

Standards

Improving Darwin Core for research and management of alien species

Quentin Groom[‡], Peter Desmet[§], Lien Reyserhove[§], Tim Adriaens[§], Damiano Oldoni[§], Sonia Vanderhoeven[‡], Steven J Baskauf[¶], Arthur Chapman[#], Melodie McGeoch[□], Ramona Walls[«], John Wieczorek[»], John R.U. Wilson[^], Paula F F Zermoglio[‡], Annie Simpson[?]

[‡] Meise Botanic Garden, Meise, Belgium

[§] Research Institute for Nature and Forest (INBO), Brussels, Belgium

[|] Belgian Biodiversity Platform, Brussels, Belgium

[¶] Vanderbilt University, Nashville, Tennessee, United States of America

[#] Australian Biodiversity Information Services, Ballan, Australia

[□] Monash University, School of Biological Sciences, Clayton, Australia

[«] Bio5 Institute and CyVerse, University of Arizona, Tucson, United States of America

[»] Museum of Vertebrate Zoology, University of California, Berkeley, United States of America

[^] South African National Biodiversity Institute, Kirstenbosch, South Africa

[‡] Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Stellenbosch, South Africa

[‡] Instituto de Ecología, Genética y Evolución de Buenos Aires (IEGEB-CONICET), University of Buenos Aires, Buenos Aires, Argentina

[?] US Geological Survey, Reston, United States of America

Corresponding author: Quentin Groom (quentin.groom@plantentuinmeise.be)

Academic editor: Gail Kampmeier

Received: 08 Jul 2019 | Accepted: 28 Sep 2019 | Published: 10 Oct 2019

Citation: Groom Q, Desmet P, Reyserhove L, Adriaens T, Oldoni D, Vanderhoeven S, Baskauf SJ, Chapman A, McGeoch M, Walls R, Wieczorek J, Wilson JR.U, Zermoglio PFF, Simpson A (2019) Improving Darwin Core for research and management of alien species. Biodiversity Information Science and Standards 3: e38084.

<https://doi.org/10.3897/biss.3.38084>

Abstract

To improve the suitability of the Darwin Core standard for the research and management of alien species, the standard needs to express the native status of organisms, how well established they are and how they came to occupy a location. To facilitate this, we propose:

1. To adopt a controlled vocabulary for the existing Darwin Core term *dwc:establishmentMeans*
2. To elevate the pathway term from the Invasive Species Pathways extension to become a new Darwin Core term *dwc:pathway* maintained as part of the Darwin Core standard
3. To adopt a new Darwin Core term *dwc:degreeOfEstablishment* with an associated controlled vocabulary

These changes to the standard will allow users to clearly state whether an occurrence of a species is native to a location or not, how it got there (pathway), and to what extent the species has become a permanent feature of the location. By improving Darwin Core for capturing and sharing these data, we aim to improve the quality of occurrence and checklist data in general and to increase the number of potential uses of these data.

Keywords

establishment means, invasive species, non-native, biodiversity, data standards, Essential Biodiversity Variables, invasion pathway, invasion stage

Context

To improve the management and reduce the spread of alien species, data are needed on an ongoing basis on the current occurrences of those species, their statuses, how they are spreading and where they originated (McGeoch et al. 2016, Wilson et al. 2018). Data-driven exercises, such as horizon scanning, early warning systems and impact assessment, should be conducted regularly, as part of routine monitoring (Latombe et al. 2017, Ricciardi et al. 2017). These activities provide policy-makers and other decision-makers with evidenced-based information. Horizon scanning provides a broad systematic examination of potential threats (Sutherland and Woodroof 2009); early warning systems facilitate a rapid response to invasion (Katsanevakis et al. 2015) and impact assessment is typically conducted to prioritize control and prevention (Turbé et al. 2017). Still, data are collected and maintained by a wide variety of people and organizations. These data sources are often segregated taxonomically by habitat, methods, date and geography. As a result, alien species data are captured and shared using a variety of distinct data structures and values, which makes combining these data sources time-consuming and prone to information loss or misinterpretation. Manual intervention is required to transform data to a single format. In the process, information is often lost and the meanings of standard terms are distorted or broadened to make data acceptable. In the Biodiversity Information Standards (TDWG) Questions & Answers Site for Darwin Core (dwc) (<https://github.com/tdwg/dwc-qa/tree/master/data>), one can see many examples of the wide variety of values and formats for data, such as for *dwc:establishmentMeans* (Suppl. material 1).

Occurrences of biodiversity, including alien species, are primarily communicated using the Darwin Core (dwc) standard, notably by the [Global Biodiversity Information Facility](#) (GBIF). Darwin Core standard is a collection of terms and definitions that describe taxa and their occurrence in nature (Wieczorek et al. 2012). A Darwin Core Archive (DwC-A) is a self-contained dataset of one or more delimited text files where the rows are records and the columns are defined by Darwin Core terms (Remsen et al. 2010). An Archive might also contain an XML metadata file that describes the contents. Darwin Core Archives generally have a central file with the core elements of the record, but may also contain extension files linked by a unique record identifier. In this way many additional types of data can be linked to the record and, if necessary, these extensions can have a one-to-many relationship with the core data.

If alien species monitoring and research are to be made routine and reliable then data collection needs to be standardized and data handling and aggregation must be automated. Therefore, standards and formats need to converge to capture relevant information and simplify this process, or, at least, there should be an overall framework onto which the current multitude of structures and values can be mapped.

Improved data interoperability would accelerate the process of biodiversity monitoring, reduce the time to produce actionable evidence, and also reduce the costs. In addition to monitoring invasive species, similar situations exist in, for example, the assessment of conservation status (Rodrigues et al. 2006) and the monitoring of wild game animals. Such information is also crucial for large-scale biogeographic or ecological studies that assess patterns of species' distributions or relationships to climate, as these studies often assume biogeographic origins from current distributions. In all these cases, the lack of machine-readable resources and inadequate standards prevent the automation of research and monitoring processes.

Basic pieces of information are required for risk assessment, horizon scanning, species management and monitoring. In previous work, we identified four species properties that are needed. These are the introduction pathway, the degree of establishment, the species status and the impact mechanism (Groom et al. 2017a, Groom et al. 2017b). Similarly, Latombe et al. (2017) identified three "Essential Variables For Invasion Monitoring" that they determined are critical to slowing the spread of alien species and reducing their negative impacts. These essential variables were alien species occurrence, species' status as an alien and alien species impact. They also identified four supplementary variables, which included the pathways of introduction and spread. Wilson et al. (2018) expanded on this with 20 indicators designed to monitor biological invasions at a national level, in particular by considering indicators that track the effectiveness of interventions.

Although impact mechanism was identified as important in all these studies, it is not treated here, because it is derived information about many aspects of the organism's biology and thus not generally included in original occurrence records. Therefore, we focus on the introduction pathway, the degree of establishment and the species status. We also divide species status into two concepts, firstly whether the taxon is present or absent and secondly whether the taxon is native or alien (non-native). It should also be noted that the

term "invasive species" is a source of confusion. In the biological sense, it refers to any species that is rapidly extending its range. However, its definition from a political perspective, notably in the Convention on Biological Diversity, restricts the term to those alien species that may have a negative impact (Secretariat of the Convention on Biological Diversity 2009). It is, for example, used in this sense in the Global Register of Introduced and Invasive Species (Pagad et al. 2018). In practice, the distinction may be difficult to make, as impact must often be assumed from presence of the organism. For information on classifying species impacts, readers are referred to Blackburn et al. (2014).

A recurrent issue when considering alien species data types is their scope. An invasion is ultimately a population-level phenomenon. A species can be classified as introduced to a particular region only if individuals have been brought in and are present outside of their native range. If such individuals reproduce and spread, then the population (or populations) in that locality may be considered "invasive". This means that from the perspective of a particular country, there might be both alien and native populations of a species present. Furthermore, this issue of scope also pertains to time, as populations can expand, shrink, become extinct and be reintroduced at different periods.

