

# Wikipedia Cultural Diversity Observatory:

Empowering Communities to Bridge the Culture Content Gaps

[<https://meta.wikimedia.org/wiki/WCDO>]

**Marc Miquel, PhD**

{marcmiquel@gmail.com}

Username:marcmiquel

Pompeu Fabra University, Barcelona, Catalonia

Amical Wikimedia (Catalan Wikipedia)

WG Diversity – Strategy 2030

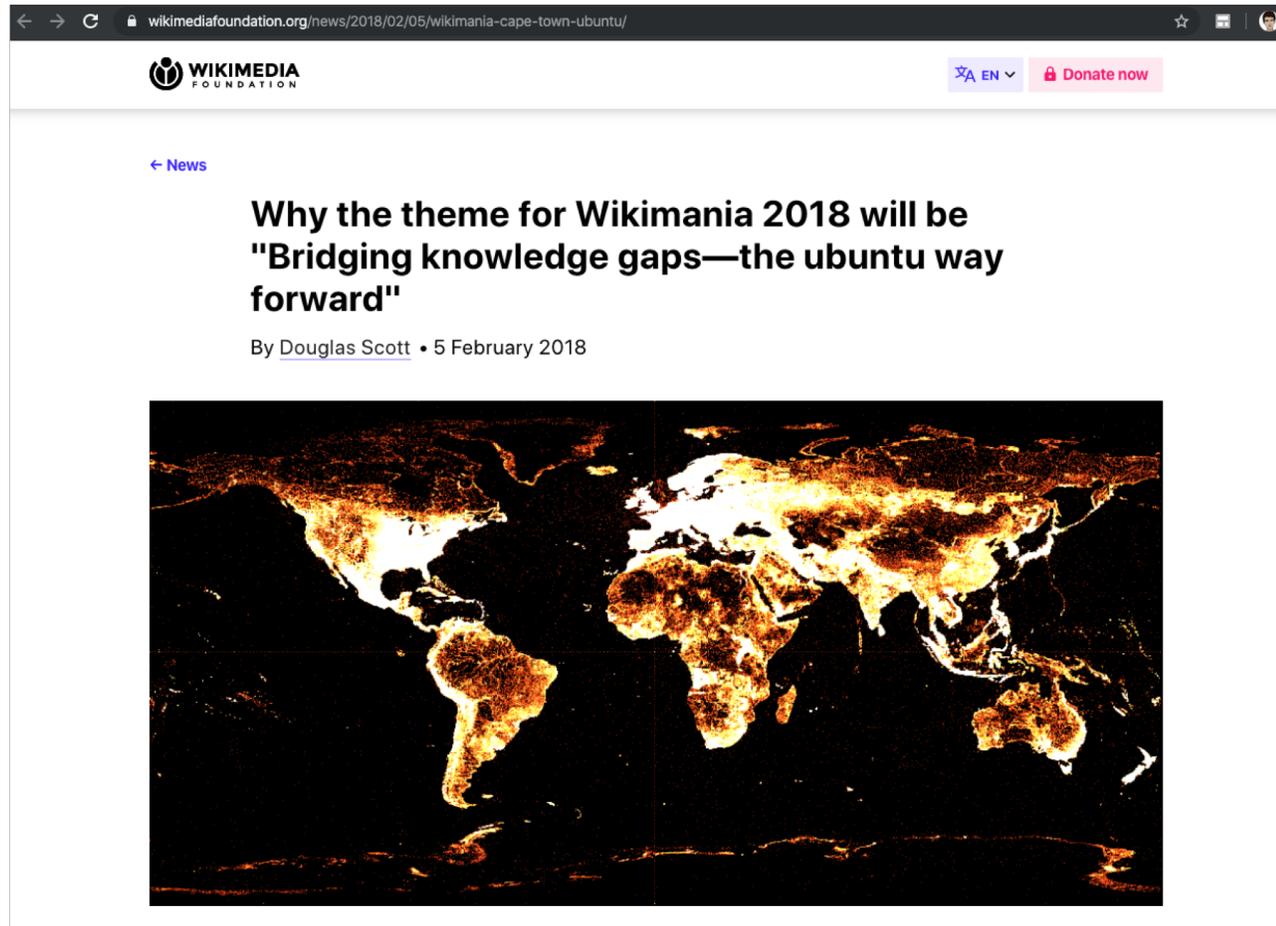


**WIKIMANIA**  
**STOCKHOLM**

AUGUST 14-18, 2019

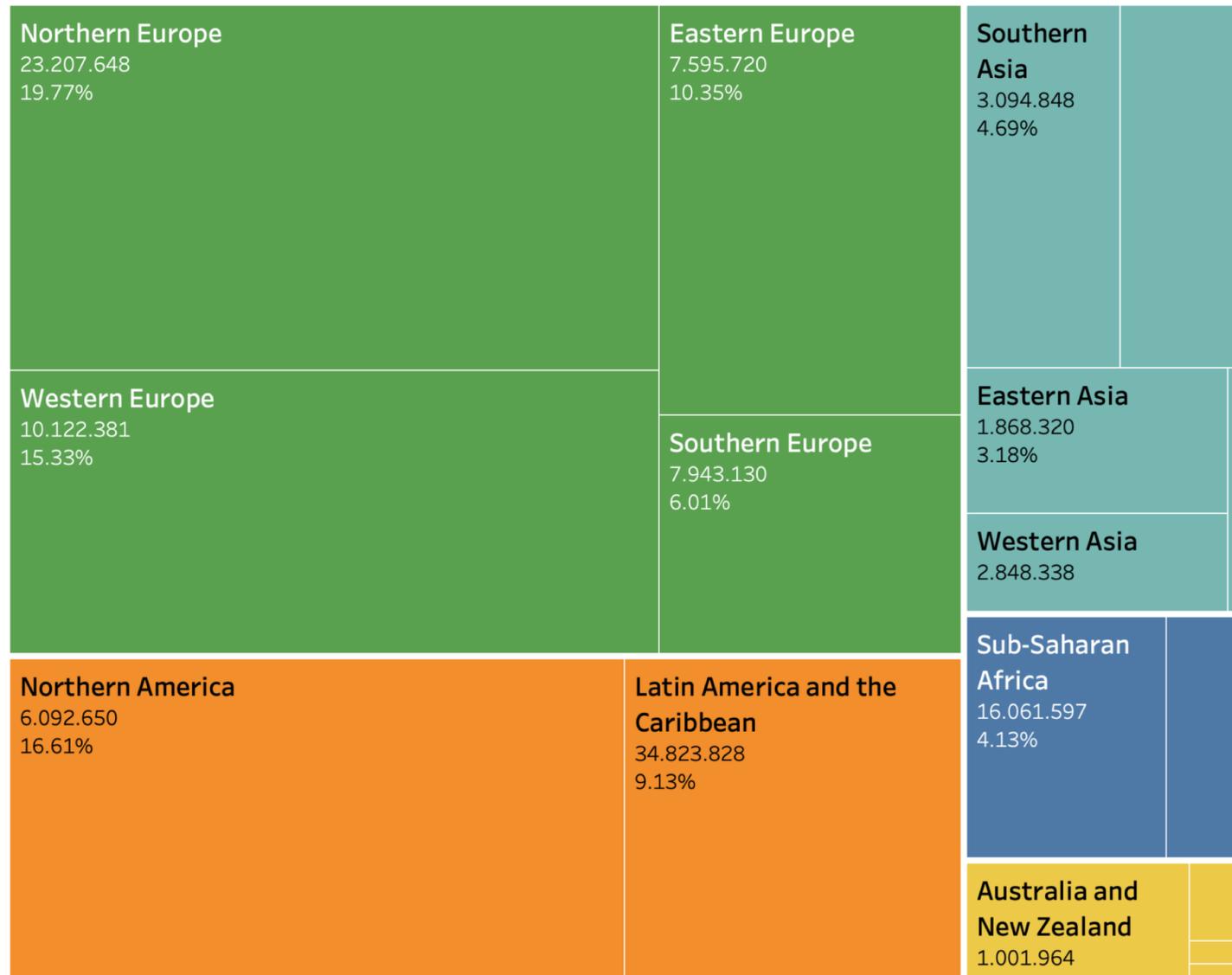


