Normas para el manejo e intercambio de información sobre biodiversidad

Francisco Pando
GBIF España

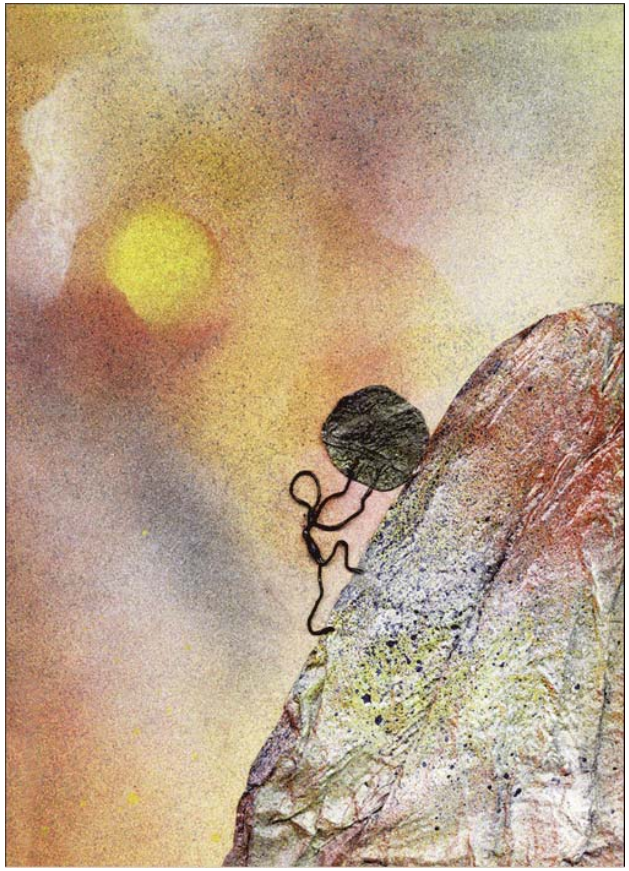
Ciudad de la Habana, Cuba
Del 8 al 12 de noviembre, 2010

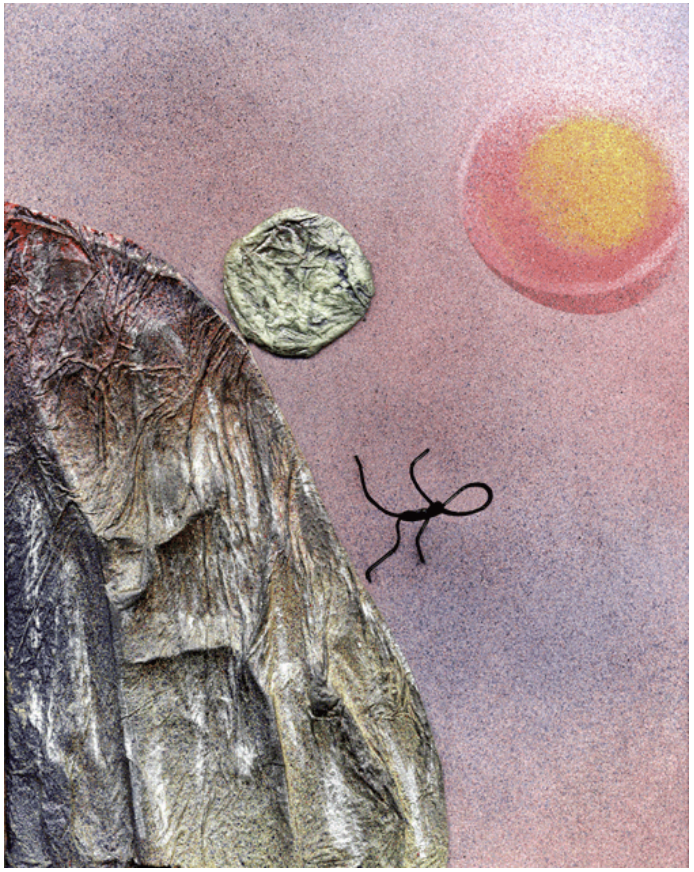
Taller “GBIF Cuba 2010”
--Proyecto de “Mentoring” de GBIF--

Sumario

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

Compartir para que





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

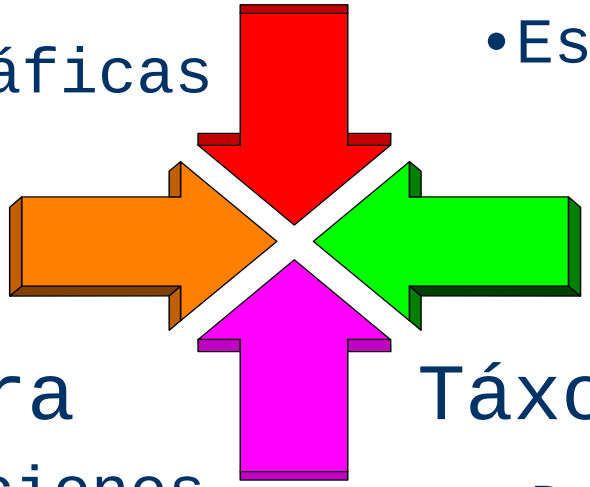
La naturaleza de la información en biodiversidad

Datos primarios

- Especimenes
- Observaciones
- Citas bibliográficas

Nombres

- válidos & sinónimos
- Información de tipos
- Esquemas taxonómicos



Literatura

- Publicaciones
- TL2, BPH, etc
- Palabras clave

Táxones

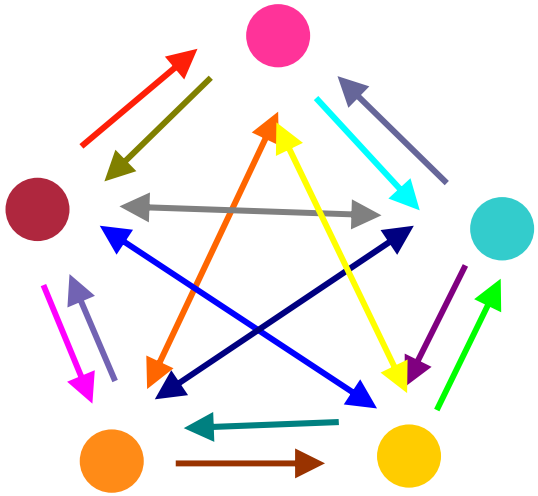
- Descripciones, claves de identificación, conservación, usos, distribución, hábitat, etc.

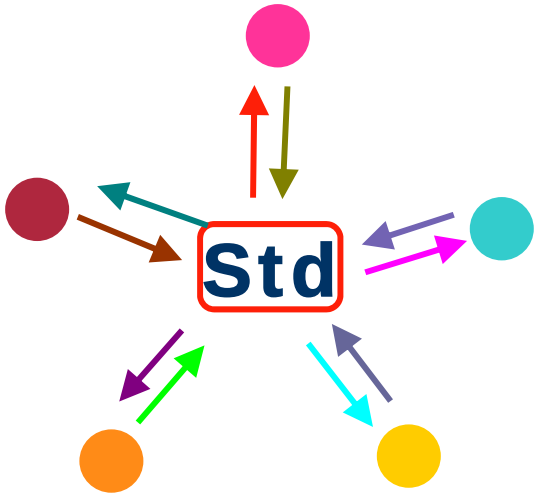
Adaptado de:
Leenhouts, Regnum Veg. 58. 1968.

