From liberating imprisoned data in publications to publishing linked open data

Donat Agosti
Plazi
agosti@plazi.org; @myrmoteras
ORCID: 0000-0001-9286-1200

Swiss Research Data Day 2018, 12 June 2018, Zürich

DOI: 10.5281/zenodo.1287269

How many species are on planet Earth?

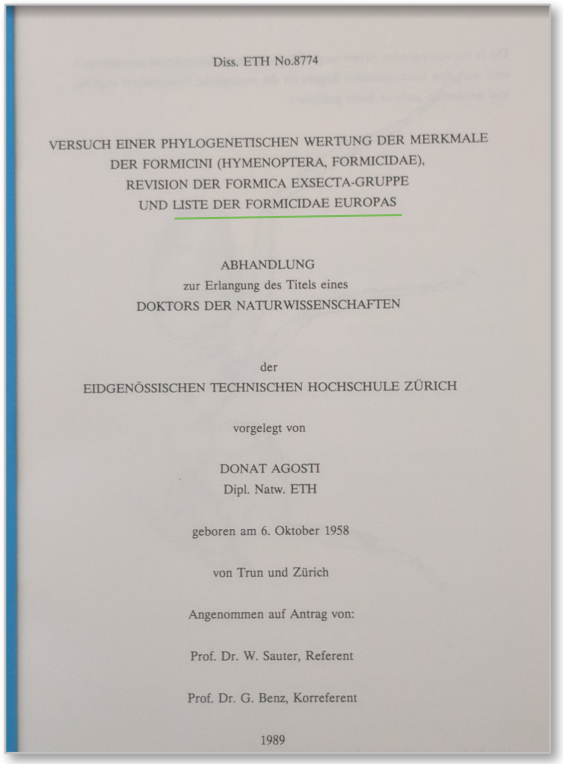
1989 How many ants in Europe?
1992 Rio Earth Summit: How many species on planet Earth?
2002 Online catalogue of the ants of world
2018 Half Earth: how many species?

A scientific challenge as well as a crucial question for a sustainable life on planet Earth.

We depend on biodiversity.

We depend on the understanding of species and its traits.

We don't know how many species there are.
We don't know, how many we know.



Diss. ETH No.8774

VERSUCH EINER PHYLOGENETISCHEN WERTUNG DER MERKMALE DER FORMICINI (HYMENOPTERA, FORMICIDAE), REVISION DER FORMICA EXSECTA-GRUPPE UND LISTE DER FORMICIDAE EUROPAS

ABHANDLUNG
zur Erlangung des Titels eines
DOKTORS DER NATURWISSENSCHAFTEN

der
EIDGENÖSSISCHEN TECHNISCHEN HOCHSCHULE ZÜRICH

vorgelegt von

DONAT AGOSTI
Dipl. Natw. ETH

geboren am 6. Oktober 1958

von Trun und Zürich

Angenommen auf Antrag von:

Prof. Dr. W. Sauter, Referent

Prof. Dr. G. Benz, Korreferent

1989

The challenge

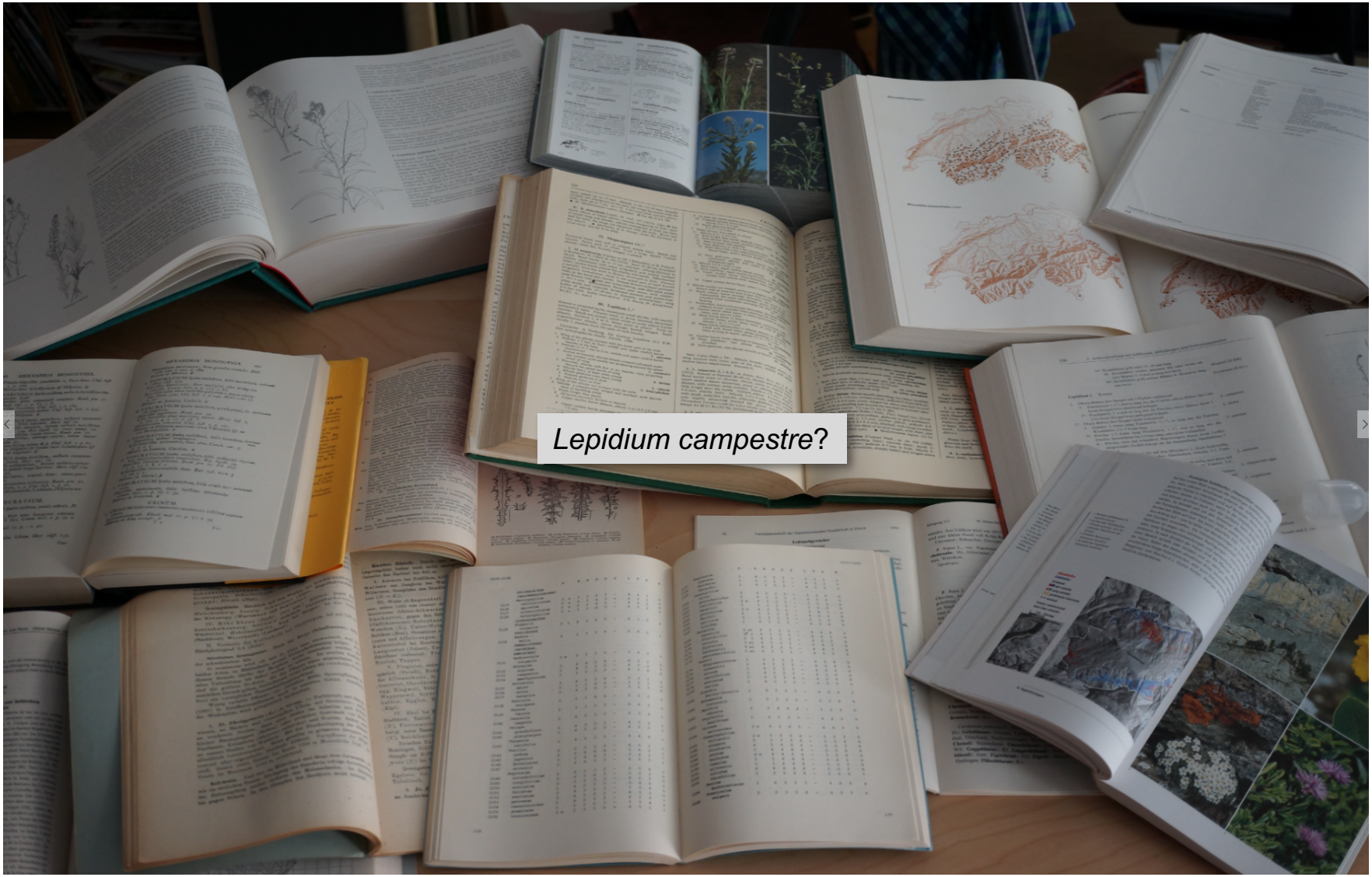
500,000,000+ printed pages
1,900,000 species described
20,000,000+ species treatments
18,000 new species discovered / year

BUT: Data are hidden

- Incomplete digitization
- Publications are not semantically enhanced
- Data are not linked
- Most data are not open

How can we query our accumulated biodiversity knowledge?

Biodiversity Knowledge: 20 years ago, and still today



TreatmentBank: Visualization of data from one article

PLAZI
TAKING CARE OF FREEDOM

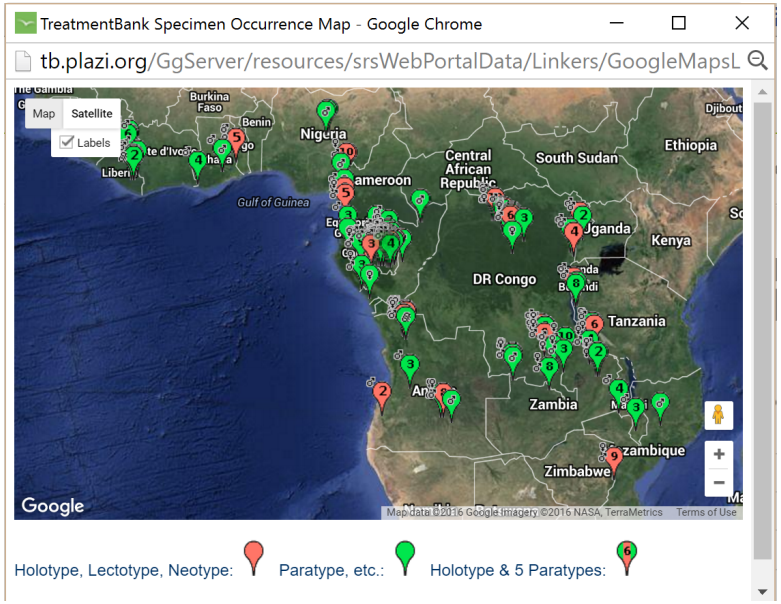
News | Resources | Projects | Activities | Cooperation | Developers | About | Search

Dijkstra, Klaas-Douwe B., Kipping, Jens & Mézière, Nicolas, 2015, Sixty new dragonfly and damselfly species from Africa (Odonata)

Dijkstra, Klaas-Douwe B., Kipping, Jens & Mézière, Nicolas, 2015, Sixty new dragonfly and damselfly species from Africa (Odonata), Odonatologica 44 (4), pp. 447-678

publication ID	http://dx.doi.org/10.5281/zenodo.35388
link to original document	http://dx.doi.org/10.5281/zenodo.35388
document provided by	Donat (2015-12-11 15:22:31, last updated by Admin 2015-12-18 12:58:08)

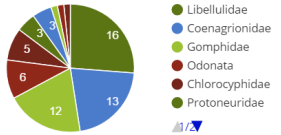
Treatments (61)



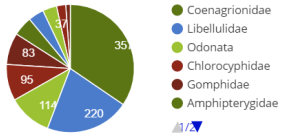
Distribution Map

Specimens

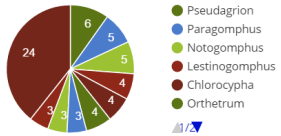
Article: Treatments by Family (n=61)



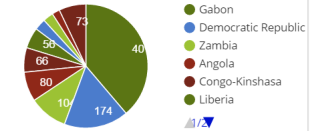
Article: Specimens by Family (n=1034)

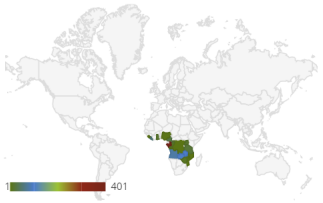


Article: Treatments by Genus (n=61)

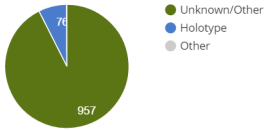


Article: Specimens by Country (n=1034)

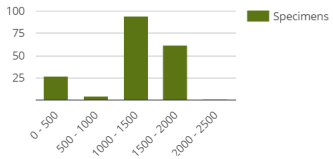




Article: Specimens by Type Status (n=1034)



Article: Specimens by Elevation (n=1034)



Explicit links

Profundiconus puillandrei sp. nov.
urn:lsid:zoobank.org:act:55807BF8-D984-42C1-B039-F3EBD8EB2C13
Figs 12A–J, 13A–D, 14

Conus ikedai – Poppe 2008: pl. 615, fig. 1a–b (non *C. ikedai* Ninomiya, 1987).
Conus smirna – Marshall 1981: 499, fig. 3j (non *C. smirna* Bartsch & Rehder, 1943).
Conus sp. C – Röckel *et al.* 1995b: 585, fig. 49.
Profundiconus n. sp. g. – Puillandre *et al.* 2014: Supplementary Material 1 (unfigured).

Etymology

This new species is dedicated to Dr. Nicolas Puillandre, mollusc curator at the MNHN. Dr. Puil
has a long and outstanding trajectory in the study of the phylogeny of the Conoidea. The nam
the new species after him recognises his important contributions to the taxonomy of Conoidea
molecular level.

Type material examined

Holotype
NEW CALEDONIA: 43.2 × 18.0 mm, R/V *Alis*, NORFOLK 1 Expedition, st. DW 1707, Banc J
Est, Norfolk Ridge, 23°43' S, 168°16' E, 381–493 m (MNHN IM-2000-30771; Fig. 12A).

Paratypes
NEW CALEDONIA: 39.6 × 17.3 mm, R/
Norfolk Ridge, 25°21' S, 168°57' E, 100

28

MNHN-IM-2007-30772

```
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/
  <rdf:type rdf:resource="http://plazi.org/vocab/treatment
  <dwc:collectionCode>MNHN</dwc:collectionCode>
  <dwc:collectorName>V Alis</dwc:collectorName>
  <dwc:country>New Caledonia</dwc:country>
  <dwc:elevation>1003</dwc:elevation>
  ▼<dwc:httpUri>
    http://coldb.mnhn.fr/catalognumber/mnhn/im/2007-30772
  </dwc:httpUri>
  <dwc:latitude>-25.35</dwc:latitude>
  ▼<dwc:location>
```

MUSÉUM
NATIONAL D'HISTOIRE NATURELLE

MNHN / Molluscs (IM) / 2007-30772

Conus frigidus Reeve, 1848

SPECIMEN

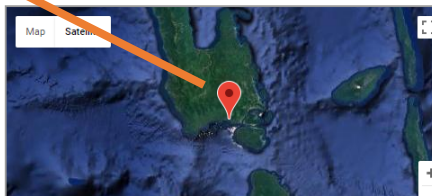
ALCOOL95	MNHN-IM-2007-30772
Size	37.3

TAXONOMY

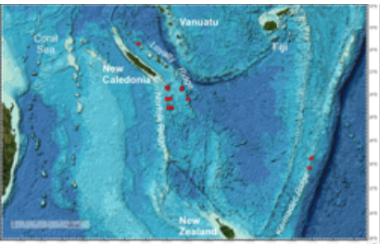
Phylum	Mollusca
Class	Gastropoda
Order	Neogastropoda
Family	Conidae
Genus	Conus
Species	Conus frigidus
Name	*Conus frigidus* Reeve, 1848

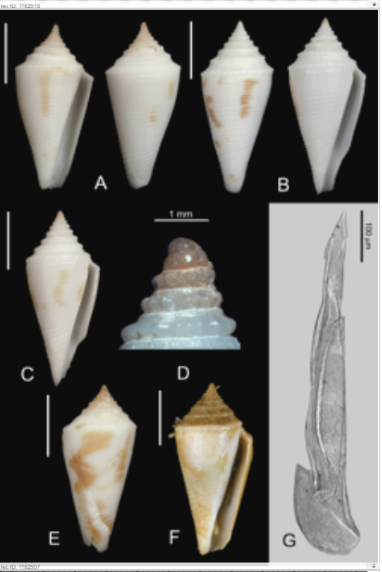
ORIGIN

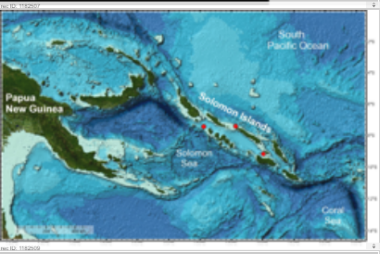
Country label	VANUATU
Locality	Baldwin Core
Expedition	SANTO 2006
Station number	VM02
Collection date	2006-09-10
Depth (meters)	0 -1
Geographic	15° 34' 51.9996" S ; 167° 2' 23.5248" E



