Perspective

"Blind Trust is Not Enough": Considering Practical Verifiability and Open Referencing in Wikipedia

Cal Murgu and Krisandra Ivings

ABSTRACT

This article draws attention to the often-unseen information inequalities that occur in the way that Wikipedia content is referenced. Drawing on digital information control and virtual gatekeeping scholarship, we contend that by not considering the degree to which references are practically accessible and verifiable, Wikipedia editors are implicitly promoting a control mechanism that is limiting the potential of Wikipedia to serve, as Willinsky puts it, "as a gateway to a larger world of knowledge." We question the widespread practice of referencing peer-reviewed literature that is obscured by prohibitive paywalls. As we see it, two groups are disadvantaged by this practice. First, Wikipedia editors who lack access to the most current scholarly literature are unable to verify a reference or confirm the veracity of a fact or figure. Second, and perhaps most important, general readers are left with two unsatisfactory options in this scenario: they can either trust the authority of the citation or pay to access the article. Building on Don Fallis' work on epistemic consequences of Wikipedia, this paper contends that promoting practically verifiable references would work to mitigate these inequalities and considers how relatively new initiatives could be used to improve the quality and utility of Wikipedia more generally.

INTRODUCTION

"The news is broken, but we figured out how to fix it," declared the founder of Wikipedia, Jimmy Wales, in an April 2017 video announcing the launch of a new crowd-sourced journalism platform, Wikitribune. "On the internet, nobody is guarding the gate," he continued, "it's time to rethink the gatekeeper."[1] Wales' response to the diminishing quality of online, ad-financed journalism in an age of "alt-facts" and "misinformation" is to bring together high-quality journalism, volunteer fact-checkers, and grammarians on a wiki platform and make the product freely available to all. This initiative adopts the most celebrated qualities of Wikipedia, in particular, principles of transparency, openness, and the awesome potential of crowdsourced content creation. Indeed, Wales' description of Wikitribune is not that far removed from how he described Wikipedia in 2004: "Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge, that's what we're doing."[2]

Since 2001, Wales' flagship product, Wikipedia, has become the largest encyclopedia known to humankind. The English-language Wikipedia includes over 5.4 million unique articles, with over 31 million registered users, 129,237 of which are active contributors.[3] Wikipedia has become a trusted source for learning about unfamiliar concepts, historical figures, and other general knowledge. It can also be a convenient starting point for personal research. As John Willinsky aptly stated in 2007, "Wikipedia stands alongside open source software (such as Linux and Apache) as a new breed of impossible public goods that have been made possible by the Internet."[4] Moreover, what makes Wikipedia possible is the unexpected rigor of its contributors; we can infer from the surprising quality of many Wikipedia articles that they are keen and committed to the Wikipedia project.[5] The wiki platform has also been lauded for facilitating democratic

---

[1] Anthony Cuthberson, "Wikipedia Founder Jimmy Wales Believes He Can Fix Fake News with Wikitribune Project," *Newsweek,* April 25, 2017, https://www.newsweek.com/wikipedia-jimmy-wales-fake-news-wikitribune-crowdfunding-589226.

[2] "Wikipedia Founder Jimmy Wales Responds," *Slashdot,* July 28, 2004, https://slashdot.org/story/04/07/28/1351230/wikipedia-founder-jimmy-wales-responds

[3] "Wikipedia:Statistics," *Wikipedia,* accessed November 10, 2017, https://en.wikipedia.org/wiki/Wikipedia:Statistics

[4] John Willinsky, "What Open Access Research Can Do For Wikipedia," *First Monday* 12 (2007). https://firstmonday.org/ojs/index.php/fm/article/viewArticle/1624/1539.

[5] Several studies over the last decade have compared Wikipedia to other encyclopedias, like the *Encyclopedia Britannica.* See, for example, Jim Giles, "Internet Encyclopedias go Head to Head," *Nature* 438 (2005): 900-901; Imogen Casebourne, Chris Davies, Michelle Fernandes, and Naomi Norman, "Assessing the Accuracy and Quality of Wikipedia Entries Compared to Popular Online Encyclopaedias: A Comparative Preliminary Study Across Disciplines in English, Spanish and Arabic," https://upload.wikimedia.org/wikipedia/commons/2/29/EPIC_Oxford_report.pdf (accessed

participation in a digital information age. As Matthew Barton writes of wikis, "[they] are not weapons of hegemonic domination but tools for democratic liberation. Their power comes not from above but from below; not from publishers but plebeians; not from a single proud tyrant but from a thousand humble citizens."[6] Indeed, if the success of Wikipedia is any measure it seems that Wales' rallying call has already happened: there are very few gatekeepers on Wikipedia.

Or are there? In this paper we draw attention to the often-unseen information control that occurs in the way that Wikipedia content is created and referenced. We focus on Wikipedia's policy of verifiability and contend that in its current form and practice it does not do enough to encourage open citation practices. Specifically, we question the referenced literature blocked by prohibitive paywalls without considering and including links to open resources, a practice that, as we suggest, is encouraged by Wikipedia content creation policies. In a context of increased growth of open access content, we argue that this practice runs counter to the principles of transparency and openness which together underpin Wikipedia.[7] As a result, Wikipedia is impeded from becoming, as Willinsky puts it, "a gateway to a larger world of knowledge."[8] We find that two groups are disadvantaged by this practice. First, Wikipedia editors who lack access to the most current scholarly literature are unable to verify a reference or confirm the veracity of content. While publishers like Taylor & Francis and Elsevier have distributed free accounts to prolific editors to increase the quality of content in Wikipedia, these special dispensations have in effect created another level of inequality among editors; those who meet the criteria, and those who do not. Second, and perhaps most important, general readers are left with two unsatisfactory options in this scenario: they can either trust the authority of the citation or pay to access the article.

In putting this perspective forward, we do not mean to suggest that everything must be made open, nor are we promoting a Wikipedia editorial process that privileges

---

February 17, 2019). It is worth noting, however, that studies such as these often consider traditional encyclopedias (whether available electronically or not) and Wikipedia as analogous products. Yet as Morgan Currie argues, "research that assesses articles [in Wikipedia] for accuracy and comprehensiveness too often glosses over the fact that articles are constructed over time and on another medium entirely. Wikipedia is a unique reference source…" See Morgan Currie, "The Feminist Critique: Mapping Controversy in Wikipedia," in *Understanding Digital Humanities,* ed. David Berry (Houndmills, Basingstoke, UK: Palgrave Macmillan, 2012), 227.

