

Participatory Journalism and Asia: From Web Logs to Wikipedia

Andrew Lih

*Journalism and Media Studies Centre
University of Hong Kong
alih@hku.hk*

Paper for the
13th Asian Media Information & Communications Centre Annual Conference:
ICT & Media Inputs & Development Outcomes
Impact of New & Old Media on Development in Asia
July 1-3, 2004

Abstract

With a wide range of media environments and press freedoms, Asia provides a rich setting for studying the emergence of *participatory journalism*. This paper describes the origins and evolution towards one-to-many and many-to-many online participatory journalism, by focusing on the evolution of web logs and the largest project to date - Wikipedia. This multilingual, Internet-based electronic encyclopedia has been created by thousands of ordinary Internet users. It has achieved this by presenting a low barrier to participation, allowing anyone to edit any page in the Web site. In just three years, it has gained a large following and is available in over 50 different languages. The paper examines the evolution of participatory journalism projects and the implications for the Asia region.

Table of contents

Abstract.....	1
Table of contents.....	2
Participatory Journalism	3
Personal publishing – First Wave	5
Origins.....	5
Statistics	6
Open Content and Wikis – The Second Wave.....	8
Introduction.....	9
Keeping it social and neutral.....	11
Wikipedia Operations	13
Working as Humans.....	16
Foreign Language Wikipedias	18
Japanese Wikipedia.....	20
Chinese Wikipedia	21
Conclusions.....	23
About the Author	25
Acknowledgements.....	25

Participatory Journalism

Over the last decade, the Web has become an essential tool for journalists and researchers. While information can be found quickly and easily using search tools such as Google or Yahoo, the problem is often not a lack of content, but rather the large volumes of stale and questionable information. Determining the accuracy of search results is a challenge for any Internet user. However, the advent of *participatory journalism* has provided a unique solution to this problem – it engages the news audience to participate in the process of rationalizing Web content, crafting the news and contributing knowledge into the *media ecology*.

In their work on the subject, *We Media, How audiences are shaping the future of news and information*, Shayne Bowman and Chris Willis define participatory journalism as¹:

The act of a citizen, or group of citizens, playing an active role in the process of collecting, reporting, analyzing and disseminating news and information. The intent of this participation is to provide independent, reliable, accurate, wide-ranging and relevant information that a democracy requires.

This tradition started with the USENET² relayed discussion boards of the early Internet, predating the World Wide Web. The USENET system allowed users from all over the world to engage in discussion topics using text messages, much like today's bulletin board systems. However, it also spawned a form of persistent content – the frequently asked questions or FAQs, which were collected and stored from among the ephemeral bulletin board messages. The FAQ concept is now well known in cyberspace, and is perhaps the earliest forms of participatory journalism on the Internet.

Web logs emerged in 1998, shortly after the rise of the Internet, and allowed greater numbers of people to create persistent Internet content. American writer Jorn Barger is widely considered to have coined the term *weblog* in December 1997 to describe web sites which provided a location for individuals to publish their own opinions and Internet links, typically used to put current events in context or to document observations in daily life³.

In places with restricted press environments, web logs have become an important “reality check” to the mainstream and state controlled media, allowing ordinary users to help craft the news and keep traditional media sources accountable. This is especially true in the Asia region, where free media environments are the exception rather than the rule. Variants of the web log concept in Korea (OhMyNews.com) and Malaysia (Malaysiakini.com) have played crucial roles in local politics and media ecologies. The web log Back to Iraq was created by an independent journalist to document the U.S.-Iraq war in March 2003, in contrast to the embedded American journalists. When readers donated nearly US \$15,000 dollars to send him to Iraq, he became the “Web's first fully reader-funded journalist-blogger⁴.”

Newest to the Internet ecology are *wiki* Web sites. Wikis allow users to immediately edit and directly change any page with one click of the mouse. Contrary to nearly all previous assumptions about having to maintain editorial and structural control over content in order for it to flourish, wikis have taken the other extreme – complete openness and changeability. It is this wiki technology that has produced the largest form of participatory journalism to date – Wikipedia.

Personal publishing – First Wave

Participatory journalism today is most widely realized through personal publishing on web logs, usually defined as “updated online journals, with reverse-chronological entries and numerous links.⁵” Many web logs, commonly called blogs, offer the chance for readers to leave feedback or to engage in a bulletin board conversation about individual web log entries, fostering a sense of community among readers.

Origins

Conceptually, web logs are not very different from personal publishing that occurred at the dawn of the Internet in the early 1990s. Web sites such as Cool Site of the Day, Suck.com and Drudge Report (the latter responsible for breaking the Monica Lewinsky scandal) were established daily dispatches during the dot-com era and had loyal readers around the world. However, because hosting a Web site and managing the content were involved tasks, publishing was not a popular activity for average Internet users. The breakthrough for the popularity of web logs was the development of content management systems (CMS) which provided publishing functions to ordinary users, without the need to learn Web site management, complex formatting or programming commands. One pioneer in the field was Dave Winer, whose company, UserLand, created the earliest tools for publishing Web sites through simple Web-based interfaces. In the 2000s, the wide availability of content management software (both commercial and open source) and free web log hosting sites supported rapid growth in this form of publishing⁶. The Moveable Type software package from Six Apart Ltd, San Mateo, California, is among the most popular and feature rich web log systems available for free non-commercial use. Web

log hosting from Blogspot, Blogger and Typepad have allowed individuals to start up their own free personal blogs within minutes, using a very simple interface.

