Wiki Wiki! The Rise of Virtual Internet Communities

John Favaro

Viareggio, 25 March 2006

Introduction: a new word

It's not easy for a new word to be created in the English language. But that's just what happened last year, as a lexicographer working on current English dictionaries for Oxford University Press (publishers of the *Oxford English Dictionary*) wrote in a letter to Howard "Ward" Cunningham: "We're preparing a new edition of our largest current English dictionary and *wiki* is one of the new words we intend to include …"

It all started in Hawaii, where Cunningham arrived in Honolulu on his honeymoon in 1995. There he was directed by an agent at the airport to the shuttle bus service that connected the terminals with each other. The name of that service was (and still is today) the "Wiki Wiki Bus." "Wiki" is the word in the Hawaiian language for "quick."

Back from his honeymoon, Cunningham returned to his work as a software engineer. His career had been an eclectic one. Living and working in Oregon, in the Far West of America, he exhibited an appropriately pioneering spirit from the beginning, turning out a number of innovations together with his colleague and fellow pioneering spirit Kent Beck.

One of those innovations concerned something that originally had nothing to do with the world of computers, but rather with the world of architecture, where yet another pioneer named Christopher Alexander had introduced a radical new perspective in a 1977 book called *A Pattern Language*. In this book Alexander, a professor at the University of California at Berkeley, had presented a series of 253 "patterns."

What were these "patterns"? They were solutions to architectural problems, large and small – from designing a kitchen ("have a table at waist-height for putting down utensils") to designing an entire city ("around a central square"). This book had been quite a hit in its day, especially in the do-it-yourself home-building community, and is still considered one of the most practical books on architecture ever written.

Beck and Cunningham saw other possibilities in the patterns concept for their own work in software engineering. Anybody who has ever written a computer program knows how difficult it is – it seems like with every new program you have to start all over. In fact, that's what *did* happen all too often: people simply started over, not learning from previous experience. One reason for this wasteful and redundant behavior was that previous experience had never been written down: nobody had ever captured and documented the "patterns" of solutions to common software development problems for others to study and use in building their software programs, just like Alexander's architects could use his patterns in building their houses.

John Favaro
Via Gamerra, 21
56123 Pisa - Italy

Tel +39 050 55 60 74
Fax +39 050 6143 1145
john@favaro.net

www.favaro.net/john

Cunningham's plan was to collect these patterns from all over the software development world and deposit them in what he dubbed the Portland Pattern Repository. But he couldn't do this alone. He needed the *collaboration* of the entire software community, contributing patterns that they or others had used.

Collaboration over the Internet

But how to organize this collaborative effort? Fortunately, the software community was already extensively using the first medium perfectly suited to global collaboration – the Internet. But the Internet was just the *medium* for communication. By itself it provided no organized means to manage a collaborative effort. Being the mid-1990s, Cunningham decided to try using another new technology that was becoming famous in those years: the World Wide Web.

The Web was the perfect way to organize and display information to a broad public – even a global public. After it exploded onto the scene in 1993 with the introduction of the Mosaic browser, there were thousands of web sites making enormous amounts of information available on all kinds of topics. And so it seemed like the natural way to organize and make available the Portland Pattern Repository that Cunningham wanted to create.

But there was a problem. The World Wide Web is naturally a *broadcast* medium, "look but don't touch." Like a newspaper, or television, or radio, or any other broadcast medium, there is somebody who creates the information and broadcasts it to the rest of the world; the others can only look at it. But that's not what Cunningham wanted. He wanted *everybody* to be able to contribute information on his web site, not just himself. In short, he wanted an *interactive* medium, not a broadcast medium.

Now, traditional broadcast media like newspapers do usually have a minimal capability for interaction; for example, a reader can write a letter to the editor that might be published in a subsequent edition. Cunningham might have simply been content to do the same. As "webmaster" he could have just collected contributions that arrived from others, say, via e-mail, and published them in updated editions of the web site. But that was not good enough, for a couple of reasons:

◊ Firstly, it would create a bottleneck: everything would have to go through a single point, the webmaster himself, just like everything that arrives at a newspaper has to go through the editor before being published. This would have been quite a burden.

