

Calidad de datos: nomenclatura y taxonomía

Taller sobre calidad en bases de
datos sobre biodiversidad

25-26 septiembre 2008. Madrid.

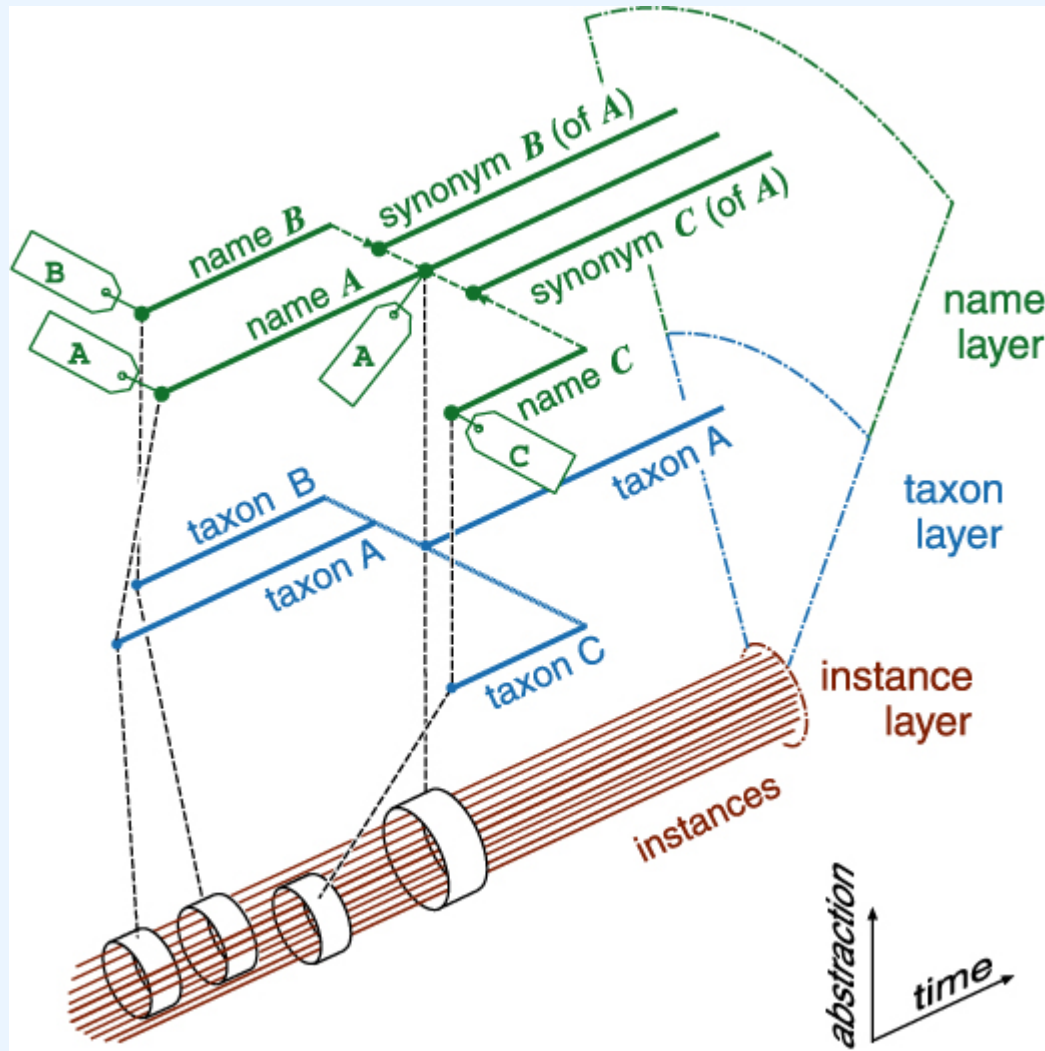
Francisco Pando
Unidad de Coordinación GBIF-España



- **De que estamos hablando**
- **Conocimiento explícito e implícito**
- **Donde estamos**
- **Metadatos de identificaciones**

De que estamos hablando

- Identificaciones, nombres, conceptos (táxones)



De Nozomi Ytow & al.

Ejemplo: el helecho macho



Conocimiento explícito e implícito

HORTUS REGIUS MATRITENSIS (MA-Fungi)

Martellia sp.

ESP. CIUDAD REAL: Viso del Marqués, Arroyo de la Poveda, bajo Halimium sp., 23-XI-1993, T. Pérez Jarauta, det. F.D. Calonge

MA-FUNGI 32269



Zelleromyces piennensis

Horno-Arroyo, Viso
F. Calonge

Rev. *José S. Vidal*

Dat. *25-5-2000*

GenBank Acc. No. AF215649
MA-Fungi 32269

nombre

quien

cuando

...

Donde estamos

- Establecer el significado de una identificación es en el mejor de los casos un ejercicio de exegesis basado en conocimiento implícito y asunciones
- Esto es un tremendo obstáculo para que la colección sea útil a un mayor número de usuarios y por tanto para su importancia

Explicitar el conocimiento...

