COPY WITH COMMENTS - JUNE 14 -(DO NOT REMOVE ANY COMMENT SO WE DO NOT MISS ANYTHING)

Properties of hierarchical Events in Humboldt Extension for ecological inventories

Title: Properties of hierarchical Events in Humboldt Extension for ecological inventories

Date version issued: 2023-xx-xx

Date created: 2023-xx-xx

Part of TDWG Standard: http://www.tdwg.org/standards/450

This version: http://rs.tdwg.org/eco/doc/hierarchy/2023-xx-xx

Latest version: http://rs.tdwg.org/eco/doc/hierarchy/

Abstract:

Contributors: Yi Ming Gan, Wesley Hochachka, John Wieczorek, Yanina V. Sica, Peter Brenton, Anahita J.N. Kazem, Steven J. Baskauf

Creator: TDWG Humboldt Core Task Group

Bibliographic citation: Humboldt Core Task Group. 2023. Properties of hierarchical Events Guidelines. Biodiversity Information Standards (TDWG). http://rs.tdwg.org/eco/doc/xxxx/2023-xx-xx

1 Introduction (non-normative)1.1 Status of the content of this document

Section 3 of this document is normative. The other sections are non-normative.

1.2 RFC 2119 keywords

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

1.3 Namespaces and terminology

The namespace eco: abbreviates http://rs.tdwg.org/eco/terms/ and is used with terms minted for the Humboldt Extension for ecological inventories. dwc: abbreviates http://rs.tdwg.org/dwc/terms/, terms in the main Darwin Core vocabulary namespace. Words in code markup are term IRIs or literal values. The word "organism" is used colloquially and is not used in the technical sense of the dwc:Organism class unless specifically presented as "dwc:Organism".

2 Rationale (non-normative)

Ecological inventories in the context of Darwin Core are considered types of dwc:Events. These inventories or actions, which occur at a specific location over a defined period of time, can take multiple forms ranging from singular observations of individual taxa (1 event: 1 observation, Example 1 in Figure 1) to highly structured and deeply nested observations inside other observations (1 event: 2 sub-events, each sub-event: 2 sub-sub-events, Example 4 in Figure 1).
Such highly nested sampling designs can be translated into a relational database schema of parent-child relationships (1 parent to many children, 1 event to many sub-events) where the parent is the entity on the "one" side of a relation, and each child is an entity on the "many" side of a relation (Figure 1).

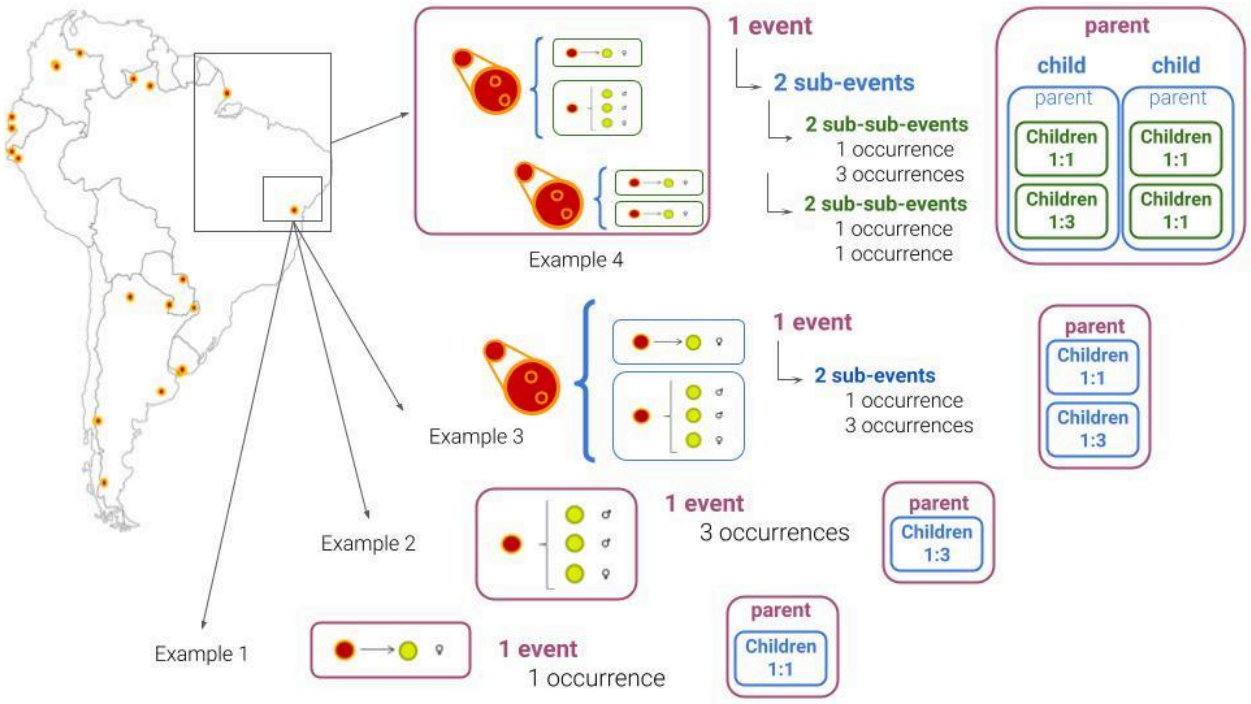

Figure 1: Visual representation of an ecological inventory illustrating examples of occurrence data associated with events being nested within parent events.

