



Wikipedia:Wikipedia Signpost/2021-02-28/Recent research

< [Wikipedia:Wikipedia Signpost](#) | [2021-02-28](#)

Recent research

Take an AI-generated flashcard quiz about Wikipedia; Wikipedia's anti-feudalism

By [Tilman Bayer](#) and [Miriam Redi](#)



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

"WikiFlash: Generating Flashcards from Wikipedia Articles"

Reviewed by [Tilman Bayer](#)

Flashcards are a popular method for memorizing information. A paper^[1] by six Zurich-based researchers, presented earlier this month at the annual AAAI conference, describes a tool to automatically extract flashcards from Wikipedia articles, aiming "to make independent education more attractive to a broader audience."

A proof-of-concept version is available online (<https://flashcard.ethz.ch/>), with results available for export in a format that can be used with the popular flashcard software Anki. User can choose from four different variants based on either the entire Wikipedia article or just its introductory section.

The researchers emphasize that "generating meaningful flashcards from an arbitrary piece of text is not a trivial problem" (also concerning the computational effort), and that there is currently no single model that can do this. They separate the task into four stages, each making use of existing NLP techniques:

- summarization, to first extract the most relevant information from Wikipedia (the user can also choose to have this step skipped and instead generate flashcards based on the full text)
- answer identification, where a model extracts answer statements from a given sentence based on context information from the surrounding paragraph

- question generation, where a model constructs a question from the statement generated in the previous step, again taking context information from the surrounding paragraph into account
- To improve quality, these are followed by a final filtering step, where a question-answering model tries to reconstruct the answer based on the paragraph from which the question was extracted, and the generated flashcard is discarded if the reconstructed answer does not overlap enough with the pre-generated answer.

Apart from evaluating the results using quantitative text measures, the researchers also conducted a user study to compare the output of their tool to human-generated flashcards from two topic areas, geography and history, rated by helpfulness, comprehensibility and perceived correctness. The "results show that in the case of geography there is no statistically meaningful difference between human-created and our cards for either of the three aspects. For history, the difference for helpfulness and comprehensibility is statistically significant ($p < 0.01$), with human cards being marginally better than our cards. Neither category revealed a statistically significant difference in perceived correctness." (However, the sample was rather small, with 50 Mechanical Turk users split into two groups for geography and history.)

A quick test of the tool with the article Wikipedia (introduction only) yielded the following result (text reproduced without changes):

Question: What does Wikipedia use to maintain it's [sic] content?

Answer

wiki-based editing system

Question: In 2021, where was Wikipedia ranked?

Answer

13th

Question: What language was Wikipedia initially available in?

Answer

English

Question: How many articles are in English version of Wikipedia [sic] as of February 2021?

Answer

6.3 million

Question: Who hosts Wikipedia?

Answer

Wikimedia Foundation

Question: Whose vision did *Time* magazine believe made Wikipedia the best encyclopedia in the world?

Answer

Jimmy Wales

Question: What is a systemic bias on Wikipedia?

Answer

gender bias

Question: What did Wikipedia receive praise for in the 2010s?

Answer

unique structure, culture, and absence of commercial bias

Question: What two social media sites announced in 2018 that they would help users detect fake news by suggesting links to related Wikipedia articles?

Answer

Facebook and YouTube

Briefly

- See the [page of the monthly **Wikimedia Research Showcase**](#) for videos and slides of past presentations.
- [@WikiResearch](https://twitter.com/WikiResearch) (<https://twitter.com/WikiResearch>), the Twitter feed associated with this monthly research update, [celebrated](https://twitter.com/WikiResearch/status/1365991638867144705) (<https://twitter.com/WikiResearch/status/1365991638867144705>) its ninth anniversary today. Over the past 9 years, we have shared on average 1.9 tweets per day about Wikimedia-related research. The feed is also available in syndicated form on [Facebook](https://www.facebook.com/WikiResearch/) (<https://www.facebook.com/WikiResearch/>) and [Mastodon](https://mastodon.social/@wikiresearch) (<https://mastodon.social/@wikiresearch>).

Other recent publications

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

Compiled by Tilman Bayer and Miriam Redi

Wikipedia's "sophisticated democracy" resists the "implicit feudalism" of online communities

A paper in *New Media & Society*^[2] argues that

"[...] an 'implicit feudalism' informs the available options for community management on the dominant platforms for online communities. It is a pattern that grants user-administrators absolutist reign over their fiefdoms, with competition among them as the primary mechanism for quality control, typically under rules set by platform companies.

