

[TITLE]

Adam Wight
awight@wikimedia.org
Wikimedia Foundation

ACM Reference Format:

Adam Wight. 2018. [TITLE]. In *Proceedings of Understanding the political economy of digital technology (WebSci'18)*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

[THESIS]

[WP is both embedded within capitalism and affected by it, and works against it through free knowledge dissemination, non-vocational education, and consensus principles. At the end of each section, some suggestions are made about how we might fight to expropriate the resource.]

2 RESOURCE LIFECYCLES

This section is intended to give an outline of our main resource streams, their flow, and who holds the control.

2.1 Money

This story begins with money only to say that it's not a dominant force in the Wikimedia movement. The majority of the Wikimedia Foundation's revenue, for example, comes from individual donations under \$30. Almost none [cn] of this revenue goes towards directly producing content.

2.1.1 Where does the money come from?

- Shadow editor employers: Unknown.¹
- Declared paid editor employers: Unknown, see above.
- Individual donors giving \$100 or less: \$65M or about 75% of the WMF budget.^{2 3}
- Unrestricted large grants of \$100 or more: \$22M or about 25% of the WMF budget.
- Restricted, large grants: Assets of \$3.5M⁴

2.1.2 *Where's the money?* The Wikimedia Foundation holds \$113M in net assets.³

Wikimedia chapters have independent finances. To take two examples of large chapters, the German chapter has roughly \$6M in assets, and \$6M in revenue. The Swiss chapter has \$1.8M in assets and \$6M in revenue.

¹This would be an important number to estimate. One possible approach is to write an algorithm which can match paid promotional edits with reasonable precision, then multiply the number of articles being edited by the average price per paid article, currently something like \$400.

²https://www.mediawiki.org/wiki/User:Adamw/Contribution_distribution

³https://upload.wikimedia.org/wikipedia/foundation/d/da/Wikimedia_Foundation_Audit_Report_-_FY16-17.pdf

⁴WMF Audit FY16-17, section (5)

2.1.3 *Who controls the money?* The Wikimedia Foundation's Board has the final say over how their donor money will be spent. This Board consists of 10 members, 5 are self-appointed, 2 members are chosen by chapters, and 3 are nominally elected by a vote where about 4% of active editors normally turn out.

A separate, elected entity called the Funds Dissemination Committee makes recommendations covering roughly \$1M of grants.

Each Wikimedia chapter is legally independent and responsible for its own budget. Most have a board and an executive staff.

2.1.4 *Where does the money go?* The Wikimedia Foundation's major spending is \$33.7M in payroll, and \$11.2M in grants to aligned projects. \$24.2M more is spent on a mixed bag of services and other supporting costs.

Paid editor funds behave like payroll, and don't do further work in the wiki ecosystem.

2.1.5 *"Fight for zero" against capitalist money.* There are efforts to get paid advertising out of Wikipedia, and undeclared paid advertising is already prohibited. Specifically, a rule on English Wikipedia reads: "If you are paid in any way for contributing to Wikipedia, you **must** disclose it."⁵ Another policy reads, "Do not edit Wikipedia in your own interests or in the interests of your external relationships."⁶ Rules exist to be broken, so paid editing continues, both declared and undeclared.

2.1.6 *Fighting capitalist influence on the Wikimedia Foundation.* The low proportion of restricted grants and complete lack of advertising money is a great start for protecting the Wikimedia Foundation's independence.

2.2 Knowledge

Knowledge is the bread and butter of Wikipedia and its sister wiki projects[cn]. This is our currency and the most unique aspect of our worth. One estimate is that the replacement cost of Wikipedia would be \$6.6B, and the consumer benefit is in the hundreds of billions of dollars.⁷

2.2.1 *What is knowledge?* Good question. For our purposes, it has two forms, one is its written form and the other when it's retained in a living human. Knowledge can be true or false. It can be useful or irrelevant to the learner, depending on context. Knowledges can be contradictory and yet coexist. Knowledge has value. Its use value can be decoupled from its exchange value.

The discussions on "talk" pages are the primary means of coordinating and communicating content creation, but we'll ignore as it's an intermediate product of knowledge production.

