11   The Most Important Laboratory for Social Scientific and Computing Research in History

Benjamin Mako Hill and Aaron Shaw

Wikipedia's founders could not have dreamed they were creating the most important laboratory for social scientific and computing research in history, yet Wikipedia has had an enormous effect on academic research.

Twenty years ago, Wikipedia's founders could not have dreamed they were creating the most important laboratory for social scientific and computing research in history. And yet that is exactly what has happened. Wikipedia and its sister projects have launched a thriving scholarly literature. How thriving? Results from Google Scholar suggest that over six thousand scholarly publications mention Wikipedia in their title and over 1.7 million mention it somewhere in their text. For comparison, the phrase "Catholic church"—an organization with a nearly two-thousand-year head start—returns about the same number of mentions in publication titles. In under twenty years, Wikipedia has become one of the most heavily studied organizations of any kind. To the extent that Wikipedia research is a field of study, what major areas of investigation have been pursued in the field so far? What are the big discoveries? The most striking gaps? This essay addresses these questions and considers some of the most important directions Wikipedia research might take in the future.

The State of Wikimedia Research

In 2008, Mako Hill was about to start his first year as a social science graduate student at the Massachusetts Institute of Technology where he hoped to study, among other things, organizational processes that had driven Wikipedia's success. Mako felt it would behoove him to become better connected to the recent academic scholarship on Wikipedia. He was also looking for a

topic for a talk he could give at Wikipedia's annual community conference, called "Wikimania," which was going to be hosted by the Library of Alexandria in Egypt. Attempting to solve both problems at once, Mako submitted a session proposal for Wikimania suggesting that he would summarize all of the academic research about Wikipedia published in the previous year in a talk entitled "The State of Wikimedia Scholarship: 2007–2008."

Happily, the proposal was accepted. Two weeks before Wikimania, Mako did a Google Scholar search to build a list of papers he needed to review. He found himself facing nearly eight hundred publications. When Mako tried to import the papers from the search results into his bibliographic management software, Google Scholar's bot detection software banned his laptop. Presumably, no human could (or should!) read that many papers.

Mako never did read all the papers that year, but he managed to create a talk synthesizing some key themes from the previous year in research. Since then, Mako recruited Aaron Shaw to help create new versions of the talk on a yearly basis. Working together since 2008, the two authors of this chapter have collaborated on a "State of Wikimedia Scholarship" talk nearly every year. With a growing cast of collaborators, we sort through the huge pile of published papers with the term "Wikipedia" in their title or abstracts from the past year. Increasingly, we incorporate papers that analyze other communities supported by the Wikimedia Foundation. Each time around, we select five to eight themes that we think capture major tendencies or innovations in research published in the previous year. For the presentation, we summarize each theme and describe an exemplary paper (one per theme) to the Wikimania audience.

Over the first twenty-years of the project's life, Wikipedia research has connected researchers who have formed a new interdisciplinary field. We have each coordinated the program of the International Symposium on Open Collaboration (OpenSym), a conference started in 2005 as WikiSym. As part of this work, we helped coordinate papers in a track dedicated to "Wikipedia and Wikimedia research." Each year the Web Conference (formerly WWW) hosts a workshop that focuses on Wikipedia and Wikimedia research. Since 2011, volunteers have helped create a monthly "Wikimedia Research Newsletter" which is published in English Wikipedia's newsletter *The Signpost* and provides a sort of monthly version of our annual talk. The Wikimedia Foundation runs a monthly "research showcase" where
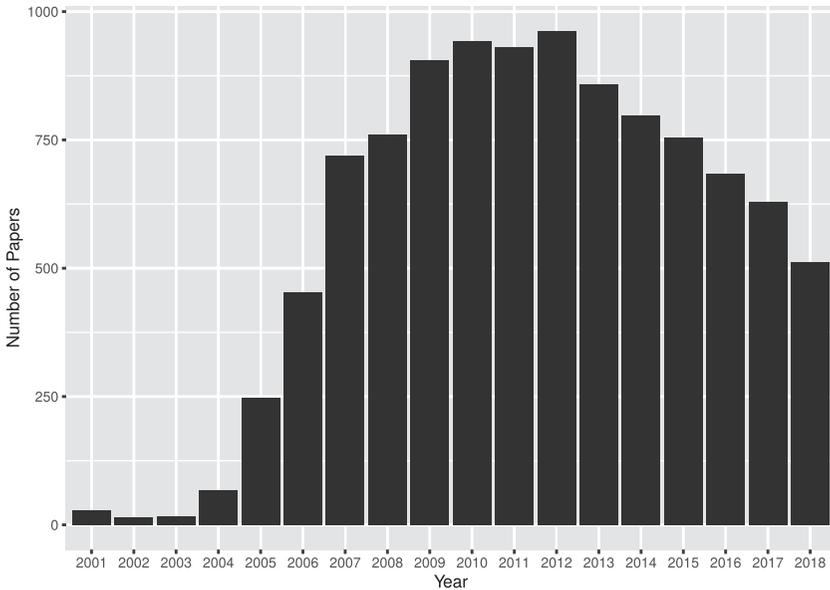
Figure 11.1