◊ Second, this single point of centralized editorial authority was not what Cunningham was looking for. He was looking for a genuinely democratic arrangement in which anybody could directly contribute, without the intervention of an editorial "censor."

Being a programmer himself, Cunningham did what came naturally: he set about creating the technology that was needed to make an interactive Web. The actual programming needed was surprisingly little, and the result was a web site where anybody could not only *read* the web pages (as in any other site) but *change* them, just like in a word processor, or even delete them or introduce new pages, in a quick and straightforward manner.

When thinking about a name to give his new technology, Cunningham first came up with the name Quickweb, to reflect the quick and easy manner in which it could be used. But then he remembered that honeymoon trip to Hawaii, and came up with the name by which it was finally known: the WikiWikiWeb.

The Wiki Wiki Web

In the WikiWikiWeb, there was no longer a single, centralized "webmaster." Rather, *everybody* was a webmaster, and everybody shared equal rights and equal responsibilities for the maintenance of the information. In effect, with the WikiWikiWeb, Cunningham had created one of the first *virtual communities* on the Internet.

With everybody sharing equal rights and responsibilities, the idea is for communities to grow up around a Wiki Wiki Web. For example, in the original Wiki for the Portland Pattern Repository, people could contribute their "patterns" directly by simply creating their own new web pages on the Wiki. Not only that, but others could then comment directly on what they had written, or even change or improve it, like some kind of gigantic community blackboard with everybody gathered around, contributing ideas and discussion.

But can this really work in practice? It's all very well if the people in the community are of good will and want to collaborate, as we might expect in a 'friendly" community like the software engineering community. But what happens if, for example, an argument starts? Suppose somebody doesn't agree with something that another person contributed to the Wiki? He could just change it to suit himself, or even remove it. Then the first person might put it right back up; and the other would rip it out again. And so on. Or, as another example, some vandal or spammer might simply start writing silly or obscene things in pages and destroying the serious things that other people are trying to do. With no central authority, no "police department" to step in and enforce order in the community, there is a serious danger of degeneration into chaos. What happens then?

The answer is that it depends on the community. The good functioning of a virtual community is similar to the good functioning of a real community. And here, the so-called "**Broken Window Theory**" is operative. This theory was first expounded in a 1982 article in the *Atlantic Monthly*. The authors wrote the following:

> … at the community level, disorder and crime are usually inextricably linked, in a kind of developmental sequence. Social psychologists and police officers tend to agree that if a window in a building is broken and is left unrepaired, all the rest of the windows will soon be broken. This is as true in nice neighborhoods as in rundown ones. Window-breaking does not necessarily occur on a large scale because some areas are inhabited by determined window-breakers whereas others are populated by window-lovers; rather, one unrepaired broken window is a signal that no one cares, and so breaking more windows costs nothing.

In a good neighborhood, if somebody sees a broken window, he will immediately fix it as a sign that people care. If an ill-intentioned person sees that whenever he vandalizes the site by writing something offensive or malicious, a person from the community will immediately erase it (like removing graffiti from the train station), then soon he will tire of it and move on.

To a great extent this has turned out to be true in Wikis established in "good" communities like the software engineering community. IBM did a study where they determined that vandalism to many Wiki sites was generally "repaired" by other users within only five minutes. But in the world at large, it has sometimes worked out differently. In one well-known experiment, the Los Angeles Times decided in 2005 to use a Wiki in order to promote discussions on their editorial pages. Anybody could chime in with an opinion, and the hope was that a vigorous and open debate would be fostered in a virtual community of interested readers. But the Wiki was so constantly devastated by vandals that the editors of the Times decided to shut it down and terminate the experiment.

Despite this particular failure, the Wiki concept has had enormous success, and surely the greatest and most publicized success has been an initiative to create the largest encyclopedia in the world.

