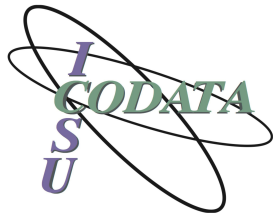


One Repository, Two Implementations, & A World of Legal Interoperability Opportunities and Challenges

Donat Agosti, Gail Clement, Willi Egloff, and Tom Morrell



RDA 9th Plenary, Barcelona, Spain. April 7, 2017 DOI: 10.5281/zenodo.439627



Legal Interoperability of Research Interest Group



Session Outline

- Introduction to Invenio repository platform
- Caltech Data landscape
- Plazi Data landscape
- Comparison of Legal Interop Issues & Challenges
- Discussion



1

One Repository: Invenio



Digital Library Framework

<http://invenio-software.org/> | <http://opendata.cern.ch/>

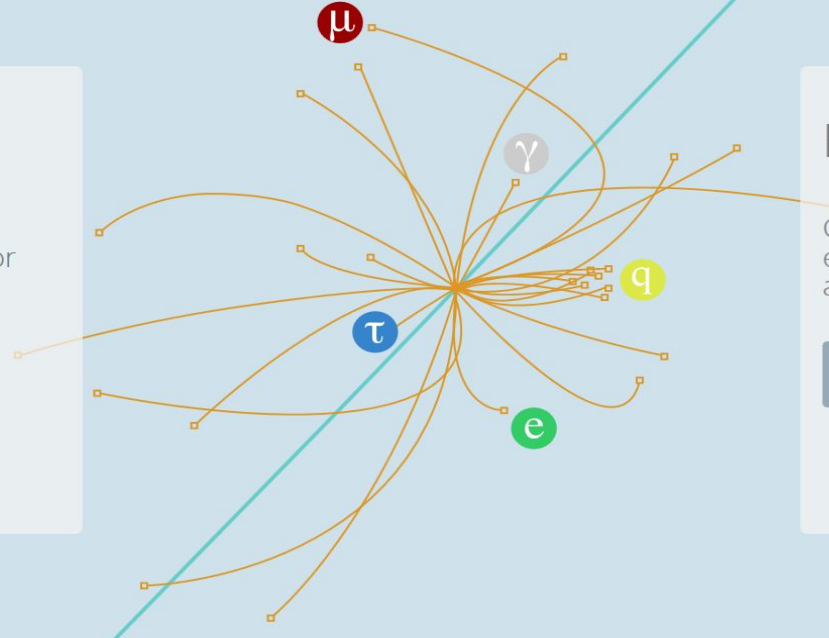
opendata
CERN

ABOUT SEARCH EDUCATION RESEARCH

Education

Visualise events, check reconstructed data, run tools or build your own!

Start learning



Research

Get the genuine working environments, virtual machines and datasets to start your research

Start analysing



<http://inveniosoftware.org/>



Flexible data model

Use JSON Schema to describe articles, books, photos, videos, data, and software. Several popular master metadata formats are supported, such as MARC21 with BibTeX, DataCite, Dublin Core, EndNote, RefWorks.



Configurable workflows

Organise document corpus in community collections. Configure user and robot ingestion workflows. Attribute community moderators.



Extensible packages

Invenio is composed of hundreds of independent pluggable packages that collaborate via rich APIs. Pick the packages you need and use the full power of Python to extend their capabilities.



Powerful search engine

Very fast response for repositories of up to several million records. Customisable query language and second-order search operators. Configurable UI and facets. Combined metadata, fulltext and reference search in one go. Citation networks.



Collaborative communities

Organise users in groups and teams. Share documents of interest in annotable baskets. Configure automated email and RSS notification alerts.



Open standards

We love open access, open source, and open standards. DOI, JSON Schema, Memento, OAI-PMH, ORCID, OpenAIRE, REST, XML... you name it.



Invenio Repository

- Free open source digital repository software
- Harvest datasets, analysis code, virtual machine environment, configuration and knowledge information.
- Visualise data in the browser.
- Rerun preserved code on the cloud.



<http://inveniosoftware.org/>

The DataCite Metadata Schema is a list of core metadata properties chosen for the accurate and consistent identification of a resource for citation and retrieval purposes, along with recommended use instructions.

The screenshot shows the GitHub repository page for `inveniosoftware/datacite`. The repository is a Python API wrapper for the DataCite Metadata Store API, with a link to <http://datacite.readthedocs.org>. The repository has 46 commits, 1 branch, 5 releases, and 7 contributors. The current branch is `master`. The file list includes:

File Name	Description	Last Commit
<code>datacite</code>	version: post-release bump	4 months ago
<code>docs</code>	docs: README badges fix	4 months ago
<code>tests</code>	global: datacite schema v4.0 support	4 months ago
<code>.coveragerc</code>	datacite: initial package release	2 years ago
<code>.editorconfig</code>	global: package alignment	a year ago
<code>.gitignore</code>	datacite: initial package release	2 years ago
<code>.lgtn</code>	docs: MAINTAINERS file and LGTM configuration	6 months ago
<code>.travis.yml</code>	release: v0.2.0	a year ago
<code>AUTHORS.rst</code>	global: datacite schema v4.0 support	4 months ago

Table 1: DataCite Mandatory Properties

<i>ID</i>	<i>Property</i>	<i>Obligation</i>
1	Identifier (with type sub-property)	M
2	Creator (with name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	PublicationYear	M

