Estándares y vocabularios controlados para información sobre biodiversidad
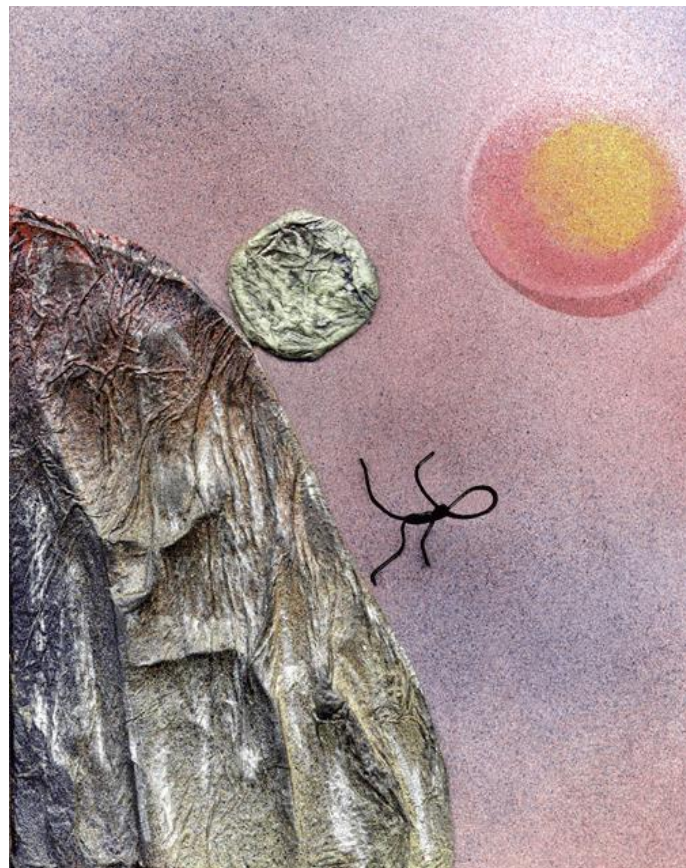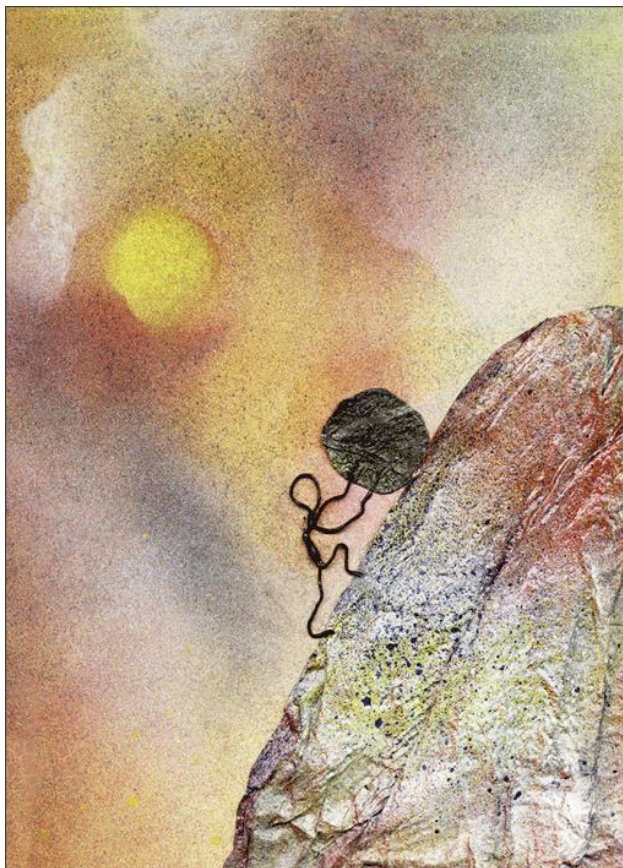
gbif.es

Francisco Pando
Unidad de Coordinación – GBIF.ES
Nodo Nacional de Información en
Biodiversidad en España

Sumario

- Compartir, para qué
- Estandarizar, qué
- Estandarización, para que
- Tipos de estándares
- Que hay
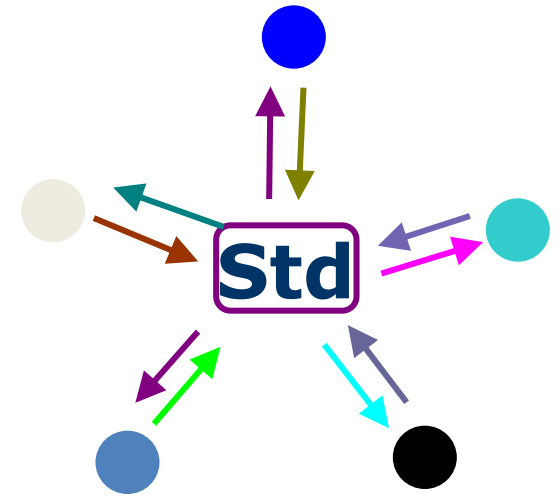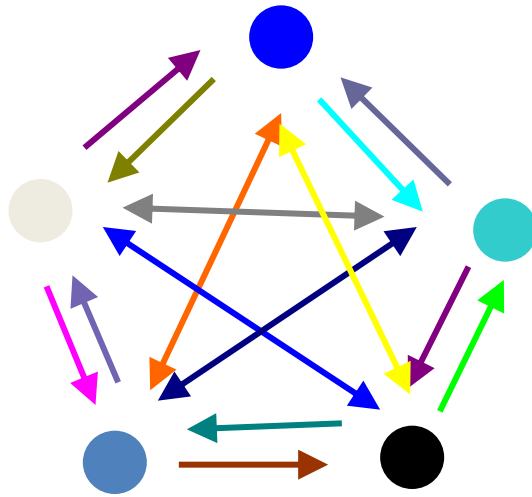- Que va a haber
- Consideraciones finales

Compartir para que

© The Art of
Steven Nelson

HEYWOOD, V.H. (1974). Systematics-the stone of Sysyphus. Biol. J. Linn. Soc. 6(2): 169-178.

