**Taller GBIF.ES: Estandarización, documentación y publicación de datos de seguimientos de biodiversidad**

28 – 30 de octubre de 2025, Universidad de Sevilla

# Estándares para la información de la biodiversidad

**Francisco Pando**
**GBIF España**

SECEM
Sociedad Ibérica para la Conservación y Estudio de los Mamíferos

IUCN SSC
GRUPO DE ESPECIALISTAS DE ESPECIES ESPAÑA

UNIVERSIDAD DE SEVILLA

GBIF.ES

CSIC

# Guión

- Datos y ciencia
- Ciencia abierta y estándares
- Tipos de estándares
- TDWG
- Estándares en GBIF
  - Darwin Core
  - Humboldt Core
  - Plinian Core
- Darwin Core Archive
- GBIF new data model
- Darwin Core – DP

# Elementos que hacen la ciencia abierta

- Metadatos
- Estándares
- Identificadores
- Licencias
- Citación
- El cómo se sirve el dato

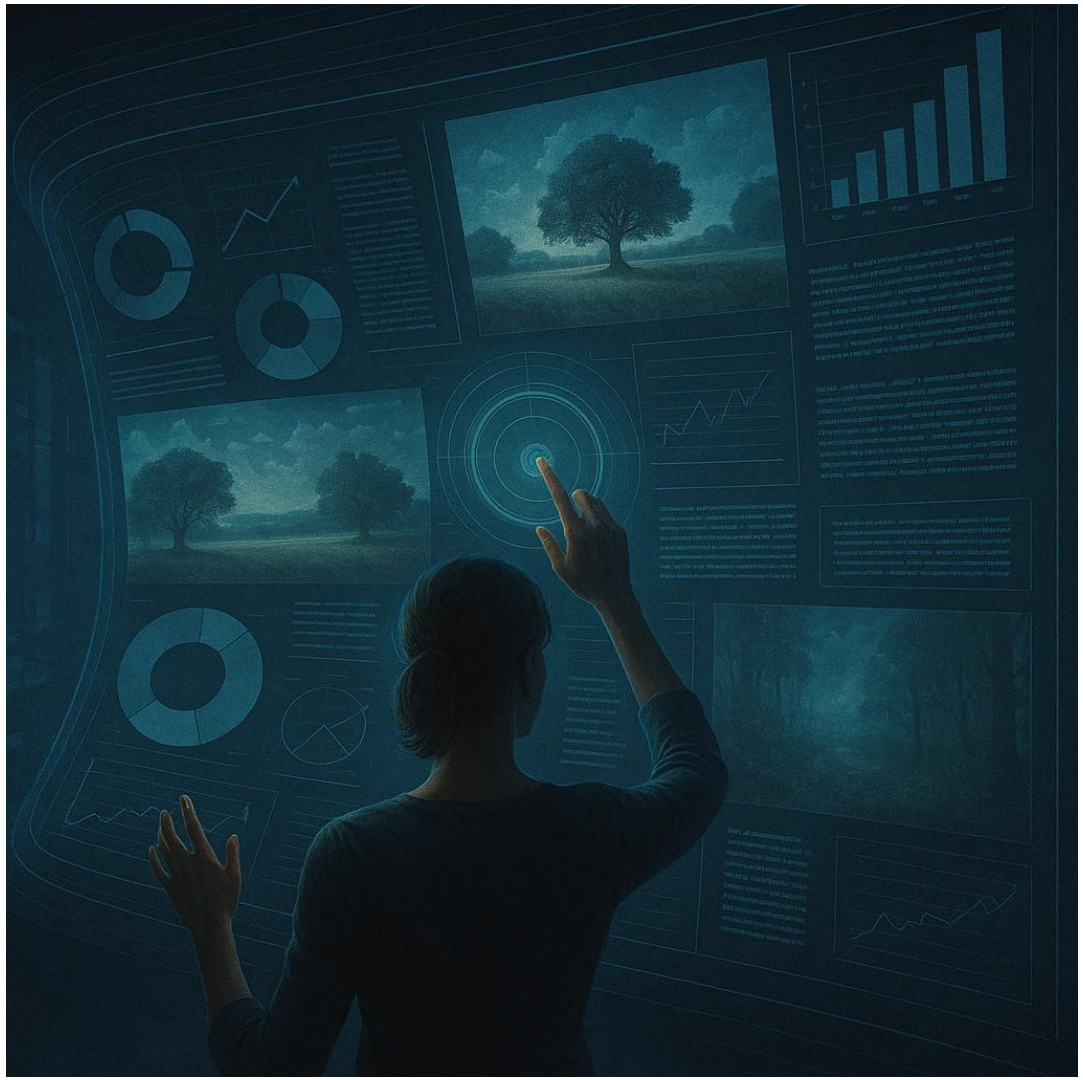
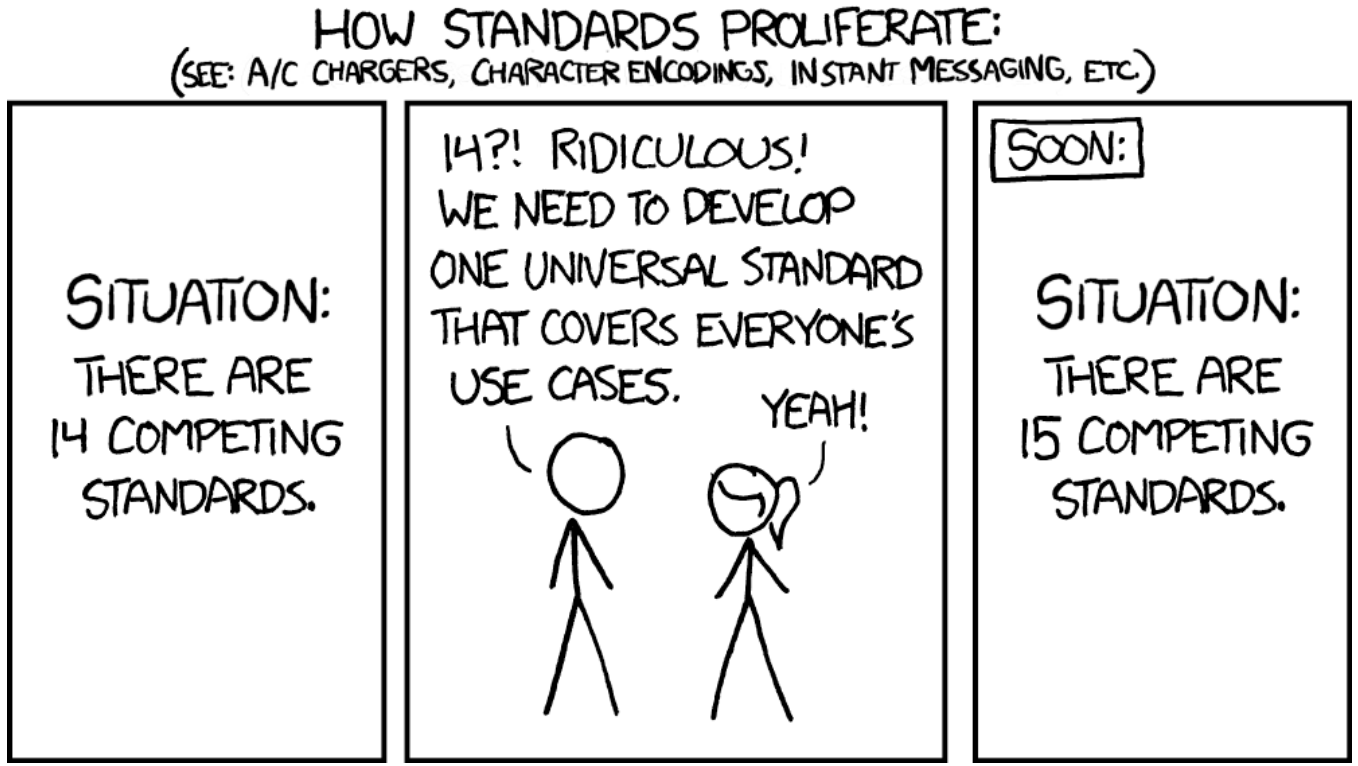


Imagen generada con IA

# Estándares



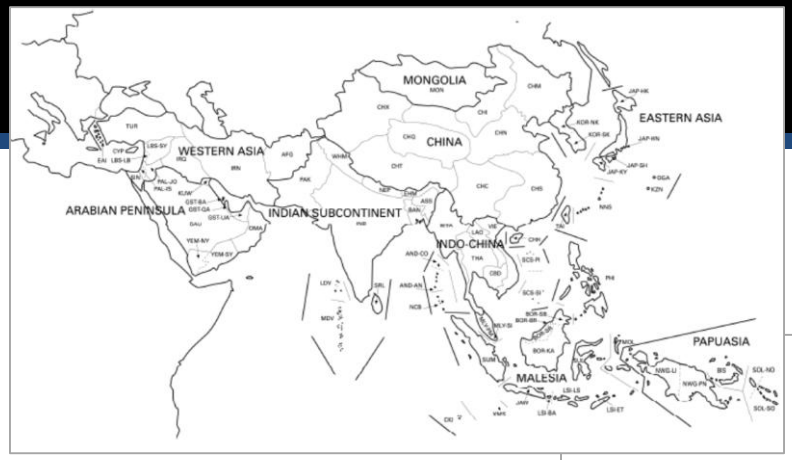

- Si lo cambias, ya no es un estándar
- Si solo lo usas tu, no es un estándar
- Los estándares son dinámicos y están, en constante evolución: revisiones y ediciones

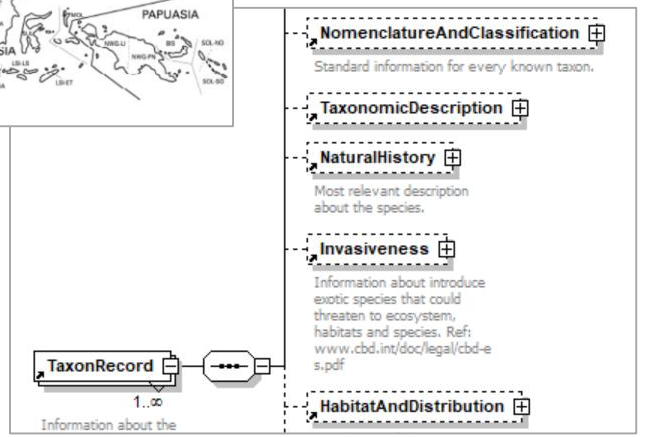
https://xkcd.com/927/

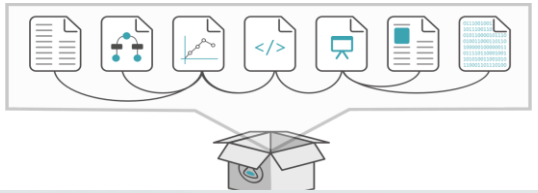
“pensaba que después del COVID no tendría que contar estas cosas…”