[6] Matthew Barton, "Is There a Wiki in this Class?: Wikibooks and the Future of Higher Education," in *Wiki Writing: Collaborative Learning in the College Classroom,* eds. Matthew Barton and Robert Cummings (Ann Arbor, MI: University of Michigan Press, 2008), 184.

[7] For recent OA numbers, see https://www.sspnet.org/community/news/delta-think-releases 2018-updated-open-access-market-sizing-analysis/

[8] Willinsky, "What Open Access Research Can Do For Wikipedia."

open access work over paywalled content. We are cognizant, for example, of the disciplinary differences in the way that research is published, disseminated, challenged, and refuted; publisher embargo policies correlate with disciplinary publishing practices, including peer-review and differing rates of publication. We are arguing that in a context where openly licensed versions of paywalled materials are increasingly available — in absolute terms never before has so much scholarship been openly available — every effort should be made to identify and link to those resources.

To this end, we begin our analysis by discussing information control and gatekeeping theories in the context of digital crowdsourced knowledge platforms. Drawing on Jan van Dijk's work on information control in digital environments and Karine Barzilai-Nahon's research on virtual gatekeeping, we show how gatekeeping mechanisms, including editorial, channeling, and regulatory actions manifest in Wikipedia's editorial communities. The direct impacts of these gatekeeping mechanisms are clearly seen when considering the apparent gender bias in Wikipedia content. Subsequently, we consider the impact of Wikipedia's core content creation policies on another aspect of knowledge production and control: content verifiability. Specifically, using the discussions surrounding "WikiGate" as a springboard, we consider Wikipedia's position on content verifiability and the technical and practical accessibility of linked references in Wikipedia. Ultimately, drawing from scholarship on information control in digital environments, this paper contends that a greater emphasis on including references to openly licensed scholarship within Wikipedia would work to mitigate information inequalities, and considers how new initiatives could be used to improve the quality and accessibility of Wikipedia more generally.[9]

PRACTICES OF INFORMATION CONTROL IN DIGITAL ENVIRONMENTS AND WIKIPEDIA

The theoretical concept of the gatekeeper was introduced by the social psychologist Kurt Lewin in 1951. Gatekeeping refers to the practice of controlling goods, such as information, as they move from producers to users through an intermediary, and is considered by scholars of sociology, psychology, communications, and information

---

[9] Fallis asks whether Wikipedia has a positive epistemic consequence on society. Put another way, he questions whether Wikipedia, as a knowledge resource, helps or hinders society from becoming more informed. He concludes that Wikipedia produces positive epistemic consequences, while at the same time stressing the importance of improving content and processes. See Don Fallis, "Toward an Epistemology of Wikipedia," *Journal of the Association for Information Science and Technology* 59 (2009), 1662–1674.

science as an exercise of power.[10] While the concept of gatekeeping has been fully embraced in the fields of communication and social psychology, more recently information science researchers have also adopted the gatekeeper theory in their analyses.[11]

Gatekeepers in virtual communities have several "gatekeeping mechanisms" at their disposal, including censorship, editorial, channeling, and regulation mechanisms.[12] Pertinent to our analysis of knowledge production within Wikipedia are editorial, channeling, and regulatory mechanisms. Editorial mechanisms refer to the ability to modify content, and decide, directly or indirectly, as to what constitutes appropriate content. Channeling mechanisms, on the other hand, refer to the process of directing users' attention to particular pieces of information or resources. Examples of channeling mechanisms include making certain issues more easily accessible than others, either through hyperlinks or favorable ratings. The final mechanism, regulation, refers to explicit or implicit rules or social norms that give the community a sense of order.[13] In the context of Wikipedia, the "core content policies" serve as the main regulatory mechanism, in addition to other community-centered social norms. Approaching Wikipedia from a perspective informed by gatekeeping theory enables us to identify the often-subtle ways in which information is created, privileged, made available, or subverted.

Other information scholars have written about methods of information control in digital environments. For Jan van Dijk, "control is the enactment of… authority in several...categorical inequalities"; that is, control is possible when one user has increased access to digital technologies or content over another user.[14] For a power dynamic of control to exist, one individual must have both more information and a higher societal position than another.[15] Offline socio-economic relationships seem to have crept in the digital landscape as well, leading to major inequalities in terms of access. Extended far enough into the future, van Dijk's conclusion suggests that far from being a socio-cultural equalizer, the Internet and its many offerings, including leisure, gaming, social interaction,

---

[10] Karine Barzilai-Nahon, "Gatekeeping: A Critical Review," *Annual Review of Information Science and Technology* 43 (2009): 1; see also, Karine Nahon, "Toward a Theory of Network Gatekeeping: A Framework," *Journal of the Association for Information Science and Technology* 59 (2008): 1493–1512, and Karine Nahon, "Gatekeepers, Virtual Communities and the Gated: Multidimensional Tensions in Cyberspace," *International Journal of Communications, Law and Policy* 11 (2006): 1-28.

[11] Adetoun A. Oyelude and Alice A. Bamigbola, "Libraries As the Gate: 'Ways' and 'Keepers' in the Knowledge Environment," *Library Hi Tech News* 29 (2012): 7-10.

[12] Karine Barzilai-Nahon, "Gatekeepers, Virtual Communities, and the Gated," 4-5.

[13] Ibid., 4.

[14] Jan A.G.M. van Dijk, *The Deepening Divide: Inequality in the Information Society* (Thousand Oaks, London, New Delhi: Sage Publications, 2005), 19.