Table 2: DataCite Recommended and Optional Properties

ID	Property	Obligation
6	Subject (with type and term sub-properties)	R
7	Contributor (with type, name identifier, and affiliation sub-properties)	R
8	Date (with type sub-property)	R
9	Language	O
10	ResourceType (with general type description sub-property)	R
11	AlternateIdentifier (with type sub-property)	O
12	RelatedIdentifier (with type and relation type sub-properties)	R
13	Size	O
14	Format	O
15	Version	O
16	Rights	O
17	Description (with type sub-property)	R
18	GeoLocation (with point and box sub-properties)	R

ID	DataCite-Property	Occ	Definition	Allowed values, examples, other constraints
7.1	contributorType	1	The type of contributor of the resource.	<p>If Contributor is used, then contributorType is mandatory.</p> <p>Controlled List Values: ContactPerson DataCollector DataCurator DataManager Distributor Editor Funder HostingInstitution Producer ProjectLeader ProjectManager ProjectMember RegistrationAgency RegistrationAuthority RelatedPerson Researcher</p> <p>RightsHolder Sponsor Supervisor</p>

ID	DataCite-Property	Occ	Definition	Allowed values, examples, other constraints
16	Rights	0-n	Any rights information for this resource.	<p>Free text. ***</p> <p>Provide a rights management statement for the resource or reference a service providing such information. Include embargo information if applicable.</p> <p>Use the complete title of a license and include version information if applicable.</p> <p>Example: Creative Commons Attribution 3.0 Germany License</p>

Legal Interop support in Invenio

Contributors

Contributor

Contributor Name

Contributor Affiliation

Contributor Identifier

Identifier Type

Contributor Role

Contributor Email

Embargo

Embargoed Status

Embargoed Access Open Access

Embargo Date

License/Walver

License/Walver Name *

- Apache License, 2.0
- BSD 3-Clause "New" or "Revised" License (BSD-3-Clause)
- Caltech Research Data Repository License
- Creative Commons Attribution
- Creative Commons Attribution-NonCommercial
- Creative Commons CCZero**
- GNU General Public License (GPL)
- GNU Lesser General Public License (LGPL)

Related Identifier(s)

Related Identifier(s) +

This CaltechDATA Record

Related Identifier

Identifier Type

2

Two Implementations Caltech & Plazi

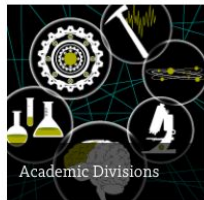


Inventing Tools for Detecting Life Elsewhere

At Caltech, in the Exoplanet Technology Laboratory, or ET Lab, researchers have been developing new ways to search exoplanets for biosignatures.



Discover More



Academic Divisions



Research



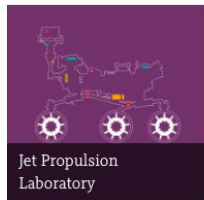
Admissions



Leadership



The Caltech Campaign



Jet Propulsion Laboratory



Campus Life



Library



Major Research Institutes



Caltech Today

Caltech is small but prizes excellence and ambition.



The contributions of Caltech's faculty and alumni have earned national and international recognition, including 35 Nobel Prizes.

The Institute manages the Jet Propulsion Laboratory (JPL) for NASA, sending probes to explore the planets of our solar system and quantify changes on our home planet; owns and operates large-scale research facilities such as the Seismological Laboratory and a global network of astronomical observatories, including the Palomar and W. M. Keck Observatories; and cofounded and comanages LIGO, which, in 2016, observed gravitational waves for the first time.

Getting Started

Tutorials

Data

Events

Bulk Data

Timelines

My Sources

Software

GPS ↔ UTC

About LIGO

Data Analysis
Projects

Acknowledgement

Data release for event GW150914

This page has been prepared by the LIGO Scientific Collaboration (LSC) and the Virgo Collaboration to inform the broader community about a confirmed astrophysical event observed by the gravitational-wave detectors, and to make the data around that time available for others to analyze. There is also a [technical details](#) page about the data linked below, and feel free to [contact us](#). This dataset has the Digital Object Identifier (doi) <http://dx.doi.org/10.7935/K5MW2F23>

Summary of Observation

The event occurred at GPS time 1126259462 == September 14 2015, 09:50:45 UTC. The false alarm rate is estimated to be less than 1 event per **203,000 years**, equivalent to a significance of **5.1 sigma**. The event was detected in data from the [LIGO Hanford](#) and [LIGO Livingston](#) observatories.

- There are [Science Summaries](#), covering the information below in ordinary language.
- There is a [one page factsheet about GW150914](#), summarizing the event.

How to Use this Page

- **Click on the section headings below to show available data files.**
 - [\(click to Open/Close all sections\)](#)
- There are lots of data files available in the sections below, look for the word **DATA**.
- Click on each thumbnail image for larger image.
- See the papers linked below for full information, references, and meaning.
- Many of the data files linked below have heterogeneous formatting; if you have any questions, please [contact us](#).

The G150914 detection paper:

Observation of Gravitational Waves from a Binary Black Hole Merger

For full details see [LIGO DCC](#), [arXiv](#), or [Phys. Rev. Letters](#)
This paper and all the companion papers can also be found at papers.ligo.org.

