**DTU Library**

# Wikipedia in the eyes of its beholders: A systematic review of scholarly research on wikipedia readers and readership

**Okoli, Chitu; Mehdi, Mohamad; Mesgari, Mostafa; Nielsen, Finn Årup; Lanamäki, Arto**

Link back to DTU Orbit

# Wikipedia in the eyes of its beholders:
# A systematic review of scholarly research on Wikipedia readers and readership

**Chitu Okoli**
John Molson School of Business, Concordia University, Montreal, Canada
Chitu.Okoli@concordia.ca

**Mohamad Mehdi**
Computer Science, Concordia University, Montreal, Canada
mo_mehdi@encs.concordia.ca

**Mostafa Mesgari**
John Molson School of Business, Concordia University, Montreal, Canada
mmesgari@jmsb.concordia.ca

**Finn Årup Nielsen**
DTU Compute, Technical University of Denmark, Kongens Lyngby, Denmark
fn@imm.dtu.dk

**Arto Lanamäki**
Department of Information Processing Science, University of Oulu, Oulu, Finland
arto.lanamaki@oulu.fi

## Abstract

Hundreds of scholarly studies have investigated various aspects of the immensely popular Wikipedia. Although a number of literature reviews have provided overviews of this vast body of research, none of them has specifically focused on the readers of Wikipedia and issues concerning its readership. In this systematic literature review, we review 99 studies to synthesize current knowledge regarding the readership of Wikipedia and also provide an analysis of research methods employed. The scholarly research has found that Wikipedia is popular not only for lighter topics such as entertainment, but also for more serious topics such as health information and legal background. Scholars, librarians and students are common users of Wikipedia, and it provides a unique opportunity for educating students in digital

literacy. We conclude with a summary of key findings, implications for researchers, and implications for the Wikipedia community.

**Keywords:** Wikipedia, systematic review, literature review, readers, readership, encyclopedias, knowledge sources, health information, news sources, website ranking, website popularity, credibility, information literacy, students, Web references

# Introduction

For many years, Wikipedia has been the world's most popular reference source on the Web. Since at least 2006, it has been consistently ranked as one of the top ten websites globally (Perez, 2007), with around 500 million unique visitors per month as of February 2014[1]. Not only because of its popularity, but more so because of the influence of such a widely-disseminated information source, an enormous body of scholarly research has developed where scholars have investigated diverse questions to better understand the nature of Wikipedia, an icon of the contemporary Internet age. A number of literature reviews have followed to provide overviews of this large body of research. These reviews have typically focused on Wikipedia contributors (Jullien, 2012; Martin, 2010; Yasseri & Kertész, 2013), on the content of Wikipedia articles (Jullien, 2012; Martin, 2010), or on scholars using Wikipedia as a corpus for text-based research (Medelyan, Milne, Legg, & Witten, 2009), but none of these has specifically focused on the readers of Wikipedia and various aspects of its readership. Given that the readers are by far the most numerous part of the Wikipedia community—and we might argue its very raison d'être—it is necessary to specifically review scholarly findings on this important aspect of Wikipedia.

Online participation research has often characterized Internet readers as "lurkers", even free-riding parasites who benefit from others' contributions but who contribute little themselves (Kollock & Smith, 1996; Nonnecke, Andrews, & Preece, 2006; Preece, Nonnecke, & Andrews, 2004). Often, when readers are considered valuable, it is mainly for their potential to be converted to contributors (Rafaeli, Ravid, & Soroka, 2004; Schneider, von Krogh, & Jäger, 2013).

In contrast, this present review adopts the perspective of those who take a decidedly favorable perspective on non-contributory Wikipedia readership (Antin & Cheshire, 2010). In general, non-contributing readers constitute 90% or more of any given discussion forum, online community, or social website (Arthur, 2006; Li & Bernoff, 2011); a 2011 Readership Survey showed that the number for Wikipedia was 94%[2]. Although often called a "participation inequality" (J. Nielsen, 2006), such statistics indicate that reading constitutes the norm, whereas contribution is the anomaly (albeit a necessary one). Although we recognize the crucial importance of Wikipedia contribution (Okoli, Mehdi, Mesgari, Nielsen, & Lanamäki, 2012, sec. Participation), the breadth of studies we review in this present article offer a rich variety of perspectives of Wikipedia readers and their reading habits, demonstrating that non-contributory readership is not an illness in need of a cure; such readership is valuable in itself.

Specifically, this review offers two major contributions. First, we synthesize current knowledge regarding the readership of Wikipedia, spanning a broad range of research questions. Beyond merely recognizing that Wikipedia is popular, what specific categories of information are more popular than

---

[1] http://reportcard.wmflabs.org

[2] http://meta.wikimedia.org/wiki/Research:Wikipedia_Readership_Survey_2011/Results

others? Is Wikipedia used mainly for information on leisure topics, or is it also used for serious purposes? It is widely assumed that scholars and academics are generally uncomfortable about Wikipedia; how true is this perception? Is Wikipedia helpful for students or does it harm their education? Although Wikipedia is run by a not-for-profit corporation, does it have any commercial value or usefulness? Many of these important questions have been answered by scholarly research, though some only preliminarily. In this review we will examine how current knowledge has come up with answers on such questions examining different aspects of Wikipedia readership.

The second major contribution of this review is, besides synthesizing current knowledge, to provide an overview of research methods for Wikipedia readership for the benefit of both new and experienced researchers of Wikipedia. Thus, we analyze the studies in detail and compile various characteristics of the studies of particular interest to researchers. For instance, we note the various research designs, data collection approaches, and units of analysis that Wikipedia readership research has employed. This analysis would help researchers quickly determine the various research approaches that have been adopted, which would be helpful in guiding their research decisions.

This review of Wikipedia readership is part of a much larger review of scholarly research on all topics related to Wikipedia. In an overview paper of the entire project, we described our systematic review methodology in detail (Okoli et al., 2012), in which we identified over 500 scholarly studies on Wikipedia published in English using the methodology specified by Okoli and Schabram (2010). In brief, we mainly focused on identifying peer-reviewed journal publications and doctoral theses and tried to be as exhaustive as possible in identifying these. However, we did include over 100 influential and important conference papers. We had to restrict our search time frame to June 2011, after which time the Wikimedia Foundation (WMF) launched their Wikimedia Research Newsletter[3], which summarizes current scholarly research on Wikipedia and other WMF projects. We analyzed and parsed research details from 476 studies identified from the systematic search. These research details are published on an interactive website which permits visitors to conduct various analyses on them (WikiLit: http://wikilit.referata.com). In addition to the systematic search, we included many other scholarly studies that did not match the strict systematic criteria, but which we found relevant and worthy of inclusion, giving a total of over 500 studies.

This present review covers the subset of the larger review focused on Wikipedia readership. Although all Wikipedia contributors are of course also readers, we restrict our focus here to topics that concern reading Wikipedia without substantial treatment of contribution. We cover issues related to Wikipedia contributors and participation in depth in a separate review (Okoli et al., 2012, sec. Participation).

This review is organized as follows: after this introduction, we organize all the identified Wikipedia readership studies by topic, summarize them and comment on them. Next, we describe various pertinent research trends as analyzed from data on the WikiLit website. We then conclude this review with a summary of key findings, implications for researchers and implications for the Wikipedia community.

## Findings from Scholarly Research on Wikipedia Readership

Table 1 displays the topic categories of studies in our sample, with the number of studies in each category. Our categorizations have been determined by literary warrant; that is, we carefully noted the topic of each study identified, and then grouped them into topic categories and subcategories such that each category had between three and twenty studies. Note that the numbers of articles in the leaf (terminal) categories do not add up to any meaningful total, since most articles cover more than one topic

---

[3] http://meta.wikimedia.org/wiki/Research:Newsletter

category and are thus counted multiple times. The WikiLit website has 91 studies systematically identified and analyzed in detail. On the website, we report methodological details for each of these studies that we are not able to include in this present article. In addition, as we explain in our detailed methodology (Okoli et al., 2012), we identified another 8 studies that we also summarize and discuss in this review, for a total of 99 studies of Wikipedia readership. As we describe in our detailed methodology paper, all the coauthors were involved in several rounds of assigning and verifying the topic categories for each study.

The major categories here cover studies about Wikipedia's ranking and popularity compared to other knowledge sources; the use of Wikipedia as a general source of knowledge on the Internet; various topics related to students as readers of Wikipedia;  the extent to which Wikipedia readers consider it credible; software tools targeted to helping Wikipedia readers; and commercial aspects of Wikipedia content. In each topic section, we summarize each study that treated the topic. At the end of each section, we make some general comments about all the studies on the topic. Later, in the discussion section of this review, we highlight key findings from all the studies.

**Table 1. Categorization of topics of Wikipedia research in WikiLit website**

```
► Readership (91*)
  - Ranking and popularity (12)
  ► Knowledge source (31)
    - News source (4)
    - Health information source (11)
    - Knowledge source for scholars and librarians (16)
  ► Student readership (34)
    - Student information literacy (13)
    - Domain-specific student readership (12)
    - Cross-domain student readership (17)
  - Reader perceptions of credibility (21)
  ► Software for readership (12)
    - Computational estimation of trustworthiness (9)
    - Reading support (3)
  - Commercial aspects (10)
```

*The numbers in parentheses are the numbers of distinct studies in each subtopic. Since a study might cover multiple topics, these numbers are not additive.*

**Ranking and Popularity**

Since 2007, Wikipedia has been consistently ranked one of the world's top ten websites; it is the world's top reference website. However, beyond examining mere popularity statistics, various scholarly studies have investigated more nuanced aspects of Wikipedia's Web rankings. This category of 12 studies (13%) includes studies that compared the use of Wikipedia with other knowledge sources for getting information, as well as studies that investigated the popularity of topics within Wikipedia. These studies have consistently confirmed that Wikipedia is a premier source of knowledge on the Internet.

Some studies compared Wikipedia's ranking with that of other important websites. Höchstötter and Lewandowski (2009) compared search results of four major search engines: Google, Yahoo, Live.com and Ask. They found that Wikipedia is the most frequently represented website in all search engines. However, there are some differences in how different search engines rank Wikipedia pages: "Yahoo and MSN place the most Wikipedia results on their results pages. Google boost Wikipedia result mostly on

first position but shows less Wikipedia links in total [*sic*]" (2009, p. 1810). DiStaso and Messner (2010) examined the rankings of Wikipedia articles of ten companies over four years on three search engines. In 2006, the Wikipedia links for these companies were in the top 20 for all three search engines; in 2010, these links had all risen to the top ten, which led to the conclusion that Wikipedia articles about corporations are quite important. Lewandowski and Spree (2011) noted that Wikipedia results shown on search engines are quite dependent on the quality of articles.

In a more focused study, Bar-Ilan (2006) studied a case of "Google-bombing," where the top results to the search keyword "Jew" yield the Wikipedia article and an anti-Semitic website. She observed that the Google ranking is primarily due, not to pages that actually discuss these two websites, but rather to links from discussions and blogs purposely inserted to influence search engine rankings.

Four studies examined what is popular on Wikipedia. Ratkiewicz et al. (2010) provided a quantitative analysis of the dynamics of online popularity of Wikipedia content. They found that the dynamics of popularity are characterized by "bursts, displaying characteristic features of critical systems such as fat-tailed distributions of magnitude and inter-event time" (2010, p. 1). Spoerri (2007a) examined which were the most popular articles and topics on Wikipedia. He found that over half of the most-visited pages are related to entertainment and sexuality, and that popularity of Wikipedia pages is related to search behavior on the Web. He further found that search engines—especially Google—fuel Wikipedia's growth, and thus shape what is popular on Wikipedia. He also examined the 100 most visited Wikipedia articles for five consecutive months, finding that 40% of these—mostly related to sexuality and entertainment—were highly visited in all five months, and 25% were highly visited only in a single month (2007b). Waller (2011) investigated the search queries that directed Australians to Wikipedia pages, and found that they search more for lighter topics such as entertainment rather than for more serious information.

In a study not strictly related to comparative ranking, Koolen et al. (2009) found that high-frequency Web search queries often directly relate to Wikipedia pages. Within a large sample of web queries, 38% exactly matched the title of a Wikipedia page. The content and context of the matched Wikipedia page could then be used to expand the query. Wikipedia pages can also form an intermediary between a user query and a collection of books being searched.

Some articles compare Wikipedia's ranking with other sources of health information; we discuss these in the topic "Health Information Source" (Johnson, Chen, Eng, Makary, & Fishman, 2008; Laurent & Vickers, 2009; Mühlhauser & Oser, 2008). In addition, another article that dealt with issues related to Wikipedia's ranking and popularity is described in the commercial aspects section of this review (Langlois & Elmer, 2009).

We believe that because of Wikipedia's role as the world leading reference source, scholarly studies on finer aspects of Wikipedia's popularity are important. Existing studies have found that Wikipedia's popularity is very much driven by search engines—which are in turn driven by user clicks of page links across the Web. The most commonly consulted pages were related to entertainment, which corresponds to the nature of other highly ranked websites (e.g. Facebook and Twitter for social networking and YouTube for videos). Thus, we suggest that it would be valuable to investigate the popularity of webpages not so much by website as by genre, as the leading websites necessarily cut across various genres. Nonetheless, Wikipedia's function as a general encyclopedia assures that it has an ample representation of all subjects of human interest, and thus will likely remain a leading source of Web information for the foreseeable future.

**Knowledge Source**

In this section, we discuss the use of Wikipedia as a source of various kinds of knowledge, including current news, health information, as a resource for scholars and librarians, and for other information

purposes. This covers 31 peer-reviewed studies (34%) as one of the main categories representing scholarship on Wikipedia readership.

*News Source*

Wikipedia is a popular source of background information on current events; in fact, it is sometimes linked to from major news websites. Moreover, because it is freely editable, it is often used as a public source of reporting breaking news. However, because of its "No Original Research" policy[4], Wikipedia only permits posting of news that has already been published in other news outlets that can then be cited in Wikipedia articles. In fact, related to these reasons, in late 2004 the Wikimedia Foundation launched a sister project, Wikinews, dedicated to collaborative reporting of current events including original news reports, unlike Wikipedia (Wikipedia contributors, 2013). Nonetheless, Wikipedia itself, described as "the largest form of participatory journalism to date" (Lih, 2004, p. 1), remains a popular source of news (Messner & South, 2010).