3 Usage guidelines (normative)

The objectives of the nested Event hierarchy are to:

- aggregate information as you move upwards through a hierarchy (*i.e.*, from child Events upward to parent Events);
- disaggregate/disambiguate information as you move downwards through a hierarchy (*i.e.*, from parent Events to child Events); and
- ensure that the interpretation of data contained within a set of data (child Events) below a given node level (parent Event) is accurate and comprehensive.

Implementation principles:

1. The value(s) for a given term in a parent Event SHOULD encompass the scope/range of values for that term contained in all of its child Events (see section 4 of this document to see how this is implemented for different types of values).
2. The value of a term in a parent Event may constrain the scope/range of possible values for that term in its child Events.
3. Following Darwin Core principle 4, data providers should filled as many terms as possible (https://dwc.tdwg.org/simple/#5-are-there-any-rules-normative).

For terms in geospatial, temporal and taxonomic scope, it is recommended to complete as much as possible if the scopes were defined *a priori*. In the cases where values for terms included in child Events are not explicitly included in parent Events, it is recommended NOT to infer its value from the child events.

In the cases where values for terms included in child Events are not explicitly included in parent Events, it is reasonable to assume aggregated values in the parent Event based on values from its child Events. However, the inverse is not necessarily true, *i.e*. a parent Event MAY include a statement of scope which is not completely fulfilled by values in all of its child Events.

4 Examples (non-normative)

For the following types of values, these rules apply for point 2 of the implementation principles:

eventID	parentEventID	samplingProtocol	eventRemarks
EVT			voyage
EVT_001	EVT	trawl	
EVT_002	EVT	trawl	
EVT_003	EVT	trawl	

1. Boolean terms: the values for a given term in a parent Event MUST be true if *at least one* of the child Events has a 'true' value.

eventID	hasMaterialSample
EVT	true
EVT_001	false
EVT_002	true
EVT_003	true

2. Non-boolean, non-quantitative & non-hierarchical, life stage.the values for a given term in a parent Event MUST list all values for that term that are contained in the total set of child Events.

eventID	targetTaxonomicScope	targetLifeStageScope
EVT	Myctophidae | Macrouridae	all | larvae & juvenile
EVT_001	Myctophidae	all
EVT_002	Macrouridae	larvae & juvenile

eventID	materialSampleTypes
EVT	
EVT_001	
EVT_002	tissue
EVT_003	stomach content

3. Taxonomy, geography i.e. hierarchical

eventID	siteCount	siteNestingDescription
EVT	125	sampling stations in total for this voyage
EVT_001	1	sampling station
EVT_002	1	sampling station
EVT_003	1	sampling station

4. Terms that are non-boolean & quantitative: the values for a given term in a parent Event MUST include the range of values for that term that are contained in the total set of child Events.

eventID	eventDuration	eventDurationUnit
EVT	45	days
EVT_001	15	minutes
EVT_002	15	minutes
EVT_003	15	minutes

5. Wesley's example: where value appropriate for Child is not applicable at Parent level…

eventID	samplingEffortValue	samplingEffortUnit
EVT		
EVT_001	4.5	km/h
EVT_002	4.5	km/h
EVT_003	4.5	km/h

THE FOLLOWING TEXT WAS MOVED FROM THE NORMATIVE SECTION TO THE NON NORMATIVE SECTION (VERIFY IF IT STILL APPLIES)

A parent Event should reflect the characteristics of its child Events. For example, a scope of a property in a parent Event should encompass the individual values of that property in all of its child Events.

For example, if a parent Event has a child whose eco:targetTaxonomicScope = "Canidae" and another child whose eco:targetTaxonomicScope = "Hyaenidae", then the eco:targetTaxonomicScope for the parent Event would be eco:targetTaxonomicScope = "Canidae | Hyaenidae".

The same basic principle should be applied to boolean properties. For example, if any child of a parent Event has material samples (eco:hasMaterialSamples=true), then the parent Event resulted in material samples and must have eco:hasMaterialSamples = true. If no child Event has eco:hasMaterialSamples = true, then the parent Event MUST have eco:hasMaterialSamples = false.