

Darwin Core proposals to support alien species research

Quentin Groom

Peter Desmet



Consultation on the proposal

Alien Challenge COST Action European Information System for Alien Species - WG4: Data standardisation and harmonisation

Chuck Bargo (USA)

Niki Chartosia (Cyprus)

Fabio Crocetta (Italy)

Keith Douce (USA)

Anna Gazda (Poland)

Milka Glavendekic (Serbia)

Quentin Groom (Belgium)

Alberto Inghilesi (Italy)

Jana Medvecká (Slovakia)

Jan Pergl (Czech Republic)

Olivera Petrovic-Obradovic (Serbia)

Jodey Peyton (United Kingdom)

Gareth Richards (United Kingdom)

Elena Tricarico (Italy)

Kath Turvey (United Kingdom)

Sonia Vanderhoeven (Belgium)

GBIF - Task Group on Data Fitness for Use in Research on Invasive Alien Species

Melodie A. McGeoch (Chair), Monash University, Australia

Shyama Pagad, IUCN SSC Invasive Species Specialist Group

Varos Petrosyan, Russian Academy of Science

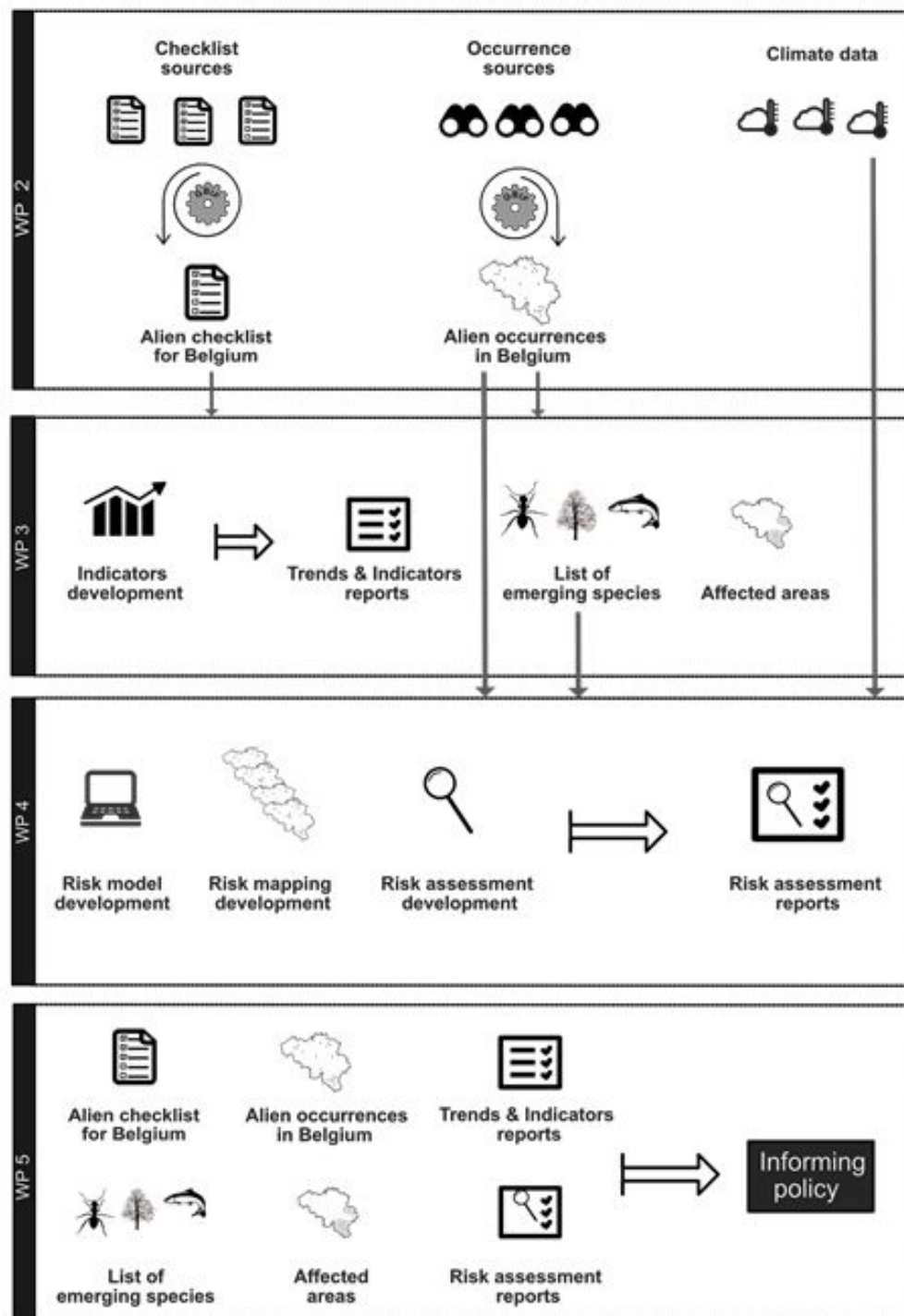
Gregory Ruiz, Smithsonian Environmental Research Centre, U.S.A.