Estandarizar para qué

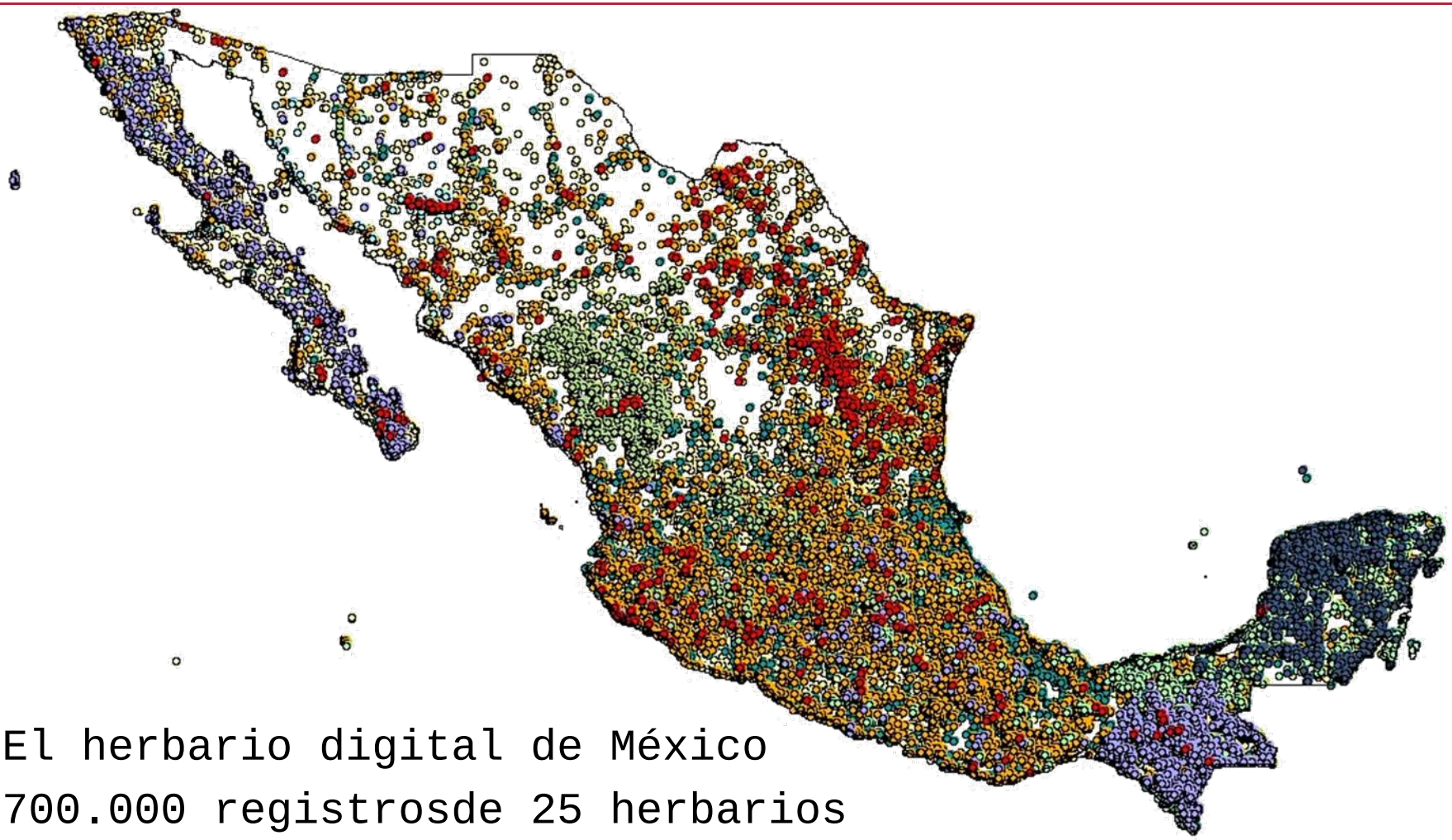
- Estandarizar para intercambiar
- Estandarizar para utilizar
- Estandarizar no para producir

Estandarizar para intercambiar





Estandarizar para utilizar



El herbario digital de México

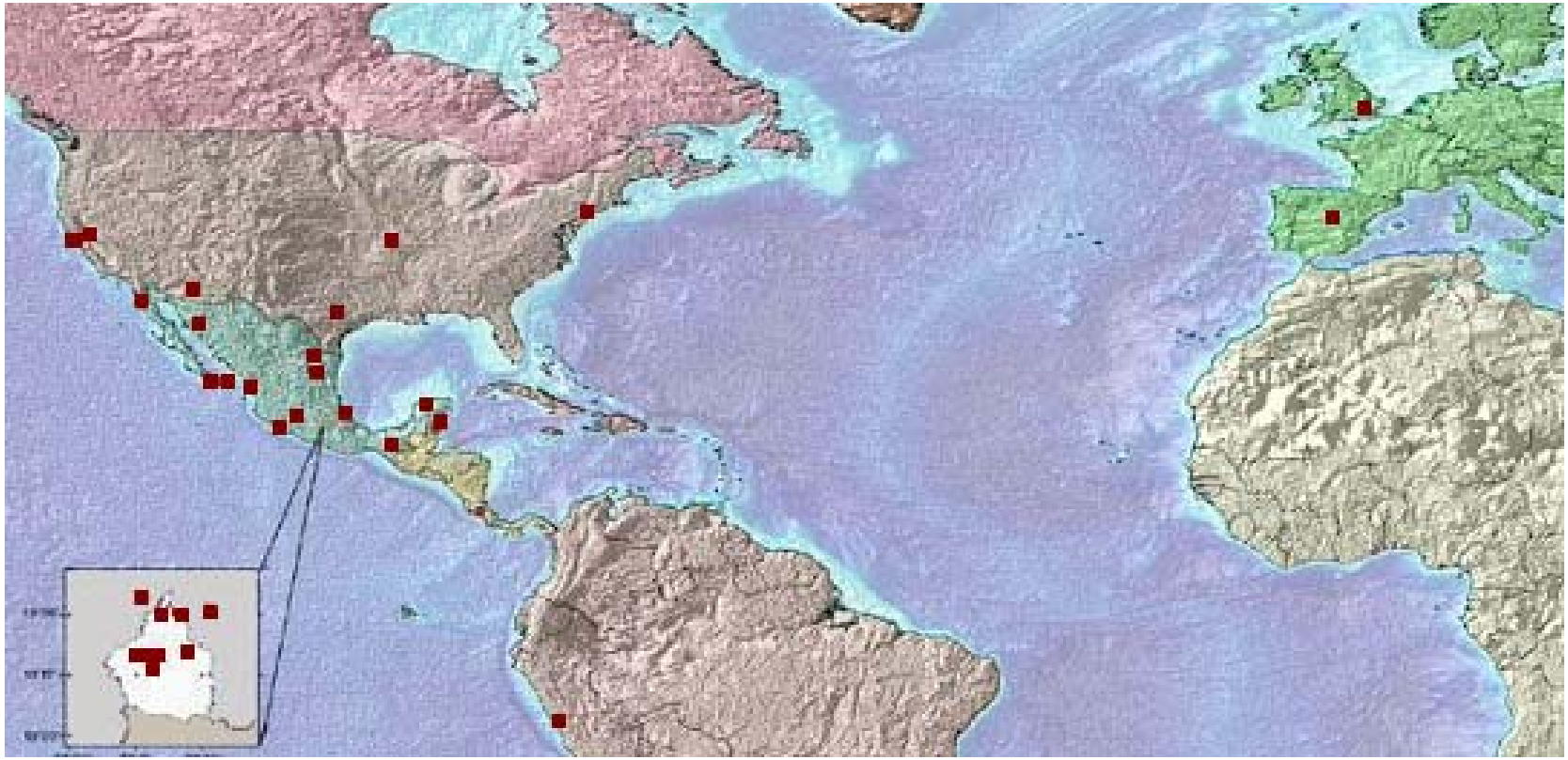
700.000 registrosde 25 herbarios

En México y E.E.U.U.

