### Digital repositories for FAIR data management Examples from Zenodo/InvenioRDM

Lars Holm Nielsen / CERN - IT Department CS3 Conference 2024



This work is licensed under Creative Commons Attribution 4.0 International.



# Make it Open FAIR





l Open FAIR

### How?





# DMPs & Repositories

Search records Q Communities My dashboard	• +• Lars.holm •		
Featured communities			
Biodiversity Literature Repository A community to share publications related to bio-systematics. DOI 10.5281/zenod0.10641396	>		
DOI       10.5201/201000.100410590         Becent uploads         March 13, 2024 (vi)       Image: Image	<ul> <li>Why use Zenodo?</li> <li>Safe — your research is stored safely for the future in CERN's Data Centre for as long as CERN exists.</li> <li>Trusted — built and operated by CERN and OpenAIRE to ensure that everyone can join in Open Science.</li> <li>Citeable — every upload is assigned a Digital Object Identifier (DOI), to make them citable and trackable.</li> <li>No waiting time — Uploads are made available online as soon as you hit publish, and your DOI is registered within seconds.</li> <li>Open or closed — Share e.g. anonymized clinical trial data with only medical professionals via our restricted access mode.</li> <li>Versioning — Easily update your dataset with our versioning</li> </ul>		
March 13, 2024 (v1) Dataset	<ul> <li>feature.</li> <li>GitHub integration — Easily preserve your GitHub repository in Zenodo.</li> <li>Usage statistics — All uploads display standards compliant usage statistics</li> </ul>		
Occurrence data on PCNs in food and feed and dietary surveys per country and age group available in the EFSA Comprehensive Database, considered in the exposure assessment	Newsletter		
Uploaded on March 12, 2024 Part of Knowledge Junction	Receive updates on our latest developments, projects and upcoming webinars sent quarterly.		
March 13, 2024 (v1) Dataset 🖨 Open	your-email@example.com		
Annex C - Human dietary exposure assessment	Name Your name		



Export	
JSON	Export
JSON	
JSON-LD	
CSL	
DataCite JSON	
DataCite XML	
Dublin Core XML	
MARCXML	
BibTeX	
GeoJSON	
DCAT	
Codemeta	
Citation File Format	J

# Taxonomic treatments

Describe the discovery of new biological species

**Example:** Journal article describing 22 new millipedes, published in European Journal of Taxonomy





Abstract. Twenty-two new species of the genus Eviulisoma Silvestri, 1910, from the Eastern Arc Mountains, Tanzania, are described: E. acaciae sp. nov., E. aequilobatum sp. nov., E. akkariae sp. nov., E. angulatum sp. nov., E. articulatum sp. nov., E. biquintum sp. nov., E. breviscutum sp. nov., E. cetafi sp. nov., E. chitense sp. nov., E. commelina sp. nov., E. coxale sp. nov., E. ejti sp. nov., E. grumslingslak sp. nov., E. kalimbasiense sp. nov., E. navuncus sp. nov., E. nessiteras sp. nov., E. ottokrausi sp. nov., E. paradisiacum sp. nov., E. sternale sp. nov. and E. zebra sp. nov. from the Udzungwa Mts, E. culter sp. nov. from the Rubeho Mts and E. kangense sp. nov. from the Kanga Mts. Eviulisoma kwabuniense Kraus, 1958, and E. dabagaense Kraus, 1958, both from the Udzungwa Mts, are redesribed based on new material. Notes are provided on E. iuloideum (Verhoeff, 1941) based on type material. Eoseviulisoma Brolemann, 1920, is synonymized under Eviulisoma, based on newly collected material of E. julinum (Attems, 1909), type species of Eoseviulisoma. New material of Suohelisoma ulugurense Hoffman, 1964, type species of Suohelisoma Hoffman, 1964, has revealed that the gonopod structure is more similar to that of Eviulisoma than originally thought, but Suohelisoma is retained as a valid genus. Four species groups are recognized among Eviulisoma species from the Udzungwa Mts, but the need for a revision of the entire genus is emphasized. Two types of epizootic fungi are recorded from Eviulisoma spp., and an enigmatic amorphous mass, which may be a kind of plugging substance, is recorded from the gonopod tips and excavated sixth sternum of several species.