Context


Coral Sea
Vanuatu
Fiji
New Caledonia
New Zealand

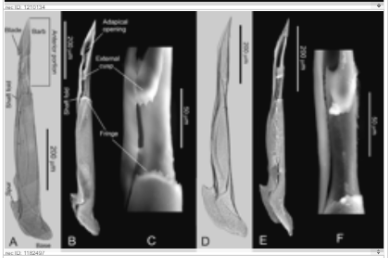

A
B
1 mm
C
D
E
F
G

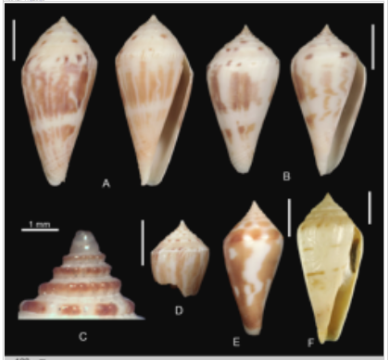

South Pacific Ocean
Papua New Guinea
Solomon Islands
Solomon Sea
Coral Sea


100 µm
A

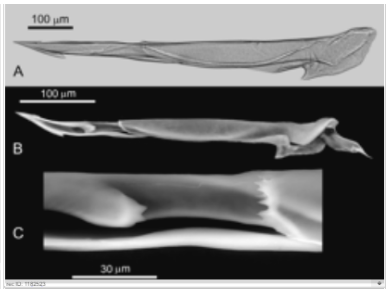

100 µm
B

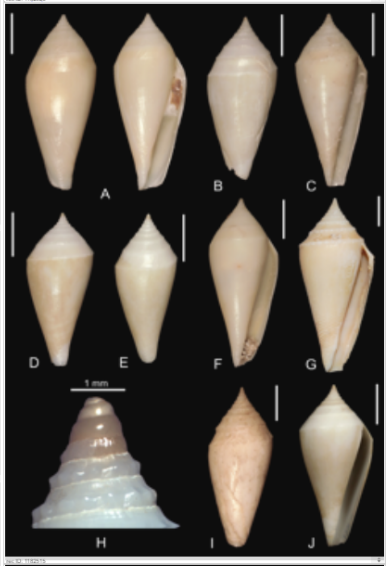

© MNHN Paris, France


Adapical opening
External cord
Fringe
A
B
C
D
E
F

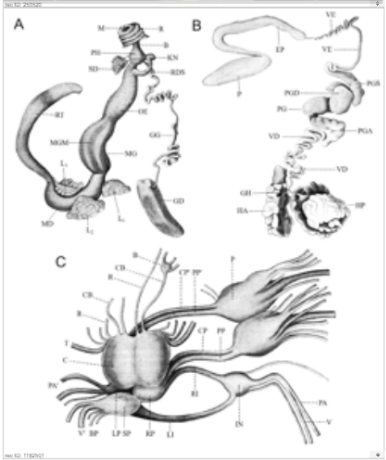

A
B
1 mm
C
D
E
F

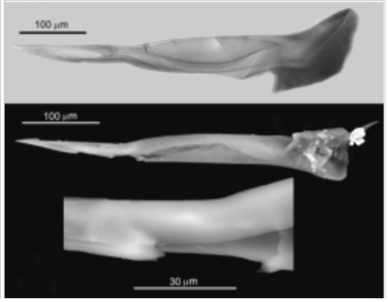

100 µm
G
100 µm

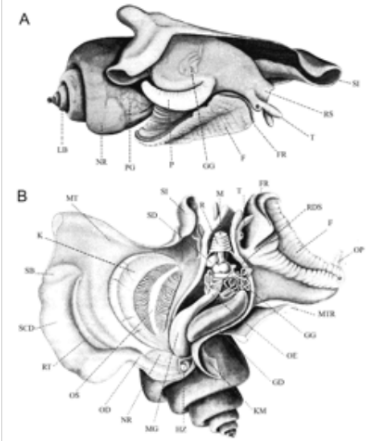

100 µm
A
100 µm
B
C
30 µm

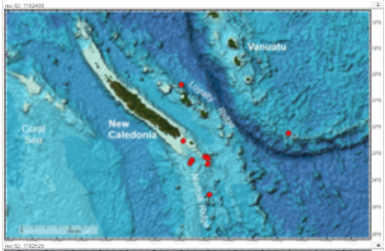

A
B
C
D
E
F
G
1 mm
H
I
J

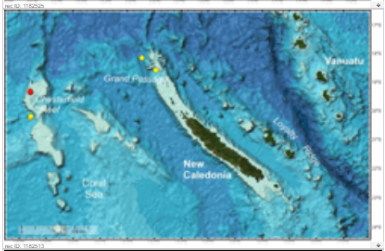

A
B
C
1 mm
D
E
F

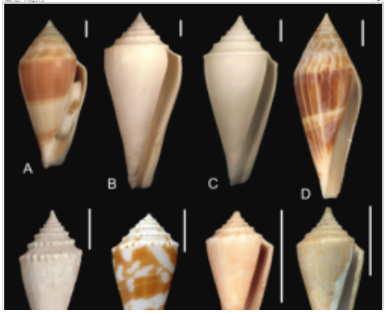

A
B
C


100 µm
100 µm
30 µm


A
B


Vanuatu
New Caledonia
Coral Sea


Vanuatu
Grand Passage
New Caledonia
Coral Sea


A
B
C
D

Dissemination

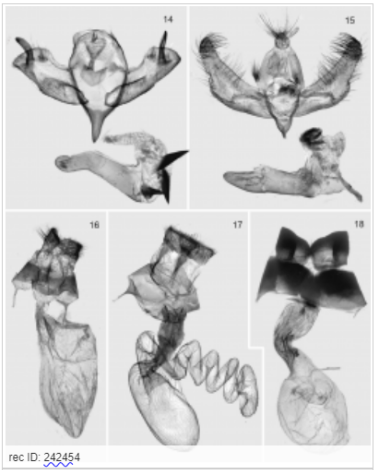
Reuse of **subarticle** parts

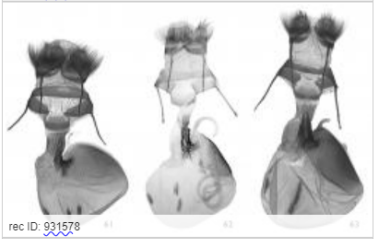
Novel applications, novel access to publications

