4D4Life Base Schema ERM v0.2

This document presents an Entity Relationship Model (ERM) of the base schema as an abstract and conceptual representation of its data, that is, the entities, their attributes and the relationships between them.

Entity   A distinguishable object of the real or conceptional world.
Attribute   Property of an entity.
**Relationship** Logical function between two or more entities.

Legend

Taxon

A taxon represents a group of one or more organisms adjudged to be a unit, and having a correct scientific name. This entity is used to represent the taxa that ranks at any level. In this diagram the taxon concept is split into the two sub-concepts of higher and lower taxon with the purpose of emphatizing the difference between them in terms of describing data.

Higher taxon

A taxon that ranks above the species level (e.g. genus, family, order, phyllum, kingdom, …).
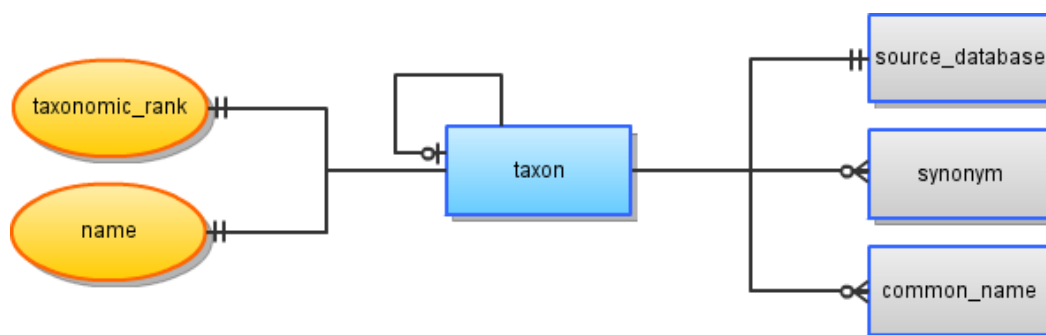
Relationships

taxon *(zero or one)*
An optional reference to the parent taxon.

source_database *(exactly one)*
A mandatory source database.

synonym *(zero or more)*
Synonyms of this taxon.

common_name *(zero or more)*
Common names for this taxon.

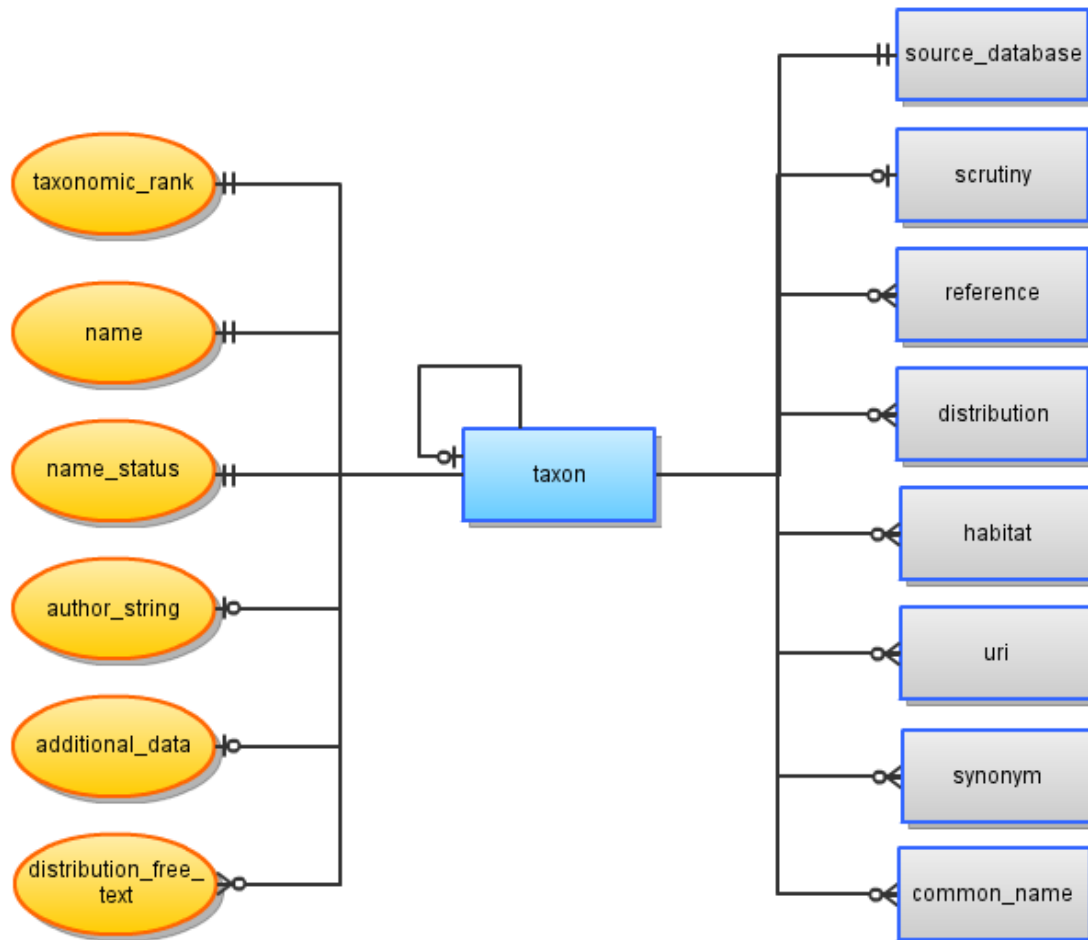**taxonomic_rank** *(exactly one)*
Taxonomic rank of the higher taxon, i.e. Superfamily, Order, Class, etc.

