**Darwin Core Terms**

|  |
| --- |
| **Record Level** |
| Term | Notes |
| type | * The Type Vocabulary used in Darwin Core consists of two parts, vocabulary to describe the record in terms consistent with the Dublin Core Type vocabulary (using the dcterms:type term) and vocabulary to describe the specific biodiversity-related content for a record (using the basisOfRecord term).
* Type terms
	+ PhysicalObject
	+ StillImage
	+ MovingImage
	+ Sound
	+ Event
* The nature or genre of the resource. For Darwin Core, recommended best practice is to use the name of the class that defines the root of the record.
 |
| modified | * The most recent date-time on which the resource was changed. For Darwin Core, recommended best practice is to use an encoding scheme, such as ISO 8601:2004(E).
* Example: YYYY-MM-DD
* Is time of day important or do we just care about the day?
 |
| references | * A related resource that is referenced, cited, or otherwise pointed to by the described resource.
* Another record’s occurrenceID would go here, but since occurrenceID isn’t well defined this is difficult.
 |
| institutionID | * An identifier for the institution having custody of the object(s) or information referred to in the record.
* I suspect this should be a globally unique identifier of some sort. Imagine AMNH could be American Museum of Natural History or African Museum of Natural History.
* See institutionCode
 |
| collectionID | * An identifier for the collection or dataset from which the record was derived. For physical specimens, the recommended best practice is to use the identifier in a collections registry such as the Biodiversity Collections Index
* Example: urn:lsid:biocol.org:col:34818
 |
| institutionCode | * The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record.
* Examples: "MVZ", "FMNH", "AKN-CLO", "University of California Museum of Paleontology (UCMP)".
 |
| collectionCode | * The name, acronym, coden, or initialism identifying the collection or data set from which the record was derived.
* Eamples: "Mammals", "Hildebrandt", "eBird", “Krau Bats”
 |
| basisOfRecord | * The specific nature of the data record - a subtype of the dcterms:type. Recommended best practice is to use a controlled vocabulary such as the Darwin Core Type Vocabulary
* Terms
	+ **Occurrence** -- A resource describing an instance of the Occurrence class.
	+ **Event** – A non-persistent, time-based occurrence. For Darwin Core, a resource describing an instance of the Event Class.
	+ **Location** – A resource describing an instance of the Location class.
	+ **Taxon** -- A resource describing an instance of the Taxon class.
	+ **PreservedSpecimen** – A resource describing a preserved specimen.
	+ **FossilSpecimen** – A resource describing a fossilized specimen.
	+ **LivingSpecimen** – A resource describing a living specimen.
	+ **HumanObservation**—A resource describing an observation made by one or more people.
	+ **MachineObservation** – A resource describing an observation made by a machine.
	+ **NomenclaturalChecklist** -- A resource describing a nomenclatural checklist.
 |
| informationWithheld | * Additional information that exists, but that has not been shared in the given record.
* Examples: "location information not given for endangered species", "collector identities withheld", "ask about tissue samples".
 |
| **Occurrence** |
| Term | Notes |
| occurrenceID | * The occurrenceID is supposed to (globally) uniquely identify an occurrence record, whether it is a specimen-based occurrence, a one-time observation of a species at a location, or one of many occurrences of an individual who is being tracked, monitored, or recaptured. Making it globally unique is quite a trick, one for which we don't really have good solutions in place yet, but one which ontologists insist is essential.
* Doi (digital object identifier) proposed by DataCite and STM
 |
| catalogNumber | * An identifier (preferably unique) for the record within the data set or collection.
* Examples: "2008.1334", "145732a", "145732"
* catalogNumber is an unfortunate name for this term because it suggests a catalog, which suggests a specimen. The definition tries to ameliorate the potential bias by saying that it is a number to identify an occurrence record within a data set or collection. So, it could be a specimen catalog number or it could be a unique identifier for a record within an observation or animal movement data set.
 |
| recordNumber | * For specimens this is the equivalent of a collector's number - the identifier the collector given to a specimen or sample in the field and which is likely to have been written in associated field notes. The same idea applies to original numbers applied to any observation or sample in the field, though the terminology in a given discipline might be distinct. The recordNumber isn't the same as the catalogNumber, which is usually only applied once the specimen gets accessioned into a collection.
* An identifier given to the Occurrence at the time it was recorded. Often serves as a link between field notes and an Occurrence record, such as a specimen collector's number.
 |
| recordedBy | * A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence. The primary collector or observer, especially one who applies a personal identifier (recordNumber), should be listed first.
