Unflattening Knowledge Graphs

Marieke van Erp
DHLab
KNAW Humanities Cluster
Amsterdam, the Netherlands

merpeltje

Let’s talk about coffee

Let’s talk about coffee

ESWC Conferences @eswc_conf · May 28

Did you sign yours yet? ✏️

#ESWC2023 attendees are offered ESWC mugs. Personalise it to remember the 20th edition of ESWC

Let’s talk about coffee

Let’s talk about coffee

“A cup of coffee”

Let’s talk about coffee

“A cup of coffee”

Let’s talk about coffee

“A cup of coffee”

Let’s talk about coffee

“A cup of coffee”

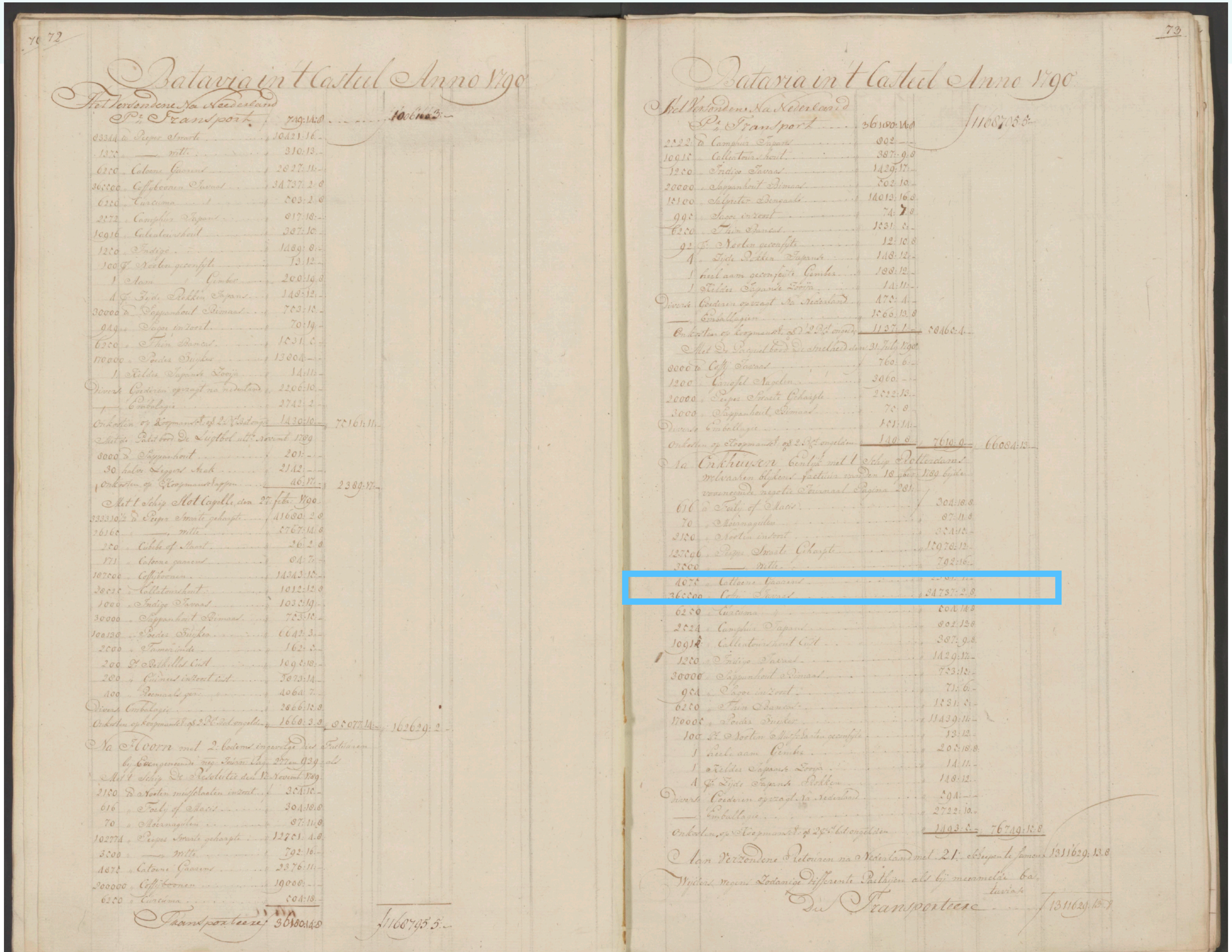
Let’s talk about coffee

“Let’s have coffee”

Let’s talk about coffee

17th Century Coffee House

Let’s talk about coffee



70 72

Batavia in t Castiel Anno 1790

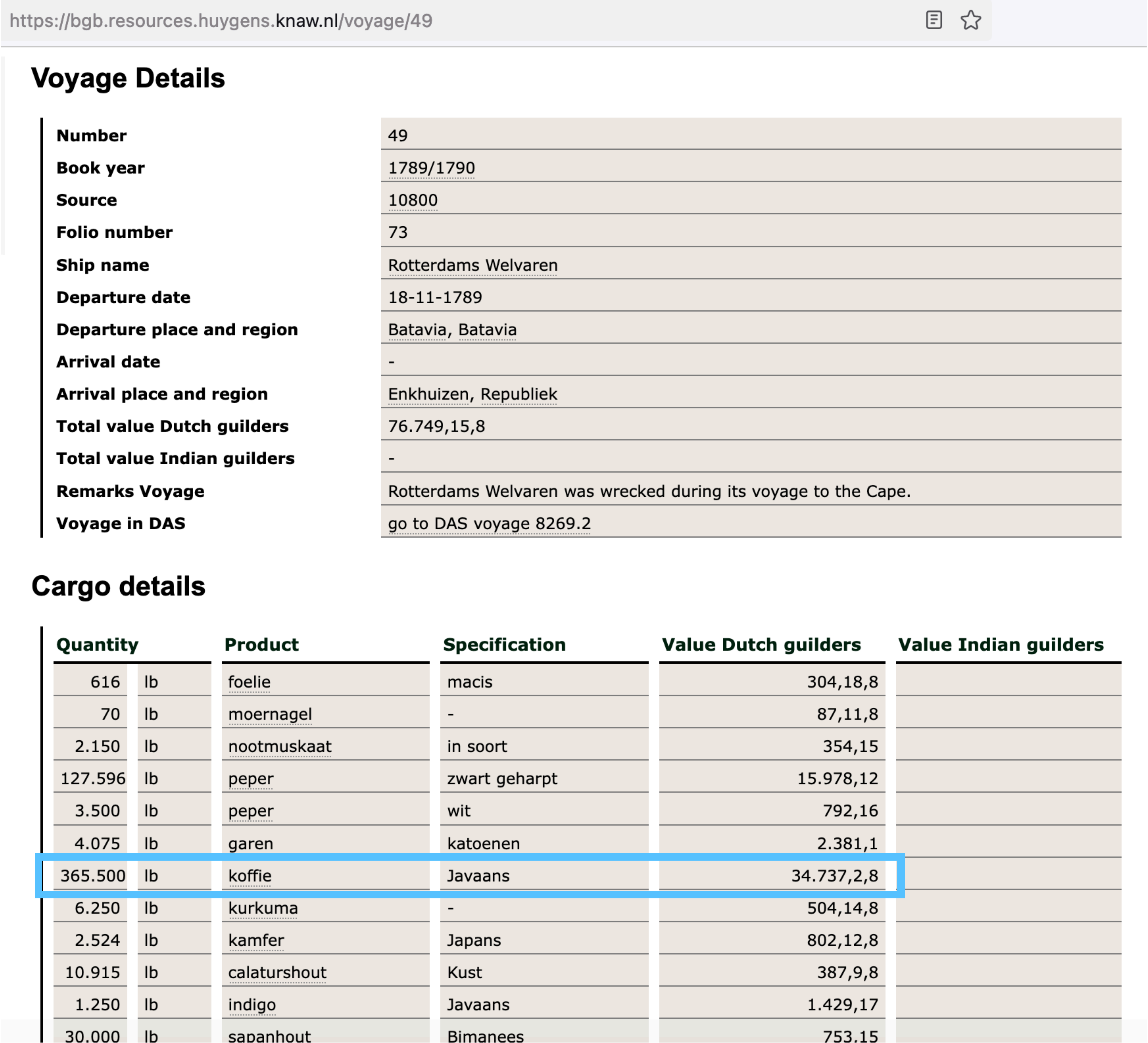
	Het Verzondene Na Nederland P. Transport	749:14:8		1006663:-
83344	₠ Peeper Swarte	10421:16.-		
1372	— witte	310:13.-		
6250	Catoene Gaarens	2827:11.-		
365500	Coffyboonen Javaat	34737:2:8		
6250	Curcuma	503:2:8		
2572	Camphur Japans	817:18:-		
10916	Calcatourshout	387:10:-		
1250	Indigo	1489:8:-		
100	F. Nooten geconfyte	13:12.-		
1	Aam Gember	200:19:8		
4	F. Zyde Rokken Japans	148:12:-		
30000	₠ Sappanhout Bimaas	753:15.-		
949	Sagoe in zoort	70:19.-		
6250	Thin Bancas	1531:5.-		
170000	Poeder Suyker	13004:-.-		
1	Kelder Japanse Zooja	14:11:-		
Diverse	Goederen opzagt na nederland	2206:10.-		
—	Embalagie	2742:2.-		
	Onkosten op Koopmansch. a 2 Pt Batong.	1430:10.-	75161:11.-	
	Met de Paketboot De Zugtbol ult. Novemb. 1789			
8000	₠ Sappanhout	201:-.-		
30	halve Leggers Arak	2142:-.-		
	Onkosten op Koopmanschappen	46:17.-	2389:17.-	
	Met t Schip Slot Capelle den 27: febr: 1790			
333310½	₠ Peeper Swarte gehaspte	41680:2:8		
26160	— witte	5767:14:8		
250	Cubebe of Staart	26:2:8		
171	Catoene garens	84:7:-		
187500	Coffyboonen	14343:15:-		
28525	Callatourshout	1012:12:8		
1000	Indigo Javaas	1030:19:-		
30000	Sappanhout Bimaas	753:15:-		
100138	Poeder Suyker	6642:3.-		
2500	Tamerinde	162:5.-		
200	P. Bethilles Cust	1095:18:-		
280	Guinees inzoort Cust	5073:14.-		
100	Roemaals ger.	4064:7.-		
Diverse	Embalagie	2866:15:8		
	Onkosten op koopmansch. a 2 Pt. Bat. ongelden	1660:3:8	95077:14:-	162629:2.-
	Na Hoorn met 2: bodems ingevolge dies Factuuren by Evengenoemde neg. Journ. Pag. 277 en 939 als			
	Met t Schip De Resolutie den 12. Novemb. 1789.			
2150	₠ Nooten muskaaten inzoort	354:15.-		
616	Foely of Macis	304:18:8		
70	Moernagulen	87:11:8		
102774	Peeper Swarte gehaspte	12751:4:8		
3500	— witte	792:16.-		
4875	Catoene Gaarens	2376:11:-		
200000	Coffyboonen	19000:-.-		
6250	Curcuma	504:18.-		
	Transporteere	36180:14:8		f1168795.5.-

73

Batavia in t Casteel Anno 1790

	Het Verzondene Na Nederland P. Transport	36180:14:8		f1168795:5:-
2522	₠ Camphur Japans	802:-.-		
10915	Calleatourshout	387:9:8		
1250	Indigo Javaas	1429:17.-		
20000	Sappanhout Bimaas	502:10.-		
15100	Salpeter Bengaals	14013:16:8		
995	Sagoe inzoort	74:7:8		
6250	Thin Bancas	1531:5.-		
92	P. Nooten geconfyte	12:10:8		
4	Zyde Rokken Japanse	148:12.-		
1	heel aam geconfeyte Gember	188:12.-		
1	Kelder Japanse Zooja	14:11:-		
Diverse	Goederen opzagt Na Nederland	475:4.-		
—	Emballagien	1566:13:8		
	Onkosten op koopmansch. a 2 Pct ongeld.	1137:1.-	58465:4.-	
	Met De Pacquet boot De Snelheid den 31: July 1790			
8000	₠ Coffij Javaas	760:6:-		
1200	Garioffel Nagelen	3960:-.-		
20000	Peeper Swarte Gehaspte	2522:13.-		
3000	Sappanhout Bimaas	75:8.-		
Diverse	Emballagie	151:14.-		
	Onkosten op Koopmansch. a 2 Pct ongelden	149:8.-	7619:9.-	66084:13.-
	Na Enkhuysen Eentyk met t Schip Rotterdams welvaaren blykens factuur van den 18 9ber 1789 byde voornoemde negotie Journaal Pagina 281.			
616	₠ Foely of Macis		304:18:8	
70	Moernagulen		87:11:8	
2150	Nooten inzoort		354:15.-	
127596	Peeper Swarte Gehaspte		15970:12.-	
3500	— witte		792:16.-	
4875	Cattoene Gaarens		2381:11.-	
365500	Coffij Javaas		34737:2:8	
6250	Curcuma		504:14:8	
2524	Camphur Japans		802:12:8	
10912	Calleatourshout Cust		387:9:8	
1250	Indigo Javaas		1429:17.-	
30000	Sappanhout Bimaas		753:15.-	
954	Sagoe in Zoort		71:6.-	
6250	Thin Bancas		1531:5.-	
170000	Poeder Suyker		11439:11.-	
100	P. Nooten Muscaaten geconfyte		13:12.-	
1	heele aam Gember		205:18:8	
1	Kelder Japanse Zooja		14:11.-	
4	P. Zyde Japanse Rokken		148:12.-	
Diverse	Goederen opzagt Na Nederland		594:-.-	
—	Emballagie		2722:10.-	
	Onkosten op Koopmansch. a 2 Pct bat. ongelden		1493:5.-	76749:15:8
	Aan Verzondene Retouren na Nederland met 21: Scheepen te samen			1311629:13:8
	Wyders wegens Zodanige differente Parthyen als bij meermelde bataviase			
	De Transporteere			f1311629:13:8

Let’s talk about coffee



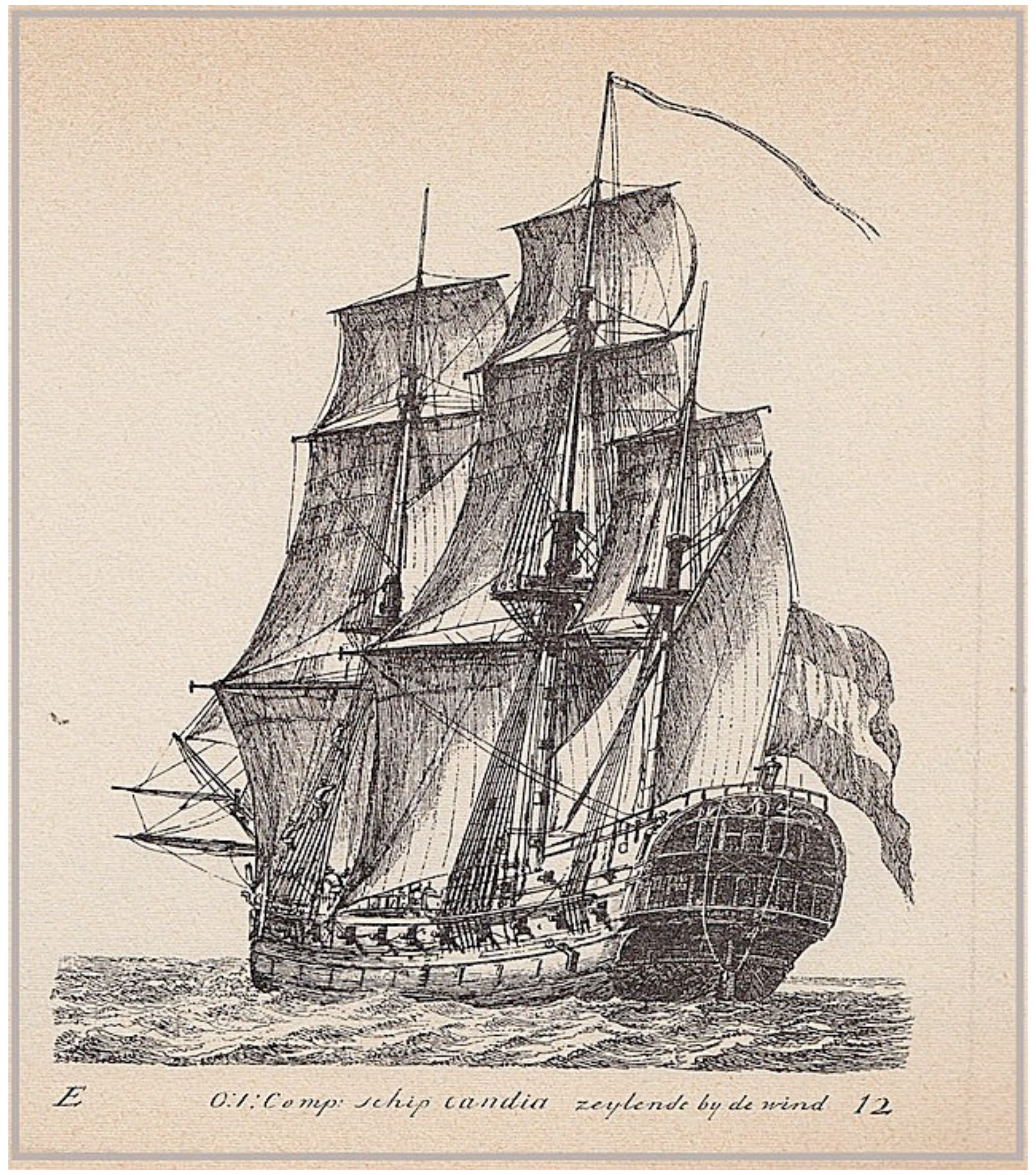
https://bgb.resources.huygens.knaw.nl/voyage/49

Voyage Details

Number	49
Book year	1789/1790
Source	10800
Folio number	73
Ship name	Rotterdams Welvaren
Departure date	18-11-1789
Departure place and region	Batavia, Batavia
Arrival date	-
Arrival place and region	Enkhuizen, Republiek
Total value Dutch guilders	76.749,15,8
Total value Indian guilders	-
Remarks Voyage	Rotterdams Welvaren was wrecked during its voyage to the Cape.
Voyage in DAS	go to DAS voyage 8269.2