In one of the earliest academic studies of Wikipedia, Lih (2004) gave an introductory sketch of the then three-year-old endeavor. He analyzed Wikipedia articles cited in the news in a thirteen-month period to compare their quality before and after citation in the press. He used computational quality measures based on the numbers of edits and of contributors, and his analysis showed that increases in his quality measure were correlated with citation in the press.

Wikipedia was used as one of many sources of documentation on details of the delayed response in 2005 to Hurricane Katrina in the United States (Chua, Kaynak, & Foo, 2007). Thelwall and Stuart (2007) found that "Web 2.0 resources such as Wikinews, the Wikipedia, and the Flickr picture sharing site" (2007, p. 523) were important secondary sources of information provision and sharing in the event of natural disasters and similar crises. Nonetheless, traditional mass media remained the predominant sources of information.

Some studies have examined how newspapers frame Wikipedia and use it as a source. Shaw (2008) reported that *Philadelphia Inquirer* instructed journalists never to use Wikipedia "to verify facts or to augment information in a story," and that one reporter complained: "there is no way for me to verify the information without fact-checking, in which case it isn't really saving me any time." Some other news organizations, such as *Los Angeles Times*, did occasionally permit citation of Wikipedia as a source. Messner and South (2010) reported that although newspapers had not referenced Wikipedia very much in the past, their reliance on this source had recently been increasing, and they tended to present it as a generally accurate source.

Although these studies indicate that Wikipedia is indeed included in the collection of resources that journalists draw from, their attitude towards it is generally more cautious (Shaw, 2008) than that of scholars and librarians, as we describe elsewhere in this review. We suggest that a likely reason for this caution is that Wikipedia articles about current events witness an unusually high pace of edits, including correction of errors. In fact, articles about current events are often flagged with a special template to warn about the unstable (and perhaps inaccurate) nature of the page contents[5]. Thus, scholars' general esteem of Wikipedia might apply to more "ripened" articles, whereas journalists might be more cautious about those currently in flux.

---

[4] http://en.wikipedia.org/wiki/Wikipedia:No_original_research

[5] http://en.wikipedia.org/wiki/Template:Current

*Health Information Source*

Many articles discussed the use of Wikipedia as a source of health information for the general public as well as for health professionals. This is distinct from its use by medical students, which we cover in Domain-Specific Student Readership. The majority of these articles critically examined the factual accuracy of Wikipedia's health information, though some other articles examined other aspects of health information.

A popular area of investigation has been how Wikipedia compares with other online sources of healthcare information. Since this involves not only popularity, but also the responsible provision of information that could affect people's health, accuracy was of primary concern in these comparisons. However, we do not attempt here a comprehensive review of rigorous studies of Wikipedia's accuracy for health information—we discuss this aspect in the Reliability section of a separate review (Mesgari, Okoli, Mehdi, Nielsen, & Lanamäki, 2014). Our focus here is on studies that have specifically considered Wikipedia as a source for public health information.

Probably the best general overview thus far published on this topic was coauthored by 19 members of the WikiProject Medicine[6], mainly consisting of medical doctors (Heilman et al., 2011). They reviewed the literature in this area, finding that Wikipedia is an extremely popular health information source; Wikipedia's up-to-datedness is a major strength; its variable article quality is a major weakness for health information; health articles contain few errors but are often incomplete; and that among the multitude of medical wikis, Wikipedia has the best potential for providing a unified platform of public dissemination of health information. They used the article as an opportunity to call on medical professionals to contribute in increasing the quality of Wikipedia's health information. In the rest of this section, we discuss various specific studies on this general topic.

In a study on the efficiency of Web resources for identifying medical information for clinical questions, Wikipedia failed to give the desired answer in around one third of the cases, whereas Web search engines, especially Google, were much more effective. However, Wikipedia was more efficient than medical sites such as UpToDate and eMedicine in terms of failed searches and number of links visited, and it proved to be the most frequent "end site" that provided the ultimate answer from a Google search (Johnson et al., 2008). Mühlhauser and Oser (2008) found the German Wikipedia comparable in quality to the websites of two major German statutory health insurance providers for content and presentation of patient information. However, in their assessment based on the standards of evidence-based medicine, none of the three sources proved satisfactory. Yermilov et al. (2008) compared the quality of Internet sources of surgery information. Wikipedia's average information quality was less than that of professional societies, government and hospital sites, but it was higher than the average quality of universities and manufacturer or pharmaceutical sites.

Using search engine optimization techniques, Laurent and Vickers (2009) investigated the Google ranking of the English Wikipedia for health topics. Queries based on 1,726 keywords from an index of the American MedlinePlus, 966 keywords from a NHS Direct Online index and 1,173 keywords from the United States National Organization of Rare Diseases, they compared Wikipedia to .gov domains, MedlinePlus, Medscape, NHS Direct Online and a number of other domains. They found the English Wikipedia as the website with the most top rankings. Using data from stats.grok.se for June and January 2008, they also examined health-related topics with probable seasonal effects, such as frostbite, hypothermia, hyperthermia and sunburn. They found a clear effect in the page views. They also analyzed the page view statistics of three articles describing melamine, salmonella and ricin. These examples were associated with official health alerts in 2008, and page view statistics showed a marked increase correlating with the timing of announcements.

---

[6] http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine

Beyond investigating the quality of Wikipedia's health information, two studies have examined the practical use of Wikipedia in awareness of existing concerns. Hughes et al. (2009) examined how Web 2.0 tools like Wikipedia were being used in clinical contexts. They found that although medical practitioners were aware of credibility deficiencies of Wikipedia, they employed different strategies to cope with the risk while meeting their background information needs. Younger (2010) studied the potential of wikis and Wikipedia as an information source for nurses. She argued that although it is not likely for Wikipedia to replace the traditional printed valid information sources, it would be a promising starting point for nurses in searching for evidence-based patient-related information.

Two studies focused on how Wikipedia's wiki technology facilitates dialogue on public health matters. Hickerson and Thompson (2009) examined Wikipedia and WikiHealth as case studies for promoting health information online, particularly these wikis' ability to engage the target public in productive dialogue. Cimini (2010) investigated the impact of online dialogues on the meaning of Down's syndrome and the extent to which these dialogues can change the way that "disability" is theorized.

In summary, most studies on this topic have found Wikipedia useful for general health information, especially because of its dialogic wiki characteristics and its broad popularity. Nonetheless, unsurprisingly, most medical scholars do not consider it a reliable source for healthcare decisions. Wikipedia, for its part, explicitly disclaims giving medical advice[7], and beyond attempting to provide generally useful information, has no goal of being a primary source for medical decisions. In contrast, Medpedia (http://www.medpedia.com) was launched in 2009 to meet the need for a wiki-driven medical knowledge base. It only permits editing by certified physicians or biomedical researchers, though it accepts suggestions from the general public. As of 2013, it was labelled as "beta"; only time will tell if it will eventually achieve its goal of becoming the Web's primary source of authoritative medical information, replacing commercial sources.

*Knowledge Source by Scholars and Librarians*

Perhaps surprisingly, quite a body of studies has been built up that investigate Wikipedia's use for their own research purposes by scholars and librarians, as distinct from their role as student educators. While it is true that Wikipedia's most vociferous critics hail from these ranks, a very large number of academicians in fact have quite positive, if nuanced, perceptions of Wikipedia's value.

**Source for Scholarly Research.** A Wikimedia Foundation survey has found researchers to be generally quite positive towards Wikipedia: Over 90% of 1743 self-selected respondents were "very favorable" or "somewhat favorable" (Moeller, 2009). Among Public Library of Science (PLoS) authors, the result was 96%. To the question, "Would you be in favor of efforts to invite scientists to add or improve Wikipedia articles?", 68% answered, "Yes, on a large scale." Such results are very positive for Wikipedia, but may be biased due to the self-selection of respondents and because the publisher web site with initial reference to the survey was open access—the responding researchers were likely quite favorable to open information resources in general.

Three studies have surveyed scholars to understand their usage and attitudes towards Wikipedia. Dooley (2010) surveyed 105 university faculty members and found that 54.4% considered Wikipedia to be moderately or very credible, 26.6% considered it having some credibility, and 20% considered that it had "no credibility." 45 of 105 respondents said they used Wikipedia moderately to frequently in their teaching or research, 40 only occasionally, and 20 said they never used Wikipedia for teaching or research. Despite controversies with student citation of Wikipedia, many professors and other researchers do in fact cite Wikipedia. Concerning citations of Wikipedia, Dooley examined 250 research reports published in 2009 and early 2010 from the Academic OneFile electronic database that contained

---

[7] http://en.wikipedia.org/wiki/Wikipedia:Medical_disclaimer

"Wikipedia" in their text. She found that 27 of the papers featured Wikipedia as the main topic and 62 had brief mentions of Wikipedia. 249 of these papers cited Wikipedia as a source.

Chen (2010) found that although academics extensively use online information resources and databases for teaching and research purposes, they are often concerned about credibility of Wikipedia content. Those who use Wikipedia are more likely to also be Wikipedia contributors. Eijkman (2010) observed how academics cautiously use Wikipedia along with other sources of knowledge. He found that although they are aware that it disrupts their traditional power as knowledge providers, academics are not generally as antagonist towards Wikipedia as is commonly assumed.

Page (2010) compared the current state of Wikipedia's documentation of biological species with E. O. Wilson's vision of an "encyclopedia of life." He contended that in its dominance of search result rankings, contributor size, and potential linkage to other data, Wikipedia is currently the closest achievement of this vision. Curiously, he made no mention whatsoever of WikiSpecies, the Wikimedia Foundation's project whose goal is more closely aligned to that vision.

In his article on Israel-Lebanon conflict on Wikipedia and Wikinews, Hardy (2007) argued that while "Wikipedia is not a threat to the peer reviewed publications of academia, it is definitely a competitor, not only by dint of the number of people who consult it, but the quality of some of the articles in their own right" (2007, p. 22). Other scholars were more cautious in their appraisal of Wikipedia, and pointed out some challenges that its use presents scholars. In her paper about conducting art history research, Chen (2009) recommended caution in using Wikipedia: "the content of the Wikipedia article can be used as tips for possible approaches to this object, but not as a source for the actual paper" (2009, p. 123). Knapp (2008) discussed the challenges that arise from citing amorphous web sources such as Wikipedia, where the source materials are constantly changing. She explored possible resolutions, such as a "virtual bookshelf" which includes electronic attachments of source materials along with a published work.

In a separate article (Okoli et al., 2012), we discussed in more depth several studies that go beyond scholars reading and using Wikipedia to examine direct contribution to Wikipedia by scholars. Here we will more briefly comment on some such studies, emphasizing their implications for scholarly use of scholar-contributed Wikipedia content. Davis et al. (2010) noted that Wikipedia could function as a central hub of scholarly information for industrial ecology. They also illustrated how DBpedia, an online service with information from Wikipedia restructured in database format (Bizer et al., 2009), can be used to query Wikipedia with industrial ecology research questions. Huss et al. (2010) reported on Portal:Gene Wiki, an organized grouping of gene-related information on the English Wikipedia.

In addition to the studies on scholarly use of Wikipedia that we have discussed here, we later discuss Kubiszewski et al. (2011) in the section on Reader Perceptions of Credibility, since they empirically examined scholarly use from that perspective.

**Source for Librarians.** Several studies discussed issues particularly pertinent to librarians. In our WikiLit website, we identify all library-related research in the "Library science" domain[8]—we identified over 30 such studies. However, we normally assigned such articles in other topic categories if they are more specific about the focus of the article, but here we discuss those that are essentially focused on librarianship. These studies unanimously called on librarians to consider Wikipedia a positive phenomenon, and to take advantage of it.

Some studies considered the use of Wikipedia inevitable, and so called on librarians to thus embrace it and even use it as an opportunity to teach information literacy. Choolhun (2009) documented that Wikipedia is increasingly being used as the first source for legal information inquiries by lawyers and law students. She thus called for increased engagement with Web 2.0 use by legal librarians. Gunnels and

---

[8] http://wikilit.referata.com/wiki/Category:Library_science

Sisson (2009) cautioned against avoiding Wikipedia and other Web 2.0 tools for research, but rather urged teaching students to be critical about the information found on these sources and how to validate them through reliable sources. East (2010) examined the future of subject encyclopedias in the age of Wikipedia. He concluded that if librarians want to keep the use of subject encyclopedias alive, they must make them available online, easily searchable, and well cross-linked with other resources.

Three studies went further to regard Wikipedia as a unique opportunity to promote libraries and librarianship, and to spearhead their relevance to the forefront of the information age. Belden (2008) presented a case of how a university library was able to gain "dramatic increases in Web usage and reference requests by harnessing the power of social networks such as Wikipedia and MySpace" (2008, p. 99). She explained that sites such as Wikipedia "provide the tools to allow dynamic, interactive means of sharing information and helping connect the dots," and that the skills needed in these activities are "the very abilities that librarians and scholars hope to inculcate in our educational endeavors" (2008, p. 110). Luyt et al. (2010) found that many librarians view Wikipedia as an opportunity for the profession rather than a threat. They considered this finding as an opportunity to connect with non-Western users and content, and also as an opportunity for librarians to forge a leading role in the emerging information society. Jacobs (2009) commented on Hahn's (2009) study of student use of Wikipedia on iPods (described in the section on Cross-Domain Student Readership). She examined how new information technologies like Wikipedia influence librarians and academic libraries, suggesting that librarians engage in and promote the new technologies to the library world.

**Source for Legal Studies.** In addition to general scholarly studies, three legal studies examined the judiciary use of Wikipedia and discussed the controversy of using Wikipedia as an authority. Stoddard traced an increasing trend of usage over time (Stoddard, 2009). Breinholt (2008) classified the different uses of Wikipedia into four categories:

1. Wikipedia as a dictionary. For example, Wikipedia was used to answer what "candy striper" means.
2. Wikipedia as a source of evidence. In the most perilous uses of Wikipedia, judges rely on Wikipedia for evidence (for example, to determine whether or not the United States Interstate 20 passes through California).
3. Wikipedia as a rhetorical tool. This involved innocuous uses, such as for literary allusions.
4. Judiciary commentary about Wikipedia. One case involved a judge cautioning against citing Wikipedia in an appellant brief.

