SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

INFORMACJA I KOMUNIKACJA

What and how can you learn from great bibliographies?

prof. dr hab. Adam Pawłowski
University of Wrocław

dr inż. Tomasz Walkowiak
Wrocław University of Science and Technology
28th November 2023

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Presentation plan

- **SHOW FROM WEB [http://phc.uni.wroc.pl/wbw]**
- **DATA**
  Bibliography as a source of data
  Structured data and MARC format
  National Library databases
- **METHOD**
  Automatic language processing (NLP)
  Digital humanities and data science
- **RESULTS**
  Bibliomap as a historical cross-section of the development of Poland
  The world in Polish publications
  Gender vs topics and genres of publications
  Analysis of publishing dynamics in Polish cities

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Data - types of bibliography

Colloquially, a bibliography is a short appendix to a document.

For a computer scientist, it is a structured corpus of data.

The basic division of bibliography types includes:

- general (national);
- special (limited range, scope, type of publication)
- field (subject)
- mixed

Joking in the Orwellian style, one should say:
all bibliographies are equally interesting, but national bibliographies are more interesting.

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Bibliography as a source of data

Traditionally, the recipient of bibliographic information is the user who seeks publications on a given topic.

De facto, therefore, it receives not information, but metainformation.

BUT...

A sufficiently comprehensive, structured corpus of data is also a source of knowledge.

Thus, in the world of digital humanities, large bibliographies are also a primary source of information and knowledge about the world.

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Formats

This type of database is dominated by the MARC, or Machine-Readable Cataloging, family of formats, ISO 2709:2008 standard.

The first version was created at the Library of Congress as early as in the 1960s, and has been the standard in the US since 1971.
The most widely used version today is MARC 21.

Most of mankind's written heritage is thus described by a standard that is now more than 60 years old.

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Digression: how many years longer?

INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

MARC 21 format – example of instructions

041 Język dokumentu (P)

Wskaźniki

Pierwszy - wskaźnik przekładu

0 dokument nie jest przekładem i nie zawiera przekładów

1 dokument jest przekładem lub zawiera przekład

Drugi - źródło kodu

@ *MARC Code List for Languages.*

Kody podpól

$a Język tekstu (P)
$b Język streszczeń (P)
$h Język tekstu oryginalnego i/lub pośredniego (P)

Zawartość pola

Pole 041 zawiera informację dotyczącą języków dokumentu, jeżeli nie jest możliwe umieszczenie pełnej informacji o językach dokumentu w polu 008/35-37. Ma zastosowanie w opisie dokumentów wielojęzycznych (także jeśli jednym z języków jest język migowy) lub zawierających tłumaczenie oraz jeśli język streszczeń, abstraktów lub materiałów towarzyszących jest inny niż język tekstu głównego.

Informację o języku podaje się w postaci trzyliterowego kodu według *MARC Code List for Languages*.

Pole 041 używane jest w połączeniu z pozycjami 35–37 pola 008. Kod podany w tych pozycjach powinien być identyczny z kodem wpisanym w pierwszym podpolu **$a** pola 041.

Informacje o języku mogą być podawane także w formie słownej w polu 546.

Znaki umowne

Na końcu pola 041 nie stawia się kropki.

Wskaźniki

Pierwszy wskaźnik - wskaźnik przekładu

0 - dokument nie jest przekładem i nie zawiera przekładów

008/35-37 eng

041 0@ $a eng **$a** fre

dokument nie jest przekładem, tekst jest w językach angielskim i francuskim

1 - dokument jest przekładem lub zawiera przekład

008/35-37 eng

041 1@ $a eng **$h** rus

dokument jest przekładem z języka rosyjskiego na angielski

Drugi wskaźnik - źródło kodu

@ - *MARC Code List for Languages*
Wartość @ (spacja) oznacza, że kody podawane w polu 041 są przejmowane z *MARC Code List for Languages*.

Podpola

$a - Język tekstu

Podaje się trzyliterowe kody języków tekstu dokumentu podane w porządku ich dominacji w tekście lub w porządku alfabetycznym, jeżeli żaden z języków nie dominuje. Dla każdego kodu języka tekstu powtarza się podpole **$a**. Pierwszy kod powinien być zgodny z kodem wymienionym w polu stałej długości (008/35-37). Jeżeli tekst jest w bardzo wielu językach (np. więcej niż sześciu), a w pozycjach 35-37 pola 008 podano kod „mul", w pierwszym podpolu **$a** podaje się kod języka tytułu dokumentu lub pierwszego tytułu (jeżeli jest on w więcej niż jednym języku), w drugim podpolu **$a** podaje się kod „mul" (wiele języków).

Jeżeli dokument zawiera tekst oryginalny i towarzyszący mu przekład, w powtórzonym podpolu $a podaje się zarówno kod języka przekładu, jak i oryginału. Język oryginału wymienia się także w podpolu **$h.**

008/35-37 pol
041 0@ $a pol **$a** eng **$a** rus
546 @@ $a Tekst pol., ang., ros.
tekst w języku angielskim, polskim, rosyjskim z przewagą języka polskiego

$b - Język streszczeń

Podaje się w porządku alfabetycznym kody języków streszczeń różnych od języka tekstu głównego. Dla każdego kodu języka streszczeń powtarza się podpole **$b**.

041 0@ $a eng **$a** mul **$b** eng **$b** fre

tekst w więcej niż sześciu językach z przewagą języka angielskiego, streszczenia w języku angielskim i francuskim

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

MARC 21 format

The description of all fields of the MARC 21 format is contained in 175 pages of the manual.

Or in many

thick volumes.

INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Structured data

Structured data contain a repetitive arrangement of fields. They are easy to process, but shallow in information.

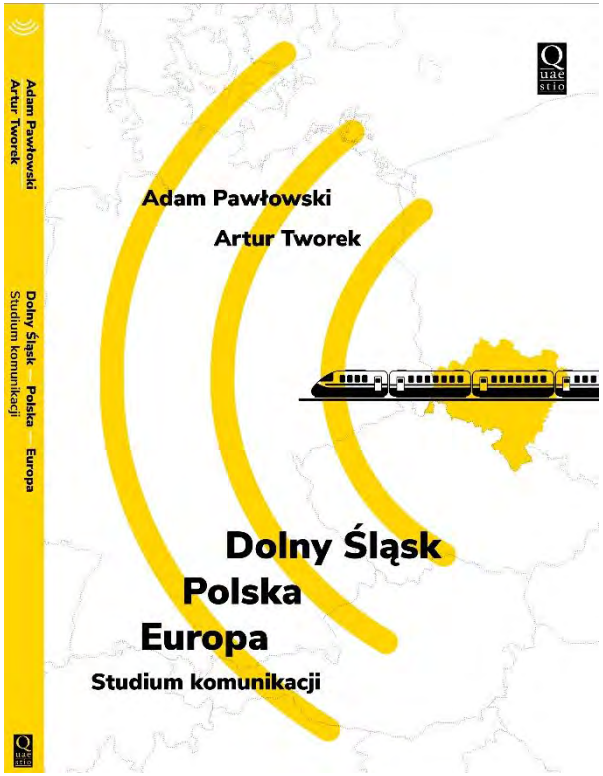
010 – Library of Congress control number
020 – International standard book number
037 – Source of acquisition
040 – Original source of cataloging $c Inputting library
041 – Language code
050 – Library of Congress call number
082 – Dewey decimal call number
090 – Locally assigned Library of Congress call number
092 – Locally assigned Dewey call number
049 – Local holdings
1xx – Author main entry (personal, corporate or conference)
245 10 Title proper : $b Remainder of title / $c Statement of responsibility
246 – Varying form of title
250 – Edition statement
260 – Place of publication : $b Publisher, $c Date
300 – Extent of item : $b Other physical details (illustrations, etc.) $c Dimensions
490 – Series statement as transcribed from item
500 – General notes
504 – Bibliography, etc. note
505 – Contents note
520 – Summary, abstract, annotation, scope, etc. note
590 – Local notes
6xx – Subject access fields (added entries)
700-740 Added entry-Uncontrolled related/analytical title
800-740 Series added entry as established

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Marc 21 record as it is read by machines

{"id":5675461,"createdDate":"2017-08-07T13:50:59.000+02:00","updatedDate":"2017-11-06T14:54:28.000+01:00","language":"polski","subject":"Teoria gier","subjectPlace":"","subjectTime":"","subjectWork":"","isbnIssn":"9788380885943 9788380885950","author":"Binmore, Ken (1940-). Konarzewska, Iwona. Wydawnictwo Uniwersytetu Łódzkiego.","placeOfPublication":"Łódź : Polska","location":"","title":"Teoria gier / Game theory : a very short introduction, Krótkie Wprowadzenie ; 8","udc":"519.83 02","publisher":"Wydawnictwo Uniwersytetu Łódzkiego. Wydawnictwo Uniwersytetu Łódzkiego.","kind":"książka","domain":"Filozofia i etyka","formOfWork":"Książki Publikacje popularnonaukowe","genre":"Opracowanie","timePeriodOfCreation":"2007","audienceGroup":"","demographicGroup":"","nationalBibliographyNumber":"PB 2017/27081","publicationYear":"2017","languageOfOriginal":"angielski","fixedFields":[{"label":"LANG","value":"pol","display":"Polish","id":"24"},{"label":"COUNTRY","value":"pl ","display":"Polska","id":"89"},{"label":"CAT DATE","value":"2017-08-07","id":"28"},{"label":"CREATED","value":"2017-08-07T11:50:59Z","id":"83"},{"label":"MARCTYPE","value":" ","id":"107"},{"label":"REVISIONS","value":"9","id":"85"},{"label":"SUPPRESS","value":"b","id":"31"},{"label":"SKIP","value":"0","id":"25"},{"label":"REC TYPE","value":"b","id":"80"},{"label":"MAT TYPE","value":"a","display":"Book","id":"30"},{"label":"COPIES","value":"0","id":"27"},{"label":"PDATE","value":"2017-08-30T10:55:06Z","id":"98"},{"label":"BIB LVL","value":"m","display":"Monograph","id":"29"},{"label":"AGENCY","value":"1","id":"86"},{"label":"UPDATED","value":"2017-11-06T13:54:28Z","id":"84"},{"label":"RECORD #","value":"5675461","id":"81"},{"label":"LOCATION","value":"multi","id":"26"}],"varFields":[{"fieldTag":"a","marcTag":"100","ind1":"1","ind2":" ","subfields":[{"tag":"a","content":"Binmore, Ken"},{"tag":"d","content":"(1940-)."},{"tag":"e","content":"Autor"}]},{"fieldTag":"b","marcTag":"700","ind1":"1","ind2":" ","subfields":[{"tag":"a","content":"Konarzewska, Iwona."},{"tag":"e","content":"Tłumaczenie"}]},{"fieldTag":"b","marcTag":"710","ind1":"2","ind2":" ","subfields":[{"tag":"a","content":"Wydawnictwo Uniwersytetu Łódzkiego."},{"tag":"e","content":"Wydawca"},{"tag":"4","content":"pbl"}]},{"fieldTag":"d","marcTag":"380","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Książki"}]},{"fieldTag":"d","marcTag":"380","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Publikacje popularnonaukowe"}]},{"fieldTag":"d","marcTag":"388","ind1":"1","ind2":" ","subfields":[{"tag":"a","content":"2001-"}]},{"fieldTag":"d","marcTag":"650","ind1":" ","ind2":"4","subfields":[{"tag":"a","content":"Teoria gier"}]},{"fieldTag":"d","marcTag":"655","ind1":" ","ind2":"4","subfields":[{"tag":"a","content":"Opracowanie"}]},{"fieldTag":"d","marcTag":"658","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Filozofia i etyka"}]},{"fieldTag":"g","marcTag":"015","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"PB 2017/27081"}]},{"fieldTag":"i","marcTag":"020","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"9788380885943"}]},{"fieldTag":"i","marcTag":"020","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"9788380885950"},{"tag":"q","content":"e-ISBN"}]},{"fieldTag":"j","marcTag":"080","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"519.83"}]},{"fieldTag":"l","marcTag":"998","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"ik"}]},{"fieldTag":"n","marcTag":"504","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Bibliografia na stronach 195-200. Indeks."}]},{"fieldTag":"n","marcTag":"530","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Dostępne także jako e-book."}]},{"fieldTag":"n","marcTag":"536","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Publikacja sfinansowana ze środków Wydawnictwa Uniwersytetu Łódzkiego"}]},{"fieldTag":"p","marcTag":"260","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Łódź :"},{"tag":"b","content":"Wydawnictwo Uniwersytetu Łódzkiego,"},{"tag":"c","content":"2017."}]},{"fieldTag":"r","marcTag":"300","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"206, [1] strona :"},{"tag":"b","content":"ilustracje, fotografie, wykresy ;"},{"tag":"c","content":"21 cm."}]},{"fieldTag":"s","marcTag":"490","ind1":"1","ind2":" ","subfields":[{"tag":"a","content":"Krótkie Wprowadzenie ;"},{"tag":"v","content":"8"}]},{"fieldTag":"s","marcTag":"830","ind1":" ","ind2":"0","subfields":[{"tag":"a","content":"Krótkie Wprowadzenie ;"},{"tag":"v","content":"8"}]},{"fieldTag":"t","marcTag":"245","ind1":"1","ind2":"0","subfields":[{"tag":"a","content":"Teoria gier /"},{"tag":"c","content":"Ken Binmore ; tłumaczenie Iwona Konarzewska."}]},{"fieldTag":"u","marcTag":"246","ind1":"1","ind2":" ","subfields":[{"tag":"i","content":"Tytuł oryginału:"},{"tag":"a","content":"Game theory :"},{"tag":"b","content":"a very short introduction,"},{"tag":"f","content":"2007"}]},{"fieldTag":"y","marcTag":"008","ind1":" ","ind2":" ","content":"170807s2017 pl aod |001 0 pol nam i "},{"fieldTag":"y","marcTag":"040","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"WA N"},{"tag":"c","content":"WA N"}]},{"fieldTag":"y","marcTag":"041","ind1":"1","ind2":" ","subfields":[{"tag":"a","content":"pol"},{"tag":"h","content":"eng"}]},{"fieldTag":"y","marcTag":"046","ind1":" ","ind2":" ","subfields":[{"tag":"k","content":"2007"}]},{"fieldTag":"y","marcTag":"336","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Tekst"},{"tag":"b","content":"txt"},{"tag":"2","content":"rdacontent"}]},{"fieldTag":"y","marcTag":"337","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Bez urządzenia pośredniczącego"},{"tag":"b","content":"n"},{"tag":"2","content":"rdamedia"}]},{"fieldTag":"y","marcTag":"338","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"Wolumin"},{"tag":"b","content":"nc"},{"tag":"2","content":"rdacarrier"}]},{"fieldTag":"y","marcTag":"920","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"978-83-8088-594-3"}]},{"fieldTag":"y","marcTag":"920","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"978-83-8088-595-0 (e-ISBN)"}]},{"fieldTag":"y","marcTag":"999","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"zkd"},{"tag":"b","content":"eoaw"},{"tag":"x","content":"33"},{"tag":"y","content":"17"}]},{"fieldTag":"y","marcTag":"084","ind1":" ","ind2":" ","subfields":[{"tag":"a","content":"02"}]},{"fieldTag":"_","content":"00000nam a2200517 i 4500"}],"marc":{"leader":"00000nam a2200517 i 4500","fields":[{"001":"b56754619"},{"008":"170807s2017 pl aod |001 0 pol nam i "},{"015":{"ind1":" ","ind2":" ","subfields":[{"a":"PB 2017/27081"}]}},{"020":{"ind1":" ","ind2":" ","subfields":[{"a":"9788380885943"}]}},{"020":{"ind1":" ","ind2":" ","subfields":[{"a":"9788380885950"},{"q":"e-ISBN"}]}},{"040":{"ind1":" ","ind2":" ","subfields":[{"a":"WA N"},{"c":"WA N"}]}},{"041":{"ind1":"1","ind2":" ","subfields":[{"a":"pol"},{"h":"eng"}]}},{"046":{"ind1":" ","ind2":" ","subfields":[{"k":"2007"}]}},{"080":{"ind1":" ","ind2":" ","subfields":[{"a":"519.83"}]}},{"084":{"ind1":" ","ind2":" ","subfields":[{"a":"02"}]}},{"100":{"ind1":"1","ind2":" ","subfields":[{"a":"Binmore, Ken"},{"d":"(1940-)."},{"e":"Autor"}]}},{"245":{"ind1":"1","ind2":"0","subfields":[{"a":"Teoria gier /"},{"c":"Ken Binmore ; tłumaczenie Iwona Konarzewska."}]}},{"246":{"ind1":"1","ind2":" ","subfields":[{"i":"Tytuł oryginału:"},{"a":"Game theory :"},{"b":"a very short introduction,"},{"f":"2007"}]}},{"260":{"ind1":" ","ind2":" ","subfields":[{"a":"Łódź :"},{"b":"Wydawnictwo Uniwersytetu Łódzkiego,"},{"c":"2017."}]}},{"300":{"ind1":" ","ind2":" ","subfields":[{"a":"206, [1] strona :"},{"b":"ilustracje, fotografie, wykresy ;"},{"c":"21 cm."}]}},{"336":{"ind1":"

Book records vs. digital humanities



Title	Dolny Śląsk - Polska - Europa: studium komunikacji / Adam Pawłowski, Artur Tworek.
Creator	Pawłowski, Adam (historia) Autor > Tworek, Artur (1965-) Autor > Oficyna Wydawnicza Atut Wydawca >
Format	334 strony : faksymilia, fotografie, ilustracje, mapy, wykresy ; 29 cm + mapa.
Publisher	Wrocław : Oficyna Wydawnicza Atut - Wrocławskie Wydawnictwo Oświatowe
Creation Date	2021
Note	Bibliografia na stronach 235-246. Indeksy. Współfinansowanie: Uniwersytet Wrocławski
Form and type	Książki Publikacje naukowe
Subject	1801-1900 > 1901-2000 > Komunikacja społeczna > Regionalizm > Transport > Dolny Śląsk > Województwo dolnośląskie (1999-) >
Genre	Opracowanie
Discipline and Approach	Historia Media i komunikacja społeczna Socjologia i społeczeństwo Transport i logistyka
Identifier	ISBN : 9788379777082
UKD	913(438) 316.77(438) 656
Language	Polish

Legenda:
- – toponim (rzutowanie na mapę)
- – antroponim (rzutowanie na bazę bibliograficzną)
- – słowo kluczowe (rzutowanie na ontologię)
- – gatunek piśmienniczy
- – numer klasyfikacji UKD (rzutowanie na słowniki i bazy leksykalne)

9324 characters

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Data: National Library databases

Poland does not currently have a complete National Bibliography.

