

Community sustainability in Wikipedia: a review of research and initiatives

...

Stuart Geiger, ethnographer/postdoc
UC-Berkeley Institute for Data Science (BIDS)

@staeiou | stuart@stuartgeiger.com
github.com/staeiou/wiki-stat-notebooks

Some resources & references

Wiki Research Portal:

meta.wikimedia.org/wiki/Research

research newsletter, mailing list, data and code
resources, ongoing research projects, and more!

Evaluation Portal:

meta.wikimedia.org/wiki/Grants:Evaluation

WikiLit

wikilit.referata.com/

Itinerary

The “oh shit!” graph

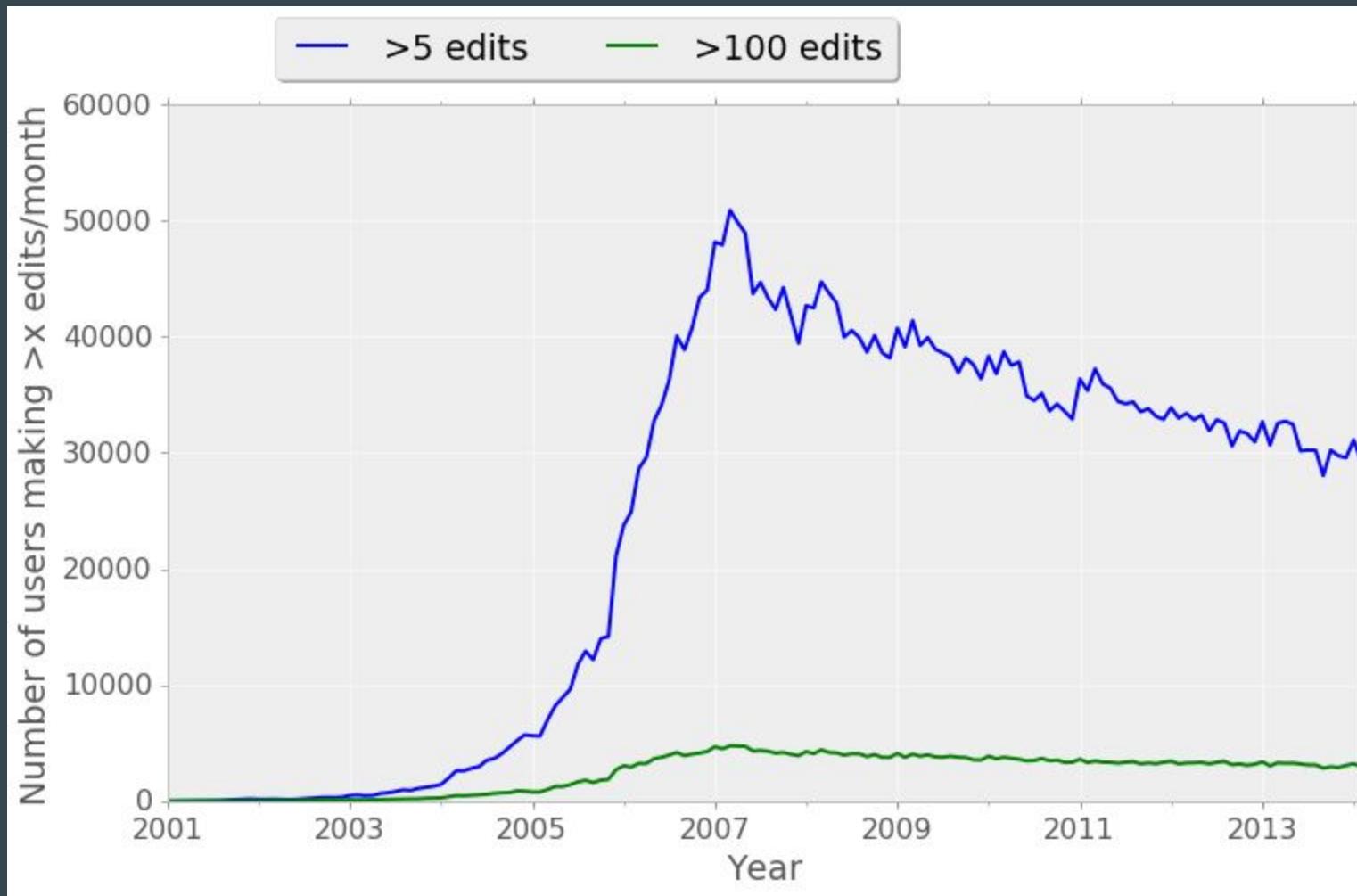
The history of the Wikipedia community

Newcomer socialization and governance

Diversity, inclusion, and representation

Initiatives and future steps

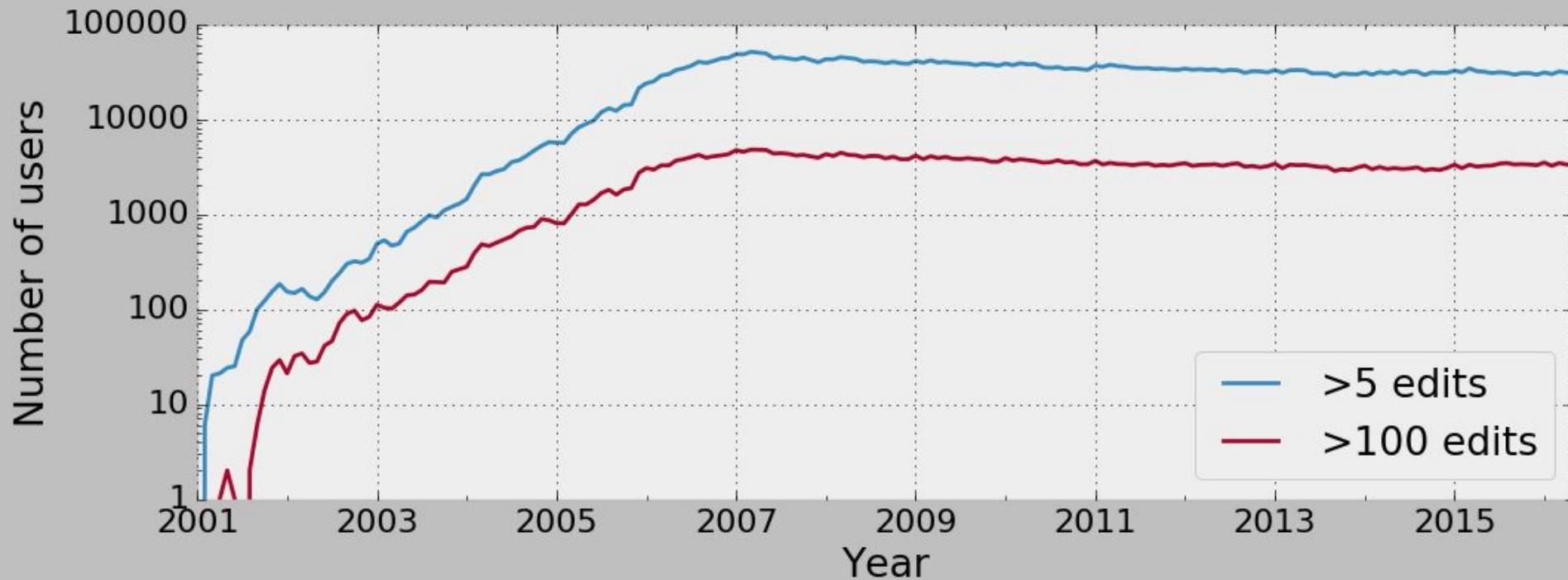
Number of active Wikipedia editors



Data from Erik Zachte, <https://stats.wikimedia.org>

Jupyter notebook at <https://github.com/staeiou/wiki-stat-notebooks>

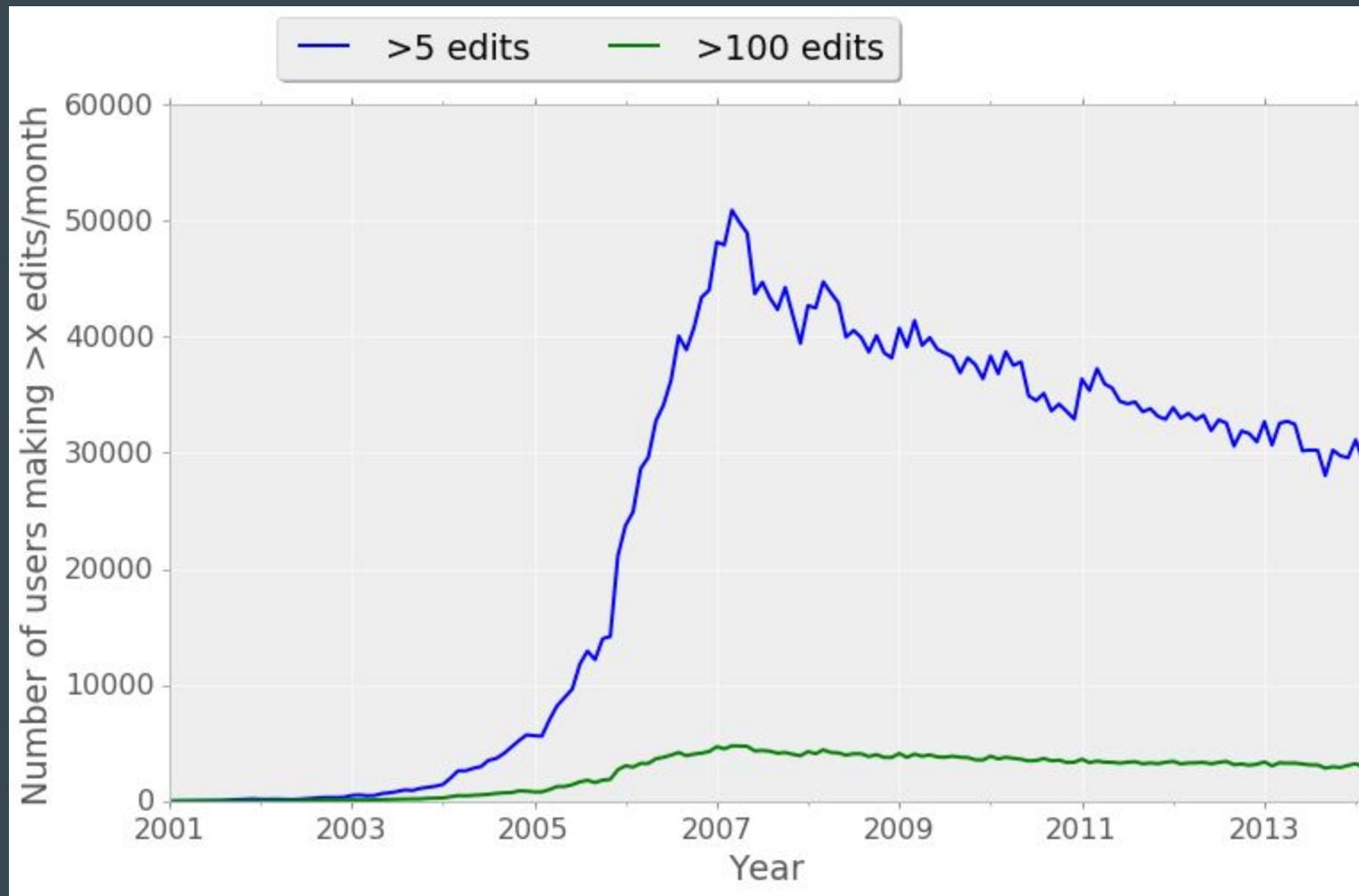
Number of active Wikipedia editors



Data from Erik Zachte, <https://stats.wikimedia.org>

Jupyter notebook at <https://github.com/staeiou/wiki-stat-notebooks>

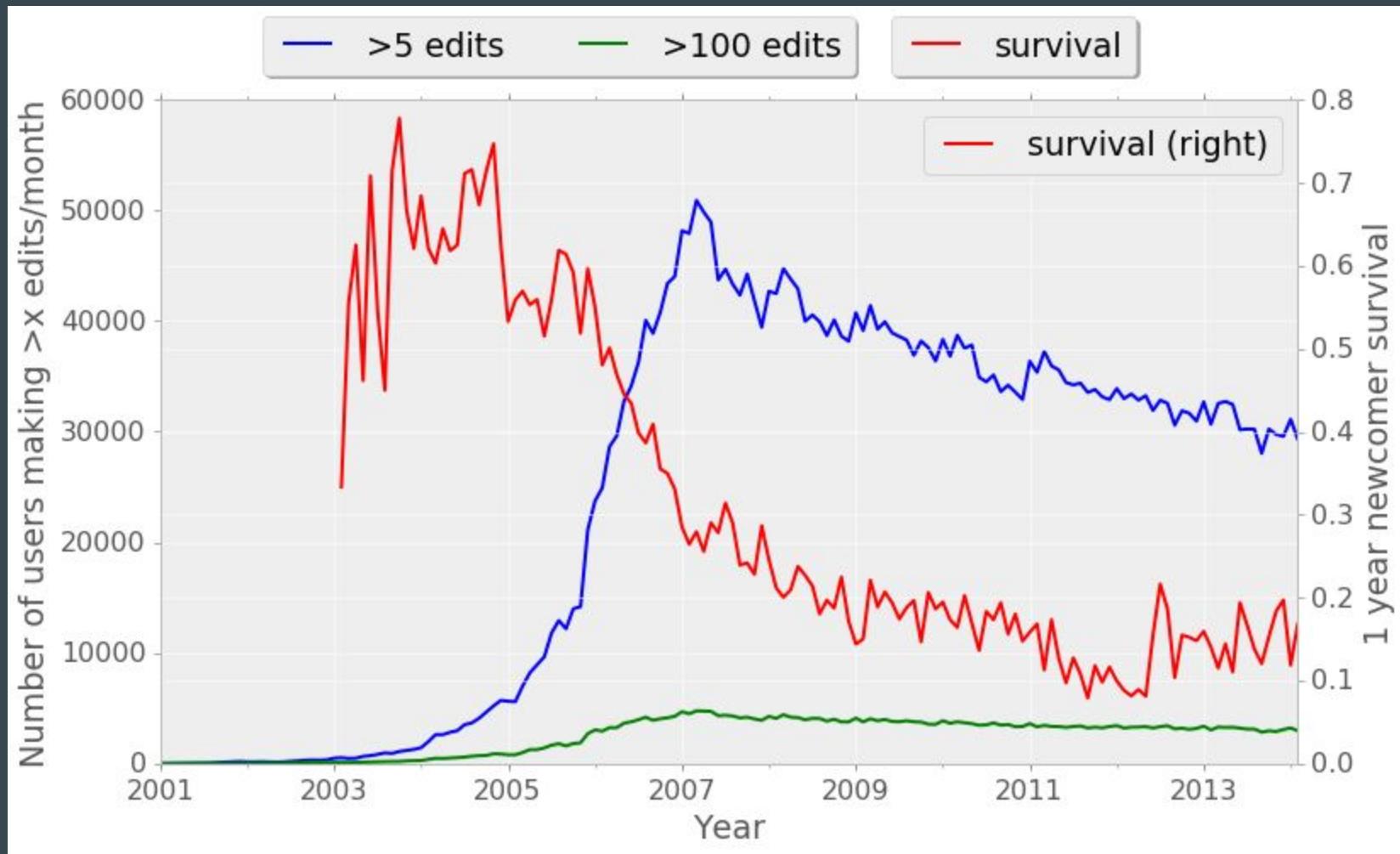
Number of active Wikipedia editors



Data from Erik Zachte, <https://stats.wikimedia.org>

Jupyter notebook at <https://github.com/staeiou/wiki-stat-notebooks>

The “oh shit!” graph



Data from Erik Zachte, <https://stats.wikimedia.org> and me
Jupyter notebook at <https://github.com/staeiou/wiki-stat-notebooks>

Wikipedia's community history

Some great books on this:

The Wikipedia Revolution

by Andrew Lih

Good Faith Collaboration

by Joseph Reagle

Wikipedia and the Politics of Openness

by Nate Tkacz

Common Knowledge

by Dariusz Jemielniak

2001-2004: The early days