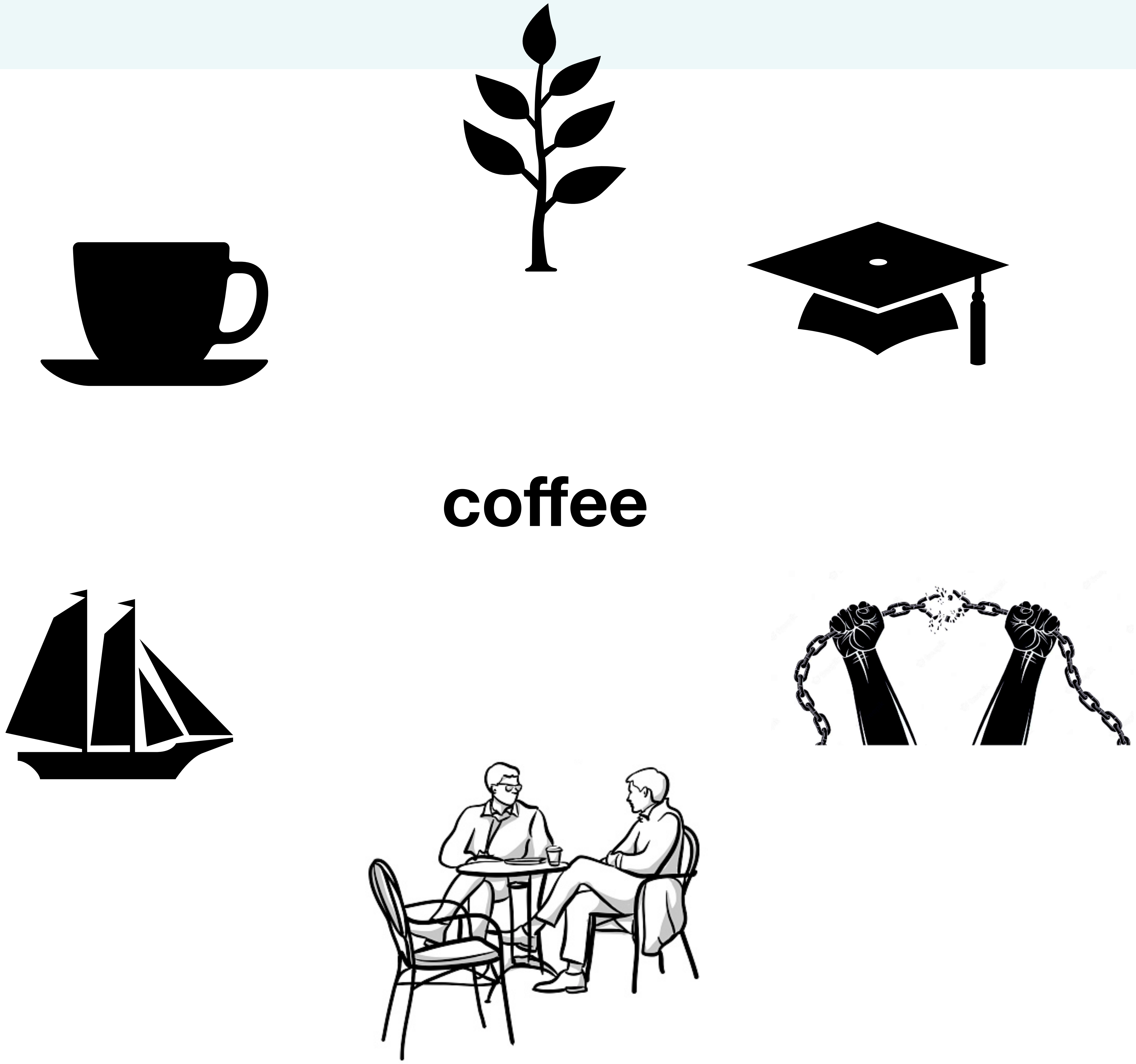
Cargo details

Quantity		Product	Specification	Value Dutch guilders	Value Indian guilders
616	lb	foelie	macis	304,18,8	
70	lb	moernagel	-	87,11,8	
2.150	lb	nootmuskaat	in soort	354,15	
127.596	lb	peper	zwart geharpt	15.978,12	
3.500	lb	peper	wit	792,16	
4.075	lb	garen	katoenen	2.381,1	
365.500	lb	koffie	Javaans	34.737,2,8	
6.250	lb	kurkuma	-	504,14,8	
2.524	lb	kamfer	Japans	802,12,8	
10.915	lb	calatursthout	Kust	387,9,8	
1.250	lb	indigo	Javaans	1.429,17	
30.000	lb	sapanhout	Bimanees	753,15	



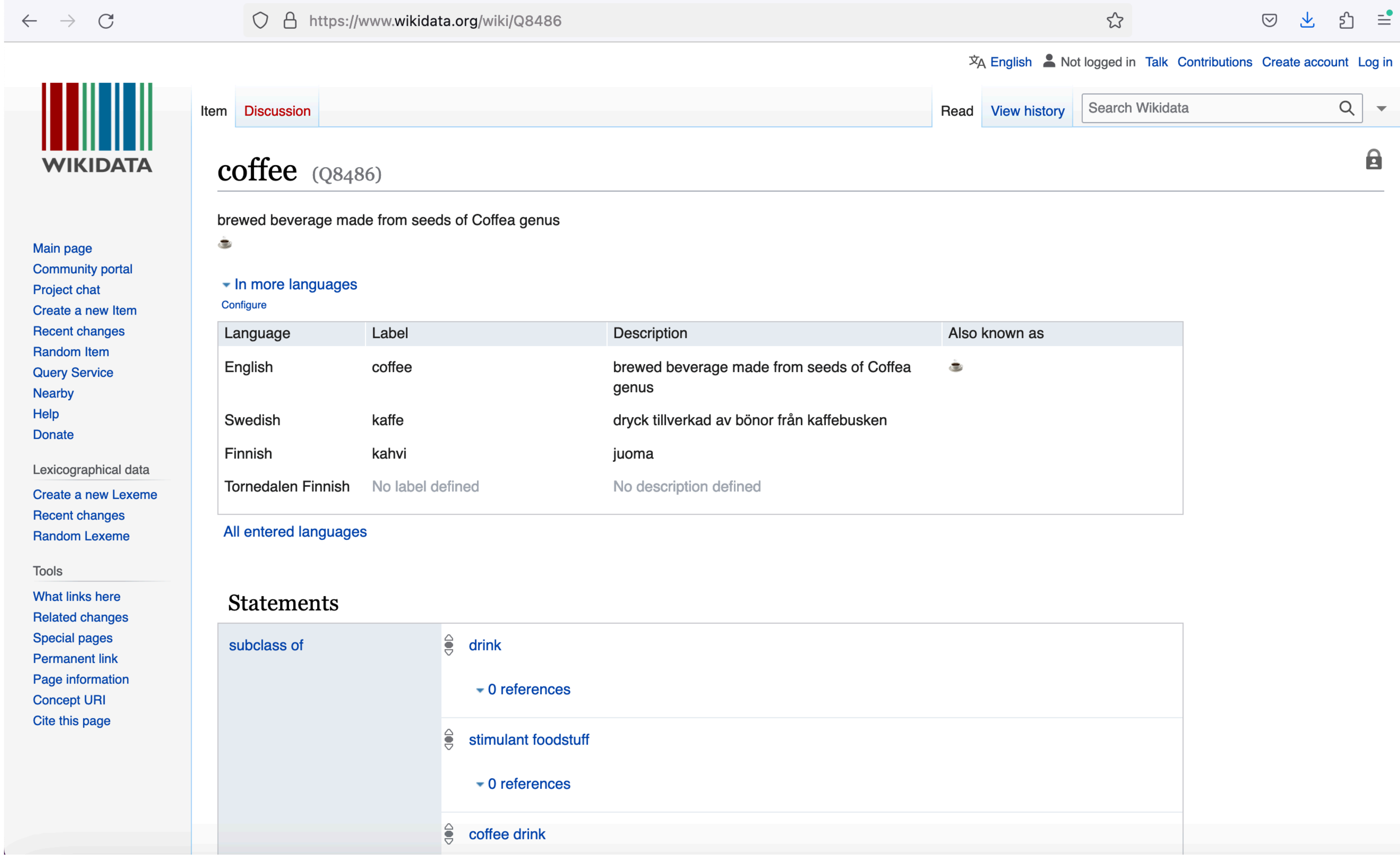
Let’s talk about coffee

PICKING COFFEE.



coffee

Querying for coffee



https://www.wikidata.org/wiki/Q8486

English Not logged in Talk Contributions Create account Log in

Item Discussion Read View history Search Wikidata

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

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

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

coffee (Q8486)

brewed beverage made from seeds of Coffea genus
☕

In more languages
Configure

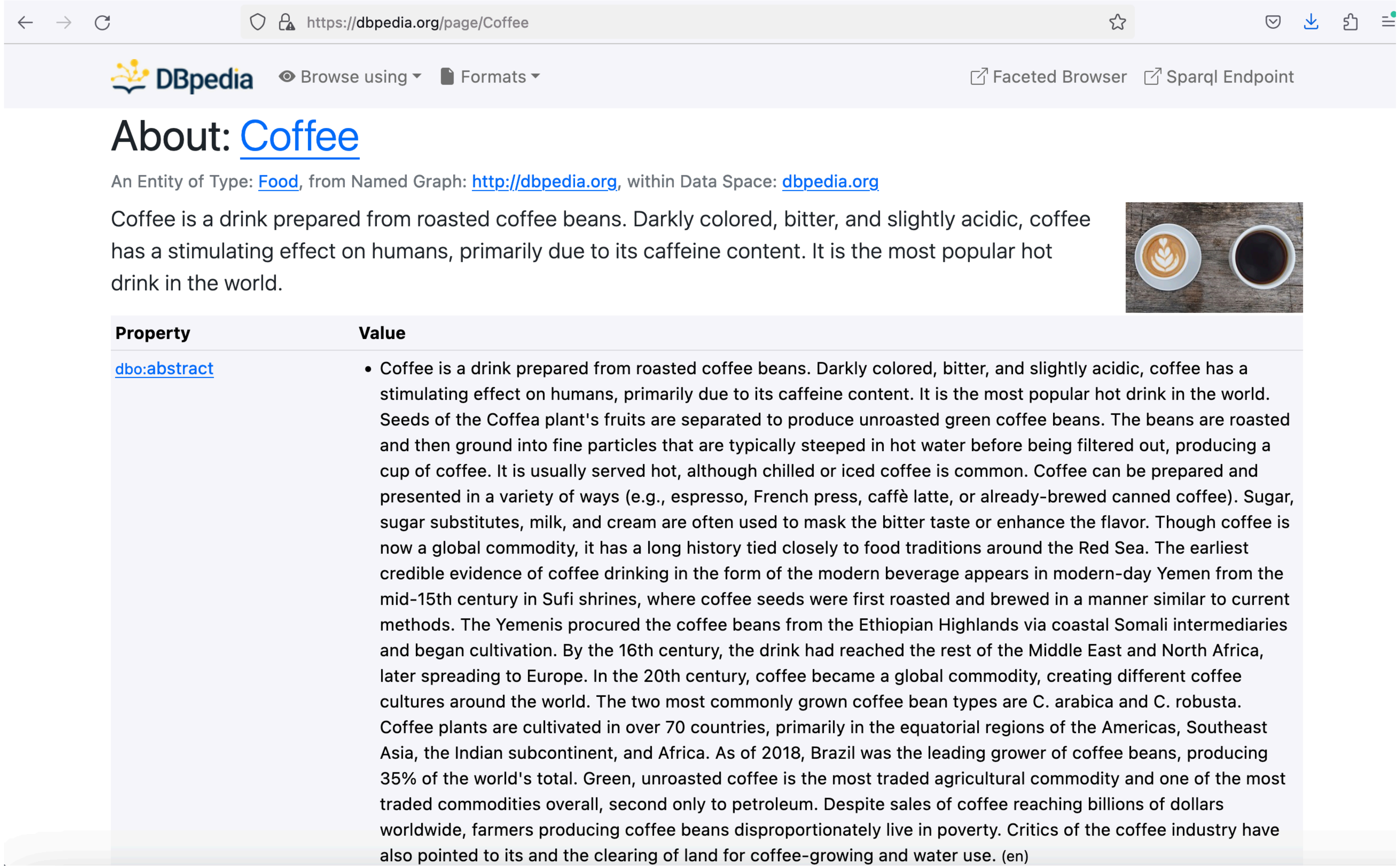
Language	Label	Description	Also known as
English	coffee	brewed beverage made from seeds of Coffea genus	☕
Swedish	kaffe	dryck tillverkad av bönor från kaffebusken	
Finnish	kahvi	juoma	
Tornedalen Finnish	No label defined	No description defined	

All entered languages

Statements

subclass of	drink 0 references stimulant foodstuff 0 references coffee drink

Querying for coffee



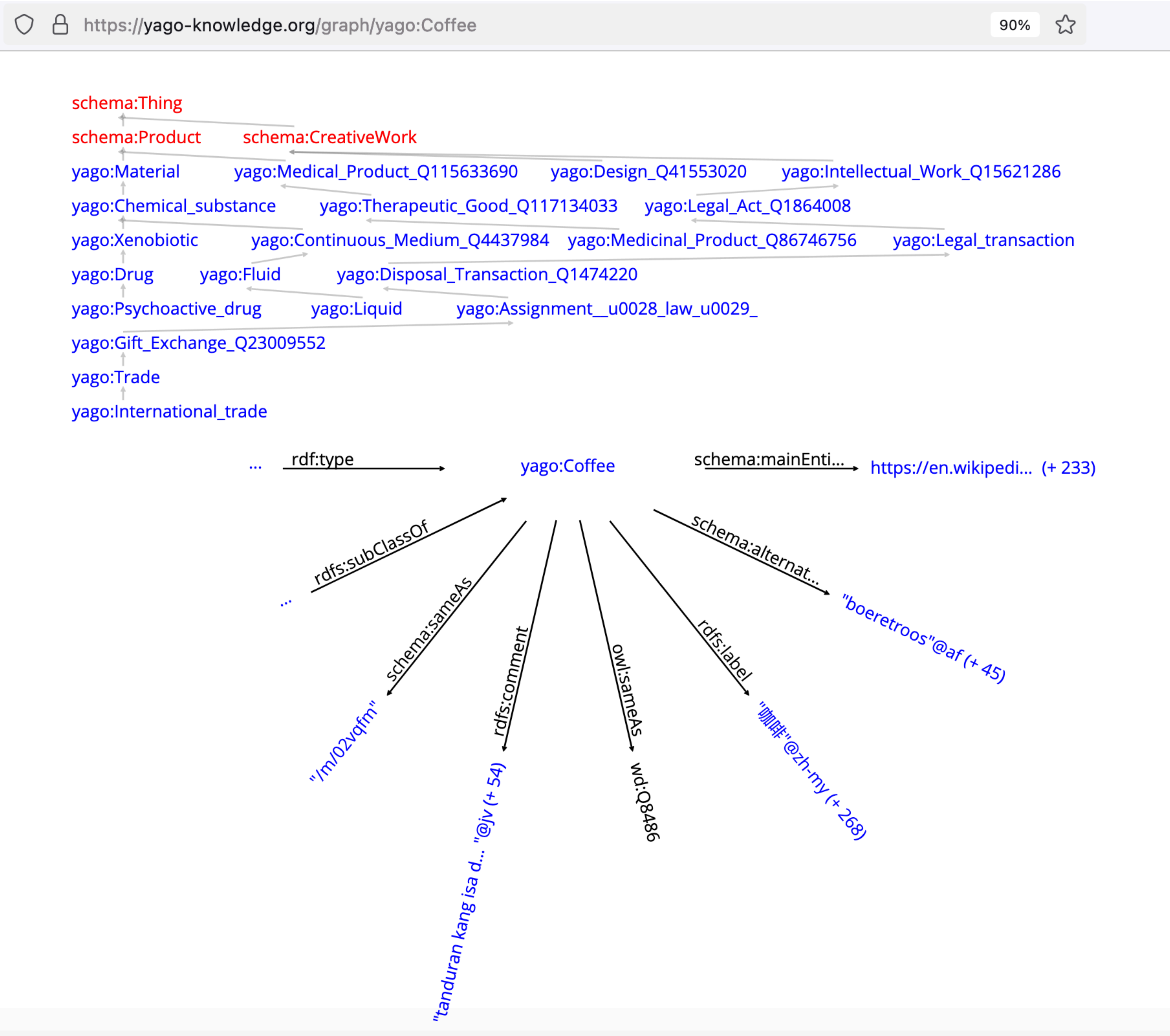
About: Coffee

An Entity of Type: Food, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org

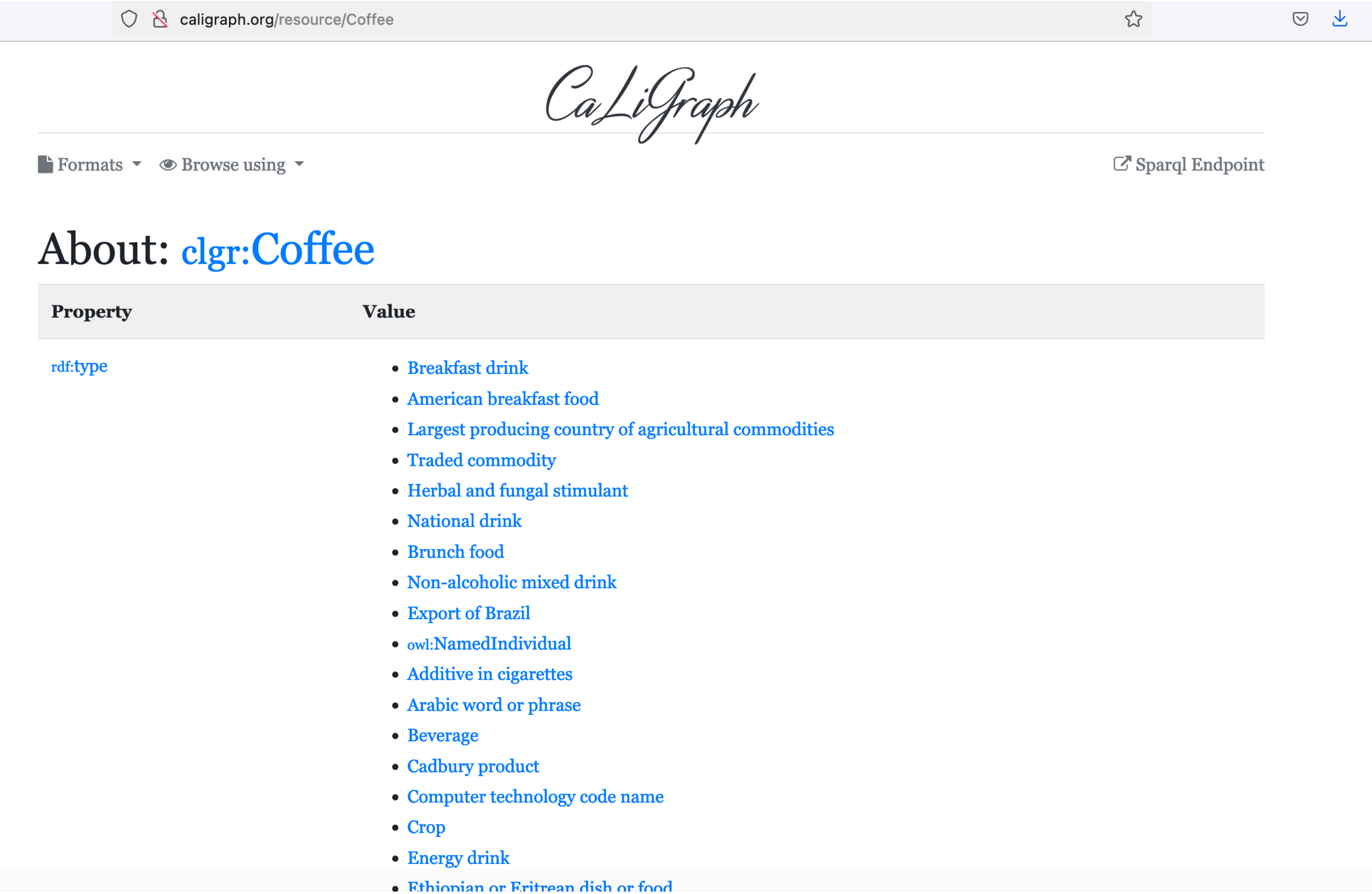
Coffee is a drink prepared from roasted coffee beans. Darkly colored, bitter, and slightly acidic, coffee has a stimulating effect on humans, primarily due to its caffeine content. It is the most popular hot drink in the world.

