

Google and Wikipedia: 'Feeling Lucky' or 'Citation Needed'?

[Creative Commons Attribution- Non Commercial - Share Alike 4.0 International License](#)

An Indispensable Resource: The World Wide Web Frames Our World

You're headed to the store to buy a new toothbrush. Or you need an air conditioner. You need the exact address of that new Chinese restaurant where you're supposed to meet some friends. You vaguely recall that it's possible to cook a potato in the microwave, but you have no idea how.

You open your search engine. Or, increasingly, you pull up an app on your phone, or a new browser tab's address bar, and immediately input the beginnings of your query. You're likely to get a list of suggested search terms even before you finish typing the words in. You hit enter, and within split seconds you have pages of possible results before you. How often do you really feel the need to click beyond the first, or maybe second, page of results? How much do you bank on the convenience and presumed relevance of the first few links? How deeply do you take for granted that you're getting the whole picture?

Upon its introduction, Google's original search algorithm represented an attempt to automate a complex and wide-reaching process of peer review. The history of the Internet already depended on various replications of the process of peer review (Akerman, 2006; Editors of the New Atlantis, 2006). Still, when Larry Page and Sergey Brin designed a centralized tool that could easily handle the ever-growing data of the World Wide Web, it was rapidly adopted, as users with very little computer experience could now track down information through a relatively simple and straightforward interface (Segev, 2005). This classic need to sort and classify information grew more and more crucial as the reserve of sources and contributors to the Internet grew larger. The "sea of information" that the World Wide Web had become was now navigable from a seemingly central point (Roberts, 1999). The Google search engine creates a continually updated "picture" of the content of the World Wide Web, mapping out content behind the scenes and returning links based on their relevance to the search terms and on how often they're linked by other sites on the Web (Vaidhyanathan, 2011).

Page and Brin's proposed algorithm looked like a simple tweak from the outside at first, but it was a tool that would subtly contribute to an entirely different way of navigating the repository of information that we call the World Wide Web. However, that targeted arrow of the ranking algorithm, which legitimizes some sources over others, might make *too much* "sense" of the Internet. Now, it might be so narrowly focused that it is eroding the very promise of the World Wide Web: to provide anyone who has online access the potential platform to have their voice heard. Particularly with the monetization of search engine results, in which Google more prominently displays paid advertising, the scope begins to narrow, and results pages become repetitious and representative of significant source biases ([Goldman, 2006](#)).

When the PageRank algorithm was still young, it operated on a relatively new frontier, and there were thus many opportunities for sites and sources to gain increasing recognition

through back-linking and communal promotion. As its methodology has grown stronger and more entrenched, Google's indirect endorsement of certain sources over others becomes a self-fulfilling prophecy: sites are heavily linked to because they are the first results that come up on Google. van Dijck (2010) uses the specific example of Google Scholar, which focuses on academic articles, to emphasize that a search engine exists as "a nodal point of power, while the mechanism of knowledge production is effectively hidden in the coded mechanisms of the engine" (p. 579). With the contemporary World Wide Web's scale of operation, narratives have become solidified in much the same way as any dominant societal narrative of truth.

Wikipedia as a Rising Source of Presumed Authority

Consistently, the top search result in Google search is the Wikipedia entry for that term or topic ([Metz, 2009](#)). For many day-to-day inquiries, this is incredibly convenient, and has fueled a sort of synergy between the two sites. This arrangement becomes problematic, though, if users begin to equate Wikipedia with "the truth" on any given subject. Since a landmark study in 2005, Wikipedia has consistently surprised researchers with its standard of accuracy compared to traditional academic encyclopedias ([Terdiman, 2005](#)). In fact, it's irresistible to cite Wikipedia's own thorough and well-rounded page addressing the wide range of research that has been done on its accuracy ([Reliability of Wikipedia, 2015](#)). However, even the site itself warns visitors that Wikipedia pages should be used as the first step in research, and not the end destination (Baron, 2009). By learning to follow the hyperlinks to Wikipedia's cited sources, readers can continue to exercise critical literacy in tracking users' edits to their origin points ([Rosenzweig, 2006](#)).

