THE GENERATION OF KNOWLEDGE · KNOWLEDGE PRODUCTION

Crowdsourcing Knowledge: Wikipedia and Beyond

How crowdsourcing produces knowledge, from Wikipedia to open innovation contests and online research panels, what the evidence says about its accuracy, the coordination and diversity problems it faces, and what it teaches about combining many contributions into reliable knowledge.

ModQu research article · 1,059 words · 11 references (APA 7th edition)

In 2001 a small group launched an online encyclopaedia that anyone could edit. Many observers expected it to fail: how could anonymous volunteers, with no editors checking their credentials, produce reliable knowledge? Two decades later, Wikipedia is one of the most visited websites in the world. Its success is the best-known example of **crowdsourcing**, the production of knowledge and other goods by large numbers of people, usually volunteers, coordinated through the internet. This article examines how crowdsourcing works, what the evidence says about its reliability and limits, and what it means for organisations that gather knowledge from many contributors.

What crowdsourcing is

The journalist Jeff Howe coined the term in a magazine article in 2006, describing how companies were using the internet to outsource tasks to large, undefined groups of people rather than to employees or contractors (Howe, 2006). Brabham (2013) defined crowdsourcing more precisely as an online, distributed problem-solving and production model in which an organisation sets a task and an online community carries it out, with benefits for both.

Benkler (2006) placed crowdsourcing within a broader development he called **commons-based peer production**. Cheap computers and communication networks, he argued, allowed large numbers of people to cooperate in producing information goods, such as software and encyclopaedias, without the prices of a market or the commands of a manager.

The idea draws also on the **wisdom of crowds**. Surowiecki (2004) argued that the aggregated judgements of a large group are often more accurate than those of individual experts, provided that the members of the group are diverse, independent and decentralised, and that there is a good method for combining their contributions.

How accurate is Wikipedia?

The question of Wikipedia's reliability was tested early. Giles (2005) reported a study for the journal *Nature* in which experts reviewed pairs of articles on scientific topics from Wikipedia and from the *Encyclopaedia Britannica*, without being told which came from which. The reviewers found four serious errors in each source. They also found 162 smaller factual errors, omissions or misleading statements in the Wikipedia articles and 123 in the Britannica articles, an average of about four per Wikipedia article and three per Britannica article.

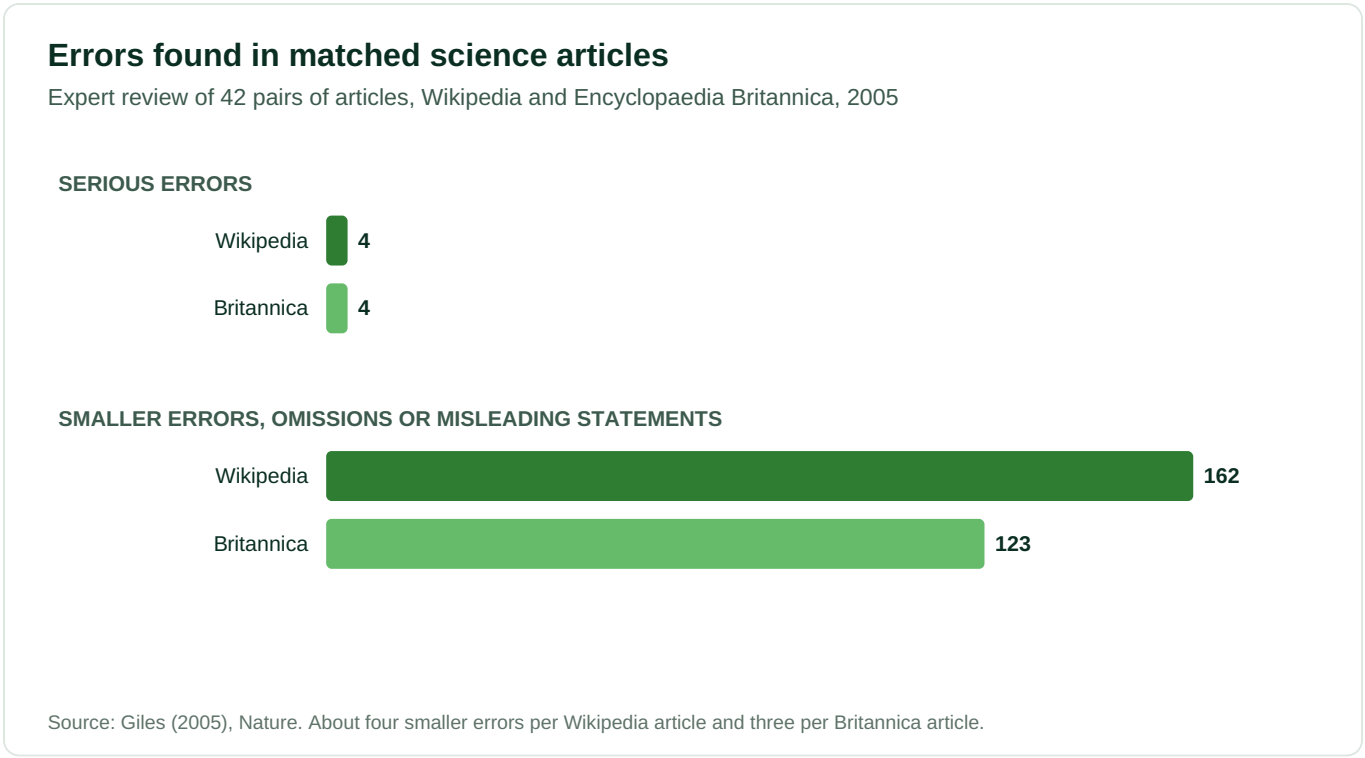


Figure 1. Errors found by expert reviewers in matched science articles from Wikipedia and Britannica. Data from Giles (2005); chart by ModQu.