name *(exactly one)*
The accepted, valid or correct scientific name currently accepted for the species as a taxon.

Lower taxon

A taxon that ranks at the species level and below (e.g. subspecies).

Relationships

taxon *(zero or one)*
An optional reference to the parent taxon.

source_database *(exactly one)*
An mandatory source database.

scrutiny *(zero or one)*
An optional reference to the latest taxonomic scrutiny.

reference *(zero or more)*
Optional link to one or more references that accept this taxon in the same taxonomic status, and with the

same name.

distribution *(zero or more)*
Zero or more references to the standarized distribution of the taxon.

habitat *(zero or more)*
Zero or more references to the standarized habitat of the taxon.

uri *(zero or more)*
Optional link to one or more URIs, e.g. LSID, URL.

synonym *(zero or more)*
Synonyms of this taxon.

common_name *(zero or more)*
Common names for this taxon.

Attributes

taxonomic_rank *(exactly one)*
Taxonomic rank of the higher taxon, i.e. Superfamily, Order, Class, etc.

name *(exactly one)*
The accepted, valid or correct scientific name currently accepted for the taxon.

name_status *(exactly one)*
Two variants of name status are possible: 'Accepted name' or 'Provisionally accepted name'.

author_string *(zero or one)*
Name of author(s), who described this species or published current combination, and the year when appropriate.
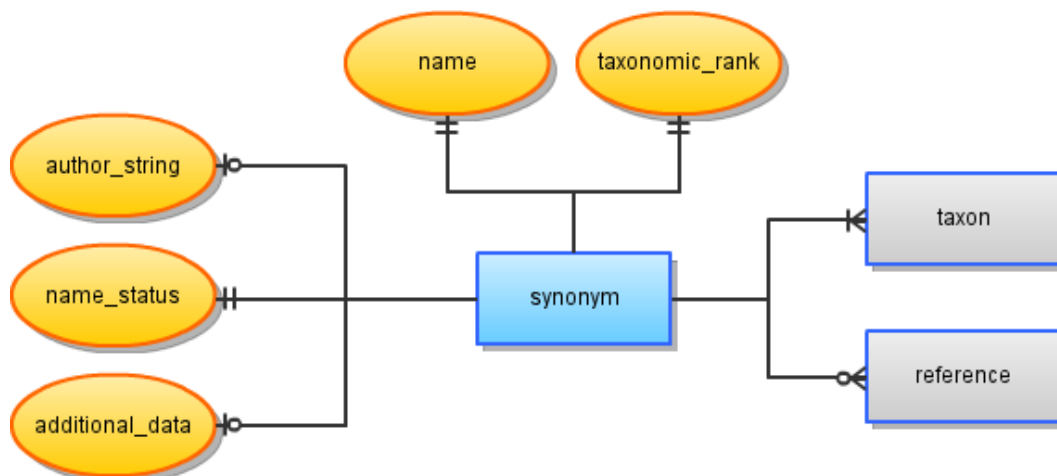
additional_data *(zero or one)*
Optional extra information about the taxa in form of free text.