Estimated source parameters

Quantity	Value	Upper/Lower error estimate	Unit
Primary black hole mass	36	+5 -4	M sun
Secondary black hole mass	29	+4 -4	M sun
Final black hole mass	62	+4 -4	M sun
Final black hole spin	0.67	+0.05 -0.07	
Luminosity distance	410	+160 -180	Mpc
Source redshift, z	0.09	+0.03 -0.04	
Energy radiated	3	+0.5 -0.5	M sun

TABLE I. Estimated source parameters for GW150914. We report the median value as well as the range of the 90% credible interval. Masses are measured in the source frame; to convert masses to detector frame, multiply by $(1 + z)$. The source redshift assumes standard cosmology.



California Institute of Technology

Research Data Repository

 Search 11 records 



Data Sets



Software



Submit a Data Set

[Terms >](#)

[Managed by Caltech Library](#) [Terms](#) [Report a Problem](#) [Contact](#)

Powered by



data.caltech.edu

Long-tail datasets in CaltechData



Identifying and Quantifying Mineral Abundance through VSWIR Microimaging Spectroscopy: A Comparison to XRD and SEM

- Dataset
- 2017-03-13
- CaltechDATA

Graduate Student

Faculty | JPL Researcher

Data copied from another server

US federal funder

US private funder

Data associated with published paper

Details

Authors
Leask, Ellen K. Division of Geological & Planetary Sciences, California Institute of Technology
Ehlmann, Bethany L. Division of Geological & Planetary Sciences, California Institute of Technology
0000-0002-2745-3240 ORCID

Description
Other:
Sample: A rock was collected from an travertine conglomerate, containing ophiolite clasts. The rock was collected in Oman near the Samail ophiolite (collected Jan 2012; B. Ehlmann; analyzed in Leask & Ehlmann, 2016). Data included are: 1) UCIS (Ultra-Compact Imaging Spectrometer; B. Van Gorp et al) data cube. Data near ends of sensor may be suspect [e.g. under 500 nm, over 2500 nm]. {om12L_001_UCIS_cube_masked} 2) EDS-SEM mosaic image cube (4x downsampled to despeckle, using a nearest neighbour algorithm). Acquired Caltech October 2015. Chemical data are in atomic %. {om12L_001_SEM_4x_masked} 3) EDS-SEM data warped to UCIS cube [degree 4 convolution, using built-in ENVI warping algorithm and based on control points]. Three UCIS images were mosaicked together to build the quantitative mineralogy Mars analogue petrology

DOI
10.22002/D1.222

Version
1.0

Related Identifier(s)
IsIdenticalTo (URL): ftp://ftp.gps.caltech.edu/pub/ehlmann/Leask_and_Ehlmann_2016_OM12001_hyperspectralcube.zip
ISupplementTo (URL): <http://resolver.caltech.edu/10.22002/D1.222#20758>

License
cc-zero

Funding
NASA: NNX12AB42G
JPL-RTD/PDF
Rose Hills Foundation

Language
eng

Related publications
Identifying and Quantifying Mineral Abundance through VSWIR Microimaging Spectroscopy: A Comparison to XRD and SEM
Leask, Ellen K., Division of Geological & Planetary Sciences, California Institute of Technology
Ehlmann, Bethany L., Division of Geological & Planetary Sciences, California Institute of Technology



Identifying and Quantifying Mineral Abundance through VSWIR Microimaging Spectroscopy: A Comparison to XRD and SEM

- Dataset
- 2017-03-13
- CaltechDATA

Variety of data types

Filename	Size	Download
Macro_OM12L_001.JPG	1.1 MB	Download
om12L_001_SEM_4x_masked.hdr	717 Bytes	Download
OM12L_001_smoothface_2.JPG	898.7 kB	Download
om12L_001_SEM_4x_masked	17.4 MB	Download
OM12L_001_UCIS_cube_masked.hdr	3.2 kB	Download
OM12L_001_UCIS_cube_masked	194.4 MB	Download
OM12L_001_warp_SEM_to_UCIS.hdr	15.7 kB	Download
om12L_001_XRD.csv	191 Bytes	Download
read_me.txt	2.2 kB	Download
OM12L_001_warp_SEM_to_UCIS	313.7 MB	Download