Recently moblogs, or mobile logs, have been developed to capitalize on the 2.5G and 3G wireless data networks and camera phones, allowing users to post pictures directly to Web sites from their mobile handsets.

Statistics

A Pew Internet and American Life study performed March 12 to May 20, 2003 recently studied user contributions to Internet content. The polling project, *The Material People Contribute to the Online World* is one of the few detailed studies on the phenomenon of Internet users involved in content creation. Its study of Web sites, web cams, web logs, file sharing and newsgroup participation, found that 44% of adult American Internet users (roughly 53 million people) had contributed content online. This is a significant shift as individual users are becoming content creators, using text, pictures and even video⁷. For web logs, the study found 2% to 7% of Internet users published one, and only 10% of those update their web logs daily. These numbers are particular to American Internet users, as it is likely the numbers on a worldwide scale are lower.

However, the Asia region does have an array of web logs, ranging from expatriates living abroad, to locals documenting life in their country. The year 2003 marked the first Asia-region web log awards, and was run by expatriate web logger Phil Ingram in Hong Kong, of Flyingchair.net.

Top Web Logs in Asia, by Region (2003)

Region/Community	Web Logs
People's Republic of China	博客堂 - http://blog.joycode.com/ 2SIMPLE - http://www.xjyk.com/datou/ 南大百合 Blog - http://lilybbs.net/blogall
Hong Kong	Gweilo Diaries - http://www.gweilodiaries.com/ BWG - http://www.bigwhiteguy.com/ Simon World - http://simonworld.mu.nu/
Taiwan	3q2u - http://3q2u.com/ zonble's @rsenal - http://zonble.twbbs.org/ echeng.com - http://www.echeng.com/
Singapore	Xiaxue - http://www.xiaxue.blogspot.com/ Mr. Brown - http://www.mrbrown.com/ Wishful Thinking - http://angelicgal.blogspot.com/
Malaysian	sultanmuzaffar - http://sultanmuzaffar.blogspot.com/ gaharasiber MURING-WIEN cyberathaus - http://hitme.at/muringwienvotes Jongkang Jongket - http://jongkang.blogspot.com/
Thailand	Thailand Life - http://www.thailandlife.com/ pressurize.net - http://pressurize.net/ The Voices In My Head - http://starryseas.blogspot.com/
Indonesian	kamingjantan[dot]com - http://www.kamingjantan.com/blog.html nasgorkam - serendipityq - http://serendipityq.com/ Weblog Wannabe - http://www.wannabegirl.org/
Japanese	Kind of Crap - http://www.kindofcrap.com/ Sushicam - http://www.sushicam.com/ Tokyo Damage Report - http://www.harmful.org/homedespot/ADIARY.htm
Korean	The greatest Aspirin - http://blog.naver.com/kickthebaby Korea Life Blog - http://korealife.blogspot.com/ The Marmot's Hole - http://marmot.blogs.com/korea/
Philippines	It's True! It's True! - http://mistervader.blogspot.com/ Mind's Zen Garden - http://carlos.ricebowljournals.com/ CrazyPundit - http://crazypundit.typepad.com/
India	Turbanhead.com - http://www.turbanhead.com/mt ASHIFICATION - http://ashforyou.rediffblogs.com/ LimeIce - http://www.limeice.net/
Bangladesh	The 3rd world view - http://rezwanul.blogspot.com/
Iraqi	Baghdad Burning - http://riverbendblog.blogspot.com/ healing iraq - http://healingiraq.blogspot.com/ Salam Pax - http://dearraed.blogspot.com/
Iranian	Iran-Sare2008 Group - http://sare2008.persianblog.com/ Zahra HB - http://zahra-hb.com/ NORTHERN BOY - http://irannorth.persianblog.com/

Source: Asiablog Awards, Flyingchair.net, 2003⁸

In the China Internet community, even with close oversight by the central authorities, bulletin boards have played an important role in the media ecology by providing commentary on news events and government policies. Since 2002, it has also seen a significant rise in the number of web logs, with an estimated 10,000 active web loggers and 600 web logs⁹. Web logs are referred to in Chinese as 博客 (bo ke) which is phonetically similar to the word “blog” but also has a literal meaning of *rich* or *abundant traveler*. Among the most popular web log sites include BlogCN, BlogChina, Blogbus.com and CNBlog.org. In March 2004, there was a temporary halt to the more popular web log sites Blogbus and BlogCN by the central authorities, including blocks to all Typepad-related sites. The two China blog sites have since been revived, but with reports of content being “cleaned” for political correctness¹⁰.

Open Content and Wikis – The Second Wave

Personal publishing through web logs has allowed the one-to-many model of publishing to be extended to average Internet users, largely as a result of ubiquitous, easy to use publishing tools. Similarly, many-to-many collaborative projects have followed the same type of evolution, building on the wide availability of open source tools.

The open content movement draws upon the spirit of the open source software (OSS) movement, which created the increasingly popular Linux operating system and GNU software tools. The idea for OSS was to have software be free, in both senses of the word – free cost to distribute and free to be modified by anyone. The result today is that the Linux operating system has provided a challenge to the dominance of the commercial Microsoft Windows operating system in many sectors, especially in corporate and industrial environments. In fact, the three largest countries in East Asia

– China, Japan and Korea – have agreed to cooperate on open source development to provide an alternative to Microsoft’s products^{11,12}.

Open content draws upon the same themes – allowing free dissemination of the content and freedom by anyone to change the content. The concept of free dissemination of content is not new – on the Internet, Project Gutenberg¹³ is a well known effort to make public domain texts widely available free of charge. However, only recently have multiple authors been able to collaborate effectively on content in a shared online forum. Wikipedia provides the most successful and largest scale example to date.

