How accurate is Wikipedia's content?

TIMES NEWS NETWORK [WEDNESDAY, FEBRUARY 01, 2006 12:51:51 AM]

NRIs Rang De Basanti FREE DVD

It's not easy being Wikipedia, a free web encyclopedia created and edited by anonymous contributors. Just ask founder Jimmy Wales, who has seen his creation come under fire in just a few short months as the site fends off vandalism and charges of inaccurate entries. "Wikipedia has always been in a state of change," says Wales, in defense of his product.

That's putting it mildly.

On November 29, journalist John Seigenthaler, Sr., once a member of Robert Kennedy's staff, penned an op-ed piece in USA Today noting that an article on Wikipedia had incorrectly linked him to the assassination of Robert Kennedy and John F Kennedy. The article, which stayed on the site for four months, stated that "John Seigenthaler, Sr. was the assistant to Attorney General Robert Kennedy in the early 1960s.

For a brief time, he was thought to have been directly involved in the Kennedy assassinations of both John and his brother Bobby. Nothing was ever proven." Wikipedia eventually deleted the inaccurate information, and now even contains an entry entitled, "John Seigenthaler, Sr. Wikipedia biography controversy", explaining the history of its own error.

And last month, former MTV VJ and podcasting pioneer Adam Curry was accused of editing out Wikipedia's references to Kevin Marks, another early podcasting luminary. In a December 2 blog entry, Curry owned up to the switch and apologised.

Today, popular or potentially controversial Wikipedia entries, such as one on George W Bush, greet readers with the following: "In response to recent vandalism, editing of this page by new or anonymous users has been temporarily disabled. Please discuss possible changes or request unprotection."

Wikipedia, founded in 2001 as a non-profit organization and supported mainly by donations, allows anyone with internet access to edit its articles. The premise is that collective knowledge, which some call 'open source' content, is every bit as valuable as professionally edited content.

As a result, Wikipedia has become a hybrid encyclopedia/current events site that raises a number of interesting questions, including: Just how accurate is free content, given recent events at Wikipedia? Does the aggregate 'wisdom of the crowd' trump the expertise of knowledgeable individuals? Does Wikipedia's policing mechanism work? Does the controversy over Wikipedia merely reflect further tension between old and new media?

The answer to all those questions, according to experts at Wharton, boils down to two words: It depends. According to Eric Clemons, professor of operations and information management, it is foolish to take Wikipedia's definitions as gospel, but the site is still worth reading from time to time. Marketing professor Peter Fader notes that Wikipedia shows there is wisdom in crowds, but a better user rating system would filter out those who post bogus information. Joel Waldfogel, professor of business and public policy, agrees that much of the concern about Wikipedia is just a new spin on whether old media (printed encyclopedias, in this case) stand a chance against the new breed of instantly updatable online media.

For Kendall Whitehouse, senior director of information technology at Wharton, the most interesting question is "whether the wisdom of the crowd is ultimately a better approach compared to scholarly review and edited content." As Whitehouse observes, "Wikipedia's strength is that it has thousands of eyes looking at it. The hope is that errors will be quickly caught and corrected." But Whitehouse also points out that as the scale of Wikipedia's content expands, even a thousand eyes can miss certain details — which is apparently what happened in the case of the Seigenthaler material.

Wales doesn't dismiss concerns about Wikipedia, and notes that he has implemented changes to prevent abuse, such as only allowing registered users to make certain types of changes. He adds, however, that Wikipedia has been slammed for mistakes simply because it's relatively new, and he compares the attention swirling around the site to stories that focused on quirky items for sale on eBay in 1999 and 2000. "The media loves a story," he says.

Growing Pains

Just the fact that Wikipedia has attracted so much attention indicates the clout it has achieved since being created almost five years ago. Indeed, a December 14 study of 42 entries by Nature magazine puts Wikipedia almost on a par with Britannica in

terms of accurate science coverage. Nature found that the average science entry in Wikipedia had four errors while Britannica had three. The biggest complaint against Wikipedia was that the entries were confusing and poorly structured. Overall, concluded the Nature study, problems such as the Seigenthaler incident were the exception, not the rule.

Wikipedia has also grown at a rapid clip. There are 13,000 active contributors working on 1.8 million articles in more than 100 languages. On the English-language sites, 800 to 1,000 editors edit most of the entries, says Wales. According to Alexa, a web traffic tracking site, Wikipedia is the 37th most highly visited site on the internet.

But with its clout comes additional scrutiny, says Fader, who views the most recent flaps about Seigenthaler, Curry and suggested policy changes as "tempests in a teapot. Wikipedia is as credible as anything else I find on the web. It comes up a lot... but it doesn't mean that I stop my research there." Indeed, the key may be putting Wikipedia into a broader context. It can be a research tool, but it's far from definitive.

Wikipedia, itself, acknowledges this fact, noting on its site that "it differs from a paper-based reference source in some important ways. In particular, mature articles tend to be more comprehensive and balanced, while other (often fledgling) articles may still contain significant misinformation, unencyclopaedic content or vandalism. Users need to be aware of this in order to obtain valid information and avoid misinformation which has been recently added and not yet removed."

