

Who Writes Wikipedia?

September 4, 2006

[Original link](#)

Wikimedia 2006 Elections

Part 1: [Wikimedia at the Crossroads](#)

Part 2: **Who Writes Wikipedia?**

Part 3: [Who Runs Wikipedia?](#)

Part 4: [Making More Wikipedians](#)

Part 5: [Making More Wikipedias](#)

Part 6: [Code, and Other Laws](#)

Translations: , [Español](#), [Deutsch](#), [Français](#) ([add](#))

Vote for me in the election for the Wikimedia Foundation's Board of Directors.

I first met Jimbo Wales, the face of Wikipedia, when he came to speak at Stanford. Wales told us about Wikipedia's history, technology, and culture, but one thing he said stands out. "The idea that a lot of people have of Wikipedia," he noted, "is that it's some emergent phenomenon — the wisdom of mobs, swarm intelligence, that sort of thing — thousands and thousands of individual users each adding a little bit of content and out of this emerges a coherent body of work."† But, he insisted, the truth was rather different: Wikipedia was actually written by "a community ... a dedicated group of a few hundred volunteers" where "I know all of them and they all know each other". Really, "it's much like any traditional organization."

The difference, of course, is crucial. Not just for the public, who wants to know how a grand thing like Wikipedia actually gets written, but also for Wales, who wants to know how to run the site. "For me this is really important, because I spend a lot of time listening to those four or five hundred and if ... those people were just a bunch of people talking ... maybe I can just safely ignore them when setting policy" and instead worry about "the million people writing a sentence each".

So did the Gang of 500 actually write Wikipedia? Wales decided to run a simple study to find out: he counted who made the most edits to the site. "I expected to find something like an 80-20 rule: 80% of the work being done by 20% of the users, just because that seems to come up a lot. But it's actually much, much tighter than that: it turns out over 50% of all the edits are done by just .7% of the users ... 524 people. ... And in fact the most active 2%, which is 1400 people, have done 73.4% of all the edits." The remaining 25% of edits, he said, were from "people who [are] contributing ... a minor change of a fact or a minor spelling fix ... or something like that."

Stanford wasn't the only place he's made such a claim; it's part of the standard talk he gives all over the world. "This is the group of around a thousand people who really matter", he told us at Stanford. "There is this tight community that is actually doing the bulk of all the editing", he explained at the Oxford Internet Institute. "It's a group of around a thousand to two thousand people," he informed the crowd at GEL 2005. These are just the three talks I watched, but Wales has given hundreds more like them.

At Stanford the students were skeptical. Wales was just counting the number of edits — the number of times a user changed something and clicked save. Wouldn't things be different if he counted the amount of text each user contributed? Wales said he planned to do that in "the next revision", but was sure "my results are going to be even stronger", because he'd no longer be counting vandalism and other changes that later got removed.

Wales presents these claims as comforting. Don't worry, he tells the world, Wikipedia isn't as shocking as you think. In fact, it's just like any other project: a small group of colleagues working together toward a common goal. But if you think about it, Wales's view of things is actually much *more* shocking: around a thousand people wrote the world's largest encyclopedia in four years for free? Could this really be true?

Curious and skeptical, I decided to investigate. I picked an article at random ("Alan Alda") to see how it was written. [Today the Alan Alda page](#) is a pretty standard Wikipedia page: it has a couple photos, several pages of facts and background, and a handful of links. But [when it was first created](#), it was just two sentences: "Alan Alda is a male actor most famous for his role of Hawkeye Pierce in the television series MASH. Or recent work, he plays sensitive male characters in drama movies." How did it get from there to here?

Edit by edit, I watched the page evolve. The changes I saw largely fell into three groups. A tiny handful — probably around 5 out of nearly 400 — were "vandalism": confused or malicious people adding things that simply didn't fit, followed by someone undoing their change. The vast majority, by far, were small changes: people fixing typos, formatting, links, categories, and so on, making the article a little nicer but not adding much in the way of substance. Finally, a much smaller amount were genuine additions: a couple sentences or even paragraphs of new information added to the page.

Wales seems to think that the vast majority of users are just doing the first two (vandalizing or contributing small fixes) while the core group of Wikipedians writes the actual bulk of the article. But that's not at all what I found. Almost every time I saw a substantive edit, I found the user who had contributed it was not an active user of the site. They generally had made less than 50 edits (typically around 10), usually on related pages. Most never even bothered to create an account.