* Example: "Oliver P. Pearson; Anita K. Pearson" where the value in recordNumber "OPP 7101" corresponds to the number for the specimen in the field catalog (collector number) of Oliver P. Pearson.
 |
| individualID | * individualID is meant for any records that need to identify individuals for whom there may be more than one record. Banded birds, marine mammal photos allowing individual identification, individual trees resampled overtime, periodic biopsies on the same individuals, etc. could all use this term to group the records corresponding to individuals.
* An identifier for an individual or named group of individual organisms represented in the Occurrence. Meant to accommodate resampling of the same individual or group for monitoring purposes. May be a global unique identifier or an identifier specific to a data set.
* Examples: "U.amer. 44", "Smedley", "Orca J 23"
 |
| individualCount | * The number of individuals represented present at the time of the Occurrence.
* Examples: “1”, “25”
 |
| sex | * The sex of the biological individual(s) represented in the Occurrence. Recommended best practice is to use a controlled vocabulary.
* Examples
	+ Unknowable
	+ Undetermined
	+ Female
	+ Male
	+ Hermaphrodite
 |
| lifeStage | * The age class or life stage of the biological individual(s) at the time the Occurrence was recorded. Recommended best practice is to use a controlled vocabulary.
* Examples: "egg", "eft", "juvenile", "adult", "2 adults 4 juveniles"
 |
| reproductiveCondition | * The reproductive condition of the biological individual(s) represented in the Occurrence. Recommended best practice is to use a controlled vocabulary.
* Examples: "non-reproductive", "pregnant", "in bloom", "fruit-bearing"
 |
| behavior | * A description of the behavior shown by the subject at the time the Occurrence was recorded. Recommended best practice is to use a controlled vocabulary.
* Examples: "roosting", "foraging", "running"
 |
| occurrenceStatus | * A statement about the presence or absence of a Taxon at a Location. Recommended best practice is to use a controlled vocabulary.
* Vocabulary:
	+ Present - There is at least one well documented record of the taxon's presence in the area.
	+ Absent - There is evidence to document the absence of a taxon in the area.
	+ Common - The taxon has been observed frequently in the area.
	+ Irregular - The presence of the taxon varies episodically in the area.
	+ Rare - The taxon has been observed infrequently in the area.
	+ Doubtful - The taxon is presumed present in the area, but there is doubt over the evidence, including taxonomic or geographic imprecision in the records.
* **Notes:** Use only "present" or "absent" as the possible values for occurrenceStatus of particular Events. Other values of the vocabulary are permissible for taxon distribution records.
 |
| preparations | * A list (concatenated and separated) of preparations and preservation methods for a specimen.
* Examples: Examples: "skin; skull; skeleton", "whole animal (ETOH); tissue (EDTA)", "fossil", "cast", "photograph", "DNA extract"
 |
| otherCatalogNumbers | * A list (concatenated and separated) of previous or alternate fully qualified catalog numbers or other human-used identifiers for the same Occurrence, whether in the current or any other data set or collection.
* Examples: "FMNH:Mammal:1234", "NPS YELLO6778; MBG 33424"
 |
| associatedMedia | * A list (concatenated and separated) of identifiers (publication, global unique identifier, URI) of media associated with the Occurrence.
* Examples: "http://www.sciencemag.org/cgi/content/abstract/322/5899/261", "Christopher J. Conroy, Jennifer L. Neuwald. 2008. Phylogeographic study of the California vole, Microtus californicus Journal of Mammalogy, 89(3):755-767."
 |
| associatedSequences | * A list (concatenated and separated) of identifiers (publication, global unique identifier, URI) of genetic sequence information associated with the Occurrence.
* Example: "GenBank: U34853.1"
* A semicolon delimited list of sequence identifiers with an optional prefix to indicate their origin. If the prefix is omitted it should be a well known identifier format from one of the INSDC databases (International Nucleotide Sequence Database Collaboration), see <http://www.ddbj.nig.ac.jp/sub/acc_def-e.html>
* Example: AB425960;AB425963;DQ286547
* For other sequence identifiers a prefix indicating the source or a full URL is desirable. The individual INSDC sequence read archives should be prefixed with the following
	+ SRA: NCBI Sequence Read Archive
	+ DRA: DDBJ Sequence Read Archive
	+ ERA: EMBL Sequence Read Archive
 |
| associatedTaxa | * A list (concatenated and separated) of identifiers or names of taxa and their associations with the Occurrence.