A screenshot of a web browser window displaying the Wikipedia homepage. The URL in the address bar is <https://nostalgia.wikipedia.org/wiki/HomePage>. The page title is "HomePage". Below the title, it says "From Wikipedia". A navigation bar includes links for "HomePage", "Recent changes", "View source", "Discuss this page", "Page history", and "Log in". There is also a "Special pages" dropdown menu and a "Go" button. A message at the top right encourages users to copy and paste text from Wikipedia into their own documents. The Wikipedia logo and tagline "The Free Encyclopedia" are visible on the right side. Below the main content area, there are links for "Printable version", "Disclaimers", and "Privacy policy".

Welcome to [Wikipedia](#), a collaborative project to produce a complete encyclopedia from scratch. We started in January 2001 and already have **over 19,000 articles**. We want to make over 100,000, so let's get to work--*anyone* can edit any article--copyedit, expand an article, write a little, write a lot. See the [Wikipedia FAQ](#) for information on how to edit pages and other questions.

The content of Wikipedia is covered by the [GNU Free Documentation License](#), which means that it is free and will remain so forever. See [open content](#) and [free content](#) for background.

Current Events and Breaking News

Encyclopedia articles about topics behind the news.

[2001 U.S. Attack on Afghanistan](#) - [More current events](#)

2001-2004: The early days

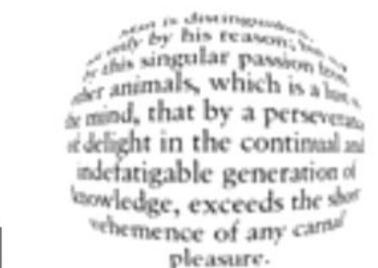
Rules to consider

(Redirected from [RulesToConsider](#))

[HomePage](#) | [Recent changes](#) | [View source](#) | [Discuss this page](#) | [Page history](#) |

[Log in](#) | [Special pages](#) ▾ Go