[15] Ibid., 145.

commercialism, and information gathering, will reflect the socio-economic and cultural inequalities of the offline world.[16] Drawing on Weber's theory of stratification, Massimo Ragnedda complements van Dijk's work through his definition of digital stratification, in which he argues that rather than reducing inequalities, new information and communications technologies (ICTs) "add to those already existing, in a circular and cumulative process."[17] The inequalities produced by this digital stratification can affect a user's access to ICTs and their content, but can also affect their usage. In other words, even if users gain access, they may be dissuaded from using them, whether due to a lack of digital literacy skills or technology skills, or of not feeling a sense of belonging to the e-community.[18] These inequalities then produce and circularly reinforce a power dynamic that becomes embedded into the structure of ICTs and digital content.[19]

We argue below that Wikipedia editors participate in these forms of information control. Despite this, it is a measurably successful platform. Millions of users flock to Wikipedia for information on a daily basis, and thousands of editors are actively collaborating on what is the largest crowd-sourced project in existence. In a short period, communities that once considered it anathematic have embraced Wikipedia. It is not uncommon to see Wikipedia in university and college classrooms, for example, where it is used as a learning crutch by most or as an intricate and experiential pedagogical exercise by others.[20] The process by which Wikipedia content is created is also worthy of acclaim. As Helen Nissenbaum and Yochai Benkler suggest, Wikipedia has facilitated an environment where global communities of anonymous writers and editors connect and collaborate. To create content in this participatory environment, users engage in a "user-run quasi-formal mediation and arbitration, rather than on mechanical control of

---

[16] Alexander van Deursen and Jan van Dijk, "The Digital Divide Shifts to Differences in Usage," *New Media & Society* 16 (2014): 507.

[17] Massimo Ragneeda, *The Third Digital Divide: A Weberian Approach to Digital Inequalities* (London, UK: Routledge, 2017), 21.

[18] Ibid., 21.

[19] Ibid., 67.

[20] Alison Head and Michael Eisenberg, "How Today's College Students Use Wikipedia for Course-Related Eesearch," *First Monday* 15 (2010), https://firstmonday.org/article/view/2830/2476; Rob Nelson and Heidi L.M. Jacobs, "History, Play, and the Public: Wikipedia in the University Classroom" *History Teacher* 50 (2017): 483-500; Frances Di Lauro and Rebecca Johinke, "Employing Wikipedia for Good not Evil: Innovative Approaches to Collaborative Writing Assessment," *Assessment & Evaluation in Higher Education* 42 (2017): 478-491; Samuel Kai Wah Chu, Yin Zhang, Katherine Chen, Chi Keung Chan, Celina Wing Yi Lee, Ellen Zou, and Wilfred Lau, "The Effectiveness of Wikis for Project-Based Learning in Different Disciplines in Higher Education," *The Internet and Higher Education* 33 (2017): 49-60.

behavior."[21] In theory, content can only be developed if a level of consensus among editors is met, which has established a standard of quality that has contributed to Wikipedia's popularity and success. Nevertheless, Wikipedia has been criticized by different camps for disadvantaging certain communities over others.

In recent years, Wikipedia has come under fire for its apparent gender bias. Despite the growing concerns over its existence and promises to address it, the bias persists.[22] Wikipedia's articles about women are fewer, less extensive, and more focused on family and relationships than those about men, and women must therefore be more "notable" than men in order to be assigned their own page on Wikipedia.[23] Gender bias also appears behind the scenes, with the overwhelming majority of Wikipedia editors being male, and with edits made by women editors being fewer, smaller, and more likely to be contested or reverted.[24] The same inequities exist for content concerning other groups in the "epistemic minority," such as indigenous communities.[25]

In these ways, Wikipedia's gender bias is measurable and concrete, but the exclusion of female perspectives in Wikipedia can also be attributed to the platform's infrastructure. Ford and Wajcman argue that Wikipedia's infrastructure — composed of the determination of "acceptable" knowledge, of software and coding knowledge requirements, and of the technical and legalistic policy framework of Wikipedia — produces power relations that disempower women in particular.[26] Wikipedia prides itself on its amateurism, openness, and the freedom for all to edit its content, which leads readers to see its content as "natural and obvious,"[27] but this rhetoric hides a power

---

[21] Hellen Nissenbaum and Yochai Benkler, "Commons-Based Peer Production and Virtue," *The Journal of Political Philosophy* 14 (2006): 397.

[22] "Gender Bias on Wikipedia," Wikipedia, accessed November 3, 2017, https://en.wikipedia.org/wiki/Gender_bias_on_Wikipedia.

[23] Wikipedia, "Gender bias"; Heather Ford and Jody Wajcman, "'Anyone Can Edit', Not Everyone Does: Wikipedia's Infrastructure and the Gender Gap," *Social Studies of Science* 47 (2017); Claudia Wagner, David Garcia, Mohsen Jadidi, Markus Strohmaier, "It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia," *Proceedings of the Ninth International AAAI Conference on Web and Social Media* (2015); Wager et al., "It's a Man's Wikipedia?"; Wikipedia, "Gender Bias."

[24] Ford and Waljcan, "Anyone Can Edit," 512; Alana Cattapan, "(Re)Writing 'Feminism in Canada': Wikipedia in the Feminist Classroom," *Feminist Teacher* 22, no. 2 (2012): 128-129; Wikipedia, "Gender Bias"; Wagner et al., "It's a Man's Wikipedia?"; Shyong (Tony) K. Lam et al., "WP:Clubhouse? An Exploration of Wikipedia's Gender Imbalance," in *Proceedings of the 7th International Symposium on Wikis and Open Collaboration,* Mountain View, CA, October 3-5, 2011, 2-3.

[25] For an excellent discussion see Henry Farrell and Melissa Schwartzberg, "Norms, Minorities, and Collective Choice Online," *Ethics & International Affairs* 22 (2008): 359.

[26] Ford and Wajcman, "Anyone Can Edit," 522.

[27] Ibid., 512.

imbalance between the men who currently control the site and the women and other marginalized communities who seek representation in both article content and in the editing process. Wikipedia's problem runs deeper than the surface-level absence of articles and of female editors; it is rooted in the policies and rhetoric of the site's foundation.