distribution_free_text *(zero or more)*
Zero to many non-standarized distributions as text.

Synonym

A synonym can be a name which points unambiguously at one species, a name which is ambiguous because it points at the current species and one or more others, e.g. homonyms, pro-parte synonyms, or a name that has been wrongly applied to the current species although may also be correctly applied to another species.

taxon *(one or more)*
A mandatory reference to the taxon to whose the synonym applies to.

reference *(zero or more)*
Optional link to one or more references that accept this species in the same taxonomic status, and with the same name.

Attributes

taxonomic_rank *(exactly one)*
Taxonomic rank of the higher taxon, i.e. Superfamily, Order, Class, etc.

name *(exactly one)*
The accepted, valid or correct scientific name currently accepted for the species as a taxon.

name_status *(exactly one)*
Three variants of name status are possible: 'Synonym', 'Ambiguous synonym' or 'Misapplied name'.

**author_string** *(zero or one)*
Name of author(s), who described this species or published current combination, and the year when appropriate.

additional_data *(zero or one)*
Optional extra information about the synonym in form of free text.

Hybrid

Hybrid refers to offspring resulting from the interbreeding between two species of different taxa. Assuming that a hybrid can be either a synonym or an accepted name, it is defined by the combination of two (or more) of those entities.
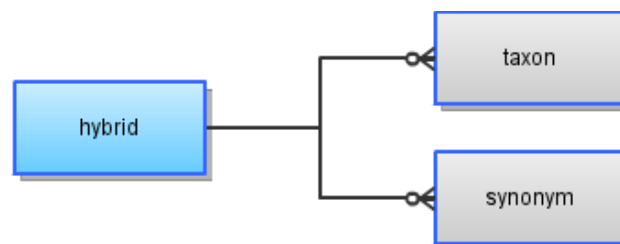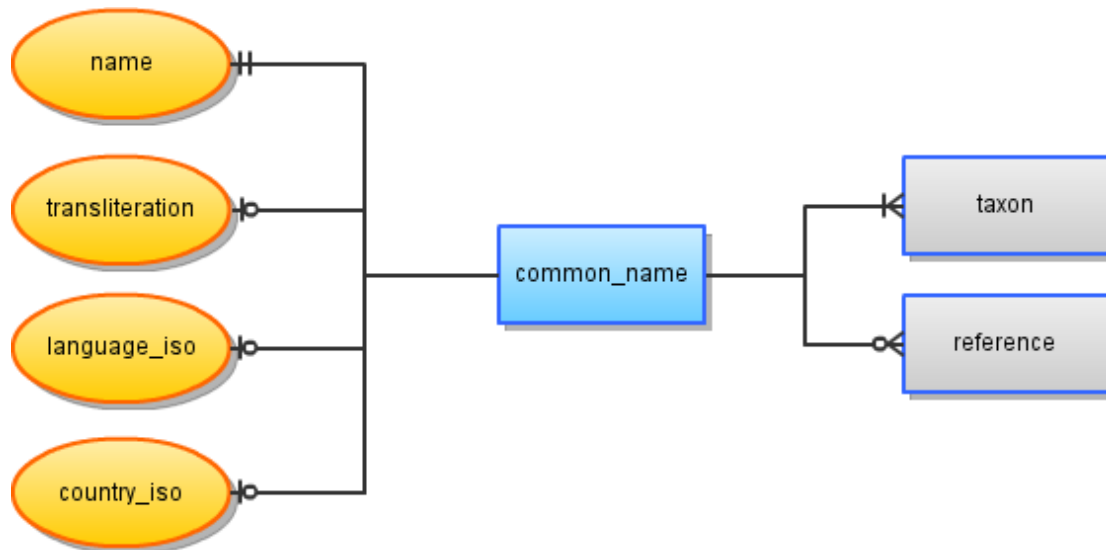