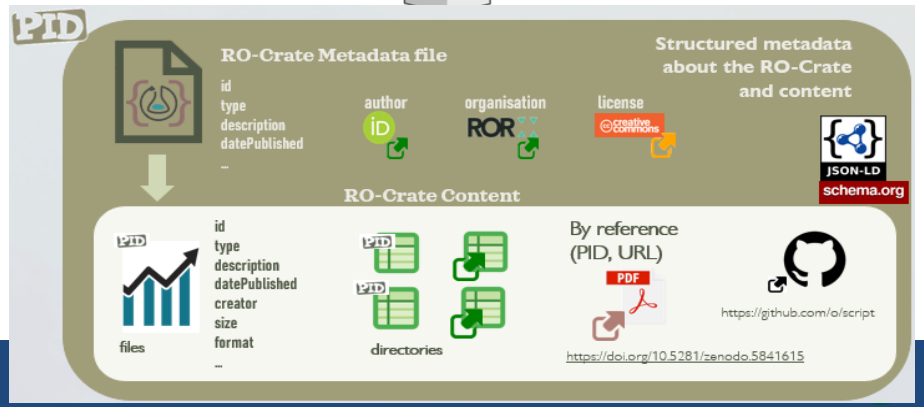
# Estándares

- Vocabularios controlados (léxicos)
  - Que valores son válidos
  - (listas de géneros, de países…)
- Definiciones (semánticos)
  - para humanos, para máquinas
  - (v.gr.: Darwin Core [terms])
- Estructurales (sintácticos)
  - Como se codifica la información
  - (v.gr.: Frictionless Data Packages)





# “Estándares orquestados”

- CC es un vocabulario controlado
- Dublin Core es un estándar de términos
- XML es un estándar sintáctico

Xml

```xml
<metadata>
  <dc:title>Example Document</dc:title>
  <dc:creator>John Doe</dc:creator>
  <dc:subject>Metadata Example</dc:subject>
  <dc:description>This is an example of a metadata section in XML.</dc:des
  <dc:publisher>Example Publisher</dc:publisher>
  <dc:date>2025-01-23</dc:date>
  <dc:type>Text</dc:type>
  <dc:format>application/pdf</dc:format>
  <dc:identifier>http://example.com/document</dc:identifier>
  <dc:language>en</dc:language>
  <dc:rights>Creative Commons Attribution-NonCommercial 4.0 International
</metadata>
```

# Biodiversity Information Standards (TDWG)

We are a non-profit organization and a community dedicated to developing biodiversity information standards.

https://www.tdwg.org

## ABCD

Access to Biological Collection Data (ABCD) Schema

## Audiovisual Core

Audiovisual Core Multimedia Resources Metadata Schema

## Darwin Core

Darwin Core

## GUID Applicability Statements

GUID and Life Sciences Identifiers Applicability Statements

## Latimer Core

Latimer Core, a standard for describing collections

## SDS

TDWG Standards Documentation Standard (SDS)

## TAPIR

TDWG Access Protocol for Information Retrieval (TAPIR)

## TCS

Taxon Concept Schema (TCS)

## VMS

Vocabulary Maintenance Standard (VMS)

## WGSRPD

World Geographical Scheme for Recording Plant Distributions (WGSRPD)

## Interest and Task Groups

- Access to Biological Collections Data
- Annotations
- Attribution
  - People in Biodiversity Data
- Audiovisual Core
  - 3D Imagery and Data
  - Views Controlled Vocabularies
- Biodiversity Data Quality
  - Framework on data quality
  - Data quality tests and assertions
  - Data quality use cases
  - Best practices for development of vocabularies of values ("Vocabularies")
- Biodiversity Informatics Curriculum
- Biodiversity Services and Clients
- Biological Interactions Data
- Citizen Science
- Collection Descriptions
  - Latimer Core
  - Minimum Information about a Digital Specimen (MIDS)
  - Modelling Research Expeditions
- Darwin Core
  - Invasive Species
  - Plant Phenology
  - Realm and Biome
  - Sensitive Species Extension
- Earth Sciences and Paleobiology
  - Darwin Core Chronometric Age Extension
  - Extension for Geosciences (EFG)
  - Mineralogy Extension
- Genomic Biodiversity
  - Environmental Samples and eDNA
  - Sustainable DarwinCore MIxS Interoperability
- Geographical Schemes
  - Marine Geo Schemes Task Group
  - Terrestrial GeoSchemes TG
- Machine Observations
- Observations & Specimens
  - How Did It Die?
  - Humboldt Extension
  - Material Sample
- Process
- Species Information
  - Invasive Organism Information
  - Plinian Core
- Taxon Concept Schema (TCS) Maintenance Group
  - Taxonomic Concept Schema (TCS2) Task Group
- Technical Architecture Group
  - Standards Mapping Task Group

# Biodiversity Information Standards (TDWG)

We are a non-profit organization and a community dedicated to developing biodiversity information standards.

https://www.tdwg.org

## ABCD

Access to Biological Collection Data (ABCD) Schema

## Audiovisual Core

Audiovisual Core Multime ources Metadata Schema

## Darwin Core

Darwin Core

## GUID Applicability St ments

GUID and Life Sciences Iden pplicability Statements

## Latimer Core

Latimer Core, a standard for describing collections

## SDS

TDWG Standards Documentation Standard (SDS)

## TAPIR

TDWG Access Protocol for Information Retrieval (TAPIR)

## TCS

Taxon Concept Schema (TCS)

## VMS

Vocabulary Maintenance Standard (VMS)

## WGSRPD

World Geographical Scheme for Recording Plant Distributions (WGSRPD)

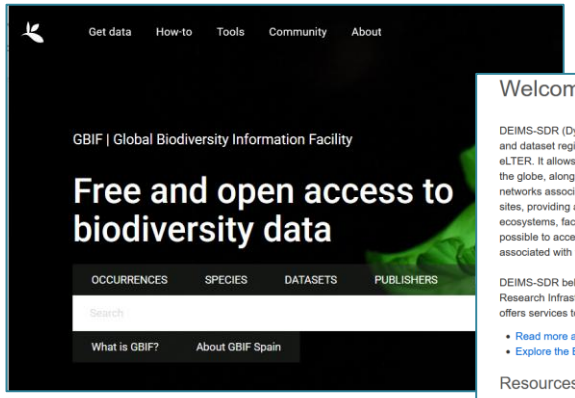
## Interest and Task Groups

- MG Access to Biological Collections Data
- IG Annotations
- IG Attribution
  - TG People in Biodiversity Data
- MG Audiovisual Core
  - TG 3D Imagery and Data
  - TG Views Controlled Vocabularies
- IG Biodiversity Data Quality
  - TG Framework on data quality
  - TG Data quality tests and assertions
  - TG Data quality use cases
  - TG Best practices for development of vocabularies of values ("Vocabularies")
- IG Biodiversity Informatics Curriculum
- IG Biodiversity Services and Clients
- IG Biological Interactions Data
- IG Citizen Science
- IG Collection Descriptions
  - MG Latimer Core
  - TG Minimum Information about a Digital Specimen (MIDS)
  - TG Modelling Research Expeditions
- MG Darwin Core
  - TG Invasive Species
  - TG Plant Phenology
  - TG Realm and Biome
  - TG Sensitive Species Extension
- IG Earth Sciences and Paleobiology
  - TG Darwin Core Chronometric Age Extension
  - TG Extension for Geosciences (EFG)
  - TG Mineralogy Extension
- IG Genomic Biodiversity
  - TG Environmental Samples and eDNA
  - TG Sustainable DarwinCore MIxS Interoperability
- IG Geographical Schemes
  - TG Marine Geo Schemes Task Group
  - TG Terrestrial GeoSchemes TG
- IG Machine Observations
- IG Observations & Specimens
  - TG How Did It Die?
  - TG Humboldt Extension
  - TG Material Sample
- IG Process
- IG Species Information
  - TG Invasive Organism Information
  - TG Plinian Core
- MG Taxon Concept Schema (TCS) Maintenance Group
  - TG Taxonomic Concept Schema (TCS2) Task Group
- IG Technical Architecture Group
  - TG Standards Mapping Task Group

# EML (Ecological Metadata Language)

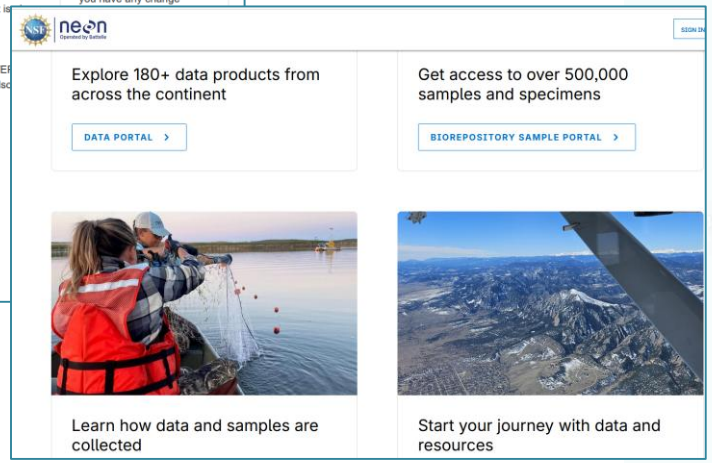
>> estándar de código abierto y basado en XML, diseñado para documenta datos de investigación en ciencias ambientales y ecológicas. Permite describir el contenido, alcance espacial y temporal, métodos y protocolos los datos

