Wikipedia's Influence on the Evolution of Encyclopedia

Sara Librenjak, student
Department of Information Sciences
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, 10 000 Zagreb, Croatia
sara.librenjak@gmail.com

Zdenko Jecić
Lexicographic Institute Miroslav Krleža
Frankopanska 26, Zagreb, Croatia
zdenko.jecic@lzmk.hr

Damir Boras
Department of Information Sciences
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3, 10 000 Zagreb, Croatia
dboras@ffzg.hr

Summary

This article deals with the users of Wikipedia and their usage and opinions regarding Wikipedia in comparison to printed encyclopedias. A representative sample of high school and university students from Zagreb was surveyed.
It is investigated why students tend to use Wikipedia significantly more than printed encyclopedias, even though they are considered more accurate and reliable, which is important in educational usage.
The results of the survey may serve as guidelines for both users and professionals, i.e. encyclopedists. Various issues with Wikipedia which also need attention in this context, e.g. the development of Croatian Wikipedia, are briefly addressed.
The goal of this work is to broaden the awareness of recent phenomena in encyclopedic, or generally, information science, in order to aid in improvement of encyclopedic products.

Key words: traditional encyclopedia, online encyclopedia, encyclopedics, survey, Wikipedia

Introduction

In recent years we are witnessing substantial growth of digital media usage. What follows from this fact is that in our school or work environment printed media is used less frequently. This migration to the "digital world"[1] has brought various changes to our perception and usage of information and knowledge. Thus, it is a relevant topic in both information science and encylopedics[2].

Shores defines encyclopedics as "the art and science of selecting and disseminating the information most significant to mankind" [4]. Although almost half a century old, this definition stresses the *significance* of information as the criteria for its selection. Since the first (known) encyclopedias in the Roman time, the selection of information has been made by the authors. Traditional encyclopedia is compiled and edited by a group of experts, which has, especially in past times, caused some subjectivity in the decision making process: which are those information that are significant to mankind? This problem is less present since the development of the objective science work in encyclopedics, but it is always difficult to speak in the name of the *whole mankind* when a minority is making this decision.

The issue with traditional encyclopedia which is more difficult to mend is due to objective reason: it is considerably difficult to update information in the paper format. Traditional encyclopedia takes years to develop and print, and reprints and corrections are expensive and time consuming. This can cause recurring errors which are hard to spot and improve, and outdated information which appears in subsequent editions.

Digital format, especially online published, grants significant improvement in this field. Data is easily updated, checked and compared. It allows more editors to work at the data, at the same time, and even the usage of language technologies in spell and grammar checking (we may expect more of this field in the future).

The most prominent example of online encyclopedia is omnipresent Wikipedia[3]. It embodies the good sides of the digital format, as well as some drawbacks.

---

[1] See McLuhan [2] for theory of printed media influence ("Gutenberg galaxy") and early predictions of "electronic age" emergence.

[2] A branch of information science; scientific discipline which deals with principles and practices of assembling an encyclopedia.

[3] Mituzas, one of the Wikipedia's system administrators shared the following metrics about Wikipedia usage in 2008 [3]:
  50,000 http requests per second
  80,000 SQL queries per second
  7 million registered users
  18 million page objects in the English version
  250 million page links
  220 million revisions
  1.5 terabytes of compressed data

We can view the matter of open editing, one of the frequent complaints about Wikipedia, from two perspectives. It can remove the subjectivity and control the mistakes, because the mistakes and inappropriate content will be removed faster. On the other hand, it is prone to vandalism and, more subtly, it is also biased in a different way. At the moment, most of the Wikipedians[4] belong to the (sub)culture of Internet users, and as participants of specific culture, the information selection and presentation might also be culturally biased. Unlike the paper formatted encyclopedia, Wikipedia is up-to-date, but not all topics are updated as promptly as recent news and popular culture, in which Wikipedia resembles a news portal.

This research will try to answer why Wikipedia is so popular among users, in what manner and why are its shortcomings ignored, is it replacing the traditional encyclopedia or is the format of encyclopedia just evolving with the social changes.

