



Wikipedia:Wikipedia Signpost/2019-02-28/Recent research

< [Wikipedia:Wikipedia Signpost](#) | [2019-02-28](#)

Recent research

Research finds signs of cultural diversity and recreational habits of readers

By [Andrew Krizhanovsky](#), [Bri](#) and [Tilman Bayer](#)



A monthly overview of recent academic research about Wikipedia and other Wikimedia projects, also published as the [Wikimedia Research Newsletter](#).

Wikipedia is a second screen for Internet TV watchers

Reviewed by [Bri](#)

In this study^[1] it was shown that Wikipedia pageviews follow Internet TV coverage (technically, there was a high [Pearson correlation](#) between [viewership](#) and pageviews). The researchers were able to use data mining of anonymized [AbemaTV](#) logs to correlate individual viewers to Wikipedia pageviews changes. They could even determine [second screen](#) behavior: viewers were watching programs and reading related articles in near-realtime after 23:00 (viewer's time) – taking measures to exclude [channel surfing](#) viewers. This reviewer finds the use of Wikipedia as an adjunct or research tool for television watching is somewhat at odds with a naive assumption of an either-or media consumption model, and could help explain some of our media-centric [Top 25](#) article views.

Wikidata calculates cultural diversity

Reviewed by [Andrew Krizhanovsky](#)

This research^[2] deals with **Cultural Context Content** (CCC), that is articles of Wikipedia "related to the editors' geographical and cultural context (i.e. their places, traditions, language, agriculture, biographies, etc.)". CCC is about 25% of articles in one Wikipedia edition. These articles are exclusive and have no equivalence across language editions.

There are thousands of Wikidata properties. So, this research paper can be useful for learning Wikidata, because it describes Wikidata properties grouped by country, location, language, author, affiliation and several taxonomic relations (part of, has part). The future research of this paper could be directed to the visualization of some results via Wikidata Query Service (<https://query.wikidata.org>) (see examples in the Wikiversity course "Research in programming Wikidata").

Very interesting results of this work are

- The mapping of language to territories (https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Language_Territories_Mapping).
- Number of unique articles and intersections with other Wikipedias ([https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Culture_Gap_\(spread\)](https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Culture_Gap_(spread))).
- Nearest Wikipedia or which Wikipedias cover this Wikipedia ([https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Culture_Gap_\(coverage\)](https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Culture_Gap_(coverage))) (e.g. for the Tamil Wikipedia, the five other language editions which cover most of its articles are, in descending order, English, Hindi, Arabic, Simple English, and Urdu, as of October 2018).

See also below for a list of other papers related to the "Wikipedia Cultural Diversity Observatory", by the same authors

Briefly

By Tilman Bayer

Wikimedia research whitepapers

The Wikimedia Foundation's Research team has announced (<https://wikimediafoundation.org/2019/02/14/research-directions-towards-the-wikimedia-2030-strategy/>) three white papers outlining research plans and priorities for the next five years, on the strategic themes of "knowledge gaps", "knowledge integrity" and "foundations".

Dario Taraborelli and Erik Zachte leave Wikimedia Foundation

Two longtime employees known for their work on statistics and research about Wikimedia projects departed the Wikimedia Foundation this month. Erik Zachte, who in 2003 began to work as a volunteer on what would become known as Wikistats (<https://stats.wikimedia.org/>), and in 2008 took up paid work as data analyst for WMF, announced (http://infodisiac.com/back_to_volunteer_mode.htm) his retirement (<https://lists.wikimedia.org/pipermail/wiki-research-l/2019-February/006607.html>). Dario Taraborelli, the Foundation's head of research (also co-founder of this research report and its associated @wikiresearch Twitter feed (<https://twitter.com/wikiresearch>)).

er.com/wikiresearch)) announced (<https://lists.wikimedia.org/pipermail/wiki-research-l/2019-February/006655.html>) his departure after eight years to do open science work at a different organization.

Conferences and events

See the research events page on Meta-wiki for upcoming conferences and events, including submission deadlines, and the page of the monthly Wikimedia Research Showcase for videos and slides of past presentations.

Other recent publications

Other recent publications that could not be covered in time for this issue include the items listed below. Contributions are always welcome for reviewing or summarizing newly published research.