...para hacer los datos sean útiles en un número de casos \Rightarrow tenga más calidad. Hablamos de metadatos

- ✓ **Quien hace la identificación**
- ✓ **Cuando**
 - **Tipo de identificación (fiabilidad)**
 - **Calificadores de la identificación (precisión)**
 - **Exactitud (\Rightarrow sistema de referencia \Rightarrow taxonomía)**

Tipo de identificación

From: Australian National Fish Collection (in use since 1993)

Level 1: *Highly reliable identification*

Specimen identified by (a) an internationally recognised authority of the group, or (b) a specialist that is presently studying or has reviewed the group in the Australian region.

Level 2: *Identification made with high degree of confidence at all levels*

Specimen identified by a trained identifier who had prior knowledge of the group in the Australian region or used available literature to identify the specimen.

Level 3: *Identification made with high confidence to genus but less so to species*

Specimen identified by (a) a trained identifier who was confident of its generic placement but did not substantiate their species identification using the literature, or (b) a trained identifier who used the literature but still could not make a positive identification to species, or (c) an untrained identifier who used most of the available literature to make the identification.

Level 4: *Identification made with limited confidence*

Specimen identified by (a) a trained identifier who was confident of its family placement but unsure of generic or species identifications (no literature used apart from illustrations), or (b) an untrained identifier who had/used limited literature to make the identification.

Level 5: *Identification superficial*

Specimen identified by (a) a trained identifier who is uncertain of the family placement of the species (cataloguing identification only), (b) an untrained identifier using, at best, figures in a guide, or (c) where the status & expertise of the identifier is unknown.

Fiabilidad



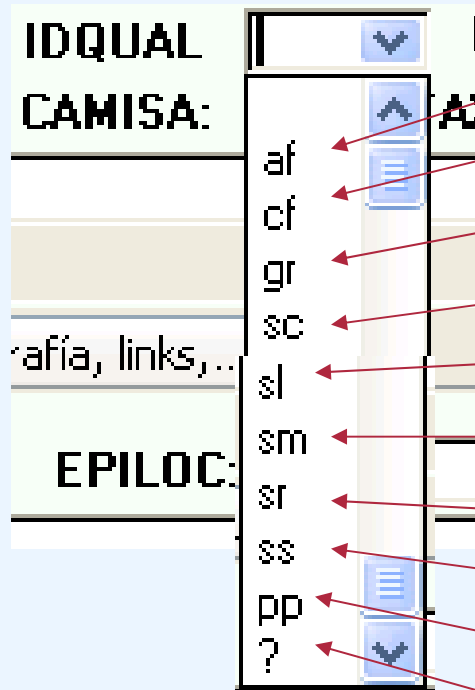
Tipo de identificación

From: Chapman (2005) Principles of Data Quality. GBIF

Suggestion:

- **identified by** World expert **in the taxa with** high certainty
- **identified by** World expert **in the taxa with** reasonable certainty
- **identified by** World expert **in the taxa with** some doubts
- **identified by** regional expert **in the taxa with** high certainty
- **identified by** regional expert **in the taxa with** reasonable certainty
- **identified by** regional expert **in the taxa with** some doubts
- **identified by** non-expert **in the taxa with** high certainty
- **identified by** non-expert **in the taxa with** reasonable certainty
- **identified by** non-expert **in the taxa with** some doubt
- **identified by** the collector **with** high certainty
- **identified by** the collector **with** reasonable certainty
- **identified by** the collector **with some doubt.**

Calificadores de la identificación



afín

comparar

grupo

sección

sensu lato

species multae

serie

sensu stricto

pro parte

duda

Precisión

*basado en ITF2, un estándar del TDWG (www.tdwg.org)

Exactitud (ref. a taxonomía)

Herbario del Real Jardín Botánico

MODIFICACIÓN DE
ESPECIMENES

NÚMERO DE HERBARIO:

261

1

Lista Det.

GRUPO Rosaceae

GENERO: Geum

HIB

ESPECIE: chiloense

Nombres

AUT_ESPEC: Balb. in DC.

INFRANK: _

INFRA: _

AUT_INFRA: _

IDQUAL af

DETERM: C. Navarro

MESAN: _

ES_TIPO

CAMISA: S

TAXONOMIA Fl.iber.

OBSERV: _

Fl.iber.

Castroviejo, S.(ed.) (1983). Flora iberica

Registro: 1 de 2

Localidad, habitat, fecha

Formato, Exsiccata, pliegos...

Iconografía, links,...

