21   Wikipedia Has a Bias Problem

Jackie Koerner

The leader of free knowledge has a bias problem. Wikipedia contributors strive for neutrality, but the reality is they are distorting knowledge equity. It's all their fault. But that doesn't mean they can't be the solution.

Wikipedia started my love affair with free knowledge and open education. Like any relationship, it had its problems. Over the years, I felt hurt, inspired, embarrassed, hopeful, and unsafe. First, in 2001, I found a new way to discover information. In 2008, I received harassment on-wiki for the first time. In 2009, the dean resented my passion for open education, and as a result, my graduate degree sat in limbo. In 2014, I used Wikipedia as a springboard for my research. In 2017, thousands of academics laughed at me as I stood on stage and said, "I edit Wikipedia." The uncertainty fed doubts I held about Wikipedia. This caused me to stop and reflect on what Wikipedia meant to me.

Initially, I knew nothing about educational inequities.[1] I grew up in a privileged part of the United States. My school received sufficient funding enabling it to easily meet the educational needs of students.[2] I only knew what I experienced. That changed in 2001 when I became acquainted with Wikipedia. I learned what free knowledge means. I admit I initially consumed content and contributed nothing in return. That changed in 2016 when Wikipedia helped me when I needed it the most. It gave me a purpose when I felt I had none. Quickly, I realized how much Wikipedia needed me too.

Working with Wikipedia is tricky. Working with my peers to improve access to knowledge brings me joy. But it pains me to see denigration. For example, to make more room in the category "American novelists," contributors removed women novelists from the category.[3] The women novelists were placed in a subcategory named "American women novelists." The people

moving women from one category to the other thought they solved a problem but they exacerbated another. They didn't consider why they moved the women and not the men nor the marginalization they were perpetuating. This example illuminates how our actions affect Wikipedia in subtle ways.

These subtle ways affecting Wikipedia are bias. Bias creates an unwelcoming environment for people and content on Wikipedia. Our biases influence societal structures, practices, and principles. It's no different for Wikipedia. After twenty years of development, Wikipedia still prevents the very thing it set out to change.

What Is Knowledge Equity?

In 1948, the Universal Declaration of Human Rights listed education as a human right.[4] Despite the simplicity of this notion of *equity meaning everyone*, its realization remains out of reach: more than 262 million youth do not attend school, with six out of ten youths struggling to obtain basic literacy—leading to 750 million illiterate adults.[5]

Wikipedia has been a radical force in providing material for education. Examples include Wiki Education and their work on improving student learning outcomes using Wikipedia in the United States (chapter 20); the Wikipedia Library, provisioning access to paywalled databases for Wikipedia contributors (chapter 8); and the Wikipedia+Libraries: Better Together program, strengthening the relationship between public libraries and Wikipedia (chapter 6).[6]

Yet Wikipedia's aspiration of sharing the "sum of all human knowledge" falls short. While Wikipedia has dramatically increased the accessibility of knowledge, the type of knowledge available remains incomplete.

The Wikimedia 2030 project envisions free knowledge as truly representative of human diversity. Nine teams with over a hundred community members, including myself, work to outline the services and structures necessary for greater participation and representation. The Wikimedia 2030 project declared that

> as a social movement, we will focus our efforts on the knowledge and communities that have been left out by structures of power and privilege. We will welcome people from every background to build strong and diverse communities. We will break down the social, political, and technical barriers preventing people from accessing and contributing to free knowledge.[7]

This is knowledge equity—the participation and presence of all people is the only way we can achieve equity in the knowledge presented on Wikipedia.

Where Wikipedia Fails Knowledge Equity

Uneven participation and representation on Wikipedia reproduce knowledge inequality. These structures of power and privilege survive without intentional efforts to disrupt them. The community structures of power and privilege perpetuated by bias act to interrupt Wikipedia's potential knowledge equity. Not challenging bias is worse than ignorance. We know it exists and choose to do nothing. Bias disrupts everything, and allowing it to spread uncontrollably will lead to Wikipedia's demise.