Estandarizar para qué

- Estandarizar para intercambiar

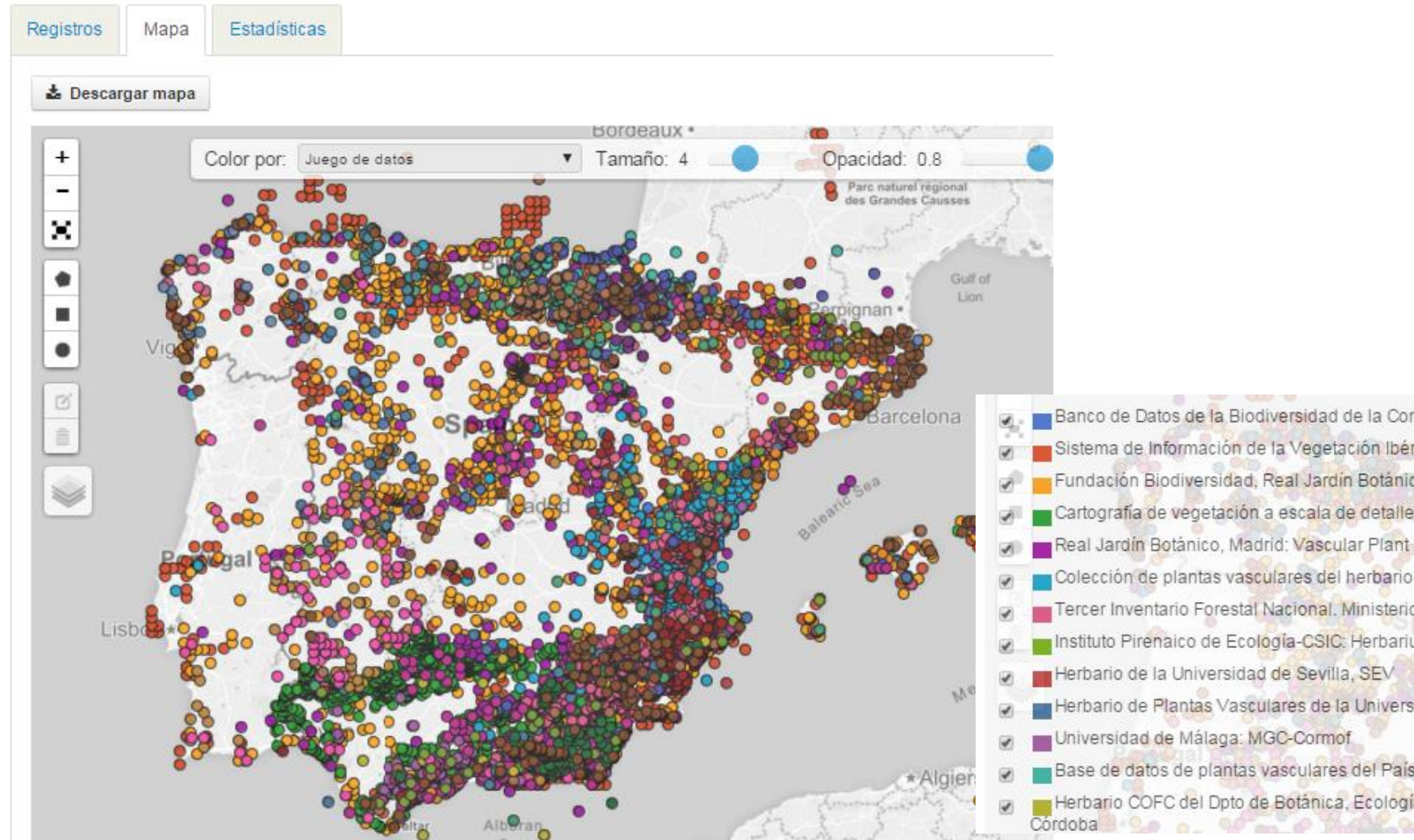- Estandarizar para utilizar

- Estandarizar no para producir

Estandarizar para intercambiar

Estandarizar para utilizar

http://datos.gbif.es/generic-hub/occurrences/search?taxa=anthyllis#tab_mapView
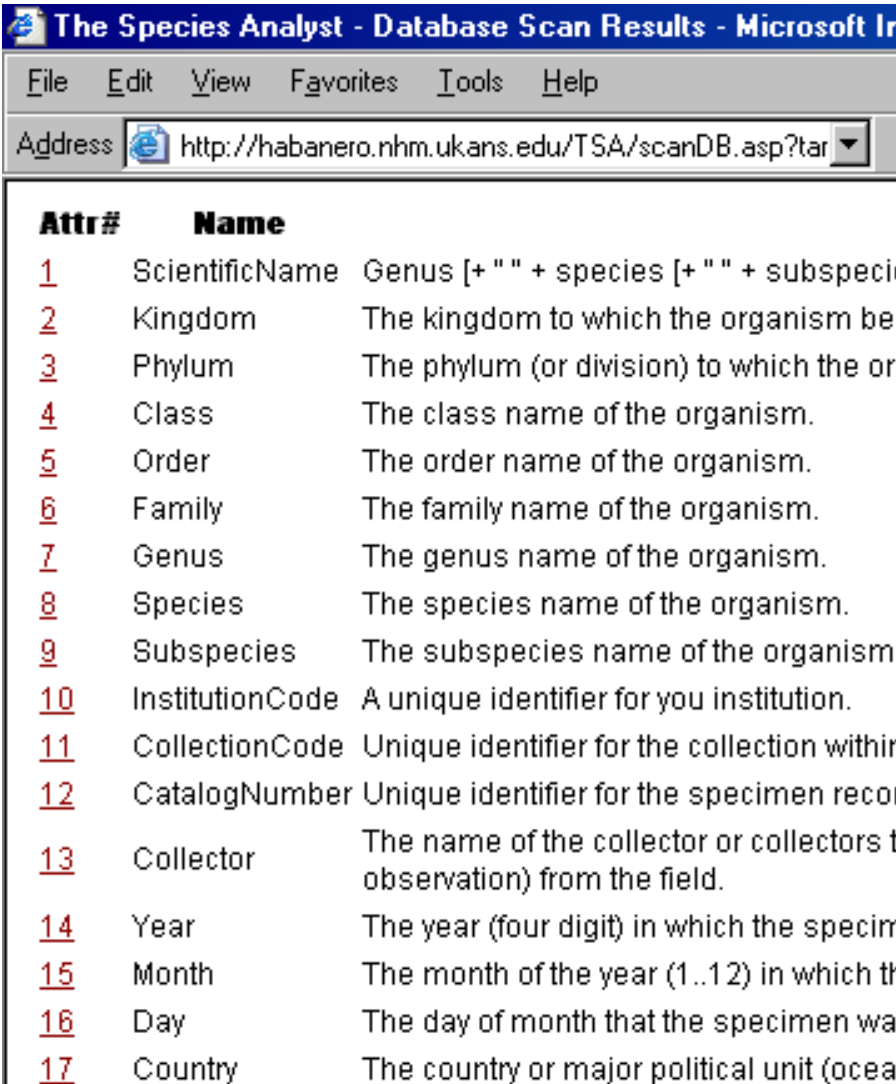
Estandarizar no para producir

Un estándar de intercambio no es una buen guía para diseñar una base de datos

- ¿Dos sitios para el nombre científico?
- ¿Quien ha hecho la identificación?
- ¿Imágenes?
- ¿Enlaces a datos moleculares?
- ¿Datos de hábitat?

Una base de datos tiene que servir a nuestros fines primero, Después viene compartir y diseminar

The Species Analyst - Database Scan Results - Microsoft I

File   Edit   View   Favorites   Tools   Help

Address   http://habanero.nhm.ukans.edu/TSA/scanDB.asp?tar

Attr#	Name	
1	ScientificName	Genus [+ " " + species [+ " " + subspeci
2	Kingdom	The kingdom to which the organism be
3	Phylum	The phylum (or division) to which the or
4	Class	The class name of the organism.
5	Order	The order name of the organism.
6	Family	The family name of the organism.
7	Genus	The genus name of the organism.
8	Species	The species name of the organism.
9	Subspecies	The subspecies name of the organism
10	InstitutionCode	A unique identifier for you institution.
11	CollectionCode	Unique identifier for the collection within
12	CatalogNumber	Unique identifier for the specimen reco
13	Collector	The name of the collector or collectors t observation) from the field.
14	Year	The year (four digit) in which the specim