Property	Value
dbo:abstract	• Coffee is a drink prepared from roasted coffee beans. Darkly colored, bitter, and slightly acidic, coffee has a stimulating effect on humans, primarily due to its caffeine content. It is the most popular hot drink in the world. Seeds of the Coffea plant's fruits are separated to produce unroasted green coffee beans. The beans are roasted and then ground into fine particles that are typically steeped in hot water before being filtered out, producing a cup of coffee. It is usually served hot, although chilled or iced coffee is common. Coffee can be prepared and presented in a variety of ways (e.g., espresso, French press, caffè latte, or already-brewed canned coffee). Sugar, sugar substitutes, milk, and cream are often used to mask the bitter taste or enhance the flavor. Though coffee is now a global commodity, it has a long history tied closely to food traditions around the Red Sea. The earliest credible evidence of coffee drinking in the form of the modern beverage appears in modern-day Yemen from the mid-15th century in Sufi shrines, where coffee seeds were first roasted and brewed in a manner similar to current methods. The Yemenis procured the coffee beans from the Ethiopian Highlands via coastal Somali intermediaries and began cultivation. By the 16th century, the drink had reached the rest of the Middle East and North Africa, later spreading to Europe. In the 20th century, coffee became a global commodity, creating different coffee cultures around the world. The two most commonly grown coffee bean types are C. arabica and C. robusta. Coffee plants are cultivated in over 70 countries, primarily in the equatorial regions of the Americas, Southeast Asia, the Indian subcontinent, and Africa. As of 2018, Brazil was the leading grower of coffee beans, producing 35% of the world's total. Green, unroasted coffee is the most traded agricultural commodity and one of the most traded commodities overall, second only to petroleum. Despite sales of coffee reaching billions of dollars worldwide, farmers producing coffee beans disproportionately live in poverty. Critics of the coffee industry have also pointed to its and the clearing of land for coffee-growing and water use. (en)

Querying for coffee



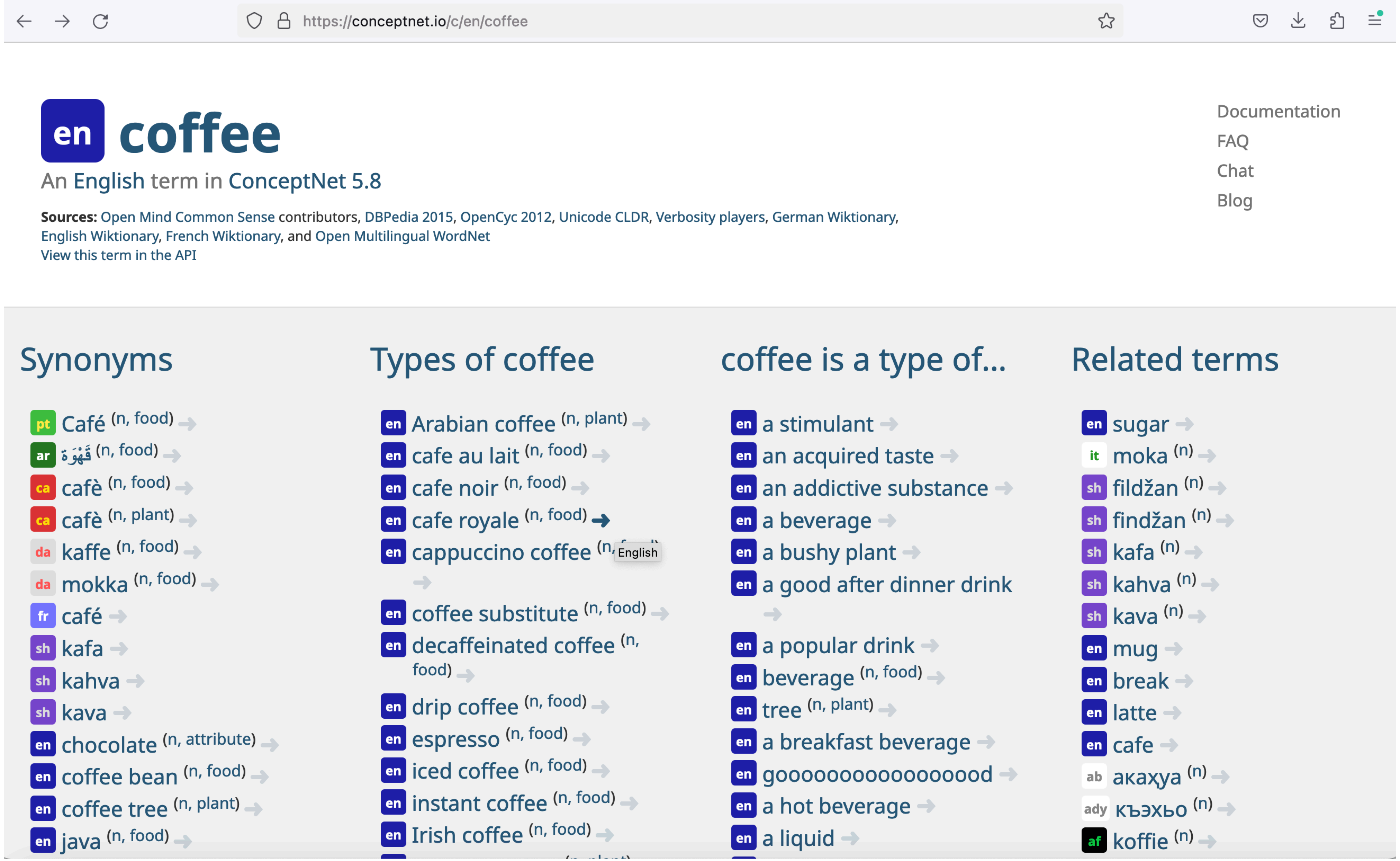
Querying for Coffee



About: clgr:Coffee

Property	Value
rdf:type	• Breakfast drink • American breakfast food • Largest producing country of agricultural commodities • Traded commodity • Herbal and fungal stimulant • National drink • Brunch food • Non-alcoholic mixed drink • Export of Brazil • owl:NamedIndividual • Additive in cigarettes • Arabic word or phrase • Beverage • Cadbury product • Computer technology code name • Crop • Energy drink • Ethiopian or Eritrean dish or food

Querying for coffee



https://conceptnet.io/c/en/coffee

en **coffee**

An English term in ConceptNet 5.8

Sources: Open Mind Common Sense contributors, DBPedia 2015, OpenCyc 2012, Unicode CLDR, Verbosity players, German Wiktionary, English Wiktionary, French Wiktionary, and Open Multilingual WordNet

View this term in the API

Documentation
FAQ
Chat
Blog

Synonyms

- pt Café (n, food)
- ar قَهْوَة (n, food)
- ca cafè (n, food)
- ca cafè (n, plant)
- da kaffe (n, food)
- da mokka (n, food)
- fr café
- sh kafa
- sh kahva
- sh kava
- en chocolate (n, attribute)
- en coffee bean (n, food)
- en coffee tree (n, plant)
- en java (n, food)

Types of coffee

- en Arabian coffee (n, plant)
- en cafe au lait (n, food)
- en cafe noir (n, food)
- en cafe royale (n, food)
- en cappuccino coffee (n, English
- en coffee substitute (n, food)
- en decaffeinated coffee (n, food)
- en drip coffee (n, food)
- en espresso (n, food)
- en iced coffee (n, food)
- en instant coffee (n, food)
- en Irish coffee (n, food)

coffee is a type of...

- en a stimulant
- en an acquired taste
- en an addictive substance
- en a beverage
- en a bushy plant
- en a good after dinner drink
- en a popular drink
- en beverage (n, food)
- en tree (n, plant)
- en a breakfast beverage
- en gooooooooooooooood
- en a hot beverage
- en a liquid

Related terms

- en sugar
- it moka (n)
- sh fildžan (n)
- sh findžan (n)
- sh kafa (n)
- sh kahva (n)
- sh kava (n)
- en mug
- en break
- en latte
- en cafe
- ab акаҳуа (n)
- ady къэхьо (n)
- af koffie (n)

Why is this a problem

- Entity Linking is often employed in KG creation
- Imprecise links can create imprecise analyses
- Not only an issue for historical use cases

Car Manufacturers

- April 2006:
  - production of Polo from Spain to Eastern-Europe because of social problems in Volkswagen - Pamplona and maybe to Volkswagen -Vorst in Belgium
- July 2006:
  - Polo production in Vorst, no jobs lost in Spain but extra jobs in Belgium.
- August 2006:
  - Fewer Golfs produced in Vorst, maybe more Polos. ‘If not, we have a problem’, says a union representative.....Chances that Vorst will not make any Polos next year are minimal, because the factory invested this year in a special new welding installation specific for Polo cars.
- November 2006:
  - Volkswagen stops the production of Golf in Vorst: 3,500 jobs are lost plant renamed to Audi-Brussels
- November 2009:
  - Audi plant in Vorst stops the production of Polo: 300 jobs lost

Audi-Brussels present in DBpedia
Volkswagen Pamplona linked to Volkswagen
Volkswagen closes Volkswagen Pamplona ≠ dbp:Volkswagen closes dbp:Volkswagen

Marieke van Erp, Filip Ilievski, Marco Rospocher, and Piek Vossen. "Missing Mr. Brown and Buying an Abraham Lincoln-Dark Entities and DBpedia." In NLP-DBPEDIA@ ISWC, pp. 81-86. 2015.

Commodities

Sugar is an energy source

Sugar is a treat

Sugar caused ecological crises

Sugar enabled enslavement

Sugar is a health hazard

Ulbe Bosma (2023) The World of Sugar, Harvard University Press.

Smells

Jewellery is a status symbol

Jewellery can convey a message

Scented jewellery can ward off disease

Pasquale Lisena, Daniel Schwabe, Marieke van Erp, Raphaël Troncy, William Tullett, Inger Leemans, Lizzie Marx, and Sofia Colette Ehrich. "Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information." In *The Semantic Web: 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29–June 2, 2022, Proceedings*, pp. 387-405. Cham: Springer International Publishing, 2022.

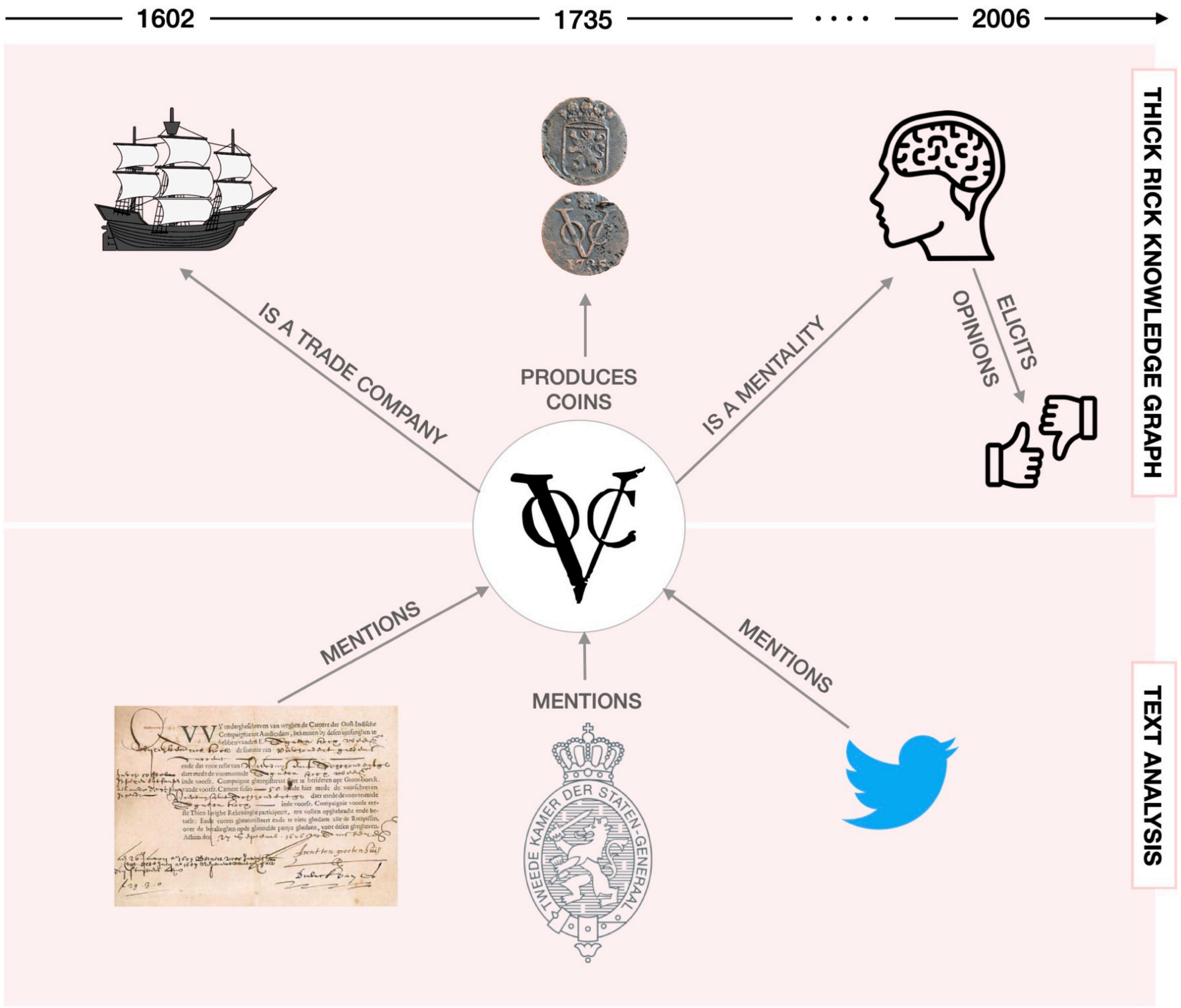
Unflattening KGs

- Identity - concepts are multidimensional and have ***coreferring***, ***non***- and ***near***-identity relations to other concepts

- Change - concepts evolve over time

- Long tail - go beyond popular concepts and entities

- Provenance - keep track of where the information comes from



Identity, non-identity & near-identity

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

*Bill said **Alice** would arrive soon, and **she** did.*

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

*Bill said **Alice** would arrive soon, and **she** did.*

- Non-identity: when two entities are distinctly different

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

*Bill said **Alice** would arrive soon, and **she** did.*

- Non-identity: when two entities are distinctly different

*Bill said **Alice** would arrive soon, then **Jane** arrived.*

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

*Bill said **Alice** would arrive soon, and **she** did.*

- Non-identity: when two entities are distinctly different

*Bill said **Alice** would arrive soon, then **Jane** arrived.*

- Near-identity: when two entities share most but not all feature values

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity, non-identity & near-identity

- Co-referring relationship: when two entities are the same (owl:sameAs)

*Bill said **Alice** would arrive soon, and **she** did.*

- Non-identity: when two entities are distinctly different

*Bill said **Alice** would arrive soon, then **Jane** arrived.*

- Near-identity: when two entities share most but not all feature values

***The United States** has officially restored diplomatic relations with Yugoslavia . . .*
***The White House** said the United States will provide 45 million dollars in food aid.*

Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.

Identity @ESWC2023

Refining Large Integrated Identity Graphs using the Unique Name Assumption

Shuai Wang[1] ✉ [0000-0002-1261-9930], Joe Raad[2] [0000-000
Bloem[1] [0000-0002-0189-5817], and Frank van Harmelen[1] [0

[1] Department of Computer Science, Vrije Universiteit Amster
{shuai.wang | p.bloem | frank.van.harmelen
[2] LISN, University of Paris-Saclay, Orsay, Fr
joe.raad@lisn.fr

Abstract. The Unique Name Assumption (UNA) su
terms with distinct identif

SAP-KG: Synonym Predicate Analyzer across Multiple Knowledge Graphs

azmand[1,2] [0000-0001-8194-8079] and Maria-Esther
Vidal[1,2,3] [0000-0003-1160-8727]

ation Centre for Science and Technology, Hannover, Germany
eibniz University of Hannover, Germany
Emetis.Niazmand@tib.eu
[3] L3S Research Center, Germany
Maria.Vidal@tib.eu

mo paper presents SAP-KG, a knowledge graph ag-
trate the benefits of identifying the synonym predi-
omplementary information; they are used for query
e query answer completeness. SAP-KG proposed a
he percentage of overlap between pairs of synonym
and capture the

DBpe-
nature
knowl-
modify

Transformer based Semantic Relation Typing for Knowledge Graph Integration

Sven Hertling[0000-0003-0333-5888] and Heiko Paulheim[0000-0003-4386-8195]

Data and Web Science Group, University of Mannheim, Germany
{sven,heiko}@informatik.uni-mannheim.de

More and m
ains. Applic
of those
or ontolog
tween clas
ass, etc. In
or **Seman**
n holds
ween equ
nd no r
to refin
es. The
graphs
can be
how th
solve

Entity Typing with Triples using Language Models

Aniqa Riaz[1], Sara Abdollahi[2] [0000-0001-7752-146X], and Simon Gottschalk[2] [0000-0003-2576-4640]

[1] Universität Bonn, Germany
s6anriaz@uni-bonn.de
[2] L3S Research Center, Leibniz Universität Hannover, Germany
{abdollahi,gottschalk}@L3S.de

Abstract. Entity Typing is the task of assigning a type to an entity in a
In this paper, we propose ETwT (Entity Typing with
only its label, de-