Introduction

Wikipedia is an Internet-based, volunteer-contributed encyclopedia that has become a popular online reference (Figure 1) in just three years of existence. It has thousands of international contributors and is the largest example of an *open content* wiki. (The Hawaiian word for “quick,” WikiWiki, is the basis for the wiki name.) The goal of Wikipedia is to create an encyclopedia that can be shared and copied freely while encouraging people to change and improve the content. Each and every article has an “Edit this page” button, allowing anyone, even anonymous passersby, to add or delete any content on any page. What would surely seem to create chaos has actually produced increasingly respected content which has been evaluated and revised by the thousands of visitors to the site over time.

Wikipedia Traffic Rank vs. Other Encyclopedias (2003-2004)

Source: www.alexa.com

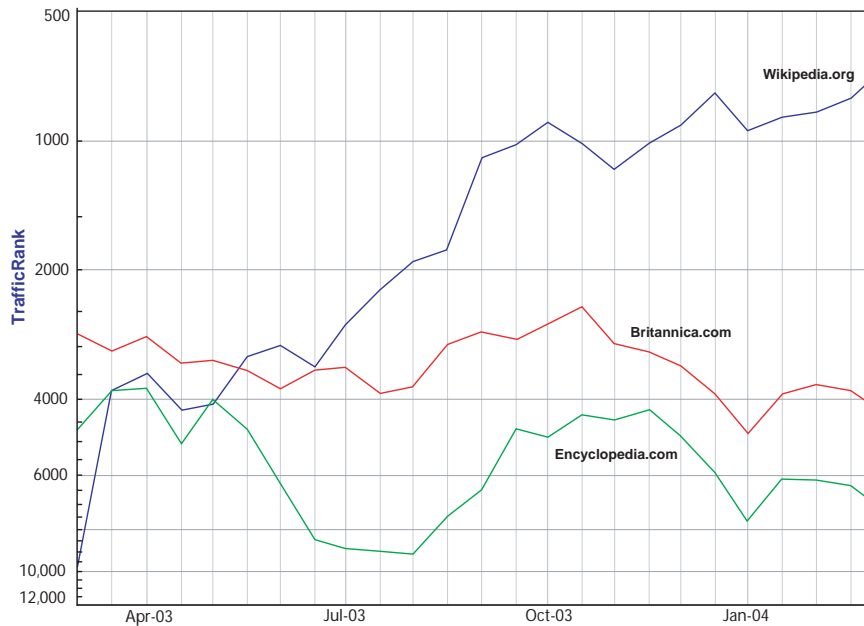


Figure 1 Alexa.com weekly Traffic Rank for major online encyclopedias

The project was started by Jimmy Wales, head of Internet startup Bomis.com, after his original concept for a strictly controlled, PhD-edited free encyclopedia ran out of money and resources, resulting in only a few hundred articles. At the time, the project was called Nupedia¹⁴ and in March 2000 had one full-time employee, Larry Sanger, who was editor-in-chief and co-coordinator for the project. Volunteers were solicited on the Internet, but there was a fairly complex working structure, including formal positions for writers, editors, peer reviewers, copyeditors and translators. Nupedia was built on the traditional structure of peer-reviewed academic publications, and according to their policy, "We wish editors to be true experts in their fields and (with few exceptions) possess Ph.D.'s."¹⁵ After the project failed to take off, Sanger ceased being a paid staff member and the project came towards a close.

Not wanting the already-created content to stagnate, Wales and Sanger put the content onto their Web site in the form of a wiki in January 2001, and invited visitors to edit or add to the collection. It became a popular site shortly after, after gaining

mentions on the influential Slashdot technical community. In the first year it found a loyal following, and generated over 20,000 English language articles¹⁶ and spawned over a dozen language translations, despite the fact that it did not fully support internationalization. After two years, it reached a milestone of 100,000 English articles and in February 2004 at the three year mark, it exceeded 200,000 articles in English and 500,000 articles in 50 languages (Figure 2). In February 2004, it was adding articles at a rate of roughly 2,000 articles a day across all the various languages.

