



UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Sum of all knowledge: Wikipedia and the encyclopedic urge

Salor, F.E.

[Link to publication](#)

Citation for published version (APA):

Salor, F. E. (2012). Sum of all knowledge: Wikipedia and the encyclopedic urge

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<http://dare.uva.nl>)

3. Introducing Wikipedia

Not to be too dramatic about it, but, 'to create and distribute a free encyclopedia of the highest possible quality to every single person on the planet in their own language,' that's who I am. That's what I am doing. That's my life goal.

Jimmy Wales
Speaking to *The New York Times* in 2008

3.1 Introduction

The previous chapters have followed the encyclopedic ideal from its Platonic origins to the end of the 20th century. Under many names and in various organizational schemes, evolving societies have produced works that embody differing conceptions of knowledge and authority, depending on an ever-changing media landscape. As the conclusion of the preceding chapter argued, at the end of the 20th century, the encyclopedic ideal was facing a struggle to relevance, which established practices were unable to resolve. While numerous prominent thinkers from diverse fields looked into possible ways to reshape encyclopedias throughout the 20th century, their propositions were mostly related to organizational schemes and they usually did not consider questioning fundamental assumptions regarding the nature of knowledge and authority.

This chapter will introduce *Wikipedia*, starting with the social and technological outlook surrounding its inception, and with a brief overview of the personal goals of its founders. Before moving on to the creation and growth of *Wikipedia*, a review of wiki software and its underlying principles will be presented. Some of the most current research on *Wikipedia* will be instrumental in presenting a thorough analysis of what has become the largest encyclopedia ever compiled and the greatest online collaborative work of any kind. Before the in-depth evaluation of the key concepts that consists the core of this study can take place, such a thorough and multifaceted understanding of *Wikipedia* is vital to avoid either oversimplification of its workings or being caught in hyperbole about its achievements and promise.

3.2 Nupedia and Origins

Similar to the *Encyclopédie*, *Wikipedia* emerged as a revolutionary encyclopedia almost by accident. While the original intent of the publishers of the *Encyclopédie* was to translate Chambers's *Cyclopaedia* to French, founders of *Wikipedia* started with the aim of building an online encyclopedia that would conform to most of the established norms.

Observing the success of open-source software movement and the Project Gutenberg, Internet entrepreneur Jimmy Wales saw the potential to build an encyclopedia with similar tools and, more importantly, built on principles of openness and collaboration. Born and raised in Huntsville, Alabama, near the Marshall Space Flight Center-nexus of activity during the cold-war space race-Wales harbored an early enthusiasm for science, technology and encyclopedias (Lih, 18). Both Lih and Wales's *Wikipedia* page emphasizes the importance of the *Encyclopedia Britannica* and more importantly the *World Book Encyclopedia* for the young Wales.¹⁷ After pursuing college degree in Finance, Wales briefly worked in options trading before starting up his own online business in 1996, starting a male-oriented online portal called *Bomis*. With the relative financial ease of owning a profitable Internet company, around year 2000, Wales started considering the idea of building an online encyclopedia. According to Lih, being a fan of Ayn Rand and objectivism accounts for an important portion of Wales's interest in encyclopedias in general and his belief in the possibility of building one collaboratively (32).

In early 2000, Wales appointed Larry Sanger, a PhD student in Philosophy whom Wales knew through objectivist discussion forums, as the editor-in-chief and given him freedom to build the guidelines of the new encyclopedia, dubbed *Nupedia*. At the time, Sanger was finalizing his dissertation titled *Epistemic Circularity: An Essay on the Justification of Standards of Justification* and one of his primary areas of focus in shaping *Nupedia*, and later *Wikipedia*, was maintaining neutrality. Given the financial backing of *Bomis*, Sanger went on to draft his vision of *Nupedia* and proposed a multi-tiered and

¹⁷ *World Book Encyclopedia* is a general encyclopedia set first published in the United States of America in 1917. Among the distinctive features of the *World Book Encyclopedia* is its pedagogical focus, emphasizing topics that might be of interest to adolescents and highlighting parts of articles in order to assist navigation and its inclusion of numerous illustrations and diagrams.

tightly controlled review structure. Given his enormous respect and love for the academic world, Sanger was also adamant in drafting the best possible writers and editors for each topic. Editors, who would be authoritative experts in their fields, would assign topics to volunteering writers who are required to hold a PhD in their respective topics. These general principles led to a seven-stage process for the preparation of each article:

1. Assignment
2. Finding a lead reviewer
3. Lead review
4. Open review
5. Lead copyediting
6. Open copyediting
7. Final approval and markup

The project was announced and went online in March 2000 and the first article to clear the reviewing process went online in September (“Atonality”). According to Lih, the project was forecasting growth by tens of articles for the year which, given the incredibly rigorous review process, is not entirely surprising (40). Designed by Sanger according to his personal stance towards the nature of expertise, the review system of *Nupedia* created too many bottlenecks for the creation of content and rendered volunteering for the project very unappealing. Even Jimmy Wales, the co-founder and financier of the project, realized this when he attempted to write an article on options trading, a topic he had studied in graduate school and later worked at professionally. Even though his knowledge on the topic was adequate, Wales admits to having been intimidated by the fact that his entry would be sent to finance professors for review: “it felt like homework” he explained (Wales qtd. in Schiff).

During its first year, *Nupedia* managed to produce more than two dozen articles and, by all accounts, the shortcomings of the project were evident. It was clear that a fundamental change was necessary in order to realize Wales’s dream of producing a comprehensive encyclopedia with open and free production tools. That change would transform the ailing *Nupedia* into the burgeoning *Wikipedia* that is still growing today.

3.3 The Wiki

The ideas of openness and sharing that constituted the foundations of the web in the beginning have been covered in the previous sections. However, the rapid popularization and mass adoption of the Internet during the 90's also came to imply a web that is almost entirely read-only. Publishing a website increasingly became a minority occupation and the overwhelming majority of new users coming to the web used their access to browse existing pages. While still representing a remarkable potential for disseminating and sharing information, a mostly read-only Internet would fall short of the ideals of not only the web pioneers like Berners-Lee, but Nelson and Bush as well, who all envisioned the web as an empowering creative medium for all users.