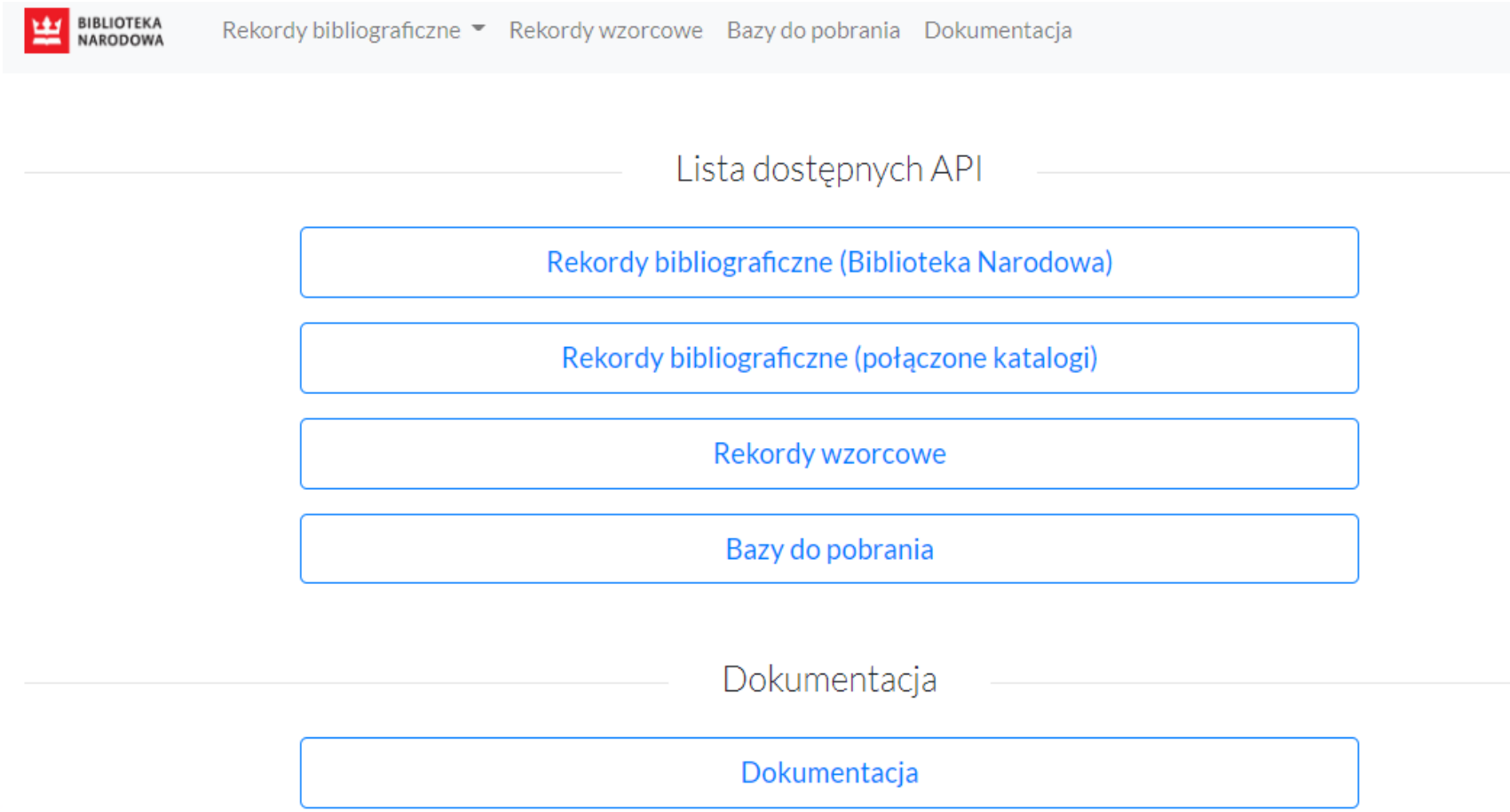
It does, however, have the catalogs of the National Library, which are almost complete after 1918.

BN catalog = magazine content

National Bibliography = anything that meets the criteria (but doesn't have to be physically in stock)

BN Catalog < National Bibliography

Data: National Library databases



SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Data: National Library databases

Bazy bibliograficzne

Biblioteka narodowa

- **bibs-all.marc** 11,3 GB, 10 paź 08:02
- **bibs-artykul.marc** 4,3 GB, 10 paź 08:03
- **bibs-audiobook.marc** 33,5 MB, 10 paź 08:03
- **bibs-czasopismo.marc** 231 MB, 10 paź 08:03
- **bibs-dokument_elektroniczny.marc** 10,8 MB, 10 paź 08:03
- **bibs-druk_ulotny.marc** 688 MB, 10 paź 08:03
- **bibs-e_book.marc** 203 MB, 10 paź 08:03
- **bibs-film.marc** 93,2 MB, 10 paź 08:03
- **bibs-fotografia.marc** 222 MB, 10 paź 08:03
- **bibs-grafika.marc** 167 MB, 10 paź 08:03
- **bibs-ksiazka.marc** 4,38 GB, 10 paź 08:03
- **bibs-mapa.marc** 257 MB, 10 paź 08:03
- **bibs-nagrania_dzwiekowe.marc** 230 MB, 10 paź 08:03
- **bibs-nuty.marc** 220 MB, 10 paź 08:03
- **bibs-obiekt_trojwymiarowy.marc** 1,92 MB, 10 paź 08:03
- **bibs-rekopis.marc** 76,1 MB, 10 paź 08:03
- **bibs-rysunek.marc** 42,9 MB, 10 paź 08:03
- **bibs-starodruk.marc** 155 MB, 10 paź 08:03

SBP
STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH

Methods

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

3 methodological pillars

- **Programming for databases, data linking and visualization (especially on maps).**
- **Natural Language Processing (NLP)**
- **Digital humanities (interpretation, topic selection, holistic and integrative approach)**

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

What is NLP?

A collection of algorithms and resources that allow automatic analysis of language (text).

- recognition of basic forms (*domowi* → *dom*)
- recognition of syntactic structures
- quantitative analyses
- semantic analyses

Application: analysis of text fields of a record (e.g., titles).

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Digital Humanities

A system of interrelated resources, functionalities and cognitive practices, created by transferring to the digital realm and creatively expanding the heritage of the humanities growing out of print culture.

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Results part 1

– projection of records onto a map
– historical cross-section

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Where to get the coordinates from?

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

One name - many places

Cieplice | all countries | search | [advanced search]

286 records found for "Cieplice"

	Name	Country	Feature class	Latitude	Longitude
1	Cieplice Śląskie Zdrój Bad Warmbrunn,Cieplice Slaskie Zdroj,Cieplice Slaskie-Zdroj,Cieplice Śląskie Zdrój,Cieplice Śląskie-...	Poland, Lower Silesia Jelenia Góra > Jelenia Góra	populated place population 15,000	N 50° 51' 55''	E 15° 41' 1''
2	Teplice Cieplice,Teplice,Teplice-Sanov,Teplice-Šanov,Teplicė,Teplitse,Teplitz,Teplitz-Schoenau,Teplitz-Schön...	Czechia, Ústecký kraj Teplice District > Teplice	populated place population 51,223	N 50° 38' 25''	E 13° 49' 28''
3	Cieplice Cieplice,Kolonia Cieplice	Poland, Lesser Poland Powiat miechowski > Gmina Gołcza	populated place	N 50° 21' 13''	E 19° 54' 59''
4	Cieplice	Poland, Subcarpathia Powiat przeworski > Adamówka	populated place	N 50° 15' 6''	E 22° 37' 41''
5	Trenčianske Teplice Trencanske Teplice,Trench'janske Teplice,Trench'janske-Teplice,Trenchinske Tjeplice,Trencianske Tepl...	Slovakia, Trenčiansky kraj Trenčín District > Trenčianske Teplice	populated place population 4,438, elevation 267m	N 48° 54' 38''	E 18° 10' 0''
6	Teplice District Daerah Teplice,Distretto di Teplice,District de Teplice,Districte de Teplice,Distrikto Teplice,Okres...	Czechia, Ústecký kraj Teplice District	second-order administrative division population 128,610	N 50° 38' 0''	E 13° 50' 0''
7	Duchcov Duchcov,Dukhcov,Dux,doufutsuofu,dshkwf,du he zuo fu,duhochopeu,Духцов,دشكوف,ドゥフツォフ,杜赫佐夫,두호초프	Czechia, Ústecký kraj Teplice District > Duchcov	populated place population 8,937	N 50° 36' 13''	E 13° 44' 46''
8	Dubí Dubi,Eichwald,doubi,du bi,dubi,Дуби,ドゥビー,杜比,두비	Czechia, Ústecký kraj Teplice District > Dubí	populated place population 7,591	N 50° 41' 8''	E 13° 47' 8''
9	Krupka Graupen,Krupka,ke lu pu ka,Крупка,克鲁普卡	Czechia, Ústecký kraj Teplice District > Krupka	populated place population 13,687	N 50° 41' 4''	E 13° 51' 29''
10	Bílina Bilin,Bilina,bi li na,Билина,比利纳	Czechia, Ústecký kraj Teplice District > Bílina	populated place population 15,738	N 50° 32' 54''	E 13° 46' 31''
11	Cieplice	Poland, Warmia-Masuria Powiat elbląski > Elbląg	populated place	N 54° 15' 20''	E 19° 20' 26''

INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

One name - many places

Wołowiec | all countries | search | [advanced search]

6 records found for "Wołowiec"

	Name	Country	Feature class	Latitude	Longitude
1	Volovec Wołowiec,Wołowiec	SK,PL	mountain elevation 2063m	N 49° 12' 27"	E 19° 45' 46"
2	Wołowiec	Poland, West Pomerania Powiat goleniowski > Gmina Nowogard	populated place	N 53° 46' 1"	E 15° 9' 14"
3	Wołowiec	Poland, Lesser Poland Powiat gorlicki > Gmina Sękowa	populated place	N 49° 31' 10"	E 21° 21' 9"
4	Wolowiec Lake	Canada, Manitoba	lake	N 56° 7' 24"	W 101° 23' 3"
5	Volovets Voloc,Volovec,Volovec',Volovets,Volóc,Wołowiec,Wołowiec,Воловец,Воловець	Ukraine, Transcarpathia Volovets Raion	seat of a second-order administrative division population 5,128, elevation 492m	N 48° 42' 39"	E 23° 11' 6"
6	Wołowiec	Poland, Lower Silesia Wałbrzych County > Jedlina-Zdrój	mountain elevation 776m	N 50° 44' 2"	E 16° 18' 23"

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

One name - many places

Stanisławów | all countries

Feature Class: city, village,...

Continent: Europe

fuzzy search : ☐

search

62 records found for "Stanisławów"

	Name	Country	Feature class	Latitude	Longitude
1	Ivano-Frankivsk IFO,Ivano-Frankivs'k,Ivano-Frankivsk,Ivano-Frankovsk,Iwano-Frankiwsk,Stanislau,Stanislav,Stanislavov...	Ukraine, Ivano-Frankivsk Ivano-Frankivsk Raion	seat of a first-order administrative division population 238,196, elevation 260m	N 48° 55′ 23"	E 24° 42′ 44"
2	Stanisławów Stanislavuv,Stanislawow,Stanisławów,Станиславув	Poland, Mazovia Mińsk County > Gmina Stanisławów	seat of a third-order administrative division population 2,200	N 52° 17′ 21"	E 21° 32′ 54"
3	Stanisławów Skrzański Stanislawow,Stanislawow Skrzanski,Stanisławów,Stanisławów Skrzański	Poland, Mazovia Gostynin County > Gostynin	populated place	N 52° 23′ 7"	E 19° 29′ 36"
4	Stanisławów Studziński Stanislawow,Stanislawow Studzinski,Stanisławów,Stanisławów Studziński	Poland, Łódź Voivodeship Tomaszów County > Czerniewice	populated place	N 51° 40′ 28"	E 20° 5′ 45"
5	Stanisławów-Kolonia Kolonie Stanislawow,Kolonie Stanisławów	Poland, Mazovia Mińsk County > Gmina Stanisławów	section of populated place	N 52° 17′ 18"	E 21° 34′ 43"
6	Stanisławów Kolonia Stanislawow,Kolonia Stanisławów,Stanislawow,Stanisławów	Poland, Lublin Radzyń County > Gmina Ulan-Majorat	populated place population 111	N 51° 48′ 48"	E 22° 22′ 12"
7	Stanisławów Stary Stanislawow Stary,Stanisławów Stary	Poland, Łódź Voivodeship Pabianice County > Lutomiersk	populated place	N 51° 47′ 21"	E 19° 12′ 14"
8	Stanisławów Nowy Stanislawow Nowy,Stanisławów Nowy	Poland, Łódź Voivodeship Pabianice County > Lutomiersk	populated place	N 51° 48′ 5"	E 19° 11′ 23"
9	Stanisławów Lipski Stanislawow Lipski,Stanisławów Lipski	Poland, Łódź Voivodeship Tomaszów County > Czerniewice	populated place	N 51° 41′ 57"	E 20° 8′ 39"
10	Stanisławów Stanislawow,Stanisławów	Poland, Mazovia Kozienice County > Głowaczów	populated place	N 51° 34′ 18"	E 21° 23′ 32"

INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

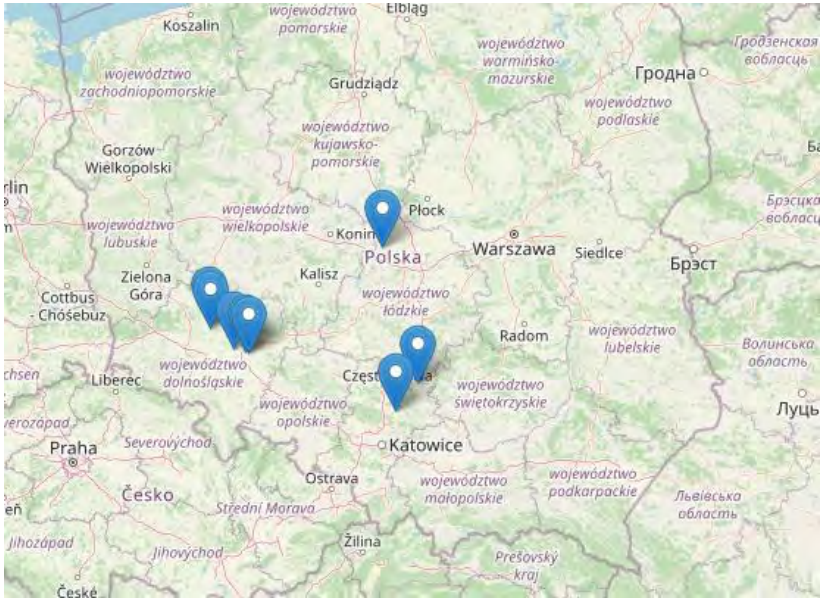
One name - many places

	Name	Country	Feature class	Latitude	Longitude
1	Ligota Ligota	Poland, Silesia Bielsko-Biała County > Czechowice-Dziedzice	populated place population 4,454	N 49° 53′ 54″	E 18° 57′ 3″
2	Ligota Woźnicka Ligota Woznicka,Ligota Woźnicka	Poland, Silesia Lubliniec County > Woźniki	populated place	N 50° 36′ 39″	E 19° 2′ 36″
3	Ligota Tułowicka Ligota Tulowicka,Ligota Tułowicka	Poland, Opole Voivodeship Opole County > Tułowice	populated place	N 50° 34′ 54″	E 17° 39′ 11″
4	Ligota Rybińska Ligota Rybinska,Ligota Rybińska	Poland, Greater Poland	populated place	N 51° 22′ 25″	E 17° 42′ 33″
5	Ligota Prószkowska Ligofa Proszkowska,Ligofa Prószkowska,Ligota Proszkowska,Ligota Prószkowska	Poland, Opole Voivodeship Opole County > Gmina Prószków	populated place	N 50° 33′ 15″	E 17° 49′ 53″
6	Ligota Zamecka Ligota Zamecka	Poland, Opole Voivodeship Kluczbork County > Gmina Kluczbork	populated place	N 50° 58′ 28″	E 18° 14′ 24″
7	Ligota Turawska Ellguth,Ellguth-Turawa,Ligota Turawska	Poland, Opole Voivodeship Opole County > Gmina Turawa	populated place	N 50° 47′ 12″	E 18° 10′ 57″
8	Ligota Oleska Ligota Oleska	Poland, Opole Voivodeship Olesno County > Radłów	populated place	N 50° 57′ 29″	E 18° 32′ 6″
9	Ligota Dolna Ligota Dolna	Poland, Opole Voivodeship Kluczbork County > Gmina Kluczbork	populated place	N 50° 58′ 21″	E 18° 10′ 30″
10	Ligota Bialska Ligota Bialska	Poland, Opole Voivodeship Prudnik County > Gmina Biała	populated place	N 50° 24′ 39″	E 17° 40′ 48″
11	Ligota Ligota	Poland, Opole Voivodeship Krapkowice County > Gmina Krapkowice	populated place	N 50° 26′ 47″	E 17° 59′ 29″
12	Ligota Mała Ligota Mala,Ligota Mała	Poland, Lower Silesia Dzierżoniów County > Gmina Niemcza	populated place	N 50° 40′ 21″	E 16° 48′ 47″
13	Ligota Książęca	Poland, Opole Voivodeship Namysłów County > Namysłów	populated place	N 51° 1′ 11″	E 17° 34′ 11″
14	Ligota Dobrodzieńska	Poland, Opole Voivodeship Olesno County > Gmina Dobrodzień	populated place	N 50° 43′ 58″	E 18° 24′ 32″
15	Ligota	Poland, Łódź Voivodeship Łask County > Widawa	populated place	N 51° 28′ 53″	E 18° 55′ 54″
16	Ligota Wielka	Poland, Opole Voivodeship Nysa County > Gmina Otmuchów	populated place	N 50° 29′ 11″	E 17° 6′ 37″
17	Ligota Górna	Poland, Opole Voivodeship Kluczbork County > Gmina Kluczbork	populated place	N 50° 58′ 7″	E 18° 15′ 42″
18	Ligota Czamborowa	Poland, Opole Voivodeship Strzelce County > Izbicko	populated place	N 50° 34′ 20″	E 18° 10′ 52″
19	Ligota Wielka Ligota Wielka	Poland, Lower Silesia Dzierżoniów County > Łagiewniki	populated place	N 50° 45′ 31″	E 16° 47′ 29″
20	Ligota Dolna Ligota Dolna,Nieder Ellguth	Poland, Opole Voivodeship Strzelce County > Gmina Strzelce Opolskie	populated place	N 50° 29′ 27″	E 18° 7′ 14″
21	Ligota Ligota	Poland, Greater Poland Ostrzeszów County > Kobyla Góra	populated place	N 51° 22′ 17″	E 17° 48′ 18″
22	Ligota	Poland, Greater Poland Ostrów County > Gmina Raszków	populated place	N 51° 47′ 19″	E 17° 41′ 48″
23	Ligota Tworkowska Ligota Tworkowska,Tworkowska	Poland, Silesia Wodzisław County > Gmina Lubomia	populated place	N 50° 1′ 26″	E 18° 15′ 57″
24	Ligota Polska Alt Ellguth,Ligota Polska	Poland, Lower Silesia Oleśnica County > Oleśnica	populated place	N 51° 14′ 48″	E 17° 32′ 38″
25	Ligota Ligota	Poland, Lower Silesia Góra County > Góra	populated place	N 51° 42′ 1″	E 16° 36′ 30″
26	Ligota-Panewniki	Poland, Silesia Katowice > Katowice	section of populated place	N 50° 13′ 49″	E 18° 58′ 18″
27	Ligota Mała	Poland, Lower Silesia Oleśnica County > Oleśnica	populated place	N 51° 8′ 11″	E 17° 21′ 45″

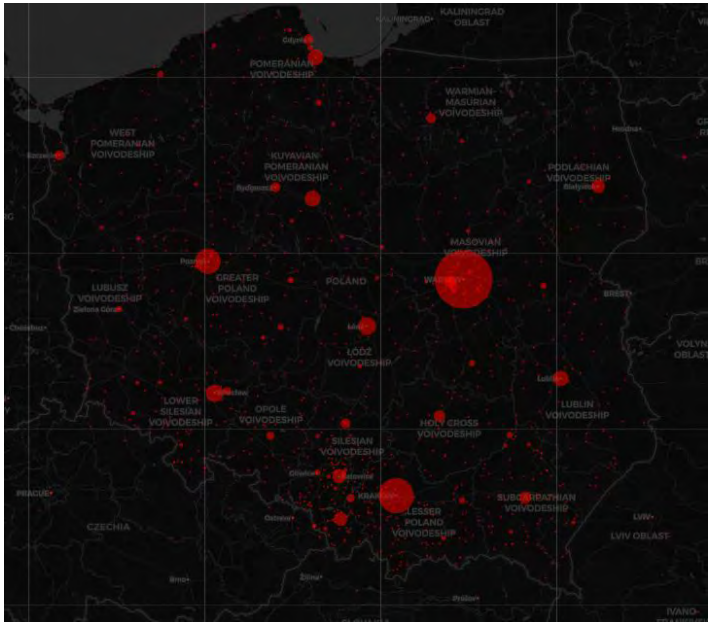
FORMACJA
OMUNIKACJA

Visualization

In this way one can show 5 points...



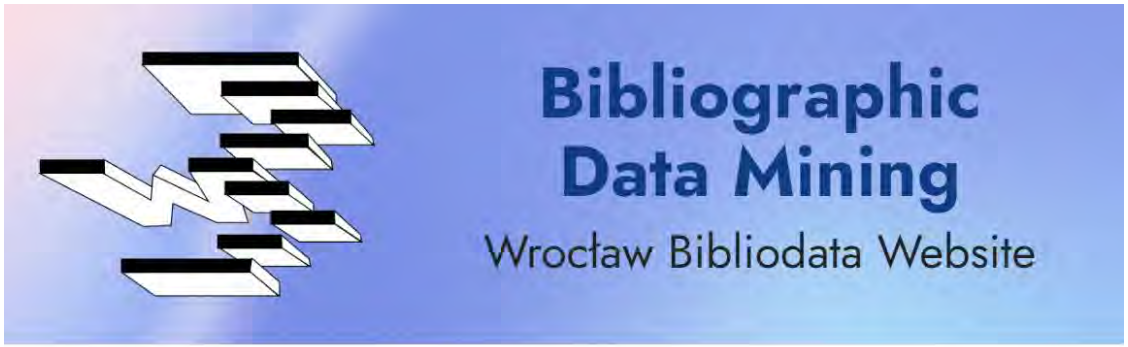
...and in this way 50,000.



SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Where to find us?

http://phc.uni.wroc.pl/wbw/



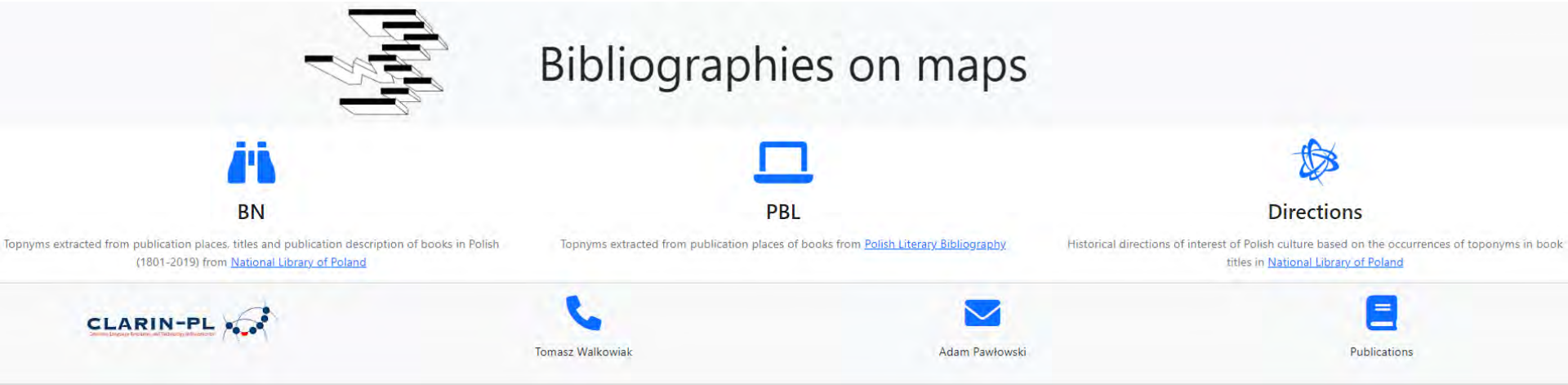
Large bibliographies have been created for decades wherever nations or communities of communication have produced a significant body of writing. They are not just simple publication registers serving the immediate information needs of society. The scale effect and the (meta)data contained in the records make large bibliographies synthetically reflect trends in culture, politics and civilization. The development of digital technologies has made it possible to extract and visualize this information, and the Internet has become an excellent platform for its dissemination.

Wrocław Bibliodata Website presents, in the form of an interactive application, the effects of projecting toponyms (local names) contained in bibliographic data onto maps, as well as visualizations of other information extracted from large bibliographies using NLP and text mining methods. Taking into account the requirements of big data methods, as well as commonly used historical periodizations, the so-called late modern and contemporary times, the conventional beginning of which is the year 1800.

The scientific base of **the Wrocław Bibliodata Website** is the Institute of Scientific Information and Library Science of the University of Wrocław, while technological support is provided by the CLARIN-PL consortium team and the Wrocław University of Science and Technology. The originators and contractors of the project are prof. Ph.D. Adam Pawłowski (UWr) and Dr. Eng. Tomasz Walkowiak (PWr).

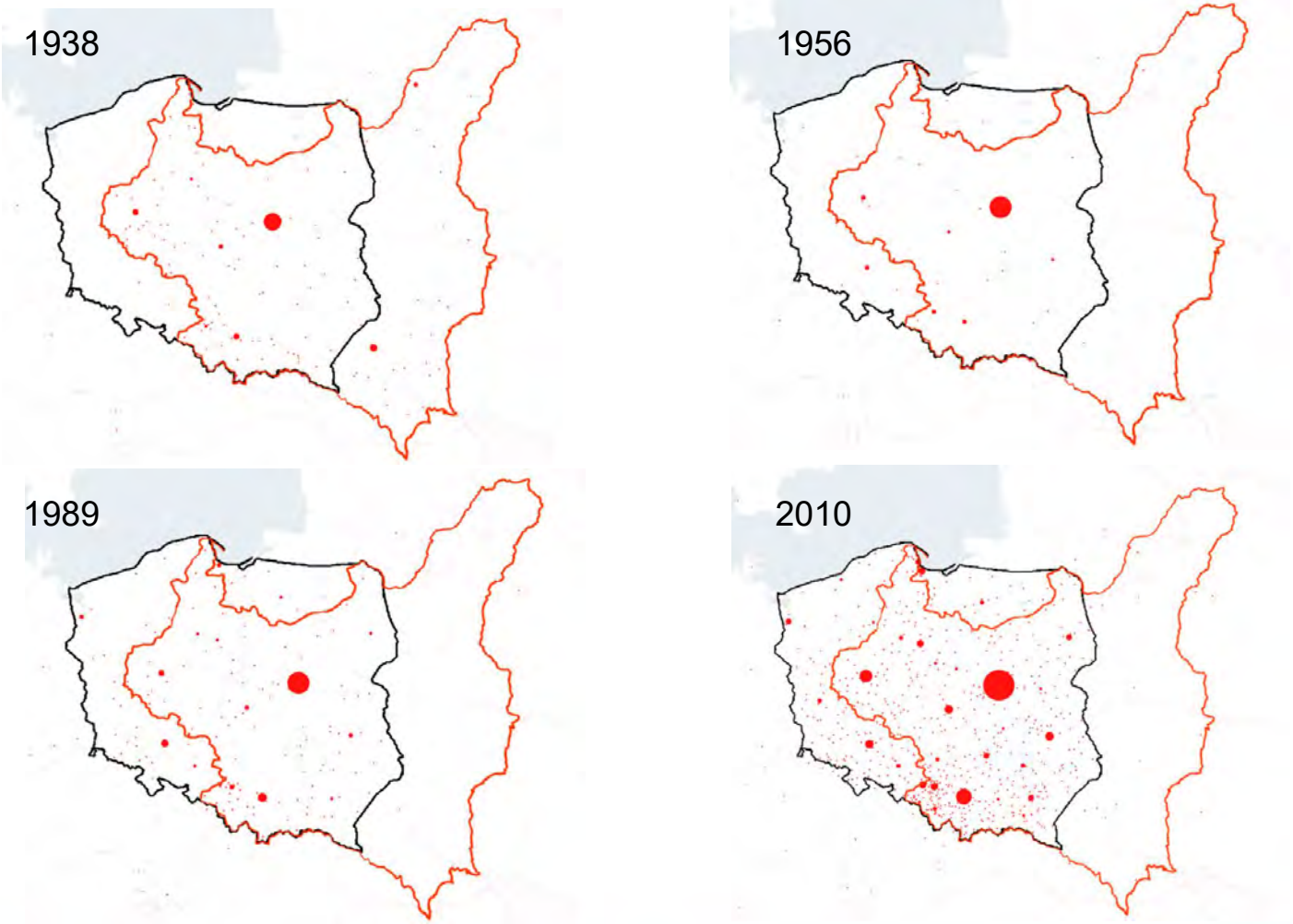
Where to find us?

https://bibgeos.clarin-pl.eu/



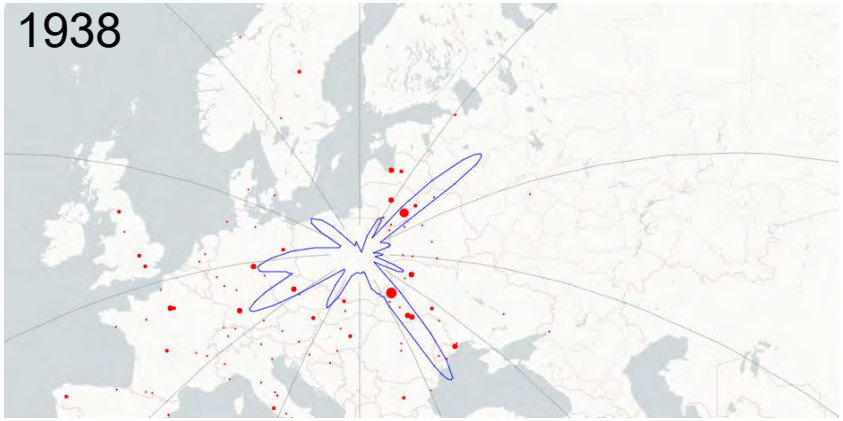
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

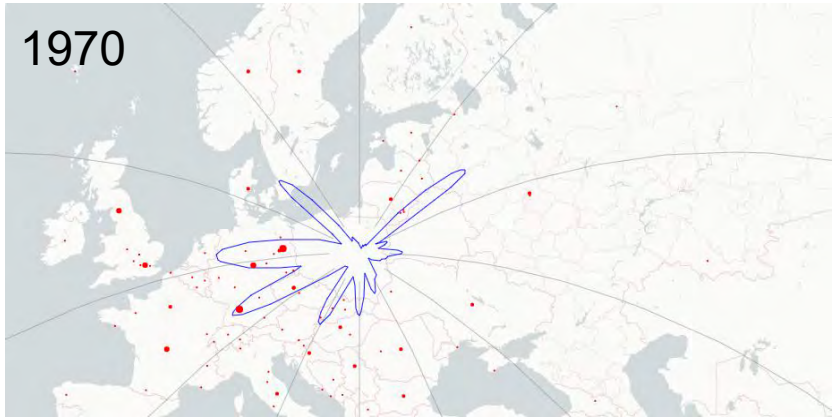
“Bibliographic frontiers” outline

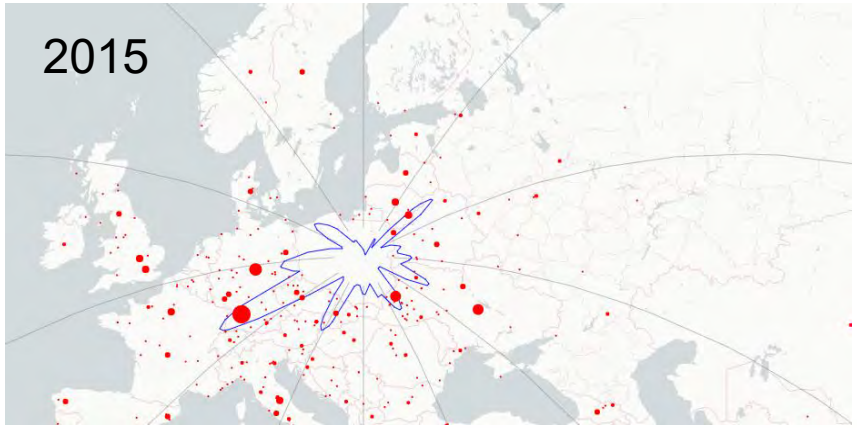


SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Directions of interest and influence







SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Why are the lines kinking?