The Wikimedia 2030 commitment describes what Wikipedia hopes to achieve for knowledge equity. The action to achieve such, however, has not yet been defined. Wikipedia is the platform to support knowledge equity, and of course, a provider of free knowledge should practice knowledge equity. The people building the encyclopedia need to practice it too.

What Is Bias?

Bias consists of the thoughts and beliefs we have about society. We learn these biased thoughts and beliefs from family, friends, and the media. Bias is not based on facts, and it is socially constructed. Learning bias is not conscious or deliberate.

We all have biases. People are not bad for having biases. Bias influences our actions, beliefs, relationships, and even our work. The most common biases people think about when they hear the word bias are gender, sexual orientation, and racism. People feel if they are not acting in overtly sexist or racist ways, they are not biased. Acting on our biases is completely unconscious. Just like the example of the contributors moving women novelists, we generally do not intend to act in biased ways. It's often completely unintended, but that does not mean the result is any less harmful.

Learning bias is unavoidable and completely unconscious, but this does not mean we get to absolve all responsibility. Recognizing our own biases is hard work. It's easier to identify bias in others than it is in ourselves. We tend to join groups and seek information that confirms our thoughts and beliefs.

There is much work being done regarding gender bias on Wikipedia. While the gender bias does largely imbalance Wikipedia, this is not the only bias working to misrepresent knowledge.

Bias Is a Problem for Wikipedia

While we try to be neutral, our work on Wikipedia will always involve bias. Bias can appear in many areas, like Wikipedia's policies, practices, content, and participation.

Bias leads to barriers to inclusion. These barriers mean imbalanced participation and distorted knowledge. The most recognizable barriers relate to contributor retention, emerging communities, and content exclusion.

Disruption of bias is hard. The most common example is demonstrated by the harried response that proposed changes to policy or practice receive. Contributors who unquestioningly defend policy or practice make it difficult to implement inclusive changes. This happens because they are not seeking to understand but rather to be heard.

Confirmation bias occurs when people feel reaffirmed in their beliefs due to their interpretation of information. This happens when contributors read other discussion comments that support their perspective or see how often the policy works rather than considering where it does not work. This behavior maintains the problematic power dynamics within Wikipedia's community and prevents the project from encompassing underrepresented knowledge.

Examining one's own bias is difficult. Here is an example. At Wikimania 2017, I presented a session about bias.[8] When I completed my presentation and asked for questions, one person stood up. They asked, how you tell someone they are wrong when they tell you that you acted in a biased manner? I was delighted when the room filled with chatter and murmurs.

When a person is called out for bias, it's usually warranted. This person felt they were wronged when someone brought their bias to their attention. I invited people from the audience to answer the question. Many responded with content from the session, and some even shared personal stories of how bias tormented them and disrupted their work.

After thoughtful responses, encouragement, and honest vulnerability, the person still refused the possibility that they could be biased. This

frustrated them so terribly that for the remainder of the multi-day event they tried to convince me of my faulty assumption.

This resistance to addressing personal bias still haunts the contributor, as they posted about the interaction nearly two years later. They wrote a post in a discussion on Wikipedia explaining their experience, expressing the feeling the audience in the room that day judged them unfairly. The person went on to ask a similar question: how to tell someone from a "minority group" they are wrong about encyclopedia writing, without them thinking it is a white man abusing his power privilege. Unfortunately, this is a white man abusing his power privilege. They choose to remain moored in their ways about encyclopedia development, knowledge curation, and equity instead of asking questions like, "How are we excluding people and their knowledge by doing things this way?" Asking this question might end up being a real eye-opener for a lot of people and could advance Wikipedia toward achieving knowledge equity.

Acknowledging bias is hard, and while it is painful work, it is critical. Wikipedia grapples with bias, and we need to be honest about our role in it. We all need to be aware of the problem and take action to reduce the influence of bias.

Where Bias Shows Up in Policy and Practice

Wikipedia policies and practices largely follow Westernized traditions of knowledge sharing and information publishing. More inclusive changes to the policies and practices are difficult to undertake. This sends the message that quality means Westernized practices and excludes anything and anyone not following these arbitrary principles.

