

Bits and Bias: The Case of Women in Religion

Clifford B. Anderson

This paper introduces the *Women in Religion* project as a grassroots movement to identify gaps in Wikipedia's coverage of significant women religious figures by contributing articles and other forms of information to Wikipedia and related projects. The existence of gender gaps on Wikipedia is widely noted by both researchers and the community of editors. In the *Community Insights 2021 Report*, a self-study conducted annually by the Wikimedia Foundation, only 15 % of editors reported being women, non-binary, or another non-male gender.¹ This gender gap among editors roughly correlates with the percentage of biographies of women (20 %) versus men (80 %) on Wikipedia.² The Wikimedia Foundation, along with many affiliated movements, is actively striving to address these imbalances. A challenge to correcting bias for women in religion is the underlying need to ascertain and impute sensitive information about the gender and religious affiliation of subjects. In the case of the *Women in Religion* project, I argue that the relational epistemology at the center of the movement provides a means, at least provisionally, to handle such sensitive information judiciously.

In the first part of this study, I examine the multifaceted nature of the gender bias that the *Women in Religion* movement seeks to address as it manifests itself across Wikimedia projects. I then place the movement in relationship to kindred groups while highlighting its distinctive approach to addressing bias. My hope is to show that, by setting its editing activities within an encompassing circle of scholar-practitioners of theology and religious studies, members of the *Women in Religion* community respect both editorial guidelines and the sensitivity of the information they contribute about their subjects.

1 See [https://meta.wikimedia.org/wiki/Community_Insights/Community_Insights_2021_Report/Collaboration,_Diversity_%26_Inclusion_\(2021\)](https://meta.wikimedia.org/wiki/Community_Insights/Community_Insights_2021_Report/Collaboration,_Diversity_%26_Inclusion_(2021)) (April 30, 2025).

2 See https://meta.wikimedia.org/wiki/WikiGap/Virtual_WikiGap (April 30, 2025).

1. *The Universe of Wikipedia*

Wikipedia has established itself as a secondary source of information about nearly every topic of interest since its inception on January 15, 2001. Wikipedia provides robust articles on ranging from arcane mathematical theories to the latest Japanese anime series. These articles are written, edited, and reviewed by volunteers. From its humble origins as an experimental internet encyclopedia launched by Jimmy Wales and Larry Sanger, Wikipedia has grown to a top ten website, the only nonprofit among giants like Google, YouTube, and Facebook.³

In casual parlance, people speak of “Wikipedia” as if it were a unitary online encyclopedia. Following Alfred North Whitehead, let us label this way of speaking as a “fallacy of misplaced concreteness.”⁴ This fallacy frequently appears in the popular press about Wikipedia. “Too often, you read ‘Wikipedia says...’ or ‘Wikipedia decided...’ or ‘Wikipedia rejected my edit,’” remark Ian A. Ramjohn and LiAnna L. Davis. “News coverage of Wikipedia often speaks about it as an amorphous entity. But it is not a faceless collective – Wikipedia is a collection of individuals with varying motivations and approaches to interacting with other humans.”⁵ The discovery that Wikipedia consists of multiple language editions, as well as many interlocking projects, proves equally surprising to casual users.⁶ Just learning how the parts of Wikimedia interconnect and how to work effectively across platforms proves daunting for newcomers.

2. *Wikimedia Foundation*

The Wikimedia Foundation is a California-based nonprofit that supports the technical infrastructure of Wikipedia and its related projects. The Wikimedia Foundation provides administrative, financial, legal, and technical support for Wikimedia projects and also conducts outreach to the

3 See <https://www.alexa.com/topsites> (April 30, 2025).

4 Alfred N. Whitehead, *Process and Reality* (Free Press, 2010), 18, <https://books.google.com/books?id=uJDEx6rPu1QC>.

5 See <https://wikipedia20.pubpub.org/pub/teosjfcn/release/1> (April 30, 2025).

6 The Wikimedia Foundation led a concerted effort to re-brand itself as the “Wikipedia” foundation in 2019 and 2020, but ultimately failed to garner sufficient support from its community; see <https://wikimediafoundation.org/news/2019/02/26/leading-with-wikipedia-a-brand-proposal-for-2030/> (April 30, 2025).

volunteer community as well as the public. At present, the Wikimedia Foundation maintains thirteen projects, including Wikipedia, Wikidata, and Wikimedia Commons. It also maintains the underlying MediaWiki software and coordinates activities and events related to these projects and the wider movement on a site called Meta-Wiki.⁷

The Wikimedia Foundation also recognizes affiliates that align with its mission. These affiliates may take the form of chapters, thematic organizations, and user groups.⁸ Chapters and thematic groups are independently incorporated nonprofits, while user groups are unincorporated. These affiliates provide opportunities for Wikipedians to form connection either geographically or on matters of topical interest and to instruct and socialize prospective editors in the arcana of editing Wikimedia projects. As we shall see in the second half of this paper, these affiliates also advocate for changes in editorial policy and practice as well as serve as informal liaisons between Wikimedia and other professional communities.

3. Wikipedia

Wikipedia has been termed the “World’s Greatest Encyclopedia.”⁹ By any quantitative measure, this appellation appears appropriate. By sheer size alone, Wikipedia towers over its competition. In spring 2022, the English language Wikipedia contained approximately six and a half million articles.¹⁰ By way of comparison, the largest print encyclopedia musters one million articles.¹¹

There are 325 language editions of Wikipedia at present.¹² The language editions range in size from the very active English-language Wikipedia (with ~130,000 active users) to the relatively quiescent like the Palatine German edition (13 active users) and Tahitian edition (16 active users). These language editions are self-governing, following contextual rules and norms. “Each Wikipedia language edition is created in a decentralized

7 See <https://meta.wikimedia.org/> (April 30, 2025).

8 See https://meta.wikimedia.org/wiki/Wikimedia_movement_affiliates (April 30, 2025).

9 Andrew Lih, *The Wikipedia Revolution: How a Bunch of Nobodies Created the World’s Greatest Encyclopedia* (Hachette Books, 2009).

10 See https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia (April 30, 2025).

11 See https://en.wikipedia.org/wiki/Wikipedia:Size_comparisons (April 30, 2025).

12 See https://en.wikipedia.org/wiki/List_of_Wikipedias (April 30, 2025).

way,” note Marc Miquel-Ribé and David Laniado in “Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions,” “as a result, editors themselves may not always be aware of the global product.”¹³ What makes a good article in one language edition may not translate successfully into another edition. When translating an article from one language edition to another, editors may discover that the topic is not considered sufficiently notable in the target language edition or requires supplemental references to secondary literature.

There are also differences between Wikipedia editors. In theory, anyone can edit Wikipedia. You do not even need to log into Wikipedia to make edits. In practice, Wikipedia distinguishes between categories of editors based on the longevity of their accounts as well as the number of edits they have contributed. For example, a user account that has “both existed for more than 30 days and made at least 500 edits” is numbered among the “extended confirmed users.” These editors represent the primary cadre of contributors. On the English-language Wikipedia, there are approximately 60,000 extended confirmed users. Editors on Wikipedia not only create new articles and revise existing articles; they also review articles, merge related articles, and, in some cases, nominate articles for deletion. A much smaller group of editors who have been elected as administrators carries out tasks like protecting pages, blocking users, and deleting articles.