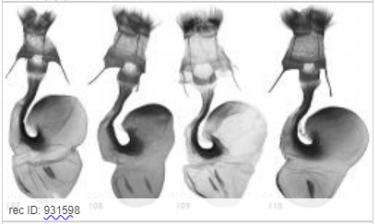
OCELLUS

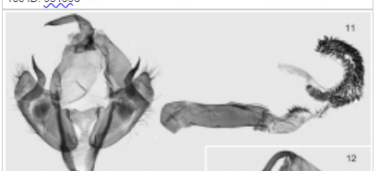
noctuidae+genitalia GO

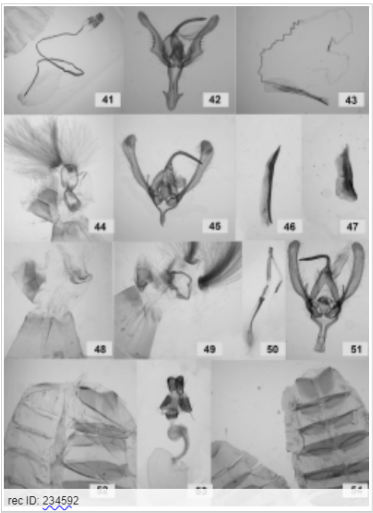
13111 records found

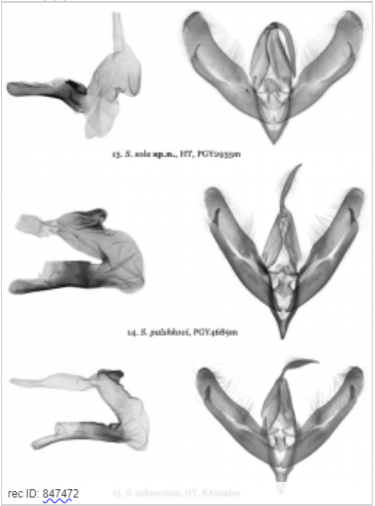


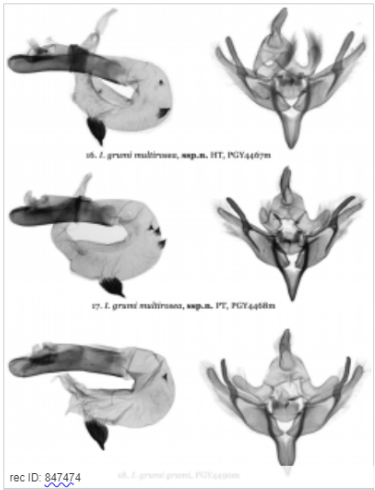


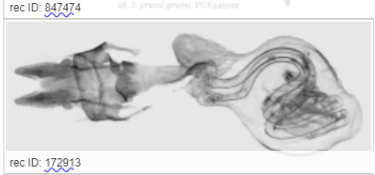


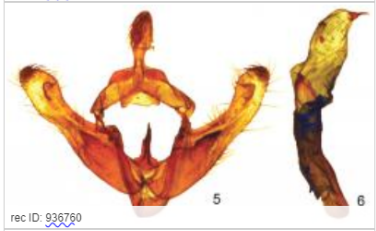


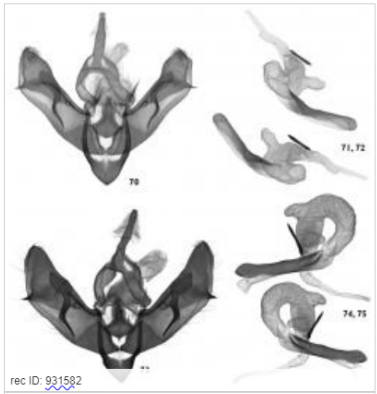


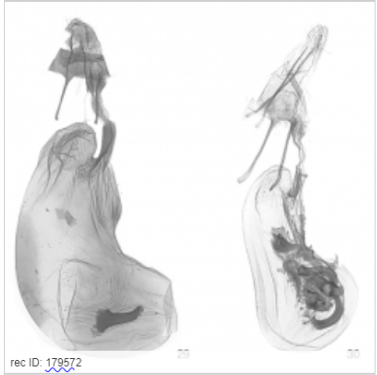


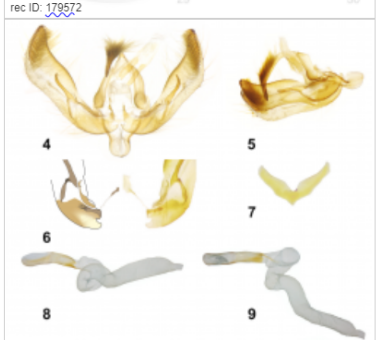




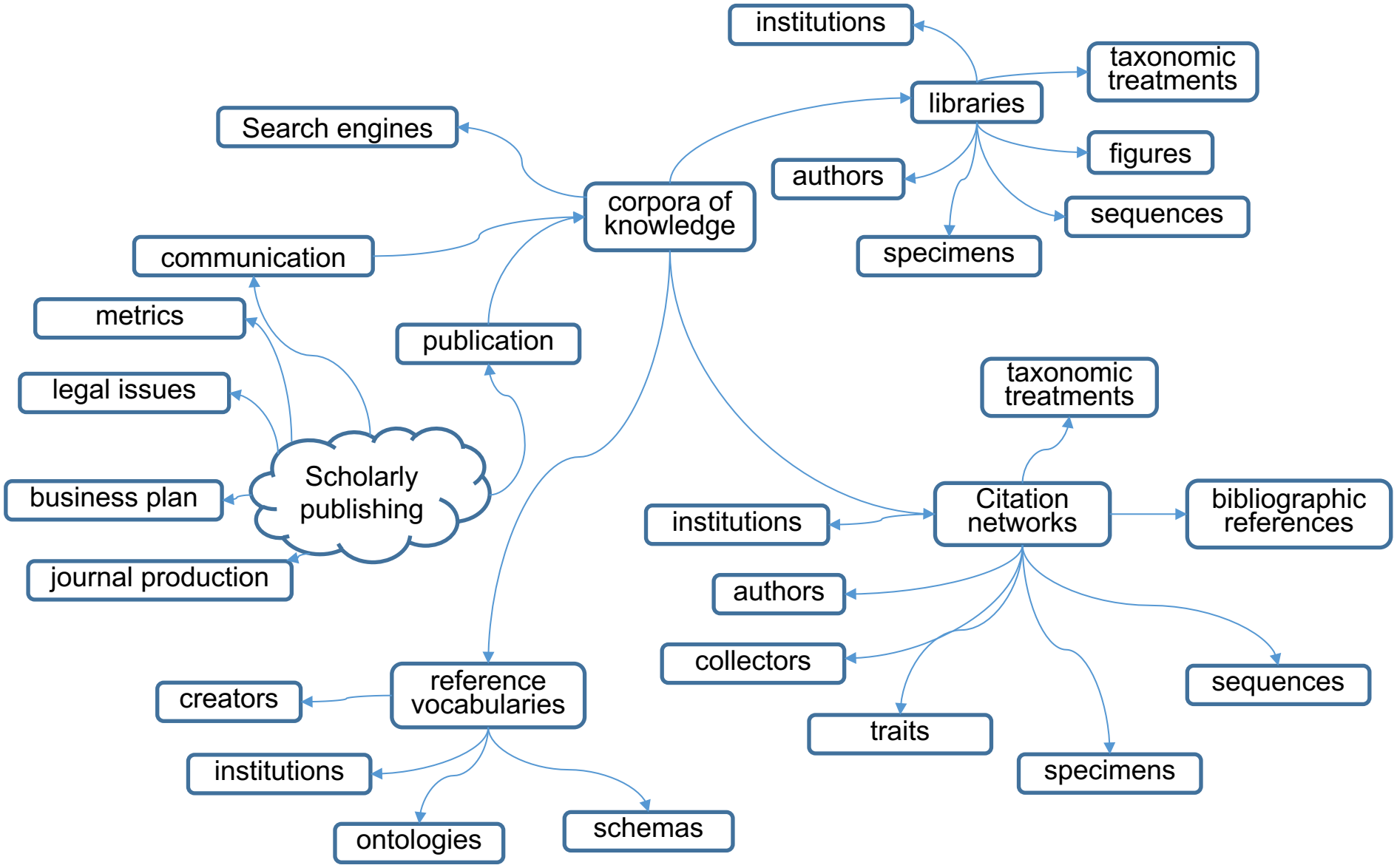




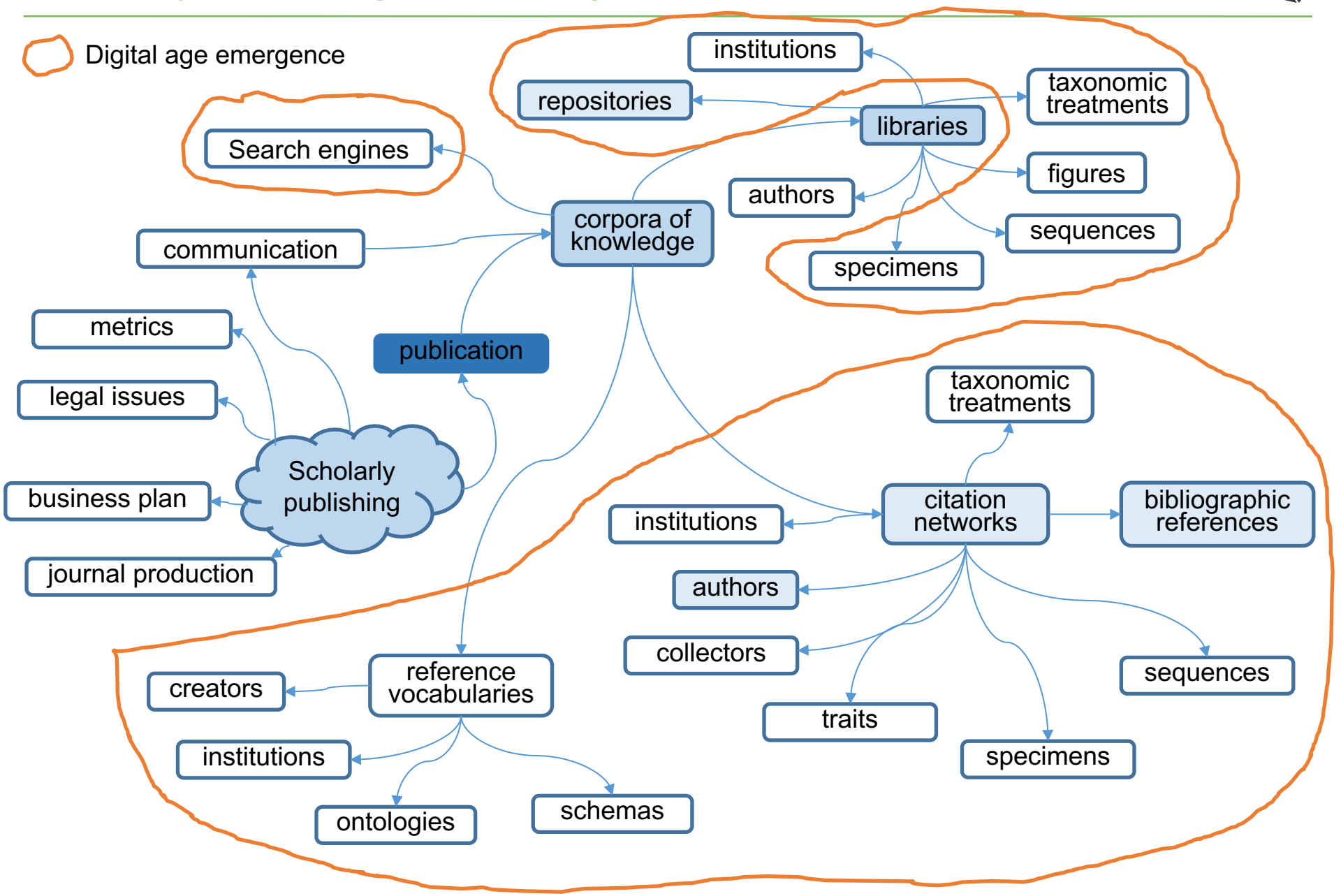




Scholarly publishing



Scholarly publishing: a daunting task



The PDF prison

Highly **structured** research data

publishing

Semantically **unstructured** text

data mining

Semantically enhanced, **structured** data

Beyond the PDF – unstructured to structured publications …

From an article to data and knowledge


```
▼<rdf:RDF xmlns:dwc="http://rs.tdwg.org/dwc/terms/" xmlns:cnt="http://www.w3.org/2011/content#" xmlns:spm="http://rs.tdwg.org/ontology/voc/SpeciesProfileModel"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns
 xmlns:dwcFP="http://filteredpush.org/ontologies/oa/dwcFP#">
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5">
   <rdf:type rdf:resource="http://plazi.org/vocab/treatment#Treatment"/>
   <trt:definesTaxonConcept rdf:resource="http://taxon-concept.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5"/>
   <trt:publishedIn rdf:resource="http://dx.doi.org/10.5281/zenodo.33990"/>
   <spm:hasInformation rdf:resource="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#section_1"/>
   <spm:hasInformation rdf:resource="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#section_2"/>
   <spm:hasInformation rdf:resource="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#section_3"/>
   <trt:hasMaterialCitation rdf:resource="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#materialCitation_1"/>
 </rdf:Description>
▼<rdf:Description rdf:about="http://dx.doi.org/10.5281/zenodo.33990">
 ▼<dc:title>
    New and remarkable leafhoppers and planthoppers (Hemiptera: Auchenorrhyncha) from Switzerland
   </dc:title>
   <dc:creator>Valeria Trivellone</dc:creator>
   <dc:creator>Eva Knop</dc:creator>
   <dc:creator>Tabea Turrini</dc:creator>
   <dc:creator>Line Andrey</dc:creator>
   <dc:creator>Jean-Yves Humbert</dc:creator>
   <dc:creator>Gernot Kunz</dc:creator>
   <rdf:type rdf:resource="fabio:JournalArticle"/>
 ▼<bibo:journal>
    Mitteilungen Der Schweizerischen Entomologischen Gesellschaft
   </bibo:journal>
   <bibo:pageEnd>284</bibo:pageEnd>
 </rdf:Description>
▼<rdf:Description rdf:about="http://taxon-concept.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5">
   <rdf:type rdf:resource="http://filteredpush.org/ontologies/oa/dwcFP#Taxon"/>
   <dwc:authority>Say, 1830</dwc:authority>
   <dwc:authorityName>Say</dwc:authorityName>
   <dwc:authorityYear>1830</dwc:authorityYear>
   <dwc:box>[153,494,999,1022]</dwc:box>
   <dwc:class>Insecta</dwc:class>
   <dwc:family>Acanaloniidae</dwc:family>
   <dwc:genus>Acanalonia</dwc:genus>
   <dwc:higherTaxonomySource>CoL</dwc:higherTaxonomySource>
   <dwc:kingdom>Animalia</dwc:kingdom>
   <dwc:order>Hemiptera</dwc:order>
   <dwc:pageId>1</dwc:pageId>
   <dwc:pageNumber>274</dwc:pageNumber>
   <dwc:phylum>Arthropoda</dwc:phylum>
   <dwc:rank>species</dwc:rank>
   <dwc:species>conica</dwc:species>
 </rdf:Description>
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03
   <rdf:type rdf:resource="spm:InfoItem"/>
 ▼<spm:hasContent>
    First record from Switzerland: Ticino, Agno, green area[
   </spm:hasContent>
 </rdf:Description>
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03
   <rdf:type rdf:resource="spm:InfoItem"/>
   <spm:hasContent>Distribution:USA, Northern Italy, Switzerl
 </rdf:Description>
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03
   <rdf:type rdf:resource="spm:InfoItem"/>
 ▼<spm:hasContent>
    Remarks:  A. conicais for the moment the only representat
    (Wilson & Lucchi 2001). In Europe it was recorded for the
   </spm:hasContent>
 </rdf:Description>
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03
   <rdf:type rdf:resource="cito:CitationAct"/>
   <cito:hasCitingEntity rdf:resource="http://treatment.plazi
   <cito:hasCitedEntity rdf:resource="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#material_1"/>
   <cito:hasCitationCharacterization rdf:resource="cito:citesAsEvidence"/>
 </rdf:Description>
▼<rdf:Description rdf:about="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0BFC7BFDE0FAC5#material_1">
```



```
▼<rdf:Description
 rdf:about="http://treatment.plazi.org/id/03FEB44DFF91FFB1FF0B
   <rdf:type rdf:resource="http://plazi.org/vocab/treatment#Ma
   <dwc:collectingDate>2014-07-26</dwc:collectingDate>
   <dwc:collectingMethod>mouth aspirator</dwc:collectingMethod
   <dwc:collectorName>Valeria Trivellone</dwc:collectorName>
   <dwc:country>Switzerland</dwc:country>
   <dwc:determinerName>Valeria Trivellone</dwc:determinerName>
   <dwc:elevation>272</dwc:elevation>
   <dwc:latitude>45.995697</dwc:latitude>
   <dwc:location>Agno, green area</dwc:location>
   <dwc:longLatPrecision>1</dwc:longLatPrecision>
   <dwc:longitude>8.902558</dwc:longitude>
   <dwc:pageId>1</dwc:pageId>
   <dwc:pageNumber>274</dwc:pageNumber>
   <dwc:specimenCount>1</dwc:specimenCount>
   <dwc:specimenCount-male>1</dwc:specimenCount-male>
   <dwc:stateProvince>Ticino</dwc:stateProvince>
 </rdf:Description>
</rdf:RDF>
```


One article = thousands of facts

Modeling scientific domains

Development of

- Standards vocabularies
- Ontologies
- Schemas

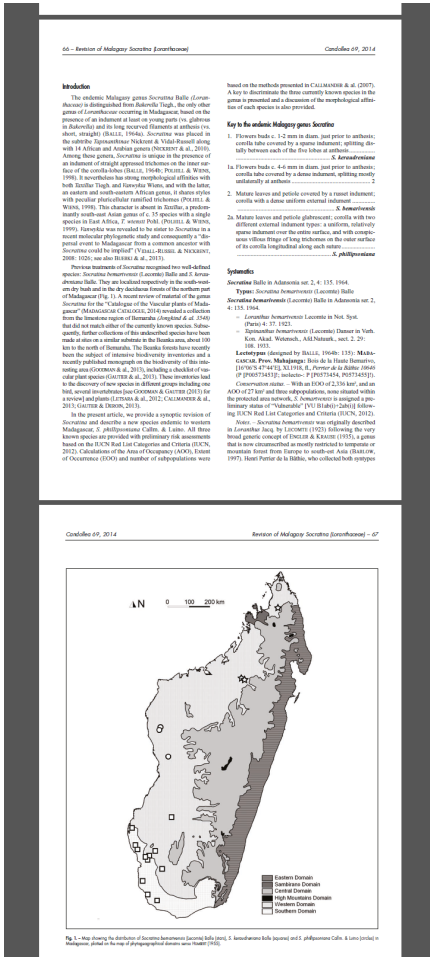
- Tools to mine and semantically enhance publications

- Repositories for semantical enhanced entities to make them FAIR

- Search engines

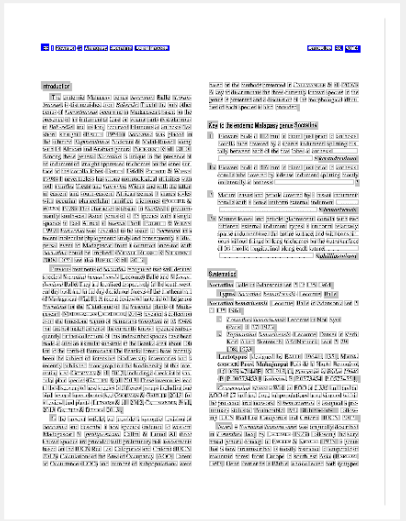
Plazi: conversion: extracting text streams and other elements

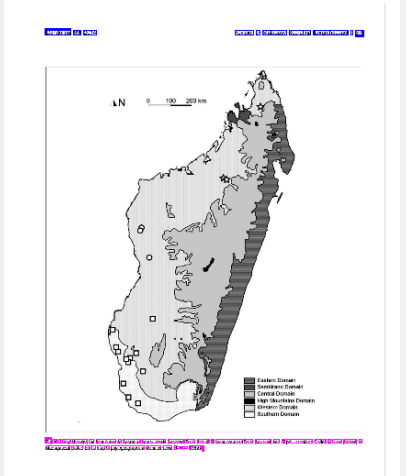
Re-creating a semantically enhanced publication