Keywords. Taxonomy, new species, epizootic fungi, copulatory plug

445: 1-90. https://doi.org/10.5852/eit.2018.445



Engloff H. 2018. A mountain of millipedes VII: The genus Eviulisoma Silvestri, 1910, in the Udzungwa Mountains, Tanzania, and related species from other Eastern Arc Mountains. With notes on Sese Xiulisoma Brolemann, 1920, and Suohelisoma Hoffman, 1963 (Diplopoda, Polydesmida, Paradoxosomatidae). European Journal of Taxonomy 445: 1–90. https://doi.org/10.5852/ejt.2018.445

ropean Journal of Taxonomy 445: 1-90

ISSN 2118-9773 www.europeanjournaloftaxonomy.eu 2018 · Enghoff H

This work is licensed under a Creative Commons Attribution 3.0 Licens

Monograph urn:lsid:zoobank.org:pub:852A3F68-B728-413A-B12E-56F306D56C35

### A mountain of millipedes VII: The genus Eviulisoma Silvestri, 1910, in the Udzungwa Mountains, Tanzania, and related species from other Eastern Arc Mountains. With notes on Eoseviulisoma Brolemann, 1920, and Suohelisoma Hoffman, 1963 (Diplopoda, Polydesmida, Paradoxosomatidae)

### Henrik ENGHOFF

Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 København Ø, Denmark

Email: henghoff@snm.ku.dk

### urn:lsid:zoobank.org:author:FB09A817-000D-43C3-BCC4-2BC1E5373635

Enghoff H. 2018. A mountain of millipedes VII: The genus Eviulisoma Silvestri, 1910, in the Udzungwa Mountains Tanzania, and related species from other Eastern Arc Mountains. With notes on *Eoseviulisoma* Brolemann, 1920 and Suohelisoma Hoffman, 1963 (Diplopoda, Polydesmida, Paradoxosomatidae). European Journal of Taxonomy

## Treatments: Data in disguise

Geographic coordinates

Date of collection

Material (total: 3 중중)

Holotype

TANZANIA: *A*, Mwanihana Forest, above Sanje, 1650 m a.s.l., pitfall trap, 18 Aug. 1982, M. Stoltze and N. Scharff leg. (ZMUC).

### **Paratypes**

TANZANIA: 1 3, Morogord Region, Kilombero District, Udzungwa Mts National Park, forest below Mwanihana Peak, 7°49' S, 36°50' E, 1800 m a.s.l., sifted from leaf litter, 20 Aug. 2017, T. Pape leg. (ZMUC); 1 3, Morogoro Region, Udzungwa Mts National Park, Mito Mitatu, above Mang'ula, 07°49′3″ S, 36°52′58″ E, 1487 m a.s.l., 16 Dec. 2016,

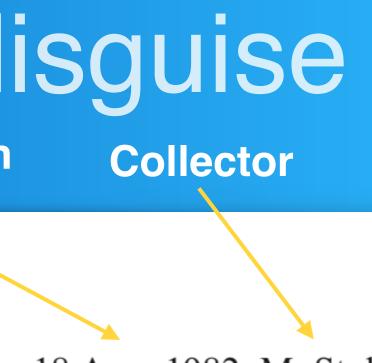
### **Host collection**



In accordance with the European Journal of Taxonomy's FAIR & Open Science policy, the formatting guide for entomology, zoology and palaeozoology material citations is provided below (guidelines for botany available soon). Authors are encouraged to prepare their manuscripts according to this model prior to submission, but they will also be given the opportunity to comply upon acceptance of the article.

While EJT strongly recommends that authors adhere to the guidelines given below, the fine-grain formatting of the meterial citations is not compulsory, if an outhor decides not to comply or the m





### EUROPEAN JOURNAL OF TAXONOMY MATERIAL CITATIONS FORMATTING GUIDE

# Locked up data

**Unanswerable questions:**  How many species have been described by my collection? • Give me a list of all new species? Retrieve all images for a given taxon? What's known about a geographic region?

### •Treatments:

- Past 260 years: ~10+ millions published
- Every year: ~17k new / ~130k augmented