15	Month	The month of the year (1..12) in which th
16	Day	The day of month that the specimen wa
17	Country	The country or major political unit (ocea

Tipos de estándares

- ## Vocabularios controlados (léxicos)
  - Que valores son válidos
  - (listas de géneros,...)

- ## Definiciones (semánticos)
  - para humanos, para máquinas
  - (v.gr.: HISPID, Darwin Core [terms])

- ## Estructurales (sintácticos)
  - Como se codifica la información
  - (v.gr.: ABCD, Darwin Core Archive)

Vocabularios controlados

SiB — Sistema de Información sobre Biodiversidad de Colombia

Herramientas de administración de datos

Enlaces

Descargas

\ Metadatos \ Archivos de au

Tesauros

- ENTRAR AL TESAURO
- ---> Áreas del conocimien
- ---> Biogeografía
- ---> Conservación
- ---> Gacetero
- ---> Métodos y atributos
- ---> Unidades biológicas
- ---> Uso y saberes locales
- -----> Grupo humano
- -------> Barí

Ficha nomenclatural

Realizar nueva búsqueda

Eremophilus mutisii

Autor:	Humboldt, 1805
Reino:	Animalia
Nivel taxonómico:	Especie
Estado taxonómico:	Válido
Taxones relacionados:	
Estado de revisión/verificación:	Revisado
Fecha de revisión/verificación:	

Jerarquía taxonómica: En rojo aparecen los nombres no válidos / no aceptad

Reino: Animalia
 Phylum: Chordata
  Subphylum: Vertebrata
   Superclase: Osteichthyes
    Clase: Actinopterygii
     Subclase: Neopterygii
      Infraclase: Teleostei
       Superorden: Ostariophysi
        Orden: Siluriformes
         Familia: Trichomycteridae
          Género: Eremophilus
           Especie: Eremophilus mutisii

Referencias

Fuente de la descripción :

