**, 976–982 (2018).
49. Kampe, T. U., Johnson, B. R., Kuester, M. & Keller, M. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *J. Appl. Remote Sens.* **4**, 043510 (2010).
50. Asner, G. P. et al. Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science* **355**, 385–389 (2017).
51. Jetz, W. et al. Monitoring plant functional diversity from space. *Nat. Plants* **2**, 16024 (2016).
52. Schimel, D. S., Asner, G. P. & Moorcroft, P. Observing changing ecological diversity in the Anthropocene. *Front. Ecol. Environ.* **11**, 129–137 (2013).
53. Stavros, E. N. et al. ISS observations offer insights into plant function. *Nat. Ecol. Evol.* **1**, 0194 (2017).
54. Asner, G. P. et al. Amazonian functional diversity from forest canopy chemical assembly. *Proc. Natl Acad. Sci. USA* **111**, 5604–5609 (2014).
55. Asner, G. P. & Martin, R. E. Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Front. Ecol. Environ.* **7**, 269–276 (2009).
56. Cornelissen, J. H. C. et al. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust. J. Bot.* **51**, 335–380 (2003).
57. Pérez-Harguindeguy, N. et al. New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **61**, 167–234 (2013).

58. Moretti, M. et al. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Funct. Ecol.* **31**, 558–567 (2017).
59. Salgado-Negret, B. *La Ecología Funcional Como Aproximación al Estudio, Manejo y Conservación de la Biodiversidad: Protocolos y Aplicaciones* (Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, 2015).
60. Costello, M. J. et al. Biological and ecological traits of marine species. *PeerJ* **3**, e1201 (2015).
61. Park, C. A. et al. The Vertebrate Trait Ontology: a controlled vocabulary for the annotation of trait data across species. *J. Biomed. Semant.* **4**, 13 (2013).
62. Baird, D. J. et al. Toward a knowledge infrastructure for traits-based ecological risk assessment. *Integr. Environ. Assess. Manag.* **7**, 209–215 (2011).
63. Mungall, C. J. et al. Integrating phenotype ontologies across multiple species. *Genome Biol.* **11**, R2 (2010).
64. Hoehndorf, R. et al. The flora phenotype ontology (FLOPO): tool for integrating morphological traits and phenotypes of vascular plants. *J. Biomed. Semant.* **7**, 65 (2016).
65. Walls, R. L. et al. Ontologies as integrative tools for plant science. *Am. J. Bot.* **99**, 1263–1275 (2012).
66. Hughes, L. M., Bao, J., Hu, Z.-L., Honavar, V. & Reecy, J. M. Animal trait ontology: the importance and usefulness of a unified trait vocabulary for animal species. *J. Anim. Sci.* **86**, 1485–1491 (2008).
67. Kattge, J. et al. A generic structure for plant trait databases. *Methods Ecol. Evol.* **2**, 202–213 (2011).
68. Costello, M. J., Michener, W. K., Gahegan, M., Zhang, Z.-Q. & Bourne, P. E. Biodiversity data should be published, cited, and peer reviewed. *Trends Ecol. Evol.* **28**, 454–461 (2013).
69. Penev, L. et al. Strategies and guidelines for scholarly publishing of biodiversity data. *Res. Ideas Outcomes* **3**, e12431 (2017).
70. Michener, W. K. Ecological data sharing. *Ecol. Inform.* **29**, 33–44 (2015).
71. RDA-CODATA Legal Interoperability Interest Group *Legal Interoperability of Research Data: Principles and Implementation Guidelines* (Research Data Alliance, 2016).
72. Carroll, M. W. Creative commons and the new intermediaries. *Michigan State Law Rev.* **2006**, 45–65 (2006).
73. Stucky, B. J. et al. The Plant Phenology Ontology: a new informatics resource for large-scale integration of plant phenology data. *Front. Plant Sci.* **9**, 517 (2018).
74. *Global Biodiversity Change Indicators: Model-Based Integration of Remote-Sensing and In Situ Observations that Enables Dynamic Updates and Transparency at Low Cost* (GEO BON Secretariat, 2015).
75. van Vliet, A. J. H. et al. The European Phenology Network. *Int. J. Biometeorol.* **47**, 202–212 (2003).
76. Bush, A. et al. Connecting Earth observation to high-throughput biodiversity data. *Nat. Ecol. Evol.* **1**, 0176 (2017).
77. Lowndes, J. S. S. et al. Our path to better science in less time using open data science tools. *Nat. Ecol. Evol.* **1**, 0160 (2017).