Wikipedia Growth (English language) 2001-2004

Source: http://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia

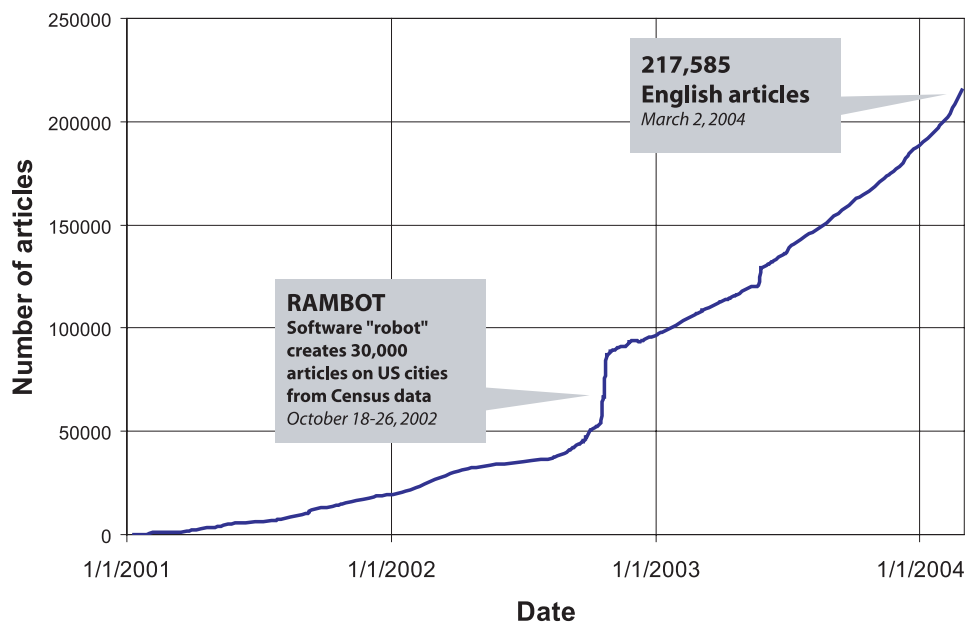


Figure 2 Wikipedia growth over its three year history

Keeping it social and neutral

The wiki concept is somewhat counterintuitive because the technical implementation itself provides no gate keeping function to ensure quality material is being contributed. Unlike typical creative efforts, no proof of identity or qualifications is needed to participate and a reputation tracking system is not used

within the community. Given the description of how a wiki works, visitors to Wikipedia are often surprised the site works at all.

What allows this completely open editing system to function? Because wikis provide the ability to track the status of articles, review individual changes, and discuss issues, they function as *social software*, acting to foster communication and collaboration with other users. A wiki also tracks and stores every version of an article edited, so no operation is ever permanently destructive. As a foil to malicious contributors, it takes much more effort to vandalize a page than to revert an article back to an acceptable version¹⁷. While it may take five or ten seconds to deface one article, it can be quickly undone by others with just one click of a button. This crucial asymmetry tips the balance in favor of productive and cooperative members of the wiki community, allowing quality content to emerge.

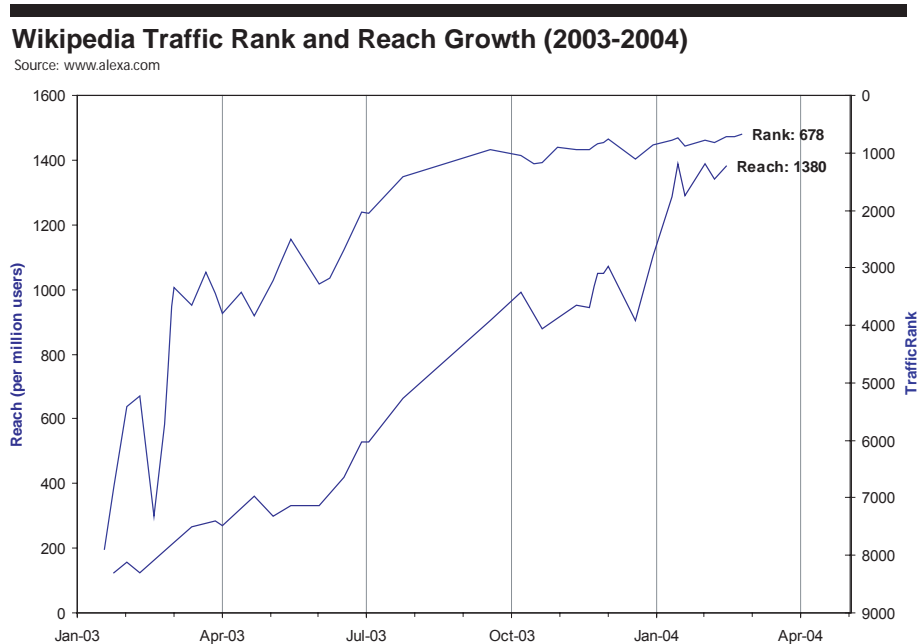


Figure 3 Alexa.com Traffic Rank and Reach (users per 1 million Internet users)

However, while the technical means for managing the information space and user interaction provides the mechanisms for the Wikipedia community to operate,

they are not enough on their own. Founder Wales created an editorial policy of maintaining a neutral point of view (NPOV) as the guiding principle. “NPOV is an absolute non-negotiable requirement of everything that we do,” he says, and according to Wikipedia’s guidelines, “The neutral point of view attempts to present ideas and facts in such a fashion that both supporters and opponents can agree.” Guided by this policy, the grassroots project has implicitly adopted the same types of operational policies facing modern news operations – sticking to the facts, attributing sources and maintaining balance. Some of the decisions are strikingly similar to other professional news organizations. For example, the Wikipedia community’s tendency to avoid the use of the word “terrorist” is similar to a policy adopted by the Reuters news agency. Whether to reveal the name of the woman involved in the Kobe Bryant rape trial has been a subject of intense debate within the community.

Wikipedia Operations

Wikipedia’s operations are fairly simple because there is only one basic object within the system – the wiki page. Editing a page is much like editing an ordinary text document, with some special formatting “markup” (Figure 5) to make bold, italic, headings, hyperlinks and lists.