Original: Humboldt, F.H.A. von (1805) Mémoire sur l´ Eremophilus et Astroblepus,

Vocabularios controlados

Name	Last modified	Size	Description
Parent Directory		-	
ac/	2015-08-27 17:01	-	
basic/	2015-08-27 17:01	-	
dcterms/	2015-08-27 17:01	-	
dwc/	2015-08-28 16:01	-	
eml/	2015-08-27 17:01	-	
eol/	2015-08-27 17:01	-	
gbif/	2015-08-28 18:01	-	
germplasm/	2015-08-27 17:01	-	
ggbn/	2015-08-27 17:01	-	
gisin/	2015-02-13 22:36	-	
iso/	2015-08-27 17:01	-	
iucn/	2015-08-27 17:01	-	
mixs/	2015-08-27 17:01	-	
sn2000/	2015-08-27 17:01	-	
un/	2015-08-27 17:01	-	

http://rs.gbif.org/vocabulary/

Estructurales (sintácticos)

```
<record>

    <dwc:DateLastModified>Mar 7, 2004</dwc:DateLastModifie

    <dwc:InstitutionCode>Univeristy of Alaska Museum of th
North</dwc:InstitutionCode>

    <dwc:CollectionCode>Mammal specimens</dwc:CollectionCo

    <dwc:CatalogNumber>15985</dwc:CatalogNumber>

    <dwc:S
    <dwc:K
    <dwc:Pl
    <dwc:C
    <dwc:O
    <dwc:F
    <dwc:G
    <dwc:Sp
    <dwc:Su
    <dwc:S
    <dwc:Y
    <dwc:M
    <dwc:D
    <dwc:C
    <dwc:C
```


Darwin Core Archive Components

TypesSpecimens.txt
VernacularNames.txt
Basis of Resource (Taxon or Occurrence)
speciesProfile.txt
References.txt
Distribution.txt
Extensions
DarwinCore "Star Schema"
XML Descriptor file
EML XML Metadata Document
+ → ZIP DwC Archive
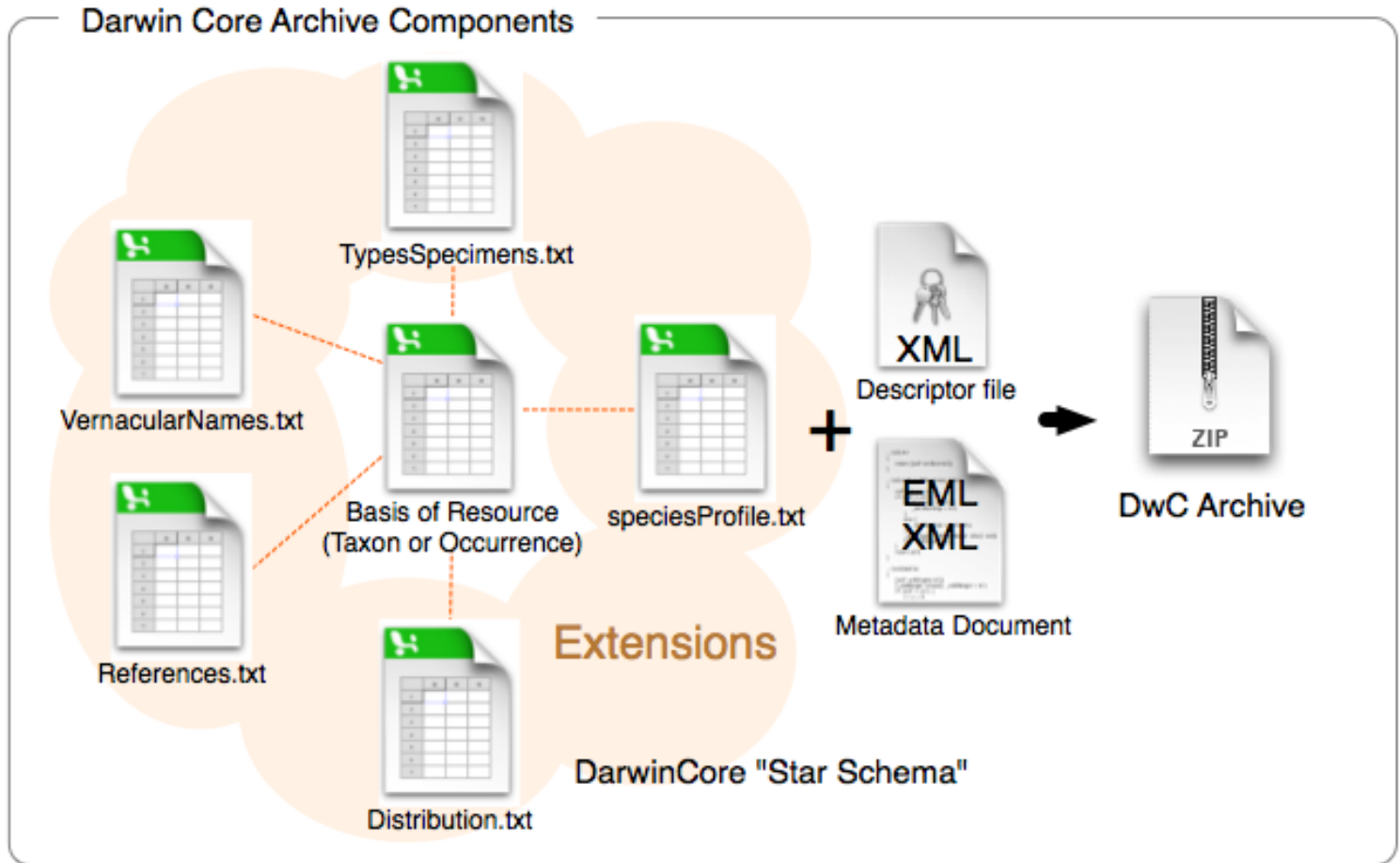
Definiciones (semánticos)

Biodiversity Information Standards TDWG

Introduction
References

Darwin Core Terms: A quick reference guide

Record-level Terms

dcterms:type | dcterms:modified | dcterms:language | dcterms:license | dcterms:rightsHolder | dcterms:accessRights | dcterms:bibliographicCitation | dcterms:references

institutionID | collectionID | datasetID | institutionCode | collectionCode | datasetName | ownerInstitutionCode | basisOfRecord | informationWithheld | dataGeneralizations | dynamicProperties