Wikipedia was built in the early 2000s. The internet was very different back then. People published anything they wanted on their websites. Making online purchases seemed risky. Teachers laughed when students suggested doing online research. Policies and practices developed on Wikipedia responded to the problems the internet dealt with at the time. These problems still exist, but we have learned a lot considering the internet is so integrated into our lives. Although the internet has changed over the past twenty years, the policies and practices on Wikipedia have kept their same rigid beginnings.

Wikipedia materialized through predominantly Westernized cisgender male voices, opinions, and biases. The awareness in the community at that time illustrated a rather singular point of view and developed policies and practices accordingly. This foundation is difficult to break. Preference on Wikipedia concerning changes or inclusion is still very singular and causes diverse participants to have to work within the dominant culture.[9]

In the example I gave in the previous section, I feel the contributor was telling themselves the story of: "This is the policy. They are not following the policy. I will educate them about this policy." I hoped the Wikipedia contributor would have listened and reflected on the information and vulnerability being given in that room. We knew this contributor meant no ill will but saw they were stuck viewing the world through their perspective and their bias.

If they had reflected on the interaction, what they would have taken away from the session would have been very different. Perhaps they would see how, while not meaning to do so, they were applying their biased perspective on the situation and telling others in the community how things should be done.

We are all victims of the stories we tell ourselves. The response in situations like this should not be holding our policies so tightly that we cannot figure out how to listen to concerns. We should adapt our policies to a more inclusive and equitable world.

This narrow and inflexible behavior functions within the Wikipedia community to oppress and exclude. Simply because experience and history have been traditionally told from white, cisgender, male perspectives, these voices and perspectives within society are taken as fact when often they are opinions or interpretations. We all experience life from our lived experiences; the Wikipedia community is no different.[10]

By infusing homogeneous points of view into policies and practices of a community, a disservice is being done. Content and people are being removed and excluded if they do not fit into the policies and practices designed by the existing cohort of contributors.

How Reliable Sources Are a Bad Thing

One important policy is a good example of a well-intended bias perpetuating knowledge inequity. The reliable sources policy limits the sources and

forms of information used on Wikipedia. This policy developed out of a need to keep people from posting unsubstantiated claims to the encyclopedia. Requiring reliable sources is a good thing, but the implementation on Wikipedia is the opposite.

Defining materials to fit the policy means limitations. As of this publication, knowledge from published, written materials with a preference toward academic and peer-reviewed publications epitomizes reliability. The reliable sources policy limits knowledge equity by ignoring knowledge that falls outside of the rules.

The knowledge available in published, written materials is biased. People and knowledge published in written materials are largely white and male. The way the current reliable sources policy is written and followed leads to an information imbalance on Wikipedia. There is far less content about women on Wikipedia than there is about men: as of October 2019, only 18 percent of biographies on Wikipedia are about women.[11]

The bias toward Westernized publications and knowledge-sharing practices exaggerates the lack of diverse content on Wikipedia. If there is no source about a person (or a topic) to meet the standards of the Wikipedia community, then no article will be written. That person is excluded from history. By following policies like reliable sources, contributors are replicating and magnifying the bias already depicted by published sources.

Contributors use their personal beliefs to determine, design, and follow policies. The dilemma grows when those in power within the Wikipedia community deny agency to those challenging the policies and practices of the Wikipedia community.

The policies around reliability are often applied in a way that removes anything varied or diverse. We should aim for balance in content. We should provide knowledge from diverse sources. Instead, we are refusing to listen to one another. This has to stop. Information is not accurately represented. Contributors are pushed away. Knowledge is lost.

What Went Wrong?

In the Wikipedia community, people are not listening to each other. Collaboration devolves into combative discourse. Discussions surrounding knowledge equity, reliability, verifiability, and neutrality draw their energy from bias. Communities and knowledge remain excluded.