[...] the online encyclopedia Wikipedia operates through a sophisticated democracy among active volunteers. Wikipedia also possesses a widely acknowledged benevolent dictator in the person of founder Jimmy Wales [...] Implicit feudalism has reigned over the dominant platforms for online communities so far, from the early BBSes to AI-enabled Facebook Groups. Peer-production practices surrounding free/open-source software and Wikipedia also exhibit it.

[...] The feudal pattern has by and large been written into the default behaviors of online-community platforms. Exceptions like Wikipedia and Debian have required considerable, intentional effort to counteract the implicit feudalism of their tools' defaults."

"Most scientific articles cited by Wikipedia articles are uncited or untested by subsequent studies"

From the abstract:^[3]

"Using a novel technique, a massive database of qualitatively described citations, and machine learning algorithms, we analyzed 1 923 575 Wikipedia articles which cited a total of 824 298 scientific articles in our database and found that most scientific articles cited by Wikipedia articles are uncited or untested by subsequent studies, and the remainder show a wide variability in contradicting or supporting evidence. Additionally, we analyzed 51 804 643 scientific articles from journals indexed in the Web of Science and found that similarly most were uncited or untested by subsequent studies, while the remainder show a wide variability in contradicting or supporting evidence."

"HopRetriever: Retrieve Hops over Wikipedia to Answer Complex Questions"

From the abstract:^[4]

"Collecting supporting evidence from large corpora of text (e.g., Wikipedia) is of great challenge for open-domain Question Answering (QA). Especially, for multi-hop open-domain QA, scattered evidence pieces are required to be gathered together to support the answer extraction. In this paper, we propose a new retrieval target, hop, to collect the hidden reasoning evidence from Wikipedia for complex question answering. Specifically, the hop in this paper is defined as the combination of a hyperlink and the corresponding outbound link document."

(See also the above review of the "WikiFlash" paper presented at the same conference)

"Structured Knowledge: Have we made progress? An extrinsic study of KB [knowledge base] coverage over 19 years"

From the abstract:^[5]

"... we employ question answering and entity summarization as extrinsic use cases for a longitudinal study of the progress of KB coverage. Our analysis shows a near-continuous improvement of two popular KBs, DBpedia and Wikidata, over the last 19 years, with little signs of flattening out or leveling off."

See also the [video recording](https://www.youtube.com/watch?v=TK6O9GP8Jk4) (https://www.youtube.com/watch?v=TK6O9GP8Jk4) of a talk by the authors at Wikidata Workshop 2020.

"A Review of Public Datasets in Question Answering Research"

Presented at the ACM Special Interest Group on Information Retrieval (SIGIR) forum last December, this paper^[6] found that the majority of Question Answering (QA) datasets are based on Wikipedia data.

Wikipedia has "become more popular in research on knowledge representation and natural language processing" in recent years

From the "Evaluation" section of an AAAI'21 paper titled "Identifying Used Methods and Datasets in Scientific Publications":^[7]

"Figure 4c shows the absolute amount of publications for the top four extracted datasets. [...] Another trend is visible for Wikipedia, which has become popular in research on knowledge representation and natural language processing."

"SF-QA: Simple and Fair Evaluation Library for Open-domain Question Answering"

The contributions of this paper^[8] include

"a hub of pre-indexed Wikipedia [dumps, of the English and Chinese language versions] at different years with different ranking algorithms as public APIs or cached results". The authors note that "Opendomain QA datasets are collected at different time, making [them depend] on different versions of Wikipedia as the correct knowledge source. [...] Our experiments found that a system's performance can vary greatly when using the wrong version of Wikipedia. Moreover, indexing the entire Wikipedia with neural methods is expensive, so it is hard for researchers to utilize others' new rankers in their future research."

"The Truth is Out There: Investigating Conspiracy Theories in Text Generation"

This preprint^[9] includes a dataset consisting of 17 conspiracy theory topics from Wikipedia (including e.g. the articles Death of Marilyn Monroe, Men in black, Sandy Hook school shooting) and comes with a content warning ("Note: This paper contains examples of potentially offensive conspiracy theory text").