John Wilson, Stellenbosch University, South Africa

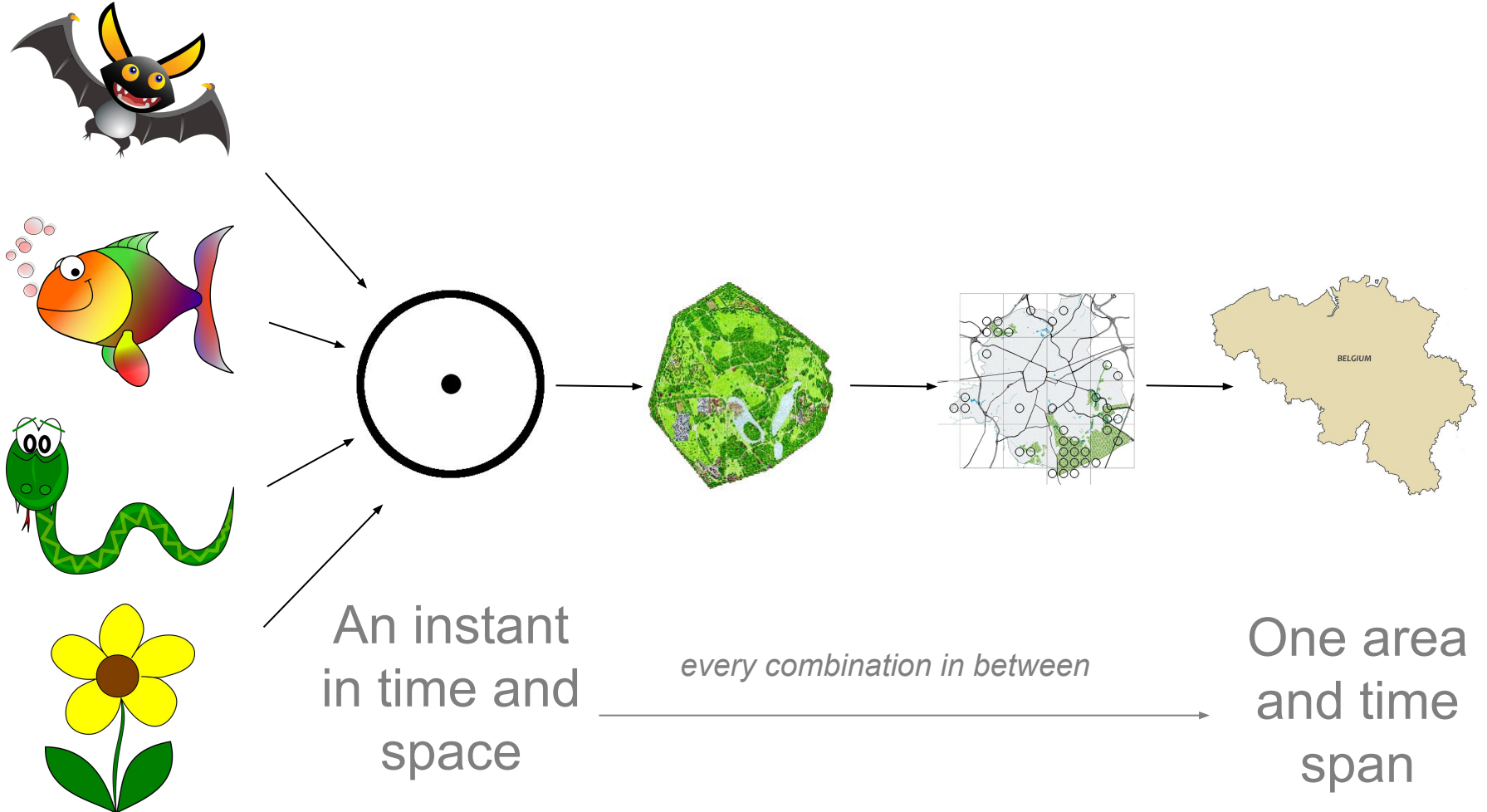
Dmitry Schigel, GBIF

Taxacom & IUCN Invasive Species Specialist Group

Use case



Observations versus checklists



dwc:establishmentMeans

*“The process by which the biological individual(s) represented in the **occurrence** became established at the location.”*

DwC & GBIF currently suggest the following vocabulary:

- native
- introduced
- naturalised
- invasive
- managed
- uncertain

See <http://terms.tdwg.org/wiki/dwc:establishmentMeans> and http://rs.gbif.org/vocabulary/gbif/establishment_means.xml

dwc:establishmentMeans

A new suggested vocabulary

Propose to adopt framework by Hulme et al. (2008):

Hulme, P.E., Bacher, S., Kenis, M., Klotz, S., Kuhn, I., Minchin, D. et al. (2008) Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *Journal of Applied Ecology*: 45, 403–414.