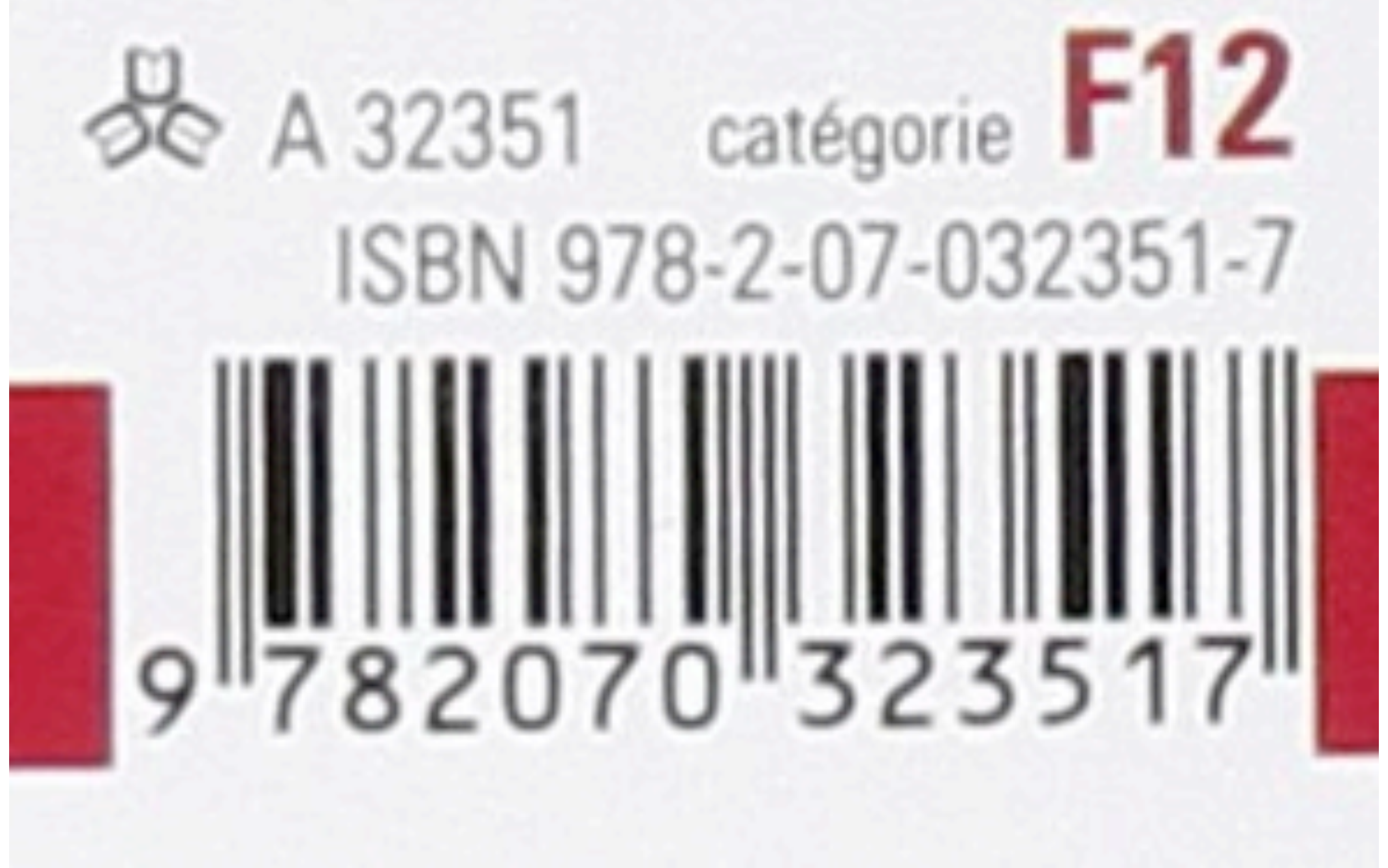
NASTyLinker: NIL-Aware Scalable Transformer-based Entity Linker

Nicolas Heist[0000-0002-4354-9138] and Heiko Paulheim[0000-0003-4386-8195]

Data and Web Science Group, University of Mannheim, Germany
{nico,heiko}@informatik.uni-mannheim.de

Abstract. Entity Linking (EL) is the task of detecting mentions of entities in text and disambiguating them to a reference knowledge base. Most prevalent EL approaches assume that the reference knowledge base is complete. In practice, however, it is necessary to deal with the case of linking to an entity that is not contained in the knowledge base (NIL entity). Recent works have shown that, instead of f

Identity



Entity Linking & Entity Spaces

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Entity Linking & Entity Spaces

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Entity Linking & Entity Spaces

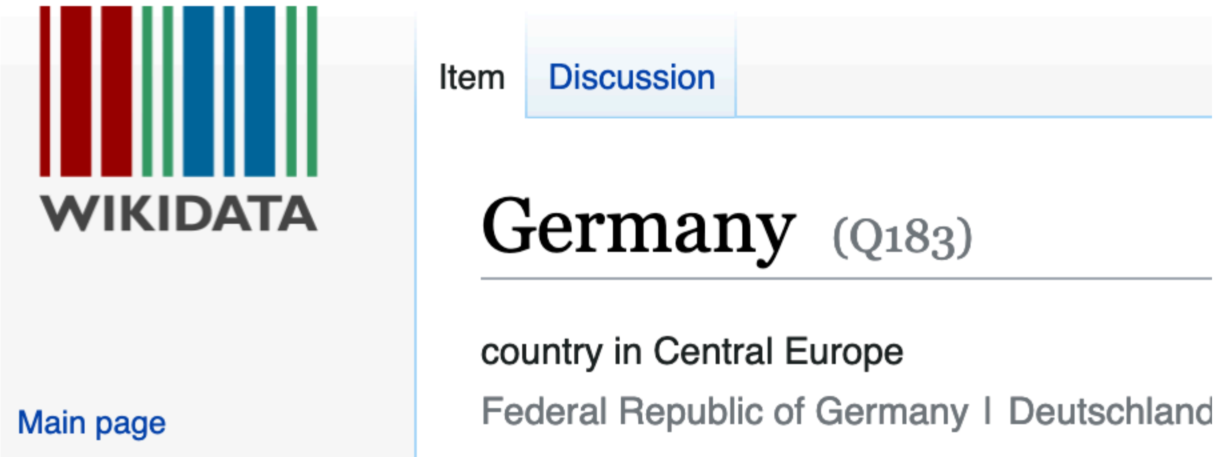
http://en.wikipedia.org/wiki/Germany {

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

http://en.wikipedia.org/wiki/Germany_Davis_Cup_team

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Entity Linking & Entity Spaces



WIKIDATA

Main page

Item | Discussion

Germany (Q183)

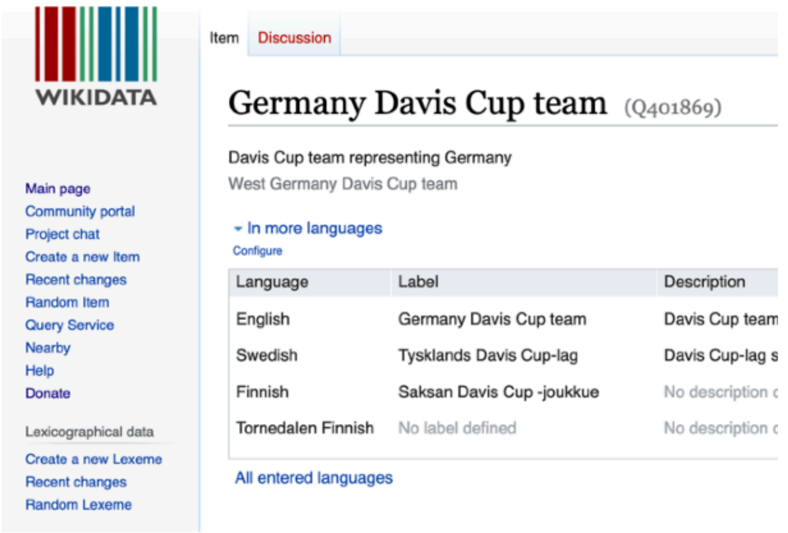
country in Central Europe

Federal Republic of Germany | Deutschland

http://en.wikipedia.org/wiki/Germany

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

http://en.wikipedia.org/wiki/Germany_Davis_Cup_team



WIKIDATA

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

Lexicographical data

Create a new Lexeme
Recent changes
Random Lexeme

Item | Discussion

Germany Davis Cup team (Q401869)

Davis Cup team representing Germany

West Germany Davis Cup team

In more languages

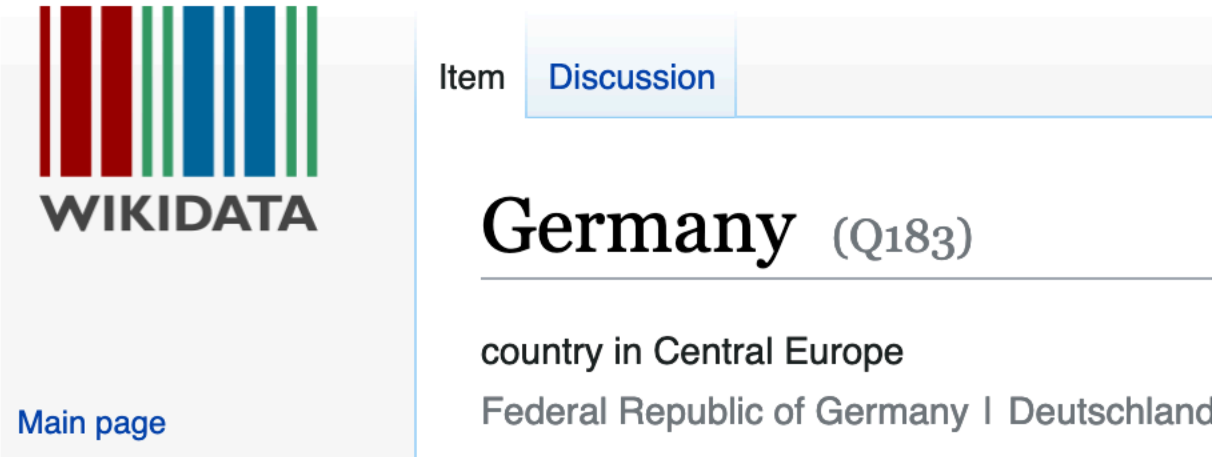
Configure

Language	Label	Description
English	Germany Davis Cup team	Davis Cup team
Swedish	Tysklands Davis Cup-lag	Davis Cup-lag s
Finnish	Saksan Davis Cup -joukkue	No description c
Tornedalen Finnish	No label defined	No description c

All entered languages

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Entity Linking & Entity Spaces



Germany (Q183)

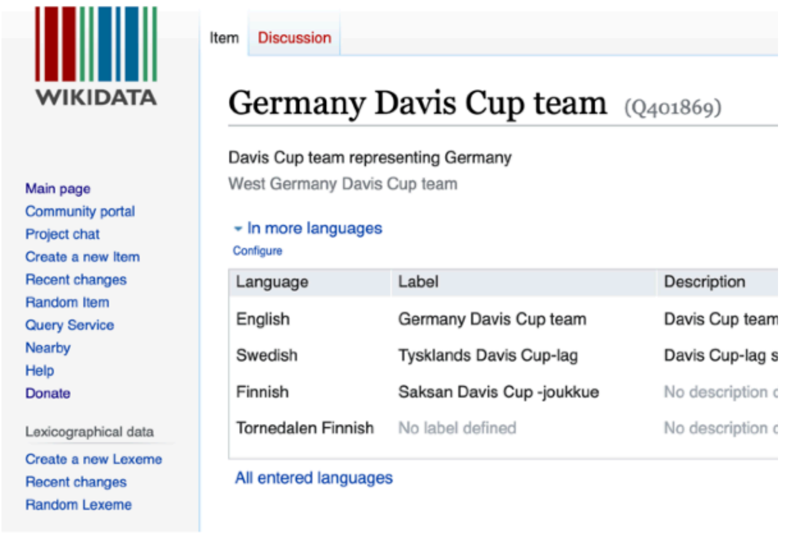
country in Central Europe

Federal Republic of Germany | Deutschland

http://en.wikipedia.org/wiki/Germany

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

http://en.wikipedia.org/wiki/Germany_Davis_Cup_team



Germany Davis Cup team (Q401869)

Davis Cup team representing Germany

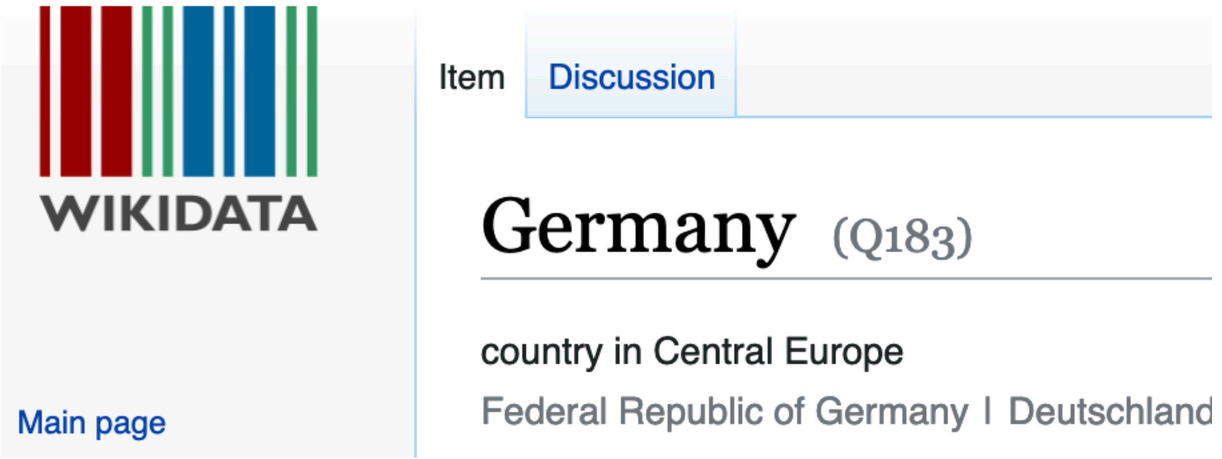
West Germany Davis Cup team

Language	Label	Description
English	Germany Davis Cup team	Davis Cup team
Swedish	Tysklands Davis Cup-lag	Davis Cup-lag s
Finnish	Saksan Davis Cup -joukkue	No description c
Tornedalen Finnish	No label defined	No description c

Germany 1996 Davis Cup Team

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Entity Linking & Entity Spaces



Germany (Q183)

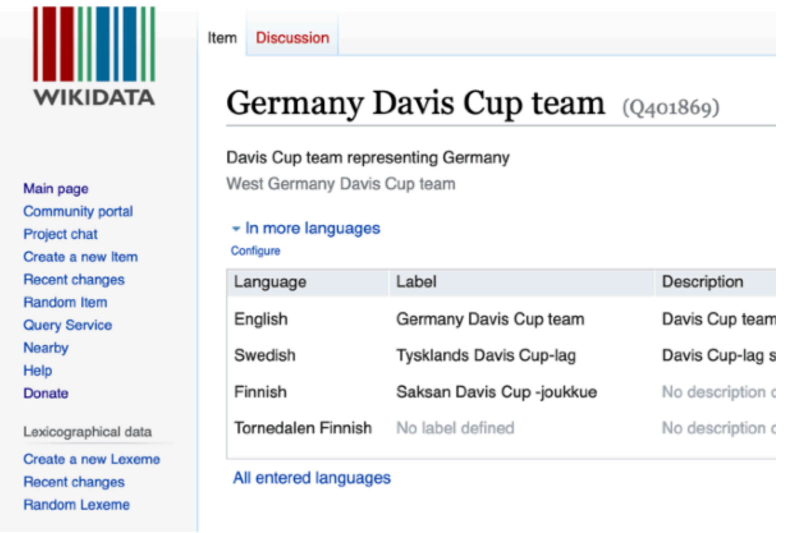
country in Central Europe

Federal Republic of Germany | Deutschland

http://en.wikipedia.org/wiki/Germany

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

http://en.wikipedia.org/wiki/Germany_Davis_Cup_team



Germany Davis Cup team (Q401869)

Davis Cup team representing Germany

West Germany Davis Cup team

In more languages

Configure

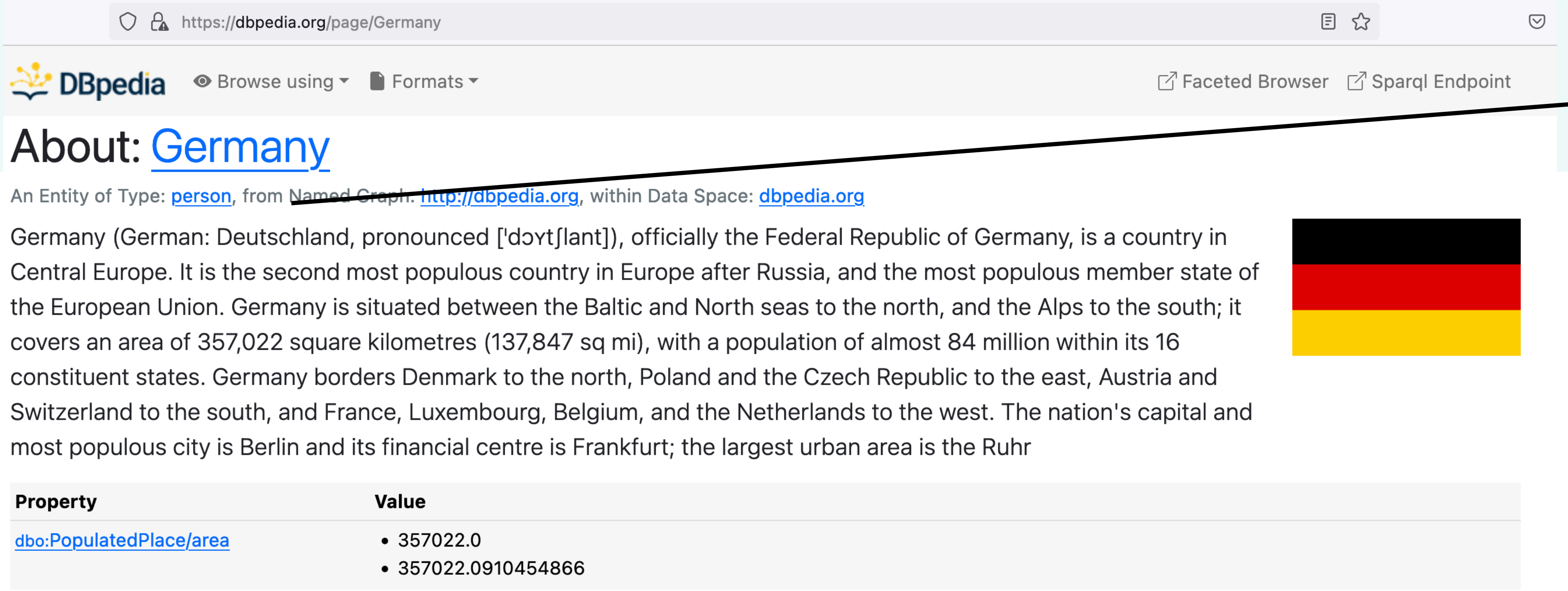
Language	Label	Description
English	Germany Davis Cup team	Davis Cup team
Swedish	Tysklands Davis Cup-lag	Davis Cup-lag s
Finnish	Saksan Davis Cup -joukkue	No description c
Tornedalen Finnish	No label defined	No description c

