

LexBib

a corpus and digital bibliography
of Lexicography and Dictionary
Research



David Lindemann david@lexbib.org

Introduction

Digital bibliography platforms

Amazing progress

- publication metadata
 - structured: DC, bibtex, RIS, etc.
- content describing metadata
 - subject indexation, citation networks, extracted terms, topic models, ...

One for each community

- More than the big platforms can deliver
- DBLP, ACM, econbiz

Lexicography: a small discipline

- much to digitize, much to make accessible
- LexBib > Elexis

DARIAH WG Bibliographical data

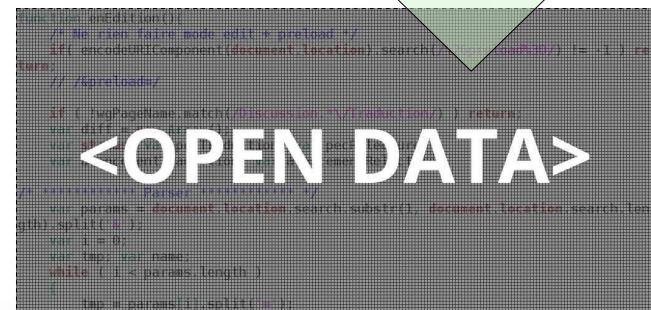
- Platform for knowledge transfer (training schools, datathons, etc.)
- Guidelines and best practices for bibliography digitization and digital biblio-platforms

21923 Voillat, François: Le "Glossaire des patois de la Suisse romande" (GPSR). In: Actes du XVIII^e Congrès International de Linguistique et de Philologie Romanes [...] Tome VII, 1989¹, 338–345.

15475 Nyhlén, Lars-Olaf: Wie sagt man in Österreich? In: Moderna Språk 68, 1974, 275–281. [Bes. zu: Jakob Ebner: Duden. Wie sagt man in Österreich? Wörterbuch der österreichischen Besonderheiten. Mannheim 1969 (Duden Taschenbücher 8)].

15397 Norden 11.–14. maj 1995. Red. av Anna Garde og Pia Jarvad. Oslo 1995 (Skrifter udgivet av Nordisk Forening for Leksikografi. Skrift nr. 2). [Daraus: 363, 1003, 2755, 3101, 3870, 5222, 5251, 5390, 6058, 6758, 9206, 9743, 10158, 11521, 11553, 11843, 12600, 12940, 12974, 13146, 13399, 14192, 15316, 15345, 15402, 15474, 15651, 15944, 17322, 20818, 21371].

15427 Novikov, L. A.: K probleme omonimii. [Zum Problem der Homonymie]. In: Leksikografičeskij sbornik 1960, 93–102.



LexBib Zotero Collection

zotero

Q Title, Creator, Year

Group Libraries

- LexBib
 - ABOUT_LEXBIB
 - BIBLIOGRAPHIES
 - COLLECTIVE_VOLS
 - CONFERENCES
 - Euralex_2018_experiment
 - JOURNALS
 - MONOGRAPHS
 - OBELEX-Meta-Part

Export

Item Type	Title	Creator	Date
Conference Paper	A Call for a Corpus-Based Sign Language Dictionary: An ...	Majetić and Bago	2018
Conference Paper	A Good Match: a Dutch Collocation, Idiom and Pattern Di...	Colman and Tiberius	2018
Conference Paper	A lexicon of Albanian for natural language processing	Kabashi	2018
Conference Paper	A Sample French-Serbian Dictionary Entry based on the ...	Marjanović et al.	2018
Conference Paper	A Universal Classification of Lexical Categories and Gram...	Saurí et al.	2018
Conference Paper	A Workflow for Supplementing a Latvian-English Dictiona...	Deksne and Veisbergs	2018
Conference Paper	Advances in Synchronized XML-MediaWiki Dictionary De...	Hämäläinen and Rueter	2018
Conference Paper	An Overview of FieldWorks and Related Programs for Co...	Baines	2018
Conference Paper	Analyzing User Behavior with Matomo ¹ in the Online Info...	Ripp and Falke	2018