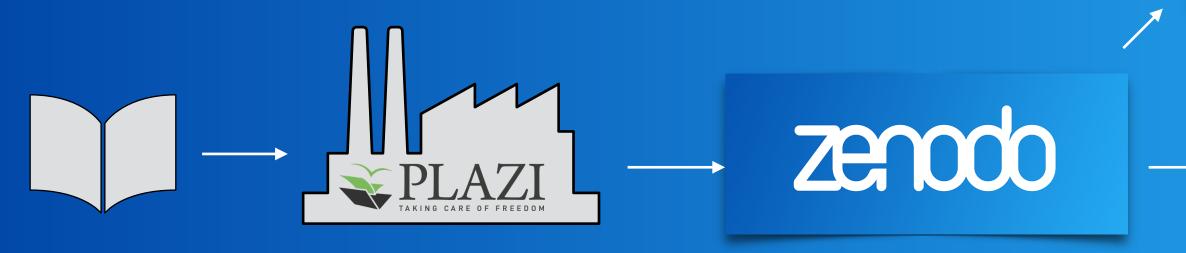








### It takes an ecosystem



### Article Text and data mining Preserve & disseminate





### B≞LS® S

LITERATURE SERVICE

**BiodiversityPMC** 

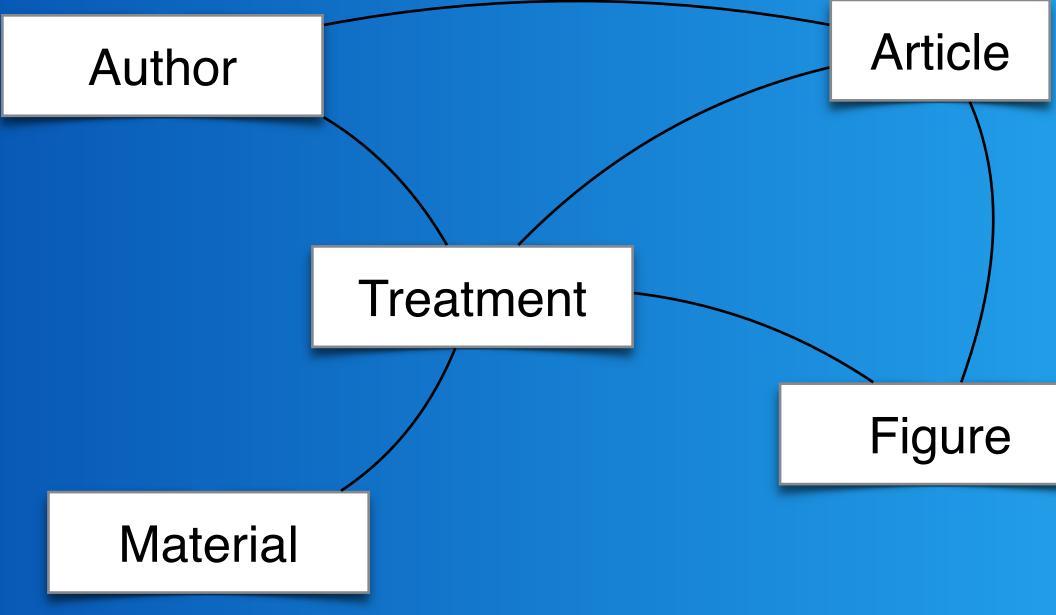


Global Biodiversity



### Search, retrieve & linking

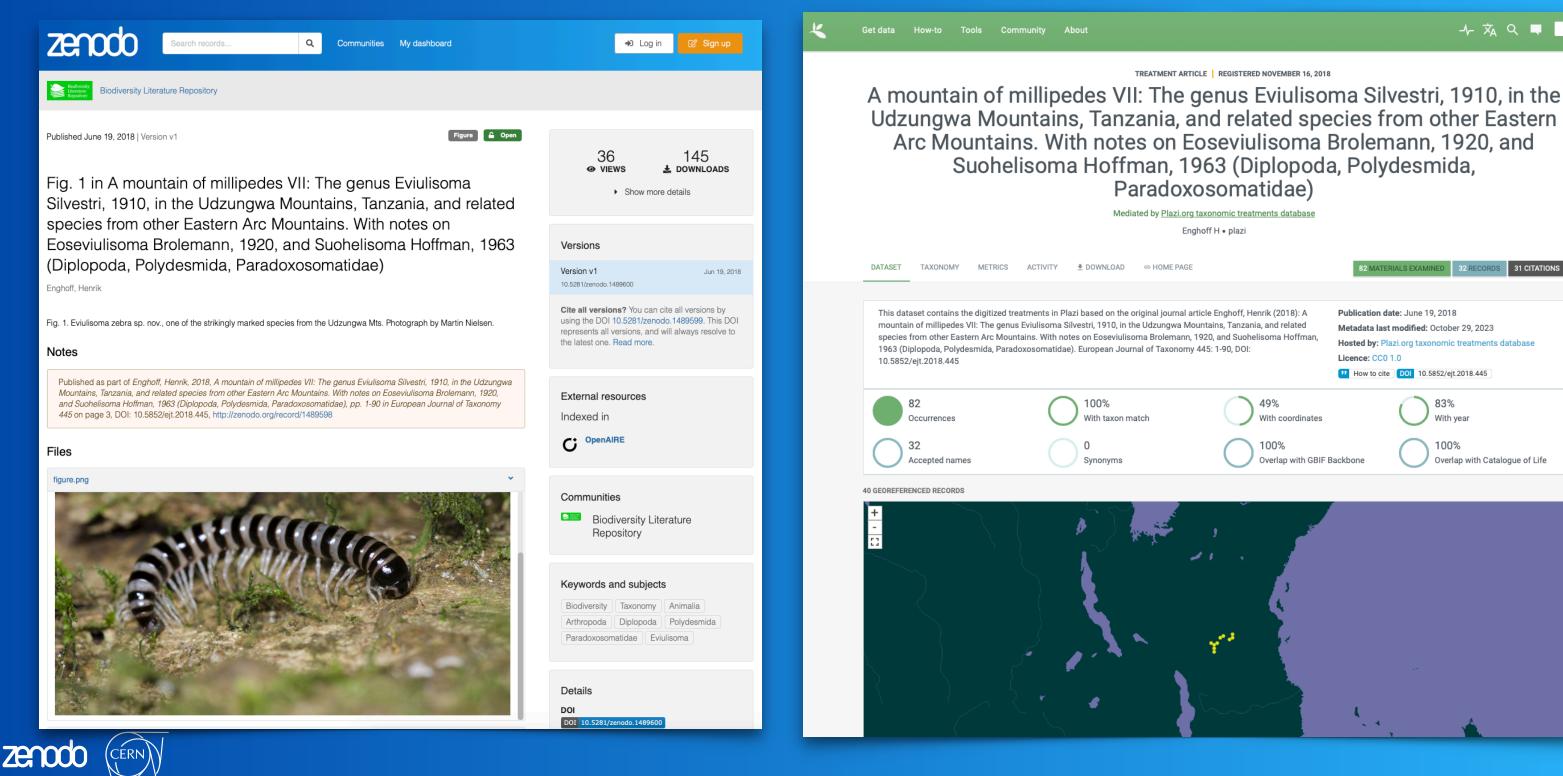
### It takes an ecosystem







# FAIR Data



-/- 📩 Q 🔲 🛛 Login

TREATMENT ARTICLE | REGISTERED NOVEMBER 16. 2018

Udzungwa Mountains, Tanzania, and related species from other Eastern Arc Mountains. With notes on Eoseviulisoma Brolemann, 1920, and Suohelisoma Hoffman, 1963 (Diplopoda, Polydesmida, Paradoxosomatidae)

Mediated by Plazi.org taxonomic treatments database

Enghoff H • plazi

PAGE	82 MATERIALS EXAMINED 32 RECORDS 31 CITATIONS
nal article Enghoff, Henrik (2018): A a Mountains, Tanzania, and related ann, 1920, and Suohelisoma Hoffman, nomy 445: 1-90, DOI:	Publication date: June 19, 2018         Metadata last modified: October 29, 2023         Hosted by: Plazi.org taxonomic treatments database         Licence: CC0 1.0         Image: How to cite         Doi: 10.5852/ejt.2018.445
49% With coordinates	83% With year
0verlap with GBIF B	ackbone 100% Overlap with Catalogue of Life



# **Domain-specific metadata**

### Searchable metadata

Add field and continue
Add field and close

Search field names	Q, in: A	All domains -
	Biodiversity	Family 🗹
Basis of record The specific nature of the data record, e.g., PreservedSp		Paradoxosomatidae
Ø Biodiversity: dwc i≣ Multiple value field I Tex		Genus 🗹
Catalog number		Eviulisoma
An identifier for the record within the catalog or collection  Blodiversity: dwc		Kingdom 🗹
		Animalia
Class The taxonomic class of the organism.		Order 🔽
Biodiversity: dwc		Polydesmida
Collection code		Phylum 🗹
A unique identifier assigned to the collection from which Biodiversity: dwc		Arthropoda
Country		Scientific name authorship 🗹
The name of the country or major administrative unit whe Blodiversity: dwc		Silvestri
		Taxon rank 🗹
County The name of the county, shire, or equivalent where the sp		genus
Biodiversity: dwc		-
Date identified		Taxonomic concept label
The date when the specimen was identified or determine		Eviulisoma Silvestri, 1910 sec. Enghoff, 2018