Figure 4 Basic Wikipedia article

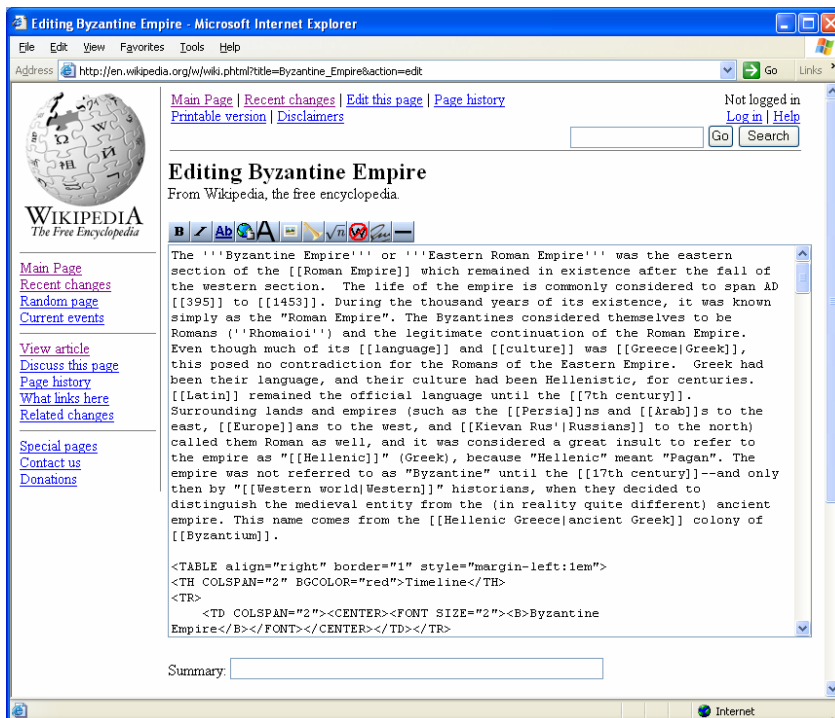


Figure 5 Editing a basic Wikipedia article, with special markup commands

The metadata contained in the system about a page is limited to only the most basic historical information about the content. Associated with each page is:

- **Edit history.** This is a chronological log of every change that has been made to the page since its inception (Figure 6). Users can select any two versions, and Wikipedia will use the “diff” utility to highlight the syntactic differences between any two versions (Figure 7). This is especially useful when trying to determine what information has been changed by others.

Giant Panda

Revision history

View (previous 100) (next 100) ([20](#) | [50](#) | [100](#) | [250](#) | [500](#)).

Legend: (cur) = difference with current version, (last) = difference with preceding version, M = minor edit

- (cur) (last) .. m [21:21, 25 Feb 2004](#) .. [Robbot](#) (*Andre Engels - robot Adding:fr*)
- (cur) (last) .. m [03:49, 25 Feb 2004](#) .. [Hadal](#) (+*Status*)
- (cur) (last) .. m [15:08, 11 Feb 2004](#) .. [Tannin](#)
- (cur) (last) .. m [15:01, 11 Feb 2004](#) .. [Tannin](#) (*trim surplus from caption*)
- (cur) (last) .. m [01:23, 15 Nov 2003](#) .. [Baldhur](#) (*de:*)
- (cur) (last) .. [22:46, 9 Aug 2003](#) .. [Munford](#)
- (cur) (last) .. [22:20, 14 Jul 2003](#) .. [Vicki Rosenzweig](#) (*taxonomy still under debate; size of wild population; some copyediting*)
- (cur) (last) .. [20:57, 14 Jul 2003](#) .. [218.221.176.60](#) (+*ja*)
- (cur) (last) .. m [01:25, 20 Jun 2003](#) .. [Timwi](#)
- (cur) (last) .. [23:15, 19 Jun 2003](#) .. [Digi](#) (*link to Polish Wiki article*)
- (cur) (last) .. m [15:49, 19 Jun 2003](#) .. [TeunSpaans](#) (+*nl:*)
- (cur) (last) .. m [15:47, 19 Jun 2003](#) .. [TeunSpaans](#) (+*nl:*)
- (cur) (last) .. m [02:53, 19 Jun 2003](#) .. [Timwi](#) (*prevent ugliness in Opera 6.05 :)*)
- (cur) (last) .. m [23:37, 14 Jun 2003](#) .. [Rotem Dan](#) (*caption for second photo*)
- (cur) (last) .. m [23:35, 14 Jun 2003](#) .. [Rotem Dan](#) (*Putting the second photo lower on the page*)
- (cur) (last) .. m [23:08, 14 Jun 2003](#) .. [Jimbleak](#)

Figure 6 Edit history of the Giant Panda article

Giant Panda

(Difference between revisions)

Revision as of 03:50, 18 May 2003 Tannin (Talk contribs)	Revision as of 11:56, 20 May 2003 Menchi (Talk contribs)
--	--

<p>Line 20:</p> <pre><tr><th align="center">" Ailuropoda melanoleuca"</th></tr> </table></pre> <div style="background-color: yellow; padding: 5px; margin: 5px 0;"> <p>The "Giant Panda" (&#29066;&#35987; [[pinyin]] xiong2mao1), "Ailuropoda melanoleuca" ("black-and-white cat-foot") is a [[mammal]] in the [[bear]] family [[Ursidae]].</p> </div> <p>Despite being taxonomically a [[carnivore]], its diet is overwhelmingly vegetarian. In fact, it lives almost entirely on [[bamboo]], although, like most animals, pandas have been known to eat eggs, and they consume some [[insect]]s along with their bamboo diet.</p>	<p>Line 20:</p> <pre><tr><th align="center">" Ailuropoda melanoleuca"</th></tr> </table></pre> <div style="background-color: #e0ffe0; padding: 5px; margin: 5px 0;"> <p>The "Giant Panda" (&#29066;&#35987; [[pinyin]] xiong2mao1), "Ailuropoda melanoleuca" ("black-and-white cat-foot") is a central [[Chinese China]] [[mammal]] in the [[bear]] family [[Ursidae]]. Its Chinese name means "[[bear]]-[[cat]]," and is one of also be read in reverse to mean the same thing. Its [[Western]] name originates in the [[Himalaya]] but its meaning is uncertain.</p> </div> <p>Despite being taxonomically a [[carnivore]], its diet is overwhelmingly vegetarian. In fact, it lives almost entirely on [[bamboo]], although, like most animals, pandas have been known to eat eggs, and they consume some [[insect]]s along with their bamboo diet.</p>
---	--

Figure 7 A "diff" between versions, showing changes in red

- **Talk page.** Because article pages are meant to contain finished copy, Talk pages for each page are used for discussing editorial issues.