Peoples (2009) examined the quality of the Wikipedia articles cited by American judicial opinions. He found that the "majority of citations to Wikipedia entries in cases were not significant to the case but were merely collateral references" (2009, p. 27). Peoples proposed a number of best practices for citing Wikipedia. He presented some cases and scenarios where Wikipedia should not be cited and others where citing Wikipedia could be deemed appropriate.

Overall, the studies about the use of Wikipedia as a source for scholars and for librarians show that, contrary to widespread assumptions, these professional creators and disseminators of knowledge more often consider Wikipedia a positive phenomenon than otherwise. Scholars and librarians not only widely use Wikipedia themselves, but often even promote it as a valuable knowledge resource. While such a positive attitude is by no means universal, it indicates that even the traditional gatekeepers of knowledge have come to terms with the significance and value of a publicly-generated knowledge source in the dawn of the 21st century.

**Student Readership**

34 studies (37%) investigated how students use Wikipedia, both as a general source of information and in student projects where they were assigned work that explicitly involved reading Wikipedia articles. These studies treated students in secondary school, undergraduate institutions and in post-graduate education.

Articles we classify here mainly involve student information literacy in critically reading and using information from Wikipedia articles. In contrast, we classify articles concerning projects where students are assigned to contribute to and develop Wikipedia articles as Student Contribution, which we discuss in a separate review (Okoli et al., 2012). Since all Student Contribution activities necessarily require students to read Wikipedia, we normally discuss here only those articles that have no substantial component involving the students contributing content to Wikipedia. We categorize the following kinds of student readership articles: those that dealt with general matters of student literacy; those that treated students and Wikipedia articles across various domains of knowledge or field of study; and those that are restricted to students in a specific domain.

### *Student Information Literacy*

A popular research stream has involved studies of how Wikipedia relates to student information literacy. These studies unanimously called on teachers and professors to embrace rather than ban Wikipedia, urging them to seize the opportunity to educate students in information literacy skills needed for the 21st century.

Some studies discussed information literacy in general. Rand (2010) argued that students can learn critical thinking skills using Wikipedia. Jennings (2008) argued the necessity of information literacy skills for 21st century students to become lifelong learners in using all information resources. He highlighted the importance of Wikipedia, as it facilitates teaching and learning such skills. Gunnels (2007) proposed using Wikipedia as a starting point towards a new way of teaching information literacy skills which enhances the quality of both the users and creators of information resources. Judd and Kennedy (2011) looked into how medical students use Wikipedia and other online resources to acquire their needed information. They concluded that higher emphasis on information literacy skills training is required to make sure students are able to locate and use the best available information.

Three studies described teaching experiences engaging Wikipedia for information literacy. Harouni (2009) reported using Wikipedia in a literacy class to teach students critical reading skills. After the lessons, students could clearly articulate their reference choices, and they were able to discern and use more comprehensive and unbiased Wikipedia articles. Patch (2010) argued that as students are already using Wikipedia, writing teachers should follow suit and incorporate Wikipedia into their teaching. She described some of her experiences with students and Wikipedia, and argued that "many students are 'underprepared' to consume and use online texts responsibly" (2010, p. 282). She concluded that by employing Wikipedia, students can "have an easier time making the leap to higher-level inquiry and responsible scholarship" (2010, p. 282). In sharing their respective experiences in teaching about Wikipedia in computer science and anthropology courses, Aycock and Aycock (2008) proposed using Wikipedia to teach students not only about the use and interpretation of information resources, but also about management of rapidly changing collaborative information resources.

Three studies featured in-depth investigation of how students handle Wikipedia information. Calkins and Kelley (2009) examined history students' perception of Wikipedia credibility and collaborative work. Although the students were aware of factual errors in Wikipedia, they nonetheless believed that it was getting better as more people contribute and correct errors; they mostly favoured accuracy and collaboration on Wikipedia. Sundin and Francke (2009) investigated how secondary school students negotiate the credibility of information in their learning process. Although these students used Wikipedia information, they were uncertain about its credibility because they employed traditional methods for credibility assessment based on authorship and origin, neither of which is clear in Wikipedia. Sundin and Francke suggested that the students need to update their credibility assessment approaches. We describe a subsequent related study by Francke et al. (2011) in the section on Reader Perceptions of Credibility.

Whereas most of the articles we discussed here involve helping students to read Wikipedia critically, Chandler-Olcott (2009) argued that constructive writing is a crucial aspect of digital literacy, and so teachers should encourage students to write in digital collaborative environments such as Wikipedia.

Two other articles related to student information literacy more directly discuss how librarians should leverage Wikipedia as a resource for students, and so we describe them in the section of this review of Wikipedia as a Source for Librarians (Chandler-Olcott, 2009; Choolhun, 2009; Gunnels & Sisson, 2009).

As we have noted, the studies on this topic all call for engaging students with Wikipedia as an essential part of their digital literacy education. The Internet has unleashed a flood of accessible information as never before known in human history; Wikipedia is only one manifestation of this. Because of its explicit attempt to provide neutral information, Wikipedia is an excellent context for educating young people in the critical skills required to make sense of the information deluge.

### *Domain-Specific and Cross-Domain Student Readership*

A large body of studies focused on student use of Wikipedia articles, either focusing on students within a specific knowledge domain, or on students from various disciplines.

**Humanities and social sciences.** One of the perennially controversial questions about Wikipedia is whether or not it should be allowed as a citation, especially for student work. In early 2007, the history department at Middlebury College decided to hold students responsible for using Wikipedia as a source after a batch of students had used erroneous information on Wikipedia about certain topics in the history of Japan (Waters, 2007). Media reports implied that the department of Neil Waters, the teacher of the class involved, "was at war with Wikipedia itself." However, Waters himself actually told students "that Wikipedia is a fine place to search for a paper topic or begin the research process." The department adopted the following policy:

> Whereas Wikipedia is extraordinarily convenient and, for some general purposes, extremely useful, it nonetheless suffers inevitably from inaccuracies deriving in large measure from its unique manner of compilation. … Students are responsible for the accuracy of information they provide, and they cannot point to Wikipedia or any similar source that may appear in the future to escape the consequences of errors. (Read, 2007)

This policy is actually in line with the opinion of the Wikimedia Foundation. However, Jimmy Wales, founder of Wikipedia, later said that he saw no problem in younger students using Wikipedia as a reference, and that it should be used as a stepping stone to other sources (Coleman, 2007).

Some researchers examined how and why journalism and mass communication university students use Wikipedia. Lim (2009) affirmed many other research findings that although students commonly used Wikipedia for finding background information with acceptable qualities, they were quite aware of its quality concerns and did not use it blindly. However, they did not verify the information on Wikipedia, but rather used its sources and links to get further information. Lim and Kwon (2010) compared student usage of Wikipedia by gender. They found that while male students used Wikipedia more frequently and had a positive attitude towards it, "female students displayed more cautious or conservative attitudes, emotions, and behaviors."

**Healthcare.** Many studies considered usage by medical students, who are being trained to make life-and-death decisions based on their evaluation of information. Judd and Kennedy (2010, 2011) studied Australian biomedical students' on-campus use of internet web sites. Wikipedia's use increased "from only 2% of sessions in 2005 to 16% in 2008 and 2009" (2010, p. 1568). They concluded, among other issues, that students "are increasingly reliant on generalist information retrieval tools, particularly Google and Wikipedia, to support their learning activities" (2010, p. 1570). Fiore (2011) found that 47% of 186 medical students who recently completed psychiatric clinical clerkship used Wikipedia as one of the primary sources for preparing for psychiatry exams. Question books (88%) and the peer-reviewed website Up-to-Date (59%) were more frequently used, but textbooks (10%) less used. Among the students using Wikipedia, 84% also used question books.

Lavsa et al. (2011) evaluated Wikipedia's appropriateness for pharmacy students. They found that all the examined Wikipedia articles provided incomplete or inaccurate drug information, and thus recommended against its use. In contrast, Haigh (2010) concluded that Wikipedia is an appropriate resource for use of nursing students as a large sample of health articles cited reputable sources for their information.

**Other sciences and mathematics.** Korosec et al. (2010) found that students of chemical thermodynamics preferred the German Wikipedia over the chemistry encyclopedia Rompp Online because they considered Wikipedia more comprehensive and readable. They concluded that while both resources were good for initiating research, students should learn how to use both peer-reviewed and non-peer-reviewed material in their learning. Wedemeyer et al. (2008) asked students to evaluate Wikipedia biochemistry articles. One third responded that they never used Wikipedia. Among the remaining two thirds, 12% used Wikipedia as their primary source and 31% used their textbook and Wikipedia equally. The remaining 57% used Wikipedia only as a supplement. The majority of the students preferred Wikipedia to the textbook. Jancarik and Jancarikova (2010) examined the appropriateness of Wikipedia material for preparing teachers of mathematics and biology in Czech. They observed that whereas the English Wikipedia properly covered the topics with highly detailed articles, the Czech Wikipedia, whose scientific topics mostly consisted of English translations, was less detailed and comprehensive, and so inadequate as an e-learning resource. Schweitzer (2008) examined the coverage of psychology-related topics on Wikipedia. These were not only well covered, but the articles also displayed on top of the major search engines. Students were found to use Wikipedia for personal and school-related activities, but generally not for their academic citations.

**Cross-domain student readership.** Many articles treated student readership in general, regardless of their domain of knowledge or field of study. Some of these generally investigated how students use Wikipedia. Tann and Sanderson (2009) examined the web-based information-seeking practices of university students. They found that many queries that have been previously considered informational by past research have since taken on a more navigational nature. Moreover, IMDb and Wikipedia have both accumulated a sufficient level of information to address the users' information needs. Head (2007) examined how students used online and offline resources in their research inquiries. She found that they used hybrids of both types of sources, though they used Wikipedia less because of their concerns of its reliability. Head and Eisenberg (2010) later discovered that when university students use Wikipedia, they are generally aware of its limitations in credibility and depth. It is more often used in the initial stages of research to obtain background information, and then is complemented with scholarly resources. Students mainly appreciate Wikipedia for its coverage, currency, comprehensibility, and convenience. Luyt et al. (2008) interviewed young people concerning their perception and use of Wikipedia. They found that Wikipedia played only a minor role in the lives of their interviewees, but these young people were quite aware of its drawbacks when they did use it. The researchers thus concluded that common concerns about Wikipedia's negative effects on young people are exaggerated. However, their study is hard to generalize, as only 15 subjects were interviewed, apparently all in Singapore.

Maehre (2009) explored various pedagogical principles to encourage instructors to allow the use of Wikipedia in their students' projects. He argued in favor of producing and engaging in information creation. Moreover, Maehre promoted the focus on the content of a resource rather than the credibility of authors. According to Maehre, the world outside of universities and colleges classrooms is an interactive world. Thus, educators should work towards having information creators rather than information readers or finders.

In two unique studies, Hahn (2009, 2010) observed that undergraduates running the iPod Wikipedia app mainly searched for recreational and for short factual information. However, the students were all satisfied with the experience and found it useful for preparing a research paper.

A notable study in this topic area is Antin and Cheshire's (2010) survey of participation in Wikipedia. Although they only surveyed students, their study is notable as an examination of the relationship

between readership and contribution of edits to Wikipedia. Consistent with their argument that readership is in itself a crucial aspect of the community effort, they found that frequent readers of Wikipedia were better aware of its inner workings, sometimes more so than frequent editors.

In addition to these studies described here, many other articles related to domain-specific and cross-domain student readership are described elsewhere in this review in sections where we discuss studies on Wikipedia as a Source for Scholars (Eijkman, 2010), Reader Perceptions of Credibility (Kubiszewski et al., 2011), and Student Information Literacy (Aycock & Aycock, 2008; Gunnels, 2007; Harouni, 2009; Jennings, 2008; Patch, 2010; Rand, 2010; Sundin & Francke, 2009).

Amidst all these studies on student readership of Wikipedia, a few generalities stand out. First, as can be expected, students widely use Wikipedia. Second, perhaps not as expected, even students who use Wikipedia for academic assignments are generally aware of its limitations as a reference source. Notably, for disciplines where accurate information can be life-critical (healthcare) or a matter of professional reputation (journalism), students were more likely to use Wikipedia as a source for citations to more "reputable" sources. Accordingly, healthcare scholars have sometimes recommended against the use of Wikipedia for student learning, though some have recommended its use. In one case in reaction to students misusing Wikipedia, a history department has formally cautioned against its uncritical use. Nonetheless, despite some professors' reservations, Wikipedia is popular among students because of its comprehensiveness in information included and accessibility in its readability, at least compared to some alternate more traditionally authoritative sources.

Rather than recommending against Wikipedia's use even in cases of inappropriate use by students, we reiterate the calls of scholars and librarians who emphasize the need to actively educate students in the critical assessment and usage of not only Wikipedia, but of any Web resource. Regardless of professors' attitudes towards it, Wikipedia is a major information resource for students, and they would be best served in being taught how to critically and profitably use it, rather than being cautioned to avoid it.

**Reader Perceptions of Credibility**

21 studies (23%) examined the credibility of Wikipedia from its readers' perspectives. We categorize articles here that examined readers' perceptions of credibility without attempting some kind of objective evaluation of reliability, such as by subject experts—we discuss those in a separate review in the Reliability section (Mesgari et al., 2014). This is also distinct from Wikipedians' insider-view perceptions of credibility of the articles they collaboratively create; we discuss those in a separate article as Contributor Perceptions of Credibility (Okoli et al., 2012).