Number of items returned for Google Scholar for publications containing "Wiki-pedia" in the title by year of publication.

researchers from the around the world can present their work. There is an active mailing list for Wikimedia researchers.

As the graph in figure 11.1 suggests, these venues capture only a tiny fraction of Wikimedia research. Our attempt to characterize this body of research in this chapter draws from our experience preparing the annual Wikimania talk each year and from our experience in these other spaces. Like our Wikimania talk, this chapter remains incomplete and aims to provide a brief tour of several important themes. Others have published literature reviews of Wikipedia and Wikimedia research which make attempts to provide more comprehensive—although still limited—approaches.[1] Our experience watching Wikipedia scholarship grow and shift has led to one overarching conclusion: Wikipedia has become part of the mainstream of every social and computational research field we know of. Some areas of study, such as the analysis of human computer interaction, knowledge management, information systems, and online communication, have undergone profound shifts in the past twenty years that have been driven

by Wikipedia research. In this process, Wikipedia has acted as a shared object of study that has connected a range of disparate academic fields to each other.

Wikipedia as a Source of Data

Perhaps the most widespread and pervasive form of Wikipedia research is not research "about" Wikipedia at all, but research that uses Wikipedia as a convenient data set to study something else. This was the only theme that showed up every single year during the nine years that we presented the "State of Research" review.

In 2017, Mohamad Mehdi and a team published a systematic literature review of 132 papers that use Wikipedia as a "corpus" of human-generated text.[2] Most of these papers come from the engineering field of information retrieval (IR) where the goal is to devise approaches for calling up particular information from a database. Wikipedia is useful for a wide range of tasks in IR research because it provides a vast database of useful knowledge that is tagged with categories and metadata—but not in the typically "structured" way required by databases.

Another large group of examples comes from the field of natural language processing (NLP), which exists at the intersection of computer science and linguistics. NLP researchers design and evaluate approaches for parsing, understanding, and sometimes generating human-intelligible language. As with IR, Wikipedia presents an opportunity to NLP research because it encompasses an enormous, multilingual data set written and categorized by humans about a wide variety of topics. Wikipedia has proven invaluable as a data set for these applications because it is "natural" in the sense that humans wrote it, because it is made freely available in ways that facilitate computational analysis, and because it exists in hundreds of languages. Nearly half of the papers in Mehdi's review study a version of Wikipedia other than English, and more than a third of the papers look at more than one language edition Wikipedia.

Recently, Wikipedia has spawned a large number of "derivative" data sets and databases that extract data from Wikipedia for studying a wide variety of topics. Similarly, a large body of academic research has focused on building tools to transform data from Wikipedia and to extract specific subsets of data. One of the newest Wikimedia projects, Wikidata, extends these

benefits by creating a new layer of structured data that is collaboratively authored and edited like Wikipedia but that formally represents underlying relationships between entities that may be the topics of Wikipedia articles. As Wikipedia and Wikidata continue to grow and render ideas and language more amenable to computational processing, their value as a data set and data source to researchers is also increasing.

The Gender Gap

In 2008, the results of a large opt-in survey of Wikipedia editors suggested that upward of 80 percent of editors of Wikipedia across many language editions were male. The finding sent shockwaves through both the Wikipedia editor and research communities and was widely reported on in the press. Both the Wikimedia Foundation and Wikipedia community have responded by making "the gender gap" a major strategic priority and have poured enormous resources into addressing the disparity. Much of this work has involved research. As a result, issues related to gender have been a theme in our report on Wikipedia research nearly every year since 2012.

One series of papers have aimed to characterize the "gender gap." This work adopted better sampling methods, adjusted for bias in survey response, and in at least one case, commissioned a nationally representative sample of adults in the United States who were asked about their Wikipedia contribution behavior.[3] Some recent projects have also begun to unpack the "gap" by looking at the ways in which it emerges.[4] Although this follow-on work presented a range of different estimates of the scope of the gap in participation between male and female editors, none of the work overturned the basic conclusion that Wikipedia's editor base appears largely, if not overwhelmingly, made up of men.