Gerard Mercator (1512 - 1594)

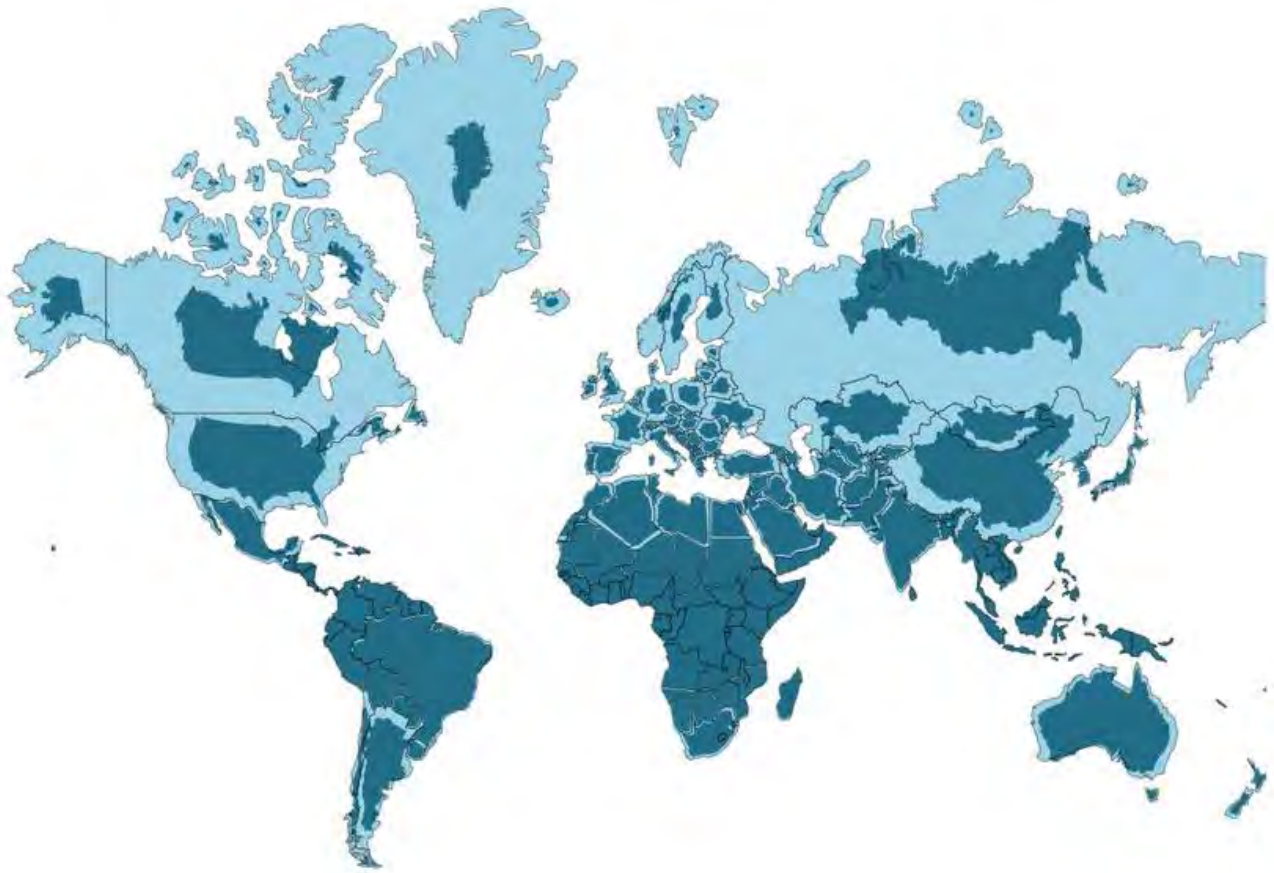
A 16th century geographer, cosmographer and cartographer who came from Flanders. His greatest fame came with the creation of a world map in 1569, based on a new projection that depicted fixed-direction sailing routes (known as rhumb lines) as straight lines. this innovation is still used in nautical charts today.

Mercator was a pioneer of cartography. Along with Gemma Frisius and Abraham Ortelius, he is widely regarded as one of the founders of the Dutch school of cartography and geography. He is also widely regarded as the most significant figure of this school. In his time he was a well-known maker of globes and scientific instruments.

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Is the earth definitely round?

World Mercator projection with true country size and shape added



SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Results part 2

Gender of authors vs. number and topics of publications (Poland vs. Germany)

General considerations

Writing and publishing a book is, in statistical terms, an indicator that aggregates other specific indicators.

- education → average education, level of education
- wealth level → average income
- prestige
- social acceptance

The assumption that the number of books published by a given social group is an approximation of the measure of emancipation is therefore correct.

In addition: certain economic measures are impossible to calculate in Central Europe during the time covered by the study (1800 onwards).

Gender of PL/DE authors: purpose and hypotheses

Specific Objectives:

- show the proportions of women and men among authors of books published between 1801 and 2021 in two European countries (communities of communication): Poland and Germany
- extract themes (topics) characteristic of authors from book titles
- identify the content that distinguishes the writing of both sexes

Approach and methodology:

We applied NLP techniques, mathematical statistics and programming databases to large bibliographic databases.

This made it possible to conduct the study in the perspective of the so-called "*longue duree"* (in the sense of the research methodology of the Annales school).

Gender of authors PL/DE: data

Data: catalog of the National Library and Deutsche Nationalbibliothek

Format: MARC 21

Additionally: Poland the 19th Cent. - Estreicher Bibliography (to the letter k)

Filtering data from both databases:

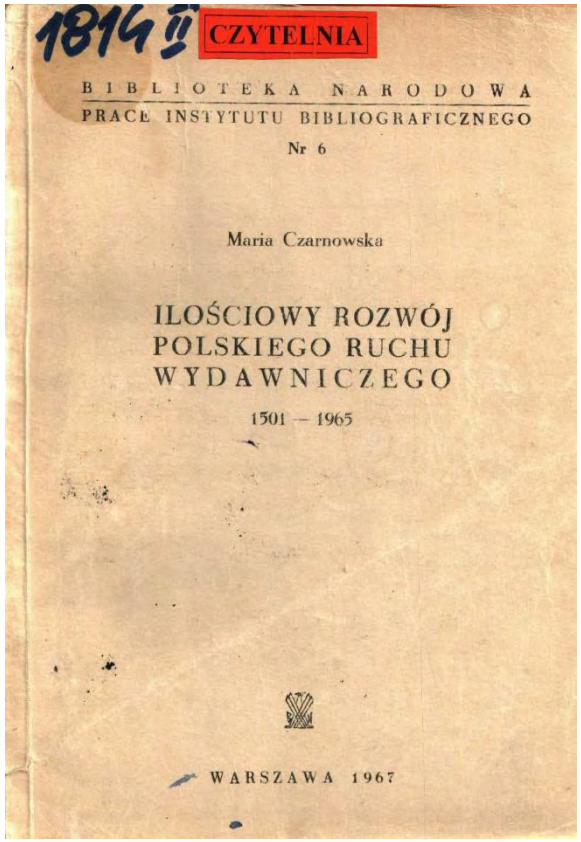
- PL and DE language included
- non-empty fields 'title', 'author', 'release date' are included
- non-text records (maps, sheet music, recordings, etc.) and periodicals were omitted

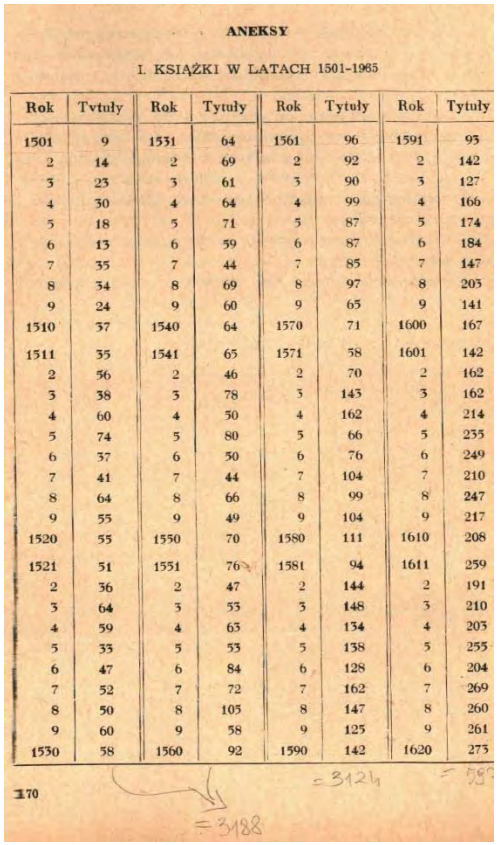
Data cleaning:

- automatic separation of first name from last name
- Unification of the coding of umlauts (can be two-characters)
- Removal of duplicates from the Estreicher Bibliography

Final volume: PL = ca 1.4 million records, DE = ca 7.5 million records

Glory for the pioneers: Maria Czarnowska





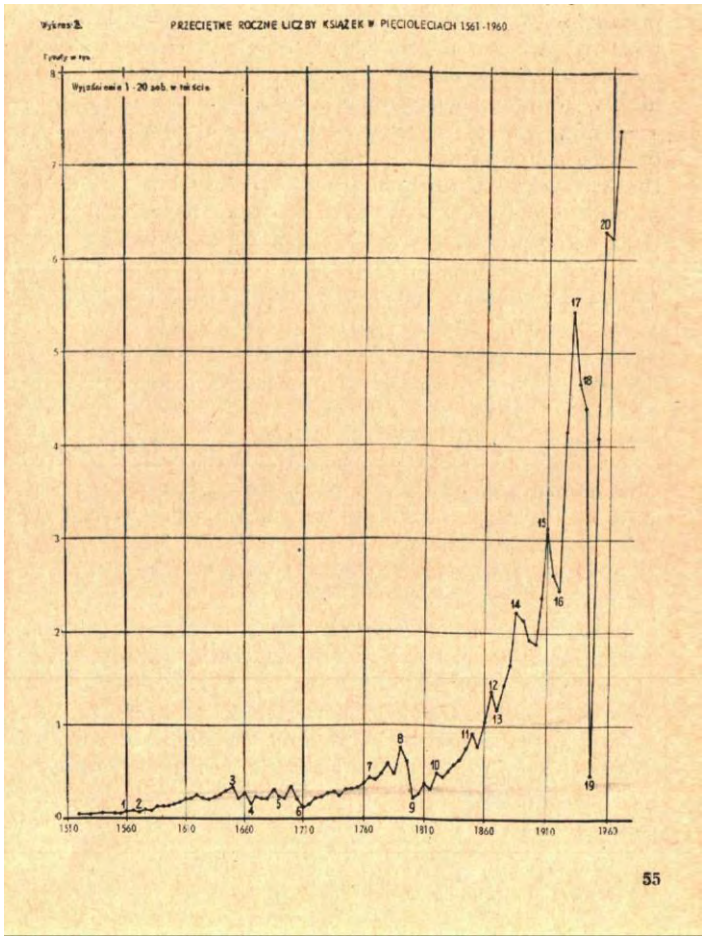
ANEKSY

I. KSIĄŻKI W LATACH 1501-1965

Rok	Tytuły	Rok	Tytuły	Rok	Tytuły	Rok	Tytuły
1501	9	1531	64	1561	96	1591	93
2	14	2	69	2	92	2	142
3	23	3	61	3	90	3	127
4	30	4	64	4	99	4	166
5	18	5	71	5	87	5	174
6	13	6	59	6	87	6	184
7	35	7	44	7	85	7	147
8	34	8	69	8	97	8	203
9	24	9	60	9	65	9	141
1510	37	1540	64	1570	71	1600	167
1511	35	1541	65	1571	58	1601	142
2	56	2	46	2	70	2	162
3	38	3	78	3	143	3	162
4	60	4	50	4	162	4	214
5	74	5	80	5	66	5	235
6	37	6	50	6	76	6	249
7	41	7	44	7	104	7	210
8	64	8	66	8	99	8	247
9	55	9	49	9	104	9	217
1520	55	1550	70	1580	111	1610	208
1521	51	1551	76	1581	94	1611	259
2	36	2	47	2	144	2	191
3	64	3	53	3	148	3	210
4	59	4	63	4	134	4	203
5	35	5	53	5	138	5	255
6	47	6	84	6	128	6	204
7	52	7	72	7	162	7	269
8	50	8	105	8	147	8	260
9	60	9	58	9	125	9	261
1530	58	1560	92	1590	142	1620	273

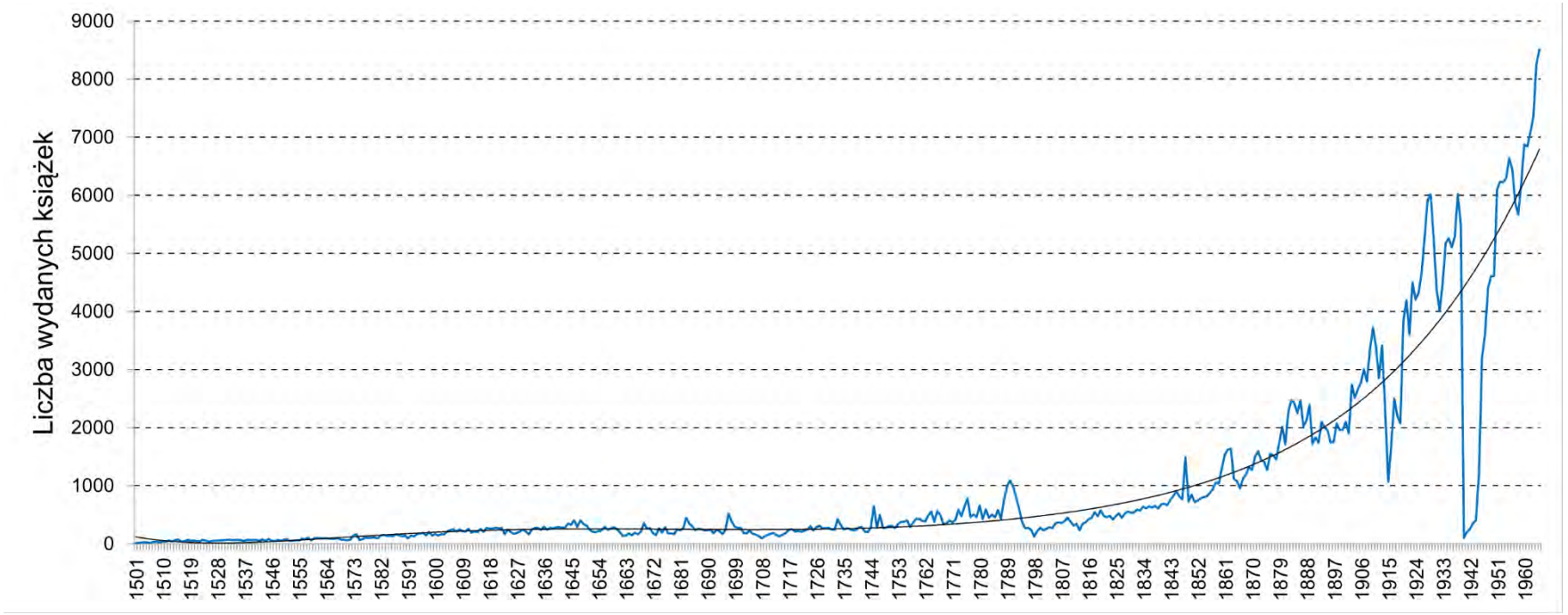
170

=3188



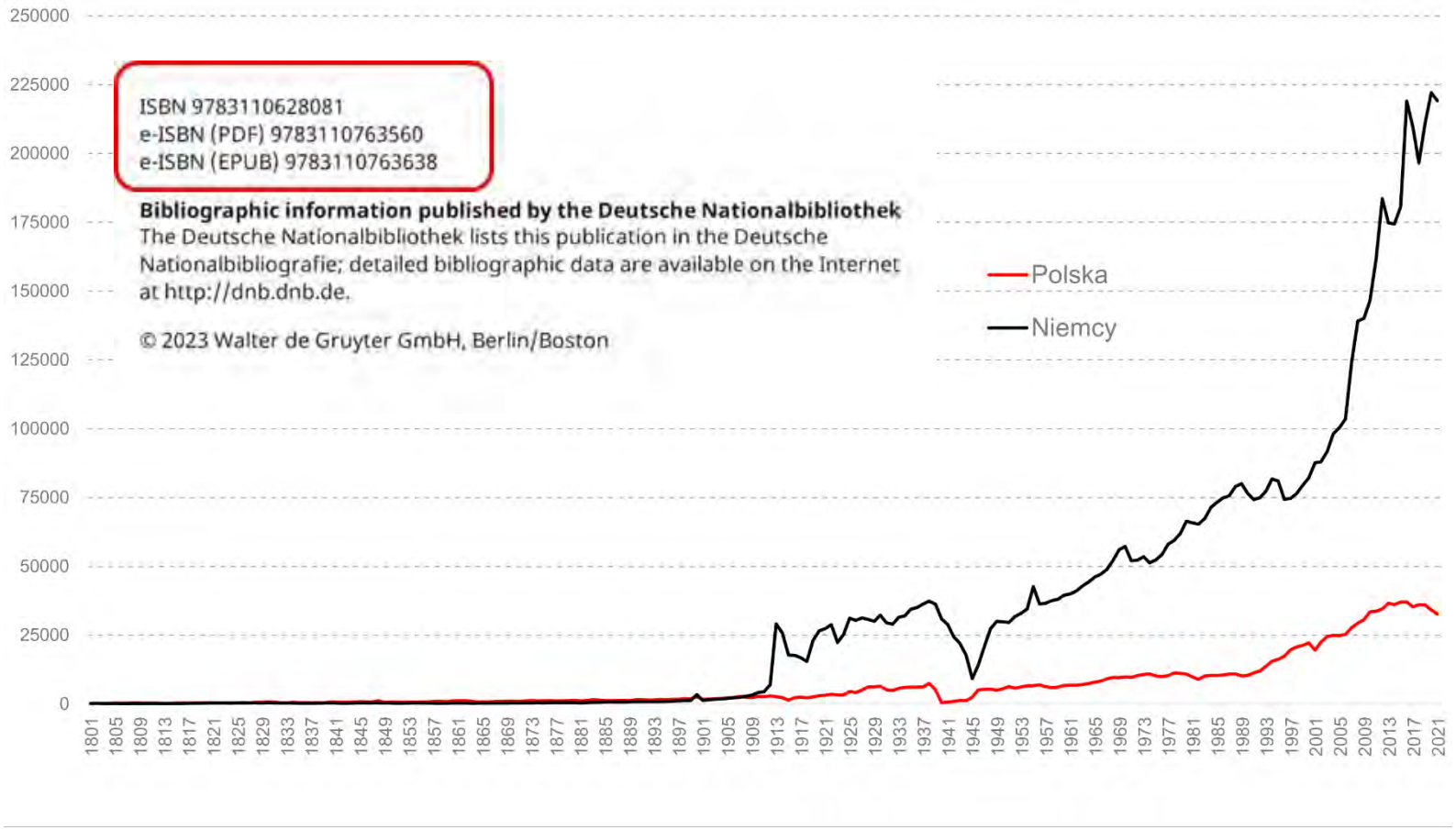
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Glory for the pioneers: Maria Czarnowska