Relationships

taxon *(zero or more)*
Reference(s) to the taxon (or taxa) of which this hybrid is derived from.

synonym *(zero or more)*
Reference(s) to the synonym(s) of which this hybrid is derived from.

Common name

A common name is a name of a species or organism in general use, contrasted with its scientific name.

Relationships

taxon *(one or more)*
The reference(s) to the taxon(s) to which the common name is applied.

reference *(zero or more)*
Link to existing references for this common name.

Attributes

name *(exactly one)*
The accepted, valid or correct scientific name currently accepted for the species as a taxon.

transliteration *(zero or one)*
Transcription of a foreign alphabet into English, when applicable.
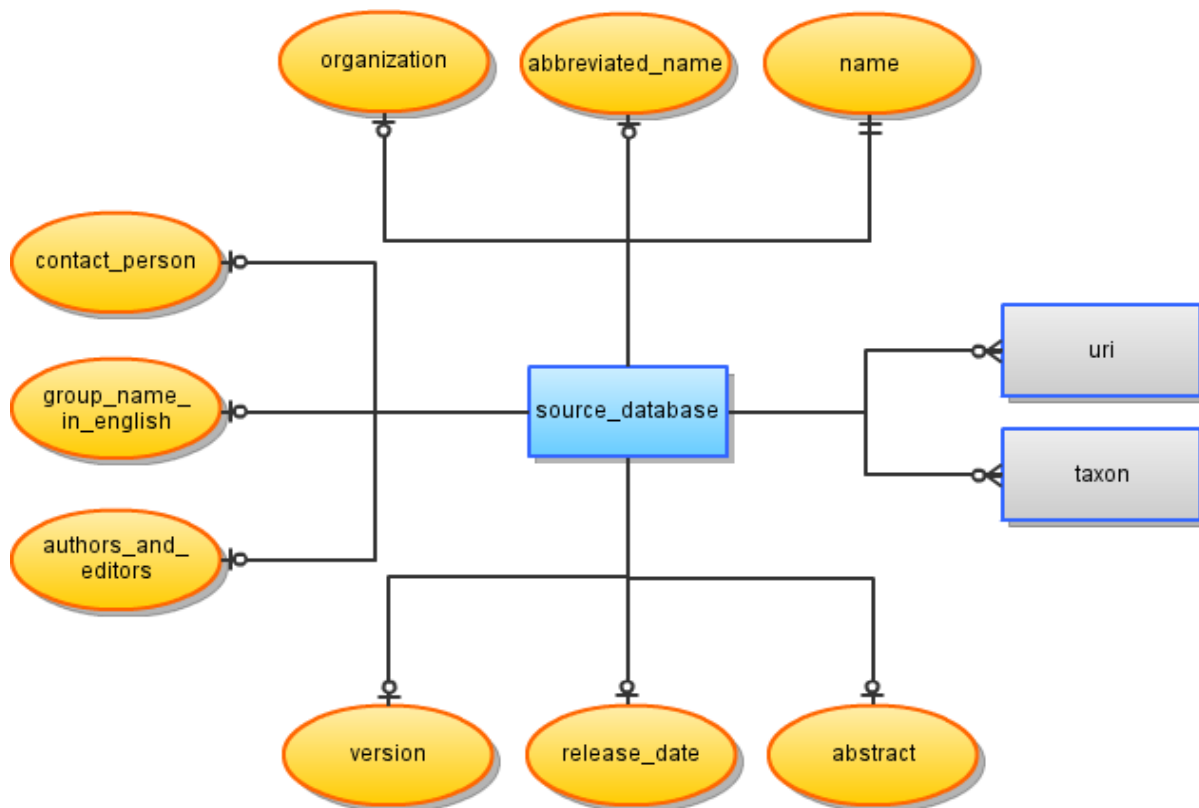
language_iso *(zero or one)*
Optional *ISO 639-3* (three-letter) code of the language of the common name.

country_iso *(zero or one)*
Optional *ISO 3166-1 alpha-2* (two-letter) code of the country where the common name is used.