# **Wikimania 2018**



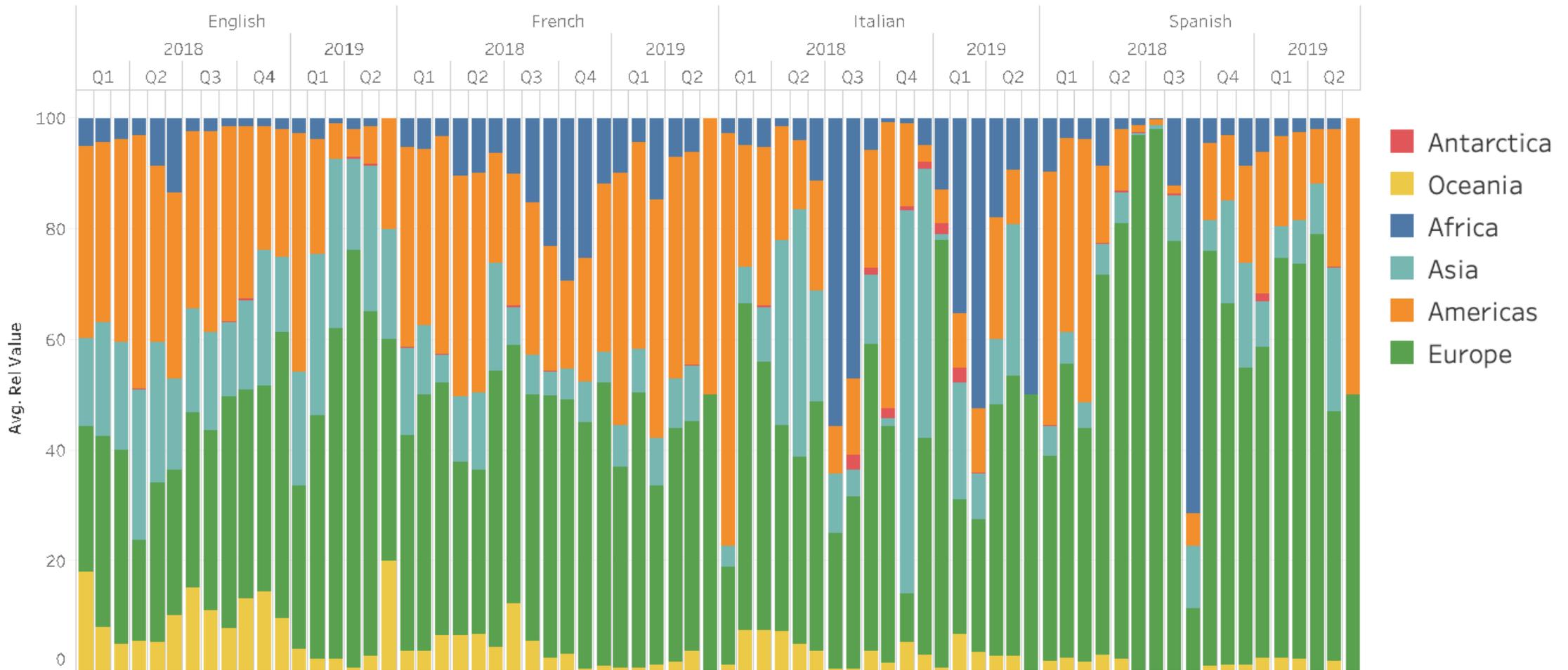
**Has the Wikimania 2018 theme really lead to bridging the knowledge gaps? (especially, the African one)**

# Wikidata geolocated items August 2019



- Antarctica
- Oceania
- Africa
- Asia
- Americas
- Europe

# What happened with African articles before and after the Wikimania?

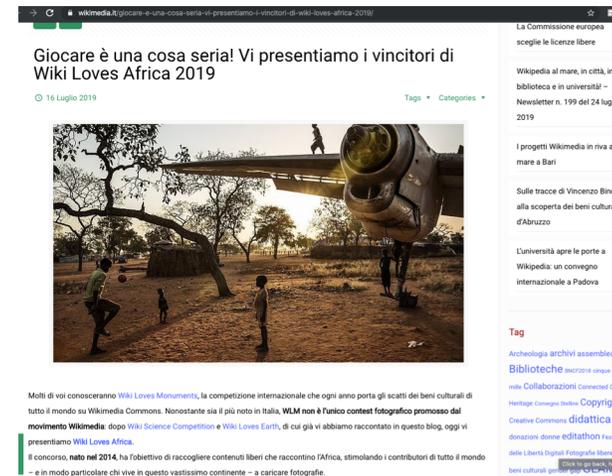


Wikimania (July 2018, Q3) has had different impact on major European languages.

