



Multilinguality in Wikidata

Lucie-Aimée Kaffee
kaffee@soton.ac.uk



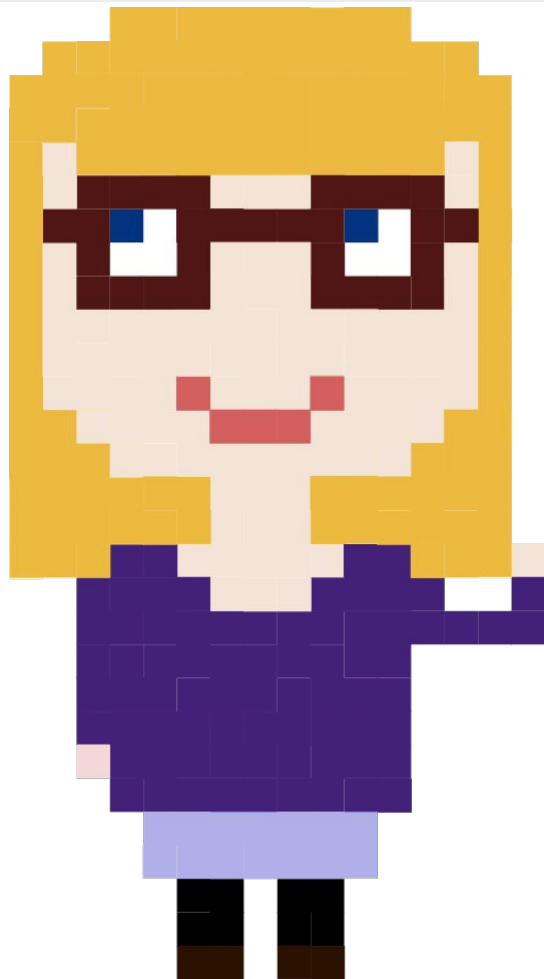
About Me

PhD Student
WAIS, University of Southampton

Previously worked as a Software
Developer at Wikimedia Deutschland, in
the Wikidata team

Interest in (under-resourced) languages

From Berlin, Germany





What we will talk about

Wikidata

Multilinguality in Wikidata

My work

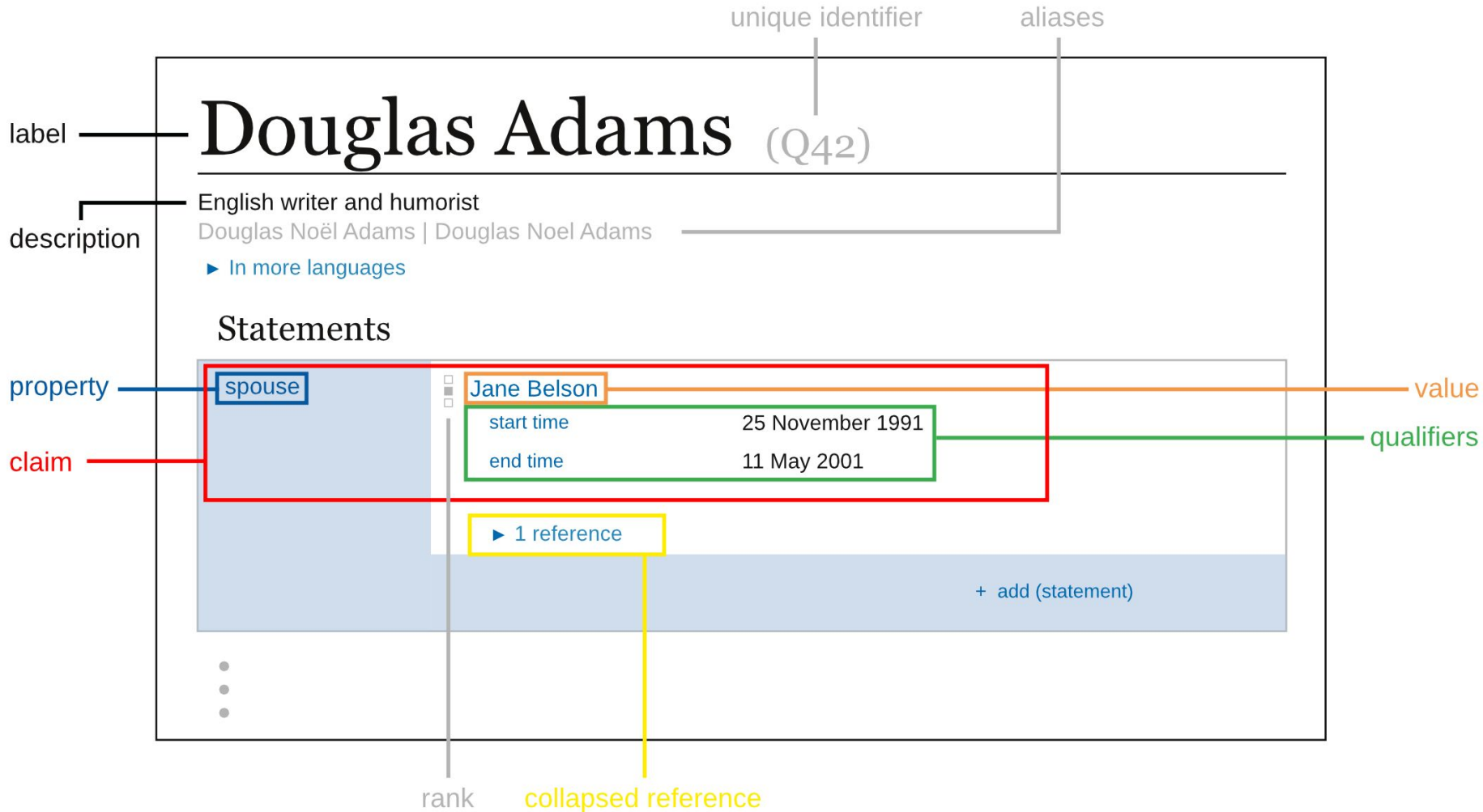


WIKIDATA



Wikidata

- Free knowledge base by the Wikimedia movement
- Started ~4 years ago
- To support Wikipedia
- Linked data
- Persistent, opaque identifiers
- > 400 languages
- People, places, events, and more
- Powered by Wikibase (open source MediaWiki extension)
- Data licensed under CC-0



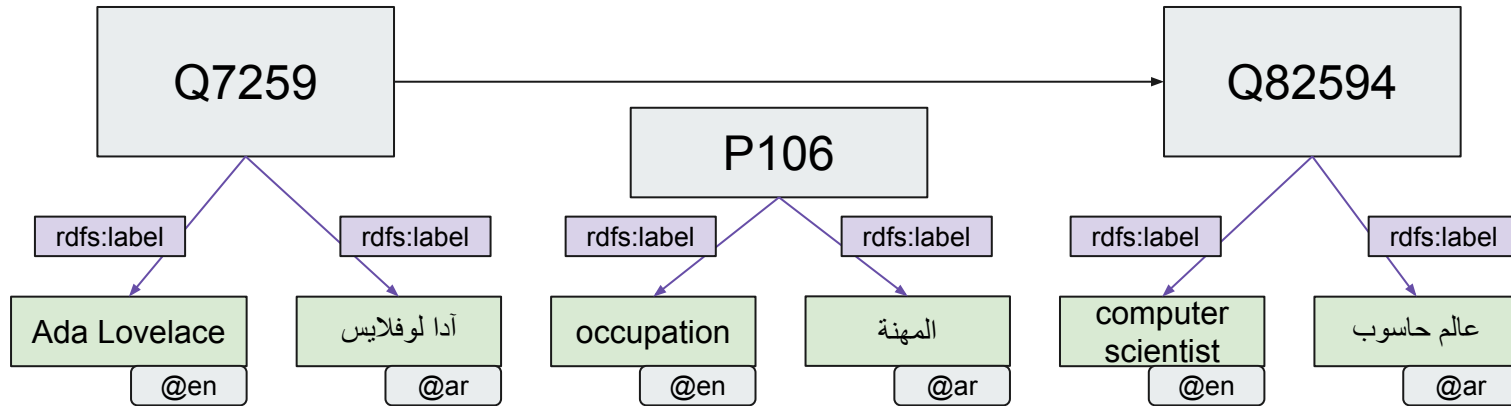
Multilinguality in Wikidata

A Glimpse into Babel: An Analysis of Multilinguality in Wikidata

Lucie-Aimée Kaffee, Alessandro Piscopo, Pavlos Vougiouklis, Elena Simperl, Leslie Carr, Lydia Pintscher

OpenSym 2017

Multilinguality in Wikidata





Multilinguality in Wikidata - Why do we care?

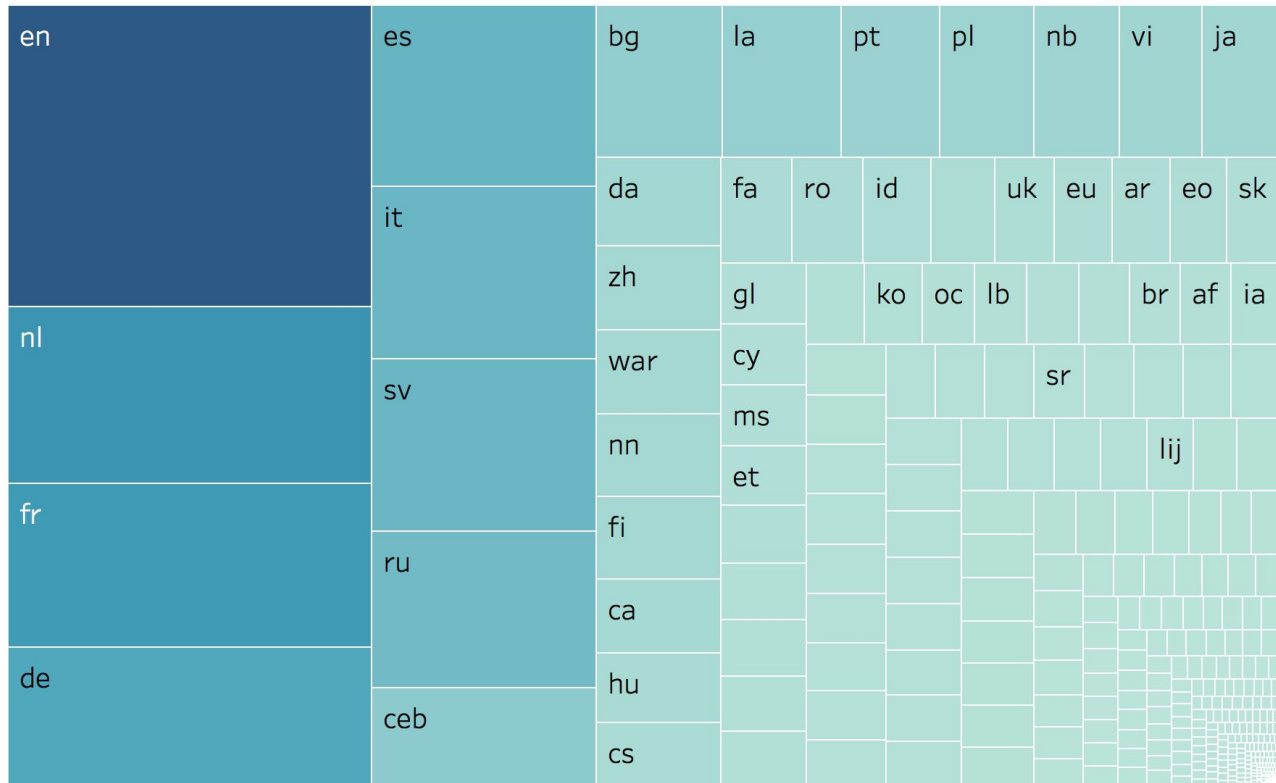
- Labels are the access point for humans
- Give language communities access to existing knowledge
- Central storage for translations for (under resourced) languages
- Semantic Web in NLP and NLG
- Wikipedia, translation, question answering, chat bots, ...