4. *Wikimedia Commons*

Wikimedia Commons describes itself as a “media file repository” of public domain and openly-licensed educational content.¹⁴ Wikimedia Commons plays a crucial role in Wikipedia and related projects by storing and serving the audiovisual media for those projects. By collecting media in a single location, Wikimedia Commons makes it possible for an editor of the English-language Wikipedia to illustrate an article on, e.g., the Taiwanese President Tsai Ying-Wen using an image contributed by a Wikipedian in Taiwan. But Wikimedia Commons serves a broader purpose than providing a media repository for Wikipedia. Wikipedia hosts formats like PDF documents (for example, conference presentations as well as scans of documents) and

13 Marc Miquel-Ribé and David Laniado, “Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions,” *Frontiers in Physics* 6 (2018): 1, doi:10.3389/fphy.2018.00054.

14 See <https://commons.wikimedia.org/wiki/Commons:Welcome> (April 30, 2025).

some kinds of datasets (e.g., GeoJSON data for maps and CSV for tabular datasets).

Wikimedia Commons follows stringent guidelines to protect intellectual property. When uploading material under copyright to Wikimedia Commons, users must attest that they either created the media themselves and therefore own the copyright or that the copyright holder has agreed to share the material under an applicable license. These policies mean “contributors to Commons must also navigate intricate policies concerning copyright and permissions, reoccurring vandalism, problems with other users, and photography critiques,” observe Menking et al. “Similar to Wikipedia, these factors create a sociotechnical environment that can be intimidating for new users.”¹⁵ They note that the difficulty of negotiating these barriers has significant consequences because readers judge articles without accompanying images to be of lesser quality.¹⁶

Representational distortions become perspicuous on Wikimedia Commons. If a woman does not have an image in Wikimedia Commons, editors will not be able to add a photograph to the infobox of a prospective article on Wikipedia. Does the absence of a photograph on Commons also subtly suggest that she may not merit an article on Wikipedia? There may be good reasons, however, that someone might not want to have a photograph of themselves on Wikimedia Commons. As Menking et al. note, survey participants were concerned about potential violations of privacy. “In particular, respondents expressed reluctance to share their location and/or to share photos of people.”¹⁷ If you upload a photograph of yourself to Wikimedia Commons, you explicitly give up the right to control its downstream uses, making it possible for a stalker or other nefarious actor to use that photograph against you—for example to create a deepfake or nonconsensual pornography.

15 Amanda Menking, Vaibhavi Rangarajan, and Michael Gilbert, “Sharing small pieces of the world: Increasing and broadening participation in Wikimedia Commons,” *Proceedings of the 14th International Symposium on Open Collaboration*, 2018, 1, doi:10.1145/3233391.3233537.

16 Ibid.

17 Ibid., 8.

5. Wikidata

Wikidata was established in October 2012 to provide a central repository of factual assertions or claims in the various language editions of Wikipedia. Denny Vrandečić articulated his vision for Wikidata in a series of workshops and articles.¹⁸ At its heart, Wikidata is composed of quasi-factual statements. These statements are built out of items, properties, and values. Items receive identifiers that begin with an uppercase Q: Q2 (“the earth”), Q2664039 (“Portrait of Madame X” by John Singer Sargent), or Q64009487 (“Laurel Schneider,” a Christian theologian). Properties start with an uppercase P: P106 (*occupation*), P734 (*family name*), and P21 (*sex or gender*). A statement relates an item (i.e., a subject) with a property (i.e., a predicate) to a value (i.e., an object). A statement may connect two items by way of a property, e.g., “Rosemary Radford Ruether” (Q198346) was *educated at* (P69) “Claremont School of Theology (Q5126331). Or a statement may connect an item to a quantitative value such as a number, string, or date, e.g., “Claremont School of Theology” (Q5126331) has an *inception* date of “1885.” An item on Wikidata will typically have many statements that relate it to other items and, conversely, be the object of many other items on Wikidata.

Who decides what statements should be made? In “Wikidata: A Free Collaborative Knowledgebase,” Vrandečić outlined a series of design principles of Wikidata. Among them is the stipulation that “Wikidata gathers facts published in primary sources, together with references to these sources.”¹⁹ Vrandečić elsewhere contends that statements on Wikidata do not claim to be true but only verifiable.²⁰ Here verifiability has the specific sense of being grounded in a reliable source. This distinction between truth and verifiability allows Wikidata to maintain conflicting statements. As long as both statements have reliable sources, Wikidata allows contradictory assertions to stand alongside each other.²¹

18 Denny Vrandečić, “Wikidata,” *Proceedings of the 21st International Conference Companion on World Wide Web – WWW ’12 Companion*, 2012, 1063–64, doi:10.1145/2187980.2188242; Denny Vrandečić, “The Rise of Wikidata,” *IEEE Intelligent Systems* 28, no. 4 (2013): 90–95, doi:10.1109/mis.2013.119; Denny Vrandečić and Markus Krötzsch, “Wikidata: a free collaborative knowledgebase,” *Communications of the ACM* 57, no. 10 (2014): 78–85, doi:10.1145/2629489.

19 Vrandečić and Krötzsch, “Wikidata: a free collaborative knowledgebase,” 79.

20 Vrandečić, “The Rise of Wikidata,” 90.

21 *Ibid.*, 92.

Wikidata's goal of providing "secondary data" about the world implies that every factual claim should also be accompanied by a reference to a primary source. But an examination of Wikidata's statements reveals this is not the case. Currently, there are 1,247,700,645 statements on Wikidata. Of these statements, approximately 68 % have an external reference and 6 % have an internal reference to Wikipedia. 26 % of statements on Wikidata thus do not have warrants.²²

Where do these unwarranted statements come from? While it is difficult to generalize about missing provenance, many of these claims appear to originate from editors' direct acquaintance with their subjects. For instance, the Wikidata item that describes me (Q38099106) asserts that I am an "instance of human" and that I am a citizen of the United States, among other claims. The first claim is implied by my authorship of scholarly books and articles (though I could turn out to be an artificial intelligence). It is difficult, however, to point to an explicit reference for my humanity. The warrants that do exist for "instance of human" seem by and large to stem from Wikipedia, simply pushing back the question of how to reference this kind of tacit knowledge. Why am I assumed to be a citizen of the United States? I actually added that bit of information myself, but I did not provide any reference to substantiate the claim.

Returning to the difference between language editions of Wikipedia, Wikidata provides a way of tracking articles across editions and keeping information consistent. An item in Wikidata will contain references to Wikipedia editions with corresponding articles. Wikidata, therefore, provides the global perspective that Marc Miquel-Ribé and David Laniado noted as lacking among Wikipedia editors.²³

6. *Analyzing Bias*

Commentators have noted the presence of bias in Wikipedia since the early days of the project. "Wikipedia's authors do not come from a cross-section of the world's population," observed Roy Rosenzweig in 2006. "They are

²² See <https://wikidata-todo.toolforge.org/stats.php> (April 30, 2025).

²³ Miquel-Ribé and Laniado, "Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions."

more likely to be English-speaking, males, and denizens of the Internet.”²⁴ In the more than fifteen years since Rosenzweig made this observation, the profile of Wikipedia editors has not changed fundamentally.

What causes this skew toward English-speaking North American internet-savvy males? In article, “(Re)triggering Backlash: Responses to News About Wikipedia’s Gender Gap,” Stine Eckert and Linda Steiner conduct a qualitative analysis of Wikipedian’s perspectives on the gender imbalance among Wikipedia’s editors.²⁵ They cataloged the responses that different categories of people—journalists, bloggers, and online commentators—offered by way of explanation of the gap. On the one hand, some thought the gender disparity mirrored gender gaps elsewhere; in other words, there is nothing special about the gap on Wikipedia. By contrast, another school of thought chalked up the difference to problems with Wikipedia, indicting its weakness as an encyclopedia as a reason why women prefer not to participate.²⁶ In truth, a combination of both factors seems to be at play.

