

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

Francisco Pando
Real Jardín Botánico-CSIC. Spain



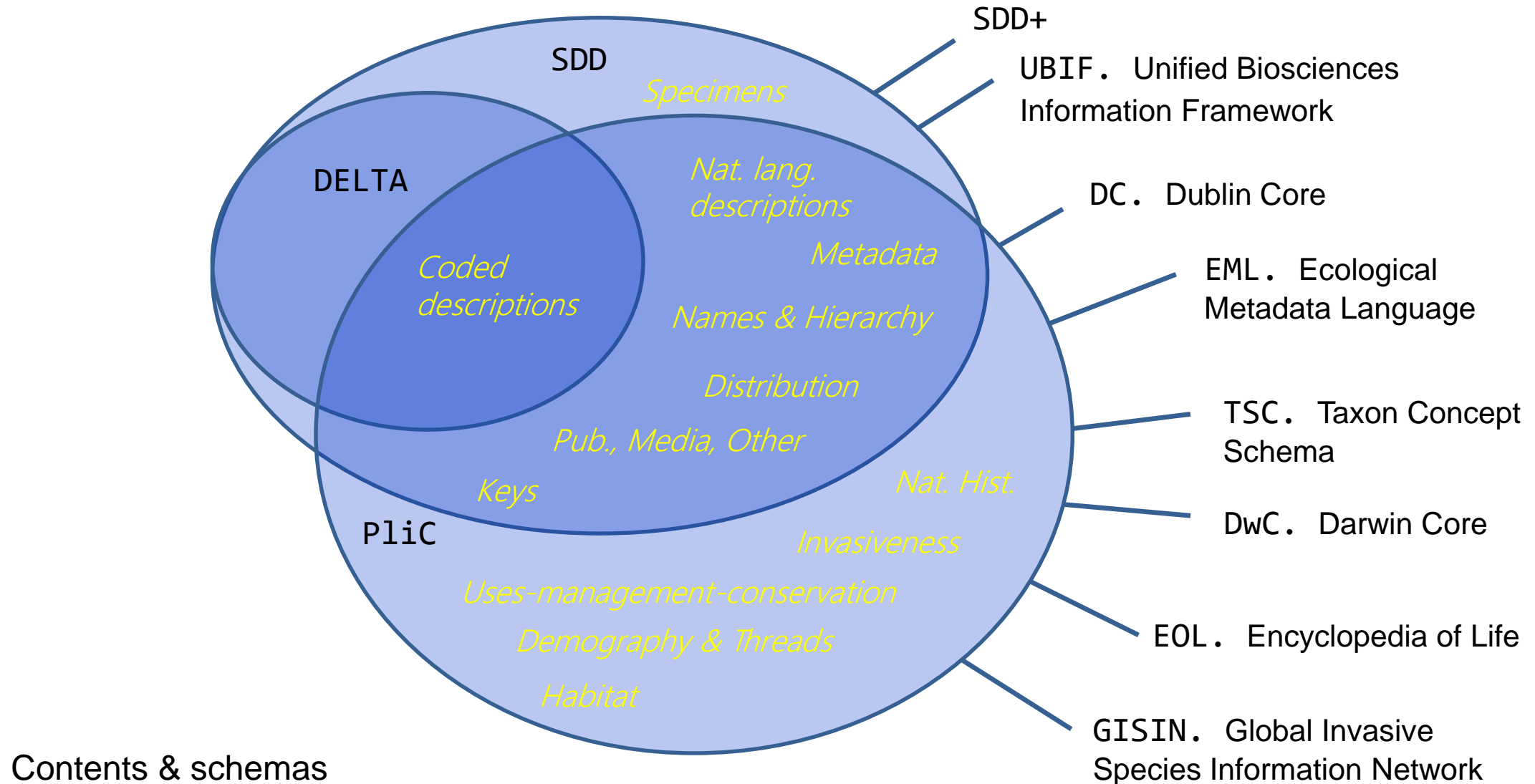
PLINIAN CORE



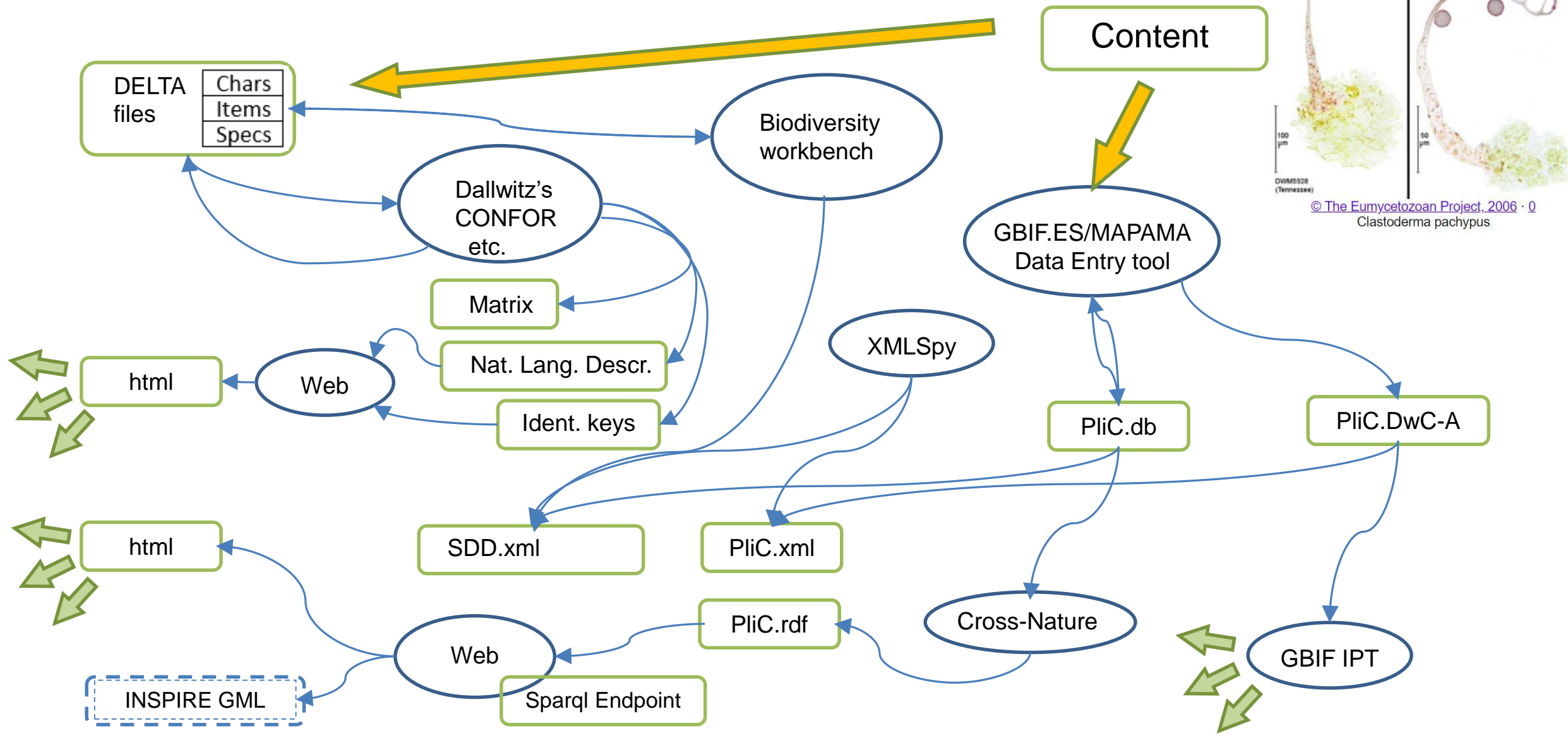
Summary

- Species-level standards Universe
- Metodology
- 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

