

Building a European Literary Bibliography

Challenges and Strategies in Integrating and Presenting Multilingual Bibliographical Data

Ondřej Vimr & Vojtěch Malínek

Czech Academy of Sciences, Institute of Czech Literature, Czech Literary Bibliography
Linked Open Data and Literary Studies, Freie Universität Berlin, 19th November 2024



European Literary Bibliography

- Bibliographical data as an essential resource
(Lahti et al. 2019, Umerle et al. 2022)
- Mission: ecosystem for **harmonising, enriching and presenting open literary bibliographical datasets**
- Developed jointly by Czech Literary Bibliography and Polish Literary Bibliography
<https://literarybibliography.eu>



Historical Overview

2019 – first demoversion containing of Czech and Polish Literary Bibliographies records

2021-2023 – DARIAH.lab project of IBL PAN

2022 – European Literary Bibliography 1.0 realised

2022 – inclusion of Spanish and Finnish data

2023 – work on BOSLIT data

2024 – European Literary Bibliography 2.0 under preparation + Czech National Bibliography data processed

ELB - Challenges



Bibliographical data aggregation (literature, CCo/BY, live data)



Multilingual semistructured data (diverse formats, national vs. universal identifiers, subject heading ontologies, etc.)



Data enrichment and LOD-ification (geolocation, WikiData interconnection, etc.)



Data presentation

Data Sources

Bibliographical (catalogue) data

- Spanish National Library
- Finnish National Library
- Czech Literary Bibliography (CLB)
- Polish Literary Bibliography (PBL)
- Bibliography of Scottish Literature in Translation (BOSLIT)

Bibliodata formats

```
=LDR 00821nam a2200265 c 4500
=001 a5933628
=005 20190125005347.6
=008 170403s2016\\\sp\\\\\\\\\\000\\spa\\\
=015 \\$aMON16
=017 \\$az 1435-2016$bOficina Depósito Legal Zaragoza
=020 \\$a978-84-946134-0-1
=035 \\$a(OCoLC)1027700728
=040 \\$aSpMaBN$bspa$cSpMaBN$erdc
=080 \\$a821.134.2-14"19/20"(082.2)
=245 00$aAntiaérea II :$bpoesía encontrada en Zaragoza$c[Ajo ... et al.]
=250 \\$a1ª ed.
=260 \\$aZaragoza$bPregunta$c2016
=300 \\$a104 p.$c18 cm
=336 \\$aTexto (visual)$2isbdcontent
=337 \\$asin mediación$2isbdmedia
=380 \\$iMajor genre$aLiterature$leng
=381 \\$iMajor genre$aLyrical poetry$leng
=650 \7$aPoesías españolas$yS.XX$2embne
=650 \7$aPoesías españolas$yS.XXI$2embne
=650 \7$aPoetry$2ELB-g
=650 \7$aSpanish literature$2ELB-n
=655 \7$aPoesías$2tgfbne
=655 \7$aAntologías$2tgfbne
=700 0\\$aAjo$4aut
=995 \\$aBiblioteca Nacional de España
```

```
{
  "creator": [
    "Casado Velarde, Manuel",
    "Vilarnovo, Antonio"
  ],
  "date": [
    "2003"
  ],
  "format": [
    "application/pdf"
  ],
  "identifier": [
    "https://dialnet.unirioja.es/servlet/oaiart?codigo=798509",
    "(Revista) ISSN 0210-8178"
  ],
  "language": [
    "spa"
  ],
  "publisher": [
    "Universidad de Extremadura: Servicio de Publicaciones"
  ],
  "rights": [
    "[...]"
  ],
  "source": [
    "Anuario de estudios filológicos, ISSN 0210-8178, Vol. 26, 2003, pags. 5-11"
  ],
  "title": [
    "Eugenio Coseriu: \"In memoriam\""
  ],
  "type": [
    "text (article)"
  ]
},
```

Data Sources

Bibliographical (catalogue) data

- Spanish National Library
- Finnish National Library
- Czech Literary Bibliography (CLB)
- Polish Literary Bibliography (PBL)
- Bibliography of Scottish Literature in Translation (BOSLIT)

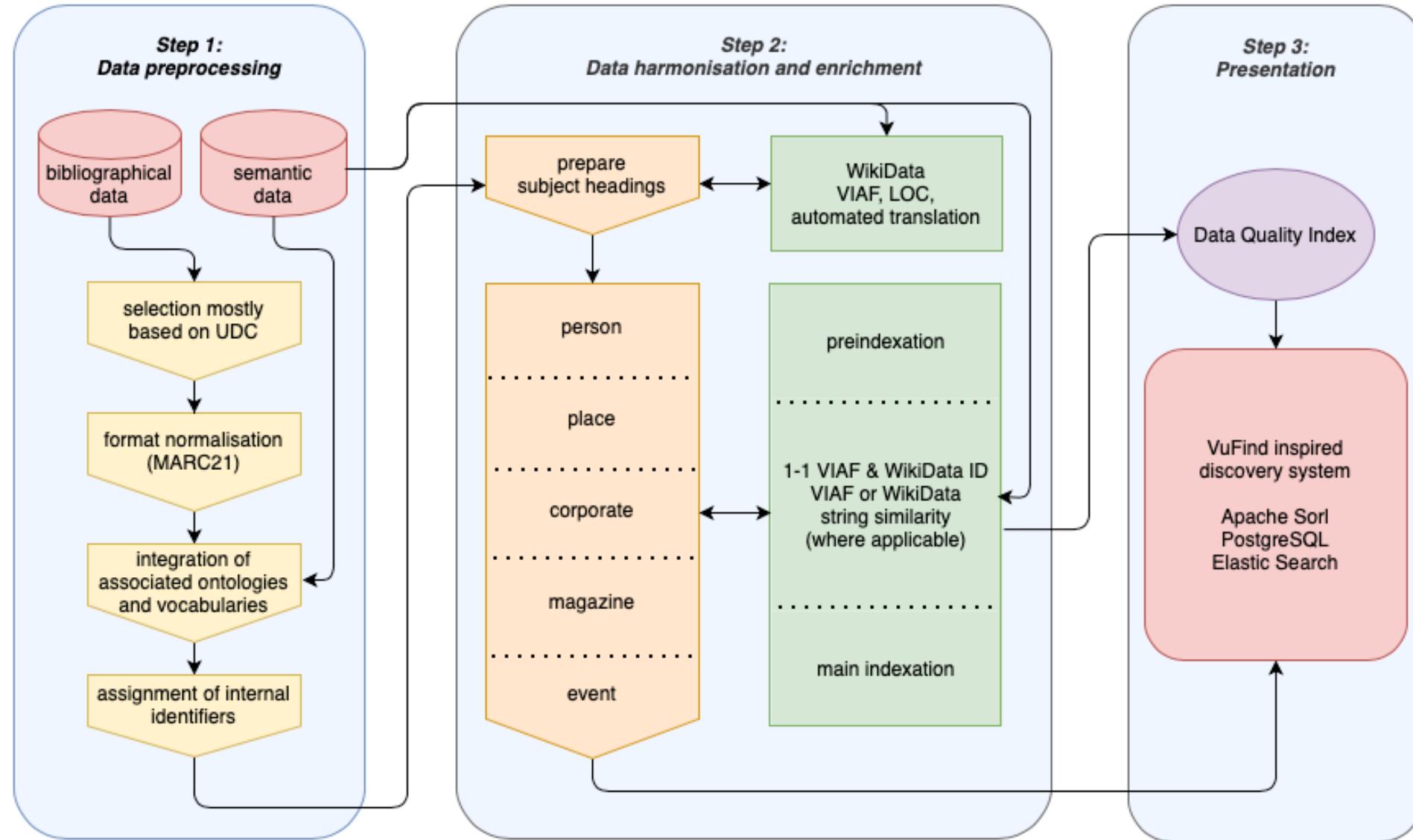
Semantic data

- Library reference ontologies (see [example](#))
- Library of Congress ID (subject headings) (see [here](#))
- Virtual International Authority File ([VIAF.org](#))
- WikiData ([wikidata.org](#))

