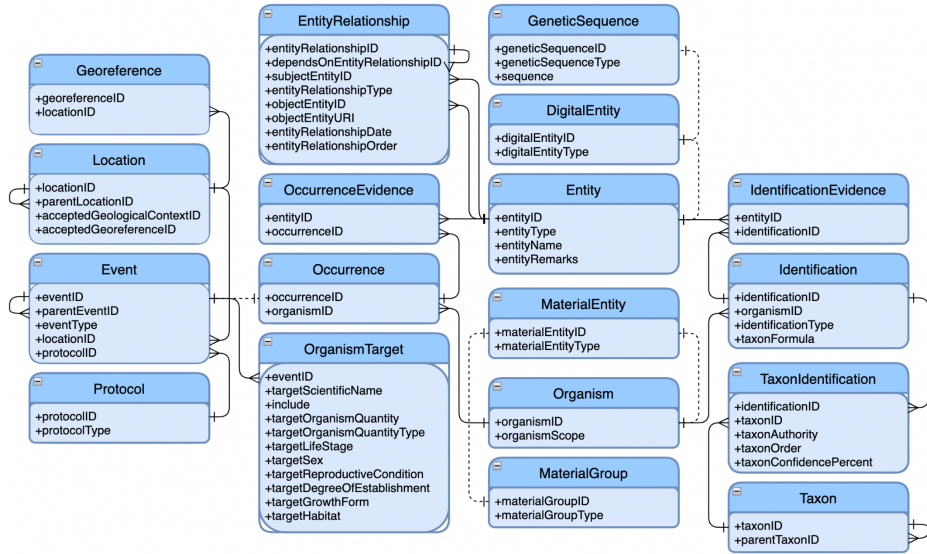


2023-01-25

Meeting:	Humboldt Core Task Group weekly meeting		
Date:	2023-01-23	Time:	0800 EST
Note taker:	Ming	Location:	Zoom
Meeting Objective(s)			
<ul style="list-style-type: none">Review Humboldt absence use case for the GBIF data model			
Attendees			
<ul style="list-style-type: none">Yanina Sica, Map of Life, Yale University, yanina.sica@yale.eduJohn Wieczorek, VertNet - Darwin Core Maintenance Interest Group, gtuco.btuco@gmail.comSteve Baskaufs, Vanderbilt University - Darwin Core Maintenance Interest Group, steve.baskauf@vanderbilt.edu (libraries, involved in dev and ratification of other vocabularies)Peter Brenton, CSIRO, Atlas of Living Australia, &TDWG Citizen Science Interest Group and CSA Data & Metadata Working Group, peter.brenton@csiro.au (data collection apps, developed biocollect).Dmitry Schigel, GBIF Secretariat, dschigel@gbif.org (scientific officer, interfacing with ecology and eDNA data streams)Zachary Kachian, Field Museum, zkachian@fieldmuseum.orgYi-Ming Gan, Antarctic OBIS, Royal Belgian Institute for Natural Sciences, ymgan@naturalsciences.beWesley Hochachka, eBird			
Prior Action Item(s)			
<ul style="list-style-type: none"><<Action item>> [status]			
Agenda and Notes, Decisions, Issues			
Topic	Discussion		

Current conceptual model



- Description in use case [document](#)
- Addition of OrganismTarget table.
- Occurrence is now a subtype of Event.

OrganismTarget table

- OrganismTarget is not planned to be in the Publishing Model.
- Alternative way of listing all entries in OrganismTarget table is to provide a checklist of
 - Target species
 - Target exclusion species
- OrganismTarget has a rule where all combinations of the targets have to be unique.
- Peter: How about having OrganismTargetID under the Occurrence table?
- 14:18:28 From Dmitry Schigel (GBIF Secretariat) to Everyone:
 - +1 for event or study level declaration, in my experience this is how people think about targets
- 14:18:56 From Peter to Everyone:
 - I appreciate the power and flexibility of this model, but I'm not sure of the use cases for target exclusions at this level of detail.
 - Ming - target exclusions exist in this dataset because the adult fish will be at the bottom of the ocean but the larvae will be in the pelagic zone that is being sampled. The other situation is for another taxa, the adult fish will be able to sense the net, so they will swim away. But because larvae cannot swim, they will be caught.

How absences are being represented in this model

row	eventID	include	targetOrganismQuantity	targetScientificName	targetLifeStageScope
6	event2	True	17	Electrona antarctica	
7	event2	False		Electrona antarctica	adult
8	event2	True	13	Electrona antarctica	larva
9	event2	True	2	Electrona antarctica	postmeta

- What happens if you don't have the information?
 - Example above means 2 electrona antarctica have other life stage which is not adult in event2 (because adult is excluded)

