

# Foreword

## Wikipedia and the Sum of All Human Information

Heather Ford

---

### Bio

*Dr. Heather Ford is a writer, researcher and activist working on issues relating to public participation in knowledge creation, diversity and representation in online spaces, as well as the algorithmic mediation of facts and knowledge. Prior to her current position at Leeds University's School of Media and Communication, she worked with numerous non-profits including Creative Commons, iCommons, the Electronic Frontier Foundation, the Wikimedia Foundation, Ushahidi and the Association for Progressive Communications. Heather undertook her postgraduate studies at Stanford University, the University of California, Berkeley's School of Information and at the University of Oxford's OII (Oxford Internet Institute).*

### Wikipedia turns fifteen

As Wikipedia celebrates its fifteenth birthday this year, many will applaud the project's phenomenal growth in scale and authority in such a relatively short amount of time. With 250 language versions and 500 million users a month, Wikipedia is now the seventh biggest website in the world (Alexa, 2016).

---

*Heather Ford, Dr.  
School of Media and Communication,  
The University of Leeds, United Kingdom  
hfordsa@gmail.com*

Sheer numbers, however, do not fully capture why Wikipedia has become an important object of study. Wikipedia is important because it is regularly consulted by millions of people in the course of everyday life: for settling bets, looking up facts, discovering new facts, and learning about what is happening around us. Wikipedia is important because it has become entangled in our everyday lives. Because of its ubiquity, ease of use and centrality to the Web experience, Wikipedia has become a marker of importance, a symbol of notability, a site of information power. So much so that it has become a symbol of success if a person has a Wikipedia article about them, and an object of concern if one has not proven worthy enough for consideration (Newman, 2014).

Despite the centrality of Wikipedia to our information ecosystem, serious critical analysis of Wikipedia's representations, its governance mechanisms, and its information politics is still highly limited. The first wave of Wikipedia critiques were largely constituted by ill-informed misunderstandings of the ways the site works (see, for example, Keen, 2007). Commentators critiqued Wikipedia for its lack of quality, its reliance on unnamed crowds, and its ability to be edited by anonymous "nobodies." The predominant theory was that Encyclopedias were books containing facts, and only professional experts could accurately represent them.

By the time of Wikipedia's tenth birthday in 2011, much had changed. Its main rival, *Encyclopedia Britannica*, had just stopped producing print copies

due to competition from Wikipedia. The world had moved into an era dominated by the politics of openness (Tkacz, 2015) and the logic of crowd wisdom (Lanier, 2011; Surowiecki, 2005). Wikipedia had originally been positioned as David to Big Media's Goliath, but Wikipedia was beginning to be recognised as a Goliath in its own right. The question about whether Wikipedia was worthy or not was no longer relevant. Like a self-fulfilling prophecy, Wikipedia had become authoritative because it was used as a reference source by millions of people. And because it was so widely used, it became authoritative.

In 2011, a project called *Critical Point of View*, (CPOV) led by Web critics Geert Lovink and Nathaniel Tkacz established a new wave of Wikipedia critique and provided a space for scholarship that was treating Wikipedia not as a novelty but as a feature of everyday life. CPOV broke new ground (Lovink et al., 2012). Until then, those who criticized the encyclopedia were branded as luddites or ridiculed for opposing the values of openness that Wikipedia espoused. With CPOV, however, a group of insiders were trying to hold the public encyclopedia up to greater standards, or at least the standards represented by its own articulated goals.

### **Wikipedia research today**

Five years later, CPOV's critical spirit needs to take root at a broader level. Now more than ever it has become important to interrogate whose facts are being represented on Wikipedia and what are the structural constraints that favor certain viewpoints above others. In the past few years, Wikipedia's centrality has become even more entrenched. The greatest contributing factor to this entrenchment is signalled by the changing role of Wikipedia in the ecology of the Internet. If we think about who controls the representation of facts online, we're seeing a consolidation in the governing forces behind those facts. While Google and other search engines have privileged Wikipedia articles in search results about people, places and things to a significant extent in the past (see, for example, Silverwood-Cope, 2012). Search engines now establish Wikipedia as a key source in their presentation of answers to users' queries in prominent fact boxes rather than only search results. The Google Knowledge Graph, announced in 2012, presented users with "key facts" about the query

drawn from 'public sources' including Wikipedia (Singhal, 2012).

The power of factual representation is becoming increasingly centralized in the hands of a few key institutions. As a result, Wikipedia, as an open platform in which actors wrestle for control over facts through argumentation, technical virtuosity and lobbying power (Ford, 2013; Halfaker, Geiger, Morgan, & Riedl, 2013; Pinsker, 2015), is becoming an important site of information politics. It is a community now seeing the emergence of often invisible and unidentified PR agents and Search Engine Optimization (SEO) professionals joining its ranks (Owens, 2013; Shapiro, 2015).

