Large bibliographies in the digital humanities. Data acquisition, processing, applications and limitations (2)

prof. dr hab. Adam Pawłowski
Institute of Information and Media Sciences, University of Wrocław

Interreg, Digital Humanities for the Future, University of Olomouc, 16-17.01.2023

LECTURE PROGRAMME

- Bibliographic data as a subject of research
- Data for research (sources)
- Potential hypotheses and objectives of the research
- Data processing methods
- Sample results and interpretation

Data, methods

Uniwersytet Wrocławski
CLARIN
Common Language Resources and Technology Infrastructure
Politechnika Wrocławska

RESEARCH DATA

1. **The digital catalog of the National Library (BN)**
   – period covered: 1800-2010
   – volume: ca 1.9 million records (after filtering ca 1.25 million).

2. **The digital catalogs of the Deutsche Nationalbibliothek**
   – period covered: 1800-2010
   – volume: ca 7.5 million records (after filtering 5 million).

3. **The type of content studied**
   – book, that is, what has an ISBN number.

HYPOTHESES AND OBJECTIVES OF THE RESEARCH

Hypotheses:

- Gender participation in intellectual life can be measured by the number of publications.
- Internal variation in a country's level of development can be measured by the number of publications and visualized by projecting the data onto a map.

General assumption:

- Book publications are a synthesis of many specific indicators (education level, social status, wealth).

TOPONYM EXTRACTION FROM TITLES – METHODS

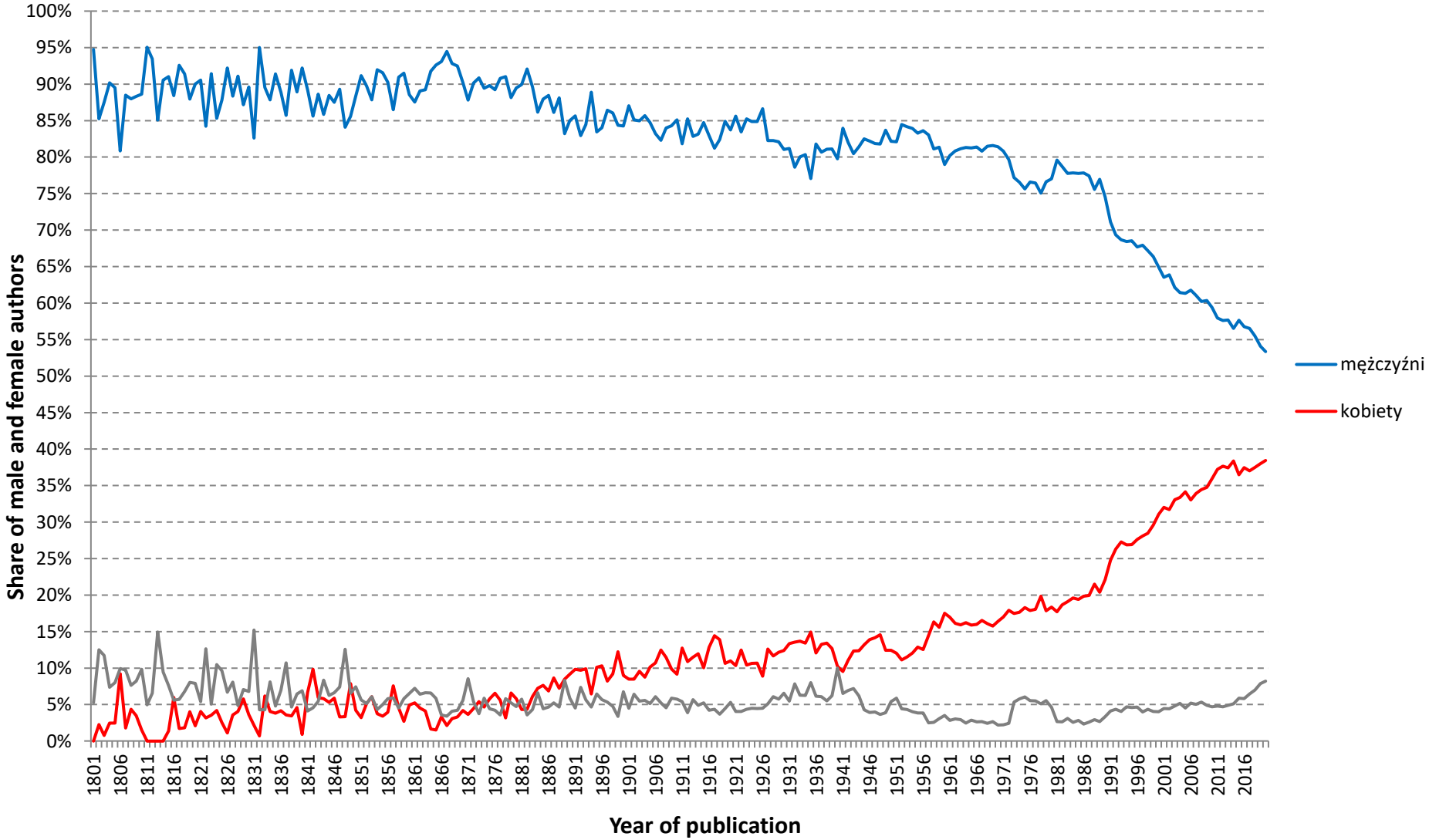
- Title (text) -> list of place names
- Named entity recognition tool for Polish
  - machine learning tool, not a dictionary
  - **detection of n.e.r** – PolDeepNer
    - XMLRoberta architecture
    - starting language model HerBERT (large BERT for Polish)
    - trained to detect NE boundary detection and 82 categories
    - 20 types are place names (i.e., country, city, lake, mountain)
  - **lemmatization of detected name entities (NE)**
    - Polish is highly inflectional
    - *Niemieckiej Republice Ludowej -> Niemiecka Republika Ludowa*
    - tagger (recurrent neural network) + Polem for multiword NE

Are men from Mars and women from Venus?

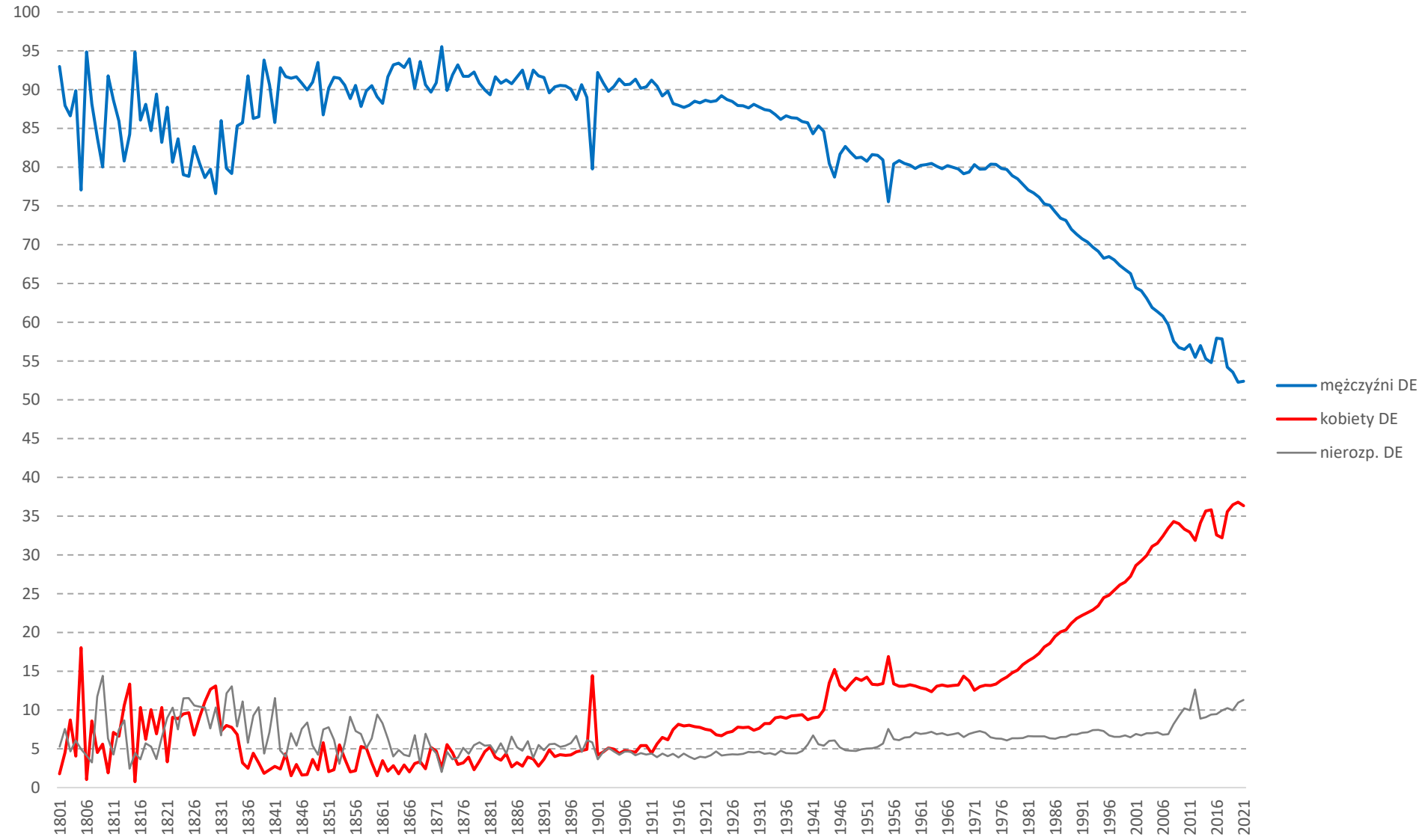
Interreg, Digital Humanities for the Future,
University of Olomouc, 16-17.01.2023