Other important tools for users include:

- **Recent changes.** A record of all changes is written serially to the Recent changes log, and people can inspect this to find out the latest modifications to articles.
- **Watchlists.** Users can place specific pages to “watch” on their list, like Web bookmarks. The status of these pages can be seen from one single Watchlist page particular to each user.

Working as Humans

One key to the operation of Wikipedia, and wikis in general, is its primary role as social software, allowing human contributors to work in the most simple and straightforward manner as possible. As such, all the orientation is towards tracking, editing and formatting documents, and communicating with others about the content. By design, there is very little *metadata*¹⁸ or associated “behind the scenes” information about the articles or organization of the content. So while editors within the system can make hyperlinks among all the cities in Europe, or sports of the Olympics, or the dynasties of China, no additional information in the database is kept relating any two articles. All information about an article and linkages to other articles and are completely revealed and managed in the human-readable articles themselves, and not hidden among database elements. The simplicity of this model has been a main reason for the success of the project.

Therefore, wikis can be viewed as systems that provide the most generic and basic tools necessary for humans to compose and build a information frameworks.

These basic features include:

- **Simple text markup and formatting.** Most Web page design uses HTML, or Hypertext Markup Language, which was originally meant to be a computer-oriented code. Its complexity makes it prone to typos and misformatting. MediaWiki markup is friendlier and uses a simple set of commands, making text readable and less intimidating. For example, creating a link can be done by simply double bracketing a word, such as [[Genghis Khan]].
- **Structure by convention, not enforced by the software.** Formats of articles are created by human contributors. They are not dictated by forms, templates and software, and can be changed at any time.
- **Soft security, ubiquitous access.** All users of the site have read and write access to the information, including anonymous and first time users. Special “administrators” have marginally more capabilities and no more editorial authority than other users.
- **Transparency and edit history.** Every action in the Wikipedia environment is logged to the database and can be inspected by any user. This allows users to easily track activity and changes to articles of interest. That no action is clandestine adds to the openness and accountability within the community.

Wikipedia has intentionally chosen not to impose strict process, consistency checking or automation of tasks, mainly because it would dictate policies and procedures within the system. Instead, wikis emphasize ease of use and providing a human-friendly environment. Writer Clay Shirky has observed:

Process is an embedded reaction to prior stupidity... Wikis dispense with all that -- all of it. A wiki in the hands of a healthy community works. A wiki in the hands of an indifferent community fails. The software makes no attempt to add 'process' in order to keep people from doing stupid things. Instead, it provides *more* flexibility, a crazy amount of flexibility, and intoxicating amount of flexibility...¹⁹

Wikipedia works largely by consensus, with users discussing changes, rallying users for certain causes, or “being bold” and making edits unilaterally. When consensus is not clear, other means such as polling or trial periods are used. In the worst cases, edit wars with constant additions and deletions may require intervention by other community members to help mediate and arbitrate issues.

Foreign Language Wikipedias

When Wikipedia started in January 2001, it was largely an English language project. As word of the project spread, Internet users were interested in creating translations into other languages and by May 2001, twelve different language editions had started, including French, German, Spanish and a very basic Chinese site. More languages were added over time, and the software was extended to support the needs of right-to-left text (for Hebrew and Arabic) and both types of Chinese characters (simplified and traditional). The project grew to use Unicode UTF-8 encoding²⁰, which has been applied to many, but not all, of the sites.

Ten Largest Wikipedias by Language (March 1, 2004)

Rank	Language	Articles	Growth ¹	Share
1.	English	217016	+17%	41.770%
2.	German (Deutsch)	55783	+32%	10.737%
3.	Japanese (日本語)	33415	+51%	6.431%
4.	French (Français)	28213	+29%	5.430%
5.	Polish (Polska)	24861	+36%	4.785%
6.	Swedish (Svenska)	22348	+22%	4.301%
7.	Dutch (Nederlands)	21688	+25%	4.174%
8.	Spanish (Castellano)	18722	+35%	3.603%
9.	Danish (Dansk)	16244	+13%	3.127%
10.	Esperanto	11235	+10%	2.162%
-	All Wikipedias	519556	+30%	100.000%

Source: Wikipedia:Multilingual Statistics²¹

¹Growth measured from December 19, 2003 to March 1, 2004

Table 1 Largest Wikipedias by Language

The initial software used for the project was a simple “script” that was designed for fairly small sites. As part of its growth, the Wikipedia project started developing its own special software system for its operation, using a combination of open source programming tools and customized programs. The system, called MediaWiki, currently uses a combination of popular programs from the open source community, the so-called **LAMP** tools – Linux operating system, Apache web server, MySQL database and PHP scripting language.

The software is currently referred to as MediaWiki Phase III and is managed using the open source community site SourceForge.net, allowing anyone to inspect and suggest changes to the software. While the MediaWiki software is open source and freely available, it is fairly complex system to set up compared to most wiki implementations. Its features are also quite specialized to the needs of Wikipedia, therefore other sites that have employed the MediaWiki software tend to use it for “encyclopedic” oriented projects such as Wikitravel or Disinfopedia.