Another group of studies examines different gender gaps, including gaps in content coverage. For example, research has found that women and people of color are systematically less likely than similarly notable white men to have articles.[5] Other work has shown that Wikipedia's content tends to suffer a range of gender biases and gaps as well—for example, by using terms and images that tend to reflect existing gender bias.[6]

Some work has also connected explanations of the gender gap among contributors to inequality and bias in articles. Existing Wikipedia communities may deter women and others from editing and may define and

enforce criteria for article creation in ways that differentially impact articles about or of interest to women.[7]

The work on the gender gap in Wikipedia began with a strong focus on gender inequality within Wikipedia and among Wikipedia editors. More recent work has sought to understand how Wikipedia content may reflect underlying inequalities and patterns of stratification in the world in other ways. This work has shown that, by studying gendered and other types of inequality in Wikipedia, we can learn about some of the mechanisms of social stratification more broadly.

Content Quality and Integrity

Research into content quality and integrity on Wikipedia has also been an enduring focus of Wikipedia research. In a 2005 piece that is one of the most widely discussed examples of Wikipedia research (see chapters 2 and 12), Jim Giles at *Nature* ran an informal study distributing a set of Wikipedia and *Encyclopædia Britannica* articles to experts and asking them to identify errors in each.[8] The expert coders found about the same number of errors in each group, leading to the conclusion—surprising at the time—that Wikipedia articles might be comparable with those produced by professionals and experts. The early *Nature* study has been reproduced in larger samples with results suggesting that, over time, Wikipedia typically surpasses general encyclopedias like *Britannica*.[9] Perhaps more influentially, the template of the Giles study has been repeated over and over again in various knowledge domains that include drug information, mental disorders, and otolaryngology—just to name several topics in medicine.[10]

Of course, quality itself is much more complicated and multidimensional than the sum of factual errors in a sample of articles. A number of studies have tried to assess quality in other terms. Some consider the relative neutrality of articles on contentious topics.[11] Others look for the absence of important information. Wikipedians regularly evaluate the quality of their own articles in terms of comprehensiveness, writing style, the number and reliability of references, and adherence to Wikipedia's own policies. There have been a series of attempts to adapt these types of quality measures quantitatively. This work seems to indicate that although Wikipedia is enormous, many topics are covered in ways that are superficial.[12]

Overall, this body of research has shown the quality of the material that is covered is high.

Some of the most exciting work on these issues has examined the social processes that lead to relatively higher or lower article quality. For example, although quality and viewership of articles are related, a few recent studies have measured the degree to which topics are "underproduced" relative to readers' interest.[13] Another paper shows that articles on contentious topics edited by more ideologically polarized editors tend to become higher quality than those with less diverse editor groups.[14] Other work has sought to understand how readers of Wikipedia perceive quality.[15] In an era where factual information is increasingly contested and polarized, this line of inquiry offers the promise of general insights into the means of producing and sustaining reliable, high quality public knowledge resources.

Wikipedia and Education

Early on in its ascendance, many viewed Wikipedia as a threat to educational authority and a source of dubious information. Initial research on Wikipedia in education documented the ways that students used Wikipedia and, in general, suggested that students were relying on Wikipedia heavily as a first stop for information on a given subject. For many teachers, Wikipedia's open editing policy made its content inherently problematic, if not inherently incompatible, with formal institutions of teaching and learning.

The study of Wikipedia in education has evolved enormously. In part, educators have changed their attitudes about the site, and some studies have attempted to document these shifts.[16] The focus of academic writing about the pedagogical role of Wikipedia is no longer on the questions of *if* students use Wikipedia or how to discourage them from doing so. Instead, researchers of Wikipedia in education now focus on how to engage students in contributing to Wikipedia as part of course work.

Partly, this change seems driven by the success of the Wiki Education Foundation—a spin-off of the Wikimedia Foundation that supports instructors of higher education in incorporating Wikipedia into their classes (see chapter 20). Numerous papers and book chapters now document these experiences. One example from psychology describes the way that ninety-three students in an introductory human development course helped to improve Wikipedia coverage of basic information on human development.[17]

Viewership

The large majority of research on Wikipedia has focused on its content and the social systems that produce it. But Wikipedia isn't only an enormous corpus created by millions, it is also one of the top ten most popular websites on earth—visited by billions of people each year. In 2007, the Wikimedia Foundation started publishing data that summarized what visitors to Wikipedia have looked at. This data has now led to a large body of research on the viewership of the encyclopedia.