# Wikimedia Italy has raised awareness for some time on Africa.



2006. WikiAfrica is launched by lettera27 Foundation in collaboration with Wikimedia Italia.



They make it visible in their communication.

# **Problem of Cultural Diversity**

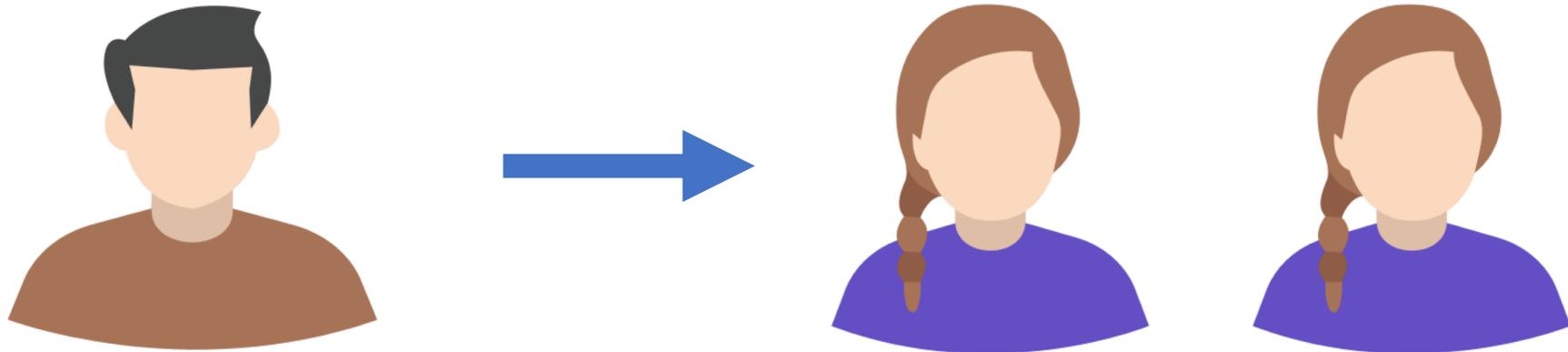
## The Problem

Wikipedia project does not reflect enough the world's cultural diversity.



What can we do?

**We cannot solve cultural diversity like we solve the content gender gap, creating two women for every man.**



## Proposed Solution

### Wikipedia Cultural Diversity Observatory (WCDO).

“a joint space for **researchers** and **activists** to study and **fight against the knowledge gaps** and increase cultural diversity in contents”.

Its work lines are:

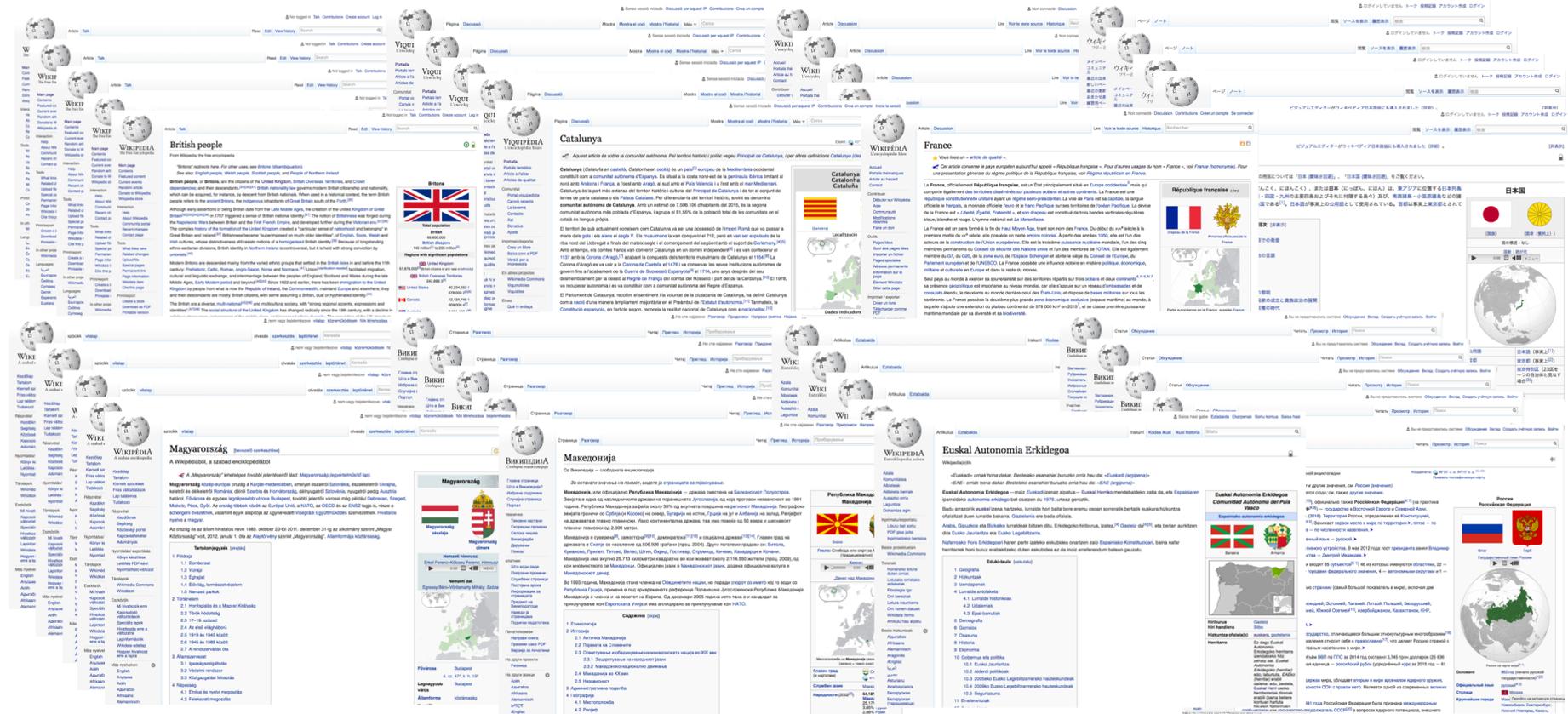
- Discourse (concepts and discussions)
- Awareness (metrics and observations)
- Organization (events and tools)
- Strategy (goals and priorities)