A-Z	Hierarchy	Groups	New and Deprecated
	<ul style="list-style-type: none"> - information - information overload - information technology architecture - initiatives - insurance - interviews - irony - judgements (philosophy) - judicial objects - justification - Karelian question - language issue - laws of thought - learning opportunities - legal cases - legal order - lifeworld - lines of action - linguistic objects - literature <ul style="list-style-type: none"> - African American literature - Afrikaans language literature - Akkadian language literature - Albanian language literature - Altai language literature - Amharic language literature - Arabic language literature - Aramaic language literature - Armenian language literature - Asian American literature - Assamese language literature - Awadhi language literature - Azerbaijani language literature - Bashkir language literature - Basque language literature - Belarusian language literature - Bengali language literature - Bosnian language literature - Braj Bhasha language literature - Breton language literature - Bulgarian language literature - Buriat language literature - Burmese language literature - Catalan language literature - chapbooks - children's literature - Chinese language literature - Chuvash language literature 		<p>objects > abstract objects > societal objects > literature > African American literature</p> <p>PREFERRED TERM</p> <p>African American literature ⓘ</p> <hr/> <p>TYPE</p> <p>General concept</p> <hr/> <p>BROADER CONCEPT</p> <p>literature</p> <hr/> <p>RELATED CONCEPTS</p> <p>minority literature</p> <hr/> <p>BELONGS TO GROUP</p> <p>87 Literature</p> <hr/> <p>BELONGS TO ARRAY</p> <p>literature by population group</p> <hr/> <p>IN OTHER LANGUAGES</p> <p>afroamerikkalainen kirjallisuus Finnish</p> <p>afroamerikálaš girjjálašvuhta Northern Sami</p> <p>afroamerikansk litteratur Swedish</p> <hr/> <p>URI</p> <p>http://www.yso.fi/onto/yso/p9797 ⓘ</p> <hr/> <p>DOWNLOAD THIS CONCEPT:</p> <p>RDF/XML TURTLE JSON-LD</p> <p>Created 9/10/98, last modified 2/8/23</p> <hr/> <p>Ehdota muutosta käsittelyseen</p> <hr/> <p>CLOSELY MATCHING CONCEPTS</p> <p>African American literature www.wikidata.org</p> <hr/> <p>EXACTLY MATCHING CONCEPTS</p> <p>African American literature KOKO Ontology</p> <p>afroamerikansk litteratur (sv) Allärs - General thesaurus in Swedish</p> <p>afroamerikkalainen kirjallisuus (fi) YSA - General Finnish thesaurus</p>

B //

Finnish ontology (example)



LDR 01151nam a2200325 c 4500
001 a6078069
003 SpMaBN
005 20210326123920.6
008 171214s2016|||||spa|||||||000\f\spa\\
015 \\|a MON16
017 \\|a B 15417-2016 |b Oficina Depósito Legal Barcelona
020 \\|a 978-84-01-01893-0
035 \\|a (OCOlc)1055846156
040 \\|a SpMaBN |b spa |c SpMaBN |e rdc
041 1\\|a spa |h fin
080 \\|a 821.511.111-31"19"
100 1\\|a Waltari, Mika |d (1908-1979) |4 aut |0 xx965743
240 1 |a Sinuhe, egyptiläinen |l Español
245 1 |a Sinuhé, el egipcio |c Mika Waltari , traducción de Manuel
250 \\|a 1ª ed.
260 \\|a [Barcelona] |b Plaza & Janés |c 2016
300 \\|a 600 p. |c 23 cm
336 \\|a Texto (visual) |2 isbdcontent
337 \\|a sin mediación |2 isbdmedia
380 \\|i Major genre |a Literature |l eng
381 \\|i Major genre |a Fiction |l eng
594 \\|a Título original: Sinuhe, egyptiläinen
655 \7|a Novelas históricas |2 tgfbne
700 1\\|a Bosch Barrett, Manuel |d (1895-1961) |e traductor |4 trl
995 \\|a Biblioteca Nacional de España

{
• 381: [1 item],
• sourceFile: "es_ksiazki_08-02-2024.mrk",
• editTime: "2024-05-28 16:07:01",
• LEADER: "01151nam a2200325 c 4500",
• majorFormat: "Book",
• rawId: "a6078069",
• id: "es.a6078069",
• language: {3 items},
• isbn: [1 item],
• OCoLC: [1 item],
• ctrlNum: "(OCOlc)1055846156",
• subject: {3 items},
• wikiQ: {
◦ Q193111: {
▪ mainAuthor: 1
},
◦ Q1492: {
▪ publicationPlace: 1
},
◦ Q93427686: {
▪ coAuthor: 1
}
},
• origin_labels: [3 items],
• persons: {3 items},
• titleShort: "Sinuhé, el egipcio",
• title: "Sinuhé, el egipcio",
• StatementOfResp: [],
• edition: {2 items},
• places: {2 items},
• corporates: {2 items},
• publicationDateStr: [1 item],
• publicationYear: [1 item],
• publishedIn: [1 item],
• major_genre: [1 item],
• sourceDB: {1 item}
}

B///

Own data structure (enriched, LODified): [record here](#)

Discovery system

BETA version

European Literary Bibliography

Any field

Bibliographic records
4,583,397

Persons
142,609

Places
15,250

Česká literární bibliografie
Czech Literary Bibliography

Ústav pro českou literaturu AV ČR
Institute of Czech Literature of the CAS

BIBLIOTHECA

PBL
POLSKA BIBLIOTECJA LITERACKA

DIGITAL HUMANITIES CENTRE

based on VuFind



Any field

mika waltari

X Q ↗



Narrow Search

▼ ▲ <

Bibliographic records | Total results: 3.282, showing: 1 - 20

Results per page: 20 ▾

view: list ▾

Sorting: Relevance ▾

Source Database



- National Library of Finland 2.902 >
- Polish Literary Bibliography 180 >
- Czech Literary Bibliography 141 >
- Biblioteca Nacional de España 59 >

 Select all Download Print 1

Mika Waltari mielikuvituksen jätiläinen

In: WSOY, 1982

Published: Porvoo ; Helsinki ; Juva : WSOY, 1982 (Juva)

Book

 2Mika Waltari Sinuhe egyptiern : utställning i Uppsala universitetsbibliotek i samarbete med Helsingfors universitetsbibliotek 16.11.1979 - 2.2.1980 : catalogue of the exhibition at the University(...)

In: Uppsala universitetsbibliotek, 1979

Published: [Uppsala] : Uppsala universitetsbibliotek, 1979.

Book

 3

Mika Waltari ja taiteilijaystäväät kirjailija kuvataiteen kuraattorina

In: WSOY, 2008

Published: Helsinki : WSOY, 2008 (Porvoo : WS Bookwell)

Book

Major Genre

 "Traditional view of bibliographic records" for [Mika Waltari](#).