J. Soberón, 2004

Remib

Red mundial de información sobre biodiversidad (CONABIO)

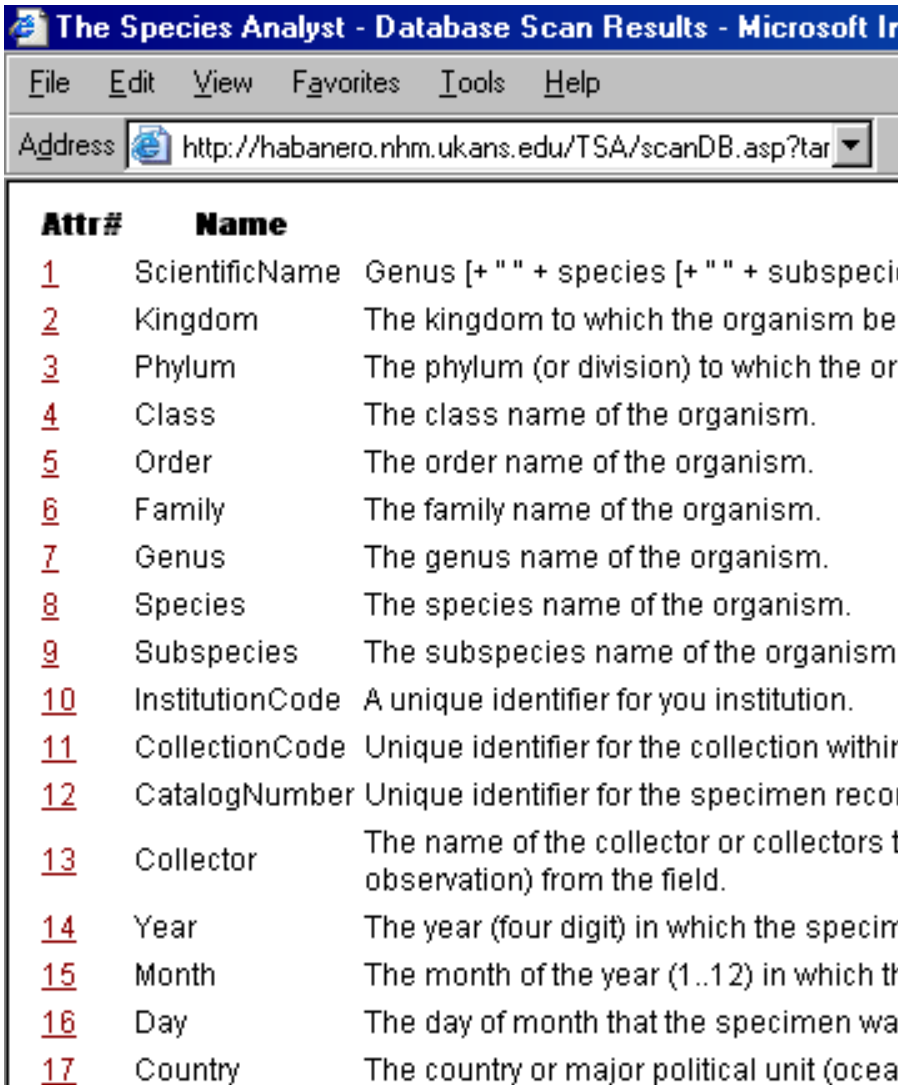


Estandarizar no para producir

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



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

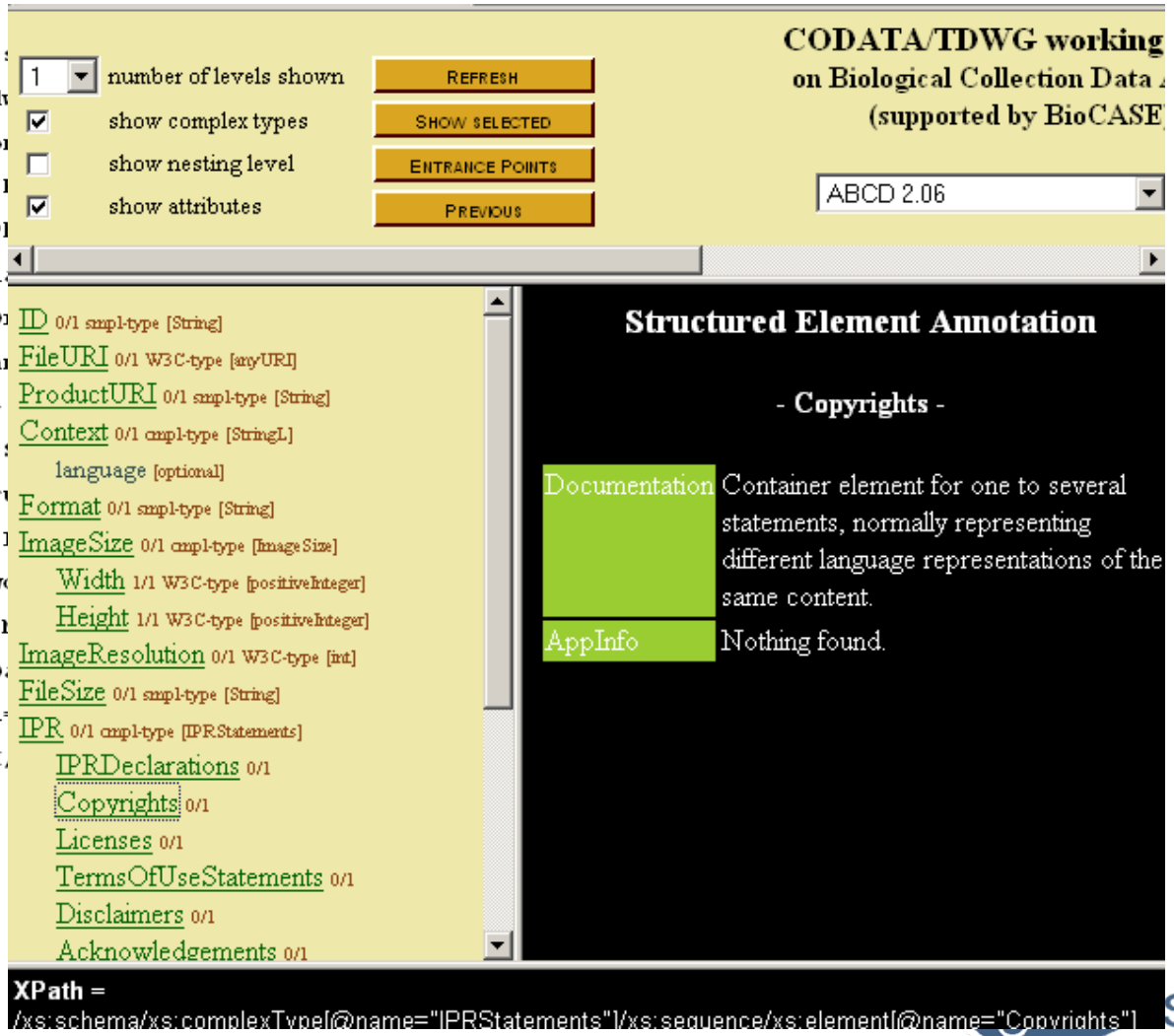
Tipos de estándares

- Vocabularios controlados (léxicos)
  - Que valores son válidos
  - (listas de géneros,…)
- Estructurales (sintácticos)
  - Como se codifica la información
  - (v.gr.: DwC , ABCD)
- Definiciones (semánticos)
  - para humanos, para máquinas
  - (v.gr.: HISPID)