Info Notes Tags Attachments Related Show Empty Fields

Item Type Conference Paper

Title A Sample French-Serbian Dictionary Entry based on the ParCoLab Parallel Corpus

Author Marjanović, Saša

Author Stosic, Dejan

Author Miletic, Aleksandra

Editor Čibej, Jaka

Editor Gorjanc, Vojko

Editor Kosem, Iztok

Editor Krek, Simon

Abstract

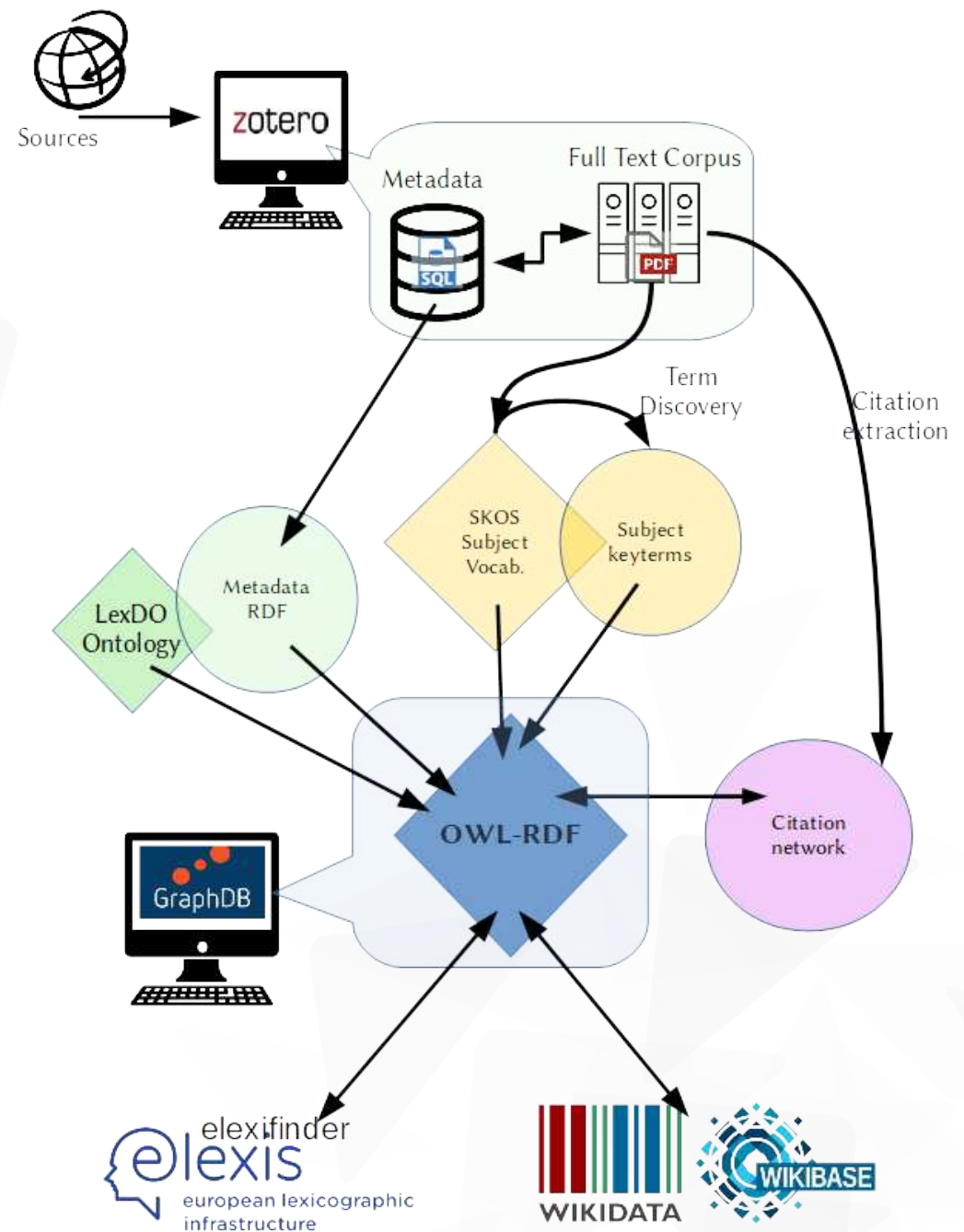
It has already been shown in the state- of the art in lexicography that the bilingual dictionary making process can be improved by relying on parallel corpora. The aim of this paper is to present such an application of the ParCoLab parallel corpus, a searchable trilingual 11 million token electronic database of aligned texts in French, Serbian and English, developed at the University of Toulouse (France) in cooperation with the University of Belgrade (Serbia). In this paper, we first point out the shortcomings of the leading general French-Serbian dictionaries, which were made using traditional lexicographic methods. We pay special attention to