Like much of the Internet, Wikipedia's primary pool of vocal contributors consists of a majority of white males, which can create an unconscious, but self-replicating, racial and gender bias ([Lapowsky, 2015](#)). If a single demographic is largely responsible for composing articles on issues of identity politics, racism, and sexism, there could be a troubling outcome in which the voices of those affected by these topics are functionally erased. If those who are privileged enough to escape the effects of racism and sexism are left to explain the topics, it could be assumed that these issues do not exist at all, or at least not to the insidious degree which is often experienced by insiders to these communities.

In 2013, 90% of Wikipedia's active editors self-reported as male, but since then, the Wikimedia Foundation has been pushing tirelessly to invite a wider demographic to participate, so that the reserve of editors more accurately reflects the distribution of the general population ([Meyer, 2013](#)). This gender bias reflects even at the more casual level of Wikipedia content contribution, as content and topics traditionally targeted towards men are significantly more likely to have coverage on the site than those that are targeted toward women ([Cohen, 2011](#); [Pappas, 2013](#)). Still, it is not as though diverse groups have been explicitly excluded from Wikipedia – rather, there seems to be a self-selection participation bias that is further reinforced by communicative norms on the site's "Talk" pages that may feel hostile to those outside its usual demographics ([Ciffolilli, 2003](#)).

On Wikipedia, anyone with Internet access can literally rewrite history – or at least attempt to. Most recently, with uproars about racially charged police brutality in the United

States, it was uncovered that IP addresses traceable to the NYPD precinct had been used, consistently and anonymously, to remove information from Wikipedia about violent police incidents ([Cheney-Rice, 2015](#)). One of Wikipedia's greatest battles is the fight to monitor and reduce this contributor bias, while maintaining a minimal framework of rules and granting a certain degree of near-anonymity ([Rosenzweig, 2006](#)). Still, the site's ability to combat bias is undermined by the dwindling participation in direct edits and content contribution, even as the site's readership grows drastically ([Simonite, 2013b](#)). The greatest hope for Wikipedia is the emphasis on the “peer” in peer review, combined with the expectation that readers will parse the site's content with a critical eye (Anderson, 2006).

Conscientization and Online Sources: ‘Feeling Lucky’ or ‘Citation Needed’?

If we accept the premise that search engines provide one of the central paths to uncovering and accepting narratives of truth in modern society, it follows that individuals must learn to learn to critically evaluate the results of their search inquiries. Community Psychology owes a great debt to the theorist and practitioner Paolo Freire, and in particular his concept of conscientization, in which oppressed communities come to recognize the ways in which they are oppressed, in order to fight back and pursue liberation (Freire, 1974; Findlay, 1994).

The general division of the stages of conscientization includes *magical consciousness*, in which circumstances of oppression are taken for granted as the status quo; *naive consciousness*, in which suspicion begins to arise regarding how and why resources and power are distributed the way they are; and *critical consciousness*, in which communities strategize and actively pursue liberation from these oppressive structures (Freire, 1974). Critical consciousness is not a stage which is arrived at, but a reflexive process which must be continually worked through (Roberts, 2000).

Freire’s concept of conscientization specifically applies to recognizing our positions within the concrete material systems of oppression and privilege that affect post-colonial societies (Roberts, 2000). However, as layers of technology become increasingly instrumental in shaping and sustaining those systems, it could be suggested that conscientizing ourselves about the unseen mechanisms of the technologies we use every day would be an equally crucial step toward empowered action. Unfortunately, it is somewhat disingenuous to divorce Freire’s powerful concept from its context in these anti-oppressive, post-colonial settings. Since the expanding commercial coercion of what appears as “true” or legitimate on the Internet affects the entire global population, I hope the appropriation of this term can be forgiven.

