

Comparison of species information TDWG standards from the point of view of the Plinian Core specification

Francisco Pando
Real Jardín Botánico-CSIC. Spain



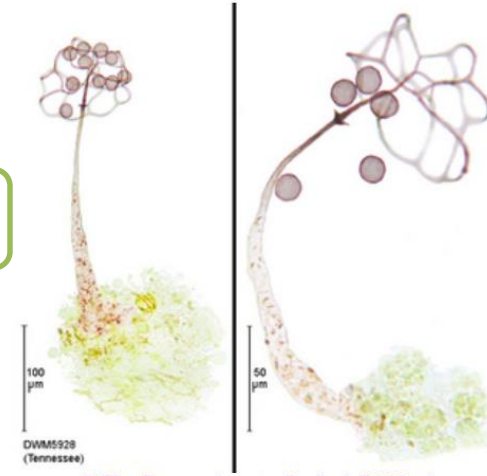
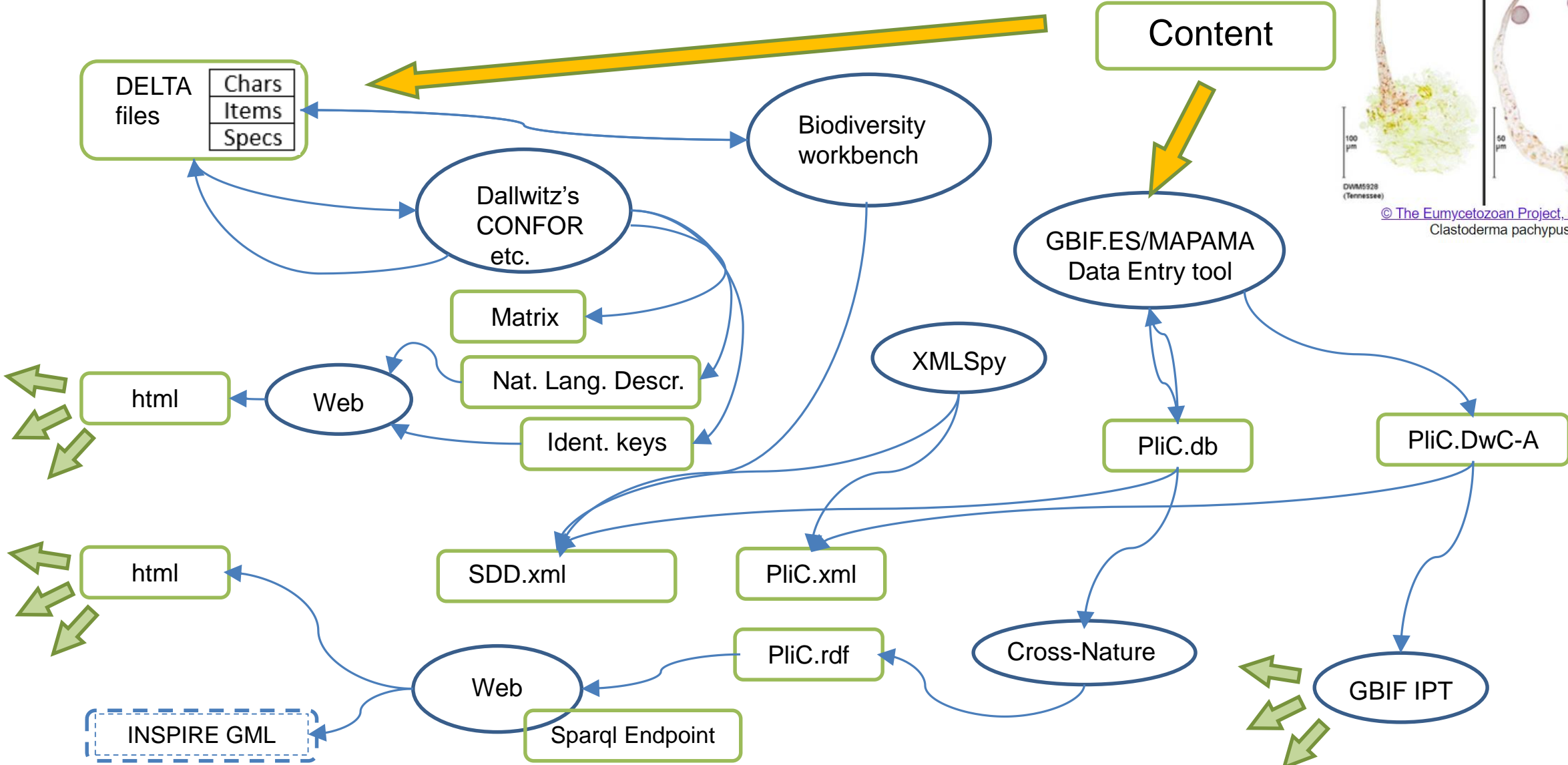
PLINIAN CORE