Uniwersytet Wrocławski
CLARIN
Common Language Resources and Technology Infrastructure
Politechnika Wrocławska

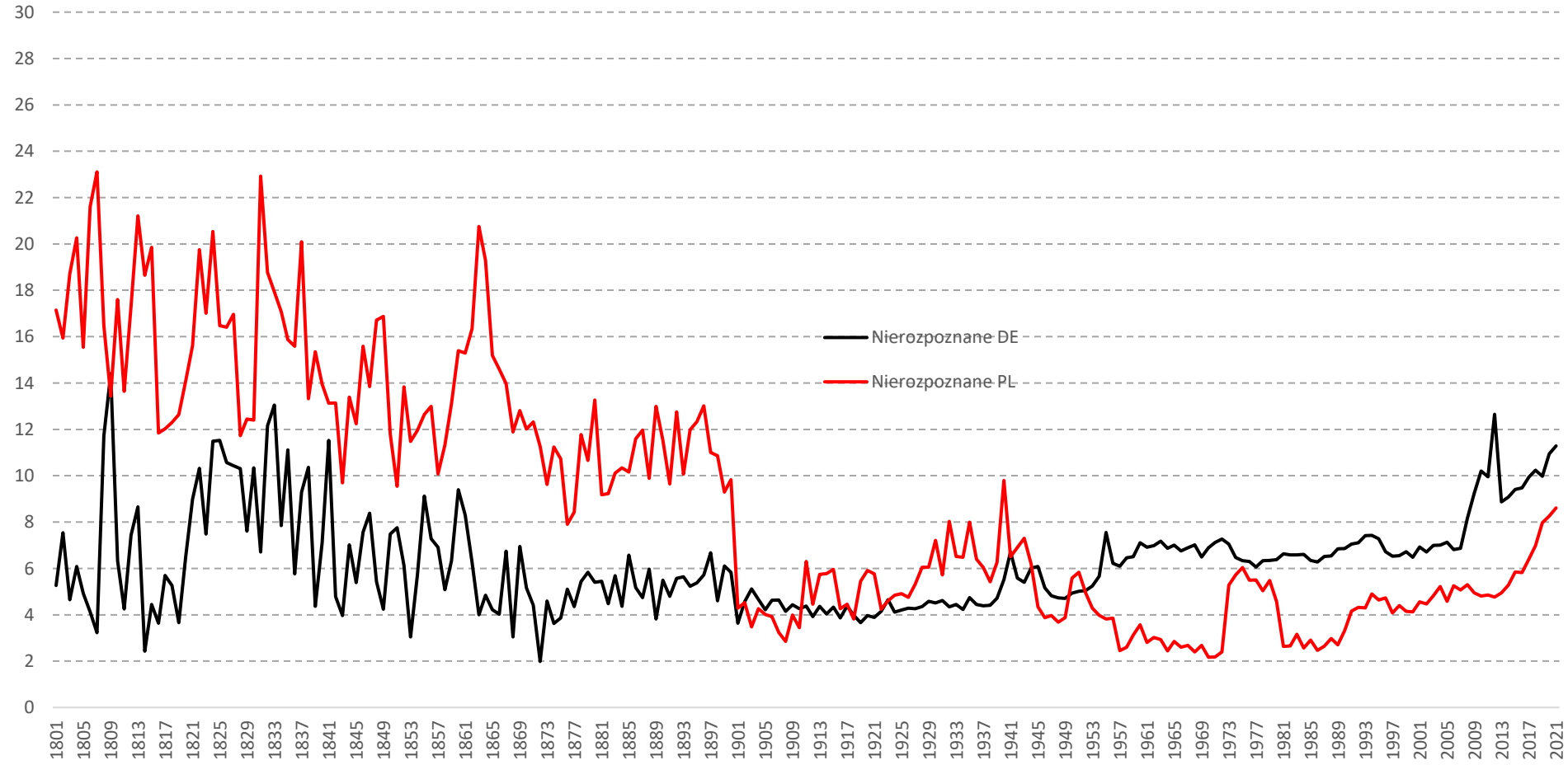
Share of women and men among authors of Polish publications (1801-2020)



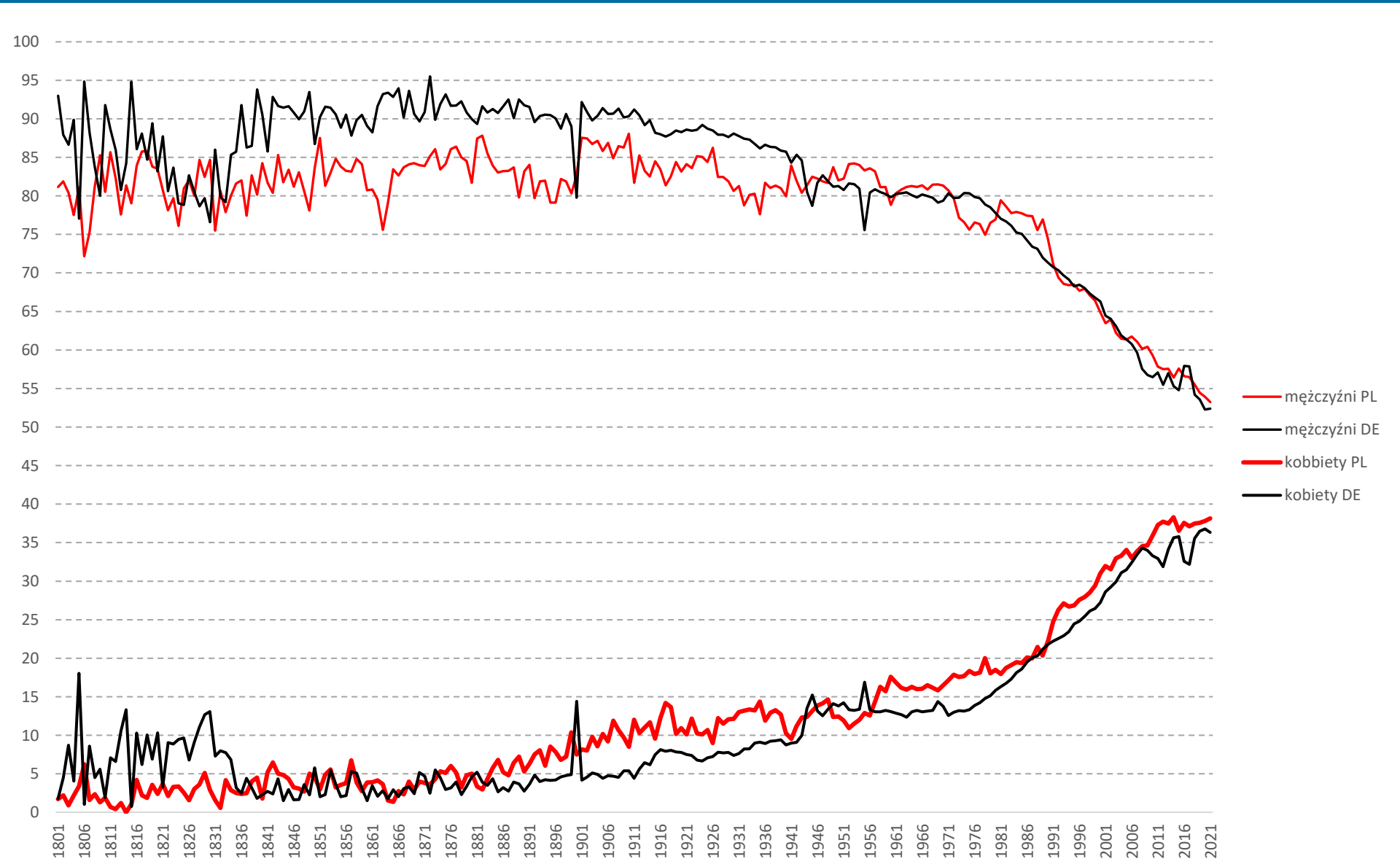
Share of women and men among authors of German publications (1801-2020)