Big Science datasets in CaltechData



TCCON data from Ascension Island (SH), Release GGG2014.R0

Dataset
2017-03-30
CaltechDATA



Download



Edit

Details

Authors	Feist, D. G. Max Planck Institute for Biogeochemistry, Jena (DE) 0000-0002-5890-6687 ORCID Arnold, S. G. Max Planck Institute for Biogeochemistry, Jena (DE) John, N. Ariane Tracking Station, Ascension Island (SH) Geibel, M. C. Stockholm University, Stockholm (SE) 0000-0002-7369-0781 ORCID
Contributors	Roehl, C. M. DataCurator California Institute of Technology, Pasadena, CA (US) 0000-0001-5383-8462 ORCID
Description	Abstract: The Total Carbon Column Observing Network (TCCON) is a network of ground-based Fourier Transform Spectrometers that record direct solar absorption spectra of the atmosphere in the near-infrared. From these spectra, accurate and precise column-averaged abundances of atmospheric constituents including CO ₂ , CH ₄ , N ₂ O, HF, CO, H ₂ O, and HDO, are retrieved. This data set contains observations from the TCCON station on Ascension Island.
Publication Date	2017-03-30
Subject(s)	"atmospheric trace gases", "CO ₂ ", "CH ₄ ", "CO", "N ₂ O", "column-averaged dry air mole fractions", "remote sensing", "FTIR spectroscopy", "TCCON"
DOI	10.14291/ttcon.ggg2014.ascension01.R0/1149285
Version	GGG2014.R0
Format	application/x-netcdf
Related Identifier(s)	IsPartOf (URL): http://ttcon.ornl.gov/ IsDocumentedBy (DOI): 10.14291/ttcon.ggg2014.documentation.R0/1221662 IsCitedBy (DOI): 10.5194/amt-9-683-2016 IsCitedBy (DOI): 10.1002/2015JD023389 IsCitedBy (DOI): 10.5194/acp-16-1653-2016 IsCitedBy (DOI): 10.5194/amt-9-3491-2016 IsCitedBy (DOI): 10.5194/amt-9-2381-2016 IsCitedBy (DOI): 10.1002/2015JD024157
Relevant Dates	Created: 2014-10-01 Updated: 2017-03-01 Collected: 2012-05-22
License	other
Funding	Max Planck Society Max Planck Institute for Biogeochemistry
Language	eng
Geographic coverage	Ariane Tracking Station (AC) -7.9165 -14.3325

Creators from outside Caltech/US

Contributor from Caltech

References Documentation

Data cited by other outputs

German funders



TCCON data from Ascension Island (SH), Release GGG2014R0

Dataset

Data documentation refers to external site for usage policy

Download

Filename	Size	
README.txt	12.1 KB	Download
ae20120522_20161221.public.nc	10.6 MB	Download

log in

TCCON - TCCON Data and Network Policies - Data Use Policy

Data Use Policy

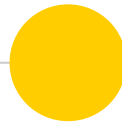
TCCON data can be downloaded from the TCCON archive¹, hosted by CDIAC² as per our Data Protocol. Please acknowledge the use of the data as described below. Citation files in BibTeX (or the concise version), RIS, and XML format has been created for your convenience that include all TCCON data citations. More detailed information on the data can be found on the Data Description page. Auxiliary data, including column averaging kernels and a priori profiles are contained in the netCDF files provided for GGG2014. Instructions on how to properly use the averaging kernels and a priori profiles for the TCCON data are on the Auxiliary Data wiki page.

For updates on our data, sign up to our data users list at https://lists.gps.caltech.edu/mailman/list/ttcon_users. Note that the website is self-signed. You can safely add an exception!

1. TCCON Acknowledgement
2. Data Protocol
 - 2.1. Use of Data
 - 2.2. Reciprocity
 - 2.3. Warnings
3. References and Contact Information
 - 3.1. Park Falls, Lamont, JPL, and Caltech
 - 3.2. Darwin and Wollongong
 - 3.3. Bremen and Hy Alesund
 - 3.4. Bialystok
 - 3.5. Orleans
 - 3.6. Lauder
 - 3.7. Tsukuba
 - 3.8. Garmisch
 - 3.9. Eureka

Σ Caltech Data

- = Institutional Repository using Invenio
- = World of Legal Interop challenges!





Search TreatmentBank



Hit enter to submit fulltext search

[Extended search](#)

News

07. March 2017

"Copyright and the Use of Images as Biodiversity Data" published

There is considerable uncertainty within the taxonomic community as to how to re-use images that were included in taxonomic publications, especially in regard to whether copyright applies. This article deals with the principles...[\[more\]](#)

Events

17. November 2016

April 7, 2017. 09:00-10:30 RDA 9th Plenary, Barcelona, Spain

On April 7, CalTech and Plazi will demonstrate their applications of the RDA/CODATA Legal Interoperability Principles in their repositories at CalTech and Biodiversity Literature Repository at Zenodo/CERN respectively. Both...[\[more\]](#)

Plazi is a small organization with broad expertise and track record

Life sciences

Legal affairs

Library sciences

Computer sciences

Semantic enhancement

Publishing

Organizing scientific meetings

ZooKeys 434: 109-120 (08 Jun 2014)
https://doi.org/10.3896/zooleaves.434.117.1

Open exchange of scientific knowledge and European copyright: The case of biodiversity information

Willi Egloff¹, David J. Patterson¹, Donat Agosti¹, Gregor Hegdorn^{1,2}

¹ Fraunhofer IPT, 32077 Berlin, Germany
² MaxPlanck Institute, 30559 Hannover, Germany

Received 28 April 2014 | Accepted 28 May 2014 | Published 08 June 2014

Abstract

BMC Research Notes

HOME ABOUT ARTICLES SUBMISSION GUIDELINES

CORRESPONDENCE OPEN ACCESS

Taxonomic information exchange and copyright: the Plazi approach

Download PDF Export citations

NCBI Resources How To

Bookshelf Books Browse Titles Limits Advanced Search

Journal Article Tag Suite Conference (JATS-Con) Proceedings 2010 [Internet].

Journal Article Tag Suite Conference (JATS-Con) Proceedings 2010 [Internet].

Search this book

TaxPub: An Extension of the NLM/NCBI Journal Publishing DTD for Taxonomic Descriptions

Terence Catapano.

Author Information

TaxPub is an extension of the NLM/NCBI Journal Publishing DTD for the description of biological specimens.