Summary

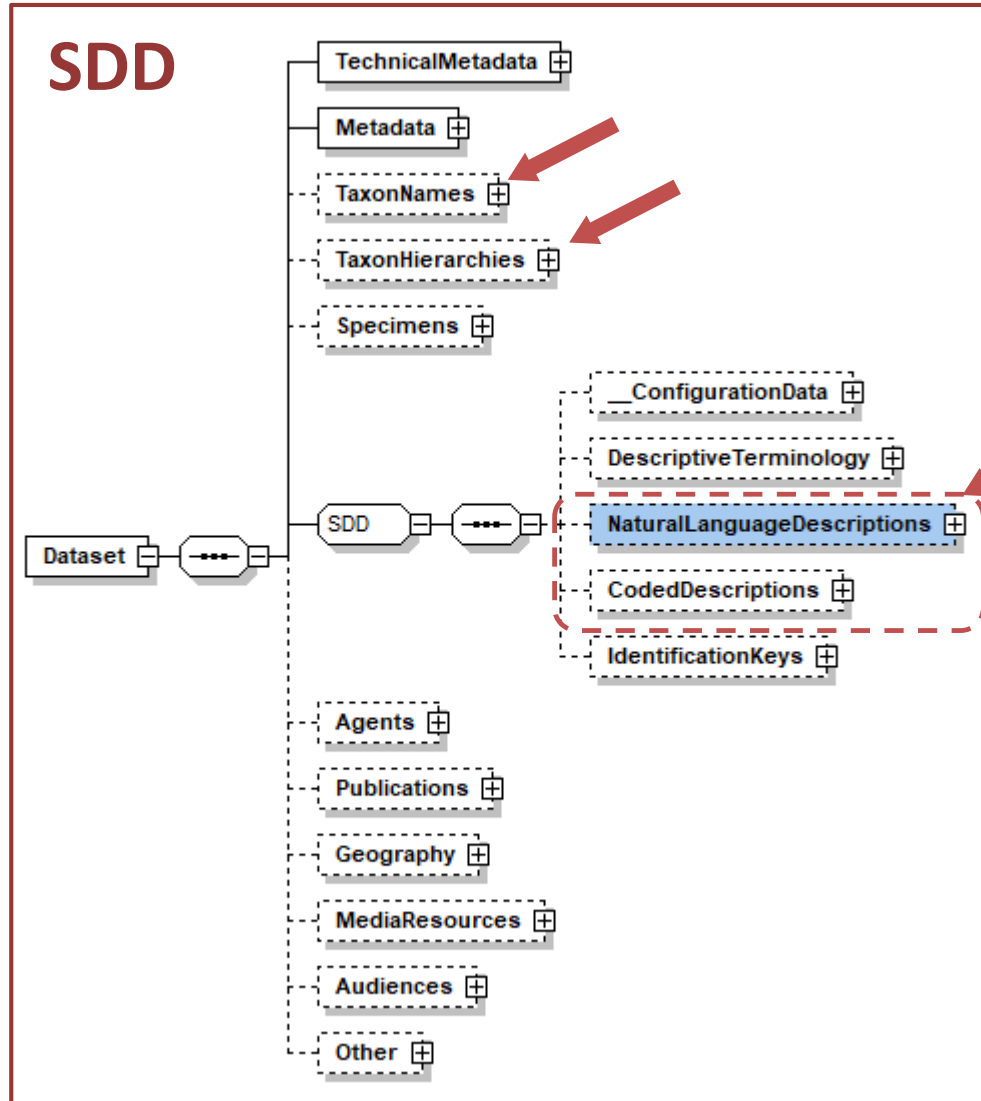
- Species-level standards Universe
- Methodology
- Coded descriptions in DELTA/SDD and PliC
 - Conclusions 1
- Natural Language descriptions
 - Conclusions 2
- Tools , flavours and outputs in DELTA/SDD and PliC
 - Conclusions 3

