

Bridging Wikipedia's gender gap, one article at a time

February 25 2022, by Alina Ladyzhensky



[Editing Wikipedia](#). Credit: Teemu Perhiö for [Wikimedia Finland](#), [CC BY-SA 2.0](#)

As the world's largest and most-used information resource, Wikipedia is home to 6.4 million articles and counting. But despite how comprehensive it seems, 90% of the site's editors are men, and women

are vastly underrepresented as subjects in the encyclopedia. The problem is particularly glaring when it comes to biographical information. Of the 1.5 million biographical articles on the site, less than 20% are about women.

A new study co-authored by Isabelle Langrock, a Ph.D. candidate at the Annenberg School for Communication, and Annenberg Associate Professor Sandra González-Bailón evaluates the work of two prominent feminist movements, finding that while these movements have been effective in adding a large volume of biographical content about [women](#) to Wikipedia, such content remains more difficult to find due to structural biases.

When it comes to research on gender gaps in digital information projects like Wikipedia, many studies focus on measuring and mapping the problem, in order to understand its extent and severity. But Langrock, who studies how groups work to create equitable public information online, wants not only to spotlight the problem, but also to offer solutions—including how to make existing feminist efforts more successful.

Langrock and González-Bailón's study in the *Journal of Communication*, "The Gender Divide in Wikipedia: Quantifying and Assessing the Impact of Two Feminist Interventions," looks at two non-profit groups with similar missions: Art+Feminism is dedicated to adding content about women and nonbinary artists to Wikipedia, while 500 Women Scientists, a nonprofit that aims to improve representation and inclusivity in STEM, creates and edits Wikipedia pages for women scientists as part of its public outreach. Both groups add and update Wikipedia content through "edit-a-thon" events held in library and museum archives, universities, and similar spaces, enabling them to gather as much information as possible from both digital and physical reference materials.

In the study, the researchers measured the outcomes of this work by analyzing more than 11,000 biographical articles, including 3,000 articles that were edited or created at the "edit-a-thons." In order to measure the interventions' impact, they compared these articles with 8,000 biographical entries not connected with the edit-a-thons, including profiles of men in professions covered by the interventions (artists and scientists), and women and men in professions with no associated feminist intervention (athletes and politicians).

They then looked at two different outcomes.

In the first, Langrock and González-Bailón measured how many new articles the edit-a-thons created, as well as those articles' length, quality, and pageviews.

What they found was that the interventions were successful both in creating new articles about women and increasing article views.

While Wikipedia pages about men tend to be longer and receive more views, the intervention flipped the script. The edit-a-thons created more extensive biographical articles for women, including 250 entirely new entries, that averaged more views than either men's pages or non-intervention women's pages.

The second outcome is how the articles were connected to the entire network of content—in other words, how easy they were to stumble upon. On that measure, the edit-a-thon content fell short.

The researchers found that the intervention articles about women used fewer infoboxes. Infoboxes are indexed summaries that appear on the top right corner of Wikipedia articles and offer quick links and metadata. They help build connections to related articles, which increase the likelihood that people will find that content. Adding infoboxes to

biographies, along with identifying and linking related pages—for example, a scientist's mentor or an artist's collaborator—builds the importance of biographical pages in the network of links that connects Wikipedia's articles.

"These features are important for thinking about how Wikipedia data permeates across the internet, and how people use the site to find information," Langrock says. "An estimated 20% of Wikipedia traffic is driven through these knowledge network links, which is really interesting to consider because it's often hidden under other inequality measures."

"The divides we analyze in this article have repercussions beyond Wikipedia," adds González-Bailón, who directs the Digital Media, Networks, and Political Communication Group at Annenberg. "They have an impact on social perceptions of knowledge, but they can also propagate beyond Wikipedia as its contents are leveraged to correct misinformation, feed content to AI devices, or improve search engine results."

Artists and scientists have fewer infoboxes than the comparison groups, and when infoboxes do exist, women's are not as comprehensive. Women are also less represented in articles beyond their own biographies—for example, articles about institutions or mentors. This makes them less visible in the network of links that connect pages. As a result, readers aren't as likely to stumble onto women's biographies when spontaneously hopping from page to page.

"This puts them on the fringes of the knowledge network," Langrock says. "If you start at any given article on Wikipedia, you're much less likely to eventually reach an article about a woman artist than you are about a male artist—and this was true for women across the board."

As the authors note, these structural aspects haven't been a major focus

of prior efforts to close Wikipedia's gender gap. Adding new content and longer articles about women addresses one aspect of the disparity, but doesn't improve biases and inequities on other parts of the platform.

To this end, Langrock and González-Bailón encourage activist groups and Wikipedia editors to improve the coverage of infoboxes and increase the number of links pointing to newly created content. They also recommend that future work on these inequalities—and related gaps, like racial inequities—distinguish different types of bias, which may require different types of interventions.

Langrock also presented these findings last summer to the Wikimedia Foundation, in the hopes that the knowledge would help ongoing efforts to reduce gender gaps.

Researchers can learn a lot from the work of online activist movements, Langrock emphasizes, both in terms of what tactics are effective and where changes would help. Studying these interventions can illuminate different aspects of inequality on Wikipedia and how to better target them.

"Focusing only on the dramatic gender gaps implies that no one's working to solve the problem and that there isn't a solution," she says. "A lot of groups are actively trying things out, and as researchers, we can help them determine what's working and what isn't. We need to help activist groups by highlighting their successes and building tools to help them do better at integrating women's pages into the knowledge network as a whole."

More information: Isabelle Langrock et al, The Gender Divide in Wikipedia: Quantifying and Assessing the Impact of Two Feminist Interventions, *Journal of Communication* (2022). [DOI: 10.1093/joc/jqac004](https://doi.org/10.1093/joc/jqac004)

Provided by University of Pennsylvania

Citation: Bridging Wikipedia's gender gap, one article at a time (2022, February 25) retrieved 18 April 2024 from <https://phys.org/news/2022-02-bridging-wikipedia-gender-gap-article.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.