Format

Journal article

3.353.494

Book

1.181.495

Book chapter

108.624

Other

7

Author

Main author

Co-author

Corporate Author

Events



□ 2

Werfels

Search in Source Document

11K

Lidové noviny

120.552

Venkov

82.853

Právo lidu

47.928

České slovo

44.830

Zvon

35.719

Národní Osvobození

27.507

Source Document

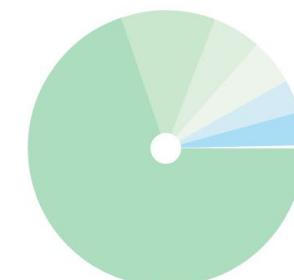
noviny



Sort: Result count Alphabetical

□ 4

<input type="checkbox"/>	Lidové noviny	120.552
<input type="checkbox"/>	Lidové noviny	19.094
<input type="checkbox"/>	Literární noviny	9.766
<input type="checkbox"/>	Literární noviny	9.318
<input type="checkbox"/>	Národní noviny	6.717
<input type="checkbox"/>	Zemědělské noviny	6.715



ELB facets and related tools

PieGraph includes about 3.7% of all results.



A black and white photograph of Mika Waltari, a Finnish writer. He is shown from the waist up, wearing a dark suit jacket, a light-colored shirt, a dark tie, and a fedora hat. He is looking slightly to his left.

00:00 | 00:00

Finnish writer

1099 597 G

Sex or Gender: male

Date of Birth: 1908-09-19

Place of Birth: Гельсингфорс (at present: Helsinki)

Date of Death: 1979-08-26

Place of Death: Хельсинки (at present: Helsinki)

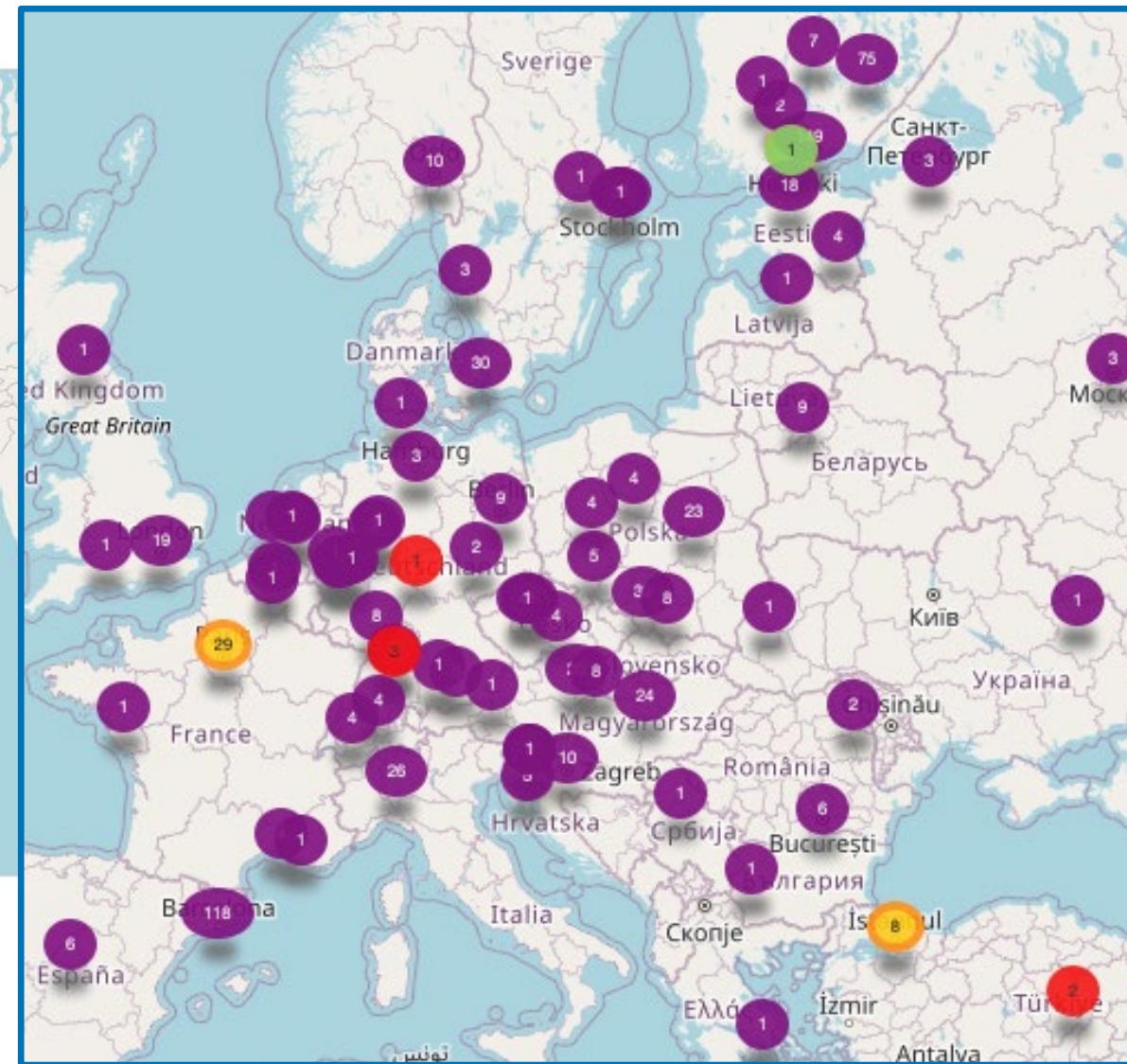
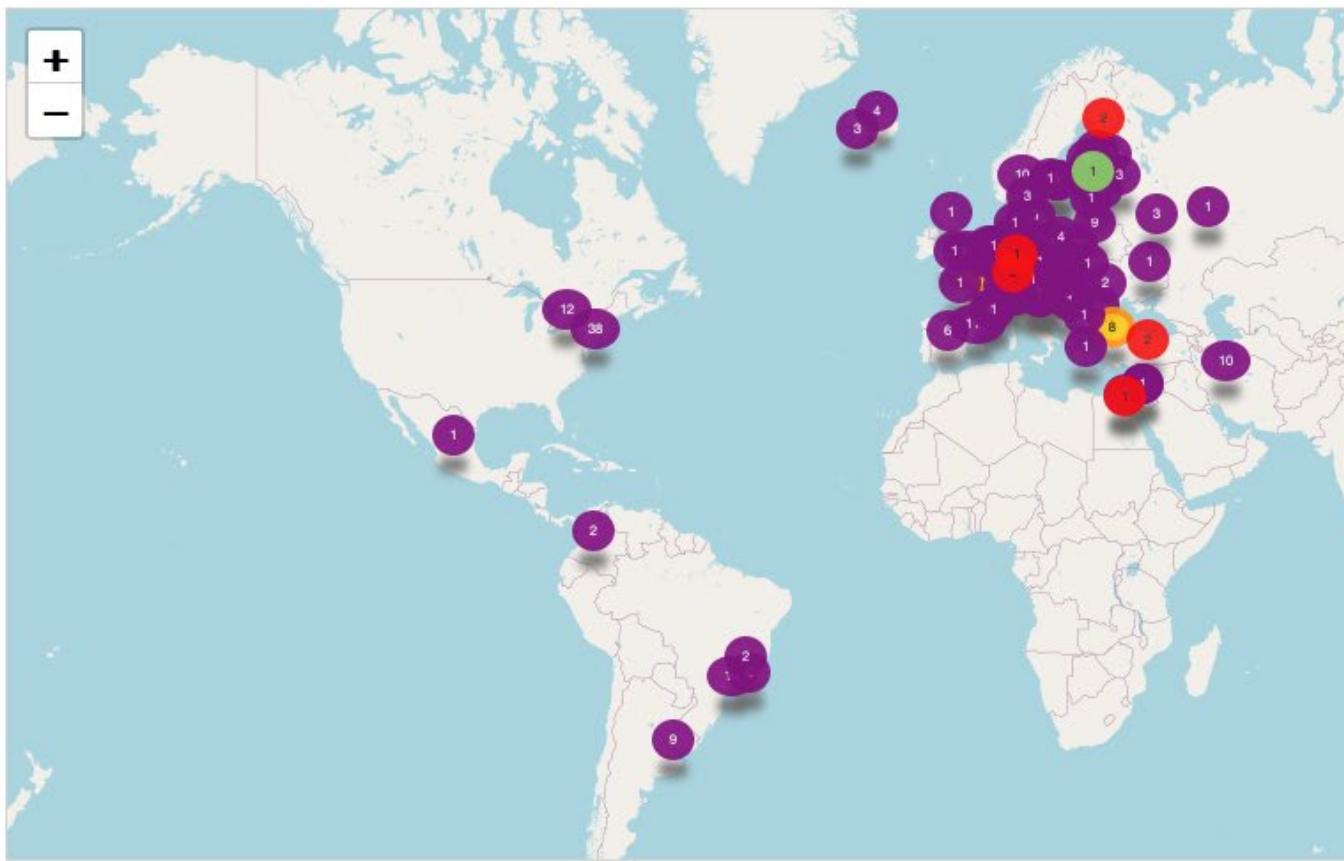
Country of citizenship: Finland