All entered languages

Germany 1996 Davis Cup Team

Germany 1993 Davis Cup Team

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

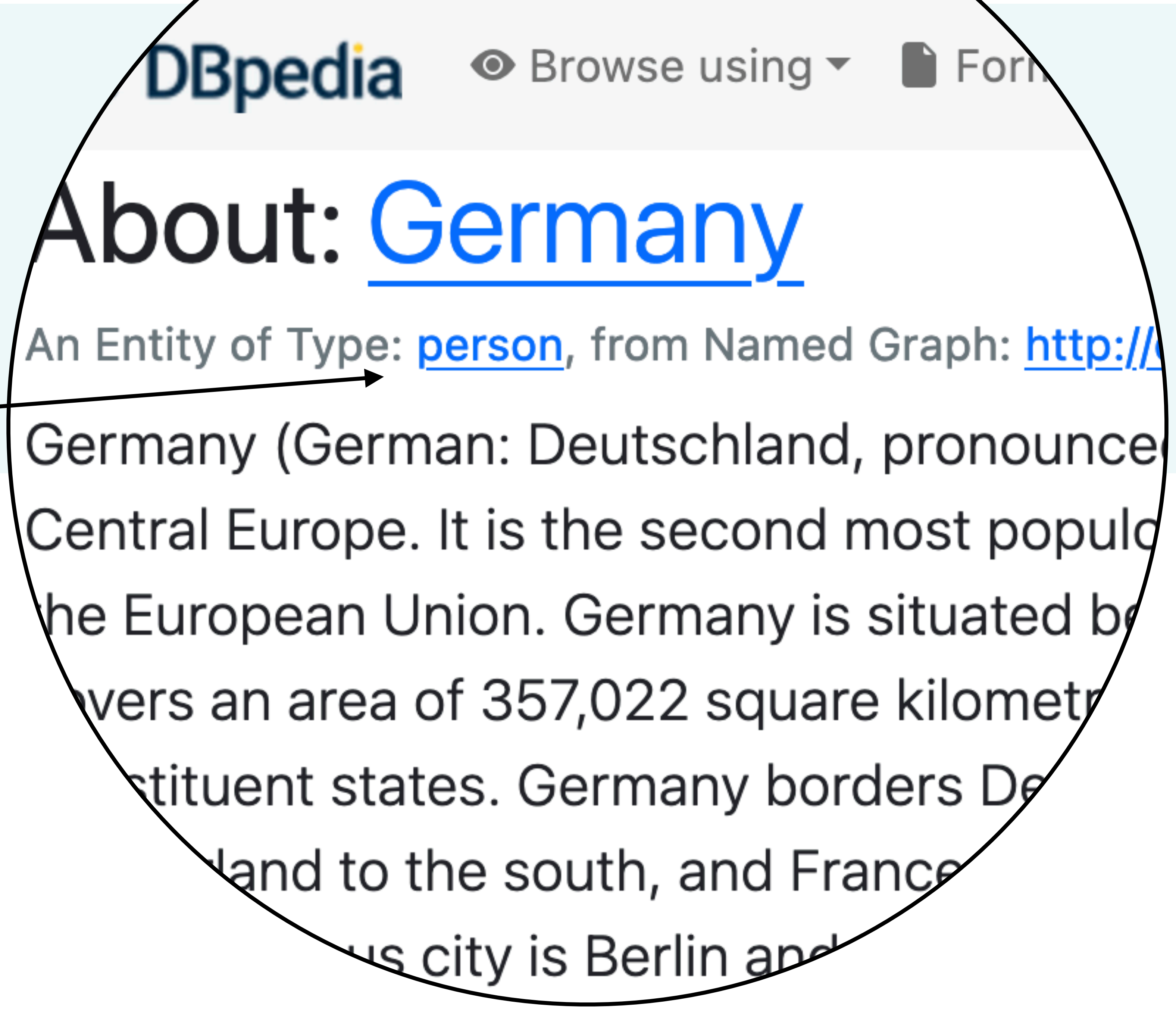


About: Germany

An Entity of Type: person, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org

Germany (German: Deutschland, pronounced [ˈdɔʏtʃlant]), officially the Federal Republic of Germany, is a country in Central Europe. It is the second most populous country in Europe after Russia, and the most populous member state of the European Union. Germany is situated between the Baltic and North seas to the north, and the Alps to the south; it covers an area of 357,022 square kilometres (137,847 sq mi), with a population of almost 84 million within its 16 constituent states. Germany borders Denmark to the north, Poland and the Czech Republic to the east, Austria and Switzerland to the south, and France, Luxembourg, Belgium, and the Netherlands to the west. The nation's capital and most populous city is Berlin and its financial centre is Frankfurt; the largest urban area is the Ruhr

Property	Value
dbo:PopulatedPlace/area	• 357022.0 • 357022.0910454866



About: Germany

An Entity of Type: person, from Named Graph: http://

Germany (German: Deutschland, pronounce
Central Europe. It is the second most popul
he European Union. Germany is situated b
vers an area of 357,022 square kilometr
stituent states. Germany borders De
land to the south, and France
s city is Berlin an

?

http://en.wikipedia.org/wiki/Germany

- Germany imported 47,600 sheep from Britain last year, nearly half of total imports.
- German July car registrations up 142 pct yr/yr.
- Australia last won the Davis Cup in 1986, but they were beaten finalists against Germany three years ago under Fraser's guidance.

http://en.wikipedia.org/wiki/Germany_Davis_Cup_team

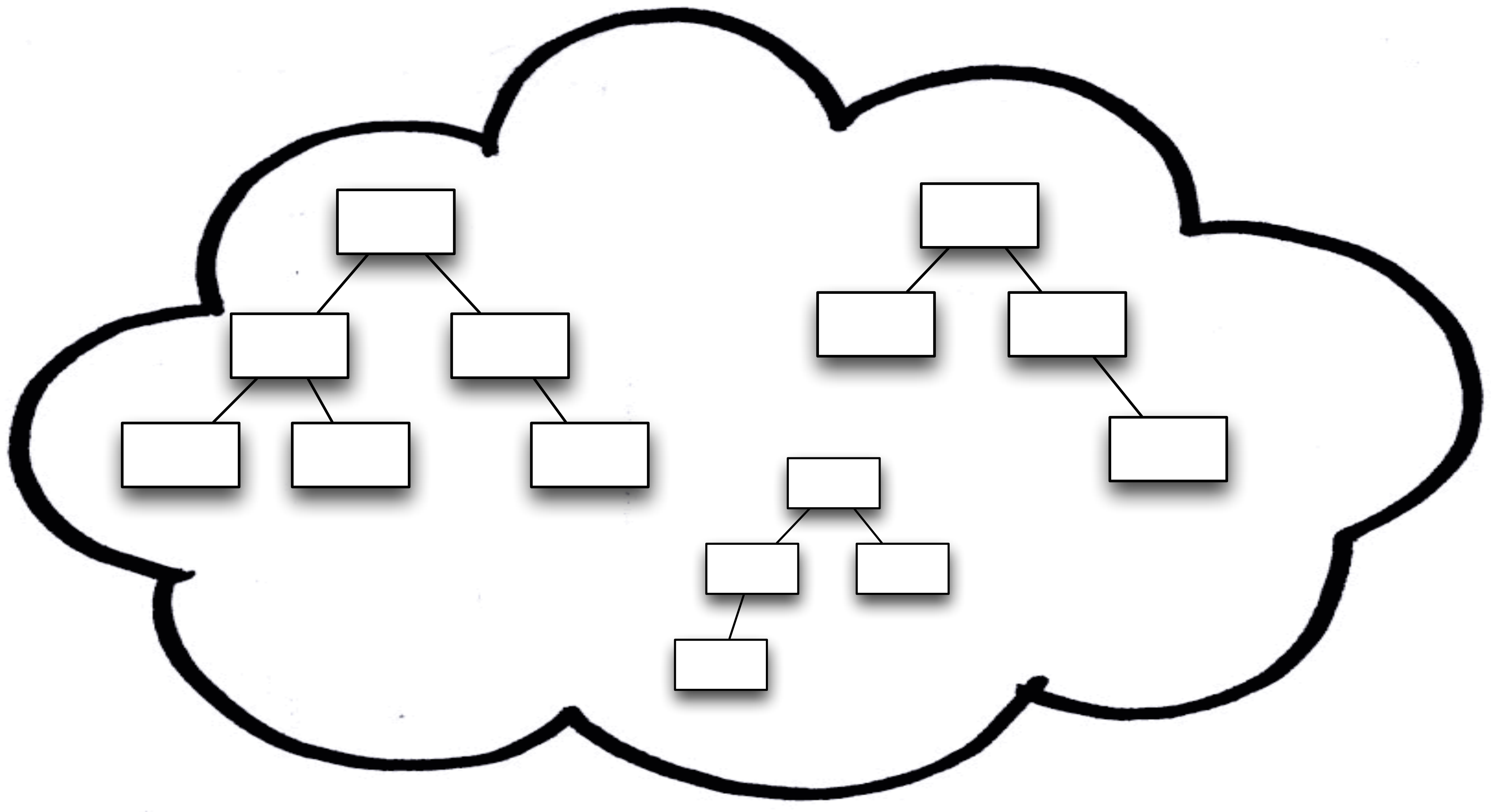
Germany 1996 Davis Cup Team

Germany 1993 Davis Cup Team

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

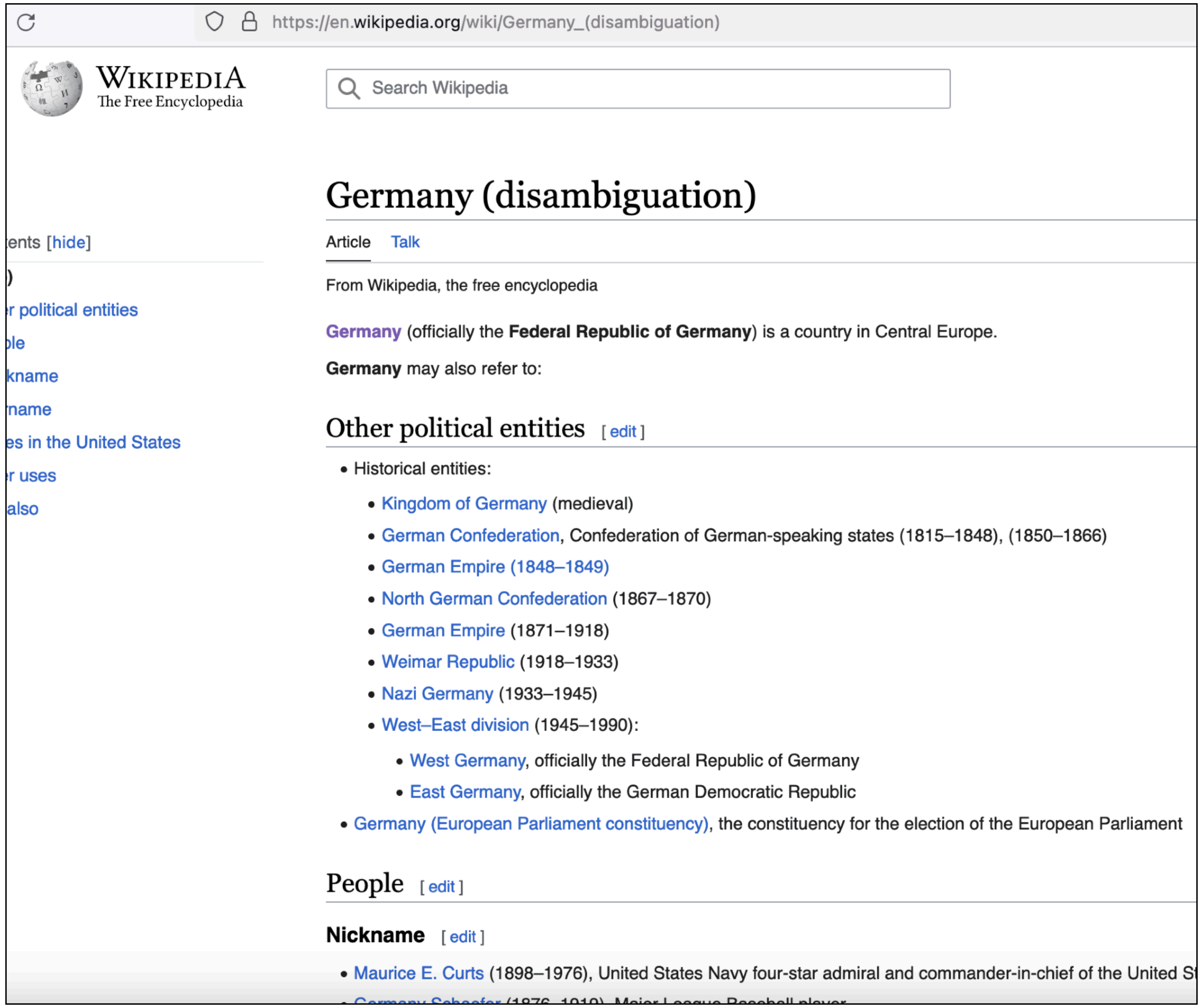
Entity Spaces

An **Entity Space** is an explicit representation of a set of entities in a knowledge base that have a strong near-identity relationship and whose linguistic labels can be used interchangeably in certain contexts.



Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Disambiguation Pages as Entity Space Proxies



Germany (disambiguation)

From Wikipedia, the free encyclopedia

Germany (officially the **Federal Republic of Germany**) is a country in Central Europe.

Germany may also refer to:

Other political entities [edit]

- Historical entities:
    - Kingdom of Germany (medieval)
    - German Confederation, Confederation of German-speaking states (1815–1848), (1850–1866)
    - German Empire (1848–1849)
    - North German Confederation (1867–1870)
    - German Empire (1871–1918)
    - Weimar Republic (1918–1933)
    - Nazi Germany (1933–1945)
    - West–East division (1945–1990):
        - West Germany, officially the Federal Republic of Germany
        - East Germany, officially the German Democratic Republic
- Germany (European Parliament constituency), the constituency for the election of the European Parliament

People [edit]

Nickname [edit]

- Maurice E. Curts (1898–1976), United States Navy four-star admiral and commander-in-chief of the United St
- Germany Schaefer (1876–1919), Major League Baseball player

Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France. European Language Resources Association.

Polyvocality

- Multiple dimensions also means multiple perspectives

- All data is biased

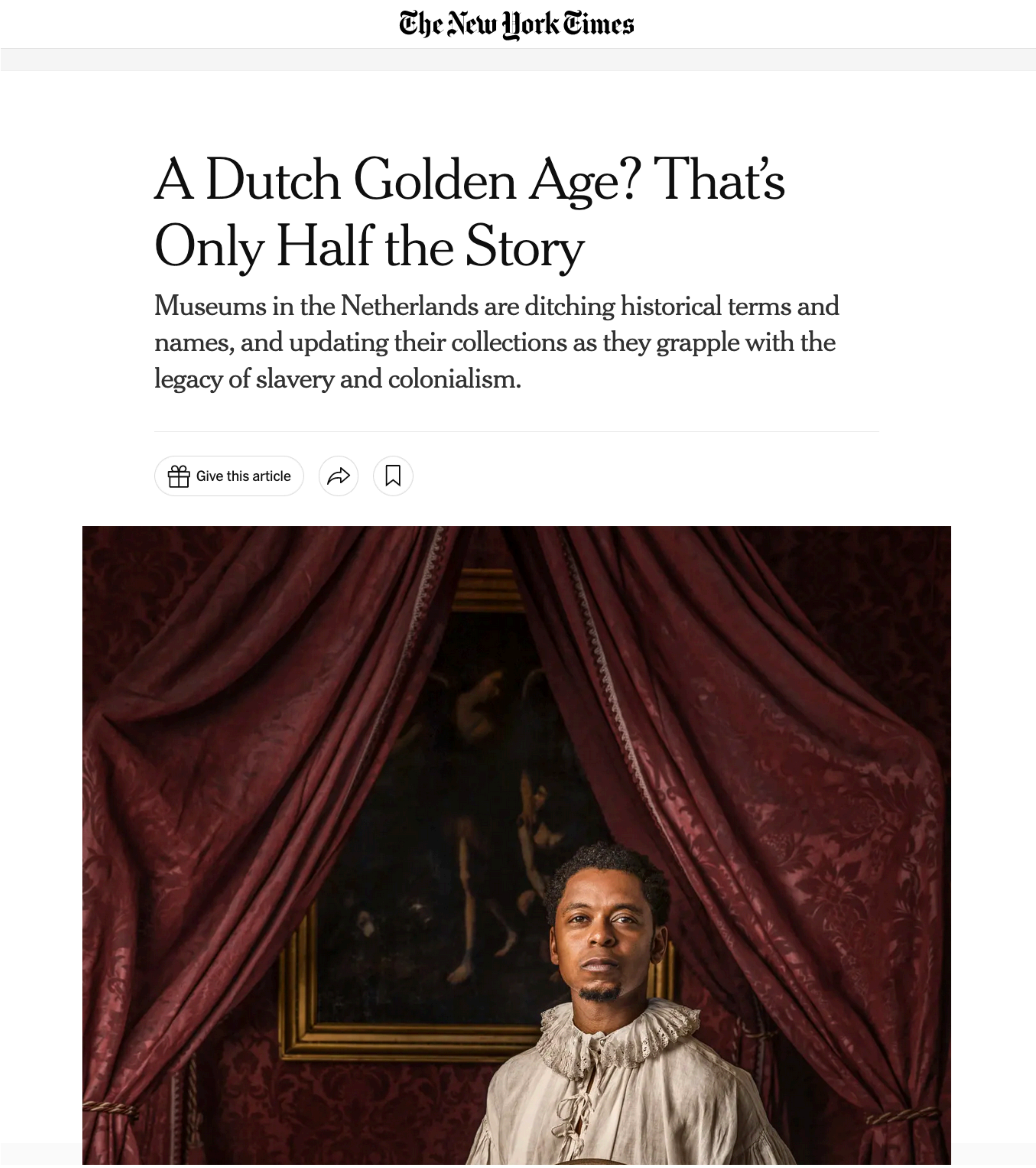
- KGs ***can*** represent multiple perspectives

Hendrick Cornelis Vroom Een aantal Oostindiëvaarders voor de kust
Rijksmuseum SK-A-3108

Marieke van Erp and Victor de Boer. "A polyvocal and contextualised semantic web." In *The Semantic Web: 18th International Conference, ESWC 2021*, Virtual Event, June 6–10, 2021, pp. 506-512. Springer International Publishing, 2021.

Words Matter

- GLAMs & society are rethinking their vocabularies
- Change is slow
- What does this mean for SW resources?