Vocabularios controlados

Estructurales (sintácticos)

```
<record>
    <dwc:DateLastModified>Mar 7, 2004</dwc:DateLastModifie
    <dwc:InstitutionCode>Univeristy of Alaska Museum of th
North</dwc:InstitutionCode>
    <dwc:CollectionCode>Mammal
    <dwc:CatalogNumber>15985</dw
    <dwc:ScientificName>Puma co
    <dwc:Kingdom>Animalia</dwc:K
    <dwc:Phylum>Chordata</dwc:Ph
    <dwc:Class>Mammalia</dwc:Cl
    <dwc:Order>Carnivora</dwc:Or
    <dwc:Family>Felidae</dwc:Fam
    <dwc:Genus>Puma</dwc:Genus>
    <dwc:Species>concolor</dwc:S
    <dwc:Subspecies xsi:nil='tru
    <dwc:ScientificNameAuthor>(K
    <dwc:YearCollected>1985</dwc
    <dwc:MonthCollected>1</dwc:M
    <dwc:DayCollected>27</dwc:Da
    <dwc:ContinentOcean xsi:nil=
    <dwc:Country>UNITED STATES</
```



Definiciones (semánticos)

HISPID

Herbarium Information Standards and Protocols for Interchange of Data

Genus Name

Transfer code: gen

TDWG Short name: GENUS

Description: The name of the genus of a plant or ,if part of a hybrid formula or intergrade, then the nam
'parent' of that formula or intergrade, entered in full.

Relevant standards: ABIS, ITF, MFN, Bisby (1994).

Domain/Range/Values: Alpha; any valid genus name, capitalisation of the first letter only.

Comments: This field, combined with the following **Species Epithet** and **Infraspecies Epithet** fields, co
minimum taxonomic information for the identity of the specimen.

This field may contain an embedded hyphen, in which case the second word is in lowercase

1. This field *must* contain one of the following:

1.1 A validly published generic name under the ICBN or a manuscript name.

If the **Hybrid Flag field is** *not* filled (not transferred), then this field must contain:

1.1.2 A nonhybrid name, validly published under the ICBN or a manuscript name.

If the name of the genus of the plant is unknown, then:

gbif.es

Que hay (la situación mundial)

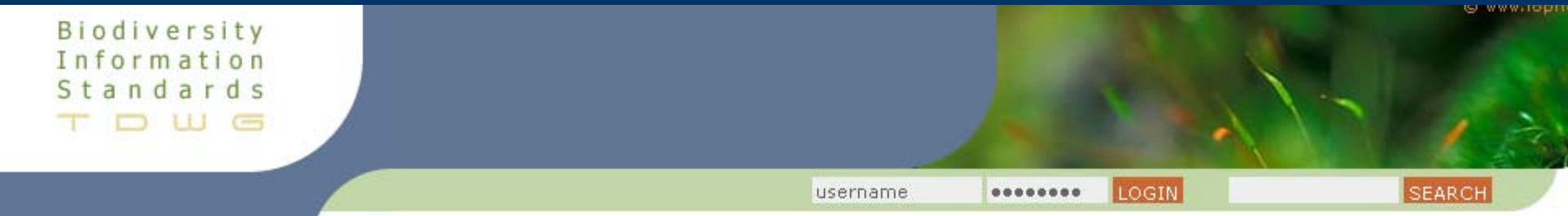
- **TDWG**

  “Taxonomic Databases Working Group”

  --Biodiversity Information Standards–

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

TDWG: Biodiversity Information Standards



Home
About TDWG
Getting Started
Standards
Activities
Membership

conference2010

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

Wiki
OJS
Mailing Lists
Glossary

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.

We Need Your Support

Please consider becoming a **member**. TDWG needs everyone with a strong interest in biodiversity information to participate in developing effective standards for sharing that information. We need IT professionals, taxonomists, ecologists, geoscientists, and librarians. We need members from institutions and agencies that do or could use our

Latest News

02-Sep-2010 Scholarships at Royal Museum for Central Africa

The Royal Museum for Central Africa (RMCA) from Belgium has launched a new call for scholarships for scientific study visits on biodiversity in African Zoology and Wood Biology. The programme targets African scientists and...

[more]

25-Aug-2010 Biofinity Workshop following

gbif.es

TDWG Current Standards (*)

http://www.tdwg.org/standards/

Title	Activity (Task or Interest Group)	Category	Status	Date Published	
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 Descriptions Interest Group	Technical Specification	Current (2005)	16-Sep-2005	Download
Taxonomic Concept Transfer Schema	Taxonomic Names and Concepts Interest Group	Technical Specification	Current (2005)	16-Sep-2005	Download
Definition of the Delta Format		Technical Specification	Current (2005)	01-Oct-1986	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 (**)
GUID and Life Sciences Identifiers Applicability Statements		Applicability Statements	Draft	26-Oct-2009	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

Sobre Darwin Core

Biodiversity Information Standards TDWG

TDWG Wiki > DarwinCore

UserName ••••••••• LOGIN SEARCH

TDWG Home
Wiki Home

DarwinCore Home
Changes
Index
Search
Notifications
Statistics
Users
Preferences

Group Wikis
Collections Data (ABCD)
Collections Descriptions
Darwin Core
Biological Descriptions
Geospatial
Global Identifiers (GUID)
Images
Invasive Species
Literature
Names & Concepts
Observational
Phylogenetics

Edit Attach Printab

You are here: TWiki > DarwinCore Web > WebHome

r46 - 09 Oct 2009 - 15:58:49 - JohnWiecz

Historical DarwinCore wiki site. Deprecated.

Note: These Wiki pages are for historical purposes, they **do not** reflect the content of the current standard, which can b found at

* http://rs.tdwg.org/dwc/index.htm

Table of Contents

Definition

The Darwin Core (sometimes abbreviated as DwC) is a standard designed to facilitate the exchange of information about the geographic occurrence of species and the existence of specimens in collections. Extensions to the Darwin Core provi a mechanism to share additional information, which may be discipline-specific, or beyond the commonly agreed upon sco of the Darwin Core itself.

- Background and History
  - Darwin Core Versions
- Design and Purpose
- Darwin Core and Extensions (Concept Lists)

gbif.es

Daewin Core: Extensiones

Definition

The Darwin Core (sometimes abbreviated as DwC) is a standard designed to facilitate the exchange of information about the geographic occurrence of species and the existence of specimens in collections. Extensions to the Darwin Core provide a mechanism to share additional information, which may be discipline-specific, or beyond the commonly agreed upon scope of the Darwin Core itself.