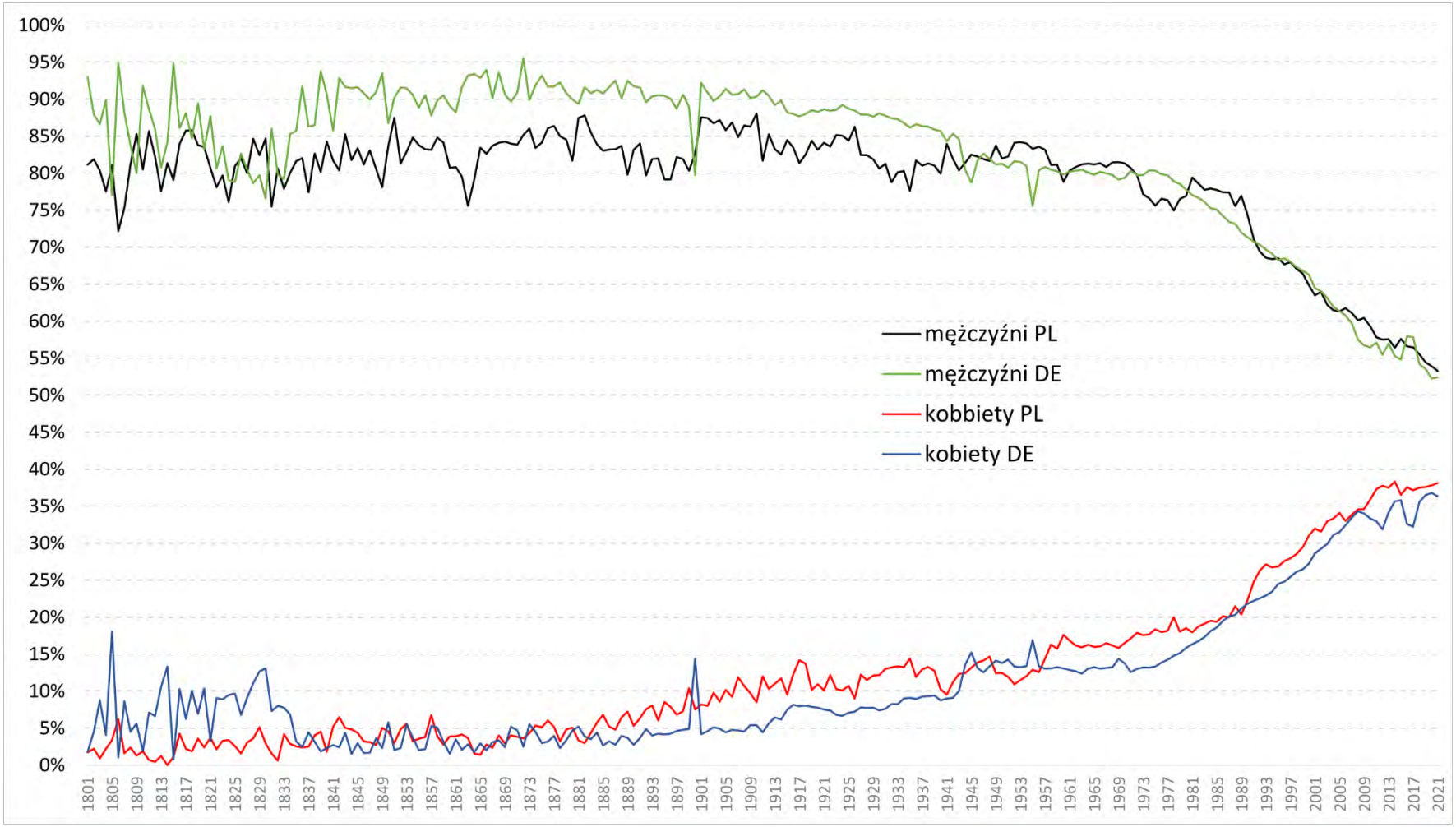
INFORMACJA I KOMUNIKACJA

Number of publications in PL and DE databases



SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

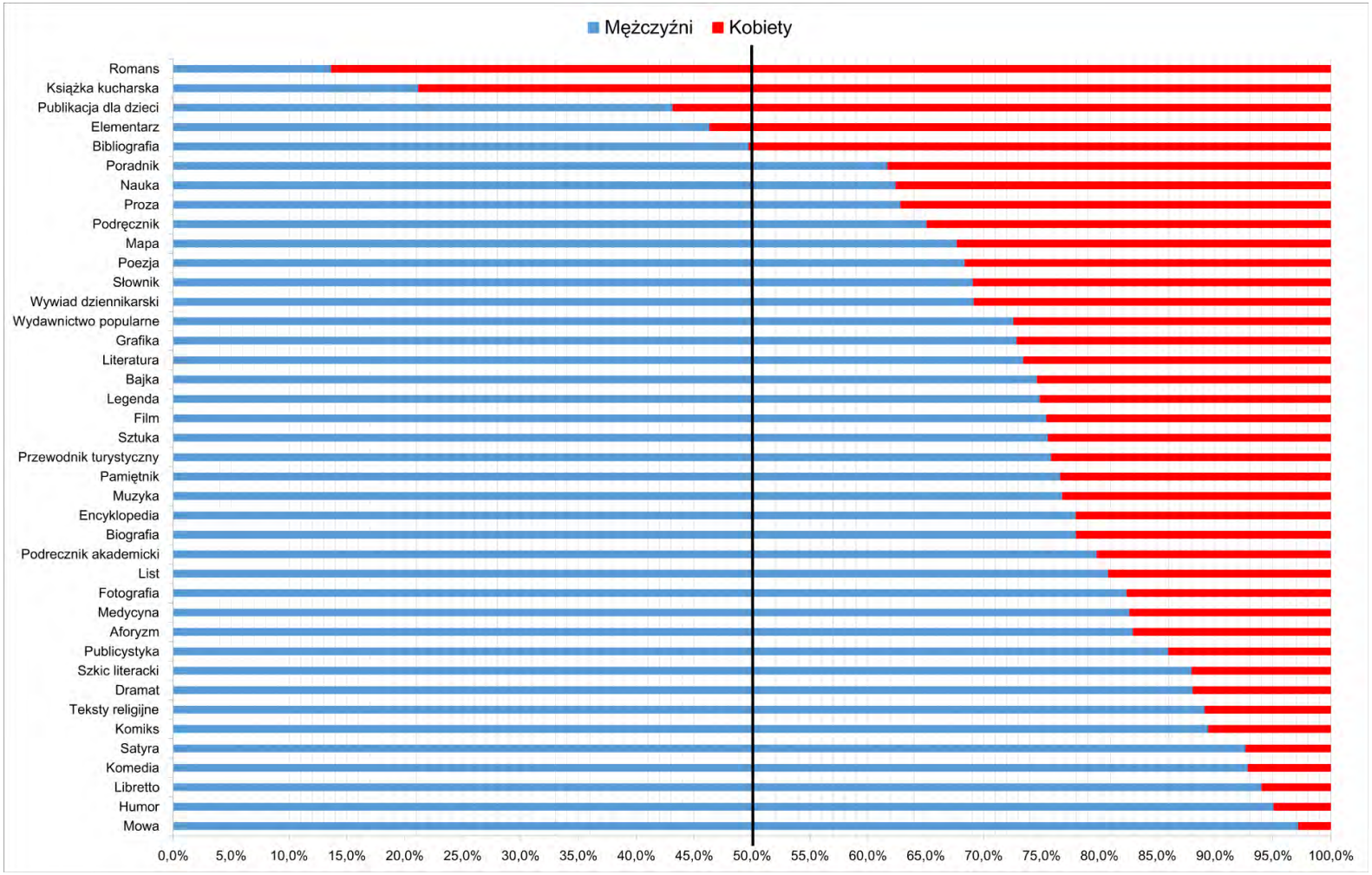
Different countries and languages – one process



INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Writing genres (Poland)



INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Distinctive topics (Poland)

Period	Women's Topics (C)	Men's Topics (C)
1801-1910	WOMAN (5,7) YOUTH (4,8) NOVEL (2,8)	PHD DEFENSE (8.8) REPORT (5,6) ACADEMIC TEXTBOOK (4,3)
1911-1945	YOUTH (4,9) ROMANCE/NOVEL (3,1) NOVEL (2,1)	LAW (3.9) LECTURE/SPEECH (3,1) ACADEMIC TEXTBOOK (2,2)
1946-1980	CHILD (4,7) LANGUAGE (2,6) SCHOOL (2,6)	ELECTRICITY (4.7) DESIGN/CONSTRUCTION (4.1) TRANSPORTATION (4.0)
1981-1999	SCHOOL (4,2) CHILDREN (3,8) ROMANCE/NOVEL (3.3)	TRANSPORTATION (4.8) ELECTRICITY (4.4) DESIGN/CONSTRUCTION (3,8)
2000-2021	SCHOOL (6,3) CHILDREN (3,6) WOMAN (3,6)	MACHINE (4,8) WAR (4,7) COMPUTER SCIENCE (3.9)

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Distinctive topics (Germany)

Period	Women's Topics (C)	Men's Topics (C)
1801-1910	NOVEL (4,6) STORY OF LOVE (2,5) LETTERS (1,9)	GERMANY (3.3) SPEECH/REPORT (3,1) RELIGION (3.0)
1911-1945	NOVEL (3.5) TALE (3,3) CHILDREN (3,2)	TAXES (5.7) ECONOMY (3.0) LAW (2.9)
1946-1980	ROMANCE/NOVEL (3.3) WOMAN (3,1) CHILDREN (2,8)	SCIENCE (3,6) LAW (3,2) ELECTRICITY (2.8)
1981-1999	CHILDREN (2,8) ROMANCE/ NOVEL (2.3) MEDICINE (2.0)	COMPUTER SCIENCE (3.2) MATHEMATICS (2.8) APPLIED SCIENCES (2.8)
2000-2021	ROMANCE/NOVEL (3.1) CHILDREN (2,8) COOKING (2,1)	SCIENTIFIC RESEARCH (2,3) FINANCES (2,1) RELIGION (1,8)

Gender distribution of authors in scientific articles (PL)

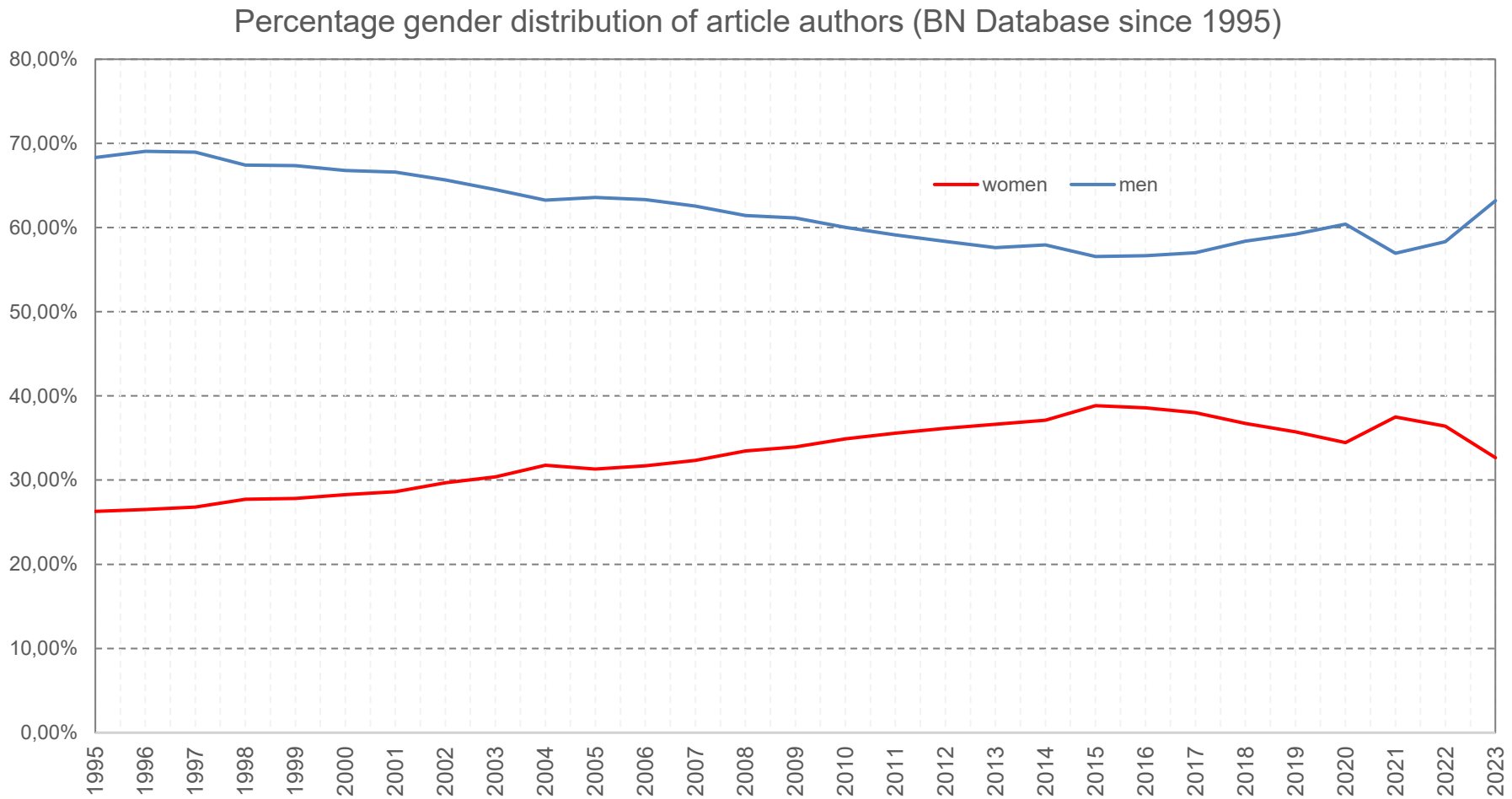
Source: data from the National Library

Volume:

women	944.478	33,98%
men	1.695.587	61,01%
unrecognized	139.241	5,01%
total	2.779.306	

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Gender distribution of authors in scientific articles (PL).



INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH

Results part 3

Dynamics of publishing in Polish cities

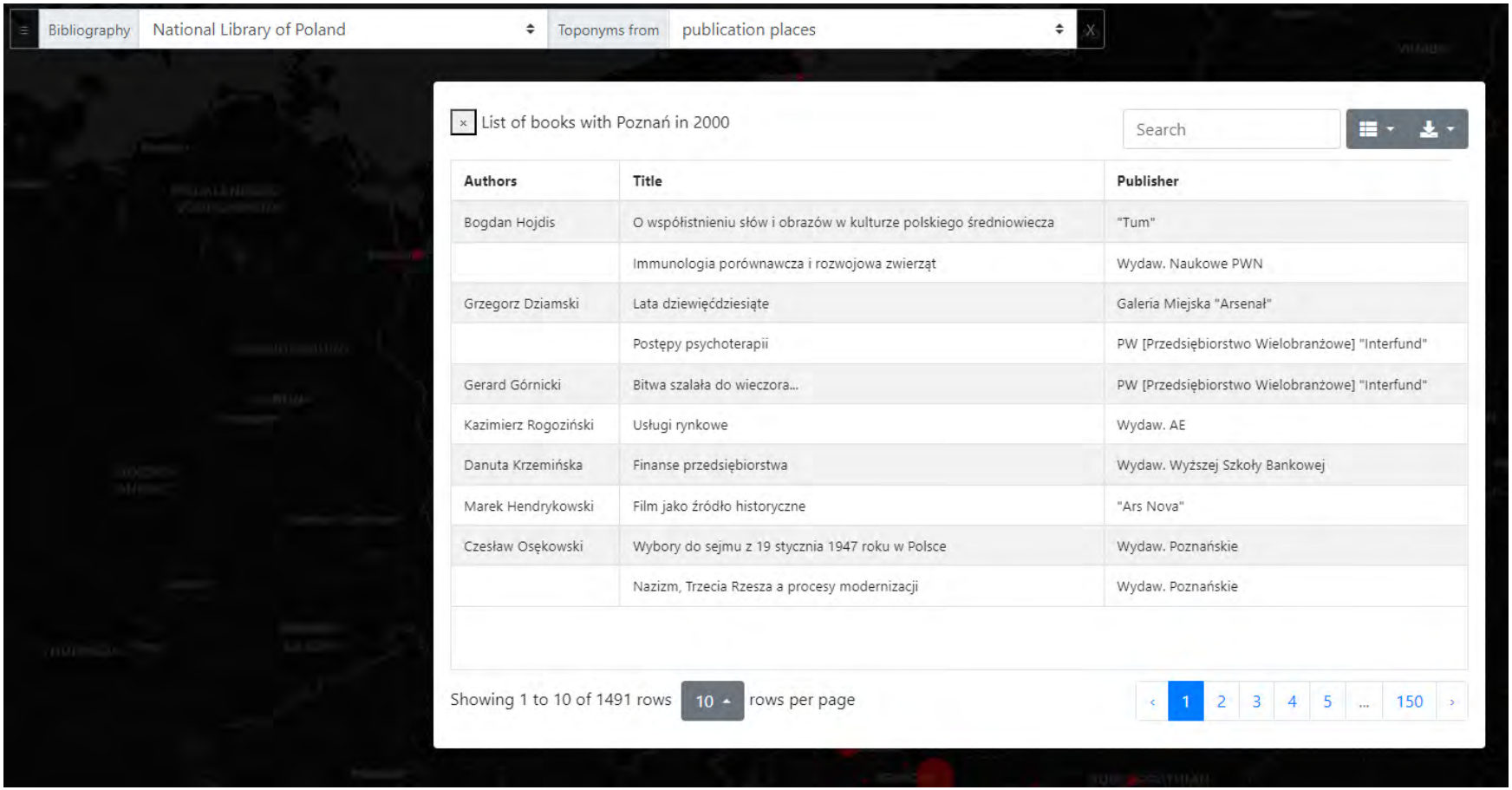
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Polish cities vs. publications

1) Large and medium-sized cities with Warsaw (title numbers / proportions)
2) Large and medium-sized cities without W-wa (title numbers / proportions)
3) Wroclaw and Poznan (title numbers / proportions)
4) Warsaw and Krakow (title numbers / proportions)
5) Warsaw, Cracow, Lviv 1801-1939 (proportions)
6) East and west of Poland (proportions)

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Source of the data



Authors	Title	Publisher
Bogdan Hojdis	O współistnieniu słów i obrazów w kulturze polskiego średniowiecza	"Tum"
	Immunologia porównawcza i rozwojowa zwierząt	Wydaw. Naukowe PWN
Grzegorz Dziamski	Lata dziewięćdziesiąte	Galeria Miejska "Arsenał"
	Postępy psychoterapii	PW [Przedsiębiorstwo Wielobranżowe] "Interfund"
Gerard Górnicki	Bitwa szalała do wieczora...	PW [Przedsiębiorstwo Wielobranżowe] "Interfund"
Kazimierz Rogoziński	Usługi rynkowe	Wydaw. AE
Danuta Krzemińska	Finanse przedsiębiorstwa	Wydaw. Wyższej Szkoły Bankowej
Marek Hendrykowski	Film jako źródło historyczne	"Ars Nova"
Czesław Osękowski	Wybory do sejmu z 19 stycznia 1947 roku w Polsce	Wydaw. Poznańskie
	Nazizm, Trzecia Rzesza a procesy modernizacji	Wydaw. Poznańskie

Showing 1 to 10 of 1491 rows 10 rows per page

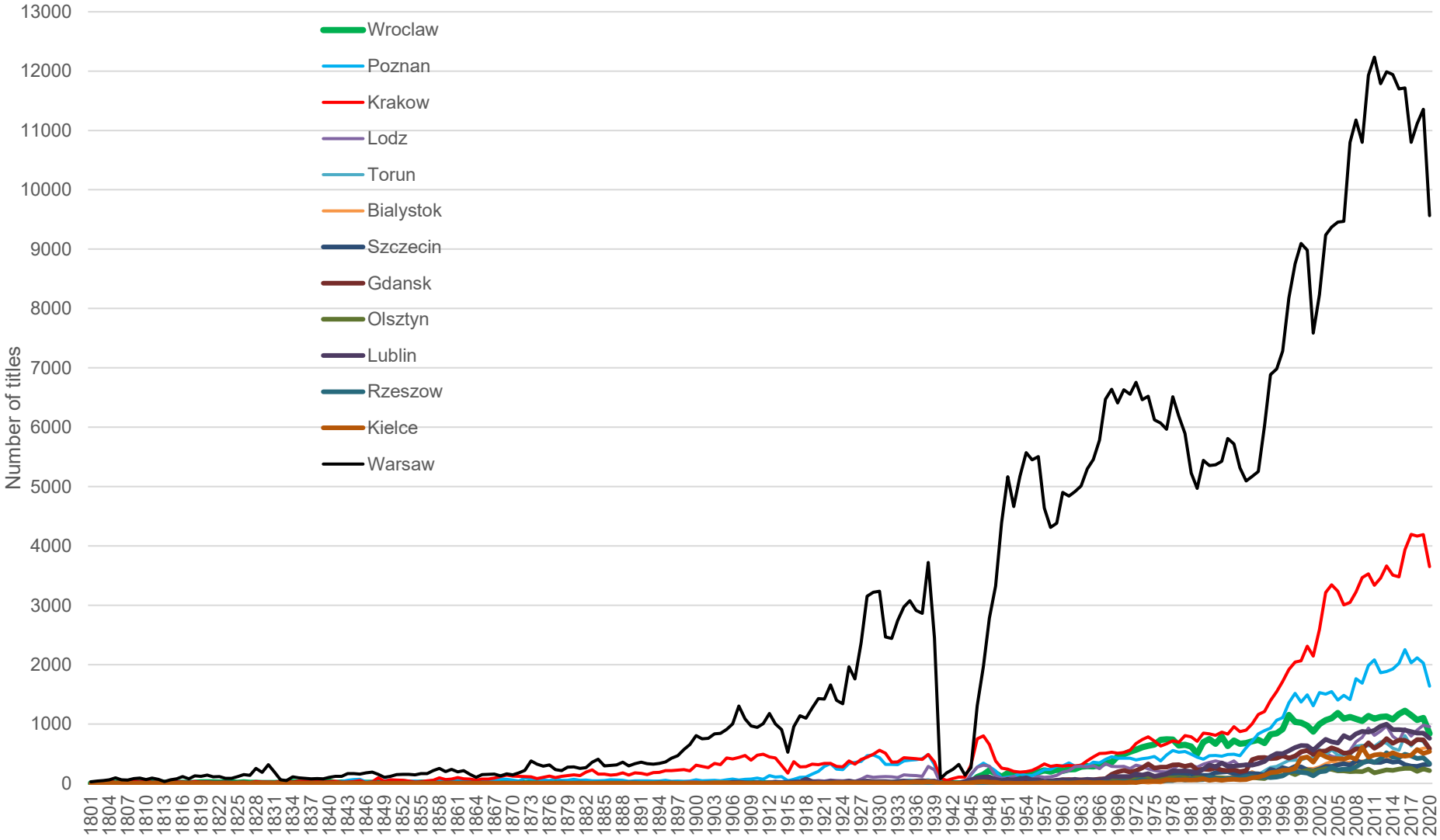
‹ 1 2 3 4 5 ... 150 ›

INFORMACJA I KOMUNIKACJA

Source of the data

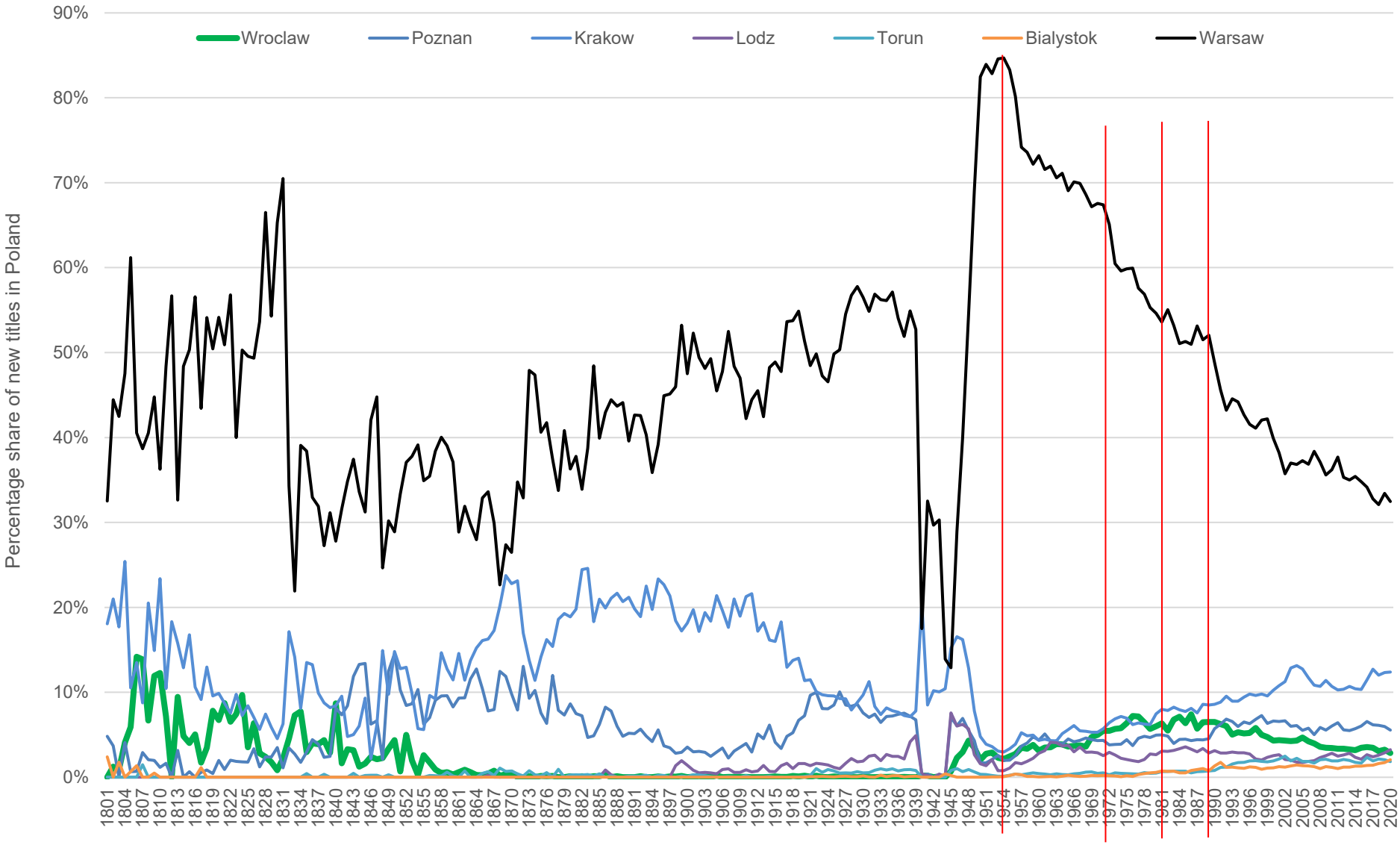
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Large and medium-sized cities with Warsaw (numbers)



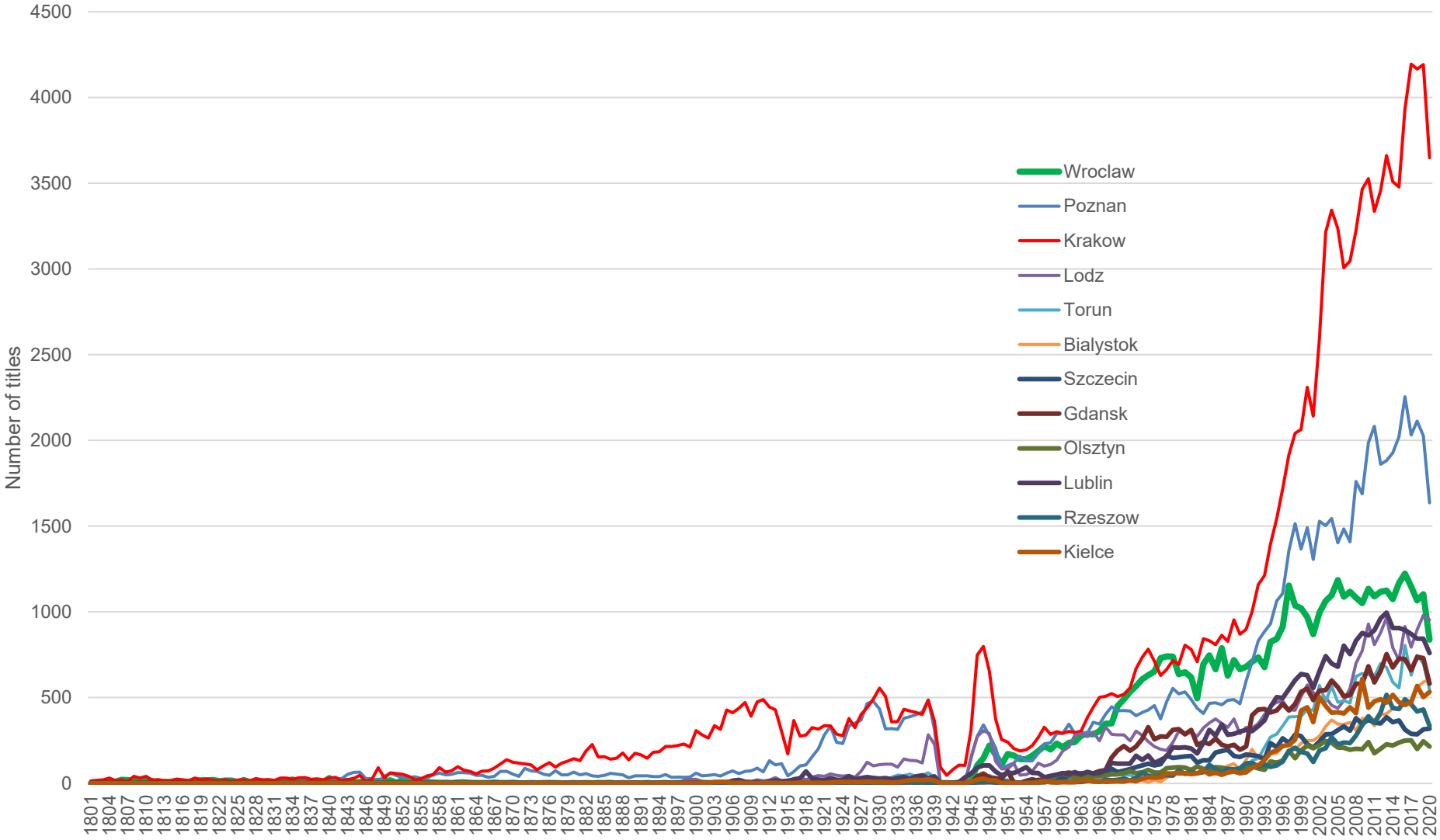
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Large and medium-sized cities with Warsaw (proportions)