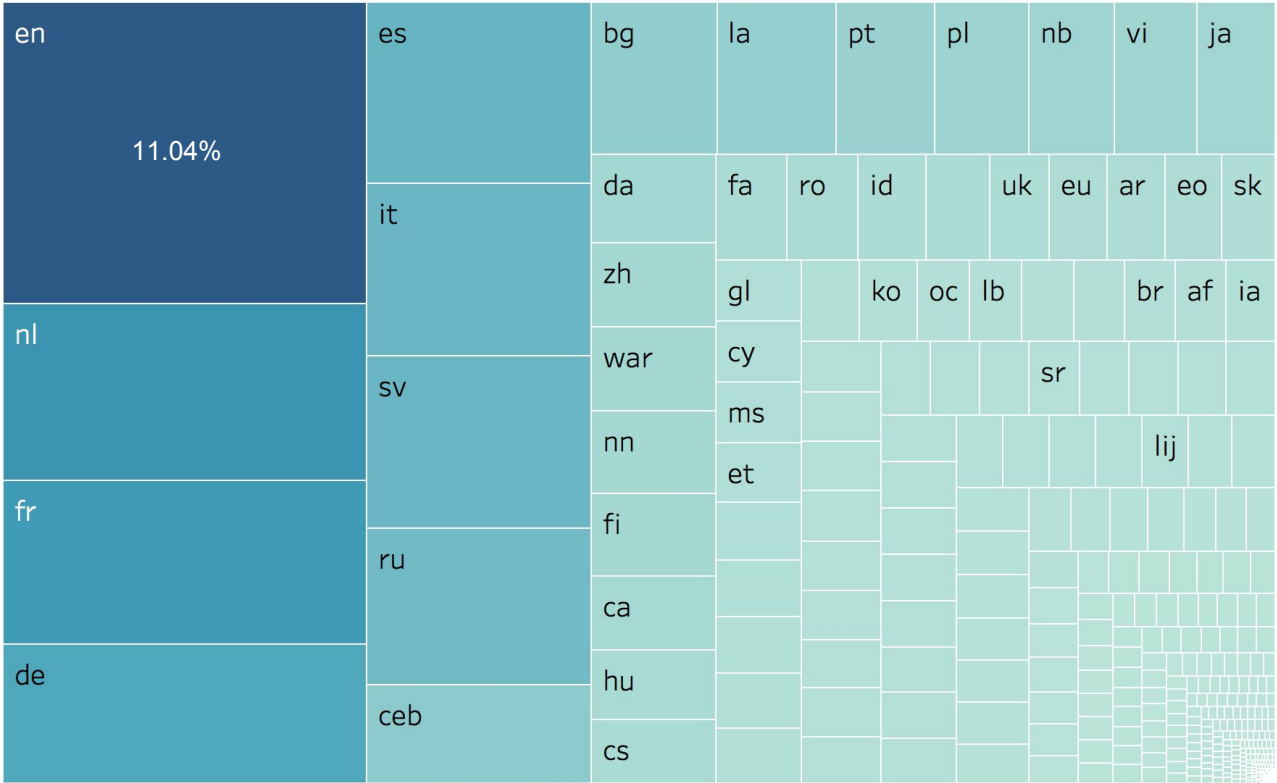
Research Questions

- What is the state of Wikidata with regard to multilinguality?
- How does Wikidata's label distribution relate to the real world and Wikipedia's language distribution?
- Is there a difference in the multilinguality of the properties, compared to the overall multilinguality of the knowledge base?