[Printable version](#) | [Disclaimers](#) | [Privacy policy](#)



WIKIPEDIA
The Free Encyclopedia

Ignore all rules. If rules make you nervous and depressed, and not desirous of participating in the wiki, then ignore them entirely and go about your business.

Supporters include: [Larry Sanger](#), [WojPob](#), [Jimbo Wales](#), [AyeSpy](#), [OprgaG](#), [Invictus](#), and [Koyaanis Qatsi](#), [Pinkunicorn](#), [sjc](#), [mike dill](#), [Taw](#), [GWO](#)

Opponents include: [tbc](#), [AxelBoldt](#) (deliberately breaking them is fine; ignoring them is not -- ignorance is bad.)

2004-2006: Wikipedia grows up

External factors:

- Wikipedia higher in Google results
- Major mass media coverage of Wikipedia
- More concern about Wikipedia's accuracy
 - *Nature* study
 - Seigenthaler incident
 - Stephen Colbert's Wikiality segment

2004-2006: Wikipedia grows up

Internal milestones:

- 500k articles by 2005, 1m by 2006
- First global Wikimania conference
- Wikimedia Foundation officially incorporates as non-profit, hires staff
- First major fundraiser for \$100,000
- More and more intense controversies

WIKIMANIA '06

IDENTIFYING AND ELIMINATING TROLLS



Photo by Gmaxwell, © GFDL 1.2
https://enwp.org/File:Troll_school.jpg

2006-2008: Wikipedia formalizes

Increased rules, policies, and
bureaucracy (esp. for biographies)

Automated tools for reviewing and
reverting low-quality contributions

New attitudes toward newcomers


WIKIPEDIA
 The Free Encyclopedia

[Project page](#) [Talk](#)
[Read](#) [Edit source](#) [View history](#) [More](#) [Page](#) [TW](#) [Search](#)


Wikipedia:List of policies [\[edit source\]](#)

From Wikipedia, the free encyclopedia

"WP:NUTSHELL" redirects here. For information about what a Wikipedia nutshell is, see [Template:Nutshell](#).