"Spontaneous versus interaction-driven burstiness in human dynamics: The case of Wikipedia edit history"

From the abstract:^[10]

"[We analyze] the Wikipedia edit history to see how spontaneous individual editors are in initiating bursty periods of editing, i.e., spontaneous burstiness, and to what extent individual behaviors are driven by interaction with other editors in those periods, i.e., interaction-driven burstiness. We quantify the degree of initiative (DOI) of an editor of interest in each Wikipedia article by using the statistics of bursty periods containing the editor's edits. The integrated value of the DOI over all relevant timescales reveals which is dominant between spontaneous and interaction-driven burstiness. We empirically find that this value tends to be larger for weaker temporal correlations in the editor's editing behavior and/or stronger editorial correlations. These empirical findings are successfully confirmed by deriving an analytic form of the DOI from a model capturing the essential features of the edit sequence."

(See also our earlier coverage (<https://meta.wikimedia.org/wiki/Special:Search?search=burstiness&pre fix=Research%3ANewsletter%2F20&fulltext=Search+past+issues&fulltext=Search&ns0=1&ns12=1&ns200=1&ns202=1>) of research on editors' burstiness)

References

1. Yuang Cheng, Yue Ding, Damian Pascual, Oliver Richter, Martin Volk and Roger Wattenhofer: WikiFlash: Generating Flashcards from Wikipedia Articles (https://tik-db.ee.ethz.ch/file/130cdc2003615c74c07f9830158ee345/Flashcards_AAAI_21_AIEd_Camera_Ready.pdf). AAAI 2021 Workshop on AI Education (<https://sites.google.com/view/tipce-2021/home>), at the 35th AAAI Conference on Artificial Intelligence, February 9, 2021. Poster (<https://drive.google.com/file/d/1QktFja7xCRmZJOAueDfHNwIN450-Xg5R/view?usp=sharing>), presentation video (<https://drive.google.com/file/d/1n8UbdHVGZSp5Et-8VNA-1Zulne0Ub9f/view?usp=sharing>), online prototype (<https://flashcard.ethz.ch/>)
2. Schneider, Nathan (2021-01-07). "Admins, mods, and benevolent dictators for life: The implicit feudalism of online communities". *New Media & Society*. **24** (9): 1965–1985. doi:10.1177/1461444820986553 (<https://doi.org/10.1177%2F1461444820986553>). ISSN 1461-4448 (<https://www.worldcat.org/issn/1461-4448>). S2CID 234132111 (<https://api.semanticscholar.org/CorpusID:234132111>).  Preprint (<https://mediarxiv.org/sf432>)
3. Nicholson, Joshua M.; Uppala, Ashish; Sieber, Matthias; Grabitz, Peter; Mordaunt, Milo; Rife, Sean C. (2020-10-20). "Measuring the quality of scientific references in Wikipedia: an analysis of more than 115M citations to over 800 000 scientific articles" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8060352>). *The FEBS Journal*. **288** (14): 4242–4248. doi:10.1111/febs.15608 (<https://doi.org/10.1111%2Ffebs.15608>). ISSN 1742-4658 (<https://www.worldcat.org/issn/1742-4658>). PMC 8060352 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8060352>). PMID 33089957 (<https://pubmed.ncbi.nlm.nih.gov/33089957>). 
4. Li, Shaobo; Li, Xiaoguang; Shang, Lifeng; Jiang, Xin; Liu, Qun; Sun, Chengjie; Ji, Zhenzhou; Liu, Bingquan (2020-12-31). "HopRetriever: Retrieve Hops over Wikipedia to Answer Complex Questions". arXiv:2012.15534 (<https://arxiv.org/abs/2012.15534>) [cs.CL (<https://arxiv.org/archive/cs.CL>)]. (Accepted at AAAI 2021)
5. Razniewski, Simon; Das, Priyanka (2020-10-19). "Structured Knowledge: Have we made progress? An extrinsic study of KB coverage over 19 years" (<https://doi.org/10.1145/3340531.3417447>). *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. CIKM '20. New York, NY, USA: Association for Computing Machinery. pp. 3317–3320. doi:10.1145/3340531.3417447 (<https://doi.org/10.1145%2F3340531.3417447>). ISBN 9781450368599.  Author's copy (http://simonrazniewski.com/wp-content/uploads/2020_CIKM_KB_Coverage.pdf)
6. B. Barla Cambazoglu, Mark Sanderson, Falk Scholer, Bruce Croft: A Review of Public Datasets in Question Answering Research (<http://www.sigir.org/wp-content/uploads/2020/12/p07.pdf>). SIGIR Forum, December 2020, Volume 54 Number 2
7. Michael Färber, Alexander Albers, Felix Schüber: "Identifying Used Methods and Datasets in Scientific Publications" (https://www.aifb.kit.edu/images/2/2e/Identifying_Methods_Datasets_SDU2021.pdf). In Proceedings of the AAAI-21 Workshop on Scientific Document Understanding (SDU'21)@AAAI'21, Virtual Event, 2021
8. Lu, Xiaopeng; Lee, Kyusong; Zhao, Tiancheng (2021-01-06). "SF-QA: Simple and Fair Evaluation Library for Open-domain Question Answering". arXiv:2101.01910 (<https://arxiv.org/abs/2101.01910>) [cs.CL (<https://arxiv.org/archive/cs.CL>)]. Data and code (<https://github.com/soco-ai/SF-QA>)
9. Levy, Sharon; Saxon, Michael; Wang, William Yang (2021-01-02). "The Truth is Out There: Investigating Conspiracy Theories in Text Generation". arXiv:2101.00379 (<https://arxiv.org/abs/2101.00379>).
10. Choi, Jeehye; Hiraoka, Takayuki; Jo, Hang-Hyun (2020-11-03). "Spontaneous versus interaction-driven burstiness in human dynamics: The case of Wikipedia edit history". arXiv:2011.01562 (<https://arxiv.org/abs/2011.01562>).

