

Benjamin Mako Hill Tilman Bayer Reem Al-Kashif Aaron Shaw Wikimania 2017, Montréal August 11, 2017

State of Wikimedia Research Introduction 2

7-08 . R



I've been doing this for many years. I started in 2008 and have done this almost every single year since.

This began as an excuse for me to make sure I was up to date on Wikimedia Research.

"This talk will try to [provide] a quick tour – a literature review in the scholarly parlance – of the last year's academic landscape around Wikimedia and its projects geared at non-academic editors and readers. It will try to categorize, distill, and describe, from a birds eye view, the academic landscape as it is shaping up around our project."

– From my Wikimania 2008 Submission

State of Wikimedia Research Introduction ₽ 2017-08-

"This talk will try to (provide) a quick tour - a fittentive review in the scholarly particular, - of the last year's academic landscape around Willimedia and its projects general at non-academic editors and readers. I will try to categoria, distill, and describe, from a birds eye view, the academic landscape as it is shaping up around our project."

- From my Wikimania 2008 Submissio

Back in Wikimania 2008, I set out to run a session at Wikimania that would provide a comprehensive literature review of articles in Wikipedia published in the last year.

"This talk will try to [provide] a quick tour – a literature review in the scholarly parlance – of the last year's academic landscape around Wikimedia and its projects geared at non-academic editors and readers. It will try to categorize, distill, and describe, from a birds eye view, the academic landscape as it is shaping up around our project."

– From my Wikimania 2008 Submission

Then, about two weeks before Wikimania, I did the scholar search so I could build the literature.

"This talk will try to [provide] a quick tour - a literature review in the scholarly parlance - of the last year's academic landscape around Wikimedia and its projects geared at non-academic editors and readers. It will try to categorize, distill, and describe, from a birds eye view, the academic landscape as it is shaping up around our project."

- From my Wikimania 2008 Submission

Google	allintitle: wikipedia	•	
Scholar	About 800 results (0.03 sec)		
<mark>Articles</mark> Legal documents	peoon Blogs. Wikipedia, Second Life, and beyond: From production to produsage <u>A Bruns</u> - 2008 - books.google.com We-the users turned creators and distributors of content-are TIME's Person of the Year 2006, and AdAge's Advertising Agency of the Year 2007. We form a new Generation C. We have MMSnace YouTube and OurMadia: we run social software, and drive the		
Any time Since 2012 Since 2011 Since 2008 Custom range 2008 — 2009 Search	Cited by Oplace, Tod tobe, and Contentia, we full solvaria solvaria, and drive term Cited by Oplace, Tod tobe, and Contentia, we full solvarias - Library Search - All 11 versions Learning to link with wikipedia D Milne Proceedings of the 17th ACM conference on, 2008 - dLacm.org Abstract This paper describes how to automatically cross-reference documents with Wikipedia, the largest knowledge base ever known, it explains how machine learning can be used to identify significant terms within unstructured text, and enrich it with links to the Cited by 240 - Related articles - All 19 versions An effective low-cost measure of semantic relatedness obtained from Wikipedia links		

	State of Wikimedia Research		year's academic landscape around Wikimedia and its projects geared at non-academic editors an readers. It will try to categoriae, distil, and describe, from a birds eye view, the academic landsca as it is shaping up around our project."	
N		Google	- From my Wikimania 2008 Submise alirette wikipedia	
4		Scholar	About BDC results (BAB and	
2-08		African Legit documents	none (Eng., Millipedia, Sacota L.K., and Jenoral, Trum and action to a sub-stage damage 1-bits: Another projections We the uses the hold-bits and distribution docubed - as TMM is reactor/time two (2016), and diagnal Anothering Sparses for the Tara 2020" Units of two as ward-meaning to C. We have Mighters. You Date, and Collindia are not include them a ward-meaning to C. We have Mighters. The Date, and Collindia are not include them a two of two the C. B. (and Collindia advance). On the MI is marked. The Date Tara 2020" Million of the Colling Date of the Colling of the Colling of the Colling of the Colling State of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling of the Colling of the Colling of the Colling State of the Colling of the Colling State of the Colling of the Colli	
201		Bana 2010 Gence 2010 Bance 2000 Custom tange JOBN - JOBN Search	Learning to this with without a Diggs Present of the CP and the Additional the second second second second second second and Methods for second se	

I tried to import the whole list into Zotero and managed to get banned for abusing the Google Scholar because they thought that no human being could realistically consume the amount of material published on Wikipedia that year. So anyway, I had a 45 minute talk so it worked out to 3.45 seconds to per paper... And believe it or not, this year is even bigger.

