



## **Willi Egloff and Donat Agosti**

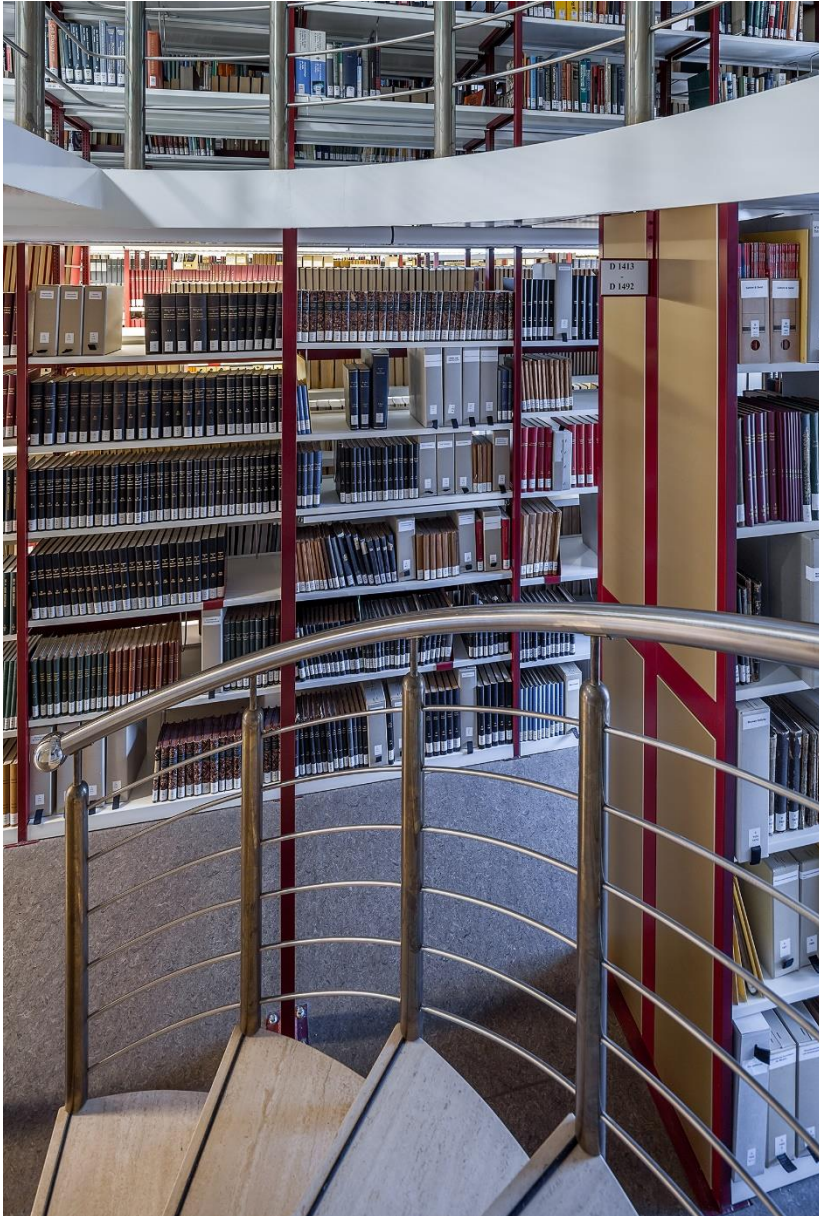
September 26, 2016, International Congress of Entomology, Orlando  
Symposium 156, #0672



# **Legal and Ethical Reasons for Free and Open Biodiversity Information**





**Willi Egloff & Donat Agosti  
Plazi, Bern**

(<http://plazi.org>)



500,000,000+ printed pages  
5,000 journals with taxonomic content  
1,900,000 species described  
20,000,000+ species treatments

***BUT: The facts are hidden***

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

**Plazi solution:  
Linking through  
taxonomic treatments**



## Taxonomic Treatment

Formica obsoleta Linnaeus, 1758: 580

name

obsoleta. ♀. F. supra nigra, subtus testaceo rufa, abdomine subglobofo.