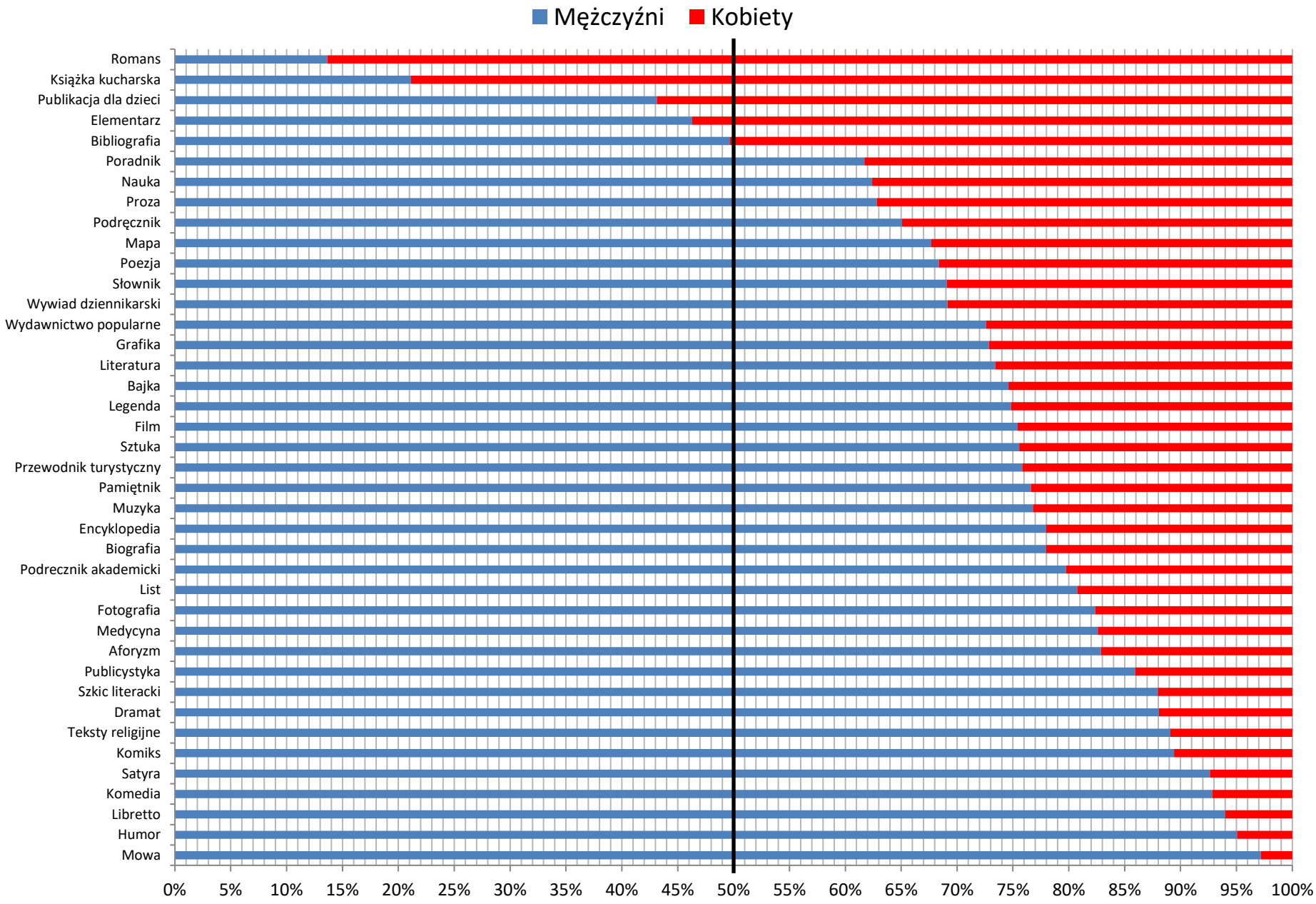
SHARE OF AUTHORS WITH UNRECOGNIZED GENDER (1801-2020)



TOTAL SHARE OF WOMEN AND MEN AMONG BOOK AUTHORS (PL AND DE, 1801-2020)



Writing genres by gender



TOPICS RECOGNIZED BY GENDER (PL 1801-1910)

1,065946739	szkół	języka	użytku	nauki	nauka	młodzieży	grammatyka	średnich	gramatyka	podręcznik
1,208173195	przewodnik	podróży	podróż	wspomnienia	listy	pamiętniki	nowele	czyli	nad	leczenie
-4,199277114	dr	geschichte	universität	herausgegeben	erlangung	nebst	inaugural	dissertation	polnischen	ueber
-1,402691415	św	kazania	przez	roku	kościele	ks	maryi	kościoła	świętych	czyli
-1,45697841	ludu	polskiego	czyli	narodu	polska	polacy	pieśni	polsce	polski	kilka
-1,938827795	historya	dzieje	dziejów	historyi	polski	wieku	historyczne	czasów	polskiego	polsce
-8,784333818	et	de	dissertatio	ad	quam	inauguralis	universitate	ordinis	the	di
-2,290138004	pisma	dzieła	mickiewicza	adama	słowackiego	juliusza	krasickiego	sienkiewicza	ignacego	henryka
1,347564629	poezye	poemat	poezje	poezyi	wiersze	literackie	wybór	wierszem	prozą	poema
-1,079334881	roślin	czyli	zwierząt	bydła	oraz	uprawa	gospodarzy	hodowla	gospodarstwa	nauka
-1,576453229	przez	tom	biblioteka	wydanie	przekład	przełożył	ii	katalog	polski	polskich
2,801818191	powieść	historyczna	współczesna	powieści	tomach	romans	wieku	opowiadanie	czasów	historyczny
1,1694024	obrazki	obrazek	obrazy	szkice	życia	szkic	obraz	historyczne	album	natury
-3,90054051	wykład	zasady	filozofii	czyli	nauka	filozofia	nauki	fizyki	studyum	ekonomii
5,663818634	kobiet	kobieta	kobiety	czyli	romans	matki	maryi	pani	matce	przez
4,856487028	dzieci	młodzieży	opowiadania	opowiadanie	bajki	powiastki	książeczka	powieści	książeczek	powiastek
-1,356698357	de	la	et	par	le	les	en	sur	pologne	pour
-5,565029655	przez	towarzystwa	dnia	roku	prawa	przy	członka	nauk	posiedzeniu	sprawozdanie
-2,090754413	dnia	roku	mowa	jana	przez	miana	stanisława	króla	kościele	augusta
-1,264957381	aktach	akcie	dramat	komedya	opera	tragedya	wierszem	pięciu	jednym	komedja

Topics recognized by gender DE 1801-1910

-2,977560595	kirche	evangelischen	luthers	christlichen	martin	heiligen	jesu	christliche	evangelische	kirch
2,565262911	roma	roman	römische	römischen	scott	walter	geschicht	bd	rom	sprüch
-2,663346713	geschichte	beitr	nebst	band	archiv	rudolf	virchow	pathologische	physiologie	anatomie
-3,094389154	190	189	jahre	187	188	geh	186	184	jahren	juni
-1,649554065	geschichte	alten	gegenwar	geschicht	allgemeine	kunst	zeit	völker	beiträge	of
1,853779698	gedicht	gedichte	poem	dichtunge	other	poe	the	and	of	dichtung
4,668728134	erzählunge	erzählun	erzählungen	novelle	of	jugen	the	geschichten	erzählung	romane
1,213167879	brief	alphons	dürr	leipzi	jügel	gotthelf	briefe	kumme	dür	paul
-3,340130264	deutschen	deutsche	deutschland	geschichte	deutscher	deutschlands	deutsches	deutschlan	berliner	berli
1,457417486	the	ueber	familie	de	of	ak	stammbaum	novell	or	schwank
-1,809608439	anleitung	handbuch	nebst	lehrbuch	sowie	unterricht	praktische	leitfaden	grammatik	system
-1,812945482	de	et	quae	platons	praeter	anatomie	qua	nova	evulgavit	dominicus
-1,337851875	werk	friedrich	sämmtliche	johann	heinrich	gesammelte	schrifte	wilhelm	ludwig	leben
-1,625769333	bedeutung	recht	de	pharmazeutische	beiträge	abhandlunge	lehre	entwicklung	aufgaben	untersuchungen
-1,65391821	geschichte	nr	jahre	regiments	1870	de	infanterie	jahrhundert	österreichischen	seit
-1,331977878	schrifte	of	de	the	and	et	la	bücher	catalogue	gesammelte
1,942097888	korresponden	korrespondenz	herrman	eva	novelle	postkart	ilia	brief	erinnerunge	theate
1,083131579	le	akte	en	aufz	repre	sente	tre	la	actes	fois
1,029512534	werk	sämtliche	gesammelte	sämmtliche	ausgewählte	teile	lessings	werke	13	reuters

TOPICS RECOGNIZED BY GENDER (PL 1911–1945)