- Access:
 - either: <http://lexbib.org/zotero>
 - or: <https://www.zotero.org/groups/lexbib>
- All metadata present, searchable
- structured metadata for local citation managers, export options
- Collaborative editing possible
 - Status: [Check here](#)



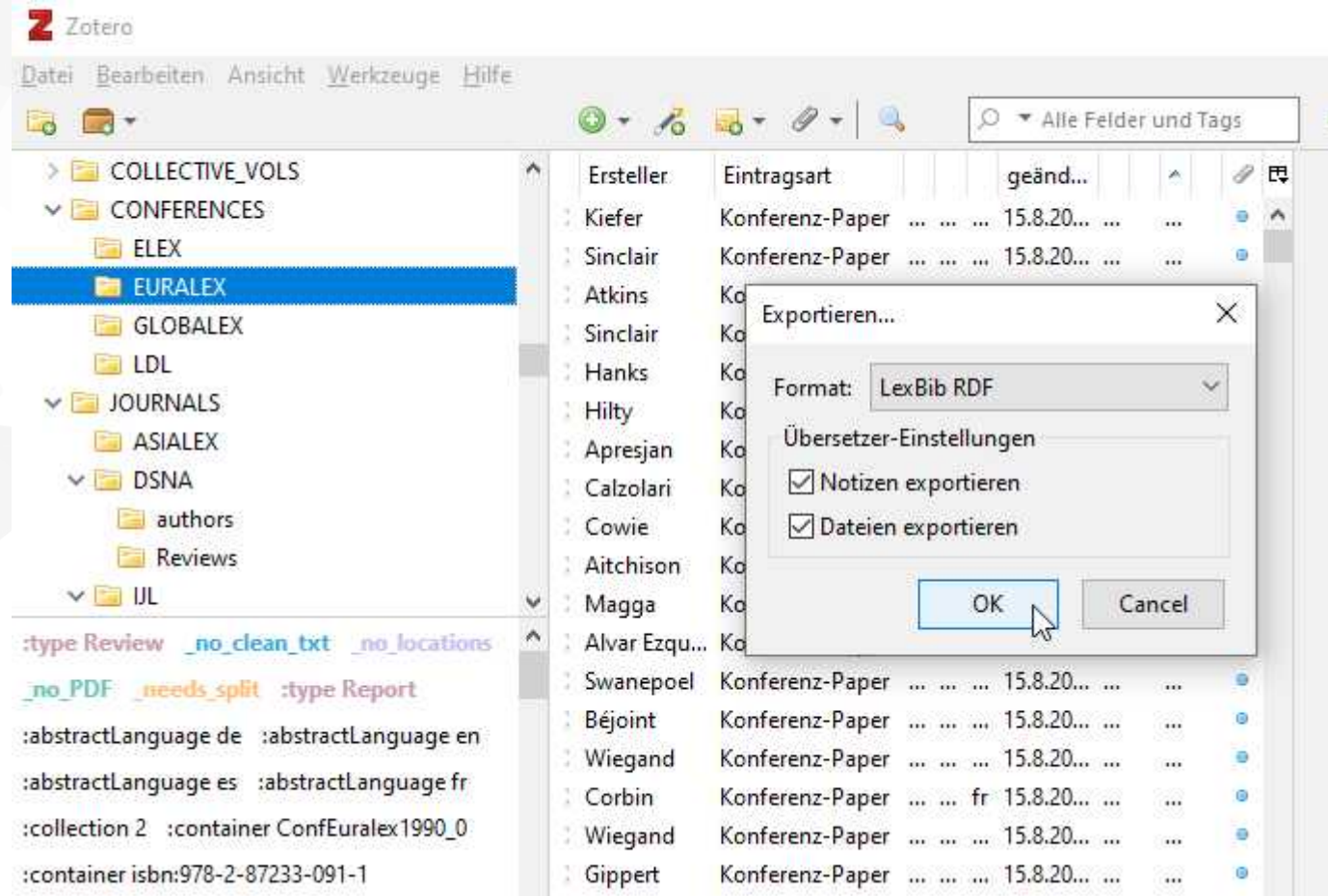
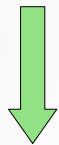
Workflow / Modules