Compiled by Tilman Bayer

"Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions"

From the abstract:^[3] "...we developed a computational method to identify articles that can be related to the editors' cultural context associated to each Wikipedia language edition. We employed a combination of strategies taking into account geolocated articles, specific keywords and categories, as well as links between articles. [...] The results show that about a quarter of each Wikipedia language edition is dedicated to represent the corresponding cultural context. Although a considerable part of this content was created during the first years of the project, its creation is sustained over time." (see also Wikidata calculates cultural diversity" above)

"Identity-based motivation in digital engagement: the influence of community and cultural identity on participation in Wikipedia"

From the abstract:^[4] "By analysing data from 15 language editions, I find that editors develop a community identity in Wikipedia and at the same time they consistently create content representing their cultural identities. Such content occupies around a quarter of each Wikipedia in number of articles, and even more in terms of edits. When editors increase their participation or become administrators, they still prefer editing content imbued with identity-based meanings, which suggests their centrality in the editing process." (see also "Wikidata calculates cultural diversity" above)

"Cultural Identities in Wikipedias"

From the abstract: ^[5] "... we developed a computational method to identify articles related to the cultural identities associated to a language and applied it to 40 Wikipedia language editions. The results show that about a quarter of each Wikipedia language edition is dedicated to represent the corresponding cultural identities. The topical coverage of these articles reflects that geography, biographies, and culture are the most common themes,

although each language shows its idiosyncrasy and other topics are also present. [...] An analysis of how this content is shared among language editions reveals special links between cultures." (see also "*Wikidata calculates cultural diversity*" above)

Students favor Wikipedia in English even when it's not their main language

From the abstract:^[6] "Seventy-seven first-year audiovisual communication students [most with Catalan and/or Spanish as their main languages] made contributions to Wikipedia as part of the assessed work in the first year course titled 'Digital Culture.' Before and after writing Wikipedia articles, the students responded to two questionnaires that enquired about their language-related habits when using the site and about their language choice for contributing to it. ... Students favor the English edition of Wikipedia when consulting it despite the fact that this is the language they assess themselves as being less proficient at in reading. More generally, our research shows that multilingual Wikipedia users move seamlessly from one language edition to another, thus refuting the cliché that relates minority languages with exclusively local and self-referential topics."

"Studying the Effect of Network Position on Efficiency"

From the abstract:^[7] "We gathered data about 2978 Wikipedia featured article editing history. We use degree centrality for both article affiliation network and editor's affiliation network. [...] This study finds that article degree centrality [has a] negative effect the collaboration process, that suggests that article linkages with other article attracts diverse knowledge bases and editors, that require time and consensus for further moving the editing process. Editors maximum centrality in a single article editor affiliation network have a positive effect on efficiency of an article, while editor maximum degree centrality in a multi-article affiliation network have negative effect on editing process efficiency."

"Collaborative Approach to Developing a Multilingual Ontology: A Case Study of Wikidata"

From the abstract:^[8] "In this article, Wikidata has been taken as an example to understand how community-driven approach is used to develop a multilingual ontology and in the subsequent building of a knowledge base."

"Exploring Translators' Expectations of Wikipedia: A Qualitative Review"

From the abstract: ^[9] "This paper's goal is to explore potential uses that translators could expect from Wikipedia. [...] We have concluded that translators might use Wikipedia expecting to find linguistic, semantic, terminological, lexicographic and cultural information."

"Wikipedia as a translation zone. A heterotopic analysis of the online encyclopedia and its collaborative volunteer translator community"

This article^[10] is a case study focused on the construction of the English Wikipedia article about Tokyo.

"Discussing Wartime Collaboration in a Transnational Digital Space: The Framing of the UPA and the Latvian Legion in Wikipedia"

From the abstract:^[11] "[This book chapter] investigates the different framing strategies used to represent wartime memories in Wikipedia, the ways these strategies are developed by local editors' communities, and the reception of Wikipedia's representations of the past by national and transnational audiences. The chapter concludes with a discussion of the different forms of consensus used in Wikipedia for dealing with contentious past as well as the promises and dangers of using digital media for transnational history writing."


"Translation and the Production of Knowledge in 'Wikipedia': Chronicling the Assassination of Boris Nemtsov"






From the abstract:^[12] "Based on a set of articles about the assassination of Russian politician Boris Nemtsov from nine different editions of the encyclopaedia, the article examines the place of translation in Wikipedia and the role it plays in knowledge production. Each of the articles is likely to use a number of different information sources, including other Wikipedia articles that are already in existence, with translation contributing to knowledge production as each new article evolves. ..."