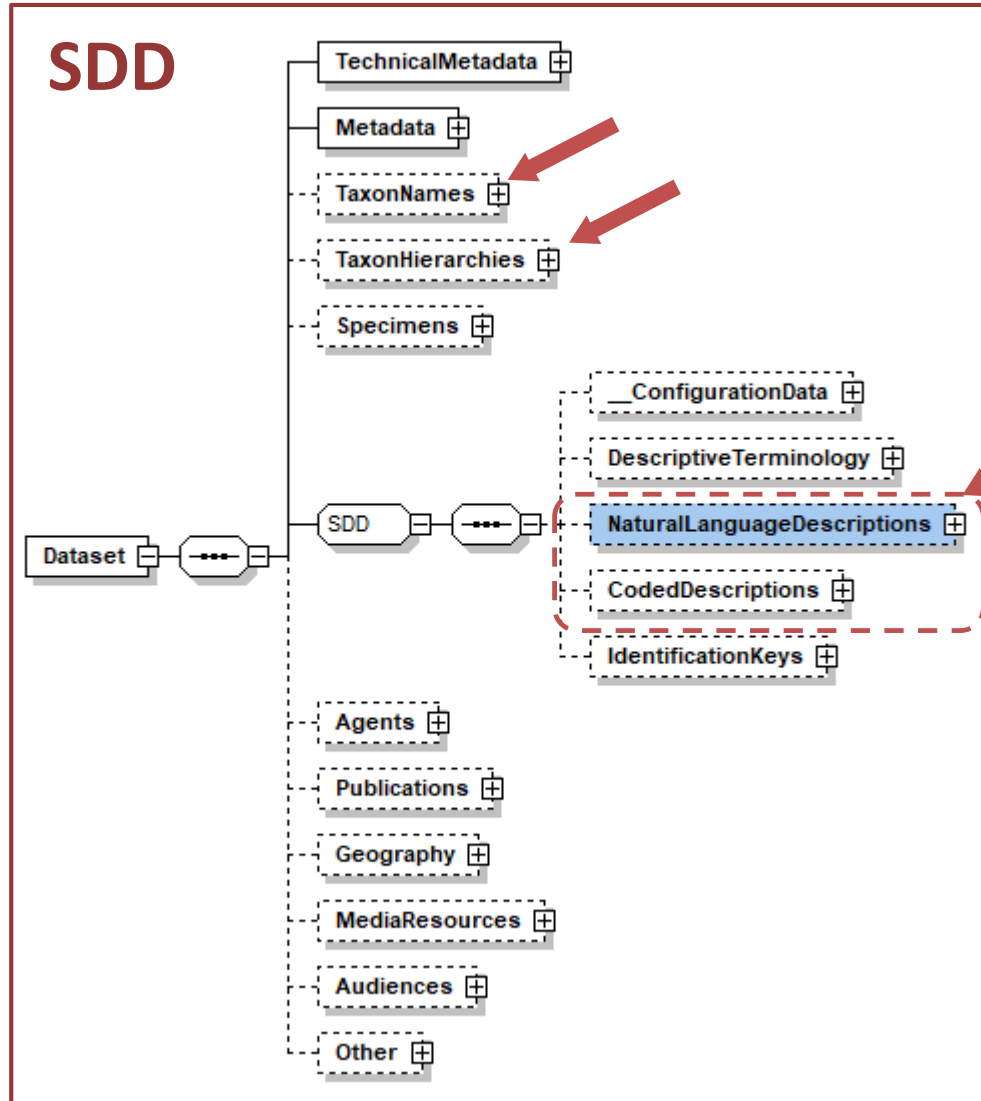
Species level standards Universe



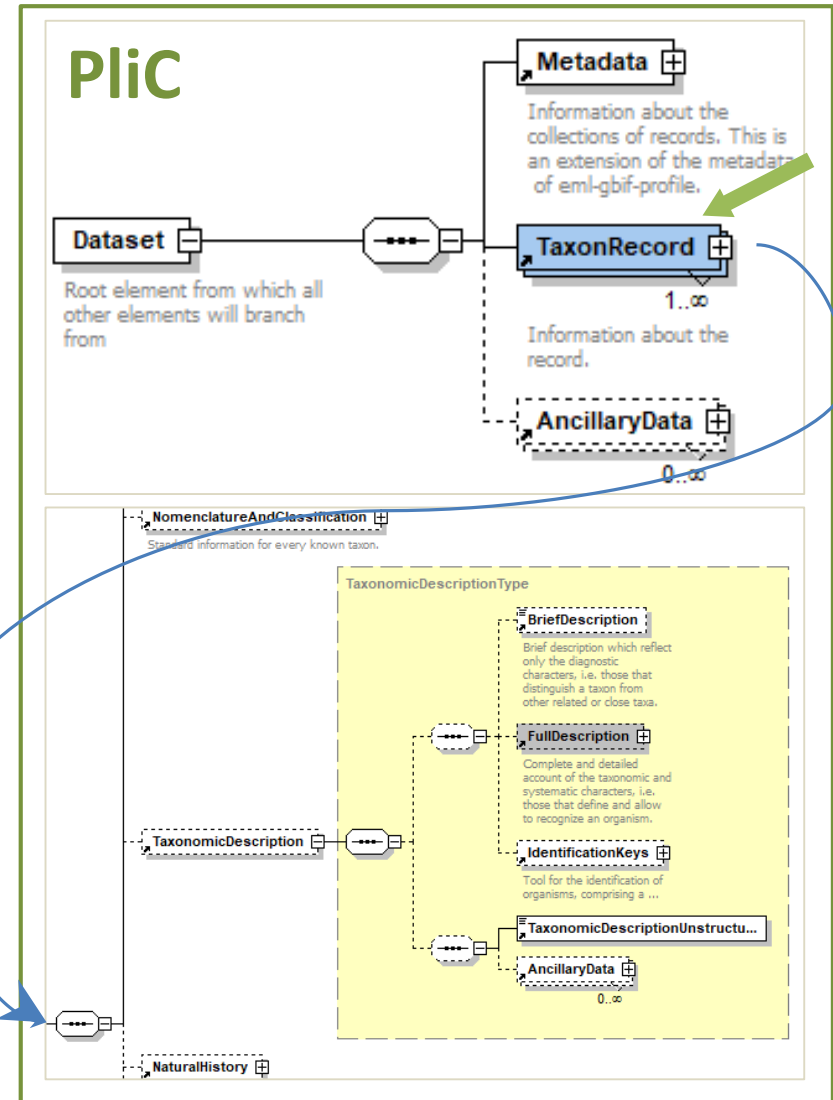
