SELF-GOVERNING ONLINE COMMUNITIES IN WEB 2.0: PRIVACY, ANONYMITY AND ACCOUNTABILITY IN WIKIPEDIA

Shun-Ling Chen

Thank you for being here, and many thanks for the Journal for organizing this Symposium. I'm going to talk about self-governance of online communities in the era of Web 2.0, and I will use Wikipedia as an example.

First I'll give you an overview of the presentation. I will go through some basic information on Wikipedia and explain how it works. I will mainly discuss how the Wikipedia community tries to find a balance between some potentially conflicting principles. For example, how it strives to be a credible encyclopedia project while allowing contributors to remain anonymous. I will offer some incidents which had lead to criticism of Wikipedia and explain how the community responded to criticisms. I will also discuss the role of the Wikimedia Foundation as an institutional buffer to help the community to negotiate for more room and more time for their experiment between them and the rest of society. I will also address briefly how the community's self-government structure deals with copyright violation in large-scale collaborative projects.

So let me first give some basic information about Wikipedia. It is a free and online encyclopedia, and it's based on a wiki. Twoconcepts are particularly important here, one is free and one is Wiki. Wikipedia is "free" in the sense of free software- although each contributor claims copyright of each contribution, by adopting a free license, contributors guarantee every user the freedom to copy, to distribute, to modify the work. And in Wikipedia, the free license is also a "copyleft" one, meaning it requires all the modified works to be released under exactly the same license. This provides the legal framework for the community's online collaboration.

The second is "wiki". A Wiki is a set of editable web pages that are heavily linked to one another, and this provides the technical

platform for the community's online collaboration. Wikipedia is based on MediaWiki, which is developed by the community based on a wiki software that was also released as free software.

Wikipedia started in 2001. Initially it was started as a feeder project of Nupedia, which is another Internet free encyclopedia project started by Jimmy Wales in 2000. Nupedia followed a more traditional process of making reference works, where its articles were written by experts and were vetted by a rigorous editing process. But Nupedia only produced twenty-five articles by the end of the first year and it didn't scale. So wiki was introduced to the project as a way to produce content which would be eventually introduced into Nupedia's production process. However, Wikipedia soon took off and began to have a life of its own, generating twenty-thousand articles within the first year. Unlike Nupedia, whose administrators controlled the content of the web page with a more centralized structure, Wikipedia operates an interactive platform in which users are invited to participate and can easily alter the content without endorsement of the website administrators. This interactivity, user generated content and the decentralized collaboration are key features of Web 2.0.

I'll give you an anecdote. At the party of Wikimania 2006, which is an annual meeting, there was a mock pitch for already outdated Web 1.0 business ideas, judged by real venture capitalists from the dot com era. When Jimmy Wales, the founder of the Wikipedia Project, was urged to the stage by the crowd he gave a single word proposal: Nupedia. He stepped down immediately in the expected cheers of the crowd, which consisted of people who had been contributing and running the Wikipedia.

Wikipedia is now hosted on the servers of the Wikimedia Foundation, which is a non-profit organization founded in 2003 after Wikipedia had already gained some fame. Wikipedia has become very popular. It has now three million articles in the English version alone, and it has probably two hundred plus projects in different languages. And it's the sixth most visited website in the world; it only comes after major search engines such as Yahoo!, Google, Bing, as well as YouTube and Facebook.

Although some studies have shown that the quality of Wikipedia articles is actually quite good, because of its open structure, its quality can be somewhat uneven. Some articles are better than others. Nevertheless, it has been recognized as a

major public source of information. For example, National Institute of Health has been working with the Wikipedia community to train its staff to provide up-to-date health-related information on Wikipedia.

Many people are often surprised when they hear that a website as big and as popular as Wikipedia is run by volunteers. Even the Wikimedia Foundation whose servers host the project has only thirty employees now and had only two to five employees until 2007. The Foundation has only one general counsel who collaborates with outside counsels in some cases, but mostly with volunteers who enforce community policies and rules. Occasionally, he has legal interns, like me, who was a legal intern in 2008 in the summer.