Some studies examined characteristics of articles' presentation that affect readers' perceptions of their credibility. Veltman (2005) argued that access to the entirety of knowledge is becoming feasible with the open source and open content movements with information widely available on the Internet. Tracing a very brief history of encyclopedic compilation, she criticized Wikipedia because "critical tools concerning variants, certainty, authority, and significance are lacking" (2005, p. 23). She highlighted the central importance of quality along with accessibility in terms of quantity and proposed techniques for presenting knowledge in general on the Internet that facilitate readers' rapid assessment of its credibility. Korosec et al. (2010) compared chemical thermodynamics topics on German Wikipedia and the chemistry encyclopedia Rompp Online based on 30 articles. After evaluating the two references in terms of varied aspects of content quality, they report that both encyclopedias obtained very good marks and performed nearly equally with regard to their accuracy. Kubiszewski et al. (2011) performed an experiment on "whether certain webpage characteristics affect academics' and students' perception of the credibility of information presented in an online article" (2011, p. 659). They concluded that "compared to Encyclopedia Britannica, article information appearing in both Encyclopedia of Earth and Wikipedia is perceived as significantly less credible" (2011, p. 664). They also found that the appearance of a biased sponsor lowered credibility.

Chen (2009) examined how information technology professionals used Wikipedia information for work-related purposes. He found that they treated Wikipedia as a ready reference for general information, but did not consider it sufficiently developed for professional use. They considered that Wikipedia needs to improve its contribution and editorial process in order to raise its quality.

Arazy and Kopak (2011) developed an instrument for measuring information quality, using Wikipedia articles as a dataset. They deliberately excluded expert evaluation of the quality of the articles, but rather focused on the quality of the information as evaluated by non-expert student readers. Although their test was not meant to be a comprehensive evaluation of Wikipedia's quality, they found that the readers generally judged the sample articles to be accurate, objective and representative, but only moderately complete; overall, the readers considered the articles to be of fairly high quality.

Francke et al. (2011) conducted an ethnographic study of a class of 29 upper-secondary students to examine how, after being specifically trained on information literacy, they navigated and assessed the credibility of online sources. In addition to the Swedish, English and German Wikipedias, the students were required to use a national Swedish encyclopedia, Greenpeace's website and a few other sources; they were free to use other sources as well. They found that the students assessed credibility mainly from four perspectives: some valued the apparent control or authoritativeness of the sources; some the balance of viewpoints presented; some the sources' commitment to a cause or opinion viewpoint; and some valued collaboratively compiled information—most notably Wikipedia.

Flanagin and Mitzger (2011) conducted two surveys among young people and adults to examine their use of Wikipedia and to what extent they trust its content. A small experiment was also embedded in the survey to compare the participants' credibility assessments of three online encyclopedias, namely Wikipedia, Citizendium, and Britannica. The survey's results revealed the unpreparedness of people to "fully relinquish traditional models of information provision" (p. 371). It also showed that both young users and adults preferred content generated by experts. However, young users preferred user-generated content when they were unaware that it was user-generated. This study demonstrated a slow shift in the perception of user- and expert-generated content.

Because of the diversity of studies on readers' perceptions of Wikipedia's credibility, it is difficult to draw any kind of general conclusions from this body of research. In fact, most such articles normally treated other subjects more substantially; thus, we describe them elsewhere in this review (Calkins & Kelley, 2009; H. Chen, 2010; Dooley, 2010; Eijkman, 2010; A.J. Head & Eisenberg, 2010; Kaplan & Haenlein, 2010; Lim & Kwon, 2010; Lim, 2009; Luyt et al., 2010, 2008; McGuinness et al., 2006; Messner & South, 2010; Page, 2010; Sundin & Francke, 2009; Zeng, Alhossaini, Ding, Fikes, & McGuinness, 2006). Nonetheless, in general, we can observe that although Wikipedia is widely read, most surveyed readers note that they are conscious that not everything they read might be accurate. We note "surveyed readers" because the enormous amount of readers might suggest that most of them in practice consider most of what they read to be true, though when explicitly asked in a survey they would emphasize their cautiousness more than their de facto acceptance of Wikipedia as a worthwhile reference source. Despite readers' expressed guardedness, they evidently consider it sufficiently credible to be a major reference source.

**Software for Readership**

Whereas most of the scholarly studies on Wikipedia readership examined social human factors, there were a few streams of research that adopted a computer science approach to investigate software specifically developed to help Wikipedia readers. Among these 12 studies (13%), some of these programs attempted to alert Wikipedia readers to the trustworthiness of articles, and others were designed to enhance the usefulness of external content by automatically identifying relevant Wikipedia content.

*Computational Estimation of Trustworthiness*

A number of studies developed computational methods for estimating the trustworthiness of articles, mainly to help readers assess whether articles were more or less reliable. Although the data used in these studies are usually independent of the reader and perform the computational modeling based on content and contributors, we discuss these articles here since the tools are directed to Wikipedia readers. This topic is distinct from human evaluations of articles' accuracy—we discuss those in a separate article as studies on the Reliability of Wikipedia (Mesgari et al., 2014). Over the years, many studies have noted and suggested various means for rapidly estimating the quality of an article, usually through the observation of reliable proxies. We discuss these techniques generally in chronological progression.

Zeng et al. (2006) developed a method to predict trustworthiness of Wikipedia articles based on the revision history of the articles, validated using featured articles. They concluded that Wikipedia is generally trustworthy, and that visualizations of article trustworthiness can enable users to access the more trustworthy versions of the articles and to avoid vandalism and malicious content. They also designed and implemented a trust management layer for collaborative information repositories in general, and Wikipedia in particular (McGuinness et al., 2006). Dondio and Barrett (2007) later developed a different method to predict trustworthiness using computational trust techniques by specifically analyzing the quality of the content and the collaborative editing contexts. They validated their method by differentiating featured articles from others using the method.

Cross (2006) offered a text-colorizing software as "a visual cue that enables [users] to see what assertions in an article have ... survived the scrutiny of a large number of people, and what assertions are relatively fresh, and may not be as reliable."

Korfiatis et al. (2006) investigated the development of quality articles in Wikipedia by using social network analysis to determine the authoritativeness of articles. They developed an approach to calculating social network measures such as centrality. They used a Web crawler (before these software agents were banned on Wikipedia because of the excess server load they cause). They argued that as Wikipedia keeps growing, it will be more challenging to keep the content reliable.

Hu et al. (2007) proposed three models for assessing quality of Wikipedia articles based on the interaction data between articles and their contributors. They found that simple article length often improved model performance. Similarly, Blumenstock (2008) found that the word count of an article performs surprisingly well as a predictor for article quality, at least when distinguishing between featured and random articles, with an error rate of around 96% on a corpus of 1,554 featured and 9,513 randomly selected articles. He suggested setting a cut-off at 2,000 words between the two sets. However, the best model that Hu et al. (2007) developed, called "ProbReview", superiorly predicted quality irrespective of article length. We suggest that the number of editors having an article on their watch list could possibly also make a good indicator of the article quality.[9]

The Wikiganda, formerly available from www.wikiwatcher.com, used automated text analysis to detect biased edits (Chandy, 2009). The sentiment analysis technique uses a lexicon of over 20,000 words from General Inquirer and Wiebe wordlists so that each revision can get a Propaganda Score labeled as negative, positive or "vague" propaganda. In conjunction with the WikiTrust system and evaluated against 200 manually labeled revisions, the system showed a precision/recall performance of 52%/63%.

A related body of work includes studies on Wikipedia reputation systems, which compute the trustworthiness of Wikipedia contributors. Although the reputation of an article's authors could be considered an indirect measure of the trustworthiness of the article itself, this topic is more directly related to Wikipedia contributors than it is to readers; hence, we discuss such studies in the Reputation

---

[9] See https://en.wikipedia.org/w/index.php?title=Wikipedia&action=info#mw-pageinfo-watchers.

Systems section of a separate review dedicated to studies of participation in and contribution to Wikipedia (Okoli et al., 2012).

Overall, the studies we have reviewed here have employed a wide variety of measures to computationally estimate article reliability for the benefit of Wikipedia readers: article revision history, longitudinal readership of portions of text, social network analysis, computational trust techniques, word counts, article-contributor interaction data, and wordlists from external corpuses. These techniques have varying degrees of value, but Wikipedia has not incorporated any of these or any other computational reliability tools into their software. Various tags (such as the ubiquitous "citation needed") signal to readers that discernment is needed in assessing the reliability of what they are reading, but there has been no attempt to implement an automatic objective estimate of article reliability. Perhaps such tools are best employed in external reader software that accesses Wikipedia articles and then presents them to readers with enhanced features, including reliability estimates.

### *Reading Support*

Encountering knowledge gaps while reading is an issue that people face daily. This problem has motivated some researchers to develop reading support tools using Wikipedia to fill these gaps. In other words, Wikipedia articles were extracted to fill the missing information resulting from knowledge gaps. The studies we discuss here do not support reading of Wikipedia articles; rather, they employ Wikipedia articles to support the reading of other texts.

Jordan and Watters (2009) designed a prototype to bring up the single most relevant Wikipedia article when a user selects part of a text in a separate software text reader. The most successful model could accurately find the best article in 70% of the cases and help readers to fill the gap in their personal knowledge using Wikipedia articles. As an application for the previous study, Jordan (2009) proposed a system to help people reading academic abstracts to be able to highlight part of them, and then a pop-up would appear with a single Wikipedia article explaining the highlighted part. The system tries to suggest the most related article based on understanding the context of the abstract and article categories.

With the increase of blogs and social networks comes the need for a support system to fill the content holes. Nadamoto et al. (2010) suggested a new method to search for these content holes, defined as "the user's unawareness of information." Wikipedia articles were used to extract and present the holes in community-type content. Their proposed method differs from otherwise related information retrieval tasks in that it searches for different information instead of similar one.

### Commercial Aspects

Wikipedia is an open content project maintained by a not-for profit foundation. Nonetheless, 10 studies (11%) have investigated commercial aspects of Wikipedia. These range from those that investigate reasons why Wikipedia chose to adopt a not-for-profit path early on, those that consider commercial enterprises' responses to Wikipedia's entries on them, and those that investigate the use of Wikipedia content for commercial benefit.

Some studies investigated characteristics of Wikipedia that relate to its not-for profit direction. In his dissertation that covered numerous other themes, Gehl (2010) discussed Wikipedia's early (2002) consideration of featuring profit-garnering advertisements. However, the fork of the Spanish Wikipedia and other widespread community protest killed that idea. Wikipedia was thus firmly steered in a not-for-profit direction, for the primary interest of its users (both readers and contributors) rather than commercial enterprises such as Jim Wales's Bomis, Inc. Although Wikipedia is not itself a commercial project, Cedergren (2003) argued that its production of useful articles mirrors a commercial value chain in creating valuable resources for readers. Although not strictly a commercial perspective, Rahman (2006, 2008) conducts economic analyses of Wikipedia as a public good. He concluded that Wikipedia's

uniqueness "as a public good, combined with free-riding and free-editing help to maintain the [large size and] reliability of Wikipedia" (Rahman, 2008, p. 96) relative to other open source systems.

Although Wikipedia is itself a not-for profit endeavour, its contents include information about commercial enterprises. Wikipedia articles about corporate entities have high prominence on the web, and are often perceived by readers, whether justifiably or not, as a less-biased source than the company's own website. As such, some researchers argue that corporations should be interested in Wikipedia's contents relating to them.

Hickerson and Thompson (2009) considered the potential of Wikipedia as a tool for public relations, investigating how "wiki sites uphold dialogic principles and encourage dialogue" (2009, p. 9). The fact that the site is open, free, and does not serve financial interests of any single party at the expense of others, contributes to participants' feeling of partial ownership. This, in turn, "may encourage repeat visits to the site and an increased investment in the organisation" (2009, p. 9).

Kaplan and Haenlein (2010) introduced businesses to using social media. They noted that Wikipedia is very restrictive in permitting commercial participation in its community, yet urged businesses to pay attention to it because "although not everything written on Wikipedia may actually be true, it is believed to be true by more and more Internet users" (2010, p. 62). However, they warned that trying to gloss corporate image by getting third parties to edit Wikipedia articles is probably futile at best and at worst, could likely backfire.

DiStaso and Messner (2010) analyzed 10 companies in 2006, 2008 and 2010. They found that over this period the search engine prominence of their Wikipedia articles increased, the tone changed for some companies in a negative direction, and the percentage of topics on controversial issues (e.g., "legal concerns/scandals") increased. They concluded that "the monitoring of Wikipedia in public relations should be included in all social media plans".

Another commercial perspective on Wikipedia hinges on the fact that its Creative Commons Attribution-ShareAlike license explicitly permits commercial reuse of its content. Some studies investigated phenomena that try to leverage or take advantage of this permission.

Two studies investigated aspects of how companies might profit directly from Wikipedia. Langlois and Elmer (2009) investigated how Wikipedia content is being used anywhere across the Internet. They found that it is mostly used for generating commercial content or for increasing traffic through search engines links. Plaza (2011) investigated how Wikipedia entries can get traffic to a tourism website in comparison with other traffic sources like Google. She found that Wikipedia entries are quite effective in getting people to visit and navigate through the sample website she studied.

In related work, Rubin and Rubin (2010) hypothesized that the degree of Web activity about a company correlates with the extent to which investors are generally informed about their companies. To test this, they investigated the frequency of edits of Dow Jones Industrial firms' entries on Wikipedia in relation to analysts' forecasts and recommendations, and confirmed that Wikipedia edit frequencies are indeed correlated with the accuracy of corporate analysts' forecasts.

In general, these studies on the commercial aspects of Wikipedia found that although Wikipedia itself is not commercial, it has important implications for commercial enterprises. Wikipedia's broad readership makes it a primary public source of information about companies. Although readers do not necessarily believe everything they read on Wikipedia, they generally consider it a more reliable source on a company than the biased information that the company itself might provide to the public. Thus, enterprises are advised to pay attention to their Wikipedia image and to work within the community guidelines to ensure that their information on Wikipedia is not inaccurate, and that positive factual information is prominently included. Other studies have begun to explore how commercial enterprises can exploit Wikipedia's rich information base to help earn more revenue for themselves. We expect such uses

of Wikipedia to increase as enterprises gain expertise in the commercial use of social media and user-generated content in general.