Source database

A source database represents a GSD (Global Species Database) or an RSD (Regional Species Database).

Relationships

taxon *(zero or more)*
References to all the taxa provided by the source database.

uri *(zero or more)*
Zero or more URIs, e.g. URL to the source database's website.

Attributes

name *(exactly one)*
The qualified name of the source database.

abbreviated_name *(zero or one)*
An abbreviation to the name of the source database.

organization *(zero or one)*
The organization behind the source database.

contact_person *(zero or one)*
The name of the source database's person of contact.

group_name_in_english *(zero or one)*
The name of the taxonomic group(s) covered by the database, in English.

authors_and_editors *(zero or one)*
The names of the authors and editors of the contents provided by the source database.

version *(zero or one)*
The database version.

release_date *(zero or one)*
The release date of the current database version.

abstract *(zero or one)*
Summarized description of the content in the source database.

Reference

A reference represents a Nomenclatural Reference (just one reference which contains the original - validating - publication of taxon name or new name combination), a Taxonomic Acceptance Reference (one or more bibliographic references that accept this species in the same taxonomic status, and with the same name) or a Common Name Reference (one or more bibliographic references that contain common names).
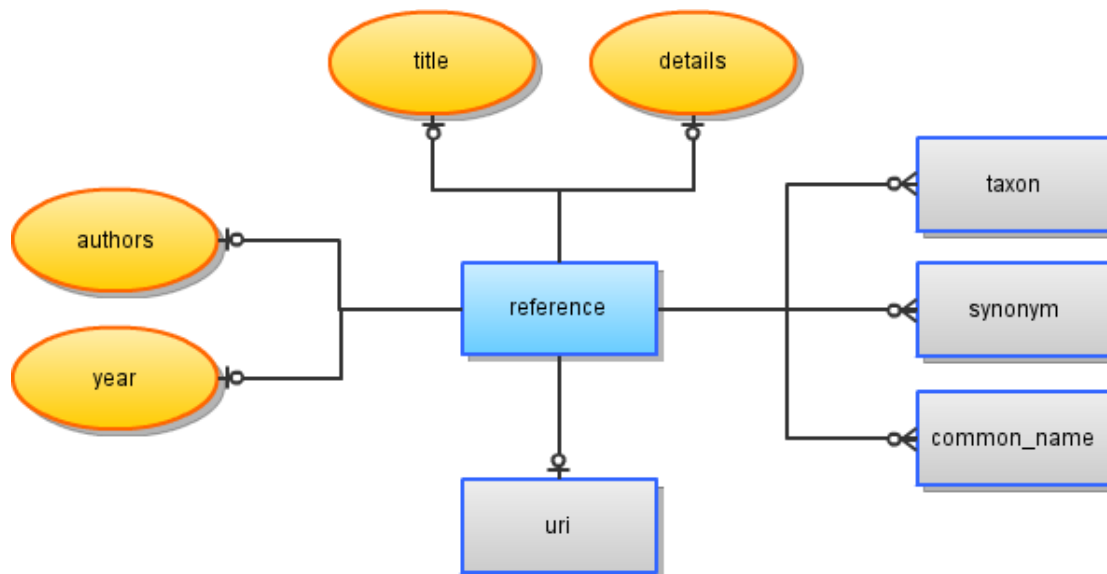
Relationships

taxon *(one or more)*
One or several links to the taxon(s) to which the reference belongs.

synonym *(one or more)*
One or several links to the synonym(s) to which the reference belongs.

common_name *(one or more)*
One or several links to the common name(s) to which the reference belongs.

uri *(zero or one)*
Optional link to one or more URIs, e.g. URL.