For any given place and period of time, we need basic information to answer at least the following four questions (cf. Essl et al. 2018):

1. **Within the area and time period in question, does the organism live there?** Is the organism present in or absent from an area and over what time period?
2. **Is the species native or alien?** Definitions of what is native vary depending on historical circumstances. Decisions on what is native and what is not, can be as much political as they are scientific. Science informs decision-makers about the history of a taxon in a region, but cannot make a decision on where the cut-off dates should be. Nonetheless, we value a region's native organisms because they provide a unique character to different areas and habitats. So we need information on the native status of an organism in an area to make conservation assessments and direct invasive species policy.
3. **How well established is the organism in that location?** Their degree of establishment in a particular location ranges from those that are temporary visitors, such as migrating birds resting en route to their summer or winter range, to those whose continuing presence is dependent on human assistance, to well established invasive species.
4. **How did the organism get to that location?** Invasion biologists refer to this as the "pathway" of introduction. Understanding and managing introduction pathways are important if the introduction of alien species is to be limited to those that pose an acceptable risk of invasion. Introduction pathways are the focus of target 9 of the Aichi Biodiversity Targets (Ad Hoc Open-ended Working Group on Review of Implementation of the Convention 2011).

Information to answer these questions is frequently collected in species checklists and occurrence observations datasets, and published to [GBIF](#) using appropriate standard terms in the Darwin Core (Wieczorek et al. 2012, Parr et al. 2012). However, when trying to create and use data formatted in Darwin Core, it is difficult to express information to answer these four questions. The pertinent terms that already exist in Darwin Core (`dwc:establishmentMeans` and `dwc:occurrenceStatus`) are not always sufficient to capture the needed information. There are other questions we may ask, particularly those related to the impact of the species on other organisms, but these four basic questions are the foundation upon which other questions rest, and we intend to return to impact-relevant data elements at a later stage. However, it is currently difficult to communicate all these concepts within Darwin Core, either due to a lack of terms or to the lack of clear advice on suitable controlled vocabularies.

Many Darwin Core terms were created to describe the details of biological specimens (e.g. `dwc:sex`). Specimens frequently consist of all or part of a single organism from a single location on a single collection event, sometimes referred to as a gathering. Darwin Core terms usually also perform well when applied to field observations, though the application of certain terms becomes more difficult, when field observations and some specimens consist of multiple individuals. In recent years, Darwin Core has become more frequently used for ecological survey data (Wieczorek et al. 2014, Guralnick et al. 2017) and checklists (Remsen et al. 2012). This can put strain on the definitions of Darwin Core terms because the scope of the term may range from individuals through whole populations as captured in the concept of `dwc:Organism`. When working on the terms in this document we have tried to consider whether these terms are applicable across such a broad scope. In the current context of Darwin Core Archive, they are intended to be used as terms for Occurrence records (http://rs.gbif.org/core/dwc_occurrence_2015-07-02.xml: a `dwc:Organism` at a place and time) and Species Distribution extension records (<http://rs.gbif.org/extension/gbif/1.0/distribution.xml>: a `dwc:Taxon` at a place and time).

Darwin Core provides the essential elements of an observation, however there are several extensions that have been created to expand the data that can be incorporated (e.g. Endresen and Knüpffer 2012). These extensions are not part of the standard itself, but provide a means to accommodate the needs of specialist communities. In this paper, we propose changes to the central Darwin Core standard to resolve some of the problems mentioned above. Darwin Core is maintained by the [TDWG](#) organization. Any changes will be handled according to Section 3.3 of the TDWG Vocabulary Maintenance Specification (VMS) (<http://hdl.handle.net/1803/9512>) (Baskauf et al. 2017a) and having metadata according to Section 4.5.4 of the Standards Documentation Specification (Baskauf et al. 2017b). Each controlled value term will be identified by a Uniform Resource Identifier (URI) and have an associated controlled value string for use in spreadsheets or text tables. Publication of our proposals here is a step in the process of getting these changes to the standard adopted.

Current terms, proposed changes and a new term

In the following section, details of the proposed changes to Darwin Core are explained.

dwc:establishmentMeans

Current dwc:establishmentMeans

Currently, dwc:establishmentMeans is defined in the Darwin Core documentation as “*The process by which the biological individual(s) represented in the Occurrence became established at the location.*” (Biodiversity Information Standards (TDWG) 2018).

The vocabulary recommended by GBIF for dwc:establishmentMeans includes the categories and subcategories in Table 1, column 1. The term establishmentMeans and its definition give the impression, at least to an invasion biologist, that the data concern the introduction pathway. That is, the means by which invasive species are moved, intentionally or unintentionally, into new areas, such as a horticultural escape. However, the examples given for the field in Biodiversity Information Standards (TDWG) (2018) make it clear that this is not the case. The recommended vocabulary answers the question of whether a species is native or alien, but also conflates this with how well established an organism is, by including subclasses such as invasive.

Table 1.

A proposed controlled vocabulary for dwc:establishmentMeans based on the vocabularies used by [GBIF](#) and the International Union for Conservation of Nature ([IUCN](#)) to express whether a species is native or alien. Hierarchical levels are indicated with colons, synonyms are in parentheses. Appropriate URIs will be assigned upon adoption of the controlled vocabulary.

GBIF establishmentMeans	IUCN origin	Proposed human readable label for establishmentMeans	Proposed controlled value string for establishmentMeans
native (indigenous, reintroduced)	native	native (indigenous)	native
	reintroduced	native: reintroduced	nativeReintroduced
introduced (exotic, alien)	introduced	introduced (alien, exotic, non-native, nonindigenous)	introduced
introduced: naturalised			
introduced: invasive			
introduced: managed (cultivative, captive)			
	assisted colonisation	introduced: assisted colonisation	introducedAssistedColonisation
	vagrant	vagrant (casual)	vagrant
uncertain (unknown)	origin uncertain	uncertain (unknown, cryptogenic)	uncertain

Unlike many fields in Darwin Core, GBIF encourages conformity in the field establishmentMeans by flagging records as "distribution invalid" if the value is not in the GBIF vocabulary for this term. GBIF also uses a lookup dictionary to interpret some unambiguous values for values found in the vocabulary (Suppl. material 2).

The term `dwc:establishmentMeans` is well entrenched in the biodiversity informatics community and is widely used and validated (e.g. Aedo and Pando 2017, Marchand et al. 2017). The term, to some extent, answers the question of whether an occurrence is native or alien, but it does lack the necessary nuance, for example it lacks the ability to communicate that a species was reintroduced, as a subcategory of native. A similar vocabulary is used by the International Union for Conservation of Nature (IUCN) under the name "*origin*" (Table 1, column 2) (IUCN 2018). The IUCN additionally includes the classes "reintroduced", "vagrant" and "assisted colonisation". The "reintroduced" class could be considered a subclass of native and assisted colonisation as a subclass of introduced. Vagrant is a term used for natural occurrences of organisms outside their normal ranges and also for human-aided introductions where the degree of establishment is minimal. The "introduced" subclasses "naturalised", "invasive" and "managed" are deprecated and we recommend expressing this information in the new term `dwc:degreeOfEstablishment`, respectively as "established", "invasive" and "cultivated". Naturalised is currently used for introduced organisms that are established. Invasive species are often a subset of naturalised species (those that have spread from their point of introduction), but in some cases naturalised has been reserved for situations where the degree of invasiveness is either minimal or undefined.

Proposed changes to `dwc:establishmentMeans`

As `dwc:establishmentMeans` and its vocabulary are frequently used, deprecating it would either result in confusion or be ignored by the community. A more helpful approach is to maintain backward compatibility of the use of `dwc:establishmentMeans`, while augmenting the vocabulary with additional terms, deprecating redundant terms and providing an additional Darwin Core term to express the degree to which a taxon is established. Preexisting data in GBIF with an `establishmentMeans` of "naturalised", "invasive" or "managed" could be mapped to the term proposed below, `degreeOfEstablishment`.