* Example: "host: Quercus alba"
 |
| **Event** |
| Term | Notes |
| samplingProtocol | * The name of, reference to, or description of the method or protocol used during an Event.
* Example: "UV light trap", "mist net", "bottom trawl", "ad hoc observation", "point count", "Penguins from space: faecal stains reveal the location of emperor penguin colonies, http://dx.doi.org/10.1111/j.1466-8238.2009.00467.x", "Takats et al. 2001. Guidelines for Nocturnal Owl Monitoring in North America. Beaverhill Bird Observatory and Bird Studies Canada, Edmonton, Alberta. 32 pp.", "http://www.bsc-eoc.org/download/Owl.pdf".
 |
| samplingEffort | * The amount of effort expended during an Event.
* Example: "40 trap-nights", "10 observer-hours; 10 km by foot; 30 km by car".
 |
| eventDate | * The eventDate term can be used to capture specific dates at various levels of resolution as well as periods of time with distinct beginnings and endings. It is for this reason that the legacy beginning and ending dates were not retained from previous versions of Darwin Core. Note that eventDate cannot be used to capture information about events on geological time scales.
* Further explanation of the International Standards Organization standard ISO 8601 can be found at <http://en.wikipedia.org/wiki/ISO_8601>
* Examples
	+ Year only: 2010
	+ Year and month: 2010-01
	+ Year, month, and day: 2010-01-17
	+ Year, month, day, and UTC time: 2010-01-17T09:26Z
	+ Year, month, day, and local time: 2010-01-17T09:33:59-0300
 |
| habitat | * A category or description of the habitat in which the Event occurred.
* "oak savanna", "pre-cordilleran steppe"
 |
| fieldNumber | * An identifier given to the event in the field. Often serves as a link between field notes and the Event.
* Example: “S100001”
 |
| eventRemarks | * Comments or notes about the Event.
* Example: "after the recent rains the river is nearly at flood stage"
 |
| **Location** |
| Type | Notes |
| continent | * The name of the continent in which the Location occurs. Recommended best practice is to use a controlled vocabulary such as the Getty Thesaurus of Geographic Names or the ISO 3166 Continent code.
* Example: AS = Asia
* <http://en.wikipedia.org/wiki/List_of_countries_by_continent_%28data_file%29>
 |
| island | * The name of the island on or near which the Location occurs. Recommended best practice is to use a controlled vocabulary such as the Getty Thesaurus of Geographic Names.
* Example: "Isla Victoria"
 |
| country | * The name of the country or major administrative unit in which the Location occurs. Recommended best practice is to use a controlled vocabulary such as the Getty Thesaurus of Geographic Names.
* <http://www.iso.org/iso/home/standards/country_codes/country_names_and_code_elements_txt.htm>
* ISO 3166-1-alpha-2
 |
| stateProvince | * The name of the next smaller administrative region than country (state, province, canton, department, region, etc.) in which the Location occurs.
* Examples: "Montana", "Minas Gerais", "Córdoba"
 |
| locality | * The specific description of the place. Less specific geographic information can be provided in other geographic terms (higherGeography, continent, country, stateProvince, county, municipality, waterBody, island, islandGroup). This term may contain information modified from the original to correct perceived errors or standardize the description.
* Example: "Bariloche, 25 km NNE via Ruta Nacional 40 (=Ruta 237)"
 |
| minimumElevationinMeters | * The lower limit of the range of elevation (altitude, usually above sea level), in meters.
* Example: 100
 |
| maximumElevationinMeters | * The upper limit of the range of elevation (altitude, usually above sea level), in meters.
* Example: 200
 |
| locationRemarks | * Comments or notes about the Location.
* Example: "under water since 2005"
 |
| decimalLatitude | * The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.
* decimalLatitude and decimalLongitude are always in decimal degrees. The Spatial Reference System for these coordinates must be given in the geodeticDatum term.
* Example: "-41.0983423"
 |
| decimalLongitude | * The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.
* decimalLatitude and decimalLongitude are always in decimal degrees. The Spatial Reference System for these coordinates must be given in the geodeticDatum term.
* Example: Example: "-121.1761111"
 |
| geodeticDatum | * Ideally one should use a standard European Petroleum Survey Group (EPSG) code to a Spatial Reference System as a value for this term. The recommended controlled vocabulary is to use a value consisting of the string "EPSG:" followed by a valid EPSG code corresponding to the spatial reference system used for the geographic coordinates in the terms decimalLatitude and decimalLongitude.
