CEE Spring:
why you want less but better articles

By Mykola Kozlenko (NickK)
Wikimedia Ukraine

CEE
WIKIMEDIA MEETING
TBILISI 2023

CONNECT
TO LEARN!
შეუერთდი სწავლას!

If you don't know what CEE Spring is…

- It is an annual contest for writing articles about countries and communities of Central and Eastern Europe
- Born during CEE Meeting 2014 in the Bochka bar in Kyiv
- First organised in spring 2015, annual since then
- Based (but not limited to) on lists of important articles about every community created by that community
- First centered around countries only, grew to cover minority communities (such as Romani, Sorbian or Western Armenian)
- Sometimes joined by non-CEE communities writing about CEE (English, Italian, Uzbek)…
- And it's a key example of CEE collaboration

CEE Spring in Ukraine

Almost always first in the CEE
- 2015: 630 articles

Series of first places:
- 2016: 1491 created, 91 improved
- 2017: 1118 created, 222 improved
- 2018: 2147 articles (vs 894 for 2nd)
- 2019: 2064 articles (2nd at 1223)
- 2020: 4059 articles (2nd at 1276)
- 2021: 1757 articles (2nd at 1566)

After a break in 2022, first place again:
- 2023: 945 articles (2nd at 799)

OK result, we had bigger contests before

Best contest ever, happy for the 1st place
We wanted to get more improved articles
Easy 1st place, but how to judge them all?
It's a mess, results only in December
Covid madness, we have too many articles!
Let's make it shorter to get less articles

We really want less but better articles!

What went wrong?

1. Article lists

3. Topics selection

5. Quality and judging

And what we added?

2. Adding sister projects

4. Partnerships for outreach

1. Article lists: what goes wrong?

Article lists are supposed to have around 100 most important articles about the community

- The notion of 100 most important articles is unclear: usually 'vital' articles are already covered, so lists are based on some other criteria
- For many communities, lists contain significantly more than 100 articles, often hitting the technical limit (~400 per page)
- It is unclear why some articles are even on the list, Wikimedians can translate a stub about an obscure Bulgarian scientist or Estonian writer with zero interwikis, but why is it important?
- On the other hand, for smaller languages even key articles are often missing. As a result, we got an article in Bashkir about a book by Taras Shevchenko but still no article about Taras Shevchenko himself
- What else?

1. Article lists: how we fixed it for us?

Level 1
- Limit ourselves to 100 articles. Every year there are community ideas for many more, they go to Extended lists.
- Try to pick articles that have a decent version explaining why the subject is notable in at least one major language, preferably English
- Define criteria or theme, and target distribution by category before picking articles, this helps to get consistent lists
- Do rotation, avoid having any article more than once in a few years

Level 2
- For each language, we picked top-10 red links and top-10 stubs
- Main inspiration: list of 500 vital articles about Ukraine
- Tools for gaps: PetScan, Pageviews, Wikipedia Diversity Observatory

Other ideas?

1. Article lists: how we fixed it for us?

Red links for Ukraine [edit]

These articles are important and popular (most vital topics, lots of pageviews, interwikis etc.), but are absent (red) in given wikis

az Azerbaijani [edit]

№ ⟺	Article ⟺	en	az	ba	be	be -t	bs	bg	crh	cs	de	el	eo	et	hr	hu	hy	hyw	ka	kk
1	President of Ukraine	+	–	+	+	+	+	+	+	+	+	+	–	+	+	–	+	–	+	+
2	Prime Minister of Ukraine	+	–	–	+	–	–	+	–	+	–	+	–	+	–	–	+	–	+	+
3	History of Ukraine	+	–	+	+	+	–	+	–	+	+	+	+	+	+	+	+	–	–	–
4	Ukrainian diaspora	+	–	+	+	+	–	+	–	–	–	–	–	+	–	–	–	–	–	–
5	Ivan Franko (1856–1916)	+	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	–	+	–
6	Saint Sophia Cathedral 🌐	+	–	–	+	+	–	+	–	+	+	+	+	+	+	+	+	–	+	–
7	varenyky	+	–	–	+	+	–	+	–	+	+	–	+	–	+	–	–	–	–	+
8	Hryhorii Skovoroda (1722–1794)	+	–	–	+	+	–	+	+	+	+	–	+	+	+	+	+	–	–	+
9	bandura	+	–	–	+	+	–	+	–	+	+	–	+	+	+	+	–	–	+	+
10	Peter Mogila (1596–1647)	+	–	–	+	+	–	+	–	–	+	–	–	–	+	–	–	–	–	–
Σ	→ Autolist ⧉ → map gallery ⧉ → timeline ⧉	10	1	4	10	9	2	10	2	8	8	5	6	7	9	4	7	0	5	5

1. Article lists: how we fixed it for others?

Level 1
- Copy lists from Meta to Ukrainian Wikipedia. We adapted the Lua module so that it displays links to Ukrainian Wikipedia, red or blue
- For red links, Wikidata label in Ukrainian is displayed, thus we added Ukrainian labels for as many articles as possible

Level 2
- For each country, add top-10 red links and top-10 stubs that are not a part of contest lists
- Both are based either on articles that have many pageviews or interwikis in source Wikipedia, or have many incoming red links / are among popular articles but are still stubs in Ukrainian.