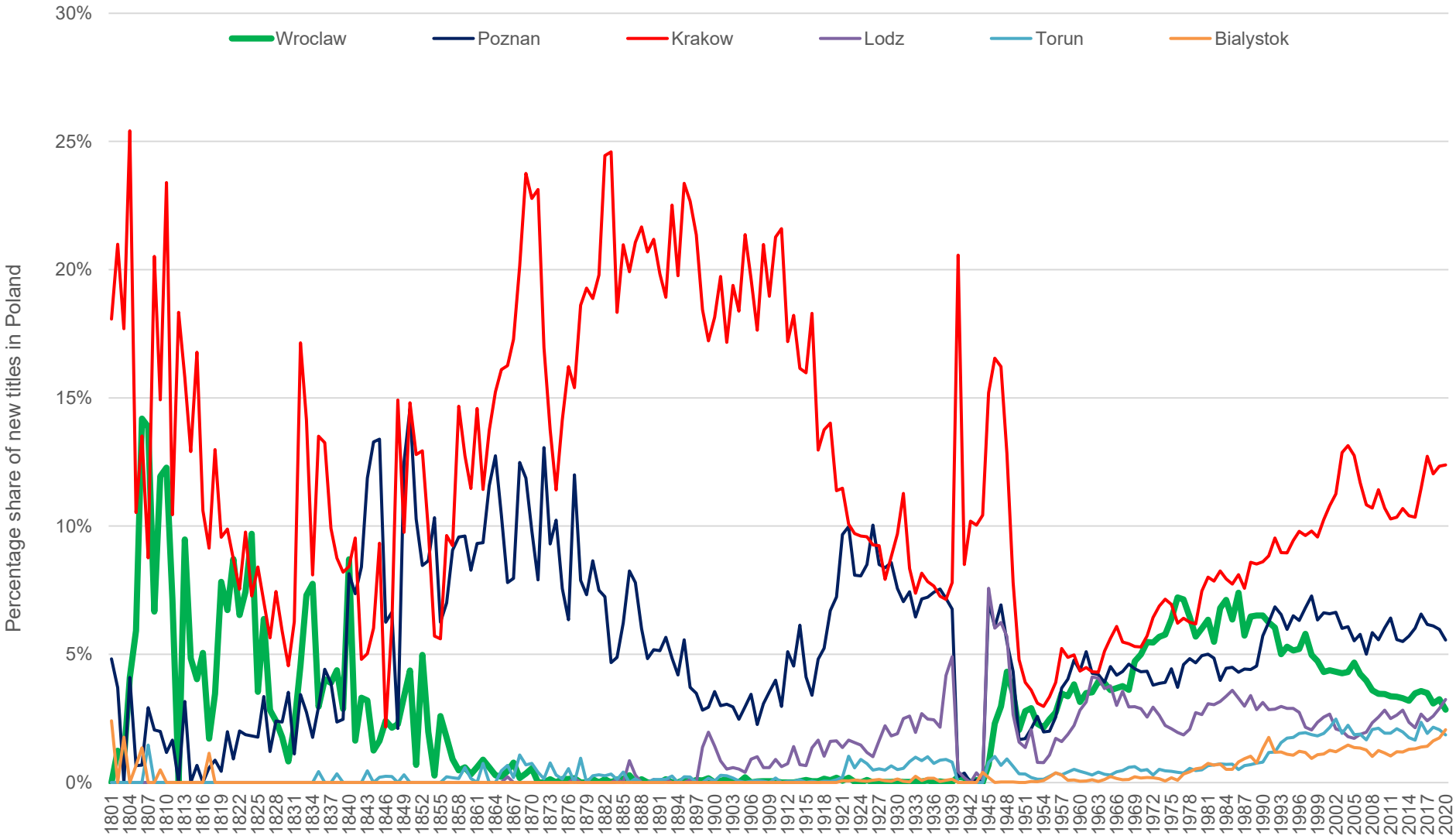
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Large and medium-sized cities excluding Warsaw (number of publications)



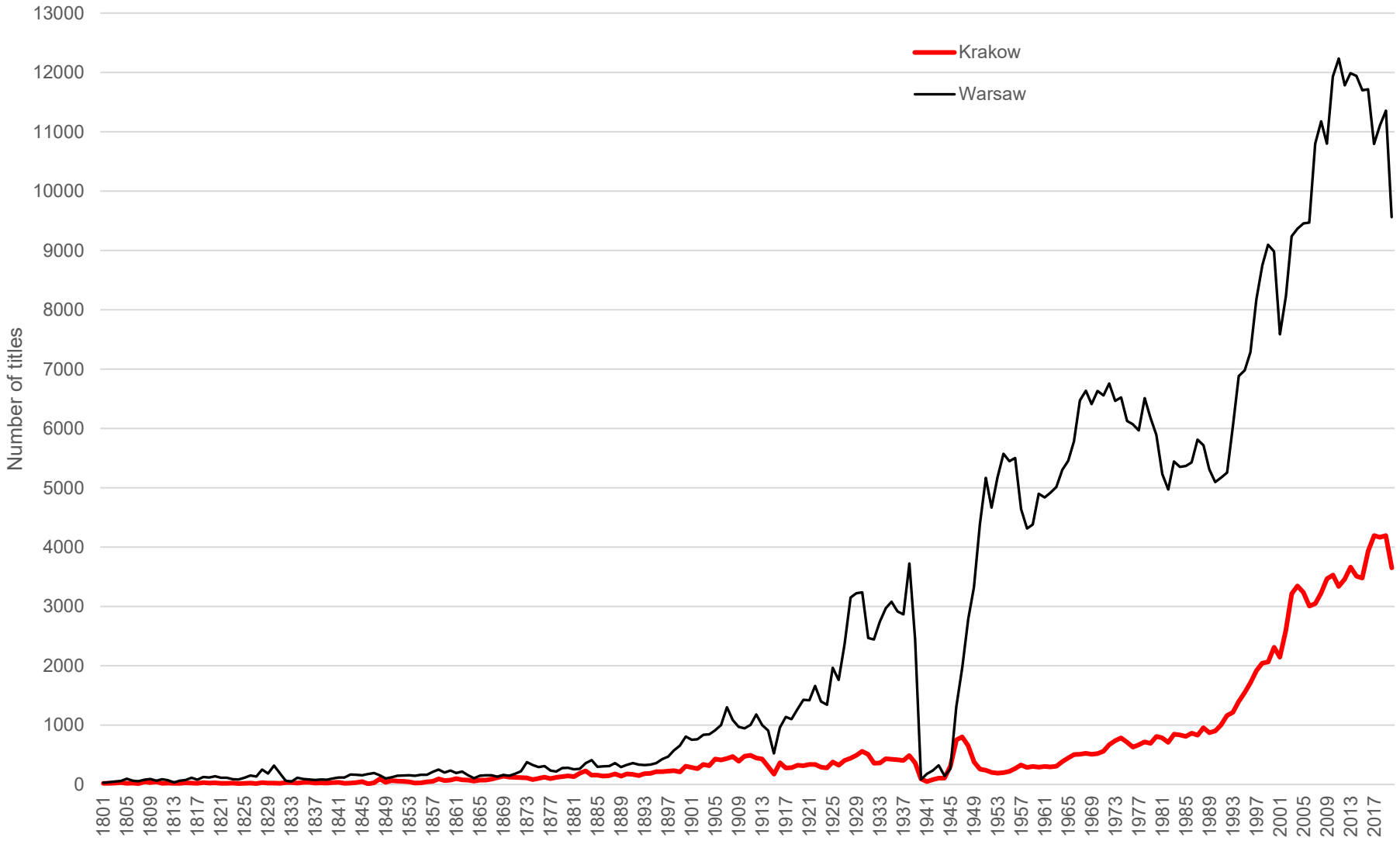
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Large and medium-sized cities without Warsaw (proportions)



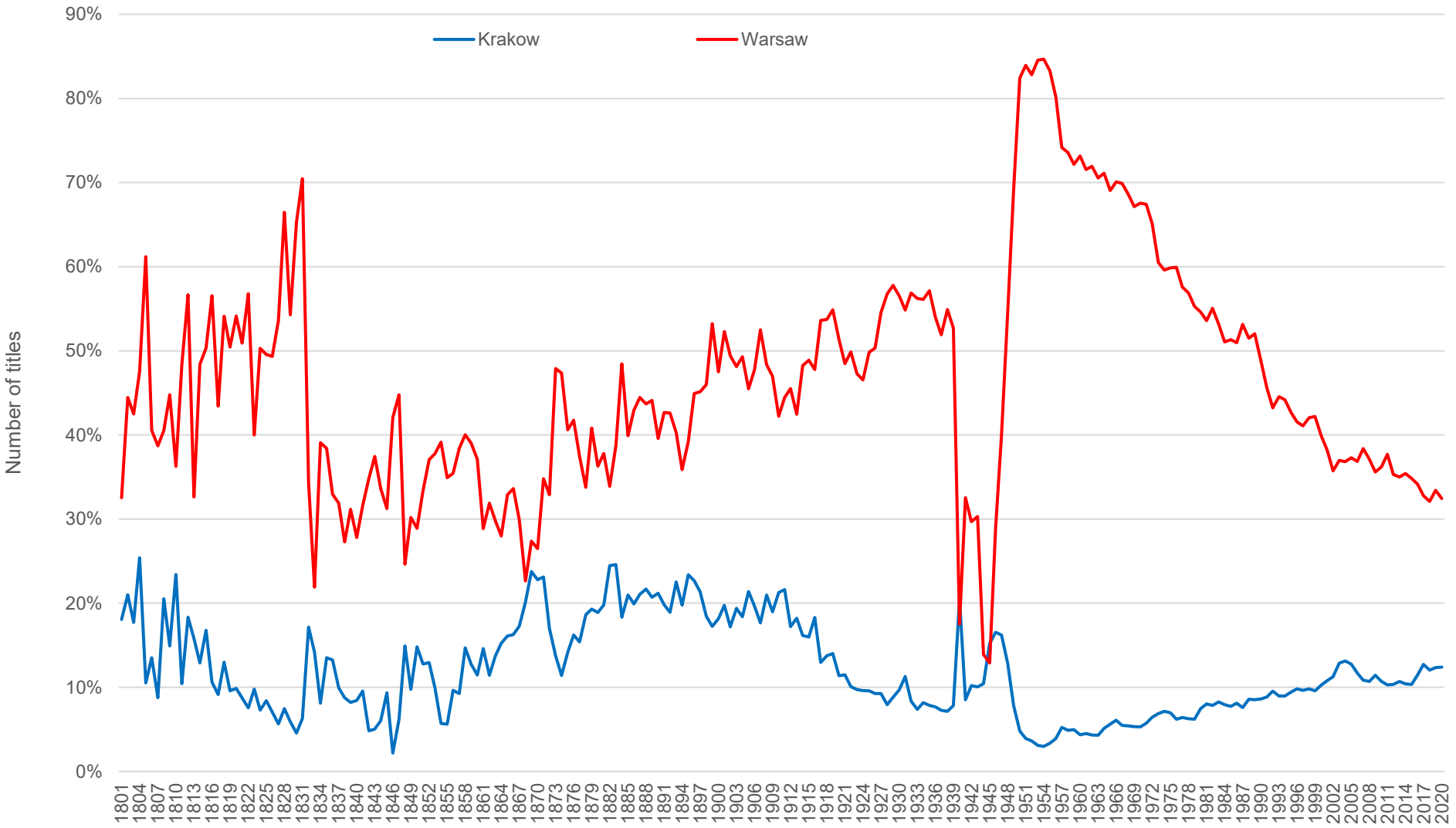
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Warsaw and Krakow (title numbers)

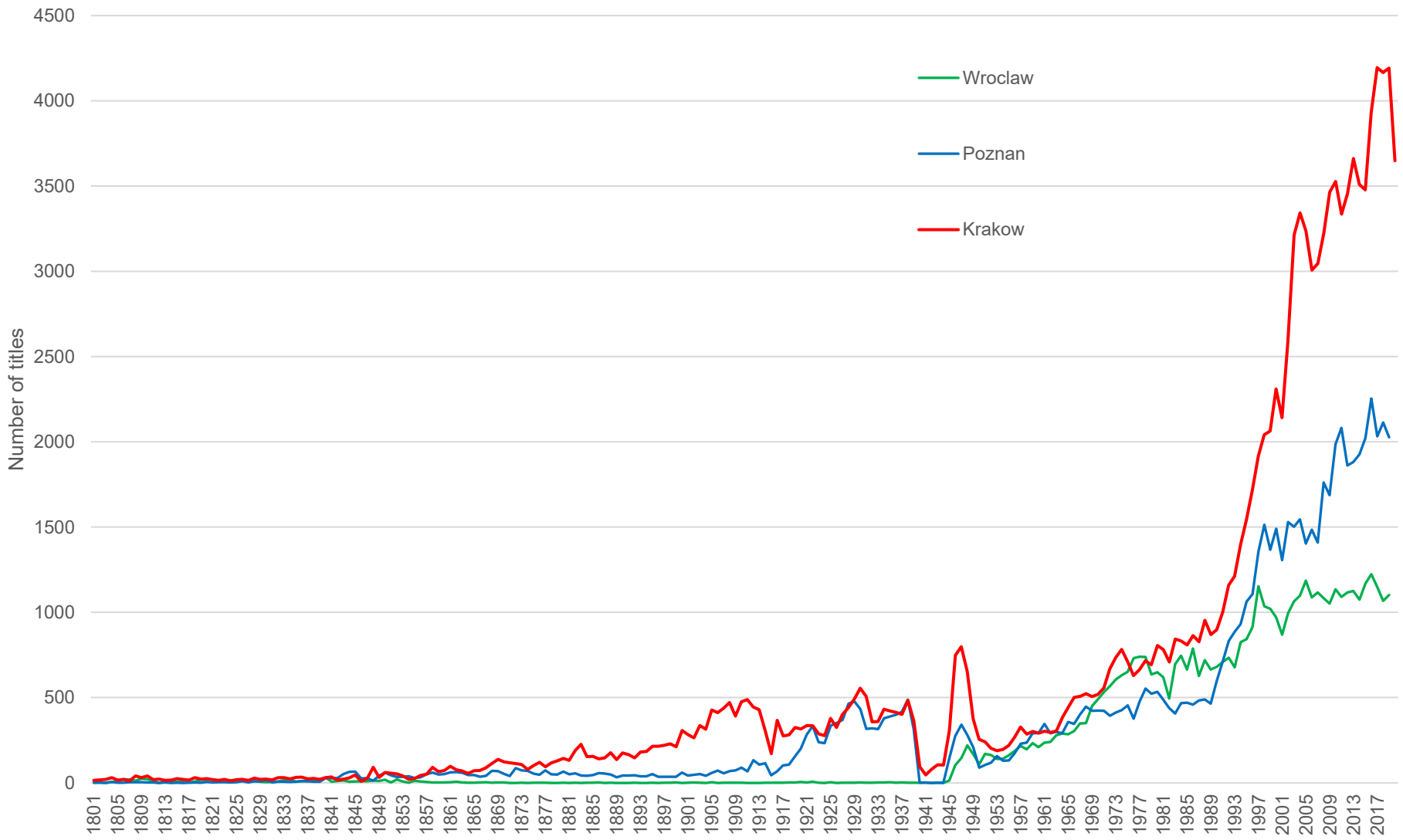


SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Warsaw and Krakow (proportions)

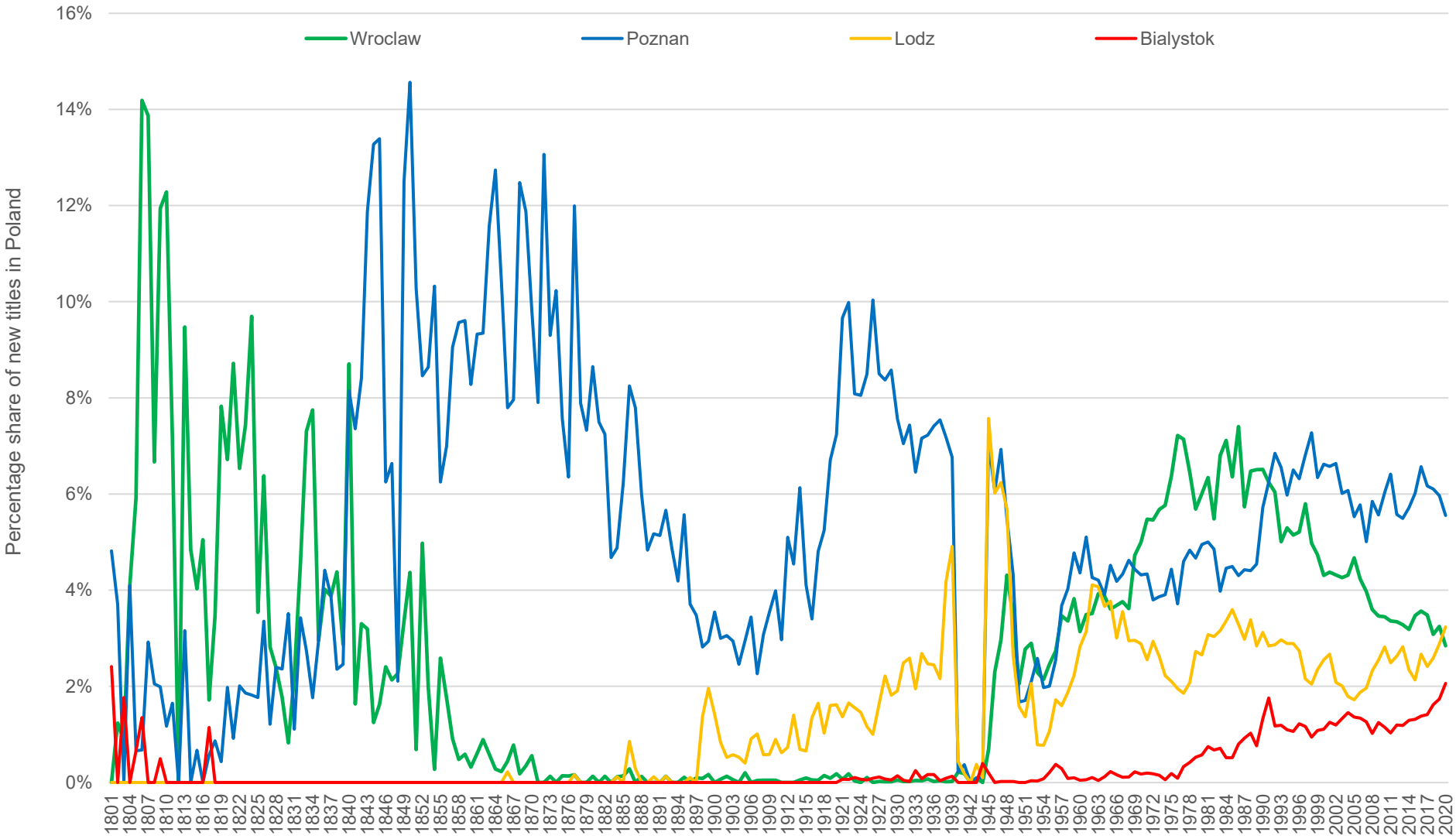


Krakow, Poznan, Wroclaw (title numbers).