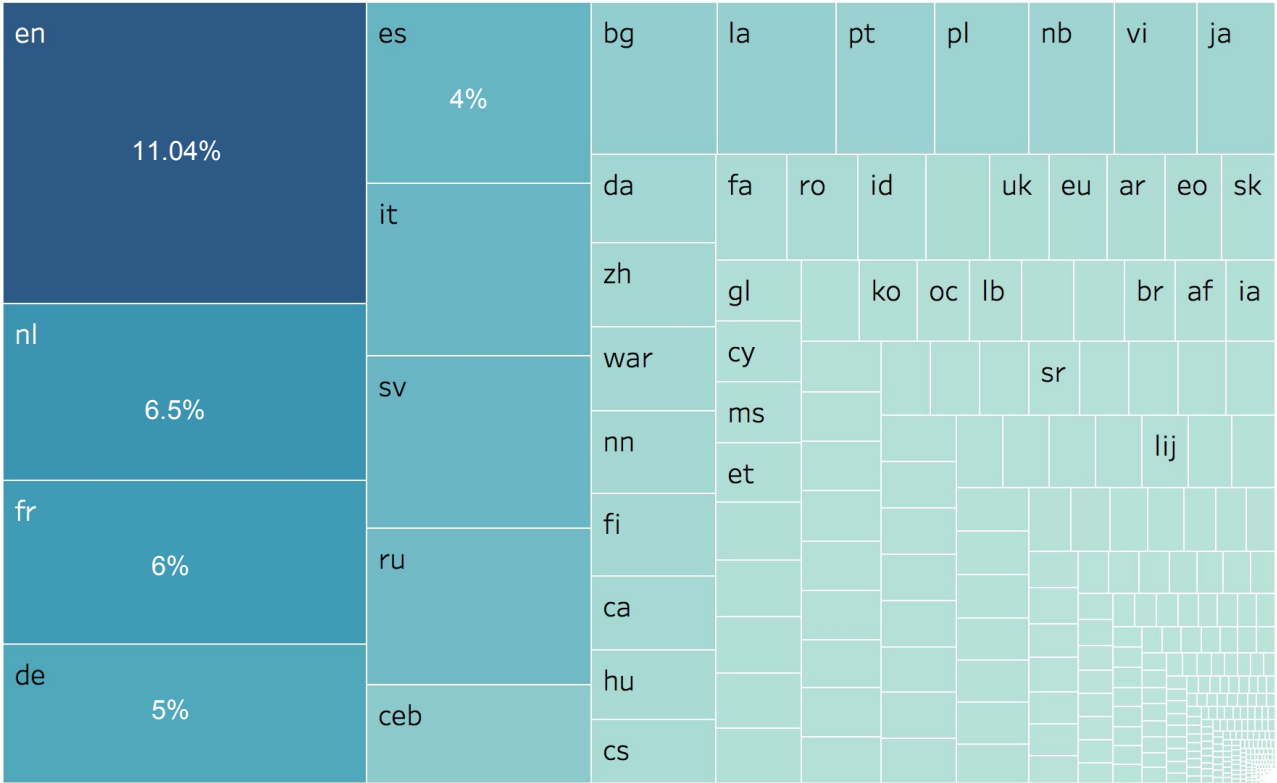
Wikidata

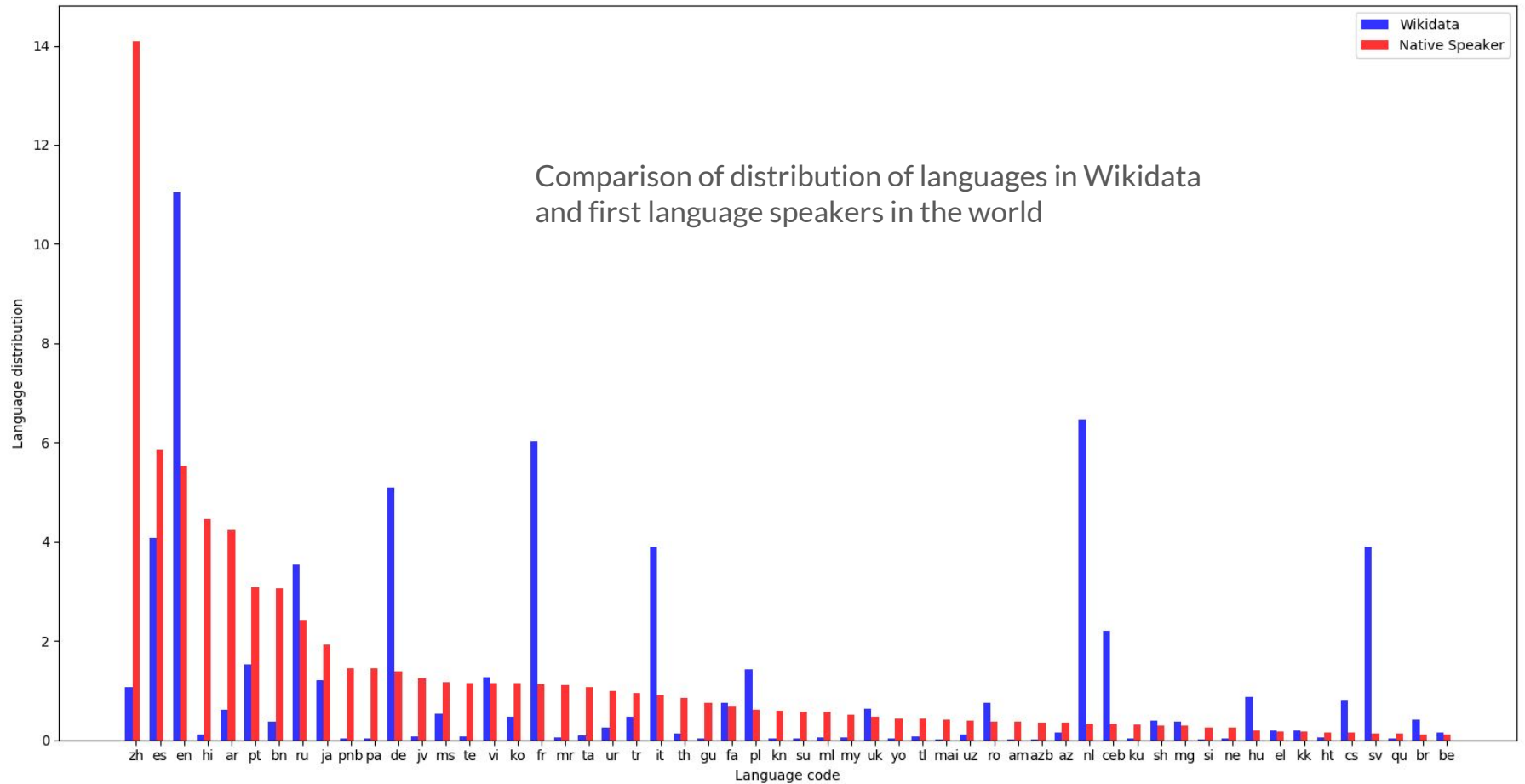


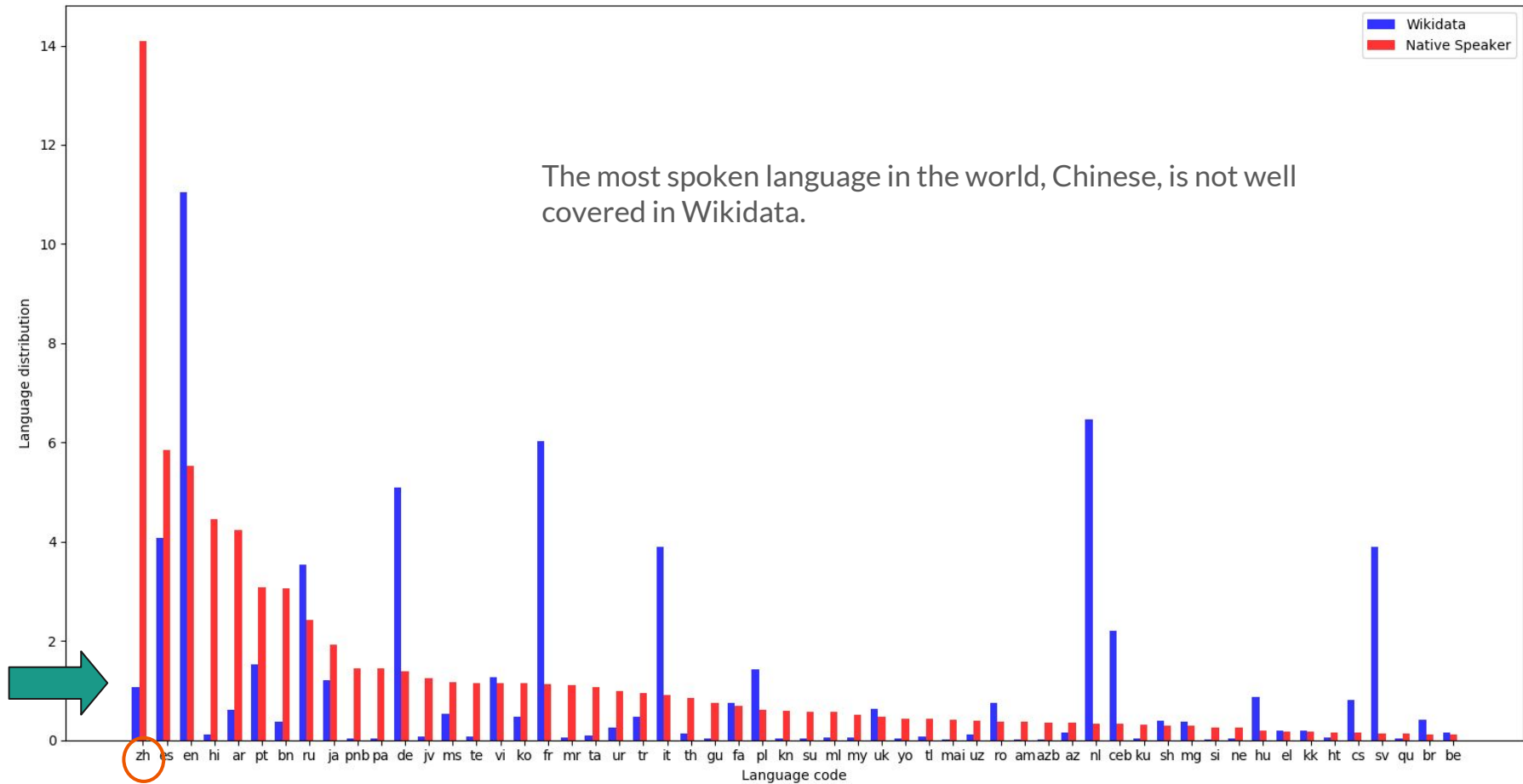
Wikidata

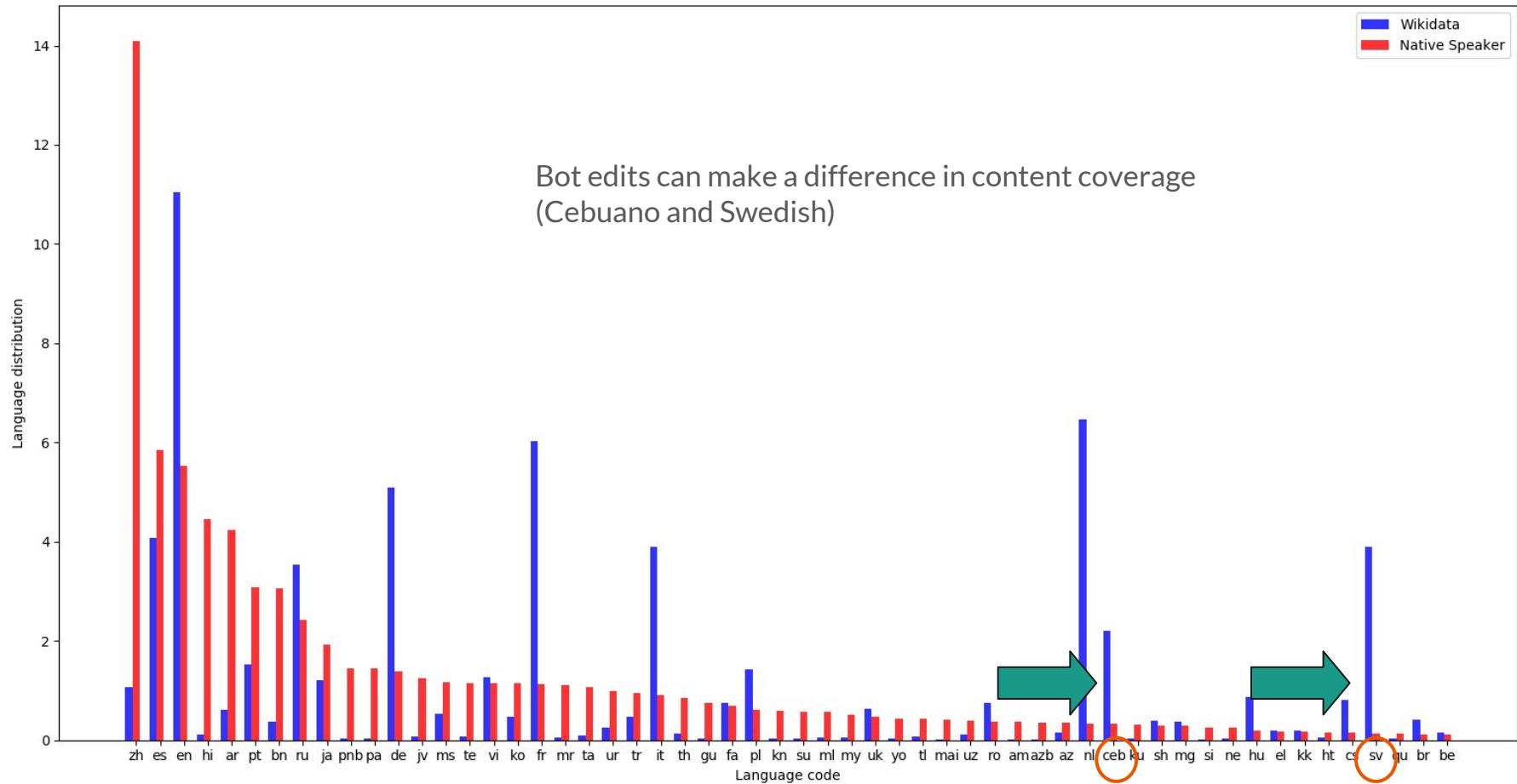


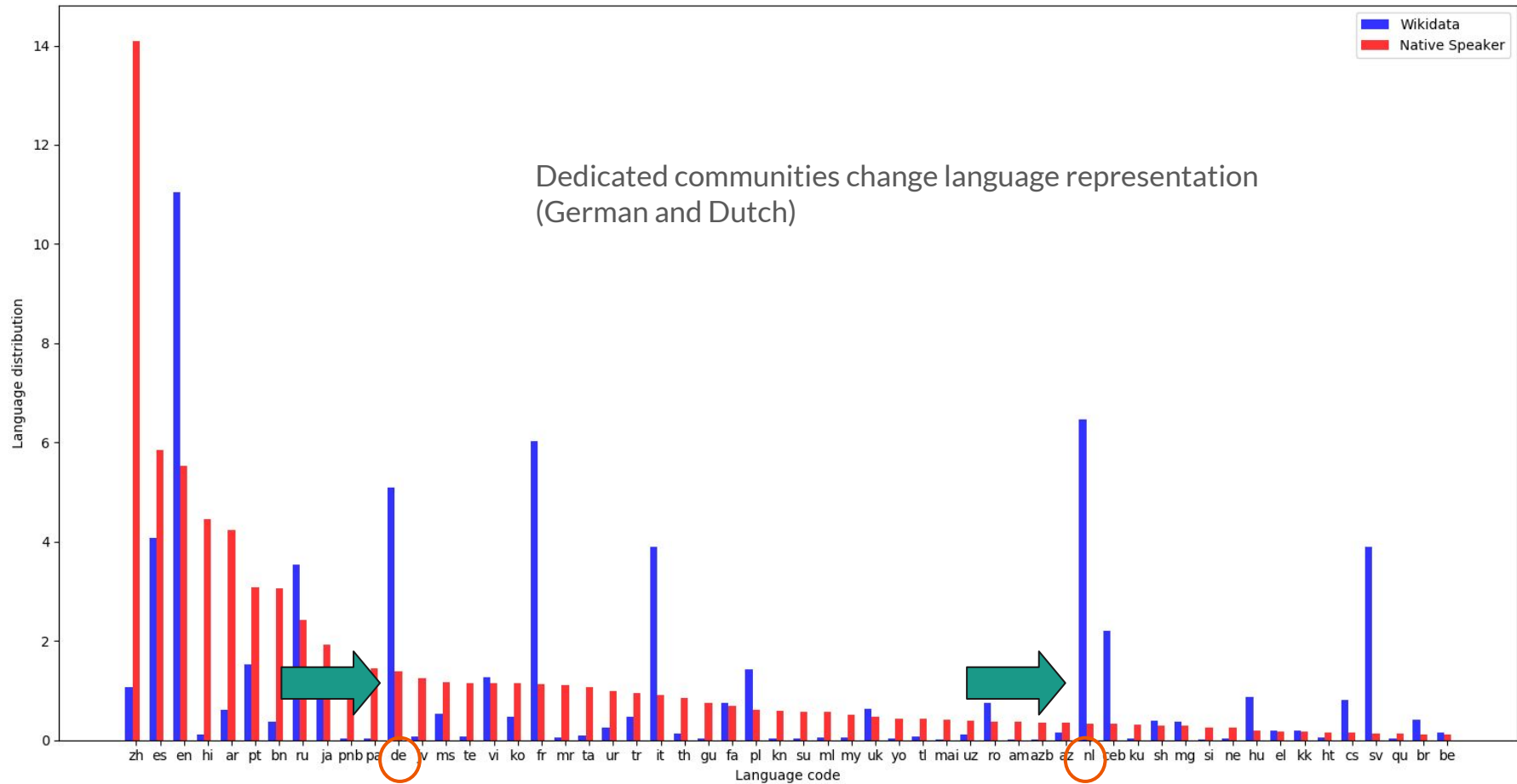
Wikidata



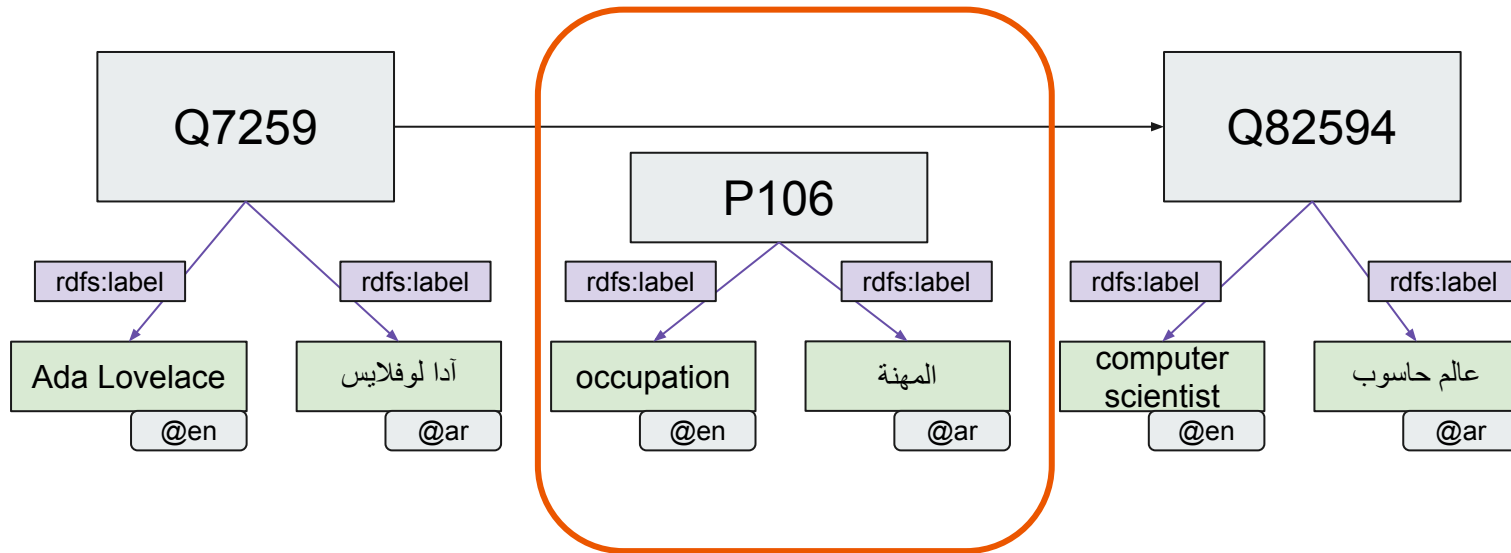








Wikidata Properties





Ranking of number of

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



Ranking of number of
Wikipedia articles by language,

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



Ranking of number of
Wikipedia articles by language,
all labels in Wikidata,

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



Ranking of number of
Wikipedia articles by language,
all labels in Wikidata,
and labels for properties in Wikidata

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



German is widely used in Wikipedia and Wikidata

High coverage through active community

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



As German, active community that brings high coverage of labels

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