### Keywords and subjects Suggest from EuroSciVoc phys Languages Add phys Search for a language by name (e.g "eng (EuroSciVoc) So Dates (EuroSciVoc) Ph (EuroSciVoc) Pa Date YYYY-MM-DD or YYYY-MM-DD/Y (EuroSciVoc) So

System files (27.1 kB)

Name

application/vnd.plazi.v1+xml

md5:b0dd7dd2e94c0a018634bbc5054e1369 3



× Cancel

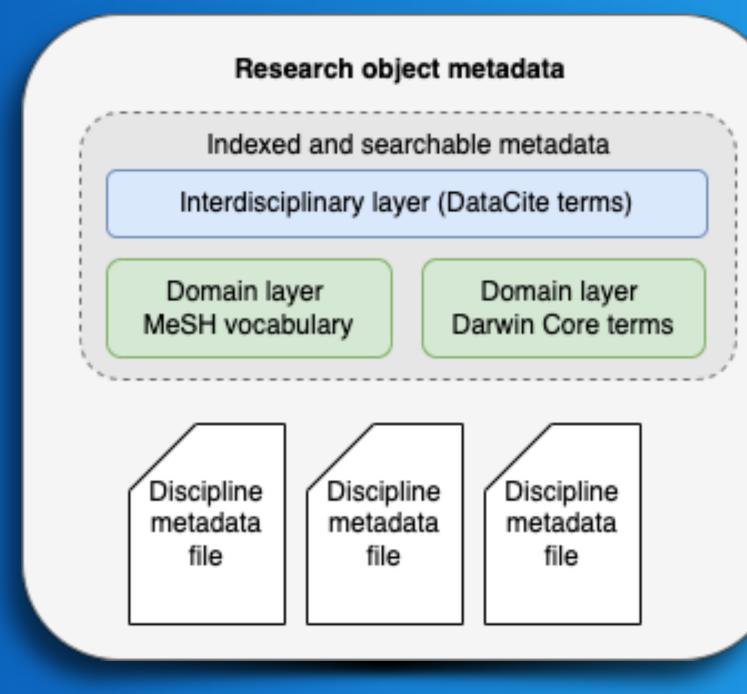
### Vocabularies

	•
lar physics	
ysical cosmology	
rticle physics	
lid-state physics	

### Metadata files

	~
Size	Download all
27.1 kB	L Download

# Domain-specific metadata



Zenodo (CERN)

# Annotations

One new genus and nineteen new species of ground spiders (Araneae: Gnaphosidae) from Iran, with other taxonomi

Annotations

### **e** | 0

≡

Diagnosis The new species is most similar to B . venatrix ( O . Pickard-Cambridge , 1874 ) and B . koponeni... subSubSection part 3

2

Berlandina Dalmas , 1922

taxonomicName

order: Araneae

family: Gnaphosidae

genus: Berlandina

Berlandina artaxerxes

taxonomicName

order: Araneae

family: Gnaphosidae

genus: Berlandina

species: artaxerxes

venatrix ( O Picka

European Journal of Taxonomy 751: 68-114 (2021)

### Type material

### Holotype

IRAN • ♂; Ilam Province, Abdanan: 32°53' N, 47°30' E; Jul. 2015; A. Zamani leg.; MHNG

### Paratypes

IRAN • 1  $\phi$ ; same collection data as for holotype: MHNG • 8  $\sigma\sigma$ ; Kermanshah Province, Zarrin Vil.; Apr. 2017; M. Amiri leg.; MHNG • 1  $\phi$ ; Kurdistan Province, Marivan; May 2017; A. Zamani MHNG • 1  $\sigma$ ; Fars Province, 5 km N of Persepolis; 1970; K. Bilek leg; NHMW.

### Description

### Male (holotype)

Habitus as in Fig. 2A. Total length 4,50. Carapace 1.92 long, 1.50 wide. Eye sizes and interdistan PMEs: AME = 0.09, ALE = 0.12, PME = 0.12, PLE = 0.12, PME–PME = 0.03. Carapace, chelic sternum, maxillae and labium light reddish brown. Chelicera with 2 pro- and 5 retromarginal teeth. slightly lighter than carapace, without annulations, and with numerous spines (III and IV). Abdomen brown, without any markings and with a small scutum anteriorly. Spinnerets light gray. Measuren of legs: I: 7.29 (1.77, 1.33, 1.66, 1.50, 1.03), II: 5.35 (1.41, 0.87, 1.13, 1.08, 0.86), III: 4.41 (1.19, 0.82, 1.01, 0.71), IV: 7.07 (1.81, 1.06, 1.67, 1.71, 0.82).