https://www.gbif.org



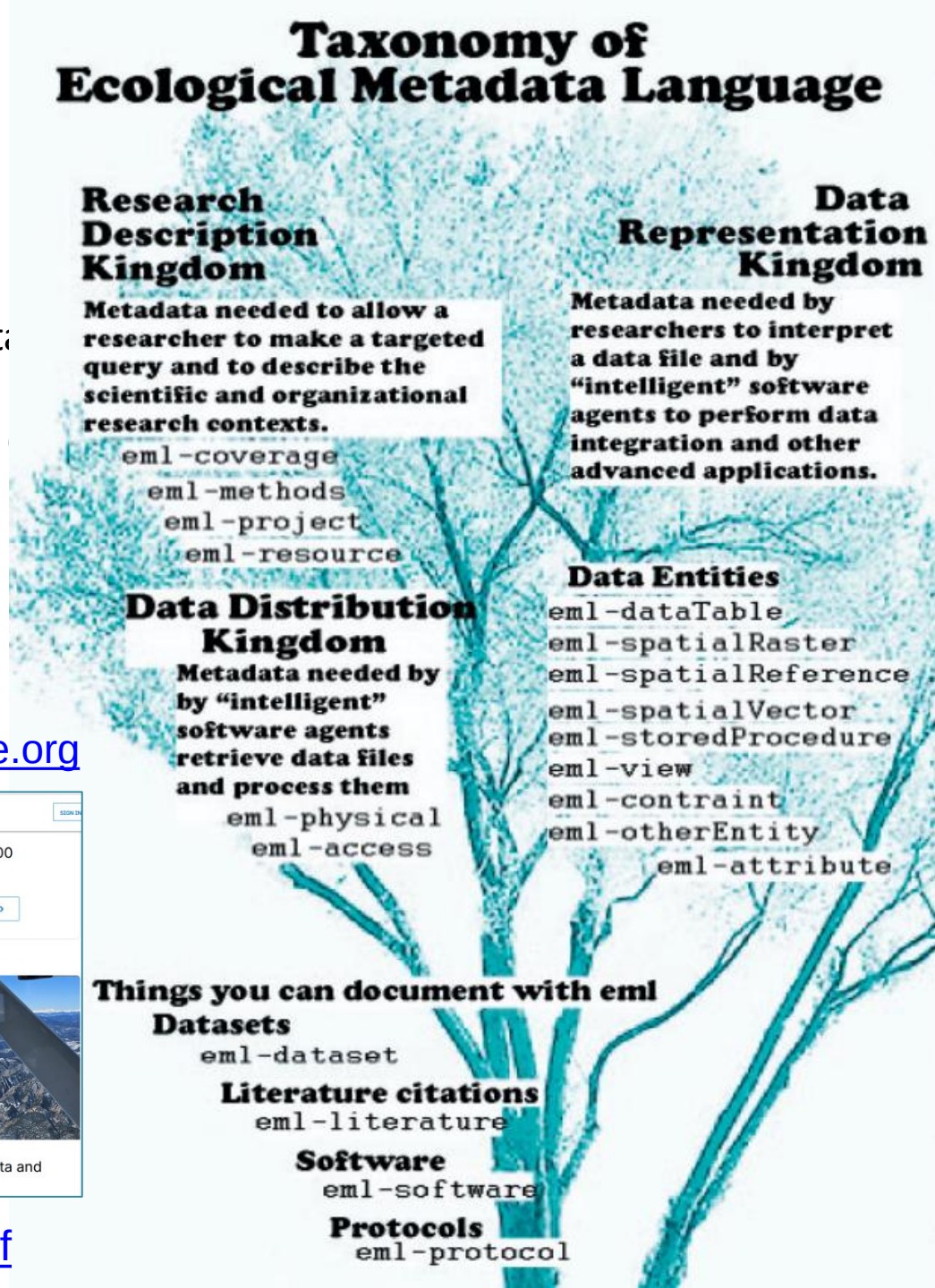

https://deims.org


www.neonscience.org




https://eml.ecoinformatics.org/

https://lternet.edu/wp-content/uploads/2010/12/EMLHandbook-2.pdf



# Darwin Core

| | | |
|---|---|---|
| Record-level Terms | Dublin Core terms, institutions, collections, nature of data record | Simple Darwin Core (flat) |
| Occurrence | evidence of species in nature, observers, behavior, associated media, references. | |
| Event | sampling protocols and methods, date, time, field notes | |
| Location | geography, locality descriptions, spatial data | |
| Identification | linkage between Taxon and Occurrence | |
| Taxon | scientific names, vernacular names, names usages, taxon concepts, and the relationships between them | |
| GeologicalContext | geologic time, chrono-stratigraphy, biostratigraphy, lithostratigraphy | |
| ResourceRelationship | explicit relationships between identified resources (e.g., one organism to another, taxon to location, etc.) | Generic Darwin Core (relational) |
| MeasurementOrFact | measurements, facts, characteristics, assertions, references | |

https://www.researchgate.net/figure/Darwin-Core-Categories-Simple-Darwin-Core-is-comprised-of-seven-categories-of-terms_fig3_221741592

TDWG Darwin Core

**Darwin Core Quick Reference Guide**

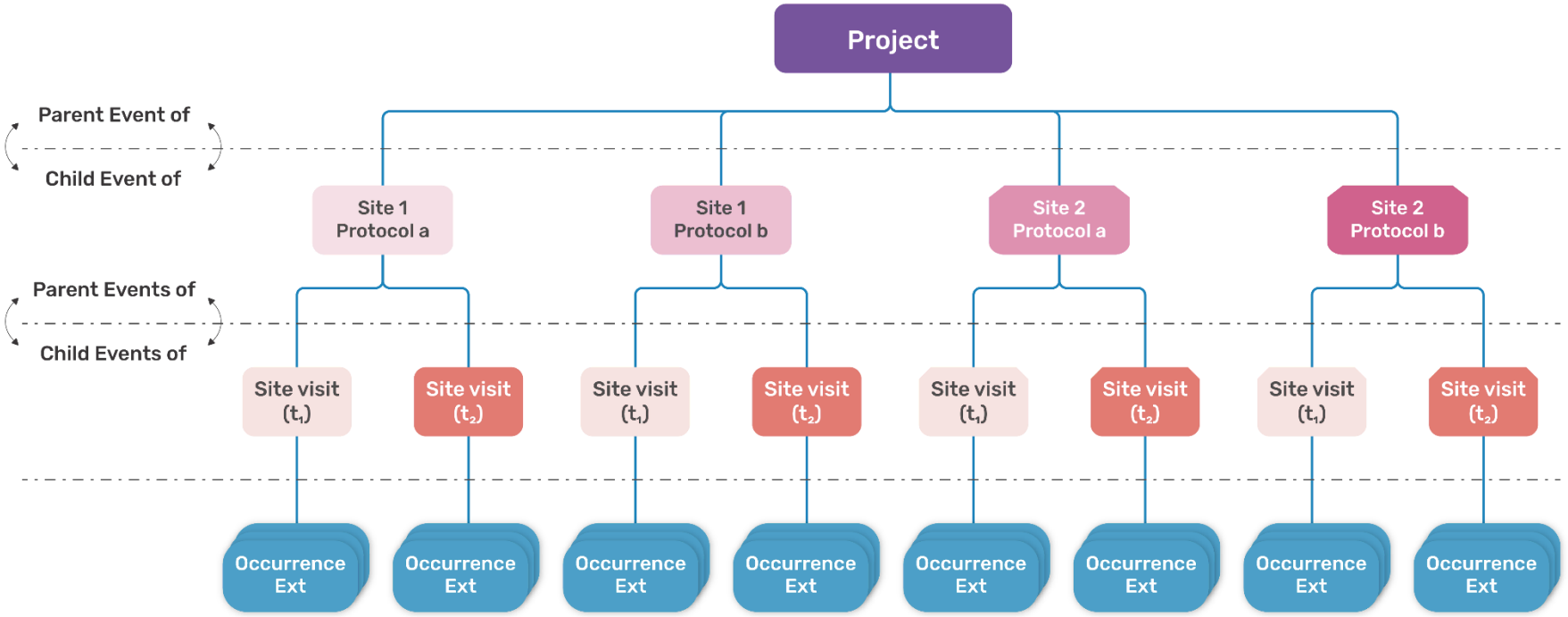
https://dwc.tdwg.org/terms/

Record-level
Occurrence
Organism
MaterialEntity
MaterialSample
Event
Location
GeologicalContext
Identification
Taxon
MeasurementOrFact
ResourceRelationship
UseWithIRI
LivingSpecimen
PreservedSpecimen
FossilSpecimen
MaterialCitation
HumanObservation
MachineObservation