Granted the caveat of reorienting Freire’s terminology, conscientization is an ideal lens for awareness of the influence of these pre-determined algorithms on the conclusions we are led to reach in our searches. When we fail to acknowledge the inner workings of computers, cell phones, and the other devices we use daily, we perpetuate a new kind of “magical consciousness.” What’s more, just as the magical consciousness described by Freire made it easier for oppressive powers to enact subtle control and make large populations feel powerless, the magical consciousness of search engine users who fail to consider the steps between hitting enter and clicking the top result allows corporations like Google to seamlessly manipulate what

we call the truth. As Michel Foucault emphasized, the “truth” is merely a powerful narrative projected to, and regurgitated by, the masses, whether it is dictated by the King, the Spanish Inquisition, or the most widely endorsed voices of the Internet.

The magical consciousness of a Google user is epitomized by their clever placement of the “I’m feeling lucky” button next to the “search” button on their home page. Selecting “I’m feeling lucky” will auto-direct the user to the first search result link. This option entirely eradicates the process of evaluating the excerpts Google returns from its range of search results. Naive consciousness could perhaps be encompassed in the first stages of suspicion regarding how Google reaches the conclusions it does: users may ask themselves “how is it that Google always seems to read my mind?” Or, alternately, “why does Google always point me toward the same databases and mega-sites for every inquiry?” Finally, critical consciousness of search algorithms would involve not only the evaluation of search results, but vigilance regarding the data that Google collects from its users in order to lead to those collections of search results. If the power is in the algorithms, and the algorithms are kept secret, the end user can, at best, maintain awareness that there is a formula below the surface, and that it is based entirely on profiling and even manipulating the activity of users.

Who's Aware and Who's in Charge?

Technological literacy is not just about the ability to turn on the computer and type a query into a search engine. Critical computer literacy looms as the crucial cultural competency in the current age (Boehme, 2002). Critical computer literacy allows users to identify the difference between targeted ads and legitimate search results, to avoid potentially harmful downloads and viruses, and to fact-check suspicious information and sources (Lessig, 2006). Advocating for the free software movement, Stallman (2002) emphasized that engaging with the underlying layer of technology helps empower users, and overcomes the narrative disseminated by mainstream companies like Windows and Apple which imply computers are, at their core, confusing and difficult. Popular operating systems continue to grow more simplistic, and deny the user the ability to personalize their computer’s interface or adjust its default settings.

Companies like Google design their terms and conditions around the assumption that most users won’t be aware of, or interested in, the intensive level of tracking and data collection the company engages in. In the end, Google doesn’t just tell you what the Internet says, it tells you what it thinks you want to hear. A recent study showed that, after searching for information online, people were more likely to mistakenly identify that they themselves had already known that information (Fisher, Goddu, & Keil, 2015). A popular media source colloquialized this report as “Google makes people think they are smarter than they are,” but perhaps that’s not quite the right phrasing ([Knapton, 2015](#)). A more apt summary might instead be “Google makes people think its ideas are their own thoughts.”

The circular bias created by Google’s increasingly personalized results leads to a “filter bubble” effect, in which we see exactly what we want to see on the Internet, and lose touch with other points of view and other realities ([Pariser, 2011](#)). As this happens, it may be comforting,

but it also undermines the original vision of the World Wide Web as a radically free and open platform to share information and transcend the boundaries of our fixed locations on this planet.

Wikipedia puts a different spin on this vision, and it is possible to see this site as an enactment of this dream of free and egalitarian communication. Wikipedia's rise has been linked to “the death of the expert,” and the domination of crowdsourced knowledge has opened the ranks of academia to a wide chorus of new voices ([Bustillos, 2011](#)). However, it also leads to a crisis of trust, in which readers must carefully trace and examine the original sources cited by Wikipedia editors if they want to ensure that the entries' content reflects unbiased views (Arms, 2006). This shows that even the crowdsourced realm of Wikipedia still relies on classically hierarchical academic proof in order to legitimize its content.