WIKIPEDIA AND THE CRUX OF VERIFIABILITY

An additional episode works to illustrate Wikipedia's unequal playing field with respect to who and in what way editors can participate in Wikipedia's knowledge production. In 2015, Elsevier "donated" 45 ScienceDirect accounts to "top Wikipedia editors" to help them create comprehensive content more efficiently.[28] This number has since risen to over 5,000 accounts which give access to publisher databases to over 2,600 editors.[29] Dubbed "WikiGate," this development prompted widespread debate amongst academics, librarians, and advocates over the ethical and epistemic implications of Wikipedia collaborating with Elsevier.[30] The implications of the largest for-profit publisher handpicking "top" editors to include additional links to paywalled ScienceDirect articles was not lost on interested parties. Michael Eisen, a founder of the open access movement, contended that Elsevier's actions were not magnanimous; instead, he suggested that Elsevier were using Wikipedia as a vehicle to market their content.[31] This practice has since grown to include free accounts for Wikipedia editors to pay-to-read services for other publishers.

Opinions on WikiGate varied, some viewing it as unimportant while others offered more vicious critiques of the new partnership. One commenter noted that "[WikiGate is being] blown out of proportion. Wikipedia lists all kinds of sources, no matter if they are open or not. And they rightfully should do so. If this donation helps to get more sources into Wikipedia, it's a good thing. Making this a political/ideological thing

---

[28] Sacha Boucherie, "Elsevier Access Donations Help Wikipedia Editors Improve Science Articles," September 10, 2015, https://www.elsevier.com/connect/elsevier-access-donations-help-wikipedia-editors-improve-science-articles. "Top Wikipedia editors" in this case refers to Wikipedia editors with accounts older than 6 months with a minimum of 500 edits.

[29] Nicole Askin, "Meet the Wikipedia Library… At Your Library," *Computers in Libraries* 36, no. 7 (2016). Retrieved from http://www.infotoday.com/cilmag/sep16/Askin--Meet-The-Wikipedia-Library-at-Your-Library.shtml.

[30] Glynn Moody, "'WikiGate' Raises Questions About Wikipedia's Commitment to Open Access," *Ars Technica,* October 14, 2015, https://arstechnica.co.uk/science/2015/09/wikigate-raises-questions-about-wikipedias-commitment-to-open-access/.

[31] Per article prices vary but average around $15 USD.

seems just stupid to me."[32] Others were critical of the partnership. Another offered the following:

> Even fully trusting one's fellow editors, they are human and make mistakes. [Wikipedia's] usefulness relies on the accuracy of its content for trust. I would sincerely hope that editors using paywalled sources will include other, accessible, alternative citations alongside them. Blind trust is not enough.

The Wikipedia Library "Why not OA" page justifies its practice thus:

> … we have to serve our readers and editors as best we can and that means giving them as much access to the best research as possible *today*. Collaboration with publishers is a compromise with mutual benefit: we'd rather have our editors summarize paywalled content for our readers than for that content to not be represented on Wikipedia at all, even if readers may hit a paywall when they click-through… We cannot just wait for the publishing industry to transform while readers are relying on us today. [Further], accessibility is actually not a factor in determining the reliability or verifiability of a source.

The final sentence summarizes Wikipedia's position by differentiating between two services that are often conflated:

> Wikipedia aims to be an open access summary of all reliable knowledge—not a summary of only open access knowledge.[33]

Ultimately, the main critique of Wikipedia's sometimes overt relationship with for-profit publishers is that it contradicts the underlying mission statement of the popular online encyclopedia. In particular, commentators argue that the practice seems to go against Wikipedia's position on verifiability. Indeed, if verifiability means that "other people using the encyclopedia can check that the information comes from a reliable source," and that "all material in Wikipedia mainspace, including everything in articles, lists and captions, must be verifiable" (as Wikipedia's own policies define it), then encouraging references to point to pay-walled knowledge under the guise of it being "the best research as

---

[32] These, and hundreds of other comments, can be found in the comments section of the preceding *Ars Technica* article.

[33] "Wikipedia:Why Not OA", *Wikipedia,* last modified 25, September 2017, https://en.wikipedia.org/wiki/Wikipedia:The_Wikipedia_Library/OA.

possible *today*" goes directly against this principle.[34] Put another away, the tension rests with the potential consequences of forcing readers into trusting an editor's summary, as well as two disparate interpretations of what constitutes a verifiable reference. Considering the level of accessibility of references in Wikipedia reveals an additional level of information inequity affecting the platform.

For new content to pass quality assurance checks in Wikipedia it must abide by several "core content policies", those being "verifiability," "no original research," and "neutral point of view." Verifiability underpins content creation on Wikipedia. Given that subjective interpretations of events or people are discouraged and that allegations must be backed up by sources that can be confirmed.[35] Therefore, the policies of "no original research" and "neutral point of view" are predicated on there being verifiable references present to support written content. These policies try to address the inherent challenges faced by an online resource attempting to summarize the entirety of the world's knowledge in a succinct, impartial, and accessible format. However, Wikipedia's current policy on verifiability privileges some over others; in particular, it privileges editors and readers with access to research databases, whether through an affiliated research library or by means of special access from publishers.

For Wikipedia, verifiability "means people reading and editing the encyclopedia can check that the information comes from a reliable source".[36] Simply counting the number of references listed at the bottom of a Wikipedia article is not an appropriate measure of verifiability, as this measure disregards accessibility as a prerequisite factor in the verification process. Harder et al. make this point explicit as they argue that "providing citations and references does not automatically guarantee verifiability," for access "is just as important as providing the reference or citation in the first place."[37] They provide further nuance by defining two levels of verifiability on Wikipedia: "technical verifiability" and "practical verifiability."[38] Technical verifiability refers to having adequately linked metadata available to automatically process and locate sources (links to catalogues, ISBNs, and DOIs, for example). Practical verifiability, on the other hand, refers to the extent to which referenced material is accessible to someone encountering the reference.

---

[34] "Wikipedia: Verifiability," *Wikipedia,* last modified October 26, 2017,
https://en.wikipedia.org/wiki/Wikipedia:Verifiability.

[35] Olof Sundin, "Janitors of Knowledge: Constructing Knowledge in the Everyday Life of Wikipedia Editors," *Journal of Documentation* 67 (2011): 855.

[36] "Wikipedia:Core Content Policies," *Wikipedia,* accessed November 20, 2017,
https://en.wikipedia.org/wiki/Wikipedia:Core_content_policies.