<http://wcdо.wmflabs.org>

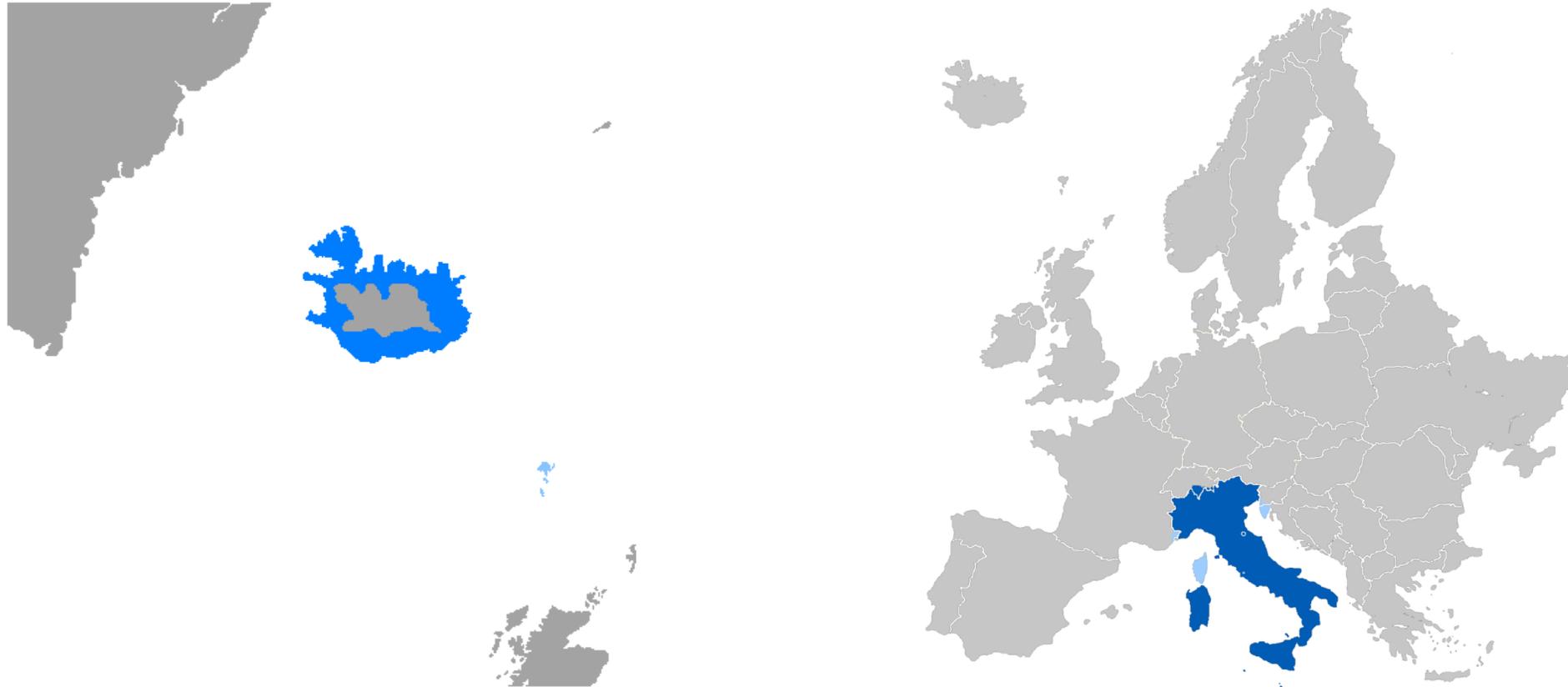




# Cartography



For each Wikipedia language edition, we aim at selecting the **Cultural Context Content (CCC)**, i.e. traditions, language, politics, agriculture, biographies, places, events, etcetera, related to the territories where the language is spoken.



**Icelandic Cultural Context** only relates to concepts from Iceland.

**Italian Cultural Context** includes articles about everything related to Italy, San Marino, Vaticano, Canton Ticino, Istria among others.

# Why is Wikipedia failing at gathering the human cultural diversity?

- Some languages, contexts and their concepts are not in Wikipedia. (**Representation**)
- Some are in Wikipedia but remain exclusive to some language editions. (**Sharing**)

				ሂ					
			ᄃ	ᄄ		ᄆ	ᄇ		
ᄈ	ᄉ	ᄊ	ᄋ	ᄌ	ᄍ	ᄎ	ᄏ	ᄐ	ᄑ
B	ᄒ	ᄓ	ᄔ	ᄕ	ᄖ	ᄗ	ᄘ	ᄙ	ᄚ
ᄛ	ᄜ	ᄝ	ᄞ	ᄟ	ᄠ	ᄡ	ᄢ	ᄣ	ᄤ
ᄥ	U	ᄦ	ᄧ	ᄨ	ᄩ	ᄪ	ᄫ	Ä	典
				ᄬ	V	ᄭ			
				ᄮ	ᄮ	ᄮ			