Methodology

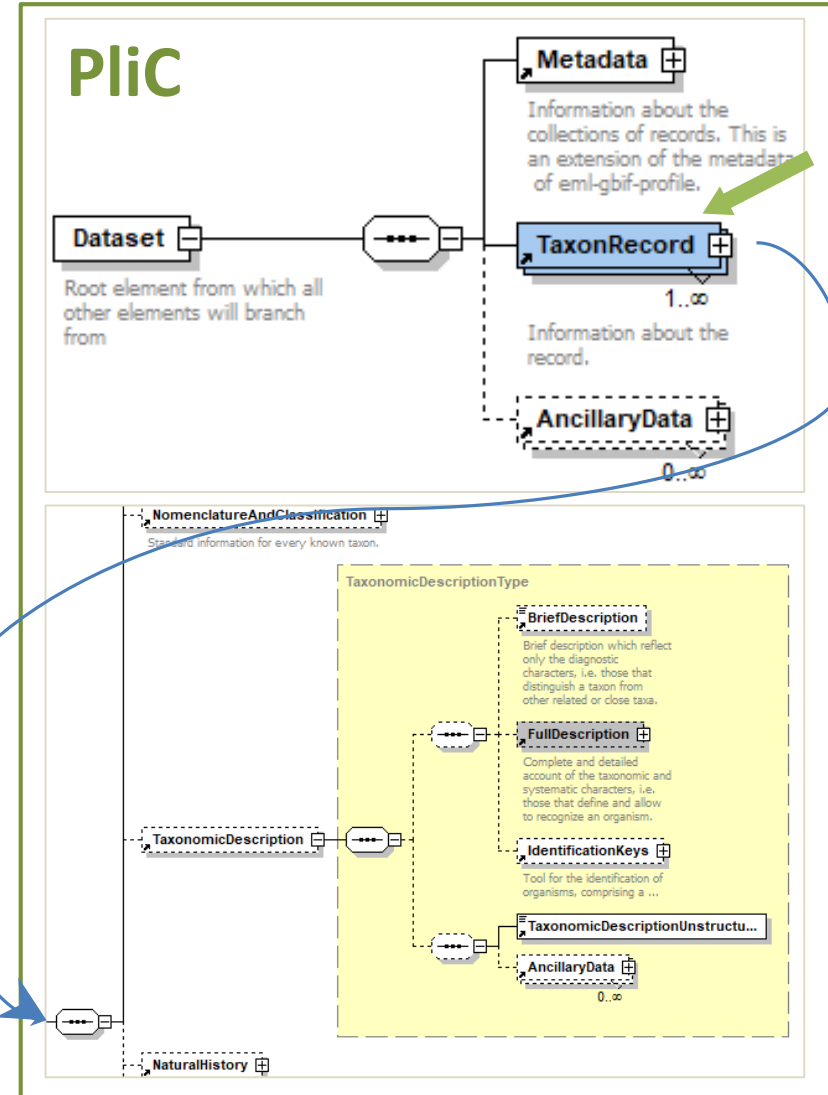


© The Eumycetozoan Project, 2006 · 0
 Clastoderma pachypus

Conceptual approaches



The unit is the Dataset

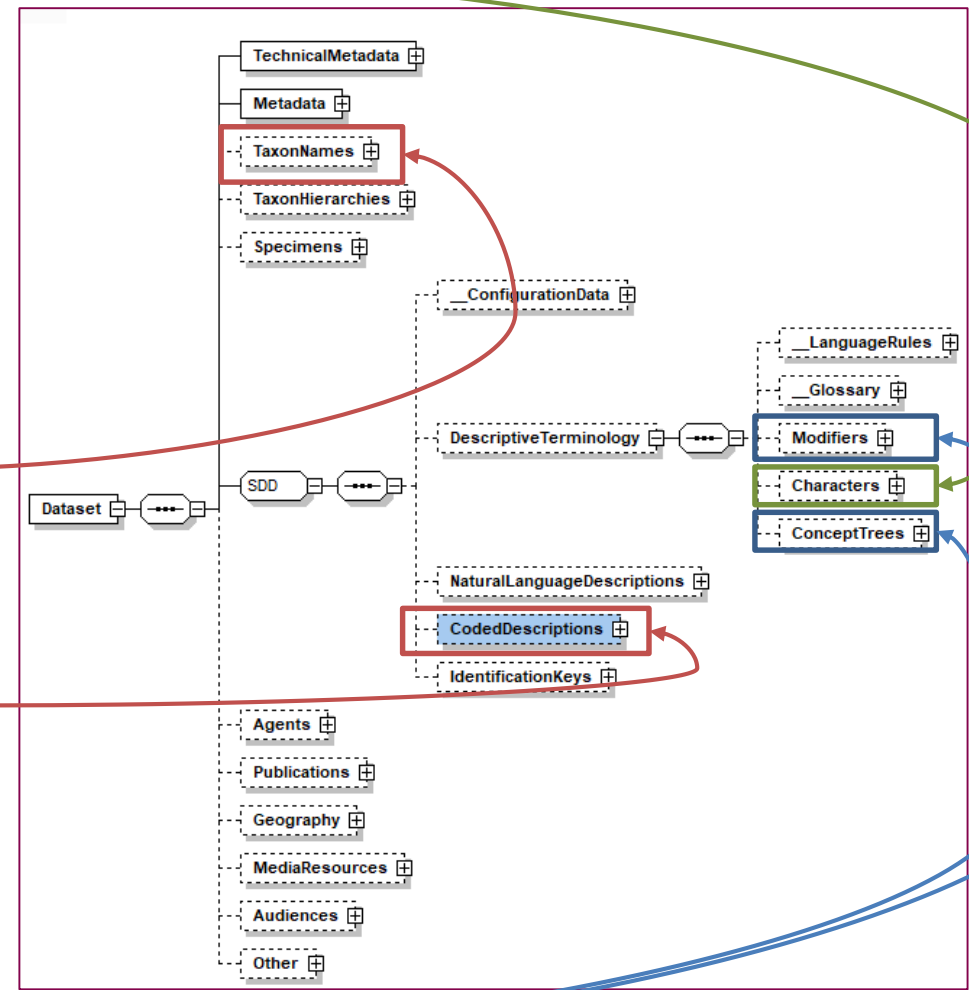


The unit is the TaxonRecord

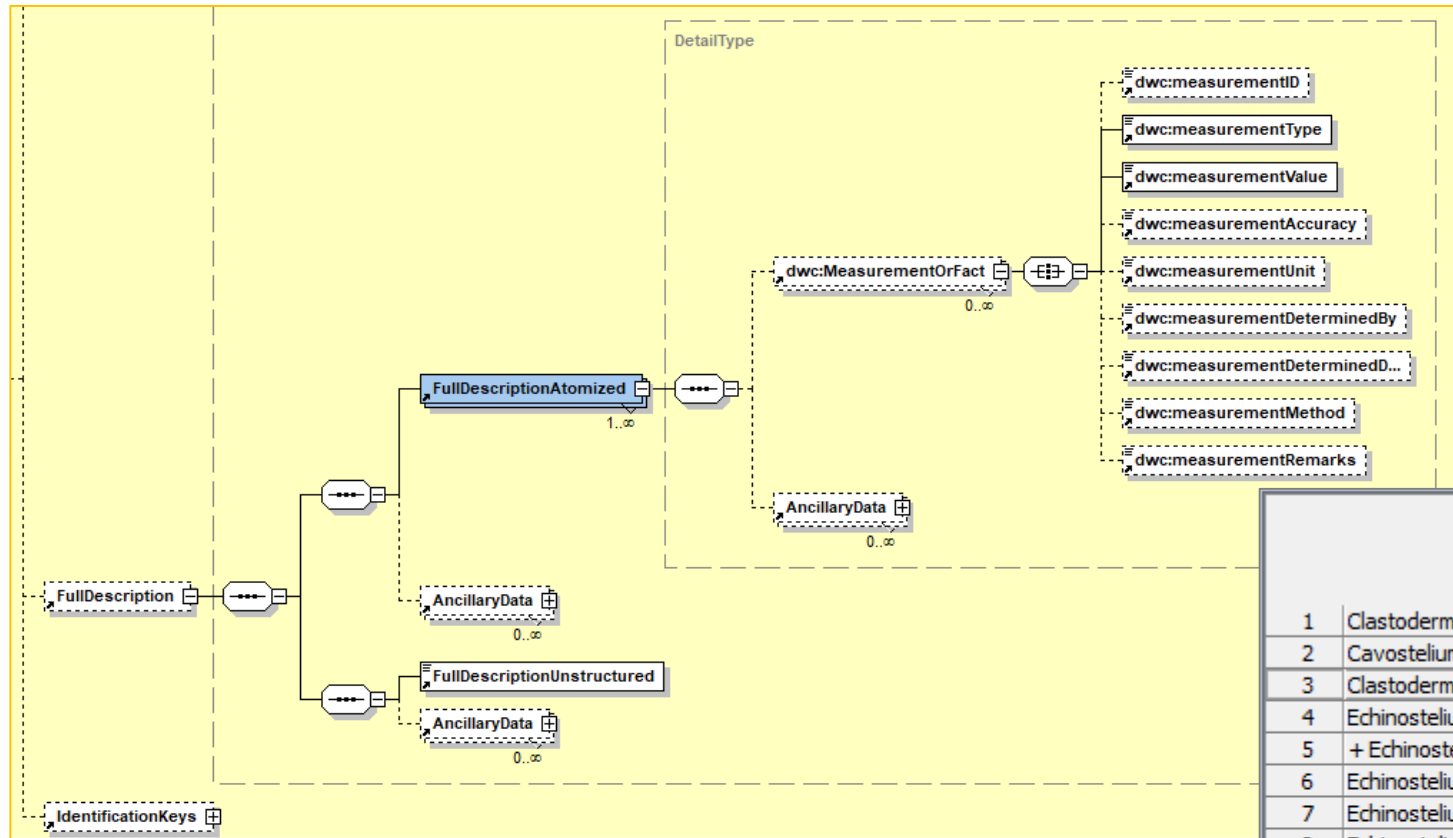
DELTA / SDD (Coded descriptions)

```
208 #30. capillitium <development>/
209 1. absent/
210 2. rudimentary <less than tree nodes>/
211 3. scanty <between three and six nodes>/
212 4. well developed <with over six nodes>/
213
214 #31. capillitium <polarity>/
215 1. isotropic; with the same features all over it/
216 2. showing some zonal differentiation <lower-upper/inne
217
218 #32. capillitium <type>/
219 1. made of lime tubules/
220 2. present as hollow tubules/
221 3. made of solid threads <apparently>/
222 4. of free elaters/
223 5. present as pseudocapillitium/
224 6. absent/
225
226 #33. capillitium <branching>/
227 1. very branched <over four nodes per branch>/
228 2. fairly branched <between four and three nodes>/
229 3. unbranched <one to two>/
230
231 *DEPENDENT CHARACTERS
232 4,4/5:18:19:20:21:22:23:24:25:26:27:28:29:30 8,3:12 16,5:17 28,1:29
233 32,2/4:15 32,5:15:33:35 36,1/2:37 38,6:36:37:39 48,2:49:50:51:52:53
234 54,3:55 58,1/2:60 63,1/2/3/4/5/6/7:73 63,9:61:62:66:67:73 64,1/2/4:65
235 66,1/2/3/4:68:69:70:71:72 73,2:74:75 78,1/2/3:80:81:82 88,4:89
236
237 * CHARACTER RELIABILITIES
238 3,6 5,0 6,3 7,2 8,7 9,6 10,7 11,6 13,5 14,7 15,7 20,1 21,1 22,1 23,4 24,7 25,3
239 26,2 27,2 28,1 29,7 30,2 31,7 32,7 35,2 36,6 37,8 38,0 40,7 48,7 49,6 53,2
240 69,0 71,6 88,0 89,0
```

```
13
14 # Clastoderma pachypus <Nann.-Bremek.>/
15 1,5 2,2-3 3,(213-)250-410(-460) 4,1 5,14 6,10 7,3 8,(76-)90-125 11,1
16 15,30-45(-80) 16,1<though some have a rather wide base> 17,2 18,2
17 the granular deposits are> 22,2 23,1 24,3-4<usually> 26,2 27,1<not
18 32,2 33,3 34,1/2 35,1 36,5 39,2 40,4<in a few poorly developed sp
19 the most>/5 44,4<which are few>/5 45,16-24<down to 8 in DWM 4356>
20 52,2 53,1-2 54,1 55,3 56,2 57,1 58,3 60,1
21
22 # Echinostelium apitectum <K.D. Whalley>/
23 1,5 2,2-3 3,150-250 4,1 5,2 6,2 7,3 8,40-65 11,5 12,1 13,(120-)150
24 20,1/2 22,2 23,1 24,4 26,3 27,1 28,3-6 30,1 41,1 42,4 45,0-4.5 46,
25 54,1 55,1 60,1 64,1<usually containing the columella, rarely over
```