[37] Reed H. Harder, Alfredo J. Velasco, Michael S. Evans, Daniel N. Rockmore, "Wikipedia Verification Check: A Chrome Browser Extension," in *Proceedings of the 26th International Conference on World Wide Web Companion*, ed. Rick Barrett (New York: Association for Computing Machinery, April 2017), 1620.

[38] Ibid., 1620-21.

For example, if a DOI is present but refers to a paywalled journal article, then the information it supports is *technically* verifiable but *practically* unverifiable to someone without the additional means to access the supporting content. To this end, Harder et al. analyzed the technical and practical verifiability of citations in the top 5,000 most viewed articles in English Wikipedia. Their results found that while Wikipedia's technical verifiability is relatively high, only 5,275 citations to journal articles out of 41,244 (12.8%) included a link to an openly licensed version, while 30,632 (74.3%) contained some digital identifier but were not confirmed to be open. For referenced e-books, out of the 10,922 working Google Books links, 7,749 (71.0%) were partially viewable with limited samples, while only 1,359 (12.4%) were fully viewable and 1,814 (16.6%) were not viewable at all.[39]

   Despite Wikipedia's claims to be an open, equal access tool, its system of knowledge production encompasses and reinforces the inequalities outlined by van Dijk and Ragnedda precisely through subtle "gatekeeping mechanisms."[40] The result is that Wikipedia is more open to some constituencies over others. While anyone can suggest changes to article content on Wikipedia, these changes are all overseen and can be overruled by editors using the aforementioned gatekeeping mechanisms in a variety of ways. Some seasoned editors receive privileged access to scholarly content while general users and content producers do not, and editors have the power to accept or reject content changes suggested by users on many grounds, thereby channeling attention towards specific topics or sources.[41] In this way, certain editors have both increased access to information and a higher position in Wikipedia's structure, fulfilling van Dijk's criteria for a dynamic of control. This dynamic creates a hierarchy among users that Wikipedia claims to have erased. While Wikipedia claims its platform has removed the gatekeeping aspect of knowledge production, it has only shifted its form.

WIKIPEDIA AND SCHOLARSHIP

All of this is worth exploring because Wikipedia's success has impacted scholarship more broadly. Wikipedia's success can be contributed in part to its growing reputation as a reputable resource. But for Wikipedia to continue to have, as Fallis argues, "positive epistemic consequences," the online platform must continue to nurture its image as being

---

[39] Harder et al., "Wikipedia Verification Check," 1621. Note that Harder et al. only considered links to PubMed Central and arXiv as constituting an open access version.

[40] "Wikipedia Founder Jimmy Wales Responds," *Slashdot*.

[41] For a recent example of how Wikipedia editors marginalize women, see article on now Nobel prize winning Donna Strickland: Leyland Cecco, "Female Nobel Prize Winner Deemed Not Important Enough for Wikipedia Entry," *The Guardian*, October 3, 2018, https://www.theguardian.com/science/2018/oct/03/donna-strickland-nobel-physics-prize-wikipedia-denied (accessed February 17, 2019).

a trustworthy resource. Considering the degree to which reference lists are verifiable in Wikipedia articles contributes directly to Wikipedia's credibility. As Olaf Sundin argues, "Wikipedia is stabilized and gains credibility through established media, including popular science, encyclopedias, and scholarly literature. Referring to external sources also creates trust in the knowledge claims in relation to other editors in the on-going discussions of what Wikipedia should contain." Thus, carefully curated, verifiable reference lists "create trust for Wikipedia among its many users." [42] Not all references are created equal, however. Sundin suggests that there is a "hierarchy" or "order of precedence" of reference lists, closely associated with Wikipedia's policies of "no point of view" and "no original research," that encourage greater levels of trust. Unsurprisingly, peer-reviewed literature is atop the hierarchical order while blogs and cross-references to other Wikipedia pages are seen as the least trustworthy.

This hierarchy has made Wikipedia a trusted source for scholars as well, with recent studies confirming that Wikipedia is playing a larger role in scholarship. Robert Tomaszewski and Karen MacDonald's study reveals that citations to Wikipedia in scholarly publications have been increasing universally across disciplines since 2002. "This research indicates that the trend is universal," for it "cannot be attributed to any particular academic discipline; to OA publication; to research published in poor-quality journals, from poor-quality academic institutions, or from economically disadvantaged countries." [43] The increase in the number of references to Wikipedia suggests that authors, reviewers, and journal editors trust and endorse Wikipedia as an authoritative source of information. However, the same trend cannot be observed in the reverse. A recent article by Kayvan Kousha and Mike Thelwall sought to investigate "if citations from encyclopedia articles to academic publications may reflect the transfer of knowledge from the scholarly domain into a format that is accessible to, and perhaps used by, a wider public." They analyzed the number of citations to 302,328 articles and 18,735 monographs in English from 2005 to 2012. Their results indicate that approximately 5% of citations direct to a scholarly article or monograph. [44] Taken together, these recent findings confirm that editors of the English Wikipedia act as "distillers" of a relatively limited number of scientific articles to a public audience. However, there is a clear opportunity for Wikipedia "to act as an 'amplifier' for the (already freely available) OA

---

[42] Sundin, "Janitors of Knowledge," 857.

[43] Robert Tomaszewski and Karen I. MacDonald, "A Study of Citations to Wikipedia in Scholarly Publications," *Science & Technology Libraries* 35 (2016): 248.

[44] Kayvan Kousha and Mike Thelwall, "Are Wikipedia Citations Important Evidence of the Impact of Scholarly Articles and Books?" *Journal of the Association for Information Science and Technology* 68 (2016): 762-779. See also Aida Pooladian and Ángel Borrengo, "Disseminating Knowledge Beyond Scholarly Journals: Coverage of Library and Information Science Literature in Wikipedia," in *24th BOBCATSSS Conference Proceedings & Abstracts* (2016), 333-342.

literature by preferentially broadcasting its findings to millions." [45] Harnessing this potential would be most beneficial for users, editors, and scholars.

