

NewScientistTech

[Home](#) | [News](#) | [Forums](#) | [Special Reports](#) | [Subscribe](#) | [Search](#) | [RSS](#)

Wikipedia 2.0 - now with added trust

20 September 2007

NewScientist.com news service

Jim Giles

WIKIPEDIA'S entry on [Albert Einstein](#) looks good. Covering each phase of the physicist's life, from childhood to death, it tells readers about his politics, religion and science. Honours named after him and books and plays about his life are listed. But there is one snag: there is no way to tell whether the information is true.

It is a problem that dogs every Wikipedia entry. Because anyone can edit any entry at any time, users do not generally know if they are looking at a carefully researched article, one that has had errors mischievously inserted, or a piece written by someone pushing their own agenda. As a result, although Wikipedia has grown in size and reputation since its launch in 2001 - around [7 per cent](#) of all internet users now visit the site on any given day - its information continues to be treated cautiously.

That could be about to change. Over the past few years, a series of measures aimed at reducing the threat of vandalism and boosting public confidence in Wikipedia have been developed. Last month a project designed independently of Wikipedia, called [WikiScanner](#), allowed people to work out what the motivations behind certain entries might be by revealing which people or organisations the contributions were made by (see "Who's behind the entries?"). Meanwhile the Wikimedia Foundation, the charity that oversees the online encyclopedia, now says it is poised to trial a host of new trust-based capabilities.

The changes could help transform the encyclopedia from a rough guide into a trusted authority. But they might also erode the very freedoms that encourage people to contribute to the encyclopedia in the first place. Either way, the stage appears set for Wikipedia 2.0.

News of the plans came to light last August when Wikipedia co-founder Jimmy Wales announced [changes to the editing restrictions](#) on the German-language version. However, implementing those changes turned out to be more difficult than anticipated and has still not happened. Now *New Scientist* has learned that Wikimedia plans to start the first trial of the changes this month.

The shift is a dramatic one for the encyclopedia. For now, edits to an entry can be made by any user and appear immediately to all readers. In the new version, only edits made by a separate class of "trusted" users will be instantly implemented.

To earn this trusted status, users will have to show some commitment to Wikipedia, by making 30 edits in 30 days, say. Other users will have to wait until a trusted editor has given the article a brief look, enough to confirm that the edit is not vandalism, before their changes can be viewed by readers.

This is sure to ease some readers' doubts. Most malicious edits involve crude acts of vandalism, such as the deletion of large chunks of text. Now such changes will rarely make it into articles.

These benefits will come at a price, though. New users could be deterred from participating, since they will lose the gratification that comes from seeing their edit instantly implemented. That could reduce the number of editors as well as creating a

Advertisement

class system that divides frequent users from readers. The trusted editors, likely to number around 2000, may also find that articles are being changed too fast for them to monitor.

Not all versions of the encyclopedia will follow this route, says Erik Müller of the Wikimedia Foundation. While editors on the German version are happy with a hierarchy of contributors, the English editors favour a more egalitarian approach. So English readers are likely to continue to see the latest version of an entry, with a page that has been certified as vandalism-free by trusted editors available via a link.

For edits that are more subtly inaccurate, perhaps because they have been designed to promote an agenda, another tool is in store. It allows select groups of editors, probably associated with specific subject areas, to vote on whether an article should be flagged as high quality. Readers would still see the latest version of an article by default, but a link to a high-quality version, if it exists, would also be available.

As well as relying on trusted editors, Wikipedia's upgrade will involve automatically awarding trust ratings to chunks of text within a certain article. Müller says the new system is due to be incorporated into Wikipedia within the next two months, as an option for the different language communities.

The [software](#) that will do this, created by Luca de Alfaro and colleagues at the University of California, Santa Cruz, starts by assigning each Wikipedia contributor a trust rating using the encyclopedia's vast log of edits, which records every change to every article and the editor involved. Contributors whose edits tend to remain in place are awarded high trust ratings; those whose changes are quickly altered get a low score. The rationale is that if a change is useful and accurate, it is likely to remain intact during subsequent edits, but if it is inaccurate or malicious, it is likely to be changed. Therefore, users who make long-lasting edits are likely to be trustworthy. New users automatically start with a low rating.