High coverage in labels on Wikidata
through high number of bot-imported
Wikipedia articles, however low number of
community edited properties

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh



Even more extreme than Swedish: Not in top 25 of community-edited properties by language

Rank	Wikipedia	Wikidata	WD properties
1	en	en	en
2	ceb	nl	nl
3	sv	fr	fr
4	de	de	ru
5	nl	es	mk
6	fr	it	de
7	ru	sv	es
8	it	ru	pl
9	es	ceb	ca
10	war	bg	it
11	pl	la	sr
12	vi	pt	hu
13	ja	pl	pt
14	pt	nb	nb
15	zh	vi	ko
16	uk	ja	fa
17	ca	da	da
18	fa	zh	cs
19	ar	war	ja
20	no	nn	sv
21	sh	fi	be
22	fi	ca	el
23	hu	hu	ar
24	id	cs	uk
25	cs	fa	zh

Users in Wikidata

(Work in Progress)

Babel user information

de-N

Diese Benutzerin spricht **Deutsch** als Muttersprache.

en-4

This user has **near native speaker** knowledge of **English**.

es-2

Esta usuaria tiene un conocimiento **intermedio** del **español**.

tr-1

Bu kullanıcı **temel düzeyde Türkçe** bilir.

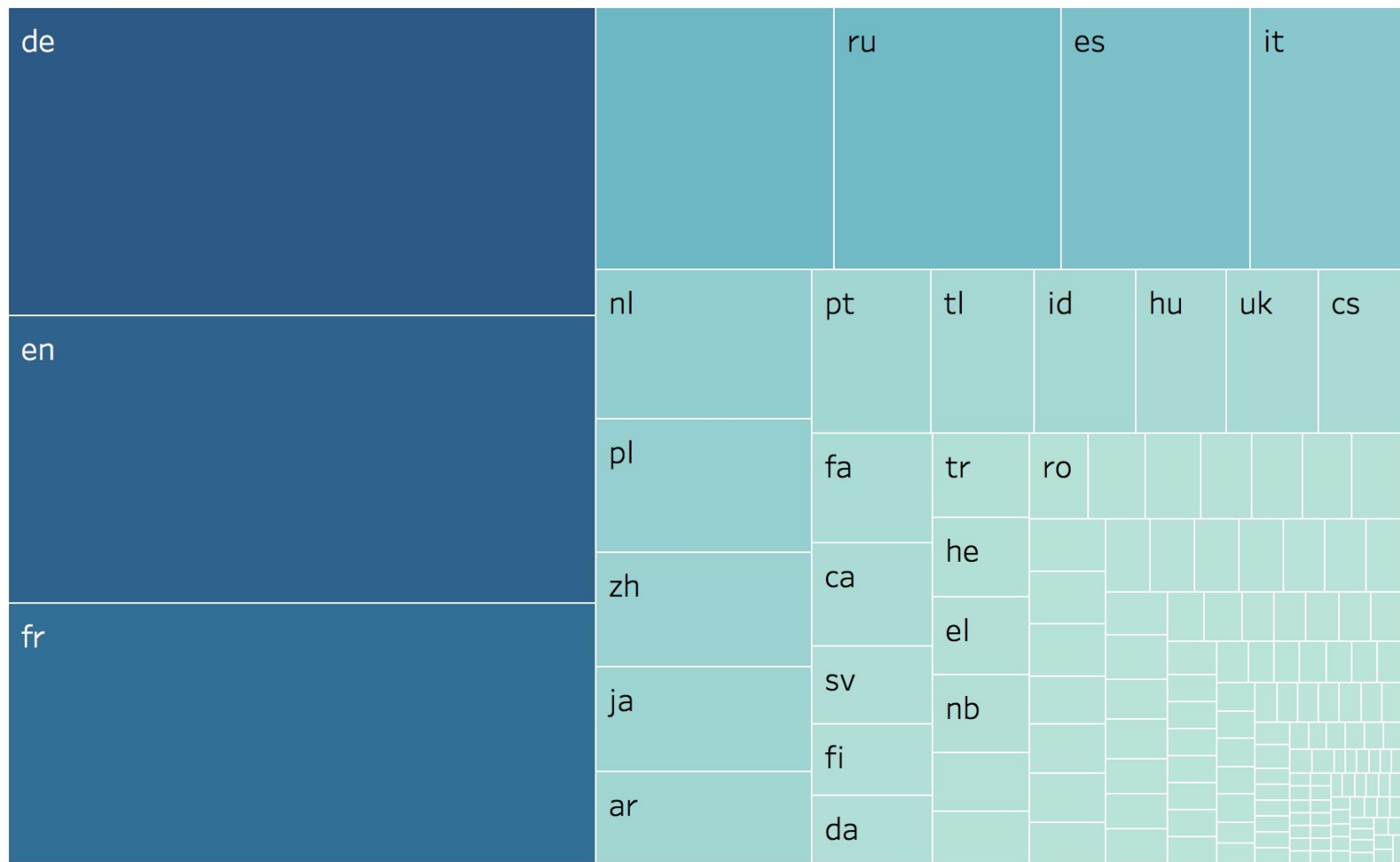
fr-1

Cette utilisatrice dispose de connaissances **de base** en **français**.