Japanese Wikipedia

The Japanese language Wikipedia (<http://ja.wikipedia.org>) is the largest of all the non-Western language Wikipedias, and is an impressive third largest overall with 33,000 articles (March 1), trailing only English and German. It was one of the original twelve foreign language Wikipedias started in May 2001, but because the software at the time did not handle Chinese-Japanese-Korean (CJK) languages, the pages were written in Romaji, which uses Latin characters to spell out the “phoneticized” Japanese words.

They have roughly two to three dozen core contributors (those with more than 100 edits per month)²². Participation in Japanese Wikipedia increased after it was publicized in the Wired News Japanese edition on January 31, 2003, when the number of registered users jumped from seven to 66 in one month, and the number of articles from 68 to over 2,000.

Largest Asian Language Wikipedias (March 1, 2004)

Rank	Language	Articles	Growth ¹	Share
1.	English	217016	+17%	41.770%
3.	Japanese (日本語)	33415	+51%	6.431%
12.	Chinese (Hànyǔ)	6675	+91%	1.285%
14.	Hebrew (עברית)	4806	+103%	0.925%
25.	Russian (Русский)	1774	+53%	0.341%
28.	Malay (Bahasa Malaysia)	1102	+55%	0.212%
29.	Korean (Hangukeo)	975	+79%	0.188%
38.	Indonesian (Bahasa Indonesia)	416	new	0.080%
41.	Hindi	221	+183%	0.043%
42.	Tamil	206	new	0.040%
44.	Turkish (Türkçe)	173	+312%	0.033%
45.	Arabic (عربي)	119	+164%	0.023%
49.	Kurdish	75	new	0.014%
52.	Vietnamese (Tiếng Việt)	65	new	0.013%
53.	Farsi (فارسی)	43	new	0.008%
55.	Malayalam	38	+6%	0.007%
57.	Tatar	31	new	0.006%
59.	Albanian (Shqip)	21	new	0.004%
60.	Nauruan	16	new	0.003%
63.	Filipino/Tagalog	11	new	0.002%

Source: Wikipedia:Multilingual Statistics

¹Growth measured from December 19, 2003 to March 1, 2004

Table 2 Asian Language Wikipedias

Chinese Wikipedia

The Chinese language Wikipedia (<http://zh.wikipedia.org>) is still relatively small, with just over 6,500 articles, and ranks as the 12th largest just behind Esperanto and Italian (as of March 1, 2004). In Chinese, the term for encyclopedia is 百科全书 (bai ke quan shu, literally “hundred sections complete book”) which was used to construct the Chinese name for Wikipedia – 维基百科 (wei ji bai ke). Phonetically, it is similar to “wiki” and directly translated means “maintain a foundation of a hundred sections.”

In common use, it is referred to as ZH (short for zhongwen), the two letter code designated for the project.

Given the number of Chinese speakers worldwide, it is curious that the number of dedicated contributors is quite small, with only about a half-dozen regular writers for the site. However, they come from backgrounds representing the breadth of the worldwide Chinese diaspora – PRC residents from Beijing, Guangdong and Hunan; Taiwanese in Canada; Hong Kong residents; and Chinese speakers living in Singapore, United States, Canada, France and Germany. Because of this diversity, the content reflects this as well – it uses a mix of both simplified (PRC and Singapore) and traditional (Hong Kong, Taiwan, United States and others) Chinese characters, mainly because the user community is so small there is not enough critical mass to create two separate versions. Nevertheless, simplified characters and colloquial Putonghua used in the PRC tend to dominate the site.

Although it was started in 2001, ZH Wikipedia only began full operation in 2002 when the original main page was created on October 27, 2002 by a Beijing-based user, Mountain. With the help of friends, he helped configure Wikipedia's software to support Unicode, which enabled the mixed use of both simplified and traditional Chinese.

A bright spot for ZH is its status as one of the fastest growing Wikipedias, and the existence of a large untapped set of contributors in the PRC who are increasingly becoming Internet-savvy. At just over 80 million Internet users, China trails only the United States and Japan, and by 2010 experts predict China will be the world's largest Internet market²³.

Since Wikipedia is hosted on a server outside the PRC, one concern is whether it is subject to blocking by government Internet filtering and firewalls. So far, to the

surprise of many, it has not been subject to any blanket restrictions. Whether it's the current small size and low profile of ZH Wikipedia, or a long term comfort with the existence of Wikipedia as an online source, the case of Wikipedia will be an interesting indicator of how the PRC sees this and other forms of participatory journalism affecting the mainland's media ecology.

Other Wikipedias

Other than Japanese and Chinese, most of the East and Southeast Asian language Wikipedias are still in the early stages of development. Only Malay, Korean and Indonesian Wikipedias are in the vicinity of 1,000 articles. For the Indian subcontinent, Hindi, Tamil (Tamil Nadu and Sri Lanka) and Malayalam (Kerala state) versions have been created. However, the language versions for several prominent countries have not yet been started, such as the Burmese (Myanmar) and Khmer (Cambodia) Wikipedias. To assist these newer arrivals, the Simple Wikipedia project (<http://simple.wikipedia.org>) is an English language version using basic information in the articles and only simple English, which provides an easy starting point for translation to other languages.

Conclusions

Wikis are still in the early stages of generating credible collaborative content. Perhaps the toughest part of Wikipedia's future is how to manage its own success. While Wikipedia has recorded impressive accomplishments in three years, its articles have a mixed degree of quality because they are, by design, always in flux, and always editable. That reason alone makes people wary of its content. First time visitors are typically impressed with what the community has developed, considering the decentralized nature of the effort and the usefulness of its content. But there is still

a large number of disputes among community members making ever more persistent stands on different issues. Entire sets of articles relating to Israel-Palestine, Polish and German names for Gdansk/Danzig, North Korea's government, medical science, scientific skepticism, AIDS and religion are constantly under assault and have resulted in heated debates and even in individuals leaving the community in resentment. However, the vast majority of Wikipedia articles have settled into a state of equilibrium, as less than 0.10% of the pages are actively disputed with a "neutral point of view" warning²⁴.