Methodology



Conceptual approaches

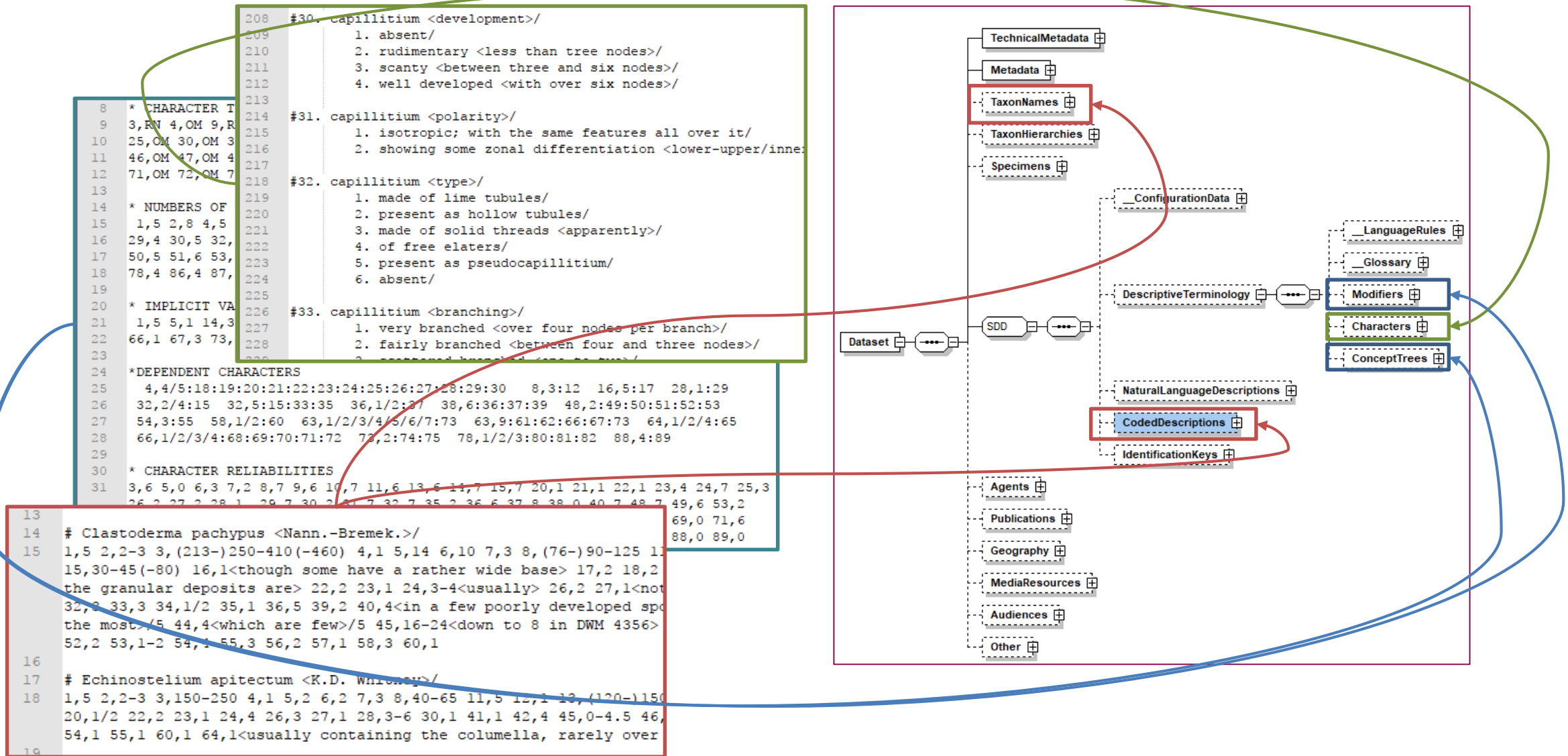


The unit is the Dataset