3,091923177	miłości	miłość	kobieta	kobiety	pani	romans	matka	potęga	matki	panna
-3,860609217	prawo	pracy	oraz	przepisy	prawa	polsce	podręcznik	kodeks	ustawy	organizacja
-1,645343911	języka	słownik	polsko	gramatyka	polskiego	polski	samouczek	niemieckiego	wyrazów	słowniczek
1,385231303	pozorna	śmierć	nowele	nad	pustyni	serce	sonety	słońce	ziemia	treny
-1,947704585	dziejów	polski	dzieje	historja	polsce	polskiej	wieku	polskiego	zarys	historji
-1,751744126	wojny	pułku	wojennej	wojna	1914	piechoty	historji	1918	zarys	roku
-1,017421225	aktach	akcie	dramat	komedja	odsłonach	sceniczny	obrazek	ch	śpiewami	sztuka
-2,252870263	zarys	zadań	podręcznik	wykładów	geografja	zbiór	teorja	arytmetyka	zasady	teorji
-1,324906832	józef	rocznicę	jan	piłsudski	pan	historyczna	tadeusz	czasów	powieść	jana
1,468293725	szkół	podręcznik	powszechnych	klasy	szkoły	średnich	powszechnej	nauki	kl	szkole
2,08849997	powieść	współczesna	historyczna	młodzieży	tomach	wieku	powieści	życia	tajemnica	sensacyjna
1,155403373	ludzie	powieść	bracia	krzyżacy	żydzi	polacy	badacze	pisarze	podróżnicy	zdobywcy
-1,195142725	roślin	uprawa	hodowla	hodowli	zwierząt	drzew	uprawy	bydła	polsce	owocowych
-1,466890492	przewodnik	podróży	wrażenia	podróż	polska	przygody	droga	przez	drogi	wycieczki
-1,548822948	św	katechizm	kościoła	ks	jezusa	chrystusa	czyli	świętej	kazania	kościół
1,021913888	poezje	wiersze	wybór	poezye	dzieła	poezji	poemat	mickiewicza	poezyj	wierszy
-1,051497403	człowiek	romans	świat	człowieka	dom	pan	pamiętnik	jan	antek	kruzoe
4,922654166	dzieci	baśń	bajki	fantastyczna	opowiadanie	opowiadania	młodzieży	bajka	przygody	powieść
-3,079459393	wygłoszony	dnia	dn	referat	odczyt	roku	dniu	przemówienie	warszawie	wygłoszone
-1,8148266	de	la	badania	nad	les	leczenie	et	sur	the	choroby

Topics recognized by gender (DE 1911–1945)

-1,481710512	stadt	europa	europ	preußischen	polen	de	geschichte	entwicklung	jahre	preuße
1,575133717	fall	beitrag	fälle	frage	univ	behandlung	berücksichtigung	kenntnis	beim	besonderer
1,810697433	kriminalroma	roman	novell	liebe	novelle	leben	lieb	tod	geheimnis	de
-1,682916656	beitr	geschichte	deutschen	dt	berücks	geh	berlin	bes	auftr	jahre
-2,895687463	recht	rechtliche	bürgerlichen	gesetz	berücks	stellung	rechts	rechtsprechung	gesetze	bes
1,033679163	untersuchungen	experimentelle	studien	untersuchung	versuche	frage	vergleichende	beim	beiträge	wirkung
-1,549321482	wirtschaft	wirtschaftliche	problem	wirtschaftlichen	soziale	bedeutung	sozialen	entwicklung	probleme	organisation
1,111970507	ueber	fall	univ	behandlung	wirkung	einfluß	einfluss	verhalten	vorkommen	beim
-2,428762372	werk	gesammelte	sämtliche	ausgewählte	dramatische	werke	zeichnunge	arbeitsrech	arbeitsberich	ausgewählter
-1,461691833	landwirtschaftlichen	landwirtschaft	landwirtschaftliche	landwirtschaf	landwirtschaftl	obst	gemüse	prakt	landwirt	anbau
-1,820070231	eisen	erz	stahl	untersuchungen	mitteilg	aluminium	beton	bestimmung	kupfer	kenntnis
-1,277485881	abenteuer	berge	land	führer	welt	geologie	natur	geologische	erlebnisse	wild
-2,637153988	elektrische	elektrischen	theorie	siemens	untersuchungen	messung	berechnung	elektrischer	technik	bestimmung
1,801147832	gedicht	gedichte	dichtung	dichtunge	dichter	dichtun	neue	mundar	ausgewählte	gedich
1,371034099	schwank	ak	aufz	schwei	spie	kleine	guil	lustige	american	spiel
2,112753077	brie	bries	krie	effi	buch	arie	baugeschichte	pariser	offener	familie
-1,228690158	geschichte	alten	geschicht	zeit	mittelalter	alte	gegenwar	jahrhundert	bilder	erinnerungen
-1,297429123	schweizerischen	schweiz	schweizer	19	jahrhundert	österreich	wiener	schweizerische	österreichischen	18
1,106698839	heima	familie	heimat	geschichte	stadt	hannove	rhein	stimme	goeth	hessen
1,019963283	akte	aufz	schausp	vier	drama	schauspiel	bilder	fünf	akten	komödie
-2,018134512	193	191	192	geh	194	jahre	juni	rede	juli	mai
-1,085715347	akte	drei	aufz	akten	komödie	operette	lustsp	oper	schausp	lustspiel
3,169254351	kinder	kinde	jugend	kind	kindern	mutter	kindes	eltern	kindesalte	jugen
1,085988379	buch	briefe	roman	text	schrifte	de	verf	werke	brief	bilder
3,494789484	roma	kriminal	liebe	orig	detektiv	frau	glück	mann	abenteuer	heiterer
-1,653381023	deutschen	deutsche	deutschland	deutscher	deutsches	deutschlands	deutschlan	deutsch	berli	geschichte
-1,175775916	psychologie	philosophie	gedanken	seele	philosophi	leben	kunst	geist	psychologi	wesen
-1,14298026	friedrich	heinrich	wilhelm	johann	leben	goethe	goethes	ludwig	franz	werk
1,022366203	behandlung	therapie	klinik	univ	beitrag	klinische	medizin	krankheiten	med	chirurgischen
1,000567047	provinz	büttner	rechenbuch	geschichte	bismarck	bd	abc	bayer	beitr	ballade
-1,776781504	lehrbuch	unterricht	schule	höhere	lehranstalte	lehr	schulen	handbuch	selbstunterrich	lehrer
3,2934648	erzähl	erzählunge	erzählg	erzählun	geschichten	märchen	geschichte	erzählungen	märche	erzählung
1,204499693	korresponden	postkart	brief	märche	rechenbuc	lebe	kleine	fragebogen	novelle	beantwortung
1,082102616	erz	ak	de	roman	geschichte	beitr	aufz	kenntnis	leben	beiträge
-3,04188361	haftung	buchführung	berücks	beim	recht	schutz	versicherung	bes	bgb	aktiengesellschaf
-1,654933731	anleitung	nebst	leitfaden	prakt	anleitg	berechnung	praktische	ratgeber	sowie	aufgaben
1,032026431	wirkung	kenntnis	einfluß	einwirkung	beiträge	beitrag	beeinflussung	vitamin	synthese	versuche
-1,775296171	krieg	1914	weltkrieg	kriege	weltkriege	kriegs	nr	regiment	krieges	gegenwar
-5,691527297	steuer	einkommensteuer	steuern	besteuerung	steuerrech	einkommen	steuerrecht	umsatzsteuer	neuen	nebst
-1,481578501	kirche	jesu	gottes	gott	predigt	heiligen	jesus	heilige	christus	kirch

Topics recognized by gender (PL 1946–1980)