The community often reacts to questions about the policies, practices, and community norms in a defensive way. Notable people cannot be documented because nothing about them exists in an acceptable published format.[12] Information is discredited, even when quoted from the subject in an oral history.[13] Women scientists only become notable because of an award while their male colleagues were notable before any such acknowledgment.[14]

Although social groups with power possess the privilege to address imbalances caused by bias, the responsibility for abolishing ignorance unfairly lies with the excluded or oppressed. This emotional labor taxes an already overtaxed individual and community. This is no different on Wikipedia. "Be bold!" But being bold can be risky. Anyone has the power to enact change, but power structures privilege long-time contributors, administrators, and policy writers. Within the Wikipedia community, these groups work together to deny change.

For example, when discussing information gatekeeping as a worrisome practice, another contributor disagreed with me. Instead of engaging in the discussion and trying to create a solution, they chose to "read all forty-eight pages" of my website and sent me a message about being "great enemies" if I disagreed with them. The person justified this behavior by mentioning the countless hours and thousands of dollars they contributed to Wikipedia.

This uncomfortable experience was mild compared with that of others, but they all have the same intentions: to silence diverse voices. The rampantly unchecked power dynamics within the community function to silence the voices aiming to address bias in content and policy on Wikipedia.

The Wikipedia we need must eliminate narrow policies and practices and elevate the culture to become inclusive. If we are not reaching the people who need our service, we are practicing inequity.[15]

There Is Hope

Education serves as a great tool for social mobility and stability. Wikipedia paints an illustration of daily acts of human decency. Here the Wikipedia community works tirelessly for societal good via an unmatched source of volunteerism. Festering imbalances exist due to the rather homogeneous composition of the contributor pool, the restrictive policies they created, and the inconsistent way in which the policies are practiced. The sum of all human knowledge cannot be built under these conditions where logic

is designed out of the illogical. But there is hope. Wikipedia is just turning twenty.

How to initiate this change is for no one person or homogeneous group to undertake. The curation of knowledge, development of policies, and denial of change decided in such groups stunts the growth of Wikipedia. It serves no purpose or benefit to the longevity or growth of Wikipedia. Continuing to accept community consensus developed through homogeneous groups will keep Wikipedia in a perpetually sophomoric state. We should be concerned about quality sources, and we need to listen to what quality sources mean across the world. We should volunteer, and we should make space at the table for more people to volunteer too. There is not a limited amount of equity.

Equity comes from actions people take against oppressive and imbalanced policies and practices in society. We can change the world if we choose to enact equitable policies and practices on Wikipedia, refuse to manipulate discourse to squash diverse perspectives, and acknowledge that change is not scary but rather impressive. Without change, we continue to inflict and deepen wounds opened by oppression, exclusion, and continued ignorance.

What Can Be Done

The content in this chapter might sound familiar to you, but it might also be different in many ways. By listening to each other's lived experiences, we change together as a community. We must not only accept what we find acceptable for ourselves but accept what is needed for knowledge equity.[16]

In addressing knowledge equity and implementing these changes to reduce the effect of societal challenges on Wikipedia, we must proceed with care. Some areas need more structured support than others—such as setting stronger cultural norms and being empowered to act on bad behavior.[17] We should encourage growth through methods of listening, witnessing, and advocating. Growth this way can change the environment for Wikipedia and knowledge equity.

Wikipedia as a community and an encyclopedia has accomplished some amazing things in the first twenty years of its life. We learned what it means to collaborate online to build an encyclopedia. Much information has been developed about online communities, online collaboration,

and information sharing. Educators and knowledge professionals began using Wikipedia to teach information literacy, regardless of the concerns and issues colleagues expressed. The Wikipedia community has come so far, but there is so much more to be done. If there is any hope for truly achieving the sum of all human knowledge, the next chapter in Wikipedia's life needs to meaningfully address the inequities perpetuated by bias. Although unfinished, rather progressive in some circles, and a little rough around the edges, even at twenty Wikipedia is the experimental educational equalizer and the solution to knowledge equity. We just have to stop preventing its success.