A possible solution to create easily editable websites came from Ward Cunningham in the shape of wikis in 1994. Inspired by Bush's Memex and Apple's *Hypercard*, Cunningham's wiki, meaning "quick" in Hawaiian, allowed each page to be edited within a standard web browser with a very simple markup language.¹⁸ In addition, all the edits made within a single page were individually archived and made available for further browsing, therefore allowing instant revisions and reverts. Since any change can be easily reverted and any mistake rapidly rectified, wiki software encouraged its users to contribute without fear of damaging what was already done. Providing such a radical degree of openness and transparency, the WikiWikiWeb started to grow in popularity among programmers, who, by the nature of their work and interests, have been on the forefront of most collaborative projects this far. In 2001, Larry Sanger was introduced to the wiki idea by, Ben Kovitz and instantly found the idea behind it very appealing. The ailing state of Nupedia was leading its founders to look for alternatives to improve article creation and speed of development.

On January 10 2001 Larry Sanger sent a message to the Nupedia mailing list with the title "Let's make a Wiki", wherein he detailed the idea and proposed its implementation:

No, this is not an indecent proposal. It's an idea to add a little feature to Nupedia. Jimmy Wales thinks that many people might find the idea objectionable, but I think not.

¹⁸ *Hypercard* was an application software created for the Apple Macintosh in 1987 that allowed users to create hypertext and hypermedia projects. It is credited as being among the first widespread and user-friendly application of hypertext and related concepts.

...what it means is a VERY open, VERY publicly-editable series of web pages. For example, I can start a page called EpistemicCircularity and write anything I want in it. Anyone else (yes, absolutely anyone else) can come along and make absolutely any changes to it that he wants to. (The editing interface is very simple; anyone intelligent enough to write or edit a Nupedia article will be able to figure it out without any trouble.) On the page I create, I can link to any other pages, and of course anyone can link to mine. The project is billed and pursued as a public resource. There are a few announced suggestions or rules. As to Nupedia's use of a wiki, this is the ULTIMATE "open" and simple format for developing content. We have occasionally bandied about ideas for simpler, more open projects to either replace or supplement Nupedia. It seems to me wikis can be implemented practically instantly, need very little maintenance, and in general are very low-risk. They're also a potentially great source for content. So there's little downside, as far as I can see. We would not integrate the Nupedia wiki into the rest of Nupedia (though wiki pages could link to regular Nupedia pages, there wouldn't be links back). It would be a completely separate part of the website. On the front page of the Nupedia wiki we'd make it ABSOLUTELY clear that this is experimental, that Nupedia editors don't have control of what goes on here, and that the quality of articles, discussion, etc., should not be taken as a reflection of the quality of articles, review, etc. on the main part of the Nupedia website. (Sanger)

The academic community of Nupedia was lukewarm to the idea of Wiki based editing and wanted to keep the experiment separate from the main site. The disdain of the Nupedia's experts towards the idea of mingling with potentially anonymous editors on a relatively unmoderated space was telling. During the majority of its ten years in existence, *Wikipedia* bore a stigma from the academic community and the supposed anti-authoritarian stance of its editors became a basis of some of the harshest criticisms directed at the project.

Thanks in part to outreach programs of the Wikimedia foundation, vocal advocacy of students and a limited number of academics and mostly due to its enduring and overwhelming popularity, the academic community is slowly coming to terms with the open and free encyclopedia in recent years. However, despite the highly skeptical stance of the existing expert community of *Nupedia*, the new site is named *Wikipedia* and went online 15 January 2001 ("History of *Wikipedia*").

The previous chapter has discussed the impact of various technologies on the eventual character and community practices of *Wikipedia* today. While it is clear that a project like *Wikipedia* owes its existence to the Internet and the ubiquity of personal computers, the influence clearly goes much deeper. The open source software movement in general and Ward Cunningham in particular all shared certain principles regarding access to

knowledge and the terms under which one's work should be shared. While these proved revolutionary in the computer industry and transformed the manner in which many companies conducted their core business, an equally fundamental effect can be observed on established notions concerning knowledge and authorship. As explained above, the implementation of wiki software transformed the nature of the project first attempted by Wales and Sanger and necessitated a different organizational and editorial scheme. The remaining sections of this chapter will follow the growth of *Wikipedia*, charting its popularization and by drawing upon the most current research and observations as a means of presenting a comprehensive picture of the project as it stands after a decade in continuous revision.

3.4 Charting the Growth of Wikipedia

The growth of *Wikipedia*, unlike that of Nupedia, was immediate and explosive from the beginning. While the project has attracted much attention since its inception, a watershed moment was being reported on *Slashdot.net* on July 2001, an immensely popular website frequented by the most devoted and knowledgeable technology enthusiasts. The sudden influx of knowledgeable and engaged users directed to *Wikipedia* via Slashdot was a boon to the early growth of *Wikipedia* and contributed to the prevailing culture of the community. Following technology news sites, the first mainstream media report on the fledgling project was on the *New York Times* in 20 September 2001 (Meyers). These and other media profiles acted as pointers to the site and brought fresh volunteers with possibly new interest groups. Numbers provided by the "History of *Wikipedia*" article illustrates the growth rate adequately: "The project passed 1,000 articles around 12 February 2001, and 10,000 articles around 7 September. In the first year of its existence, over 20,000 encyclopedia entries were created—a rate of over 1,500 articles per month. On 30 August 2002, the article count reached 40,000."

Shortly after the launch of *Wikipedia* the so-called dot-com bubble has burst leading many Internet companies to bankruptcy and strictly limited the financial levity of Bomis

towards Wales's encyclopedia projects.¹⁹ As a result Wales had to let go all of the staff hired for Nupedia, including Larry Sanger in December 2001 ("Nupedia"). Sanger stayed on the project as a volunteer with the hope that better economic times would see the return of his employment. However, the incompatibility of Sanger's academic and supposedly authoritarian attitude with the unique and rapidly evolving features of *Wikipedia* and its community would only lead to increased tension over time, ending with an embittered Sanger leaving the project for good in 2002 (Reagle, 41).

Meanwhile, a remarkable development took place with the launch of the non-English *Wikipedias* which began almost immediately after the initial January launch in English. Within months, German, Spanish, Japanese, French, Dutch and many other languages were online ("History of *Wikipedia*"). "In January 2002, 90% of all *Wikipedia* articles were in English. By January 2004, less than 50% were English, and this internationalization has continued to increase. As of 2007, around 75% of all *Wikipedia* articles are contained within non-English *Wikipedia* versions (ibid.)".