<http://doi.org/10.1111/j.1365-2664.2007.01442.x>



**Convention on
Biological Diversity**

dwc:establishmentMeans

Proposed vocabulary

Release	Escape	Contaminant	Stowaway	Corridor
biological control	agriculture	contaminant nursery material	angling/fishing equipment	interconnected waterways/basins/seas
erosion control/ dune stabilization	aquaculture / mariculture	contaminated bait	container/bulk	tunnels and land bridges
fishery in the wild	botanical	food contaminant	hitchhikers in or on airplane	unaided
Hunting	garden/zoo/aquaria	contaminant on animals	hitchhikers on ship/boat	
landscape/flora/fauna	pet/aquarium/terrarium species	parasites on animals	machinery/equipment	
“improvement” in the wild	farmed animals	contaminant on plants	people and their luggage/equipment	
introduction for conservation purposes or wildlife management	Forestry	parasites on plants	organic packing material	
release in nature for use	fur farms	seed contaminant	ship/boat ballast water	
other intentional release	Horticulture	timber trade	ship/boat hull fouling	
	ornamental purpose other than horticulture	transportation of habitat material	Vehicles	
	research and ex-situ breeding		other means of transport	
	live food and live bait			
	other escape from confinement			

dwc:establishmentMeans

current usage on GBIF

value	occurrences	% occurrences	datasets	% datasets
null	695,217,928	98.79%	14,157	98.13%
native	8,403,205	1.19%	111	0.77%
managed	61,403	0.01%	106	0.73%
introduced	40,920	0.01%	10	0.07%
uncertain	21,550	0.00%	37	0.26%
naturalised	2,669	0.00%	3	0.02%
invasive	659	0.00%	3	0.02%
total	703,748,334	100.00%	14,427	100.00%

Stats collected from GBIF on 2016-12-02

dwc:origin

A new term

“To express the concept that a **Taxon** has either been introduced by human activity or occurs naturally in a **Location** and has done so before the impact of modern humans.”

dwc:origin

A new suggested vocabulary

- Native
- Reintroduced
- Introduced
 - introduced before the modern era
 - Introduced during the modern era
- Vagrant
- Unknown

<http://bit.ly/1NhL5dm>



dwc:occurrenceStatus

“A statement about the presence or absence of a **Taxon** at a Location. Recommended best practice is to use a controlled vocabulary.”

DwC & GBIF currently suggest the following vocabulary:

- present
- absent
- common
- irregular
- rare
- doubtful
- excluded (GBIF voc only)

See

<https://terms.tdwg.org/wiki/dwc:occurrenceStatus>

and

http://rs.gbif.org/vocabulary/gbif/occurrence_status.xml

The current use of dwc:occurrenceStatus on GBIF

- Presence and absence
- Abundance (common, frequent, rare)
- Conservation status (EX, EW, CR, EN, VU, NT, LC, DD, NE)
- Health condition (damaged, infested, defoliated)
- Breeding (presumably for birds)

dwc:occurrenceStatus

A new suggested vocabulary

- Extant
- Possibly extant
- Possibly extinct
- Extinct (post 1500)
- Presence uncertain
- Absent

<http://tinyurl.com/hsg46th>



Example for a checklist dataset (VASCAN)

Distribution codes used in VASCAN:

<http://data.canadensys.net/vascan/about#distribution>

occurrence Status	establishment Means	establishment Means	origin	occurrence Status
present	native		native	extant
present	introduced		introduced	extant
irregular	introduced		introduced	vagrant
excluded				absent
doubtful				presence uncertain
absent				absent

Brouillet et al. 2016 <http://doi.org/10.5886/zw3aqw>

Controlled Vocabularies, Definitions and Guidance

Darwin Core makes **few recommendations** on the content and format of terms, e.g. how should hybrids be represented? How should taxonomic authorities be written?

Laissez-faire or Procrustean

- + Loose definitions make it easier to format and publish data.
- Tight definitions improve interoperability, but limit expressiveness.

The users want guidance, but they can choose to ignore it when they want.