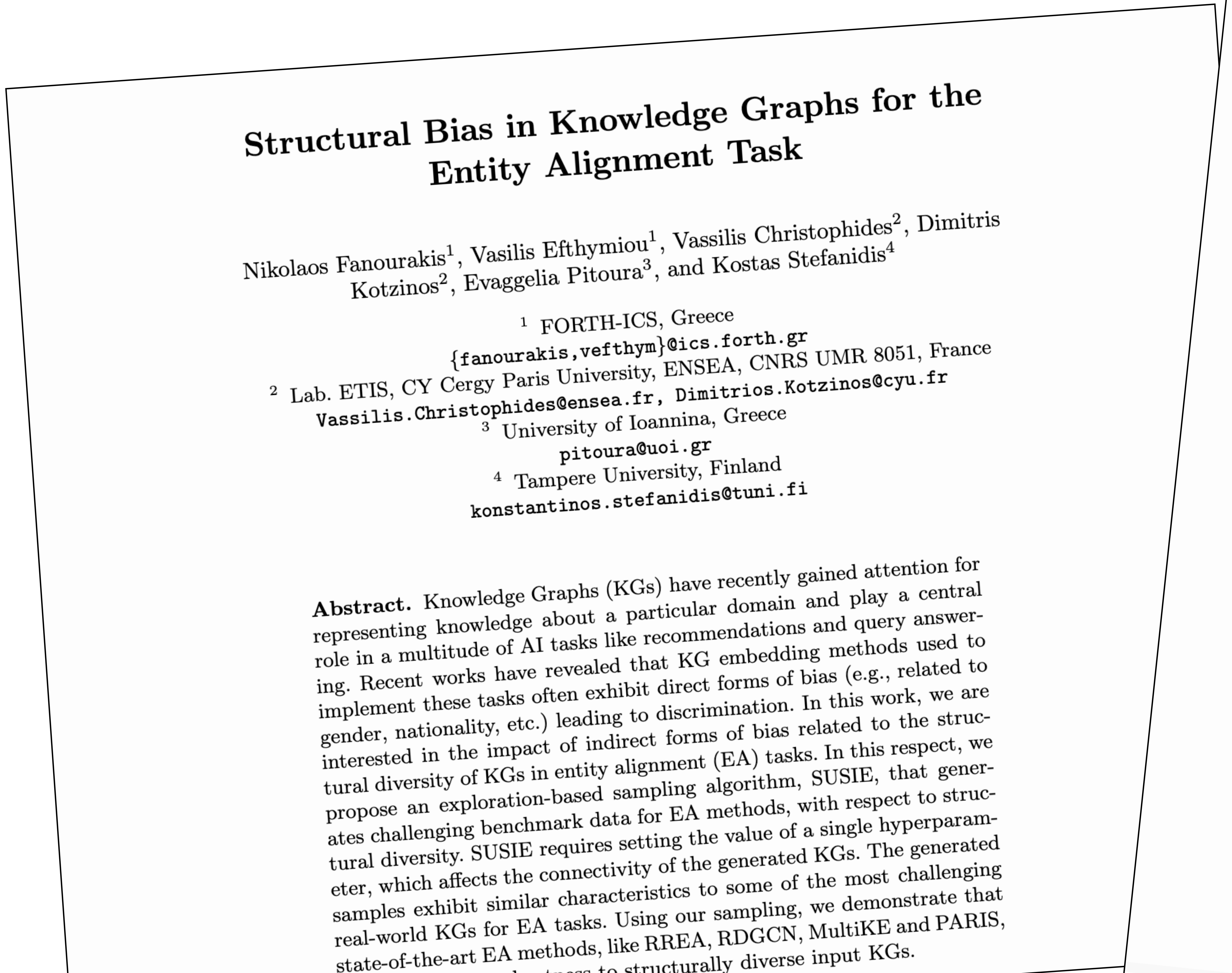


The New York Times

A Dutch Golden Age? That's Only Half the Story

Museums in the Netherlands are ditching historical terms and names, and updating their collections as they grapple with the legacy of slavery and colonialism.

Bias @ESWC2023

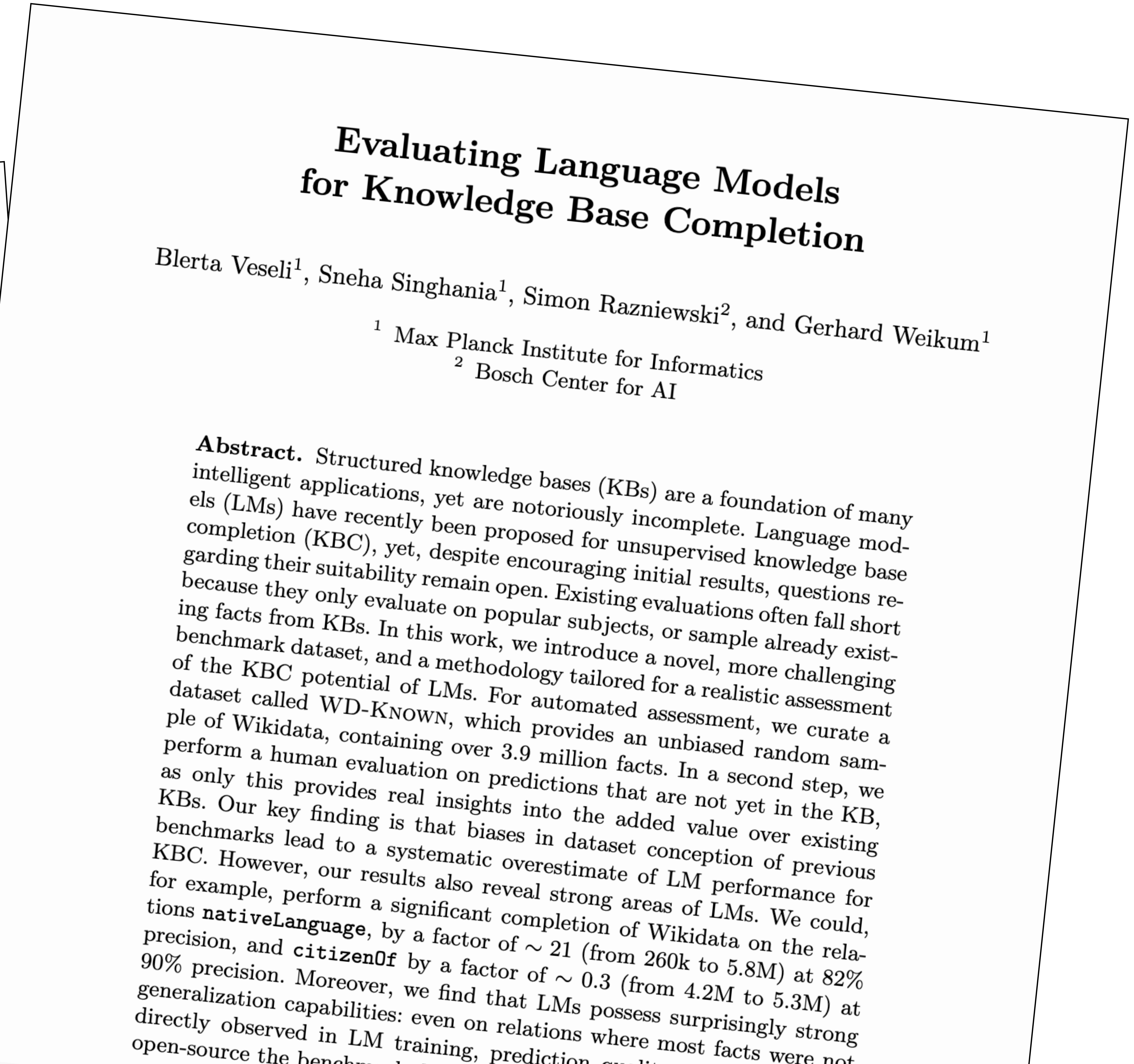


Structural Bias in Knowledge Graphs for the Entity Alignment Task

Nikolaos Fanourakis[1], Vasilis Efthymiou[1], Vassilis Christophides[2], Dimitris Kotzinos[2], Evaggelia Pitoura[3], and Kostas Stefanidis[4]

[1] FORTH-ICS, Greece
{fanourakis,vefthym}@ics.forth.gr
[2] Lab. ETIS, CY Cergy Paris University, ENSEA, CNRS UMR 8051, France
Vassilis.Christophides@ensea.fr, Dimitrios.Kotzinos@cyu.fr
[3] University of Ioannina, Greece
pitoura@uoi.gr
[4] Tampere University, Finland
konstantinos.stefanidis@tuni.fi

Abstract. Knowledge Graphs (KGs) have recently gained attention for representing knowledge about a particular domain and play a central role in a multitude of AI tasks like recommendations and query answering. Recent works have revealed that KG embedding methods used to implement these tasks often exhibit direct forms of bias (e.g., related to gender, nationality, etc.) leading to discrimination. In this work, we are interested in the impact of indirect forms of bias related to the structural diversity of KGs in entity alignment (EA) tasks. In this respect, we propose an exploration-based sampling algorithm, SUSIE, that generates challenging benchmark data for EA methods, with respect to structural diversity. SUSIE requires setting the value of a single hyperparameter, which affects the connectivity of the generated KGs. The generated samples exhibit similar characteristics to some of the most challenging real-world KGs for EA tasks. Using our sampling, we demonstrate that state-of-the-art EA methods, like RREA, RDGCN, MultiKE and PARIS, to structurally diverse input KGs.



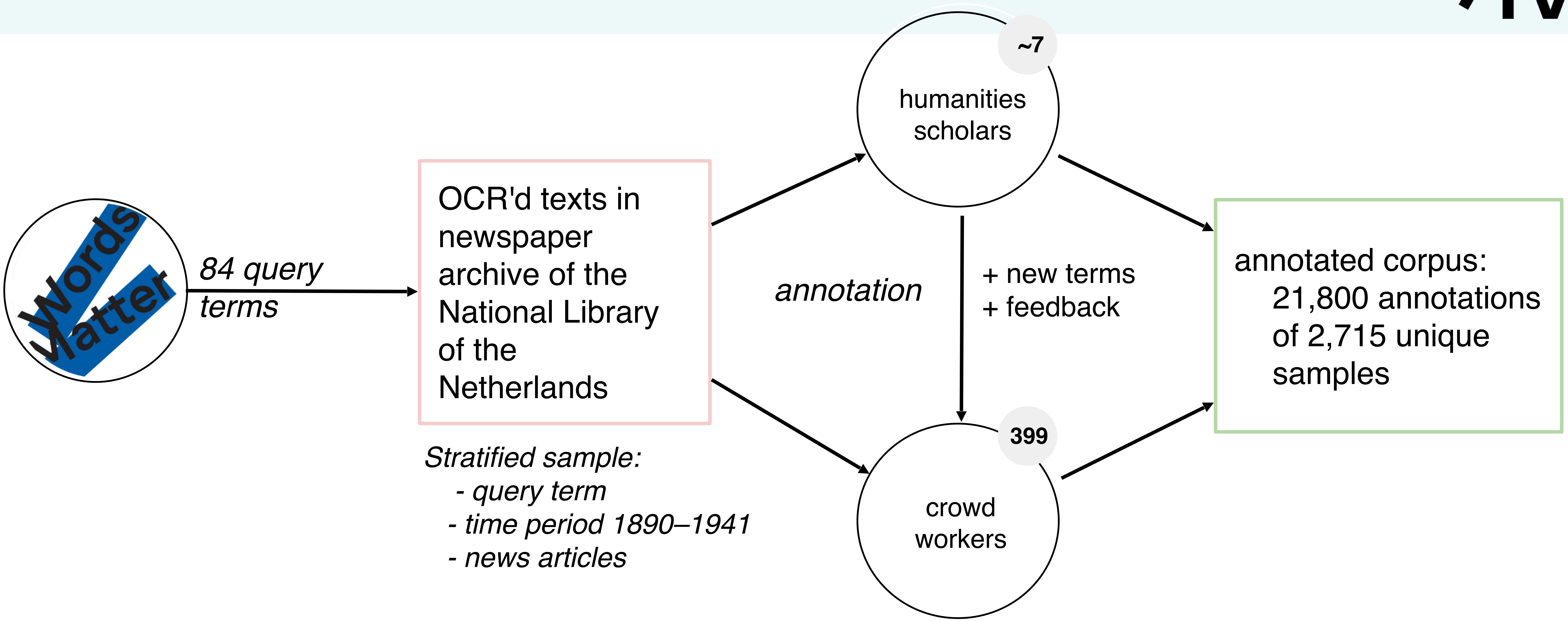
Evaluating Language Models for Knowledge Base Completion

Blerta Veseli[1], Sneha Singhania[1], Simon Razniewski[2], and Gerhard Weikum[1]

[1] Max Planck Institute for Informatics
[2] Bosch Center for AI

Abstract. Structured knowledge bases (KBs) are a foundation of many intelligent applications, yet are notoriously incomplete. Language models (LMs) have recently been proposed for unsupervised knowledge base completion (KBC), yet, despite encouraging initial results, questions regarding their suitability remain open. Existing evaluations often fall short because they only evaluate on popular subjects, or sample already existing facts from KBs. In this work, we introduce a novel, more challenging benchmark dataset, and a methodology tailored for a realistic assessment of the KBC potential of LMs. For automated assessment, we curate a dataset called WD-Known, which provides an unbiased random sample of Wikidata, containing over 3.9 million facts. In a second step, we perform a human evaluation on predictions that are not yet in the KB, as only this provides real insights into the added value over existing KBs. Our key finding is that biases in dataset conception of previous benchmarks lead to a systematic overestimate of LM performance for KBC. However, our results also reveal strong areas of LMs. We could, for example, perform a significant completion of Wikidata on the relations **nativeLanguage**, by a factor of ∼ 21 (from 260k to 5.8M) at 82% precision, and **citizenOf** by a factor of ∼ 0.3 (from 4.2M to 5.3M) at 90% precision. Moreover, we find that LMs possess surprisingly strong generalization capabilities: even on relations where most facts were not directly observed in LM training, prediction
open-source the

Constructing ConConCor



Words Matter

84 query terms

OCR'd texts in newspaper archive of the National Library of the Netherlands

Stratified sample:
- *query term*
- *time period 1890–1941*
- *news articles*

annotation

~7

humanities scholars

+ new terms
+ feedback

399

crowd workers

annotated corpus:
21,800 annotations of 2,715 unique samples

Ryan Brate, Andrei Nesterov, Valentin Vogelmann, Jacco Van Ossenbruggen, Laura Hollink, and Marieke Van Erp. "Capturing contentiousness: Constructing the contentious terms in context corpus." In Proceedings of the 11th on Knowledge Capture Conference, pp. 17-24. 2021.

How did crowd workers annotate the word "exotic"?

"De vrouw tegenover hem was nog maar een meisje, twintig naar schatting.

Een nauwsluitend zwart manteltje en rok, witte satijnen blouse, een kleine, chique, zwarte toque, modieus gedragen op één oor.

*Ze had een mooi, **exotisch** gezichtje, mat-witte huid, groote bruine oogen, git-zwart haar.*

Ze rookte een sigaret in een langen houder.

Haar gemanicuurde handen hadden donkerroode nagels."

“The woman opposite him was a mere girl—twenty at a guess.

A tight-fitting little black coat and skirt, white satin blouse, small chic black toque perched at the fashionable outrageous angle

*She had a beautiful **foreign-looking** face, dead white skin, large brown eyes, jet black hair.*

She was smoking a cigarette in a long holder.

Her manicured hands had deep red nails."

7 or 8 annotators per sample

4: contentious, 3: not contentious, 1: I don't know

Ryan Brate, Andrei Nesterov, Valentin Vogelmann, Jacco Van Ossenbruggen, Laura Hollink, and Marieke Van Erp. "Capturing contentiousness: Constructing the contentious terms in context corpus." In Proceedings of the 11th on Knowledge Capture Conference, pp. 17-24. 2021.

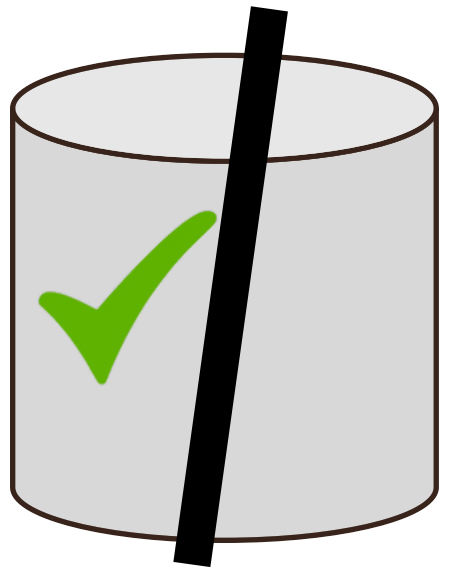
Lessons learned

Inter-rater agreement is low
- α = 0.54 among experts
- α = 0.31 for crowd workers

but can be improved (to α = 0.50)
by filtering out underperforming annotators:
- using control questions?
- using pairwise agreement between annotators?

First experiments demonstrate that the corpus can be used to train a model to predict contentiousness
baseline: balanced accuracy = [0.54-0.55]
model: balanced accuracy = [0.76-0.78]

Multiple annotators helps to get reliable data:
on half of the samples, over 80% of annotators agreed with each other.



Context is necessary to judge contentiousness:
most words are sometimes contentious and sometimes not contentious.

Cultural AI
a lab for
culturally
valued AI

CWI
Research institute for mathematics &
computer science in the Netherlands

KNAW
Humanities
Cluster

VU

NWO

A knowledge graph of contentious terminology for inclusive representation of cultural heritage

Andrei Nesterov
Laura Hollink
Marieke van Erp
Jacco van Ossenbruggen

Thursday 1 June
Session 4B: KGS in the Real World II 11:40 – 12:00
Orpheas

The Extended
Semantic Web Conference
Hersonissos, Greece
June 1st 2023

eswc23

Sneak Peek

- Characterising charged terms
- Identifying similarly behaving terms
- Accepted to: LDK 2023, 12-15 September, Vienna, Austria

Contextual Profiling of Charged Terms in Historical Newspapers

Ryan Brate and **Marieke van Erp**
KNAW Humanities Cluster, DHLab
Oudezijds Achterburgwal 185
1012 DK Amsterdam, Netherlands
{ryan.brate,marieke.van.erp}
@dh.huc.knaw.nl

Antal van den Bosch
Utrecht University
Institute for Language Sciences
Utrecht, the Netherlands
a.p.j.vandenbosch@uu.nl

Abstract

We extract nouns and corresponding co-occurrent targeted context features from a large corpus of Dutch language newspaper articles, from 1950s through the 1990s. Applying a well-established approach for scoring context feature and centre word associativity, we explored using the scores in the task of identifying key characteristics of known–charged terminology. Then use these features to draw parallels between known–charged and other terms. In the context of the very current decolonisation efforts amongst museum institutions, such approaches offer an opportunity to condense large quantities of data into the most-significant, salient information for digestion by heritage professionals. The methods were found to indeed yield insights into known and candidate charged terms.

Disclaimer: This paper contains derogatory words and phrases. They are provided solely as illustrations of the research results and do not reflect the opinions of the authors or their organisations. In-text examples of derogatory and potentially offensive are presented in ***"quotes, boldfaced and italicised"***.