The survey: demographics and background

In an online survey[5], a sample of 123 Wikipedia users answered questions about their usage of Wikipedia, stated opinions about its reliability and compared it to a traditional encyclopedia. The sample consisted of high-school (13%) or university students (59%) and employed young people (28%). It represented both genders equally, with 48% of men and 52% of women subjects. Most of the subjects resided in Zagreb (84%), and the others were from large (10%) or smaller cities (6%). Subjects are all Internet users, who use it every day for various purposes. They all had some or considerable experience with Wikipedia. Graphs 1 and 2 summarize general statistics about Internet and Wikipedia.
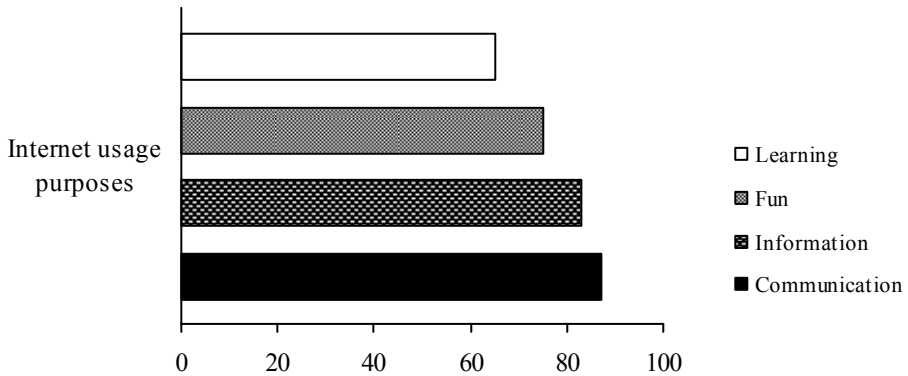
The subjects are from the demographic group which uses Internet the most, and to whom the education plays one of the central roles in life. Since encyclopedia's primary purpose is to inform and educate, the sample consisted mostly of students. Also, since Wikipedia is an Internet phenomena, chosen subjects range from average to experienced Internet users. It is observed that English Wikipedia is used more than Croatian, which will be commented later in the article, but at this point informs us of young people's English proficiency.

The results of the survey are divided in three topics. Firstly, general opinions of Wikipedia are addressed: reasons and manner of usage, reliability and its role in personal education. Secondly, the subjects compared Wikipedia with the traditional encyclopedia. The last topic is the opinion on the development of the Croatian version of Wikipedia, which is connected to the traditional vs. online encyclopedia debate.
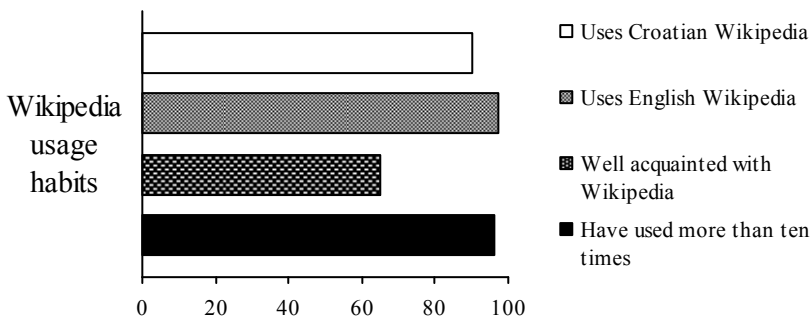
---

[4] Term used to denote people who write and edit articles on the Wikipedia.

[5] The survey was assembled, conducted and processed by the authors on the address http://osnovephp.kset.org/varsara/anketa, which is no longer available online.

Graph 1: Internet usage purposes

Graph 2: Wikipedia usage habits

Opinions on Wikipedia, its usage and role

The survey showed that most subjects use Wikipedia because of the vast information available, and facilitated availability. Users stated that they mostly use the search box and click the links in the article when trying to find some information on the Wikipedia. Average surveyed subject does not use Wikipedia's category system, and most users were satisfied with searching and data retrieval options. Table 1 sums up these answers with statistical data[6].

---

[6] For both questions, more then one answer was accepted. Answers are sorted in ascending order according to the percentage of each answer.

Table 1: Reasons for Wikipedia usage and browsing manners

Why do you use Wikipedia?	
Wide choice of subjects	74%
Ability to find desired information quickly	70%
Free of cost	68%
Available in multiple languages	35%
Open editing	28%
Quality of content	26%
How do you search for the information in Wikipedia?	
Using the search box	81%
Following the links in the articles	51%
Browsing categories	19%
I rarely use Wikipedia for searching or browsing	9%

Wikipedia also seems to play a role in subjects' education. But, according to the answers, it is used more as a simple introduction to a topic, then to broaden knowledge about some previously known facts. 81 percent of the subjects agree that Wikipedia is useful as an introduction to some topic. 74% uses it as an aid in writing a school or university paper (more or less reluctantly), and 52% uses Wikipedia for learning. Although they are using it, that does not mean that the users are considering Wikipedia reliable. 26 percent said they consider Wikipedia absolutely reliable, 3% stated they do not consider it reliable at all, and the majority (71%) answered that Wikipedia is more or less reliable. More specific answers to the reliability issue can be found if we question what (un)reliability entails in the case of Wikipedia, and what the reasons for the lesser quality of article content are.