In this fiscal year, the total budget of the Foundation is $9.4 million, and the budget for the legal department alone is $130,000. This includes costs not just for litigation, but for various kinds of registrations like trademark registration and defense. Also, because it's a non-profit organization, it also needs to do a state-by-state registration for its fund-raising status. Without the self-organization among community members and somewhat effective policy enforcement within the community, the Foundation is not likely to operate with such a low legal budget. In this talk, I would like to explain why Wikipedia can be seen as a prime example of community self-governance in the era of Web 2.0.

Before I go on, just a little survey of this room. Can you show your hand if you have consulted Wikipedia? Okay. Can you show your hand if you have edited Wikipedia? Oh, a few of you. It's worth the time to explain a little bit in detail how it actually works.

So go into any Wikipedia page you will first see the main article, but what I am showing here is not a main article page. Every article has a history page, and it has a log of all the edits that has been done to this particular page. And you can compare any two versions. This is what I'm doing here. Compare the two versions and you can see on this slide the differences between the two versions are highlighted. You can see the anonymous editor who was connecting to the internet from the displayed IP address.[*] It was a disruptive edit, and it was almost immediately

* *See* Chaco Culture National Historical Park, WIKIPEDIA, *available at* http://en. wikipedia.org/w/index.php?title=Chaco_Culture_National_Historical_Park&acti on=historysubmit&diff=322477265&oldid=322477248.

reverted by this registered user Trevor MacInnis.

Note that "Trevor" doesn't have to be his real name, it can just be a pseudo name. And when an edit is made by a user account, that user's IP address is not shown. As you know, an IP address may show a physical location of a user to an extent, a logged-in user actually enjoys a higher level of privacy on Wikipedia.

This is a page from a featured article, so on the top you can see there's a star which means the article is voted by the community to have met its highest standard of content. You can participate in a discussion. You can also edit the page. And if you want to edit, you don't really need to register. But if you do want to have a username you don't have to provide your real name—you don't even have to provide an e-mail address. Users' privacy is highly respected and considered as essential for free speech on Wikipedia. With an open platform like this, the Wikipedia community has been developing a mechanism for quality control since the project has started. This mechanism is reflected in the community's rules and policies, and is enforced by trusted members of the community who are elected by the community after a long process of scrutiny.

The community rules have evolved with the growth of the project and have been adjusted to reflect the community's self-needs to improve the quality-control mechanism or to respond to external criticisms. It has a set of content policies, a few of them are considered to be essential: "No original research," Wikipedia only seeks to provide a summary of existing scholarship; and "Verifiability", it requires editors to provide sources so that other

readers or editors may verify the material independently; and "Neutral point of view"—note that this is not to find the one superior and unbiased perspective of a topic, but to present fairly all significant viewpoints.

Rules are also developed to enforce these content policies. When a topic of an article appears to be controversial and leads to "edit war" or vandalism, administrators may "protect" the page, meaning they can temporarily prevent unregistered, or sometimes even registered, users from editing. When a user account has been making disruptive edits, administrators may decide to block that account. And in cases of mass vandalism, a range of IP addresses from which vandalism edits were being made can also be "range blocked". Users can add articles of their interests to their watch list to monitor changes on that page, so that when disruptive edits were made, they can respond immediately.

There are also robotic programs known as "bots" to automatically patrol and identify vandalism edits and revert articles to their previous versions. I will show you another example, which is the same article as in the previous slide. We saw Trevor reverted the disruptive edits to the previous version of "Cluebot". Cluebot is actually a robotic program, not a human. And it also reverted almost immediately a previous disruptive edit. So in those examples, vandalism edits were reverted immediately by human or bot "vandal fighters," as they call it. Sometimes it takes longer to detect a vandalism edit, especially when it is made to look like a legitimate edit.

The Wikipedia community has a team of volunteers called the OTRS Team, and they handle inquires, complaints and comment e-mails from the public. For example, negative criticisms of a living person or companies are likely to lead to complaints, although in some cases these negative criticisms are not without sources. We will look into one example later. The OTRS Team either responds to these e-mails directly or forwards questions and criticisms to other trusted users to look into the issue and to respond to the person who made the complaints.