In light of this, we argue that more needs to be done by Wikipedia and associated groups to increase the practical verifiability of references. In the current context of misinformation, "fake news," and "alternative facts," Wikipedia's policy of verifiability must be extended to consider the potential for users to actually access and verify cited sources. Indeed, as evidenced by the WikiGate controversy, there is at present a clear disconnect between Wikipedia's and users' understanding of "verifiability." Opening up access to verifiable references, when openly licensed and available, will benefit users interested in confirming or exploring topics in greater depth and in the process help to mitigate the information inequity stemming from Wikipedia's subtle practice of information control. In an environment where an increasing amount of scholarship is being published under an open license, this is more convenient than ever before.[46] Tools which would allow Wikipedia to do this already exist and are readily available, which we briefly discuss and amplify in the following section.

INCREASING PRACTICAL VERIFIABILITY ON WIKIPEDIA

Developed by the Committee for the Accessibility of Publications in Sciences and Humanities, Dissem.in is a web application that harvests metadata from many sources, including Crossref (a non-profit that provides publishing services, including DOI registration and metadata indexing), Bielefeld Academic Search Engine, SHERPA/RoMEO, and Zotero, and subsequently scans for the full text availability of publications.[47] One use case exemplifies the tool's potential. Federico Leva, of Wikimedia Italia, frustrated by the number of paywalled articles referenced in English and Italian Wikipedia, began contacting authors that had been cited. Leva identified over a quarter of a million

---

[45] Misha Teplitskiy, Grace Lu, and Eamon Duede, "Amplifying the Impact of Open Access: Wikipedia and the Diffusion of Science," *Journal of the Association for Information Science and Technology* 68 (2017): 2122.

[46] Madian Khabsa, and C. Lee Giles, "The Number of Scholarly Documents on the Public Web," *PLOS ONE*, 9 (2014), https://doi.org/10.1371/journal.pone.0093949; Richard Van Noorden, "Half of 2011 Papers Now Free to Read," *Nature News,* August 20, 2013, http://www.nature.com/news/half-of-2011-papers-now-free-to-read-1.13577; Heather Piwowar et al., "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles," *PeerJ Preprints* 5:e3119v1 (2017): https://doi.org/10.7287/peerj.preprints.3119v1.

[47] See https://dissem.in/faq#what-is-dissemin. For more on alternative tools, like Unpaywall, see Dalmeet Singh Chawla, "Unpaywall finds free versions of paywalled papers," *Nature News* (2017), https://www.nature.com/news/unpaywall-finds-free-versions-of-paywalled-papers-1.21765.

depositable articles by more than a million authors.[48] Leva contacted authors one by one, informing them that they had been cited in a Wikipedia article, and asked them to deposit a pre-print or post-print to an institutional or subject repository. According to Leva, over 15% of authors engaged with the email by following a link and over 3000 papers were deposited via Dissem.in within two months.[49] Leva's approach is not unlike that of librarians tasked with populating local institutional repositories; most of the legwork is completed before soliciting the author. Importantly, Leva draws on Wikipedia's content verifiability policy to bolster his approach. The following is an excerpt from the boilerplate email solicitation:

> I need an open access copy [of your paper] to be able to discuss it with fellow editors and make sure the Wikipedia article provides an accurate and neutral overview under a free license for everybody to use. We also want every user who reads Wikipedia to be able to verify its content by consulting its primary sources.

This solicitation email appeals to the author directly to make available a copy of their research. This approach works to inform researchers of how their publications are being referenced on Wikipedia, as well as drawing attention to usage restrictions placed on their work. As we have shown, Wikipedia's current verifiability policy does not purport that all Wikipedians be able to verify content themselves. This narrow interpretation of verifiability is, in our view, passing on an opportunity to expand the utility of Wikipedia: from a functionally limited crowdsourced encyclopedia, to a platform that facilitates critical information literacy.[50]

To its credit, the Wikipedia Library (WPL) has done much to promote critical information literacy and has increased its efforts to increase the practical accessibility of citations available on Wikipedia. Indeed, one of WPL's main goals is to help Wikipedians understand "how Wikipedia is made, how to use it within a research process, and other key critical thinking and digital research skills."[51] The WikiProject Open Access, for example, attempts "to help Wikipedians cooperate, organize, and share ideas on improving open access-related articles on Wikipedia," and to "improve coverage to open

[48] The project has garnered a significant following. An archive of Dissem.in discussion lists is available here: https://lists.wikimedia.org/pipermail/openaccess/.

[49] Michele Marchetto, "How Wikimedia Helped Authors Make Over 3000 Articles Green Open Access Via Dissemin," OKIB, October 26, 2017, https://blog.okfn.org/2017/10/26/how-wikimedia-helped-authors-make-over-3000-articles-green-open-access-via-dissemin/.

[50] Heidi L.M. Jacobs, "Posing the Wikipedia 'Problem': Information Literacy and the Praxis of Problem-Posing in Library Instruction," in *Critical Library Instruction: Theories and Methods,* eds. M. Accardi, E. Drabinski, and A. Kumbier (Duluth, MN: Library Juice Press, 2010), 179–197.

[51] "Wikipeida:The Wikipedia Library," *Wikipedia,* last updated September 19, 2017, https://en.wikipedia.org/wiki/Wikipedia:The_Wikipedia_Library/Research_libraries

access material."[52] Signalling OA is one such initiative supported by WikiProject Open Access. It aims to build an infrastructure which signals whether citations on Wikipedia are openly licensed. Signalling OA provides the tools and infrastructure necessary for editors to provide article-level licensing information as developed by the Directory of Open Access Journals and Crossref. Signaling OA badges indicate to users whether they can access the cited resource, rather than encounter a paywall after clicking on a reference. Furthermore, if the source is open access and has been published under a permissible license, images and accompanying data could also be reused.[53] Finally, the Wikipedia Library's OAbot leverages the potential of crowdsourcing to increase the number of links to open versions of paywalled articles. Drawing on OAdoi and dissem.in APIs, the OAbot scans Wikipedia citations that link to paywalled articles, then looks up and presents open versions for volunteers to review.[54] The aforementioned projects are just some of the initiatives that provide a clear way forward for individuals interested in improving practically verifiable references. Despite a growing awareness of these tools, they have only been adopted by editors interested in enlarging the number of practically verifiable sources in Wikipedia.