Some work on viewership takes advantage of Wikipedia's general usefulness and uses those pages that people visit as an index of how people allocate their attention. For example, the Snowden revelations led to chilling effects where people became systematically less likely to look at certain sensitive topics.[18] Other studies have used Wikipedia viewership data to predict the prevalence of illnesses and influenza, box office revenue, election results in a number of countries, or simply to capture a zeitgeist.[19]

Scholars have also combined data on Wikipedia viewership with editing data to understand the relationship between the consumption and production of knowledge. Some early work in this area considered whether viewership related to participation in editing and content quality.[20] Others have tried to model relatively complex dynamics through which viewers become editors to help produce the encyclopedia.[21]

Organization and Governance

When Wikipedia was first founded, one of the most urgent areas of inquiry focused on the organization and governance of the project. Seminal work by Yochai Benkler, author of chapter 3, suggested that Wikipedia used technology to organize knowledge production in transformative ways. Since then, research on the organization of Wikipedia has grown steadily, often in an attempt to explain its arguably shocking success.[22]

Research has sometimes treated Wikipedia as a community of communities to investigate collaborative processes. For example, both article-level collaborations and organized editing efforts in the form of WikiProjects have attracted extensive research. Perhaps not surprisingly, WikiProjects appear to struggle with many of the same kinds of organizational challenges that affect collaborative efforts elsewhere.[23] Many studies of organization within

Wikipedia have found creative ways to document and describe otherwise familiar patterns and have sometimes revealed distinctions between more familiar organizational practices and those pursued in a large, distributed, online volunteer effort like Wikipedia.

We have been involved in some related work that challenges the "stylized facts" about Wikipedia's organization and that has suggested some of the ways that Wikipedia's mode of organization and governance may be limited.[24] We also advocated for comparative studies that look beyond Wikipedia—and English Wikipedia in particular—to draw more general understandings of the organizational processes involved.[25] Wikipedia includes hundreds of more-or-less completely distinct language communities with different experiences and with different degrees of success. For instance, several papers—ours and others'—undermine the widespread perception that Wikipedia's style of organizing does not entail hierarchies or other patterns of entrenchment among early community leaders.[26] A small number of studies have engaged in comparative work that studies Wikipedia across numerous language editions, illustrating the diversity of collaborative dynamics.[27]

As a large population of organizations, Wikipedia offers a data source of exceptional granularity. Nevertheless, scholars continue to struggle to understand how Wikipedia is like and unlike more traditional organizations. We still know little about when the experience of traditional organizations will be instructive to Wikipedia. For example, in our own work we found that an attempt to import newcomer socialization practices with a long history of success in traditional organizations seemed to have little effect on newcomer retention in Wikipedia.[28] In a related sense, we still know little about when the things we learn about organization in Wikipedia will—or will not—translate into other spaces.

Wikipedia in the World

The metaphor of a laboratory that we used in our introduction depicts Wikipedia as somehow isolated from the rest of the world. However, Wikipedia *affects* the world in important ways as well. Some exciting studies have investigated specific aspects of this relationship.

The earliest versions of this work simply documented the ways that Wikipedia became increasingly integrated into many people's everyday

lives. One striking example from 2009 described the growing rate at which legal opinion and published law relied on citations to Wikipedia to establish facts about the world in hundreds of legal opinions in the US district courts and courts of appeals.[29] Other work looks at how Wikipedia content is increasingly syndicated into other places and suggests that an enormous portion of all successful internet searches would be failures if Wikipedia did not exist.[30]

Given its prominence in search engine rankings, a group of scholars—primarily economists—have come to Wikipedia as a platform on which to run experiments *on the world.* For example, one group improved a random set of articles about small European cities and showed that tourism traffic increased relative to a control group whose articles were not improved.[31] Another study showed that improving a randomly selected set of Wikipedia articles about scientific studies tends to increase the citations to the studies mentioned in the articles and tends to shape the language that subsequent research studies use when they describe the cited work.[32]

These studies do more than show that Wikipedia is important, although they certainly do that. They provide important evidence in favor of particular theories of information diffusion, and they document the way that knowledge is created and spreads. In this way, Wikipedia provides not only a laboratory for studying social processes but acts as a key piece of laboratory equipment for studying social behavior "in the wild."