This page includes a summary of [official policies](#) on the English Wikipedia which are set out in detail elsewhere.

Policies have wide acceptance among editors and are considered standards that all editors should follow. When editing this page, please ensure that your revision is consistent with the underlying policies. When in doubt, discuss it on the [talk](#)



page. Where [a discrepancy exists](#), the policy page itself overrides. Changing this page does not change policy. Likewise, adding a page to this summary does not elevate it to policy status. Policies are promoted through consensus. See [Wikipedia:Polices and guidelines](#) for more.

Shortcuts:
[WP:LOP](#)
[WP:NUTSHELL](#)

All policies are listed at [Category:Wikipedia policies](#). A list that also includes guidelines is at [List of policies and guidelines](#). This page is broken into the following categories:

- [Content](#), which defines the scope of the encyclopedia and the material that is suitable for it
- [Conduct](#), which describes how editors can successfully collaborate and what behavior is acceptable
- [Deletion](#), which explains the processes by which pages, revisions, and logs may be deleted
- [Enforcement](#), which accounts for various means by which standards may be enforced
- [Legal](#), which includes rules influenced by legal considerations, and remedies for their misuse
- [Procedural](#), which documents various processes by which the English Wikipedia operates

[Contents](#) [\[show\]](#)

Content [\[edit source\]](#)

Pages currently in [Category:Wikipedia content policies](#):

[Policy shortcut:](#)
[WP:CONPOL](#)

Article titles

The ideal title for a Wikipedia article is recognizable to English speakers, easy to find, precise, concise, and consistent with other titles.

Biographies of living persons

Articles about living persons, which require a degree of sensitivity, must adhere strictly to Wikipedia's content policies. Be very firm about high-quality references, particularly about details of personal lives. "Unsourced or poorly sourced contentious material—whether negative, positive, or just questionable—about living (or sometimes recently deceased) persons should be removed immediately and without discussion from Wikipedia articles, talk pages, user pages, and project space."

Image use policy

Generally avoid uploading non-free images; fully describe images' sources and copyright details on their description pages, and try to make images as useful and reusable as possible.

Neutral point of view

Everything that our readers can see, including articles, templates, categories and portals, must be written neutrally and without [bias](#).

No original research

Policies and guidelines

Principles

[Five pillars](#) · [Ignore all rules](#) ·

[Core content policies](#)

Content policies

[Neutral point of view](#) · [Verifiability](#) ·

[No original research](#) ·

[Biographies of living persons](#) · [Article titles](#) ·

[Image use](#) ·

[What Wikipedia is not \(Not a dictionary\)](#)

Conduct policies

[Civility](#) · [No personal attacks](#) · [Harassment](#) ·

[No legal threats](#) · [Consensus](#) ·

[Dispute resolution](#) · [Username policy](#) ·

[Clean start](#) · [Vandalism](#) · [Editing policy](#) ·

[Edit warring](#) · [Ownership of content](#) ·

[Non-discrimination policy](#)

Other policy categories

[Deletion](#) · [Enforcement](#) · [Legal](#) · [Procedural](#)

Directories

[List of policies](#) · [List of guidelines](#) ·

[Manual of Style contents](#)

V·I·E

Core content policies

[Neutral point of view](#)

[No original research](#)

[Verifiability](#)

[Other](#)

Wikipedia's policy on speedy deletion, one of two process created to streamline the existing deletion process.

<https://enwp.org/WP:CSD>

Bots and “cyborgs”

Applications Places System Firefox Hoggle

Hoggle

System Queue Go to Revision Page User Browser Help

88 edits per minute, 1 reverts per minute

Page February History

User 86.130.203.224 Contribs

Revert this revision, and issue a user warning [Q]

Filtered edits 39 items

- Leeds Bradford Interna...
- Ramalinga Swamigal
- Hollywood Tonight
- Fusion power
- Xbox Live
- Business process outso...
- Rok Kronaveter
- Michel Vautrot
- Marian Mihail
- Saliva (band)
- Krisnan Inu
- List of Autobots
- British Rail Class 379
- Ariel (The Little Mermai...
- Nissan 350Z
- Zubin Potok
- List of programs broad...
- Foreign exchange mark...
- Talk:2011 Libyan uprisi...
- Talk:Banana equivalent ...
- Christian Schwarz-Schill...
- Iponan Local Network
- Louise Schroeder

February

Revision as of 11:52, 19 March 2011 (edit)
86.130.203.224 (talk)
(←History)

February was named after the Latin term "februum", which means "purification", via the purification ritual "[Februalia]" held on February 15 (full moon) in the old lunar [[Roman calendar]]. January and February were the last two months to be added to the Roman calendar, since the Romans originally considered winter a monthless period. They were added by [[Numa Pompilius]] about 713 BC. February remained the last month of the calendar year until the time of the [[decemvirs]] (c. 450 BC), when it became the second month. At certain intervals February was truncated to 23 or 24 days; and a 27-day intercalary month, [[Mercedonius|Intercalaris]], was inserted immediately after February to realign the year with the [[seasons]].

Under the reforms that instituted the [[Julian calendar]], Intercalaris was abolished, leap years occurred regularly every fourth year, and in leap years February gained a 29th day. Thereafter, it remained the second month of the calendar year, meaning the order that months are displayed (January, February, March, ..., December) within a year-at-a-glance calendar. Even during the Middle Ages, when the numbered [[Anno Domini]] year began on March 25 or December 25, the second month was February whenever all twelve months were displayed in order. The [[Gregorian calendar]] reforms made slight changes to the system for determining which years were leap years and thus contained a 29-day February. **Also some one very specil was born in this month.**

Historical names for February include the [[Old English language|Old English]] terms Solmonath (mud month) and Kale-monath (named for [[cabbage]]) as well as [[Charlemagne]]'s designation Hornung. In [[Finnish language|Finnish]], the month is called "helmikuu", meaning "month of the pearl"; when snow melts on tree branches, it forms droplets, and as these freeze again, they are like pearls of ice. In [[Polish language|Polish]] and [[Ukrainian language|Ukrainian]], respectively, the month is called "luty" or "лютій", meaning the month of ice or hard frost.