The Bouchout Declaration for Open Biodiversity Knowledge Management

The purpose of the Bouchout Declaration is to help make digital data about our biodiversity open members of the biodiversity community a way to demonstrate their commitment to open science

Biodiversity Literature Repository

Decision on the Biodiversity Literature Repository

As signatories of the Biodiversity Literature Repository, we agree to the following principles: [...]

read more

RIO

Research Notes and Open Access | v21592 (06 Mar 2017)
https://doi.org/10.3896/ris.21592.001

Copyright and the Use of Images as Biodiversity Data

David Patterson, Jeremy A. Miller

Specimen data in taxonomic literature are among the highest quality primary biodiversity data. Innovative cyber-taxonomic journals are using workflows that maintain data structure and disseminate electronic content to aggregators and other users; such structure is lost in traditional taxonomic publishing. Legacy taxonomic literature is a vast repository of knowledge about biodiversity. Currently, access to that resource is cumbersome, especially for non-specialist data consumers. Markup is a mechanism that makes this content more accessible, and is especially suited to machine analysis. Fine-grained XML (Extensible Markup Language) markup was applied to all 137 open access articles published in the Journal ZooKeys containing taxonomic treatments, either singly or in aggregate. The data can be filtered by several fields (including journal, taxon, institutional collection, country, collector, author, article and treatment) to query particular aspects of the data. We demonstrate here that XML markup using GoldenThread can address the challenge presented by unstructured legacy data, can extract structured primary biodiversity data which can be aggregated with and jointly queried with data from other Darwin Core-compatible sources, and show how visualization of these data can communicate key information contained in Biodiversity Literature. We complement recent studies on aspects of biodiversity knowledge using XML structured data to explore 1) the time lag between species discovery and description, and 2) the prevalence of rarely in species descriptions.

Keywords: Araneae, Biodiversity informatics, Data mining, Open access, Spiders, Taxonomy, XML, markup

Introduction

Biodiversity Data Journal

Articles About

Abstract Permalink Books EBooks Blog Journals

Integrating and visualizing primary data from prospective and legacy taxonomic literature

Jeremy A. Miller, Donat Agosti, Ludovico Pavesi, Guido Sauer, Todor Georgiev, Terry Catapano, David Patterson, David King, Serrano Peters, Rutger Aldo Vis, Soraya Sierra

Abstract

Specimen data in taxonomic literature are among the highest quality primary biodiversity data. Innovative cyber-taxonomic journals are using workflows that maintain data structure and disseminate electronic content to aggregators and other users; such structure is lost in traditional taxonomic publishing. Legacy taxonomic literature is a vast repository of knowledge about biodiversity. Currently, access to that resource is cumbersome, especially for non-specialist data consumers. Markup is a mechanism that makes this content more accessible, and is especially suited to machine analysis. Fine-grained XML (Extensible Markup Language) markup was applied to all 137 open access articles published in the Journal ZooKeys containing taxonomic treatments, either singly or in aggregate. The data can be filtered by several fields (including journal, taxon, institutional collection, country, collector, author, article and treatment) to query particular aspects of the data. We demonstrate here that XML markup using GoldenThread can address the challenge presented by unstructured legacy data, can extract structured primary biodiversity data which can be aggregated with and jointly queried with data from other Darwin Core-compatible sources, and show how visualization of these data can communicate key information contained in Biodiversity Literature. We complement recent studies on aspects of biodiversity knowledge using XML structured data to explore 1) the time lag between species discovery and description, and 2) the prevalence of rarely in species descriptions.

Keywords: Araneae, Biodiversity informatics, Data mining, Open access, Spiders, Taxonomy, XML, markup

Introduction

TreatmentBank

What is TreatmentBank?

Scientists describe and communicate the discovery of new biological species with taxonomic treatments that are bound to the names used to refer to these taxa. Often they are very rich in content and increasingly linked to external resources. TreatmentBank is a resource that stores and provides access to the treatments and data therein.

Search TreatmentBank

TreatmentBank can be searched using fulltext search or the extended search allowing more precise searches related to taxonomic names, biogeographic records or observation records.

read more

Desktop data mining and extraction Visualization

Biodiversity Literature Repository

Organizers



Community: 22 contributors
DOI recipients: 4 journals
Number of journals: 960

Goal: enriching publications and included data by making it findable, accessible, interoperable, and re-usable

The screenshot shows the Zenodo interface for the Biodiversity Literature Repository. At the top, there is a blue header with the Zenodo logo, a search bar, and navigation links for 'Upload' and 'Communities'. The repository name 'Biodiversity Literature Repository' is displayed below the header. A 'Recent uploads' section features a search bar and three entries. Each entry includes a date (December 31, 2017), a category (Journal article or Figure), and an access status (Closed Access or Open Access). The first entry is a journal article titled 'Scapanoclypeus bicoloratus new species from Hardap, Namibia (Coleoptera: Scarabaeidae: Melolonthinae: Tanyproctini)' by Richard Sehnaal, with a DOI and upload date. The second entry is a figure titled 'FIGURE 3 in Scapanoclypeus bicoloratus new species from Hardap, Namibia (Coleoptera: Scarabaeidae: Melolonthinae: Tanyproctini)' by Sehnaal, Richard, with a detailed description of the figure's content and an upload date. The third entry is another figure titled 'FIGURES 2 A - E. Scapanoclypeus carinatus Evans, male. A in Scapanoclypeus bicoloratus new species from Hardap, Namibia (Coleoptera: Scarabaeidae: Melolonthinae: Tanyproctini)' by Sehnaal, Richard, with a detailed description of the figures and an upload date. On the right side, there is a community sidebar with the Biodiversity Literature Repository logo, a description of the community's goal, a list of four principles (open access, DOI, citation, and search), and a 'Curated by: plazi-admin' section with a 'Curation policy' list of four points.