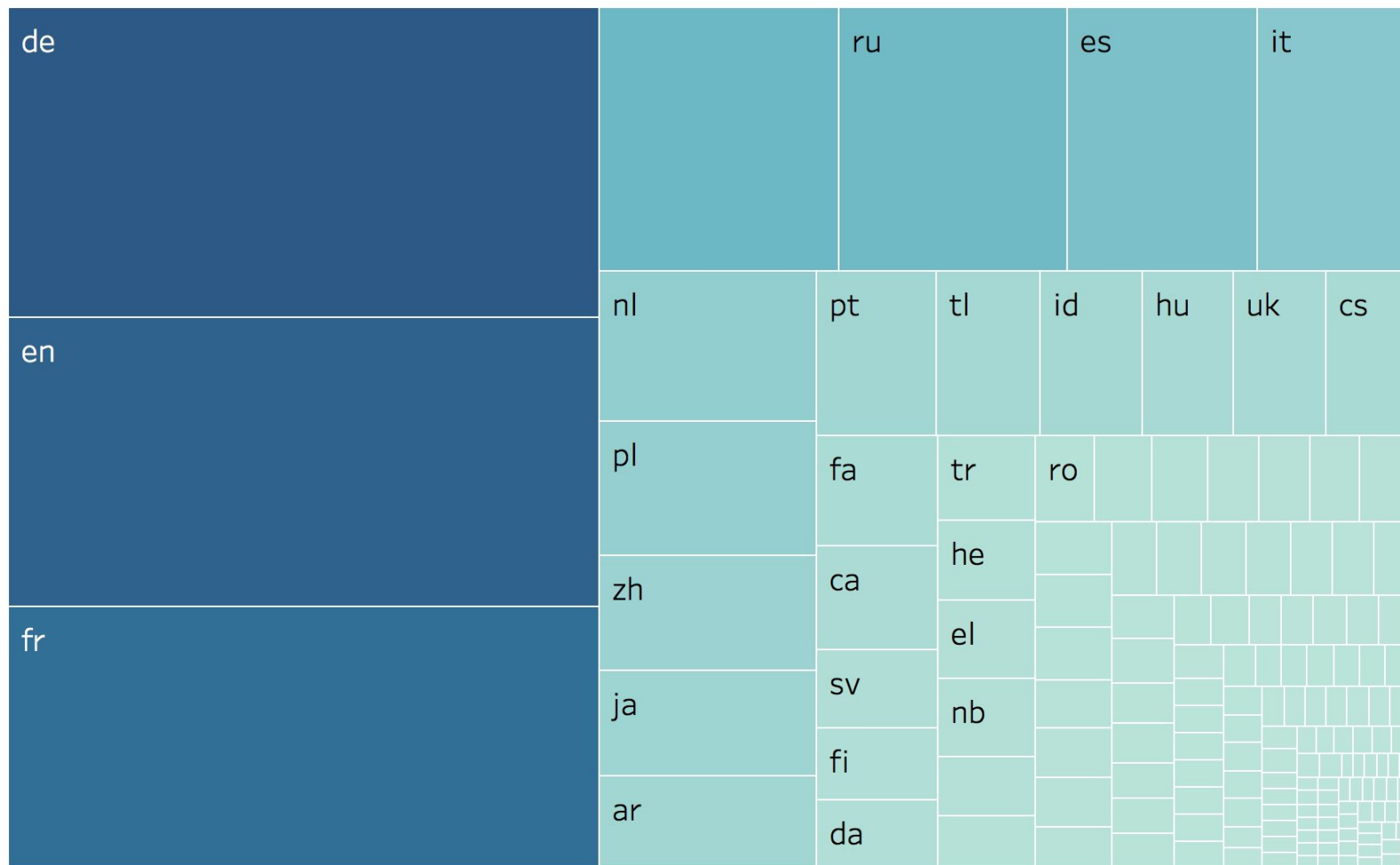
ar-0

هذه المستخدمة **ليس** لديها معرفة **بالعربية** (أو تفهمها بصعوبة بالغة).

Users by language



Native Languages of Wikidata users



Language coverage of labels does not reflect in languages Wikidata's users speak

Native Languages of Wikidata users

Learning to Generate Text from Wikidata for Wikipedia's ArticlePlaceholder



Wikipedia available in 285 languages, but content is unevenly distributed
Wikidata is cross- and multilingual

Page Statistic	Esperanto	Arabic	English	Wikidata
Articles	241,901	541,166	5,483,928	37,703,807
Avg edits/page	11.48	8.94	21.11	14.66
Active users	2,849	7,818	129,237	17,583
Vocab. size	1.5M	2.2M	2.0M	—

Recent page statistics and number of unique words (vocab. size) of Esperanto, Arabic and English Wikipedias in comparison with Wikidata



Content From Wikidata to Wikipedia



WIKIDATA



WIKIPEDIA
The Free Encyclopedia

Learning to Generate Wikipedia Summaries for Underserved Languages from Wikidata

Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest,
Jonathon Hare, Elena Simperl

NAACL 2018

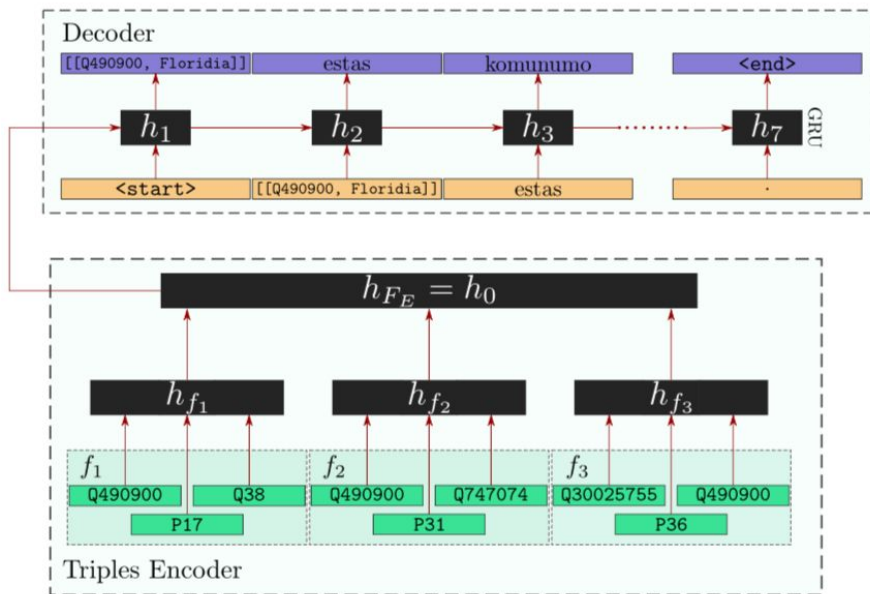
Neural Text Generation

Arabic and Esperanto output text

Feed-forward architecture encodes Wikidata triples into vector of fixed dimensionality

RNN-based decoder generates text summaries, one token at a time

Article-Placeholder Triples	f_1 : Q490900 (Florida)	P17 (ŝtato)	Q38 (Italio)
	f_2 : Q490900 (Florida)	P31 (estas)	Q747074 (komunumo de Italio)
	f_3 : Q30025755 (Florida)	P1376 (ĉefurbo de)	Q490900 (Florida)
Textual Summary	Florida estas komunumo de Italio.		
Vocab. Extended Summary	[[Q490900, Florida]] estas komunumo de [[P17]].		





Automatic Evaluation

- Baselines
 - Machine Translation, Information Retrieval Based, Kneser-Ney
- Automatic Evaluation
 - BLEU 1, BLEU 2, BLEU 3, BLEU 4, METEOR, ROUGE

Model		BLEU 1		BLEU 2		BLEU 3		BLEU 4		ROUGE _L		METEOR	
		Valid.	Test	Valid.	Test	Valid.	Test	Valid.	Test	Valid.	Test	Valid.	Test
Arabic	MT	31.12	33.48	19.31	21.12	12.69	13.89	8.49	9.11	29.96	30.51	31.05	30.1
	TP	41.39	41.73	34.18	34.58	29.36	29.72	25.68	25.98	43.26	43.58	32.99	33.33
	TP _{ext}	49.87	48.96	42.44	41.5	37.29	36.41	33.27	32.51	51.66	50.57	34.39	34.25
	Ours	53.18	52.94	45.86	45.64	40.38	40.21	35.7	35.55	57.9	57.99	39.22	39.37
Esperanto	MT	5.35	5.47	1.62	1.62	0.59	0.56	0.26	0.23	4.67	4.79	0.66	0.68
	TP	43.01	42.61	33.67	33.46	28.16	28.07	24.35	24.3	46.75	45.92	20.71	20.46
	TP _{ext}	52.75	51.66	43.57	42.53	37.53	36.54	33.35	32.41	58.15	57.62	31.21	31.04
	Ours	56.51	56.96	47.72	48.1	41.8	42.13	37.24	37.52	64.36	64.69	28.35	28.76

Results of the automatic evaluation: **Our network outperforms all baselines**

Mind the (Language) Gap: Generation of Multilingual Wikipedia Summaries from Wikidata for ArticlePlaceholders

Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest, Jonathon Hare, Elena Simperl

ESWC 2018

Currently deployed on 11
Wikipedias

[illegible]



[masquer]

dinosaure herbivore de la famille des cératopsidés

taxon fossile

Triceratops

auteur taxonomique: Othniel Charles Marsh

date de description scientifique: 1889

genre

Ceratopsinae

النَّزِيمِيُّ اتَّوَيْسَ

טריצרטופס

Трицератопс

三角龍

Maastrichtien

Maastrichtien

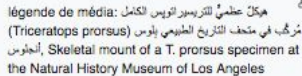
3 Mèt (mezi)

8 500 Kilogram

9 Mèt (mezi)

Triceratops

Triceratops



identifiant bibliothèque du Congrès [sh880002](#)
[54^{\[1\]}](#)

identifiant Encyclopædia Britannica
en ligne

identifiant Encyclopædia Britannica
en ligne

identifiant de la Grande Encyclopédie russe en ligne 4204642

identifikan sou Freebase [/m/0f6mr](#)
[2]

identifiant Fossilworks 38862^[3]

identifiant Encyclopædia Universalis [triceratop](#)
en ligne [s](#)

identifiant Gemeinsame Normdatei 4331835-6

identifiant Quora d'un sujet Tricerato

Referans

1. † bibliothèque du Congrès, 15 janvye 2018
2. † sauvegarde de la base de données Freebase, 28 oktòb 2013
3. † Fossilworks, 13 me 2015
4. † Interim Register of Marine and Nonmarine Genera, 19 avril 2018

Working with Arabic and Esperanto

[illegible]



Community Study

- Two 15 days online surveys, aimed at readers and editors in Esperanto and Arabic
- Aiming to test our work with the actual Wikipedia community, outreach on Wikipedia platforms
- Reader:
 - ◆ Fluency: Is the text understandable and grammatically correct?
 - ◆ Appropriateness: Does the summary 'feel' like a Wikipedia article?
- Editor:
 - ◆ Editors were asked to edit the article starting from our summary (2-3 sentences)
 - ◆ How much of the text was reused?