And my talk is even shorter.

OL 1 () A / 1 1

"This talk will try to [provide] a quick tour – a literature review in the scholarly parlance – of the last year's academic landscape around Wikimedia and its projects geared at non-academic editors and readers. It will try to categorize, distill, and describe, from a birds eye view, the academic landscape as it is shaping up around our project."

- From my Wikimania 2008 Submission



	State of Wikimedia Research
N	Introduction
8-1	
2-0	
201	



I tried to import the whole list into Zotero and managed to get banned for abusing the Google Scholar because they thought that no human being could realistically consume the amount of material published on Wikipedia that year. So anyway, I had a 45 minute talk so it worked out to 3.45 seconds to per paper... And believe it or not, this year is even bigger.

And my talk is even shorter.



Number of citation, per year, with the term "wikipedia" in the title.

(Source: Google scholar results. Accessed: 2016-06-24)

State of Wikimedia Research

2017-08-



Academics have written a lot of papers about Wikipedia. There are more than 500 papers published about Wikipedia each year and although we've reached and moved past a peak it seems, it's not slowing by much.

State of Wikimedia Research

 6,967 Wikipedia-related publications in the Scopus database as of this morning (August 11, 2017)
 154 recent publications covered in the 12 issues of the Wikimedia Research Newsletter from June 2016 to May 2017 (and hundreds more on our list!)

- 6,967 Wikipedia-related publications in the Scopus database as of this morning (August 11, 2017)
- 154 recent publications covered in the 12 issues of the Wikimedia Research Newsletter from June 2016 to May 2017 (and hundreds more on our list!)

The newsletter aims to be comprehensive, but mostly ignores papers that use Wikipedia as a corpus only (which is popular e.g. in NLP research).



In selecting papers for this session, the goal is always to choose examples of work that:

- Represent important themes from Wikipedia in the last year.
- ► Research that is likely to be of interest to Wikimedians.
- ► Research by people who are not at Wikimania.
- ► ... with a bias towards peer-reviewed publications

State of Wikimedia Research	Programmer and the state is the state of the state o
3	 with a bias towards peer-reviewed publications

This is my disclaimer slide...

Within these goals, the selections are incomplete, and wrong.

Gender Gap in Participation

State of Wikimedia Research

Gender Gap in Participation

Reem

₽

2017-08-

There has been a lot of work on the different manifestations of gender bias on Wikipedia. In some of this work, the trend to suggest including more women editors would help increasing women biographies on Wikipedia. However, this is not always the case.

Gender Gap in Participation

Nicolaes, Feli. 2016. "Gender Bias on Wikipedia: An Analysis of the Affiliation Network." Bachelors Thesis, Amsterdam, The Netherlands: University of Amsterdam. https://esc.fnwi.uva.nl/thesis/ centraal/files/f1270649307.pdf.

```
State of Wikimedia Research
```

Data:

- English Wikipedia editors and the pages they edit.
- Tracking of editing behavior of both self-identified male and female editors on Wikipedia.

Nicolaes, Fell. 2016. "Gender Blas on Wikipedia: An Analysis of the Affiliation Network: Bachelors Thesis, Amsterdam, The Netherlands University of Amsterdam. https://erc.fnei.ova.nl/thesis/ contrass/fises/fis70649307.pdf.



Figure 10: The percentage of pages from people born in a certain birthyear



(a) Percentages when looking at all arrows in the (b) Percentages when only looking at arrows with a network weight of more than 200 [Figure 11 in Nicolae 2017]

State of Wikimedia Research

.-80-/

. R



- · Women editors are not focused on female biography pages.
- Opposed to men editors, women editors who edit a single biography article more than 200 times are more likely to put this effort in biography articles about men.

Gender Gap in Content

State of Wikimedia Research

Gender Gap in Content

Reem

There has been a lot of research on how the content of Wikipedia favors men over women. This research presents another way this bias can be visible to readers.

Gender Gap in Content

State of Wikimedia Research

<u>∩</u>

17-08-

. R Gender Gap in Content

Gender Gap in Content

Zagovora, Otga, Fabian Flock, and Claudia Wagner. 2017. "(Weltergeletet von Journalistin): The Endered Presentation of Professions on Wikipedia: In Proceedings of the 2017 ACM on Web Science Conference (WebSci '17), 83–82. New York, New York: ACM. http://dx.doi.org/10.1016/3031478.2031488.