Conclusion

Insights about how the largest volunteer effort in the world have managed to produce the largest encyclopedias in history will continue to advance the frontiers of scientific knowledge. Understanding how Wikipedia and projects like it work can help us organize other parts of social life more effectively.

We conclude with an invocation to researchers to think about Wikipedia even more and in even broader ways. Wikipedia is the most influential and widely accessed free information resource on the internet as well as the most widely used information platform in human history. As such, Wikipedia merits comparisons to other epochal transformations in how humans collect, organize, store, and disseminate ideas. It deserves the scholarly attention it has received. In particular, understanding how and why communities like Wikipedia manage to mobilize vast numbers of

volunteers and sustain such high quality, large-scale information resources means looking beyond the boundaries of Wikipedia to conduct comparisons, impact evaluations, and more. That ought to keep us all busy for at least another twenty years.

Notes

1. Chitu Okoli, Mohamad Mehdi, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki, "Wikipedia in the Eyes of Its Beholders: A Systematic Review of Scholarly Research on Wikipedia Readers and Readership," *Journal of the Association for Information Science and Technology* 65, no. 12 (2014): 2381–2403, https://doi.org/10.1002/asi.23162; Mohamad Mehdi, Chitu Okoli, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki, "Excavating the Mother Lode of Human-Generated Text: A Systematic Review of Research That Uses the Wikipedia Corpus," *Information Processing & Management* 53, no. 2 (March 2017): 505–529, https://doi.org/10.1016/j.ipm.2016.07.003.

2. Mehdi et al., "Excavating the Mother Lode of Human-Generated Text."

3. Eszter Hargittai and Aaron Shaw, "Mind the Skills Gap: The Role of Internet Know-How and Gender in Differentiated Contributions to Wikipedia," *Information, Communication & Society* 18, no. 4 (2015): 424–442, https://doi.org/10.1080/1369118X.2014.957711; Benjamin Mako Hill and Aaron Shaw, "The Wikipedia Gender Gap Revisited: Characterizing Survey Response Bias with Propensity Score Estimation," *PLoS ONE* 8, no. 6 (2013): e65782, https://doi.org/10.1371/journal.pone.0065782; Aaron Shaw and Eszter Hargittai, "The Pipeline of Online Participation Inequalities: The Case of Wikipedia Editing." *Journal of Communication* 68, no. 1 (2018): 143–168, https://doi.org/10.1093/joc/jqx003.

4. For example, Shaw and Hargittai, "The Pipeline of Online Participation Inequalities."

5. For example, Joseph Reagle and Lauren Rhue, "Gender Bias in Wikipedia and Britannica," *International Journal of Communication* 5 (2011): 1138–1158, http://ijoc.org/index.php/ijoc/article/view/777; Julia Adams, Hannah Brückner, and Cambria Naslund, "Who Counts as a Notable Sociologist on Wikipedia? Gender, Race, and the 'Professor Test,'" *Socius* 5 (2019), https://doi.org/10.1177/2378023118823946.

6. Max Klein and Piotr Konieczny, "Wikipedia in the World of Global Gender Inequality Indices: What the Biography Gender Gap Is Measuring," in *OpenSym '15: Proceedings of the 11th International Symposium on Open Collaboration*, ed. D. Riehle

(New York: ACM, 2015), 16:1–16:2, https://doi.org/10.1145/2788993.2789849; Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier, “It’s a Man’s Wikipedia? Assessing Gender Inequality in an Online Encyclopedia,” in *Proceedings of the Ninth International AAAI Conference on Web and Social Media (ICWSM ’15)* (Palo Alto, CA: AAAI, 2015), 454–463, https://www.aaai.org/ocs/index.php /ICWSM/ICWSM15/paper/view/10585; Olga Zagovora, Fabian Flöck, and Claudia Wagner, “‘(Weitergeleitet von Journalistin)’: The Gendered Presentation of Professions on Wikipedia,” in *WebSci ’17: Proceedings of the 2017 ACM on Web Science Conference* (New York: ACM, 2017), 83–92, https://doi.org/10.1145/3091478 .3091488.