CONCLUSION

Over the years, Wikipedia has grown to be one of the most popular websites on the Internet. Despite its resounding success, the platform has been criticized for several shortcomings with respect to content and the community that it has cultivated. The Wikipedia gender gap is one such failing that must be addressed. Another, in our view, is the disconnect between Wikipedia's policy of verifiability and the low number of links to openly licensed versions of scholarship. Both of these cases require changes in policy as well as a shift in the way that editors produce knowledge on Wikipedia. Approaching these tangible issues from a perspective informed by information control and virtual gatekeeping theory has enabled us to critically examine why Wikipedia is structured in this way, and how control is manifested. Far from being a virtual place "without gatekeepers," Wikipedia consists of a community bound to a set of content policies and a socio-technical structure that work to cement digital inequalities. As a result, Wikipedia users find themselves in one of two distinct camps: those who have access, or those who

---

[52] "Wikipedia:WikiProject OA," *Wikipedia,* last updated September 23, 2017,
   https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Open_Access.
[53] "WikiProject Open Access/Signalling OA-ness/Report," Wikipedia, last updated December 2, 2015, https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Open_Access/Signalling_OA-ness/Report.
[54] See "Wikipedia:OABOT," Wikipedia, Accessed October 27, 2017,
   https://en.wikipedia.org/wiki/Wikipedia:OABOT.

do not. Wikipedia's information access inequalities therefore mimic those of the offline world. Encouraging editors to include access points to open access and openly licensed versions of paywalled literature would enable Wikipedia to move towards its own stated goals of enabling "free access to the sum of all human knowledge." Increasing the level of practically verifiable sources would work to diminish information gatekeeping practices that currently disadvantage general readers and editors who lack access to research libraries or proprietary databases.

Fortunately, projects such as the Wikimedia Italia and Dissem.in collaboration, Signalling OA, and WikiLibrary's OAbot are easy options to turn to for identifying and locating open scholarship. For that reason, we urge practitioners to implement them as standard elements of Wikipedia edit-a-thons and experiential learning involving Wikipedia. In particular, popular initiatives such as *#1Lib1Ref* should highlight these tools, as well as stressing the importance of practical verifiability to Wikipedia citations. As evidenced by the "Wikigate" controversy, there is a clear disconnect between Wikipedia's and users' understanding of "verifiability." We argue that whenever possible, the former must be extended to match the latter.

Our objective in this article has been to demonstrate how Wikipedia's policies and editing practices promote a subtle form of information gatekeeping, despite its stated goals of openness and transparency. In simulating the citation and referencing practices of scholarly publications without a mandate to link to available open access resources, Wikipedia and its editors offer readers technical verifiability without considering the consequence of practical verifiability. Ultimately, this practice works to diminish the value of the platform — from one which holds the potential to encourage critical engagement with scholarship from its readers, to one which simply offers prescribed knowledge. A greater emphasis on open access materials and the tools which enable them could mitigate the information inequalities Wikipedia claims to already address, and create the truly open platform Wikipedia already purports itself to be.

BIBLIOGRAPHY

Askin, Nicole. "Meet the Wikipedia Library… At Your Library." *Computers in Libraries* 36, no. 7 (2016). Retrieved from http://www.infotoday.com/cilmag/sep16/Askin--Meet-The-Wikipedia-Library-at-Your-Library.shtml.

Barton, Matthew. "Is There a Wiki in this Class?: Wikibooks and the Future of Higher Education." In *Wiki Writing: Collaborative Learning in the College Classroom*, edited by Matthew Barton and Robert Cummings, 175-193. Ann Arbor, MI: University of Michigan Press, 2008.

Barzilai-Nahon, Karine. "Gatekeeping: A Critical Review." *Annual Review of Information Science and Technology* 43 (2009): 1-79.

Boucherie, Sacha. "Elsevier Access Donations Help Wikipedia Editors Improve Science Articles." *Elsevier Connect*. September 10, 2015. Retrieved from https://www.elsevier.com/connect/elsevier-access-donations-help-wikipedia-editors-improve-science-articles.

Casebourne, Imogen, Chris Davies, Michelle Fernandes, and Naomi Norman. "Assessing the Accuracy and Quality of Wikipedia Entries Compared to Popular Online Encyclopaedias: A Comparative Preliminary Study Across Disciplines in English, Spanish and Arabic." (2012). Retrieved from https://commons.wikimedia.org/wiki/File:EPIC_Oxford_report.pdf.

Cattapan, Alana. "(Re)Writing 'Feminism in Canada': Wikipedia in the Feminist Classroom." *Feminist Teacher* 22, no. 2 (2012): 125-136.

Cecco, Leyland. "Female Nobel Prize Winner Deemed Not Important Enough for Wikipedia Entry." *The Guardian*, October 3, 2018. https://www.theguardian.com/science/2018/oct/03/donna-strickland-nobel-physics-prize-wikipedia-denied (accessed February 17, 2019).

Chawla, Dalmeet Singh. "Unpaywall finds free versions of paywalled papers." *Nature News*. April 4, 2017. https://www.nature.com/news/unpaywall-finds-free-versions-of-paywalled-papers-1.21765.

Chu, Samuel Kai Wah, Yin Zhang, Katherine Chen, Chi Keung Chan, Celina Wing Yi Lee, Ellen Zou, and Wilfred Lau. "The Effectiveness of Wikis for Project-Based Learning in Different Disciplines in Higher Education." *The Internet and Higher Education* 33 (2017): 49-60.

Currie, Morgan. "The Feminist Critique: Mapping Controversy in Wikipedia." In *Understanding Digital Humanities*, edited by David Berry, 224-248. Houndmills, Basingstoke, UK: Palgrave Macmillan, 2012.

Cuthberson, Anthony. "Wikipedia Founder Jimmy Wales Believes He Can Fix Fake News with Wikitribune Project," *Newsweek*, April 25, 2017, https://www.newsweek.com/wikipedia-jimmy-wales-fake-news-wikitribune-crowdfunding-589226.

Di Lauro, Frances and Rebecca Johinke. "Employing Wikipedia for Good not Evil: Innovative Approaches to Collaborative Writing Assessment." *Assessment & Evaluation in Higher Education* 42 (2017): 478-491.