Zagovora, Olga, Fabian Flöck, and Claudia Wagner. 2017. "'(Weitergeleitet von Journalistin)': The Endered Presentation of Professions on Wikipedia." In *Proceedings of the 2017 ACM on Web Science Conference (WebSci '17)*, 83–92. New York, New York: ACM. http:dx.doi.org/10.1145/3091478.3091488.

Journalist

(Weitergeleitet von Journalistin)



Journalisten interviewen sich



Journalisten bei der Fußball-Europameisterschaft 2008



Fotografen beim Fußbal



Pressetribüne im niedersächsischen Landtag

deutscher Korresponder (1929-2010)



Heinrich Heine, Dichter

und Journalist

5

sus den USA



Fotojournalisten bei der Fußball-Europameisterschaft 2008





Islerischen Journaldimen

State of Wikimedia Research Paper Summaries

4

2017-08-



- List of professions names.
- All articles from the category "professions" (DE:"Beruf") in German
- Images of people in these articles
- Numbers of mentions of males and females in the articles.
- · Numbers of men and women in professions.
- · Google search results for male and female professions.

Gender Gap in Content: Zagovora et al. Results

- Most of pages about professions have male titles even when the profession is dominated by females.
- Disproportionate distribution of male images even in female dominated professions.
- Articles mention men more than women (4k men and 800 women). Same is true even in female dominated professions.

Why is this important? Because it affects the readers' perception of these professions by either perpetuating existing biases & stereotype or establishing new ones.

Gender Gap in Content: Zagovora et al. Results

State of Wikimedia Research

Paper Summaries

2

2017-08-

professions

 Most of pages about professions have male titles ever when the profession is dominated by females

 Disproportionate distribution of male images even in female dominated professions.
 Articles mention men more than women (4k men and 800

women). Same is true even in female dominated

State of Wikimedia Research

4

2017-08-

Aaron

Fake News!

Fake News!

13/33

Kumar, Srijan, Robert West, and Jure Leskovec. 2016. "Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes." In Proceedings of the 25th International Conference on World Wide Web (WWW '16), 591–602. Geneva, Switzerland: International World Wide Web Conferences Steering Committee. http://dx.doi.org/10.1145/2872427.2883085.

State of Wikimedia Research Paper Summaries 7-08 -Fake News!

2

2

Kumar, Srijan, Robert West, and Jure Leskovec. 2016 Disinformation on the Web: Impact. Characteristics. and Detection of Wikipedia Hoaxes." In Proceedings of the 25th International Conference on World Wide Web (WWW '16), 591-602. Geneva, Switzerland: International World Wide Web Conferences Steering Committee, http://dx.doi.org/10.1145/2872427.2883085

Fake News!

Historically, many papers study whether and how WP produces accurate content. This paper looks at hoaxes more closely in a way that provides some really great insights.

Authors collect the set of all en:WP articles ever flagged as hoaxes. Of these, 21,218 removed.

- How do hoaxes perform? How efficiently are they flagged, removed, viewed etc.?
- How do hoaxes that are removed compare against non-hoaxes of various kinds?
- Machine classification performance (against human classification)?

Disinformation on the Web

State of Wikimedia Research

The paper is filled with interesting points about hoax articles! Here are some that I enjoyed learning:

 Most hoaxes are removed within a few hours. 1,175 survive more than a day. 1% survive over a year (!)

Disinformation on the Web

- Articles flagged as hoaxes lack features associated w good content (infoboxes, links, templates).
- Articles flagged falsely lack these features at a higher rate than articles flagged correctly!
- Hoaxes a re about topics that have been mentioned before, but often by fewer people and less frequently than non-hoaxes.
- Machine classifier performs really well (91% accuracy overall). Beats Mturk raters when shown hoax/non-hoax pairs (86% vs. 66% accuracy).

explain figure

[Kumar, West, & Leskovec, 2016]

Disinformation on the Web





(a) Prior mentions (b) 1st prior mention (c) 1st-men. creator

State of Wikimedia Research Paper Summaries ₽ 17-08-Disinformation on the Web



The paper is filled with interesting points about hoax articles! Here are some that I enjoyed learning:

- Most hoaxes are removed within a few hours. 1,175 survive more than a day. 1% survive over a year (!)
- Articles flagged as hoaxes lack features associated w good content (infoboxes, links, templates).
- Articles flagged falsely lack these features at a higher rate than articles flagged correctly!
- Hoaxes a re about topics that have been mentioned before, but often by fewer people and less frequently than non-hoaxes.
- Machine classifier performs really well (91% accuracy overall). Beats Mturk raters when shown hoax/non-hoax pairs (86% vs. 66% accuracy).

explain figure

[Kumar, West, & Leskovec, 2016]

. R

Using Wikipedia for Prediction

State of Wikimedia Research

Using

Wikipedia for

Prediction

Aaron

Using Wikipedia for Prediction

Smith, Benjamin K., and Abel Gustafson. 2017. "Using Wikipedia to Predict Election Outcomes: Online Behavior as a Predictor of Voting." *Public Opinion Quarterly*, nfx007. http://dx.doi.org/10.1093/poq/nfx007. State of Wikimedia Research

Using Wikipedia for Prediction

Smith, Benjamin K., and Abel Gustatson. 2017. "Using Wikipedia to Predict Election Outcomes: Online Behavior as a Predictor of Voting." Public Opinion Quarterly, http://dx.doi.org/10.1093/pog/nfx007.

Continued growth of using Wikipedia for prediction. This selection is about election forecasting.

Authors test two claims:

- 1. Are WP pageviews associated with electoral outcomes?
- 2. Do WP pageviews improve the performance of standard forecasting models?

predict vote share in 104 senatorial elections in the US in 2008, 2010, 2012 and en:WP pageviews 200 days prior.

Using Wikipedia to Predict Election Outcomes



Figure 2. Absolute Errors for Each Projection Type.

State of Wikimedia Research

Using Wikipedia to Predict Election Outcomes



Explain the figure.

2017-08-12

Opinion polls and fundamentals still both better than pageviews alone. However, pageviews helps improve the aggregate model.

Great example of how interactions w Wikipedia can help shed light on different kinds of behavior in the world in ways that complement existing data sources.

[Smith & Gustafson, 2017 18 / 33

Syndication

State of Wikimedia Research Paper Summaries 4 2017-08-

Mako:

There was an increase in studies that look at how Wikimedia content - including WikiData - is being reused in different places.

Syndication

McMahon, Connor, Isaac L. Johnson, and Brent J. Hecht. 2017. "The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship between Peer Production Communities and Information Technologies." In *International AAAI Conference on Web and Social Media (ICWSM 2017)*, 142–151. Palo Alto, California: AAAI. http://brenthecht.com/publications/ icwsm17_googlewikipedia.pdf. State of Wikimedia Research

McMahon, Connor, Inaac L. Johnson, and Brenil. Hecht. 2017. "The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship between Peer Production Communities and Internation Tachnologies." In International AAAI Contenence on Web and Social Media (ICVSM 2017, 142–151. Paol. Abo. California: AAAI. http://beanthe.ht.eom/publicstion/ International-AAAI.

Syndication

A very cool study was an experimental study that attempted to look at the interdependence between Google and Wikipedia. You can tell from the title of the paper that the authors believe that the interdependence is "substantial." This is a guestion that Dario in particular, and Wikimedians in general have been

asking for years.



State of Wikimedia Research

Experiment:

20

- Several dozen people who all used Chrome and Google installed a browser extension. They were *not* told what the extension would do!
- Secretly and quietly, the extension modified their search engine results to remove content from Wikipedia from search results. This happened in three ways:
 - It removed links/hits to WP from search results
 - it removed things from the knowledge graph if they had come from Wikipedia. This include thing in the two areas shown above. We know that much of this data is taken from WikiData and other places.
- See what happened in terms of how much people click through (an answer of search engine effectiveness)
- See what happen to Wikipedia viewership.



Syndication: Results from McMahon et al.

- Google depends on Wikipedia: Removing

Wikipedia links decreases click-through rate by $\tilde{8}0\%$ (26.1% \rightarrow 14.0%)

State of Wikimedia Research

Syndication: Results from McMahon et al.

Results...

₽

2017-08-

Obvious implications for Wikipedia:

- Research has suggested that WP relies on an influx of traffic for production.
- Obviously also important for donations.

 Google depends on Wikipedia: Removing Wikipedia links decreases click-through rate by 80% (26.1% → 14.0%)

Syndication: Results from McMahon et al.