Other ideas?

1. Article lists: how we fixed it for others?

Популярні статті [ред.]

Ці статті популярні (багато переглядів, редакторів, інтервіки) у відповідному мовному розділі, але короткі або відсутні в українському.

Відсутні

№	Стаття	укр.	англ.	пол.	нім.	рос.	be	be-t	bg	crh	cs	hr	hu	ro	sk	sr	tr	sq	Σ	Вікідані	st.	
1	Національний театр Албанії	-	+	+	+	-	-	-	-	-	-	-	-	+	-	-	+	+	6	Q597573	9	+
2	Петро Марко	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	4	Q3301264	12	+
3	Мітруш Кутелі	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	6	Q3302469	16	+
4	Арілена Ара	+	+	+	+	+	-	-	-	-	-	-	+	+	-	-	+	+	10	Q10417885	14	+
5	Монастирський конгрес	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	9	Q1290184	6	+
6	Перша албанська школа	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	3	Q13043261	11	+
7	Камза	+	+	+	+	+	-	-	-	-	+	-	-	-	-	-	+	+	11	Q605160	15	+
8	Томорр	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	7	Q2305090	10	+
9	Військово-повітряні сили Албанії	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	7	Q2085043	10	+
10	Іглі Таре	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	5	Q472159	15	+
Σ	→ Автосписок → карта галерея → по часу	2	10	7	10	6	0	0	4	0	3	1	5	3	0	2	5	10	68	сер: 7\4\40%	118	

Короткі

№	Стаття	укр.	англ.	пол.	нім.	рос.	be	be-t	bg	crh	cs	hr	hu	ro	sk	sr	tr	sq	Σ	Вікідані	st.	
1	Ісмаїл Кадаре	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	15	Q179695	25	+
2	Блеона Кереті	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	14	Q22137038	18	+
3	Мечеть Етем Бея	+	+	+	+	+	+	+	-	+	+	-	-	-	-	+	+	+	9	Q798610	12	+
4	Прізренська ліга	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	15	Q829597	3	+
5	Ельбасан	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	16	Q114257	15	+
6	Берат	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	16	Q170244	26	+
7	Буна	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	15	Q372490	15	+
8	Еді Рама	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	15	Q316901	32	+
9	Тоскський говір	+	+	+	+	-	-	-	+	+	+	+	+	-	+	+	+	+	13	Q180937	11	
10	Прапор Албанії	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	16	Q48409	9	+
Σ	→ Автосписок → карта галерея → по часу	10	10	10	10	10	9	8	9	0	10	9	8	8	4	9	10	10	144	сер: 14\8\85%	166	

2. Sister projects: why we want them?

By default, CEE Spring happens only on Wikipedia.
- We want to use it as an opportunity to develop sister projects
- It is a relatively easy topic to encourage new users join the project
- While we bridged most of the CEE gap on Wikipedia, CEE is still a gap on many sister projects, and we encourage existing users to bridge it.
- CEE topics are popular on sister projects, especially on Wikivoyage.

Concerned sister projects:
- Wikiquote, since 2019. Top result: 13 participants, 92 new and 31 improved (all 2021)
- Wikivoyage, since 2019. Top result: 5 participants (2023), 39 new (2020), 7 improved (2023)
- Wikisource: difficult as two copyrights apply (source and translation)
- Wikimedia Commons: discussed but failed because of unclear scope.

3. Topics selection: what goes wrong?

Ideally we want:
- Addressing major content gaps
- And improving popular but lower-quality articles

What we see after eight editions of CEE Spring:
- Owing to CEE Spring, some countries are not a gap anymore, while others still are (e.g. Poland is well covered, while Serbia is not)
- 'Easy' articles (typically easy to translate but not very important, e.g. a random member of Olympic relay) are popular among participants
- Lists and list-like articles (e.g. sports tables) are also popular
- Hard to encourage people to improve existing articles
- Even if you make a list of vital articles, no guarantee it will be written (e.g. still no article on Armenian diaspora, and Chacha still a stub)

Anything else?