- Background and History
  - Darwin Core Versions
- Design and Purpose
- Darwin Core and Extensions (Concept Lists)
  - Darwin Core
  - Curatorial Extension
  - Geospatial Extension
  - Paleontology Extension
  - Interaction Extension
- Darwin Core Group Task Group
  - Draft Charter
- Pending Issues
- GBIF Circa Discussion (archive 29 Sep 2005 to 09 Jun 2006)
- Distributed Generic Information Retrieval (DiGIR)

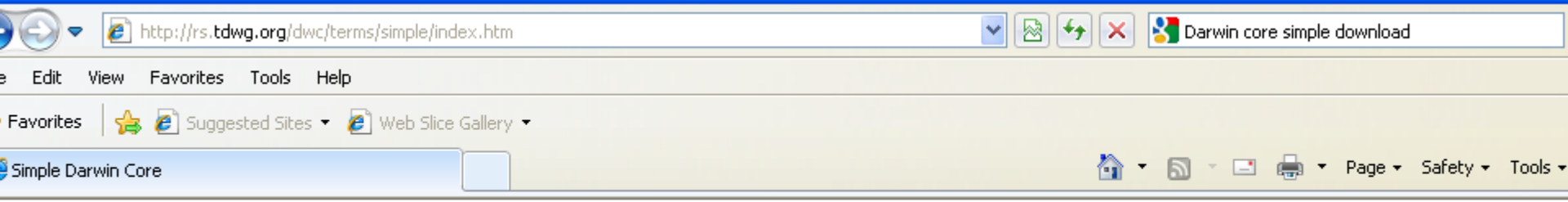
DwC: múltiples versiones

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)

The normative form of any Darwin Core version is its XML schema document (*.xsd).



Biodiversity Information Standards TDWG

Simple Darwin Core

Title: Simple Darwin Core

Date Issued: 2009-04-21

Date Modified: 2009-12-07

Abstract: This document is a reference for the Simple Darwin Core standard.

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/2009-12-07/terms/simple/

Latest Version: http://rs.tdwg.org/dwc/terms/simple/

Replaces: http://rs.tdwg.org/dwc/2009-09-23/terms/simple/

Document Current Standard

gbif.es

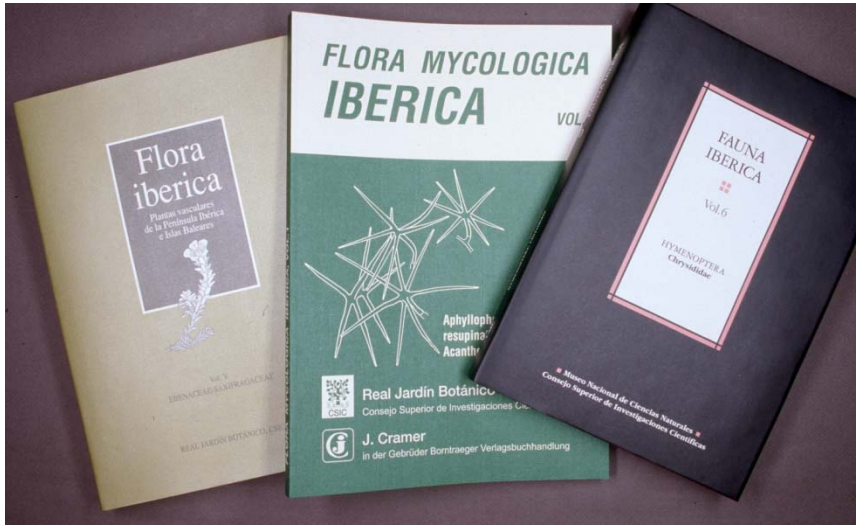
Estándares temáticos y/o nacionales

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



SiB Sistema de Información sobre Biodiversidad de Colombia

Enlaces

- Archivos de autoridad taxonómica

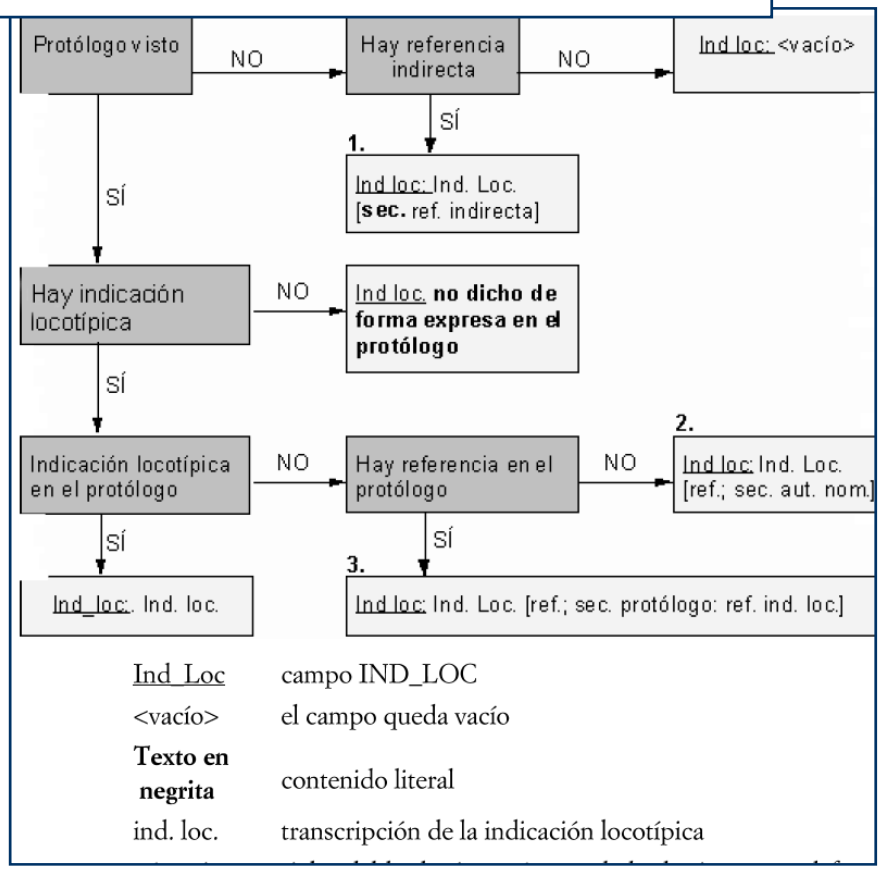
Enlaces

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 Saénz 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 (Hymenoptera neotrópico
8. Rodríquez-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, Psyc Rodriguezia, Schlimia y Selenipedium (Orchidaceae) para Colomb
10. Brummitt R. K. (1992) Vascular plant families and genera: a visit