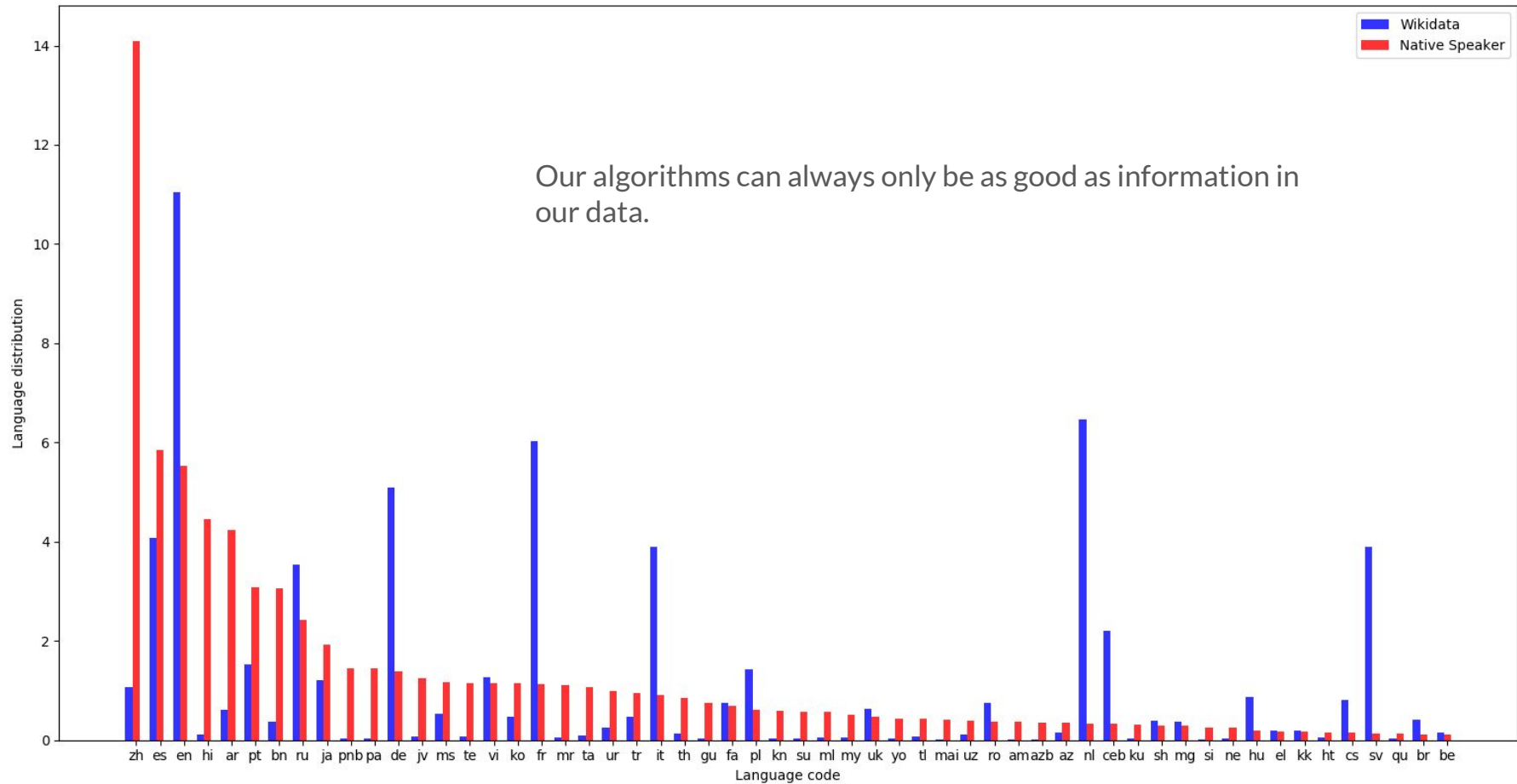


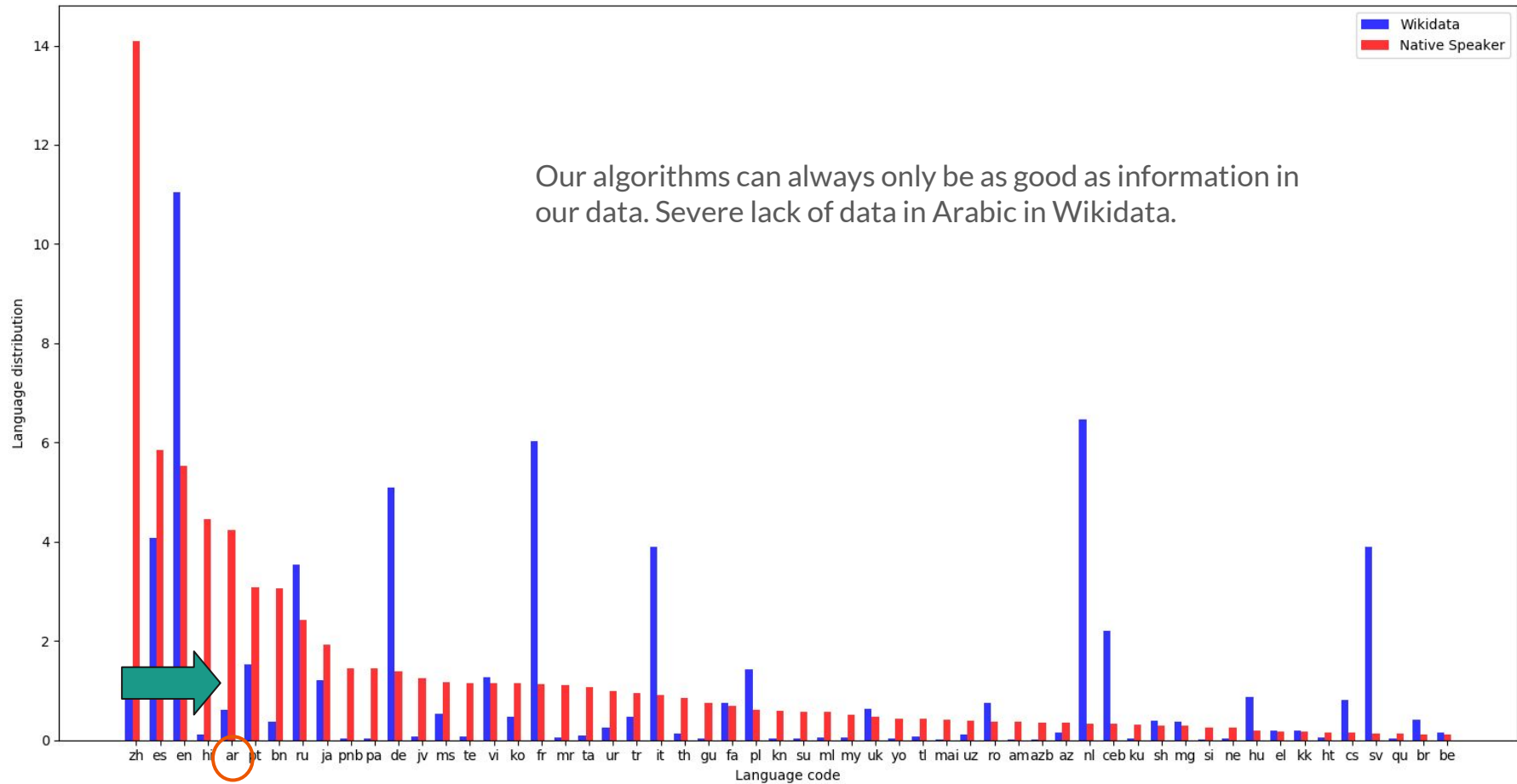
		Fluency		Appropriateness	
		Mean	SD	Part of Wikipedia	
Arabic	Ours	4.7	1.2	77%	
	Wikipedia	4.6	0.9	74%	
	News	5.3	0.4	35%	
Esper.	Ours	4.5	1.5	69%	
	Wikipedia	4.9	1.2	84%	
	News	4.2	1.2	52%	

Results of the reader study: We generate sentences of comparable fluency, that “feel” like Wikipedia sentences

		Category	Examples	%
Arabic	WD		خماسي كلوريد الزرنيخ مركب كيميائي له الصيغة (كلمة ناقصة)، ويكون على شكل بلورات بيضاء. A B	45.45%
			خماسي كلوريد الزرنيخ هو مركب كيميائي له الصيغة (AtClu2085)، ويكون على شكل بلورات بيضاء. A B	
	PD		بيتش باتوم أوهايو (بالإنجليزية (كلمة ناقصة) Ohio) هي منطقة سكنية تقع في الولايات المتحدة في (كلمة ناقصة). D E	33.33%
			بيتش باتوم (بالإنجليزية: Beach Batom) هي قرية تقع في الولايات المتحدة الأمريكية في بروك كاونتري. D E	
Esperanto	WD		دير علا هي بلدة تقع في جنوب غرب إيران. F	21.21%
			دير علا، أو بيتر، هي قرية أردنية F	
	PD		Zederik estas komunumo en la nederlanda provinco Zuid-Holland. G	78.98%
			Zederik estas komunumo en la nederlanda provinco Zuid-Hooland kaj estas ĉirkaŭata de la municipoj Lopik kaj Zederik. G	
Esperanto	PD		Nova Pádua estas municipo en la brazila subŝtato Suda Rio-Grando. H	15.79%
			Nova Pádua estas municipo en la brazila subŝtato Suda Rio-Grando. H	
Esperanto	ND		Ibiúna estas municipo de la brazila subŝtato San-Paŭlio, kiu taksis (manka nombro) enloĝantojn en (jaro). I J K L	5.26%
			Ibiúna estas brazila [[municipo]] kiu troviĝas en la administra unuo [[San-Paŭlo]]. I J K L	

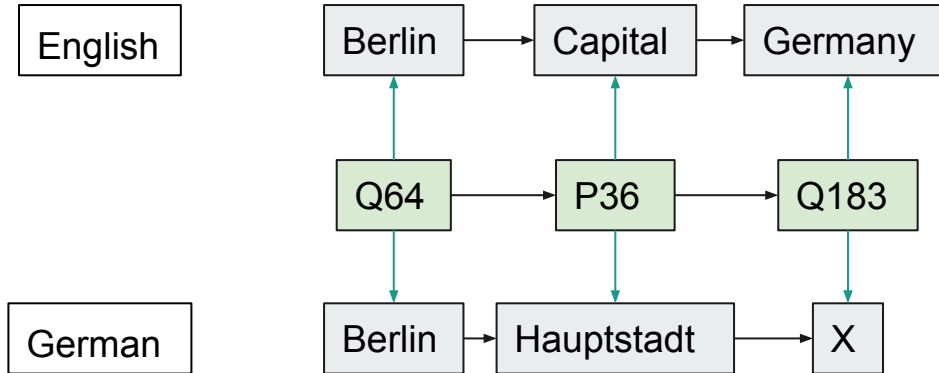
Results of the editor study: **We generate sentences that are highly reused by editors**