-3,968187752	transportu	pojazdów	samochodów	ruchu	samochodowych	samochodowego	transport	transporcie	obsługa	kolejowego
2,636117882	języka	język	słownik	polski	gramatyka	rosyjski	rosyjskiego	polsko	pisowni	gramatyki
1,080109401	wieku	dziejów	historia	xvi	xviii	dzieje	xvii	polski	polsce	historii
-2,871309563	prawo	prawa	prawie	prawne	przepisy	karnym	karnego	zagadnienia	cywilnego	postępowanie
1,037644675	studentów	podręcznik	szkół	metodyczny	skrypt	przewodnik	wyższych	wydziału	studiów	podręcznika
1,756377275	wychowania	nauczania	młodzieży	nauczycieli	szkole	wychowanie	szkół	kształcenia	uczniów	metodyka
-2,564603742	technologia	metali	materiałoznawstwo	obróbka	stali	obróbki	materiałów	drewna	towaroznawstwo	wyrobów
-1,745350178	zadań	zbiór	ćwiczenia	fizyki	teorii	chemii	matematyki	zadania	analizy	statystyki
1,550653064	wystawa	katalog	wystawy	malarstwo	rysunek	muzeum	sztuka	sztuki	grafika	portret
1,302480423	choroby	chorób	leczenie	badania	leczenia	układu	chirurgia	chorobach	krwi	płuc
-2,769571015	1939	1945	wojny	światowej	1944	wojna	walki	wspomnienia	armii	wojnie
-1,193520643	pracy	organizacja	pracowników	organizacji	higieny	higiena	bezpieczeństwo	płac	bezpieczeństwa	spółdzielni
-2,386226476	przedsiębiorstw	organizacja	ekonomika	rachunkowość	przedsiębiorstwie	planowanie	handlu	analiza	organizacji	kosztów
-1,091086694	przewodnik	okolice	turystyczny	podróż	droga	informator	podróże	podróży	wędrówki	wyprawa
1,342966353	słońce	niebo	noce	lato	noc	dnie	światło	ziemi	ziemia	słońca
-1,121894161	rozwoju	problemy	produkcji	gospodarki	przemysłu	krajach	polsce	krajów	latach	ekonomiczne
-3,128855851	węgla	wody	gospodarka	kopalniach	ścieków	wód	eksploatacja	kopalń	kopalni	gazu
1,390371837	opowiadania	inne	bajki	opowieści	dzieła	baśnie	bajka	opowiadanie	przygody	opowieść
-1,672070192	św	pisma	modlitewnik	katechizm	świętego	rozważania	komunii	świętej	kościoła	spowiedzi
-1,025896586	of	badania	the	nad	wpływ	własności	and	metody	związków	oraz
-2,654192351	chemia	ogólna	mechanika	geometria	fizyka	matematyka	organiczna	anatomia	analityczna	arytmetyka
1,919074862	maria	miłość	pani	ziemia	tajemnica	wielka	złota	miłości	polska	księga
1,175762123	człowiek	don	pamiętnik	poradnik	człowieka	cichy	dom	czas	świat	elementarz
1,283166865	roślin	uprawa	rośliny	drzew	lasu	gleb	leśnych	warzyw	uprawy	ochrona
4,694637169	dzieci	dziecka	dziecko	chłopiec	przygody	niemowląt	dzieciństwo	dzieciństwa	wieku	przedszkolu
2,569681922	kl	liceum	ogólnokształcącego	klasy	szkoły	podręcznik	podstawowej	język	technikum	11
1,103737552	powieść	historyczna	opowieść	młodzieży	powieści	tomach	czasów	wieku	lalka	współczesna
1,087714784	jan	opowieść	pan	józef	życie	król	stanisław	mickiewicz	aleksander	karol
1,299002624	zwierząt	bydła	hodowla	chów	hodowli	trzody	chlewnej	drobiu	gospodarskich	zwierzęta
-1,916190549	zarys	filozofii	teorii	zagadnienia	psychologii	socjologii	teoria	psychologia	filozofia	ekonomia
2,373645456	bibliografia	biblioteki	literatury	wybór	pism	literatura	polskiej	bibliografii	katalog	polska
-4,689216939	elektryczne	elektrycznych	elektrotechniki	urządzenia	podstawy	układy	elektrotechnika	układów	sieci	elektroenergetycznych
-4,139234587	budownictwie	budowlanych	budowlane	konstrukcje	konstrukcji	budownictwa	budownictwo	budowli	robót	budynków
1,291759643	nowele	listy	trzy	wspomnienia	dramaty	szkice	pamiętniki	cztery	rozmówki	polskie
1,220897296	ludzie	chłopcy	chłopi	polscy	bracia	polacy	powieść	bezdomni	żołnierze	czterej
-1,264661775	wiersze	poezje	wybór	wybrane	poezji	wierszy	utwory	poematy	szkice	poemat
-1,161345339	nad	wyspa	morza	morze	morskie	potop	rzeki	niemnem	rzek	morskich
-3,623971528	maszyn	maszyny	produkcji	urządzeń	urządzenia	części	procesów	programowanie	automatyki	podstawy
-1,792604474	latach	ludowej	partii	referat	1945	1939	1918	polsce	wygłoszony	pzpr
1,091958491	aktach	komedia	sztuka	teatr	dramat	teatru	odsłonach	tragedia	akcie	dramatyczne

Topics recognized by gender (DE 1946–1980)

1,020447853	kriminalroma	kriminalroman	mord	kriminal	tod	mörder	kommissar	maigret	the	detektiv
-2,214872042	berechnung	verfahren	anwendung	of	bestimmung	theorie	tab	methoden	methode	analyse
-1,404018271	friedrich	beitr	geschichte	heinrich	wilhelm	johann	leben	werk	goethe	goethes
-2,865515749	untersuchungen	untersuchung	messung	of	elektronen	bestimmung	elektrische	anwendung	elektronik	elektrischen
-1,004052658	psychologie	problem	psychologi	psycholog	soziale	sozialen	psychologische	denken	de	theorie
-2,339843253	untersuchungen	eisen	stahl	untersuchung	beitrag	eigenschaften	nickel	kupfer	metall	herstellung
2,050455145	roma	kriminal	frauenroma	wildwest	fiction	western	science	wildwestroma	wild	frauen
1,710622995	erzählunge	erzählun	geschichten	geschichte	erzähl	märchen	erzählung	erzählungen	erzählg	märche
-2,26978899	beitrag	einfluss	einfluß	berechnung	beim	untersuchungen	bestimmung	untersuchung	verhalten	messung
-1,336074875	synthese	darstellung	reaktionen	untersuchungen	eigenschaften	chemie	verbindungen	kenntnis	synthesen	versuche
-1,257677397	drei	akte	akten	komödie	oper	schauspiel	aufzüge	dritte	aufz	lustspiel
1,39533447	untersuchungen	experimentelle	tierexperimentelle	vergleichende	untersuchung	wirkung	frage	beim	studien	nachweis
1,134313817	gedicht	gedichte	dichtung	dichter	dichtunge	dichtun	poe	mundar	lyrik	poesie
1,871221763	no	leben	wer	roman	geheimnis	mädchen	de	mensch	mann	kleine
-1,311301204	abenteuer	land	welt	natur	reise	weste	wald	de	roman	führer
2,066664828	behandlung	untersuchungen	patienten	ergebnisse	klinik	berücksichtigung	besonderer	klinische	beim	chirurgischen
1,212212981	roma	roman	römischen	römische	rom	italien	röm	italie	de	römer
-3,238868274	recht	gesetz	rechtsprechung	berücks	rechts	kommentar	erl	kommenta	abs	beitr
1,166996029	abc	bach	sebastian	johann	the	bibe	de	geschichte	beitr	bd
-2,752387904	besteuerung	buchführung	grundlagen	leitfaden	betriebswirtschaftliche	steuerliche	probleme	ratgeber	betriebliche	aufgaben
1,125759948	de	novell	the	leben	abenteuer	könig	peter	komödi	insel	robinson
1,31475421	wel	kleine	schwei	the	de	zei	märche	abenteuer	auswah	märchen
1,131019532	leben	de	la	letzte	grosse	lebe	tag	glück	the	sieben
1,782460175	wirkung	einfluss	einfluß	beeinflussung	verhalten	untersuchungen	ratte	beim	veränderungen	serum
2,770419345	kinder	kinde	kindern	kind	jugend	kindesalte	eltern	kindes	jugendliche	jugendlichen
-1,016788055	korresponden	brief	postkart	kleine	schrifte	märche	drame	faus	heid	kapita
-1,817598855	kirche	gott	gottes	jesus	jesu	christ	christus	heiligen	kirch	heilige
-1,126181947	buch	werke	schrifte	literatur	gesammelte	roman	text	briefe	selbstzeugnissen	of
1,446883791	therapie	klinik	behandlung	medizin	klinische	therapi	medizi	arzt	patienten	medizinischen
-1,620018824	marx	entwicklung	sozialismus	revolution	19	sozialistischen	politik	geschichte	the	jahrhundert
-1,327550139	195	196	197	194	jahre	jahren	193	191	1945	192
-1,36503095	geschichte	alten	geschicht	jahrhundert	gegenwar	mittelalter	alte	jahrhunder	zeit	kunst
-1,535127273	deutschen	deutsche	deutschland	bundesrepublik	deutschlan	german	deutsch	berlin	demokratischen	deutschlands
-2,316402172	entwicklung	planung	probleme	landwirtschaft	sozialistischen	industrie	organisation	landwirtschaftlichen	beispiel	wirtschaft
-3,587560763	philosophie	physik	mathematik	lehrbuch	einführung	philosophi	physi	mathemati	theorie	chemi
3,338253484	liebe	lieb	herz	liebesroma	roman	lieben	herzen	ehe	eh	liebesgeschichte
3,041414887	roma	liebe	heiterer	herz	frau	glück	nacht	mann	utop	mädchen
-1,115418902	schule	unterricht	lehrbuch	erziehung	ausbildung	einführung	lehr	handbuch	anleitung	praxis

Topics recognized by gender (PL 1981–1999)

