

From the Academy to the Wiki: Practical Applications of Scholarly Research on Wikipedia

Chitu Okoli^{*†}

*Department of Decision Sciences and MIS, Concordia University, 1455 De Maisonneuve Blvd. West,
Montreal, Quebec, H3G 1M8, Canada*

Kira Schabram

*Department of Management, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal,
Quebec, H3G 1M8, Canada*

Bilal Abdul Kader

*Department of Finance, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Quebec,
H3G 1M8, Canada*

SUMMARY.

To date, over 400 peer-reviewed scholarly studies have researched various aspects of Wikipedia. These studies contribute valuable knowledge in understanding the inner workings of Wikipedia and can serve to continuously improve it. In this presentation, we offer a coherent synthesis of the scholarly research that has been conducted on Wikipedia, highlighting the main research trends, summarizing the key findings, and identifying the gaps and unanswered questions. Based on these findings, we apply the research conclusions to the practical functioning of Wikipedia and highlight implications for policy and administration for Wikipedia contributors, administrators, and the Wikimedia Foundation.

KEY WORDS: Wikipedia, literature review, open content, open source, Management, Human Factors, Theory.

1. From the Academy to the Wiki

Wikipedia's phenomenal success has attracted the interest of scholars who desire to understand the inner workings of this exemplary open content application. The goals of their studies have varied widely, ranging from attempting to understand open content creation mechanisms, to investigating Wikipedia itself as a web-based encyclopedia, to using Wikipedia's abundant data as a source for various kinds of research. Much of this research can prove valuable in guiding Wikipedia contributors and administrators, on developing policies and best practices to improve the quality, performance, and overall value of Wikipedia.

Employing rigorous and replicable academic standards, this proposed presentation will offer a comprehensive review and analysis of the abundant body of scholarly work on Wikipedia to date, offering the layman a clear and easy overview of the scholarly research.

* Correspondence to: Chitu Okoli, Department of Decision Sciences and MIS, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada

† E-mail: chitu.okoli@concordia.ca

Although the Wikimedia Foundation website informally catalogues a large number of scholarly publications, these lists currently include a mere fraction of existing studies; furthermore, there is no attempt at thorough classification and analysis. Only three brief literature reviews have thus far attempted to sum up the body of academic inquiry (Ayers 2006; Miller 2007; Okoli 2009); however, none of these has offered an in-depth critical analysis of the entire field.

In response to this shortage, this proposed presentation will analyze particular trends in research and offer the basic groundwork for practitioners to leverage the research findings to build a better Wikipedia. While there has been very valuable research and publications carried out by non-academics, this presentation will focus only on peer-reviewed scholarly work, that is, research that has been vetted by fellow academics as being of high quality and of valuable contribution.

This review of studies is being executed with the highest standards of academic rigour, to ensure that no relevant study is missed, and that meaningful and useful analyses and conclusions are drawn. We began our exhaustive review of research on Wikipedia by searching over 480 article databases, each comprising hundreds of scholarly journals and conference proceedings (though with many overlaps across databases). The databases we searched covered business, fine arts, humanities, science and engineering, and social sciences. This search resulted in approximately 400 peer-reviewed studies. While we have hundreds of articles in our review, here we highlight only a few samples in some major categories to illustrate the nature of our presentation.

A large body of research investigates various aspects of how Wikipedia works. Schroer and Hertel (2009), Wagner (2006), and Nov (2007) unearthed various motivations for participating in Wikipedia such as perceived task characteristics like autonomy, skill variety, and self reported engagement; however, they found that ideological and altruistic motives were the most relevant. Such findings are important in attracting and retaining contributors. Research on the psychology and sociology of Wikipedia has studied occurrences of team work (Salz 2007; Winder 2005), volunteer motivation (Allen 2005; Amichai-Hamburger *et al.* 2008; Nov 2007), vandalism (Potthast *et al.* 2008; Riedl 2008), authoritarian tendencies in an egalitarian structure (Reagle 2007) and the social networking of niche groups such as sports fans (Ferriter 2009). Such studies are valuable in learning how to build the camaraderie among Wikipedia contributors, and developing a sense of community among them. On the technological side, Wikipedia research includes analysis of basic statistics and of semantic structure (Holloway *et al.* 2007), coding (Ponzetto and Strube 2007), content tagging (Yang *et al.* 2009) and topic mapping (Junghoon *et al.* 2007). Such studies are invaluable for charting and developing the technological infrastructure of Wikipedia.