Inequity on Wikipedia goes beyond gender. The disparity in Wikipedia’s coverage of notable women versus notable men has garnered the most scholarly attention. But, as Jackie Koerner remarks, “While [the gender bias] does largely imbalance Wikipedia, this is not the only bias working to misrepresent knowledge.”²⁷ A major subject of discussion these days, for example, is the predominance of English across the movement and on Meta-Wiki, which privileges English speakers over contributors from other regions of the world. Bias exists beyond Wikipedia as well.²⁸ Researchers have also documented patterns of bias in Wikidata related to racial, ethnic, and national groups. In “Analyzing Race and Citizenship Bias in Wikidata,” Zaina Shaik et al. showed evidence of under-representation of racial groups and nationalities. Shaik and her co-authors drew on external reference

24 Roy Rosenzweig, “Can History Be Open Source? Wikipedia and the Future of the Past,” *Journal of American History* 93, no. 1 (2006): 127, doi:10.2307/4486062.

25 Stine Eckert and Linda Steiner, “(Re)triggering Backlash: Responses to News About Wikipedia’s Gender Gap,” *Journal of Communication Inquiry* 37, no. 4 (2013): 284–303, doi:10.1177/0196859913505618.

26 *Ibid.*, 289.

27 Jackie Koerner, “Wikipedia Has a Bias Problem,” in *Wikipedia @ 20: Stories of an Incomplete Revolution* (The MIT Press, 2020), 314, doi:10.7551/mitpress/12366.003.0026.

28 See “English as a lingua franca of the Wikimedia movement: how do we ensure people’s inclusion?” August 25, 2021: <https://www.youtube.com/watch?v=2X6UJ25TiN8> (April 30, 2025).

sources to compute demographic baselines, then used a combination of SPARQL queries and hand-coding to quantify the racial backgrounds and nationalities of items for people in STEM fields (e.g., “scientists, software developers, and engineers.”) The research team found that white Europeans and white North Americans are overrepresented in Wikidata while members of other racial groups and nationalities are underrepresented.

“While white people make up 17.80 % of the world population, in Wikidata, white scientists made up 83.95 %, white software developers 44.08 %, and white engineers – 70.74 %. In terms of the country of citizenship of scientists, software developers, and engineers, Wikidata is skewed towards European and North American countries while underrepresenting other continents.”²⁹

The authors do not address how they treated missing statements about ethnicity (P172) and citizenship (P27) in their analyses. There is a consensus among editors that P172 should not be used to describe human beings without sources. A note in the description of the property states, “consensus is that a VERY high standard of proof is needed for this field to be used. In general this means 1) the subject claims it themselves, or 2) it is widely agreed on by scholars, or 3) is fictional and portrayed as such.”³⁰ While this note does not represent official policy, it does underscore the difficulty of measuring bias on Wikimedia projects. If editors do not contribute data about gender, nationality, and other sensitive information, how can researchers measure and compare their rates of participation? We shall return below to possible methods for handling such information sensitively.

During the past decade, the question of bias in machine learning and artificial intelligence has likewise become a central topic of concern and inquiry. Cathy O’Neill’s *Weapons of Math Destruction*³¹ and Safiya Umoja Noble’s *Algorithms of Oppression*³² show how mathematical models and search engine algorithms produce biased results and have deleterious im-

29 Zaina Shaik, Filip Ilievski, and Fred Morstatter, “Analyzing Race and Citizenship Bias in Wikidata,” *2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2021, 2, doi:10.1109/mass52906.2021.00099.

30 See <https://www.wikidata.org/wiki/Property:P172> (April 30, 2025).

31 Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown, 2016), <https://books.google.com/books?id=cbwvDwAAQBAJ>.

32 Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (NYU Press, 2018), <https://books.google.com/books?id=-ThDDwAAQBAJ>.

pacts on marginalized groups of people. In an analogous manner, Virginia Eubanks demonstrated the real-world harms that emerge from the social-political application of biased models in *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*.³³ How to address latent bias in machine learning models has become a major subject of scholarly research during the past decade.

What does the problem of bias in machine learning have to do with inequities on Wikipedia? As it happens, the presence of bias in Wikimedia projects has an outsize effect because software engineers draw on Wikipedia to power many other systems. These days, Wikipedia as a data source extends far beyond Wikipedia as an encyclopedia. As Florian Meier remarks,

“Wikipedia is significantly shaping our digital experience. Most state-of-the-art machine learning algorithms are trained on Wikipedia content, millions of searchers are dependent on Google surfacing Wikipedia links to satisfy their information needs, tech giants are using human workforce and Wikipedia for fact checking on their platforms, Wikipedia links are among the most prominent URLs on community question answering sites like Reddit or Stack Overflow, voice-operated virtual assistants like Siri and Alexa rely heavily on Wikipedia and for many users Wikipedia is a stepping stone or gateway to the larger web itself.”³⁴

As early as 2009, researchers noted that “Knowledge derived from Wikipedia has the potential to become a resource as important for NLP [Natural Language Processing] as WordNet.”³⁵ Data from Wikipedia continues to serve as a key source for the latest machine learning models. In 2019, for example, Jacob Devlin and his colleagues at Google AI introduced BERT (Bidirectional Encoder Representations from Transformers). As they noted in the paper introducing the model, “For the pre-training corpus we use the BooksCorpus (800M words) and English Wikipedia (2,500M

33 Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin’s Publishing Group, 2018), <https://books.google.com/books?id=pn4pDwAAQBAJ>.

34 Florian Meier, “TWikiL – The Twitter Wikipedia Link Dataset,” (2022). ArXiv:2201.05848; internal references omitted.

35 Simone Paolo Ponzetto and Michael Strube, “Extracting world and linguistic knowledge from Wikipedia,” 2009, 7, doi:10.1145/1620950.1620954.

words).³⁶ We could cite dozens of other machine learning models trained using Wikipedia as a source.

7. Information Quality

The problem of bias also raises the question of information quality. Wikipedia has faced the charge of being a low-quality source of information from the beginning. In 2006, Elizabeth Svoboda published a balanced overview of Wikipedia in the pages of *IEEE Spectrum*. “Many publishers and academics...have criticized the Wikipedia model on the grounds that it generates the informational equivalent of sludge,” she reported. “The lack of formal gatekeeping procedures, they say, ensures that the lowest common denominator will prevail—and since no experts or editors are hired to vet articles, no clear standards exist for accuracy or writing quality.”³⁷ As a source, she cited Robert McHenry, who had previously been the editor-in-chief at *Encyclopedia Britannica*. An analysis published in *Nature* the previous year had already undermined McHenry’s claim that Wikipedia’s “content and accuracy will tend toward the mediocre.”³⁸ Jim Giles, who reported on the results of the *Nature* study, noted that “the difference in accuracy [between the two encyclopedias] was not particularly great.”³⁹ In the years since that study, the major difference between Wikipedia and *Encyclopedia Britannica* is the difference in their scale. At present, the English-language edition of Wikipedia contains 6.5 million articles, while the *Encyclopedia Britannica* contains 120,000.⁴⁰ The dynamic and evolving nature of Wikipedia means that volunteers will also write and edit articles about newly emerging topics of public interest, from COVID-19 to the War in Ukraine, long before these topics receive coverage in traditional reference sources. To be fair, Wikipedia editors’ acknowledgment of women’s professional accomplishments may lag behind the recognition of their peers and the public. In recent years, for instance, Wikipedia has come under

36 Jacob Devlin et al., “BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding,” *CoRR* abs/1810.04805 (2018): 4175. arXiv:1810.04805.