→ Missing pieces

# Method to collect Cultural Context Content

We created a method (Miquel-Ribé, 2017; Miquel-Ribé & Laniado, 2019) that requires (i) creating a database with **Language-Territories Mapping** and (ii) employing **different retrieval strategies** to extract content from each language edition and label it as CCC.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	territoryname	territorynameNative	QitemTerritory	languageName	Wiki	demon	demon	ISO3166	ISO31662	region	country	indl	lan	official	nu
1	Afar	Qafar	Q193494	Afar	aa			ET	ET-AF	yes	Ethiopia	yes	2	regional	0
2	Somali	Q202800	Q202800	Afar	aa			ET	ET-SO	yes	Ethiopia	yes	2	regional	0
3	Amhara	Q203009	Q203009	Afar	aa			ET	ET-AM	yes	Ethiopia	yes	2	regional	0
4	Ali Sabieh	Q821008	Q821008	Afar	aa			DJ	DJ-AS	yes	Djibouti	yes	5	no	0
5	Arta	Q705941	Q705941	Afar	aa			DJ	DJ-AR	yes	Djibouti	yes	5	no	0
6	Obock	Q844929	Q844929	Afar	aa			DJ	DJ-OB	yes	Djibouti	yes	5	no	0
7	Dikhil	Q283979	Q283979	Afar	aa			DJ	DJ-DI	yes	Djibouti	yes	5	no	0
8	Debubawi K'eyih	Q27728	Q27728	Afar	aa			ER	ER-DU	yes	Eritrea	yes	5	no	0
9	Semenawi K'eyi B	Q27910	Q27910	Afar	aa			ER	ER-SK	yes	Eritrea	yes			
10	Abkhazia	Аҧсны	Q23334	Abkhaz	ab	Abkhaz		GE	GE-AB	yes	Georgia	yes	2	regional	1
11	Aceh	Acèh	Q1823	Aceh	ace			ID	ID-AC	yes	Indonesia	yes	6	no	0
12	Sumatera Utara	Sumatra Baròh	Q2140	Aceh	ace			ID	ID-SU	yes	Indonesia	yes	6	no	0
13	Republic of Adygə	Адыгəй	Q3734	Adyghe	ady			RU	RU-AD	yes	Russian Federation	yes	2	regional	1
14	Krasnodar Krai	Краснодар край	Q3680	Adyghe	ady			RU	RU-KDA	yes	Russian Federation	yes	2	regional	1
15	Karachay-Cherkessia	Къарачае-Черкес	Q5328	Adyghe	ady			RU	RU-KC	yes	Russian Federation	yes	2	regional	1
16	South Africa	Suid-Afrika	Q258	Afrikaans	af	South Afri	Suid-Afrika	ZA		no	South Africa	yes	1	national	1
17	Central	Sentraal distrik	Q57525	Afrikaans	af			BW	BW-CE	yes	Botswana	yes	5	no	1
18	Ghanzi	Ghanzi	Q57571	Afrikaans	af			BW	BW-GH	yes	Botswana	yes	5	no	1
19	Kgalagadi	Kgalagadi	Q57581	Afrikaans	af			BW	BW-KG	yes	Botswana	yes	5	no	1
20	Kgatleng	Kgatleng	Q57593	Afrikaans	af			BW	BW-KL	yes	Botswana	yes	5	no	1
21	Southern	Suid distrik	Q57609	Afrikaans	af			BW	BW-SO	yes	Botswana	yes	5	no	1
22	Botswana	Botswana	Q963	Afrikaans	af	Motswana;Botswana	BW			no	Botswana	yes	5	no	1
23	Ghana	Ghana	Q117	Akan	ak	Ghanaian		GH		no	Ghana	yes	3	no	1
24	Switzerland	Schweiz	Q39	German, Swiss	als	Swiss		CH		no	Switzerland	yes	5	no	0
25	Vorarlberg	Vorarlberg	Q38981	German, Swiss	als			AT	AT-8	yes	Austria	yes	5	no	0
26	Champagne-Arde	Champagne-Ardenne	Q14103	German, Swiss	als			FR	FR-G	yes	France	yes	6	no	0
27	Lorraine	Lothringen	Q1137	German, Swiss	als			FR	FR-M	yes	France	yes	6	no	0
28	Alsace	Elsass	Q1142	German, Swiss	als			FR	FR-A	yes	France	yes	6	no	0
29	Baden-Württemb	Baden-Württemberg	Q985	German, Swiss	als			DE	DE-BW	yes	Germany	yes	5	no	0
30															

**Language Territories mapping spreadsheet with 1783 rows.**

(i) Wikidata Language Qitem, Language name, Language name in Native language, the ISO 639 code, the associated territories at country level (ISO 3166 code, English name, Native language name, demonym, Qitem) or at first subdivision (ISO 3166-2 code, English name, Native language name, demonym, Qitem) according to the information generated by Ethnologue.

[[https://wcd0.wmflabs.org/language\\_territories\\_mapping](https://wcd0.wmflabs.org/language_territories_mapping)]

- **Geolocation**
- **Category Graph**
- **Keywords in title**

- **Some Wikidata properties**

Location properties (location, located in administrative,...).

Country properties (country of citizenship, of origin).

Language properties (official language, native language...).

Affiliation properties (member of, educated at, employer,...).

Has part (contains administrative entity, has part).

Language properties (language of work, language used,...).

- **Link features:**

Number and percentage of Inlinks/Outlinks (incoming/outgoing links) to CCC is very explicative on how an article is needed to expand CCC or is dedicated to CCC.

**Machine Learning Classifier:**

We introduce it to a Random Forest classifier to obtain the final CCC dataset for each language edition. The manual assessment (blind) determined a 5%-5% false positive and false negatives.



**Japanese CCC**

# Cartography



For each Wikipedia language edition, we aim at selecting the **Cultural Context Content (CCC)**, i.e. traditions, language, politics, agriculture, biographies, places, events, etcetera, related to the territories where the language is spoken.

## Awareness (metrics and visualizations): extent of CCC

Taking into account the largest Wikipedia language editions, CCC is in average about a quarter of each Wikipedia (Miquel-Ribé and Laniado 2016).

CCC articles tend to be more developed in number of Bytes, images, and categories (Miquel-Ribé, 2017). Every language should represent their context properly (extent). This is healthy.

We have a **problem of representation** considering that the CCC extent in non-western languages (African and Asian) is on average much smaller (Miquel-Ribé and Laniado, 2019).



## Awareness (metrics and visualizations): gap between languages

About a 60% of the content language gaps are due to CCC (Miquel-Ribé and Laniado 2018). **Culture gap.**

Big languages like English or geographically close languages are the ones covering best the smaller languages (Miquel-Ribé and Laniado 2016).

We have a **problem of sharing** considering a Wikipedia language edition cultural diversity as the coverage of all the others' CCC.



## **Strategy (goals and priorities)**

**It is impossible to bridge all the knowledge gaps between languages.**

In the Cultural Diversity Observatory we propose every Wikipedia has 100 articles about every other language's cultural and geographical content.