1 Introduction

Museums of the World,[1] a database of cultural heritage institutions, records

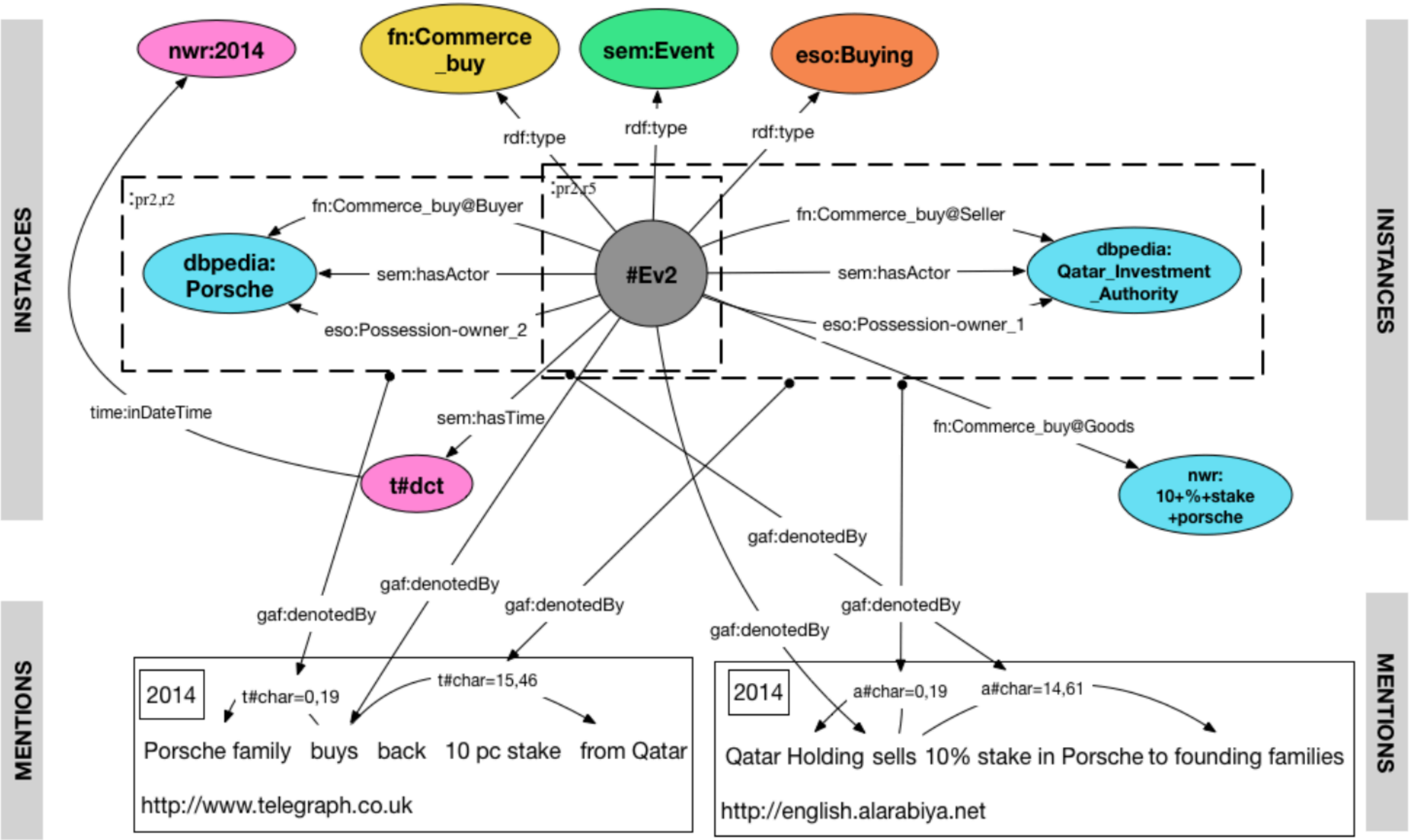
Many museum collections originate from the colonial period, with metadata and object portrayals stemming from the particular world of the time. There is now a growing movement of *decolonisation* in western museums aimed at the acknowledgement and accommodation of previously marginalised voices to combat biases propagated by the advancement of narrow viewpoints (Odumosu, 2020). Part of the decolonisation effort centres around greater sensitivity and reconsideration of the terminology and language used in item metadata. This is more complicated than wholesale removal of terminology from metadata and items from collections. To handle the complexities properly, there needs to be greater contextual understanding of a term's implied characterisation in context. For instance, many terms nowadays considered problematic are ambiguous, also in their contentiousness: calling a plant *exotic* is different from calling a person the same.

In this paper, we aim to explore the contextual profiles of a reference set of known charged collective nouns, reflective of some people group and identify the contextual features that distinguish them. Specifically, we consider four complementary context feature types: verbs for which the noun is the agent, verbs for which the noun is the patient, adjectives, and compound word modifiers

Change

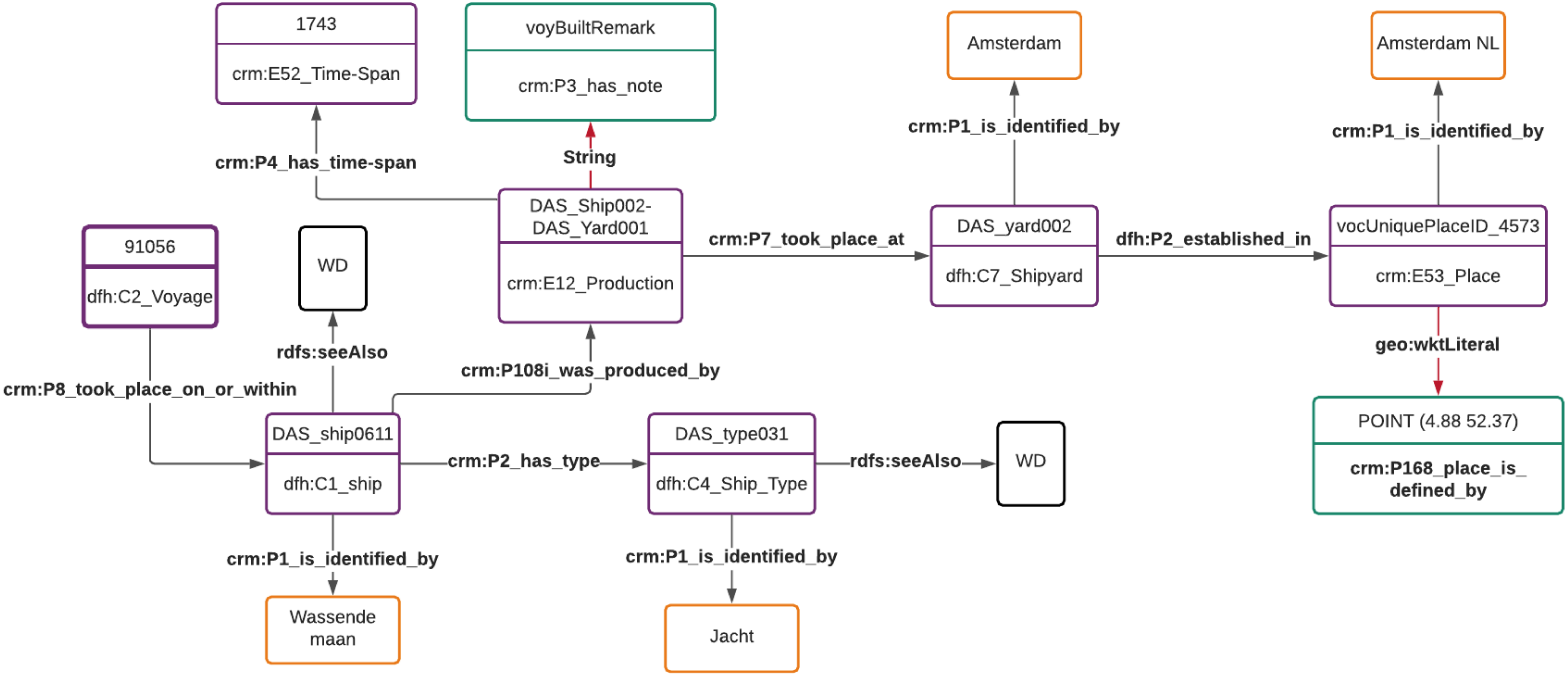
https://freesvg.org/img/cacao.png

Event-Centric KGs



Marco Rospocher, Marieke Van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. "Building event-centric knowledge graphs from news." Journal of Web Semantics 37 (2016): 132-151.

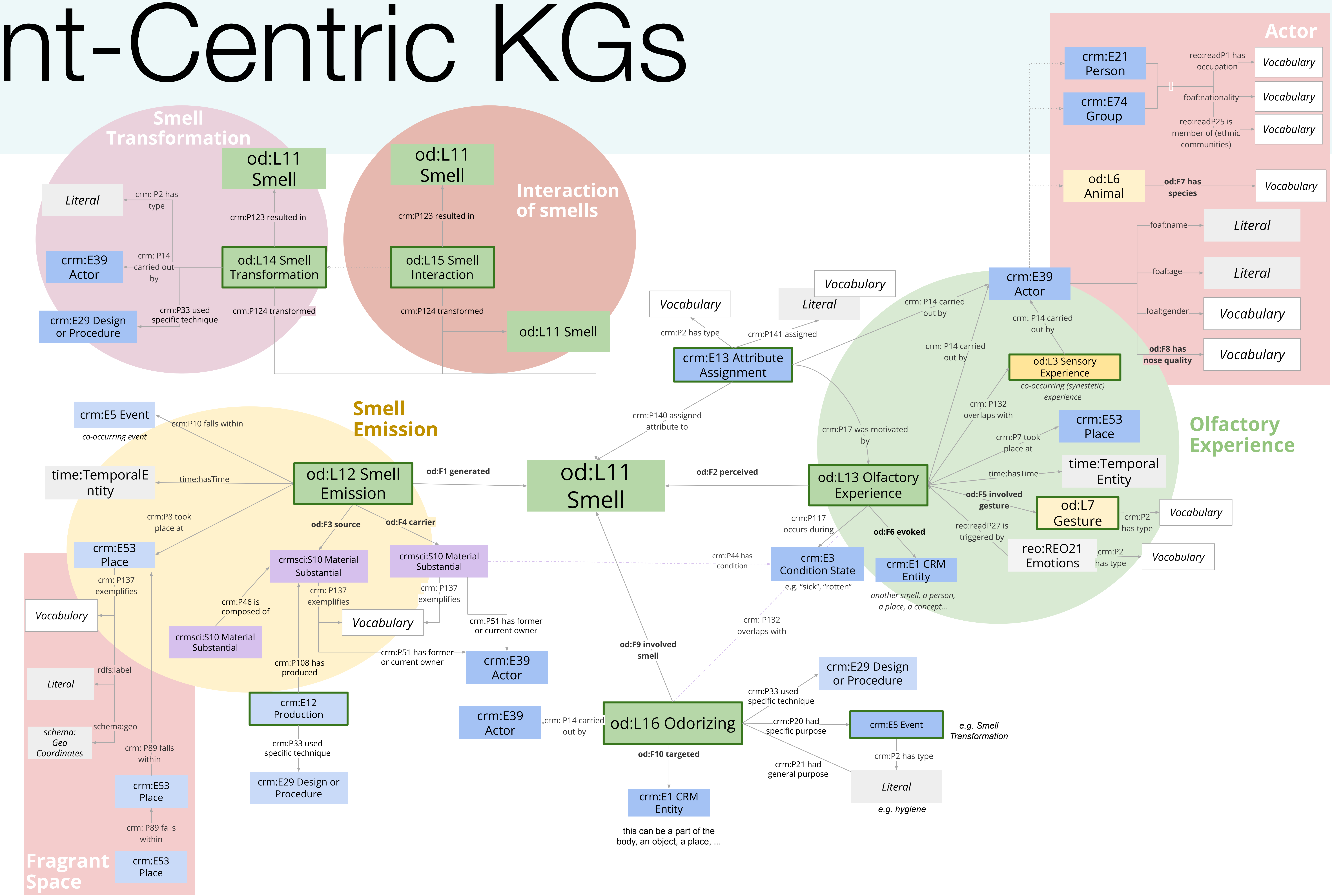
Event-Centric KGs



Stijn Schouten,, Victor De Boer, Lodewijk Petram, and Marieke Van Erp. "The wind in our sails: developing a reusable and maintainable Dutch maritime history knowledge graph." In Proceedings of the 11th on Knowledge Capture Conference, pp. 97-104. 2021

Event-Centric KGs

Odeuropa



Smell Transformation

od:L11 Smell

crm:P123 resulted in

od:L14 Smell Transformation

Literal

crm: P2 has type

crm:E39 Actor

crm: P14 carried out by

crm:E29 Design or Procedure

crm:P33 used specific technique

crm:P124 transformed

Interaction of smells

od:L11 Smell

crm:P123 resulted in

od:L15 Smell Interaction

crm:P124 transformed

od:L11 Smell

Smell Emission

crm:E5 Event

co-occurring event

crm:P10 falls within

time:TemporalEntity

time:hasTime

od:L12 Smell Emission

od:F1 generated

od:L11 Smell

crm:P8 took place at

crm:E53 Place

od:F3 source

od:F4 carrier

crmsci:S10 Material Substantial

crmsci:S10 Material Substantial

crm:P46 is composed of

crmsci:S10 Material Substantial

crm: P137 exemplifies

crm: P137 exemplifies

Vocabulary

crm:P51 has former or current owner

crm:P51 has former or current owner

crm:E39 Actor

crm:P108 has produced

crm:E12 Production

crm:P33 used specific technique

crm:E29 Design or Procedure

Fragrant Space

crm: P137 exemplifies

Vocabulary

rdfs:label

Literal

schema:geo

schema: Geo Coordinates

crm: P89 falls within

crm:E53 Place

crm: P89 falls within

crm:E53 Place

Actor

crm:E21 Person

crm:E74 Group

reo:readP1 has occupation

Vocabulary

foaf:nationality

Vocabulary

reo:readP25 is member of (ethnic communities)

Vocabulary

od:L6 Animal

od:F7 has species

Vocabulary

foaf:name

Literal

foaf:age

Literal

foaf:gender

Vocabulary

od:F8 has nose quality

Vocabulary

Olfactory Experience

Vocabulary

Literal

Vocabulary

crm:P2 has type

crm:P141 assigned

crm:E13 Attribute Assignment

crm:P140 assigned attribute to

crm: P14 carried out by

crm:E39 Actor

crm: P14 carried out by

crm: P14 carried out by

od:L3 Sensory Experience

co-occurring (synestetic) experience

crm:P17 was motivated by

crm: P132 overlaps with

crm:E53 Place

crm:P7 took place at

time:Temporal Entity

time:hasTime

od:F2 perceived

od:L13 Olfactory Experience

od:F5 involved gesture

od:L7 Gesture

crm:P2 has type

Vocabulary

reo:readP27 is triggered by

reo:REO21 Emotions

crm:P2 has type

Vocabulary

crm:P117 occurs during

od:F6 evoked

crm:P44 has condition

crm:E3 Condition State

e.g. "sick", "rotten"

crm:E1 CRM Entity

another smell, a person, a place, a concept...

crm: P132 overlaps with

od:F9 involved smell

crm:E39 Actor

crm: P14 carried out by

od:L16 Odorizing

crm:P33 used specific technique

crm:E29 Design or Procedure

crm:P20 had specific purpose

crm:E5 Event

e.g. Smell Transformation

crm:P21 had general purpose

crm:P2 has type

Literal

e.g. hygiene

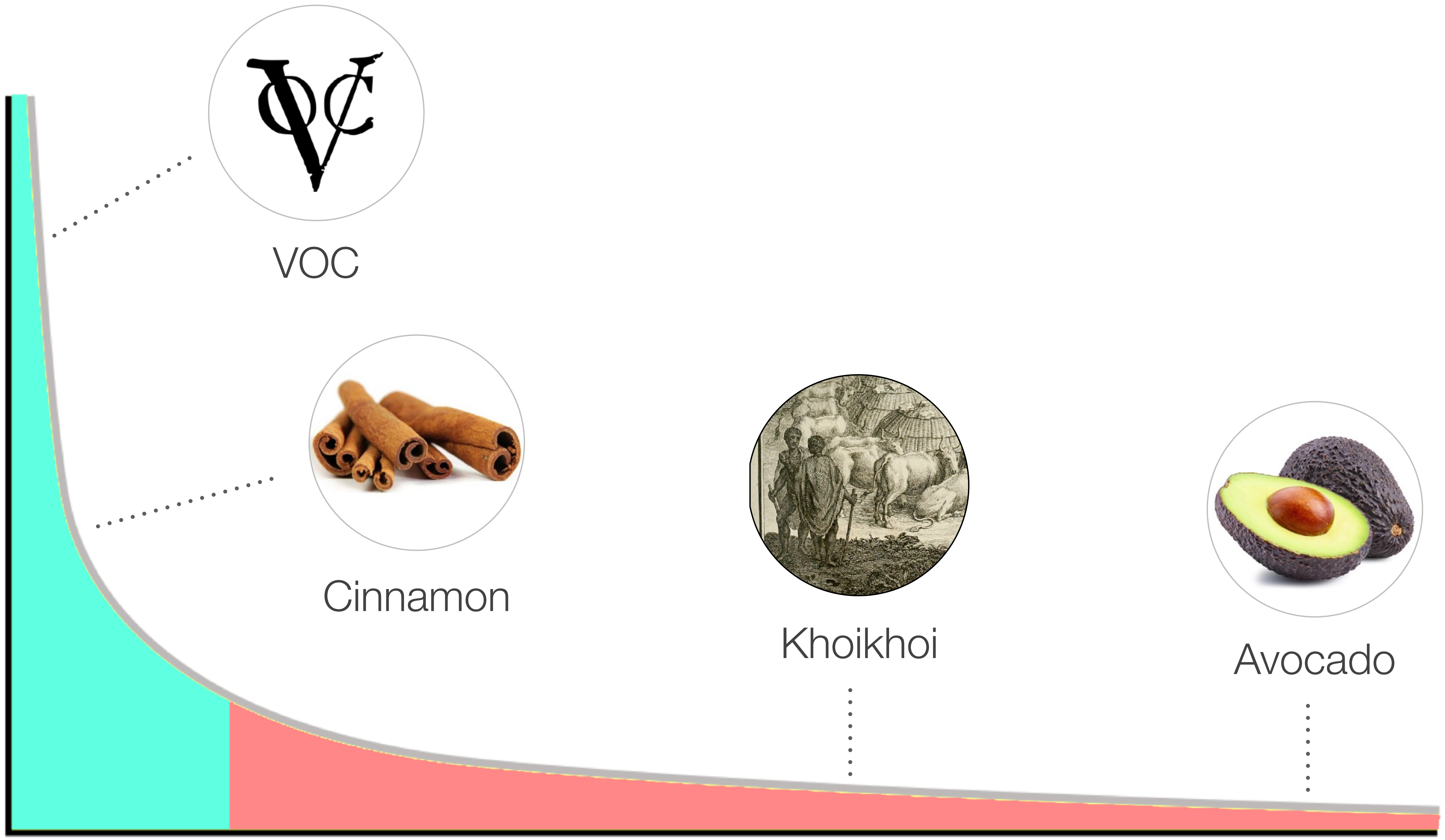
od:F10 targeted

crm:E1 CRM Entity

this can be a part of the body, an object, a place, ...

Pasquale Lisena, Daniel Schwabe, Marieke van Erp, Raphaël Troncy, William Tullett, Inger Leemans, Lizzie Marx, and Sofia Colette Ehrich. "Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information." In *The Semantic Web: 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29–June 2, 2022, Proceedings*, pp. 387-405. Cham: Springer International Publishing, 2022.

The Long Tail



VOC

Cinnamon

Khoikhoi

Avocado

Datasets

Dataset	Type	Domain	Doc length	Format	Encoding	License
AIDA-YAGO2	news	general	medium	TSV	ASCII	Agreement
2014/2015 NEEL	tweets	general	short	TSV	ASCII	Open
OKE2015	encyclopaedia	general	long	NIF/RDF	UTF8	Open
RSS-500	news	general	medium	NIF/RDF	UTF8	Open
WES2015	blog	science	long	NIF/RDF	UTF8	Open
WikiNews	news	general	medium	XML	UTF8	Open

Marieke van Erp, Pablo N. Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo, and Jörg Waitelonis. "Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job."