78. Worm, B. et al. Rebuilding global fisheries. *Science* **325**, 578–585 (2009).
79. Coll, M. et al. Ecological indicators to capture the effects of fishing on biodiversity and conservation status of marine ecosystems. *Ecol. Indic.* **60**, 947–962 (2016).
80. Welcomme, R. L. et al. Inland capture fisheries. *Phil. Trans. R. Soc. B* **365**, 2881–2896 (2010).
81. Krause, S. et al. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front. Microbiol.* **5**, 251 (2014).
82. Skidmore, A. K. et al. *Remote Sensing Enabled Essential Biodiversity Variables* (Convention on Biological Diversity, 2018).
83. Kays, R., Crofoot, M. C., Jetz, W. & Wikelski, M. Terrestrial animal tracking as an eye on life and planet. *Science* **348**, aaa2478 (2015).
84. Houborg, R., Fisher, J. B. & Skidmore, A. K. Advances in remote sensing of vegetation function and traits. *Int. J. Appl. Earth Obs. Geoinform.* **43**, 1–6 (2015).
85. Asner, G. P. et al. Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sens. Environ.* **112**, 1942–1955 (2008).
86. Skidmore, A. K. et al. Forage quality of savannas — Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sens. Environ.* **114**, 64–72 (2010).
87. Butler, E. E. et al. Mapping local and global variability in plant trait distributions. *Proc. Natl Acad. Sci. USA* **114**, E10937–E10946 (2017).
88. Greve, M. et al. Realising the potential of herbarium records for conservation biology. *S. Afr. J. Bot.* **105**, 317–323 (2016).
89. La Salle, J., Williams, K. J. & Moritz, C. Biodiversity analysis in the digital era. *Phil. Trans. R. Soc. B* **371**, 20150337 (2016).
90. Merchant, N. et al. The iPlant collaborative: cyberinfrastructure for enabling data to discovery for the life sciences. *PLoS Biol.* **14**, e1002342 (2016).
91. Violle, C. et al. Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).
92. Lausch, A., Erasmi, S., King, D., Magdon, P. & Heurich, M. Understanding forest health with remote sensing -part I—a review of spectral traits, processes and remote-sensing characteristics. *Remote Sens.* **8**, 1029 (2016).
93. Galetti, M. et al. Functional extinction of birds drives rapid evolutionary changes in seed size. *Science* **340**, 1086–1090 (2013).
94. Madin, J. S., Bowers, S., Schildhauer, M. P. & Jones, M. B. Advancing ecological research with ontologies. *Trends Ecol. Evol.* **23**, 159–168 (2008).
95. 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).
96. Bruskiewich, R. M. et al. Linking genotype to phenotype: the International Rice Information System (IRIS). *Bioinformatics* **19**, i63–i65 (2003).
97. Smith, C. L. & Eppig, J. T. The mammalian phenotype ontology: enabling robust annotation and comparative analysis. *Wiley Interdisc. Rev. Syst. Biol. Med.* **1**, 390–399 (2009).
98. Buttigieg, P. L. et al. The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. *J. Biomed. Semant.* **7**, 1–12 (2016).
99. Schimel, D., Hargrove, W., Hoffman, F. & MacMahon, J. NEON: a hierarchically designed national ecological network. *Front. Ecol. Environ.* **5**, 59 (2007).
100. Cooper, L. & Jaiswal, P. in *Plant Bioinformatics: Methods and Protocols* (ed. Edwards, D.) 89–114 (Springer, New York, 2016).

Acknowledgements

This paper emerged from a workshop of the Horizon 2020 project GLOBIS-B (Global Infrastructures for Supporting Biodiversity research; http://www.globisb.eu/). Financial support came from the European Commission (grant 654003). We thank J. Konijn for administrative support and the members of the BIOMAC lab (https://www.biomac.org/) for discussion. L.M.N. is supported by the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118). N.R. was funded by a research grant from Deutsche Forschungsgemeinschaft DFG (RU 1536/3-1).

Author contributions

W.D.K. coordinated the study and wrote the draft manuscripts. R.W., A.B., M.O.J., J.K., R.P.G. and W.D.K. led the writing of sections. All authors provided substantial input into ideas and text, and commented on draft manuscripts.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0667-3.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to W.D.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.