A refined definition of `dwc:establishmentMeans`:

"A statement about whether an organism or organisms have been introduced to a given place and time through the direct or indirect activity of modern humans."

The concept of nativeness is fluid and depends upon the temporal, taxonomic and geographic perspective. We refer to modern humans here to avoid defining nativeness within the definition of `dwc:establishmentMeans`, but also to acknowledge that these terms refer to comparatively recent biogeographic changes.

dwc:occurrenceStatus

The dwc:occurrenceStatus is defined in the Darwin Core standard as “*A statement about the presence or absence of a Taxon at a Location*” (Biodiversity Information Standards (TDWG) 2018).

This term helps us answer our question as to whether an organism occurs in a defined location and time frame. To express the absence of an dwc:Organism, dwc:occurrenceStatus should only be used where there are defined temporal and spatial boundaries. An assertion of absence has no meaning or use for specimens or point observations where presence is explicit (MacKenzie et al. 2002). Point observations have no spatial boundaries, even if the observer has provided a measurement of precision or uncertainty of the coordinates (dwc:coordinatePrecision, dwc:coordinateUncertaintyInMeters). For distribution modelling, the first step is to assign point presence observations to a grid to give them spatial boundaries. To grid presence data, a variety of assumptions are made about the accuracy of point coordinates, but these assumptions do not hold for point absences.

Nevertheless, presence and absence are particularly useful when bounded by a time period and location. As absence can never be proven, it can only ever be derived from a reasoned analysis of the evidence, and this has to be bounded. Darwin Core terms suitable for establishing these limits are found under categories Event (e.g. dwc:eventDate) and Location (e.g. dwc:country).

dwc:occurrenceStatus is a useful term because combined with dwc:establishmentMeans, dwc:occurrenceStatus allows the user to express whether an organism is native or alien to an area and whether it still exists there. Yet currently, dwc:occurrenceStatus is not universally used on GBIF or it is mistakenly used to express different types of information, such as the breeding status of birds or the IUCN threat status of the organism. For breeding status, the term dwc:reproductiveCondition is more appropriate, and for threat status the term "threatStatus" is available in the Species Distribution extension (<http://rs.gbif.org/extension/gbif/1.0/distribution.xml>). Darwin Core extensions have been created to provide additional functionality for specific communities and to allow more experimentation with terms outside the formal governance of the standard (Wieczorek et al. 2012). The diverse and disjunct content of data labeled dwc:occurrenceStatus indicates the need for these proposed terms, and for improved guidance in their documentation. We propose adding notes to the documentation of dwc:occurrenceStatus, to point users to other status fields that might be appropriate for their needs.

dwc:pathway

Pathways are the means by which invasive species surmount the biogeographic barriers to dispersal and are introduced into new places. Some of these pathways are literal pathways to introduction, such as waterways and bridges, while others are figurative pathways, such as agricultural and trading practices. It is also worth noting that multiple alien and native species are dispersed through individual pathways, though this is less evident in the case

of native species where individuals arrive at a destination where their species is already present. Even if a species has already established, policies to eliminate its pathway stop other species from using the same route to introduction. Therefore, improved information on introduction pathway informs policy on trade, agriculture and environmental management (Leung et al. 2014, Keller et al. 2011).

Current dwc:pathway

The species introduction term "pathway" is only available through the Invasive Species Pathways extension to Darwin Core (<http://rs.gbif.org/sandbox/extension/issg-pathway.xml>). However, we argue that this knowledge is so fundamental to biodiversity information that it needs to be part of the Darwin Core standard, classified under the class Occurrence, as a term dwc:pathway. It should also be added to the Species Distribution extension so that it can be used in taxon-based checklists. Pathway information is not only relevant to alien species, but to any taxon, native or alien.

Proposed definition of dwc:pathway

"The process by which an Organism came to be in a given place at a given time."

Recommended vocabulary

Hulme et al. (2008) published a framework for a pathway vocabulary, which has since been adopted and refined by the Convention on Biological Diversity and by the IUCN Species Survival Commission Invasive Species Specialist Group (IUCN SSC ISSG). This is the vocabulary already recommended to be used with the Invasive Species Pathways extension to Darwin Core (Pagad et al. 2015, Scalera et al. 2016). The recommended vocabulary for pathway includes six major categories: "release", "escape", "transport-contaminant", "transport-stowaway", "corridor", and "unaided". Under these categories are 44 subcategories that can be used to further specify the pathway. This pathway vocabulary can be used for individual occurrences, when it is known, such as when a tree has been planted or a released animal has been tagged. However, they can be used more broadly in checklists to express the pathways by which the organism arrives (Fig. 1, Harrower et al. 2017). In the case of a checklist, if there are multiple pathways, these can be expressed by having a one-to-many relationship between the entries in the taxon file and the entries in the Species Distribution extension file, each with a specific pathway. In the latter case, this can also help describe the temporal changes in pathways of introduction. A text file is available in supplementary files containing the full vocabulary, including controlled value strings that can be used to implement the vocabulary according to the TDWG Vocabulary Maintenance Specification (Suppl. material 3).

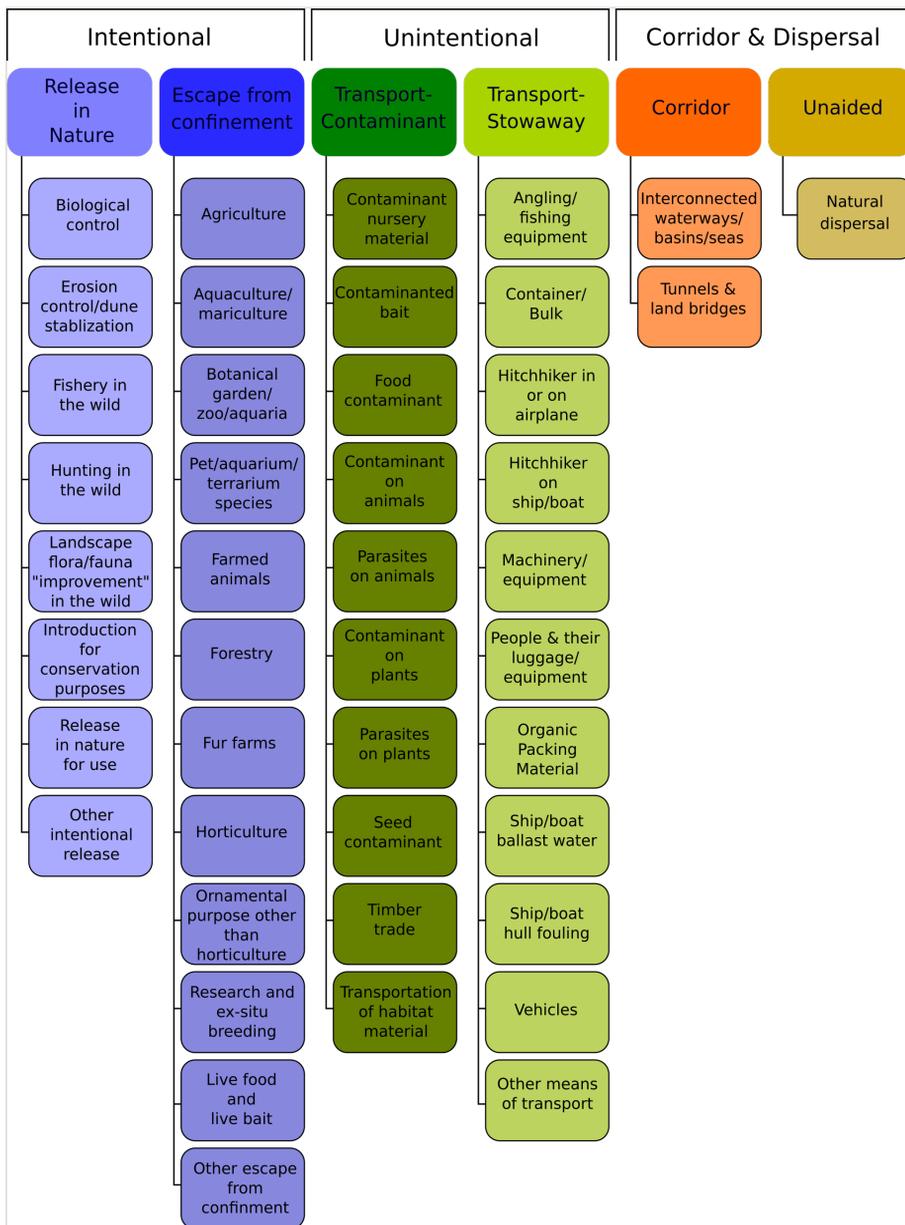


Figure 1.