37 E. Svoboda, “One-click content, no guarantees [online encyclopedia reliability],” *IEEE Spectrum* 43, no. 5 (2006): 65, doi:10.1109/mspec.2006.1628827.

38 *Ibid.*

39 Jim Giles, “Internet encyclopaedias go head to head,” *Nature* 438, no. 7070 (2005): 900, doi:10.1038/438900a.

40 See https://en.wikipedia.org/wiki/Wikipedia:Largest_encyclopedia (April 30, 2025).

criticism for not having articles on women (or worse, having deleted them) who went on to win major scientific prizes.⁴¹

Concerns about information quality extend beyond Wikipedia. What is the quality of the data in Wikidata? Given the nearly 100 million items (as of February 2022), there is no straightforward way to survey all items and individual claims to check for correctness and accuracy. Researchers have developed a number of metrics to assess the quality of Wikidata, however. In “A Study of the Quality of Wikidata,” Kartik Shenoy et al. outlined several methods of assaying data quality: measuring how many times volunteers on Wikidata (or the bots they create) update items, counting the number of so-called “deprecated statements,” that is, claims that were once considered valid but have been marked as no longer valid, and violations of “property constraints,” in other words, claims that break the formal rules for what they purport to describe.⁴²

In theory, nearly every claim on Wikidata should be accompanied by a warrant, that is, a reference to an authority that provides the epistemic justification for making a claim. Or, in less fancy terms, you can footnote any statement on Wikidata. This feature distinguishes Wikidata from other data sources, which typically provide data without context. In “Familiar Wikidata: The Case for Building a Data Source We Can Trust,” Constance Crompton et al. foreground the significance of these references: “While not all assertions have citations, and perhaps not all assertions come from sources scholars are ready to trust, the Wikidata platform, Wikibase, is designed to cite the sources of assertions in a way that supports trustworthiness.”⁴³ In practice, editors frequently leave out references, making it impossible to trace the provenance of their claims.

It is possible to place a constraint on properties to make references mandatory. The “property constraint” (P2302) property provides a mechanism for upholding data integrity and flagging problematic data. By applying this constraint when defining a property, editors can require that statements using that property exclude particular values. For example, the “citation needed constraint” (Q54554025) requires a property to have at least one reference. This constraint is applied in a variety of contexts, ranging from

41 See <https://www.marketplace.org/shows/marketplace-tech/why-its-so-hard-for-biographies-about-women-to-stay-on-wikipedia/> (April 30, 2025).

42 Kartik Shenoy et al., “A Study of the Quality of Wikidata,” *Journal of Web Semantics* 72 (2022): 2, doi:10.1016/j.websem.2021.100679.

43 Constance Crompton et al., “Familiar Wikidata: The Case for Building a Data Source We Can Trust,” *Pop! Public. Open. Participatory* 2 (2020), doi:10.54590/pop.2020.002.

properties such as “personal pronoun” (P6553) to “birthday” (P3150). Another property (P2316) allows for specifying two levels of constraint: “suggestion constraint” and “mandatory constraint.” If you add a statement without a reference when the property requires a “suggestion constraint,” you will receive a warning that a reference is necessary. If you try the same when the property stipulates a “mandatory constraint,” you will not be able to add the statement at all.

Property P21 (“sex or gender”) is a “property that may violate privacy” (Q44601380). The description of this item indicates that “when this property is used with items of living people it may violate privacy; statements should generally not be supplied unless they can be considered widespread public knowledge or openly supplied by the individual themselves” (Q44601380). You might imagine that adding a statement about a living person’s “sex or gender” to Wikidata would mandate adding a reference. However, the imposition of a constraint on this property has provoked significant debate among editors.⁴⁴ While the arguments are manifold, the bottom line is a trade-off between efficiency and sensitivity. On the one hand, gender is a central category for human experience and, as such, is hard to leave out without distorting biographical description. On the other, gender is a sensitive category that should not simply be inferred from markers like first name or personal pronoun. As a participant in the discussion asks, what would constitute a proper reference for gender? The identification of a subject’s gender might ideally require self-attestation but, at present, such systems are primarily informal, and such information is typically imputed by others.

8. Notability

A key contributing factor to the existence of bias on Wikimedia projects arises from the criteria for notability. While Wikipedia may seem to have an article for every topic, Wikipedia’s editorial policies place limits on acceptable subjects for articles. If one reviews the list of articles proposed for deletion, non-notability frequently appears as a motivation for the removal of articles. For academics, the bar of notability may come across as higher than anticipated. A full professor at a research university who has

⁴⁴ See, e.g., https://www.wikidata.org/wiki/Property_talk:P21/Archive_1#Gender_is_central_for_forming_grammatical_sentences_and_as_such_needed (April 30, 2025).

published peer-reviewed articles and monographs would not necessarily be considered notable from the standpoint of Wikipedia's editors; the level of references in secondary literature, especially in the press, proves more crucial to determining notability. But notability and coverage in secondary literature are not analytically equivalent. "Notable people cannot be documented," asserts Jackie Koerner, "because nothing about them exists in an acceptable published format."⁴⁵

Why do Wikipedians uphold such a high bar for notability when Wikipedia is a digital encyclopedia that, in theory, could contain articles about nearly any topic? In fact, a debate exists on Wikipedia between the so-called 'inclusionists' and 'deletionists'.⁴⁶ Broadly speaking, inclusionists advocate for a lower bar to notability on Wikipedia, favoring the retention of articles of apparently marginal significance in order to expand the encyclopedia's coverage of less known and even obscure topics. The deletionists, by contrast, contend that the barrier should remain high to prevent the encyclopedia from being populated with low-quality articles that may, among other problems, be self-promotion or spread misinformation. The paucity of active editors and administrators on Wikipedia suggests the need to impose limits on the number of articles to avoid spreading editorial attention too thin. Contributors to Wikipedia already find, to their frustration, that the "article for submission" process, which peer-reviews articles prior to their publication, may take longer than expected. That said, a degree of arbitrariness pervades the review process for new articles on Wikipedia. If an inclusionist reviews an article, it may be accepted with a suggestion that additional references are required; by contrast, a deletionist may nominate the article for deletion. Depending on who reviews their initial articles, new editors may find their contributions immediately rejected, an experience that then shapes their perspective on Wikipedia. The battles over deletions of biographical articles about women have also received attention from the scientific trade press.⁴⁷

Is the threshold for notability higher for women than men on Wikipedia? This question is difficult to answer directly because the criteria for notability are debatable and, to a degree, circular. In a groundbreaking

45 Koerner, "Wikipedia Has a Bias Problem," 318.

46 See https://en.wikipedia.org/wiki/Deletionism_and_inclusionism_in_Wikipedia (April 30, 2025).

47 See <https://www.chemistryworld.com/news/female-scientists-pages-keep-disappearing-from-wikipedia-whats-going-on/3010664.article> (April 30, 2025).

study titled “Women Through the Glass Ceiling: Gender Asymmetries in Wikipedia,” Claudia Wagner et al. developed a methodology that relies on internal and external proxies for judging notability.⁴⁸ As an external proxy, they rely on the ranking of a person in Google Trends. As an internal proxy, they count how many language editions feature an article about that person. Given these proxies for notability, are women more or less notable than men on Wikipedia? If we look at articles about contemporary women (born in the 20th century or later), it turns out that women are 13 % *more* notable than men on average.⁴⁹ “These findings suggest the existence of a subtle glass-ceiling effect that makes it more difficult for women to be included in Wikipedia than for men.”⁵⁰

Wagner et al. contend that Wikipedia should adopt policies to address this notability bias. “One way to mitigate the glass-ceiling effect is by relaxing notability guidelines for women, in order to include women who are locally notable, and for whom secondary sources might be hard to find,” they argue. “We acknowledge that this is not easy, because relaxing notability guidelines can open the door for original research, which is not allowed in Wikipedia.”⁵¹ As we shall see below, the *Women in Religion* project has forged a path around this conundrum by fostering original research about significant women in religion *off wiki* to establish their notability *on wiki*, so to speak.