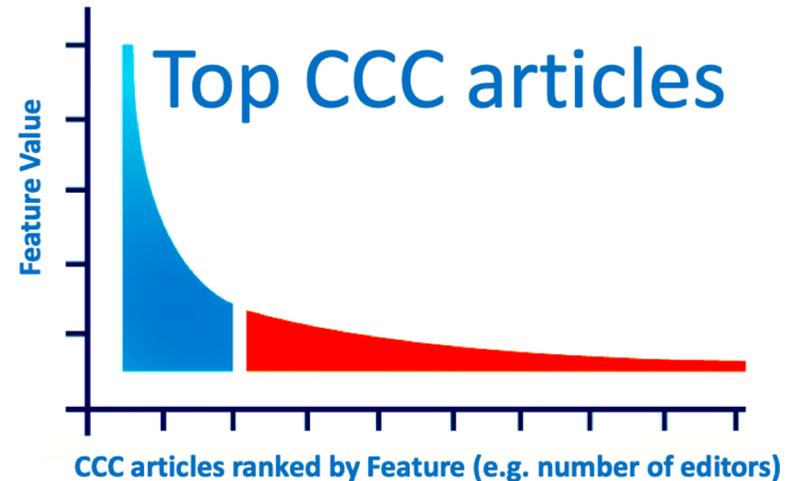
Every language should have this minimum. This is cultural diversity.

**28-30 thousand articles to cover a minimum of Wikipedia cultural diversity.**

## Organization (events and tools): Top CCC articles lists

From each language, those articles from their cultural context which are the **most relevant according to specific features**:

- List = [editors, featured, geolocated, keywords, women, men, created\_first\_three\_years, created\_last\_year, pageviews, discussions]
- Country\_origin (optional) = ISO3166 code
- Lang\_origin = wikicode
- Lang\_target = wikicode



[http://wcdo.wmflabs.org/top\\_ccc\\_articles/?list=men&lang\\_origin=pl&lang\\_target=uk](http://wcdo.wmflabs.org/top_ccc_articles/?list=men&lang_origin=pl&lang_target=uk)

[http://wcdo.wmflabs.org/top\\_ccc\\_articles](http://wcdo.wmflabs.org/top_ccc_articles)

# Top 500 CCC articles list "Women" from Swedish CCC in Catalan Wikipedia

Select the parameters

[Download Table \(Excel\)](#)

List:  × ▾

Language origin:  × ▾

Country origin:

Language target:  × ▾

Nº	Swedish Title	Edits	Editors	Pageviews	Bytes	References	Wikidata Properties	Interwiki Links	Inlinks from CCC	Creation Date	Other Languages	Catalan Title
1	<a href="#">Astrid Lindgren</a>	2269	928	1539	68.6k	96	94	97	546	2003-01-23	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Astrid Lindgren</a>
2	<a href="#">Selma Lagerlöf</a>	2225	926	1732	29.0k	78	107	125	426	2003-01-23	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Selma Lagerlöf</a>
3	<a href="#">Mona Sahlin</a>	1512	598	366	29.7k	62	34	28	309	2003-09-13	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<b><a href="#">Mona Sahlin</a></b>
4	<a href="#">Carola Häggkvist</a>	1384	646	644	42.8k	31	36	39	376	2004-03-14	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Carola Häggkvist</a>
5	<a href="#">Drottning Kristina</a>	1057	484	613	58.6k	41	62	63	581	2002-12-05	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Cristina de Suècia</a>
6	<a href="#">Amy Deasismont</a>	983	476	314	15.9k	34	29	24	3	2005-03-14	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<b><a href="#">Amy Deasismont</a></b>
7	<a href="#">Kronprinsessan Victoria</a>	957	454	1116	30.6k	48	37	58	161	2003-02-01	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Victòria de Suècia</a>
8	<a href="#">Loreen</a>	843	357	260	40.9k	95	39	63	244	2005-06-22	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Loreen</a>
9	<a href="#">Edda Magnason</a>	805	66	98	27.3k	55	23	9	19	2010-01-20	<a href="#">en</a> , <a href="#">fr</a>	<b><a href="#">Edda Magnason</a></b>
10	<a href="#">Agnes Carlsson</a>	770	400	362	38.7k	77	40	27	149	2005-12-02	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Agnes Carlsson</a>
undo	<a href="#">Prinsessan Madeleine</a>	754	338	650	17.2k	31	29	45	68	2001-07-28	<a href="#">es</a> , <a href="#">en</a> , <a href="#">fr</a> , <a href="#">it</a>	<a href="#">Magdalena de Suècia</a>

## Organization (events and tools): Panels to understand the coverage of Top CCC

### How do languages cover each others Top CCC articles?

These are panels to obtain a general view on the coverage and spread of the Top CCC.

- **Languages Top CCC articles coverage**

[https://wcdo.wmflabs.org/languages\\_top\\_ccc\\_articles\\_coverage/?lang=ca](https://wcdo.wmflabs.org/languages_top_ccc_articles_coverage/?lang=ca)

- **Languages Top CCC articles spread**

[https://wcdo.wmflabs.org/languages\\_top\\_ccc\\_articles\\_spread/?lang=ca](https://wcdo.wmflabs.org/languages_top_ccc_articles_spread/?lang=ca)

Lang = wikicode

# Catalan Wikipedia Top 100 CCC article lists spread across the rest of Wikipedias

This page shows some statistics that explain how well the first each Catalan Wikipedia Top 100 CCC articles (only the first 100) are spread across the other language editions.

These lists are created by ranking the articles according to specific features and sometimes giving them weights. These different features are usually based on the content type (e.g.

plain CCC or geolocated articles) or article characteristics (number of Bytes). The Top CCC articles lists are: list of CCC articles with most number of editors (**Editors**), list of CCC articles with featured article distinction (**Featured**), most bytes and references (weights: 0.8, 0.1 and 0.1 respectively), list of CCC articles with geolocation with most links coming from CCC, list of CCC articles with keywords on title with most bytes (**Bytes**), list of CCC articles categorized in Wikidata as women with most edits (**Women**), list of CCC articles categorized in Wikidata as men with most edits (**Men**), list of CCC articles created during the first three years and with most edits (**First 3Y**), list of CCC articles created during the last year and with most edits (**Last Y**), list of CCC articles with most pageviews during the last month (**Pageviews**), list of CCC articles with most edits in talk pages (**Discussions**).