A summary of the pathways categorisation scheme reproduced with permission from Harrower et al. (2017). The pathways are classified into three types:

1. intentional transport of taxa (blue)
2. unintentionally transported (green)
3. taxa moved between regions without direct transportation by humans and/or via artificial corridors (orange & yellow).

degreeOfEstablishment (New)

In the current and proposed vocabulary for `dwc:establishmentMeans` there is the explicit recognition that the occurrence of an organism can be either temporary or established. A bird may be blown off course and occur fleetingly in an area, or a seedling may germinate in an unsuitable place only to be killed a few weeks later by the conditions in that habitat, such as frost or drought. Likewise there are those organisms that are so well established that they reproduce and increase in range. Between these two extremes are different degrees of establishment. In this middle ground there are those organisms that persist in a location with no reproduction, others that reproduce, but do not have a significant population increase, and others that might reach high local densities but do not spread. There are, in essence, different routes to commonness (McGeoch and Latombe 2015). Such information is sometimes obvious at the time of observation, such as when there are numerous saplings around a mature invasive tree. Systematically recording how far such offspring have spread from the initial point of introduction provides important insights into invasion dynamics (Wilson et al. 2013). In the case of checklists, the degree of establishment is derived from the author's experience and information on the abundance, reproduction and spread of a taxon. This information is not only valuable to invasion biologists, but is also important for conservation assessments of rare species and for general wildlife management. Under conditions of environmental change, native species may also show increases in abundance or extent (Buczkowski 2010).

Currently, Darwin Core lacks an independent term to express degree of establishment. The closest term is "invasiveness" from the Invasive Species Distribution extension, but it has a limited vocabulary and, because it is restricted to invasive species, is of finite use. The vocabulary consists of the four terms, `invasive`, `notInvasive`, `uncertain` and `unspecified` and was created by the IUCN Species Survival Commission Invasive Species Specialist Group (Pagad et al. 2015).

In the case of introduced organisms, Blackburn et al. (2011) proposed a framework to describe the invasion process, which spans all degrees of establishment from species in captivity to fully invasive. This framework is from the perspective of the invasion process and as such it combines the translocation of the organism, its ability to survive in a novel location and its ability to reproduce and spread (Table 2). To some extent the same vocabulary can also be applied to native species. For example, categories B1–B3, which relate to captivity, can apply to the stocking of native fish or plantings of native tree species. Categories C1–E relate to how well the taxon is surviving and reproducing, and they are just as relevant to populations of native taxa as they are to populations of alien taxa. It is however important that `dwc:establishmentMeans` and `dwc:occurrenceStatus` are used in conjunction with `dwc:degreeOfEstablishment` to communicate the full context.

Table 2.

Proposed controlled vocabulary for dwc:degreeOfEstablishment adapted from Blackburn et al. (2011) including a simple human readable label. Populations categorised as C3–E would be considered naturalised, and populations categorised as D2 or E as invasive. Appropriate URIs will be assigned upon adoption of the controlled vocabulary.

Important Note: The definition of an invasive species by the Convention on Biological Diversity (and others) is restricted to those species that may cause economic or environmental harm or adversely affect human health. We use the term invasive here in the broader biological sense of the word.

category	definition	Proposed label and controlled value string
A	Not transported beyond limits of native range	native
B1	Individuals in captivity or quarantine (i.e. individuals provided with conditions suitable for them, but explicit measures of containment are in place)	captive
B2	Individuals in cultivation (i.e. individuals provided with conditions suitable for them, but explicit measures to prevent dispersal are limited at best)	cultivated
B3	Individuals directly released into novel environment	released
C0	Individuals released outside of captivity or cultivation in a location, but incapable of surviving for a significant period	failing
C1	Individuals surviving outside of captivity or cultivation in a location, no reproduction	casual
C2	Individuals surviving outside of captivity or cultivation in a location, reproduction is occurring, but population not self-sustaining	reproducing
C3	Individuals surviving outside of captivity or cultivation in a location, reproduction occurring, and population self-sustaining	established
D1	Self-sustaining population outside of captivity or cultivation, with individuals surviving a significant distance from the original point of introduction	colonising
D2	Self-sustaining population outside of captivity or cultivation, with individuals surviving and reproducing a significant distance from the original point of introduction	invasive
E	Fully invasive species, with individuals dispersing, surviving and reproducing at multiple sites across a greater or lesser spectrum of habitats and extent of occurrence	widespreadInvasive

It is recognised that the scheme of Blackburn et al. (2011) may not be suitable for all situations. However, users of Darwin Core are at liberty to use other controlled vocabularies if they wish. Yet, by providing a term to express these data and by providing a recommendation for a vocabulary there will be an improvement in the usefulness of data and their interoperability.

The degreeOfEstablishment term and its suggested vocabulary are proposed to be added to the Darwin Core standard, classified under the class Occurrence.

Proposed definition of dwc:degreeOfEstablishment

“The degree to which an Organism survives, reproduces, and expands its range at the given place and time.”

Example use cases

These proposed changes to Darwin Core have been tested on, and informed by, real data. Below are three examples where we have used these terms and vocabularies in datasets published to GBIF. A zoological example has also been published by Backeljau et al. (2019).

Manual of the Alien Plants of Belgium

The Manual of the Alien Plants of Belgium is a regularly updated checklist of all of the non-indigenous plants that have been found in Belgium, including those that have subsequently become extinct and those that only casually occur there (Verloove 2018). The manual includes information on the origins of the alien species, how they arrived in Belgium and information about when they were first and last seen. Each entry is based on solid evidence, particularly herbarium specimens, but also on photographs when their taxonomic identity is unequivocal. The checklist is maintained as a Microsoft Excel spreadsheet for convenience. This checklist is then published as a dataset to GBIF (Verloove et al. 2018). The conversion of the spreadsheet to Darwin Core is described in the metadata of this dataset on GBIF and the code to do the conversion is available on GitHub (<https://github.com/trias-project/alien-plants-belgium>). A small extract of the relevant fields from the checklist is shown in Table 3.

Table 3.

Example data from the Manual of Alien Plants of Belgium (Verloove 2018). The abbreviations are Mode of introduction (M/I): D=Deliberate, A=Accidental; Year of first record (FR); Year of most recent record (MRR); Presence: in Flanders, Brussels and Wallonia (fl, br, wa); Degree of naturalisation (D/N): Cas. = casual, Nat. = naturalised, Ext. = extinct; Vector of introduction (V/I): Hort.=Horticulture, Wool=propagules introduced with imported wool, Ore = propagules introduced with imported ore.