Estándares más ricos, tambíen más restringidos

Estándar proyectos *Flora iberica & Flora Mycologica Iberica*

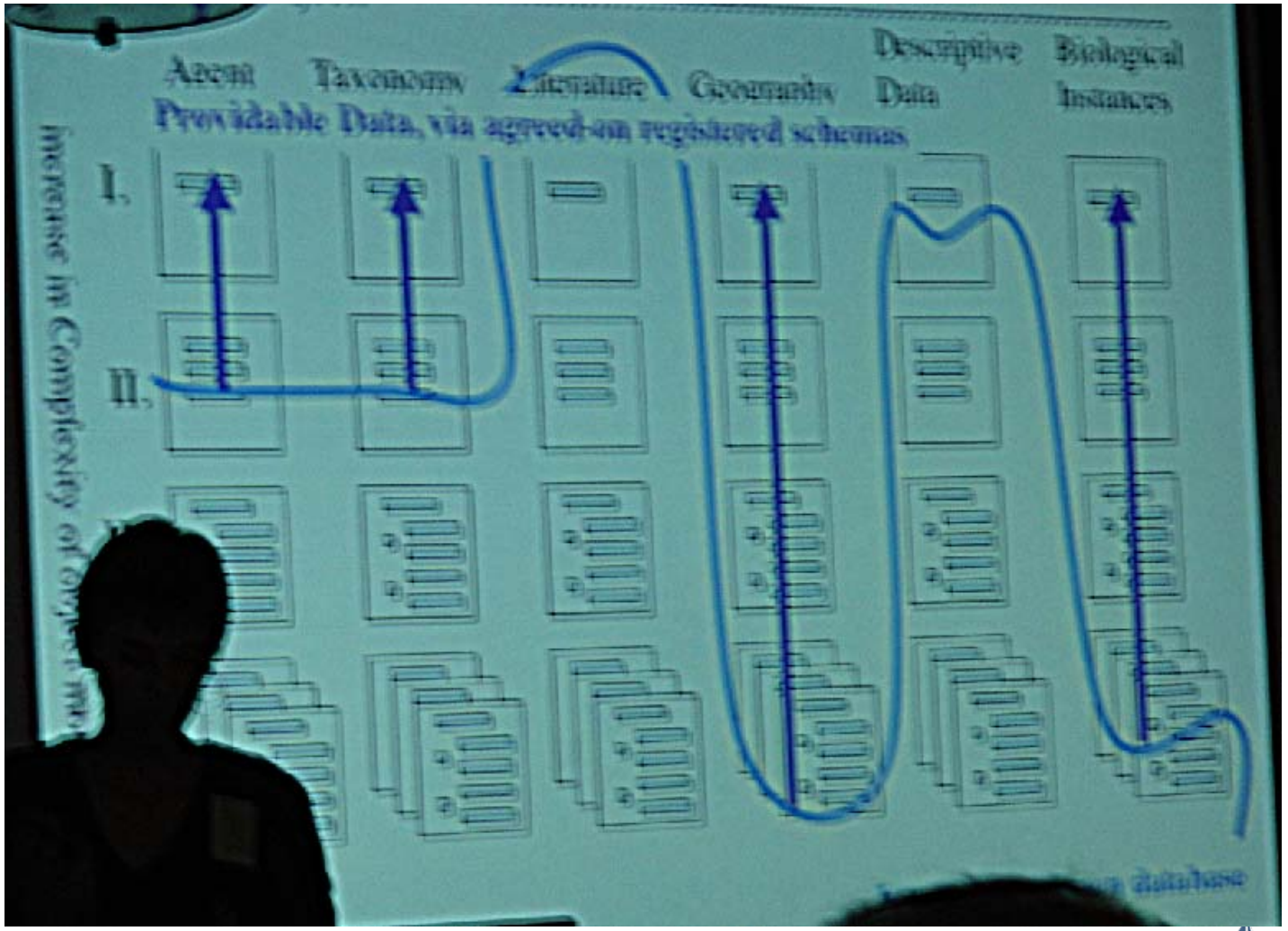


Estándar TDWG & Red de datos del GBIF

Darwin Core 1.2

Type: is the specimen a nomenclatural type or not

La clave: agregar o atomizar de manera compatible



TDWG & GBIF

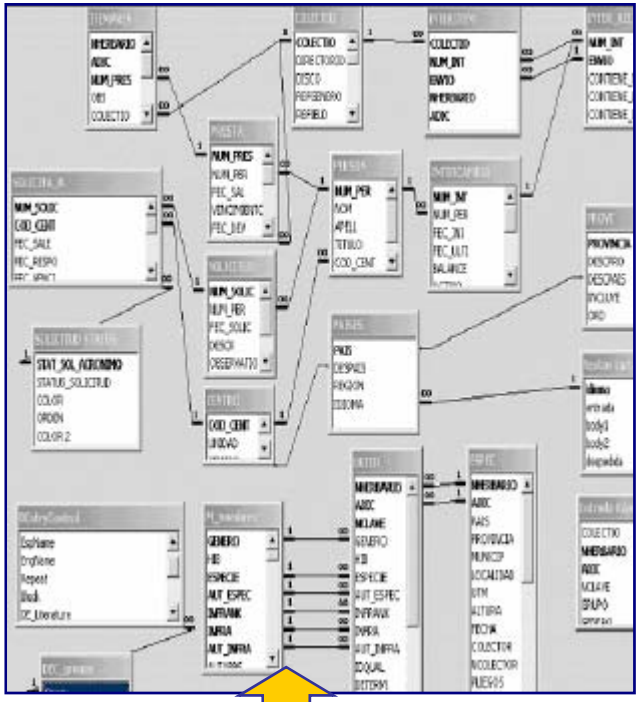
TDWG proporciona estándares y comunidad de expertos para las redes de datos del GBIF

GBIF marca la dirección y las necesidades de desarrolladores y usuarios

Que va a haber

- Web semántica
- Ontologías
- esquemas xml,
- triple storage
- ...

Web semántica: donde estamos

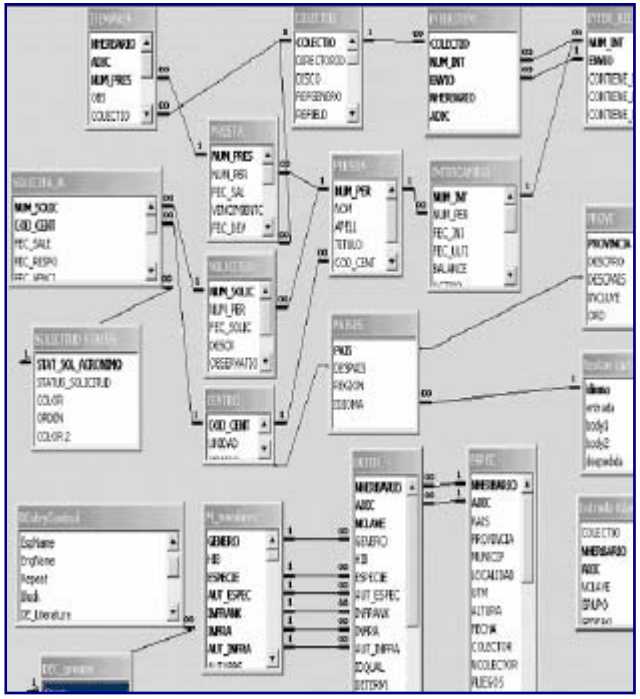


Attr#	Name
1	ScientificName
2	Kingdom
3	Phylum
4	Class
5	Order
6	Family
7	Genus
8	Species
9	Subspecies
10	InstitutionCode
11	CollectionCode
12	CatalogNumber
13	Collector
14	Year
15	Month
16	Day
17	Country
18	StateProvince
19	County
20	Locality
21	Longitude

CatalogNumber	Unique identifier for the specimen record within
Collector	The name of the collector or collectors that were observation) from the field.
Year	The year (four digit) in which the specimen was
Month	The month of the year (1..12) in which the speci
Day	The day of month that the specimen was collec
Country	The country or major political unit (ocean) from
StateProvince	The state, province or region (i.e. next political r collected.
County	The county (or shire, or next political region sma
Locality	The locality description (place name plus option specimen was collected.