"Locating foci of translation on Wikipedia"

From the abstract:^[13] "... it is generally accepted that most Wikipedia content is the product of original writing rather than being translated from another language version of the encyclopaedia. [...] The main aim of this paper is to make a number of proposals towards a possible methodology for discovering where the main foci of this new type of collaborative translation are located. Significant methods for this include the use of the encyclopaedia's list-based structure and of different features of page anatomy. The article [is] using Russian and Chinese to English translation as its main sources of examples."

References

1. Hayano, H.; Takano, M.; Morishita, S.; Yoshida, M.; Umemura, K. (December 2018). "Analysis of the Influence of Internet TV Station on Wikipedia Page Views". *2018 IEEE International Conference on Big Data (Big Data)*. 2018 IEEE International Conference on Big Data (Big Data). pp. 4328–4332. doi:10.1109/BigData.2018.8622239 (<https://doi.org/10.1109%2FBigData.2018.8622239>).  Freely available version (<https://aminer.org/archive/analysis-of-the-influence-of-internet-tv-station-on-wikipedia-page-views/5bff525d28686d2064b2da16>) on Aminer
2. Miquel-Ribé, Marc; Laniado, David (2019-01-23). "Wikipedia Cultural Diversity Dataset: A Complete Cartography for 300 Language Editions". arXiv:1901.07999 (<https://arxiv.org/abs/1901.07999>) [cs.CY (<https://arxiv.org/archive/cs.CY>)].

3. Miquel-Ribé, Marc; Laniado, David (2018). "Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions" (<https://doi.org/10.3389/fphy.2018.00054>). *Frontiers in Physics*. 6: 54. Bibcode:2018FrP....6...54M (<https://ui.adsabs.harvard.edu/abs/2018FrP....6...54M>). doi:10.3389/fphy.2018.00054 (<https://doi.org/10.3389/fphy.2018.00054>). ISSN 2296-424X (<https://www.worldcat.org/issn/2296-424X>).
4. Miquel Ribé, Marc (2017-03-24). "Identity-based motivation in digital engagement: the influence of community and cultural identity on participation in Wikipedia". *TDX (Tesis Doctorals en Xarxa)*. hdl:10230/32435 (<https://hdl.handle.net/10230/32435>).
5. Miquel-Ribé, Marc; Laniado, David (2016). "Cultural Identities in Wikipedias". *Proceedings of the 7th 2016 International Conference on Social Media & Society*. SMSociety '16. New York, NY, USA: ACM. pp. 24–1–24:10. doi:10.1145/2930971.2930996 (<https://doi.org/10.1145/2930971.2930996>). ISBN 9781450339384.  Freely available version (<https://www.researchgate.net/publication/304919040>)
6. Soler-Adillon, Joan; Freixa, Pere (2017-12-15). "Wikipedia access and contribution: Language choice in multilingual communities . A case study" (<https://analisi.cat/article/view/n57-soler-adillon-freixa>). *Anàlisi* (57): 63–80. doi:10.5565/rev/analisi.3109 (<https://doi.org/10.5565/rev/analisi.3109>). ISSN 2340-5236 (<https://www.worldcat.org/issn/2340-5236>). 
7. Khan, Naveed (February 2018). "Studying the Effect of Network Position on Efficiency: : A Case of Affiliation Network Featured Article Promotion" (<http://repository.hanyang.ac.kr/handle/20.500.11754/68965>). `{{cite journal}}: Cite journal requires |journal= (help)` (PhD thesis, Hanyang University)
8. Samuel, John (2017-11-28). *Collaborative Approach to Developing a Multilingual Ontology: A Case Study of Wikidata*. Research Conference on Metadata and Semantics Research. Communications in Computer and Information Science. Springer, Cham. pp. 167–172. doi:10.1007/978-3-319-70863-8_16 (https://doi.org/10.1007/978-3-319-70863-8_16). ISBN 9783319708621. 
9. Alonso, Elisa; Robinson, Bryan J. (2016-10-05). "Exploring Translators' Expectations of Wikipedia: A Qualitative Review". *Procedia - Social and Behavioral Sciences*. International Conference; Meaning in Translation: Illusion of Precision, MTIP2016, 11-13 May 2016, Riga, Latvia. 231: 114–121. doi:10.1016/j.sbspro.2016.09.079 (<https://doi.org/10.1016/j.sbspro.2016.09.079>). ISSN 1877-0428 (<https://www.worldcat.org/issn/1877-0428>).
10. Jones, Henry (20 December 2018). "Wikipedia as a translation zone" (https://publication.s.aston.ac.uk/id/eprint/41627/1/HETEROTOPIA_Wikipedia_as_Translation_Zone_Author_Approved_Manuscript_.pdf) (PDF). *Target*. 31: 77–97. doi:10.1075/target.18062.jon (<https://doi.org/10.1075/target.18062.jon>). S2CID 91384183 (<https://api.semanticscholar.org/CorpusID:91384183>).
11. Kaprāns, Mārtiņš; Makhortykh, Mykola (2018). "Discussing Wartime Collaboration in a Transnational Digital Space: The Framing of the UPA and the Latvian Legion in Wikipedia". *Traitors, Collaborators and Deserters in Contemporary European Politics of Memory*. Palgrave Macmillan Memory Studies. Palgrave Macmillan, Cham. pp. 169–195. doi:10.1007/978-3-319-66496-5_7 (https://doi.org/10.1007/978-3-319-66496-5_7). ISBN 9783319664958. 
12. Shuttleworth, Mark; (2018) شتورث, مارك. "Translation and the Production of Knowledge in "Wikipedia": Chronicling the Assassination of Boris Nemtsov / الترجمة وإنتاج المعرفة "على" (ويكيبيديا)): توثيق اغتيال بوريس نيمتسوف" (<https://www.jstor.org/stable/26496376>). *Alif: Journal of Comparative Poetics* (38): 231–263. ISSN 1110-8673 (<https://www.worldcat.org/issn/1110-8673>). JSTOR 26496376 (<https://www.jstor.org/stable/26496376>). 