Childrens: Satu Waltari

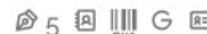
Occupation: writer
translator
journalist
novelist
screenwriter
poet
playwright

Genres: historical novel
crime novel

Viaf ID: 111008219

**B**///Bibliographical records related to PERSON ([Mika Waltari](#)) according to the respective capacity.

Milan Kundera was the subject of publications in which the main author was:



Marek Bieńczyk (1956-)

Polish writer and translator

[Go to bibliographic records](#)



Aleksander Kaczorowski

(1969-)

Polish translator, essayist and journalist

[Go to bibliographic records](#)



Jakub Češka (1971-)

Narozen 1971. Literární vědec, vysokoškolský pedagog, práce z oboru.

[Go to bibliographic records](#)



Krystyna Kofta (1942-)

Polish writer

[Go to bibliographic records](#)

Milan Kundera

They were the subject of publications in which the main author was Milan Kundera



Franz Kafka (1883-1924)

Bohemian writer from Prague (1883-1924)

[Go to bibliographic records](#)



Ernest Hemingway (1899-1961)

American author and journalist (1899-1961)

[Go to bibliographic records](#)



Johann Sebastian Bach

(1685-1750)

German composer (1685-1750)

[Go to bibliographic records](#)



Francis Bacon (1561-1626)

English philosopher and statesman (1561-1626)

[Go to bibliographic records](#)



Place of publication of works with subject place Barcelona 233

Subject place of works published in Barcelona 331

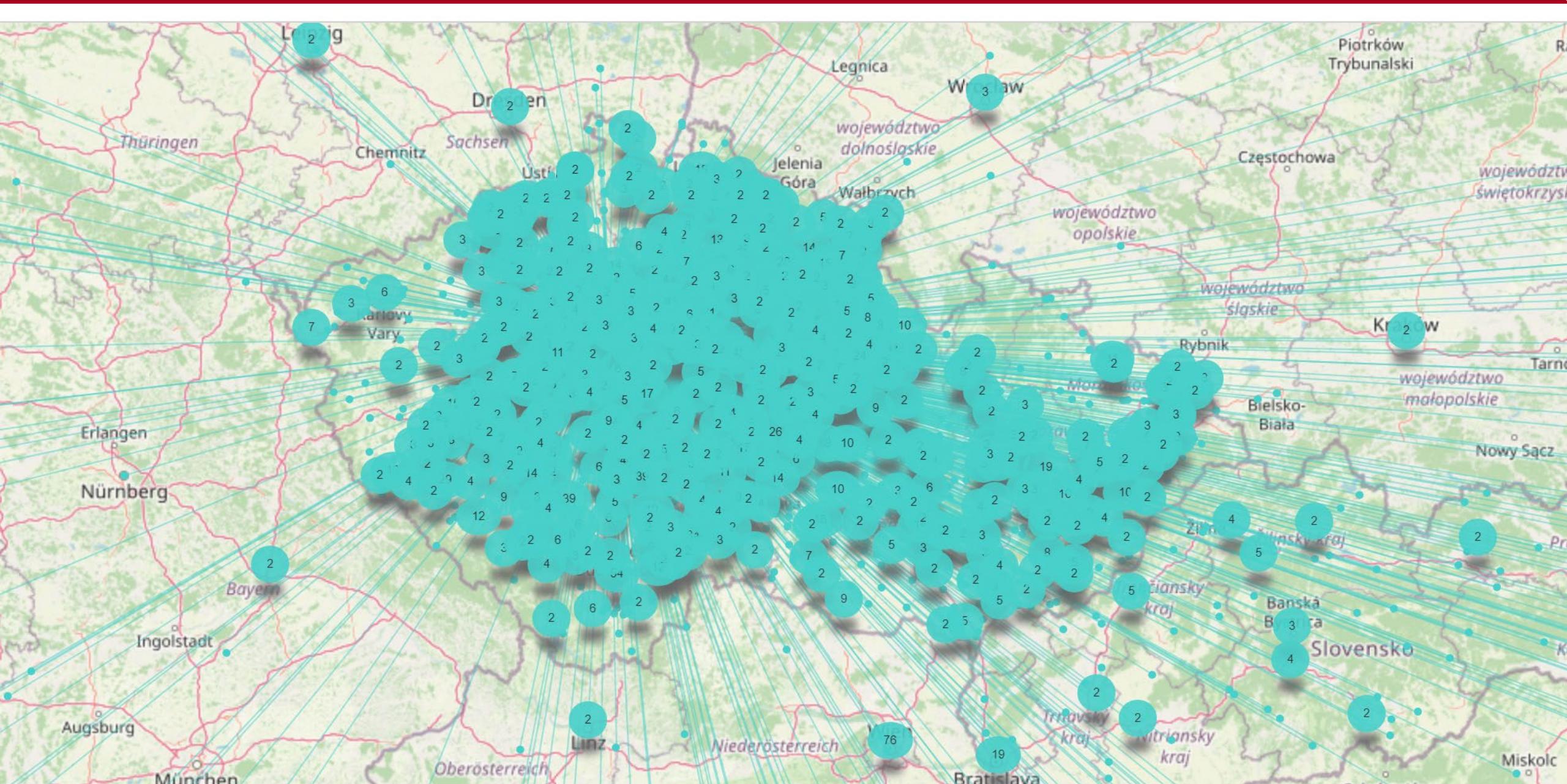
Birthplace of newcomers 24

Place of death of "emigrants" 18

to do: publication places of works of persons live-related (born/died/residence) with Barcelona 21