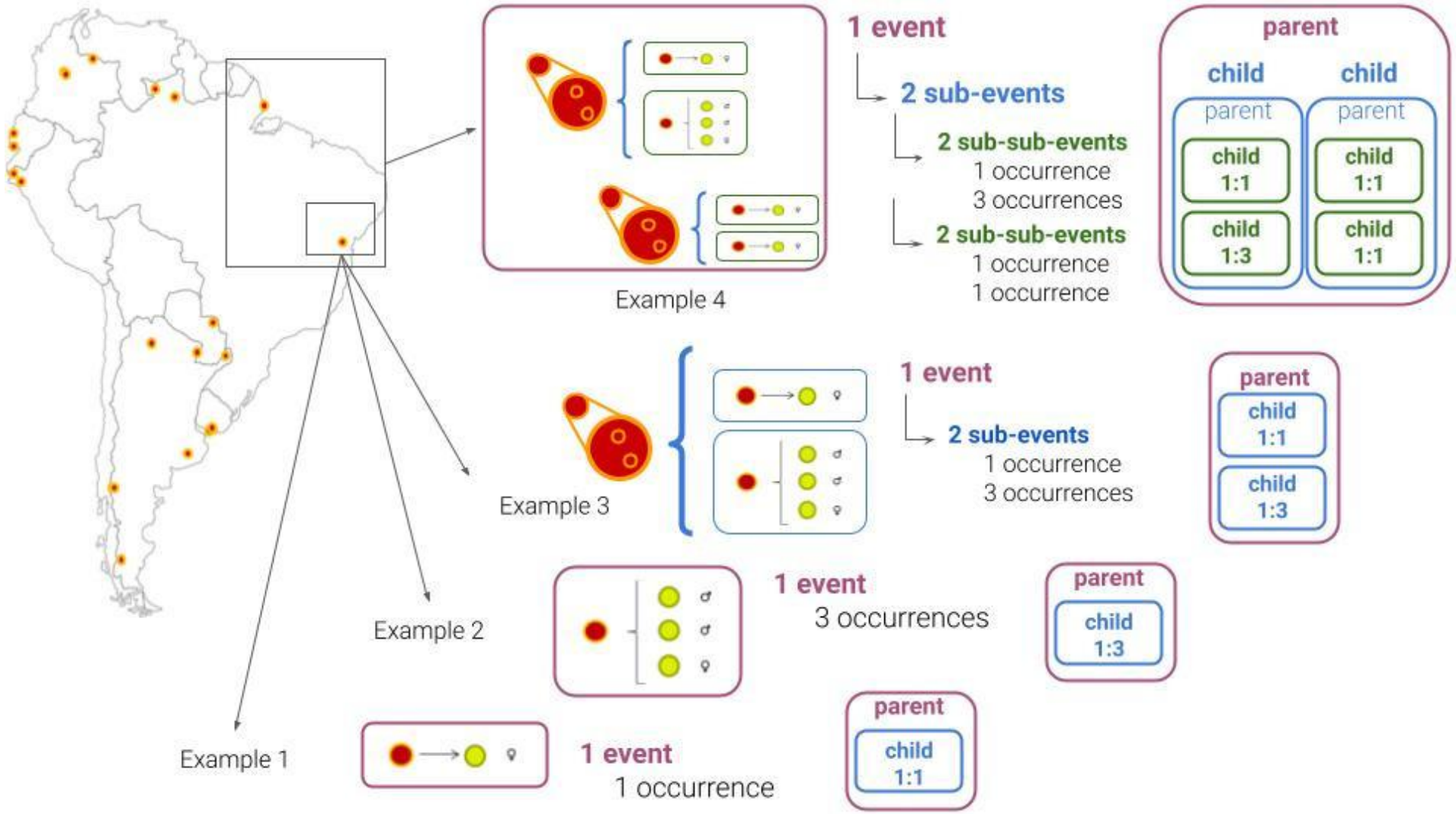
# Humboldt Core

Es una extensión del estándar Darwin Core (DwC) es un vocabulario estructurado (términos y definiciones) desarrollada para mejorar la descripción de inventarios y muestreos estructurados de biodiversidad. Sirve para describir y documentar datos tanto procedentes de muestreos tradicionales como los recogidos por redes de sensores-

- 55 elementos
- Jerárquico
- Ratificado por TDWG en 2024
- https://eco.tdwg.org
- Humboldt Extension for Ecological Inventories: User Guide
- https://docs.gbif.org/survey-monitoring-quick-start/en



# Humboldt Core

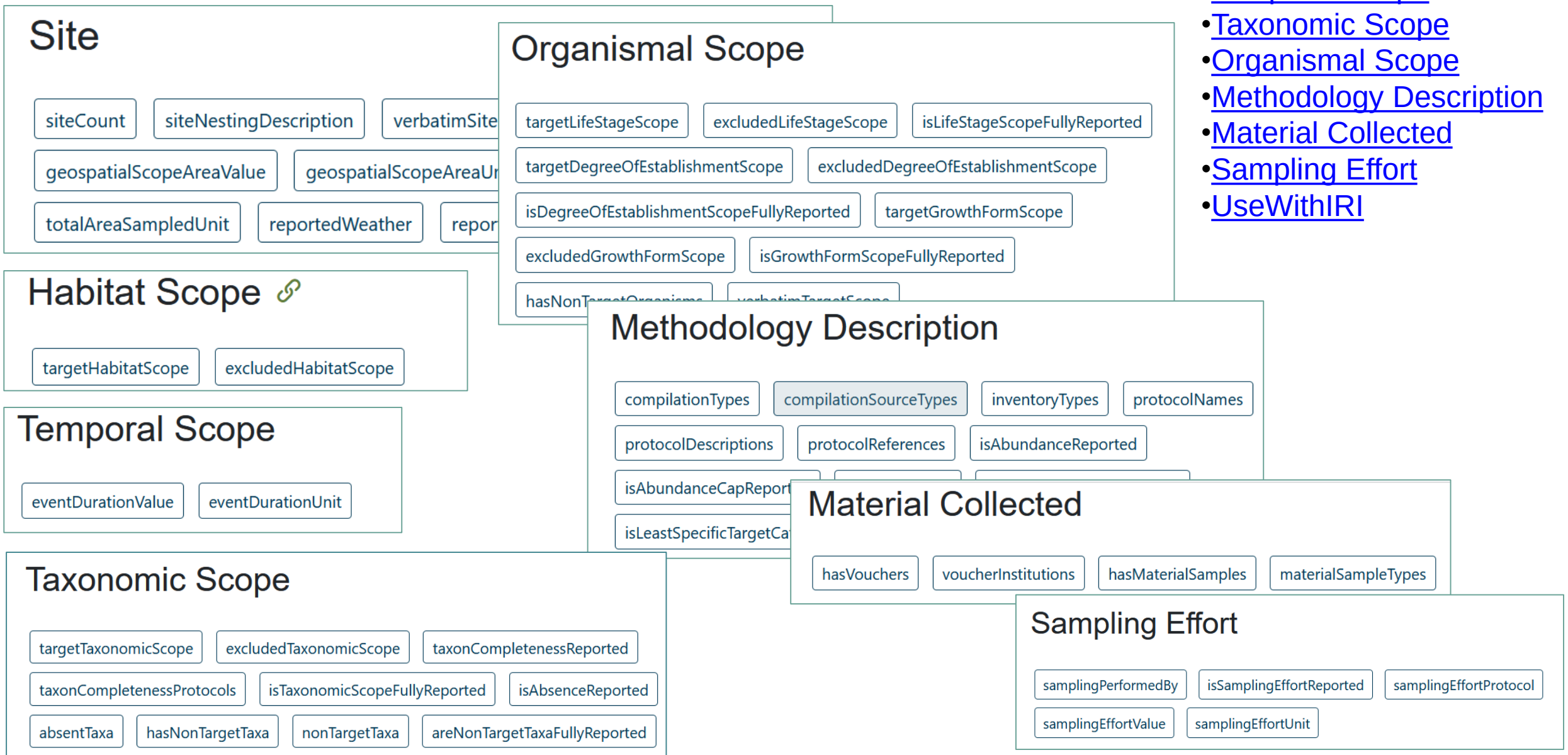



Visual representation of a biodiversity inventory illustrating the four examples of inventory data ranging from simple, individual occurrences (Example 1) to more complex, hierarchically structured data

Humboldt Extension for Ecological Inventories: User Guide

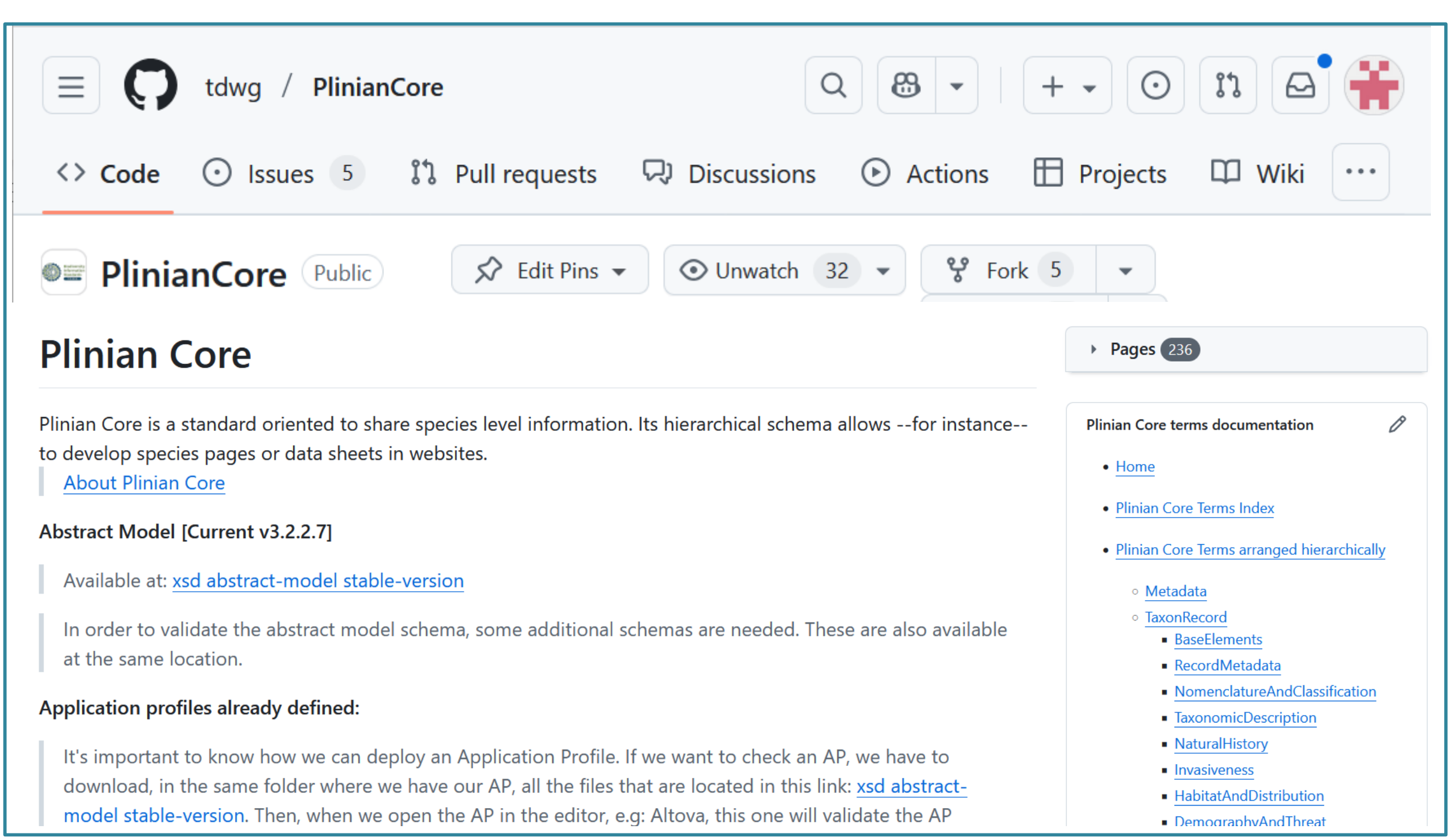
# Humboldt Core

- Site
- Habitat Scope
- Temporal Scope
- Taxonomic Scope
- Organismal Scope
- Methodology Description
- Material Collected
- Sampling Effort
- UseWithIRI