-1,209280916	habilitacyjna	badania	chemia	wpływ	właściwości	rozprawa	oraz	środowiska	wybranych	praca
-1,893586042	filozofii	filozofia	teorii	zarys	studium	teoria	myśli	człowieka	psychologia	socjologii
1,123457342	katalog	wystawy	wystawa	muzeum	malarstwo	sztuki	galeria	sztuka	fotografii	portret
-3,033781706	pawła	jana	ii	św	świętego	kościoła	ojca	kościół	ks	święty
-1,406049889	prawo	komentarz	prawa	podatek	ustawa	ustawy	orzecznictwo	kodeks	postępowanie	komentarzem
-1,615544267	pan	tadeusz	józef	stanisław	jan	generał	król	ostatni	książę	szlachecka
-1,065781251	of	the	and	de	poland	for	polish	et	la	with
3,816501228	dzieci	dziecko	dziecka	rodziców	najmłodszych	dzieciom	wieku	dziećmi	zabawy	przedszkolnym
1,680562497	młodzieży	nauczycieli	wychowania	wychowanie	szkole	edukacji	nauczania	szkół	dzieci	edukacja
-4,819772028	kw	pojazdów	naprawa	km	samochodów	samoloty	eksploatacja	samochodowych	jazdy	samolot
3,292037022	miłość	miłości	kobiety	kobieta	pani	małżeństwo	kochać	czyli	panna	żona
4,249642682	klasy	kl	szkoły	podstawowej	podręcznik	zeszyt	ćwiczeń	liceum	matematyka	klasa
-1,016320577	wiersze	poezje	wierszy	wybór	poezji	wybrane	poezja	inne	utwory	poemat
-1,22979353	rozwoju	polityka	polsce	gospodarki	gospodarczej	krajach	krajów	europejskiej	latach	polityki
1,327757354	światła	niebo	światło	słońce	słońca	cień	nieba	cienie	gwiazdy	ciemności
1,376262194	zwierzęta	zwierząt	pies	ptaki	przygody	psy	kot	kota	psów	chów
-3,7734873	konstrukcji	maszyn	podstawy	technologia	materiałów	projektowanie	projektowania	budowlanych	konstrukcje	obróbki
1,407800616	ziemia	ania	święta	polska	mała	tajemnica	czarna	zemsta	gra	wielka
-1,03187887	opowiadania	powieść	bajki	baśnie	inne	opowieści	opowieść	baśń	bajka	przygody
-1,482553961	historia	dziejów	wieku	dzieje	xviii	xvi	zarys	historii	miasta	polski
1,33564785	pracy	pracowników	poradnik	ubezpieczenia	osób	ubezpieczeń	społecznych	społeczne	zatrudnienia	organizacji
-4,394126907	laboratorium	układów	układy	elektrycznych	elektryczne	sieci	urządzenia	podstawy	metody	sterowania
1,162260551	leczenie	habilitacyjna	rozprawa	choroby	chorób	leczenia	poradnik	chorych	układu	badania
-1,365534482	ludzie	polscy	polacy	krzyżacy	żydzi	chłopi	bracia	przyjaciele	bogowie	chłopcy
-1,065609878	studentów	podręcznik	skrypt	szkół	technikum	studiów	wyższych	zasadniczej	wydziału	ćwiczeń
-1,239771855	ogniem	mieczem	czarny	czerwony	pan	człowiek	conan	powrót	ostatni	taniec
-1,217022764	przedsiębiorstw	zarządzanie	rachunkowość	zarządzania	organizacja	przedsiębiorstwa	przedsiębiorstwie	ekonomika	rachunkowości	zasady
-1,821709843	windows	programowanie	pl	wersja	pc	programowania	microsoft	języku	basic	dos
1,967206824	słownik	języka	polsko	język	polski	polskiego	angielski	gramatyka	angielsko	gramatyki
1,116190242	życie	dom	człowiek	życia	świat	człowieka	żyć	czas	domu	miasto
1,494164808	roślin	rośliny	ogród	kwiaty	drzew	uprawa	ochrona	drzewa	leśnych	warzyw
-1,988088304	rozważania	boga	jezusa	chrystusa	modlitwy	modlitwa	jezus	życie	kazania	wiary
1,520713675	bibliografia	literatura	literatury	polskiej	lat	wybór	literaturze	biblioteki	dziennik	polska
-1,008429958	dzieła	listy	wszystkie	trzy	cztery	szkice	pisma	aforyzmy	tajemnice	fraszki
-1,519999702	przewodnik	turystyczny	okolice	droga	podróż	tatry	góry	podróże	szlakiem	beskid
1,368261983	wspomnienia	pamiętnik	myśli	sen	noc	pustyni	dni	pamięci	puszczy	czas
-1,70375812	zadań	ćwiczenia	zbiór	fizyki	zadania	matematyka	chemii	teorii	fizyka	podstawy
-1,263355088	nad	wyspa	potop	rzeki	morza	rzeka	morze	morskie	wody	dolina
3,114399717	kuchnia	potrawy	kuchni	przepisów	przepisy	ciasta	mięsa	żywności	smak	dieta
-3,135655842	1939	1945	latach	1944	1918	wojny	światowej	armii	polski	polska

Topics recognized by gender (DE 1981–1999)

-1,911562712	recht	berücksichtigung	rechtsprechung	rechtliche	besonderer	gesetz	kommentar	abs	europäischen	internationalen
1,796925708	patienten	ergebnisse	therapie	klinische	retrospektive	diagnostik	behandlung	berücksichtigung	vergleich	untersuchungen
-1,480492692	friedrich	heinrich	geschichte	johann	wilhelm	werk	leben	goethes	goethe	hans
1,276413991	roma	roman	römischen	rom	römische	italie	italien	röm	italienische	römer
-1,067617315	lernen	schule	einführung	ausbildung	unterricht	training	handbuch	schüler	praxis	lehrer
-2,307659816	rechnungswesen	ratgeber	geld	buchführung	besteuerung	steuern	einkommensteuer	neuen	unternehmen	analyse
-2,130949088	untersuchungen	untersuchung	laser	mittels	messung	spektroskopie	optische	of	bestimmung	eigenschaften
-1,136363744	de	edgar	the	leben	einf	roman	insel	abenteuer	aphorisme	abenteue
-1,107967489	brie	bach	buc	basic	abc	geschichte	bgb	bibe	sebastian	the
1,041506219	synthese	charakterisierung	reaktionen	eigenschaften	untersuchungen	synthesen	darstellung	alpha	struktur	beta
1,990963571	charakterisierung	untersuchungen	nachweis	dna	expression	antikörper	patienten	humanen	zellen	vitro
1,901879273	einfluss	wirkung	einfluß	ratte	ratt	bedeutung	regulation	beeinflussung	veränderungen	beta
-2,812320313	beitrag	berechnung	untersuchungen	einfluss	untersuchung	beim	experimentelle	simulation	entwicklung	bestimmung
-1,240405009	deutschland	deutschen	bundesrepublik	deutsche	deutschlan	deutsch	berlin	deutscher	german	berli
-1,216207379	198	199	194	jahre	193	191	197	jahren	200	195
-3,242496643	computer	windows	pc	software	grundlagen	internet	einführung	entwicklung	programmierung	dos
2,271509247	roma	liebe	fiction	science	mann	nacht	heiterer	frau	schatten	haus
1,448628288	therapie	patienten	medizin	therapi	behandlung	klinik	patiente	diagnostik	klinische	praxis
-1,37450836	werk	gesammelte	sämtliche	gesamtausgab	werkausgab	ausgewählte	brief	schrifte	mainzelmännche	freund
-1,496278844	of	for	bestimmung	the	verfahren	analyse	bewertung	and	beurteilung	tabelle
-1,946089299	management	planung	gestaltung	unternehmen	organisation	marketing	entwicklung	unternehme	controlling	beispiel
-1,677771886	untersuchungen	untersuchung	eigenschaften	herstellung	struktur	metall	charakterisierung	einfluss	beitrag	eisen
1,256368282	wel	kleine	schwei	zei	knickerbocker	band	grosse	seel	de	abenteuer
1,384862525	erzählunge	erzählun	geschichten	geschichte	märchen	erzählungen	kleine	kurzgeschichte	erzähl	wahre
-1,480035545	ökologische	umwelt	ökologie	beispiel	untersuchungen	umweltschutz	natur	berücksichtigung	ökologischen	entwicklung
-1,005944389	soziale	politische	sozialen	politik	politischen	theorie	sozialer	gesellschaft	and	entwicklung
1,18563037	kriminalroma	roman	leben	tod	geheimnis	kriminalroman	mord	mann	geschichte	traum
-1,939276549	europäischen	wirtschaft	the	ökonomische	analyse	europäische	and	beispiel	of	europa
-1,089829956	werke	buch	schrifte	gesammelte	roman	texte	literatur	text	handbuch	briefe
1,353952429	gedicht	gedichte	lyrik	dichtung	mundar	poesie	pros	dichter	prosa	lyri
-2,817443523	of	for	the	and	with	linear	berechnung	finite	system	equation
1,327369834	reise	reiseführer	leben	welt	neue	roman	frau	lebe	reisen	de
2,801813878	kinder	kindern	kinde	kind	eltern	jugendlichen	jugendliche	jugend	kindheit	kindes
1,44142143	leben	gesund	gesundheit	kraft	glück	wege	lebe	seele	körper	fit
-1,223069334	geschichte	geschicht	alten	gegenwar	zeit	erinnerungen	leben	mittelalter	alte	historische
1,664535765	untersuchungen	untersuchung	experimentelle	tierexperimentelle	vitro	vergleichende	beim	berücksichtigung	wirkung	studie
-1,17812968	19	jahrhundert	20	jahrhunder	18	jahre	geschichte	österreich	jahrhunderts	17
-1,740339052	gott	kirche	gottes	jesus	jesu	theologie	glaube	glauben	christus	heiligen
1,189129786	rezepte	wanderungen	entdecken	tips	sehenswertes	erleben	schönsten	ausflüge	essen	pflege

TOPICS RECOGNIZED BY GENDER (PL 2000–2021)