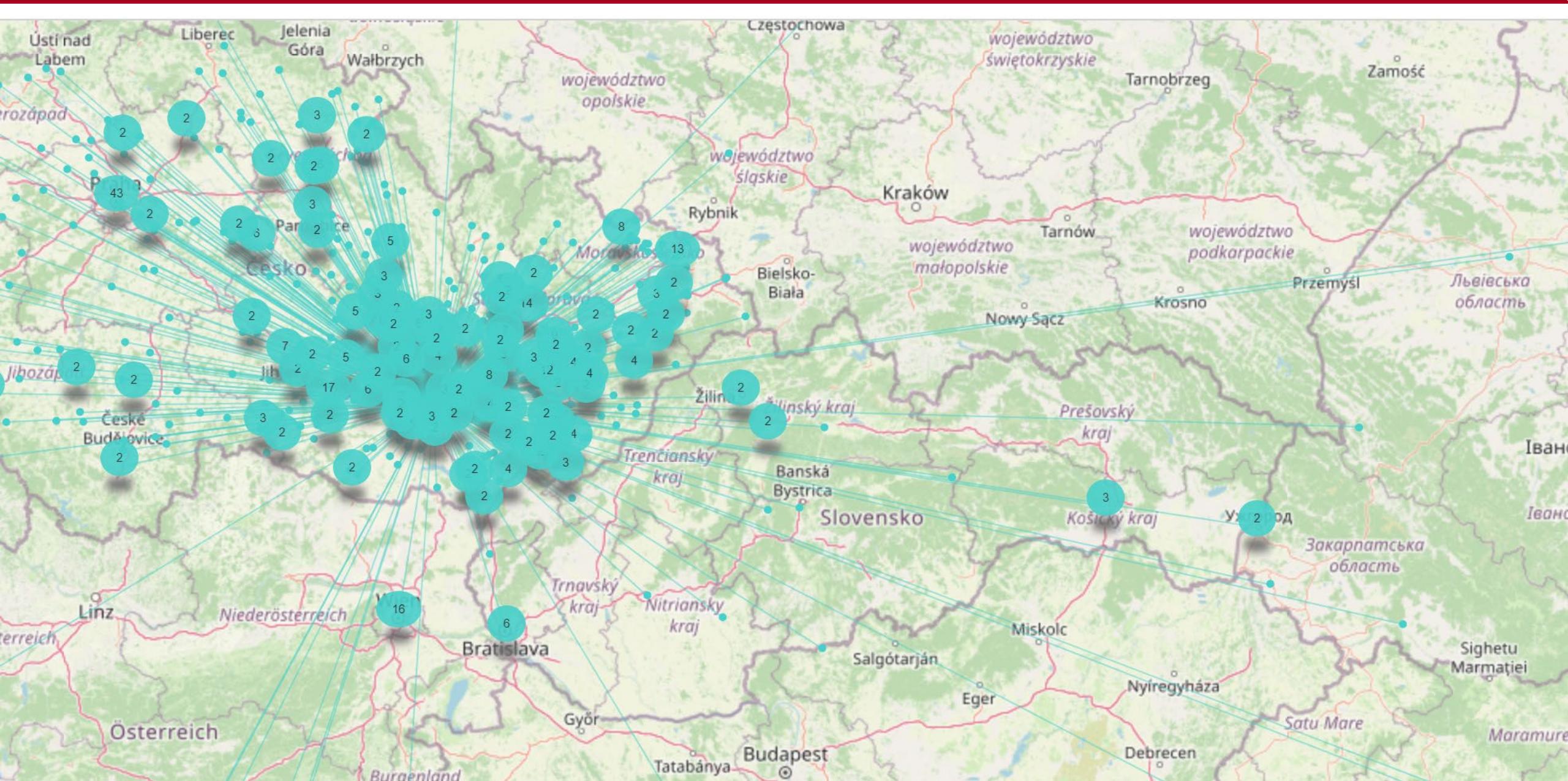
B///

Bibliographical records related to PLACE ([Barcelona](#)) according to their respective capacity.



B///

Birthplaces of the literary active people who died in [Prague](#).



B

Birthplaces of the literary active people who died in Brno

Pre-harmonisation and unification bibliographic record labels linked to this record:

Name (a)	Dates range (d)	Has VIAF	Other ID	Appearances	Certainty of fit
Cervantes Saavedra, Miguel de	(1547-1616)	✓		6.274	
Cervantes Saavedra, Miguel de	(1547-1616)	✓	jn19990001383	430	
Cervantes Saavedra, Miguel de	(1547-1616)	✓		182	
Cervantes Saavedra, Miguel de	(1547-1616)	✓		8	
Cervantes Saavedra, Miguel de	(1547-1616)	✓		3	
Cervantes.				3	
Figura, Michael				3	
we				1	
Cervantes Saavedra, Miguel de	(1547-1616)	✓	aun2006372705	1	

 Above you can see a list of all the variants of how this record was written in the bibliographic records.
The **Certainty of fit** column shows how "confident" we are that this label should be here.

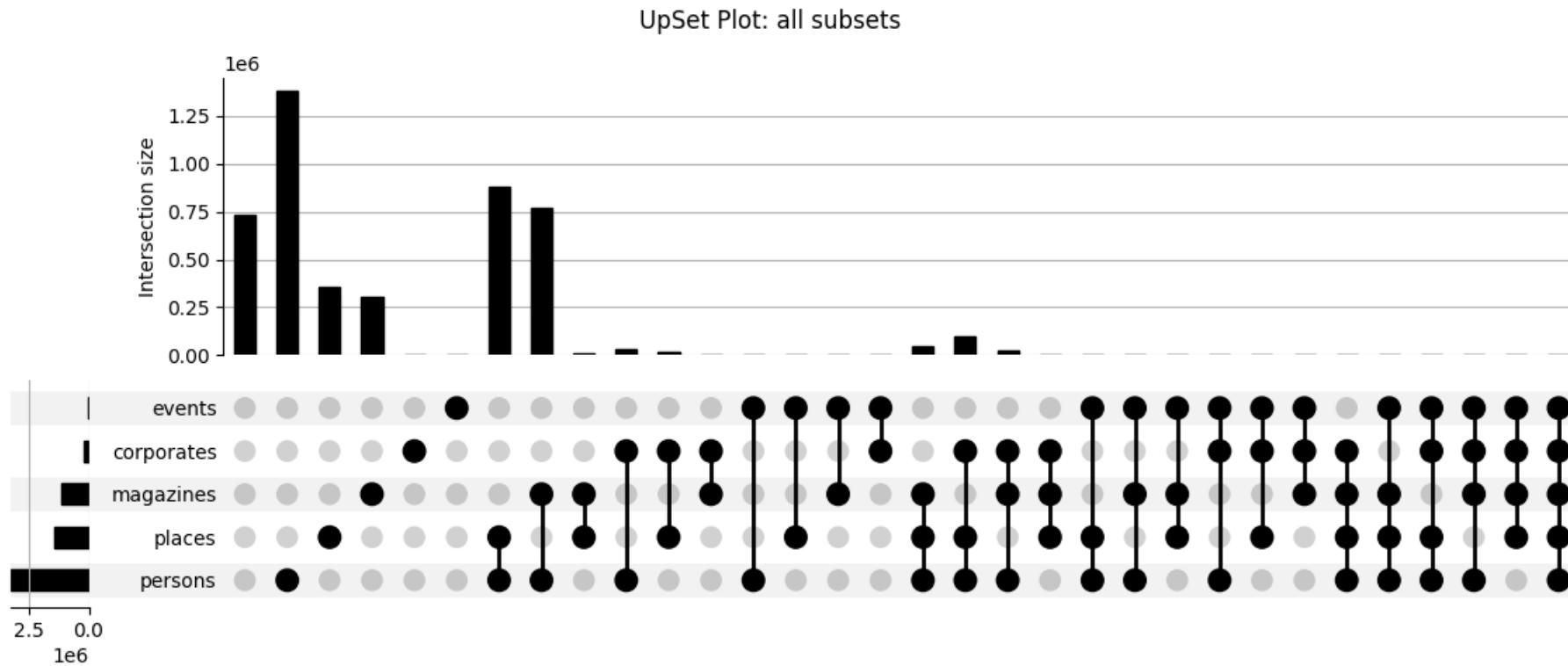
Labels on VIAF:

Cervantes Saavedra, Miguel de 1547-1616	31
Cervantes Saavedra, Miguel de	5
Cervantes, Miguel de, 1547-1616	2
Cervantes y Saavedra, Miguel de, 1547-1616	1
Сервантес Сааведра, М. де 1547-1616 Мигель де	1
Cervantes Saavedra, M. de 1547-1616 Miguel de	1
سرفانتس، میگیل دی، 1547-1616	1
1547-1616, מיגל דה, סרונטס	1
Сервантес Сааведра, Мигель Де, 1547-1616	1
Miguel de Cervantes Saavedra	1
de Cervantes Saavedra, Miguel	1
Cervantes Saavedra, M. de	1
ثربانتس ساابدرا، میجل دی، 1547-1616	1
Cervantes de Saavedra, Miguel, 1547-1616	144C

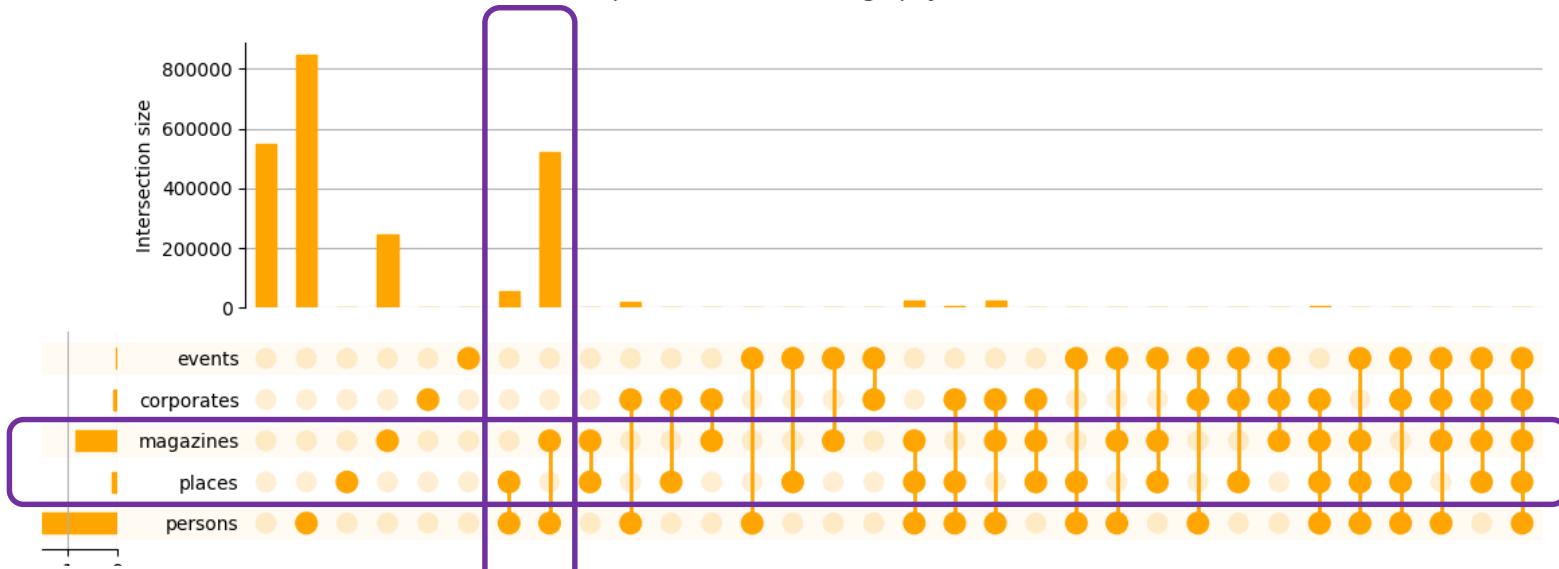
Data identification and links for 4,611,498 bibliographical entries

category	records with data	unique items	links total
Person	3,745,450 (81%)	279,259	6,381,219
Place	1,440,803 (31%)	8,787	1,784,478
Corporate	1,376,301 (30%)	7,206	1,598,087
Magazine	2,508,598 (54%)	882	2,667,997
Event	62,007 (1%)	314	56,796

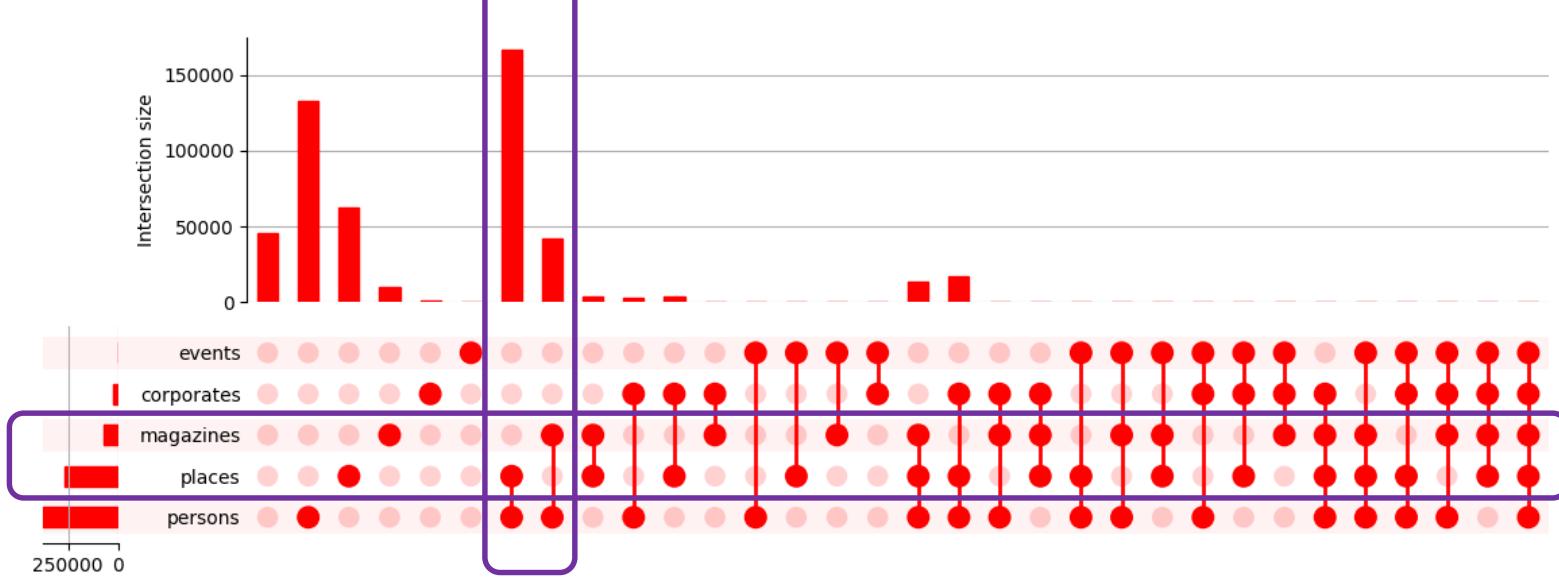
Intersections of sets and their sizes



UpSet Plot: Czech bibliography



UpSet Plot: Finnish bibliography



Issues and Challenges

- Limits of non-interventionist approaches
 - Inclusion of more datasets
 - Definition of literature (primary & secondary)
 - Curate or not curate
 - Live data and updates (original datasets, wiki, VIAF)
- Technological sustainability
- Data format backward compatibility
- Wikidata corruption (including hacking)