As I mentioned earlier, users' privacy is highly respected on Wikipedia, and this is mainly for two reasons: users' freedom of speech and personal security. It is especially so in controversial topics and for users from authoritative regimes or highly divided societies. When we think of anonymous editors sometimes we think of the "IP editors", or those who edit without user accounts.

But on Wikipedia, as I explained earlier, a registered user's IP address is recorded but not made public. A registered user with a pseudonym may have a higher level of privacy than IP editors. A Wikipedian once said, by contributing to Wikipedia you not only give your time but also your privacy.

As each edit that is logged may reveal some private information of the editor, determined data-miners may piece together these small pieces of information. Sometimes they may be able to identify and "out" an anonymous editor. There are a few examples. You might have heard of the Seigenthaler incident. John Seigenthaler, he's a writer, a journalist and former editor-in-chief of *USA Today*. An anonymous editor inserted false information about Seigenthaler in his biography page, claiming that he was involved in the assassination of John Kennedy and Robert Kennedy. The information was unsourced and would remain in the article for months before Seigenthaler became aware of the false claim and made complaints.

Wikipedia's model is somewhat borrowed from the free software model, and it has this Linus's Law that states "Given enough eyeballs, all bugs are shallow"—if there are enough people looking at it, all bugs are shallow and would be soon fixed. If a page gets fewer visits, although it may receive fewer vandalism edits, any vandalism edits that are already there are less likely to be detected and corrected in time, so as in the case of Seigenthaler's biography. After Seigenthaler's complaint was made into news, Daniel Brandt, a privacy advocate and Wikipedia critic, looked up the IP address and identified that it belonged to a small courier company in Nashville, Tennessee. The prankster confessed within hours and temporarily lost his job.

Another user Daniel Brandt identified was a user whose pseudonym is Newyorkbrad. Newyorkbrad has been active since 2005 and served as an administrator as well as a member of the "Arbitration Committee," which is the highest dispute resolution body in the English Wikipedia. Daniel Brandt usually targets anonymous editors who make vandalism edits, but he sometimes also identifies high profile editors, like Newyorkbrad, and threatens to reveal their real life identities in exchange for policy changes on Wikipedia.

Newyorkbrad is a practicing lawyer in Manhattan in real life, and he had never revealed his real name on Wikipedia before he was "outed" by Daniel Brandt. In 2008, Brandt obtained

2010]   SELF GOVERNING ONLINE COMMUNITIES   427

Newyorkbrad's IP address after Newyorkbrad had posted on a web forum, and then he identified the IP address belonged to a New York law firm. He then matched Newyorkbrad's photo, which was taken at the local Wikipedian community gathering, to the law firm's website and found out his real name. Upon this outing, Newyorkbrad quit Wikipedia and his active position and returned only after four months.

Another example is Wikiscanner. It is a tool combining two kinds of information: a list of IP addresses that have been used to edit Wikipedia and also a database that shows which IP addresses belong to which companies. Wikiscanner showed that multinational companies and governments have been using Wikipedia for their public relations purposes. However, as a registered user's IP address is not shown to the public, Wikiscanner is only able to identify the employers of IP editors, those people who edit without an account.

For those registered users, all the edits done from an account are logged and can be easily retrieved. One can find more personal information about a registered user by tracing his editing history and editing pattern. For example, a user's first edits were usually made in the field of his profession or related to his home town. A user may also reveal his religious beliefs. For example, an orthodox Jew who strictly observes the Sabbath could reflect his religious practice by showing a gap on the day of the week in his edit history. And if he travels from one time zone to the other and continues to edit Wikipedia during his travel, the edit history will reveal that he has been traveling. My roommate, for example, may reveal that he's a football fan because he sits loyally in front of TV every Sunday and Monday evenings.

Although anonymity may be abused, and there have been many proposals to prohibit IP edits or anonymous edits, so far none of these proposals have succeeded.