Occurrence

occurrenceID | catalogNumber | recordNumber | recordedBy | individualCount | organismQuantity | organismQuantityType | sex | lifeStage | reproductiveCondition | behavior | establishmentMeans | occurrenceStatus | preparations | disposition | associatedMedia | associatedReferences | associatedSequences | associatedTaxa | otherCatalogNumbers | occurrenceRemarks

Organism

organismID | organismName | organismScope | associatedOccurrences | associatedOrganisms | previousIdentifications | organismRemarks

MaterialSample | LivingSpecimen | PreservedSpecimen | FossilSpecimen

materialSampleID

Event | HumanObservation | MachineObservation

eventID | parentEventID | fieldNumber | eventDate | eventTime | startDayOfYear | endDayOfYear | year | month | day | verbatimEventDate | habitat | samplingProtocol | sampleSizeValue | sampleSizeUnit | samplingEffort | fieldNotes | eventRemarks

Location

locationID | higherGeographyID | higherGeography | continent | waterBody | islandGroup | island | country | countryCode | stateProvince | county | municipality | locality | verbatimLocality | minimumElevationInMeters | maximumElevationInMeters | verbatimElevation | minimumDepthInMeters | maximumDepthInMeters | verbatimDepth | minimumDistanceAboveSurfaceInMeters | maximumDistanceAboveSurfaceInMeters | locationAccordingTo | locationRemarks | decimalLatitude | decimalLongitude | geodeticDatum | coordinateUncertaintyInMeters | coordinatePrecision | pointRadiusSpatialFit | verbatimCoordinates | verbatimLatitude | verbatimLongitude | verbatimCoordinateSystem | verbatimSRS | footprintWKT | footprintSRS | footprintSpatialFit | georeferencedBy | georeferencedDate | georeferenceProtocol | georeferenceSources | georeferenceVerificationStatus | georeferenceRemarks

http://rs.tdwg.org/dwc/terms/

Que hay (la situación mundial)

- TDWG

  "Taxonomic Databases Working Group"

  --Biodiversity Information Standards—

- Ámbitos más reducidos, estándares más ricos (estándares nacionales, temáticos,..)

TDWG: Biodiversity Information Standards

Biodiversity Information Standards

T D W G

© www.lophoto.com

username ········ LOGIN | SEARCH

- Home
- About TDWG
- Getting Started
- Standards
- Activities
- Membership

- Conference 2012

- Biodiversity Information Projects of the World
- Biodiversity Information Networks Database
- Biodiversity Informatics Events Database

- Wiki
- OJS
- Mailing Lists

Basic Standards Recommendations

The most widely deployed formats for biodiversity occurrence data are Darwin Core (wiki) and ABCD (wiki). New deployments of these and other XML based formats should use the TAPIR exchange protocol.

The TDWG community's priority is the deployment of Life Science Identifiers (LSID), the preferred Globally Unique Identifier technology and transitioning to RDF encoded metadata as defined by a set of simple vocabularies. All new projects should address the need for tagging their data with LSIDs and consider the use or development of appropriate vocabularies.

TDWG's activities within the biodiversity informatics domain can be found in the Activities section of this website.

2012 TDWG annual conference Announcement

The Biodiversity Information Standards (TDWG) Executive Committee is pleased to announce that the 2012 TDWG annual conference will be held on October 22-26, 2012 at the China People's Palace in Beijing, China.

Latest News

30-May-2012 Deadline to apply to host TDWG meeting in 2013 expression of interest extended to the 15th of June 2012

[more]

29-May-2012 Travel Awards for African Participation in 2012 Biodiversity Information Standards (TDWG) Conference

TDWG Standards

TDWG's standards are listed below. For information about current work within TDWG, we would suggest that you look at the TDWG Wiki. For further information, please contact us.

TDWG Current Standards (*)

Title	Activity (Task or Interest Group)	Category	Status	Date Published	
GUID and Life Sciences Identifiers Applicability Statements		Applicability Statements	Current	30-Jan-2011	Download
Darwin Core	DarwinCore Task Group (DwC)	Technical Specification	Current	09-Oct-2009	Download
TAPIR - TDWG Access Protocol for Information Retrieval	TAPIR Task Group	Technical Specification	Current	09-Sep-2009	Download

TDWG Current (2005) Standards (*)

Title	Activity (Task or Interest Group)	Category	Status	Date Published	
Access to Biological Collection Data - version 2.06	Access to Biological Collections Data	Technical Specification	Current (2005)	16-Sep-2005	Download
Structured Descriptive Data	Biological	Technical	Current	16-Sep-2005	Download

TDWG Draft Standards [*]

Title	Activity (Task or Interest Group)	Category	Status	Date Published	
TDWG Standards Documentation Specification	TDWG Infrastructure Project	Technical Specification	Draft		Download (**)
Natural Collections Descriptions (NCD): A data standard for exchanging data describing natural history collections	Natural Collections Descriptions Interest Group	Technical Specification	Draft		Download (**)
MRTG Submission		Technical Specification	Draft		Download (**)
ABCDDNA ◆ DNA extension for Access to Biological Collection Data	Access to Biological Collections Data	Technical Specification	Draft		Download (**)

TDWG Prior Standards [(*)]