La compatibilizacion entre sistemas se hace “a mano”

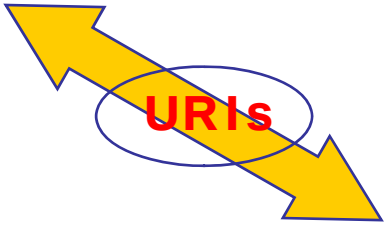
Web semántica: comunicar contenidos

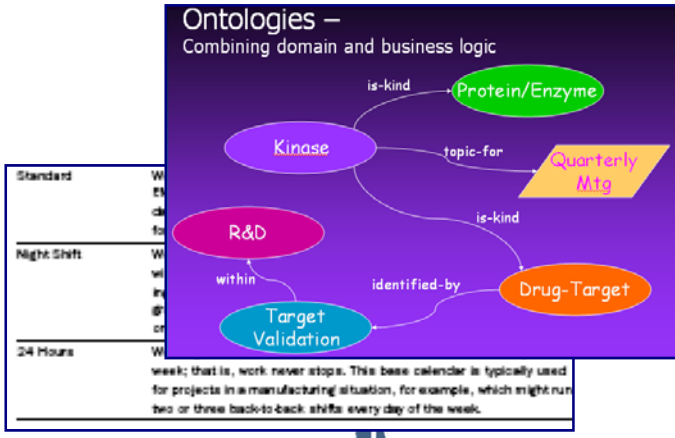


Las definiciones de las bases de datos se almacenan (o hacen uso) en un sistema:

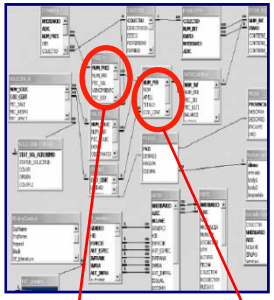
- *Accesible vía Web*
- *Ininteligible por "maquinas"*
- *Relaciona conceptos, relaciones e instancias*

→ *ontologías*





Web semántica: Triple storage



Sujeto	Predicado	objeto
25322	genero	Inga
25322	especie	alba
25322	pais	COL
COL	Pais largo	Colombia

XML

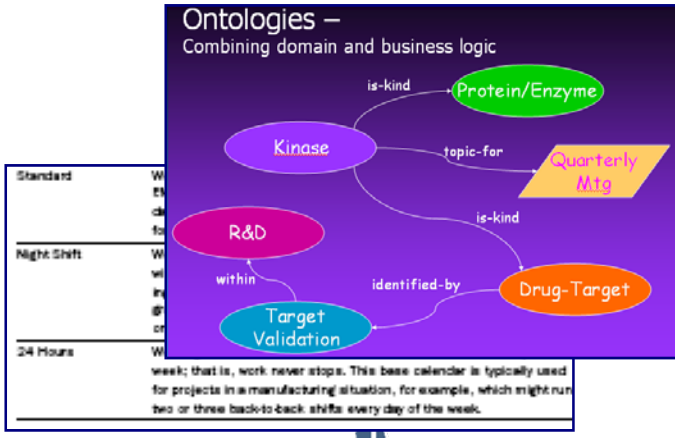
Especimenes

Campo	valor
Nmr_ejemplar	***25322***
genero	*Inga*
especie	alba
pais	COL

URIs

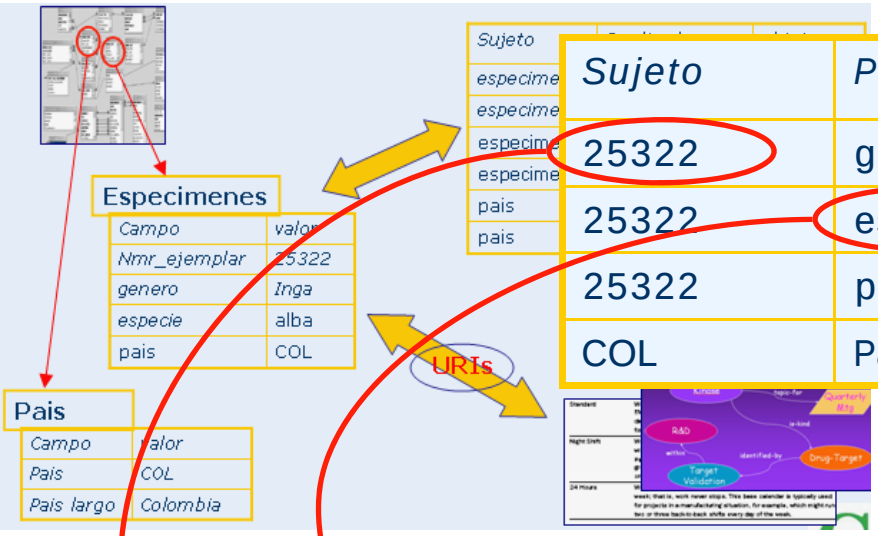
Pais

Campo	valor
Pais	***COL***
Pais largo	*Colombia*



gbif.es

Web semántica: ídentificadores únicos

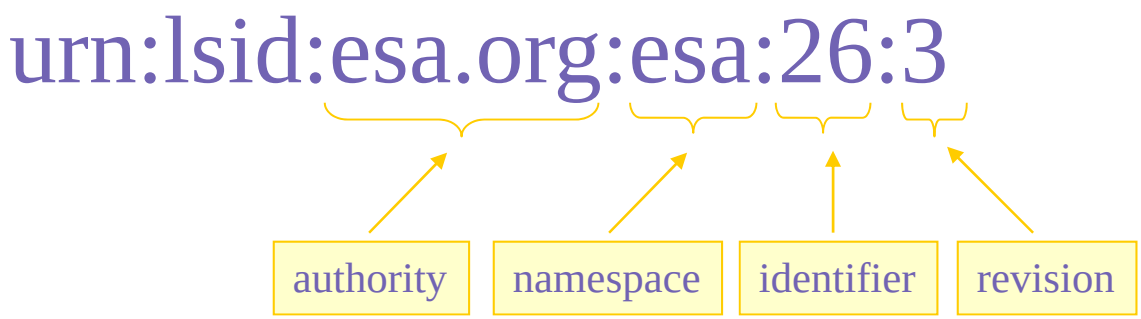


Sujeto	Predicado	objeto
25322	genero	Inga
25322	especie	alba
25322	pais	COL
COL	Pais largo	Colombia

- ***Identificadores únicos de conceptos (ontología)***
- ***Identificadores únicos de objetos***