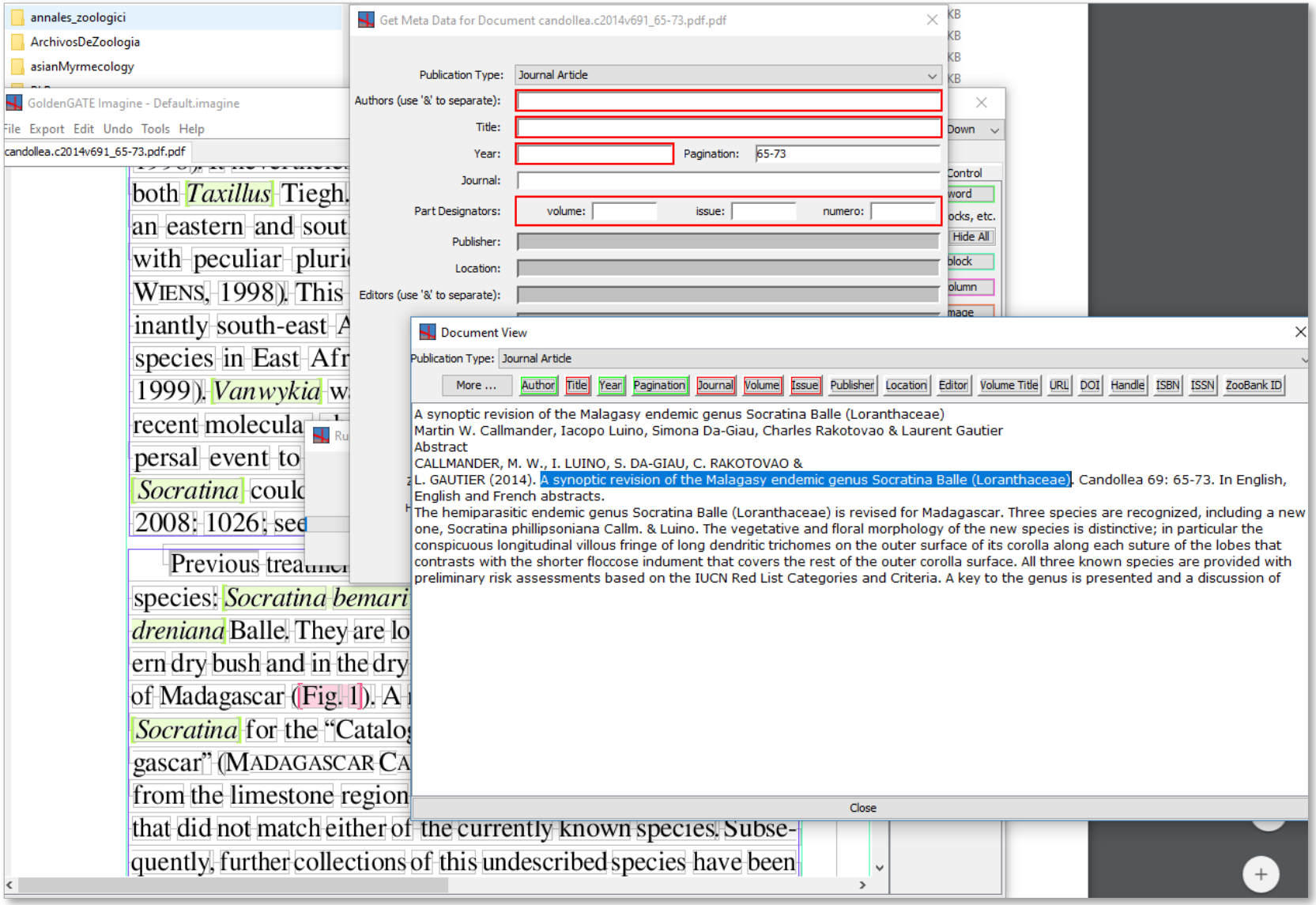
preparation

Extraction of text
text stream
various digital elements

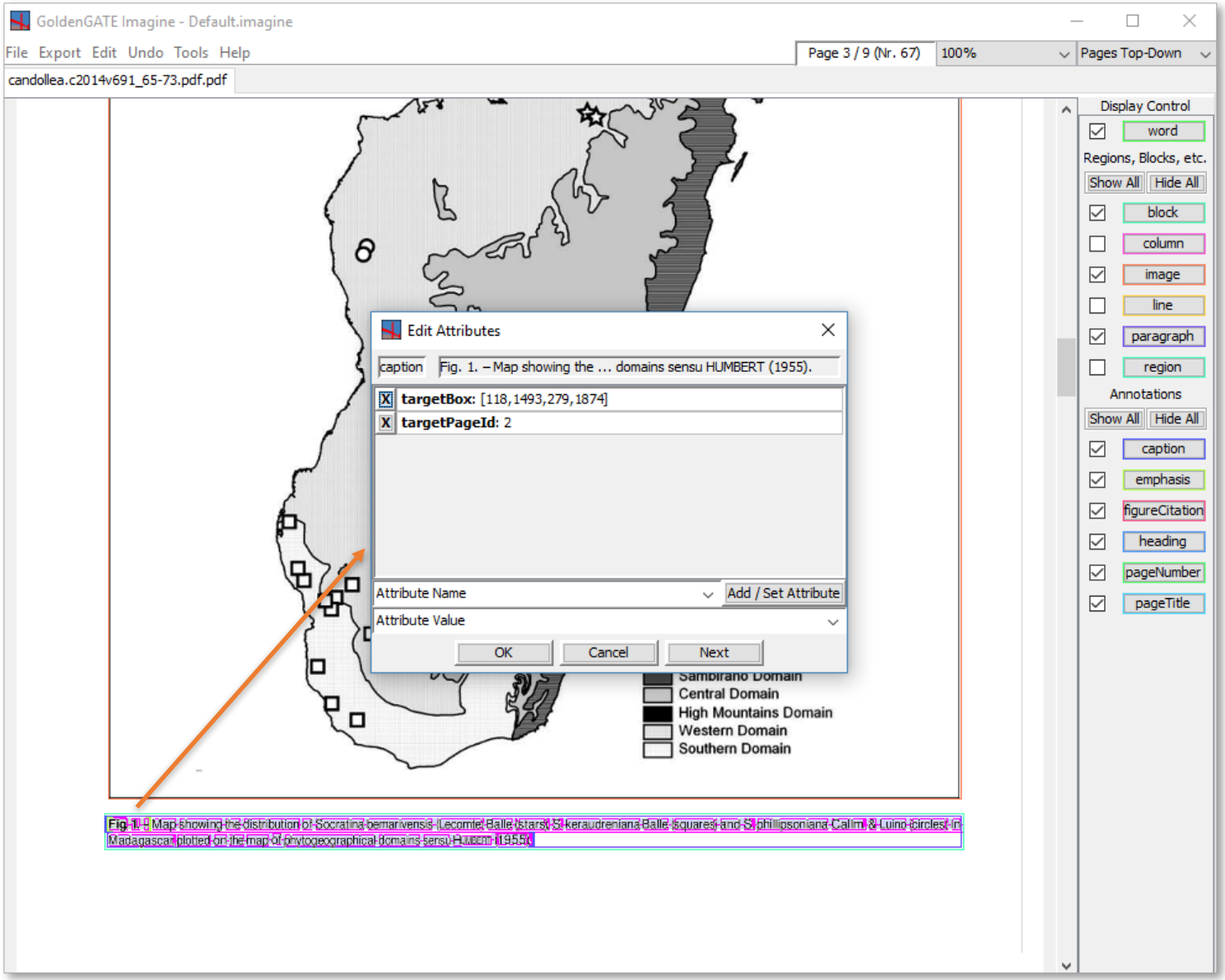




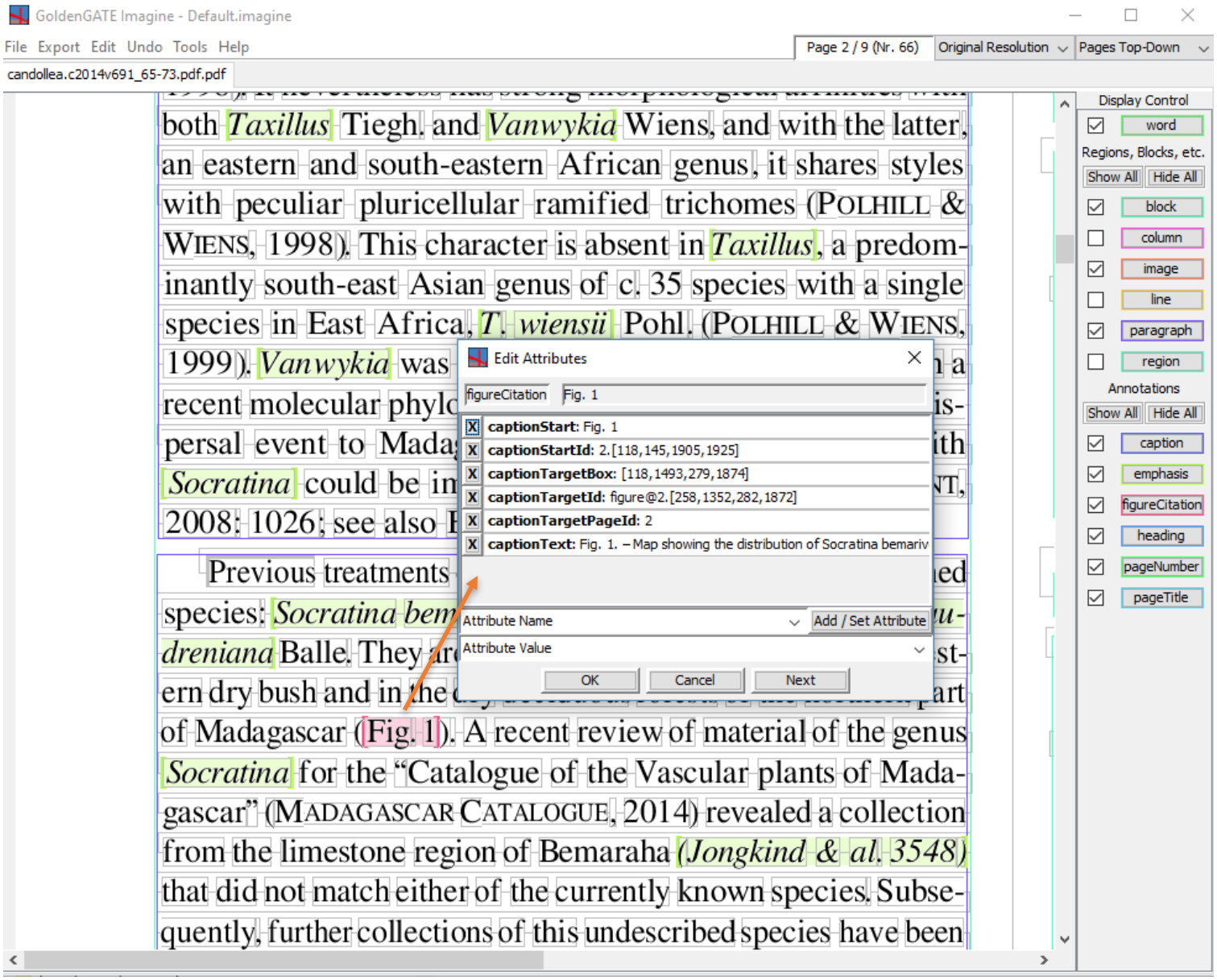
Plazi: conversion: keep the provenience



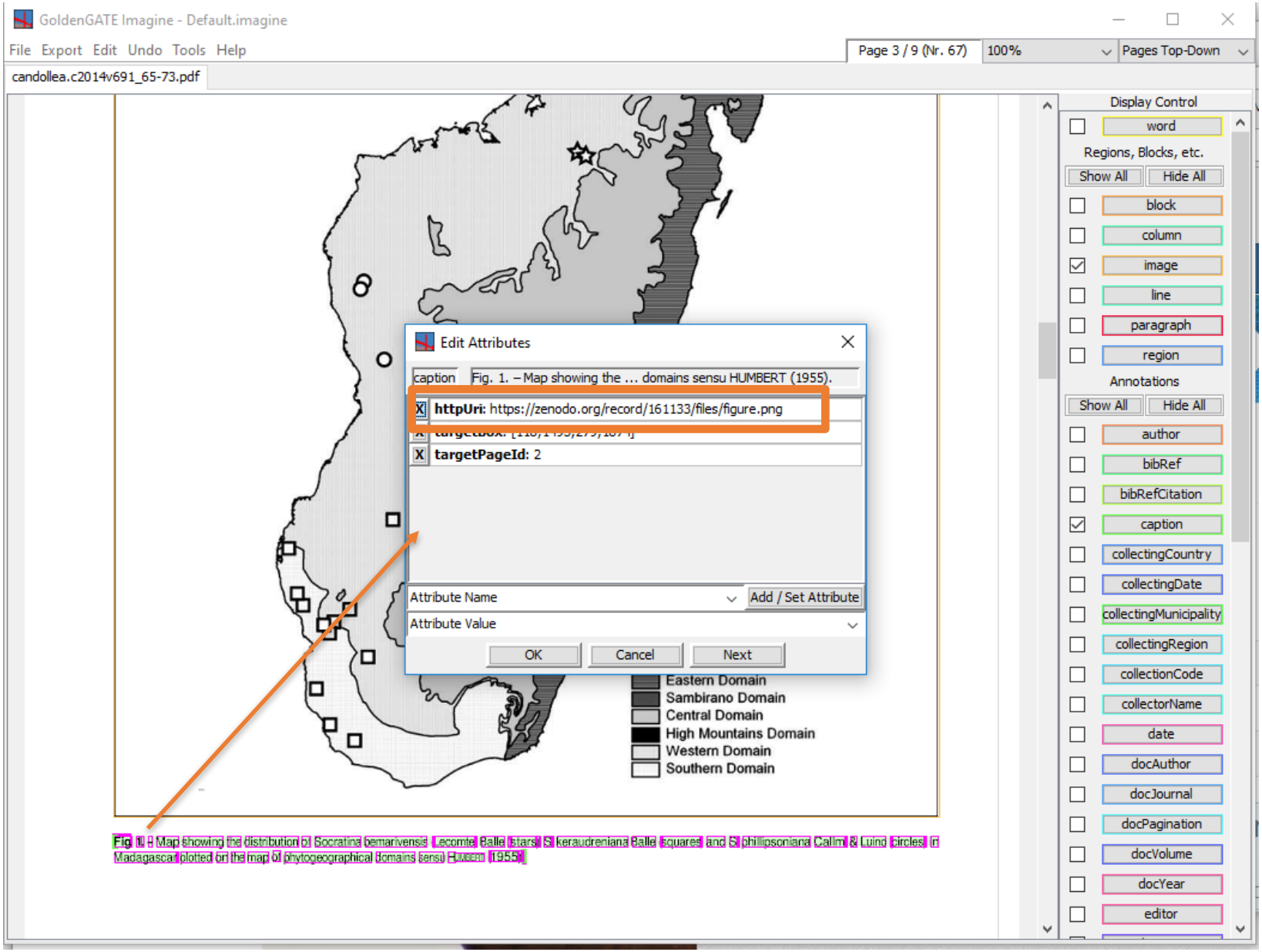
Plazi: conversion: discovering images



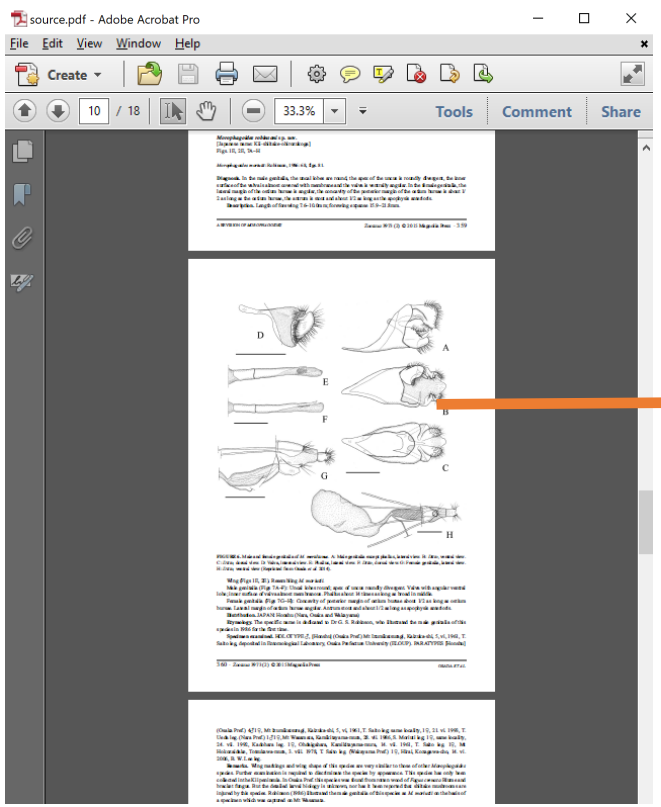
Plazi: conversion: discovering images

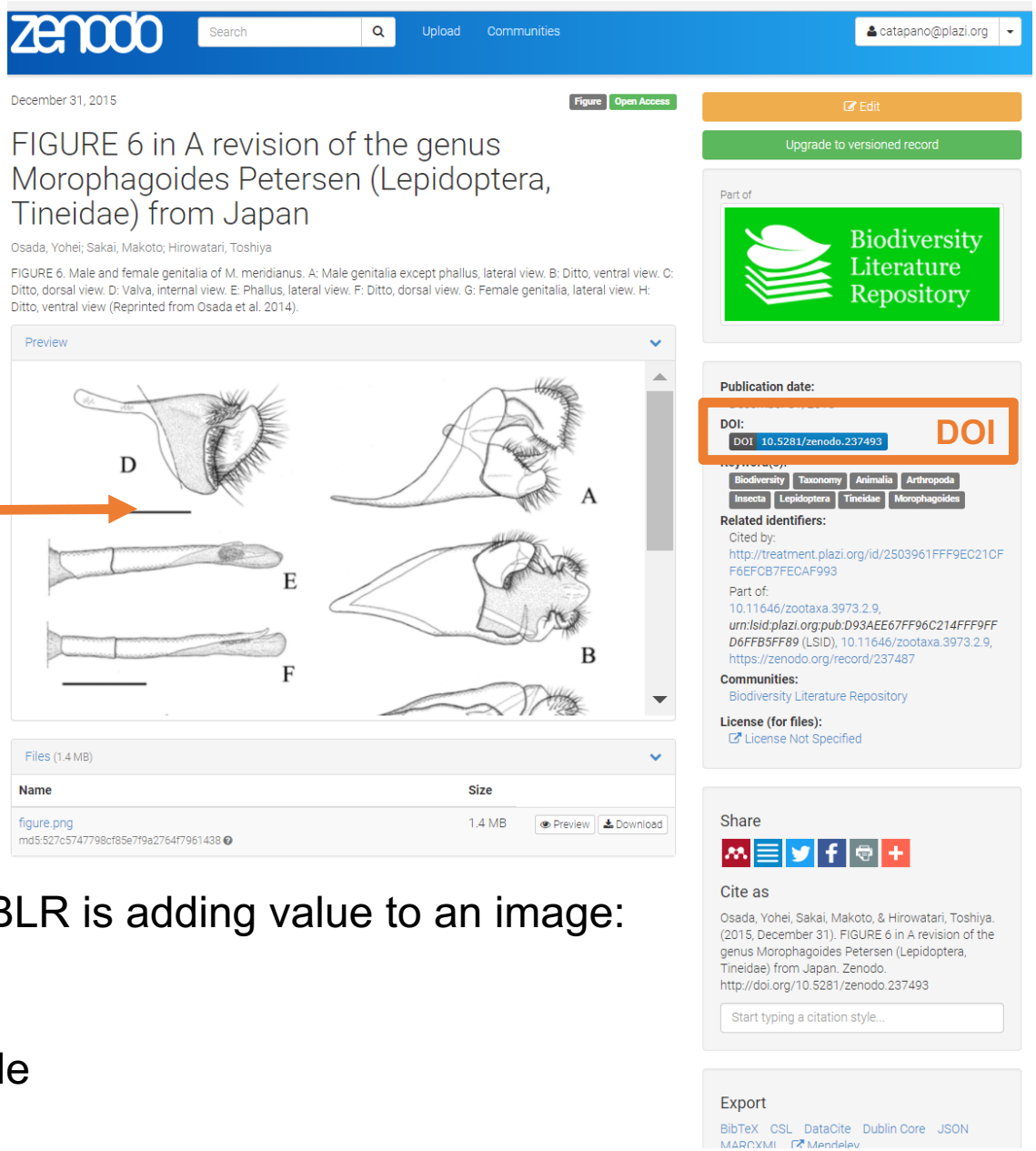


Plazi: conversion: discovering images



Plazi infrastructure: Biodiversty Literature Repositor

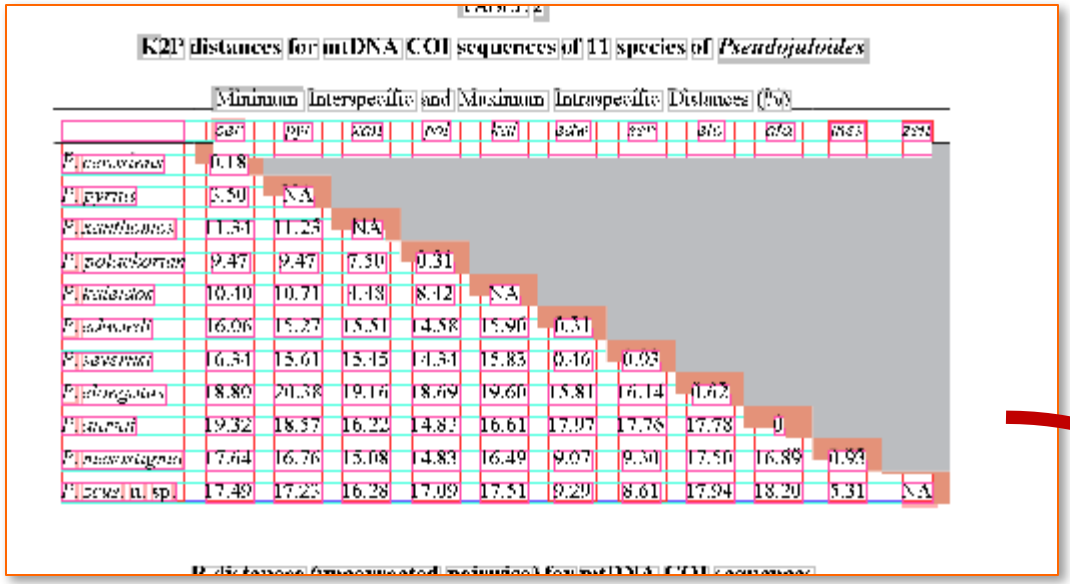




Depositing in BLR is adding value to an image:

- findable
- accessible
- interoperable
- reusable

Plazi: GoldenGATE: Table extraction



K2P distances for mtDNA COI sequences of 11 species of *Pseudojuloides*