Later research has painted a more detailed picture. Mesgari et al. (2015) systematically reviewed scholarly studies of Wikipedia's content. They found that its accuracy was generally high, often comparable to that of professional sources, but that coverage, depth and quality varied widely between topics and between languages, and that articles were sometimes incomplete or poorly referenced.

Wikipedia's own practices help explain this performance. Every edit is recorded, so changes can be compared and reversed. Articles are expected to cite published sources, and claims without sources can be challenged or removed. Disagreements are discussed openly on talk pages attached to each article. Policies require a neutral point of view and forbid original research, so that the encyclopaedia summarises what reliable sources say rather than what editors believe. These rules turn a mass of individual contributions into a process with checks, records and accountability.

Why crowdsourcing works, and when it does not

Crowds do not produce good knowledge automatically. Kittur and Kraut (2008) studied how the number of editors affected the quality of Wikipedia articles. They found that adding more editors improved quality only when the editors coordinated effectively, for example when a small group did most of the work and set the direction, or when editors discussed changes on the article's talk page. Without coordination, more contributors could make articles worse.

Crowdsourcing communities also face problems of sustainability. Halfaker et al. (2013) found that the number of active Wikipedia editors began to decline after 2007. They traced part of the cause to the tools and rules introduced to protect quality, which increasingly rejected the contributions of newcomers, including good-faith ones, and discouraged them from staying.

Diversity is another concern. If contributors are drawn from a narrow group, the knowledge they produce may reflect that group's interests and blind spots. Hill and Shaw (2013) estimated, after correcting for survey response bias, that only about one in six adult Wikipedia editors in the United States were women. Gaps of this kind can affect which topics are covered and how.

Beyond encyclopaedias

Crowdsourcing has spread to many kinds of knowledge work.

Innovation contests. Organisations post technical problems online and offer rewards for solutions. Jeppesen and Lakhani (2010) studied one such platform and found that problems were often solved by people whose expertise lay outside the field in which the problem arose. Their distance from the field gave them a fresh perspective.

Online research participants. Researchers increasingly recruit participants through online labour platforms. Buhrmester et al. (2011) found that data collected through one such platform were at least as reliable as data from traditional samples, although later research has raised concerns about inattentive respondents and the repeated participation of the same people.

Distributed analysis. Projects ask volunteers to transcribe historical documents, classify images or check facts, dividing large tasks into small pieces that many people can complete.

Lessons for knowledge-generating organisations

Crowdsourcing shows that many contributions can be combined into reliable knowledge, but only under certain conditions.

Aggregate carefully. The method of combining contributions matters as much as the contributions themselves. Averaging independent estimates, requiring agreement between several contributors and weighting by demonstrated accuracy all improve results.

Preserve independence and diversity. Crowds are wise when their members think independently and come from different backgrounds. Recruiting broadly and avoiding designs in which people simply copy each other protects these conditions.

Build in quality control without discouraging newcomers. Checks are essential, but overly harsh rules can drive away contributors.

Check who is in the crowd. Like any sample, a crowd may not represent the population. Knowing who contributed is necessary for judging what their knowledge can show.

Conclusion

Wikipedia and other crowdsourcing projects show that large numbers of people, coordinated well, can produce knowledge of impressive scale and quality. Their weaknesses are also instructive: uneven coverage, dependence on coordination, difficulty keeping contributors and the risk that the crowd represents only part of society. The lessons apply to any organisation that gathers knowledge from many people, including survey researchers: who contributes, how independently, and how contributions are combined determine how much the result can be trusted.

References

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Brabham, D. C. (2013). *Crowdsourcing*. MIT Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature, 438*(7070), 900–901. https://doi.org/10.1038/438900a

Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2013). The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist, 57*(5), 664–688. https://doi.org/10.1177/0002764212469365

Hill, B. M., & Shaw, A. (2013). The Wikipedia gender gap revisited: Characterizing survey response bias with propensity score estimation. *PLOS ONE, 8*(6), Article e65782. https://doi.org/10.1371/journal.pone.0065782

Howe, J. (2006). The rise of crowdsourcing. *Wired, 14*(6).

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science, 21*(5), 1016–1033. https://doi.org/10.1287/orsc.1090.0491

Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (pp. 37–46). Association for Computing Machinery. https://doi.org/10.1145/1460563.1460572

Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. Å., & Lanamäki, A. (2015). "The sum of all human knowledge": A systematic review of scholarly research on the content of Wikipedia. *Journal of the Association for Information Science and Technology, 66*(2), 219–245. https://doi.org/10.1002/asi.23172

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. Doubleday.