Title	Activity (Task or Interest Group)	Category	Status	Date Published	
HISPID3 - Herbarium Information Standards and Protocols for Interchange of Data	Observation and Specimen Records	Technical Specification	Prior	01-Oct-1996	Download
Economic Botany Data Collection Standard	Economic Botany Interest Group	Best Current Practice	Prior	01-Oct-1995	Not available for download
Plant Occurrence and Status Scheme		Status and Categories	Prior	01-Oct-1995	Download
Plant Names in Botanical Databases		Best Current Practice	Prior	01-Oct-1995	Download
Authors of Plant Names		Status and Categories	Prior	01-Oct-1992	Not available for download
World Geographical Scheme for Recording Plant Distributions		Status and Categories	Prior	01-Oct-1992	Download
XDF - A Language for the Definition and Exchange of Biological Data Sets		Technical Specification	Prior	01-Oct-1991	Not available for download
Botanico-periodicum-huntianum/supplementum		Status and Categories	Prior	01-Oct-1991	Not available for download
Index Herbariorum. Part I: The Herbaria of the World		Status and Categories	Prior	01-Oct-1990	Not available for download
International Transfer Format for Botanic Garden Plant Records		Technical Specification	Prior	01-Oct-1987	Download
Floristic Regions of the World		Status and Categories	Prior	01-Oct-1986	Not available for download
Taxonomic Literature, ed. 2 and its Supplements		Status and Categories	Prior	01-Oct-1976	Not available for download

Darwin Core Terms: A quick reference guide

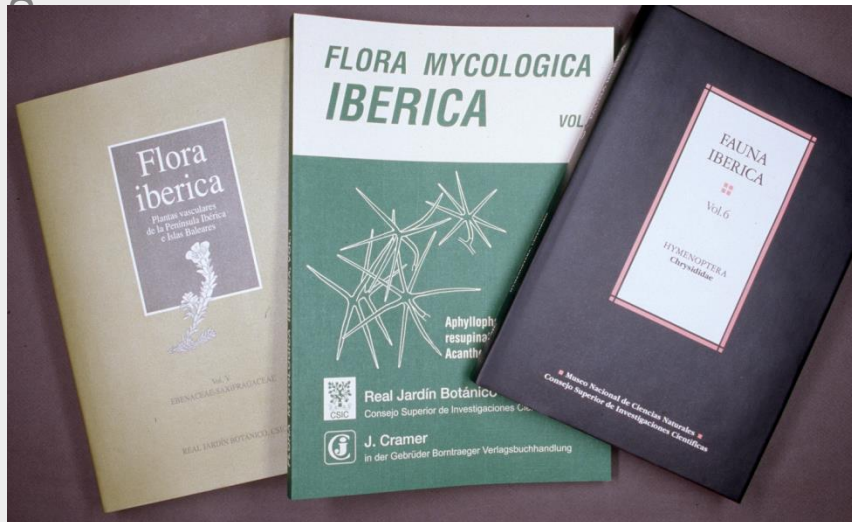|	
Title:	Darwin Core Terms: A quick reference guide
Date Issued:	2009-02-12
Date Modified:	2011-10-26
Abstract:	This document is a quick reference for all recommended Darwin Core terms. For complete historical term information, including version changes and pre-standard terms, see [HISTORY]. For a comparative table of elements from pre-standard versions of Darwin Core to the current terms in the standard, see [VERSIONS].
Contributors:	John Wieczorek (MVZ), Markus Döring (GBIF), Renato De Giovanni (CRIA), Tim Robertson (GBIF), Dave Vieglais (KUNHM)
Legal:	This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.
Part of TDWG Standard:	http://www.tdwg.org/standards/450/
Creator:	Darwin Core Task Group
Identifier:	http://rs.tdwg.org/dwc/2011-10-26/terms/
Latest Version:	http://rs.tdwg.org/dwc/terms/
Replaces:	http://rs.tdwg.org/dwc/2009-12-07/terms/
Document Status:	Current Standard

DwC: múltiples versiones

Darwin Core Versions

The following versions of the Darwin Core are of historical significance:

- DwC 1.2 schema -- first deployed version (used in GBIF, also known as DarwinCoreV2?)
- DwC 1.21 schema -- revised version (used in GBIF, MaNIS, HerpNet, OrNIS, and FishNet2)
- DwC 1.3 schema -- draft standard of the Darwin Core accepted at TDWG Meeting 2004, Christchurch, New Zealand. (unused)
- DwC 1.4 schema -- draft standard under discussion (not for use)(but used in GBIF, and recommended to use in GBIF see: http://www.gbif.org/DataProviders/HowTo)
- OBIS -- based on DwC 1.2 used in GBIF, Ocean Biogeographic Information System)
- PaleoPortal -- based on DwC 1.2 (used in The Paleontology Portal)

Biodiversity Information Standards TDWG

Introduction
| References
Quick Reference Guide
Simple Darwin Core
Type Vocabulary

Namespace Policy
XML Guide
Text Guide

Complete History
Decision History
Mapping to ABCD
Mapping to Old Versions

Darwin Core

Title:	Darwin Core
Date Issued:	2009-02-12
Date Modified:	2009-10-08
Abstract:	This document is a cover page, an entry-level document to the Darwin Core standard. It describes the purpose of the standard and orients the reader to the documents that cover specific topics within the standard, such as the quick guide to the list of terms.
Contributors:	John Wieczorek (MVZ), Markus Döring (GBIF), Renato De Giovanni (CRIA), Tim Robertson (GBIF), Dave Vieglais (KUNHM)
Legal:	This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.
Part of TDWG	http://www.tdwg.org/standards/450/

Estándares temáticos y/o nacionales

Manual de las bases de datos nomenclaturales de *Flora Mycologica Iberica* y *Flora iberica*

Francisco Pando, Félix Muñoz Garmendia & Carlos Aedo

Real Jardín Botánico

El objetivo principal de este manual —disponible aquí
está concebido para usarse con los programas NOME
el de servir de guía para fichar, en una base de datos
organismos, e información relacionada, como p
publicación, de su tipo, de su estado nomenclatural, e

http://www.rjb.csic.es/bibmaste/manu_n.html

DE DATOS

FLORA MYCOLOGICA IBERICA VOL.