Entity Overlap

Proportion of entities present in one dataset that are also present in other datasets

	AIDA-YAGO2	NEEL2014	NEEL2015	OKE2015	RSS500	WES2015	Wikinews
AIDA-YAGO2 (5,596)		5.87%	8.06%	0.00%	1.26%	4.80%	1.16%
NEEL2014 (2,380)	13.73%		68.49%	2.39%	2.56%	12.35%	2.82%
NEEL2015 (2,800)	16.11%	58.21%		2.00%	2.54%	7.93%	2.57%
OKE2015 (531)	0.00%	10.73%	10.55%		2.44%	28.06%	3.95%
RSS500 (849)	8.24%	7.18%	8.36%	1.53%		3.18%	1.88%
WES2015 (7,309)	3.68%	4.02%	3.04%	2.04%	0.16%		0.66%
Wikinews (279)	23.30%	24.01%	25.81%	7.53%	5.73%	17.20%	

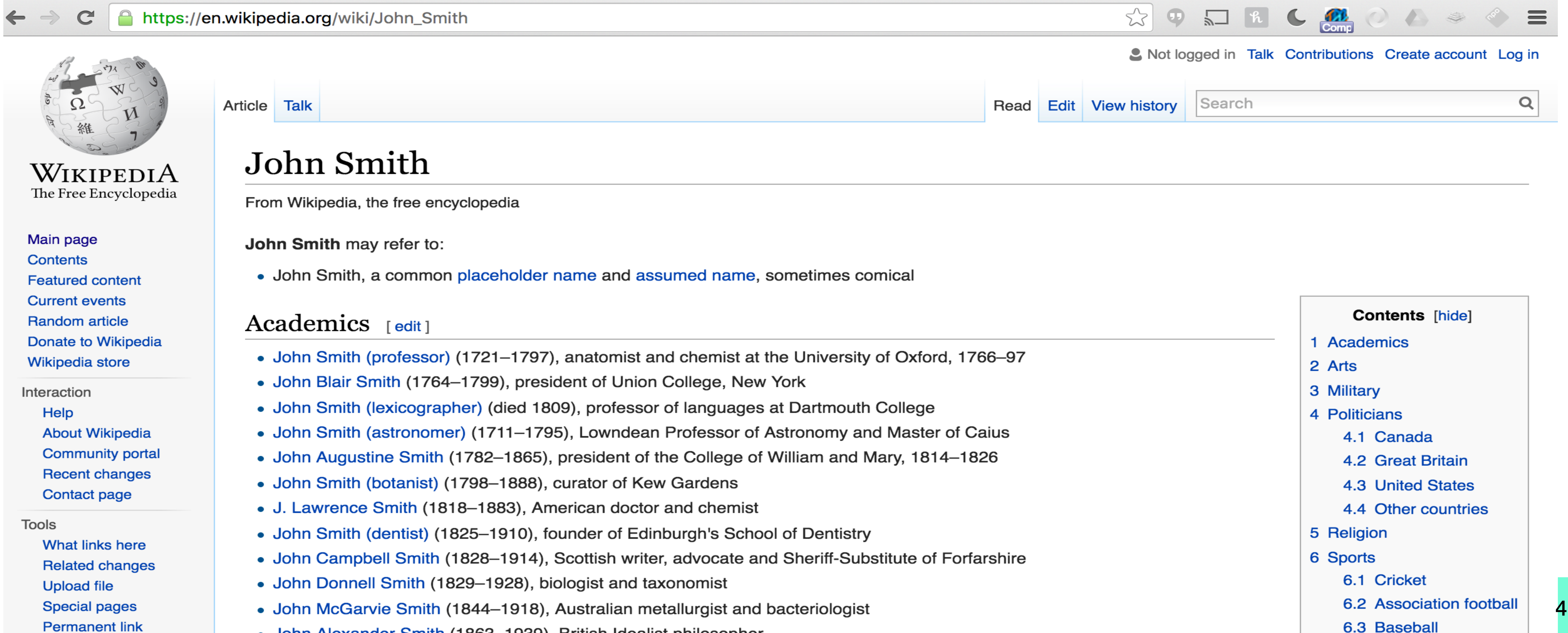
Marieke van Erp, Pablo N. Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo, and Jörg Waitelonis. "Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job."

Confusability

- The number of meanings a surface form (mention) can have

Confusability

- The number of meanings a surface form (mention) can have



John Smith

From Wikipedia, the free encyclopedia

John Smith may refer to:

- John Smith, a common placeholder name and assumed name, sometimes comical

Academics [edit]

- John Smith (professor) (1721–1797), anatomist and chemist at the University of Oxford, 1766–97
- John Blair Smith (1764–1799), president of Union College, New York
- John Smith (lexicographer) (died 1809), professor of languages at Dartmouth College
- John Smith (astronomer) (1711–1795), Lowndean Professor of Astronomy and Master of Caius
- John Augustine Smith (1782–1865), president of the College of William and Mary, 1814–1826
- John Smith (botanist) (1798–1888), curator of Kew Gardens
- J. Lawrence Smith (1818–1883), American doctor and chemist
- John Smith (dentist) (1825–1910), founder of Edinburgh's School of Dentistry
- John Campbell Smith (1828–1914), Scottish writer, advocate and Sheriff-Substitute of Forfarshire
- John Donnell Smith (1829–1928), biologist and taxonomist
- John McGarvie Smith (1844–1918), Australian metallurgist and bacteriologist
- John Alexander Smith (1863–1939), British Idealist philosopher

Contents [hide]

Confusability

Corpus	Average	Min	Max	σ
AIDA-YAGO2	1.08	1	13	0.37
2014 NEEL	1.02	1	3	0.16
2015 NEEL	1.05	1	4	0.25
OKE2015	1.11	1	25	1.22
RSS500	1.02	1	3	0.16
WES2015	1.06	1	6	0.30
Wikinews	1.09	1	29	1.03

Marieke van Erp, Pablo N. Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo, and Jörg Waitelonis. "Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job."

Dominance

Corpus	Dominance	Min	Max	σ
AIDA-YAGO2	.98	1	452	0.08
2014 NEEL	.99	1	47	0.06
2015 NEEL	.98	1	88	0.09
OKE2015	.98	1	1	0.11
RSS500	.99	1	1	0.07
WES2015	.97	1	1	0.12
Wikinews	.99	1	72	0.09

Marieke van Erp, Pablo N. Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo, and Jörg Waitelonis. "Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job."

Change @ESWC2023

A Geological Case Study on Semantically Triggered Processes

Yuanwei Qu, Eduard Kamburjan, and Martin Giese

SIRIUS Center, University of Oslo, Norway
quy,eduard,martingi@ifi.uio.no

Abstract. We present an approach to connect semantic
situations to program-based descriptions of processes. Th
anism is a semantically formalised *trigger* that initiates
demonstrate the viability of the approach by modellin
processes in petroleum geoscience.

1 Introduction

Semantic technologies are designed to build graph-based
reason about static relationships between entities and
to represent dynamic behavior and changes. Although
on formalisation of the concept of change [6] and top-l
to describe processes [1], there is still limited support f
e.g. in simulations, to build conditionals and loops to
described by the knowledge representations.

The distinction between utilizing semantic techn
edge of dynamic processes and programming languag
ics remains pronounced. The current work of [3] int
language called 'Semantic Micro Object Language'

HHT: An Approach for Representing Temporally-Evolving Historical Territories

W. Charles[1], N. Aussenac-Gilles[1], and N. Hernandez[1]

IRIT - Université de Toulouse name.surname@irit.fr

he notion of territory plays a major role in human and
. Representation of this spatio-temporal object and com-
e change
historic
tric dat
ories,w
y) to re
to kn
ks to
provi
kn
at r
ry,
har
ow

LauNuts: A Knowledge Graph to identify and compare geographic regions in the European Union*

Adrian Wilke and Axel Ngonga

DICE group, Department of Computer Science, Paderborn University
adrian.wilke@uni-paderborn.de, axel.ngonga@upb.de
https://dice-research.org/

Abstract. The *Nomenclature of Territorial Units for Statistics* (NUTS) is a classification that represents countries in the European Union (EU). It is published at intervals of several years and organized in a hierarchical system where geographical areas are subdivided according to their population sizes. In addition to NUTS, there is a further subdivided hierarchy level, named *Local Administrative Units* (LAU), whose data are updated annually by EU member states. While both datasets are published by Eurostat as Excel files, an additional RDF dataset is available for NUTS up to the 2016 scheme. With this work, we provide the Linked Data community with an up-to-date Knowledge Graph in which NUTS and LAU data are linked and which contains population numbers as well as area sizes. We also publish an Open Source generator software for future released versions that will naturally arise due to changes in population numbers. These contributions can be used to enrich other datasets and allow comparisons among regions in the European Union. All resources are available at https://w3id.org/launuts.

GLENDA: Querying RDF Archives with full SPARQL

Olivier Pelgrin[1][0000−0002−1025−9687], Ruben Taelman[2][0000−0001−5118−256X], Luis Galárraga[3][0000−0002−0241−5379], and Katja Hose[1,4][0000−0001−7025−8099]

[1] Aalborg University, Denmark, {olivier,khose}@cs.aau.dk
[2] Ghent University, ruben.taelman@ugent.be
[3] Inria, France, luis.galarraga@inria.fr
[4] TU Wien, Austria, katja.hose@tuwien.ac.at

Abstract. The dynamicity of semantic data has propelled the research on *RDF Archiving*, i.e., the task of storing and making the full history of large RDF datasets accessible. However, existing archiving techniques fail to scale when confronted with very large RDF datasets and support only simple SPARQL queries. In this demonstration, we theref
GLENDA, a system that can run full SPARQL 1.1
large RDF archives. We achieve th
based storage archit
engin

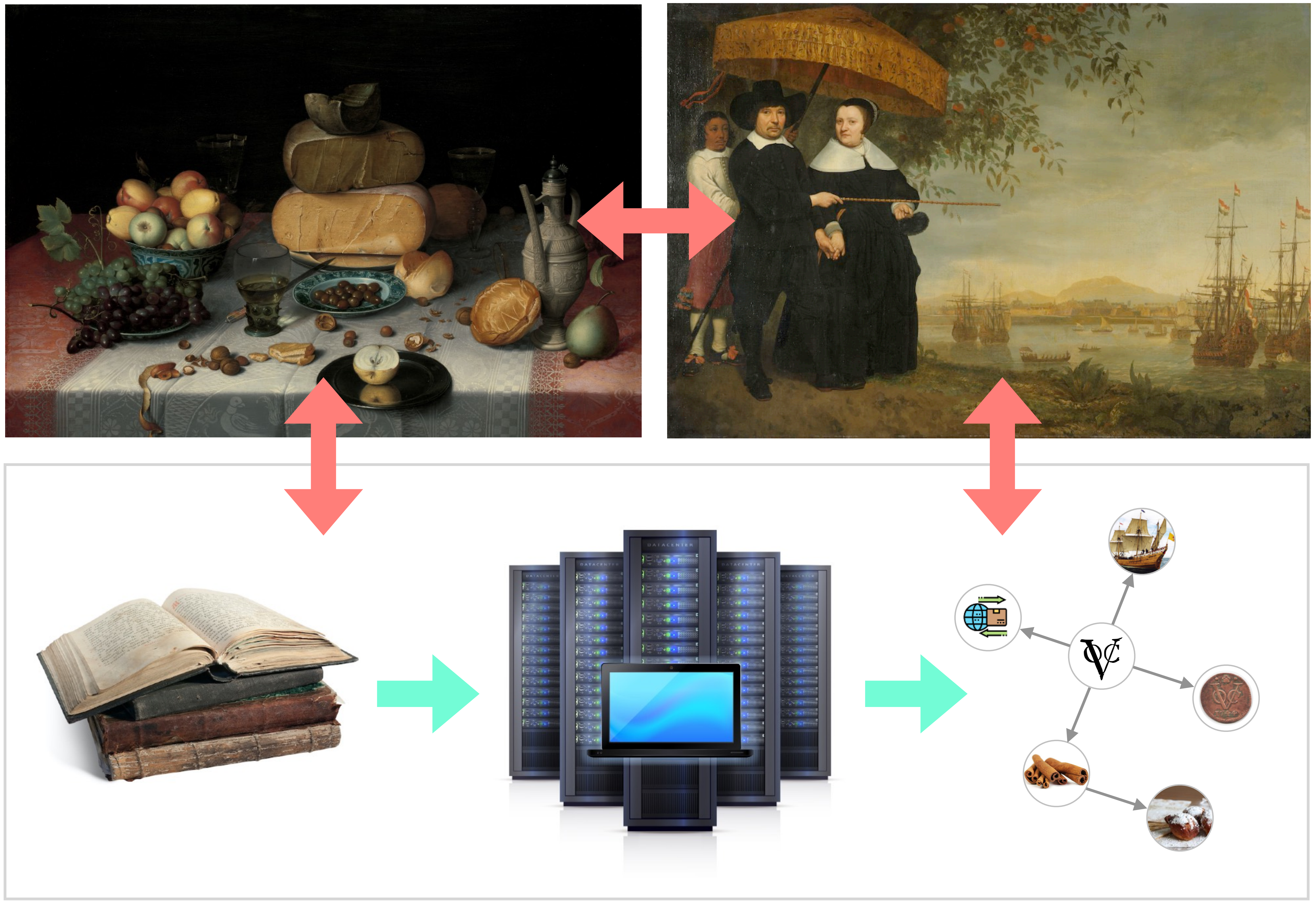
Towards polyvocal & contextualised KGs

- KGs have come a long way
- More dimensions, change, and including the long tail are the next frontier
- Benchmark datasets need fixing
- This is a community effort

Olfert Dapper (1680) Naukeurige beschryving van Asie

TRIFECTA: Capturing Identity, Change, and the Long Tail in Knowledge Graphs

- Tracing contentious entities and concepts in food and maritime history
- Combining language and semantic web technology to unflatten knowledge graphs
- Strengthening computational data-driven humanities research



trifecta.dhlab.nl

Thank you

NewsReader, Odeuropa and TRIFECTA have received funding from the European Union's FP7, Horizon 2020 and Horizon Europe research and innovation programmes under grant agreement No 316404, 101004469, and 10188548. Cultural AI has received funding from the Dutch Research Council (NWO), the Dutch Digital Heritage Network (NDE) and Europeana.

References

- Marieke van Erp, Filip Ilievski, Marco Rospocher, and Piek Vossen. "Missing Mr. Brown and Buying an Abraham Lincoln-Dark Entities and DBpedia." In NLP-DBPEDIA@ ISWC, pp. 81-86. 2015.
- Ulbe Bosma (2023) The World of Sugar, Harvard University Press.
- Pasquale Lisena, Daniel Schwabe, Marieke van Erp, Raphaël Troncy, William Tullett, Inger Leemans, Lizzie Marx, and Sofia Colette Ehrich. "Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information." In The Semantic Web: 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29–June 2, 2022, Proceedings, pp. 387-405. Cham: Springer International Publishing, 2022.
- Marta Recasens, Eduard Hovy, and M. Antònia Martí. "Identity, non-identity, and near-identity: Addressing the complexity of coreference." Lingua 121, no. 6 (2011): 1138-1152.
- Marieke van Erp and Paul Groth. 2020. Towards Entity Spaces. In Proceedings of the Twelfth Language Resources and Evaluation Conference, pages 2129–2137, Marseille, France. European Language Resources Association.
- Marieke van Erp and Victor de Boer. "A polyvocal and contextualised semantic web." In The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event, June 6–10, 2021, pp. 506-512. Springer International Publishing, 2021.
- Ryan Brate, Andrei Nesterov, Valentin Vogelmann, Jacco Van Ossenbruggen, Laura Hollink, and Marieke Van Erp. "Capturing contentiousness: Constructing the contentious terms in context corpus." In Proceedings of the 11th on Knowledge Capture Conference, pp. 17-24. 2021.
- Andrei Nesterov, Laura Hollink, Marieke van Erp, and Jacco van Ossenbruggen. "A Knowledge Graph of Contentious Terminology for Inclusive Representation of Cultural Heritage." In The Semantic Web: 20th International Conference, ESWC 2023, Hersonissos, Crete, Greece, May 28–June 1, 2023, Proceedings, pp. 502-519. Cham: Springer Nature Switzerland, 2023.
- Ryan Brate, Marieke van Erp, and Antal van den Bosch. “Contextual Profiling of Charged Terms in Historical Newspapers” Accepted to: Language, Data and Knowledge 2023, Vienna, Austria, September 2023.
- Marco Rospocher, Marieke Van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. "Building event-centric knowledge graphs from news." Journal of Web Semantics 37 (2016): 132-151.
- Stijn Schouten,, Victor De Boer, Lodewijk Petram, and Marieke Van Erp. "The wind in our sails: developing a reusable and maintainable Dutch maritime history knowledge graph." In Proceedings of the 11th on Knowledge Capture Conference, pp. 97-104. 2021
- Marieke van Erp, Pablo N. Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo, and Jörg Waitelonis. "Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job."

Image credits

Slide 3: https://i.pinimg.com/736x/8d/5a/bf/8d5abf3a0dca3fdb3dfc7771929490a3--danny-odonoghue-coffee-cups.jpg https://thecoffeevine.com/wp-content/uploads/2013/10/img_5753.jpg https://greekcitytimes.com/wp-content/uploads/2018/09/ellinikos-kafes.jpg

Slide 4: https://www.thisisathens.org/sites/default/files/2019-05/MCH-Kafeneia-CaptainMichalis3.jpg https://pxhere.com/en/photo/1453579

Slide 5: https://www.bl.uk/collection-items/drawing-of-a-london-coffee-house-c-1690-1700

Slide 6: https://www.nationaalarchief.nl/onderzoeken/archief/1.04.18.02/invnr/10800/file/NL-HaNA_1.04.18.02_10800_0039

Slide 7: https://upload.wikimedia.org/wikipedia/commons/thumb/6/67/VOC_Candia_by_G.Groenewegen_1798.jpg/640px-VOC_Candia_by_G.Groenewegen_1798.jpg

Slide 8: https://upload.wikimedia.org/wikipedia/commons/2/28/Coffee%3B_from_plantation_to_cup._A_brief_history_of_coffee_production_and_consumption._With_an_appendix_containing_letters_written_during_a_trip_to_the_coffee_plantations_of_the_East_and_through_the_%2820541585900%29.jpg

Slide 15: https://res.infoq.com/articles/graph-knowledge-base-cloud-native/en/headerimage/graph-knowledge-base-cloud-native-1586764855790.jpg

Slide 17:https://www.britishmuseum.org/collection/image/33593001

Slide 18: ttps://wartski.com/wp-content/uploads/2016/04/Pomander-16th-cenrtury-2-jpg.jpg

Slide 22: https://m.media-amazon.com/images/I/71XOTFj2nsL._AC_UF1000,1000_QL80_.jpg

Slide 27: Hendrick Cornelis Vroom Een aantal Oostindiëvaarders voor de kust. Rijksmuseum SK-A-3108

Slide 28: https://www.nytimes.com/2019/10/25/arts/design/dutch-golden-age-and-colonialism.html

Slide 48: Olfert Dapper (1680) Naukeurige beschryving van Asie https://books.google.nl/books?id=-iGqvQEACAAJ&hl=nl&pg=RA2-PA62#v=onepage&q&f=false