Web semántica: Identificadores globales -> LSID

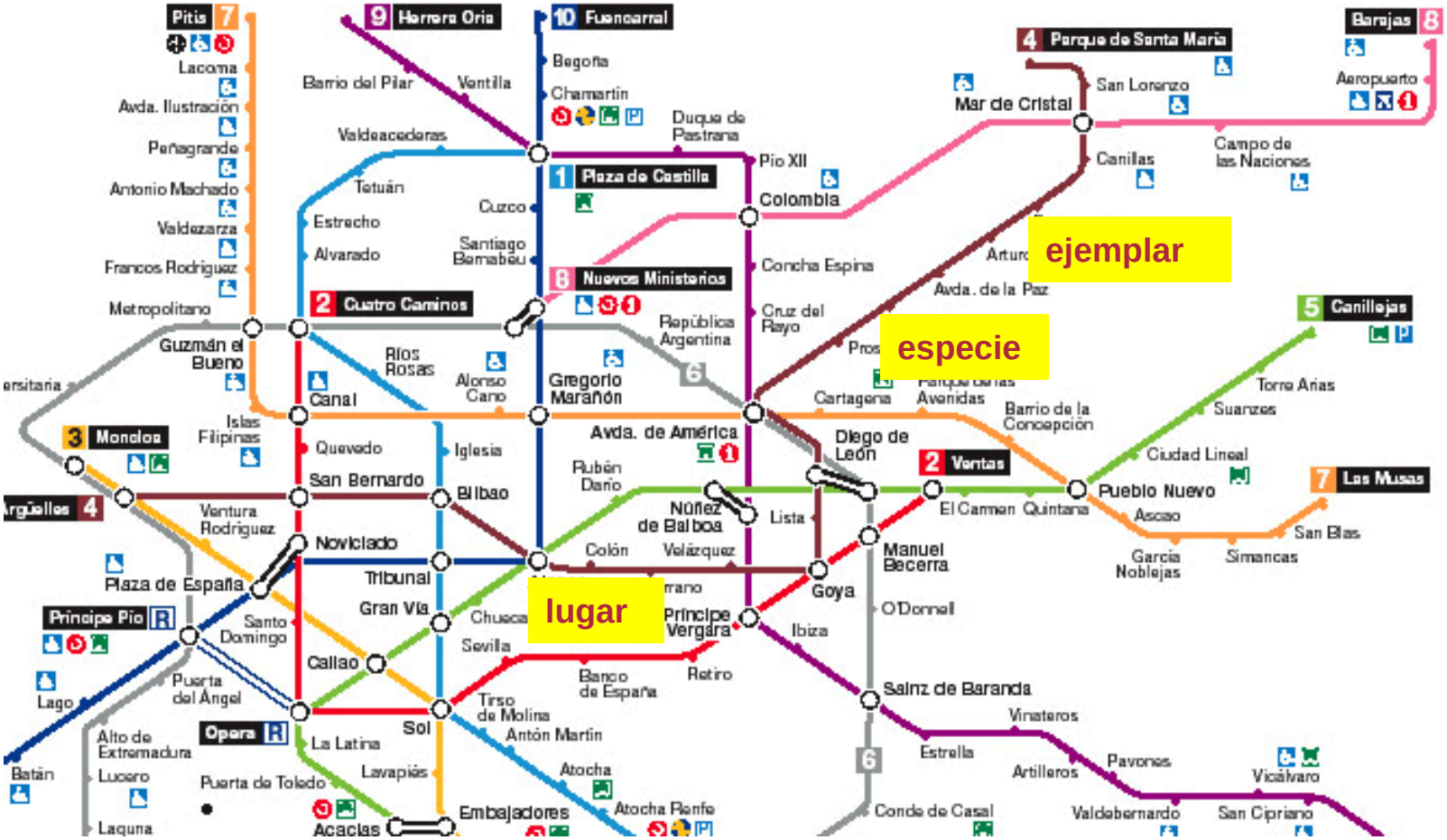
- Life Science Identifiers (LSID) son un tipo de identificadores globales
  - Facilitar comunicación e intercambio
  - Más robustos que URL
  - Se pueden resolver (esto es, los entienden las máquinas)



Ver: Adoption of Persistent Identifiers for Biodiversity Informatics

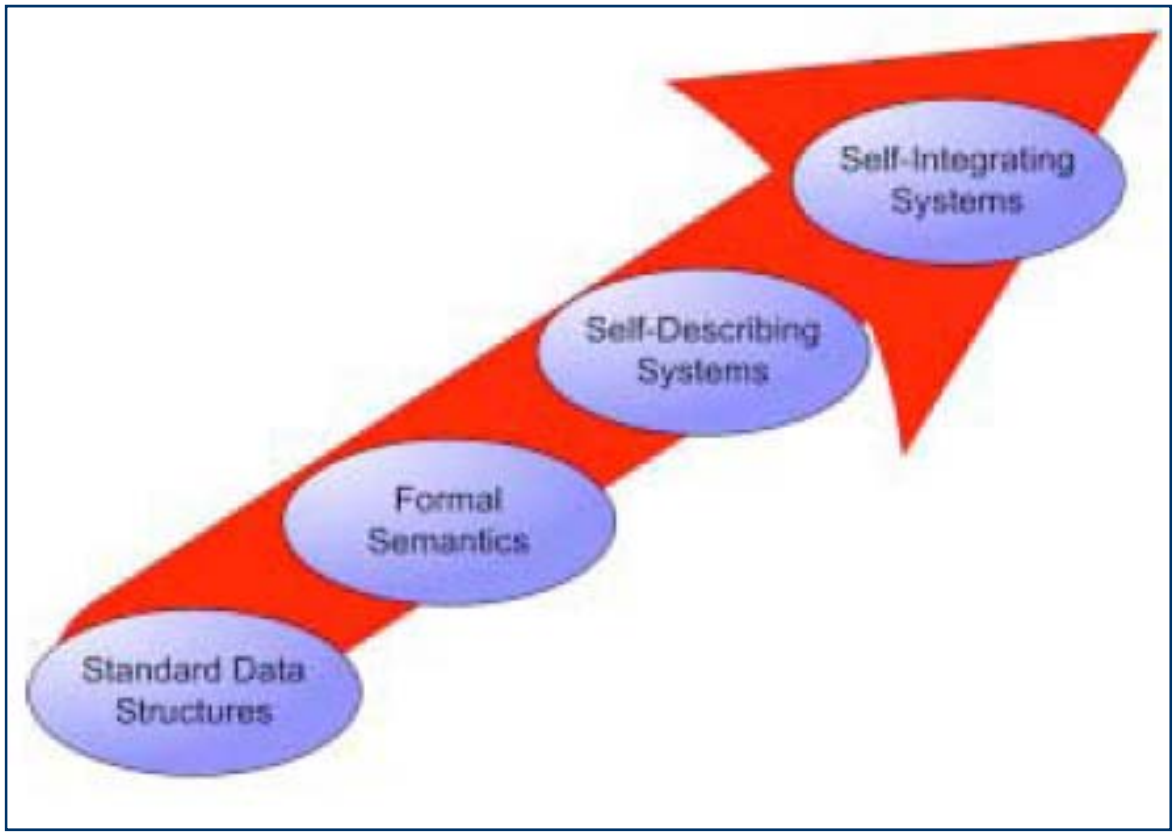
Draft recommendations of the GBIF LGTG, 18 August 2009

Web semántica: navegar los datos como el metro de una ciudad



- *Bases de datos conectadas por ontologías y la Web*
- *Conseguir que podamos inferir hechos a partir de datos distribuidos*

Hacia donde vamos



La idea es que la web funciona como un único sistema de información

Estándares y web semántica

- Los estándares del TDWG son compatibles con la Web semántica
- Resolver ontologías
  - No nos libramos de los mapeos
- Resolver LSIDs
  - Puede haber más de uno…
- Lo que hagamos debe tener en cuent el escenario de Web semántica

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

A vuestra disposición

Francisco Pando
Responsable
GBIF España, U. de Coordinación
Real Jardín Botánico-CSIC
Pza. de Murillo, 2
28014 Madrid, España

pando@gbif.es
Tel.: + 34 91 420 3017
Fax: + 34 91 429 2405
www.gbif.es
www.gbif.org