Palp as in Fig. 2B–G. Femur unmodified; patella slightly longer than tibia (not including apoph tibial apophysis as long as tibia, with swollen base and flat anterior part, margins parallel; cym 2 times as long as wide, its width as long as tibia; spermophor coiled; median apophysis (= cond sensu Panayiotou *at al.* 2010) large, sharply pointed; filamentous embolus originates almost from center of the bulb, makes 360° loop.

### Female (paratype from Ilam Province)

Carapace 3.05 long, 2.28 wide (abdomen missing). Eye sizes and interdistance of PMEs: AME = ALE = 0.15, PME = 0.13, PLE = 0.12, PME-PME = 0.05. Coloration and habitus features as in 1 Measurements of legs I: 8.79 (2.30, 1.73, 1.86, 1.00, 1.12, 1.12) II: 6.99 (1.92, 1.26, 1.55, 1.34, 0.92) 4.41 (1.19, 0.68, 0.82, 1.01, 0.71), IV: 6.04 (1.65, 1.00, 1.12, 1.39, 0.88). 7 of 47 • page0006.png



nic consider	0
	_
1 Joob	
i leg.	
nce of cerae,	
Legs	
n light ments	
, 0.68.	
iysis);	
ibium fuctor	_
m the	
0.13, male.	
2), III	

# **Beyond biodiversity**

### Research object metadata

Indexed and searchable metadata

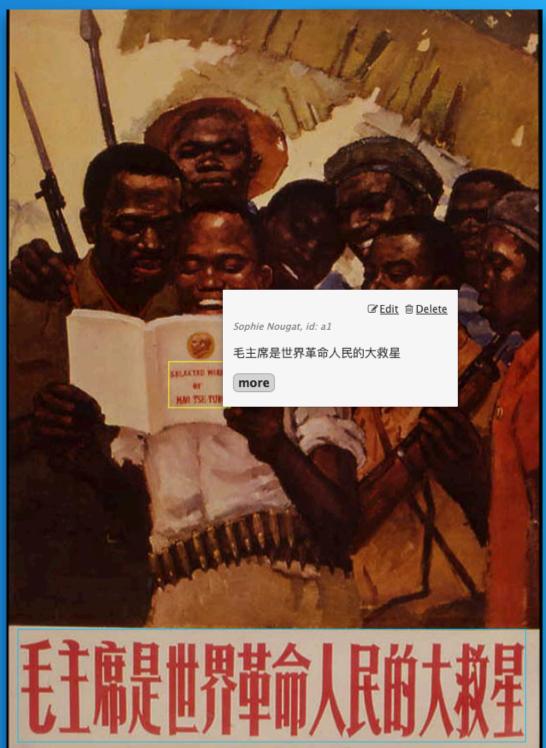
Interdisciplinary layer (DataCite terms)