To investigate more formally, I purchased some time on a computer cluster and downloaded a copy of the Wikipedia archives. I wrote a little program to go through each edit and count how much of it remained in the latest version.[†] Instead of counting edits, as Wales did, I counted the number of letters a user actually contributed to the present article.

If you just count edits, it appears the biggest contributors to the Alan Alda article (7 of the top 10) are registered users who (all but 2) have made thousands of edits to the site. Indeed, #4 has made over 7,000 edits while #7 has over 25,000. In other words, if you use Wales's methods, you get Wales's results: most of the content seems to be written by heavy editors.

But when you count letters, the picture dramatically changes: few of the contributors (2 out of the top 10) are even registered and most (6 out of the top 10) have made less than 25 edits to the entire site. In fact, #9 has made exactly one edit — this one! With the more reasonable metric — indeed, the one Wales himself said he planned to use in the next revision of his study — the result completely reverses.

I don't have the resources to run this calculation across all of Wikipedia (there are over 60 million edits!), but I ran it on several more randomly-selected articles and the results were much the same. For example, the largest portion of the Anaconda article was written by a user who only made 2 edits to it (and only 100 on the entire site). By contrast, the largest number of edits were made by a user who appears to have contributed no text to the final article (the edits were all deleting things and moving things around).

When you put it all together, the story become clear: an outsider makes one edit to add a chunk of information, then insiders make several edits tweaking and reformatting it. In addition, insiders rack up thousands of edits doing things like changing the name of a category across the entire site — the kind of thing only insiders deeply care about. As a result, insiders account for the vast majority of the edits. But it's the outsiders who provide nearly all of the content.

And when you think about it, this makes perfect sense. Writing an encyclopedia is hard. To do anywhere near a decent job, you have to know a great deal of information about an incredibly wide variety of subjects. Writing so much text is difficult, but doing all the background research seems impossible.

On the other hand, everyone has a bunch of obscure things that, for one reason or another, they've come to know well. So they share them, clicking the edit link and adding a paragraph or two to Wikipedia. At the same time, a small number of people have become particularly involved in Wikipedia itself, learning its

policies and special syntax, and spending their time tweaking the contributions of everybody else.

Other encyclopedias work similarly, just on a much smaller scale: a large group of people write articles on topics they know well, while a small staff formats them into a single work. This second group is clearly very important — it's thanks to them encyclopedias have a consistent look and tone — but it's a severe exaggeration to say that they wrote the encyclopedia. One imagines the people running *Britannica* worry more about their contributors than their formatters.

And Wikipedia should too. Even if all the formatters quit the project tomorrow, Wikipedia would still be immensely valuable. For the most part, people read Wikipedia because it has the information they need, not because it has a consistent look. It certainly wouldn't be as nice without one, but the people who (like me) care about such things would probably step up to take the place of those who had left. The formatters aid the contributors, not the other way around.

Wales is right about one thing, though. This fact does have enormous policy implications. If Wikipedia is written by occasional contributors, then growing it requires making it easier and more rewarding to contribute occasionally. Instead of trying to squeeze more work out of those who spend their life on Wikipedia, we need to broaden the base of those who contribute just a little bit.

Unfortunately, precisely because such people are only occasional contributors, their opinions aren't heard by the current Wikipedia process. They don't get involved in policy debates, they don't go to meetups, and they don't hang out with Jimbo Wales. And so things that might help them get pushed on the backburner, assuming they're even proposed.

Out of sight is out of mind, so it's a short hop to thinking these invisible people aren't particularly important. Thus Wales's belief that 500 people wrote half an encyclopedia. Thus his assumption that outsiders contribute mostly vandalism and nonsense. And thus the comments you sometimes hear that making it hard to edit the site might be a good thing.

"I'm not a wiki person who happened to go into encyclopedias," Wales told the crowd at Oxford. "I'm an encyclopedia person who happened to use a wiki." So perhaps his belief that Wikipedia was written in the traditional way isn't surprising. Unfortunately, it is dangerous. If Wikipedia continues down this path of focusing on the encyclopedia at the expense of the wiki, it might end up not being much of either.

More: [Further reading](#) | [Vote for me](#) | [Discuss on Meta-Wiki](#) | [Discuss on Reddit](#)

Follow ups: [Summary of further research](#), [summary of responses](#).

Additional information: [my collection of research on this question](#).