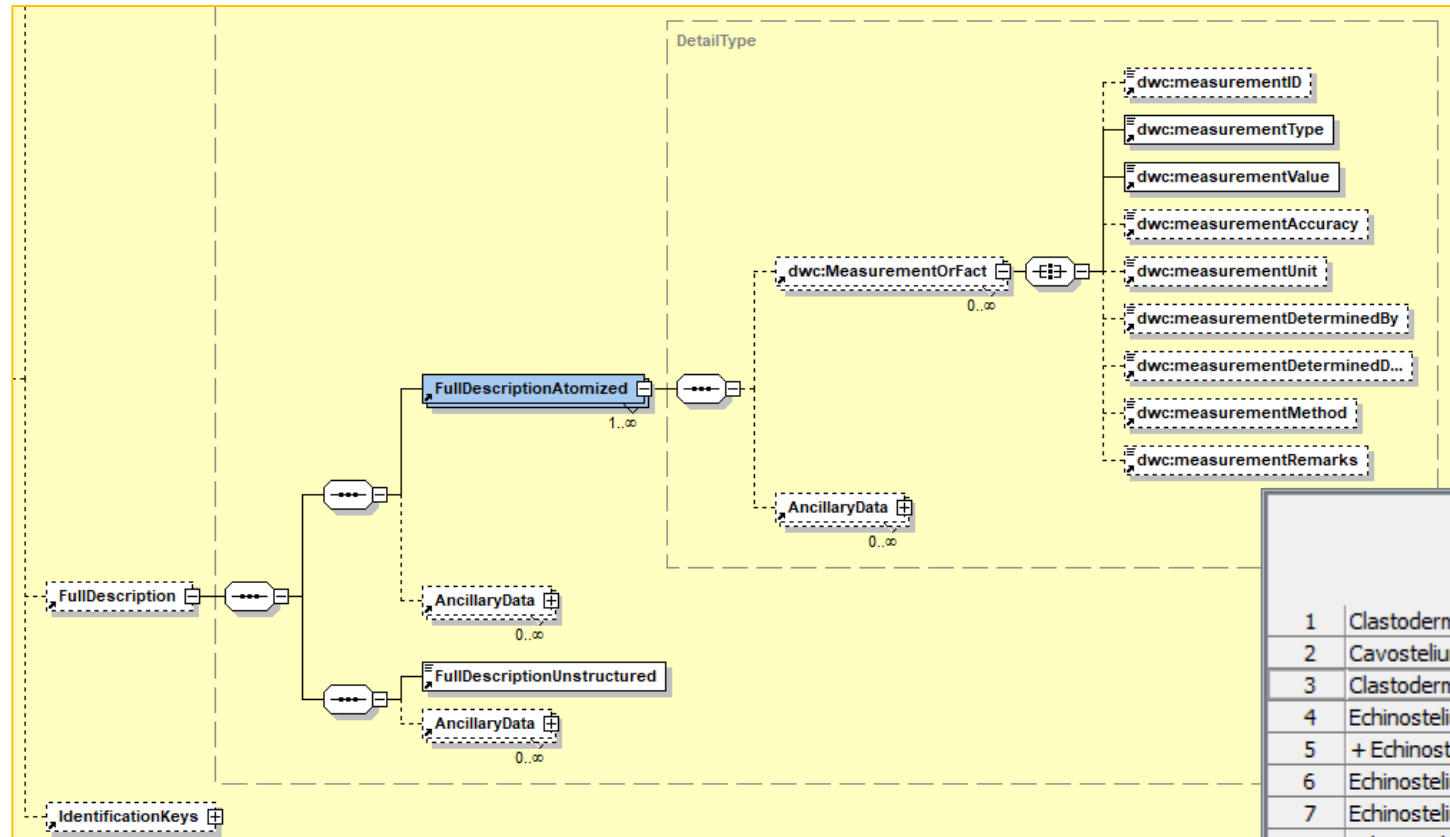


The unit is the TaxonRecord

DELTA / SDD (Coded descriptions)