Google depends on Wikipedia: Removing

Wikipedia links decreases click-through rate by $80\% (26.1\% \rightarrow 14.0\%)$

Wikipedia depends on Google: 84.5% visit to Wikipedia were attributable to Google

```
State of Wikimedia Research
       Paper Summaries
2017-08-
              -Syndication: Results from McMahon et al.
```

Syndication: Results from McMahon et al

Google depends on Wikipedia: Removing Wikipedia links decreases click-through rate by $80\% (26.1\% \rightarrow 14.0\%)$ Wikipedia depends on Google: 84.5% visit to Wikipedia were attributable to Google

Results...

2

Obvious implications for Wikipedia:

- Research has suggested that WP relies on an influx of traffic for production.
- Obviously also important for donations. ٠

Syndication: Results from McMahon et al.

• Google depends on Wikipedia: Removing

Wikipedia links decreases click-through rate by $\tilde{8}0\%$ (26.1% \rightarrow 14.0%)

- Wikipedia depends on Google: 84.5% visit to
 Wikipedia were attributable to Google
- ► Knowledge graph reduces Wikipedia traffic:

Removing knowledge graph elements increased Wikipedia visits rate (11.1% \rightarrow 20.5%)

```
State of Wikimedia Research
```

Syndication: Results from McMahon et al

 Google depends on Wilkpedia: Removing Wikipedia links decreases click-through rate by 80% (26.1% -> 14.0%)
 Wikipedia depends on Google: 84.5% visit to Wikipedia were attributable to Google
 Knowledge graph reduces Wikipedia traffic: Removing knowledge graph elements increased

Wikipedia visits rate (11.1% → 20.5%)

Results...

Obvious implications for Wikipedia:

- Research has suggested that WP relies on an influx of traffic for production.
- Obviously also important for donations.

Wikipedia and the World

State of Wikimedia Research

Wikipedia and the World

Tilman

₽

2017-08-

Several interesting studies this year looked at lasting effects of real-world changes on Wikipedia or vice versa. E.g. rising unemployment in European countries during the Great Recession from 2008 on caused increased reading and editing activity. And a new pre-print (not yet peer-reviewed) found that adding content to articles about Spanish towns increased local tourism by 9%.

Wikipedia and the World

State of Wikimedia Research

4

2017-08-

Wikipedia and the World

Wikipedia and the World

Penney, Jonathon. 2016. "Chilling Effects: Online Surveillance and Wikipedia Use." Berkeley Technology Law Journal 31 (1): 117. http://dx.doi.org/10.15779/Z385813

Penney, Jonathon. 2016. "Chilling Effects: Online Surveillance and Wikipedia Use." *Berkeley Technology Law Journal* 31 (1): 117. http://dx.doi.org/10.15779/Z38SS13

24/33

Does the awareness of potential surveillance deter Internet users from accessing sensitive content?

The "external shock" from the June 2013 Snowden revelations increased worldwide awareness that Internet communications are being monitored by the US government.

Paper examines its impact on the pageview numbers of a set of 48 terrorism-related articles on English Wikipedia.

State of Wikimedia Research Paper Summaries <u>∩</u> 7-08

-Wikipedia and the World

Does the awareness of potential surveillance deter Internet users from access samilia content? The "external shock" from the June 2013 Showden revelations increased terrorism-related articles on English Wikipedia

Wikipedia and the World

Assumptions were verified very diligently - e.g. a survey among 415 Mechanical Turk users confirmed that the article topics (derived from a list of the US Department of Homeland Security) were indeed considered sensitive.

. R



State of Wikimedia Research



Evaluated with a statistical method called "interrupted time series": 25% immediate drop-off around June 2013

Education

State of Wikimedia Research

Tilman:

Minitalk: 2 minutes

The use of Wikipedia in education, in particular for college writing assignments, continues to be the focus of many research publications. Often these are simple case studies focusing on the authors' own teaching project. Others examine the changing attitudes of faculty to Wikipedia. Some good overview articles came out this year - see the January 2017 special issue of the Research Newsletter: https:

//meta.wikimedia.org/wiki/Research:Newsletter/2017/January

Education

Shane-Simpson, Christina, Elizabeth Che, and Patricia J. Brooks. 2016. "Giving Psychology Away: Implementation of Wikipedia Editing in an Introductory Human Development Course." *Psychology Learning & Teaching* 15 (3): 268–93.

http://dx.doi.org/10.1177/1475725716653081

State of Wikimedia Research Paper Summaries ₽ 17-08--Education . R

Shane-Simpson, Christina, Eirabeth Che, and Parkia J. Brooks. 2016. "Giving Psychology Away: Implementation of Wikipedia Estiting in an Introductory Human Development Course." *Psychology Laarning & Taaching* 15 (3): 258–93. http://dx.doi.org/10.1171/4107128716663081

We picked one case study that came with some interesting results from the class survey.