PLiC (Coded descriptions)



		56	57	58	59
		spores in mass	spores by transmitted light <color>	spores <shape>	spores <diam. >
1	Clastoderma debaryanum...	8	8	5	8-10(-12)
2	Cavostelium apophysatu...				
3	Clastoderma pachypus N...		8	5	(9-)10-12(-13)
4	Echinostelium apitectum K...	3/5	1/4	1	10-12
5	+ Echinostelium apitectu...				7-8
6	Echinostelium arboreum H...	5	1/2/3	1	(5-)6-9
7	Echinostelium bisporum (...)	3	1	1	7-10
8	Echinostelium brooksii Wh...		1/4	1	10-14
9	Echinostelium coelocephal...	5	1	1/2	9-12
10	Echinostelium colliculosu...	4	1	1	9-13
11	Echinostelium corynophor...	4	1	1	(7-)9-15(-16)
12	Echinostelium cribrarioide...	5	1-3	1	8-10

B I U x² x₂ | ↶ ↷ ↻ ↵
 (9-)10-12(-13) R μm in diam

“this taxon has this state/value for this character”

Conclusions (1)

Coded descriptions on the basis of a list of characters, for which a set of states or a range of values are present in items (taxa, OTUs) being described.

Coded descriptions is a powerful tool in taxonomic research:

- They provide coherence (all taxonomic products -- descriptions, diagnoses, identification keys, etc.-- are originated from the same elements)
- They make explicit the information elements on which taxonomic decisions are taken (taxon circumscription, classifications)
- They allow for analyses

SDD (and DELTA) are very well suited to represent coded descriptions with all their nuances.

It is possible to store coded descriptions in PliC as "item-character-state/value". However, features built in DELTA (and SDD), that make description data handling easier, such as codification for:

- Implicit values
- Descriptions with multiple states in characters
- Character dependency

...are missing in PliC (though it is possible in convoluted ways, e.g. via the AncillaryData" element)

PliC is not the ideal specification to represent coded descriptions

Natural language descriptions

XML
Comment edited with XMLSpy v2013 (x64) (http://www.altova.com) by Francisco Pando (Real Jardín Botánico - CSIC)
Comment Sample XML file generated by XMLSpy v2013 (x64) (http://www.altova.com)

Dataset

- xsi:moNa... PinianCore_AbstractModel_v3.2.2.xsd
- xmns:dc http://purl.org/dc/terms/
- xmns:dwc http://rs.tdwg.org/dwc/terms/
- xmns:eol http://www.eol.org/transfer/content/0.3
- xmns:gi... http://www.gis.in/gis.in/SpeciesStatus
- xmns:tcs http://www.tdwg.org/schemas/tcs/1.01
- xmns:xsi http://www.w3.org/2001/XMLSchema-instance
- xmns:geo http://www.w3.org/2003/01/geo/wgs84_pos#

Metadata

TaxonRecord (2)

BaseElements	RecordMetadata	NomenclatureA...	TaxonomicDescription
1	BaseElements	RecordMetadata	NomenclatureA...