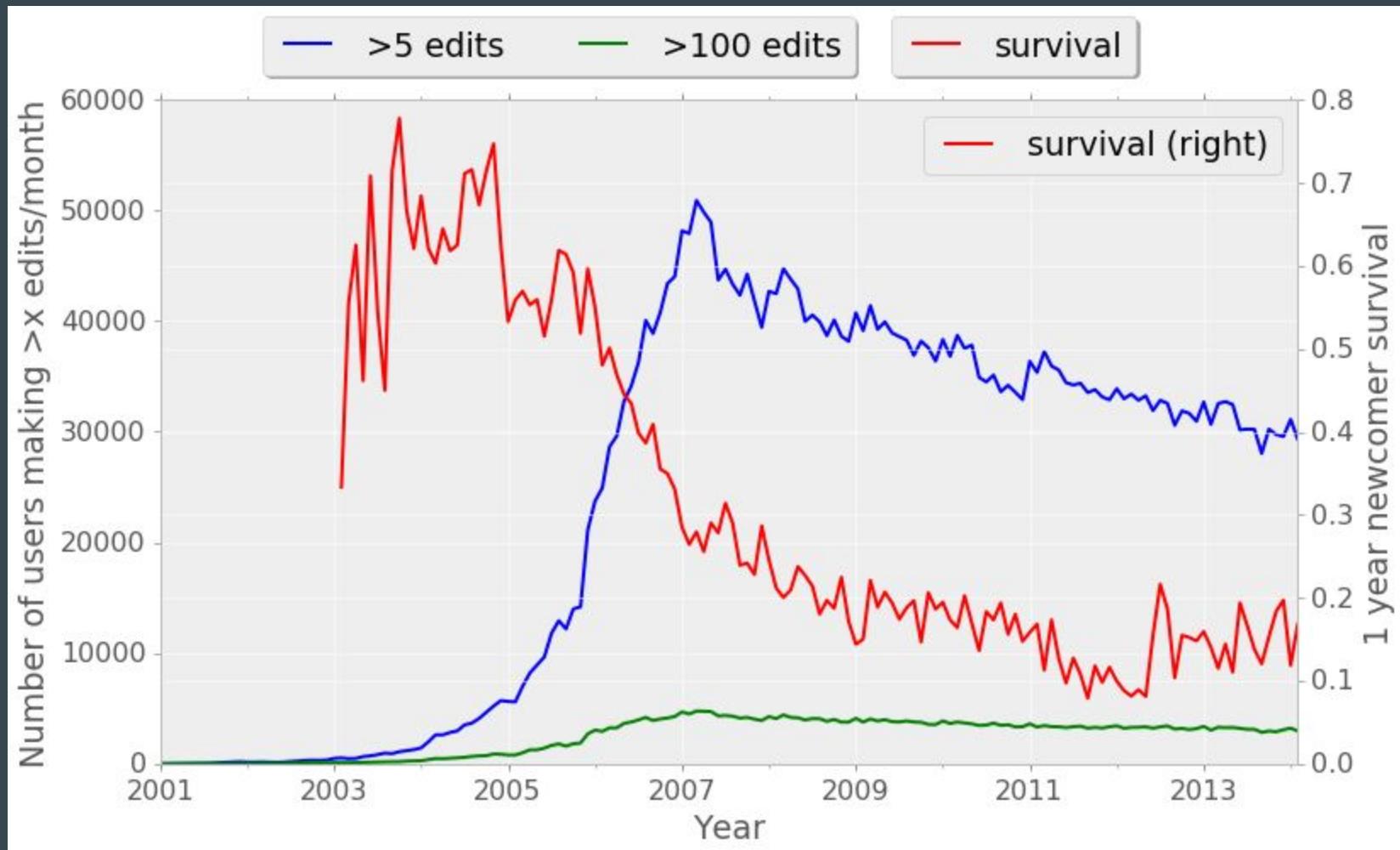
== Pronunciation ==

== Pronunciation ==

2011-03-19 11:52:35 -- Connected to IRC recent changes feed
2011-03-19 11:52:27 -- Attempting connection to IRC recent changes feed; this may take a few moments...

Inbox (34 total...) Wikipedia:Hug... @Hedgehog45... matthew@ma... Downloads matthew@ma... matthew@ma... Hoggle