- LexBib Zotero Collection
- LexDO Domain Ontology
- LexDO SKOS subject vocabulary
- Citation Network
- Graph database (triple store)
- Export for Elexifinder discovery platform
- Connection to Wikidata
- Setup of an own Wikibase

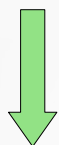


Zotero > LexBib RDF

▼ LexBib RDF export: See [Elexifinder at Github](#)



Zotero > LexBib RDF



```
22      <zotexport:pdfFile>D:/Zotero/storage/BXTR4R44/Adamska-SalaciakA_2016_fullte
23      <zotexport:txtFile>D:/Zotero/storage/A7L6MWCY/euralex_2016_084_p758.txt</z
24    </owl:NamedIndividual>
25    <owl:NamedIndividual rdf:about="https://euralex.org/publications/on-bullying-mo
26      <rdf:type rdf:resource="http://lexbib.org/lexdo/BibItem"/>
27      <lexdo:publishingPlace>Tbilisi</lexdo:publishingPlace>
28      <lexdo:publisherName>Ivane Javakhishvili Tbilisi State University</lexdo:pu
29      <bibo:pages>758-766</bibo:pages>
30      <bibo:isbn13>978-9941-13-542-2</bibo:isbn13>
31      <lexdo:eventName>XVII International EURALEX Congress</lexdo:eventName>
32      <dcterms:title>On Bullying, Mobbing (and Harassment) in English and Polish:
33      <lexdo:containerTitle>Proceedings of the 17th EURALEX International Congres
34      <lexdo:containerShortTitle>Euralex (2016)</lexdo:containerShortTitle>
35      <lexdo:creator rdf:resource="http://lexbib.org/agents/person/Adamska-Sařac
36      <bibo:authorList rdf:resource="https://euralex.org/publications/on-bullying
37      <lexdo:editor rdf:resource="http://lexbib.org/agents/person/MargalitadzeTir
38      <bibo:editorList rdf:resource="https://euralex.org/publications/on-bullying
39      <lexdo:editor rdf:resource="http://lexbib.org/agents/person/MeladzeGeorge",
40      <dcterms:date>2016-01-01T02:00:00.000Z</dcterms:date>
41      <lexdo:publicationLanguage rdf:resource="http://lexvo.org/id/iso639-3/eng",
42      <lexdo:event rdf:resource="http://lexbib.org/events/ConfEuralex2016"/>
43      <lexdo:container rdf:resource="http://worldcat.org/isbn/9789941135422"/>
44      <lexdo:collection>2</lexdo:collection>
45      <lexdo:abstract rdf:resource="https://euralex.org/publications/on-bullying-
46      <lexdo:fullTextUrl>https://euralex.org/publications/on-bullying-mobbing-and
47      <lexdo:zoteroItemUri rdf:resource="http://zotero.org/groups/1892855/items/M
48      <lexdo:zoteroItemID>M9Z45UVF</lexdo:zoteroItemID>
49      <lexdo:firstAuLoc>http://en.wikipedia.org/wiki/Poznań</lexdo:firstAuLoc>
50      <rdf:type rdf:resource="http://lsdis.cs.uga.edu/projects/semdis/opus#Artic
```