BriefDescription text

FullDescription

FullDescriptionAtomized

AncillaryData

FullDescriptionU... (1) Sporocarps(2) Scattered to gregarious, (3) (213-)250-410(-460) µm high, (4) stalked, (6) when fresh, rosaceous-brown, (7) when mature, moderate brown (58. m. Br). (8) Sporotheca globose, (9) (76-)90-125 µm in diam. (14) Spores per sporcarp: many. (16) Hypothallus inconspicuous. (18) Stalk present, (19) 115-300(-330) µm long, (20) 4-6 µm diam at the apex, (21) 30-45(-80) µm diam at the base, (22) arising from a non expanded base (though some have a rather wide base), (23) slightly flexuous, (24) tapering to the apex, (25) translucent, (26) by transmitted light brilliant yellow (83. brill. Y) or moderate yellow (87. m. Y) (in the lower section, where the granular deposits are), (28) filled, (29) with granular material, (30) up to 1/2 to 2/3, or more (usually), (31) hollow apophysis lacking. (32) Peridium persistent as fragments which cling to the tips of the capillium, (33) when mature as a collar (not always visible), (34) 15-20 µm in diam. (36) Capillium scanty to well developed, (37) showing some zonal differentiation, (39) scattered branched, (40) dichotomous or mostly splitting dichotomously, (41) peripheral capillium forming a complete net, (42) peripheral capillium with few or no free ends, (46) internal capillium sinuate, (47) internal capillium mostly anastomosed (in a few poorly developed sporocarps) or completely

PliC

XML
Comment edited with XMLSpy v2013 (x64) (http://www.altova.com) by Francisco Pando (Real Jardín Botánico - CSIC)
Comment Sample XML file generated by XMLSpy v2013 (x64) (http://www.altova.com)

Dataset

- xsi:schemaLocation http://www.tdwg.org/2005/UBIF file:///C:/Data/Projects/GBIF/GBIF%20Others/TDWG/Species%20IG/SDD%20std/SDD.xsd
- xmns http://www.tdwg.org/2005/UBIF
- xmns:n2 http://www.altova.com/samplexml/other-namespace
- xmns:xsi http://www.w3.org/2001/XMLSchema-instance

TechnicalMetadata created=2018-06-26T10:24:02 expires=2022-06-26T10:24:02

Metadata

TaxonNames

TaxonHierarchies

Specimens

ConfigurationData

DescriptiveTerminology

NaturalLanguageDescriptions (2)

id	d...	Repr...	L...	Versi...	Cus...	Defi...	Revi...	Header	NaturalLanguageData
1	CP	String	Repr...	L... a	Versi...	Cus...	Defi...	Revi...	Header

NaturalLanguageData

parsed false

newpage a

Text

(1) Sporocarps(2) Scattered to gregarious, (3) (213-)250-410(-460) µm high, (4) stalked, (6) when fresh, rosaceous-brown, (7) when mature, moderate brown (58. m. Br). (8) Sporotheca globose, (9) (76-)90-125 µm in diam. (14) Spores per sporcarp: many. (16) Hypothallus inconspicuous. (18) Stalk present, (19) 115-300(-330) µm long, (20) 4-6 µm diam at the apex, (21) 30-45(-80) µm diam at the base, (22) arising from a non expanded base (though some have a rather wide base), (23) slightly flexuous, (24) tapering to the apex, (25) translucent, (26) by transmitted light brilliant yellow (83. brill. Y) or moderate yellow (87. m. Y) (in the lower section, where the granular deposits are), (28) filled, (29) with granular material, (30) up to 1/2 to 2/3, or more (usually), (31) hollow apophysis lacking. (32) Peridium persistent as fragments which cling to the tips of the capillium, (33) when