De Alfaro has shown that the software's ratings correlate with human judgements. Using data from the Italian Wikipedia, his software assigned trust ratings to editors based on the persistence of past contributions, and then asked volunteers to rate edits by those editors. Edits made by editors with ratings in the bottom 20 per cent were up to six times more likely to be judged as bad than those with higher ratings.

Once all contributors are rated, the software then uses this information to rate chunks of text. If whole entries have been contributed by one person and left unchanged, the text inherits the rating of that person. If text has been edited several times, then its rating is calculated using the ratings of all contributors. If a modification to an entry leaves a particular chunk unchanged, that chunk will get a high rating.

The system runs the risk of penalising editors who tackle malicious changes by correcting them, because the corrections are often quickly changed back to the malicious version by the vandals. To try to minimise this, the drop in an editor's rating that occurs when their edit is changed will depend upon the rating of the other editor involved.

Once all text has been rated, the software colour-codes it, with darker shades for lower ratings. Readers will then have the option of clicking through to a colour-coded page, allowing them to immediately judge which parts of an entry to trust.

Automation introduces challenges, however. New editors could get put off when they see their text flagged as questionable by default. A high rating may also become an end in itself, leading people to come up with ways to get their text rated highly without necessarily enhancing its quality. Although de Alfaro won't publish the ratings, Wikipedia's log is public, so anyone with a copy of the algorithm could publish the results.

Müller says that ultimately the best way to make Wikipedia more trustworthy might be to combine the trusted users approach with the automated one. "We could have an icon saying that the version you're looking at is unlikely to contain vandalism and also whether it was a human or computer that made the decision," he says. "As simple as possible, that's the main goal."

Who's behind the entries

Wikipedia has a two-pronged plan for reducing vandalism and inaccuracies on its site, but an independent website launched last month might discourage another kind of bad behaviour: agenda-driven edits.

[WikiScanner](#) allows people to find out which organisations are behind contributions to Wikipedia entries by taking the IP address of the computer that submitted the entry,

which Wikipedia makes public, and looking it up in a second database that links organisations to their IP addresses.

The site has already revealed that staff at Fox News Network cut sections of an article criticising the channel's correspondents and that someone at Diebold, which manufactures voting machines, removed paragraphs questioning the machines' reliability.

The site's influence could go further, says Wikipedia co-founder Jimmy Wales. A similar system could be created that displays the name of someone's organisation as they are submitting their edits, warning them in real time that it will be clear, to anyone who wants to know, who they are. That might make them think twice about trying to distort an entry.

If it had been in place, such a tool might have restrained the employees of some prestigious institutions as they contributed to the entry on [George W. Bush](#). At the BBC, a staff member changed Bush's middle name to "Wanker", while over at *The New York Times* an employee simply contributed the words "jerk jerk jerk".

Related Articles

Wikipedia founder buys Grub for distributed search
<http://technology.newscientist.com/article/dn12387>
31 July 2007

Wikipedia ranks top for news and information
<http://technology.newscientist.com/article/dn12223>
09 July 2007

Interview: Knowledge to the people
<http://technology.newscientist.com/article/mg19325896.300>
31 January 2007

Open-source search engine gangs up on Google
<http://technology.newscientist.com/article/mg19426066.500>
30 May 2007

Weblinks

WikiScanner
<http://wikiscanner.virgil.gr/>

Wikipedia Trust-coloring Demo
<http://trust.cse.ucsc.edu/>

Albert Einstein - Wikipedia
http://en.wikipedia.org/wiki/Albert_Einstein

Can German engineering fix Wikipedia? - News.com
http://news.com.com/2100-1038_3-6108495.html

Content-Driven Author Reputation and Text Trust for the Wikipedia
http://trust.cse.ucsc.edu/UCSC_Wiki_Lab?action=AttachFile&do=get&target=wikimania07.pdf

Wikipedia traffic graph - Alexa
http://www.alexa.com/data/details/traffic_details?url=http%3A%2F%2Fwikipedia.org

From issue 2622 of New Scientist magazine, 20 September 2007, page 28-29

[Close this window](#)