The Wikipedia Community

Jimmy Wales grew up in Alabama, but after attending both college and graduate school there, he went off to Chicago to make his fortune as an options trader – and did just that, by the time he was in his early thirties. With his fortune made, and the financial security to do what he wanted, he hearkened back to one of the pastimes of his youth: reading the *World Book Encyclopedia*. His belief was that knowledge must be shared, and acting upon that belief, he started a project known as *Nupedia*, an initiative to create a free encyclopedia available to all. The Nupedia project was not really different from a standard encyclopedia project: acknowledged experts write an article on a subject, a panel of their peers review that article, and once everything has been approved, the article enters into the encyclopedia. The only difference was that this was all to be free of charge to the readers.

But soon afterwards, one of his colleagues introduced him to the idea of the Wiki web, where anybody could contribute. He thought that would be a nice way to collect articles for later review and inclusion into his Nupedia. But that's not the way it worked out: within a very short time, the Wiki-based project had far outgrown the Nupedia project, and the Nupedia project died a quiet death.

This left Jimmy Wales with a very unusual project indeed: an encyclopedia whose authors weren't acknowledged experts in their respective fields but just … anybody – anybody who wanted to write an article. And so this Wiki-based project went on to become one of the most well-known initiatives on the Web today: the **Wikipedia** project.

The Wikipedia today

Today the Wikipedia is published in more than 200 languages. The English version has more than one million articles. To give you a feeling for how large that is, note that the Encyclopedia Britannica has a "mere" 63 thousand articles.

But it's not just the sheer size of the Wikipedia that makes it different. It is also different in a way that only an online encyclopedia can be: it is *timely*. On Saturday morning, 11 February 2006, I looked into the Wikipedia. One of the articles was entitled "Opening Ceremony of the Winter Olympics in Torino." It gave a full account of everything that had happened in the opening ceremony, from start to finish. But that opening ceremony had taken place *the night before* – a mere twelve hours earlier! Only

an online encyclopedia, updated continuously every day, even every hour, could possibly achieve that kind of timeliness.

Who writes the Wikipedia?

Who are the people who actually write the Wikipedia? The short answer is "ordinary people like you and me." But who are these "ordinary people"? It turns out that there is an entire community of dedicated Wikipedia authors, who spend much of their time writing, correcting, and watching over the articles in the online encyclopedia. To give you a better idea, let's take a look at the Italian community of Wikipedia authors. There are around forty thousand of these authors, according to current estimates. They come from all walks of life, from professionals to manual laborers. They are old and young – the youngest is twelve years old.

Not all of these authors contribute regularly. The circle of "hardcore" authors probably numbers around 150. Around forty of these, in turn, have been agreed informally by the community to act as vigilantes of sorts to guarantee the quality and, above all, the neutrality of the articles that appear in the Italian version of the encyclopedia. For example, when an article is particularly controversial, it is they who may intervene to eliminate certain offensive or partisan phrases, or even go so far as to block the editing of an article altogether.

Exactly this situation has in fact arisen during this period in Italy. The presidential elections will be held here on 9 April 2006, as you know, and President Berlusconi is campaigning vigorously for re-election. The Wikipedia article on Mr. Berlusconi has become a kind of forum – some say a battleground – for a political discussion rather than the neutral exposition that you would normally expect in an encyclopedia. As a result, it has been necessary to block the article several times during this electoral season.

The hardcore Wikipedia authors say that it is so addictive that they have even invented a kind of "Wiki-dependence" test. When asked how she reconciled her normal life and her activities on the Wikipedia, one such Wikipedia addict replied, "I don't sleep very much."

But is it accurate?

By now you will inevitably have asked yourself the obvious question: "But is it accurate?" Consider the case of Mr. John Seigenthaler, which was related in an article in the *New York Times* on 4 December 2005. Mr. Seigenthaler was surprised to read in the Wikipedia one day that he "was thought to have been directly involved in the Kennedy assassinations of both John and his brother Bobby." "Nothing was ever proven," the Wikipedia article added ominously. Needless to say, Mr. Seigenthaler was dismayed, and even more so when he discover that this material – which he stresses is absolutely absurd – had been posted in the Wikipedia for a number of months. He eventually found out that the article had been written as a practical joke and was able to correct it.