7. Julia B. Bear and Benjamin Collier, “Where Are the Women in Wikipedia? Understanding the Different Psychological Experiences of Men and Women in Wikipedia,” *Sex Roles* 74, nos. 5–6 (January 4, 2016): 254–265, https://doi.org/10.1007/s11199 -015-0573-y; Amanda Menking, Ingrid Erickson, and Wanda Pratt, “People Who Can Take It: How Women Wikipedians Negotiate and Navigate Safety,” in *CHI ’19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York: ACM, 2019), https://doi.org/10.1145/3290605.3300702; Heather Ford and Judy Wajcman, “‘Anyone Can Edit’, Not Everyone Does: Wikipedia’s Infrastructure and the Gender Gap,” *Social Studies of Science* 47, no. 4 (2017): 511–527, https://doi .org/10.1177/0306312717692172.

8. Jim Giles, “Internet Encyclopaedias Go Head to Head,” *Nature* 438, no. 7070 (2005): 900–901, https://doi.org/10.1038/438900a.

9. Fernando Silvério Nifrário Rodrigues, “Colaboração Em Massa Ou Amadorismo Em Massa? Um Estudo Comparativo Da Qualidade Da Informação Científica Produzida Utilizando Os Conceitos E Ferramentas Wiki” (PhD diss., Universidade de Évora, 2012), http://massamateurism.blogspot.co.uk/p/synopsis.html.

10. Kevin A. Clauson, Hyla H. Polen, Maged N. Kamel Boulos, and Joan H. Dzenowagis. “Scope, Completeness, and Accuracy of Drug Information in Wikipedia,” *The Annals of Pharmacotherapy* 42, no. 12 (December 2008): 1814–1821, https://doi .org/10.1345/aph.1L474; Thomas J. Hwang, Florence T. Bourgeois, and John D. Seeger, “Drug Safety in the Digital Age,” *New England Journal of Medicine* 370, no. 26 (2014): 2460–2462. https://doi.org/10.1056/NEJMp1401767; Jona Kräenbring et al., “Accuracy and Completeness of Drug Information in Wikipedia: A Comparison with Standard Textbooks of Pharmacology,” *PLOS ONE* 9, no. 9 (2014): e106930, https://doi.org/10.1371/journal.pone.0106930; N. J. Reavley, A. J. Mackinnon, A. J. Morgan, M. Alvarez-Jimenez, S. E. Hetrick, E. Killackey, B. Nelson, R. Purcell, M. B. H. Yap, and A. F. Jorm, “Quality of Information Sources about Mental Disorders: A Comparison of Wikipedia with Centrally Controlled Web and Printed Sources,” *Psychological Medicine* 42, no. 8 (August 2012): 1753–1762, https://doi.org/10.1017 /S003329171100287X; Peter G. Volsky, Cristina M. Baldassari, Sirisha Mushti, and Craig S. Derkay, “Quality of Internet Information in Pediatric Otolaryngology: A

Comparison of Three Most Referenced Websites," *International Journal of Pediatric Otorhinolaryngology* 76, no. 9 (September 2012): 1312–1316, https://doi.org/10.1016/j.ijporl.2012.05.026.

11. Shane Greenstein and Feng Zhu, "Is Wikipedia Biased?" *American Economic Review* 102, no. 3 (May 2012): 343–348, https://doi.org/10.1257/aer.102.3.343; Shane Greenstein and Feng Zhu, "Do Experts or Crowd-Based Models Produce More Bias? Evidence from Encyclopedia Britannica and Wikipedia," *Management Information Systems Quarterly* 42, no. 3 (2018): 945–959, https://aisel.aisnet.org/misq/vol42/iss3/14.

12. Morten Warncke-Wang, Vivek Ranjan, Loren Terveen, and Brent Hecht, "Misalignment Between Supply and Demand of Quality Content in Peer Production Communities," in *Proceedings of the Ninth International AAAI Conference on Web and Social Media (ICWSM '15)* (Palo Alto, CA: AAAI, 2015), 493–502, https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/viewFile/10591/10532.

13. Aniket Kittur and Robert E. Kraut, "Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination," in *CSCW '09: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (New York: ACM, 2008), 37–46, https://doi.org/10.1145/1460563.1460572; Warncke-Wang et al., "Misalignment Between Supply and Demand of Quality Content in Peer Production Communities."

14. Feng Shi, Misha Teplitskiy, Eamon Duede, and James A. Evans, "The Wisdom of Polarized Crowds," *Nature Human Behaviour* 3, no. 4 (April 2019): 329, https://doi.org/10.1038/s41562-019-0541-6.

15. W. Ben Towne, Aniket Kittur, Peter Kinnaird, and James Herblseb, "Your Process Is Showing: Controversy Management and Perceived Quality in Wikipedia," in *CSCW '13: Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (New York: ACM, 2013), 1059–1068, https://doi.org/10.1145/2441776.2441896.