Community Conundrum

Despite that disclaimer, observers say Wikipedia is too often viewed as authoritative. David Winer, a software pioneer who contributed to several common standards used on the internet and published one of the first weblogs, said in a June blog entry that the biggest issue with Wikipedia is that it is widely viewed as authoritative but "can easily be manipulated by people with an axe to grind".

Ari Friedman, a senior at Wharton who founded an online journal called Copyright that plans to adopt some of Wikipedia's techniques to vet copyright research and case studies, says it would be foolish to dismiss Wikipedia's open-content approach. But, he adds, there are trade-offs with disseminating information via the crowd approach. "The primary trade-off is not really 'speed for accuracy' as some might suggest, but rather 'speed for guarantee'," says Friedman. "If the information absolutely must not be wrong for each and every person who views it, then collaborative content is not going to work very well."

Clemons notes that Wikipedia should never be the sole source of information. "I only use Wikipedia for things I know cold, looking for a terse description to share with my daughter doing homework so I don't have to write it myself. But if I don't know the subject, I would never consider Wikipedia as a source of information," he says, adding that the biggest issue is that Wikipedia provides no information on who posted the story, the writer's credentials or the motives behind changes. Waldfogel agrees. "With these sites it is buyer beware.

Anyone who is interested in the topic has to do more research. You can't think that these things are absolutely true." At issue is whether collaborative content, such as Wikipedia, is up to standards of professionally-edited content. Wales acknowledges that some of the attention given to Wikipedia of late could be attributed to the ongoing debate over old media versus new media, but adds that the online encyclopedia functions more like a traditional editor driven outlet in which writers post articles and editors check them. "The way Wikipedia operates is far more traditional than people realize," says Wales. "We have a core group of editors who do the work."

Policing Wikipedia

What succeeded in getting Wikipedia to this juncture — an open format where anyone can post articles — may not work as the online encyclopedia continues to grow, says Fader. The issue: There is no rating system to lend credibility to people who create and edit articles. For instance, a system that keeps editors of articles anonymous, yet rates them much like eBay rates merchants, could enhance Wikipedia's credibility. A contributor known for accuracy or subject expertise would be given, say, five stars. A vandal or someone shamelessly plugging a product would get one star.

In response to recent events, Wales has made a few changes to better police the site. In addition to implementing a "semi-protection" policy where pages facing vandalism problems can only be edited by registered users with accounts older than four days, site administrators will add a delay to high profile pages, such as the President Bush page that is currently closed.

The delay, similar to a "dump" button in radio that gives producers a few seconds to prevent something like crude language from hitting the airwaves, will mean that online vandals cannot quickly amend a page on a breaking news article, such as those posted when the Pope died. Wales notes that the length of the delay will be decided by the community on Wikipedia. "We have hundreds of people monitoring the site all the time. They will figure out the community norms."

Other changes are in development, but Wales didn't disclose anything specific, except to say that there are no plans to institute an eBay-like rating system such as the one Fader suggests. That type of system couldn't work and would "be destructive of the community", Wales suggests, because Wikipedia has 800 to 1,000 active editors. A rating system would be degrading as well as unrefined.

For instance, a biologist may be an expert on biology topics, but turn into a ranting madman when it comes to the Israeli-Palestinian conflict. "In that example, the person would probably be rated a three-star overall," says Wales. "But the rating would lose the human judgment element. In our system, we learn over time that you need to keep an eye on that editor" when he or she is working on topics related to the Middle East. Whitehouse acknowledges that some Wikipedia content may be inaccurate, but overall the policing system works. "Look at what Wikipedia has done in a short period. By and large, self policing has worked well."

Emerging Models

The challenge for Wikipedia will be to juggle the monitoring of highly visible pages and the low-traffic areas that harbor misinformation. "The process works when there is a back and forth that allows discussion and dialog," says Whitehouse. "The problem is policing the obscure articles no one is looking at." It's unclear how the Wikipedia model will evolve, but some potential rivals have their own ideas. A site called Digital Universe recently launched offering "stewards" to review content and providing two tiers of articles — those verified by experts and those not. The company bills itself as the "PBS of the Web" and a university of human thought and experience.

Friedman's Copyright, which will kick off in mid-February, is an open access, peer-reviewed journal on all aspects of copyright in the internet age. Authors will submit detailed research, case studies and opinion pieces vetted by rapid peer review. Copyright will seek a balance between offering access to all users, but making sure information is accurate.

According to Friedman, there is still a big role for collaborative content, but the filtering of vast amounts of information will become increasingly critical. While Wikipedia is effective, it "allows anyone to post [information] and just prays that the experts will wind up posting more than the malicious or incompetent," says Friedman. Copyright's approach will be open to all, but submissions will be vetted by, among others, Fader, Lawrence Lessig, a law professor at Stanford Law School, and William M Landes, a law professor at the University of Chicago, before being published.

Experts at Wharton agree that Wikipedia will weather its most recent flap, but its model may have to be tweaked. Whitehouse suggests that citizen-authored content is a valuable contribution to knowledge, but there is still a place for more edited information. Wikipedia's challenge: Find the middle ground.