Other issue related to Invasive Species Data and Darwin Core

How well is the species established in an area? Blackburn et al. (2011) <http://doi.org/10.1016/j.tree.2011.03.023>

Zoos and cultivation: basisOfRecord is not up to the job

Interactions between organisms: parasitism, pollination

Other sorts of data:

- Camera trap
- eDNA
- Ringing data
- GPS Tracking data
- Indirect observations (tracks, hair, faeces, galls etc)



GBIF

Global Biodiversity
Information Facility



**Botanic Garden
Meise**



ALIEN

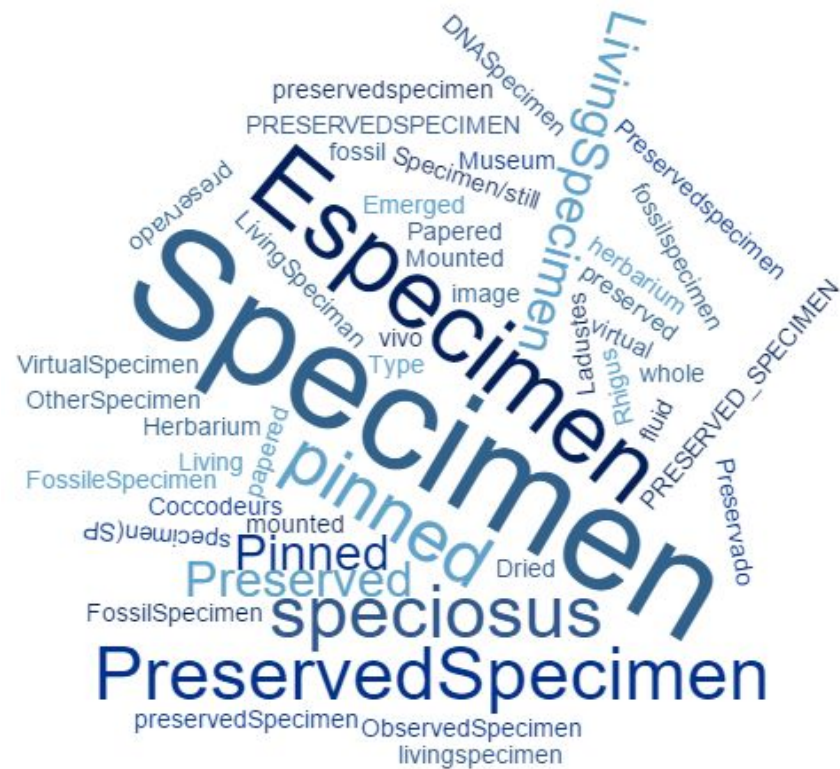


Challenge

dwc:basisOfRecord

2436 different values in GBIF: many don't follow any standard of Darwin Core

- 17 Single alphanumeric codes
- 3 Two letter codes
- 13 Three letter codes



Darwin Core & Checklists

DwC is used for this purpose, but designed for a moment in time and space. However **Darwin Core Event** fixes the concept of an observation made in an area over a period.

If the area of the checklist covers the whole range of the species then **Plinian Core** is more appropriate.

Some terms are relevant to checklists that are not so relevant to observations:

- extinct dwc:nomenclaturalStatus
- native conservation status (endangered, etc.)
- absent seasonality (resident, migrant, etc.)

Blackburn et al. (2011)

<http://doi.org/10.1016/j.tree.2011.03.023>

Table 1. A categorisation scheme for populations in the unified framework^a

Category	Definition
A	Not transported beyond limits of native range
B1	Individuals transported beyond limits of native range, and in captivity or quarantine (i.e. individuals provided with conditions suitable for them, but explicit measures of containment are in place)
B2	Individuals transported beyond limits of native range, and in cultivation (i.e. individuals provided with conditions suitable for them but explicit measures to prevent dispersal are limited at best)
B3	Individuals transported beyond limits of native range, and directly released into novel environment
C0	Individuals released into the wild (i.e. outside of captivity or cultivation) in location where introduced, but incapable of surviving for a significant period
C1	Individuals surviving in the wild (i.e. outside of captivity or cultivation) in location where introduced, no reproduction
C2	Individuals surviving in the wild in location where introduced, reproduction occurring, but population not self-sustaining
C3	Individuals surviving in the wild in location where introduced, reproduction occurring, and population self-sustaining
D1	Self-sustaining population in the wild, with individuals surviving a significant distance from the original point of introduction
D2	Self-sustaining population in the wild, with individuals surviving and reproducing a significant distance from the original point of introduction
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

^aHuman-mediated dispersal has created several novel categories of dispersal pathway (i.e. B1 and B2) [38], and human intervention has also significantly increased the frequency and duration that populations can persist in other categories (C0, C1 and C2).