As the project gained momentum and rapidly increased in magnitude, the establishment of fundamental guidelines and principles became a pressing issue, as evidenced by what is known as the Spanish Fork incident. After the termination of his paid position at *Wikipedia*, Sanger openly speculated about the possibility that Bomis, then the owner of the project, might put advertising online with the encyclopedia content in order to improve its financial situation, which might result in his reemployment (Sanger qtd. in Lih, 137). Even in 2002, just one year after the launch of *Wikipedia*, this mere mention of the possibility of advertising created uproar among the editors. While both Wales and Sanger immediately assured the editors that no actual plans for such a scheme were in place, a group of editors in the Spanish *Wikipedia*, lead by Edgar Enyedy, announced that they would be creating a fork of the Spanish *Wikipedia*, independent of Bomis and called all editors to join

¹⁹ *Dot-Com Bubble*: The sudden, and mostly speculative, increase in the value of stock markets of the Western World due to the expanding base of businesses that are related to the Internet is generally referred as the dot-com bubble. Most of the businesses were focused on rapid increases in market shares rather than actual growth and their founders have spent their earnings on their personal consumption rather than investing on the business. The speculative increase came to an abrupt end at the end of 2000 and many such businesses perished with the following recession. (http://en.Wikipedia.org/wiki/Dot_com_bubble, Accessed 20.04.2011)

them.²⁰ The resulting *Enciclopedia Libre*, hosted by the University of Seville, went online in February 2002 and the resulting divide in editors and attention was a considerable setback for the Spanish *Wikipedia* (Lih, 138). While the later growth of *Wikipedia* as a whole, and continuing assurances that the site would never run advertisements eventually unified and reinforced the Spanish volunteers, the subject of advertising continues to be an extremely sensitive topic among the *Wikipedia* community. While the terms and conditions of the discussion around the means of funding *Wikipedia* is entirely different than any encyclopedia of the last three centuries, the debate still proves that all encyclopedic projects has to depend on a sustainable economic model to develop and mature. It is striking that forming a profitable corporation around the periodical publication of encyclopedias seems to be only one way of serving this need.

In order to address this and other institutional issues, while taking into account the meteoric growth of the project, Jimmy Wales announced the establishment of the non-profit Wikimedia Foundation in June 2003 (“Wikimedia Foundation”). The foundation has the principal aim of overseeing *Wikipedia* and its sister projects, appointing key staff and, since the project is devoted to being an independent charity, and given that the idea of advertisement on the site is effectively ruled out, one of the foundation’s aims is also to organize yearly fundraising efforts. Jimmy Wales transferred all the rights over *Wikipedia* to the newly established foundation and while he is the chairman emeritus, the board of trustees is elected yearly and Wales himself usually focuses his role on public outreach and adopts a hands-off approach to most administrative issues (ibid.). The adoption of certain managerial policies by Wales, as well as the foundation in general, has a significant impact on the continuing growth of *Wikipedia* and health of its community.

²⁰ *Forks*: A term common to software engineering, “a fork happens when developers take a legal copy of source code from one software package and start independent development on it, creating a distinct piece of software. Free and open source software is that which, by definition, may be forked from the original development team without prior permission [and] without violating any copyright law”. ([http://en.Wikipedia.org/wiki/Fork_\(software_development\)](http://en.Wikipedia.org/wiki/Fork_(software_development)) Accessed 20.04.2011). The open license applied over all *Wikipedia* content allows such an action.

In addition to the Foundation as a much more formal organizing element, the *Wikipedia* community also started formulating its own policies and guidelines very early on. Arguably the most central of *Wikipedia* principles, The Neutral Point of View (NPOV) was formulated as early as 2001 and *Wikipedia*'s own Manual of Style is launched in 2002 ("History of *Wikipedia*"). The NPOV principle continues to be at the center of numerous discussions around *Wikipedia* and will be discussed more thoroughly later in this study. With these and other organizational features are established, the number of articles continues to grow at an accelerating rate and in an increasing number of languages. By 2007, *Wikipedia.org*, which became the homepage for all languages, is ranked as one of the ten most frequently visited domains on the Internet with 7.5 million articles in more than 250 languages (ibid.). The chart below shows the growth pattern for the English language encyclopedia alone.

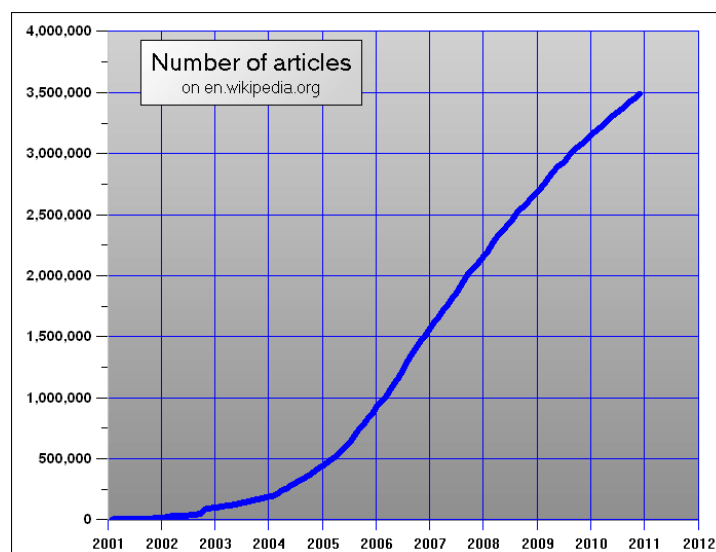


Figure 4 Growth of English Wikipedia²¹

While the Wikimedia foundation fostered the creation and growth of sister projects like Wiktionary, a wiki based dictionary and Wikimedia Commons, a media repository consisting of public domain and otherwise freely licensed images, video, sound and other educational material, as *Wikipedia* itself continues to grow. As of 2010 there are more than

²¹ <http://en.Wikipedia.org/wiki/File:EnWikipediaArt.PNG> (Accessed: 21.04.2010)

17 million articles in more than 270 languages, 3.5 million of which are in English. An often overlooked fact is, while most detractors of *Wikipedia* concentrate on the faults and shortcomings of the English version, for an overwhelming majority of the 270 languages in which *Wikipedia* is available, for all its faults, it is the only encyclopedia available.

It is remarkable that within ten years the non-profit, collaborative online project that attracted constant criticism, has managed to create an encyclopedia that surpasses, at least in scope and reach, every example in the long history covered in the early part of this study. Given these basic figures, it is not surprising that *Wikipedia* has been the subject of numerous studies and heated discussions, academic and otherwise, since its inception. The following sections will explore *Wikipedia* through multiple viewpoints, covering the established rules and principles of its community, both explicit and implicit, and highlight some of the critical issues that surround it as a credible encyclopedia. However, it is vital for these in-depth analyses to be grounded in a solid understanding of *Wikipedia*, therefore the next section will provide the fundamental policies and practices of the encyclopedia that anyone can edit.