1,582191931	ogród	roślin	rośliny	kwiaty	drzewa	ogrodzie	las	uprawa	drzew	atlas
-1,929073141	dziejów	historia	dzieje	wieku	miasta	historii	lat	latach	xix	xx
-2,362546456	rozważania	jezusa	boga	św	modlitwy	jezus	droga	kazania	chrystusa	bożej
-1,105883422	nad	wyspa	morze	wyspy	rzeka	rzeki	morza	wody	woda	morskie
-2,220897905	filozofii	filozofia	teorii	studium	teoria	filozoficzne	psychologia	zarys	człowieka	podstawy
1,190396743	trzy	cztery	listy	sekrety	tajemnice	myśli	fraszki	anioły	marzenia	pory
-1,027077529	rozwoju	polsce	rynku	europejskiej	unii	przedsiębiorstw	uwarunkowania	gospodarki	rozwój	polityki
-1,436830169	prawo	komentarz	prawa	kodeks	prawie	postępowanie	ustawa	postępowania	postępowaniu	sprawach
-1,016502857	literaturze	literatura	literatury	twórczości	szkice	bibliografia	powieść	polskiej	wieku	wybór
-3,885900522	pl	programowanie	windows	programowania	excel	microsoft	tworzenie	danych	aplikacji	sieci
3,58654561	kobieta	dziewczyna	pani	księżniczka	kobiety	córka	królowa	matka	maria	ania
6,312103852	klasa	klasy	gimnazjum	zeszyt	podstawowej	podręcznik	szkoły	ćwiczeń	szkoła	matematyka
-1,860764141	biografia	król	józef	książę	pan	jan	stanisław	człowiek	tadeusz	życie
1,243087265	tajemnica	księga	gra	zemsta	ucieczka	ostatnia	wielka	krew	złota	ziemia
1,399170433	noce	czarne	dnie	cień	białe	światło	światła	słońca	noc	nocy
1,073534633	chorych	ocena	choroby	leczenia	leczenie	układu	pacjentów	oraz	wybranych	badania
-1,439361812	właściwości	wpływ	wybranych	środowiska	of	oraz	materiałów	węgla	chemia	metody
2,309271564	miłość	miłości	małżeństwo	ślub	kobiety	małżeństwa	romans	kochać	seks	kocham
1,917230792	edukacji	młodzieży	nauczycieli	edukacja	dzieci	szkół	szkole	wychowania	uczniów	szkoły
1,696891324	zwierzęta	kot	zwierząt	ptaki	pies	psy	koty	kotek	psa	kota
-1,216644448	wspomnienia	pamiętnik	pamięci	zapiski	wspomnień	lat	pamięć	pamiętniki	dzienniki	zapomnienia
-1,131779037	ostatni	człowiek	pan	powrót	czarny	dom	nowy	dziennik	czas	anioł
1,340122006	opowiadania	opowieści	bajki	inne	baśnie	opowieść	przygody	historie	legendy	bajka
-1,440903955	ludzie	polscy	polacy	żydzi	bracia	święci	romek	tytus	tomek	którzy
1,010146659	wiersze	wierszy	poezji	poezje	wybór	wybrane	poezja	zebrane	wierszyki	inne
-2,980778318	św	ii	pawła	jana	kościoła	świętego	ks	parafii	święty	kościół
1,2767375	liceum	podręcznik	zakres	podstawowy	rozszerzony	technikum	zadań	matematyka	studentów	ogólnokształcącego
-4,672143542	1939	1945	wojny	latach	światowej	1918	1944	wojna	ii	armii
2,938816211	kuchnia	przepisów	przepisy	dieta	kuchni	potrawy	dania	smak	ciasta	kulinarne
1,071212199	poradnik	trening	zdrowie	sztuka	terapia	ćwiczenia	czyli	zdrowia	ciała	życia
-4,826059689	podstawy	maszyn	modelowanie	metody	układów	konstrukcji	sieci	wybrane	projektowanie	urządzeń
-1,117414053	malarstwo	sztuki	fotografii	wystawa	muzeum	sztuka	galeria	fotografia	wystawy	katalog
1,678498055	słownik	polsko	języka	język	polski	angielski	rozmówki	angielsko	gramatyka	leksykon
-1,463635609	the	of	and	from	de	polish	poland	for	la	industry
1,45579891	życie	życia	dom	dzień	serce	miasto	świat	czas	za	myśli
-1,899222252	polityka	latach	europejskiej	unii	wobec	polsce	1989	polska	bezpieczeństwa	rzeczypospolitej
-1,32828219	przewodnik	podróż	turystyczny	droga	podróże	podróży	przez	wyprawa	śladami	wędrówki
3,634660404	dzieci	dziecko	dziecka	rodziców	zabawy	metodyczny	lat	mały	książeczka	przewodnik
-1,189165974	zarządzanie	zarządzania	pracy	organizacji	marketing	zarządzaniu	pracowników	poradnik	organizacja	przedsiębiorstwie
1,233885389	vat	rachunkowości	rachunkowość	podatek	podatku	komentarz	zasady	podatki	dochodowy	podatkowe

Topics recognized by gender (DE 2000–2021)

-2,325477237	charakterisierung	mittels	untersuchungen	entwicklung	untersuchung	eigenschaften	herstellung	simulation	of	modellierung
1,005013932	gedicht	gedichte	lyrik	poesie	gedanken	dichter	lyri	poesi	lyrische	dichtung
-1,396089478	philosophie	kunst	philosophische	philosophi	theorie	denken	wissenschaft	ästhetik	kritik	mathematik
1,553940489	wel	kleine	abenteuer	leben	piraten	wal	wolf	seewölfe	kleinen	weltmeere
-1,069621013	buch	roman	rhodan	perry	band	literatur	heftroman	zyklus	teil	ban
-1,450030596	management	unternehmen	unternehme	business	führung	marketing	erfolgreich	grundlagen	controlling	managemen
1,368389732	warum	leben	mehr	no	immer	allein	wer	welt	ganz	mann
-1,479957159	krieg	the	politik	politische	welt	revolution	demokratie	of	politischen	jahrhundert
-2,108925031	analyse	ifrs	bilanzierung	unternehmen	gmbh	private	berücksichtigung	steuern	geld	besteuerung
-1,027026228	atlan	erotische	engel	band	heftroman	zyklus	teil	leben	abenteuer	the
1,392815153	untersuchungen	synthese	untersuchung	charakterisierung	vitro	analyse	mittels	entwicklung	nachweis	experimentelle
-1,392094579	europäischen	eu	europa	europäische	union	europ	unio	europäischer	europas	european
2,029549512	roma	historischer	erotischer	liebe	nacht	sommer	leben	schatten	haus	glück
-1,624153834	internet	online	digitale	marketing	digitalen	medien	web	social	media	kommunikation
-1,084619143	werk	gesammelte	sämtliche	brief	kleine	gesamtausgab	letzte	lebe	schrifte	nacht
-1,078203858	kriminalroma	thrille	tod	mord	krimi	krim	kriminalroman	fall	thriller	tote
1,453262592	therapie	patienten	pflege	medizin	psychotherapie	homöopathie	gesundheit	behandlung	therapi	praxi
-1,731744697	200	201	194	199	191	jahre	20	193	202	198
-1,580039494	analyse	entwicklung	bewertung	methoden	of	beispiel	verfahren	modellierung	empirische	untersuchung
-1,445423206	geschichte	alten	geschicht	erinnerungen	jahrhundert	zeit	historische	mittelalter	antike	alte
1,609842406	schule	lernen	bildung	unterricht	klasse	schüler	learning	lehrer	grundschule	förderung
-1,22869521	band	brie	ban	folge	john	sinclair	bgb	buc	teil	abc
-1,087390889	reise	reiseführer	reisen	abenteuer	touren	tipps	insider	schönsten	wanderungen	welt
-1,04593805	roma	römischen	römische	roman	rom	romana	römer	italien	italienische	italie
-1,011961141	band	fal	neue	drei	thrille	fünf	fall	the	handbuc	aufzüge
2,724475543	kinder	kindern	eltern	kind	jugendlichen	kinde	jugendliche	kindheit	kindes	jugend
-1,215044272	deutschland	deutschen	deutsche	deutschlan	deutsch	deutscher	berlin	bundesrepublik	deutschlands	berli
1,406679874	soziale	sozialen	einfluss	arbeit	sozialer	auswirkungen	bedeutung	menschen	of	and
2,018114265	einfluss	charakterisierung	expression	rolle	humanen	regulation	zellen	bedeutung	protein	untersuchungen
1,757624942	patienten	therapie	retrospektive	klinische	vergleich	ergebnisse	mittels	untersuchung	diagnostik	untersuchungen
-1,814306544	gott	kirche	gottes	jesus	religion	glaube	jesu	glauben	bibel	theologie
-1,497947679	beispiel	entwicklung	analyse	and	of	stadt	nachhaltigkeit	ökologische	nachhaltige	umwelt
1,326515585	geheimnis	magische	band	schatten	magie	fluch	teufel	dunkle	legende	geheimnisvolle
1,412093599	leben	lebe	glück	mehr	kraft	glücklich	seele	yoga	erfolg	lebens
-1,63671683	friedrich	rezension	heinrich	geschichte	johann	wilhelm	band	hans	hermann	wolfgang
-1,531007406	handbuch	grundlagen	einführung	tipps	klasse	schritt	übungen	praxis	mathematik	windows
3,141938228	liebe	lieb	love	herz	liebesroma	lieben	liebesroman	verliebt	roman	erotische
-1,784965259	recht	berücksichtigung	besonderer	rechtliche	strafrecht	abs	internationalen	untersuchung	rechtsprechung	gesetz
2,161626135	rezepte	kochbuch	kochen	ernährung	essen	rezept	küche	leckere	gesund	besten
1,27256363	geschichten	erzählunge	geschichte	erzählun	kurzgeschichte	märchen	kurzgeschichten	erzählungen	kleine	wahre

GENDER SPECIFIC TOPICS / KEY WORDS

Period	Women		Men	
	Topic	C	Topic	C
1801-1910	woman	5.7	thesis defence	8.8
	youth	4.8	society report	5.6
	novel	2.8	academic	4.3
1911-1945	youth	4.9	judiciary	3.9
	romance	3.1	lecture	3.1
	novel	2.1	academic	2.2
1946-1980	child	4.7	electrics	4.7
	language	2.6	construction	4.1
	school	2.6	transport	4.0
1981-1999	school	4.2	transport	4.8
	child	3.8	electrics	4.4
	romance	3.3	construction	3.8
2000-2021	school	6.3	machine	4.8
	children	3.6	war	4.7
	woman	3.6	software	3.9

Table 1: The most gender specific topics detected in Polish titles in selected time periods. C denotes the importance ratio, calculated as the ratio of books by women to men (or vice versa) with a given theme (topic) normalized to the ratio of books for each gender.