Probably the earliest and most contentious research stream on Wikipedia has been the debate on its reliability and quality. The literature ranges from the initial debate between famous critics (“Why you can’t cite Wikipedia in my classroom”, (Waters 2007) and steadfast proponents (Crovitz and Smoot 2009; Murley 2008). Specific areas of focus include attempts to empirically compare Wikipedia to other reference works (Chesney 2006; Greenstein 2007; Rosenzweig 2006); propose methods to improve trust (Kittur and Kraut 2008; Wilkinson and Huberman 2007); or study the effect of Wikipedia trust in particular populations, such as teenagers (Luyt *et al.* 2008). Since scholars are some of the most important producers and consumers of reference materials, and since they have a tremendous influence on their students, it is imperative to understand their concern and valuations of reliability and quality, in order to know how best to steer Wikipedia’s improvement efforts in their direction.

Various research has studied real-world applications of Wikipedia. In education, Goodman (2008) and Pollard (2008) described biology and history courses that require contribution to Wikipedia, Hoorne and van Hoorne (2008) showed how the collaborative nature of Wikipedia is exploited in legal education in the Netherlands, and Achterman (2005) explored the survival chance of high school library media programs in the face of Web 2.0. In medicine, Muelhauser and Oser (2008) examined patients and consumers who increasingly use Wikipedia as a research tool, while Mercer (2007) described the use of Wikipedia by mental health patients. Other practical applications of Wikipedia include its use by journalists (Shaw 2008), banks (Sausner 2007) and genetics research (Hoehndorf *et al.* 2006). These articles critically study models of applications and extensions of Wikipedia, helping practitioners understand how best to use Wikipedia in practice and how to model it for similar applications.

Our presentation at Wikimania would have two main focuses. We will first present a coherent synthesis of the scholarly research that has been conducted on Wikipedia, highlighting the main research trends, summarizing the key findings, and identifying the gaps and unanswered questions. Then, based on these findings, we will apply the research conclusions to the practical functioning of Wikipedia and highlight implications for policy and administration for Wikipedia contributors, administrators, and the Wikimedia Foundation.

We believe that scholarly research is a critical contributor to thoroughly understanding the workings of Wikipedia, and an invaluable tool for improving this resource to eventually becoming indisputably the best encyclopedia in the world. We believe our presentation would help bring this understanding to the people who actively construct Wikipedia, enabling them to leverage this valuable knowledge base.

1.1. Acknowledgement

This study is funded by the Social Sciences and Humanities Research Council of Canada.

REFERENCES

- Achterman D. 2005. Surviving wikipedia. *Knowledge Quest* 33(5): 38-40.
- Allen C. 2005. Wikipedia: All-volunteer instant encyclopedia. *Tennessee Libraries* 55(4): 29-36.
- Amichai-Hamburger Y, Lamdan N, Madiel R, Hayat T. 2008. Personality characteristics of wikipedia members. *CyberPsychology & Behavior* 11(6): 679-681.
- Ayers P. 2006. Researching wikipedia - current approaches and new directions. *Proceedings of the American Society for Information Science and Technology* 43(1): 1-14.
- Chesney T. 2006. An empirical examination of Wikipedia's credibility. *First Monday* 11(1).
- Crovitz D, Smoot WS. 2009. Wikipedia: Friend, not foe. *English Journal* 98.
- Ferriter MM. 2009. "Arguably the greatest": Sport fans and communities at work on Wikipedia. *Sociology of Sport Journal* 26(1): 127-154.
- Goodman R. 2008. Students contribute to a global community through improvement of Wikipedia. *American Biology Teacher* 70(3): 23.
- Greenstein S. 2007. Wagging Wikipedia's long tail. *IEEE Micro* 27(2): 6-79.
- Hoehndorf R, Pruffer K, Backhaus M, Herre H, Kelso J, Loebe F, Visagie J. 2006. A proposal for a gene functions wiki. In *On the Move to Meaningful Internet Systems 2006 Proceedings, Part I*. Berlin, Germany: 669-78.