Taxon	M/I	FR	MRR	fl	br	wa	D/N	V/I
<i>Sambucus canadensis</i> L.	D	1972	2017	X			Cas.	Hort.
<i>Verbesina alternifolia</i> (L.) Britton	D	1984	N?	X			Nat.?	Hort.
<i>Hornungia procumbens</i> (L.) Hayek	A	<1850	<1850	?	?	?	Cas.	?
<i>Bothriochloa ischaemum</i> (L.) Keng	A	1813	1916	X		X	Ext.	Wool, Ore

Each entry for the Manual of Alien Plants of Belgium describes the existence of one non-native taxon in Belgium. It gives information on the species introduction status over a period of time, from the first year that it was recorded to the present day. It also gives regional information within Belgium. To convert this into a Darwin Core Archive checklist, a taxon file is created with one record for each entry in the checklist (Table 4). A distribution extension is also required to express the multiple presences of the taxon in Belgium and its regions (Table 5). Furthermore, if the checklist states that the species has become extinct, then multiple entries in the distribution file are needed to express the status before and after the last year it was recorded.

Table 4.

The relevant Darwin Core Archive taxon core created from the Manual of Alien Plants of Belgium data in Table 3.

taxonID	scientificName
alien-plants-belgium:taxon:03206f4a769c6649658ab96839e8a016	<i>Sambucus canadensis</i> L.
alien-plants-belgium:taxon:318b79c7d62889c229128c57e61973c7	<i>Verbesina alternifolia</i> (L.) Britton
alien-plants-belgium:taxon:b27d5b74783b9add7bd6747773e91fab	<i>Hornungia procumbens</i> (L.) Hayek
alien-plants-belgium:taxon:fe1d6bc47b13c9123410610d893a17cb	<i>Bothriochloa ischaemum</i> (L.) Keng

Table 5.

The relevant Darwin Core distribution extension fields created from the Manual of Alien Plants of Belgium in Table 3. The terms `dwc:pathway` and `dwc:degreeOfEstablishment` are currently not available in Darwin Core.

taxonID	locality	occurrence-Status	establish-mentMeans	eventDate	pathway	degreeOf-Establishment
alien-plants-belgium:taxon:03206f4a769c6649658ab96839e8a016	Flemish Region	present	introduced	1972/2017	horticulture	casual
alien-plants-belgium:taxon:03206f4a769c6649658ab96839e8a016	Belgium	present	introduced	1972/2017	horticulture	casual
alien-plants-belgium:taxon:318b79c7d62889c229128c57e61973c7	Flemish Region	present	introduced	1984/2018	horticulture	established
alien-plants-belgium:taxon:318b79c7d62889c229128c57e61973c7	Belgium	present	introduced	1984/2018	horticulture	established
alien-plants-belgium:taxon:b27d5b74783b9add7bd6747773e91fab	Flemish Region	doubtful	introduced			casual
alien-plants-belgium:taxon:b27d5b74783b9add7bd6747773e91fab	Walloon Region	doubtful	introduced			casual
alien-plants-belgium:taxon:b27d5b74783b9add7bd6747773e91fab	Brussels-Capital Region	doubtful	introduced			casual
alien-plants-belgium:taxon:b27d5b74783b9add7bd6747773e91fab	Belgium	doubtful	introduced			casual

taxonID	locality	occurrence-Status	establishmentMeans	eventDate	pathway	degreeOf-Establishment
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Flemish Region	present	introduced			
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Walloon Region	present	introduced			
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Belgium	present	introduced	1813/1916	contaminant OnAnimals containerBulk	
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Flemish Region	absent	introduced			
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Walloon Region	absent	introduced			
alien-plants-belgium:taxon: fe1d6bc47b13c9123410610d893a17cb	Belgium	absent		1916/2018		

In the Manual, the first and last observation date refer to Belgium as a whole, but there is no information on the first and last observations for Flanders, Wallonia and Brussels. We had the option of either supplying no temporal boundaries for these entries or providing the same dates as for Belgium as a whole. We concluded that it was better not to provide dates for Flanders, Wallonia and Brussels, rather than give misleading information.

Catalogue of the Rust Fungi of Belgium

The Catalogue of the Rust Fungi of Belgium is a static checklist published in print in 2009 (Vanderweylen and Fraiture 2007, Vanderweylen and Fraiture 2008)(Fig. 2). As part of the TriAS project we extracted the data from this paper and published it on GBIF (Vanderweylen et al. 2018). It contains information on specimens, particularly the first and most recent observations. It also has information on whether the species are native or introduced and the host species of the fungi. Table 6 shows an excerpt of the Taxon Core file for these data (Wieczorek et al. 2012). The distributional information included uses the Species Distribution extension of Darwin Core Table 7 (<http://rs.gbif.org/extension/gbif/1.0/distribution.xml>) and the host-parasite relationship is expressed in the Darwin Core file using the Resource Relationship extension (Table 8) (http://rs.gbif.org/extension/dwc/resource_relation.xml). In this checklist, there is no information on the pathway of introduction or the degree of establishment.

Table 6.

The relevant Darwin Core Archive taxon core and terms created from the Catalogue of the Rust Fungi of Belgium in Fig. 2.

taxonID	scientificName
uredinales-belgium-checklist:taxon:e82e5b9f24dc198819ebfc25068ae51	Frommeëlla mexicana (Mains) J.W. McCain & J.F. Hennen

taxonID	scientificName
uredinales-belgium-checklist:taxon:8b039e480746ec727316c1ad56ed8759	Uromyces croci Pass.
uredinales-belgium-checklist:taxon:437376fa8fa57a92cfb2ab61d4b093f1	Duchesnea indica (Jacks.) Focke
uredinales-belgium-checklist:taxon:8867819f38b85d4669981ee9e32c9851	Crocus biflorus Mill.
uredinales-belgium-checklist:taxon:df3c9aaaf6c930d84f6a4073a6a01e7b	Puccinia argentata (Schultz) G. Winter
uredinales-belgium-checklist:taxon:0c7f30a0959d9f5fcb53e63454e9957a	Adoxa moschatellina L.

Table 7.

The relevant Darwin Core distribution extension and terms created from the Catalogue of the Rust Fungi of Belgium in Fig. 2. The distribution extension holds the `dwc:occurrenceStatus` and `dwc:establishmentMeans` data.

taxonID	locality	occurrenceStatus	establishmentMeans	eventDate
uredinales-belgium-checklist:taxon:e82e5bb9f24dc198819ebfc25068ae51	Belgium	present	introduced	2007-06-08/2007-06-12
uredinales-belgium-checklist:taxon:8b039e480746ec727316c1ad56ed8759	Belgium	doubtful	introduced	1876/1876
uredinales-belgium-checklist:taxon:df3c9aaaf6c930d84f6a4073a6a01e7b	Belgium	present	native	1898-08/1995-04-30

Table 8.

The relevant Darwin Core Archive `resourceRelationship` extension terms were created from the Catalogue of the Rust Fungi of Belgium in Fig. 2 to express the host-parasite relationship. The `resourceRelationship` extension holds information about the host species of the fungi. The host plant is in the `resourceID` column, the the fungal parasite is uniquely identified in the `relatedResourceID` column, the relationship is stated in the `relationshipOfResource` column. So, the `relatedResource` is a parasite of the resource.

resourceID	relatedResourceID	relationshipOfResource
uredinales-belgium-checklist:taxon:437376fa8fa57a92cfb2ab61d4b093f1	uredinales-belgium-checklist:taxon:e82e5bb9f24dc198819ebfc25068ae51	parasite of
uredinales-belgium-checklist:taxon:8867819f38b85d4669981ee9e32c9851	uredinales-belgium-checklist:taxon:8b039e480746ec727316c1ad56ed8759	parasite of
uredinales-belgium-checklist:taxon:0c7f30a0959d9f5fcb53e63454e9957a	uredinales-belgium-checklist:taxon:df3c9aaaf6c930d84f6a4073a6a01e7b	parasite of

Observations from Durham and Northumberland, United Kingdom

These are observations based upon those from Groom et al. (2015). This was a systematic survey of vascular plants in the counties of Durham and Northumberland in the United

Kingdom. Vascular plant taxa were surveyed in the predefined area, though for the most part, pathway, establishmentMeans and degreeOfEstablishment were not recorded by the observers. Nevertheless, here we imagine some possible values that could have been used to describe the situation of the observed plant (Table 9). Although not shown here, all original records have a date and detailed location information (Groom 2019). Note that while native to the coasts of the United Kingdom, *Cochlearia danica*, has spread inland along roads that are salted in the winter.