Helsinki Helsingfors, Helsingia, Helsingin kaupunki, Stadi, Hesa



capital and most populous city of Finland

Country: [Finland](#)

Coordinates (WGS-84): [60.17556,24.93417](#)

Country: at present

Finland

1918-05-16

Finnish Socialist Workers' Republic

1917-12-06

Russian Empire

1809-09-05

Sweden

1550

Historical names:

at present

ru Хельсинки

1917-12-05

ru Гельсингфорс

long long time ago

Source of information [Wikidata](#)

More information on [Wikipedia](#)

Preferred Forms

-  110 2 _ †a Charta 77
-  110 2 _ †a Charta 77
-  110 2 _ †a Charta 77 (Group)
-  110 2 _ †a Charta 77 (Group)
-  110 2 _ †a Charta 77 (Group)
-  110 2 _ †a Charta 77 (Group)
-  110 2 _ †a Charta 77 (Prague).(1977-1989)
-  110 2 _ †a Charta 77 (musikgrupp)
-  110 2 _ †a Charta 77 (skupina)
-  110 2 _ †a Charta 77 (skupina)
-  110 2 _ †a Charta 77 Foundation (Prague)
-  210 || †a Charta 77 †c Prague
-  110 2 _ †a Charta 77 †g Körperschaft
-  **B**  Bibliographical records related to CORPORATION ([Charta 77](#)) demonstrating corrupt data (two different entities).



Literarybibliography.eu 2.0

- Under preparation
- New modules and tools
- Works on backend -> deduplications, data enrichment
- Licences
- New datasets:
 - Czech National Bibliography (books)
 - Bibliography of Scottish Literature in Translation



European Literary Bibliography



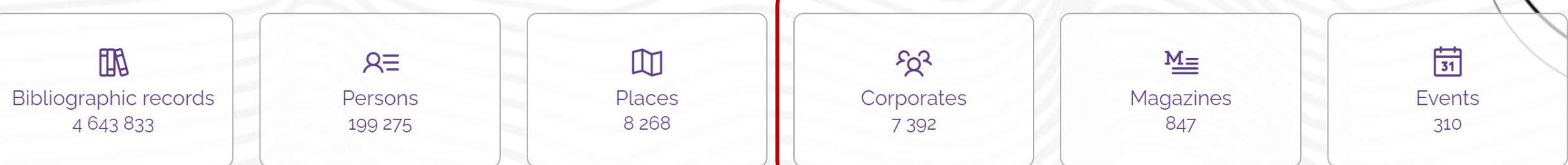
Search in bibliographic records:

Any field

▼ Search for...

Q Search

Advanced search



New ELB modules under preparation.

Sources

The Czech Literary Bibliography, created at Institute of Czech Literature of the Czech Academy of Sciences, a public research institution, is a basic research infrastructure for literary Czech studies and related humanities (e.g. philology, philosophy, history, art history and journalism). Its objective is to process, continuously edit and update and place online a unique set of bibliographical and other knowledge databases, seamlessly mapping out literary events and literary studies in the Czech lands and the circulation of Czech literature abroad from the beginning of the National Revival to the present, regardless of language, format or medium. Currently, Czech Literary Bibliography databases contain records on 2,250,000+ articles on literature and literary life in the Czech lands since 1770 up to the present days.

Licence: **CC BY-NC-SA 4.0** When using the CLB databases, you agree to acknowledge the use of the CLB resources in all your work. Example: "During the work on this research paper [book/presentation/text...] use was made of the Czech Literary Bibliography research infrastructure – <https://clb.ucl.cas.cz/> (ORJ identifier: 90243)."

Polish Literary Bibliography is created and compiled at the Institute of Literary Research of the Polish Academy of Sciences, a public research institution, that serves as a research infrastructure for literary studies and related humanities (e.g. philology, philosophy, history, art history and journalism). Its main objective is the curation of Polish literary bibliographical data. It registers, processes and updates a unique bibliography containing books published in Poland and abroad in Polish. For each year-book, PBL compiles approximately 850 periodical titles: national dailies, national literary, cultural, cultural and social periodicals, serials published by



Data licencing and datasets for download.

cz_articles0_08-02-2024.mrk
cz_articles1_08-02-2024.mrk
cz_articles2_08-02-2024.mrk
cz_articles3_08-02-2024.mrk
cz_articles4_08-02-2024.mrk
cz_books_08-02-2024.mrk
cz_chapters_08-02-2024.mrk
cz_ucla_2024-02-20update.mrk

pl_bn_articles2_08-02-2024.mrk
pl_bn_articles_08-02-2024.mrk
pl_bn_books2_08-02-2024.mrk
pl_bn_books_08-02-2024.mrk
pl_bn_chapters_08-02-2024.mrk
pl_bn_chapters_08-02-2024.mrk
pl_pbl_articles_08-02-2024.mrk
pl_pbl_books_08-02-2024.mrk

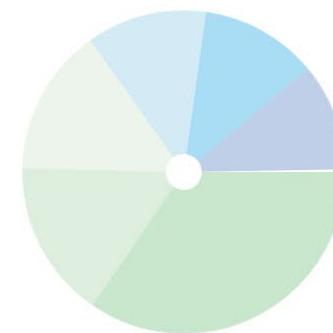
Subject

search in options

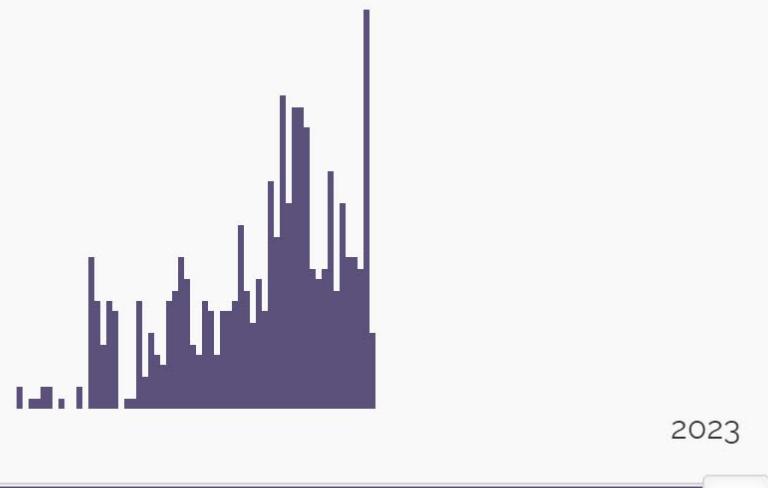


Options shown on chart:
6 /2511

čeští spisovatelé	2.582
české drama	1.181
články	1.099
rozhovory	904
zprávy	876
recenze	804
politika	787
osobnosti	696



Publishing date



Show results for range 1964 - 2023

Relations were created on the basis of entries in bibliographic records. Number of records in which the Václav Havel was found: **7610**.