13. Shuttleworth, Mark (2017-12-04). "Locating foci of translation on Wikipedia". *Translation Spaces*. **6** (2): 310–332. doi:10.1075/ts.6.2.07shu (<https://doi.org/10.1075%2Fts.6.2.07shu>). ISSN 2211-3711 (<https://www.worldcat.org/issn/2211-3711>). 

← *PREVIOUS "Recent research"*

NEXT "Recent research" →

+ *Add a comment* (https://en.wikipedia.org/w/index.php?editintro=Wikipedia:Wikipedia_Signpost/Templates/Comment-editnotice&title=Wikipedia_talk:Wikipedia_Signpost/2019-02-28/Recent_research&action=edit&preload=Wikipedia:Signpost/Templates/Signpost-article-comments-end/preload)

THESE COMMENTS ARE AUTOMATICALLY TRANSCLUDED FROM THIS ARTICLE'S TALK PAGE. TO FOLLOW COMMENTS, ADD THE PAGE TO YOUR WATCHLIST ([HTTPS://EN.WIKIPEDIA.ORG/W/INDEX.PHP?TITLE=WIKIPEDIA_TALK:WIKIPEDIA_SIGNPOST/2019-02-28/RECENT_RESEARCH&ACTION=WATCH](https://en.wikipedia.org/w/index.php?title=Wikipedia_talk:Wikipedia_Signpost/2019-02-28/Recent_research&action=watch)). IF YOUR COMMENT HAS NOT APPEARED HERE, YOU CAN TRY .

Thank you for the informative update! [Rachel Helps \(BYU\) \(talk\) 17:44, 4 March 2019 \(UTC\)](#)

IN THIS ISSUE



28 FEBRUARY 2019 (ALL COMMENTS)

- [From the editors](#)
- [News and notes](#)
- [In focus](#)
- [Discussion report](#)
- [Featured content](#)
- [Arbitration report](#)
- [Traffic report](#)
- [Technology report](#)
- [Gallery](#)
- **Recent research**
- [Essay](#)
- [From the archives](#)
- [Humour](#)

WHAT DO YOU THINK OF *THE SIGNPOST*? SHARE YOUR FEEDBACK.

[Home](#)

[About](#)

[Archives](#)

[Newsroom](#)

[Subscribe](#)

[Suggestions \(https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/Newsroom/Suggestions\)](https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/Newsroom/Suggestions)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Wikipedia:Wikipedia_Signpost/2019-02-28/Recent_research&oldid=1193878365"

■