LexDO Domain Ontology

- ▼ Documentation: <http://lexbib.org/lexdo>
- ▼ Standard RDF vocabularies
 - ▼ BIBO, PROV, Dublin Core, FOAF, Ontolex
- ▼ Top-level concepts:
 - ▼ Agent
 - ▼ Person/Organization/Software
 - ▼ Resource
 - ▼ Bibliographical resource/Lexical resource/Software
 - ▼ Subject
 - ▼ Subjects of metalexicographical items
 - ▼ Language
 - ▼ Event
 - ▼ Place

Ontology Specification Draft

Used workbench software:
VocBench 3

LexDo - Domain Ontology of Lexicogra

Release 2020-03-14T16:50:24.513+01:00

This version:

<http://lexbib.org/lexdo/1.0.2/>

Authors:

[David Lindemann](#)

Contributors:

[Christiane Klaes](#)

[Laura Giacomini](#)

Publisher:

Institut Jožef Stefan / Universität Hildesheim / UPV/EHU Uni

Download serialization:

Format [JSON LD](#) Format [RDF/XML](#) Format [N Triples](#) Format [TTL](#)

License:

License <http://creativecommons.org/licenses/by/3.0/> [License](#)

Cite as:

LexDo Domain Ontology of Lexicography and Dictionary Res
For full reference, see [About LexBib](#) at Zotero.

<http://lexbib.org/lexdo/1.0.2>



LexDO Subject Headings SKOS vocabulary

- Documentation: lexbib.org
- Subjects with labels stemming from reference resources
 - Indices from Lexicography bibliographies and handbooks brought together
- SKOS relations
 - broader / narrower
 - "exact match" (synonym)
 - related
- Modeled following **BLL-LOD** (Chiarcos et al.)

- object language
- dictionary type
 - abridged dictionary
 - analogical dictionary
 - bilingual dictionary
 - bilingualized dictionary
 - bilingual dictionary
 - collaborative dictionary
 - user-generated content
 - general dictionary
 - encyclopaedic dictionary
 - historical dictionary
 - historical dictionary
 - standard dictionary
 - overall-descriptive dictionary
 - standard-descriptive dictionary
 - ideological dictionary
 - monolingual dictionary
 - multilingual dictionary
 - normative dictionary
 - onomasiological dictionary
 - reverse dictionary
 - semantic dictionary
 - specialized dictionary
 - information type oriented dictionary
 - etymological dictionary
 - terminological dictionary
 - thematic dictionary
 - lemma type oriented dictionary
 - onomastic dictionary
 - user group oriented dictionary
 - college dictionary
 - learners' dictionary
 - variety oriented dictionary
- dictionary use
 - access
 - accessibility
 - adaptability