- Holloway T, Bozicevic M, Borner K. 2007. Analyzing and visualizing the semantic coverage of Wikipedia and its authors. *Complexity* 12(3): 30-40.
- Hoorn E, van Hoorn D. 2008. Critical assessment of using wikis in legal education. *JILT-Journal of Information Law & Technology* 1: 28-37.
- Junghoon Y, Jangwhan H, Inseok O, Mingyung K. 2007. Using Wikipedia technology for topic maps design. In *Annual Southeast Conference Winston-Salem, NC Conference Proceedings*, 106-110.
- Kittur A, Kraut RE. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *ACM San Diego, CA Conference Proceedings*, 37-46.
- Luyt B, Zanail C, Mayo OVP, Yun TS. 2008. Young people's perceptions and usage of Wikipedia. *Information Research* 13(4): 50.
- Mercer J. 2007. Wikipedia and 'open source' mental health information. *Scientific Review of Mental Health Practice* 5(1): 88-92.
- Miller N. 2007. Wikipedia revisited. *ETC: A Review of General Semantics* 64(2): 147-150.
- Mühlhauser I, Oser F. 2008. Sind medizinische und Gesundheitsinformationen auf den Internetseiten von Wikipedia evidenzbasiert? - Eine Inhaltsanalyse. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 102(7): 441-448.
- Murley D. 2008. In *Defense of Wikipedia*. *Law Library Journal* 100(3): 593-599.
- Nov O. 2007. What motivates Wikipedians? *Communications of the ACM* 50(11), 60-64.
- Okoli C. 2009. A brief review of studies of wikipedia in peer-reviewed journals. In *Third International Conference on Digital Society Cancun, Mexico Proceedings*. 155-160.
- Pollard EA. 2008. Raising the stakes: writing about witchcraft on wikipedia. *History Teacher* 42(1): 9-24.
- Ponzetto S, Strube M. 2007. Knowledge derived from Wikipedia for computing semantic relatedness. *Journal of Artificial Intelligence Research* 30: 181-212.
- Pothast M, Stein B, Gerling R. 2008. Automatic vandalism detection in Wikipedia. In *Lecture Notes in Computer Science*, Springer Verlag, Glasgow, UK: 663-668.
- Reagle Jr JM. 2007. Do as I do: Authorial leadership in Wikipedia. In *Object-Oriented Programming Systems, Languages, and Applications Montreal, Canada Conference Proceedings*, 143-156.
- Riedl J. 2008. Altruism, selfishness, and destructiveness on the social web. In *Lecture Notes in Computer Science*, Springer Verlag, Hannover, Germany: 9-11
- Rosenzweig R. 2006. Can history be open source? Wikipedia and the future of the past. *Journal of American History* 93(1): 117-146.
- Salz PA. 2007. Teams work. *EContent* 30(9): 34-40.
- Sausner R. 2007. Wouldn't wikis be wicked wonderful? *US Banker* 117(5): 22-3.
- Schroer J, Hertel G. 2009. Voluntary engagement in an open web-based encyclopedia: Wikipedians and why they do it. *Media Psychology* 12(1): 96-120.
- Shaw D. 2008. Wikipedia in the newsroom. *American Journalism Review* 30(1): 40-45.
- Wagner C. 2006. Breaking the knowledge acquisition bottleneck through conversational knowledge management. *Information Resources Management Journal* 19(1): 70.
- Waters NL., 2007. Why you can't cite Wikipedia in my class. *Communications of the ACM* 50(9): 15-17.

Wilkinson DM, Huberman BA. 2007. Cooperation and quality in wikipedia. In ACM Montreal, QC Conference Proceedings: 157-164.

Winder D. 2005. Team working. Information World Review (219): 41-2.

Yang Y, Bansal N, Dakka W, Ipeirotis P, Koudas N, Papadias D. 2009. In ACM Barcelona, Spain Conference Proceedings: 34-43.