Flora iberica
Plantas vasculares
de la Península Ibérica
e Islas Baleares

Real Jardín Botánico
Consejo Superior de Investigaciones Ci

CSIC

J. Cramer
in der Gebrüder Borntraeger Verlagsbuchhandlung

FAUNA IBERICA
Vol.6
HYMENOPTERA
Chrysididae

Museo Nacional de Ciencias Naturales
Consejo Superior de Investigaciones Científicas

SiB Sistema de Información
sobre Biodiversidad de Colombia

Enlaces

• Archivos de autoridad taxonómica

Bus

Enlaces

ón

rgas

AAT disponibles

1. Vitolo Lopez A. L. (2004) Archivo de autoridad taxonómica de esc
   (Coleoptera: Cicindelidae) de Colombia
2. Díaz-Ruiz M., Diaz-Pulido G. (2004) Archivo de autoridad taxonóm
   macroalgas verdes (Chlorophyta) del Caribe colombiano
3. Quiroga-Cárdenas S. Y., Bolaños-Rodríguez D. M. (2003) Polyclad
   Tropical Occidental
4. Martínez C. & G. Ball (2003) 1Los Platynini (Coleoptera: Carabida
5. Calderón Sáenz E., Farfán Camargo J. (2004) AAT de 13 géneros
   Colombianas
6. Murillo, J. (2003) Euphorbiaceas de Colombia
7. Sendoya S., Fernández F. (2004) AAT de hormigas (Hymenopter
   neotrópico
8. Rodríguez-P M. E. & Y. Muñoz-Saba (2004) Murciélagos de la fami
   Colombia
9. Calderón-Sáenz E., Tobón I. C. (2004) Catálogo taxonómico de l
   Cycnoches, Embreea, Lycaste, Otoglossum, Phragmipedium, Psy
   Rodriguezia, Schlimia y Selenipedium (Orchidaceae) para Colomb
10. Brummitt R. K. (1992) Vascular plant families and genera: a lisit

Estándares más ricos, tambíen más restringidos

Estándar proyectos *Flora iberica* & *Flora Mycologica Iberica*

Estándar TDWG & Red de datos del GBIF
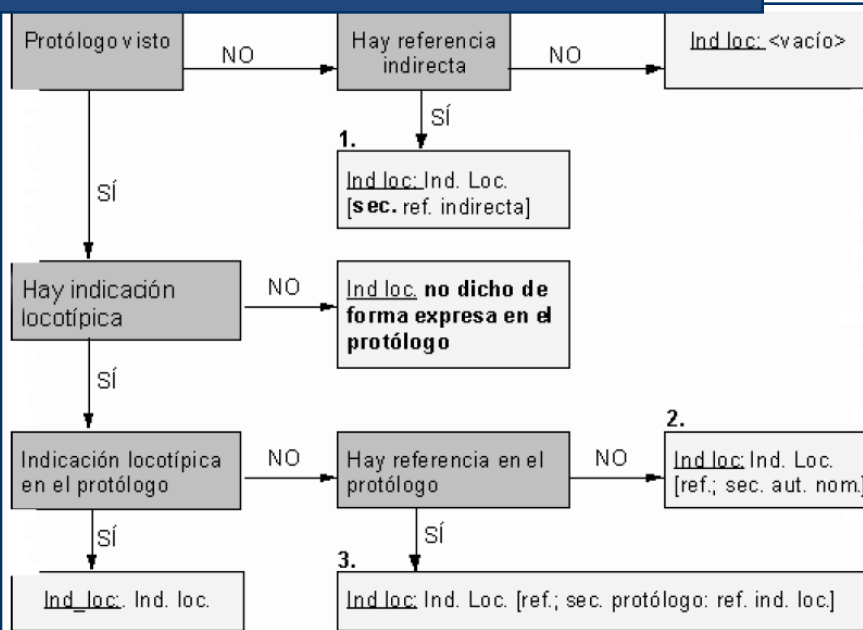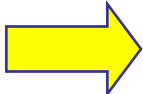
Darwin Core 1.2