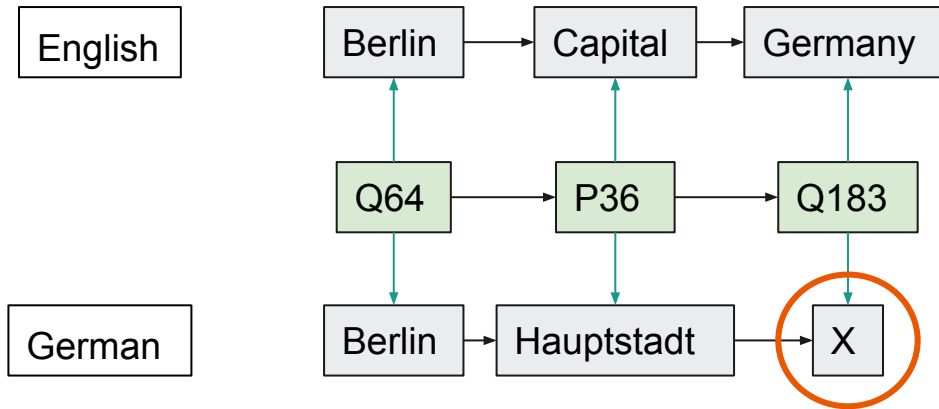


Future Work: Label Extraction From Wikipedia For Wikidata

Example Wikidata Triple



Example Wikidata Triple





WIKIPEDIA
Die freie Enzyklopädie

Hauptseite

Themenportale

Von A bis Z

Zufälliger Artikel

Mitmachen

Artikel verbessern

Neuen Artikel anlegen

Autorenportal

Hilfe

Letzte Änderungen

Kontakt

Spenden

Werkzeuge

Links auf diese Seite

Änderungen an

verlinkten Seiten

Datei hochladen

Spezialseiten

Permanenter Link

Seiteninformationen

Wikidata-Datenobjekt

Artikel zitieren

Drucken/exportieren

Buch erstellen

Als PDF herunterladen

Druckversion

In anderen Projekten

Commons

Wikinews

Wikiquote

Wikisource

★ Wikivoyage

In anderen Sprachen

Frimelle Einstellungen Beta Beobachtungsliste Beiträge Abmelden

Artikel Diskussion

Lesen

Bearbeiten

Quelltext bearbeiten

Versionsgeschichte

☆

Mehr ▾

Wikipedia durchsuchen



Koordinaten: 52° 31′ N, 13° 24′ O﻿ / ﻿



Berlin



Dieser Artikel befasst sich mit der deutschen Hauptstadt. Zu weiteren Bedeutungen siehe [Berlin \(Begriffsklärung\)](#).

Berlin [bɛʁˈlɪn] ist die **Bundeshauptstadt** der **Bundesrepublik Deutschland**^[14] und zugleich eines ihrer Länder. Die Stadt Berlin ist mit fast 3,6 Millionen Einwohnern die bevölkerungsgrößte Gemeinde Deutschlands. Sie bildet das Zentrum und der Agglomeration Berlin (4,5 Millionen Einw.). Der Stadtstaat und die Havel befinden sich im Stadtgebiet kleinere Fließgewässer s

Urkundlich erstmals im 13. Jahrhundert erwähnt, war Berlin im 19. Jahrhundert die Hauptstadt Preußens. Im 20. Jahrhundert wurde Berlin wieder gesamtdeutsche Hauptstadt und in der Folge Sitz des Bundestages, des Bundesrates sowie zahlreicher Bundesministerien.

Zu den bedeutenden Wirtschaftszweigen in Berlin gehören unter anderem die Kulturwirtschaft, die Biotechnologie und Gesundheitswirtschaft, die Informations- und Kommunikationstechnologien, die Bau- und Ingenieurwirtschaft, die Energietechnik sowie die Messe- und Kongresswirtschaft. Die Stadt ist ein Zentrum des Luftverkehrs. Berlin zählt zu den aufstrebenden, internationalen Metropolen und verzeichnet jährlich hohe Zuwachsraten bei der Zahl der Erwerbstätigen.

Berlin gilt als Weltstadt der Kultur, Politik, Medien und Wissenschaft. Es beherbergt zahlreiche Forschungseinrichtungen, Sportereignisse^[20] und Museen. Berlin ist seit 2020 ein UNESCO-Weltkulturerbe mit dem Titel *Stadt des Designs* und ist eines der meistbesuchten Zentren des Kontinents.^[21] Berlins Architektur, Festivals, Nachtleben und vielfältige Lebensbedingungen sind weltweit bekannt.^[22]



Wikidata:Q183

Englisch: Germany
country in central Europe
ist ein(e) — souveräner Staat
Gründung/Erstellung — 23. Mai 1949
ist Teil von — (en) Central Europe

Approve

Reject



Einw.)
ipree

teil
wurde

lie
die
und

den

Land Berlin



Landesflagge



Landeswappen



Basisdaten

Staatsform	parlamentarische Republik, teilsouveräner Gliedstaat eines Bundesstaates
Sprache	Deutsch ^[1]
Postleitzahlen	10115–14199
Vorwahl	030
Kfz-Kennzeichen	B
Bruttoinlandsprodukt (BIP)	129,424 Mrd. € (Februar 2017) ^[2]
BIP pro Kopf	35.627 € (Februar 2016) ^[3]
Schulden	56,935 Mrd. € (31. Dezember 2017) ^[4]
ISO 3166-2	DE-BE

Inhaltsverzeichnis [Verbergen]

- Geographie
 - Klima
 - Gliederung
 - Parkanlagen, Zoos, Schutzgebiete
 - Nachbarstädte und -gemeinden
- Geschichte
 - Namensbildung
 - Markgrafschaft und Kurfürstentum
 - Das 19. und 20. Jahrhundert



WIKIPEDIA
Die freie Enzyklopädie

Hauptseite

Themenportale

Von A bis Z

Zufälliger Artikel

Mitmachen

Artikel verbessern

Neuen Artikel anlegen

Autorenportal

Hilfe

Letzte Änderungen

Kontakt

Spenden

Werkzeuge

Links auf diese Seite

Änderungen an

verlinkten Seiten

Datei hochladen

Spezialseiten

Permanenter Link

Seiteninformationen

Wikidata-Datenobjekt

Artikel zitieren

Drucken/exportieren

Buch erstellen

Als PDF herunterladen

Druckversion

In anderen Projekten

Commons

Wikinews

Wikiquote

Wikisource

★ Wikivoyage

In anderen Sprachen

Frimelle Einstellungen Beta Beobachtungsliste Beiträge Abmelden

Artikel Diskussion

Lesen

Bearbeiten

Quelltext bearbeiten

Versionsgeschichte

☆

Mehr ▾

Wikipedia durchsuchen



Koordinaten: 52° 31′ N, 13° 24′ O﻿•﻿



Berlin



Dieser Artikel befasst sich mit der deutschen Hauptstadt. Zu weiteren Bedeutungen siehe [Berlin \(Begriffsklärung\)](#).

Berlin [bɛʁˈlɪn] ist die **Bundeshauptstadt** der **Bundesrepublik Deutschland**^[14] und zugleich eines ihrer Länder. Die Stadt Berlin ist mit fast 3,6 Millionen Einwohnern die bevölkerungsgrößte Gemeinde Deutschlands. Sie bildet das Zentrum und der Agglomeration Berlin (4,5 Millionen Einw.). Der Stadtstaat und die Havel befinden sich im Stadtgebiet kleinere Fließgewässer s

Urkundlich erstmals im 13. Jahrhundert erwähnt, war Berlin im 19. Jahrhundert die Hauptstadt des Königreichs Preußen. Nach dem Zweiten Weltkrieg wurde Berlin wieder gesamtdeutsche Hauptstadt und in der Folge Sitz des Deutschen Bundestages, des Bundesrates sowie zahlreicher Bundesministerien.

Zu den bedeutenden Wirtschaftszweigen in Berlin gehören unter anderem die Kulturwirtschaft, die Biotechnologie und Gesundheitswirtschaft, die Informations- und Kommunikationstechnologien, die Bau- und Ingenieurwirtschaft, die Energietechnik sowie die Messe- und Kongresswirtschaft. Die Stadt ist ein wichtiger Luftverkehrsknotenpunkt. Berlin zählt zu den aufstrebenden, internationalen Metropolen und verzeichnet jährlich hohe Zuwachsraten bei der Zahl der Erwerbstätigen.

Berlin gilt als Weltstadt der Kultur, Politik, Medien und Wissenschaft. Es beherbergt zahlreiche Forschungseinrichtungen, Sportereignisse^[20] und Museen. Berlin ist seit 2020 ein UNESCO-Weltkulturerbe mit dem Titel *Stadt des Designs* und ist eines der meistbesuchten Zentren des Kontinents.^[21] Berlins Architektur, Festivals, Nachtleben und vielfältige Lebensbedingungen sind weltweit bekannt.^[22]



Wikidata:Q183

Englisch: Germany
country in central Europe
ist ein(e) — souveräner Staat
Gründung/Erstellung — 23. Mai 1949
ist Teil von — (en) Central Europe

Approve

Reject



Land Berlin



Landesflagge



Landeswappen



Basisdaten

Staatsform	parlamentarische Republik, teilsouveräner Gliedstaat eines Bundesstaates
Sprache	Deutsch ^[1]
Postleitzahlen	10115–14199
Vorwahl	030
Kfz-Kennzeichen	B
Bruttoinlandsprodukt (BIP)	129,424 Mrd. € (Februar 2017) ^[2]
BIP pro Kopf	35.627 € (Februar 2016) ^[3]
Schulden	56,935 Mrd. € (31. Dezember 2017) ^[4]
ISO 3166-2	DE-BE

Inhaltsverzeichnis [Verbergen]

1 Geographie

1.1 Klima

1.2 Gliederung

1.3 Parkanlagen, Zoos, Schutzgebiete

1.4 Nachbarstädte und -gemeinden

2 Geschichte

2.1 Namensbildung

2.2 Markgrafschaft und Kurfürstentum

2.3 Das 19. Jahrhundert



WIKIPEDIA
Die freie Enzyklopädie

Hauptseite

Themenportale

Von A bis Z

Zufälliger Artikel

Mitmachen

Artikel verbessern

Neuen Artikel anlegen

Autorenportal

Hilfe

Letzte Änderungen

Kontakt

Spenden

Werkzeuge

Links auf diese Seite

Änderungen an

verlinkten Seiten

Datei hochladen

Spezialseiten

Permanenter Link

Seiteninformationen

Wikidata-Datenobjekt

Artikel zitieren

Drucken/exportieren

Buch erstellen

Als PDF herunterladen

Druckversion

In anderen Projekten

Commons

Wikinews

Wikiquote

Wikisource

★ Wikivoyage

In anderen Sprachen

Frimelle Einstellungen Beta Beobachtungsliste Beiträge Abmelden

Artikel Diskussion

Lesen

Bearbeiten

Quelltext bearbeiten

Versionsgeschichte

☆

Mehr ▾

Wikipedia durchsuchen



Koordinaten: 52° 31′ N, 13° 24′ O﻿ / ﻿



Berlin



Dieser Artikel befasst sich mit der deutschen Hauptstadt. Zu weiteren Bedeutungen siehe [Berlin \(Begriffsklärung\)](#).