Wales envisions someday a "1.0" version of Wikipedia²⁵ — a tangible product in printed form or CD-ROM as a codified snapshot and reference work targeted at those not connected to the Internet. Understandably, there has been contentious debate within the community on how to do something that is unnatural for a wiki — freeze its content or certify a particular version of an article. For many users, this amounts to a corruption of the wiki concept, while others view it as a natural and necessary progression of the project as more and more articles reach a "finished" state.

Wikipedia is a rich area of study as a forum for participatory journalism, because it is a unique implementation of social software — it attempts to maintain a balance between the goals of human usability and technological solutions.

As more developing countries get connected to the Internet, it will be important to avoid having technical expertise as a precondition to participation. Web logs and wikis attract users because they de-emphasize the "technical" aspect of technology adoption. So while the Internet has traditionally been a technocracy-driven community, the face of the global Internet is changing. Tools that can provide easy access to publishing are already bridging the digital divide, and Asia's diverse economic and social landscape will benefit from this trend.

About the Author

Andrew Lih is currently assistant professor at the Journalism and Media Studies Centre at the University of Hong Kong. He also serves as the center's director of technology. His research interests include online communities, mobile computing, collaborative and participatory journalism. Email: alih@hku.hk

Acknowledgements

I wish to thank Dan Gillmor, Jerry Michalski and Sreenath Sreenivasan for their insight on online and participatory journalism; Wikipedia developers Magnus Manske, Tim Starling and Brion Vibber for their help in tracing the early technical history of Wikipedia; and Menchi, Cindy Chan and Sun Lixia for help on researching Chinese Wikipedia and web logs.

References

- ¹ Shayne Bowman and Chris Willis, *We Media, How audiences are shaping the future of news and information*, July 2003, <http://www.hypergene.net/wemedia/>
- ² Michael Hauben and Ronda Hauben, *Netizens: On the History and Impact of Usenet and the Internet*, IEEE Computer Society Press, April 1997, <http://www.columbia.edu/~rh120/>
- ³ Dan Gillmor, *Making the News*, Draft, March 10, 2004, expected 2004
- ⁴ Back to Iraq, <http://www.back-to-iraq.com/>
- ⁵ Bowman and Willis
- ⁶ "The Material People Contribute to the Online World," Pew Internet and American Life Project, February 29, 2004, <http://www.pewinternet.org/reports/toc.asp?Report=113>
- ⁷ "Participatory Journalism Puts the Reader in the Driver's Seat," *Online Journalism Review*, J.D. Lasica, August 7, 2003
- ⁸ <http://www.flyingchair.net/awards.php>
- ⁹ "Blogger: the media future behind Diary", <http://gb.chinabroadcast.cn/41/2004/02/14/115@67443.htm>
- ¹⁰ "Blocking the Bloggers," LivinginChina.com web log, <http://www.livinginchina.com/archives/000711.html>
- ¹¹ "Japan, China, Korea plan joint open-source project," IDG News Service September 4, 2003, Martyn Williams, http://open.itworld.com/4907/030904japan/page_1.html
- ¹² "Asian trio in deal to replace Windows," CNET Asia, September 02, 2003, <http://news.zdnet.co.uk/software/linuxunix/0,39020390,39116051,00.htm>
- ¹³ "History and Philosophy of Project Gutenberg," Michael S. Hart, August 1992, <http://www.gutenberg.net/about.shtml>
- ¹⁴ Nupedia archive, http://web.archive.org/web/*/www.nupedia.com/main.shtml
- ¹⁵ "Nupedia," *Wikipedia: The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Nupedia>
- ¹⁶ "Wikipedia:Modelling Wikipedia's growth," *Wikipedia: The Free Encyclopedia*, http://en.wikipedia.org/wiki/Wikipedia:Modelling_Wikipedia's_growth
- ¹⁷ Andrea Ciffolilli, "Phantom authority, self-selective recruitment and retention of members in virtual communities: The case of Wikipedia", *First Monday*, volume 8, number 12 (December 2003), http://firstmonday.org/issues/issue8_12/ciffolilli/index.html
- ¹⁸ "Metadata," *Wikipedia: The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Metadata>
- ¹⁹ Clay Shirky, "Wikis, Grafitti, and Process," *MANY-TO-MANY: social software* web log, August 26, 2003
- ²⁰ "Unicode 4.0.1," <http://www.unicode.org/versions/Unicode4.0.1/>
- ²¹ "Wikipedia:Multilingual_statistics," http://en.wikipedia.org/wiki/Wikipedia:Multilingual_statistics
- ²² "Wikipedia Charts Japanese," <http://www.wikipedia.org/wikistats/EN/ChartsWikipediaJA.htm>
- ²³ Li Xiguang, "ICT and the demise of propaganda: China's Internet experience," *Asian Cyberactivism: Freedom of Expression and Media Censorship*, edited by Steven Gan, James Gomez and Uwe Johannsen.

²⁴ “NPOV Dispute Metrics,” study from September 2003, Andrew Lih,
http://en.wikipedia.org/wiki/User:Fuzheado/NPOV_disputes

²⁵ “Pushing to 1.0,” http://en.wikipedia.org/wiki/User:Jimbo_Wales/Pushing_To_1.0