There are also other Wiki-based free encyclopedia projects which require editors to register with their real names, such as a project called Citizendium or a German project called Wikiweise, but neither of them manage to scale.

Anonymity has been the major source of complaints Wikipedia has received. When facing external criticisms, Wikipedia sometimes responds by adjusting its policies with enforcement mechanisms, and sometimes Wikipedia makes no adjustments but just defend its model by clarifying its methods.

I will give you a few examples. It's been complained that vandalism edits are often made by anonymous editors, and in response, Wikipedia enforces better content control. And since articles about contemporary celebrities are often targets of vandalism, Wikipedia has a heightened content policy on biographies of living persons.

Some critics are concerned about whether anonymous editors may render Wikipedia untrustworthy simply because their identity may taint even their legitimate edits.

In a well-known incident, a high profile editor, whose pseudonym is EssJay, he claimed false academic credentials and used them to back his editorial judgments. In response to this, Wikipedia reiterated its support for using pseudonyms and its method that information shall be verifiable independent of the editor who actually edited it. But as EssJay has made many contributions—he actually had many good edits and was very active, even helped to solve the Seigenthaler incident—he was selected as a trusted member and has access to some nonpublic information on Wikipedia. To solve this problem, Wikipedia started to require trusted users with access to nonpublic information to identify themselves to the Wikimedia Foundation.

And with the release of Wikiscanner, which shows that some edits were clearly done by users who have a conflict of interest and may be using Wikipedia for their public relations purposes, Wikipedia in this case did not make any policy changes. It just said, oh, this is great, we are happy to have this tool here so that people will be aware of this kind of abuse.

The Wikipedia community also has other concerns aside from its external criticisms, and one of them is called "sock-puppetry." This term refers to the situation when the user has multiple accounts. Sockpuppets can cause several problems. When debating over the content of an article, sockpuppets can be used to forge consensus in the discussion in order to make a marginal view appear to be more significant than it actually is. Or when the community is electing officers or voting for policies, sockpuppets can be used to unfairly influence a community vote. And by using sockpuppets a banned user may easily come back to the project with different usernames.

To deal with this problem, server administrators with direct access to server files used to do "hand checks"—when they received reports they would hand check reported users' IP addresses to see if these user activities are problematic. An IP

checking tool called "CheckUser" was later developed, and the community assigned a new group of users who would have access to this tool. CheckUser can be used for three different things: it can be used to find out all the edits done from one IP address, the usernames that connect by an IP address, and all the IP addresses used by one user. Hence CheckUsers can get different kinds of information to triangulate, to find out, whether a username is really a sockpuppet. The tool can be used not only to detect sockpuppeting, but it can also be used to do IP range blocks in cases of mass vandalism.

We've seen that a user's activities on Wikipedia have privacy consequences. The community has developed policies to reflect their concerns for user's privacy, but at the same time deal with abuse of anonymity, which can cause harm to the project's credibility and its overall quality.

The Wikimedia Foundation also has a role in providing an environment that is friendly to Wikipedia's high expectations for privacy. Under the Communication Decency Act Section 230, as merely a conduit, not the speakers, of the content on Wikipedia, the Foundation is immune from legal liability for the content. But in civil/criminal procedures, plaintiffs may ask the Foundation to provide IP addresses of the editors that are involved in disputed articles.

The Foundation publishes its privacy policy, which was a reorganization of various privacy related policies developed by the community. The document details privacy-related consequences to all users' involvement on Wikipedia, including what kind of personal information may be revealed through what kind of activity, to whom, and if such information would be retained permanently or only for a limited period of time. Hence users can be aware of their privacy related consequences in their activities, and decide what to do and what not to do on the project.

The Foundation also has a data retention policy which states that its servers would keep as little users' private information as possible, so that in civil or criminal cases, such as libel cases, the Wikimedia Foundation would have very little user information to provide.

The Foundation also tries to limit its legal footprint, so that it would not have to respond to courts in jurisdictions that disallow anonymity or require it to retain data for a long time. For example, when it was reported that Korea would be changing its

law to require all Internet users to use real names, the Foundation's tech team and their general counsel agreed that it is safer to move the servers in South Korea to another jurisdiction, even if those servers were only doing miscellaneous things.