# Plinian Core

- Compartir información a nivel de especies (un estándar de intercambio).
- Interoperabilidad: expresado en xml, se apoya en otros estándares
- Información = propiedades y traits relativos al taxón: descripciones, nomenclatura, estado de conservación, manejo, historia natural..
- Este proyecto ha sido liderado por:
  - El Instituto Nacional de Biodiversidad de Costa Rica (INBio, Costa Rica)
  - Nodo Español de GBIF (GBIF, España)
  - La Universidad de Granada (UG, España)
  - El Instituto Alexander von Humboldt (IAvH, Colombia)
  - La Comisión Nacional del Conocimiento y el Uso de Biodiversidad de México (CONABIO)
  - La Universidad de Sao Paulo y SiBBr  Brazil)
- Auto-contenido y con capacidad de dar soporte a la integración de datos en múltiples bases de datos, con diferentes niveles de granularidad.
- Clases: 13 Genéricas, 27 Complejas y 22 Simples.
- Adoptado por el Ministerio de Agricultura….. Y Medo Ambiente de España como estándard
- Parte de los procesos del la organización “Biodiversity Information Standards” TDWG: http://www.tdwg.org/




https://github.com/tdwg/PlinianCore

# Términos

## Term index

This quick guide provides a list of all current elements of the Plinian core.
elements are organized by categories (in bold) in the index. The categorie
correspond to Plinian Core elements that are classes (elements that have
elements to describe them). The elements that describe a given class (the
properties) appear in the list immediately below the name of the category
The index provides links to the element descriptions in the table below the

### Metadata

Dataset_ID | dateStamp | citation | eml-dataset | resourceLogoUrl | Refere

### Base Elements

TaxonRecordID | taxonConceptID | GlobalUniqueIdentifier | Abstract

### Record Metadata

Language
(TargetAudiences) > Audience | AudiencesUnstructured
(Version) > Major | Minor | Modifier | DateIssued | PreferredFlag
(Revision) > associatedParty | pubDate | created

### Nomenclature and Classification

(TaxonRecordName) > ScientificName
(Synonyms) > SynonymName | SynonymStatus | SynonymsUnstructured
(CommonNames) > Name | Language | UsedIn | UsedBy |
CommonNamesUnstructured
(Hierarchy) > kingdom | phylum | class | order | family | genus | subgenus
| specificEpithet | infraspecificEpithet | higherClassification | ParentTaxon
(MiscDetails) > MeasurementOrFact | DetailUnstructured
NomenclatureAndClassificationUnstructured

### Taxonomic Description

BriefDescription
(FullDescription) > MeasurementOrFact | FullDescriptionUnstructured
(IdentificationKeys) > Keys
TaxonomicDescriptionUnstructured

### Natural History

(LifeForm) > MeasurementOrFact | LifeFormUnstructured
(LifeCycle) > MeasurementOrFact | LifeCycleUnstructured
(Reproduction) > MeasurementOrFact | ReproductionUnstructured
(AnnualCycles) > Event | starTimeInterval | endTimeInterval |
AnnualCyclesUnstructured
(Feeding) > Strategy | StrategyRemarks | FeedingUnstructured
(Dispersal) > Purpose | Type | StructureDispersed | Distance | DispersalUnstr
(Behavior) > MeasurementOrFact | BehaviorUnstructured
(Interactions) > InteractionSpecies | InteractionSpeciesType | InteractionUnstr
(MolecularData) > MeasurementOrFact | MolecularDataUnstructured
(Migratory) > Causes | Patterns | Routes | Season | MigratoryUnstructured
(EcologicalSignificance) > MeasurementOrFact | EcologicalSignificanceUnstr
(MiscDetails)
(EnvironmentalEnvelope) > MeasurementOrFact |
EnvironmentalEnvelopeUnstructured
NaturalHistoryUnstructured

### Invasiveness

origin | presence | persistence | distribution | harmful | modified | startValidDat
endValidDate | countryCode | stateProvince | county | localityName | language
citation | abundance | trend | rateOfSpread | regulatoryListing | memo |
publicationDate | localityType | WhatImpact | vector | Route | impactTarget |
impactMechanism
InvasivenessUnstructured

### Habitat and Distribution

(Habitats) > MeasurementOrFact | HabitatUnstructured
(Distribution) > DistributionScope | temporalCoverage | GeographicEntity | country |
stateProvince | county | municipality | locality | DistributionUnstructured
(Endemic) > EndemicTo | EndemicIn | EndemicUnstructured
HabitatAndDistributionUnstructured

### Demography and Threat

(Territory) > ExtentOfOccurrence | AreaOfOccupancy | TerritoryUnstructured
(PopulationBiology) > AbundanceData | DensityData | PatternDistribution | Size |
SexRatio | Fecundity | MortalityRate | BirthRate | NumberIndividualsPerObservation |
AverageDensity | PopulationTrend | Recruitment | PopulationGrowthRate | Emigration
| Immigration | DescriptionLifeStages | ProportionIndividualsPerStageLife |
CarryingCapacity | PopulationBiologyUnstructured
(ThreatStatus) > ThreatCategory | Authority | AppliesTo | ThreatStatusUnstructured
(DirectThreats) > MeasurementOrFact | DirectThreatUnstructured
(Legislation) > LegislationName | ProtectionLegalStatus | LegislationRead | Status |
Type | Norm | AppliesTo | LegislationUnstructured
DemographyAndThreatUnstructured

### Uses, Management and Conservation

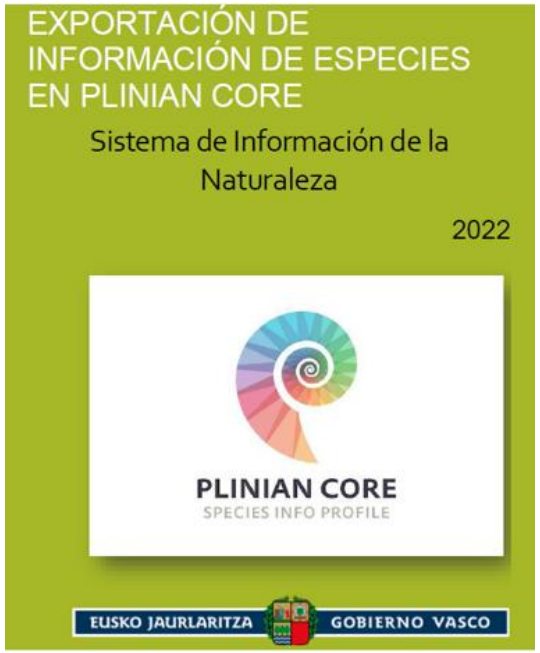
(Uses) > SourceOfInformation | Use-Value | PartUsed | Users | Organisms |
VernacularNameUseAnnotations | ProductionDetails | MeansOfApplication-
Administration | SeasonOfAvailability-Use | Conservation-ExplotationData |
UseTypeAtomized | Economics | RatingPopularity | Properties | Potential | UseNotes |
UsesUnstructured
(ManagementAndConservation) > Objetives | ManagementPlan | Actions |
HumanAndEnvironmentalrelevance | ManagementAndConservationUnstructured
UsesManagementAndConservationUnstructured