16. Hsin-liang Chen, "The Perspectives of Higher Education Faculty on Wikipedia," *The Electronic Library* 28, no. 3 (J2010): 361–373, https://doi.org/10.1108/02640471011051954; Aline Soules, "Faculty Perception of Wikipedia in the California State University System," *New Library World* 116, no. 3/4 (2015): 213–226, https://doi.org/10.1108/NLW-08-2014-0096.

17. Christina Shane-Simpson, Elizabeth Che, and Patricia J. Brooks, "Giving Psychology Away: Implementation of Wikipedia Editing in an Introductory Human Development Course," *Psychology Learning & Teaching* 15, no. 3 (2016): 268–293, https://doi.org/10.1177/1475725716653081.

18. Jonathon Penney, "Chilling Effects: Online Surveillance and Wikipedia Use," *Berkeley Technology Law Journal* 31, no. 1 (2016): 117, https://doi.org/10.15779/Z38SS13.

19. David J. McIver and John S. Brownstein, "Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time," *PLoS Computational*

Biology 10, no. 4 (2014): e1003581, https://doi.org/10.1371/journal.pcbi.1003581; Márton Mestyán, Taha Yasseri, and János Kertész, "Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data," *PLOS ONE* 8, no. 8 (2013): e71226, https://doi.org/10.1371/journal.pone.0071226; Taha Yasseri and Jonathan Bright, "Wikipedia Traffic Data and Electoral Prediction: Towards Theoretically Informed Models," *EPJ Data Science* 5, no. 1 (2016): 22, https://doi.org/10.1140/epjds /s13688-016-0083-3; Benjamin K. Smith and Abel Gustafson, "Using Wikipedia to Predict Election Outcomes: Online Behavior as a Predictor of Voting," *Public Opinion Quarterly* 81, no. 3 (2017): 714–735, https://doi.org/10.1093/poq/nfx007; Gabriele Tolomei et al., "Twitter Anticipates Bursts of Requests for Wikipedia Articles," in *DUBMOD '13: Proceedings of the 2013 Workshop on Data-Driven User Behavioral Modelling and Mining from Social Media* (New York: ACM, 2013), 5–8, https://doi.org/10 .1145/2513577.2538768; Michela Ferron and Paolo Massa, "Beyond the Encyclopedia: Collective Memories in Wikipedia," *Memory Studies* 7, no. 1 (2014): 22–45, https:// doi.org/10.1177/1750698013490590.

20. Reid Priedhorsky, Jilin Chen, Shyong (Tony) K. Lam, Katherine Panciera, Loren Terveen, and John Riedl, "Creating, Destroying, and Restoring Value in Wikipedia," in *GROUP '07: Proceedings of the 2007 International ACM Conference on Supporting Group Work* (New York: ACM, 2007), 259–268, https://doi.org/10.1145/1316624.1316663.

21. Andreea D. Gorbatâi, "The Paradox of Novice Contributions to Collective Production: Evidence from Wikipedia" (working paper, University of California, Berkeley, February 10, 2014), https://papers.ssrn.com/abstract=1949327.

22. Yochai Benkler, *The Wealth of Networks* (New Haven, CT: Yale University Press, 2006); subsequent work includes David A. Hoffman and Salil K. Mehra, "Wikitruth Through Wikiorder," *Emory Law Journal* 59, no. 1 (2009–2010): 151–210, https:// heinonline.org/HOL/P?h=hein.journals/emlj59&i=153; Piotr Konieczny, "Governance, Organization, and Democracy on the Internet: The Iron Law and the Evolution of Wikipedia," *Sociological Forum* 24, no. 1 (2009): 162–192, https://doi.org/10.1111/j .1573-7861.2008.01090.x; Piotr Konieczny, "Adhocratic Governance in the Internet Age: A Case of Wikipedia," *Journal of Information Technology & Politics* 7, no. 4 (2010): 263–283, https://doi.org/10.1080/19331681.2010.489408; Joseph Reagle, *Good Faith Collaboration: The Culture of Wikipedia* (Cambridge, MA: MIT Press, 2010); Dariusz Jemielniak, *Common Knowledge? An Ethnography of Wikipedia* (Stanford, CA: Stanford University, 2014); Emiel Rijshouwer, *Organizing Democracy: Power Concentration and Self-Organization in the Evolution of Wikipedia* (Rotterdam: Erasmus University, 2019), https://repub.eur.nl/pub/113937/.