Is Wikipedia really as open and free as it suggests? While anyone can edit Wikipedia, whose edits are actually sustained over time? Has Wikipedia become a tool for the powerful?

This special issue reflects a step towards addressing such questions. Prompted by a Wikipedia feminist edit-a-thon in Copenhagen earlier this year, this special issue provides new frameworks with which to assess Wikipedia's structure, governance and politics. Importantly, the authors note that they are not experts on Wikipedia-theirs is not a view from "inside." This should not, however, be a reason to dismiss such important views. Wikipedia's stakeholders are no longer just insiders, and because Wikipedia both systematically excludes a number of different groups from participating, insider knowledge should not be the only measure of importance. Contributors from disciplines including archival studies, art history and queer studies are therefore a welcome addition to Wikipedia scholarship that tends to be dominated by computer science, information science, and sociology.

Although the articles in this issue use a variety of lenses to understand Wikipedia's place in the world, there is a single thread that unites them. That thread is constituted by questions about Wikipedia as a repository for "all human knowledge" (Wikimedia Foundation, n.d.). Is Wikipedia truly advancing towards the sum of all human knowledge? Or is it producing a great deal of information at the expense of understanding how to engage with knowledge?

## Information vs knowledge

Analyzing questions about who is able to succeed on Wikipedia requires some benchmark, some way of evaluating Wikipedia's representation of the world. What better way of evaluating Wikipedia than by employing the language used by its creators? Centuries of thinkers have focused on questions about the nature of knowledge. Generations of social scientists have attempted to understand the social construction of knowledge (Berger & Luckmann, 1991), the influence of architecture and technologies on how some knowledges become more authoritative than others (Foucault, 1980; Latour, 1993), and the comparisons between the knowledge of different groups - adults vs. children, "civilised" vs. "primitive" peoples, and scientists vs. lay communities, for example (Jovch- elovitch, 2006).

The Internet age has heralded a new wave of literature addressing data and knowledge production. According to Brown and Duguid (2000), many tend to equate information with knowledge:

"People are increasingly eager that their perfectly respectable cache of information be given the cachet of knowledge. Such redefinitions surreptitiously extend the overlapping area where knowledge and information appear as interchangeable terms" (Brown & Duguid, 2000, p. 119).

But knowledge is not the same as information. While knowledge is indelibly linked to a person or community, information is the product of that knowledge, the symbolic representation of knowledge in the shape of words, images and sound. While knowledge is immaterial, information takes material shape on a page, wall, or in the form of digital 1s and 0s.

Ironically, although Wikipedia seeks to represent human knowledge, it asks its editors to leave their knowledge at the door. Wikipedia aims for "verifiability, not truth" (Wikipedia contributors, 2016). In other words, an editor must represent only the facts about a given subject that are represented by what the encyclopedia deems to be 'reliable sources' rather than what they know to be true about it. Verifiability is an important element of ensuring quality content on the encyclopedia-particularly when one considers that it is written by those who may not have expertise in a particular subject. But the obsessive enforcement

and logic of verifiability, as practiced by Wikipedians particularly on the English site, has had unexpected and undesirable consequences.

One of those consequences is the systematic exclusion of knowledge about the world that, for a variety of reasons, remains unexpressed in the form of (what the majority of editors regard as) reliable. A great deal of knowledge about the world remains unwritten and unpublished (Graham, Hale, & Stephens, 2011) because of issues relating to historical racial and gender inequities, or because of cultural traditions that forbid material representation of oral or indigenous knowledges (Gallert & Van der Velden, 2013).

The knowledge that ends up being represented on Wikipedia, then, is knowledge that is written and published, already available in online formats, since this is the primary way in which Wikipedians conduct research (Ford, Sen, Musicant, & Miller, 2013). Furthermore, the dominant sources and citations used by Wikipedians are highly skewed towards particular domains (including Google, the New York Times and the BBC). Editors from countries outside the United States and Western Europe face significant bias against local sources. The dominant sources of knowledge used by Wikipedians are not necessarily biased in themselves, but it is important to recognise that such sources represent particular points of view and do not, in any respect, reflect the diversity of all the world's knowledge.

There are alternative measures for determining what should be reflected on Wikipedia that go beyond conceptions of verifiability. What about quality as determined by what we *know*, as a diverse global society, rather than what we have so far *represented*? If we are truly committed to representing the sum of all human knowledge, then it becomes important to reflect the diversity of knowers among our ranks. Women make up a tiny proportion of Wikipedia's editing community and those who do edit have faced multiple obstacles. Participants from outside of North America and Western Europe are weakly represented and face similar challenges when they *do* edit (Ford, 2011; Graham, Hogan, Straumann, & Medhat, 2014). If quality is determined by the extent to which Wikipedia reflects human knowledge, then this failure wouldn't be considered unimportant, but rather essential to ensuring its continued success. I believe we should be holding Wikipedia up to its

goal of human knowledge; it is precisely, the value of diversity that, in addition to freedom and openness, we need to be advancing in our critiques.