* A good resource for determining the EPSG code from the Datum name can be found at <http://spatialreference.org/>.
* If you don't know the details of the SRS, then it is permissible to provide the name or code of the geodetic datum for the decimalLatitude and decimalLongitude coordinates. Sometimes this information is not on the original source (such as a map), while an ellipsoid name is given. In this case, provide the name of the ellipsoid. If the spatial reference system, datum, or ellipsoid are not known, use the value "unknown" for this term.
* A good reference for datum and ellipsoid names is <http://www.colorado.edu/geography/gcraft/notes/datum/edlist.html>
* Examples:
	+ EPSG:4326 = WGS84 (Common on GPS Units)
	+ EPSG:4269 = NAD83
 |
| coordinatePrecision | * A decimal representation of the precision of the coordinates given in the decimalLatitude and decimalLongitude.
* Example: "0.00001" (normal GPS limit for decimal degrees), "0.000278" (nearest second), "0.01667" (nearest minute), "1.0" (nearest degree).
 |
| **Identification** |
| Type | Notes |
| identifiedBy | * A list (concatenated and separated) of names of people, groups, or organizations who assigned the Taxon to the subject.
* Example:
	+ "James L. Patton"
	+ "Theodore Pappenfuss; Robert Macey"
 |
| dateIdentified | * The date on which the subject was identified as representing the Taxon. Recommended best practice is to use an encoding scheme, such as ISO 8601:2004(E).
* YYYY-MM-DD
 |
| identificationReferences | * A list (concatenated and separated) of references (publication, global unique identifier, URI) used in the Identification.
* Example: "Aves del Noroeste Patagonico. Christie et al. 2004."
* Example: doi:10.1000/182
 |
| typeStatus | * A list (concatenated and separated) of nomenclatural types (type status, typified scientific name, publication) applied to the subject.
* Example: "holotype of Ctenomys sociabilis. Pearson O. P., and M. I. Christie. 1985. Historia Natural, 5(37):388"
* Controlled vocab:
	+ Holotype
	+ Paratype
	+ Neotype
	+ Syntype
	+ Lectotype
	+ Paralectotype
	+ hapantotype
 |
| **Taxon** |
| Type | Notes |
| scientificName | * The full scientific name, with authorship and date information if known. When forming part of an Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the IdentificationQualifier term.
* Example: "Coleoptera" (order), "Vespertilionidae" (family), "Manis" (genus), "Ctenomys sociabilis" (genus + specificEpithet), "Ambystoma tigrinum diaboli" (genus + specificEpithet + infraspecificEpithet), "Roptrocerus typographi (Györfi, 1952)" (genus + specificEpithet + scientificNameAuthorship), "Quercus agrifolia var. oxyadenia (Torr.) J.T. Howell" (genus + specificEpithet + taxonRank + infraspecificEpithet + scientificNameAuthorship).
 |
| family | * The full scientific name of the family in which the taxon is classified.
* Example: "Felidae", "Monocleaceae"
 |
| genus | * The full scientific name of the genus in which the taxon is classified.
* Example: "Puma", "Monoclea"
 |
| subgenus | * The full scientific name of the subgenus in which the taxon is classified. Values should include the genus to avoid homonym confusion.
* "Strobus (Pinus)", "Puma (Puma)" "Loligo (Amerigo)", "Hieracium subgen. Pilosella"
 |
| specificEpithet | * The name of the first or species epithet of the scientificName.
* Examples: "concolor", "gottschei"
 |
| infraspecificEpithet | * The name of the lowest or terminal infraspecific epithet of the scientificName, excluding any rank designation.
* Example: "concolor", "oxyadenia", "sayi".
 |
| taxonRank | * The taxonomic rank of the most specific name in the scientificName. Recommended best practice is to use a controlled vocabulary.
* Vocabulary:
	+ Kingdom
	+ Subkingdom
	+ Division or phylum
	+ Subdivision or subphylum
	+ Class
	+ Subclass
	+ Order
	+ Suborder
	+ Family
	+ Subfamily
	+ Tribe
	+ Subtribe
	+ Genus
	+ Subgenus
	+ Section
	+ Subsection
	+ Series
	+ Subseries
	+ Species
	+ Subspecies
	+ Variety
	+ Subvariety
	+ Form
 |
| scientificNameAuthorship | * The authorship information for the scientificName formatted according to the conventions of the applicable nomenclaturalCode.
* Example: "(Torr.) J.T. Howell", "(Martinovský) Tzvelev", "(Györfi, 1952)"
 |