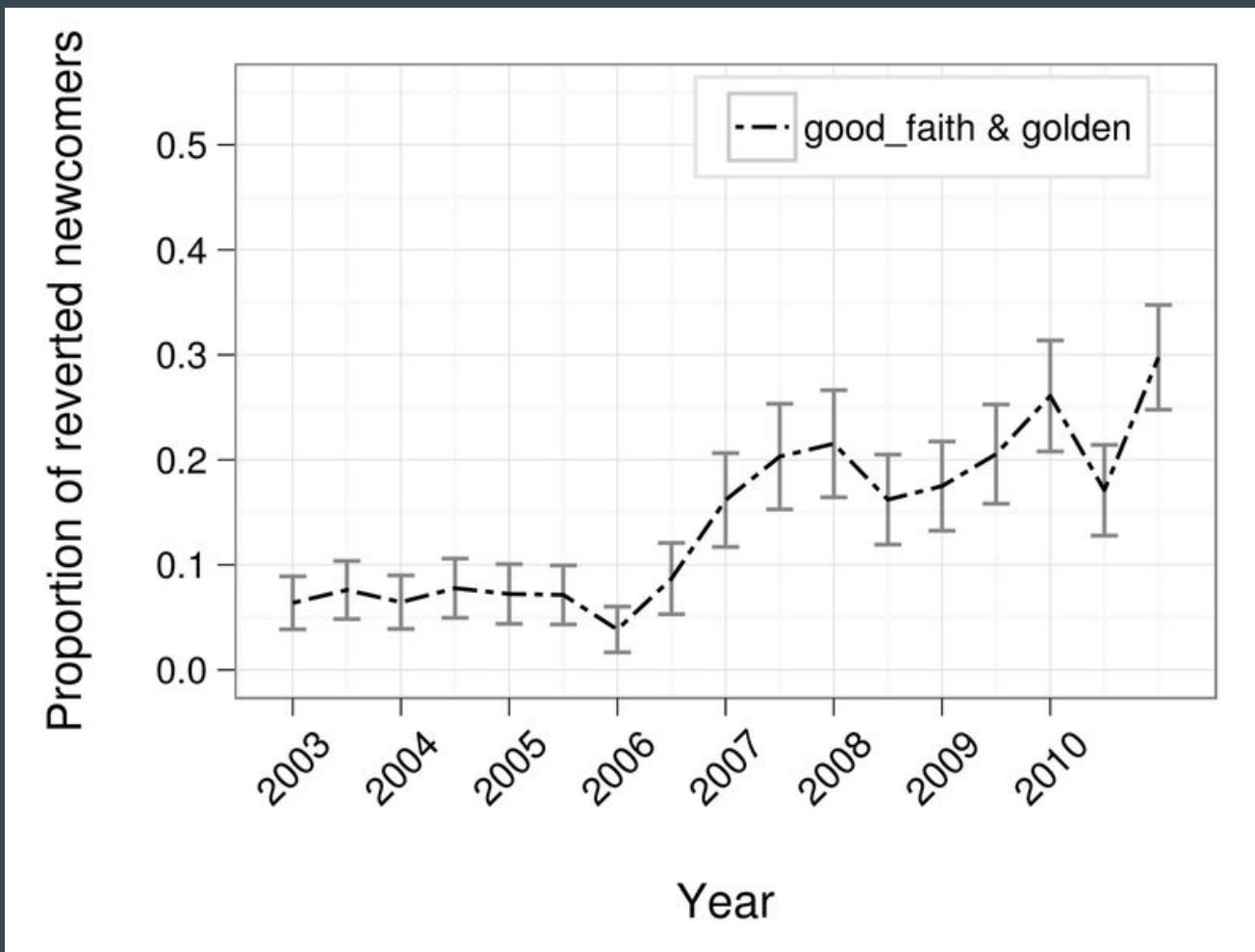
Wikipedia's user peak and plateau



We were asked:

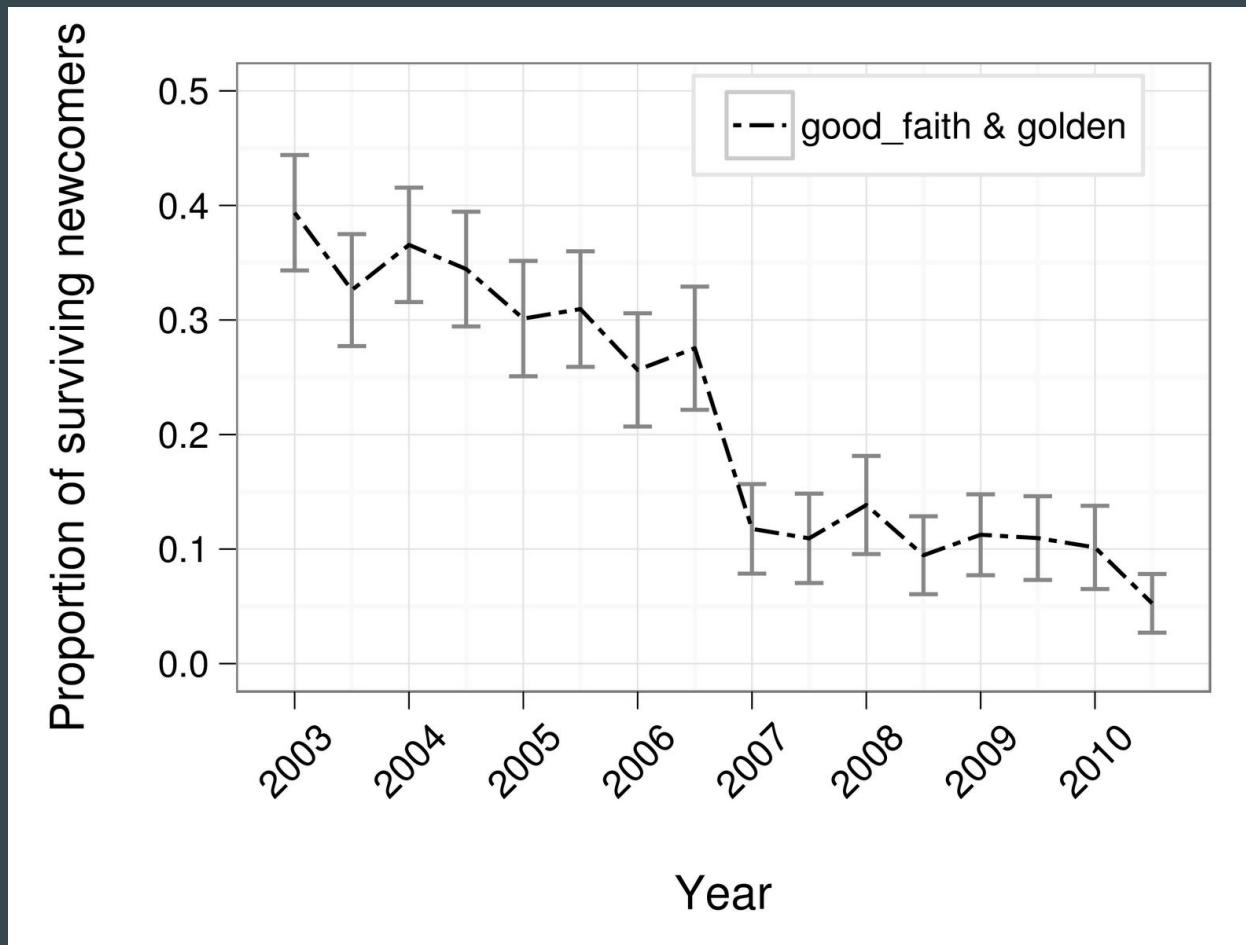
Aren't today's
newcomers just worse
than they were back in
the good old days?

Proportion of good faith newcomers reverted



Halfaker, Geiger, Morgan, and Riedl, 2012. "The Rise and Decline of an Open Collaboration System." American Behavioral Scientist.
https://meta.wikimedia.org/wiki/Research:The_Rise_and_Decline