With the immunity status and the heightened privacy related practices, the Wikimedia Foundation is able to serve as an "institutional buffer" to negotiate more room and time for the community to experiment on its method. And this is not to make Wikipedia a lawless zone. In fact, Wikipedia's policies often have higher standards than legal rules. But the community may still have room for improvement in policies themselves, technical tools or enforcement mechanisms. Sometimes, even if the community is properly enforcing its policies, the content may still make some people upset. For example, there was a case called Video Professor, which involves a company whose business model has been criticized by consumer groups. Such criticism was reflected on their Wikipedia page, and it's actually properly sourced. But the company was very upset still and went after editors. In cases like this, giving the plaintiff easy access to users' private information may cause chilling effects to users, affecting their online free speech, and prevent the community from continuing its practices, that is, having a reference work project which is not only open to experts and academics, but to anyone who wants to participate and follow good practices defined by the community.

So far I have covered the major points of Wikipedia community's practices in finding a balance between privacy and quality. I just want to quickly also talk about another issue that involves a lot of community self-policing and show how self-governance in not only Wikipedia, but the broader Wikimedia community, is able to avoid many legal disputes. Although the English Wikipedia allows Wikimedia Foundation's staff to act on an Wikipedia article in extraordinary conditions, it is the Foundation's position that the office should only take off materials in very rare cases, and that it should almost never add anything to an article. Taking off materials that violate copyright law or violate the community's copyright policy is enforced by volunteers. Enforcers tend to have high standards when deciding whether to keep a file or not in borderline cases. Because of this kind of consistent and effective policing of copyright violations, the Foundation receives only two DMCA take-down notices a year—and imagine, it is a huge project.

So let's look at one example, this is "Dali on the Rocky Steps."[**]

This picture was taken in front of the Philadelphia Museum of Art. The steps are famous for being the in Rocky movie where Rocky was doing his training. The steps and the museum are the landmarks of the city. During the Dali Exhibition in 2005, the museum used photographer Philippe Halsman's portrait of Dali (which was made in 1958, so it's still copyrighted), to prepare a surreal entrance for the exhibition. The photo was taken by Max Buten who released the photo with a free license, and posted this to the Wikimedia Commons. Commons is Wikipedia's sister project for multimedia files, and it hosts many files that are used in Wikipedia articles.

Commons has an even higher standard of copyright than Wikipedia. Wikipedia sometimes allows copyright material to be used in Wikipedia articles under the Fair Use clause. The

[**] *See* "Dali on the Rocky Steps", WIKIPEDIA, *available at* http://upload.wiki media.org/wikipedia/en/a/a8/Dali_on_the_Rocky_Steps.jpg.

Commons wouldn't allow fair use and it would only allow materials released under a free license or materials that already entered the public domain. In this case the photo was released under a free license. But some feared that this Dali photo with the copyrighted Dali portrait can be still a copyright violation. They deleted the file from the Commons and said, "Well, this cannot be considered as *de minimis* and there is no freedom of panorama for architectures in the U.S." Although some wanted to undelete the file, and the general counsel of the Foundation actually lent his support to the undeletion side, ultimately the request has failed and the high copyright standard for only including free material on the Commons was upheld. (The image can still be found in Wikipedia under the "Fair Use" rationale.)

So just to conclude, I argue that a project like Wikipedia, which is a robust community with a common goal, is likely to protect itself and to present itself as a credible project by incorporating existing legal rules and other established norms into its self-governance structure. To be recognized as a responsible project, Wikipedia has to constantly improve the system and to be responsive to public criticisms. For a young project like Wikipedia, an institutional buffer such as Wikimedia Foundation prevents early intrusion from formal institutions, which could cause chilling effects and interferes with the community's self-governing practices and structure. An institutional buffer helps the community to negotiate more room and time to adjust its rules and enforcement mechanisms to show that it is capable to meet its goals with good practices.

Thank you.