PHRAGMIDIACEAE Corda 1837						
<i>Frommeëlla</i> Cummins & Y. Hirats. 1983						
39. <i>Frommeëlla mexicana</i> (Mains) J.W. McCain & J.F. Hennen var. <i>indicae</i> J.W. McCain & J.F. Hennen, <i>Mycotaxon</i> 39 : 251 (1990).						
Synonymes : <i>Frommeëlla duchesneae</i> Arthur 1917 [basé sur un type sans stade III]. – <i>Frommeëlla duchesneae</i> Yohem, Cummins & Gilb. 1985 [basé sur un type sans stade III]. – <i>Kuehneola duchesneae</i> Arthur 1912 [type sans stade III]. – <i>Phragmidium duchesneae</i> P. Syd. & Syd. 1912 [basé sur un type sans stade III].						
Spécimens examinés : A. Vanderweylen F 560 (AV, KR), leg. A. Fraiture, Meise, 08.VI.2007, sur <i>Duchesnea indica</i> . – A. Vanderweylen F 561 (AV, BR), Auderghem, 10.VI.2007 sur <i>Duchesnea indica</i> . – A. Fraiture 3068 (BR), Meise, 12.VI.2007 sur <i>Duchesnea indica</i> .						
Hôte signalé : 0, II, III sur <i>Duchesnea indica</i> .						
Notes : Lorsqu'il décrit <i>Kuehneola duchesneae</i> , Arthur (1912) donne une description détaillée du stade III mais, en conformité avec sa philosophie, désigne comme type un spécimen à urédies. Après réexamen de ce spécimen type, McCain & Hennen (1990) concluent que, d'après le Code de nomenclature (Art. 59), le nom donné par Arthur ne peut être utilisé pour désigner le téléomorphe et ils redécouvrent ce dernier sous le nouveau nom de <i>Frommeëlla mexicana</i> var. <i>indicae</i> .						
<i>Frommeëlla mexicana</i> (Mains) J.W. McCain & J.F. Hennen n'est pas indigène en Europe. Cette espèce y a cependant déjà été observée dans différentes régions, toujours sur le même hôte. Il semble que les spécimens mentionnés ici constituent les premières récoltes de cette rouille en Belgique (Fraiture & Vanderweylen 2007).						

Figure 2.

An excerpt from (Vanderweylen and Fraiture 2007) to illustrate the format of the printed checklist, before it was converted into a digital checklist. Dates, host species and the occurrence status are highlighted, all of which had to be extracted from the text.

Table 9.

Examples of how the proposed vocabularies could be used with observations of native and alien species. These are single observations taken from survey events of a 1km² grid square made over several hours on a single day. Full occurrence data, including the dates and coordinates, are available from Groom et al. (2015).

occurrenceID	scientificName	basisOfRecord	establishment-Means	occurrence-Status	pathway	degreeOf-Establishment
2cd4p9h.24p5hq	<i>Aesculus hippocastanum</i> L.	HUMAN_OBSERVATION	introduced	present	ornamentalNon-Horticulture	cultivated
2cd4p9h.7bt1vc	<i>Cerastium fontanum</i> Baumg.	HUMAN_OBSERVATION	native	present		native

occurrenceID	scientificName	basisOfRecord	establishment-Means	occurrence-Status	pathway	degreeOf-Establishment
2cd4p9h.7qp79k	<i>Cochlearia danica</i> L.	HUMAN_OBSERVATION	introduced	present	naturalDispersal	invasive
2cd4p9h.75ycnf	<i>Heracleum mantegazzianum</i> Sommier & Levier	HUMAN_OBSERVATION	introduced	present	horticulture	invasive
2cd4p9h.7bt1ea	<i>Oxalis acetosella</i> L.	HUMAN_OBSERVATION	native	present		native
2cd4p9h.amdvmg	<i>Pinus sylvestris</i> L.	HUMAN_OBSERVATION	native	present	forestry	released
2cd4p9h.83f16f	<i>Rhododendron ponticum</i> L.	HUMAN_OBSERVATION	introduced	present	horticulture	established
2cd4p9h.62bx7w	<i>Sanicula europaea</i> L.	HUMAN_OBSERVATION	native	present		native
2cd4p9h.b2ncby	<i>Solanum lycopersicum</i> L.	HUMAN_OBSERVATION	vagrant	present	foodContaminant	casual

Summary

We have reviewed the definition and controlled vocabulary of the existing Darwin Core term `dwc:establishmentMeans`. Though its current definition and vocabulary present some difficulties for use, we feel that it is best to retain it as a term in Darwin Core, but provide a more precise definition and update the vocabulary. This will allow data to be backwardly compatible and to better answer a broader range of questions.

We have also proposed the creation of the term `dwc:pathway` in Darwin Core rather than use the non-standard term "pathway" from the in-development Invasive Species Pathways extension. This will make the term mainstream and expands its use to taxa beyond invasive species. It also will allow us to better track how humans are altering the distribution of many organisms. Finally, we propose the new term `dwc:degreeOfEstablishment` to answer the question of how well established a taxon is at a given time and place, and we propose a controlled vocabulary for this term. These proposals are summarized in Table 10.

To explain how these proposed changes to Darwin Core and its extensions can improve data sharing in the invasive species community, we also presented three use cases where sharing data through GBIF could be simplified by implementing these proposed changes to Darwin Core.

Table 10.

A summary of proposed Darwin Core changes.

Term	Proposals for term	Proposal for vocabulary
dwc:establishmentMeans	Retain term and refine definition (table 1)	Update vocabulary
dwc:pathway	Promote pathway term in Invasive Species Pathways extension to the Darwin Core standard, classified under the class Occurrence	Maintain current recommended vocabulary
dwc:degreeOfEstablishment	Add the term to the Darwin Core standard, classified under the class Occurrence	Adopt a modified vocabulary based on Blackburn et al. (2011).

Acknowledgements

These proposals have emerged from several years discussion in a number of fora and we are grateful to all those who have taken part. Some of these are mentioned below.

Alien Challenge COST Action European Information System for Alien Species - WG4: Data standardisation and harmonisation: Chuck Bargeron, Ana Cristina Cardoso, Niki Chartosia, Fabio Crocetta, Keith Douce, Anna Gazda, Milka Glavendekic, Alberto Inghilesi, Jana Medvecka, Jan Pergl, Olivera Petrovic-Obradovic, Jodey Peyton, Gareth Richards, Helen Roy, Elena Tricarico & Katharine Turvey.

GBIF - Task Group on Data Fitness for Use in Research on Invasive Alien Species: Shyama Pagad, Varos Petrosyan, Gregory Ruiz & Dmitry Schigel.

Biodiversity Information Standards Meetings (2016–2018): Lee Belbin, Matthew Blissett, Dimitry Brosens, Pier Luigi Buttigieg, Robert Guralnick, Niels Klazenga, Joel Sachs & Aaron Wilton

Funding program

Funded under the Belgian Science Policies Brain program, contract number BR/165/A1/ TriAS. Quentin Groom also acknowledges the Fonds Wetenschappelijk Onderzoek – Vlaanderen for the travel support it gave. The work of Peter Desmet, Lien Reyserhove and Damiano Oldoni is partially funded by Research Foundation - Flanders (FWO) as part of the Belgian contribution to LifeWatch.