Communities in the Wake of the Web: Who Owns Knowledge?

The Internet is largely dominated by the English language and its accompanying Global Northern ideologies, thus we can see it as the newest iteration of intellectual colonization, which must be treated in a similar fashion to the manipulation of any other sort of resource. A single example lies in the fact that English language searches for the term “Africa” produce multiple pages of results about Africa which originate from websites in North America and Europe, before finally returning primary sites derived from the continent of Africa itself (Srinivasan, 2012). Additionally, Google’s narrow news stream serves to further “intensify US and Western perceptions of the world” (Segev, 2005, p. 168). This may not be a huge change from the already existent biases of previous mainstream media sources, but Internet search engines and crowdsourced news often make suspicious claims to a more neutral or diverse representation of global circumstances.

For Foucault, “ideal truth does not liberate discourse from power, but tightens power’s control” (Lemert & Gillan, 1982, p. 65). In other words, truth is not some sort of revelation to confront and topple authority; instead, truth is labelled as such because of the illusion of authority it possesses, no matter how revolutionary it may seem on the surface. If we equate Google’s “truth” with Foucault’s epistemic “truth,” we see that it is actually a hypnotizing tale about meaning, not the end-all be-all of meaning itself. Google’s top results construct its dominant narrative, both about what is relevant to the search terms and what is accurate in the world. Google has succeeded so widely, in part, because of its ability to garner user trust regarding these results (Sanz & Stancik, 2013). For Google’s search engine, convenience and speed tighten the control of their “version” of the Internet, and reduce the impetus to question or reflect critically on these results. Similarly, the “feedback loop” created between Google and Wikipedia reinforces this centralized control, whether intentional or not (Metz, 2009).

Significant strides have already been made in establishing transparent and open source indexing of the vast World Wide Web, such as the nonprofit project Common Crawl ([Simonite, 2013a](#)). The algorithms used for undertakings like these are completely accessible, and participation is open for any interested users. Ultimately, free and open source search projects prioritize “transparency, community, quality, [...and] privacy,” while still pursuing the scalable functionality that proprietary search engines achieve (Segev, 2005, p. 42). Of course, there is no

such thing as neutral technology, and the data available through Common Crawl must be parsed with the same critical literacy as any other search engine. The crucial difference is that its transparency allows users to attempt to trace the way those results were culled from the larger Web. It's similar to showing your work on a math test: you may be able to do the problem in your head, but the teacher would much prefer to make sure you got the "right" answer for the "right" reason.

Speaking of showing your work at school, the Internet is opening many avenues for new pathways to learning outside of the traditional classroom. For example, DuoLingo is a highly successful online program which provides users free foreign language lessons in exchange for their labour in manually translating web pages ([von Ahn, 2011](#)). This is just one of many Massive Open Online Courses, new sites of education that combine lectures, interactive exercises, and many other strategies for students all over the world to engage in self-directed study ([Pappano, 2012](#)). These courses are somewhat controversial, having been both lauded as the alternative to overpriced schooling, and criticized as overly reductive and pedestrian ([Rees, 2013](#)). Still, resources like these are clearly promising to the future of the Internet, as long as users apply the same level of critical literacy that they would in any other setting.

It's not the end of the world if we take "how to microwave a potato" at face value. It will certainly be unfortunate if the information is wrong, and the potato explodes, but it's a pretty straightforward experiment. However, if we choose to click "I'm feeling lucky" when it comes to current events, or understanding diverse and unfamiliar ideologies and perspectives, that decision might in fact trigger the end of a certain type of world. This would be the end of a possible world where debates involve listening and responding thoughtfully, instead of just waiting for your turn to speak. It would be the end of a world where we feel we always have more to learn. It would be the end of a world where we're willing to admit that sometimes even we, and the increasingly monolithic Google search engine, can be wrong.