*Habitat in Europæ terra.*

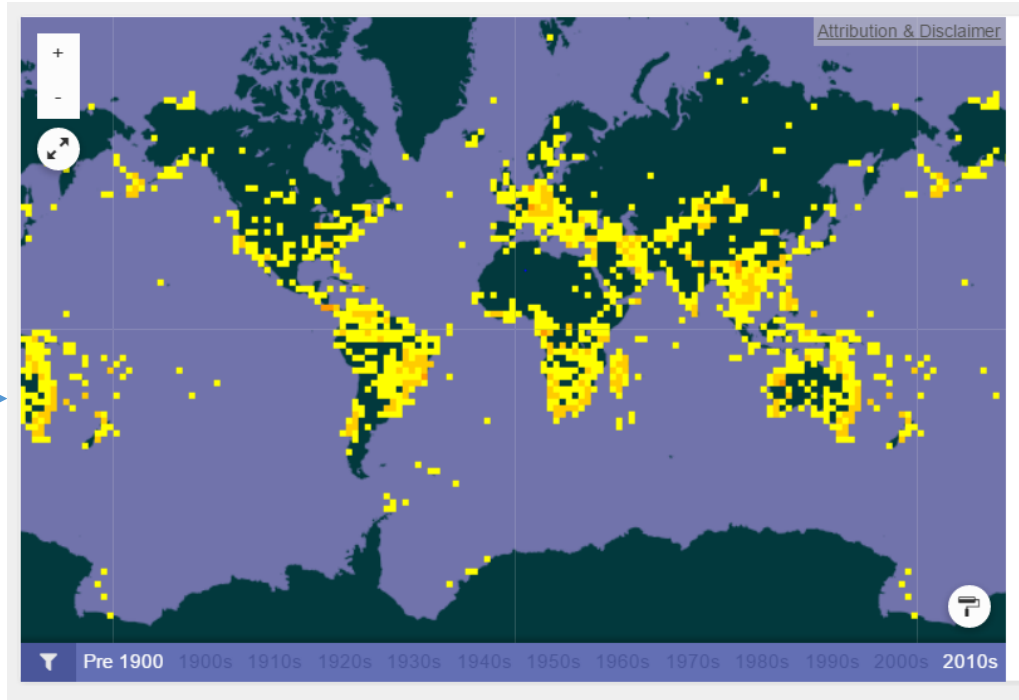
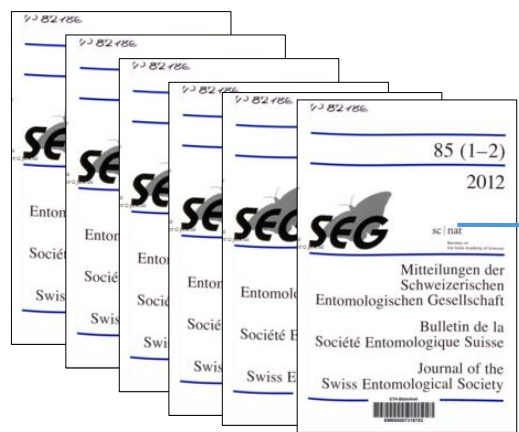
description

distribution

Linnaeus has to be credited for Latin Binomen AND Treatment.



# Plazi TreatmentBank: Mining and visualisation from treatments



## 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	<a href="http://dx.doi.org/10.5281/zenodo.35388">http://dx.doi.org/10.5281/zenodo.35388</a>
link to original document	<a href="http://dx.doi.org/10.5281/zenodo.35388">http://dx.doi.org/10.5281/zenodo.35388</a>
document provided by	Donat (2015-12-11 15:22:31, last updated by Admin 2015-12-18 12:58:08)

TreatmentBank Specimen Occurrence Map - Google Chrome

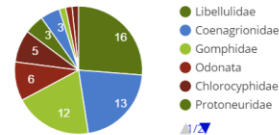
tb.plazi.org/GgServer/resources/srsWebPortalData/Linkers/GoogleMapsL

Holotype, Lectotype, Neotype: Paratype, etc.: Holotype & 5 Paratypes:

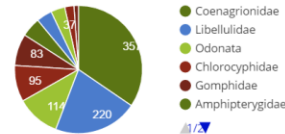
### Distribution Map

#### Specimens

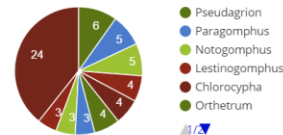
##### Article: Treatments by Family (n=61)