3.5 Notable Stubs and Other Oddities: A Wikipedia Primer

As the previous sections illustrated, the adoption of wiki software and its inherently open nature defines *Wikipedia*. On the most basic level, every article on *Wikipedia* is an independent page that can be edited by anyone at any time using a relatively simple online interface found at the top of the article titled “edit”. In addition, at the top of every article page there are links to “discussion” and “view history”. Any visitor to the site can browse the complete history of each article and compare its state in different dates and also browse some of the issues that have been highlighted by other users in the discussion section.

These features are inherent properties of wiki software and can be observed in any online platform that chose to adopt the technology. Over the last decade the *Wikipedia* community of editors, under the guidance of founder Jimmy Wales and the Wikimedia foundation, have gradually formulated the policies and guidelines for the creation of content on the open space provided by the software. Through a long process of iteration,

the long list of policies is grouped under five pillars and these constitute fundamental principles that define *Wikipedia* in the broadest sense. They are (“*Wikipedia: Five Pillars*”):

- *Wikipedia* is an online encyclopedia
- *Wikipedia* has a neutral point of view
- *Wikipedia* is free content
- *Wikipedians* should interact in a respectful and civil manner
- *Wikipedia* does not have firm rules

The first pillar is given a negative definition with a long list of things *Wikipedia* is not (a dictionary, a soapbox, a repository...) and further leads to the core content policies of the encyclopedia (“*Wikipedia: What Wikipedia is not*”, “*Wikipedia: Core content policies*”):

- Neutral Point of View (NPOV)
- Verifiability
- No Original Research

As is evident by the fact that it is featured twice, once as a pillar and once as a core content policy, NPOV can be considered as the most central and cherished value of *Wikipedia*. The policy is explained as “representing fairly, proportionately, and as far as possible without bias, all significant views that have been published by reliable sources” (“*Wikipedia: Neutral point of view*”). The unpacking of this definition takes approximately 4000 words on the dedicated policy page, however both the practical implications and the philosophical foundations of NPOV policy continues to be a hotly contested topic both among academics studying *Wikipedia* and users making edits on a day to day basis.

Moreover, the full meaning of NPOV is complemented with the other two core content policies. *Verifiability* states that “the threshold for inclusion in *Wikipedia* is verifiability, not truth: whether readers can check that material in *Wikipedia* has already been published by a reliable source, not whether editors think it is true” (“*Wikipedia: Verifiability*”). *No original research* completes this stance by stating that *Wikipedia* can not be the sole resource for any material published on it, including “any analysis or synthesis of published material that serves to advance a position not advanced by the sources” (“*Wikipedia: No original research*”). While the formulation of these core policies dates back to the earliest days of the project, their details and specific implementation has been a process of constant iteration and, although the fifth pillar reminds users that rules are

primarily enacted as best practice guidelines, the three core content policies are not subject to discussion and any group challenging them would be invited to fork the *Wikipedia* project and start their own encyclopedia.²²

The third pillar, “*Wikipedia* is free content”, refers to all the ways editors and readers can access the encyclopedia. Since the content can be edited by anyone, no single person or entity can claim ownership over the work being done. Also, the editors are invited to use images and other media from copyright-free sources in order to avoid conflict.

The fourth pillar refers to *Wikipedia’s conduct policies* that provide guidelines for community behavior. The relevant policies invite editors to be civil (“*Wikipedia: Civility*”), avoid personal attacks and edit wars and generally assume good faith (“*Wikipedia: Assume good faith*”) when discussing with fellow *Wikipedians* and always try to achieve consensus (“*Wikipedia: Consensus*”).²³

Finally, the last pillar states that almost all rules, most notable exceptions being the core content policies like NPOV and NOR, can be ignored or subjected to renegotiation if they prove to be detrimental to the project overall or hinder a particular effort (“*Wikipedia: Ignore all Rules*”).

In addition to the core content policies summarized above, *Wikipedia* have numerous other rules and guidelines. Among the most important is *notability* (“*Wikipedia: Notability*”). Notability outlines conditions a subject has to match in order to merit an article in *Wikipedia*. Due to the central role they play, notability guidelines and their specific implementations are subject to intense debate within the community and constitute the core of the *inclusionist vs. deletionist* debate, arguably the most famous disagreement among camps of editors. Briefly, editors who identify themselves as *deletionists* argue for strict enforcement of notability criteria and they argue for the greater curatorial role this imposes on the community of editors. Conversely, *Inclusionists* maintain that since the

²² There have been a few incidents where groups with certain political or religious views, become frustrated by *Wikipedia’s* NPOV rule and notability criteria, and have started their own online encyclopedias based on wiki software. Two examples are *conservapedia* (conservapedia.com) and the *creationwiki* (creationwiki.org).

²³ When individuals or groups continuously revert each others changes in a short amount of time, this behavior is called an edit war and when consensus can not be achieved, usually results in an admin locking the page for a period of time for respective parties to cool off.

marginal cost of having another article is negligible for the project, *wiki is not paper* as they point out, there is no harm in allowing what a majority of users might consider as trivial topics.

While the specifics of the debate between the two camps of editors might seem esoteric to an average visitor to the encyclopedia, it is remarkable that after centuries of evolution, the role of the editor still remains central to encyclopedic enterprise. As in other issues, the model of *Wikipedia* allows the fundamental struggles of editorship to become public and transforms the decision making into an iterative process.

Notable or otherwise, most *Wikipedia* articles start their lives as *stubs*. Mostly created by an editor as an acknowledgement that the subject probably merits an article, containing usually just a one or two sentence definition and some context. After creating a *stub*, the editor is expected to put up a notice at the top of the page indicating the state of the article as such and inviting others to contribute and expand. These notices are part of a collection of standardized warnings that aim to highlight various perceived defects in articles to visitors, so they can be cautious and to editors, as pointers for work needing to be done. Most common notices are for when an article lacks citations, when the neutrality of an article is in dispute or when the quality of the existing article is deemed inferior to the standard prescribed by the *Wikipedia* guidelines.

Striving to abide by and enforce these rules is the community of editors. Although the practice remains controversial, anonymous visitors can still edit almost all pages on *Wikipedia*. However, most editors prefer the added benefits of having and maintaining a user account. These accounts allow the editors to build a personality and prestige with their cumulative edits in the community. As the related research will illustrate in the following section, this identity building practice is a vital motivational factor for editors to sustain and improve their efforts. Selected among the most dedicated editors are the moderators who yield some extra powers mostly related to dispute resolution and maintenance. Mods can lock articles to stop heated edit wars and continuous vandalism or temporarily ban users. Representing the next line of authority in dispute resolution is the Arbitration Committee, which consists of elected members that are dedicated to resolve issues that cannot be resolved by the community. While a voting can be held to determine the outcome of a

dispute, it is regarded as a tool of last resort and for very special occasions and editors are invited to achieve consensus to the last minute (“*Wikipedia*: Polling is not a substitute for discussion”).