# Research Trends for Studies of Wikipedia Readership

Our presentation of the various readership topics in Wikipedia shows a wide variety of research topics. On our WikiLit website, we have examined each scholarly article to identify various key details of interest to researchers. In fact, the data analysis on the WikiLit website is highly interactive, and researchers can analyze these data themselves with tables and graphs[10]. In this major section of our review, we present an overview of the various research trends of studies of Wikipedia readership. For both new and experienced Wikipedia researchers, we believe it is valuable to understand how past research on Wikipedia readership has been conducted. These perspectives can help researchers determine how to conduct future research, and also identify underutilized techniques that might be appropriate for future work.

We have two important notes concerning the research trends we report here. First, not all the 99 studies that we have summarized in the preceding section on Findings were parsed for their research details in our dataset. Although we did parse most of them, because of the work involved, we only parsed details from 91 studies as mentioned in the introduction. Second, in our dataset of parsed papers on the WikiLit website, papers were frequently assigned to multiple topics. For example, as we explained earlier, many papers treated Reader Perceptions of Credibility along with other topics. Thus, although our dataset included only 91 studies, if the counts in most tables were added up, they would exceed this number. However, we publish at the bottom of each table the number of distinct studies for each criterion without duplicates.

**Table 2. Wikipedia readership studies by year**

|  | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|---|---|---|
| Commercial aspects | 1 |  |  | 1 |  | 1 | 2 | 4 | 1 |
| Knowledge source |  |  |  |  |  |  |  |  |  |
|    Health information source |  |  |  | 1 |  | 3 | 4 | 2 | 1 |
|    Knowledge source for scholars and librarians |  |  |  |  | 1 | 2 | 5 | 7 | 1 |
|    News source |  | 1 |  |  | 2 |  |  | 1 |  |
| Ranking and popularity |  |  |  | 1 | 2 | 2 | 4 | 2 | 1 |
| Reader perceptions of credibility |  |  | 1 | 2 |  | 1 | 4 | 6 | 4 |
| Software for readership |  |  |  |  |  |  |  |  |  |
|    Computational estimation of trustworthiness |  |  |  | 4 | 2 | 1 |  |  | 2 |
|    Reading support |  |  |  |  |  |  | 2 | 1 |  |
| Student readership |  |  |  |  |  |  |  |  |  |
|    Cross-domain student readership |  |  |  |  | 3 | 2 | 5 | 4 | 2 |
|    Domain-specific student readership |  |  |  |  |  | 3 | 3 | 5 | 2 |
|    Student information literacy |  |  |  |  | 1 | 2 | 6 | 2 | 2 |
| **Number of distinct studies** | **1** | **1** | **1** | **7** | **10** | **11** | **22** | **26** | **12** |

---

[10] http://wikilit.referata.com/wiki/Pivot_tables

In the rest of this section, we present detailed research trends for studies of Wikipedia Readership. In the Discussion section that follows, we highlight a few noteworthy findings from these trends, as well as from the detailed summaries in prior sections of this review.

**Readership Topics by Year**

As can be expected, studies of Wikipedia research have increased over the years, as shown in Table 2. The earliest scholarly study we have identified was published in 2003, followed by one each in 2004 and 2005. Not until 2006 did the pace of publication begin to pick up, rising to over 20 per year by 2009. (The reported studies are fewer in 2011 because we stopped counting in June 2011.) The most popular topics in recent years have been those related to Student Readership and Wikipedia as a Knowledge Source, especially as a Health Information Source and as a Knowledge Source for Scholars and Librarians.

**Table 3. Wikipedia readership studies by knowledge domain: social sciences**

| | Business | Communications | Economics | Education | Info. systems | Journalism | Knowledge management | Law | Library science | Political science | Psych. | Sociology |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Commercial aspects | 2 | 1 | 2 | | 3 | | | | | | | |
| Knowledge source | | | | | | | | | | | | |
|    Health information source | | | | | | | | | | | | |
|    For scholars and librarians | | | | 2 | | | 1 | 2 | 6 | 1 | | |
|    News source | | 3 | | | | | 1 | | | | | |
| Ranking and popularity | 1 | | | | | | | | | | | 1 |
| Reader perceptions of credibility | | 1 | | 6 | 4 | 3 | | | 2 | | 2 | |
| Software for readership | | | | | | | | | | | | |
|    Computational estim. of trust. | | | | | 4 | | | | | | | 1 |
|    Reading support | | | | | | | | | 1 | | | |
| Student readership | | | | | | | | | | | | |
|    Cross-domain student readership | | 1 | | 6 | 1 | | | | 6 | | | |
|    Domain-specific student readership | | 1 | | 5 | | | | | | | 1 | 1 |
|    Student information literacy | | | | 8 | | | | | 5 | | | |
| **Number of distinct studies** | **3** | **7** | **2** | **22** | **10** | **3** | **2** | **2** | **16** | **1** | **3** | **3** |

**Readership Topics by Domain of Knowledge**

Table 3 and Table 4 display the domains of knowledge, that is, the fields or disciplines, of the various studies of Wikipedia readership. The most common domain has been education, with 22 studies (24%), most of them concerning Student Readership. Library science is next, with 16 studies, covering Wikipedia as a Knowledge Source for Librarians and various aspects of Student Readership. Next, there have been 15 computer science studies, mainly on Software for Readership. Health studies, information

science and information systems are the other three fairly common domains covered, with 11 or 12 studies of Wikipedia readership each.

**Table 4. Wikipedia readership studies by knowledge domain: interdisciplinary, humanities and natural sciences**

| | Interdisciplinary | | | Humanities | | Natural sciences | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | Computer science | Health | Information science | History | Rhetoric | Biology | Chemistry |
| Commercial aspects | | 1 | 1 | | | | |
| Knowledge source | | | | | | | |
| Health information source | | 10 | 2 | | | | |
| Knowledge source for scholars and librarians | 1 | | | 1 | | 2 | |
| News source | | | | | | | |
| Ranking and popularity | 4 | 3 | 4 | | | | |
| Reader perceptions of credibility | 3 | | 4 | 1 | | | 1 |
| Software for readership | | | | | | | |
| Computational estimation of trustworthiness | 4 | | 1 | | | | |
| Reading support | 3 | | 1 | | | | |
| Student readership | | | | | | | |
| Cross-domain student readership | 2 | | 1 | 1 | 1 | | |
| Domain-specific student readership | 1 | 3 | | | 1 | | 1 |
| Student information literacy | | | 1 | | | | |
| **Number of distinct studies** | **15** | **12** | **12** | **3** | **2** | **2** | **1** |

**Readership Topics by Degree of Focus on Wikipedia**

Table 5 displays the extent to which studies in our dataset focused on Wikipedia. 51 of the 91 studies (56%) treated Wikipedia as the main topic of the study. 30 studies (33%) considered Wikipedia only as one case among other websites or resources that it studied. 6 studies (7%) were not explicitly focused on Wikipedia, but employed data extracted from Wikipedia to test their hypotheses or develop their tools.

**Readership Topics by Publication Type**

Table 6 displays the three different types of publication in our dataset: conference papers, journal articles and theses. Since we explicitly focused on journal articles and doctoral theses, we have 82 journal articles and three theses, and only six conference papers. Interestingly, two of the three theses treated on commercial aspects of Wikipedia. Four of the six conference papers focused on computational estimation of trustworthiness, which is not surprising since computer science topics such as this are very frequently published in conference proceedings and not in journals.

**Table 5. Wikipedia readership studies by their degree of focus on Wikipedia**

| | Main topic | One case among others | Sample data | Other |
|---|---|---|---|---|
| Commercial aspects | 5 | 5 | | |
| Knowledge source | | | | |
|     Health information source | 5 | 5 | 1 | |
|     Knowledge source for scholars and librarians | 6 | 9 | | 1 |
|     News source | 2 | | 1 | 1 |
| Ranking and popularity | 7 | 2 | 2 | 1 |
| Reader perceptions of credibility | 13 | 8 | | |
| Software for readership | | | | |
|     Computational estimation of trustworthiness | 9 | | | |
|     Reading support | | | 3 | |
| Student readership | | | | |
|     Cross-domain student readership | 11 | 5 | | 1 |
|     Domain-specific student readership | 9 | 3 | | |
|     Student information literacy | 7 | 6 | | |
| **Number of distinct studies** | **51** | **30** | **6** | **4** |

**Table 6. Wikipedia readership studies by publication type**

| | Conference paper | Journal article | Thesis |
|---|---|---|---|
| Commercial aspects | | 8 | 2 |
| Knowledge source | | | |
|     Health information source | | 11 | |
|     Knowledge source for scholars and librarians | | 16 | |
|     News source | 1 | 3 | |
| Ranking and popularity | | 12 | |
| Reader perceptions of credibility | 2 | 19 | |
| Software for readership | | | |
|     Computational estimation of trustworthiness | 4 | 5 | |
|     Reading support | | 2 | 1 |
| Student readership | | | |
|     Cross-domain student readership | 1 | 16 | |
|     Domain-specific student readership | | 12 | |
|     Student information literacy | | 13 | |
| **Number of distinct studies** | **6** | **82** | **3** |

**Readership Topics by Data Source**

Table 7 displays the source of the data that was collected in the various empirical studies. (N/A refers to non-empirical conceptual studies.) As can be expected, the most data source for Wikipedia readership research has been various kinds of Wikipedia pages (we examine the specific kinds of pages shortly). This was particularly the case in studies of software for readership, since these were directly related to enhancing and encouraging the reading of Wikipedia articles. After Wikipedia pages, the next most common source has been survey data, which has in fact been the single most popular data collection approach for studies of student readership and for perceptions of credibility. Comparisons of Wikipedia among other websites follows as the next most common data collection approach, particularly in studies of Wikipedia as a knowledge source and of Wikipedia's ranking and popularity. Experiments and interviews have also been popular. In contrast to other kinds of Wikipedia research, only very few readership studies have employed archival records, computer usage logs, direct observations, or other kinds of documents (not including Wikipedia pages).

**Table 7. Wikipedia readership studies by source of collected data**

| | Wikipedia pages | Archival records | Computer usage logs | Direct observation | Documents | Experiment | Interviews | Scholarly articles | Survey | Websites | N/A |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Commercial aspects | 1 | | | | | | 1 | | 1 | 4 | 3 |
| Knowledge source | | | | | | | | | | | |
|    Health information source | 5 | | | | | 1 | 1 | | 2 | 3 | 3 |
|    Knowledge source for scholars and librarians | 2 | 1 | | | | 1 | 1 | | 2 | 2 | 6 |
|    News source | 3 | | | | 2 | 1 | | | | 2 | |
| Ranking and popularity | 7 | | 1 | | | 2 | | | | 6 | |
| Reader perceptions of credibility | 4 | | | 2 | 1 | 3 | 3 | | 8 | 1 | 1 |
| Software for readership | | | | | | | | | | | |
|    Computational estimation of trustworthiness | 8 | | | | | 2 | 1 | | | | |
|    Reading support | 3 | 1 | | | | 3 | | | | | |
| Student readership | | | | | | | | | | | |
|    Cross-domain student readership | 2 | | | | 1 | 1 | 1 | | 8 | | 5 |
|    Domain-specific student readership | 3 | | 2 | | | 1 | | | 4 | 1 | 1 |
|    Student information literacy | 2 | | | | 1 | 1 | 1 | 1 | 1 | | 6 |
| **Number of distinct studies** | **29** | **2** | **3** | **2** | **4** | **12** | **7** | **1** | **19** | **17** | **18** |

**Readership Topics by Wikipedia Page Type**

Table 8 displays the specific different types of Wikipedia pages analyzed by those studies that carried out article-level analyses. As can be expected, of those Wikipedia readership studies that analyzed specific Wikipedia pages (45 of 91, 49%), the vast majority (37 of 45, 82%) analyzed the main encyclopedia articles, which is what Wikipedia readers actually read.

# Table 8. Wikipedia readership studies by Wikipedia page type treated

| | Article | Article:talk | History | Policy | Discussion and Q&A | Other | N/A |
|---|---|---|---|---|---|---|---|
| Commercial aspects | 2 | 1 | 1 | | | | 6 |
| Knowledge source | | | | | | | |
|    Health information source | 4 | | 1 | 1 | | | 5 |
|    Knowledge source for scholars and librarians | 7 | | 1 | | | | 9 |
|    News source | 2 | | | | | | 2 |
| Ranking and popularity | 7 | | | | | 1 | 4 |
| Reader perceptions of credibility | 9 | | | | | | 12 |
| Software for readership | | | | | | | |
|    Computational estimation of trustworthiness | 8 | | 2 | | 1 | | |
|    Reading support | 3 | | | | | | |
| Student readership | | | | | | | |
|    Cross-domain student readership | 5 | | | | | | 12 |
|    Domain-specific student readership | 5 | | | | | 1 | 6 |
|    Student information literacy | 2 | 1 | | | | | 10 |
| **Number of distinct studies** | **37** | **2** | **5** | **1** | **1** | **2** | **46** |

## Readership Topics by Unit of Analysis

Table 9 displays the units of analysis of the Wikipedia data analyzed in the studies. A third of the studies (30 studies, 33%) used individual Wikipedia articles as the unit of analysis. This was particularly the case for Ranking and Popularity studies (which often assessed the popularity of individual articles) and for Software for Readership studies (which estimated the reliability of individual articles or presented individual articles as suggestions). Second was website as a unit of analysis (22 studies, 24%), when Wikipedia was compared with other websites. Next was the individual Wikipedia user as a unit of analysis (17 studies, 19%), most frequently among studies of Reader Perceptions of Credibility, where the edit history or reputation of a contributor to an article was sometimes found to affect how readers judged an article, and among Student Readership studies, which analyzed usage by individual students as reader-users.

## Readership Topics by Time Dimension of Collected Data

Table 10 displays the time dimensions of the Wikipedia data that was collected for the studies. Over half of the studies (51 of 91, 56%) collected snapshot cross-sectional data from Wikipedia, that is, data collected at just one point in time. It is notable that all but one of the Software for Readership studies were cross-sectional, as they extracted data from the live Wikipedia at a given point in time to process for their analyses. Only 17 studies collected longitudinal data, that is, data from the same Wikipedia pages on at least two distinct dates. The remaining 23 studies were either non-empirical or they collected data from sources other than Wikipedia, yet nonetheless treated Wikipedia readership substantially.