The following table is useful in order to assess how well Catalan Wikipedia covers the Top 100 CCC articles from the lists generated from all the other language editions CCC. Languages are sorted in alphabetic order by their Wikicode, and columns present the number of articles from each list covered by English language. The last two columns, **Lists Coverage Idx.** and **Sum Covered Articles** present the percentage of articles from the lists covered and the overall sum of articles from the lists covered by Catalan Wikipedia for each language.

**The challenge is to reach 100 articles spread**

Language	Wiki	Editors	Featured
Afar	aa	0%	0%
Abkhaz	ab	4%	1%
Acehnese	ace	3%	2%
Adyghe	ady	2%	1%
Afrikaans	af	22%	7%
Akan language	ak	1%	0%
Alemannic	als	16%	4%
Amharic	am	18%	4%
Indonesian	an	71%	18%
Old English	ang	6%	3%

## Languages Top 100 CCC articles lists coverage by Catalan Wikipedia

This page shows some statistics that explain how well Catalan Wikipedia language edition covers the Top 100 of the Top CCC articles lists from other Wikipedia language editions.

These lists are created by ranking the articles according to specific features and sometimes giving them weights. These different features are usually based on the content type (e.g. plain CCC or geolocated articles) or article characteristics (number of Bytes). The Top CCC articles lists are: list of CCC articles with most number of editors (**Editors**), list of CCC articles with featured article distinction (**Featured**), most bytes and references (weights: 0.8, 0.1 and 0.1 respectively), list of CCC articles with geolocation with most links coming from CCC, list of CCC articles with keywords on title with most bytes (**Bytes**), list of CCC articles categorized in Wikidata as women with most edits (**Women**), list of CCC articles categorized in Wikidata as men with most edits (**Men**), list of CCC articles created during the first three years and with most edits (**First 3Y**), list of CCC articles created during the last year and with most edits (**Last Y**), list of CCC articles with most pageviews during the last month (**Pageviews**), list of CCC articles with most edits in talk pages (**Discussions**).

The following table is useful to assess how well Catalan Wikipedia covers the Top 100 CCC articles from the lists generated from all the other language editions CCC. Languages are sorted in alphabetic order by their Wikicode, and columns present the number of articles from each list covered by English language. The last two columns, **Lists Coverage Idx.** and **Sum Covered Articles** present the percentage of articles from the lists covered and the overall sum of articles from the lists covered by Catalan Wikipedia for each language.

**The challenge is to reach 100 articles covered (Sum Covered Articles) from each language CCC!**

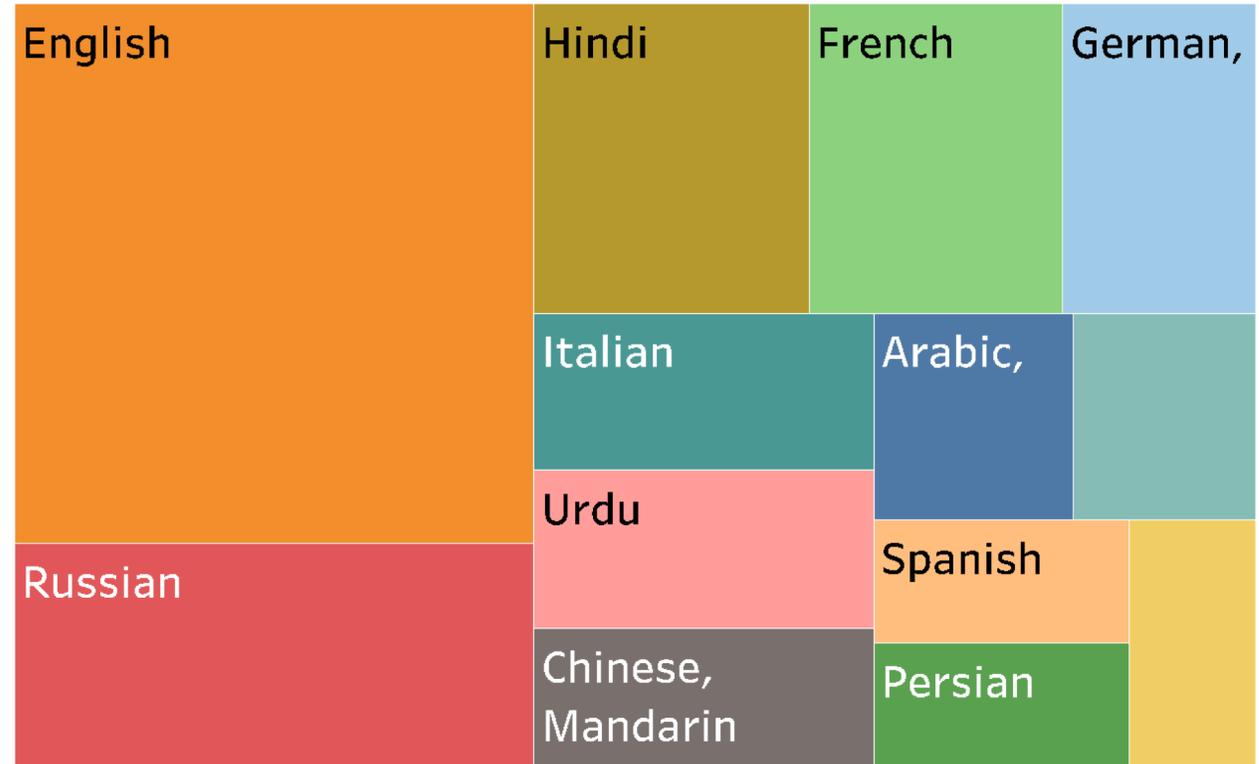
Language	Wiki	Editors	Featured	Geolocated	Keywords	Women	Men	First 3Y	Last Y	Page views	Talk Edits	List Coverage Idx.	Sum Covered Articles	World Subregion
Abkhaz	ab	7%	7%	4/87	2/13	0/0	2/11	5/9	1/20	8%	6%	12.7	8	Western Asia
Acehnese	ace	13%	0%	7%	5%	0/1	0/14	9%	0%	8%	4%	4.6	15	South-eastern Asia
Adyghe	ady	3/12	3/12	0/0	3/4	0/0	0/3	1/3	0/0	3/12	3/12	20.8	3	Eastern Europe
Afrikaans	af	79%	28%	20%	12%	15%	22%	63%	4%	39%	29%	31.1	145	Sub-Saharan Africa
Akan language	ak	11/57	11/57	4/29	2/2	1/4	4/13	7/16	0/0	11/57	11/57	29.1	11	Sub-Saharan Africa
Alemannic	als	83%	28%	88%	17%	12%	50%	79%	24%	61%	51%	49.3	295	Western Europe
Amharic	am	22%	9%	7/98	1/25	0/0	2/8	4/5	1/1	17%	12%	27.6	23	Sub-Saharan Africa
Indonesian	an	96%	69%	43%	38%	62%	89%	90%	23/61	66%	59%	65	427	Southern Europe
Old English	ang	78%	77%	18/24	5/7	5/6	37/41	29/37	2/2	81%	78%	81.2	85	Northern Europe