Fallis, Don. "Toward an Epistemology of Wikipedia." *Journal of the Association for Information Science and Technology* 59 (2009): 1662–1674.

Farrell, Henry and Melissa Schwartzberg. "Norms, Minorities, and Collective Choice Online." *Ethics & International Affairs* 22 (2008): 357-368.

Ford, Heather, David R. Musicant, Shilad Sen, and Nathaniel Miller. "Getting to the Source: Where does Wikipedia Get its Information From?" In *WikiSym13: Proceedings of the 9th International Symposium on Open Collaboration*, Hong Kong, August 5-7, 2013. DOI: 10.1145/2491055.2491064.

Ford, Heather and Jody Wajcman. "'Anyone Can Edit', Not Everyone Does: Wikipedia's Infrastructure and the Gender Gap." *Social Studies of Science* 47, no. 4 (2017): 511-527.

Giles, Jim. "Internet Encyclopaedias go Head to Head." *Nature* 438 (2005): 900-901.

Harder, Reed H., Alfredo J. Velasco, Michael S. Evans, and Daniel N. Rockmore. "Wikipedia Verification Check: A Chrome Browser Extension." In *Proceedings of the 26th International Conference on World Wide Web Companion*, edited by Rick Barrett, 1619-1625. New York: Association of Computing Machinery, April 2017.

Head, Alison and Michael Eisenberg. "How today's college students use Wikipedia for course-related research." *First Monday* 15 (2010). https://firstmonday.org/article/view/2830/2476.

Jacobs, Heidi L.M. "Posing the Wikipedia 'Problem': Information Literacy and the Praxis of Problem-Posing in Library Instruction." In *Critical Library Instruction: Theories and Methods*, edited by Maria Accardi, Emily Drabinski, and Alana Kumbier, 179–197. Duluth, MN: Library Juice Press, 2010.

Khabsa, Madian and C. Lee Giles. "The Number of Scholarly Documents on the Public Web." *PLOS ONE* 9 (2014). https://doi.org/10.1371/journal.pone.0093949.

Piwowar, Heather, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles," *PeerJ Preprints* 5:e3119v1 (2017). https://doi.org/10.7287/peerj.preprints.3119v1.

Kousha, Kayvan and Mike Thelwall. "Are Wikipedia Citations Important Evidence of the Impact of Scholarly Articles and Books?" *Journal of the Association for Information Science and Technology* 68 (2016): 762-779.

Lam, Shyong (Tony) K., Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R. Musicant, Loren Terveen, and John Riedl. "WP:Clubhouse? An Exploration of Wikipedia's Gender Imbalance." In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, Mountain View, CA, October 3-5, 2011, 1-10.

Luyt, Brendan and Daniel Tan. "Improving Wikipedia's Credibility: References and Citations in a Sample of History Articles." *Journal of the Association for Information Science and Technology* 4 (2010): 715-722.

Marchetto, Michele. "How Wikimedia Helped Authors Make Over 3000 Articles Green Open Access Via Dissemin." *OKIB*. October 26, 2017. https://blog.okfn.org/2017/10/26/how-wikimedia-helped-authors-make-over-3000-articles-green-open-access-via-dissemin/.

Moody, Glynn. "'WikiGate' raises questions about Wikipedia's commitment to open access." *Ars Technica*. October 14, 2015. https://arstechnica.co.uk/science/2015/09/wikigate-raises-questions-about-wikipedias-commitment-to-open-access/.

Nahon, Karine. "Gatekeepers, Virtual Communities and the Gated: Multidimensional Tensions in Cyberspace." *International Journal of Communications, Law and Policy* 11 (2006): 1-28.

Nahon, Karine. "Toward a Theory of Network Gatekeeping: A Framework." *ASIS&T* 59 (2008): 1493–1512.

Nelson, Rob and Heidi L.M. Jacobs. "History, Play, and the Public: Wikipedia in the University Classroom." *History Teacher* 50 (2017): 483-500.

Nissenbaum, Hellen and Yochai Benkler. "Commons-Based Peer Production and Virtue." *The Journal of Political Philosophy* 14 (2006): 394-419.

Oyelude, Adetoun A. and Alice A. Bamigbola. "Libraries As the Gate: 'Ways' and 'Keepers' in the Knowledge Environment." *Library Hi Tech News* 29 (2012): 7-10.

Ragneeda, Massimo. *The Third Digital Divide: A Weberian Approach to Digital Inequalities*. London, UK: Routledge, 2017.

Slashdot. "Wikipedia Founder Jimmy Wales Responds." Posted July 28, 2004. https://slashdot.org/story/04/07/28/1351230/wikipedia-founder-jimmy-wales-responds.

Sundin, Olof. "Janitors of Knowledge: Constructing Knowledge in the Everyday Life of Wikipedia Editors." *Journal of Documentation* 67 (2011): 840-862.

Teplitskiy, Misha, Grace Lu, and Eamon Duede. "Amplifying the Impact of Open Access: Wikipedia and the Diffusion of Science." *Journal of the Association for Information Science and Technology* 68 (2017): 2116-2127.

Tomaszewski, Robert and Karen I. MacDonald. "A Study of Citations to Wikipedia in Scholarly Publications." *Science & Technology Libraries* 35 (2016): 246-261.

van Dijk, Jan A.G.M. *The Deepening Divide: Inequality in the Information Society*. Thousand Oaks, London, New Delhi: Sage Publications, 2005.

Van Noorden, Richard. "Half of 2011 Papers Now Free to Read." *Nature News*. August 20, 2013. http://www.nature.com/news/half-of-2011-papers-now-free-to-read-1.13577.

Wagner, Claudia, David Garcia, Mohsen Jadidi, and Markus Strohmaier. "It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia." In *Proceedings of the Ninth International AAAI Conference on Web and Social Media* (2015).

Willinsky, John. "What Open Access Research Can Do For Wikipedia." *First Monday* 12 (2007). https://firstmonday.org/ojs/index.php/fm/article/viewArticle/1624/1539.

Van Deursen, Alexander JAM, and Jan AGM Van Dijk. "The Digital Divide Shifts to Differences in Usage." *New Media & Society* 16, no. 3 (2014): 507-526.