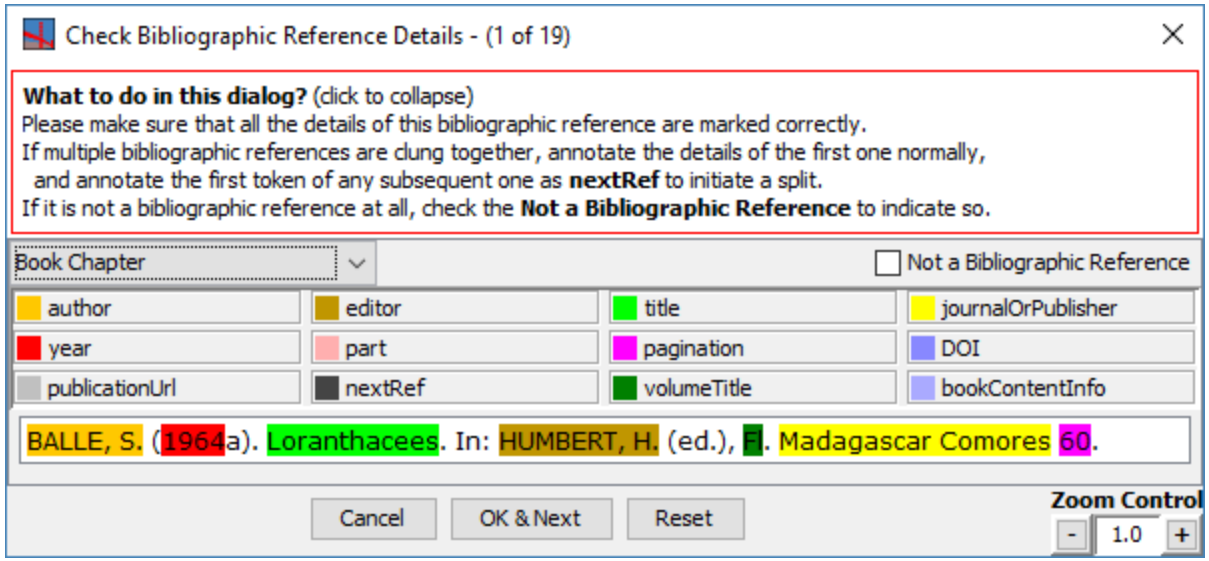
Minimum Interspecific and Maximum Intraspecific Distances (%)

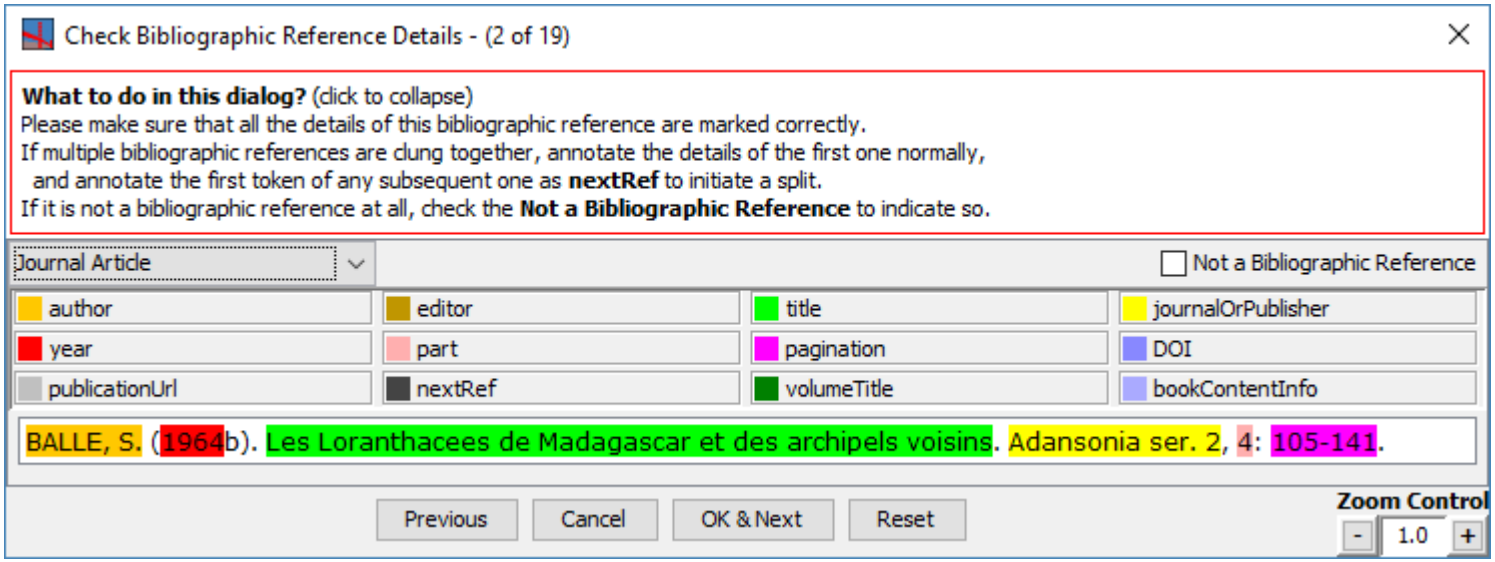
	cer	pyr	xan	pol	kal	edw	sev	elo	ata	mes	zeu
P. cerasinus	0.18										
P. pyrius	3.50	NA									
P. xanthomos	11.34	11.23	NA								
P. polackorum	9.47	9.47	7.50	0.31							
P. kaleidos	10.40	10.71	4.48	8.42	NA						
P. edwardi	16.06	15.27	15.51	14.58	15.90	0.31					
P. severnsi	16.34	15.61	15.45	14.34	15.83	0.46	0.93				
P. elongatus	18.80	20.38	19.16	18.69	19.60	15.81	16.14	0.62			
P. atavai	19.32	18.57	16.22	14.83	16.61	17.97	17.76	17.78	0		
P. mesostigma	17.64	16.76	15.08	14.83	16.49	9.07	9.30	17.50	16.89	0.93	
P. zeus, n. sp.	17.49	17.23	16.28	17.09	17.51	9.29	8.61	17.94	18.29	5.31	NA

Mark, copy and paste
(csv, text, excel)