References

- Ad Hoc Open-ended Working Group on Review of Implementation of the Convention (2011) Strategic plan for biodiversity 2011-2020. Further information related to the technical rationale for the Aichi Biodiversity Targets, including potential indicators and milestones., UNEP/CBD/COP/10/INF/12/Rev.1. Conference of the Parties to the Convention on Biological Diversity, Nagoya, 18-29 October 2010. United Nations Environment Programme, 31 pp. [In en]. URL: <https://www.cbd.int/kb/record/meetingDocument/77515?Event=COP-10>
- Aedo C, Pando F (2017) A distribution and taxonomic reference dataset of *Geranium* in the New World. Scientific Data 4: 170049. <https://doi.org/10.1038/sdata.2017.49>
- Backeljau T, Reyserhove L, Desmet P (2019) Registry of introduced terrestrial molluscs in Belgium. 1.3. Research Institute for Nature and Forest (INBO). Release date: 2019-5-17. URL: <https://doi.org/10.15468/t13kwo>
- Baskauf S, Wieczorek J, Blum S, Morris R, Rees J, Sachs J, Whitbread G (2017a) TDWG Vocabulary Maintenance Specification. <http://hdl.handle.net/1803/9512>. Accessed on: 2019-9-11.
- Baskauf S, Hyam R, Blum S, Morris R, Rees J, Sachs J, Whitbread G, Wieczorek J (2017b) TDWG Standards Documentation Specification. <http://hdl.handle.net/1803/9513>. Accessed on: 2019-9-11.
- Biodiversity Information Standards (TDWG) (2018) Darwin Core quick reference guide. <http://rs.tdwg.org/dwc/terms/>. Accessed on: 2018-11-01.
- Blackburn T, Pyšek P, Bacher S, Carlton J, Duncan R, Jarošík V, Wilson JU, Richardson D (2011) A proposed unified framework for biological invasions. Trends in Ecology & Evolution 26 (7): 333-339. <https://doi.org/10.1016/j.tree.2011.03.023>
- Blackburn T, Essl F, Evans T, Hulme P, Jeschke J, Kühn I, Kumschick S, Marková Z, Mrugała A, Nentwig W, Pergl J, Pyšek P, Rabitsch W, Ricciardi A, Richardson D, Sendek A, Vilà M, Wilson JU, Winter M, Genovesi P, Bacher S (2014) A unified classification of alien species based on the magnitude of their environmental impacts. PLoS Biology 12 (5): e1001850. <https://doi.org/10.1371/journal.pbio.1001850>
- Buczkowski G (2010) Extreme life history plasticity and the evolution of invasive characteristics in a native ant. Biological Invasions 12 (9): 3343-3349. <https://doi.org/10.1007/s10530-010-9727-6>
- Endresen DTF, Knüpffer H (2012) The Darwin Core extension for genebanks opens up new opportunities for sharing genebank datasets. Biodiversity Informatics 8 (1): 12-29. <https://doi.org/10.17161/bi.v8i1.4095>
- Essl F, Bacher S, Genovesi P, Hulme PE, Jeschke JM, Katsanevakis S, Kowarik I, Kühn I, Pyšek P, Rabitsch W, Schindler S, Kleunen Mv, Vilà M, U Wilson JR, Richardson DM (2018) Which taxa are alien? Criteria, applications, and uncertainties. BioScience 68 (7): 496-509. <https://doi.org/10.1093/biosci/biy057>
- Groom Q, Durkin J, O'Reilly J, Mclay A, Richards A, Angel J, Horsley A, Rogers M, Young G (2015) A benchmark survey of the common plants of South Northumberland and Durham, United Kingdom. Biodiversity Data Journal 3 <https://doi.org/10.3897/bdj.3.e7318>
- Groom Q, Baskauf S, Desmet P, McGeoch M, Pagad S, Schigel D, Walls R, Wilson J, Zermoglio P (2017a) Invasive Organisms Information: A proposed TDWG task group. Proceedings of TDWG 1: e20266. <https://doi.org/10.3897/tdwgproceedings.1.20266>

- Groom Q, Adriaens T, Desmet P, Simpson A, Wever AD, Bazos I, Cardoso AC, Charles L, Christopoulou A, Gazda A, Helmisaari H, Hobern D, Josefsson M, Lucy F, Marisavljevic D, Oszako T, Pergl J, Petrovic-Obradovic O, Prévot C, Ravn H, Richards G, Roques A, Roy H, Rozenberg M, Scalera R, Tricarico E, Trichkova T, Vercayie D, Zenetos A, Vanderhoeven S (2017b) Seven recommendations to make your invasive alien species data more useful. *Frontiers in Applied Mathematics and Statistics* 3 <https://doi.org/10.3389/fams.2017.00013>
- Groom Q (2019) A common plants survey of vascular plants in South Northumberland and Durham, United Kingdom. 1.9. Meise Botanic Garden. Release date: 2019-4-06. URL: <https://doi.org/10.3897/bdj.3.e7318>
- Guralnick R, Walls R, Jetz W (2017) Humboldt Core - toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* 41 (5): 713-725. <https://doi.org/10.1111/ecog.02942>
- Harrower CA, Scalera R, Pagad S, Schönrogge K, Roy HE (2017) Guidance for interpretation of CBD categories on introduction pathways. Technical note prepared by IUCN for the European Commission. IUCN, 100 pp. [In English].
- Hulme PE, Bacher S, Kenis M, Klotz S, Kühn I, Minchin D, Nentwig W, Olenin S, Panov V, Pergl J, Pyšek P, Roques A, Sol D, Solarz W, Vilà M (2008) Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *Journal of Applied Ecology* 45 (2): 403-414. <https://doi.org/10.1111/j.1365-2664.2007.01442.x>
- IUCN (2018) IUCN Red List of Threatened Species – Mapping Attributes Standards. <https://www.iucnredlist.org/resources/grid>. Accessed on: 2019-2-06.
- Katsanevakis S, Deriu I, D'Amico F, Nunes AL, Sanchez SP, Crocetta F, Arianoutsou M, Bazos I, Christopoulou A, Curto G, Delipetrou P, Kokkoris Y, Panov V, Rabitsch W, Roques A, Scalera R, Shirley S, Tricarico E, Vannini A, Zenetos A, Zervou S, Zikos A, Cardoso AC (2015) European Alien Species Information Network (EASIN): supporting European policies and scientific research. *Management of Biological Invasions* 6 (2): 147-157. <https://doi.org/10.3391/mbi.2015.6.2.05>
- Keller RP, Geist J, Jeschke JM, Kühn I (2011) Invasive species in Europe: ecology, status, and policy. *Environmental Sciences Europe* 23 (1). <https://doi.org/10.1186/2190-4715-23-23>
- Latombe G, Pyšek P, Jeschke J, Blackburn T, Bacher S, Capinha C, Costello M, Fernández M, Gregory R, Hobern D, Hui C, Jetz W, Kumschick S, McGrannachan C, Pergl J, Roy H, Scalera R, Squires Z, Wilson JU, Winter M, Genovesi P, McGeoch M (2017) A vision for global monitoring of biological invasions. *Biological Conservation* 213: 295-308. <https://doi.org/10.1016/j.biocon.2016.06.013>
- Leung B, Springborn MR, Turner JA, Brockerhoff EG (2014) Pathway-level risk analysis: the net present value of an invasive species policy in the US. *Frontiers in Ecology and the Environment* 12: 273-279. <https://doi.org/10.1890/130311>
- MacKenzie D, Nichols J, Lachman G, Droege S, Royle JA, Langtimm C (2002) Estimating its occupancy rates when detection probabilities are less than one. *Ecology* 83 (8): 2248. <https://doi.org/10.2307/3072056>
- Marchand F, Beaulaton L, Prévost E, Delanoë R, Destouches J, Gueraud F, Guilloux Y, Jeannot N, Huchet E, Lange F, Rives J, Tremblay J, Herrard N, Azam D (2017) Abundance indices and biological traits of juvenile salmon (*Salmo salar*) sampled in three rivers on the Atlantic and Channel coasts (France). *Biodiversity Data Journal* 5: e15125. <https://doi.org/10.3897/bdj.5.e15125>