This issue is being published as a precursor for the next round of edit-a-thons in 2016 and is intended to serve as useful material for both insiders and curious outsiders. It is a call to action, in addition to being an opportunity to pause for reflection. Only by involving marginalized groups in the editing of Wikipedia articles will we start to see the reflection of diverse knowledges on the platform. These brave individuals, recognizing the importance of Wikipedia while still trying to shape and change it for the better, need to be applauded in the next phase of Wikipedia's development. Maybe then we can start to see Wikipedia engaging more fully with knowledge and not just information.

## References

- Alexa (2016). *Alexa*. Alexa Internet. Retrieved from <http://www.alexa.com/siteinfo/wikipedia.org>
- Berger, PL & Luckmann, T (1991). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Penguin UK.
- Ford, H (2011). The Missing Wikipedians. In *Critical Point of View: A Wikipedia Reader*. Amsterdam: Institute of Network Cultures.
- Ford, H (2013). How Wikipedia's Dr Jekyll became Mr Hyde: Vandalism, sock puppetry and the curious case of Wikipedia's decline. In *Resistance and appropriation*. Denver, Colorado.
- Ford, H, Sen, S, Musicant, DR & Miller, N (2013). Getting to the Source: Where Does Wikipedia Get Its Information from? In *Proceedings of the 9th International Symposium on Open Collaboration* (pp. 9:1-9:10). New York, NY, USA: ACM. <http://doi.org/10.1145/2491055.2491064>
- Foucault, M (1980). *Power/knowledge : selected interviews and other writings, 1972-1977*. Hemel Hempstead: Harvester Wheatsheaf.
- Gallert, P & Van der Velden, M (2013). Reliable sources for indigenous knowledge: Dissecting Wikipedia's Catch-22. Retrieved from <http://ir.polytechnic.edu.na/jspui/handle/10628/409>
- Graham, M, Hale, S & Stephens, M (2011). *Geographies of the world's knowledge*. London: Convoco.
- Graham, M, Hogan, B, Straumann, RK & Medhat, A (2014). *Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty* (SSRN Scholarly Paper No. ID 2382617). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2382617>
- Halfaker, A, Geiger, RS, Morgan, JT & Riedl, J (2013). The Rise and Decline of an Open Collaboration System How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist*, 57(5), 664-688. <http://doi.org/10.1177/0002764212469365>
- Jovchelovitch, S (2006). *Knowledge in Context: Representations, Community and Culture* (New Ed edition). New York, NY: Routledge.
- Keen, A (2007). *The Cult of the Amateur* (First Edition First Printing edition). Doubleday.
- Lanier, J (2011). *You Are Not A Gadget: A Manifesto*. London: Penguin.
- Latour, B (1993). *We have never been modern*. Harvard University Press.
- Lovink, G, Tkacz, N, Reagle, JM, O'Sullivan, D, Liang, L, Salah, AA, ... Chen, S-L (2012). *Critical Point of View: A Wikipedia Reader* (SSRN Scholarly Paper No. ID 2075015). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2075015>
- Newman, J (2014, January 8). Wikipedia, What Does Judith Newman Have to Do to Get a Page? *The New York Times*. Retrieved from <http://www.nytimes.com/2014/01/09/fashion/Wikipedia-Judith-Newman.html>
- Pinsker, J (2015, August 11). The Covert World of People Trying to Edit Wikipedia-for Pay. *The Atlantic*. Retrieved from <http://www.theatlantic.com/>

business/archive/2015/08/wikipedia-editors-for-pay/393926/

Shapiro, N (2015, April 28). Wikidata Meets the Google Knowledge Graph. Retrieved from <http://www.searchenginejournal.com/wikidata-meets-google-knowledge-graph/130459/>

Silverwood-Cope, S (2012, February 8). Wikipedia: Page one of Google UK for 99% of searches. Retrieved from <http://www.intelligentpositioning.com/blog/2012/02/wikipedia-page-one-of-google-uk-for-99-of-searches/>

Singhal, A (2012, May 16). Introducing the Knowledge Graph: things, not strings [blog]. Retrieved from <http://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html>

Surowiecki, J (2005). *The Wisdom of Crowds*. Knopf Doubleday Publishing Group.

Tkacz, N (2015). *Wikipedia and the politics of openness*. Chicago ; London: University of Chicago Press.

Wikimedia Foundation (n.d.). Wikimedia Foundation Vision statement. Retrieved January 26, 2016, from <https://wikimediafoundation.org/wiki/Vision>

Wikipedia contributors (2016, January 26). Wikipedia Verifiability Policy [Encyclopedia]. Retrieved January 26, 2016, from <https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&oldid=701166094>