**Table 9. Wikipedia readership studies by unit of analysis of collected Wikipedia data**

| | Edit | Article | Subject | User | Website | Scholarly article | N/A |
|---|---|---|---|---|---|---|---|
| Commercial aspects | 1 | 2 | | 1 | 2 | | 4 |
| Knowledge source | | | | | | | |
|     Health information source | | 3 | 1 | 2 | 4 | | 1 |
|     Knowledge source for scholars and librarians | | 4 | | 1 | 4 | | 8 |
|     News source | | | 1 | 1 | 1 | | 1 |
| Ranking and popularity | | 7 | 2 | | 3 | | 1 |
| Reader perceptions of credibility | | 5 | | 8 | 4 | | 4 |
| Software for readership | | | | | | | |
|     Computational estimation of trustworthiness | 1 | 9 | | 1 | | | |
|     Reading support | | 3 | | | | | |
| Student readership | | | | | | | |
|     Cross-domain student readership | | 4 | | 7 | 2 | | 4 |
|     Domain-specific student readership | | 2 | | 3 | 6 | | 1 |
|     Student information literacy | | 3 | | 2 | 2 | 1 | 5 |
| **Number of distinct studies** | **2** | **30** | **4** | **17** | **22** | **1** | **21** |

**Table 10. Wikipedia readership studies by time dimension of collected Wikipedia data**

| | Cross-sectional | Longitudinal | N/A |
|---|---|---|---|
| Commercial aspects | 2 | 4 | 4 |
| Knowledge source | | | |
|     Health information source | 7 | 1 | 3 |
|     Knowledge source for scholars and librarians | 7 | 1 | 8 |
|     News source | 1 | 3 | |
| Ranking and popularity | 9 | 3 | |
| Reader perceptions of credibility | 15 | 3 | 3 |
| Software for readership | | | |
|     Computational estimation of trustworthiness | 8 | 1 | |
|     Reading support | 3 | | |
| Student readership | | | |
|     Cross-domain student readership | 9 | 3 | 5 |
|     Domain-specific student readership | 10 | 1 | 1 |
|     Student information literacy | 5 | 11 | 7 |
| **Number of distinct studies** | **51** | **17** | **23** |

**Readership Topics by Means of Wikipedia Data Extraction**

Table 11 displays the approaches employed by the studies to extracting data from Wikipedia. Nearly half of the studies (42 of 91, 46%) extracted data from the live, running Web version of Wikipedia. 10 studies, however, cloned the Wikipedia database, or at least parts of it, to conduct more extensive analysis. This was particularly the case for studies that analyzed complex longitudinal data and for those that used Wikipedia as sample data for testing complex algorithms. The other studies either did not use empirical data, or the means of extracting data from Wikipedia was not very relevant to their analysis approaches.

**Table 11. Wikipedia readership studies by their means of obtaining data from Wikipedia**

|  | Live Wikipedia | Data dump | N/A |
|---|---|---|---|
| Commercial aspects | 5 |  | 5 |
| Knowledge source |  |  |  |
|     Health information source | 7 |  | 4 |
|     Knowledge source for scholars and librarians | 7 | 1 | 8 |
|     News source | 2 |  | 2 |
| Ranking and popularity | 8 | 2 | 2 |
| Reader perceptions of credibility | 6 | 3 | 12 |
| Software for readership |  |  |  |
|     Computational estimation of trustworthiness | 5 | 4 |  |
|     Reading support | 1 | 2 |  |
| Student readership |  |  |  |
|     Cross-domain student readership | 6 | 1 | 10 |
|     Domain-specific student readership | 7 |  | 5 |
|     Student information literacy | 3 |  | 10 |
| **Number of distinct studies** | **42** | **10** | **39** |

**Readership Topics by Wikipedia Language Version**

Table 12 displays the different language versions of Wikipedia employed in the studies. Although 35 of the 91 studies explicitly mentioned that they treated the English Wikipedia, 39 did not specify what language version they treated. Since we restricted our review to studies published in English, we consider it reasonable to assume that virtually all of these unspecified studies also treated the English Wikipedia. Based on that assumption, 74 of 91 studies (81%) treated the English Wikipedia. Among the 8 other Wikipedia readership studies that explicitly treated a language version, 1 treated the Czech version, 5 treated the German, and 2 the Swedish version of Wikipedia. 11 studies were conducted such that the Wikipedia language version was irrelevant to their analysis. Although the sample size is too small to draw significant findings, we do find it interesting that studies of the German Wikipedia were predominantly interested in rankings and comparisons: two studies were about the Ranking and Popularity of Wikipedia in general, and two others specifically on its value as a Health Information Source.

**Table 12. Wikipedia readership studies by Wikipedia language version treated**

| | English | Not specified | English/Not specified | Czech | German | Swedish | Other than English | N/A |
|---|---|---|---|---|---|---|---|---|
| Commercial aspects | 3 | 5 | 8 | | | 1 | 1 | 1 |
| Knowledge source | | | | | | | | |
|   Health information source | 5 | 3 | 8 | | 2 | | 2 | 1 |
|   For scholars and librarians | 5 | 10 | 15 | | | | | 1 |
|   News source | 2 | 1 | 3 | | | | | 1 |
| Ranking and popularity | 6 | 3 | 9 | | 2 | | 2 | 1 |
| Reader perceptions of credibility | 9 | 7 | 16 | | 1 | 1 | 2 | 4 |
| Software for readership | | | | | | | | |
|   Computational estim. of trust. | 6 | 2 | 8 | | 1 | | 1 | |
|   Reading support | 2 | 1 | 3 | | | | | |
| Student readership | | | | | | | | |
|   Cross-domain | 3 | 11 | 14 | | | | | 3 |
|   Domain-specific | 5 | 3 | 8 | 1 | 1 | | 2 | 2 |
|   Information literacy | 1 | 7 | 8 | | 1 | | 1 | 4 |
| **Number of distinct studies** | **35** | **39** | **74** | **1** | **5** | **2** | **8** | **11** |

**Table 13. Wikipedia readership studies by research design**

| | Action research | Case study | Conceptual | Content analysis | Discourse analysis | Ethnography | Experiment | Grounded theory | Lit. review | Math. modeling | Phenomenology | Stat. analysis | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Commercial aspects | | 3 | 2 | 1 | | | 1 | | | 2 | | 2 | |
| Knowledge source | | | | | | | | | | | | | |
|   Health information source | | | 2 | 2 | | | 1 | 2 | | | | 3 | 1 |
|   For scholars and librarians | 1 | 2 | 6 | | 1 | | 1 | | | | | 2 | 1 |
|   News source | | | | 3 | 1 | | 1 | | | | | 1 | |
| Ranking and popularity | | | | 4 | | | 3 | | | | | 4 | |
| Reader perceptions of credibility | 2 | 3 | 1 | 1 | 1 | 2 | 3 | | | | 1 | 3 | |
| Software for readership | | | | | | | | | | | | | |
|   Computational estim. of trust. | | 1 | | | | | 3 | | | 2 | | 1 | |
|   Reading support | | | | | | | 3 | | | | | | |
| Student readership | | | | | | | | | | | | | |
|   Cross-domain | | 2 | 4 | 1 | | 1 | 2 | | | | | 2 | 1 |
|   Domain-specific | 1 | | 1 | 1 | | | | | | | | 5 | |
|   Information literacy | 1 | 2 | 5 | | | 2 | | | 1 | | | | 1 |
| **Number of distinct studies** | **2** | **8** | **16** | **9** | **2** | **2** | **12** | **2** | **1** | **4** | **1** | **32** | **4** |

**Readership Topics by Research Design**

Table 13 displays the various research designs adopted for studying Wikipedia readership. The most common approach has been various kinds of statistical analysis (32 studies, 35%), mainly involving quantitative analysis of survey data. The next most common approach (16 studies, 18%) has been non-empirical conceptual studies. Experiments have also been popular (12 studies, 13%), being the predominant approach for the computer science studies of Software for Readership.

# Discussion

**Key Findings about Wikipedia Readership**

Even without peer-reviewed academic research, it is obvious that Wikipedia is the world's most popular reference website. However, scholarly research has given us more fine-tuned understanding of the nature of what people read on Wikipedia, with various important implications for society. In this section we highlight some of the key findings from our review.

- As expected, scholarly research has affirmed that Wikipedia is an extremely popular resource. But what is most popular on Wikipedia? Studies have found that articles generally related to entertainment and sexuality top the list, covering over 40% of visits (Spoerri, 2007a, 2007b; Waller, 2011). More "serious" topics are also widely consulted, but it is not surprising that Wikipedia's primary attraction is in topics with the widest popular appeal.
- In contrast, Wikipedia is not used as a primary news source, even though news is a very popular topic on the web (Thelwall & Stuart, 2007). While it is most certainly used, as it has articles that quickly appear and develop in line with most current news topics (Chua et al., 2007; Lih, 2004), Wikipedia's collaborative editing by volunteers cannot match the instant, professional journalistic updates of numerous news organizations that publish on the Web.
- Regarding "serious" topics, Wikipedia is a widely used health information source, not only by the general public but even by health professionals (Hughes et al., 2009; Younger, 2010). However, numerous studies have cautioned that while helpful and informative, Wikipedia cannot be regarded as an authoritative health resource (Heilman et al., 2011; Laurent & Vickers, 2009).
- Another "serious" topic concerns Wikipedia's use for legal information, which has in fact been quite widespread (Breinholt, 2008; Peoples, 2009; Stoddard, 2009), sometimes perhaps too much so. Although it is valuable for providing background information surrounding legal issues, Wikipedia itself explicitly disclaims its use as a source of legal advice[11], and is not suitable for such purposes. We suggest, however, that Wikipedia articles can be appropriately used for definitions, when such definitions have been considered by many readers and multiple editors have edited the article over time—these could be considered consensus definitions.
- The most common research topics have been Student Readership and Knowledge Source for Scholars and Librarians in the education and library science domains. These topics are naturally very interesting to the professors and research librarians who have conducted much Wikipedia research. Health information studies have also been popular among professors in healthcare.
- As with all research fields where survey and experimental research is common, students are a common subject for Wikipedia research. On one hand, they are a demographic of young and

---

[11] http://en.wikipedia.org/wiki/Wikipedia:Legal_disclaimer

intelligent people, and so are highly relevant for forward-looking topics of inquiry, such as the use of Web 2.0 resources like Wikipedia. On the other hand, they provide a convenience sample, since professors can easily persuade their students to participate in studies, either purely voluntarily or for course credit, if some learning component is incorporated into the exercise.

- Contrary to popular opinion, most scholars and librarians are generally quite favorable towards Wikipedia. They sometimes not only encourage students to use it, but sometimes use it themselves in their research and publications. Even one of the most notorious public detractors of student use of Wikipedia was less categorical in his censure than the media depicted (Waters, 2007).

- There is a strong academic consensus that, like it or not, Wikipedia is a dominant form in today's information landscape and so students should be proactively trained how to critically assess and productively use Wikipedia and other contemporary Internet resources—this is a very important aspect of their education in information literacy.

- Previous research has stigmatized non-contributory readership as "free-riding" or "lurking", or valuable only in its potential to convert readers to contributors. However, almost 100 studies covering a rich diversity of Wikipedia readership demonstrate the value of readership in itself. Non-contributory readership is not a problem; it is a valuable aspect of the Wikipedia ecosystem.

## Implications for Wikipedia Researchers

Our review is mainly targeted to scholars, both new and experienced in Wikipedia research. Wikipedia is a culturally significant and revolutionary concept that has single-handedly considerably uplifted the quality of Web-based information. Not only is it in itself an important subject for research inquiry, but as the original highly successful non-software open content endeavor, it serves as a model and example for other open content projects, ranging from open educational resources to open data. Moreover, because of its open data policies, Wikipedia is uniquely rich and readily-accessible source of both quantitative and qualitative data for a very broad range of research topics. In a separate review, we identify a number of articles and numerous resources that can help newer Wikipedia researchers get started in this eclectic research platform (Okoli et al., 2012, sec. Research Platform).

In addition to these general observations, there are a few specific implications and suggestions for future research that this present review has brought to life:

- We sense a contradiction in that, although Wikipedia is such a widely used resource, surveys of Wikipedia readers have generally reported that the readers are cautious about the credibility of Wikipedia information. We would think that if they were as cautious as the surveys report, they would probably not use it as much. We fear that the surveys that ask them questions related to Wikipedia's credibility naturally incite them to report that they are more cautious than they actually might be. It might be more elucidating to investigate readers' perceptions of Wikipedia's credibility by behavior-based measures rather than by self-reporting. That is, a study might observe and compare how readers use Wikipedia information alongside other information sources, perhaps in an exercise such as preparing a report for work or school. Such observation of what information they actually use might be a more accurate reflection of to which sources they lend greater credibility. Such a study would have to carefully examine the actual source of the information used, not just the explicit citations, as people might not cite Wikipedia as often as they use it, being aware of its controversial status.

- Researchers should perhaps give more attention to other Wikimedia Foundation projects. Although Wikipedia is the Foundation's flagship project, other projects might be more closely aligned with certain disciplinary fields. For example, we were surprised that even a study of Wikipedia's representation of biological taxa (Page, 2010) made no mention of Wikispecies, the project dedicated to this purpose—indeed, we have not found any scholarly studies that examine this project. Wikinews, on the other hand, has been covered to a fair degree by journalistic studies

(for example, Bruns, 2006; Thelwall & Hellsten, 2006; Thorsen, 2008). The broad range of Wikimedia Foundation projects provides similarly promising platforms for research in various domains[12].

- Although commercial use of "the free encyclopedia" sounds oxymoronic, such a thought is in fact explicitly embedded in the open content licenses that Wikipedia has historically employed. The studies on commercial aspects of Wikipedia show that there are various approaches to monetizing Wikipedia content. However, there are many other existing approaches that have not been researched, such as the printing of books composed entirely of Wikipedia content[13] and incorporating Wikipedia pages into other websites to enhance content. This research topic could be particularly applicable to other open content projects which, though they might not have anywhere as extensive a reach as Wikipedia, might be able to provide sufficient value to both commercially sustain themselves and provide commercial opportunities to the general public.