+ *Add a comment* (https://en.wikipedia.org/w/index.php?editintro=Wikipedia:Wikipedia_Signpost/Templates/Comment-editnotice&title=Wikipedia_talk:Wikipedia_Signpost/2021-02-28/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/2021-02-28/RECENT_RESEARCH&ACTION=WATCH](https://en.wikipedia.org/w/index.php?title=Wikipedia_talk:Wikipedia_Signpost/2021-02-28/Recent_research&action=watch)), IF YOUR COMMENT HAS NOT APPEARED HERE, YOU CAN TRY .

- @HaeB: if I get 8.5/9 do I get a barnstar? I understand that you can't give everybody a barnstar - but I'm the first to claim it! [Smallbones](#)^(smalltalk) 22:02, 28 February 2021 (UTC)
 - Is that without looking at the article? I was very proud of getting 5.5/9 without looking (gave myself half a point for guessing 6.2 million) given that two are very specific statistics and at least three are not really unambiguous clearly-expressed questions. — [Bilorv](#) ([talk](#)) 23:30, 28 February 2021 (UTC)
 - I have to admit that I briefly copy edited the article, but that's not really reading for comprehension. I agree that some of the questions are ambiguous, so I answered to mysekf "If they mean W then my answer is X, if they mean Y the my answer is Z." [Smallbones](#)^(smalltalk) 23:46, 28 February 2021 (UTC)
- Re: The article "Most scientific articles cited by Wikipedia articles are uncited or untested by subsequent studies" is surprising given our favouring of secondary sources (which are more highly cited on average). Although it's higher than the literature as a whole ("28.5% of articles referenced in Wikipedia have a supporting citation vs. 11.7% of articles in Web of Science"), I wonder to what extent it is an artifact. [T.Shafee](#)([Evo&Evo](#))^{talk} 05:47, 1 March 2021 (UTC)
 - Ah, it may be something to do with the way that they define "untested by subsequent studies". In the Smart Cite system they use from scite.ai (<https://scite.ai/>), only 2.99% of citations are indicated as "Supporting citations" (i.e. "provide supporting evidence"). I suspect that most secondary sources don't get these sorts of citations as often as primary research. It'd be more interesting to separate out primary/secondary/tertiary sources cited by WP and specifically ask what percentage of those sources have Supporting citations. [T.Shafee](#)([Evo&Evo](#))^{talk} 06:02, 1 March 2021 (UTC)
 - This paper includes a "supporting evidence" section which appears to include an xls file containing a list of "retracted" sources cited on Wikipedia. Presumably we could use that list to remove sources that been retracted, but I have not opened the xls to verify. -- [GreenC](#) 16:16, 3 March 2021 (UTC)
 - @Evolution and evolvability and [GreenC](#): Ooh, that sounds like a good task for a bot, actually: Retracted citation patrolling. Perhaps with any articles found to be citing retracted papers added to a hidden tracking category, and/or templated with a cleanup notice to that effect? I wonder if the data set for that (the list of retracted papers to be flagged) could be maintained programmatically / updated periodically based on some machine-readable list of retractions, assuming there even is such a thing? -- [FeRDNYC](#) ([talk](#)) 01:09, 18 March 2021 (UTC)
 - [FeRDNYC](#), I agree retracted sources could be monitored programmatically and flagged with a trackable inline template. [WP:RSN](#) would be a good place to open