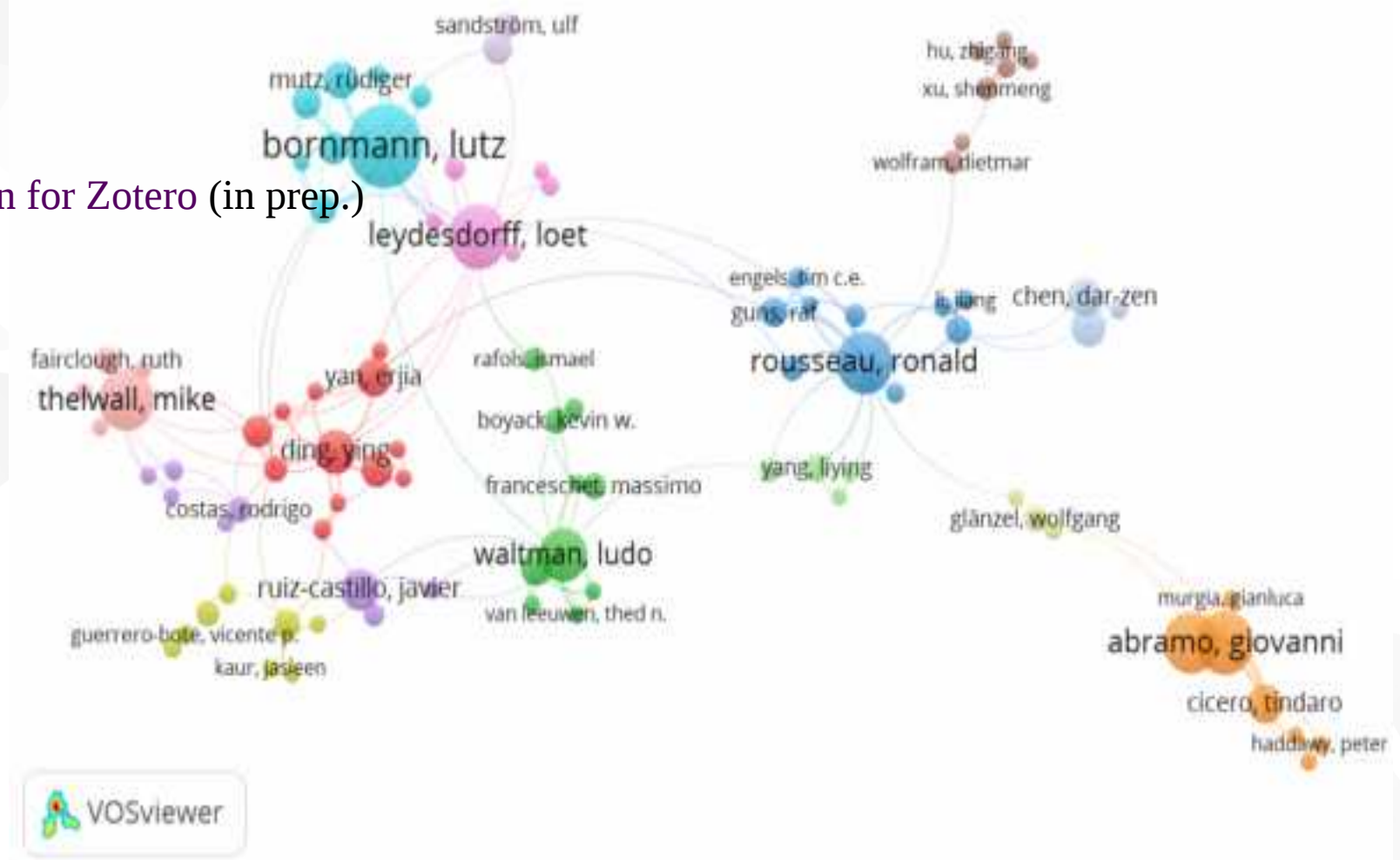


Citation Network

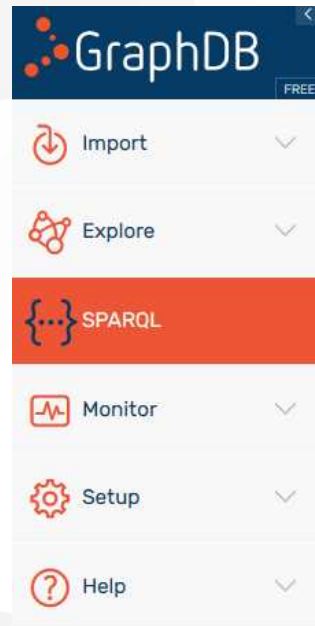
Relevant tools

- ▼ GROBID
- ▼ LocDB
- ▼ VOSviewer
- ▼ WikiCite addon for Zotero (in prep.)