Excluding the rare instances where an actual poll takes place, editors are expected to reach consensus on the best way to present any article in a way that conforms *Wikipedia* guidelines, most importantly NPOV. The wiki software provides the tools to be used in the process of building consensus. While the articles are open to editing at any time, each edit can also be reverted at any time to any other previous version. Editors are also strongly encouraged to discuss issues on the talk pages attached to every article before committing actual edits to the article itself. On the “*Wikipedia:Consensus*” page that explains the policy, a simplified diagram is offered to illustrate an idealized workflow of how consensus should be sought.

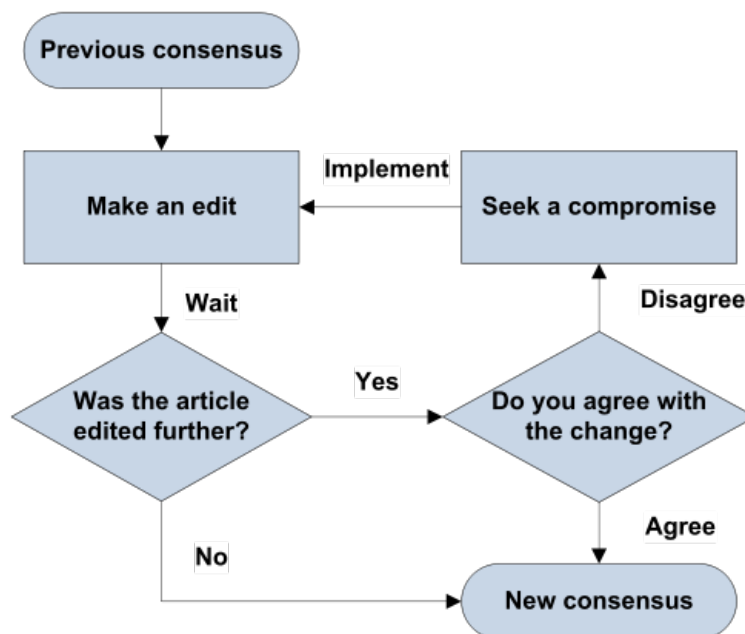


Figure 5 Consensus Flowchart by *Wikipedia*²⁴

However, it is not hard to imagine that this process not always works so smoothly. Groups of editors or individuals often become entrenched in their opinions concerning a particular aspect of an article, resulting in edit wars that usually consist of repeatedly

²⁴ http://en.Wikipedia.org/wiki/File:Consensus_Flowchart.svg

reverting each others' contributions (*Wikipedia: Edit warring*). The policy on edit wars, and consensus, prescribe that disagreements should be resolved on the discussion pages and edits committed to the article should reflect the new consensus. However, it is not uncommon to see two camps, or very dedicated individuals, continuously reverting each others' edits. To prevent such disputes from spiraling out of control and disrupting the larger community, *Wikipedia* has two crucial safeguards. First is the "three-revert rule" that limits the number of reverts that can be made by an individual editor in 24 hours by three, second is the temporary protection of the article under dispute to prevent all editing which can be imposed by a moderator in order to force the disputants cool off.

Unsurprisingly, articles related to current events and controversial topics are the most frequent scenes of edit wars. While such articles attract the attention of a great number of potential editors, they might not be the most experienced members of the community and as such, might not be familiar with the rules and guidelines that govern the article creation process, leading to friction with more seasoned members. Also problematic are the articles that cover issues like "Religion", "Abortion" or "The Holocaust", all three of which are locked at the time of writing to prevent acts of vandalism and filter all editing through discussion pages.

While the numerous policies and guidelines can be overwhelming, especially for the newcomers to the project, the fifth pillar (no firm rules) encourages visitors to not be intimidated and try to contribute while advising veteran editors to assume good faith and "do not bite newcomers" (*Wikipedia: Please do not bite the newcomers*). A survey of the most important policies of *Wikipedia* reveals the central and fundamental importance of the community to the overall character and sustained health of the project. In the next part of my analysis of *Wikipedia*, I will survey some of the research aiming to better understand this dynamic community.

3.6 A Bunch of Nobodies: Understanding the Wikipedia Community

While the technical underpinnings of *Wikipedia* are an integral part of its uniqueness and success, the community of editors that volunteer their time, skills and knowledge lies at the core of the project. This section will analyze the community of *Wikipedia* with the help of recent studies, both qualitative and quantitative, starting with theoretical works dealing not only with *Wikipedia* but with all Internet-based, community-driven efforts, and gradually move on to research papers focusing on particular aspects of *Wikipedia*.

In *The Wealth of Networks*, Yochai Benkler singles out the *Wikipedia* community for being primarily driven by social norms and common goals in contrast to open source software development communities which usually organize around technical benchmarks and more clearly defined user access protocols (2006, 72). Instead, the community of editors at *Wikipedia* relies on the established principles like *Neutral Point of View* and guidelines like *Assume Good Faith*, and tries to resolve issues through consensus building. According to Benkler, “while not entirely anarchic, the project is nonetheless substantially more social, human, and intensively discourse- and trust-based than other major projects” (72). Lih, in his overview of *Wikipedia*, also comments on the central aspect of the social norms within the community and evaluates the common proposal that the communal action of individual editors resembles ants in a colony; “there is no hierarchical command structure for ants, and the intelligent survival of the colony is in fact ‘emergent behavior’ made-up of small, simple, individual decisions and signals communicated by an ant-by-ant basis...” (82). Lih further argues that the existing principles of *Wikipedia* present a common goal-writing an encyclopedia-and set the tone-*NPOV*- so that these broad principles have allowed the community to be nimble and agile while effectively channeling their energy (ibid). Jimmy Wales reports that the French *Wikipedia* calls this phenomenon the *piranha effect*; “You start with a little tiny article and it’s not quite good enough so people are picking at it and sort of a feeding frenzy and articles grow” (Wales qtd. in Lih, 83). While *Wikipedia* might be the most visible incarnation of online collaboration, it is clear that both the tools and the social norms it is built upon are the successful descendants of the open source movement and other computing pioneers.