3. Topics selection: how we tried to fix it?

Level 1: within countries
- Adjust the scoring system to:
    - Encourage writing articles from contest lists
    - Encourage improving existing articles
    - Discourage creating list articles
- Add to article lists top-10 red links and top-10 stubs per country or community
- Promote 'interesting' articles (both articles for creation and created articles) on social media
- Work with partners on article lists suggestions for selected countries

Other ideas?

3. Topics selection: how we tried to fix it?

Level 2: geographically
- We realised that many CEE countries are not gaps anymore. Instead, they are among the best covered countries in Ukrainian Wikipedia.
- We rebranded the contest 'European Spring', and in 2023 we expanded it from CEE countries only to the entire Council of Europe.
- We introduced a coefficient per country for scoring depending on how well it is covered on Ukrainian Wikipedia, based on:
  - Number of articles on Ukrainian Wikipedia with P17 or P27 equal to this country (or articles in the category but it's less accurate)
  - Compared to the number of corresponding articles on English Wikipedia
  - And compared to the country's population

Other ideas?

3. Topics selection: how we tried to fix it?

Країна	Коеф.
Бельгія	2.0
Велика Британія	2.0
Данія	2.0
Нідерланди	2.0
Сербія	2.0
Туреччина	2.0
Азербайджан	1.7
Угорщина	1.7
Чехія	1.7
Албанія	1.5
Боснія і Герцеговина	1.5
Ірландія	1.5
Норвегія	1.5
Фінляндія	1.5

Греція	1.3
Кіпр	1.3
Німеччина	1.3
Швеція	1.3
Австрія	1.1
Іспанія	1.1
Італія	1.1
Швейцарія	1.1
Ісландія	1.0
Люксембург	1.0
Монако	1.0
Португалія	1.0
Андорра	0.7
Вірменія	0.7
Латвія	0.7
Литва	0.7

Мальта	0.7
Естонія	0.5
Північна Македонія	0.5
Сан-Марино	0.5
Франція	0.5
Чорногорія	0.5
Болгарія	0.3
Грузія	0.3
Ліхтенштейн	0.3
Молдова	0.3
Польща	0.3
Румунія	0.3
Словаччина	0.3
Словенія	0.3
Хорватія	0.3

4. Partners: why we want them?

What we want to get:
- The majority of CEE Spring participants are experienced Wikimedians who are willing to contribute on random topics. We want to diversify the audience and reach people who are interested in CEE countries but did not contribute before
- For countries where most article lists are covered or there are no article lists, partners can provide helpful suggestions
- Additional (small) prizes can attract more people

Who we worked with
- National partners, such as Austrian Kooperationsbüro or Czech Centre
- Diaspora or local cultural centres for workshops
- Communication partners (press releases, social media)

Special mention to the Ukrainian Institute for the UCDM

Other ideas?

5. Quality and judging: what goes wrong?

Ideally we want:
- Contest increasing and not reducing the quality of articles
- High-quality articles, perhaps good or featured
- Easy to score with clear feedback to participants

What CEE Spring often gives us is:
- Predominance of ContentTranslate, including badly proofread machine translations
- Very few users writing original articles with own analysis of sources
- Little to no collaboration (up to two users creating two articles about the same topic)
- Preference for easy articles, like lists or short articles
- Extreme numbers of articles making it difficult to judge

Anything else?

5. Quality and judging: how we tried to fix it?

Scoring system:
- Since the beginning, our judging is based on sum(size x quality) of all articles. This means we evaluate quality of each article.
- Additional prize for the best article (priority for good/featured articles).
- For size, articles less than 3500 bytes do not bring any points.
- In addition to the topic coefficients (see 3), our scoring gives:
  - More points for original articles
  - Less points for translations, up to zero if not proofread
- Disqualification for copyright violations
- Less focus on formatting, more focus on quality of content
- Since 2023: to encourage writing less but better articles, we multiply each article's score by 0.99^it's number (e.g. 20th : 0.99^20 = 0.818)

Other ideas?

5. Quality and judging: how we tried to fix it?

Judging:
- Ideal world (tried in early years): we want judging on the go.
  - Once the article is ready, the jury evaluates it and gives feedback to participants during the contest.
  - Participants can improve it and it submit again for final review
  - Great for helping newcomers, very difficult for judges
- Less ideal world: we have up to 4500 articles, we need to judge them
  - Big enough jury, at most 150 articles per judge
  - Transparent and straightforward scoring guidelines
  - Only score without public comments to simplify the process
  - Tool for judging (in peacetime) https://wikigrade.toolforge.org/
  - We try to provide feedback but have only capacity for newbies

Other ideas?

Free discussion...

Place to share your ideas

Or suggest another topic to discuss

Or tell about your CEE Spring experience

...and thanks!

mykola.kozlenko@wikimedia.org.ua
ceespring@wikimedia.org.ua