The advent of Wikidata has provided another roundabout way to address the gender imbalance in Wikipedia. The threshold for notability in Wikidata is significantly lower than on English-language Wikipedia. The editors of Wikidata propose a threefold test for notability.⁵² First, an item is notable if it points to an entry on another Wikipedia project (though certain exceptions apply). Second, an item is notable if it denotes a “clearly identifiable conceptual or material entity” that “can be described using serious and publicly available references.” Third, an item is notable if it satisfies a “structural need” on Wikidata. For instance, an academic who has published an article in a peer-reviewed journal would merit an item on Wikidata. If we think of Wikidata as a graph of items connected by

48 Claudia Wagner et al., “Women through the glass ceiling: gender asymmetries in Wikipedia,” *EPJ Data Science* 5, no. 1 (2016): 4–5, doi:10.1140/epjds/s13688-016-0066-4.

49 *Ibid.*, 10.

50 *Ibid.*, 20.

51 *Ibid.*

52 See <https://www.wikidata.org/wiki/Wikidata:Notability> (April 30, 2025).

properties, we discover that significant nodes are missing. If, for example, authors are represented only in the form of strings and not as items, tracing the connections between their articles, journals, co-authors, institutions, etc., becomes difficult. By creating items for those missing authors, editors improve the quality of Wikidata and reveal the shape of research networks. The WikiCite movement⁵³ aspires to create a complete graph of the scholarly publishing ecosystem within Wikidata. If the WikiCite project succeeds even partially, data from the scholarly graph should help to establish more standardized measurements for significance, at least for academics, mitigating bias and arbitrary standards for notability.

9. Activism

If bias exists in Wikipedia, how then to combat it? As Karl Marx famously remarked, “[t]he philosophers have only interpreted the world, in various ways; the point is to change it.”⁵⁴ After a pattern of prejudice has been established, how should it be addressed and corrected? As a volunteer project driven by the interests of its editors, correcting patterns of knowledge inequality requires the participation of editors. “Why does it matter who contributes to Wikipedia?” asks Marit Hinnosaar. “Who contributes to its production matters if a gender gap among contributors leads to biased content.”⁵⁵ In this section, we look at how the *Women in Religion* movement is challenging the problem of gender bias by recruiting and sustaining a community of editors on Wikipedia, Wikidata, and Wikimedia Commons.

We noted above the distinction between Wikipedia’s skewed representation of the world and existing biases in the world. On the one hand, the lack of an article about a celebrated woman artist points to a distortion in Wikipedia’s coverage of art history. The failure of critics to recognize the contributions of a creative woman artist, by contrast, suggests a bias in the art world. While distinct, the forms of bias tend to interact and become self-reinforcing. For example, if the press does not cover a female artist’s

53 See <http://wikicite.org/> (April 30, 2025).

54 Karl Marx et al., *The Revolutionary Philosophy of Marxism: Selected Writings on Dialectical Materialism* (Marxist Books, 2018), <https://books.google.com/books?id=YnJ9DwAAQBAJ>.

55 Marit Hinnosaar, “Gender inequality in new media: Evidence from Wikipedia,” *Journal of Economic Behavior & Organization* 163 (2019): 263, doi:10.1016/j.jebo.2019.04.020.

exhibitions, then editors will not find secondary sources to draw from when seeking to compose an article about her and, as a consequence, may find it difficult to establish her notability in the eyes of fellow editors. On the flip side, editors of Wikipedia may evince a pattern of disinterest in works by women artists in general, failing to watch out for blurbs, reviews, and profiles that signal the rise of new voices among women artists. In that case, Wikipedia's editors are the source of the problem. A way to address this imbalance is to raise interest in biographies of women artists among existing editors while recruiting new editors with existing interest in women artists.

The worst scenario is a negative feedback loop in which Wikipedians do not write about significant women and in which journalists, who turn to Wikipedia and also infoboxes on Google and other search engines that draw data from Wikidata for basic information about their subjects, make assumptions from their absence about the (non-)notability of those women as prospective subjects for press coverage. This anti-pattern is the subtle opposite of the 'circular reference'⁵⁶ that short-circuits fact-checking. In such cases, bias among the press reinforces bias among editors and vice versa. "A woman might be noteworthy but, without secondary sources to back that up, she is not notable by Wikipedia standards," explains Colleen Hartung. "In this way, the ubiquitous gender bias in the production of knowledge, generally noted in the production of trade books, academia, and more, is reproduced as the apparent gender bias on Wikipedia."⁵⁷ Given this kind of negative feedback loop, how should the circuit be broken? Organizing a social movement, a community of practice, provides a way of critiquing and reforming biased editorial practices.

10. Communities of Practice

To the average user, Wikipedia appears to be an entirely online project. If users move beyond reading the articles to examining the edit history and talk pages, they will see how articles have been written, sometimes linearly, sometimes dialectically, as editors discuss and at times disagree about how

56 See https://en.wikipedia.org/wiki/Circular_reporting#Examples_on_Wikipedia (April 30, 2025).

57 Colleen D. Hartung, *Claiming Notability for Women Activists in Religion*, ATLA Open Access (Atla Open Press, 2020), ix, <https://books.google.com/books?id=rIH-zQEACAAJ>.

to organize and present their content. The social dynamics of Wikipedia extend beyond talk pages, however. Editors also join projects on Wikipedia, which brings together people with shared interests to set standards for writing articles in their areas of topic concern, to set priorities for writing new articles and revising existing ones, and to assist one another as well as newcomers in learning to contribute effectively to Wikipedia.

In *Should You Believe Wikipedia? Online Communities and the Construction of Knowledge*, Amy S. Bruckman highlights the significance of communities of practice to the establishment and growth of online communities.⁵⁸ Bruckman describes how these communities serve to introduce newcomers to their activities, training them in the values, principles, and techniques that sustain them. She analogizes these pedagogical systems to a traditional apprenticeship, where the uninitiated perform straightforward tasks while also enjoying the opportunity to observe skilled practitioners at work. Drawing on Jean Lave and Etienne Wenger's concept of "legitimate peripheral participation" (LPP), Bruckman and several colleagues explored how novice editors of Wikipedia learned to become effective contributors. "Over time, our subjects were able to observe the work of others contributing to the community, and gradually take on more complex tasks themselves," they write. "As they moved from the periphery to the center of the community, they began to use more sophisticated editing tools, and to see Wikipedia as a community and not just a collection of articles. In other words, becoming a regular Wikipedia editor is a process of LPP in a knowledge-building community of practice!"⁵⁹ Given the size of the Wikimedia community, which spans multiple platforms across different linguistic communities, we would perhaps do better to speak of *communities of practice*. Connecting with a suitable community of practice is, arguably, essential to becoming an effective contributor to Wikipedia.