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)

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:

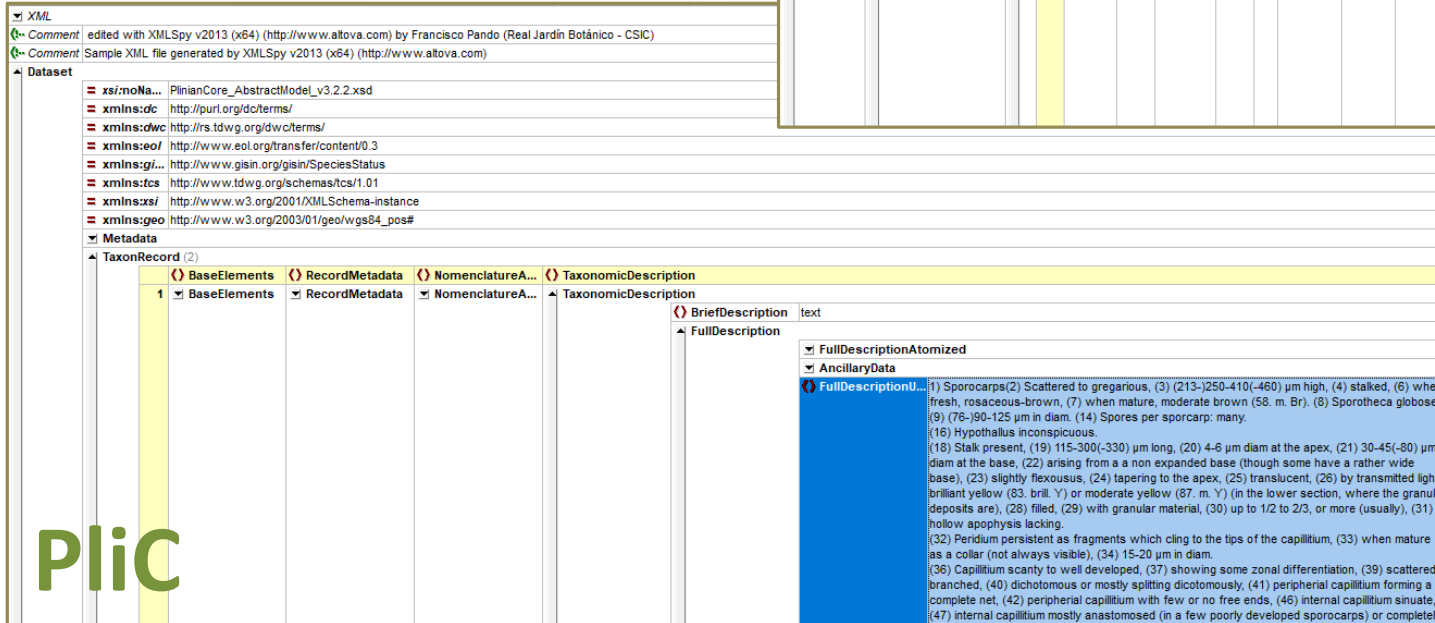
- Implicit values
- Descriptions with multiple states in characters
- Character dependency