Attributes

authors *(zero or one)*
Author(s) of the publication.

year *(zero or one)*
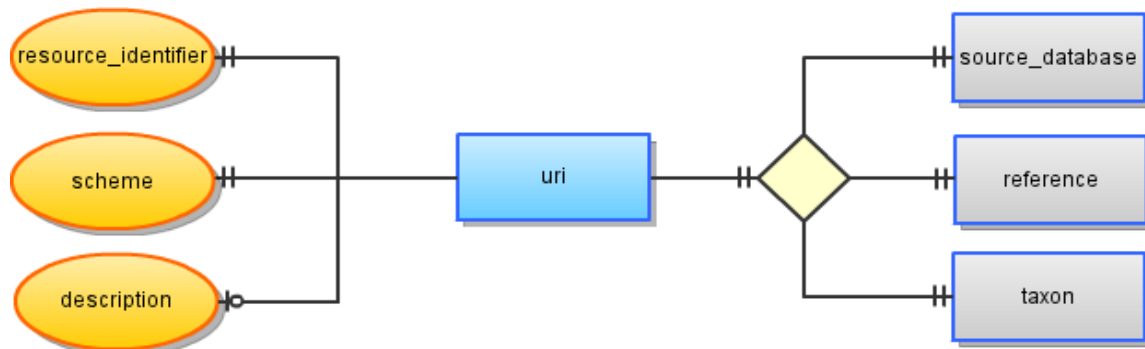Year(s) of publication.

title *(zero or one)*
The title of the paper or book.

details *(zero or one)*
Title of periodicals, volume number, and other common bibliographic details.

URI

A URI (Uniform Resource Identifier) represents a URL (Unified Resource Location), such as a web link, or a URN (Unified Resource Name), such an LSID (Life Science Identifier).

Relationships

source_database *(exactly one)*
The source database to which the URI belongs to.

reference *(exactly one)*
The reference to which the URI belongs to.

taxon *(exactly one)*
The taxon to which the URI belongs to.

Attributes

resource_identifier *(exactly one)*
The fully qualified name of the resource.
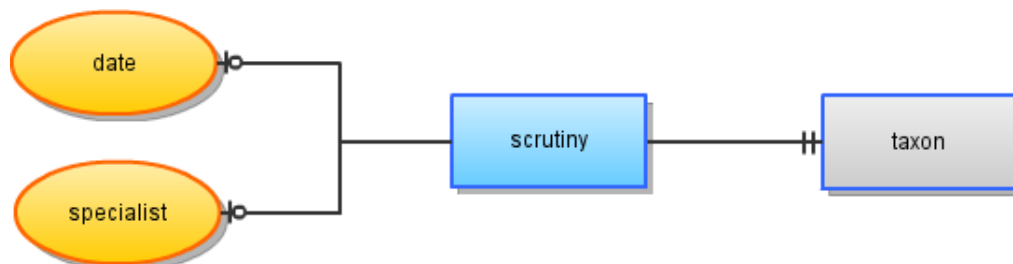
scheme *(exactly one)*
The type (schema) of resource, e.g. http, lsid.

description *(zero or one)*
The descrption text of the resource.

Scrutiny

A taxonomic scrutiny containing the name of the expert who is responsible for taxonomic quality of each species record and the date.

taxon *(exactly one)*
The taxon of which this scrutiny belogns to (latest taxonomic scrutiny).