In his extensive study on the collaborative culture of the *Wikipedia* community, Joseph Reagle offers the concept of stigmergy, coined by Pierre-Paul Grasse, to describe the collective action of wasps and termites while building complex structures (Reagle, 2008, 53). Stigmergy denotes collaborative effort where the actions of an individual are not based on communication with other participants of the community, but rather on the previous work done (ibid.). Mark Elliott, in his article “Stigmergic Collaboration: The Evolution of Group Work”, details the concept:

Stigmergy is a method of communication in which individuals communicate with one another by modifying their local environment, it is a logical extension to apply the term to many types (if not all) of Web-based communication, especially media such as the wiki. The concept of stigmergy therefore provides an intuitive and easy-to-grasp theory for helping understand how disparate, distributed, *ad hoc* contributions could lead to the emergence of the largest collaborative enterprises the world has seen. (2006)

While conceptualizations, like the ones given above, are useful in general terms, they fail to communicate the finer details of the *Wikipedia* community. Since a clearer understanding of the community of editors and their practices is crucial in understanding *Wikipedia* as an encyclopedia, the remainder of this section will discuss some of the research that focuses on the specifics of the community.

The first task in discussing the *Wikipedia*'s community of editors is to establish a basic understanding, including size and basic demographic information. As of April 21, 2011, *Wikipedia* has over 14 million named user accounts (“*Wikipedia: Wikipedians*”). However, this number, although very impressive, is almost meaningless. Not only does this number include all the members of the more than 280 language versions of *Wikipedia*, but most of these accounts are dormant or irrelevant, meaning they are created to be used for a limited time, for a specific purpose. An added complexity is the fact that an account is not required to edit, an option preferred by many contributors. Overall, even determining the exact size of the community proves to be a challenge. For these reasons, an overview of multiple studies is essential for an understanding of the authors of the world's largest encyclopedia.

A survey conducted in 2010 by Glott et al. provides one of the most comprehensive and recent data sets available. Covering 58,000 participants from 231 countries, the survey was addressed to all visitors to the site and found that 68% of respondents identified themselves as readers and 31% as active contributors. As the researchers note however, this ratio of reader to contributors is not likely to be representative, as a contributor to the site would be more likely to participate in such a survey (5). While the respondents' ages range from 10 to 85, half were under 22. Among those who identify themselves as readers, 68% were males and among only 12% reported as females (7). While these figures help in establishing a baseline for understanding the make-up of the community, they are not informative as to the particular dynamics, effectiveness or sustainability of the community. One of the most comprehensive studies to date that aimed to answer these questions was Felipe Ortega's 2009 dissertation, *Wikipedia: A Quantitative Analysis*. While developing the analytical tools with which to effectively capture the dynamics of the community, one of the central aims of Ortega's study was to outline a number of sustainability conditions.

The results of Ortega's detailed quantitative analysis show that the expansion of the encyclopedia content and the enlargement of the community have displayed similar patterns of growth in the past, ultimately leading to a recent stabilization (155). The results of the study also indicate an alarmingly high drop rate among the editors, meaning that, for a variety of reasons, more than 50% of authors abandon the project 200 days after they join (156). Perhaps more troubling is the fact that this rate is doubled (50% leave the project after 100 days) for the core members of the community who devote the most time and effort to editing. Ortega's findings place *Wikipedia* in a precarious position, since the overall number of authors has not been increasing since 2007, the burden of maintaining and improving the quality of the encyclopedia is increasingly dependent on the very active core authors (160). Given how volatile the core authors' attachment to the project can be, Ortega outlines two vital tasks for the Wikimedia Foundation in order to sustain the health of the project in the future: first, compelling new members to be involved with the project, and second, working on whatever improvements are necessary in order to compel core authors to sustain their vital efforts for longer periods of time (161).

While serving as an introduction to some of the problems experienced by the *Wikipedia* project for this study, Ortega's in-depth analysis also points towards a complex and vibrant social dynamic behind the scenes of the collaboratively written encyclopedia articles. It is noteworthy that the process of identifying areas that deserve further research is almost self evident; a review of the numerous articles and conference proceedings dealing with *Wikipedia* point to two general themes, namely, the relative influence of core users over the content with regard to casual contributors, and the identification of community dynamics and editing practices that result in the highest quality of articles. The second, which concerns article quality, also implies the corollary of determining credible metrics of quality. While the issues relating to the structure of the community, such as the distinction between core vs. casual contributors will be discussed in the remainder of this section, the following part will be dedicated to the review of the research related to community practices regarding the quality of the encyclopedic output along with a related discussion of *Wikipedia* content as a general encyclopedia.

The fact that a close-knit community of very dedicated volunteers is responsible for the vast majority of edits in *Wikipedia* is common knowledge among most observers. Jimmy Wales himself acknowledged this in many places, for example in his TED Talk in 2005. Moreover, a 2007 study conducted by Priedhorsky et al. demonstrates "that 1/10th of 1% of editors contributed nearly half of the value, measured by words read" (9). However, the case for the dominance of core users is far from conclusive. According to Kittur et al. the influence of core users was very high until 2004. Past this date however, this rate has seen a constant drop relative to the cumulative edits by the whole community (2007, 2). Despite the drop of their portion within the cumulative edits however, the study also notes that the contributions made by the core users are still sizeable and the work they put in tends to be more substantial in volume (6). While the exact influence of this core community is debated, its existence, and the role it plays over the overall health and success of the project is certain.

A better understanding of this core community is a vital research question, given its central role. One such attempt was Susan Bryant's 2005 study, which identifies *Wikipedia* as

a 'community of practice' as opposed to a formal organization according to criteria established by Wenger (Wenger, qtd. in Bryant, 2), which are:

- i. Members are mutually engaged.
- ii. Members actively negotiate the nature of the enterprise.
- iii. They have a collective repertoire of negotiable resources.

Having identified the community as such, Bryant's research investigates the dynamics of how new members move from peripheral to full participation and how this change affects their contributions to the community. According to Bryant's study, users of the encyclopedia are mostly compelled to become editors when they notice an error or find an article that relate to their interests lacking in depth (4). By making these first edits, new editors are initiated into the practice and branch out over time. Bryant's study also points out that, as users' involvement with the site becomes more mature, they become increasingly involved with 'meta' tasks, policy discussions and other issues that are more related to the efficient working of *Wikipedia* while maintaining watch lists of articles that let them keep an eye out for a number of articles, as a means of combating vandalism (9).