Class survey (N=93) regarding interactions with regular editors.

Education: Shane-Simpson et al. Results

State of Wikimedia Research

₽

2017-08-

Education: Shane-Simpson et al. Results

- 95% of students recalled beneficial interactions
- 15% recalled negative ones
- 73% of respondents were reverted
- 56% had grammar or punctuation corrected

Being reverted or corrected was often seen as beneficial.

- ▶ 95% of students recalled **beneficial** interactions
- ► 15% recalled **negative** ones
- ► 73% of respondents were **reverted**
- ► 56% had grammar or punctuation corrected

Being reverted or corrected was often seen as beneficial.

Datasets: Research that enables other research

	State of Wikimedia Research
2017-08-12	Paper Summaries

Datasets: Research that enables other research

Tilman Minitalk: 2 minutes!

Flöck, Fabian, Kenan Erdogan, and Maribel Acosta. 2017. "TokTrack: A Complete Token Provenance and Change Tracking Dataset for the English Wikipedia." In *Proceedings of the International Conference on Web and Social Media (ICWSM 2017)*. Palo Alto, California: AAAI. https://aaai.org/ocs/index.php/ ICWSM/ICWSM17/paper/view/15689/14890. State of Wikimedia Research

Filds, Fabian, Kenan Erdogan, and Maribel Acosta. 2017. "Bichtack: A Complete Tokian Provenance and Change Tracking Dataset for the English Wilkpedia: In Proceedings of the International Conference on Web and Social Media (ICWSM 2017). Pailo Alto, Galitomia: AAAJ. http://aaai.org/ocs/index.php/ 1089/1089/170801/689/14990.

Much of the existing Wikipedia research is based on the freely licensed datasets published by the Wikimedia Foundation: Content dumps, pageview numbers, Clickstream datasets, etc. See https://meta.wikimedia.org/wiki/Research:Data Some individual researchers are giving back too...

Datasets: Research that enables other research



"a dataset that contains every instance of all tokens (\approx words) ever written in undeleted, non-redirect English Wikipedia articles until October 2016, in total 13,545,349,787 instances. [...] This data would be exceedingly hard to create by an average potential user ..."

Can track each token across deletions and re-additions through the entire history. Much higher accuracy than e.g. Wikitrust.

Datasets: Research that enables other research

atasets: Research that enables other research

Firm						
Revision 10	Revision R2	Boriston \$3	Restation Bri			
		Room shad to a monthly a Mangalay are 2001 a				

undeleted, non-redirect English Wikipedia articles until October 2016, in total 13.545.349.787 instances. [...] This data would be exceedingly hard to create b an average potential user ... Can track each token across deletions and re-additions through the entire history Much higher accuracy than e.g. Wikitrust.

See also

<u>∩</u>

2017-08·

[Flöck et al. 2017

State of Wikimedia Research

Paper Summaries

https://meta.wikimedia.org/wiki/Research:Content_persistence http://f-squared.org/whovisual/#color (related tool by some of the same researchers)

More Resources

► Wikimedia Research Newsletter

[[:meta:Research:Newsletter]] / @WikiResearch

- WikiSym/OpenSym (This month in Ireland!)
- https://meta.wikimedia.org/wiki/Research:Events
- Much More



```
State of Wikimedia Research
Paper Summaries
```



Wikimedia Research Newslette

More Resources

Those are our eight exemplary studies from the past year.

There has been just tons and tons of work in this area. Trying to talk about this in 40 minutes strikes me as increasingly crazy every year we try to do it. The most important source is the Wikimedia Research Newsletter which has since 2011 been published monthly in the (English) Signpost and syndicated on the Wikimedia Research space on Meta-Wiki. (Special thanks to Dario Taraborelli and User:Masssly for finding and cataloguing new publications throughout the year!) But there are other resources as well. And I encourage you to get involved.