Communities of practice become crucially important when seeking to challenge existing editorial policy on Wikipedia. "Over one third of the explanations blamed the gender gap on Wikipedia's editing culture, technology, and rules," remark Eckhart and Steiner in *(Re)triggering Backlash*. "Commentators described interactions among Wikipedians (those most

58 Amy S Bruckman, "Should You Believe Wikipedia?" (Cambridge: Cambridge University Press, 2022) doi:10.1017/9781108780704.011.

59 *Ibid.*, 96.

active on the site) as alienating.”⁶⁰ The complexity of the rules proves challenging for new editors to negotiate, and some prospective editors walk away from editing when they receive critical comments on their talk pages about infractions: “Commentators—both women and men—usually tied their assessments to personal experience: They wanted to contribute but gave up in the face of hostility or attacks”.⁶¹ Having a community of practice to draw on when encountering criticism provides a sense of solidarity and shared purpose, allowing newcomers to contextualize and, in some cases, to discount negative feedback.

There are many communities of practice on Wikipedia that are seeking to redress the gender imbalance. Here we can only reference several of the leading initiatives. The *WikiWomen’s User Group* achieved recognition in July 2015. The goals of the group are twofold: to provide a collaborative space for women editors and to redress gender gaps on Wikimedia projects.⁶² The comparative dearth of articles about women on English-language Wikipedia and other language editions also led to the founding of the Women in Red WikiProject in 2015.⁶³ The title of the project is a wry play on words, alluding to the fact that intra-wiki links, when no article exists at that address, are rendered in a red font. When an article is created, a red link turns blue. In September 2016, Art+Feminism also gained recognition as a Wikimedia User Group.⁶⁴ Art+Feminism describes itself as “an international campaign to improve coverage of cis, and trans women, gender and the arts on Wikipedia through organizing in-person training and editing events.”⁶⁵ *Whose Knowledge?* was also recognized as a user group in September 2016 “to center the knowledge of marginalized communities,” including the contributions of women, particularly women from the majority world.⁶⁶ Among other activities, *Whose Knowledge?* Started an annual campaign in 2018 called “VisibleWikiWomen” to provide

60 Eckert and Steiner, “(Re)triggering Backlash: Responses to News About Wikipedia’s Gender Gap,” 294.

61 Ibid.

62 See https://meta.wikimedia.org/wiki/WikiWomen%27s_User_Group (April 30, 2025).

63 See https://en.wikipedia.org/wiki/Women_in_Red (April 30, 2025).

64 See https://meta.wikimedia.org/wiki/Affiliations_Committee/Resolutions/Recogniti_on_Art%2BFeminism (April 30, 2025).

65 See https://meta.wikimedia.org/wiki/Art%2BFeminism_User_Group (April 30, 2025).

66 See <https://whoseknowledge.org/about-us/> (April 30, 2025).

photographs of women and nonbinary people on Wikimedia Commons to accompany articles on Wikipedia.⁶⁷ To round out this quick list, the *Wiki Loves Women* wiki project⁶⁸ focuses on the description of notable women in African countries. *Wiki Loves Women* works across multiple Wikimedia projects, including Wikipedia, Wikimedia Commons, and Wikidata, but also Wikiquote, where its “SheSaid” campaign adds notable quotations by women.⁶⁹ This brief survey cannot do justice to the activities of these groups and other movements that seek to rectify the gender imbalances on Wikipedia, but it illustrates how many groups have formed to tackle the challenge.

11. *Women in Religion*

The *Women in Religion* WikiProject came into existence from analogous concerns that women religious leaders were underrepresented on Wikipedia. The origins of *Women in Religion* date back to an editathon held during the Parliament of the World’s Religions meeting in Toronto in 2018. That editathon, co-sponsored by the Women’s Task Force of the Parliament of World Religions and the Women’s Caucus of the American Academy of Religion and the Society of Biblical Literature, aimed to add “1,000 biographies of women religious and spiritual leaders to Wikipedia to bring attention to women’s religious and spiritual leadership world-wide.”⁷⁰ In its early stages, the initiative received crucial support from *Women in Red* and, starting in 2019, from Atla (the American Theological Library Association). Beyond hosting editathons, members of the project organized panel presentations on women’s religious biographies under the aegis of the Women’s Caucus at annual meetings of the American Academy of Religion and Society of Biblical Literature. Members also presented their work to fellow Wikipedians during conferences such as WikiConference North America in 2019 and Wikimania in 2021.

The group flourished during the pandemic due to the indefatigable leadership of its chair, Colleen Hartung, and an international group of enthusiastic participants. The group holds monthly online editathons as well as

67 See <https://whoseknowledge.org/visiblewikiwomen-2022/> (April 30, 2025).

68 See <https://www.wikiloveswomen.org/> (April 30, 2025).

69 See https://meta.wikimedia.org/wiki/Wiki_Loves_Women/SheSaid (April 30, 2025).

70 See <https://parliamentofreligions.org/program-areas/womens-dignity> (April 30, 2025).

organizational meetings. At present, members of the *Women in Religion* project are applying for user group status with the Wikimedia Foundation. In response to feedback, the group has voted to change its name from *1000 Women in Religion* to *Women in Religion*, an anodyne but arguably more accurate appellation. By becoming a user group, members hope to develop a closer relationship with allied movements and to gain greater visibility among Wikipedians for their editorial efforts. The transition to a user group also signals its maturation as a movement, growing beyond its origins in the Parliament of World Religions and the Women's Caucus of the AAR/SBL.

11.1 Lists

At the beginning of the project, members of the *Women in Religion* project assembled lists of women to foster the creation of new articles. "The heart of the 1000 Women in Religion Wikipedia Project, and the inspiration for the Women in Religion series," explains Colleen Hartung, "is a list of women important to the world's religious and wisdom traditions who should be on Wikipedia but are not."⁷¹ The list of prospective articles originated from multiple sources. An initial list was developed from the women who had spoken at one of the meetings of the Parliament of World Religions. The theory is that women who spoke at the international event must *prima facie* count as notable. Another list was put together by members of the AAR-SBL Women's Caucus and Atla. At first, these lists were maintained in the form of Excel spreadsheets. Project participants then migrated these lists to Wikipedia. A so-called "crowd-sourced list" was then developed under the aegis of the *Women in Red* project. The open-ended quality of these lists underscores the non-hierarchical nature of the project, but also sometimes produces puzzling results. Does Beyoncé belong on the list alongside Audre Lorde and Dorothy Day?

11.2 Original Research

The *Women in Religion* project also puts significant effort into original research. As a Wikipedia movement, why would it undertake this kind

71 Colleen D. Hartung, *Challenging Bias Against Women Academics in Religion*, Women in Religion (American Theological Library Association, 2021), vi, <https://books.google.com/books?id=ZTW4zgEACAAJ>.

of scholarship? After all, a guiding principle of Wikipedia is “no original research.” As an encyclopedia, Wikipedia is a tertiary source. That is, editors gather information from reliable secondary sources to write balanced and succinct article about notable topics. In principle, Wikipedia articles should not be based on primary sources, whether that means analysis of scientific data or archival research into historical documents. Editors should instead gather information from secondary sources, ranging from peer-reviewed articles to books to news reporting. If insufficient secondary source material exists, then a topic is generally considered not sufficiently notable to warrant a Wikipedia article.

A problem with the injunction against original research is that it reifies existing biases in public discourse. If women religious leaders have historically received less public attention than men, there will be fewer secondary sources to draw on for the purposes of creating a Wikipedia article. How then to overcome this historical bias when seeking to address information equity?