...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

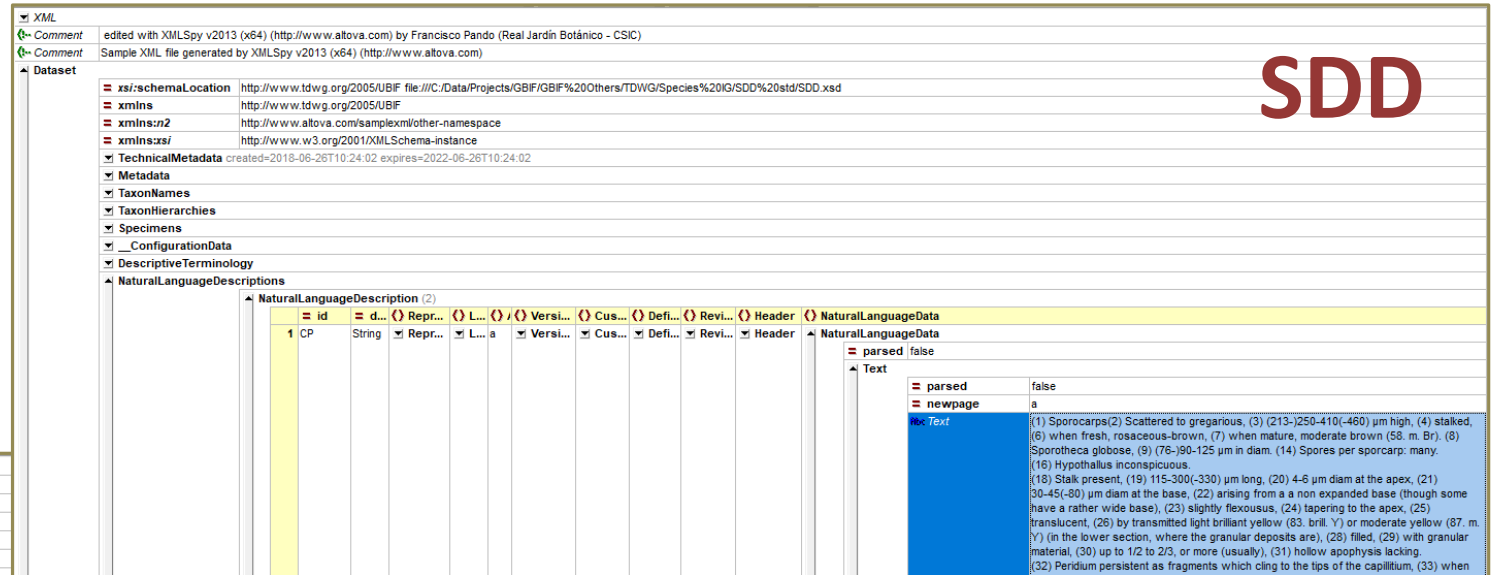
Natural language descriptions

PliC



The screenshot shows the XMLSpy v2013 (x64) interface. The main window displays a dataset with a single record. The record is expanded to show its structure, which includes a base element, record metadata, nomenclature, and a taxonomic description. The taxonomic description is a natural language text describing the morphology of a fungus, including details about sporocarps, stalks, and peridium.

SDD



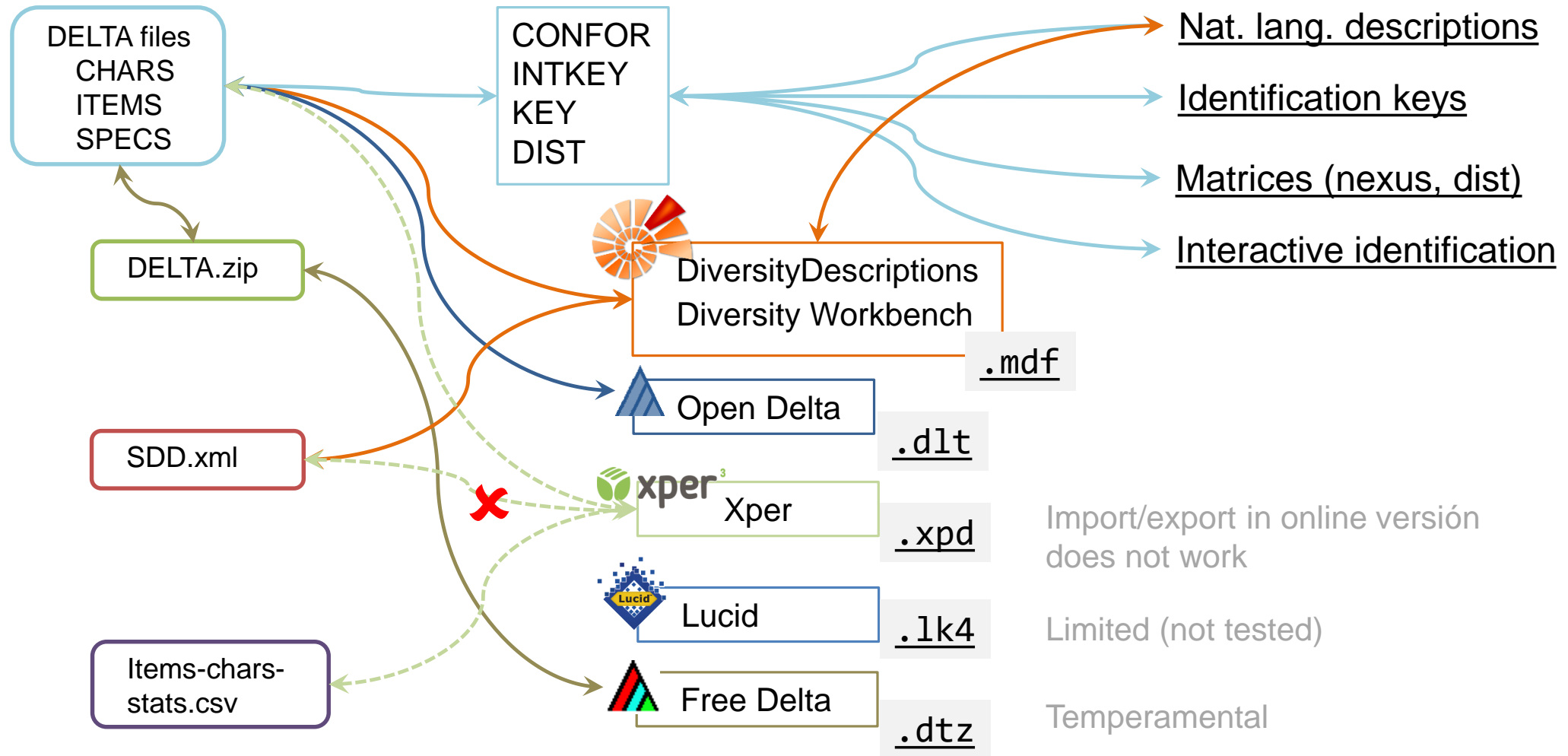
The screenshot shows the XMLSpy v2013 (x64) interface. The main window displays a dataset with a single record. The record is expanded to show its structure, which includes a base element, record metadata, nomenclature, and a natural language description. The natural language description is a text describing the morphology of a fungus, including details about sporocarps, stalks, and peridium.