PAIS: ESP

PROV: Cc

EPILOC: _

LOCALIDAD: Baños de Montemavor. Puerto de la Garrañta

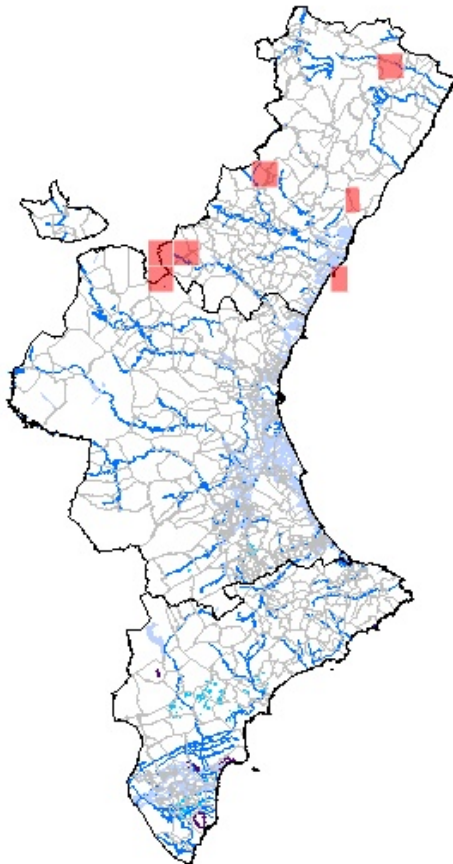
Precisión (y asunciones)

Verbascum thapsus

<http://bdb.cth.gva.es> - Mapa Especie - Mozilla Firefox

Nombre científico: **Verbascum thapsus**

Mapa de distribución:

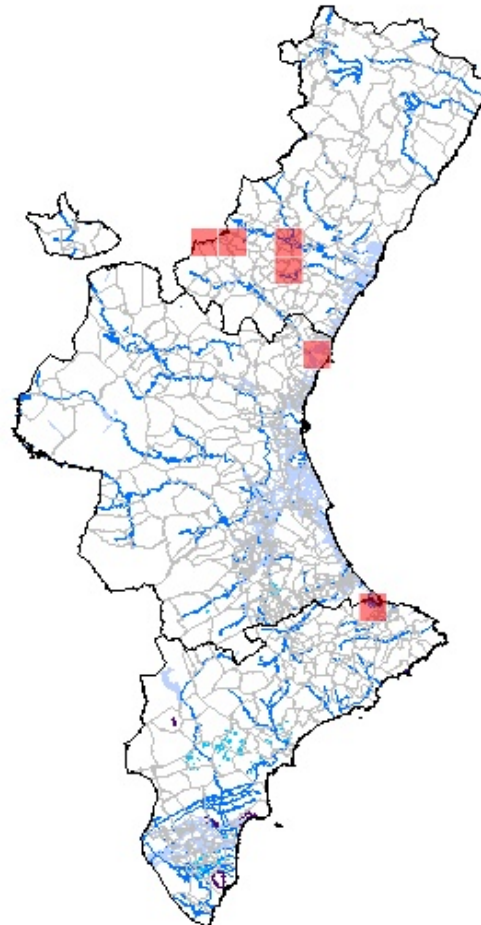


Fuente Mapa: Conser

BDBC

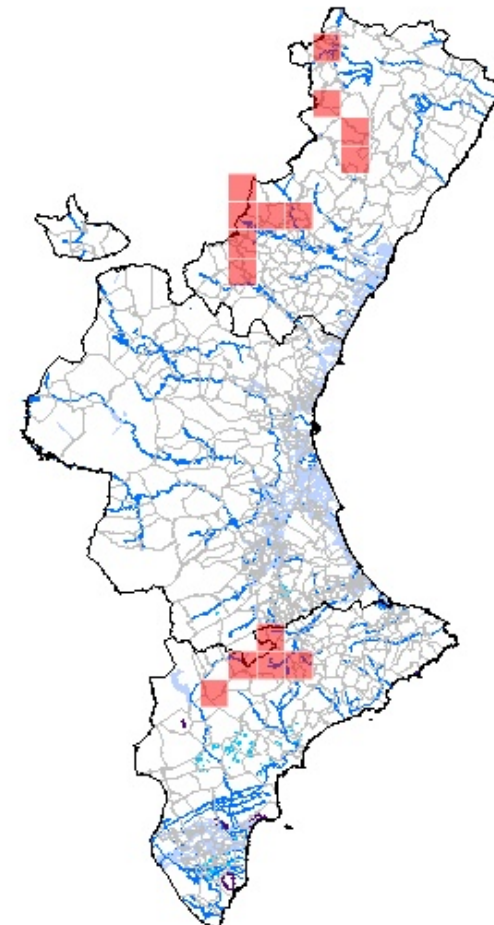
Nombre científico: **Verbascum thapsus giganteum**

Mapa de distribución:



Nombre científico: **Verbascum thapsus montanum**

Mapa de distribución:



Reflexión final

Pasar de un entorno de trabajo basado en papel a uno basado en formato electrónico supone que:

- **La información resulta mucho más accesible,**
- **La distancia entre el origen del dato y el usuario se agranda y**
- **Datos de fuentes heterogéneas son usados de manera combinada**

...todo lo cual requiere de datos más contextualizados para que sean usados apropiadamente