The leaders of the *Women in Religion* project came up with the brilliant idea of starting a series of academic presentations and publications about notable women in religion. The goal of this effort is to provide reputable secondary sources for eventual Wikipedia articles. “Documenting their noteworthy accomplishments by writing their biographies addresses the deficit of secondary sources about their lives,” writes Colleen Hartung, general editor of the publication series. “This allows us to claim their notability and, in general, creates a more inclusive and equitable understanding of notability.”⁷² The project has published two edited volumes at the time of writing—*Claiming Notability for Women Activists in Religion*⁷³ and *Challenging Bias against Women Academics in Religion*⁷⁴—with more volumes planned to follow. To maintain editorial independence, the authors of the articles in the volume do not coordinate directly with the editors who write the articles on Wikipedia. Rather, the authors provide Wikipedians with the sources they need for establishing and writing articles, listing relevant secondary sources in their biographies. In many cases, contributors to these volumes will also conduct oral histories of their subjects; while the oral histories themselves cannot be cited on Wikipedia, the biographical details that authors publish from those sources may be.

72 Hartung, *Claiming Notability for Women Activists in Religion*, vi.

73 Hartung, *Claiming Notability for Women Activists in Religion*.

74 Hartung, *Challenging Bias Against Women Academics in Religion*.

11.3 Critiquing and Seeking Notability

As noted above, standards for notability are not culturally or historically neutral. As an activist movement, *Women in Religion* engages in a feministic critique of Wikipedia's standards for notability. As a project that operates within Wikipedia's editorial ambit, participants also seek to write articles that will withstand the critical scrutiny of editors, no matter what their ideological alignment along the spectrum of inclusionism and deletionism. The push-and-pull of critique and conformity makes the *Women in Religion* project both pragmatic and powerful.

Among the project's "strategic ways for claiming notability," Colleen Hartung highlights four stratagems.⁷⁵ The first is to satisfy and surpass Wikipedia's standards for notability. While this goal seemingly sacrifices the feminist critique of those standards by capitulating to existing criteria for notability, editors write to make their articles stick. At the same time, they subtly subvert criteria for notability, forcing change by fulfilling the letter of notability while expanding its spirit. To pull off that goal, editors must scour the written record for accepted sources. Hartung terms this search "uncovering textual erasure," a second stratagem for recovering women's effaced contributions to the cultural record.⁷⁶ Editors also expand the meaning of notability by intertwining the personal and the professional. "In general, bias on Wikipedia favors global notoriety in a public and professional sphere, and these biases work against women's inclusion," remarks Hartung. "Leveraging the concept of notability includes the work of writing biographies that challenge and stretch the boundaries between the local and the global, between the private and the public, and between the domestic and the professional."⁷⁷ Connecting the personal with the public reflects the spirit of second-wave feminism. A danger is reinscribing the existing bias toward the personal in women's biographies on Wikipedia, so editors carry out this strategy by deftly avoiding dwelling on the merely quotidian; the point is to break down barriers between private and public spheres. Hartung's fourth stratagem is to highlight collaborative work alongside individual achievement. By emphasizing the significance of social movements, she and her co-authors also make the implicit case for the

75 Hartung, *Claiming Notability for Women Activists in Religion*, vii.

76 *Ibid.*, xiv.

77 *Ibid.*, xviii.

notability of the women who pulled them off, frequently overcoming nearly impossible odds to accomplish social and political change.

11.4 Information Asymmetries

A challenge when addressing one form of bias is inadvertently fomenting another. Biases tend to emerge “whack-a-mole” style, popping up unexpectedly while editors’ energies are directed elsewhere. A good example of this phenomenon is the imbalances that arose between language editions. In “Information Asymmetry in Wikipedia Across Different Languages: A Statistical Analysis,” Dwaipayan Roy et al. study the conceptual gaps between Wikipedia editions.⁷⁸ They compare eight language editions using a variety of measures, including article length, coverage of primary topics, the number of supporting references, and the quality of the infoboxes. This study leads them to conclude that “a significant information gap exists across different Wikipedia editions.”⁷⁹ Notably, these asymmetries do not all point in the same direction. “Our analysis revealed that despite being much smaller than the English Wikipedia,” the authors write, “almost 50 % of the articles present in the non-English editions did not have a corresponding article in the English edition.”⁸⁰ This pattern makes sense intuitively when one considers the differences in contextual knowledge about topics. For example, it is not surprising to find that a topic like the ‘Grand Inquisitor’ from Fyodor Dostoevsky’s *The Brothers Karamazov* receives more attention in the Russian-language edition of Wikipedia than in the English-language edition.⁸¹

A similar imbalance manifests itself within the *Women in Religion* project. While the project is international in scope, the majority of the members come from English-speaking countries. The focus of the editathons the group sponsors is naturally on creating articles on the English-language Wikipedia. In this respect, the members of the project likely do not differ substantially from other international interest groups, which tend

78 Dwaipayan Roy, Sumit Bhatia, and Prateek Jain, “Information asymmetry in Wikipedia across different languages: A statistical analysis,” *Journal of the Association for Information Science and Technology* 73, no. 3 (2022): 347–61, doi:10.1002/asi.24553.

79 Ibid., 359.

80 Ibid.

81 Ibid., 352.

to privilege English, and by extension, the English-language Wikipedia. *Australian Women in Religion*, an offshoot of the project led by Kerrie Burn, library manager at the University of Divinity, has created many articles on Wikipedia and items on Wikidata about notable religious women from Australia and New Zealand. In a different respect, the group does demonstrate a significant asymmetry by focusing more heavily on academics than activists and other movement intellectuals. As Hartung acknowledges, “Women in academia should be low-hanging fruit in relation to Wikipedia guidelines.”⁸² The movement may have started by focusing primarily on academics, but its ambitions are much broader than women in the religious academy.

11.5 Sensitive Data

As a community of scholar-practitioners, participants in the *Women in Religion* project frequently have or develop first-hand acquaintance with sensitive information about significant women religious leaders. In some cases, this information comes in the form of tacit knowledge about gender identity, preferred personal pronouns, or religious affiliation. In other cases, participants have conducted oral histories that reveal these characteristics. How should participants make use of these data?

The answer, of course, depends on the context. The authors of the chapters in the two volumes produced to date have conducted original research on their subjects and synthesized findings from secondary literature. The authors thus transmit this sensitive information in a way that the editors of Wikipedia articles can readily cite and rely on. As for the lists of notable women in religion, the situation becomes murkier. In many cases, editors who added subjects to those lists have also contributed details about their gender identity and denominations on the basis of their first-hand knowledge. From the standpoint of Wikipedia, having recourse to personal knowledge might be regarded either as a conflict of interest or a form of original research. But, as Hartung remarks, reliance on personal knowledge is not inherently problematic from a feminist perspective. “People close to the subject and the subject themselves can provide a needed, embodied point of view on aspects of a subject’s noteworthy character.”⁸³ As we have

82 Hartung, *Challenging Bias Against Women Academics in Religion*, ii.

83 *Ibid.*, x.

noted in our discussion about biographical entries on Wikidata, it is not yet feasible to enforce a strict policy about sourcing sensitive information. Drawing on personal acquaintance seems a better interim solution than making inferences from proper names or other personal characteristics. That said, Hartung notes in a personal communication to the author that “We have been very careful about outing someone as LGBTQI. In the biographies series, if it is not clear that this information was provided in one of the author’s interviews or clear from secondary sources that the subject is out publicly, we have insisted on further interview confirmation and permission.”