(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

5f81e31 Echinostelium

42935ca Echinostelium ladoi

da10921 Clastoderma debaryanu

New file

Metadata

Species Info Management in Plinian Core

Language

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

MeassurementOrFact

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

SynonymName*

SynonymStatus

Record: 1 of 1

No Filter

Search

Common Names

Name*

UsedIn

Language

UsedBy

AncillaryData

Record: 1 of 1

No Filter

Search

PliC & IPT

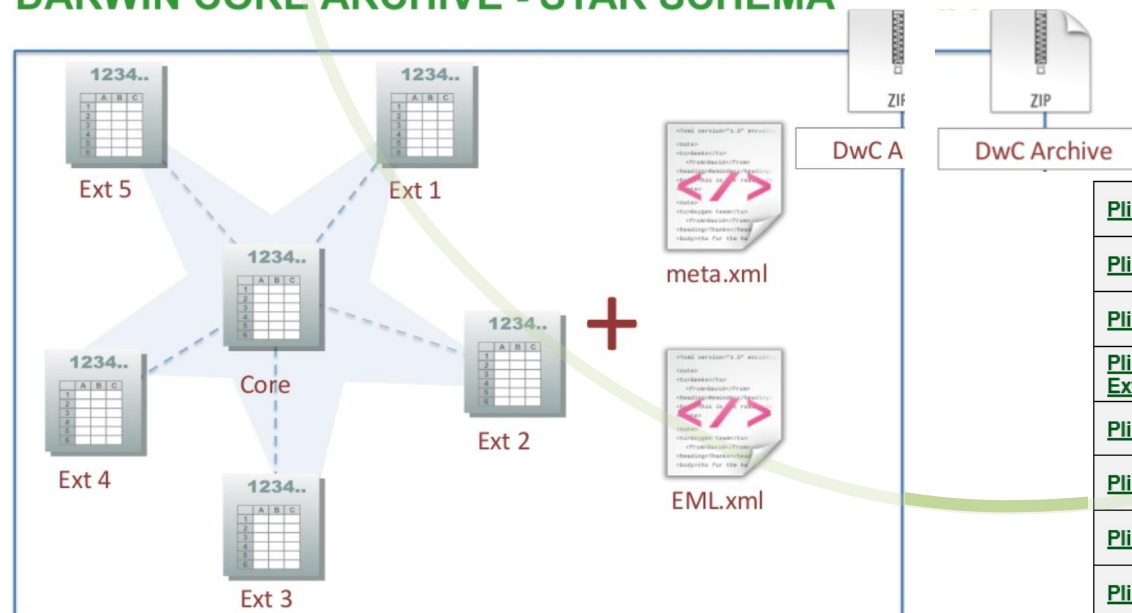
GBIF INTEGRATED PUBLISHING TOOLKIT (IPT)
free and open access to biodiversity data

Home Manage Resources About

Resources you have rights to manage

Name	Organisation	Type	Subtype	Records	Last modified
dwc-pliC-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



Darwin Core Archive Validator

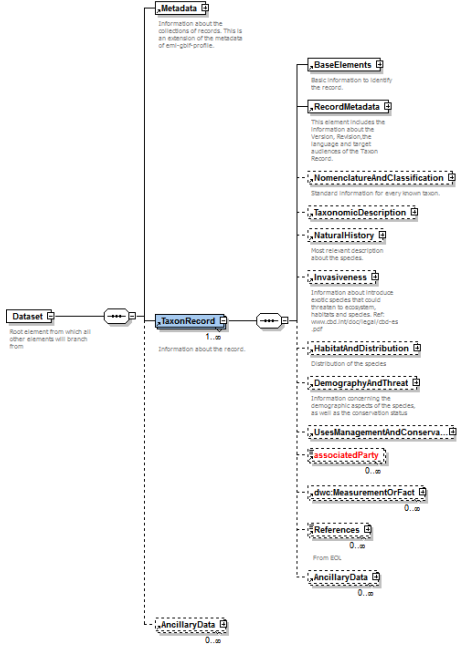
Registered Extensions

Under Development

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

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

PLIC \rightarrow RDF



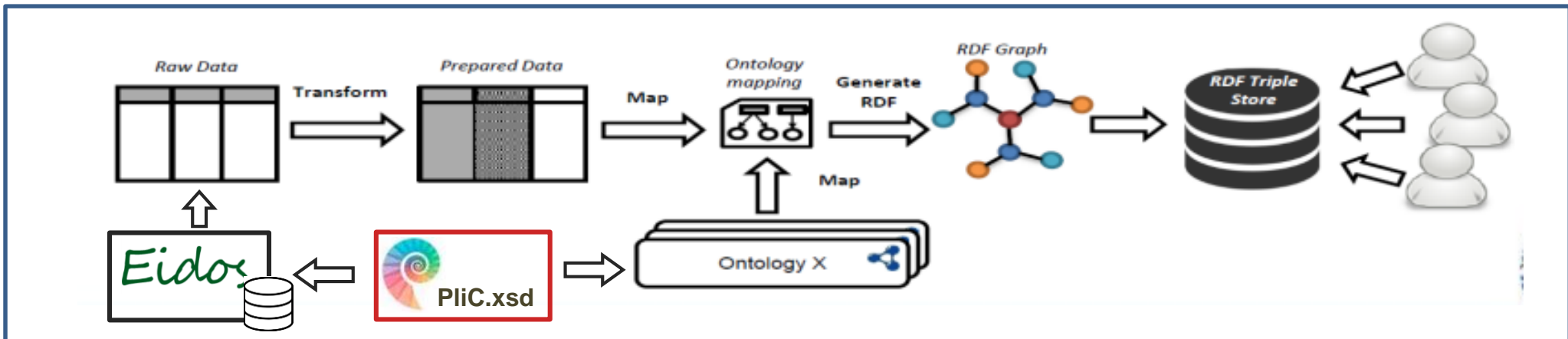
<https://github.com/tdwg/PlinianCore>



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

crossnature.eu/visor/#

Cross-Nature
CROSS HARMONIZATION & EXPLOITATION OF NATURE DATASETS

GENERAL


- Home
- Eidos
- Layouts
- + Add species
- Malla 10x10 UTM
- ☒ Canis lupus Linnaeus, 1758