**Problem: smaller language editions do not even have 100 on their cultural context to fill the lists.**

## 253 out of the 302

Wikipedia languages are from languages which coexist with other languages with a higher social status.

English and Russian are the languages which overlap most with other Wikipedia languages (87 and 37) because they are official at country level.



## The inverse situation

In some African languages, cultural context content tends to be represented in English or French rather than in their native Wikipedia.

For instance, **Luganda Wikipedia** (from Uganda) has a very low CCC (**just 3.15%**).



- We can create lists of **missing CCC articles** in minoritized languages that exist in higher status languages and encourage them to create them.

For example: **Luganda language** coexists with **English** as a higher status language (“Status 3 wider communication” and “Status 1 national” correspondingly).

Luganda Wikipedia (1171 articles) lacks articles about:

- Geography
- Traditions
- Politicians
- Etc.

**President of Uganda does exist**

**Prime Minister of Uganda does not exist**

# Missing CCC Articles\* Alpha

## Missing CCC Articles List

### Select the Wikipedia

Target Language

Luganda (lg)

Target Country (Optional)

Select a country (optional)

Target Region (Optional)

Select a region (optional)

### Filter by content

Search Type

Missing articles

CCC Segment

Select a CCC Segment

Topic

Select a topic

### Choose source of content

Source language(s)

English (en)

Order by feature

Number of Inlinks

Limit the results

100

QUERY RESULTS!

N°	Language	Title	Number of Editors	Number of Pageviews	Number of Inlinks	Number of Bytes	Label (Lang)	Qitem
1	en	<a href="#">Nile</a>	0	9131	2825	0	en Nile	<a href="#">Q3392</a>
2	en	<a href="#">Finance Trust Bank</a>	0	42	589	0		<a href="#">Q7877595</a>
3	en	<a href="#">Central Region, Uganda</a>	0	123	568	0	en Central Region	<a href="#">Q429685</a>
4	en	<a href="#">Uganda Protectorate</a>	0	444	554	0		<a href="#">Q1097943</a>
5	en	<a href="#">Operation Entebbe</a>	0	6000	513	0	en Operation Entebbe	<a href="#">Q33933</a>
	en	<a href="#">Kampala–Jinja Highway</a>	0	34	439	0		<a href="#">Q6359092</a>
	en	<a href="#">Law Development Centre</a>	0	65	404	0		<a href="#">Q6503014</a>
8	en	<a href="#">Kampala–Hoima Road</a>	0	21	396	0		<a href="#">Q25226934</a>

**Conclusion**

## Proposed Solution

### Wikipedia Cultural Diversity Observatory (WCDO).

“a joint space for **researchers** and **activists** to study and **fight against the knowledge gaps** and increase cultural diversity in contents”.

Its work lines are:

- Discourse (concepts and discussions)
- Awareness (metrics and observations)
- Organization (events and tools)
- Strategy (goals and priorities)

<http://wcd0.wmflabs.org>



## Future ideas

# Wikipedia ~~Cultural~~ Diversity Observatory (WCDO).

Diversity Stats and Metrics would guide and help in covering the unrepresented *concepts* and *points of view*, related to all groups:

- Gender
- Culture
- LGBT+
- Historically marginalized communities,
- Among others.

**WG Diversity – Strategy 2030: Diversity Metrics Recommendation**

<https://tinyurl.com/diversitymetrics>

**We should help minoritized, minority and potential languages build a minimal capacity.**



Wikimedia Summit 2019, (Back row L-R) Rebecca, Jeffrey, Simona, Dariusz, Winifred, (Front row L-R) Liang, Oscar, Michal and Sailesh

**We should encourage them not only to build a Wikipedia,**

**But to create their most unique and specific knowledge about their context: few lists of 100 articles can really make a difference.**



Arctic char fishing. By Ansgar Walk.

**We need to value their knowledge and difference.**

**Each context has some ‘uniqueness’ to contribute to human knowledge that is best explained by their native languages.**

**Wikipedia is their best strategy to pass this on.**



# **Thank you very much!**

## **Wikipedia Cultural Diversity Observatory (WCDO)**

[<https://meta.wikimedia.org/wiki/WCDO>]

**Marc Miquel**

{[marcmiquel@gmail.com](mailto:marcmiquel@gmail.com)}

Username:marcmiquel

Pompeu Fabra University, Barcelona, Catalonia

Amical Wikimedia (Catalan Wikipedia)

**Wikimedia Foundation – Project Grant**



## References (if you want to know more)

Miquel-Ribé, M., & Laniado, D. **(2019)**. Wikipedia Cultural Diversity Dataset: A Complete Cartography for 300 Language Editions. *Proceedings of the 13th International AAI Conference on Web and Social Media. ICWSM. ACM.*

Miquel-Ribé, M., & Laniado, D. **(2018)**. Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions. *Frontiers in Physics, 5, 12.* (CC BY) Open Access.

Miquel-Ribé, M. **(2017)**. *Identity-based motivation in digital engagement: the influence of community and cultural identity on participation in wikipedia* (Doctoral dissertation, Universitat Pompeu Fabra).

Miquel-Ribé, M., & Laniado, D. **(2016)**. Cultural identities in wikipedias. In *Proceedings of the 7th 2016 International Conference on Social Media & Society* (p. 24). ACM.