Notes

1. UNESCO Institute for Statistics, *Reducing Global Poverty through Universal Primary and Secondary Education*, (Policy Paper 32/Fact Sheet 44, UNESCO, June 2017), https://unesdoc.unesco.org/ark:/48223/pf0000250392.

2. Cory Turner, "Why America's Schools Have a Money Problem," *Morning Edition*, NPR, April 18, 2016, https://www.npr.org/2016/04/18/474256366/why-americas -schools-have-a-money-problem.

3. Alison Flood, "Wikipedia Bumps Women from 'American Novelists' Category," *The Guardian*, April 25, 2013, accessed July 20, 2019, https://www.theguardian.com /books/2013/apr/25/wikipedia-women-american-novelists.

4. Wikimedia, s.v. "File:The Universal Declaration of Human Rights 10 December 1948 .jpg" (United Nations), last modified March 12, 2019, https://commons.wikimedia.org /wiki/File:The_universal_declaration_of_human_rights_10_December_1948.jpg.

5. "Leading SDG4—Education 2030," UNESCO, n.d., accessed July 29, 2019, https://en .unesco.org/themes/education2030-sdg4.

6. Wikimedia Commons, s.v. "File:Student Learning Outcomes using Wikipedia-based Assignments Fall 2016 Research Report.pdf," June 1, 2017, accessed June 10, 2019, https://commons.m.wikimedia.org/wiki/File:Student_Learning_Outcomes_using _Wikipedia-based_Assignments_Fall_2016_Research_Report.pdf.

7. Meta-Wiki, s.v. "Knowledge Equity," s.v. "Strategy/Wikimedia Movement/2017/ Direction," accessed June 16, 2019, https://meta.wikimedia.org/wiki/Strategy/Wiki media_movement/2017/Direction#Knowledge_equity:_Knowledge_and_communi ties_that_have_been_left_out_by_structures_of_power_and_privilege.

8. Wikimedia Commons, s.v. "File:Koerner Implicit Bias Wikimania 2017.pdf" (Jackie Koerner), accessed on June 10, 2019, https://commons.wikimedia.org/wiki/File :Koerner_Implicit_Bias_Wikimania_2017.pdf.

9. Amanda Menking, Ingrid Erickson, and Wanda Pratt, "People Who Can Take It: How Women Wikipedians Negotiate and Nagivate Safely," in *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York: ACM, 2019), https://dl.acm.org/doi/10.1145/3290605.3300702.

10. Marti Johnson and Alex Wang, "Wikimedia Foundation Releases Gender Equity Report," Wikimedia Foundation, September 21, 2018, https://wikimediafoundation .org/2018/09/21/advancing-gender-equity-conversations-with-movement-leaders/.

11. Wikipedia, s.v. "Wikipedia: WikiProject Women in Red," accessed October 26, 2019, https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Women_in_Red.

12. Wikipedia, s.v. "Notability Is Geared towards the White Male Perspective," s.v. "Wikipedia talk: Notability/Archive 63," accessed June 10, 2019, https://en .wikipedia.org/wiki/Wikipedia_talk:Notability/Archive_63.

13. Wikipedia, s.v. "Wikipedia: Oral History," accessed on June 10, 2019, https://en .wikipedia.org/wiki/Wikipedia:Oral_history.

14. Dawn Bazely, "Why Nobel Winner Donna Strickland Didn't Have a Wikipedia Page," *The Washington Post*, October 8, 2018, https://www.washingtonpost.com /outlook/2018/10/08/why-nobel-winner-donna-strickland-didnt-have-wikipedia -page/?utm_term=.f0c748d01376.

15. Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (New York: Oxford University Press, 2007).

16. Walter Frick, "Wikipedia Is More Biased Than Britannica, but Don't Blame the Crowd," *Harvard Business Review*, December 3, 2014, https://hbr.org/2014/12 /wikipedia-is-more-biased-than-britannica-but-dont-blame-the-crowd.

17. "Your Code of Conduct," Open Source Guides, n.d., accessed June 10, 2019, https://opensource.guide/code-of-conduct/.