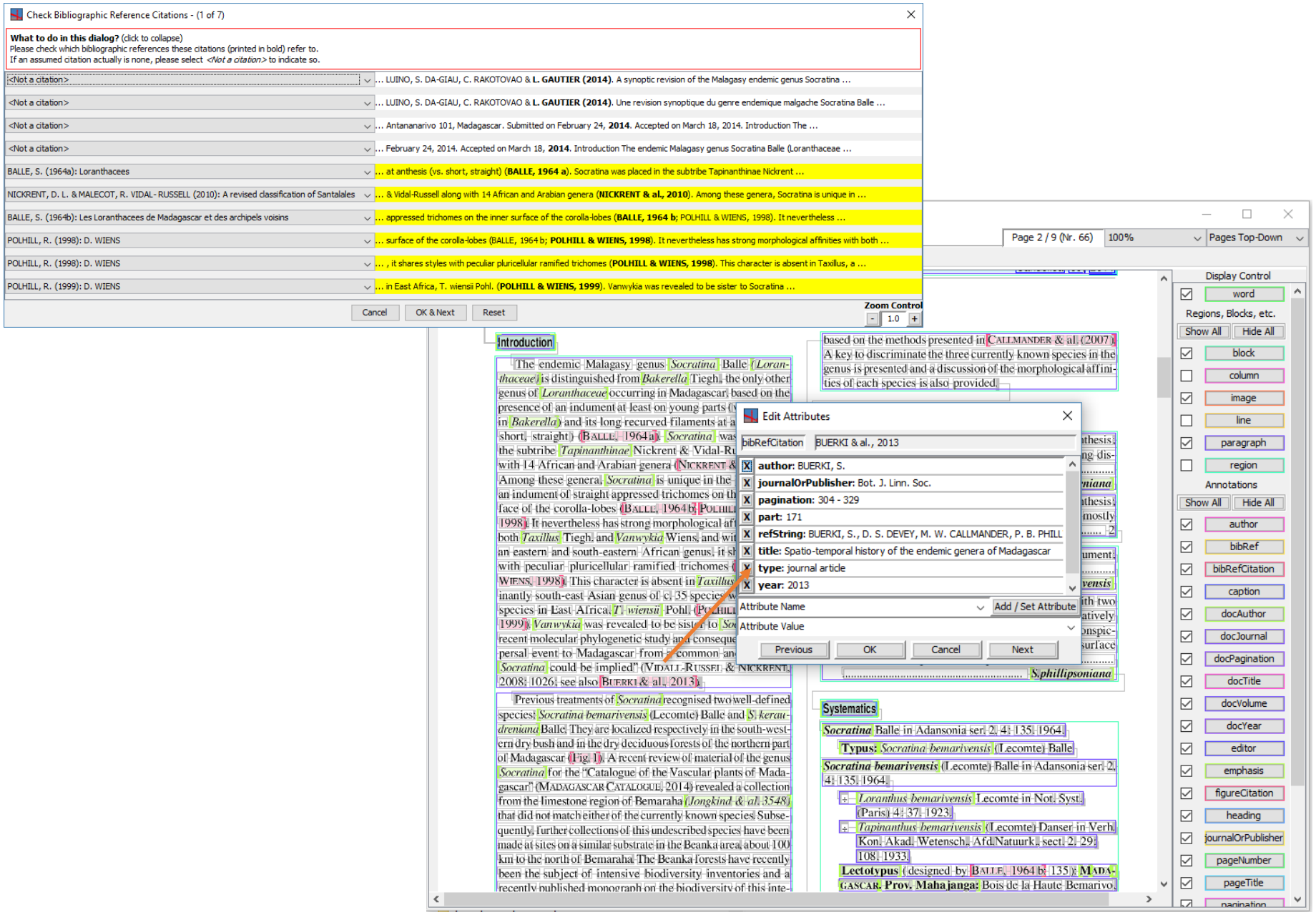
	cer	pyr	xan	pol	kal	edw	sev	elo	ata	mes	zeu
P. cerasinus	0.18										
P. pyrius	3.50	NA									
P. xanthomos	11.34	11.23	NA								
P. polackorum	9.47	9.47	7.50	0.31							
P. kaleidos	10.40	10.71	4.48	8.42	NA						
P. edwardi	16.06	15.27	15.51	14.58	15.90	0.31					
P. severnsi	16.34	15.61	15.45	14.34	15.83	0.46	0.93				
P. elongatus	18.80	20.38	19.16	18.69	19.60	15.81	16.14	0.62			
P. atavai	19.32	18.57	16.22	14.83	16.61	17.97	17.76	17.78	0		
P. mesostigma	17.64	16.76	15.08	14.83	16.49	9.07	9.30	17.50	16.89	0.93	
P. zeus, n. sp.	17.49	17.23	16.28	17.09	17.51	9.29	8.61	17.94	18.20	5.31	NA