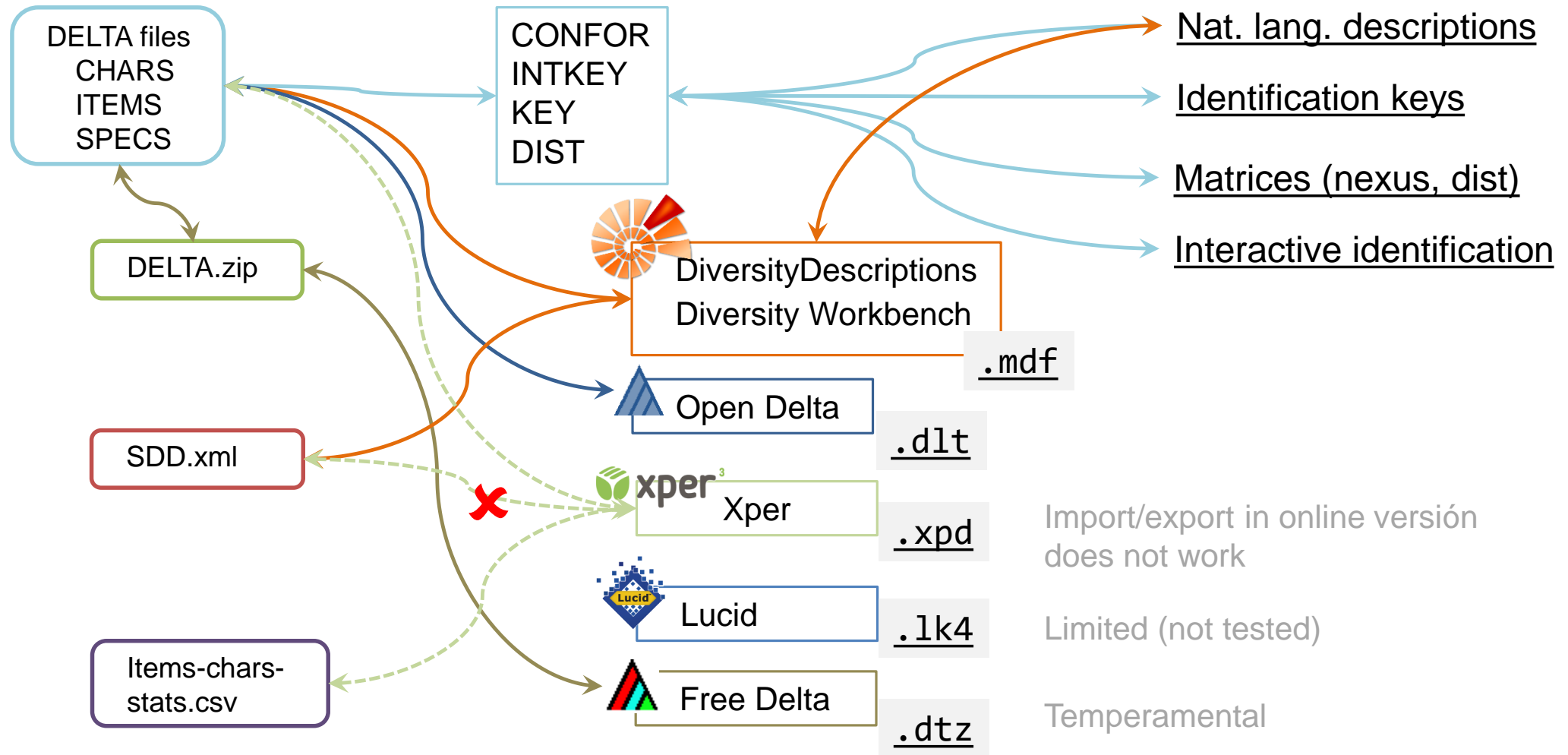
SDD

(xml)

Conclusions (2)

- Natural language descriptions (NLD) and identification keys can be easily mapped and transferred between SDD and PliC.
- However, using SDD for storing NLD is complicated (no tools available for it, more of this later) and it is a waste of the capabilities of SDD.

Tools DELTA / SDD



Tools PlIC

Export to New file Metadata Species Info Management in Plinian Core Language ⏻ ⏻

5f81e31	Echinostelium
42935ca	Echinostelium ladoi
da10921	Clastoderma debaryanu

Basic data Identifiers

IdRecord: 5f81e311-b; kingdom: Protozoa phylum: Mycetozoa class: Myxomycetes order: Echinosteliales
family: Echinosteliaceae genus: Echinostelium SpecificEpithet: n/a TaxonRank: Genero
InfraspecificEpithet: n/a Scientific name: Echinostelium language: English
Author year: de Bary 1873 Name published in: Vers. Syst. Mycetozoen: 7 Version*: 31/05/2018 12:32:27

UsesManagementAndConservation Invasiveness of the species Collaborators References MeasurementOrFact AncillaryData

NomenAndTaxoDesc NaturalHistory HabitatAndDistribution DemographyAndThreat

Scientific Description: Sporocarps 20-183-550 µm high, stalked, when fresh, hyaline or white (263. White), or yellowish white (92. y White), bright yellow (83. brill. Y), or grayish yellowish pink (32. gy. y Pink), or when mature, white, yellowish or pinkish; sporotheca globose, 30-48.21-120 µm in diam. Stalk tapering to the apex, translucent; under transmitted light, hyaline, hollow, but usually filled with granular material in its lower part. Peridium evanescent but with a collar persistent when mature, rarely with

Identification Keys: 1.Capillitium made of solid threads 2
1.Capillitium absent 8
2.Spore-like body present 1. E. apitectum
2.Spore-like body absent 3

Synonyms Estructurado No estructurado

SynonymName*
SynonymStatus

Record: 1 of 1 No Filter Search

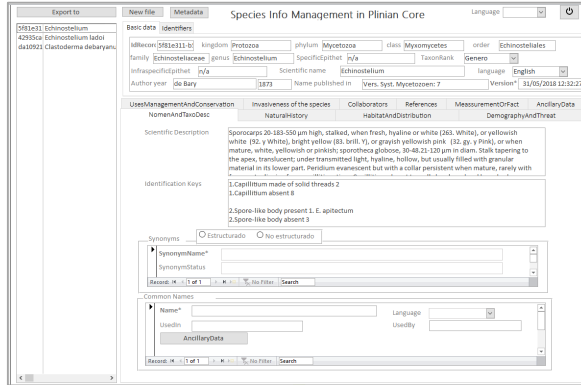
Common Names

Name* Language
UsedIn UsedBy

AncillaryData

Record: 1 of 1 No Filter Search

PlIC & IPT



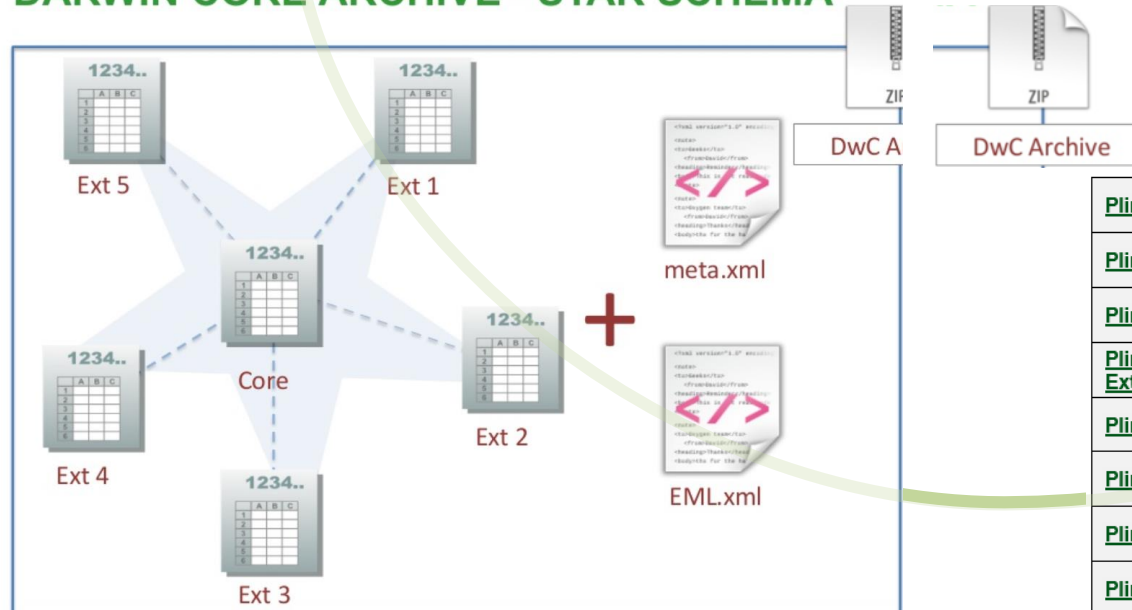
GBIF INTEGRATED PUBLISHING TOOLKIT (IPT)
free and open access to biodiversity data