Table 2 explores the issue of reliability, showing the percentage of subjects who agreed with the statements. The first issue is connected to Croatian version of Wikipedia, and shows a separate problem: less content means less quality. This does not mean that the content – quality ratio is an exponential function, because vast content does not guarantee quality, but considerable lack of content simply causes a specific version of Wikipedia to be unusable. Second issue that is mostly agreed upon is that reliability depends on the article topic. This issue is explored later. More than half of the subjects agree that there should be some content control by experts, which points out the need for a reliable method of quality control. The fact that users re-check the facts from Wikipedia in other sources seconds that conclusion.

Another question dealt with the reasons for the lack of reliability. According to the 35% of the subjects, topics concerned with subjective interest (politics, religion, history or business) are least reliable. Next reason for lower quality of the content are objective difficulties concerning the topic, which would need professional writing (18%). Scientific and technical topics fall in this category. Poor choice of style and article organization are also mentioned as one of the problems (9%).

411

Table 2: Opinions on the Wikipedia's reliability issues

English version of Wikipedia is of higher quality and I use it more frequently then Croatian version	75%
The quality of Wikipedia article depends on its topic	57%
Wikipedia should be edited and checked by professionals	52%
Whenever I use some data from Wikipedia, I check it in other sources	46%
Open editing option leads to vandalism and misinformation	39%
When writing, it is better to leave Wikipedia out of references list, even we use material from it	37%
Open editing option makes Wikipedia more reliable, because misinformation and errors can be corrected quickly	26%
Content control would considerably slow down the content growth, so I don't find that good idea	15%

The article style is one of the points where Wikipedia can be compared with the traditional encyclopedia. It also reflects how the users perceive Wikipedia – as a real encyclopedia, or as a tool for quick and simplified information about various popular subjects. Table 4 shows the answers to question about the current style of Wikipedia's articles, and the style the users think Wikipedia should employ. Most prominent points are the need for the comprehensive style of articles (which is at the moment fulfilled), and the spelling or grammar errors free article. One third of subjects also feel the style should be more like that of a professionally edited encyclopedia, and appropriate for citing. The fact that these qualities are not strongly stressed by users tells us that Wikipedia is not perceived as a counterpart to a traditional encyclopedia, but as an entirely different kind of information source.

Table 3: Style of Wikipedia's articles

How Wikipedia's articles…	are written	should be written
So that anyone could understand the content	93%	83%
Appropriate for citation in papers	15%	38%
Mostly of questionable quality	8%	-
Professional style, as in traditional encyclopedia	11%	32%
Without spelling or grammar errors, a role-model	20%	62%
The style is irrelevant	-	11%

The comparison of Wikipedia and traditional encyclopedia

The opinion that Wikipedia should not be classified as an encyclopedia, or even that Wikipedia in fact deceives its users by calling itself so, is present with some professionals. The former editor-in-chief of *Encyclopedia Britannica* and one of the Wikipedia's critics, Robert McHenry, writes [1] that an average user of encyclopedia has surprisingly low expectations. What he seeks is just a quick answer to his questions. According to McHenry, a more serious user would be satisfied only with the accurate information, but this statement is put to a test in a fast world, overloaded with information. In the survey the subjects were asked in which situation they prefer the traditional encyclopedia, which they consid-

ered more accurate and reliable, to Wikipedia. As table 4 shows, most subjects agreed that they would use a traditional paper encyclopedia when they use the encyclopedic information as a reliable knowledge source, not as a casual information source. It is shown that users tend to turn to traditional encyclopedia when they expect that the information cited will be read and examined by some other party.

Table 4: Situations in which the traditional encyclopedia is preferred

When writing a school/college paper or article	46%
When I need to check Wikipedia information	31%
In any situation	8%
Never	8%
Only if the traditional encyclopedia is new and updated	4%

When asked why they would use one or the other in specific situations, three recurring points emerged amongst various answers. Subjects explained that they do not want or are not allowed to cite a Wikipedia in a more demanding or serious writing work. Once again they stress that Wikipedia is not very useful when they need accuracy and reliability. A subject gives an example of a betting situation: he or she would turn only to a traditional encyclopedia when they need to check the answer to a general knowledge question on which they made a friendly bet. Wikipedia has almost no authority in such situations.