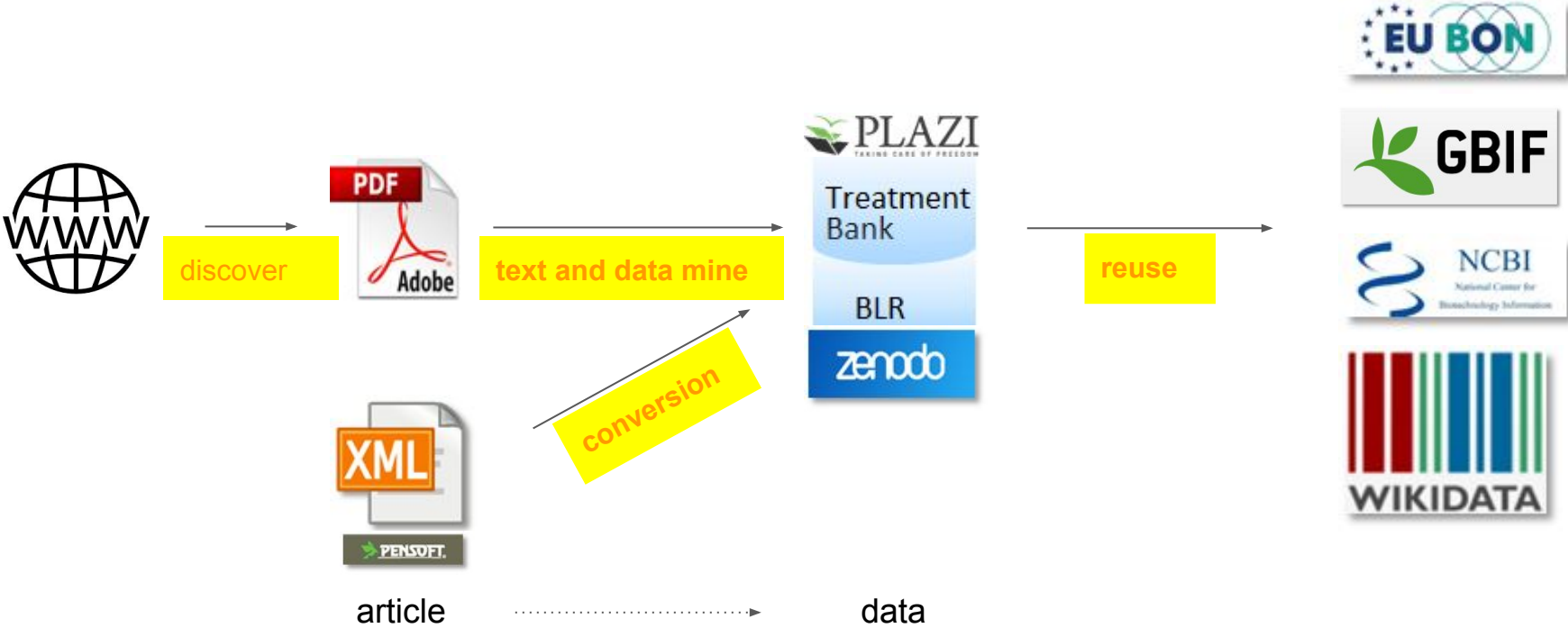
<http://biolitrepo.org>

BLR: Long tail biodiversity data at Zenodo

	Zenodo	BLR	BLR contribution
Images total	108,643	107,772	99%
Images open	108,572	107,741	99%
Images % open	100%	100%	
Publications all	90,195	21,621	24%
Publications open	43,479	5,938	14%
Publications % open	48%	27%	

(Received: March 31, 2017)

Plazi workflow: Open Access through text and data mining



December 31, 2017

Journal article Closed Access

Three new species of Ecuadorian Rhinotrugini (Coleoptera, Cerambycidae)

Bezark, Larry G.; Santos-Silva, Antonio

Bezark, Larry G., Santos-Silva, Antonio (2017): Three new species of Ecuadorian Rhinotrugini (Coleoptera, Cerambycidae). Zootaxa 4247 (5): 585-592, DOI: <https://doi.org/10.11646/zootaxa.4247.5.5>

Files

Closed Access

Files are not publicly accessible.

Access

Publication date:

December 31, 2017

DOI:DOI: [10.11646/zootaxa.4247.5.5](https://doi.org/10.11646/zootaxa.4247.5.5)**Keyword(s):**

Biodiversity Taxonomy Animalia Arthropoda

Insects Coleoptera Cerambycidae

Published in:

Zootaxa. 4247 pp. 585-592.