In another recent example, it was discovered that members of the United States Congress were modifying their own articles in the Wikipedia to reflect positively on their records. It was reported in U.S. newspapers a few months ago that the staff of Congressman Marty Meehan wiped out references to his broken term-limits pledge as

well as information about his huge campaign war chest. This and other episodes made people start to wonder, "Can we trust anything we read in the Wikipedia"?

A number of experts, especially those in library science and similar disciplines who have to deal with such dilemmas in their own work, actually defend the Wikipedia. For example, J. Stephen Bolhafner of *The St. Louis Post Dispatch* observed, "The best defense of the Wikipedia, frankly, is to point out how much bad information is available from supposedly reliable sources." In fact, by being open to revision, the Wikipedia might even be *more* reliable, some argue, because mistakes or deliberate misinformation can at least in theory be corrected with they surface.

But others argue otherwise. The Italian columnist Carlo Donati wrote that you could call the Wikipedia an enormous newspaper, a gigantic magazine, a marvelous creation of a wonderful and democratic community of enthusiasts … anything but an encyclopedia. He went on to explain that a true encyclopedia, on the contrary, is an oligarchic and authoritarian work with two principal aims: on the one hand, to supply in a systematic way the cognition relative to all human knowledge, and on the other hand, to tell the truth. (In contrast, the U.K. newspaper *The Register* observed last Thursday, the Wikipedia is based on the lovely, utopian idea that we can all be experts, and that "you can vote for the truth.") A true encyclopedia, Donati notes, guarantees what it says and consigns it to future generations as though it were written in stone. But the Wikipedia guarantees nothing: everything is fluid; anything can change at any time.

As you might imagine, traditional encyclopedias such as Britannica are particularly incensed by the Wikipedia phenomenon. Former Britannica editor Robert McHenry has called Wikipedia the "Faith based encyclopedia." It's not hard to see why Britannica is angry. Consider an investigation carried out not long ago by "Nature" magazine, one of the most prestigious scientific journals in the world. They did a relative comparison of Wikipedia and Britannica's coverage of science. Perhaps not surprisingly, they found errors in the articles written in both of the encyclopedia. But more surprisingly, they did not find much difference in their relative accuracy. Among the 42 articles that they examined, Wikipedia averaged around four errors per article, against around three in the Britannica – certainly not a scandalous difference. Yet just this week, Britannica published a devastating response, in which they effectively dismantled the study and demanded a public apology from Nature. (You can find a copy of their response on the Britannica web site.)

Given this state of matters, it is well to remember that Wikipedia is a special kind of information source, and not to be treated as your *only* source of information, and certainly not your final source. But one thing is clear: it's here to stay and thrive, a fascinating example of a virtual Internet community.

The Open Source Community

The legendary fortunes accumulated by the likes of Bill Gates of Microsoft and Steve Jobs of Apple have given many the impression that the world of computers is populated only by greedy, voracious capitalist monsters. Certainly there is plenty of that, but those fortunes of mythic proportions have obscured an equally powerful current of astonishing altruism and selfless generosity in which some of the best and the brightest in the computing world have worked tirelessly to make the fruits of their efforts freely available to everybody. These people have created what is probably the oldest virtual community on the Internet: the open source community.

One of the most remarkable of these people is Richard Stallman, a brilliant and charismatic veteran of M.I.T. who has tirelessly campaigned for many years now for a concept in which he fiercely believes: that software should be free. Given the commercial importance of software in the world today, this is an extraordinary point of view. But even more extraordinary is how many software developers share that point of view. Stallman invented a new form of copyright license, which he called **copyleft**. Why "copyleft"? Because in a sense it is the opposite of copyright: instead of *restricting* the right of redistribution of information, it *guarantees* that right. Its official name is the General Public License.

Today, a vast community of software developers works on innumerable projects of all kinds, creating software that becomes freely available to anyone who wishes to use it, usually under some form of General Public License. Internet repositories such as SourceForge host these projects, which form their own little sub-communities complete with discussion forums and the like. Eric Raymond described the way in which open source software is developed in his book with the memorable title, *The Cathedral and the Bazaar*. In the book, he compares traditional software development to the building of a cathedral, under the supervision of a single, revered architect with total control and authority. In contrast, open source software development is like a noisy, chaotic bazaar where everybody contributes, comments, and changes.