Plazi: conversion

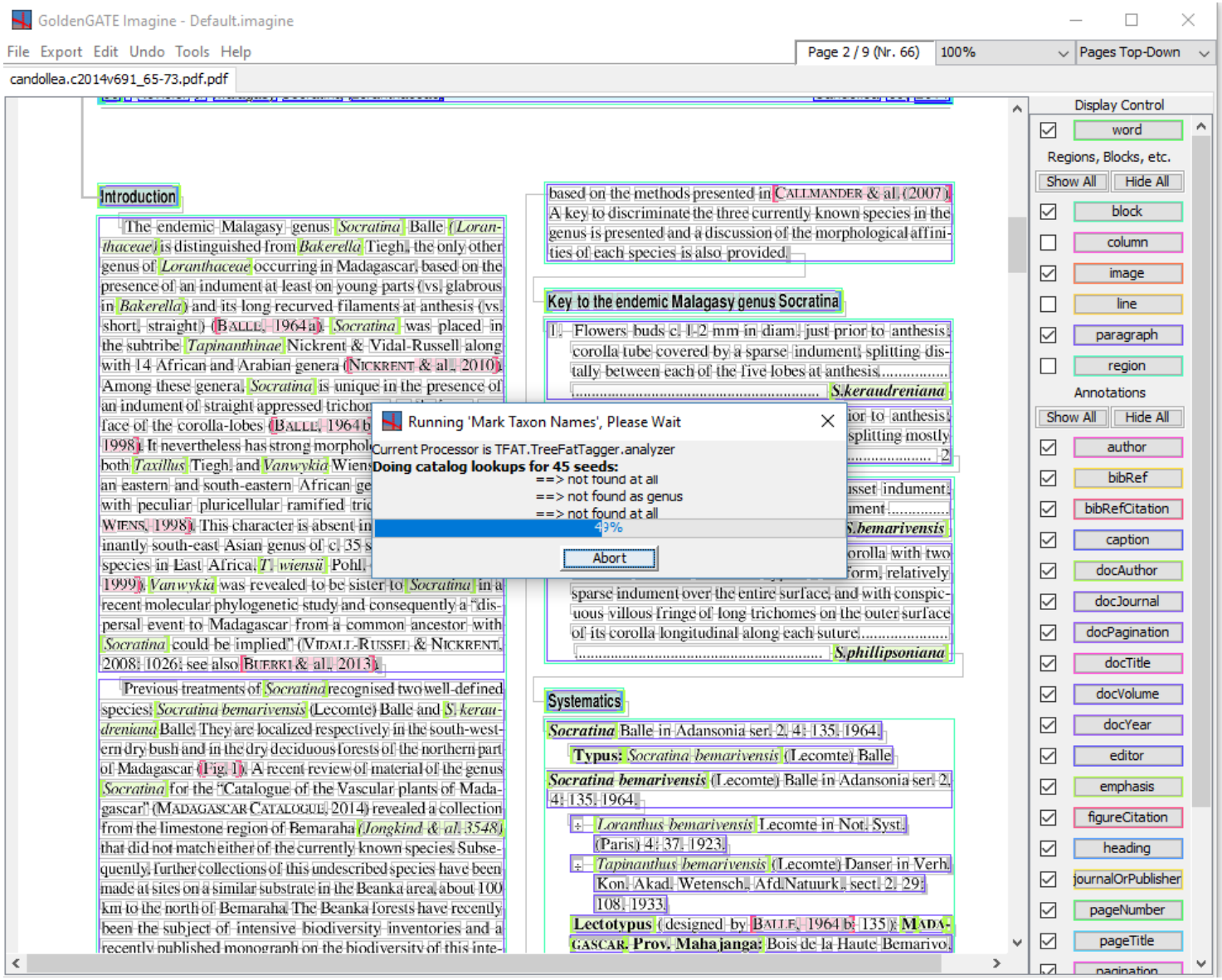




Plazi: conversion

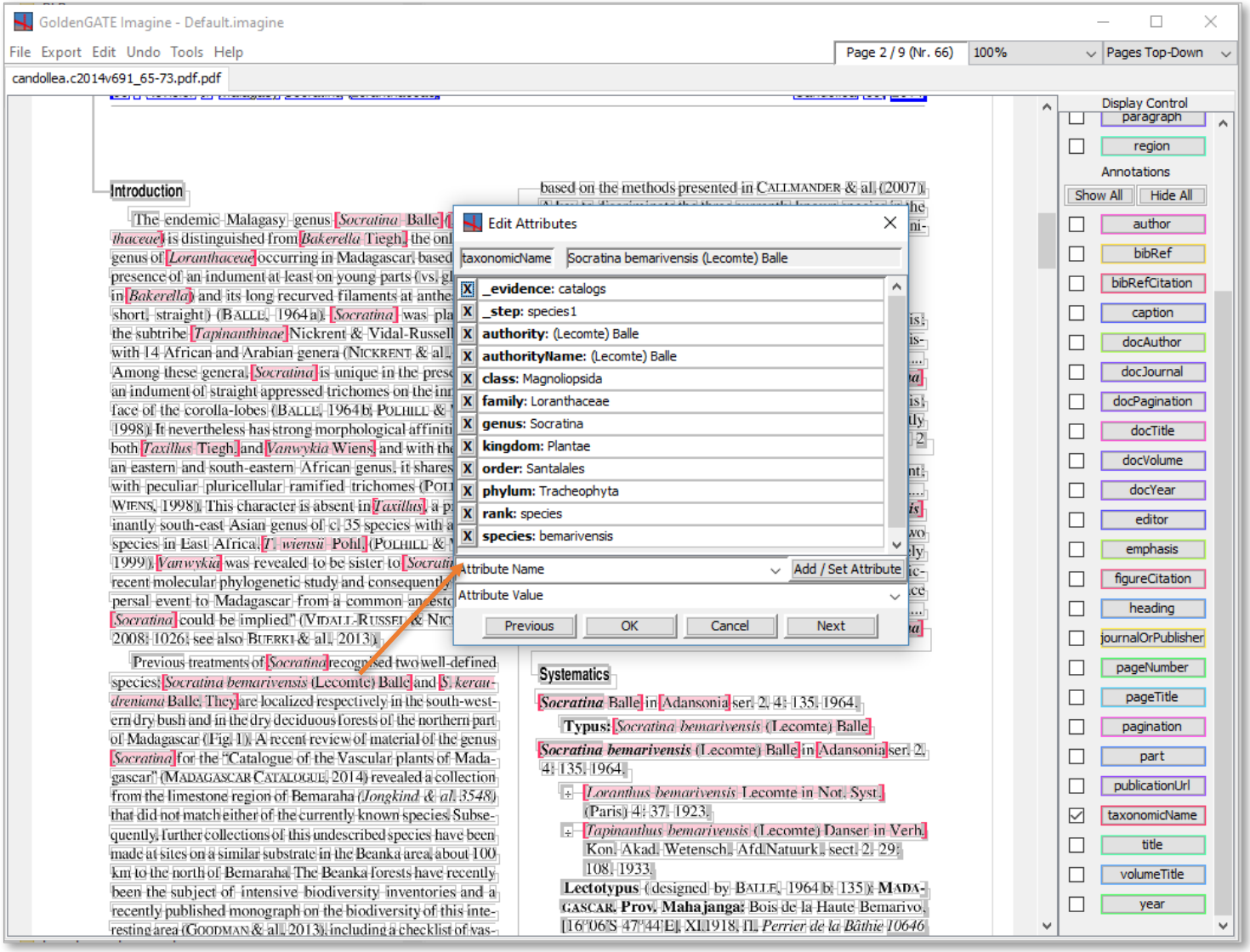


Plazi: conversion

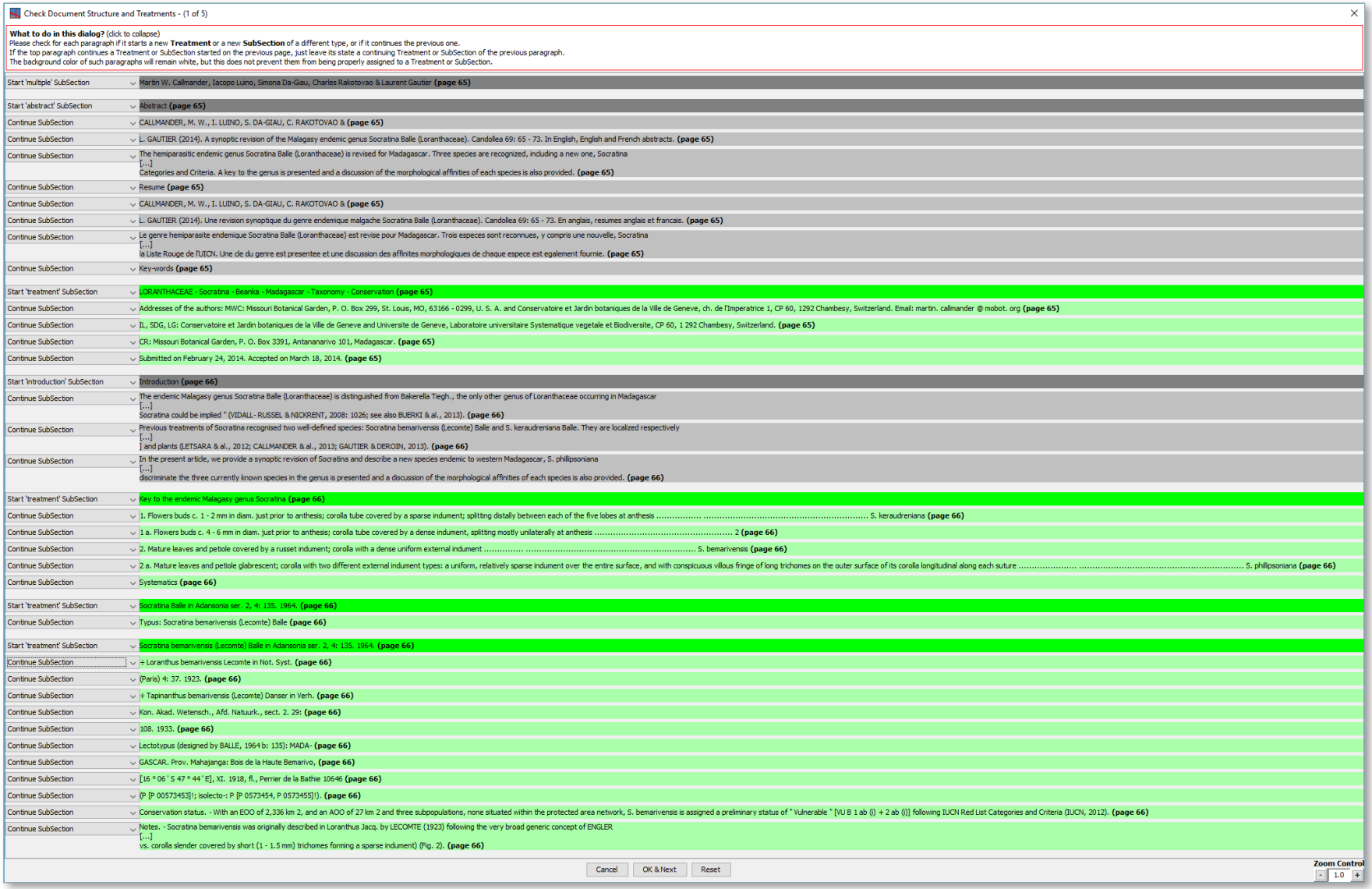


Plazi: conversion

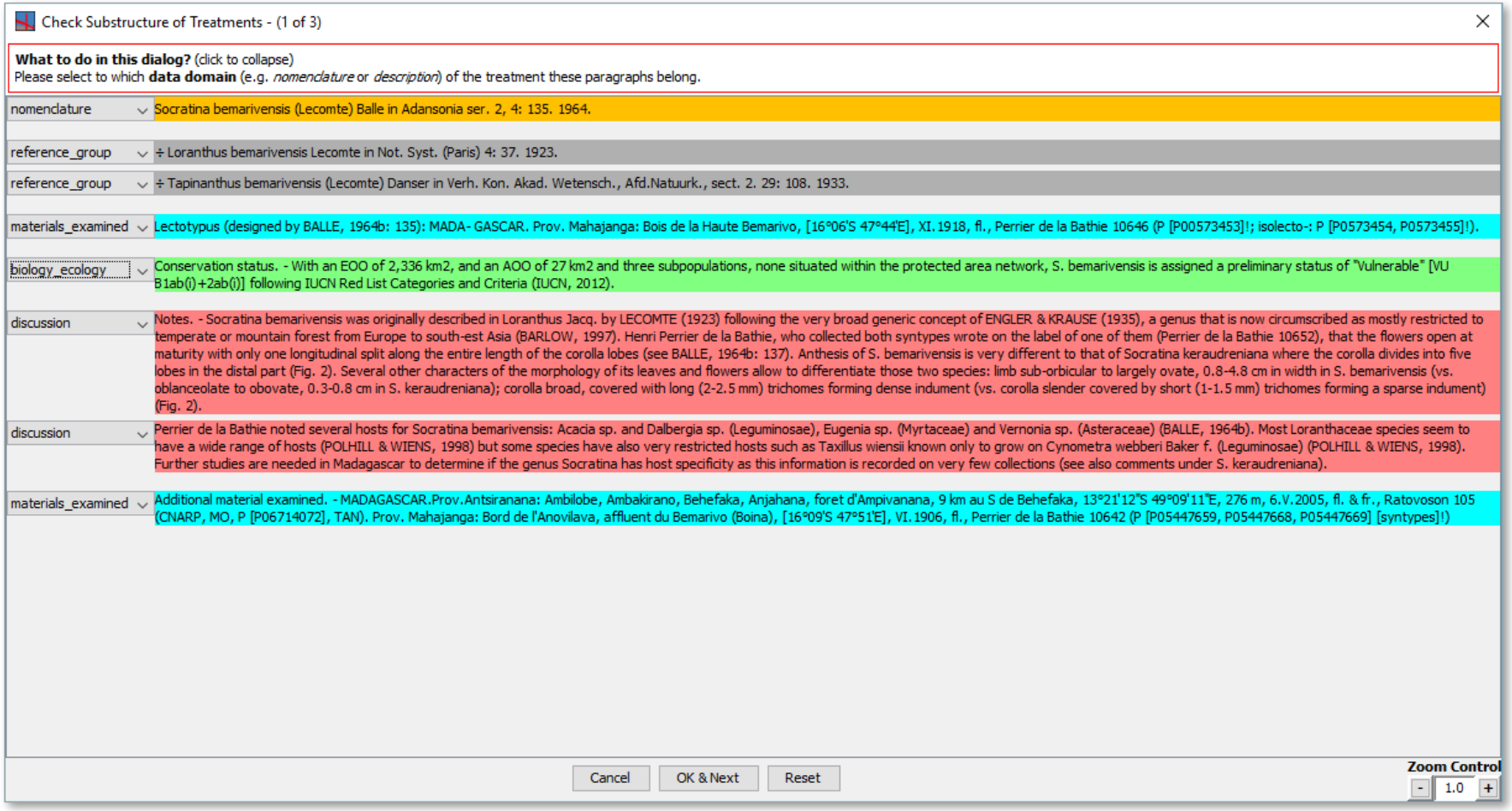
Retrieving higher ranks from *GBIF* and *Catalogue of Life* (external resources)



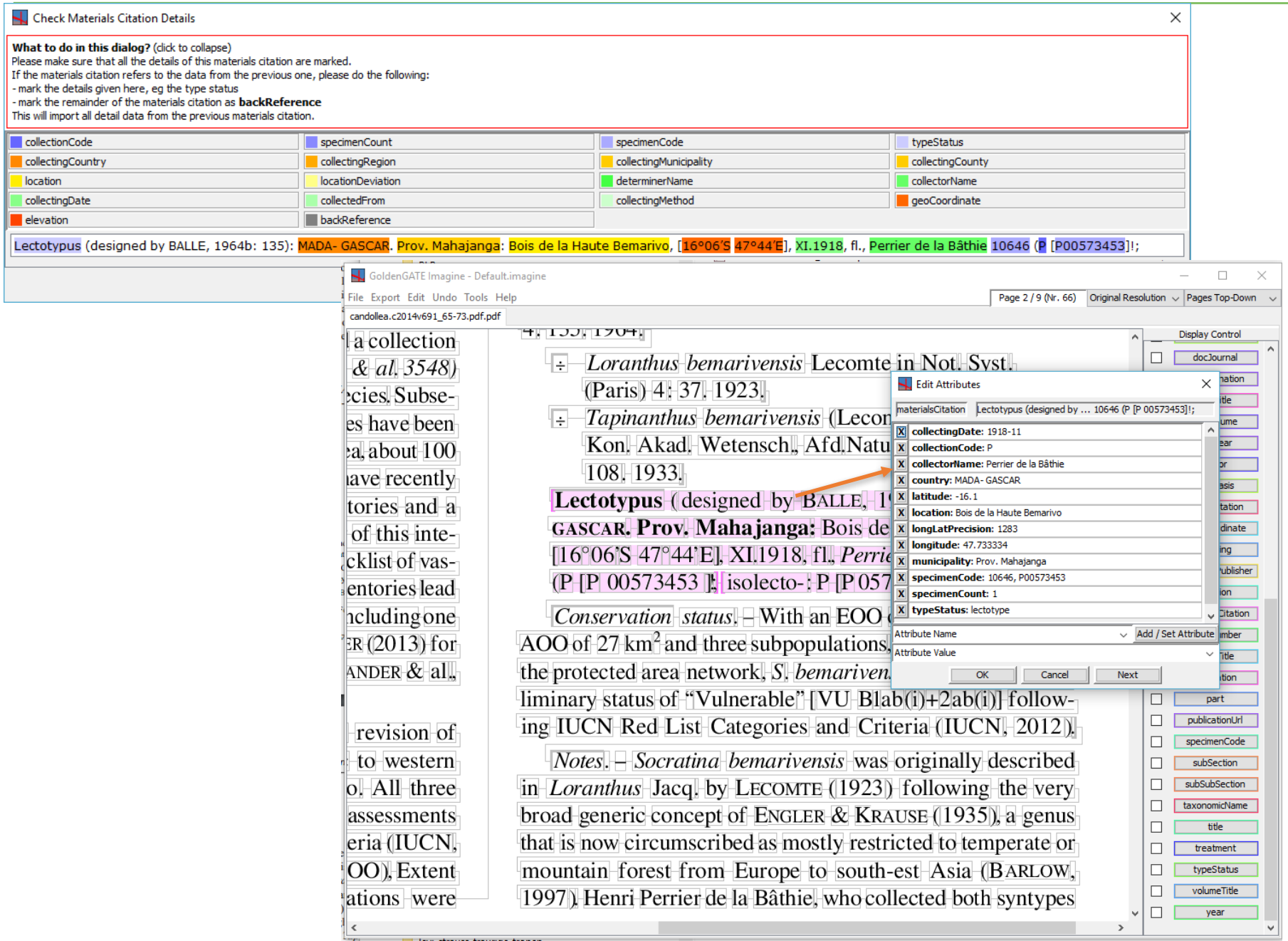
Plazi: conversion



Plazi: conversion



Plazi: conversion



Plazi infrastructure: Biodiversty Literature Repository

Access

Publication date

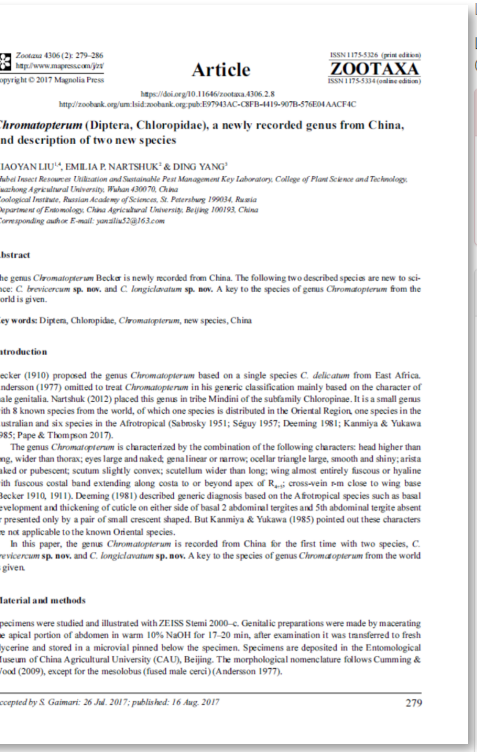
Links to other resources about the same thing

Links to related objects and source

Community(ies)

Licence

BLR: more than a PDF store: a enhanced document repository



Zootaxa 4306 (2): 279–286
http://www.mapress.com/j/zt/
Copyright © 2017 Magnolia Press

Article

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

https://doi.org/10.11646/zootaxa.4306.2.8
http://zoobank.org/urn:lsid:zoobank.org:pub:E97943AC-C8FB-4419-907B-576E04AACF4C

Chromatopterum (Diptera, Chloropidae), a newly recorded genus from China, and description of two new species

XIAOYAN LIU[1,4], EMILIA P. NARTSHUK[2] & DING YANG[3]

[1]Hubei Insect Resources Utilization and Sustainable Pest Management Key Laboratory, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
[2]Zoological Institute, Russian Academy of Sciences, St. Petersburg 199034, Russia
[3]Department of Entomology, China Agricultural University, Beijing 100193, China
[4]Corresponding author. E-mail: yanzilu52@163.com

Abstract

The genus *Chromatopterum* Becker is newly recorded from China. The following two described species are new to science: *C. brevicercum* **sp. nov.** and *C. longiclavatum* **sp. nov.** A key to the species of genus *Chromatopterum* from the world is given.

Key words: Diptera, Chloropidae, *Chromatopterum*, new species, China

Introduction

Becker (1910) proposed the genus *Chromatopterum* based on a single species *C. delicatum* from East Africa. Enderson (1977) omitted to treat *Chromatopterum* in his generic classification mainly based on the character of male genitalia. Nartshuk (2012) placed this genus in tribe Mindini of the subfamily Chloropinae. It is a small genus with 8 known species from the world, of which one species is distributed in the Oriental Region, one species in the Australian and six species in the Afrotropical (Sabrosky 1951; Séguy 1957; Deeming 1981; Kanmiya & Yukawa 1985; Pape & Thompson 2017).

The genus *Chromatopterum* is characterized by the combination of the following characters: head higher than long, wider than thorax; eyes large and naked; gena linear or narrow; ocellar triangle large, smooth and shiny; arista naked or pubescent; scutum slightly convex; scutellum wider than long; wing almost entirely fuscous or hyaline with fuscous costal band extending along costa to or beyond apex of R_{4+5}; cross-vein r-m close to wing base (Becker 1910, 1911). Deeming (1981) described generic diagnosis based on the Afrotropical species such as basal development and thickening of cuticle on either side of basal 2 abdominal tergites and 5th abdominal tergite absent or presented only by a pair of small crescent shaped. But Kanmiya & Yukawa (1985) pointed out these characters are not applicable to the known Oriental species.

In this paper, the genus *Chromatopterum* is recorded from China for the first time with two species, *C. brevicercum* **sp. nov.** and *C. longiclavatum* **sp. nov.** A key to the species of genus *Chromatopterum* from the world is given.

Material and methods

Specimens were studied and illustrated with ZEISS Stemi 2000–c. Genitalic preparations were made by macerating the apical portion of abdomen in warm 10% NaOH for 17–20 min, after examination it was transferred to fresh glycerine and stored in a microvial pinned below the specimen. Specimens are deposited in the Entomological Museum of China Agricultural University (CAU), Beijing. The morphological nomenclature follows Cumming & Wood (2009), except for the mesolobus (fused male cerci) (Andersson 1977).

Accepted by S. Gaimari: 26 Jul. 2017; published: 16 Aug. 2017

zenodo

Search | Upload | Communities | Log in | Sign up

December 31, 2017

Journal article | Closed Access

Chromatopterum (Diptera, Chloropidae), a newly recorded genus from China, and description of two new species

Liu, Xiaoyan; Nartshuk, Emilia P.; Yang, Ding

Liu, Xiaoyan, Nartshuk, Emilia P., Yang, Ding (2017): Chromatopterum (Diptera, Chloropidae), a newly recorded genus from China, and description of two new species. Zootaxa 4306 (2): 279-286, DOI: https://doi.org/10.11646/zootaxa.4306.2.8

Files

Closed Access

Files are not publicly accessible.

References

- Andersson, H. (1977) Taxonomic and phylogenetic studies on Chloropidae (Diptera) with special reference to Old World genera. Entomologica Scandinavica, 8 (Supplement), 1-200.
- Becker, T. (1910) Chloropidae. Eine monographische Studie. II. Teil. Aethiopische Region. Annales Historico-Naturales Musei Nationalis Hungarici, 8, 377-443.
- Becker, T. (1911) Chloropidae. Eine monographische Studie. III. Teil. Die indo-australische Region. Annales Historico- Naturales Musei Nationalis Hungarici, 9, 35-170.
- Cumming, J.M. & Wood, D.M. (2009) Adult morphology and terminology. In: Brown, B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E. & Zumbado, M.A. (Eds.), Manual of Central American Diptera. Uol. 1. NRC Research Press, Ottawa, pp. 9-50.
- Deeming, J.C. (1981) A contribution to the study of Afrotropical Chloropine Diptera. Journal of Natural History, 15 (5), 789- 828.
- https://doi.org/10.1080/00222938100770581
- Kanmiya, K. & Yukawa, J. (1985) Chloropidae (Diptera) of Panaitan and the Krakatau Island, Indonesia. Kontytu, 53 (3), 461- 474.
- Nartshuk, E.P. (2012) A check list of the world genera of the family Chloropidae (Diptera, Cyclorrhapha, Muscomorpha). Zootaxa, 3267, 1-43.
- Pape, T. & Thompson, F.C. (2017) Systema Dipterorum, version 2.0. Available from: http://www.catalogueoflife.org/col. (accessed 22 March 2017)
- Sabrosky, C.W. (1951) Chloropidae. Ruwenzori Expedition 1934-35. Britsh Museum (Natural History), 2(7), 711-828.
- Seguy, E. (1957) Chloropides africains du Museum (Dipt.). Revue francaise d'Entomologie, 24, 264-277.

Part of

Biodiversity Literature Repository

Publication date:
December 31, 2017

DOI:
DOI 10.11646/zootaxa.4306.2.8

Keyword(s):
Biodiversity | Taxonomy | Animalia | Arthropoda | Insecta | Diptera | Chloropidae

Published in:
Zootaxa: 4306 pp. 279-286.