MORE

- Endpoints


Map

1 results (Canis lupus Linnaeus, 1758)



Search radius (1)

Canis lupus Linnaeus, 1758



Eidos Eunis Uniprot

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/Valido

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

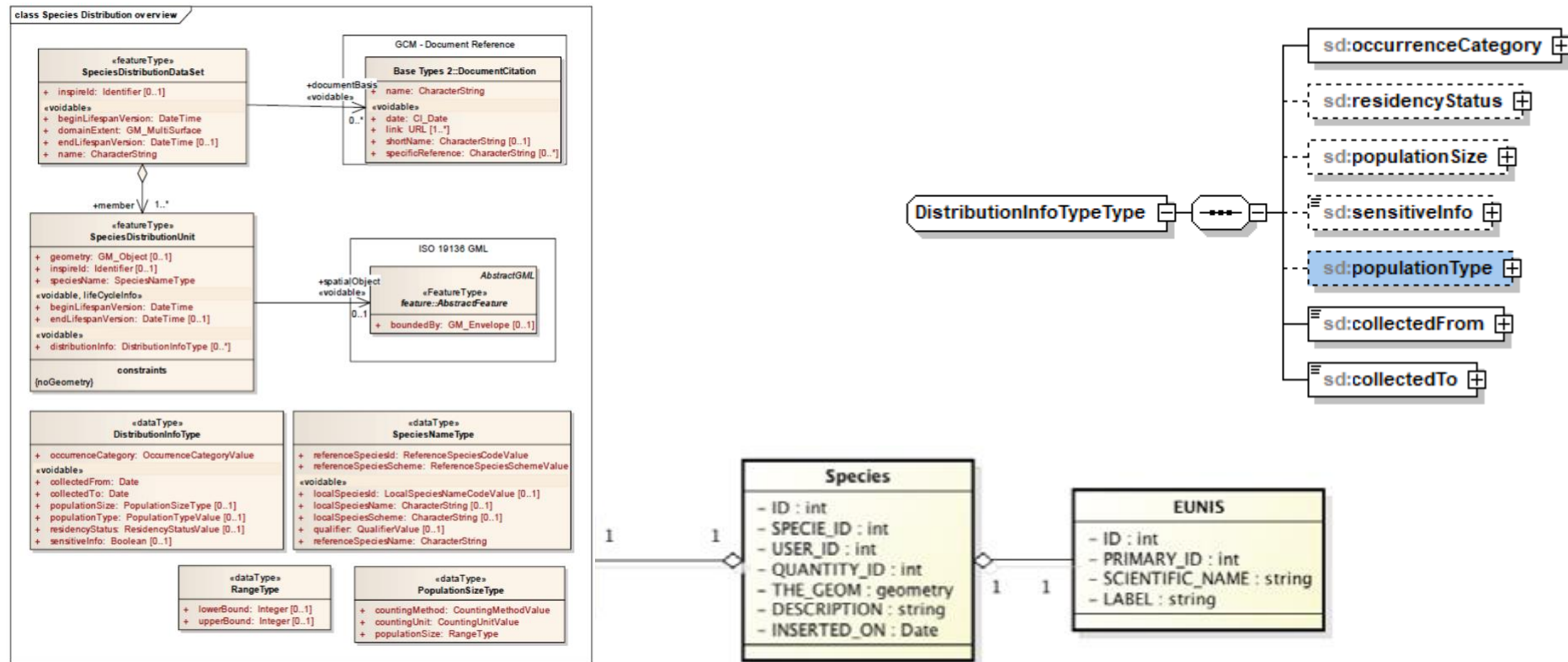
Eidos	Eunis	Uniprot
speciesCode	1367	
label	Canis lupus Linnaeus, 1758	
scientificNameAuthorship	Linnaeus, 1758	
genus	Canis	
vernacularName	Farkas, Hunt, Kurt, Lobo, Loup, Lup, Lupo, Lupa, Susi, Ujku, Uly, Varg, Vilkas, Vilks, Vlk, Vlk dravý, Volk, Vuk, Wilk, Wolf, Wölfel, Ülür	
scientificName	Canis lupus	
binomialName	Canis lupus	
sameSpecies	<ul style="list-style-type: none">http://dbpedia.org/resource/Gray_wolfhttp://lod.taxonconcept.org/ses/tknll#Species	
sameSpeciesITIS	180596	
sameSpeciesNCBI	9612	
sameSpeciesRedlist	3746	
sameSynonym	<ul style="list-style-type: none">http://dbpedia.org/resource/Gray_wolfhttp://lod.taxonconcept.org/ses/tknll#Specieshttp://rdfdata.eionet.europa.eu/itis/taxon/180596	
sameSynonymN2000	1352	
	<ul style="list-style-type: none">http://bd.eionet.europa.eu/article17/reports2012/species/summary/7	

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 3 (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)