Home **Manage Resources** About

Resources you have rights to manage

Name	Organisation	Type	Subtype	Records	Last modified
dwc-plc-test	Not registered	Checklist	--	0	2018-07-13
Flora Mycologica Iberica Project database	Real Jardín Botánico (CSIC)	Occurrence	Observation	59,235	2018-04-10
Lista de táxones de la flora vascular española	GBIF-Spain	Checklist	Inventory Regional	10,493	2018-04-10
A Distribution and Taxonomic Reference Dataset of Geranium (Geraniaceae) in the New World	Real Jardín Botánico (CSIC)	Occurrence	Specimen	8,937	2017-06-05

DARWIN CORE ARCHIVE - STAR SCHEMA



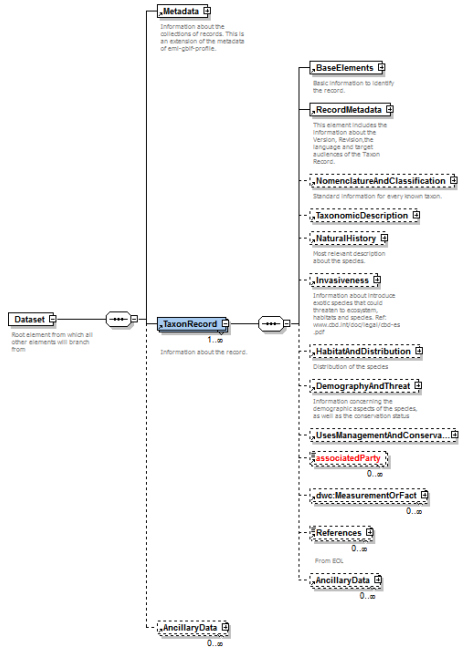
- [Plinian Distribution Extension](#)
- [Plinian Endemicity Extension](#)
- [Plinian Legislation Extension](#)
- [Plinian ManagementAndConservation Extension](#)
- [Plinian Core Simple Extension](#)
- [Plinian Synonym Extension](#)
- [Plinian ThreatStatus Extension](#)
- [Plinian Uses Extension](#)



PlIC IPT extensions need to be moved from “in development” to “stable”

<https://tools.gbif.org/dwca-validator/extensions.do>

PLIC → RDF



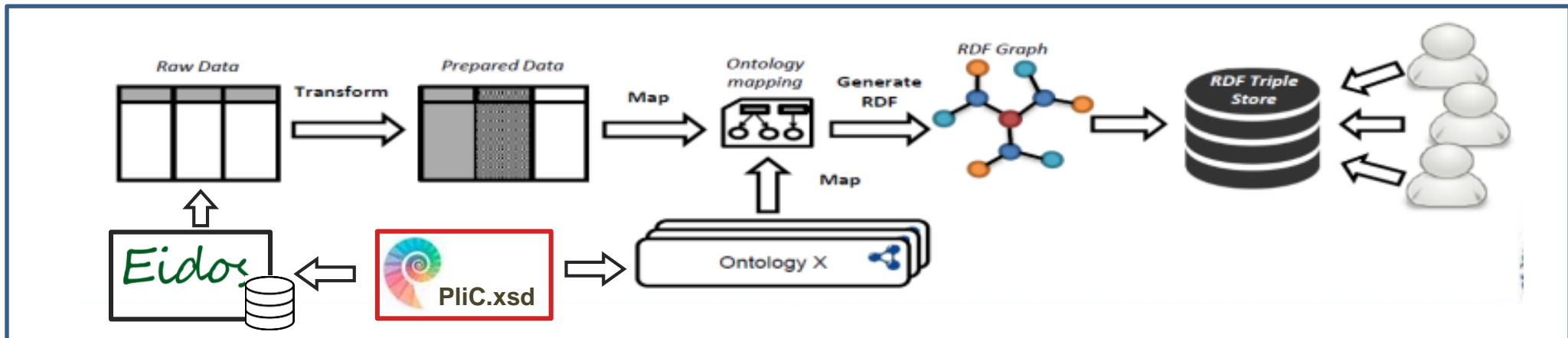
<https://github.com/tdwg/PlinianCore>



http://www.mapama.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/Eidos_acceso.aspx



Cross harmonization & exploitation of nature datasets
<https://crossnatureblog.wordpress.com/>



Combining data from three RDF repositories

Información taxonómica	
Author	Canis lupus Linnaeus, 1758
Classification	Animalia Chordata Mammalia Carnivora Canidae Canis lupus
Kingdom	Animalia
Phylum	Chordata
Class	Mammalia
Order	Carnivora
Family	Canidae
Genus	Canis
SubGenus	
Specific Epithet	lupus
InfraSpecificEpithet	
Nombres científicos	
Name	Status
Canis lupus Linnaeus, 1758	Aceptado/Válido
Nombres comunes	
Language	Name
Euskera	Otsoa
Inglés	Gray Wolf, Timber Wolf, Arctic Wolf, Grey Wolf, Mexican Wolf, Plains Wolf, Common Wolf, Tundra Wolf, Wolf
Francés	Loup, Loup Gris, Loup Vulgaire
Español, castellano	Lobo