### associatedParty, MeasurementOrFact, References, AncillaryData

https://github.com/PlinianCore/Documentation/wiki/PlinianCore_Terms

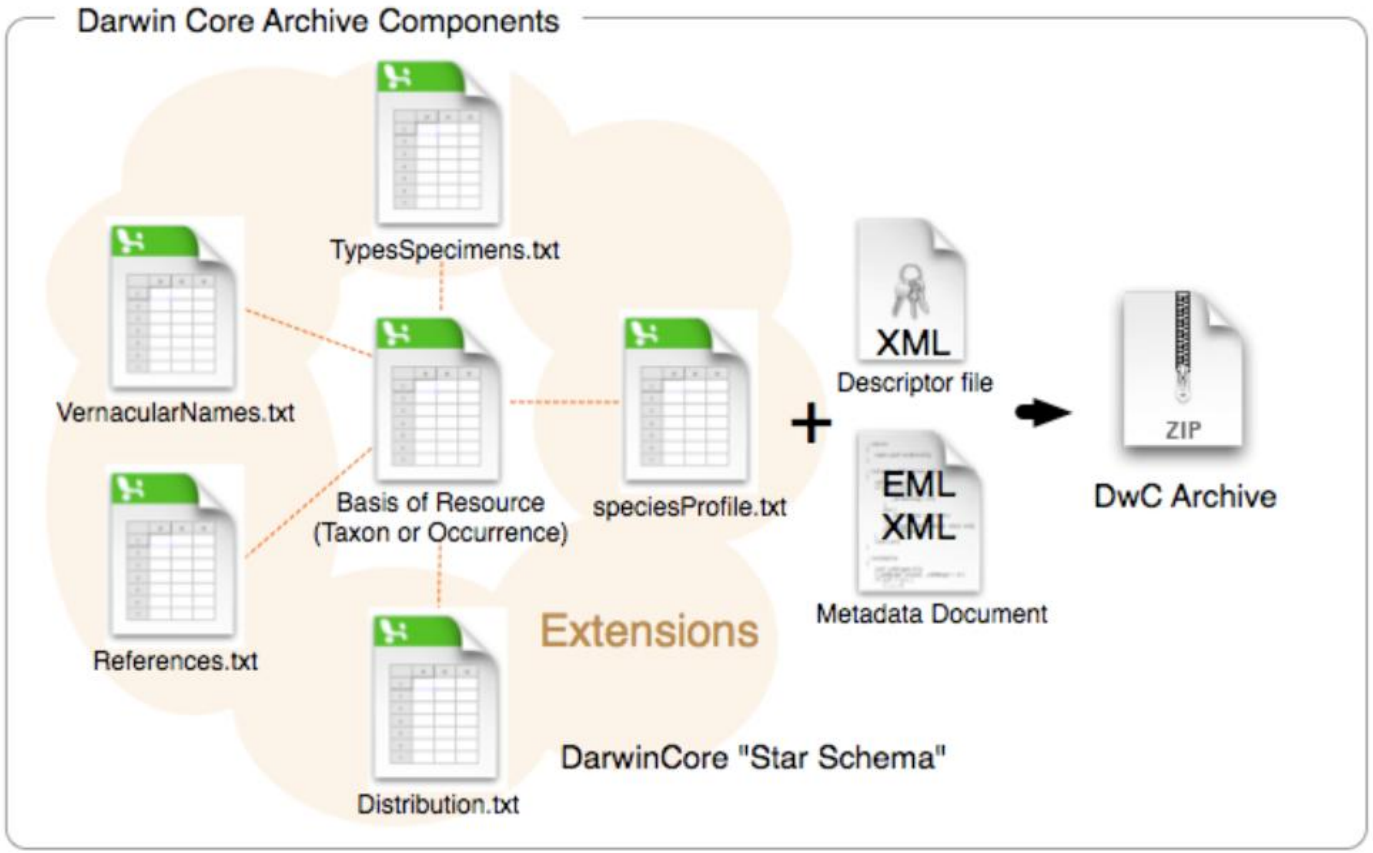
# Plinian Core is used

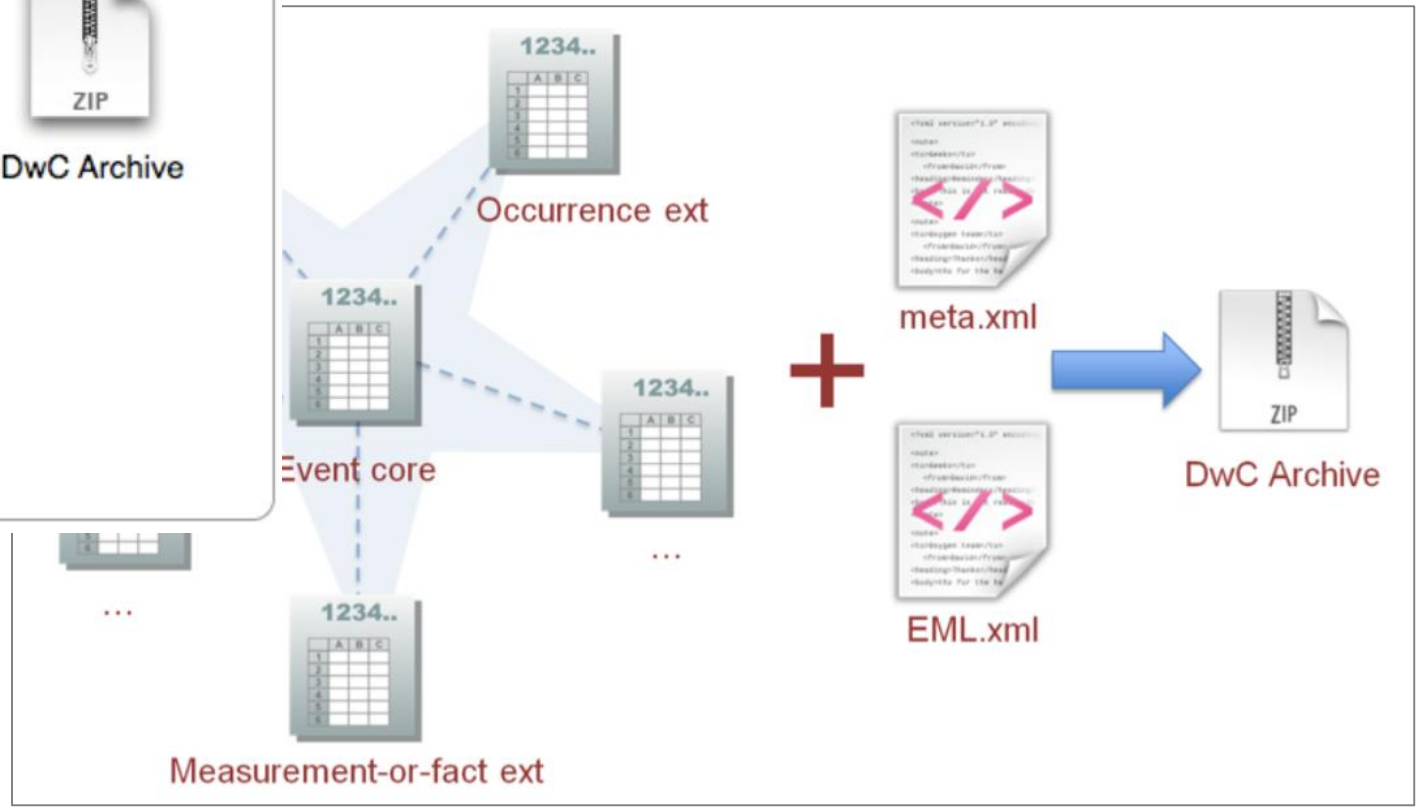


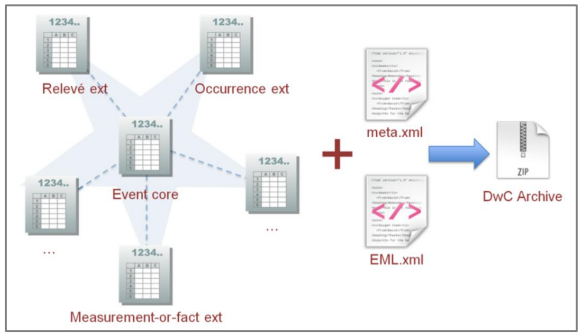


# Darwin Core. Implementación: Darwin Core Archive

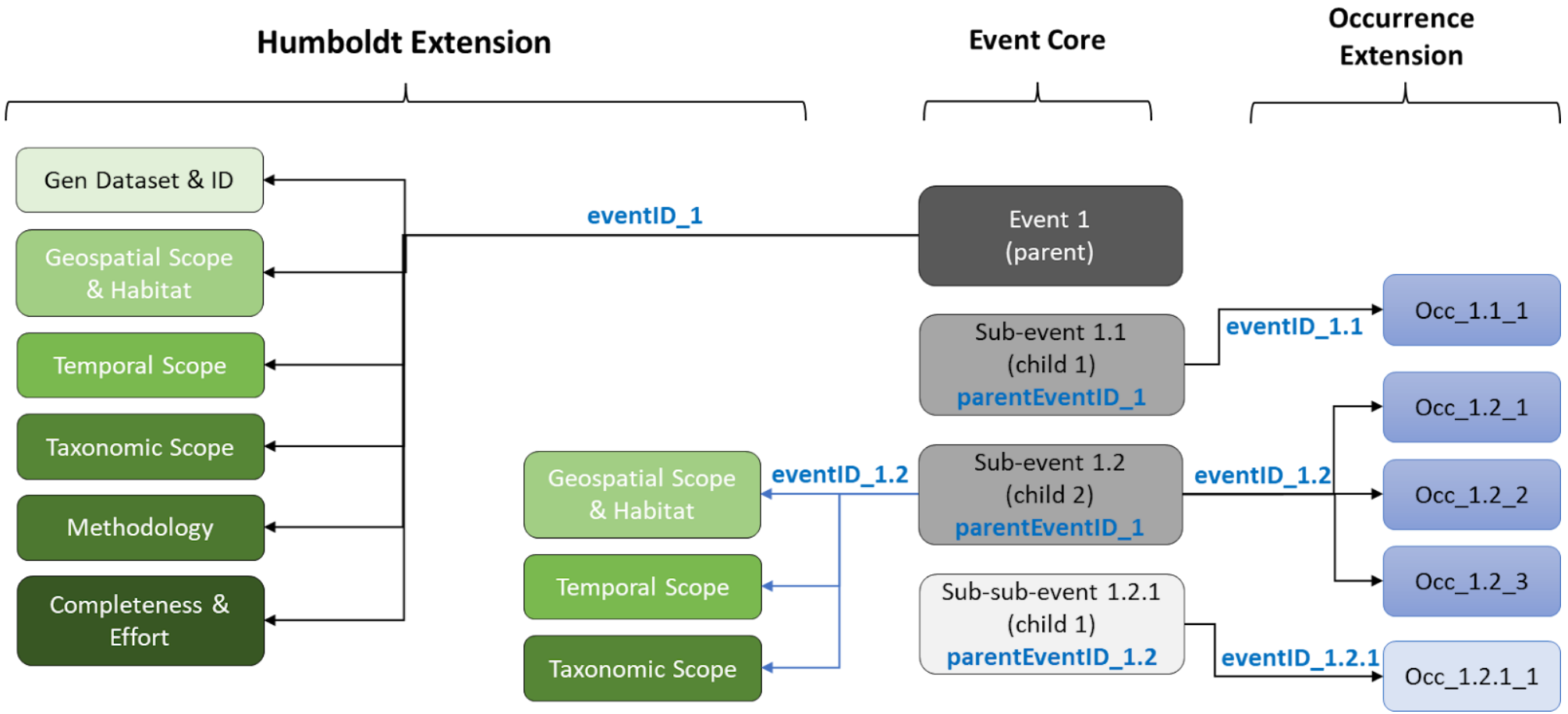



https://ipt.gbif.org/manual/en/ipt/latest/dwca-guide




# Darwin Core Archive



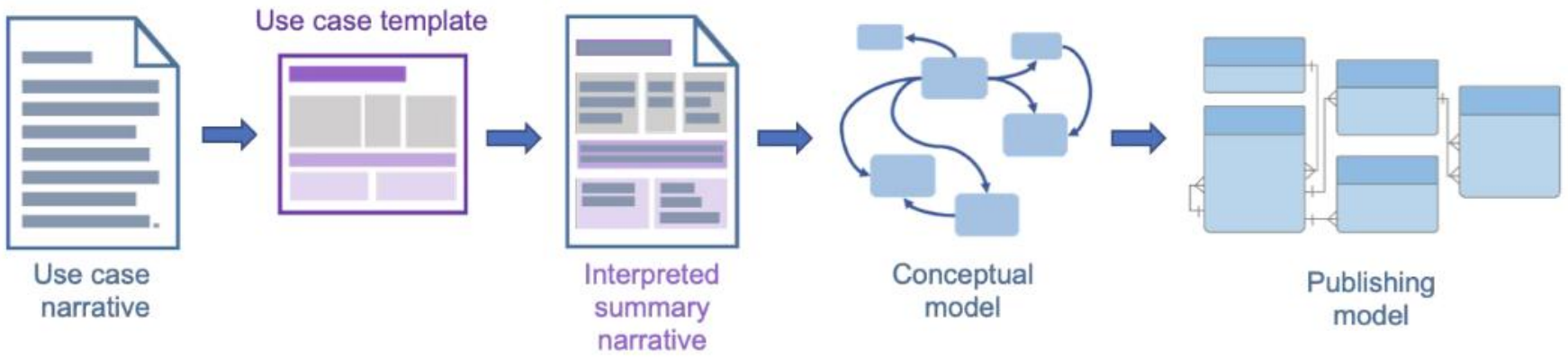

Visual representation of a structured biodiversity inventory using the Darwin Core Event class and the Occurrence and Humboldt Extensions. Information on different hierarchies is linked via each unique dwc:eventID

# Nuevo modelo conceptual para Darwin Core

>> define la estructura lógica y las relaciones entre los distintos elementos del estándar Darwin Core (DwC).Sobre el conjunto de términos o vocabulario controlado que es Darwin Core, el modelo conceptual describe cómo se conectan entre sí esos términos —es decir, la arquitectura conceptual que sustenta el intercambio de datos de biodiversidad.