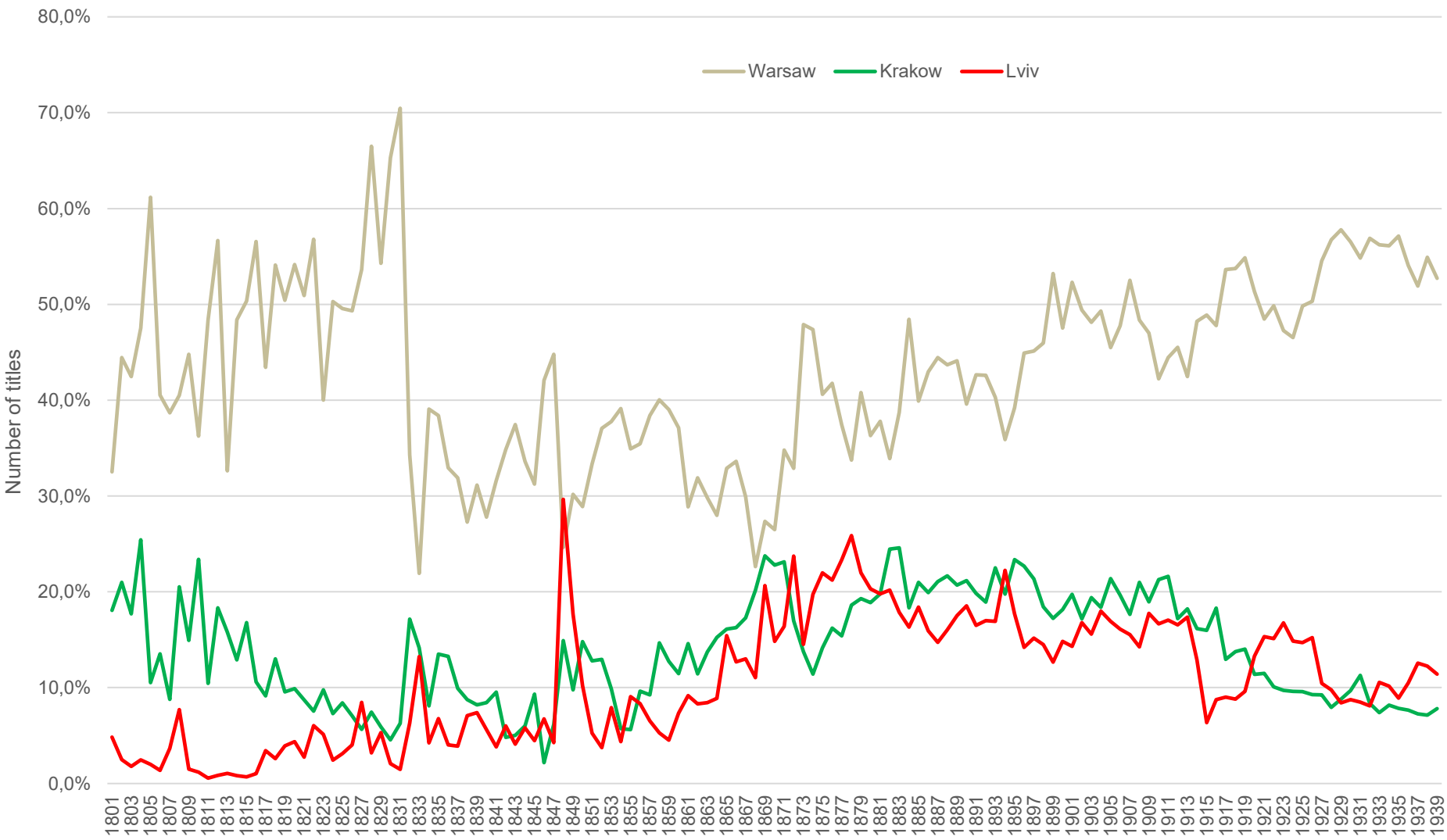
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Lodz, Bialystok, Poznan, Wroclaw (proportions)



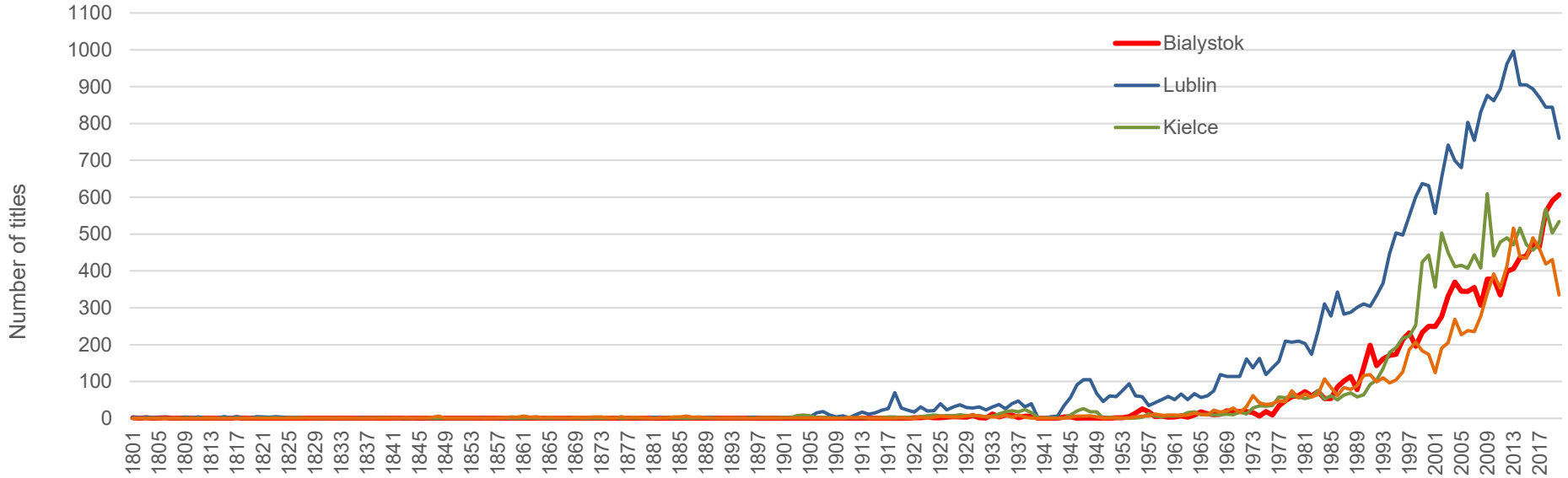
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

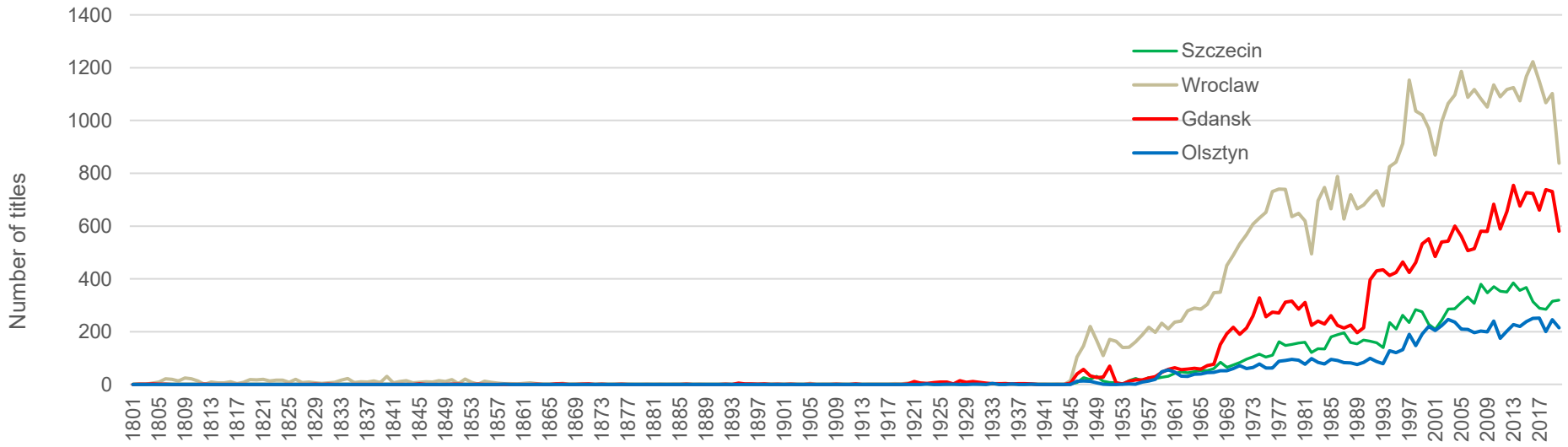
Warsaw, Krakow, Lviv (proportions to 1939)



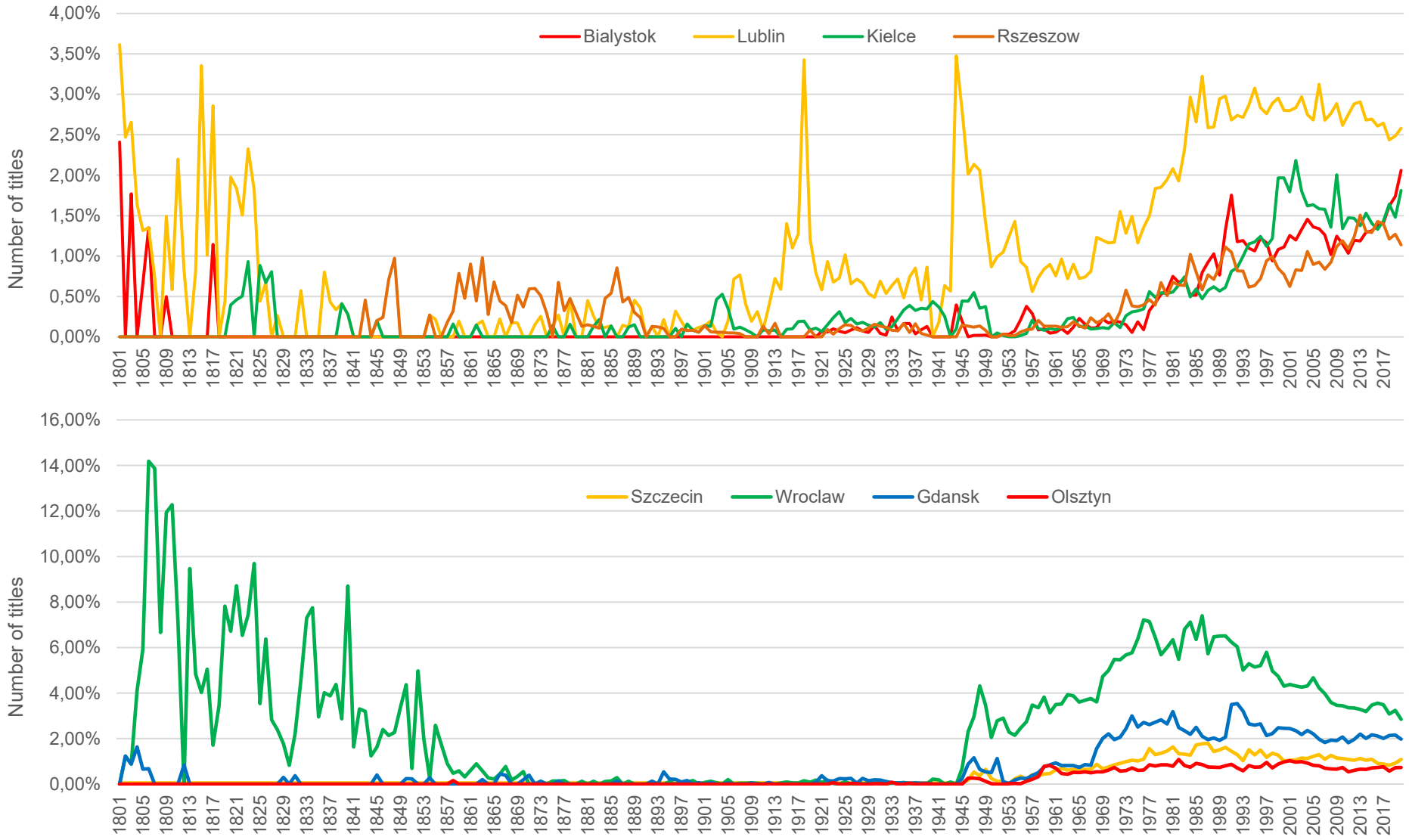
SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

East and West (?) of Poland (title numbers).



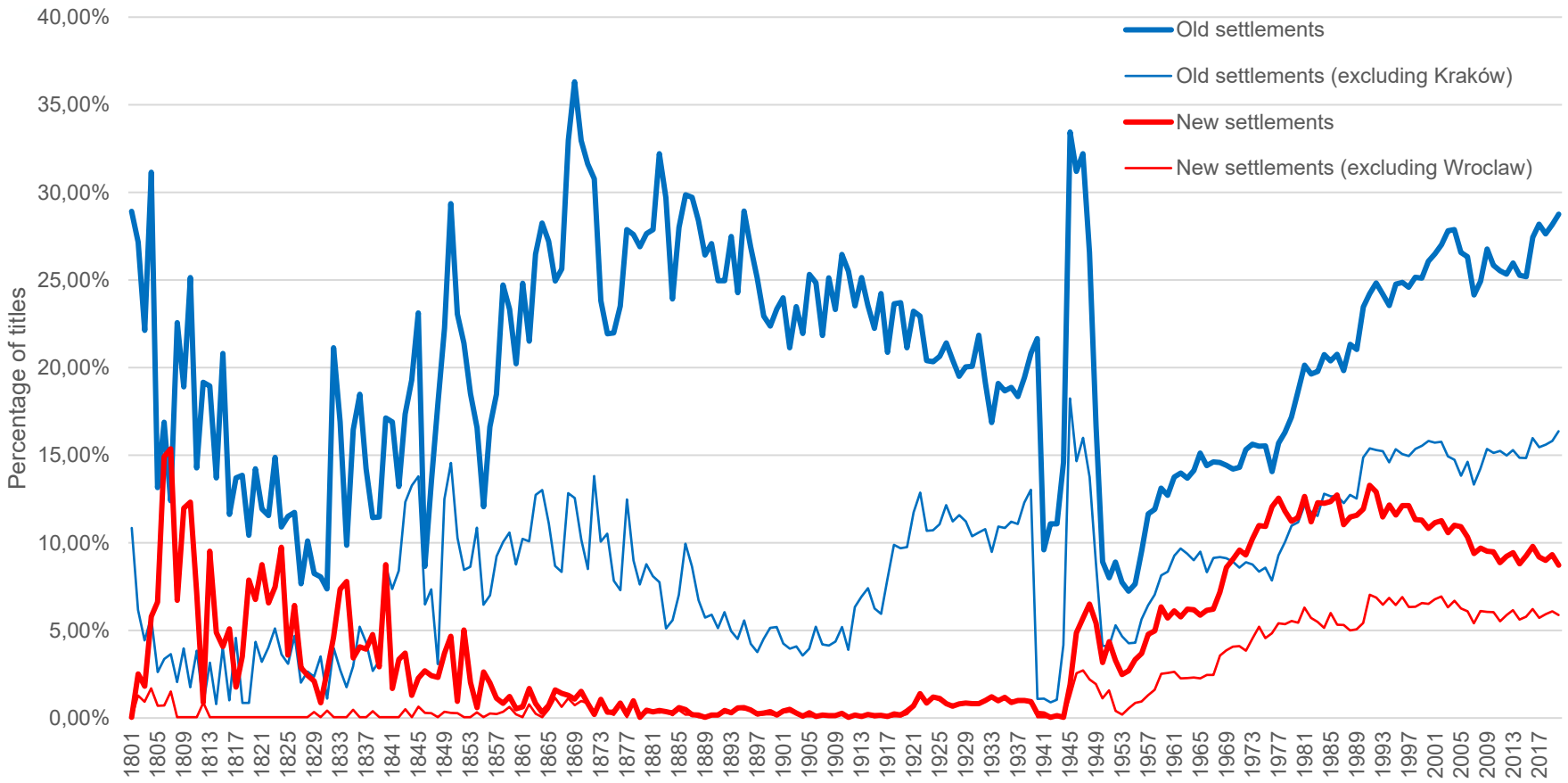


East and West (?) of Poland (title numbers).



SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Trying to explain the differences: rooting



INFORMACJA I KOMUNIKACJA

SBP
STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH

Summaries and conclusions

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Summary

- The official emergence of bibliographic data the research repertoire of NLP, AI and digital humanities
- Proof of the great potential of big bibliographic data
- Tangible effect of beneficial synergy of cooperation between computer science and humanities

but also

- A token of appreciation for the multitudes of faceless, poorly paid librarians who have painstakingly done their jobs for decades

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Barriers and problems

- Lack of understanding and appreciation for research of great bibliographies
- Lack of rooting in traditional disciplines
- The slow formation of the identity of digital humanities and humanistic strands of *data* science

SBP
STOWARZYSZENIE BIBLIOTEKARZY POLSKICH

Some self-advertising

SBP
STOWARZYSZENIE
BIBLIOTEKARZY
POLSKICH

Thank you for your attention!