Period	Women		Men	
	Topic	Co	Topic	Co
1801-1910	novel	4.6	Germany	3.3
	Rome	2.5	report, speech	3.1
	letter	1.9	religion	3.0
1911-1945	novel	3.5	tax	5.7
	story	3.3	economy	3.0
	child	3.2	judicary	2.9
1946-1980	romance	3.3	science	3.6
	woman	3.1	judicary	3.2
	child	2.8	electrics	2.8
1981-1999	child	2.8	software	3.2
	romance	2.3	mathematics	2.8
	medicine	2.0	applied science	2.8
2000-2021	romance	3.1	investigation	2.3
	child	2.8	finances	2.1
	cooking	2.1	religion	1.8

Table 2: The most gender specific topics detected in German titles.

GENDER SPECIFIC TOPICS – SUMMARY (1801–1999)

1801-1910		1911-1945		1946-1980		1981-1999		2000-2021	
1,065946739	-2,977560595	3,091923177	-1,481710512	-3,968187752	1,020447853	-1,209280916	-1,911562712	1,582191931	-2,325477237
1,208173195	2,565262911	-3,860609217	1,575133717	2,636117882	-2,214872042	-1,893586042	1,796925708	-1,929073141	1,005013932
-4,199277114	-2,663346713	-1,645343911	1,810697433	1,080109401	-1,404018271	1,123457342	-1,480492692	-2,362546456	-1,396089478
-1,402691415	-3,094389154	1,385231303	-1,682916656	-2,871309563	-2,865515749	-3,033781706	1,276413991	-1,105883422	1,553940489
-1,45697841	-1,649554065	-1,947704585	-2,895687463	1,037644675	-1,004052658	-1,406049889	-1,067617315	-2,220897905	-1,069621013
-1,938827795	1,853779698	-1,751744126	1,033679163	1,756377275	-2,339843253	-1,615544267	-2,307659816	1,190396743	-1,450030596
-8,784333818	4,668728134	-1,017421225	-1,549321482	-2,564603742	2,050455145	-1,065781251	-2,130949088	-1,027077529	1,368389732
-2,290138004	1,213167879	-2,252870263	1,111970507	-1,745350178	1,710622995	3,816501228	-1,136363744	-1,436830169	-1,479957159
1,347564629	-3,340130264	-1,324906832	-2,428762372	1,550653064	-2,26978899	1,680562497	-1,107967489	-1,016502857	-2,108925031
-1,079334881	1,457417486	1,468293725	-1,461691833	1,302480423	-1,336074875	-4,819772028	1,041506219	-3,885900522	-1,027026228
-1,576453229	-1,809608439	2,08849997	-1,820070231	-2,769571015	-1,257677397	3,292037022	1,990963571	3,58654561	1,392815153
2,801818191	-1,812945482	1,155403373	-1,277485881	-1,193520643	1,39533447	4,249642682	1,901879273	6,312103852	-1,392094579
1,1694024	-1,337851875	-1,195142725	-2,637153988	-2,386226476	1,134313817	-1,016320577	-2,812320313	-1,860764141	2,029549512
-3,90054051	-1,625769333	-1,466890492	1,801147832	-1,091086694	1,871221763	-1,22979353	-1,240405009	1,243087265	-1,624153834
5,663818634	-1,65391821	-1,548822948	1,371034099	1,342966353	-1,311301204	1,327757354	-1,216207379	1,399170433	-1,084619143
4,856487028	-1,331977878	1,021913888	2,112753077	-1,121894161	2,066664828	1,376262194	-3,242496643	1,073534633	-1,078203858
-1,356698357	1,942097888	-1,051497403	-1,228690158	-3,128855851	1,212212981	-3,7734873	2,271509247	-1,439361812	1,453262592
-5,565029655	1,083131579	4,922654166	-1,297429123	1,390371837	-3,238868274	1,407800616	1,448628288	2,309271564	-1,731744697
-2,090754413	1,029512534	-3,079459393	1,106698839	-1,672070192	1,166996029	-1,03187887	-1,37450836	1,917230792	-1,580039494
-1,264957381		-1,8148266	1,019963283	-1,025896586	-2,752387904	-1,482553961	-1,496278844	1,696891324	-1,445423206
			-2,018134512	-2,654192351	1,125759948	1,33564785	-1,946089299	-1,216644448	1,609842406
			-1,085715347	1,919074862	1,31475421	-4,394126907	-1,677771886	-1,131779037	-1,22869521
			3,169254351	1,175762123	1,131019532	1,162260551	1,256368282	1,340122006	-1,087390889
			1,085988379	1,283166865	1,782460175	-1,365534482	1,384862525	-1,440903955	-1,04593805
			3,494789484	4,694637169	2,770419345	-1,065609878	-1,480035545	1,010146659	-1,011961141
			-1,653381023	2,569681922	-1,016788055	-1,239771855	-1,005944389	-2,980778318	2,724475543
			-1,175775916	1,103737552	-1,817598855	-1,217022764	1,18563037	1,2767375	-1,215044272
			-1,14298026	1,087714784	-1,126181947	-1,821709843	-1,939276549	-4,672143542	1,406679874
			1,022366203	1,299002624	1,446883791	1,967206824	-1,089829956	2,938816211	2,018114265
			1,000567047	-1,916190549	-1,620018824	1,116190242	1,353952429	1,071212199	1,757624942
			-1,776781504	2,373645456	-1,327550139	1,494164808	-2,817443523	-4,826059689	-1,814306544
			3,2934648	-4,689216939	-1,36503095	-1,988088304	1,327369834	-1,117414053	-1,497947679
			1,204499693	-4,139234587	-1,535127273	1,520713675	2,801813878	1,678498055	1,326515585
			1,082102616	1,291759643	-2,316402172	-1,008429958	1,44142143	-1,463635609	1,412093599
			-3,04188361	1,220897296	-3,587560763	-1,519999702	-1,223069334	1,45579891	-1,63671683
			-1,654933731	-1,264661775	3,338253484	1,368261983	1,664535765	-1,899222252	-1,531007406
			1,032026431	-1,161345339	3,041414887	-1,70375812	-1,17812968	-1,32828219	3,141938228
			-1,775296171	-3,623971528	-1,115418902	-1,263355088	-1,740339052	3,634660404	-1,784965259
			-5,691527297	-1,792604474		3,114399717	1,189129786	-1,189165974	2,161626135
			-1,481578501	1,091958491		-3,135655842		1,233885389	1,27256363

Conclusions

Uniwersytet Wrocławski
CLARIN
Common Language Resources and Technology Infrastructure
Politechnika Wrocławska

Conclusions

	Women	Men
PL	1. Romance (love, woman, marriage, sex). 2. Historical romance. 3. A novel for young people. 4. Child rearing. 5. Elementary teaching. 6. Cooking, cookbooks. 7. Foreign languages (teaching, dictionaries). 8. Drama.	1. Science (history, philosophy, economics). 2. Law. 3. Management. 4. Engineering (all with computer science). 5. Medicine. 6. Political speech. 7. Religion.
DE	1. Romance (love, woman, eroticism, marriage). 2. Detective and adventure novels. 3. Fairy tales for children. 4. Elementary teaching. 5. Cooking, cookbooks. 6. Letters, correspondence. 7. In vitro treatment. 8. Medical care (therapies). 9. Personal development.	1. Science (history, philosophy, economics). 2. Law. 3. Management. 4. Finance (counseling). 5. Engineering (all domains with computer science). 6. Medicine. 7. Political speech. 8. Religion.

Conclusions

1. Large bibliographies can be processed using NLP, linguistics and text-mining methods.
2. Information generated from large bibliographies is helpful in analyzing cultural phenomena.
3. The gender of book authors is evolving towards balance.
4. The process of this evolution is very similar in PL and DE.
5. The distribution of topics in PL and DE is very similar.
6. The topics of books written by women and men are definitely different.
7. The dominance of common themes is huge (quantitatively) and shows the persistence of inherited values and/or natural traits in authors/women authors.
8. In men, the division is almost identical (the difference is made by the FINANSE category in DE).
9. In women, the difference is more noticeable.
10. Specific female topics in DE are PERSONAL DEVELOPMENT, PSYCH THERAPY, IN VITRO, CRIMINAL FICTION.
11. Specific female topics in PL are FOREIGN LANGUAGES.

THANK YOU FOR YOUR ATTENTION