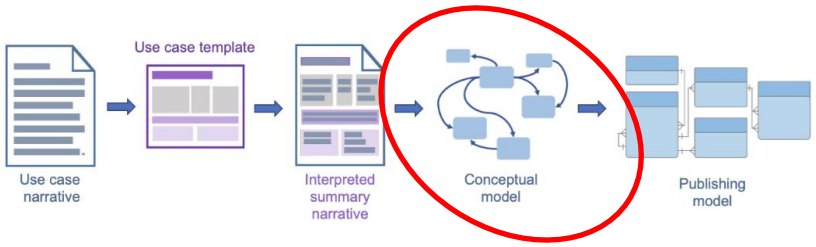
"...biodiversity data is more complicated than 'just' the occurrence of species in time and space; organisms interact, co-occur, move and evolve.
This implies a need for richer and more varied types of data than GBIF has thus far supported."

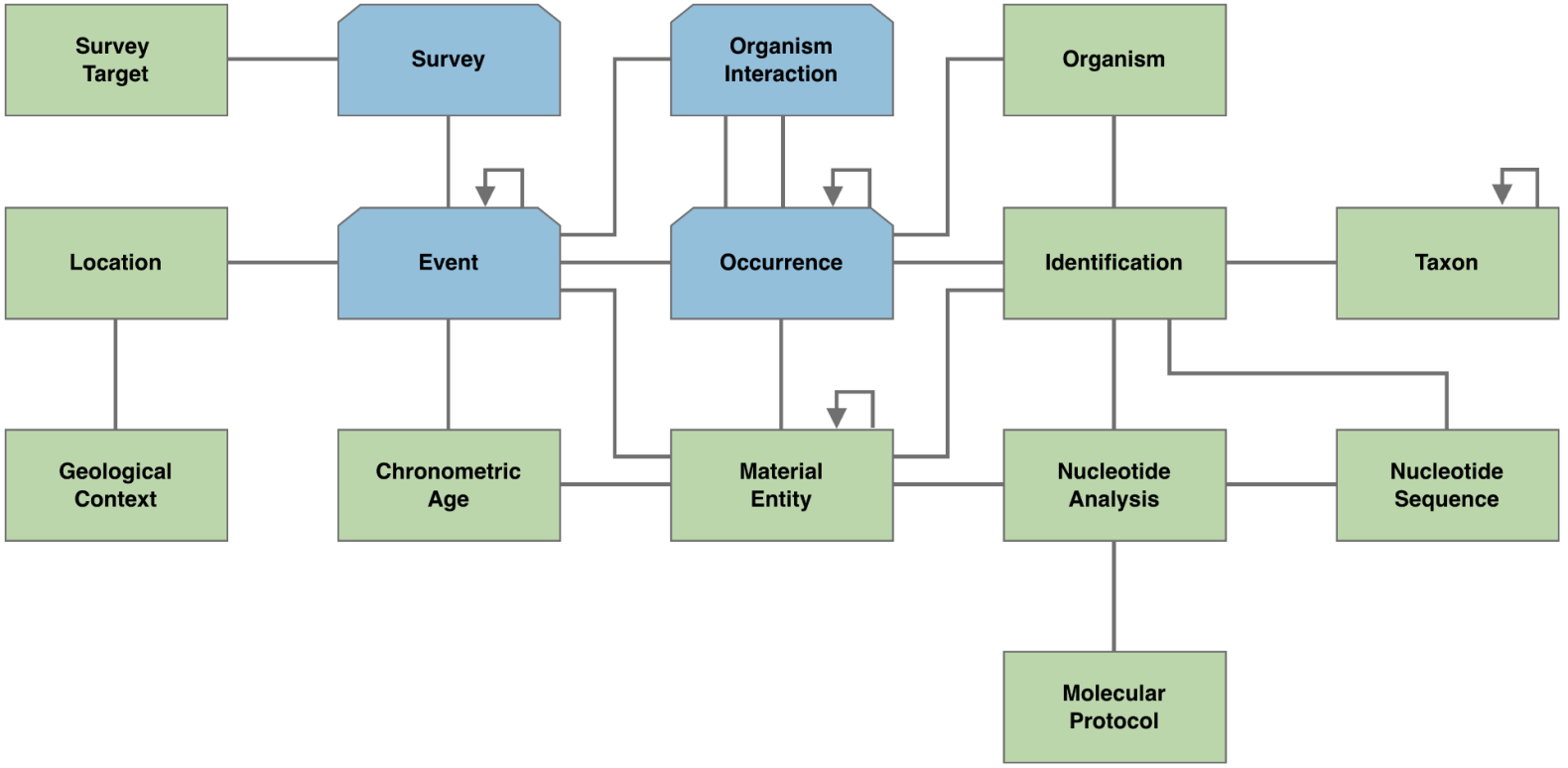
Heberling et al. (2021) Data integration enables global biodiversity synthesis https://doi.org/10.1073/pnas.2018093118

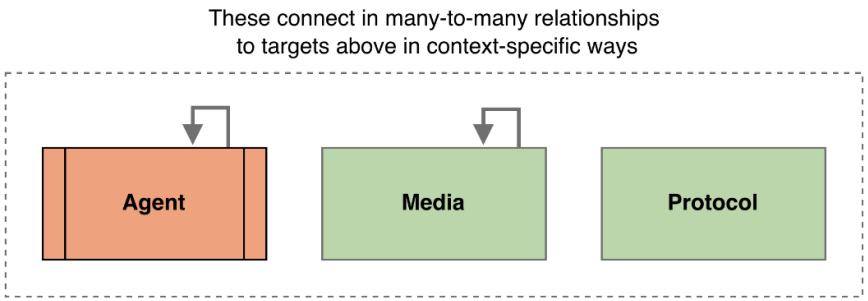


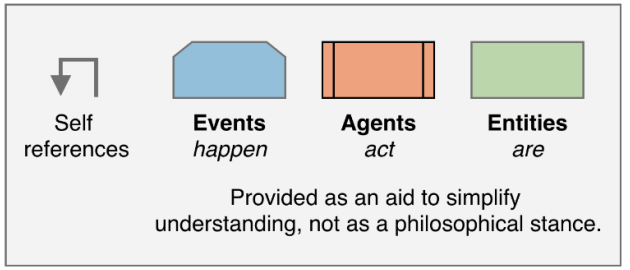
# Nuevo modelo conceptual para Darwin Core