The second point made about preference of a traditional encyclopedia falls in the domain of national encyclopedias. Quite a few subjects stated that they do not use Wikipedia when they need information about anything specifically Croatian: history, geography, famous people and so on. Croatian Wikipedia is poorly developed in comparison to English version, which is understandable and applicable to almost any smaller national Wikipedia. Since, at the moment, only several authors and editors actively work at the Croatian version, the development is slow, and Croatia related content could be richer. Users that were surveyed recognized that, and they tend to avoid Wikipedia when in need for Croatia related information.

Finally, some users stated an interesting usage for the paper encyclopedia which looks at their format as a certain advantage, not just as a setback. Although it is possible to view editing history on a certain Wikipedia article, the interface is less then comprehensive and shows the changes users have made, and not the actual history of knowledge and thought on the subject. Comparative reading of older paper encyclopedia grants us a look into history of knowledge and spirit of those times. This historical dimension is lost in the paperless world.

What about the Croatian version?

As we have seen, the users we surveyed mostly agree that the Croatian version of Wikipedia is underdeveloped, and even though 90% have used it, 75% prefer

the English version. They were asked if they think that Croatian version should be more developed, and why. 68 percent of the subjects think it is important that the Croatian version of Wikipedia develops more in the future. In addition, 41% noted that the size and quality of a national version of Wikipedia speaks about the country (or language) itself, and represents it in the digital world.

In order to be more developed, the group of motivated users/volunteers has to work on the articles. Most of the subjects did not take part in the development of Croatian Wikipedia. 38% said that they never thought about editing or writing Wikipedia articles, and 21% said they are not interested in participating in development of Croatian Wikipedia at all. 27% did write or edit an article (two thirds of that number just edited), and the rest stated that they find it too complicated, technically or content-wise. Although most surveyed users agree that we should develop Croatian Wikipedia, they feel it is a job someone else should do. This fact is in opposition to the free and open approach that Wikipedia promotes, so it is unclear why so little motivation exists. Table 5 shows most important reasons for or against the development of Croatian Wikipedia which were stated in the survey[7]. The main reason stated is to make Wikipedia more accessible to those who do not speak foreign languages, and the reasons connected to Croatian identity or Croatia-related facts follow.

Table 5: Should we develop the Croatian version of Wikipedia?

I find it important	68%
I don't find it important	32%
Yes, for those who do not speak English or other foreign language	52%
Yes, because it represents Croatian identity	26%
Yes, to add more information about Croatia	14%
Yes, because we need Croatian perspective for some events	8%
No, there is no need	11%
No, the English version will always provide more quality and content	5%

Regardless of the Croatian version issue, the points which were made about the difference between traditional encyclopedia and Wikipedia are still valid. Even those subjects who have no problem with reading the English version, and did not need Croatia-related information do not use Wikipedia as an exclusive, reliable source of general encyclopedic information, let alone the specialized information for a certain profession. This does not mean Wikipedia is useless, should be ignored or that its development should stop.

---

[7] More than one reason was acceptable.

Conclusion

The survey showed the distinct boundary in perception in usage of Wikipedia and traditional encyclopedia. On the debate which one should be used, the answer mostly depends on the motivation for the usage. But more importantly, the results of the survey show that Wikipedia is used more than traditional encyclopedia, despite the awareness of its shortcomings. When deciding which encyclopedic work to use, Wikipedia's vast choice of topics, quick searching and open access are shown to be more important than information accuracy. These results incline us to think about migrating our traditional, reliable, encyclopedia to a digital space, and learn from Wikipedia's positive sides. If there was a "traditional encyclopedia" (traditional in sense of professional editing and proof-read, accurate content) which was as accessible and updated as Wikipedia, it would satisfy both casual and demanding users. The process of evolution of encyclopedia perception and usage is visible throughout the results. Thus, it is argued that the traditional values of the encyclopedic science should be reconsidered. For example, a Wikipedia critic may argue that Wikipedia errs in calling itself an encyclopedia; but we can reply that the definition of encyclopedia might have changed, and that encyclopedist must consider this when designing his or her next encyclopedic work.

References

McHenry, Robert. The Faith-Based Encyclopedia. 11/15/04.
    http://www.tcsdaily.com/article.aspx?id=111504A (08/05/09)
McLuhan, Marshall. The Gutenberg Galaxy: The Making of Typographic Man. Toronto: University of Toronto Press, 1962.
Miller, Rich. A Look Inside Wikipedia's Infrastructure. 06/24/08.
    http://www.datacenterknowledge.com/archives/2008/06/24/a-look-inside-wikipedias-infrastructure (08/05/09)
Shores, Louis. Encyclopedics. // *Reference as the Promotion of Free Inquiry,* Littleton, CO: Libraries Unlimited, 1976, p. 162.