- Relatively few studies (17 of 91, 19%) analyzed longitudinal data over time. We note that none of the studies of Readers' Perceptions of Credibility examined changes in readers' perceptions over time. Any given Wikipedia article will generally improve over time, so it would be quite insightful for a study to examine if individual readers' perceptions might change over time. Understandably, such a study would be considerably more difficult to conduct than the cross-sectional studies that have been conducted to date.

## Implications for the Wikipedia Community

We also offer various implications of our findings to contributing members of the Wikipedia community, including the general Wikipedia-reading public.

- The Wikimedia Foundation could consider incorporating one or more of the algorithms for computational estimation of the reliability of Wikipedia articles that have been developed to help address credibility concerns, so that readers could be readily sensitized. This is consistent with the open spirit of Wikipedia's Article Feedback Tool[14], which lets readers rate articles. Moreover, articles with low estimates could be targeted for priority editing by regular contributors. There have been such efforts in the past (most notably WikiTrust[15]), but their current status is unclear.

- Educators and librarians should embrace Wikipedia as a unique opportunity for training students in digital literacy. In fact, although this review is restricted to studies of Wikipedia readership, numerous studies emphasize the writing aspect of digital literacy—we cover such studies in a separate review (Okoli et al., 2012, sec.  Student Contribution). Students, for their part, should proactively learn how Wikipedia works, and especially study its Neutral Point of View[16] and Verifiability[17] policies as prime examples of Internet-age digital literacy.

- Studies on commercial aspects of Wikipedia repeatedly echoed that organizations need to pay attention to their Wikipedia profiles (DiStaso & Messner, 2010; Hickerson & Thompson, 2009; Kaplan & Haenlein, 2010). Whether accurate or not, balanced or not, Wikipedia is often seen as a more neutral source of information than what an organization would provide on its own website. While favorably editing their own Wikipedia articles could seriously backfire, by studying and

---

[12] http://wikimediafoundation.org/wiki/Our_projects

[13] http://en.wikipedia.org/wiki/VDM_Publishing#Alphascript_and_Betascript_Publishing

[14] http://en.wikipedia.org/wiki/Wikipedia:Article_Feedback_Tool

[15] http://en.wikipedia.org/wiki/Wikitrust

[16] http://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view

[17] http://en.wikipedia.org/wiki/Wikipedia:Verifiability

understanding the policies and culture of Wikipedia, organizations can ensure that they are more fairly presented on the site.

- Since Wikipedia consistently tops search engine hit lists, website managers who want to attract visitors should learn how to appropriate Wikipedia content. This involves both learning how to appropriately get their site linked from Wikipedia (Plaza, 2011) and legally incorporating Wikipedia articles to enrich site content, as we suggested in the implications for researchers.

**Conclusion**

This systematic review has examined the scholarly research on readership of Wikipedia. We first presented summaries of scholarly studies grouped according to topics covering the ranking and popularity of Wikipedia; Wikipedia as a knowledge source for news, health information, and for scholars and librarians; student readership; reader perceptions of credibility; and commercial aspects of Wikipedia. We next discussed trends in the research details of the study, such as years of publication, knowledge domains covered, research design, and so on. We ended with a discussion of some of the most striking findings, and implications for researchers and for the Wikipedia community and readers.

Given Wikipedia's robust framework and wide popularity, it will surely continue to be a dominant force in Web reference for many years to come. The scholarly research has found that Wikipedia is popular not only for lighter topics such as entertainment, but also for more serious topics such as health information and legal background. Scholars, librarians and students are common users of Wikipedia, and it provides a unique opportunity for educating students in digital literacy. Wikipedia is an important and accessible subject for research in a wide variety of domains. Although a wide variety of subjects have already been undertaken, we believe that more nuanced studies of how readers assess Wikipedia's credibility could be conducted. We also feel that, despite a fair number of studies in that direction, there is yet much to be explored concerning commercial aspects of Wikipedia. By better understanding the nature of Wikipedia readership and its potential applications, such studies could help extend the reach of this valuable resource even further.

# Acknowledgments

# References

Antin, J., & Cheshire, C. (2010). Readers are not free-riders: reading as a form of participation on wikipedia. *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, CSCW '10 (pp. 127–130). New York, NY, USA: ACM. Retrieved October 22, 2012, from http://doi.acm.org/10.1145/1718918.1718942

Arazy, O., & Kopak, R. (2011). On the measurability of information quality. *Journal of the American Society for Information Science and Technology*, *62*(1), 89–99.

Arthur, C. (2006, July 20). What is the 1% rule? *The Guardian*. Retrieved April 3, 2013, from http://www.guardian.co.uk/technology/2006/jul/20/guardianweeklytechnologysection2

Aycock, J., & Aycock, A. (2008). Why I Love/Hate Wikipedia: Reflections upon (Not Quite) Subjugated Knowledges. *Journal of the Scholarship of Teaching and Learning*, *8*.

Belden, D. (2008). Harnessing Social Networks to Connect with Audiences : If You Build It, Will They Come 2.0? *Internet Reference Services Quarterly*, *13*(1), 99–111.

Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., & Hellmann, S. (2009). DBpedia - A crystallization point for the Web of Data. *Journal of Web Semantics*, *7*(3), 154–165.

Blumenstock, J. E. (2008). Size Matters: Word Count as a Measure of Quality on Wikipedia. *Proceedings of the 17th International World Wide Web Conference (WWW2008). April 21-25, 2008. Beijing, China*.

Breinholt, J. (2008). *The Wikipediazation of the American Judiciary*. Retrieved from http://nefafoundation.org/miscellaneous/justiceforall/nefawikipedia0108.pdf

Bruns, A. (2006). Wikinews: The Next Generation of Online News? *Scan Journal*, *3*(1). Retrieved March 18, 2013, from http://eprints.qut.edu.au/4862

Calkins, S., & Kelley, M. R. (2009). Who Writes the Past? Student Perceptions of Wikipedia Knowledge and Credibility in a World History Classroom. *Journal on Excellence in College Teaching*, *20*.

Cedergren, M. (2003). Open content and value creation. *First Monday*, *8*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/1071/991

Chandler-Olcott, K. (2009). Digital Literacies. A Tale of Two Tasks: Editing in the Era of Digital Literacies. *Journal of Adolescent & Adult Literacy*, *53*.

Chandy, R. (2009). Wikiganda: Identifying Propaganda Through Text Analysis. *Caltech Undergraduate Research Journal*, 6–11.

Chen, C.-J. (2009). Art history: a guide to basic research resources. *Collection building*, *28*(3), 122–125.

Chen, H. (2010). The perspectives of higher education faculty on Wikipedia. *Electronic Library*, Electron. Libr. (UK), *28*(3), 361–73.

Chen, H.-L. (2009). The use and sharing of information from Wikipedia by high-tech professionals for work purposes. *Electronic library*, *27*(6), 893–905.

Choolhun, N. (2009). Google: to use, or not to use. What is the question? *Legal Information Management*, *9*(03), 168.

Chua, A. Y. K., Kaynak, S., & Foo, S. S. B. (2007). An analysis of the delayed response to hurricane katrina through the lens of knowledge management. *Journal of the American Society for Information Science and Technology*, *58*(3), 391–403.

Cimini, N. (2010). Struggles online over the meaning of "Down"s syndrome': A "dialogic" interpretation. *Health (London, England: 1997)*, *14*(4), 398–414.

Coleman, A. (2007). Students `should use Wikipedia'. *BBC NEWS*. Retrieved from http://news.bbc.co.uk/2/hi/technology/7130325.stm

Cross, T. (2006). Puppy smoothies: improving the reliability of open, collaborative wikis. *First Monday*, *11*(9). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/1400/1318

Davis, C., Nikolic, I., & Dijkema, G. P. J. (2010). Industrial ecology 2.0. *Journal of Industrial Ecology*, *14*(5), 707–726.

DiStaso, M. W., & Messner, M. (2010). Forced transparency: Corporate image on Wikipedia and what it means for public relations. *Public Relations Journal*, *4*(2), 1–23.

Dondio, P., & Barrett, S. (2007). Computational trust in Web content quality: a comparative evalutation on the Wikipedia project. *Informatica*, Informatica (Slovenia), *31*(2), 151–60.

Dooley, P. L. (2010). Wikipedia and the two-faced professoriate. *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*. New York, NY, USA: ACM.

Eijkman, H. (2010). Academics and Wikipedia: Reframing Web 2.0+as a disruptor of traditional academic power-knowledge arrangements. *Campus-Wide Information Systems*, Campus-Wide Inf. Syst. (UK), *27*(3), 173–85.

Fiore, K. (2011). *APA: Med Students Cram for Exams With Wikipedia*. Retrieved from http://www.medpagetoday.com/MeetingCoverage/APA/26483

Flanagin, A. J., & Metzger, M. J. (2011). From Encyclopædia Britannica to Wikipedia. *Information, Communication & Society*, *14*(3), 355–374.

Francke, H., Sundin, O., & Limberg, L. (2011). Debating credibility: the shaping of information literacies in upper secondary school. *Journal of Documentation*, *67*(4), 675–694.

Gehl, R. (2010). *A cultural and political economy of Web 2.0*. George Mason University, United States -- Virginia. Retrieved from http://proquest.umi.com/pqdweb?did=2035889541&Fmt=7&clientId=10306&RQT=309&VName=PQD

Gunnels, C. B. (2007). Librarians on the verge of an epistemological breakdown. *Community & Junior College Libraries*, Community Jr. Coll. Libr. (USA), *14*(2), 111–20.

Gunnels, C. B., & Sisson, A. (2009). Confessions of a Librarian or: How I Learned to Stop Worrying and Love Google. *Community & Junior College Libraries*, *15*(1), 15–21.

Hahn, J. (2009). On the remediation of Wikipedia to the iPod. *Reference Services Review*, *37*(3), 272–285.

Hahn, J. (2010). Information seeking with Wikipedia on the iPod Touch. *Reference services review*, *38*(2), 284–298.

Haigh, C. A. (2010). Wikipedia as an evidence source for nursing and healthcare students. *Nurse Education Today*. Retrieved October 29, 2010, from http://www.ncbi.nlm.nih.gov/pubmed/20646799

Hardy, M. (2007). Wiki Goes to War. *AQ - Journal of Contemporary Analysis*, *79*(4), 17–22.

Harouni, H. (2009). High School Research and Critical Literacy: Social Studies with and Despite Wikipedia. *Harvard Educational Review*, *79*(3), 473–493.

Head, A. J. (2007). Beyond Google: How do students conduct academic research? *First Monday*, *12*(8). Retrieved March 7, 2013, from http://firstmonday.org/article/view/1998/1873http://firstmonday.org/article/view/1998/1873

Head, A. J., & Eisenberg, M. B. (2010). How today's college students use Wikipedia for course-related research. *First Monday*, *15*(3). Retrieved from http://firstmonday.org/article/view/2830/2476

Heilman, J. M., Kemmann, E., Bonert, M., Chatterjee, A., Ragar, B., Beards, G. M., Iberri, D. J., et al. (2011). Wikipedia: A Key Tool for Global Public Health Promotion. *Journal of Medical Internet Research*, *13*(1), e14.

Hickerson, C. A., & Thompson, S. R. (2009). Dialogue through wikis: a pilot exploration of dialogic public relations and wiki websites. *PRism*, *6*(1). Retrieved October 10, 2012, from http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Hickerson_Thompson.pdf

Hochstotter, N., & Lewandowski, D. (2009). What users see - Structures in search engine results pages. *Information Sciences*, *179*(12), 1796–1812.

Hu, M., Lim, E.-P., Sun, A., Lauw, H. W., & Vuong, B.-Q. (2007). Measuring article quality in wikipedia: Models and evaluation. *16th ACM Conference on Information and Knowledge Management, CIKM 2007, November 6, 2007 - November 9, 2007*, International Conference on Information and Knowledge Management, Proceedings (pp. 243–252). Lisboa, Portugal: Association for Computing Machinery. Retrieved November 5, 2010, from http://dx.doi.org/10.1145/1321440.1321476

Hughes, B., Joshi, I., Lemonde, H., & Wareham, J. (2009). Junior physician's use of Web 2.0 for information seeking and medical education: A qualitative study. *International Journal of Medical Informatics*, *78*(10), 645–655.

Huss, J. W., Lindenbaum, P., Martone, M., Roberts, D., Pizarro, A., Valafar, F., Hogenesch, J. B., et al. (2010). The Gene Wiki: community intelligence applied to human gene annotation. *Nucleic Acids Research*, *38*(Database issue), D633–639.

Bar-Ilan, J. (2006). Web links and search engine ranking: the case of Google and the query "Jew." *Journal of the American Society for Information Science and Technology*, *57*(12), 1581–1589.

Jacobs, M. L. (2009). Libraries and the mobile revolution: remediation = relevance. *Reference services review*, *37*(3), 286–290.

Jancarik, A., & Jancarikova, K. (2010). Wiki Tools in the Preparation and Support of e-Learning Courses. *Electronic Journal of e-Learning*, Electron. J. e-Learn. (UK), *8*(2), 123–32.

Jennings, E. (2008). Using Wikipedia to Teach Information Literacy. *College & Undergraduate Libraries*, *15*(4), 432–437.

Johnson, P. T., Chen, J. K., Eng, J., Makary, M. A., & Fishman, E. K. (2008). A Comparison of World Wide Web Resources for Identifying Medical Information. *Academic Radiology*, *15*(9), 1165–1172.

Jordan, C. (2009). *Contextual Retrieval of Single Wikipedia Articles to Support the Reading of Academic Abstracts* (PhD dissertation). Dalhousie University (Canada). Retrieved from http://proquest.umi.com/pqdlink?Ver=1&Exp=10-14-2017&FMT=7&DID=1630234261&RQT=309&attempt=1&cfc=1

Jordan, C., & Watters, C. (2009). Addressing gaps in knowledge while reading. *Journal of the American Society for Information Science and Technology*, *60*(11), 2255–2268.

Judd, T., & Kennedy, G. (2010). A five-year study of on-campus Internet use by undergraduate biomedical students. *Computers and Education*, *55*(4), 1564–1571.

Judd, T., & Kennedy, G. (2011). Expediency-based practice? Medical students' reliance on Google and Wikipedia for biomedical inquiries. *British Journal of Educational Technology*, *42*(2), 351–360.