These results also correlate with the findings of other studies; for instance, Viegas et al. demonstrated in 2007 that "administrative and coordinating elements seem to be growing at a faster pace than the bulk of the articles in the encyclopedia" (1). Talk pages are seen to be an important part of this growth, given that they are used for strategic planning for edits, policy enforcement and its relevant discussion, and other coordination tasks. Among his findings, Ortega also recognizes that the active participation of talk pages is a factor in ameliorating the extraordinary mortality rate of core users he observed (177). Regarding the initiation of new members and the distribution of activities between novice and experienced users, a 2007 study by Almeida et al. also supports the findings reported above by arguing that growth is driven by an influx of new users and points out that most users choose to contribute to existing articles rather than to create new ones (2). In the same study, researchers also pointed out that, while an editor might be involved with a diverse range of topics, in single session they tend to concentrate on a single article or category (2).

Since all of the research summarized above points to the importance of the influx of new users and continuing involvement of core members, identifying the motivational

factors that draw people to a project like *Wikipedia* and keep them contributing has importance beyond academic curiosity. However, the possible value of such findings is only matched by the challenge of obtaining it.

The survey conducted by Glott et al. mentioned above questioned its respondents on the motivations behind contributing and two overwhelmingly popular responses were “I like the idea of sharing knowledge and want to contribute to it” and “I saw an error and wanted to fix [it]”. The first response was chosen by 72% of the respondents while the second was an answer chosen by 68%, where respondents were asked to rank 4 factors by importance. The third popular choice was “Because I think information should be freely available to everyone”, chosen by 37% of respondents (9).

An earlier attempt was made by Oded Nov in 2007 with an online survey sent to a random sample of 370 *Wikipedians* selected from the *Alphabetical List of Wikipedians* page of the English *Wikipedia*, then containing 2,847 entries (63). The results were compiled from the 151 valid responses received, 140 of which were males (ibid). The motivation factors tested in the study by Nov are presented in the following chart.

Motivation	Question example
Protective	“By writing/editing in Wikipedia I feel less lonely.”
Values	“I feel it is important to help others.”
Career	“I can make new contacts that might help my business or career.”
Social	“People I'm close to want me to write/edit in Wikipedia.”
Understanding	“Writing/editing in Wikipedia allows me to gain a new perspective on things.”
Enhancement	“Writing/editing in Wikipedia makes me feel needed.”
Fun	“Writing/editing in Wikipedia is fun.”
Ideology	“I think information should be free.”

Figure 6 Motivational factors for contributing to *Wikipedia* by Oded Nov

Among the respondents, the average total time of contribution was 8.27 hours per week. According to Nov’s survey, each of the motivational factors listed above correlated positively with contribution levels where the users ranked fun and ideology highest. Nov points out that while ideology is ranked as an important factor in contributing to *Wikipedia*, it does not correlate with contribution level, meaning that while people state that ideology is high

on their list of reasons to contribute, being more ideologically motivated does not translate into increased contribution (3).

In a more recent and similarly focused study, Lai and Lang defined *Wikipedia* contribution as knowledge-sharing behavior and investigated the validity of the four possible motivational factors identified by motivational theory that are associated with this type of behavior. The factors researchers tested against were as follows (1378):

- i. *Intrinsic motivation* refers to internal satisfaction received from the process of performing behaviors.
- ii. *Extrinsic motivation* refers to performing an activity in anticipation of obtaining a return such as pay or reputation.
- iii. *External self-concept motivation* is the primary motivation for individuals to adopt an activity that is congruent with the expectations of a reference group.
- iv. *Internal self-concept motivation* refers to the force that drives individuals to pursue an activity that meets their inherent standards.

From a random sample drawn from registered users of the English *Wikipedia* 843 people were invited to participate in a survey, 235 of whom responded. It should be noted that the vast majority of registered users had a preference for not receiving emails from other members, leaving them outside the study.²⁵ The results of the survey identified the most significant factor as internal self-concept motivation, whereas the study by Nov mentioned above has pointed towards the significance of intrinsic motivation.

It is clear that such studies have considerable limitations, some of which are intrinsic to survey methodology, others stemming from peculiarities of *Wikipedia*. However such studies are still helpful in providing insights into why people choose to participate in building an encyclopedia without remuneration. This review of some of the literature clarifies the similarities between open source software communities and *Wikipedia*. In his analysis, Joseph Reagle offers an in-depth study of the connection between open source software communities and the particulars of the *Wikipedia* community. According to Reagle, *Wikipedia* can be defined as an “open content community” (71). Reagle’s definition

²⁵ Since *Wikipedia* has become a popular research topic for scholars from many disciplines, being frequently bothered by requests to participate in such studies has become a common gripe among community members.

of openness derives from an understanding established by the open source communities and implies the following (75):

- i. Open products
- ii. Transparency
- iii. Integrity
- iv. Non-discrimination
- v. Non-interference

“*Wikipedia* is Free Content”, one of the five pillars of the project, as explained above, codifies these principles and codifies the encyclopedia as an open content community. In addition, Reagle argues that an important feature of an open content community is its embrace of an intense and sustained discussion of what it means to be open (102). This self-reflexive nature of the community can be traced back to the very design of the Wiki software since anything built using a wiki also includes, by definition, the history of its own making (Reagle, 107). The way that *Wikipedia* connects with and embodies its own history is a key characteristic that sets it apart from the entire history of encyclopedias.

According to Reagle’s study, the core policy of NPOV and the particular structure of *Wikipedia* community, which Reagle calls “good faith collaboration”, constitute a complimentary stance and define the project as an encyclopedia over anything else (103). Based on his observations regarding *Wikipedia*’s blend of different leadership styles, Reagle coins the term ‘authorial leadership’ to explain how authority and decision making is propagated through the community and especially, how one can explain the somewhat unique position hold by the co-founder Jimmy Wales. According to Reagle, “authorial leaders must parlay merit resulting from ‘doing good’ into a form of authority that can also be used in an autocratic fashion, to arbitrate between those of good faith or defend against those of bad faith, with a soft touch and humor when-and only when-necessary” (138). In the case of Wales, who is occasionally called a benevolent dictator within the community, according to Reagle, the community grants such an authority to “only those leaders that tread carefully and continue to make important contributions (including, now, the judicious exercise of autocratic authority) are granted the ‘dictator’ title. According to Reagle, the roots of this conception can be traced among the open source software communities where project focused groups usually organize around such figures, famously

Linus Torvalds in the case of Linux. Regarding the term “dictator” usually adopted by such communities, *Wikipedia* included, Reagle goes on to explain that “Whereas this term might not be the most appropriate in capturing the genuine character of this role, it serves as a warning: a good-natured joke balanced on the edge of becoming a feared reality” (162). Building on Reagle’s crucial insight, the next step in understanding *Wikipedia*, after exploring its community, is to analyze it as an encyclopedia.