https://gbif.github.io/dwc-dp/cm/

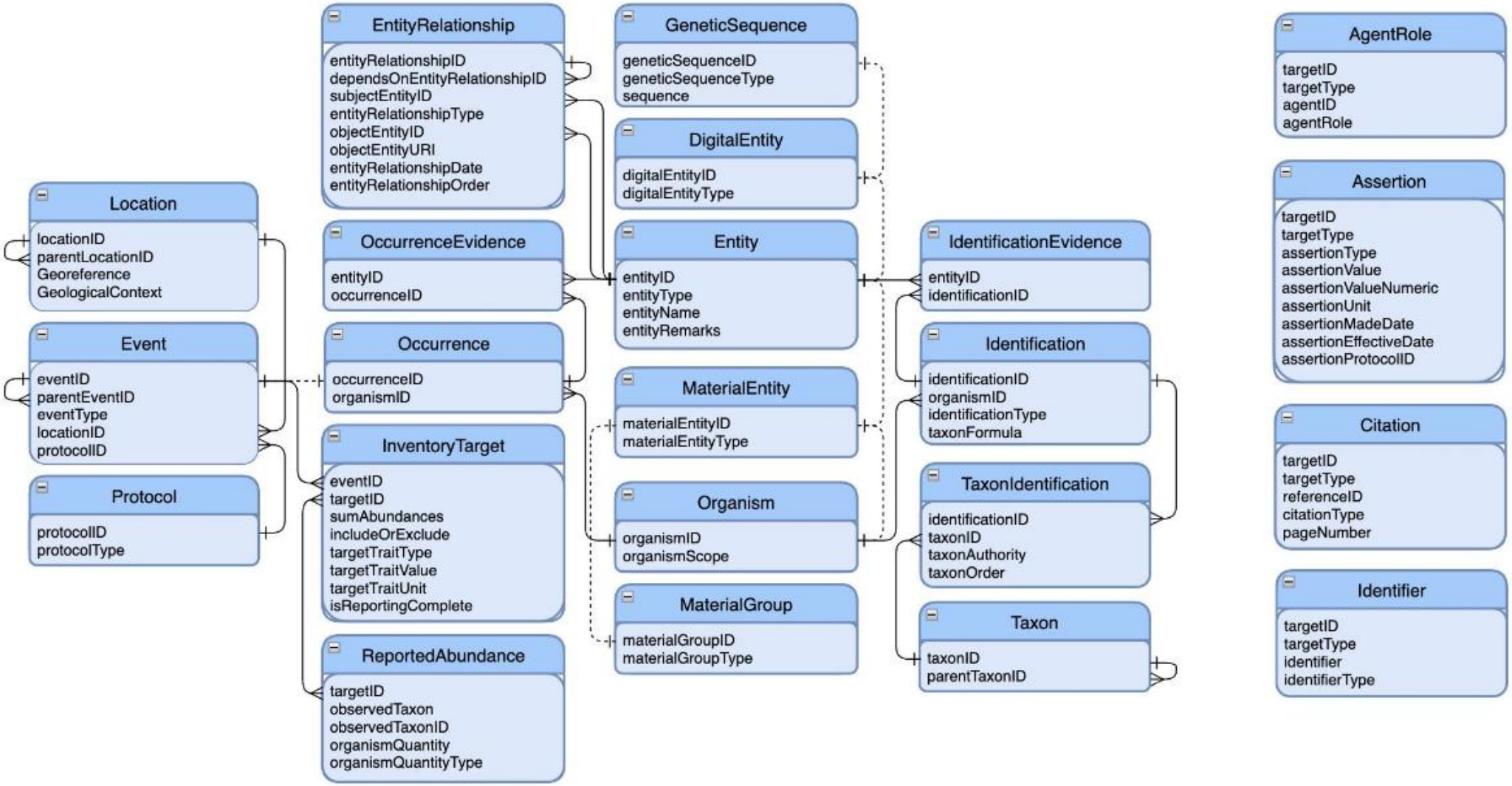


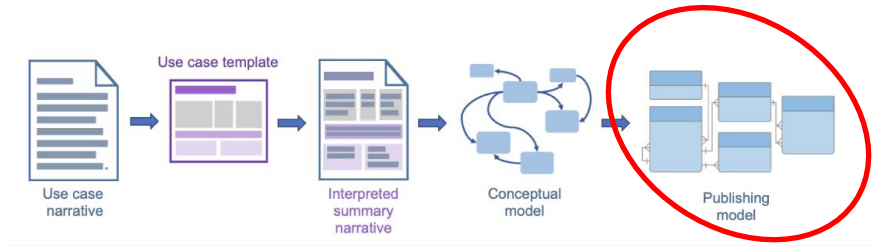



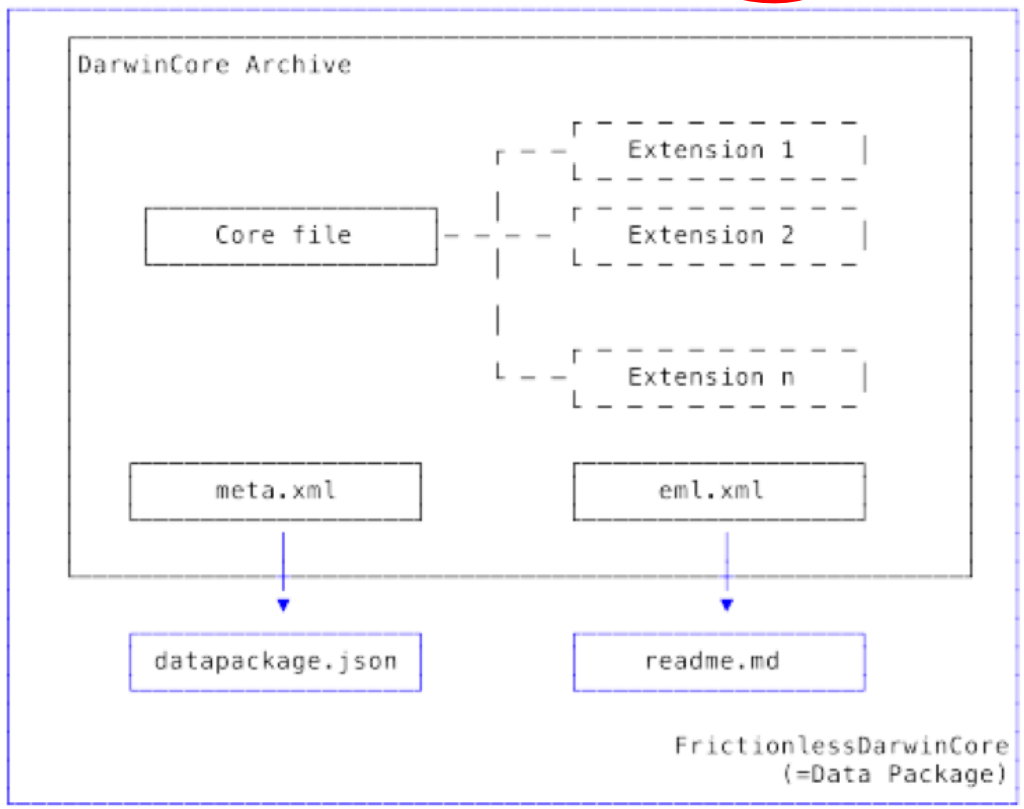



# Nuevo modelo conceptual para Darwin Core

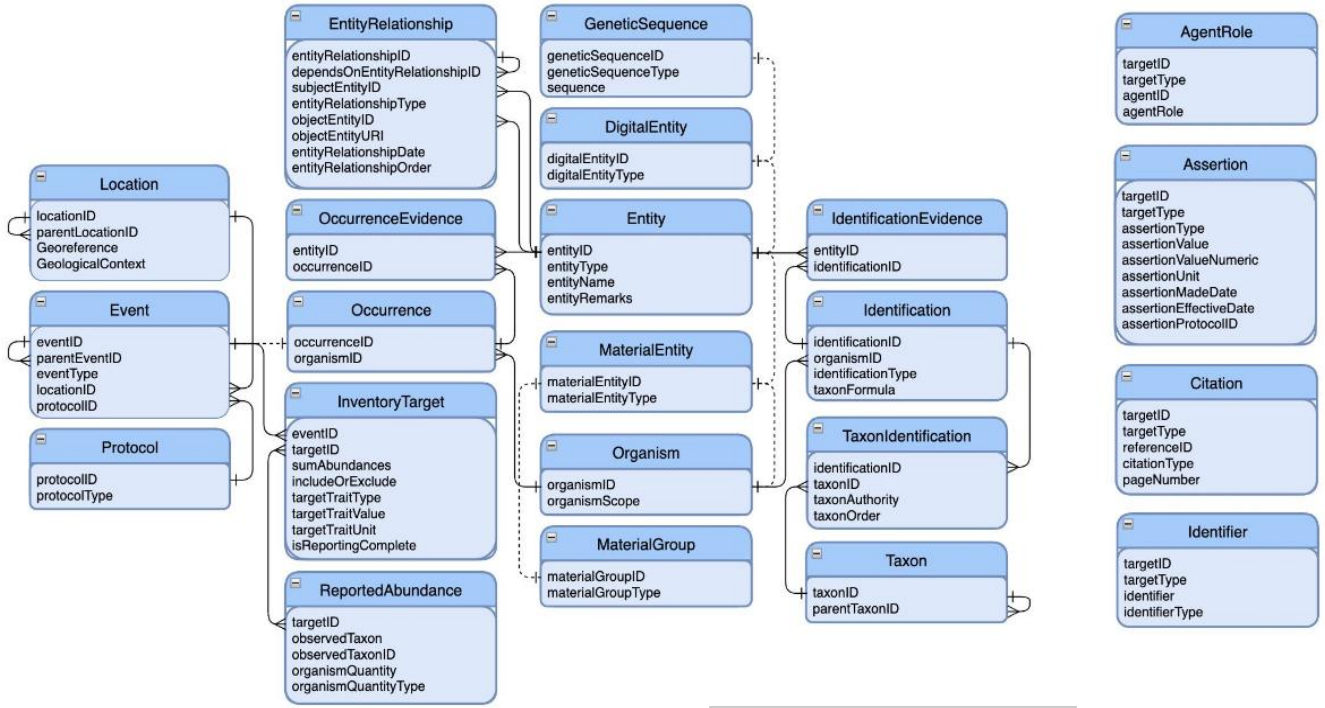


# Darwin Core – DP

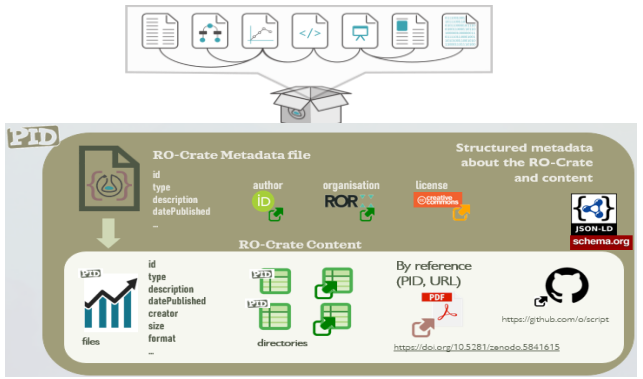



https://gbif.github.io/dwc-dp/dp/




Darwin Core – DP es una implementación de “Frictionless data Packages”



# Hemos visto…

- Datos y ciencia
- Ciencia abierta y estándares
- Tipos de estándares
- TDWG
- Estándares en GBIF
- Darwin Core
- Humboldt Core
- Plinian Core
- Darwin Core Archive
- GBIF new data model
- Darwin Core – DP

# Camtrap DP es una implementación de Darwin Core DP

https://camtrap-dp.tdwg.org/

Project / Study

Deployment
29b7d356
from T1 to T1000

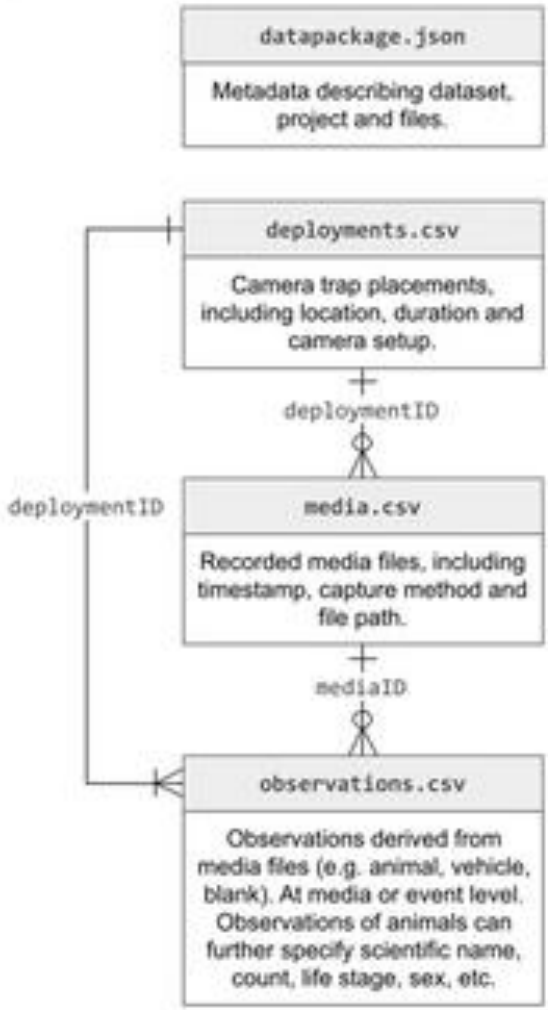
Media file
59b38bc6 at T200

Media file
cc50edaa at T201

Media file
0966e552 at T202

| observation ID | observation Level | media ID | event Start | event End | scientificName | count |
|---|---|---|---|---|---|---|
| obs1 | media | 59b38bc6 | T200 | T200 | Anas platyrhynchos | 3 |
| obs2 | media | cc50edaa | T201 | T201 | Anas platyrhynchos | 4 |
| obs3 | media | 0966e552 | T202 | T202 | Anas platyrhynchos | 4 |
| obs4 | event | | T200 | T202 | Anas platyrhynchos | 5 |

(B)



Francisco Pando
Unidad de Coordinación de GBIF.ES
CSIC
Joaquín Costa 22, 28002 Madrid

pando@gbif.es



GBIF-ES es el Nodo Nacional de Información en Biodiversidad patrocinado por el Ministerio de Ciencia, Innovación y Universidades, gestionado por el Consejo Superior de Investigaciones Científicas (CSIC). https://www.gbif.es/ https://www.gbif.es/ , https://datos.gbif.es, https://elearning.gbif.es

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

CSIC Consejo Superior de Investigaciones Científicas

GBIF.ES

https://mutis.rjb.csic.es