- McGeoch M, Latombe G (2015) Characterizing common and range expanding species. *Journal of Biogeography* 43 (2): 217-228. <https://doi.org/10.1111/jbi.12642>
- McGeoch MA, Groom QJ, Pagad S, Petrosyan V, Ruiz G, Wilson J (2016) Data fitness for use in research on alien and invasive species. GBIF, Copenhagen. URL: <http://www.gbif.org/document/82958>.
- Pagad S, Genovesi P, Carnevali L, Scalera R, Clout M (2015) IUCN SSC Invasive Species Specialist Group: invasive alien species information management supporting practitioners, policy makers and decision takers. *Management of Biological Invasions* 6 (2): 127-135. <https://doi.org/10.3391/mbi.2015.6.2.03>
- Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch M (2018) Introducing the Global Register of Introduced and Invasive Species. *Scientific Data* 5: 170202. <https://doi.org/10.1038/sdata.2017.202>
- Parr C, Guralnick R, Cellinese N, Page RM (2012) Evolutionary informatics: unifying knowledge about the diversity of life. *Trends in Ecology & Evolution* 27 (2): 94-103. <https://doi.org/10.1016/j.tree.2011.11.001>
- Remsen D, Robertson T, Döring M (2010) GBIF GNA Profile Reference Guide for Darwin Core Archives. <http://doi.org/10.5281/zenodo.3463261>. Accessed on: 2019-9-28.
- Remsen D, Knapp S, Georgiev T, Stoev P, Penev L (2012) From text to structured data: Converting a word-processed floristic checklist into Darwin Core Archive format. *PhytoKeys* 9: 1. <https://doi.org/10.3897/phytokeys.9.2770>
- Ricciardi A, Blackburn T, Carlton J, Dick JA, Hulme P, Iacarella J, Jeschke J, Liebhold A, Lockwood J, MacIsaac H, Pyšek P, Richardson D, Ruiz G, Simberloff D, Sutherland W, Wardle D, Aldridge D (2017) Invasion Science: A horizon scan of emerging challenges and opportunities. *Trends in Ecology & Evolution* 32 (6): 464-474. <https://doi.org/10.1016/j.tree.2017.03.007>
- Rodrigues A, Pilgrim J, Lamoreux J, Hoffmann M, Brooks T (2006) The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution* 21 (2): 71-76. <https://doi.org/10.1016/j.tree.2005.10.010>
- Scalera R, Genovesi P, Booy O, Essl F, Jeschke J, Hume P, McGeoch M, Pagad S, Roy H, Saul W, Wilson J (2016) Progress toward pathways prioritization in compliance to Aichi target 9. Convention on Biological Diversity, subsidiary body on scientific technical and technological advice, twentieth meeting UNEP/CBD/SBSTTA/20/INF/5 URL: <https://www.cbd.int/doc/meetings/sbstta/sbstta-20/information/sbstta-20-inf-05-en.pdf>
- Secretariat of the Convention on Biological Diversity (2009) What are invasive alien species? <https://www.cbd.int/idb/2009/about/what/>. Accessed on: 2019-6-08.
- Sutherland W, Woodroof H (2009) The need for environmental horizon scanning. *Trends in Ecology & Evolution* 24 (10): 523-527. <https://doi.org/10.1016/j.tree.2009.04.008>
- Turbé A, Strubbe D, Mori E, Carrete M, Chiron F, Clergeau P, González-Moreno P, Le Louarn M, Luna A, Menchetti M, Nentwig W, Pârâu L, Postigo J, Rabitsch W, Senar JC, Tollington S, Vanderhoeven S, Weiserbs A, Shwartz A (2017) Assessing the assessments: evaluation of four impact assessment protocols for invasive alien species. *Diversity and Distributions* 23 (3): 297-307. <https://doi.org/10.1111/ddi.12528>
- Vanderweyen A, Fraiture A (2007) Catalogue des *Uredinales* de Belgique. (1st partie). *Chaconiaceae, Coleosporiaceae, Cronartiaceae, Melampsoraceae, Phragmidiaceae, Pucciniaceae, Raveneliaceae* et *Uropyxidaceae*. *Lejeunia* 183: 1-36.
- Vanderweyen A, Fraiture A (2008) Catalogue des *Uredinales* de Belgique, 2ème partie, *Pucciniaceae* (sauf *Puccinia*). *Lejeunia, Revue de Botanique* 185.

- Vanderweyen A, Fraiture A, Groom Q, Desmet P, Reyserhove L (2018) Catalogue of the rust fungi of Belgium. 1.1. GBIF. Release date: 2018-8-15. URL: <https://doi.org/10.15468/2dboyn>
- Verloove F (2018) Manual of the alien plants of Belgium. Meise Botanic Garden. URL: alienplantsbelgium.be
- Verloove F, Groom Q, Brosens D, Desmet P, Reyserhove L (2018) Manual of the Alien Plants of Belgium. 1.6. Botanic Garden Meise. Release date: 2018-8-19. URL: <https://doi.org/10.15468/wtda1m>
- Wiecezorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglaiss D (2012) Darwin Core: An evolving community-developed biodiversity data standard. PLoS ONE 7 (1): e29715. <https://doi.org/10.1371/journal.pone.0029715>
- Wiecezorek J, Bánki O, Blum S, Deck J, Döring M, Dröge G, Endresen D, Goldstein P, Leary P, Krishtalka L, Tuama ÉÓ, Robbins R, Robertson T, Yilmaz P (2014) Meeting Report: GBIF hackathon-workshop on Darwin Core and sample data (22-24 May 2013). Standards in Genomic Sciences 9 (3): 585-598. <https://doi.org/10.4056/sigs.4898640>
- Wilson JU, Caplat P, Dickie I, Hui C, Maxwell B, Nuñez M, Pauchard A, Rejmánek M, Richardson D, Robertson M, Spear D, Webber B, van Wilgen B, Zenni R (2013) A standardized set of metrics to assess and monitor tree invasions. Biological Invasions 16 (3): 535-551. <https://doi.org/10.1007/s10530-013-0605-x>
- Wilson JU, Faulkner K, Rahlao S, Richardson D, Zengeya T, van Wilgen B (2018) Indicators for monitoring biological invasions at a national level. Journal of Applied Ecology 55 (6): 2612-2620. <https://doi.org/10.1111/1365-2664.13251>

Supplementary materials

Suppl. material 1: Distinct Establishmentmeans Values from observations GBIF 2017-02-27 [doi](#)

Authors: John Wiecezorek

Data type: tap seperated file

Brief description: Distinct values for dwc:establishmentMeans and their frequency from observations on the Global Biodiversity Information Facility on 27 February 2017. Taken from GitHub repository of the Darwin Core Questions & Answers Site (<https://github.com/tdwg/dwc-qa/tree/master/data/GBIFDistinctValues>).

[Download file](#) (50.84 kb)

Suppl. material 2: The establishmentMeans dictionary from the Global Biodiversity Information Facility as of 2018-11-02 [doi](#)

Authors: Matt Blissett

Data type: tab-delimited text

Brief description: A tab-delimited file mapping values (synonyms; orthographic and language variations) found in Darwin Core dwc:establishmentMeans to a controlled vocabulary.

Filename: establishment_means.tsv - [Download file](#) (588.00 bytes)

Suppl. material 3: dwc:pathway vocabulary [doi](#)

Authors: Adapted from Harrower CA, Scalera R, Pagad S, Schönrogge K, Roy HE (2017) Guidance for interpretation of CBD categories on introduction pathways. Technical note prepared by IUCN for the European Commission. IUCN, 100 pp. [In English].

Data type: tab delimited text file

Brief description: The Convention on Biological Diversity pathway vocabulary adapted from Harrower et al. 2017. Including proposed simple labels for these terms.

[Download file](#) (4.09 kb)