23. Andrea Forte, Vanesa Larco, and Amy Bruckman, "Decentralization in Wikipedia Governance," *Journal of Management Information Systems* 26, no. 1 (2009): 49–72, https://doi.org/10.2753/MIS0742-1222260103; Loxley Sijia Wang, Jilin Chen, Yuqing Ren, and John Reidl, "Searching for the Goldilocks Zone: Trade-Offs in Managing Online Volunteer Groups," in *CSCW '12: Proceedings of the ACM 2012 Conference*

on Computer Supported Cooperative Work (New York: ACM, 2012), 989–998, https:// doi.org/10.1145/2145204.2145351; Haiyi Zhu, Robert E. Kraut, and Aniket Kittur, "Organizing Without Formal Organization: Group Identification, Goal Setting and Social Modeling in Directing Online Production," in *CSCW '12: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (New York: ACM, 2012), 935–944, https://doi.org/10.1145/2145204.2145344; Jonathan T. Morgan, Michael Gilbert, David W. McDonald, and Mark Zachry, "Project Talk: Coordination Work and Group Membership in WikiProjects," in *WikiSym '13: Proceedings of the 9th International Symposium on Open Collaboration* (New York: ACM, 2013), 3:1–3:10, https://doi.org/10.1145/2491055.2491058.

24. Aaron Shaw and Benjamin Mako Hill, "Laboratories of Oligarchy? How the Iron Law Extends to Peer Production," *Journal of Communication* 64, no. 2 (2014): 215–238, https://doi.org/10.1111/jcom.12082.

25. Benjamin Mako Hill and Aaron Shaw, "Studying Populations of Online Communities," in *The Oxford Handbook of Networked Communication*, ed. Brooke Foucault Welles and Sandra González-Bailón (Oxford, UK: Oxford University Press, 2019), https://doi.org/10.1093/oxfordhb/9780190460518.013.8.

26. Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl, "The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline," *American Behavioral Scientist* 57, no. 5 (2013): 664–88, https://doi.org/10.1177/0002764212469365; Shaw and Hill, "Laboratories of Oligarchy?"; Nathan TeBlunthuis, Aaron Shaw, and Benjamin Mako Hill, "Revisiting 'The Rise and Decline' in a Population of Peer Production Projects," in *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York: ACM, 2018), 355:1–355:7, https://doi.org/10.1145/3173574.3173929.

27. For example, Felipe Ortega, "Wikipedia: A Quantitative Analysis" (PhD diss., Universidad Rey Juan Carlos, Madrid, Spain, 2009), http://libresoft.es/Members/jfelipe /phd-thesis.

28. Sneha Narayan, Jake Orlowitz, Jonathan T. Morgan, Benjamin Mako Hill, and Aaron Shaw, "The Wikipedia Adventure: Field Evaluation of an Interactive Tutorial for New Users," in *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (New York: ACM, 2017), 1785–1799, https://doi.org/10.1145/2998181.2998307.

29. Morgan Michelle Stoddard, "Judicial Citation to Wikipedia in Published Federal Court" (master's thesis, University of North Carolina at Chapel Hill, 2009), http://ils .unc.edu/MSpapers/3526.pdf.

30. Nicholas Vincent, Isaac Johnson, and Brent Hecht, "Examining Wikipedia with a Broader Lens: Quantifying the Value of Wikipedia's Relationships with Other Large-Scale Online Communities," in *CHI '18: Proceedings of the 2018 CHI*

Conference on Human Factors in Computing Systems (New York: ACM, 2018), 566:1–566:13, https://doi.org/10.1145/3173574.3174140; Connor McMahon, Isaac L. Johnson, and Brent J. Hecht, "The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies," in *International AAAI Conference on Web and Social Media (ICWSM 2017)* (Palo Alto, CA: AAAI, 2017), 142–151, http://brenthecht.com/publications/icwsm17_googlewikipedia.pdf.

31. Marit Hinnosaar, Toomas Hinnosaar, M. E. Kummer, and Olga Slivko, "Wikipedia Matters" (working paper, SSRN, Rochester, NY, July 14, 2019), https://papers.ssrn.com/abstract=3046400.

32. Neil Thompson and Douglas Hanley, "Science Is Shaped by Wikipedia: Evidence from a Randomized Control Trial" (MIT Sloan Research Paper No. 5238-17, February 13, 2018), https://papers.ssrn.com/abstract=3039505; Mark Zastrow, "Wikipedia Shapes Language in Science Papers," *Nature News*, September 26, 2017, https://doi.org/10.1038/nature.2017.22656.