Václav Havel appears in roles:

subjectPerson	6.917
mainAuthor	701
coAuthor	309

Show relations with:

Persons	37.010
sourceMagazine	3.265
Corporates	2.085
Places	1.738
Events	108

Persons appears in roles:

subjectPerson	17.431
mainAuthor	4.732
coAuthor	6.920

Total results **6225** showing **18**



8 4.9K 591 105 G

Milan Kundera
Franco-Czech poet and novelist
(1929-2023)

Show bibliographic record with Václav Havel and with Milan Kundera as

subjectPerson 365 mainAuthor 5 coAuthor 15



8 2.2K 818 211 G

Ludvík Vaculík
Czech publicist, politic writer and bookwriter (1926-2015)

subjectPerson 263 mainAuthor 15 coAuthor 88



8 2.2K 657 135 G

Pavel Kohout
Czech poet, playwright, exil writer, film director, politic writer and bookwriter

subjectPerson 216 mainAuthor 29 coAuthor 57

[Marc view](#)[ELB fields](#)[Map](#)[Other versions](#)

- sourceFile: "es_ksiazki_08-02-2024.mrk",

- sourceDB: {

- licence: {

- name: "CC0",

- link: "https://creativecommons.org/publicdomain/zero/1.0/legalcode".

- description: ""

- },

- name: "Biblioteca Nacional de España"

- },

- editTime: "2024-09-23 15:12:32",

- LEADER: "00949nam a2200277 4500",

- majorFormat: "Book",

- rawId: "bimo0000964468",

- prefix: "es",

- id: "es.bimo0000964468",

- language: {3 items},

- isbn: [1 item],

- OCOLC: [1 item],

- ctrlNum: "(OCOLC)431558018",

- subject: {1 item},

- wikiQ: {

- Q905: {

- mainAuthor: 1

- },

- Q1492: {

- publicationPlace: 1

- places: {

- publicationPlace: {

- Q270: {

- name: "Warszawa",

- ids: [],

- biblio_label: "Warszawa||",

- wikiQ: "Q270",

- nameML: "Warsaw|Varšava|Warszawa|Varsova|Varsovia"

- mainAuthor: {

- Q868: {

- wikiQ: "Q868",

- viaf: "7524651",

- year_death: "",

- year_born: "",

- dates: "",

- nameML: "Aristotle|Aristotelés|Arystoteles|Aristoteles|Aristóteles",

- name: "Aristóteles",

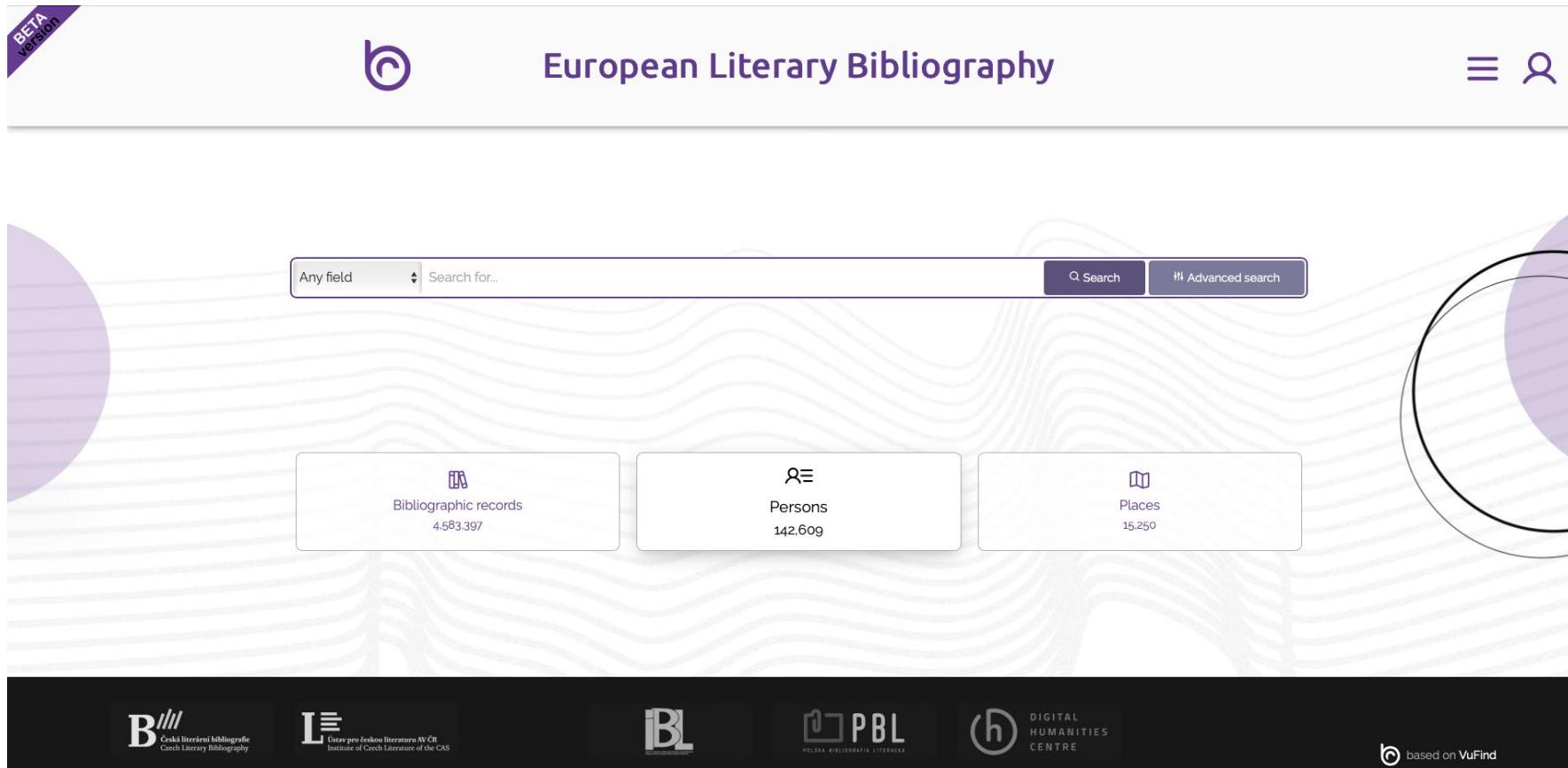
- ids: {1 item},

- role: "mainAuthor",

- roles: [],

- biblio_label: "Aristóteles||7524651||"

Please test, provide feedback and join:
<https://literarybibliography.eu/>



Mgr. Ondřej Vimr, Ph.D.
vimr@ucl.cas.cz

Mgr. Vojtěch Malínek, Ph.D.
malinek@ucl.cas.cz



Czech Academy of Sciences, Institute of Czech Literature
Na Florenci 1420/3, 110 00 Praha 1, Czech Republic
T +420 222 828 132, +420 770 136 131 E clb@ucl.cas.cz W clb.ucl.cas.cz

The Czech Literary Bibliography Research Infrastructure is fully supported by the Ministry of Education, Youth and Sports of the Czech Republic under the large research infrastructures programme (project LM2023043).