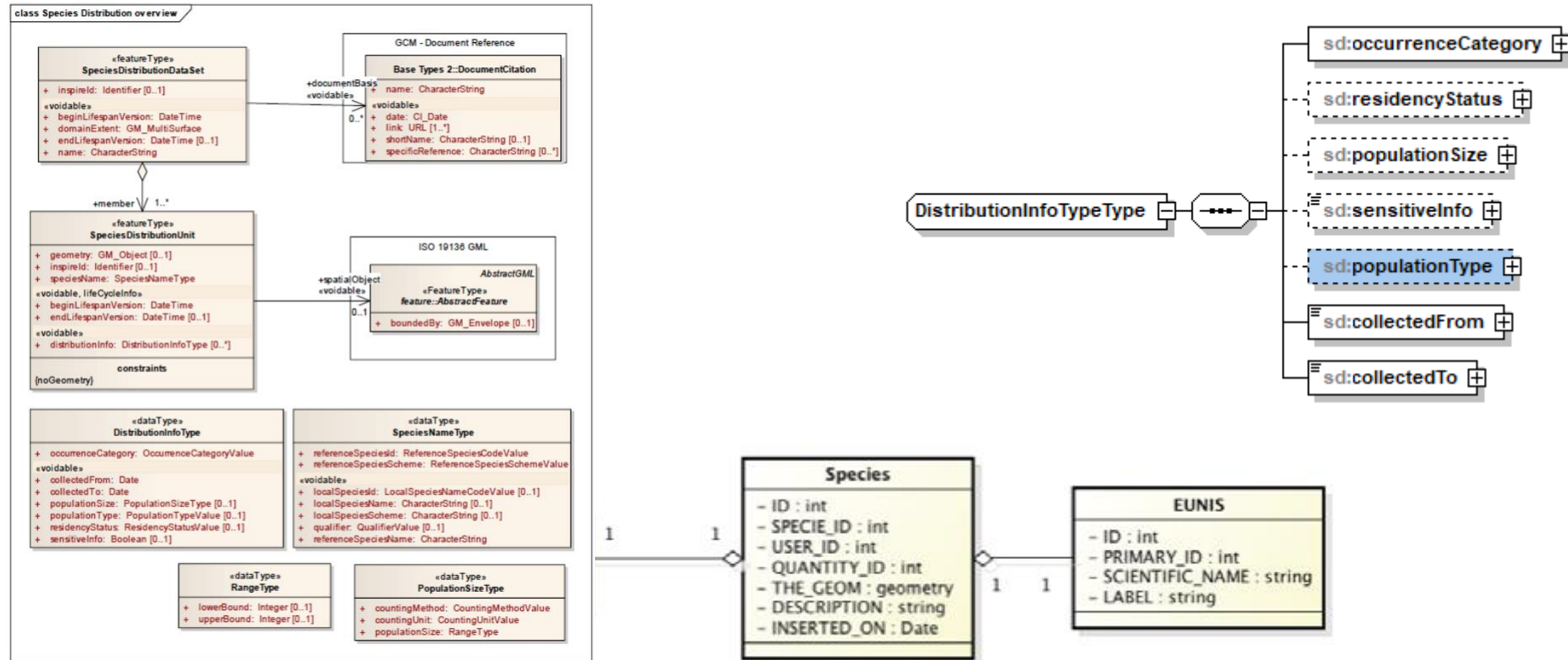
Eidos	Eunis	Uniprot
speciesCode	1367	
label	Canis lupus Linnaeus, 1758	
scientificNameAuthorship	Linnaeus, 1758	
genus	Canis	
vernacularName	Farkas, Hunt, Kurt, Lobo, Loup, Lup, Lupo, Lupa, Susi, Ujku, Ufu, Varg, Vilkas, Vilks, Vik, Vik dravý, Volk, Vuk, Wilk, Wolf, Wölfe, Ulfur	
scientificName	Canis lupus	
binomialName	Canis lupus	
sameSpecies	<ul style="list-style-type: none"> http://dbpedia.org/resource/Gray_wolf http://lod.taxonconcept.org/ses/tknlf#Species 	
sameSpeciesITIS	180596	
sameSpeciesNCBI	9612	
sameSpeciesRedlist	3746	
sameSynonym	<ul style="list-style-type: none"> http://dbpedia.org/resource/Gray_wolf http://lod.taxonconcept.org/ses/tknlf#Species http://rdfdata.eionet.europa.eu/itis/taxon/180596 	
sameSynonymN2000	1352	
	<ul style="list-style-type: none"> http://bd.eionet.europa.eu/article17/reports2012/species/summary/? 	

Eidos	Eunis	Uniprot
Canis lupus familiaris		Canine parvovirus type 2
Canis lupus familiaris		Canine distemper virus (strain Onderstepoort)
Canis lupus familiaris		Torque teno canis virus (isolate CF-TTV10)
Canis lupus familiaris		African horse sickness virus 2
Canis lupus familiaris		African horse sickness virus 5
Canis lupus familiaris		African horse sickness virus 8
Canis lupus familiaris		African horse sickness virus 9
Canis lupus familiaris		African horse sickness virus 3
Canis lupus familiaris		Mokola virus
Canis lupus familiaris		Cadivovirus A (isolate Dog/Hong Kong/209/2008)
Canis lupus familiaris		Canine oral papillomavirus (strain Y62)
Canis lupus familiaris		Parainfluenza virus 5 (strain W3)
Canis lupus familiaris		African horse sickness virus 1
Canis lupus familiaris		African horse sickness virus 4
Canis lupus familiaris		Canine adenovirus serotype 1 (strain ...)

<http://crossnature.eu/visor>

PLiC → INSPIRE

- MAPAMA gateway EIDOS > INSPIRE "species distribution" GML



EIDOS > transformation needed between grid identifiers and WKT polygons

Conclusions (and 3)

- More than an overlap or a redundancy between SDD-DELTA and PliC there is a contact point (coded descriptions).
- SDD and PliC have different strengths, and are intended for different audiences; they are complementary.
- There is no urgent need for a XSLT for transferring descriptions between these standards as there are less technologically demanding options.
- Coded descriptions (and subsequently SDD and DELTA) is information "from taxonomists, for taxonomist".
- PliC is focused on visualization, publication and post-taxonomy integration and interoperability of taxonomic information. It is "Taxonomic information (and more), for non-taxonomist".

Francisco Pando

Real Jardín Botánico - CSIC
Claudio Moyano 1, 28014 Madrid, Spain
pando@rjb.csic.es



<http://creativecommons.org/licenses/by-sa/3.0/es/>

TDWG Species Information Interest:

<https://github.com/tdwg/species-information>

Plinian Core Task Group:

<https://github.com/tdwg/PlinianCore>

Work partially supported by: EU Horizon 2020 framework programme
project DEEP-Hybrid-Datacloud (Grant Agreement number 777435)