Related identifiers:

Has part:

<http://treatment.plazi.org/id/03CE9B411F78FD13FF1D5198FFED595B>,<http://treatment.plazi.org/id/03CE9B411F7AFD15FF1D558AFB7D5E32>,<http://treatment.plazi.org/id/03CE9B411F7CFD17FF1D5222FB395CF7>, [10.5281/zenodo.439186](https://zenodo.org/record/439186), [um:lsid:plazi.org:fig:2-187-1392-196-1924](https://zenodo.org/record/439186) (LSID), [10.5281/zenodo.439187](https://zenodo.org/record/439187), [um:lsid:plazi.org:fig:6-167-1419-561-1775](https://zenodo.org/record/439187) (LSID)**Alternate identifiers:**[um:lsid:zoobank.org:pub:427F743A-5286-4023-80A6-66659E5B7B49](https://zoobank.org/pub:427F743A-5286-4023-80A6-66659E5B7B49) (LSID)**Communities:**

Biodiversity Literature Repository

Share**Cite as**

Bezark, Larry G., & Santos-Silva, Antonio. (2017). Three new species of Ecuadorian Rhinotrugini (Coleoptera, Cerambycidae). Zootaxa, 4247(5), 585-592. [http://doi.org/10.11646/zootaxa.4247.5.5](https://doi.org/10.11646/zootaxa.4247.5.5)

Start typing a citation style...

Export[BibTeX](#) [CSL](#) [DataCite](#) [Dublin Core](#) [JSON](#)
[MARCXML](#) [Mendeley](#)

Original article DOI

Keywords

Taxonomic treatment

Image

Alternative identifier

Community

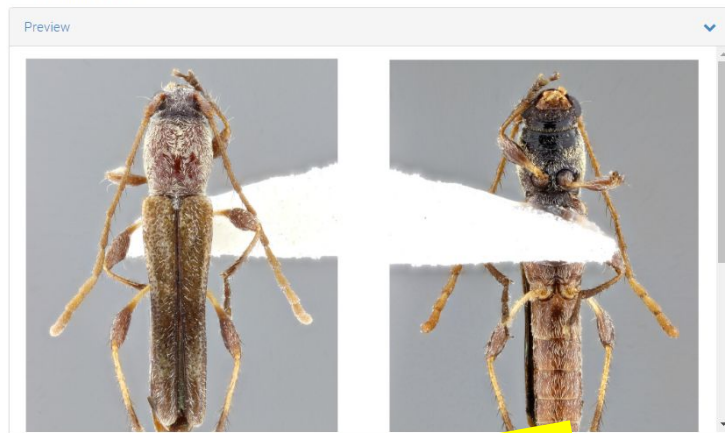
December 31, 2017

Figure Open Access

FIGURES 9 – 12 in Three new species of Ecuadorian Rhinotragini (Coleoptera, Cerambycidae)

Bezark, Larry G.; Santos-Silva, Antonio

FIGURES 9 – 12. *Lygrocharis setosus* sp. nov., holotype female: 9. Dorsal habitus; 10. Ventral habitus; 11. Lateral habitus; 12. Head, frontal view.



Files	
Name	
figure.png	5.2 MB
md5:a1b763f6de4ab303b7b96674e0f35c2b	Preview Download

Source file

Publication date:

December 31, 2017

DOI:

DOI [10.5281/zenodo.439187](https://doi.org/10.5281/zenodo.439187)

Keyword(s):

Biodiversity Taxonomy Animalia Arthropoda
Insecta Coleoptera Cerambycidae Lygrocharis

Related identifiers:

Cited by:

<http://treatment.plazi.org/id/03CE9B411F7CFD1FF1D5222FB395CF7>

Part of:

10.11646/zootaxa.4247.5.5,
urn:lsid:plazi.org/pub:FFF7E3391F79FD10FF8A55FFF25D27 (LSID)

Communities:

[Biodiversity Literature Repository](#)

License (for files):

[License Not Specified](#)

Share



Cite as

Bezark, Larry G., & Santos-Silva, Antonio. (2017, December 31). FIGURES 9 – 12 in Three new species of Ecuadorian Rhinotragini (Coleoptera, Cerambycidae). Zenodo. <http://doi.org/10.5281/zenodo.439187>

Start typing a citation style...

Export

BibTeX CSL DataCite Dublin Core JSON
MARCXML [Mendeley](#)

DataCite DOI from Zenodo

Keywords

Taxonomic treatment

Journal article DOI

Community

Licence

Export

Σ BLR Data

- = **Community Repository using Invenio**
- = **Opening up a huge corpus of imprisoned data from scientific publications**
- = **Enriching publications and its data by linking the data and publication, making it findable, accessible, inter-operable and re-usable**

3

Comparisons between Cases



Principle One:

Facilitate the lawful access to and reuse of research data

- Open access and unrestricted access and reuse
- Place data in public domain
- Apply rights waivers e.g., CC0 or PDDL
- Rights assertions and restrictions should make data available equitably, espec to disadvantaged



P1: Facilitate the lawful access to and reuse of research data

Open and unrestricted access to the over 105,000 scientific illustrations extracted from publications, as well as taxonomic treatments of species

Items with closed access remain closed

Items published on or before Dec 31 1999 are read-only access

Select types of data provided are considered to be ineligible for copyright protection (Patterson et al, 2015: The Blue List <http://plazi.org/activities/blue-list/>).