a discussion and if consensus open a bot request at [WP:BOTREQ](#). -- [GreenC](#) 01:17, 18 March 2021 (UTC)

- [@GreenC](#): Mmm, lest we imagine this is a bigger problem than it actually is, though, I *did* open that Excel file. It's a list of 50 citations (total!), divided into three categories:
 1. 15 are listed as "Acknowledges retraction", so IOW they're not the problem — there's nothing inherently wrong with referencing a retracted study, when it's done *in the context of* it being a retracted study.
 2. Another 10 are listed as "No longer referenced", which sort of undermines the title of the dataset, no?
 3. Of the remaining 25 listed as "Not acknowledged", there are actually only 13 retracted papers there. It's just that one of them happens to be cited in TWELVE different articles (and another one is cited in two). *Nearly all* (> 80%) of the articles in question are hyper-specific stubs on individual chemical compounds, like [OLIG1](#), [PTF1A](#), [MED24](#), [GCN5L2](#), etc. (Which IMHO is just further evidence that such articles have no business being part of Wikipedia in the first place, but that's just my bias talking.) -- [FeRDNYC](#) (talk) 01:33, 18 March 2021 (UTC)

[@FeRDNYC](#) and [GreenC](#): It's also worth noting the [meta:WikiCite/Shared_Citations](#) proposal as a relevant avenue for this sort of monitoring and notification. [T.Shafee\(Evo&Evo\)](#)^{talk} 02:20, 18 March 2021 (UTC)

- Based on the title of this piece, I was assuming I'd find an article about how we're biased against creating articles about nobility. signed, [Rosguill](#)^{talk} 16:16, 1 March 2021 (UTC)
 - I noticed that the entire corpus of scientific article did no better. So WP is doing in this respect about as well (or as poorly) as the world scientific community as a whole. analogous to the old finding that we were about the same as Britannica. [DGG](#) (talk) 07:41, 2 March 2021 (UTC)
 - Surely the opposite is true, if anything. I've always been surprised that [WP:GAN](#) has these two categories for history: "World history" and "Royalty, nobility and heraldry". But to each their own and there's plenty of interesting content in that category. — [Bilorv](#) (talk) 11:22, 2 March 2021 (UTC)
 - Oh I don't disagree, I just wasn't expecting an article about our "anti-feudalism" to be about our community governance. signed, [Rosguill](#)^{talk} 16:35, 3 March 2021 (UTC)
- **"Question: What does Wikipedia use to maintain its [sic] content?"** Sorry, that thud was the sound of my head hitting the keyboard. An *algorithm* came up with these? And even our **computers** aren't capable of differentiating between "its" and "it's"? Siiiiiiiiiiiiigh. Methinks they've learned to emulate humans a bit *too* well. -- [FeRDNYC](#) (talk) 01:13, 18 March 2021 (UTC)

IN THIS ISSUE



28 FEBRUARY 2021 ([ALL COMMENTS](#))

- [News and notes](#)
- [Disinformation report](#)

- [Opinion](#)
- [In the media](#)
- [News from the WMF](#)
- **Recent research**
- [Featured content](#)
- [Traffic report](#)
- [Gallery](#)

KEEP UP WITH *THE SIGNPOST* ON [TWITTER \(HTTPS://TWITTER.COM/WIKISIGNPOST\)](https://twitter.com/wikisignpost), [FACEBOOK \(HTTPS://WWW.FACEBOOK.COM/WIKISIGNPOST\)](https://www.facebook.com/wikisignpost) OR [MASTODON \(HTTPS://WIKIS.WORLD/@WIKISIGNPOST\)](https://wikis.world/@wikisignpost).

[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/2021-02-28/Recent_research&oldid=1193879895"

■