Domain layer MeSH vocabulary

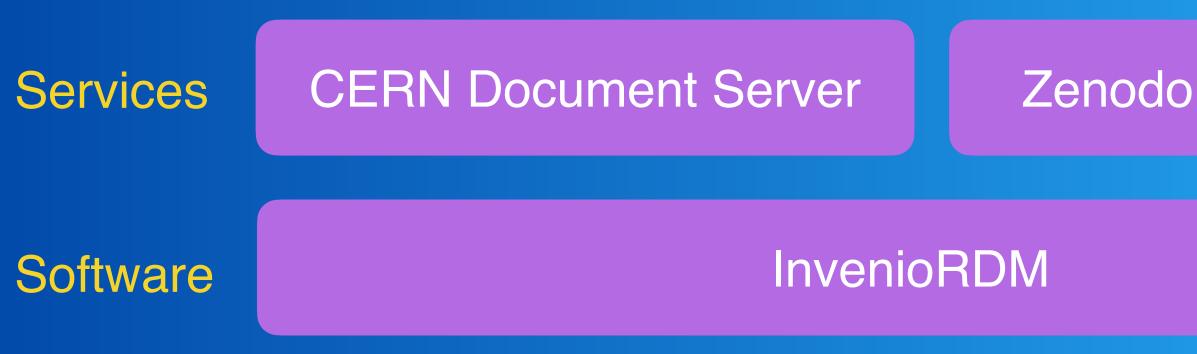
Domain layer Darwin Core terms

Discipline Discipline Discipline metadata metadata metadata file file file





# InvenioRDM





# 0

# InvenioRDM

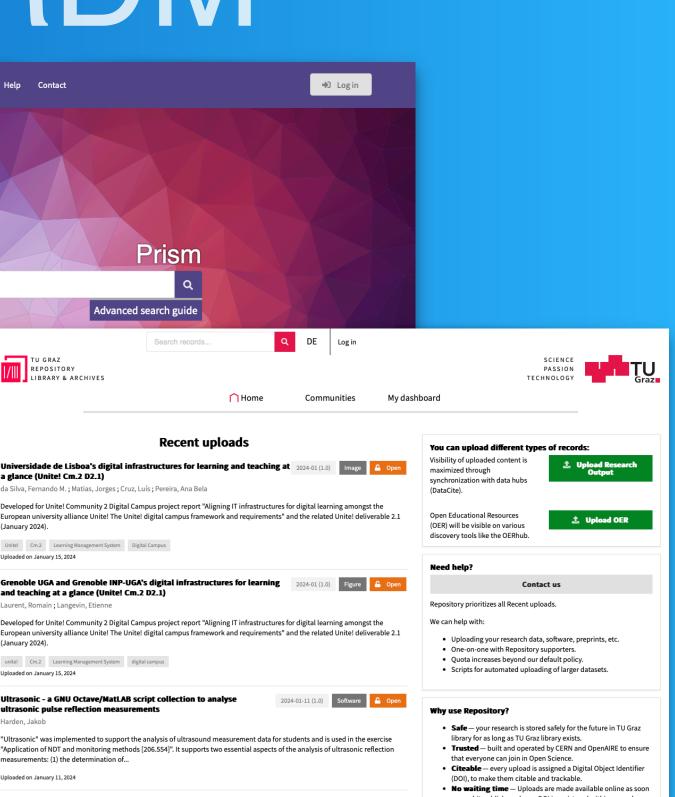
Harden, Jakob

Uploaded on January 11, 2024

measurements: (1) the determination of ...

Morthwestern Medicine Search Communities Help Contact **Caltech** DATA Communities My dashboard Log in Prism by Caltech Library Q Search Prism Advanced search guide California Institute of Technology Research Data Repository TU GRAZ /||| REPOSITORY LIBRARY & ARCHIVES Q Search Recent uploads Universidade de Lisboa's digital infrastructures for learning and teaching at 2024-01 (1.0) Image a glance (Unite! Cm.2 D2.1) da Silva, Fernando M.; Matias, Jorges; Cruz, Luís; Pereira, Ana Bela Developed for Unite! Community 2 Digital Campus project report "Aligning IT infrastructures for digital learning amongst the European university alliance Unite! The Unite! digital campus framework and requirements" and the related Unite! deliverable 2.1 (January 2024). Unite! Cm.2 Learning Management System Digital Campus Uploaded on January 15, 2024 Data Sets Software Upload Grenoble UGA and Grenoble INP-UGA's digital infrastructures for learning 2024-01 (1.0) Figure C Open and teaching at a glance (Unite! Cm.2 D2.1) Laurent, Romain ; Langevin, Etienne Developed for Unite! Community 2 Digital Campus project report "Aligning IT infrastructures for digital learning amongst the European university alliance Unite! The Unite! digital campus framework and requirements" and the related Unite! deliverable 2.1 (January 2024). unite! Cm.2 Learning Management System digital campus About CaltechDATA Uploaded on January 15, 2024 FAQ News Terms Report a Problem Contact Powered by InvenioRDM Ultrasonic - a GNU Octave/MatLAB script collection to analyse ultrasonic pulse reflection measurements





### Future directions: UX

### Curation checks

Domain apps

ZEOCO Search records	Q Communities My dashboard	
Select the community where you want to submit your record.	Select a community	
Files	~	Draft 1
Metadata-only record ③	Storage available 0 out of 100 files 0 bytes out of 50.00 G	Save draft
Drag and drop files	- or - Upload files	♥ Visibility*
		Full record Public Restricted
Basic information	~	Files only Public Restricted
Do you already have a DOI for this upload?  Yes Copy/paste your existing DOI here	○ No	Public The record and files are publicly accessible.
A DOI allows your upload to be easily and unambiguously cited. Exar	nple: 10.1234/foo.bar	Options