Conditions of reuse depend on the applicable national copyright law

Access to self-archived data, software, and other research outputs from Caltech that may (or may not be) associated with publications and theses

Items with closed access not accepted by repository

Ownership of provided data governed by Caltech IP Policy

All releases must be accompanied by required disclaimers of responsibility

Waivers or Licenses for reuse accompany each data set.
CC0 waiver is the default.
CC-BY, CC-BY-NC Licenses also available

Conditions of reuse depend on Caltech IP Policy and US national copyright law, including user exceptions e.g. Fair Use



Principle Two:

Determine the rights to and responsibilities for the data

- Establish rightsholder + rights before dissemination
- Abide by applicable rights
- Participate in law/policy that opens access to data
- Design and deliver training on rights and responsibilities with research data



P2: Determine the rights to and responsibilities for the data

The data source are exclusively publications. Publications *per se* are made accessible depending on their licences.

Metadata is exposed and enhanced with links to the exposed data.

BLR has its domicile in Switzerland and is ruled by Swiss law. Note that E.U. database protection is not applicable in Switzerland.

BLR is publishing their legal arguments and makes them as widely accessible as possible.

The BLR community is involved in lecturing on providing open access to scientific research results

Content in CaltechData generally falls under the Institute Copyright and Software Policy

For content that may be non-Caltech in origin, the depositor verifies that they have the rights to license the content and identifies/acknowledges owner(s).

The Library coordinates w/ Office of General Counsel; Research Compliance Office; Office of Technology Transfer to align repository terms and conditions.

Caltech Library has a non-exclusive, worldwide, perpetual license to use, duplicate and distribute the deposited content, and to transfer the content to any format or medium now known or later developed for archiving, preservation and access purposes.

Depositor retains right to publish and/or share deposited content as provided by the Institute Copyright and Software Policy.



Principle Three: Balance the Legal Interests

- ◉ Waive rights in publicly funded data where possible
- ◉ Limit time embargoes



P3: Balance the Legal Interests

Zenodo is publicly funded and supports the approach by the BLR community.

BLR is supported by a not for profit organization, a commercial publisher and various domain specific publishers such as scientific associations. All BLR data is openly accessible.

CaltechDATA supports embargos in order to encourage data deposit and support privacy while a publication is in preparation. However, it not an indefinite embargo and the researcher is responsible for complying with funder, publisher, or institutional requirements.

The repository will restrict access to the data until the end of the embargo period; at which time, the content will become publicly available automatically.

Users assume responsibility for complying with funder, publisher, or organizational embargo requirements (such as the Graduate School policy on thesis embargo).



Principle Four: State the Rights Clearly and Transparently

- Use standard electronic rights statements
- Consult qualified legal counsel to clarify rights
- Rightsholders inform users about terms and conditions of use

P4: State the Rights Clearly and Transparently

BLR is making use of Zenodo's rights statements.

BLR is supported by a not for profit organization, a commercial publisher and various domain specific publishers such as scientific associations. All BLR data is openly accessible.

BLR includes legal counsels to accompany the data extraction workflow, and the data policy of BLR and presents the findings in public ([Agosti and Egloff, 2009](#); [Patterson et al., 2014](#), [Egloff et al., 2017](#)).

Users are informed about their rights on the front page of BLR.

The license or waiver name is reported as part of the data record and metadata

A future goal is to include license information in the metadata deposited with the DOI registration agency (generally DataCite) for a Caltech-created dataset.

Caltech community members can access the Office of General Counsel for legal advice about data rights.

CaltechDATA does not yet have a standard mechanism for supporting special terms and conditions of use. We want to encourage researchers to select standard licenses.



Principle Five: Promote harmonization of rights in research data

- Leverage user rights in applicable statutes
- Devise workflows and best practices as useful tools to promote harmonization



P5: Promote harmonization of rights in research data

Plazi is using the BLR to demonstrate the power of data sharing to funding agencies and law makers.

The data extraction workflow has been described by [Agosti & Egloff, 2009](#); a taxonomy of data and their legal status has been compiled in the Blue List ([Patterson et al., 2014](#)) and the augmented list ([Egloff et al., 2017](#)). BLR promotes harmonization by encouraging data providers to use copyright waivers.

By limiting custom rights statements, the CaltechDATA repository encourages users to adopt standard licenses for their data.

Caltech Library applies community norms, e.g., the Association of Research Libraries' Code of Best Practices in Fair Use -- where warranted -- to make accessible via campus network any publications associated with the deposited dataset that are necessary to understand or interpret that data.



Principle Six: Provide proper attribution and credit for research data

- Establish and follow normative practices for giving credit where due



P6: Provide proper attribution and credit for research data



All data in BLR are linked to their source. Attribution is therefore guaranteed.

As follow-up of discussions in presentations on BLR, a special chapter on attribution has been included in the publication on copyright on scientific illustration ([Egloff et al. 2017](#)), accompanying BLR.

By introducing DOIs for images extracted from articles, we explicitly want to highlight the value of citing subarticle elements.

Caltech Library encourages and provides training in data attribution through Data and AuthorCarpentry workshops as well as a program on Giving Credit Where Due offered for advisors, instructors, and students

We are pursuing offering of example citations as part of the data record landing page, based on the best practices promoted by Fenner et al (2016, <http://biorxiv.org/content/early/2016/12/28/097196>) “**A Data Citation Roadmap for Scholarly Data Repositories**” but this is not currently implemented.



Thanks!

Any **questions** ?

You can find us at

gclement@library.caltech.edu

agosti@plazi.org

egloff_bader@bluewin.ch

tmorrell@library.caltech.edu