Think about it: just like in the Wikipedia, *anybody* can contribute to the development of an open source computer program. Shouldn't that lead to completely unreliable, unpredictable software? Yet open source software is generally known to be *more* rather than less reliable than commercial software. The reason? The Broken Window Theory again. As soon as any problem pops up, a bug, a misunderstanding, somebody steps in quickly and fixes it. In this way, open source software somehow tends to end up being even more robust than its commercial counterpart. (I should note in fairness that the reality of open source project development is not quite that simple. Although the basic "Wiki" principle does indeed apply, in most such projects there usually is somebody in charge of checking and filtering contributions in order to have at least some control over what is happening.)

Probably the most famous example in the realm of open source software is the Linux operating system, named after its developer Linus Torvalds, and modeled after the venerable Unix system. Building upon its success, a host of applications have been developed by programmers around the world, often in competition with their expensive commercial counterparts, particularly those of Microsoft. For example, OpenOffice is an open source alternative to Microsoft's popular Office suite of applications.

No one has ever been able to explain satisfactorily the psychology of the open source community – why millions of talented, passionate, motivated programmers around the world would rally around such a seemingly naïve and idealistic concept. Yet they have, and continue to do so in growing numbers.

The orkut Community

As a final example of a contemporary virtual Internet community, I'd like to talk about a community that stands in stark contrast to the two I have just presented. Unlike the open, inclusive Wikipedia and open source communities, this is a closed, invitation-only community, known as **orkut**.

The first question that comes to mind is "What is an orkut"? A technical gadget? An acronym? A planet in a science fiction novel? It is none of these. It is a person – a Turkish software engineer named Orkut Büyükkökten. He attended Stanford University, one of the most prestigious universities in the world for computer science. He eventually went to work for Google, the company that now dominates the world of Web searching. It was there that he developed orkut.

Orkut is what is known as a *social network*: its declared purpose is for people to meet each other. As I mentioned earlier, entry into orkut is by invitation only, from somebody who is already a member. I had that opportunity a few months ago. As many in the Anglo-Italian Club of Viareggio know, I have an interest in Brazilian popular music. While following the Usenet special interest group in Brazilian music, I discovered that one of the other participants was a member of orkut and that there is a sub-community within orkut dedicated to the music of the great Bossa Nova musician João Gilberto. That participant was kind enough to issue an invitation to me and I had the opportunity to enter orkut.

The first thing that happened was that I was presented with a long, articulated questionnaire on just about everything about me: my interests in sports, music, cinema, TV shows, and the like. But not only such run-of-the-mill information was asked of me. I was asked about my hair color; my eye color; any body art I might have (pierced navel, pierced ears, pierced tongue, "strategically placed tattoo" …), and my "turn-ons."

To help me identify my "turn-ons," I was presented with a checklist. This included: assertiveness, candlelight, erotica, intelligence, public displays of affection, sarcasm, tattoos, thunderstorms, body piercing, dancing, flirting, long hair, power, thrills, wealth, and last but not least, skinny dipping. I was also asked to describe my perfect first date.

In addition, I was asked to indicate my sexual orientation, where many more than two possibilities were offered, including a few I'd never heard of. I was invited also to describe my political orientation, whereby an equally bewildering number of possibilities were offered.

Finally, I was allowed to enter orkut, where I found myself on my "home page," with my profile available for other members to see. I was also notified how many "friends" I currently had in orkut: exactly one, the same person who had invited me. As of three days ago, I still had this one and only "friend" – it seemed that other orkut-ians were not beating a path to my door. But suddenly, late in the evening, I received a message from orkut informing me that two people had added me as a "friend." I was surprised and curious, and entered orkut to find out who my new "friends" were. They turned out to be two students from a class I teach at the University of Pisa.

My two new student "friends" were from India, but what really stands out in orkut is the huge number of members from *Brazil*. Initially I thought it might be related (although I couldn't imagine how) to my entry into a Brazilian music group. But it wasn't, of course – in fact, nobody knows why fully seventy-percent of orkut members are from Brazil. The phenomenon has been dubbed "the Brazilian invasion" by those studying orkut.