3.7 An Encyclopedia Like No Other

Wikipedia’s place as an encyclopedia has always been a precarious one. Even from the start, one of its co-founders, Larry Sanger, thought it might only be serviceable as a feeder of ideas for the more accredited *Nupedia*. After ten years, although it has established itself as the *de facto* starting point for any encyclopedic query online, in no small part thanks to being ranked first at Google searches, challenges for the open encyclopedia are far from over. This section will detail some of the major issues that are related to the encyclopedic content of *Wikipedia*, aiming to present a comprehensive view of the project along with the study of its community.

The seemingly overwhelming number of policies and guidelines that occasionally stump even the most seasoned editors are the result of a community collaboratively working out best practices over time. The fact that the complete history of this process and the results of various implementations are all publicly visible is in itself a landmark resource in encyclopedia making. This self-reflexivity finds one of its most interesting instances at the “Criticism of *Wikipedia*” page, where common topics of criticism are distinguished between those directed towards the reliability of the encyclopedia content and those that are related to the community. Various studies and opinions about the quality and reliability of *Wikipedia* as an encyclopedia are covered in the page *Reliability of Wikipedia* and, as of April 15, 2011, the page contains over 15,000 words. It is also worth mentioning that, as of this writing, the article carries the warning at the top that its neutrality is in dispute, a standard message in the encyclopedia when the adherence of a certain article to the NPOV is in question. The irony of this did not escape the editors, as user Conrthomas commented

on the relevant discussion page: “Not only is it ironic that the leading authority of *Wikipedia* quality standards is *Wikipedia* itself, but also that the top of the pages [where the NPOV warning appears] hints at the fact that the said article might be unreliable.”

Regardless, the article presents a comprehensive list of the most common issues concerning *Wikipedia* as an encyclopedic resource and defines nine categories of reliability (ibid.):

- Accuracy of information provided within articles
- Appropriateness of the images provided with the article
- Appropriateness of the style and focus of the articles
- Susceptibility to, and exclusion and removal of, false information
- Comprehensiveness, scope and coverage within articles and in the range of articles
- Identification of reputable third-party sources as citations
- Stability of the articles
- Susceptibility to editorial and systemic bias
- Quality of writing

Since the early days of *Wikipedia*, most research has been focused on the first four of these criteria. While based on a variety of methodologies, most of these sorts of studies analyze a sample of *Wikipedia* content for erroneous facts or statements, occasionally comparing these results against a more traditional encyclopedia or topical source. One of the most widely cited and well-publicized examples of such a study was conducted by *Nature* in 2004 where a random sample of scientific articles from *Wikipedia* and Encyclopedia Britannica Online were compared for the number and severity of errors they contained (Giles, 2005). When the results of the study reported that the two online encyclopedias have similar number of errors, the publishers of Britannica challenged the methodology of the study, leading *Nature* to defend its findings in response (“Fatally Flawed”, “*Encyclopedia Britannica* and *Nature*: A Response”). Meanwhile, after being highlighted by *Nature*’s study, the errors on the relevant *Wikipedia* pages were corrected almost instantly.

Following the factual correctness of articles, another common thread of criticism relates to *Wikipedia*’s coverage of topics and the notability of articles. Comparing the wide range and depth of popular culture trivia to relatively shorter articles on other topics that are usually considered to be the core content of an encyclopedia, some critics argued that *Wikipedia* embodies the cultural bias of its most devoted editors (Waldman, 2004). Such articles usually consist the core of the inclusionist vs. exclusionist debate covered above.

While studies of article accuracy are important in establishing *Wikipedia* among other reference sources, an arguably more constructive branch of research is focused on identifying the processes by which the highest quality articles are produced, in order to provide the community some guidelines to promote and behaviors to encourage.

The 2007 study by Wilkinson and Huberman builds on an analysis of an aggregate of 50 million edits that have been made to the 1.5 million articles on the English language *Wikipedia*. According to the authors, the articles that display the highest quality “are distinguished by a marked increase in the number of edits, number of editors, and intensity of cooperative behavior, as compared to other articles of similar visibility and age” (Wilkinson et al., 1). The authors point out that, while most collaborative projects usually suffer from diminishing returns as the number of participants increase, *Wikipedia* does not appear to be suffering from such an adverse effect from the tremendous growth in its community (ibid.). The data analyzed by the researchers also allows them to comment on the arguments regarding claims that a small number of individuals are responsible for the bulk of the content on the encyclopedia. The findings presented in Wilkinson and Huberman’s study confirm the analogy offered by Lih earlier in this chapter, stating that usually edits beget more edits and a quick succession of edits can be associated with a large community of users working collaboratively over a short period of time. However, some articles where the editing behavior points to constant iterations over a longer period of time, such a pattern usually indicates a single user working on an article of interest or personal attachment. Overall, the study demonstrates that the model of collaborative and open editing that has been the centerpiece of almost all discussion about *Wikipedia* is, for all its merits and faults, what really builds the encyclopedia and is not only a point for attention.

Where the quality of *Wikipedia*’s articles is concerned, a 2006 study by Rosenzweig analyzed a selection of articles on historical biographies on *Wikipedia* and compared them to other online encyclopedias as well as a dedicated resource, *American National Biography Online*. According to Rosenzweig, *Wikipedia* offers greater coverage than other online encyclopedias, but lags behind the dedicated resource written by professional biographers. Also concerning accuracy, the study found *Wikipedia* to be surprisingly accurate. However, Rosenzweig argues that the primary failing of *Wikipedia* is to be found in its use of language,

which he identifies as “choppy”(132). While the first reaction might be to blame the editing structure of *Wikipedia* for this failing, “committees rarely write well”, Rosenzweig acknowledges that *Wikipedia* merely follows the tradition set by the 20th century American encyclopedias in their quest to report the facts without strong opinions, resulting in lifeless language (ibid.).

Despite these and other academic and popular criticisms, *Wikipedia* maintains its relevancy as the immediately available and accessible knowledge source for anyone with an Internet connection. Given the active efforts of the *Wikimedia Foundation* towards promoting the development of international editions of the encyclopedia, it is safe to argue that this relevancy would transcend mere ease of access and indicate an offering beyond the capabilities of any previous encyclopedic effort. However, mere popularity and accessibility does not mean *Wikipedia* is a clearly defined and sustainable project. It is evident from numerous studies mentioned in this study and the many public appearances of Jimmy Wales that a project like *Wikipedia* requires constant care and attention of a dedicated community whose maintenance and nurturing is elemental to long term maintenance and sustainability of the world’s most comprehensive encyclopedia. Along with a study of *Wikipedia* itself as an encyclopedia, my analysis in the following chapter will focus on some of the particulars of the community that creates and maintains it.