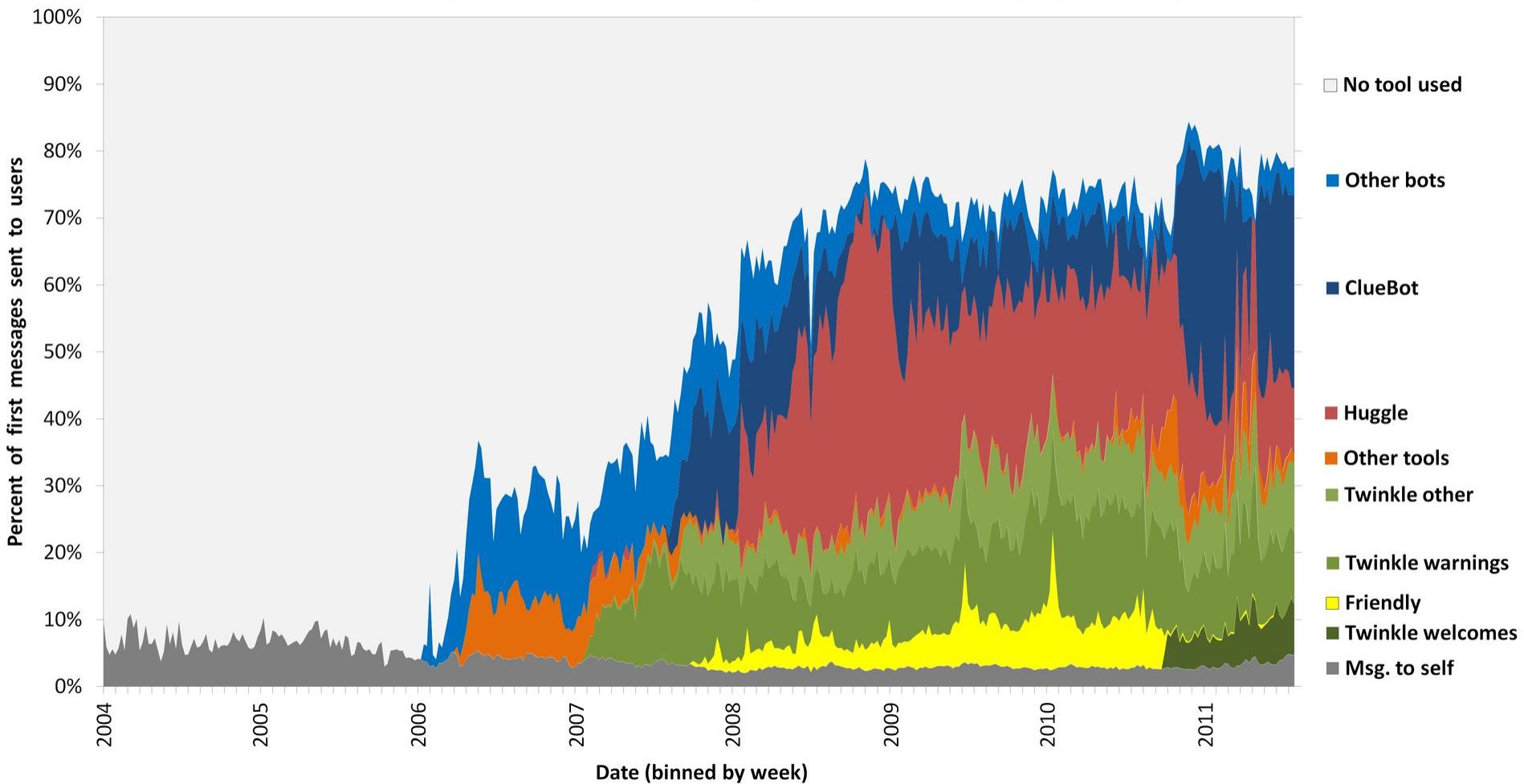
Proportion of good faith newcomers retained



Halfaker, Geiger, Morgan, and Riedl, 2012. "The Rise and Decline of an Open Collaboration System." American Behavioral Scientist.
https://meta.wikimedia.org/wiki/Research:The_Rise_and_Decline

First message sent to new users

First messages sent to new users, by tool and bot used (proportional)



Halfaker, Geiger, Morgan, and Riedl, 2012. "The Rise and Decline of an Open Collaboration System." American Behavioral Scientist.

https://meta.wikimedia.org/wiki/Research:The_Rise_and_Decline

A/B testing templated messages



Welcome to Wikipedia. Although everyone is welcome to contribute to Wikipedia, at least one of your recent edits, such as the edit you made to Science has been reverted, as it appears to be unconstructive. Use the sandbox for testing; if you believe the edit was constructive, ensure that you provide an informative edit summary. You may also wish to read the introduction to editing. Thank you. EPOCHFAIL(talk|work) 14:44, 21 September 2011 (UTC)

VS.



Hello and welcome! I edit Wikipedia too, under the username EpochFail. Wikipedia is written by people like you and me, so thank you for taking the time to participate. I wanted to let you know that I undid one of your recent contributions, such as the one you made to Science with this edit, because it didn't appear constructive to me. If you think I made a mistake, or if you have any questions about editing, you can leave me a message on my talk page. Thanks! EPOCHFAIL(talk|work) 14:44, 21 September 2011 (UTC)

A/B testing templated messages

- Many times, a new message we all thought was great would have no measurable effect
- Some messages that did have effects only worked for more experienced newcomers ($\geq 5-10$ previous edits)
- It was a great way to get volunteer contributors, researchers, and Foundation staff together

The Teahouse

Wikipedia:Teahouse

From Wikipedia, the free encyclopedia

Welcome to the **teahouse**

A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.

Do you have a question about editing?

[Get answers »](#)

Recent questions...

I'm not allowed to update a profile page for my boss, as it creates a conflict of interest. But I do have an update drafted and all of the information is accurate/factual. Will someone help me by updating the page on my behalf? The posting is in the "talk" section of my account. Thanks in advance for any help.

Gaff225 ([talk](#)) 16:15, 10 August 2016 (UTC)

[Find the answer here](#)

[...see another question](#)

Hosts are here to help »


Nihen, Teahouse host

Guests, create a profile »


Radioactivated
Editing is something I like to do sometimes to make Wikipedia better by a bit.

The Teahouse



Hi **Killerbeez3!** Thanks for contributing to Wikipedia. Be our guest at [the Teahouse!](#) The Teahouse is a friendly space where new editors can ask questions about contributing to Wikipedia and get help from peers and experienced editors. I hope to see you there! [I JethroBT \(I'm a Teahouse host\)](#)

[Visit the Teahouse](#)

This message was delivered automatically by your robot friend, [HostBot \(talk\)](#) 17:25, 24 November 2015 (UTC)

- A bot invites some newcomers who have recently made their first 10+ edits
- 4.5% participation rate from bot invites