Related identifiers:
Cites:
10.1080/00222938100770581

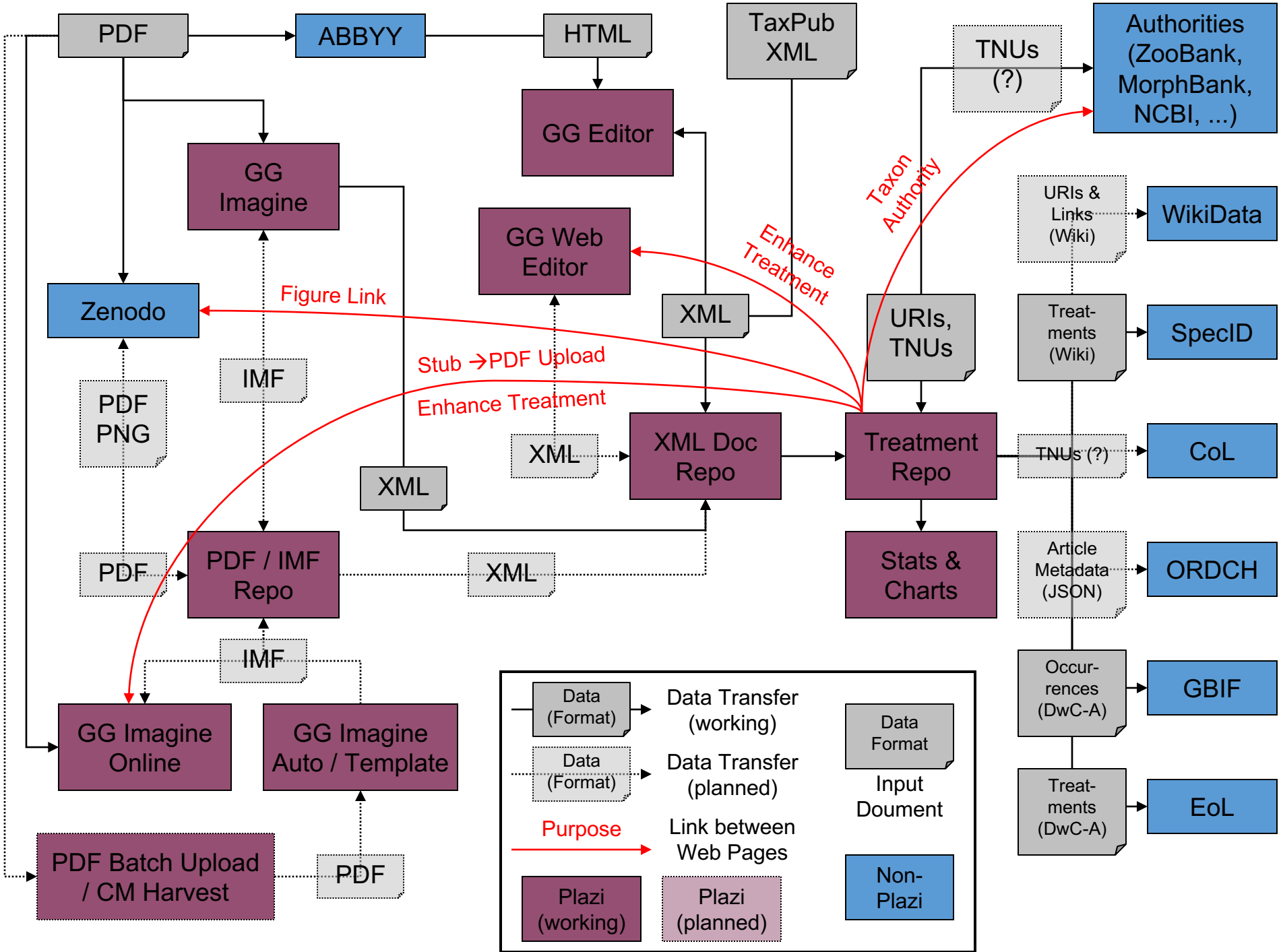
Has part:
http://treatment.plazi.org/id/039687A2FF7EFFC7FF35C69A74C8FC02,
http://treatment.plazi.org/id/039687A2FF7EFFC2FF35C2757061FDF8,
http://treatment.plazi.org/id/039687A2FF7BFFC3FF35C42C7061FD10, 10.5281/zenodo.843826, *urn:lsid:plazi.org:fig:2-378-1209-397-1420* (LSID),
http://zenodo.org/record/843826,
10.5281/zenodo.843828, *urn:lsid:plazi.org:fig:3-297-1254-183-1260* (LSID),
http://zenodo.org/record/843828,
10.5281/zenodo.843830, *urn:lsid:plazi.org:fig:3-360-1226-1422-1947* (LSID),
http://zenodo.org/record/843830,
10.5281/zenodo.843832, *urn:lsid:plazi.org:fig:5-355-1231-839-1955* (LSID),
http://zenodo.org/record/843832,
10.5281/zenodo.843834, *urn:lsid:plazi.org:fig:6-347-1236-202-1167* (LSID),
http://zenodo.org/record/843834,
10.5281/zenodo.843836, *urn:lsid:plazi.org:fig:6-354-1234-1310-1889* (LSID),
http://zenodo.org/record/843836

Alternate identifiers:
urn:lsid:zoobank.org:pub:E97943AC-C8FB-4419-

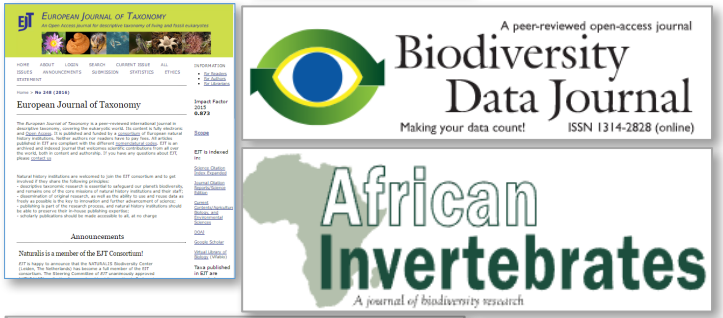
From unstructured to structured, citable

From human to machine readable

Human readable → machine readable


PDF
ABBYY
HTML
TaxPub XML
TNUs (?)
Authorities (ZooBank, MorphBank, NCBI, ...)
GG Editor
GG Imagine
GG Web Editor
Taxon Authority
URIs & Links (Wiki)
WikiData
Enhance Treatment
Figure Link
Zenodo
XML
URIs, TNUs
Treat-ments (Wiki)
SpecID
Stub →PDF Upload
IMF
PDF PNG
Enhance Treatment
XML
XML Doc Repo
Treatment Repo
TNUs (?)
CoL
XML
PDF
PDF / IMF Repo
XML
Stats & Charts
Article Metadata (JSON)
ORDCH
IMF
Occur-rences (DwC-A)
GBIF
GG Imagine Online
GG Imagine Auto / Template
Treat-ments (DwC-A)
EoL
PDF Batch Upload / CM Harvest
PDF
Data (Format)
Data Transfer (working)
Data Format
Data (Format)
Data Transfer (planned)
Input Doument
Purpose
Link between Web Pages
Plazi (working)
Plazi (planned)
Non-Plazi

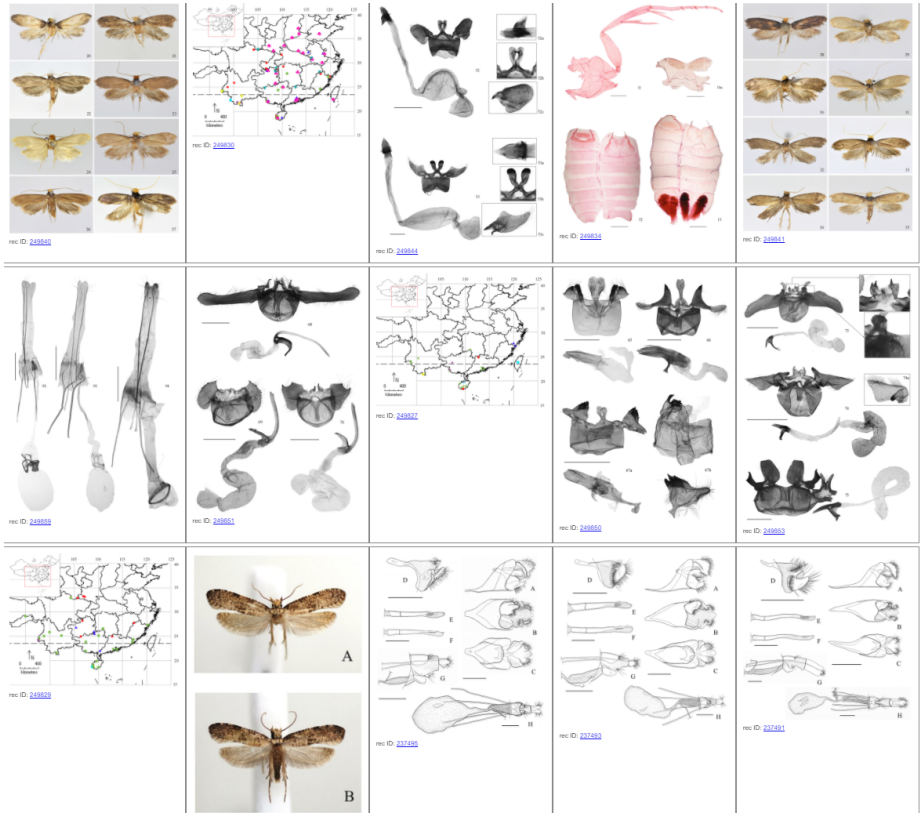
Plazi text and data mining output 2017

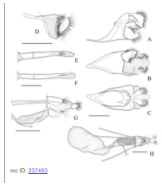


Journals	132
Articles	8,545
Taxonomic treatments	69,539
New species	6,387
Observation records	45,381
Figures	56,692
Bibliographic references	267,728
Facts	>>10 M

The value of access to subarticle elements: figures

the value of a figure





access

http://ocellus.punkish.org/?q=tineidae&page=1

Conversion and publishing services: costs

Is it worthwhile – can we afford - to semantically enhance unstructured publications?

Copyright might be one reason

Plazi workflow is based on Swiss copyright

- Copying of excerpts for internal use is a permitted act
- Individual articles are considered excerpts of journals
- Distribution of copies in institutions is a permitted act
- Data (text, figures, images) ≠ copyright protected work

→ Standortvorteil for Swiss science

(Agosti&Egloff, 2009).

Improved science communication

TDM comes at a high cost, that ought to be omitted to discover data objects in text.

This can be done using semantically enhanced publishing.

An example is Pensoft's use of the Journal Archival Tag Suit (JATS) extension Taxpub which includs all the semantic elements describing the data in taxonomy.

Pensoft publications are automatically imported, without a TDM step, into TreatmentBank.

JATS based publications can automatically be imported in PubMed Central

Removing the PDF prison

Highly **structured** research data

publishing

Semantically **unstructured** text

data mining

Semantically enhanced, **structured** data

Highly **structured** research data

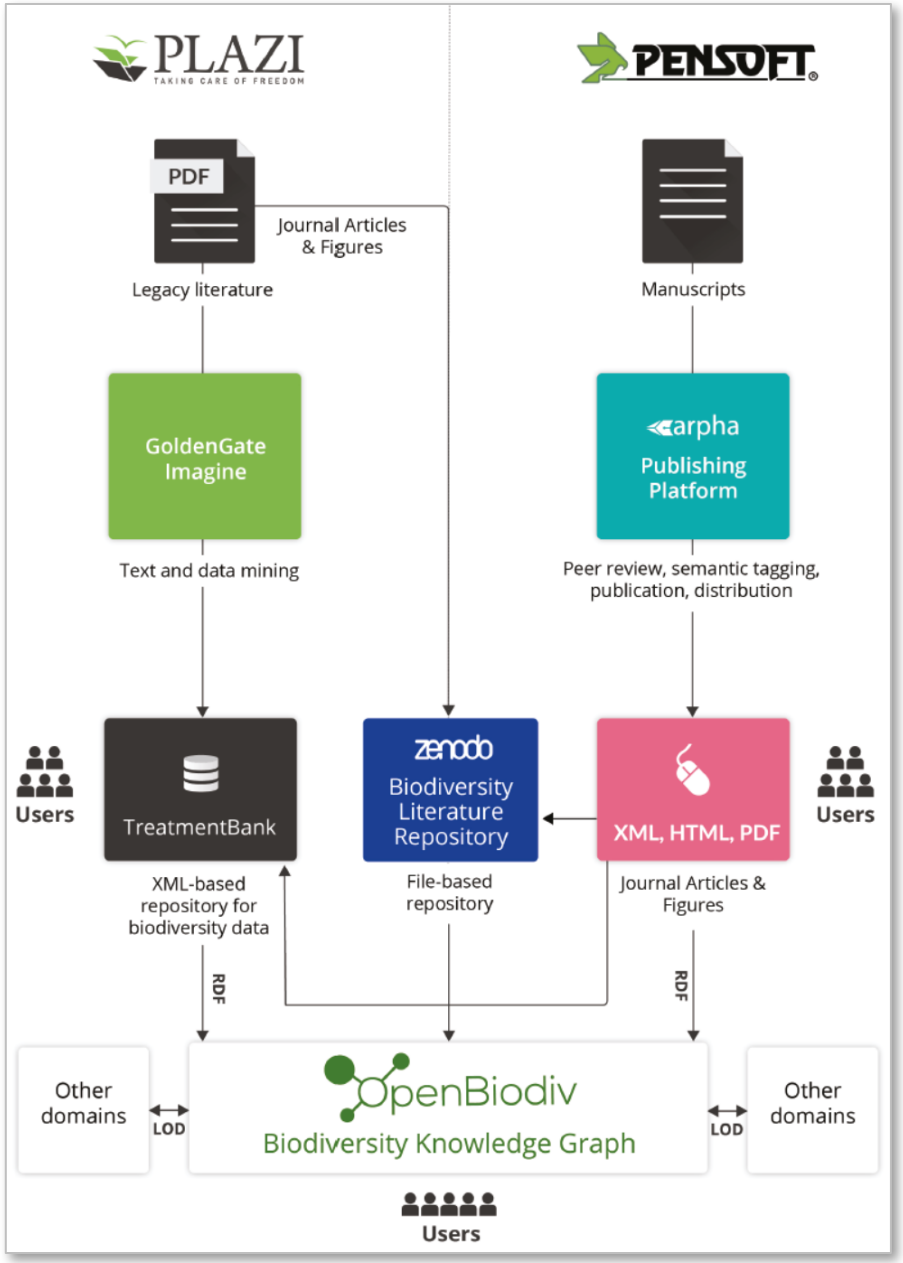
publishing

Semantically enhanced data

conversion

Semantically enhanced, **structured** data

The future will be open, semantically enhanced publishing



PLAZI
TAKING CARE OF FREEDOM

Biodiversity
Literature
Repository

zenodo

CERN

Thank you!

Donat Agosti

agosti@plazi.org