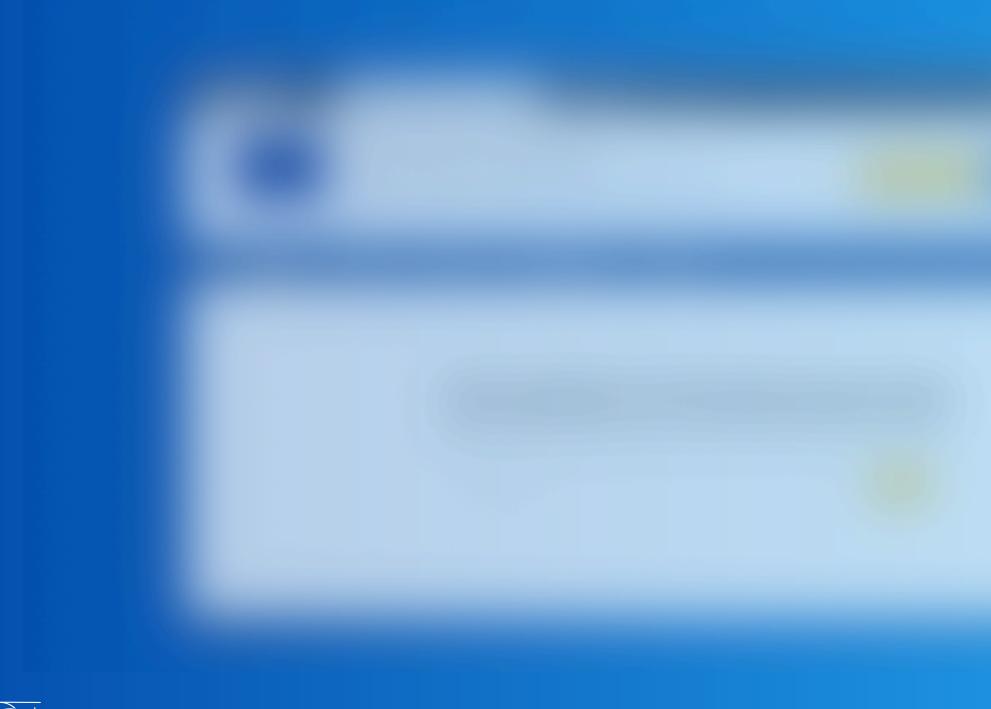
Apply an embargo ()





### Measure

### Next week



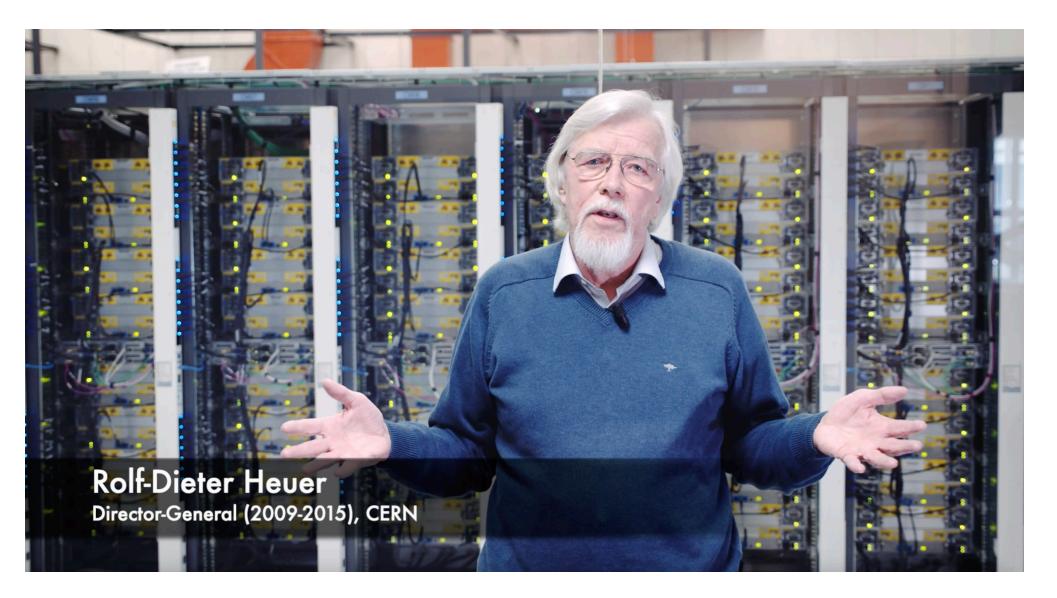






Funded by the European Union





"Why do I like Zenodo? Because Zenodo is FAIR. Fair in sense of lowercase, and FAIR in the sense of uppercase meaning, [...]" - Rolf-Dieter Heuer, Director-General (2009-2015)