The Teahouse: controlled experiment

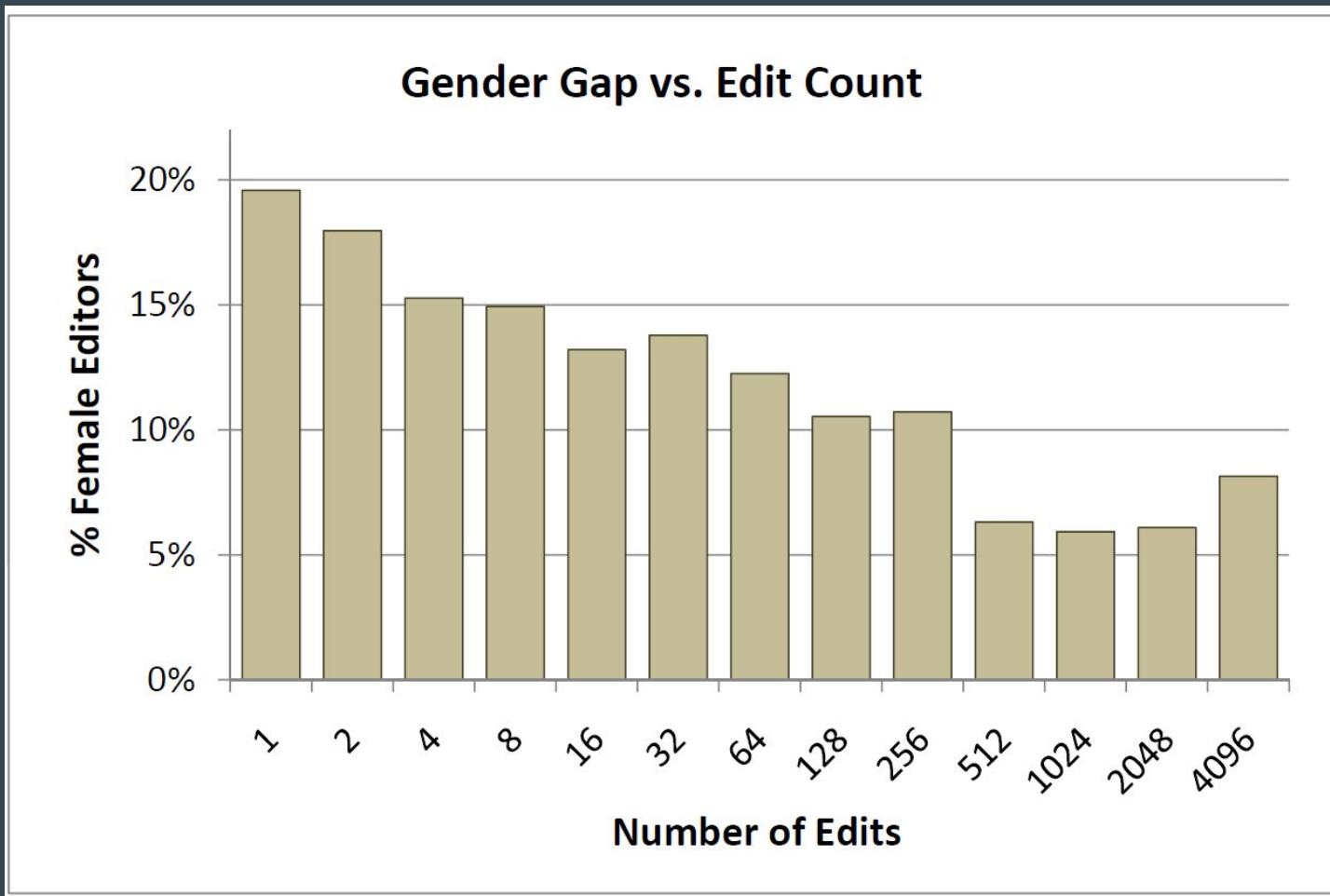
- Date: Feb 25th - May 19th, 2015
- Population: Newcomers who made ≥ 10 edits in their first 24 hours or made ≥ 20 edits in their first 4 days
- Three groups: no invite (A), invite but didn't participate (B), invited and participated (guests)

group	total editors	# recently active	% recently active	avg recent edits	avg recent article edits	# blocked	% blocked
Control A	192	17	9%	19.4	11.5	13	7%
Control B	200	21	11%	67.4	64.5	10	5%
Teahouse Guests	190	63	33%	121.0	92.1	12	6%

The Teahouse: Survey

- Overall satisfaction w/ the Teahouse:
 - 71% of newcomers stated they were satisfied or very satisfied
 - 70% of veterans stated they were satisfied or very satisfied
- 28% of newcomer respondents were women
 - vs. 8-13% of all WP editors are women

The Gender Gap -- retention



Lam, Uduwage, Dong, Sen, Musicant, Terveen, and Riedl. 2011. WP:clubhouse?: an exploration of Wikipedia's gender imbalance. In WikiSym 2011.
<http://dx.doi.org/10.1145/2038558.2038560>

The Gender Gap -- reversion

Edit #	Females		Males	
	N	Mean	N	Mean
1	6,305	6.99% ***	53,738	4.96%
2-3	4,989	6.04% ***	45,514	4.44%
4-7	3,798	4.69% **	37,272	3.98%
8-15	2,871	3.47%	30,066	3.12%
16-31	2,080	2.56%	23,798	2.66%
32-63	1,490	2.33%	18,270	2.49%
64-127	1,039	1.97%	13,850	2.27%
128-255	749	2.29%	10,355	2.07%

Lam, Uduwage, Dong, Sen, Musicant, Terveen, and Riedl. 2011. WP:clubhouse?: an exploration of Wikipedia's gender imbalance. In WikiSym 2011.
<http://dx.doi.org/10.1145/2038558.2038560>

Nine Reasons Women Don't Edit Wikipedia: In Their Own Words

By Sue Gardner, WMF ex-Executive Director

1. Difficult editing interface
2. Not having enough free time
3. A lack of self-confidence
4. Aversion to conflict and lengthy edit wars
5. Belief that their contributions are likely to be deleted
6. Some find its overall atmosphere misogynistic
7. Wikipedia culture is sexual in various ways
8. Being addressed as male is off-putting to women
9. Fewer opportunities for social relationships

<https://suegardner.org/2011/02/19/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/>

Edit-a-thons



Edit-a-thons and community outreach

- A mix of veterans and newcomers is needed
- Seeks to teach necessary tech skills and jargon to participate in Wikipedia
 - But one day is usually not enough....
- Evaluation and metrics infrastructure
 - meta.wikimedia.org/wiki/Grants:Evaluation
- New initiatives for university courses, non-anglophone countries, and long-term community engagement

A broad community effort on improving representation

- Wikiproject Countering Systemic Bias
 - enwp.org/WP:CSB
- The Gender Gap Task Force
 - enwp.org/WP:GGTF
- The Wiki Education Foundation
 - wikiedu.org
- WMF's Individual Engagement Grants
 - meta.wikimedia.org/wiki/Grants:IEG

Other efforts & projects

- Streamlined/simplified guides to policy
- A “draft” article space, with laxer rules
- Ways for newcomers to signal good faith intent
- Semi-automated tools beyond quality control, like connecting mentors and mentees (Snuggle)
- Provide ways for newcomers to take on more low-risk tasks (Article Feedback)
- Co-created machine learning projects (ORES)
- Long-term qualitative studies of contributors, especially those in marginalized groups

Thank you!

Thanks to all the Wikipedians,
Wikiresearchers, and Foundation staff!

Stuart Geiger, ethnographer/postdoc

UC-Berkeley Institute for Data Science (BIDS)

@staeiou | stuart@stuartgeiger.com

github.com/staeiou/wiki-stat-notebooks