Attributes
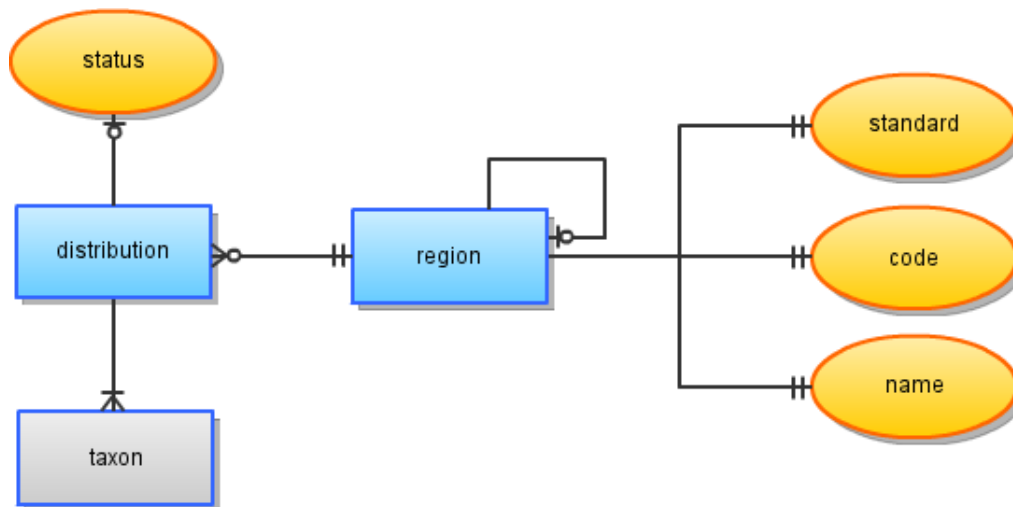
date *(zero or one)*
The date of the taxonomic scrutiny.

specialist *(zero or one)*
The name of the specialist that carried out the taxonomic scrutiny.

Distribution

Distribution is the manner in which a biological taxon is spatially arranged. It is expressed with standarized regions and statuses.

Relationships

taxon *(one or more)*
Link to one or more taxa to which this distribution applies.

region *(exactly one)*

The region to which the distribution applies. A region can belong to several distributions and it can be represented as a hierarchical taxonomy, thus allowing the definition of sub-regions.

Attributes

***distribution*.status** *(zero or one)*
A description of the state of the distribution.

***region*.standard** *(exactly one)*
The standard to which the region belongs to, e.g. TDWG Level-4.
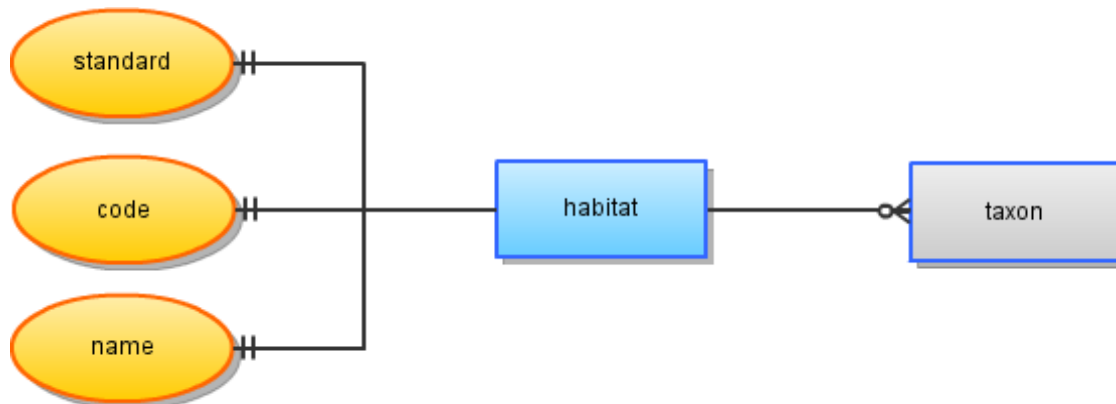
***region*.code** *(exactly one)*
The code of the standard that represents this region.

***region*.name** *(exactly one)*
The name of the region as defined by the standard code.

Habitat

The habitat is the ecological or environmental area that is inhabited by a particular species. Habitats are represented with standard codes.

Relationships

taxon *(zero or more)*
Link to zero or more taxa to which this habitat applies.

Attributes

standard *(exactly one)*
The standard to which the habitat belongs to, e.g. IUCN.

code *(exactly one)*
The code of the standard that represents this habitat.

name *(exactly one)*
The name of the habitat as defined by the standard code.