Jullien, N. (2012). What we know about Wikipedia. A review of the literature analyzing the project(s). *Available at SSRN 2053597*. Retrieved October 4, 2012, from https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2053597_code728676.pdf?abstractid=2053597&mirid=2

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, Bus. Horiz. (USA), *53*(1), 59–68.

Knapp, M. M. (2008). eBay, Wikipedia, and the Future of the Footnote. *Theatre History Studies*, *28*(1), 36–41.

Kollock, P., & Smith, M. (1996). Managing the virtual commons: Cooperation and conflict in computer communities. In S. C. Herring (Ed.), *Computer - Mediated Communication: Linguistic, Social and Cross - Cultural Perspectives* (pp. 109–128). John Benjamins Publishing.

Koolen, M., Kazai, G., & Craswell, N. (2009). Wikipedia pages as entry points for book search. *WSDM '09: Proceedings of the Second ACM International Conference on Web Search and Data Mining* (pp. 44–53). New York, NY, USA: ACM.

Korfiatis, N. T., Poulos, M., Bokos, G., & Sicilia, M.-A. (Editor. (2006). Evaluating authoritative sources using social networks : an insight from Wikipedia. *Online information review (Print)*, *30*(3), 252–262.

Korosec, L., Limacher, P. A., Lüthi, H. P., & Brändle, M. P. (2010). Chemical Information Media in the Chemistry Lecture Hall: A Comparative Assessment of Two Online Encyclopedias. *CHIMIA International Journal for Chemistry*, *64*(5), 309–314.

Kubiszewski, I., Noordewier, T., & Costanza, R. (2011). Perceived credibility of Internet encyclopedias. *Computers & Education*, *56*(3), 659–667.

Lanamäki, A., Okoli, C., Mehdi, M., & Mesgari, M. (2011). Protocol for Systematic Mapping of Wikipedia Studies. *Proceedings of IRIS 2011 – The 34th Information Systems Research Seminar in Scandinavia*. Presented at the Information Systems Research Seminar in Scandinavia, Turku, Finland.

Langlois, G., & Elmer, G. (2009). Wikipedia leeches? The promotion of traffic through a collaborative web format. *New Media & Society*, *11*(5), 773–794.

Laurent, M. R., & Vickers, T. J. (2009). Seeking health information online: does Wikipedia matter? *Journal of the American Medical Informatics Association*, *16*(4), 471–479.

Lavsa, S. M., Corman, S. L., Culley, C. M., & Pummer, T. L. (2011). Reliability of Wikipedia as a medication information source for pharmacy students. *Currents in Pharmacy Teaching and Learning*, *3*(2), 154–158.

Lewandowski, D., & Spree, U. (2011). Ranking of Wikipedia articles in search engines revisited: Fair ranking for reasonable quality? *Journal of the American Society for Information Science and Technology*, *62*(1), 117–132.

Li, C., & Bernoff, J. (2011). *Groundswell: Winning in a World Transformed by Social Technologies*. Harvard Business Press.

Lih, A. (2004). Wikipedia as participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource. *5th International Symposium on Online Journalism*. Retrieved from http://jmsc.hku.hk/faculty/alih/publications/utaustin-2004-wikipedia-rc2.pdf

Lim, S. (2009). How and why do college students use Wikipedia? *Journal of the American Society for Information Science and Technology (Print)*, *60*(11), 2189–2202.

Lim, S., & Kwon, N. (2010). Gender differences in information behavior concerning Wikipedia, an unorthodox information source? *Library & information science research*, *32*(3), 212–220.

Luyt, B., Ally, Y., NurHakimLow, & NorahBinteIsmail. (2010). Librarian Perception of Wikipedia: Threats or Opportunities for Librarianship? *Libri (Copenhagen)*, *60*(1), 57–64.

Luyt, B., Zainal, C., Mayo, O., & Yun, T. (2008). Young people's perceptions and usage of Wikipedia. *Information Research*, *13*(4). Retrieved October 29, 2010, from http://apps.isiknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=4&SID=3D9@K9HgCKnBM85J1eH&page=2&doc=82

Maehre, J. (2009). What It Means to Ban Wikipedia: An Exploration of the Pedagogical Principles at Stake. *College Teaching*, *57*(4), 229–36.

Martin, O. S. (2010). A Wikipedia Literature Review. *arXiv:1110.5863*. Retrieved October 4, 2012, from http://arxiv.org/abs/1110.5863

McGuinness, D. L., Zeng, H., Silva, P. P. da, Ding, L., Narayanan, D., & Bhaowal, M. (2006). Investigations into Trust for Collaborative Information Repositories: A Wikipedia Case Study. *Proceedings of the Workshop on Models of Trust for the Web*. Edinburgh, United Kingdom. Retrieved from http://ebiquity.umbc.edu/_file_directory_/papers/274.pdf

Medelyan, O., Milne, D., Legg, C., & Witten, I. H. (2009). Mining meaning from Wikipedia. *International Journal of Human Computer Studies*, *67*(9), 716–754.

Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. Å., & Lanamäki, A. (2014). "The sum of all human knowledge": A systematic review of scholarly research on the content of Wikipedia. *Journal of the American Society for Information Science and Technology*, *Forthcoming*.

Messner, M., & South, J. (2010). Legitimizing Wikipedia: How US national newspapers frame and use the online encyclopedia in their coverage. *Journalism Practice*. Retrieved from http://www.informaworld.com/10.1080/17512786.2010.506060

Moeller, E. (2009). *Wikipedia Scholarly Survey Results*. Wikimedia Foundation.

Mühlhauser, I., & Oser, F. (2008). Does Wikipedia provide evidence-based health care information? A content analysis. *Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen*, *102*(7), 441–448.

Nadamoto, A., Aramaki, E., Abekawa, T., & Murakami, Y. (2010). Extracting content holes by comparing community-type content with Wikipedia. *International Journal of Web Information Systems*, *6*, 248–260.

Nielsen, F. Å. (2012, February 6). Wikipedia Research and Tools: Review and Comments. SSRN. Retrieved from SSRN.

Nielsen, J. (2006). Participation inequality: Encouraging more users to contribute. *Jakob Nielsen's alertbox*. Retrieved April 3, 2013, from http://www.nngroup.com/articles/participation-inequality/

Nonnecke, B., Andrews, D., & Preece, J. (2006). Non-public and public online community participation: Needs, attitudes and behavior. *Electronic Commerce Research*, *6*(1), 7–20.

Okoli, C. (2009). A brief review of studies of Wikipedia in peer-reviewed journals. *Digital Society, 2009. ICDS'09. Third International Conference on* (pp. 155–160). Retrieved September 17, 2012, from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4782868

Okoli, C., Mehdi, M., Mesgari, M., Nielsen, F., & Lanamäki, A. (2012, October 24). *The People's Encyclopedia Under the Gaze of the Sages: A Systematic Review of Scholarly Research on Wikipedia*. SSRN Scholarly Paper, Montreal. Retrieved February 18, 2013, from http://papers.ssrn.com/abstract=2021326

Okoli, C., & Schabram, K. (2009a). Protocol for a systematic literature review of research on the Wikipedia. *Sprouts: Working Papers in Information Systems*, *9*(65). Retrieved from http://sprouts.aisnet.org/9-65

Okoli, C., & Schabram, K. (2009b). Protocol for a systematic literature review of research on the Wikipedia. *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES)* (p. 73). Presented at the International ACM Conference on Management of Emergent Digital EcoSystems (MEDES), Lyon, France: Association for Computing Machinery. Retrieved September 17, 2012, from http://dl.acm.org/citation.cfm?id=1643912

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. Retrieved September 17, 2012, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1954824

Okoli, C., Schabram, K., & Kader, B. A. (2009). From the Academy to the Wiki: Practical Applications of Scholarly Research on Wikipedia. *Proceedings of Wikimania*. Presented at the Wikimania, Buenos Aires: Wik. Retrieved September 17, 2012, from http://chitu.okoli.org/images/stories/bios/pro/research/open/Okolietal2009Wikimania.pdf

Page, R. (2010). Wikipedia as an encyclopaedia of life. *Organisms Diversity & Evolution*, *10*(4), 343–349.

Patch, P. (2010). Meeting Student Writers Where They Are: Using Wikipedia to Teach Responsible Scholarship. *Teaching English in the Two-Year College*, *37*(3), 278–85.

Peoples, L. F. (2009). The Citation of Wikipedia in Judicial Opinions. *Yale Journal of Law & Technology*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1272437

Perez, J. C. (2007, February 17). Wikipedia Breaks Into U.S. Top 10 Sites. *PCWorld*. Retrieved March 18, 2013, from http://www.pcworld.com/article/129135/article.html

Plaza, B. (2011). Google Analytics for measuring website performance. *Tourism Management*, *32*(3), 477–481.

Preece, J., Nonnecke, B., & Andrews, D. (2004). The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior*, *20*(2), 201–223.

Rafaeli, S., Ravid, G., & Soroka, V. (2004). De-Lurking in Virtual Communities: A Social Communication Network Approach to Measuring the Effects of Social and Cultural Capital. *Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 7 - Volume 7*, HICSS '04 (p. 70203–). Washington, DC, USA: IEEE Computer Society. Retrieved April 3, 2013, from http://dl.acm.org/citation.cfm?id=962755.963153

Rahman, M. M. (2006). *Essays analyzing blogs and Wikipedia*. The University of Kansas, United States - - Kansas. Retrieved from http://proquest.umi.com/pqdweb?did=1126778281&Fmt=7&clientId=10306&RQT=309&VName=PQD

Rahman, M. M. (2008). An Analysis of Wikipedia. *Journal of Information Technology Theory and Application (JITTA)*, *9*(3), 81.

Rand, A. D. (2010). Mediating at the Student-Wikipedia Intersection. *Journal of Library Administration*, *50*(7/8), 923–932.

Ratkiewicz, J., Fortunato, S., Flammini, A., Menczer, F., & Vespignani, A. (2010). Characterizing and modeling the dynamics of online popularity. *Physical Review Letters*, *105*(15). Retrieved November 5, 2010, from http://dx.doi.org/10.1103/PhysRevLett.105.158701

Read, B. (2007, February 16). Middlebury College History Department Limits Students' Use of Wikipedia. *The Chronicle of Higher Education*, *53*(24), A39.

Rubin, A., & Rubin, E. (2010). Informed Investors and the Internet. *Journal of Business Finance & Accounting*, *37*(7-8), 841–865.

Schneider, A., von Krogh, G., & Jäger, P. (2013). "What's coming next?" Epistemic curiosity and lurking behavior in online communities. *Computers in Human Behavior*, *29*(1), 293–303.

Schweitzer, N. J. (2008). Wikipedia and Psychology: Coverage of Concepts and Its Use by Undergraduate Students. *Teaching of Psychology*, *35*(2), 81–85.

Shaw, D. (2008). Wikipedia in the newsroom. *American Journalism Review*, *30*(1), 40–45.

Spoerri, A. (2007a). What is popular on Wikipedia and why? *First Monday*, *12*(4). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/1765/1645

Spoerri, A. (2007b). Visualizing the overlap between the 100 most visited pages on Wikipedia for September 2006 to January 2007. *First Monday*, *12*(4). Retrieved from http://firstmonday.org/article/view/1764/1644

Stoddard, M. M. (2009, April). *Judicial Citation to Wikipedia in Published Federal Court Opinions*. School of Information and Library Science of the University of North Carolina at Chapel Hill.

Sundin, O., & Francke, H. (2009). In search of credibility: pupils' information practices in learning environments. *Information Research: An International Electronic Journal*, *14*(4), 19 pp.

Tann, C., & Sanderson, M. (2009). Are web-based informational queries changing? *Journal of the American Society for Information Science and Technology*, *60*(6), 1290–1293.

Thelwall, M., & Hellsten, I. (2006). The BBC, Daily Telegraph and Wikinews timelines of the terrorist attacks of 7th July 2006 in London: a comparison with contemporary discussions. *Information Research*, *12*(1), 14.

Thelwall, M., & Stuart, D. (2007). RUOK? Blogging Communication Technologies During Crises. *Journal of Computer-Mediated Communication*, *12*(2007), 523–548.

Thorsen, E. (2008). Journalistic objectivity redefined? Wikinews and the neutral point of view. *New Media & Society*, *10*(6), 935–954.

Veltman, K. H. (2005). Access, claims and quality on the internet - Future challenges. *Progress in Informatics*, (2), 17–40.

Waller, V. (2011). The search queries that took Australian Internet users to Wikipedia. *Information Research*, *16*(2). Retrieved July 28, 2011, from http://informationr.net/ir/16-2/paper476.html

Waters, N. L. (2007). Why you can't cite Wikipedia in my class. *Communications of the ACM*, *50*(9), 15–17.

Wedemeyer, B., Yakubova, N., Kallenbach, J., Ekdahl, A., Lesko, L., Reed, E., & Schwartz, K. (2008). Quality of the science articles on the English Wikipedia: Preliminary results. *Wikimania 2008*. Presented at the Wikimania, Alexandria, Egypt. Retrieved from http://www.youtube.com/watch?v=B7bCZbHHeZI

Wikipedia contributors. (2013, March 1). Wikinews. *Wikipedia, the free encyclopedia*. Retrieved March 8, 2013, from http://en.wikipedia.org/w/index.php?title=Wikinews&oldid=538746746

Yasseri, T., & Kertész, J. (2013). Value production in a collaborative environment. *Journal of Statistical Physics*, *151*(3-4), 414–439.

Yermilov, I., Chow, W., Devgan, L., Makary, M. A., & Ko, C. Y. (2008). What Is the Quality of Surgery-Related Information on the Internet? Lessons Learned from a Standardized Evaluation of 10 Common Operations. *Journal of the American College of Surgeons*, *207*(4), 580–586.

Younger, P. (2010). Using wikis as an online health information resource. *Nursing Standard*, *24*(36), 49–56.

Zeng, H., Alhossaini, M. A., Ding, L., Fikes, R., & McGuinness, D. L. (2006). Computing trust from revision history. *STAR. Vol. 44* (Vol. 44). Retrieved from http://ebiquity.umbc.edu/_file_directory_/papers/302.pdf