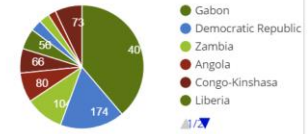
##### Article: Specimens by Family (n=1034)



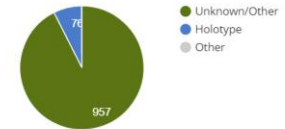
##### Article: Treatments by Genus (n=61)



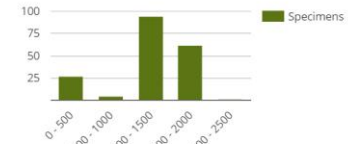
##### Article: Specimens by Country (n=1034)



##### Article: Specimens by Type Status (n=1034)



##### Article: Specimens by Elevation (n=1034)





## Plazi statistics:

### Data

- **Publications:** 17,797 (2016: 3,500)
- **Taxonomic treatments:** 165,692 (2016: 50,000)
- **Facts extracted:** > 20 Millions (2016: 8 Millions)
- **Taxonomic names:** 152,641 (2016: 35,000)

**Coverage of >25% of new described species in science** (ca 17,000 treatments annually)

**In preparation:** Backstock extraction of 1 Million treatments by end of 2017



## Mission

- Maintain a digital taxonomic literature repository to enable archiving of taxonomic treatments.
- Enhance published taxonomic treatments, extract the facts and make them widely accessible.
- Participate in the development of new models for publishing scientific articles in order to maximize interoperability with other relevant cyberinfrastructure components (e.g., name servers, biodiversity resources, etc...)





## A virtual, global organization with real people behind



**2002** first projects  
**2008** Plazi Verein  
**2012** Plazi GmbH



## Ethical principles

- *Bouchout Declaration for Open Biodiversity Knowledge Management* (June 2014);
- *RDA/CODATA Principles on the Legal Interoperability of Research Data* (September 2016);
- *GEO Data Sharing Principles* (2005 and 2015):
  - Data and metadata shall be shared as Open Data by default, without charge or restrictions on reuse;
  - Where international instruments, national policies or legislation preclude the sharing of data as Open Data, data should be made available with minimal restrictions on use;
  - Data and metadata shall be made available with minimum time delay.



## Requirements

- Access to all names that have ever been used for all taxa.
- Access to all published taxonomic treatments.
- Access to all metadata of published taxonomic literature.
- Possibility to extract and re-use data on facts (text, pictures, sounds) and metadata from all legacy literature and other biodiversity information.



## **Legal Obstacles**

**(„legislation that precludes the sharing of data as Open Data“)**

### **Restrictions based on public interests:**

- National security or public order laws
- Protection of endangered species

### **Restrictions based on private interests:**

- Copyright and other intellectual property laws
- EU-database right
- Protection of personal privacy
- Time embargoes on the release of research data



## **Plazi solution for eliminating legal obstacles:**

### **Workflow restricted to published data**

- no restrictions based on public interests;
- no restrictions based on privacy protection;
- no restrictions based on time embargoes.

### **Workflow situated outside the E.U.**

- E.U. Database protection does not apply



## **Restrictions based on copyright are often lacking legal grounds**

Copyright protects works, not data (no copyright protection by individual attribution [©-sign, CC-license])

### **„Works“ in Copyright**

- intellectual creation
- individual (new, original) with respect of the form of expression (= no standardization)

### **Permitted uses**

Exceptions and limitations / fair use



- **Names and treatments are, by default, not protected by IPR**  
*(Patterson et al., BMC Research Notes 2014, 7:79)*
- **Copyright is no legal obstacle to re-use biodiversity data**
- **Uncertainty with respect to illustrations**  
*(publication in preparation)*



## **Remaining obstacle:**

- Access to legacy literature
- Extraction of data and metadata from copyrighted texts
- Re-use of copyrightable images

## **Ways to remove obstacles:**

- Open access publishing
- Revision of copyright law in order to provide for exceptions and limitations that allow text and data mining
- Revision of E.U. Database protection in order to provide for exceptions and limitations that allow text and data mining.





## **Plazi solution:**

- Extraction of data and metadata based on copyright exceptions
- Making available of extracted data and metadata, but not of the copyrighted works (including copyrighted images)



**Thank you for your attention!**

**Willi Egloff (egloff@plazi.org)**

**Donat Agosti (agosti@plazi.org)**