Berlin [bɛʁˈlin] ist die **Bundeshauptstadt** der **Bundesrepublik Deutschland**^[1] und zugleich eines ihrer Länder. Die Stadt Berlin ist mit fast 3,6 Millionen Einwohnern die bevölkerungsgrößte Gemeinde Deutschlands. Sie bildet das Zentrum von Berlin und der Agglomeration Berlin (4,5 Millionen Einw.). Der Stadtstaat und die Havel befinden sich im Stadtgebiet kleinere Fließgewässer sind die Spree und die Havel.

Urkundlich erstmals im 13. Jahrhundert erwähnt, war Berlin im 14. Jahrhundert die Residenz- und Hauptstadt Brandenburgs, Preußens und der Stadt Hauptstadt der Deutschen Demokratischen Republik. 1990 wurde Berlin wieder gesamtdeutsche Hauptstadt und in der Folge Sitz des Bundestages, des Bundesrates sowie zahlreicher Bundesministerien.

Zu den bedeutenden Wirtschaftszweigen in Berlin gehören unter anderem die Kulturwirtschaft, die Biotechnologie und Gesundheitswirtschaft sowie die Informations- und Kommunikationstechnologien, die Bau- und Ingenieurwirtschaft, die Energietechnik sowie die Messe- und Kongresswirtschaft. Die Stadt ist ein Zentrum des Luftverkehrs. Berlin zählt zu den aufstrebenden, internationalen Metropolen und verzeichnet jährlich hohe Zuwachsraten bei der Zahl der Erwerbstätigen.

Berlin gilt als Weltstadt der Kultur, Politik, Medien und Wissenschaft. Es beherbergt zahlreiche Forschungseinrichtungen, Sportereignisse^[20] und Museen. Berlin ist seit 2020 ein UNESCO-Weltkulturerbe mit dem Titel *Stadt des Designs* und ist eines der meistbesuchten Zentren des Kontinents.^[21] Berlins Architektur, Festivals, Nachtleben und vielfältige Lebensbedingungen sind weltweit bekannt.^[22]



Wikidata:Q183

Englisch: Germany
country in central Europe
ist ein(e) — souveräner Staat
Gründung/Erstellung — 23. Mai 1949
ist Teil von — (en) Central Europe

Approve

Reject



Einw.)
ipree

teil
wurde

lie
die
und

den

Land Berlin



Landesflagge



Landeswappen



Basisdaten

Staatsform	parlamentarische Republik, teilsouveräner Gliedstaat eines Bundesstaates
Sprache	Deutsch ^[1]
Postleitzahlen	10115–14199
Vorwahl	030
Kfz-Kennzeichen	B
Bruttoinlandsprodukt (BIP)	129,424 Mrd. € (Februar 2017) ^[2]
BIP pro Kopf	35.627 € (Februar 2016) ^[3]
Schulden	56,935 Mrd. € (31. Dezember 2017) ^[4]
ISO 3166-2	DE-BE

Inhaltsverzeichnis [Verbergen]

- Geographie
 - Klima
 - Gliederung
 - Parkanlagen, Zoos, Schutzgebiete
 - Nachbarstädte und -gemeinden
- Geschichte
 - Namensbildung
 - Markgrafschaft und Kurfürstentum
 - Das 19. Jahrhundert



WIKIPEDIA
Die freie Enzyklopädie

Hauptseite

Themenportale

Von A bis Z

Zufälliger Artikel

Mitmachen

Artikel verbessern

Neuen Artikel anlegen

Autorenportal

Hilfe

Letzte Änderungen

Kontakt

Spenden

Werkzeuge

Links auf diese Seite

Änderungen an

verlinkten Seiten

Datei hochladen

Spezialseiten

Permanenter Link

Seiteninformationen

Wikidata-Datenobjekt

Artikel zitieren

Drucken/exportieren

Buch erstellen

Als PDF herunterladen

Druckversion

In anderen Projekten

Commons

Wikinews

Wikiquote

Wikisource

★ Wikivoyage

In anderen Sprachen

Frimelle Einstellungen Diskussion Beta Beobachtungsliste Beiträge Abmelden

Artikel Diskussion

Lesen

Bearbeiten

Quelltext bearbeiten

Versionsgeschichte



Mehr ▾

Wikipedia durchsuchen



Koordinaten: 52° 31′ N, 13° 24′ O﻿ / ﻿



Berlin



Dieser Artikel befasst sich mit der deutschen Hauptstadt. Zu weiteren Bedeutungen siehe [Berlin \(Begriffsklärung\)](#).

Berlin [bəˈʁɪn][ⓘ] ist die **Bundeshauptstadt** der **Bundesrepublik Deutschland**^[1] und zugleich eines ihrer Länder. Die Stadt Berlin ist mit fast 3,6 Millionen Einwohnern die **bevölkerungsgrößte Gemeinde** Deutschlands. Sie bildet das Zentrum und der Agglomeration Berlin (4,5 Millionen Einw.). Der Stadtstaat und die Havel befinden sich im Stadtgebiet kleinere Fließgewässer s

Urkundlich erstmals im 13. Jahrhundert erwähnt, war Berlin im 19. Jahrhundert die Hauptstadt des Königreichs Preußen. Im 20. Jahrhundert wurde Berlin wieder gesamtdeutsche Hauptstadt und in der Folge Sitz des Deutschen Bundestages, des Bundesrates sowie zahlreicher Bundesministerien.

Zu den bedeutenden Wirtschaftszweigen in Berlin gehören unter anderem die **Kulturwirtschaft**, die **Biotechnologie** und **Gesundheitswirtschaft** sowie die **Informations- und Kommunikationstechnologien**, die **Bau- und Immobilienwirtschaft**, die **Energietechnik** sowie die **Messe- und Kongresswirtschaft**. Die Stadt ist ein wichtiger Luftverkehrsknotenpunkt. Berlin zählt zu den aufstrebenden, internationalen Metropolen und verzeichnet jährlich hohe Zuwachsraten bei der Zahl der Erwerbstätigen.

Berlin gilt als **Weltstadt** der Kultur, Politik, Medien und Wissenschaft. Die Stadt ist Sitz zahlreicher **Forschungseinrichtungen**, Sportereignisse^[20] und Museen. Berlin ist seit 2020 ein **UNESCO-Weltkulturerbe** mit dem Titel *Stadt des Designs* und ist eines der meistbesuchten Zentren des Kontinents.^[21] Berlins Architektur, Festivals, Nachtleben und vielfältige Lebensbedingungen sind weltweit bekannt.^[22]



Wikidata:Q183

Englisch: Germany
country in central Europe
ist ein(e) — souveräner Staat
Gründung/Erstellung — 23. Mai 1949
ist Teil von — (en) Central Europe

Approve

Reject



Inhaltsverzeichnis [Verbergen]

1 Geographie

1.1 Klima

1.2 Gliederung

1.3 Parkanlagen, Zoos, Schutzgebiete

1.4 Nachbarstädte und -gemeinden

2 Geschichte

2.1 Namensbildung

2.2 Markgrafschaft und Kurfürstentum

2.3 Das 19. Jahrhundert

Land Berlin



Landesflagge



Landeswappen



Basisdaten

Staatsform	parlamentarische Republik, teilsouveräner Gliedstaat eines Bundesstaates
Sprache	Deutsch ^[1]
Postleitzahlen	10115–14199
Vorwahl	030
Kfz-Kennzeichen	B
Bruttoinlandsprodukt (BIP)	129,424 Mrd. € (Februar 2017) ^[2]
BIP pro Kopf	35.627 € (Februar 2016) ^[3]
Schulden	56,935 Mrd. € (31. Dezember 2017) ^[4]
ISO 3166-2	DE-BE



Aims for the stay at UPM

- Find possibilities for collaboration
- Discuss potential and drawbacks of collaborative KBs for ontology design
- Representing linguistic data in Wikida

- *Mind the (Language) Gap: Generation of Multilingual Wikipedia Summaries from Wikidata for ArticlePlaceholders*, Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest, Jonathon Hare, Elena Simperl, ESWC 2018
- *Learning to Generate Wikipedia Summaries for Underserved Languages from Wikidata*, Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest, Jonathon Hare, Elena Simperl, NAACL 2018
- *Property Label Stability in Wikidata*, Thomas Pellissier Tanon, Lucie-Aimée Kaffee, Wiki Workshop at The Web Conference 2018
- *A Glimpse into Babel: An Analysis of Multilinguality in Wikidata*, Lucie-Aimée Kaffee, Alessandro Piscopo, Pavlos Vougiouklis, Elena Simperl, Leslie Carr, Lydia Pintscher, OpenSym 2017
- *Provenance Information in a Collaborative Knowledge Graph: an Evaluation of Wikidata External References*, Alessandro Piscopo, Lucie-Aimée Kaffee, Chris Phethean, Elena Simperl, ISWC 2017
- *What do Wikidata and Wikipedia Have in Common?: An Analysis of their Use of External References*, Alessandro Piscopo, Pavlos Vougiouklis, Lucie-Aimée Kaffee, Christopher Phethean, Jonathon Hare, Elena Simperl, OpenSym 2017
- *Neural Wikipedian: Generating Textual Summaries from Knowledge Base Triples*, Pavlos Vougiouklis, Hady Elsahar, Lucie-Aimée Kaffee, Christoph Gravier, Frederique Laforest, Jonathon Hare, Elena Simperl, under submission



Thanks!

kaffee@soton.ac.uk
luciekaffee.github.io

Mind the (Language) Gap: Generation of Multilingual Wikipedia Summaries from Wikidata for ArticlePlaceholders, Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest, Jonathon Hare, Elena Simperl, ESWC 2018

Learning to Generate Wikipedia Summaries for Underserved Languages from Wikidata, Lucie-Aimée Kaffee, Hady Elsahar, Pavlos Vougiouklis, Christophe Gravier, Frédérique Laforest, Jonathon Hare, Elena Simperl, NAACL 2018

Property Label Stability in Wikidata, Thomas Pellissier Tanon, Lucie-Aimée Kaffee, Wiki Workshop at The Web Conference 2018

A Glimpse into Babel: An Analysis of Multilinguality in Wikidata, Lucie-Aimée Kaffee, Alessandro Piscopo, Pavlos Vougiouklis, Elena Simperl, Leslie Carr, Lydia Pintscher, OpenSym 2017