image: VOSviewer



Export for Elexifinder discovery platform



19 # (variant with author lists as JSON objects, which need to be par LexBibRDF

Table Raw Response Pivot Table Google Chart Save

Chart Config Showing results from 1 to 1,000 of 5,597. Query took 4.3s, r

uri	collection	title	article TM	zotItemUri	authorsJson
http://doi.org/10.1007/s40607-014-0001-4	9	"Bottom-up" approach in making verb entries in a monolingual Indonesian learner's dictionary	Jul 1, 2014, 2:00:00 AM	http://zotero.org/groups/1892855/items/GMPZGRY3	[{ "uri" : "http://lexbib.org/person/AmaliaDora", "name": "Amalia" }]
http://doi.org/10.1007/s40607-014-0002-3	9	Korean and English 'dictionary' questions: what does the public want to know?	Jul 1, 2014, 2:00:00 AM	http://zotero.org/groups/1892855/items/55VIRIVB	[{ "uri" : "http://lexbib.org/person/BaeSusanna", "name": "Susanna Bae" }, { "uri" : "http://lexbib.org/agents/p/NesiHilary", "name": "Hila" }]
http://doi.org/10.1007/s40607-014-0003-2	9	Towards improved coverage of Southeast Asian Englishes in the Oxford English Dictionary	Jul 1, 2014, 2:00:00 AM	http://zotero.org/groups/1892855/items/SCMFDQUD	[{ "uri" : "http://lexbib.org/person/SalazarDanica", "name": "Danica Salazar" }]
http://doi.org/10.1007/s40607-014-0004-1	9	An analysis of the smartphone dictionary app market	Jul 1, 2014, 2:00:00 AM	http://zotero.org/groups/1892855/items/Y2UBQ9QJ	[{ "uri" : "http://lexbib.org/person/WinestockChrist", "name": "Christopher Win" }, { "uri" : "http://lexbib.org/agents/p/JJeongYoungkuk", "name": "Youngkuk Jeong" }]
http://doi.org/10.1007/s40607-014-0005-0	9	Kathy Rooney (English) Kong Qi (Chinese): The English-Chinese Encarta Dictionary 2011	Jul 1, 2014, 2:00:00 AM	http://zotero.org/groups/1892855/items/GQY54PBJ	[{ "uri" : "http://lexbib.org/person/LanLi", "name": "Li" }]

Custom JSON format

Elexifinder:

<http://finder.elex.is>



Wikidata, and own Wikibase

Wikidata

- Publication metadata
- Places
- Organizations
- Persons
- Events



Own wikibase instance

- as sandbox
- as collaborative workbench
- for instant data exhibits and visualisations
- for harvesting from and feeding to wikidata
- May replace actual RDF database



Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Concept URI

Item Discussion

Euralex 1990 (Q100594659)

Academic conference about Lexicography

In more languages

Language	Label	Description
English	Euralex 1990	Academic conference about Lexicography
German	No label defined	No description defined
Spanish	No label defined	No description defined
Basque	No label defined	No description defined
Latin	No label defined	No description defined

Statements

instance of  academic conference
0 references

part of  Euralex conference series
0 references

location  Málaga
0 references



Outlook

- ▼ Elexifinder update, very soon
- ▼ Zotero collection updates
- ▼ Publicly accessible wikibase
 - ▼ Author information
 - ▼ Event information
 - ▼ { ... }
- ▼ Follow us
 - ▼ on ResearchGate
 - ▼ on lexbib.org
- ▼ DARIAH Annual Event
WG Bibliographical Data workshop
"Bibliographical Data Workflows",
November 25, 2020, 14:00-15:30 CET
- ▼ Two possible topics in one session slot
 - ▼ "Zotero etc. help you a lot - but where are the limits?"
 - ▼ "Author disambiguation using Linked Open Data"



eman ta zabal zazu



UPV EHU

This project has received funding from
IT1169-19 Consolidated Research Group of Excellence
(2019-2021), Basque Government



Contact: david@lexbib.org / david.lindemann@ehu.eus

ResearchGate: <https://www.researchgate.net/project/LexBib-Corpus-and-Bibliography>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015.

DARIAH Annual Event 2020