What is the telos of this interim solution? Should the goal be to provide published sources for this data? If so, what becomes of the community knowledge contributed from personal acquaintance? Presumably, it would have to be otherwise substantiated or removed from the record. A parallel exists from the early days of Wikipedia when unreferenced claims were benignly tolerated. As strictures against unsubstantiated claims grew, editors would add “Citation Needed” tags. Campaigns such as “iLiblRef” have succeeded in motivating information professionals to join the effort in tracking down proper sources for citations. Nowadays, editors may remove claims if references have not been provided after many years of noting their absence. From the standpoint of what Matthew A. Vetter terms the Enlightenment encyclopedia,⁸⁴ the increasing enforcement of sourcing has improved information quality. But has this increased stringency caused Wikipedia to lose a form of implicit communal knowledge? This is a question we should bear in mind as we consider the future of movements like *Women in Religion*. “The dominance of print culture plays a significant role in the marginalization of indigenous knowledge cultures,” writes Matthew A. Vetter, “especially when their knowledge is stored and transmitted orally.”⁸⁵

A similar point holds for tacit information transmitted within communities of practice. The personal ties that bind editors and subjects within communities of practice may not prove as lasting as the oral traditions sustained by indigenous communities. Personal knowledge fades as participants pass from the scene and is not typically reproduced by subsequent generations.

84 Matthew A. Vetter, “Possible Enlightenment: Wikipedia’s Encyclopedic Promise and Epistemological Failure,” in *Wikipedia @ 20: Stories of an Incomplete Revolution* (The MIT Press, 2020), doi:10.7551/mitpress/12366.003.0024.

85 *Ibid.*, 288; see Hartung, *Challenging Bias Against Women Academics in Religion*, xix.

The present thus seems like the right time to develop standards for sourcing sensitive information to communities of practice lest the movement's initial success of documenting these notable religious women become undone.

Works Cited

- Bruckman, Amy S. "Should You Believe Wikipedia?" Cambridge University Press, 2022. doi:10.1017/9781108780704.011.
- Crompton, Constance, Lori Antranikian, Ruth Truong, and Paige Maskell. "Familiar Wikidata: The Case for Building a Data Source We Can Trust." *Pop! Public. Open. Participatory* 2 (2020). doi:10.54590/pop.2020.002.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *CoRR* abs/1810.04805 (2018). arXiv:1810.04805.
- Eckert, Stine, and Linda Steiner. "(Re)triggering Backlash: Responses to News About Wikipedia's Gender Gap." *Journal of Communication Inquiry* 37, no. 4 (2013): 284–303. doi:10.1177/0196859913505618.
- Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Publishing Group, 2018. <https://books.google.com/books?id=pn4pDwAAQBAJ>.
- Giles, Jim. "Internet encyclopaedias go head to head." *Nature* 438, no. 7070 (2005): 900–901. doi:10.1038/438900a.
- Hartung, Colleen D. *Challenging Bias Against Women Academics in Religion*. Women in Religion. American Theological Library Association, 2021. <https://books.google.com/books?id=ZTW4zgEACAAJ>.
- . *Claiming Notability for Women Activists in Religion*. ATLA Open Access. Atla Open Press, 2020. <https://books.google.com/books?id=rIH-zQEACAAJ>.
- Hinnosaar, Marit. "Gender inequality in new media: Evidence from Wikipedia." *Journal of Economic Behavior & Organization* 163 (2019): 262–76. doi:10.1016/j.jebo.2019.04.020.
- Koerner, Jackie. "Wikipedia Has a Bias Problem." In *Wikipedia @ 20: Stories of an Incomplete Revolution*. The MIT Press, 2020. doi:10.7551/mitpress/12366.003.0026.
- Lih, Andrew. *The Wikipedia Revolution: How a Bunch of Nobodies Created the World's Greatest Encyclopedia*. Hachette Books, 2009. <https://books.google.com/books?id=AWuZAAAAQBAJ>.
- Marx, Karl, J. Peterson, F. Engels, A. Woods, V. I. Lenin, L. Trotsky, and M. Books. *The Revolutionary Philosophy of Marxism: Selected Writings on Dialectical Materialism*. Marxist Books, 2018. <https://books.google.com/books?id=YnJ9DwAAQBAJ>.
- Meier, Florian. "TWikiL – The Twitter Wikipedia Link Dataset," (2022). ArXiv:2201.05848.

- Menking, Amanda, Vaibhavi Rangarajan, and Michael Gilbert. "Sharing small pieces of the world: Increasing and broadening participation in Wikimedia Commons." *Proceedings of the 14th International Symposium on Open Collaboration*, 2018, 13. doi:10.1145/3233391.3233537.
- Miquel-Ribé, Marc, and David Laniado. "Wikipedia Culture Gap: Quantifying Content Imbalances Across 40 Language Editions." *Frontiers in Physics* 6 (2018): 54. doi:10.3389/fphy.2018.00054.
- Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, 2018. <https://books.google.com/books?id=-ThDDwAAQBAJ>.
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016. <https://books.google.com/books?id=cbwvDwAAQBAJ>.
- Ponzetto, Simone Paolo, and Michael Strube. "Extracting world and linguistic knowledge from Wikipedia," 2009, 7–8. doi:10.1145/1620950.1620954.
- Rosenzweig, Roy. "Can History Be Open Source? Wikipedia and the Future of the Past." *Journal of American History* 93, no. 1 (2006): 117–46. doi:10.2307/4486062.
- Roy, Dwaipayan, Sumit Bhatia, and Prateek Jain. "Information asymmetry in Wikipedia across different languages: A statistical analysis." *Journal of the Association for Information Science and Technology* 73, no. 3 (2022): 347–61. doi:10.1002/asi.24553.
- Shaik, Zaina, Filip Ilievski, and Fred Morstatter. "Analyzing Race and Citizenship Bias in Wikidata." *2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2021, 665–66. doi:10.1109/mass52906.2021.00099.
- Shenoy, Kartik, Filip Ilievski, Daniel Garijo, Daniel Schwabe, and Pedro Szekely. "A Study of the Quality of Wikidata." *Journal of Web Semantics* 72 (2022): 100679. doi:10.1016/j.websem.2021.100679.
- Svoboda, E. "One-click content, no guarantees [online encyclopedia reliability]." *IEEE Spectrum* 43, no. 5 (2006): 64–65. doi:10.1109/mspec.2006.1628827.
- Vetter, Matthew A. "Possible Enlightenments: Wikipedia's Encyclopedic Promise and Epistemological Failure." In *Wikipedia @ 20: Stories of an Incomplete Revolution*. The MIT Press, 2020. doi:10.7551/mitpress/12366.003.0024.
- Vrandečić, Denny. "The Rise of Wikidata." *IEEE Intelligent Systems* 28, no. 4 (2013): 90–95. doi:10.1109/mis.2013.119.
- Vrandečić, Denny. "Wikidata." *Proceedings of the 21st International Conference Companion on World Wide Web - WWW '12 Companion*, 2012, 1063–64. doi:10.1145/2187980.2188242.
- Vrandečić, Denny, and Markus Krötzsch. "Wikidata: a free collaborative knowledge-base." *Communications of the ACM* 57, no. 10 (2014): 78–85. doi:10.1145/2629489.
- Wagner, Claudia, Eduardo Graells-Garrido, David Garcia, and Filippo Menczer. "Women through the glass ceiling: gender asymmetries in Wikipedia." *EPJ Data Science* 5, no. 1 (2016): 5. doi:10.1140/epjds/s13688-016-0066-4.
- Whitehead, Alfred N. *Process and Reality*. Free Press, 2010. <https://books.google.com/books?id=uJDEx6rPu1QC>.