The number of different sub-communities within orkut is impressive, ranging from fashion to humor to political. And therein lies the problem. Being an invitation-only operation, it is easy to create secretive, radical sub-communities without being noticed.

Communities have been formed around neo-nazi sympathies, white supremacy, racism, and pedophilia. Although they are being continually closed down, they sprout up again just as fast – or faster – simply because it's too hard to police them all.

The Social Networking Phenomenon

Orkut is only one of many social networks. A couple of the better-known networks are Friendster and MySpace. There are also social networks specialized in various areas such as business contacts (e.g. Linked-In). Many of them are now among the busiest sites in the entire Internet, and the rise of these communities on the Internet has been a bonanza for scholars studying the phenomenon of how societies form and operate.

But even before the Internet, social networking was studied. One of the most famous experiments of this type was carried out in 1967 by the American psychologist Stanley Milgram. He was testing the so-called "small world hypothesis," which claims that the chain of acquaintances between any two people is surprisingly short. He took random pairs of people in the United States and discovered that on average, that chain had only six links. That is, all of us are related by *six degrees of separation* – a phrase that became famous when a film was made with that title. In fact, one of the very first Internet social networks took its name from that phrase, too. Its name was (it is now defunct): `www.sixdegrees.com`.

At this point I can't refrain from recounting another "small world" anecdote concerning what is known as the Erdős Number. Paul Erdős was a Hungarian mathematician who was legendary for his brilliance, prolific output, and eccentricity. He cared nothing for worldly goods and didn't even own a house. Rather, he would invariably show up on the doorstep of a colleague and stay for a few days, writing a couple of mathematical papers with him before leaving to visit another colleague. He was beloved by all and it became a badge of honor for a mathematician to be able to say that he had collaborated with Erdős. Given the nature of this talk, it is appropriate that I let the entry on Paul Erdős in the Wikipedia tell you the rest of the story:

> Because of his prolific output, friends created the Erdős number as a humorous tribute; Erdős alone was assigned the Erdős number of 0 (for being himself), while his immediate collaborators could claim an Erdős number of 1, their collaborators received Erdős number of 2, and so on. Some have estimated that 90% of the world's active mathematicians have an Erdős number smaller than 10 (not surprising in the light of the *small world* phenomenon). It is jokingly said that Baseball Hall of Famer Hank Aaron has an Erdős number of 1 because they both autographed the same baseball when Emory University awarded them honorary degrees on the same day.

Another famous rule in social networking is called the "Rule of 150," which asserts that the optimal size of a social network is about 150 members. There are a few theories about why this is the case. One theory says that it's that way because human beings can't recognize and interact with more than about 150 individuals. It is partly based on anthropological studies of the optimal sizes of villages around the world. But another theory says it's more about keeping track of people who don't pull their own weight – that is, in larger communities it becomes harder to recognize when somebody is not contributing his fair share.

A large number of measures and indicators have been developed for studying social networks, with impressive-sounding names: radiality, density, cohesion, clustering

coefficient, eigenvector centrality (a measure of the importance of a node in a network), constraint, contagion, and so forth.

But I am interested in a much simpler question, at the level of human relationships. This week's edition of TIME magazine deals with exactly that subject. Its cover story is entitled "The Multitasking Generation: Are Kids Too Wired For Their Own Good?" and describes the immersion of young people in social networks such as MySpace and Facebook. What will be the implications of the rise of such Internet virtual communities on the way that people live? Will people (young and old) leave their lives in the "real" world and live only in cyberspace? Or will they lead better, more enriched lives as a result? Nobody really knows, but right now a veritable army of sociologists around the world is scrambling to find out the answers to those questions.

Resources

Not surprisingly, much of this talk was researched on the Wikipedia itself. To see the world's largest encyclopedia go to `www.wikipedia.org`.

Go to `www.sourceforge.net` to see the open source community in action.

The Broken Window Theory was first described here: James Q. Wilson and George L. Kelling. "Broken Windows". *The Atlantic Monthly*; March 1982; Volume 249, No. 3; pages 29-38.