	<p style="text-align: center;">from the scope of the Event)</p> <ul style="list-style-type: none"> ● 14:22:26 From Dmitry Schigel (GBIF Secretariat) to Everyone: empty cells vs. informed zeros, yes ● 14:25:53 From Dmitry Schigel (GBIF Secretariat) to Everyone: There is one specific problem with absences when target is higher than species: do you then target any species within the declared target family, or all known species in this family / region. Non-detections have different meanings in these two cases, and therefore when we used target approach in Finland we ended up with a list of detectable species as a target, to be better sure in our non-detection ● 14:27:37 From John Wieczorek to Everyone: Hi Dmitry. I think that for the model to allow for either, another parameter might have to be added. Hopefully that can come out of the use case with eBird. ● 14:28:39 From Dmitry Schigel (GBIF Secretariat) to Everyone: Thanks, this is probably a borderline issue anyway, low priority ● 14:29:27 From John Wieczorek to Everyone: Also, this might be solved with two lists - one at the higher Event level with the species target checklist, and that list might take this same format. Not sure yet.
isInventoryComplete	<ul style="list-style-type: none"> ● Did you report every species you saw ● Which one was left out because: <ul style="list-style-type: none"> ○ Protocol could detect it but it is not present ○ Protocol could not detect it
Definition of “target”, absence of detection and absence of reporting in this model	<ul style="list-style-type: none"> ● Everything in the OrganismTarget is a target. ● 14:29:45 From Dmitry Schigel (GBIF Secretariat) to Everyone: Detectability is controlled by the method (size etc) and by the target ● 14:30:02 From Dmitry Schigel (GBIF Secretariat) to Everyone: Can detect / want to detect ● 14:32:44 From Yanina Sica (she/her) to Everyone: yes! but in this model how do you separate (in steve's example) between foxA being detected, foxB not being detected and deerA not being target? ● 14:33:06 From Wesley Hochachka to Everyone: Yes, I think that this is important to distinguish between "can detect" and "would report if detected".. ● 14:34:06 From Yanina Sica (she/her) to Everyone: you could get the info from the detailed protocols but was just wondering... ● 14:37:16 From Dmitry Schigel (GBIF Secretariat) to Everyone: One brutal formula can be everything > can be detected > would report is detected. Occasionally by catch is likely not quantitative not interpretable in the same way as target taxa ● 14:38:20 From Dmitry Schigel (GBIF Secretariat) to Everyone: By-catch can be interpreted as occurrences, "we happen to also observe" etc ● 14:41:54 From Wesley Hochachka to Everyone: Dmitry, that is how I think about the process of moving from collection/observation events to data in the database. "would be reported if detected" needs to be specified in the database somehow. "can be detected", to me, is an inference that is created based on data within the database (i.e, it does not necessarily need to be stored information in the database).

- 14:49:44 From Dmitry Schigel (GBIF Secretariat) to Everyone: Wesley, I agree, great to have synced thinking. re: what John just said, the more can be specified at the event or study level, the better. Then these things have downstream application all records within. Event force-populated fields in the query.

Difference between Occurrence and OrganismTarget?

Differences between OrganismTarget of an **Event** vs **Occurrence** (+Assertions)?

- Occurrence is needed for voucher/collection
- Occurrence with occurrenceStatus = present/absent can still be used, but it will not contain the information of whether it is a target

Will GBIF provide identifiers for all the entities?

Records in every table will have an identifier.

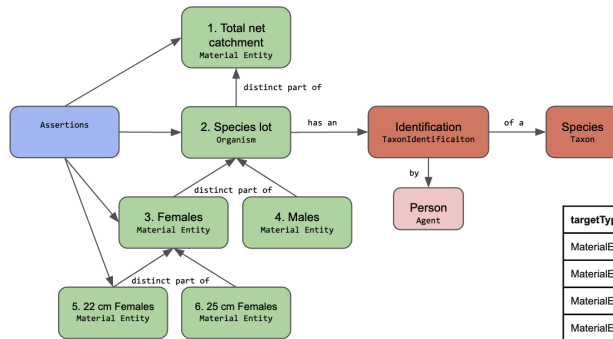
Can the unified model be collapsed into a simple table

Yes, but many fields will have complex lists.

Target body size

Different target body size is under another Community Measurements use case - the [Maine Inshore trawl use case](#)

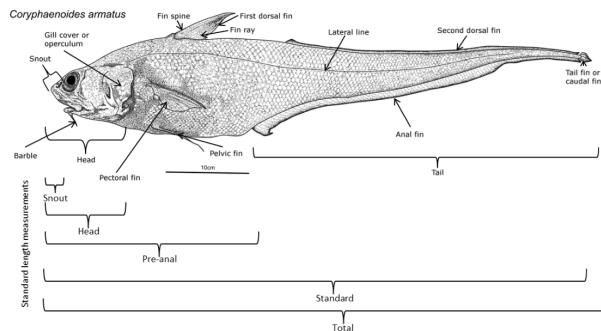
Alternative idea: Measurements recorded against Material Entity subsets



targetType	targetID	assertionType	Value	Unit
MaterialEntity	3	sex	F	
MaterialEntity	6	length	25	CM
MaterialEntity	6	frequency	3	Individuals
MaterialEntity	6	weightedFrequency	3.4	Individuals
MaterialEntity	5	length	22	CM
MaterialEntity	5	frequency	2	Individuals
MaterialEntity	5	weightedFrequency	2.7	Individuals
Organism	2	weight		KG
MaterialEntity	1	weight		KG

targetBodySize/targetBodyLengthInCm was suggested in meeting yesterday

- Can this be implied in samplingEffort (e.g. mesh size) of the sampling gear?
- As far as I know body size/length is defined differently for different taxa:
 - Bird = wing span
 - Fish = standard length, total length



14:59:26 From Wesley Hochachka to Everyone:

With birds, researchers will typically record at least 3 separate aspects of size: mass, wing length, and tarsus (lower-leg) length.

Meeting chat

Meeting chat is saved [here](#)

Current Action Item(s)