Type: is the specimen a nomenclatural type or not

La clave: agregar o atomizar de manera compatible
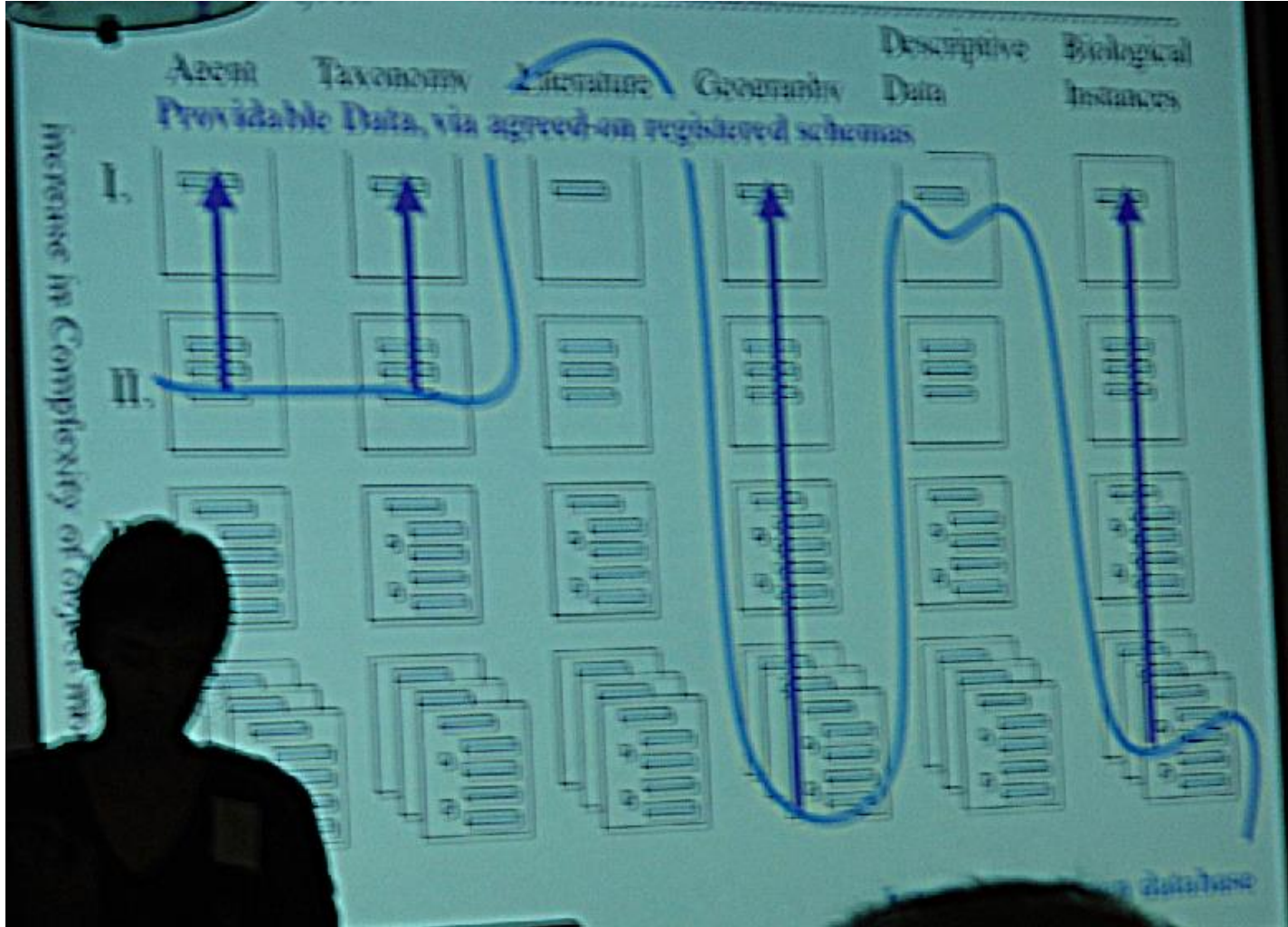
Compatibilizar estándar común y requerimientos locales: Modelo abstracto y perfiles de aplicación

Plinian Core 3

- Plinian Core Terms
  - Dataset
    - Metadata
    - TaxonRecord
      - BaseElements
      - RecordMetadata
      - NomenclatureAndClassification
      - TaxonomicDescription
      - NaturalHistory
      - Invasiveness
      - HabitatAndDistribution
      - DemographyAndThreat
      - UsesManagementAndConservation
      - associatedParty
      - References
      - AncillaryData
    - AncillaryData
- Plinian Core Classes
  - Generic
  - Complex
  - Simple
- About
- Plinian core & IPT
  - Examples

Application profiles already defined:

It's important to know how we can deploy an Applica
have to download, in the same folder where we have
this link: xsd abstract-model stable-version. Then, wh
Altova, this one will validate the AP correctly.

- **SIB_COLOMBIA:** AP_SIB-COLOMBIA_v3.2.1

- **CONABIO:** AP_CONABIO_v3.2.1

- **MAGRAMA:** AP_MAGRAMA_v3.2.1

Modelo abstracto vs Perfil de aplicación

https://github.com/pliniancore/documentation/wiki/SynonymsClass

https://github.com/PlinianCore/Documentation/wiki/SynonymsAtomizedClass
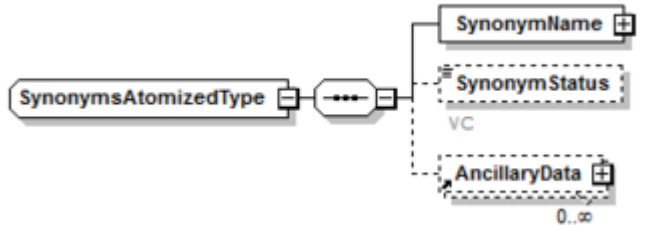
http://tdwg.napier.ac.uk/tcs_doc100/#complexType_ScientificName_Link03E58FE8

TDWG & GBIF

TDWG proporciona estándares y comunidad de expertos para las redes de datos del GBIF

GBIF marca la dirección y las necesidades de desarrolladores y usuarios, y utiliza los estándares

Consideraciones finales

- No inventar la rueda

- Es mejor un estándar ampliamente usado que uno más rico de menor implantación

- Un estándar de intercambio no debe ser tomado como una norma de implementación en un sistema de gestión/producción

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.   YEAH!

SOON:

SITUATION:
THERE ARE
15 COMPETING
STANDARDS.

Francisco Pando

Unidad de coordinación, GBIF España
Real Jardín Botánico - CSIC
Claudio Moyano 1, 28014 Madrid, Spain
pando@gbif.es
www.gbif.es , twitter.com/GbifEs , datos.gbif.es
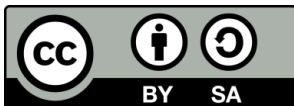
GBIF-ES es el Nodo Nacional de Información sobre Biodiversidad patrocinado por el Ministerio Español de Economía y Competitividad, gestionado por el Consejo Superior de Investigaciones Científicas

Argema mittrei