The corpus of all wikis consists of 351M articles. 13% of these are in English, and the top 6 wikis hold half of all articles.

2.2.2 *Where does knowledge come from?* Wikipedia's guidelines suggest that its knowledge should mostly come from secondary sources.⁸ Heather Ford points out that these rules would

⁵https://en.wikipedia.org/wiki/Wikipedia:Paid-contribution_disclosure

⁶https://en.wikipedia.org/wiki/Wikipedia:Conflict_of_interest

⁷<http://infojustice.org/wp-content/uploads/2013/10/band-gerafi10032013.pdf>

⁸<https://en.wikipedia.org/w/index.php?title=WP:PRIMARY>

create a bias towards scholarly works, which only cover a tiny fraction of subjects, and that the actual sources used are only estimated to be 53% secondary, with tertiary sources accounting for 13% of citations and primary sources 34%.⁹

More troubling, the geographic biases in sources are extreme. 56% of the sampled sources were from the U.S., and 13% from the U.K.

2.2.3 Where is knowledge housed? Our corpus of knowledge is the written form stored on publicly available servers. The Wikimedia Foundation hosts the dominant, editable interface to our corpus, but the raw data and forks are available, in the ballpark of [hundreds] on the Internet, and [thousands] of offline editions.

There is tremendous knowledge in the intersubjective minds of our current and future editors, and this guides their editing. According to wiki rules, editor's personal experience isn't supposed to be directly transcribed into the corpus, it can only mediate work done with other sources of knowledge.

2.2.4 Who produces knowledge? Editors and other wiki contributors are the only producers of knowledge.

2.2.5 How does knowledge flow? Content is added, edited, and consumed mostly using the MediaWiki Web interface. Roughly 5M new articles and 5M media files are added each year.

Third-party consumers will probably follow feeds, or will ingest regular full backups.

2.2.6 How is knowledge curated? Not all information added to wikis will be retained. 1,500 pages are deleted every day, and ca. 8,000 edits are reverted.¹⁰ Most wikis have no mandatory review for edits, and as soon as a change is made, it's visible to the public. A handful of automated bots follow these changes and can undo the worst vandalism almost instantaneously, then human patrollers review the most recent changes that remain, and may fix or undo the change, or mark it as good. These tools will be discussed below.

Changes that create a new article go through a more rigorous review, and the "new pages patrol" backlog has recently been the subject of an experiment in restricting the "supply" of incoming articles.¹¹

2.2.7 Where does knowledge go? Readers find and consume our written corpus. Doing so changes the reader's internal state, hopefully in wonderful ways.

Articles must be discovered in order to be read. If an article is orphaned or has few incoming links, it's unlikely to be read.

Countless AIs and other downstream software devour the corpus in real-time, as it evolves. There is no control or accounting for where this is going. Our licenses don't require anything from these consumers.

2.2.8 Fighting capitalist influence on knowledge. [see fight for zero, above]

2.2.9 Fighting freeloading capitalist use of knowledge. The CC-SA and CC0 licenses allow virtually unlimited capitalist reuse of our knowledge.

We may have the alternative to use the non-commercial CC BY-NC-SA license, although [I don't know] this might not protect

⁹[author?] [1]

¹⁰<https://tools.wmflabs.org/wmcharts/wmchart0004.php>, <https://tools.wmflabs.org/wmcharts/wmchart0008.php>

¹¹https://en.wikipedia.org/wiki/Wikipedia:New_pages_patrol, <https://en.wikipedia.org/w/index.php?title=WP:ACTRIAL>

from indirect uses like training AIs owned by capitalists [cite showing that this might help, but also repeat lessons learned about for-profit unis, etc.]

With no requirement to give back or be accountable to the public, we don't know what the social impact of capitalist reuse might be.

2.2.10 Expand the types and sources of knowledge. If we believe in the wiki model, then more types of knowledge should benefit from it. For example, oral histories could be captured digitally and made available as primary sources.¹²

2.3 Labor

Most editing is done by unpaid editors, whose work is worth an estimated \$490M per year.¹³

2.3.1 Who are the workers? Currently, there are 2,076 paid editors on English Wikipedia who use the recommended disclosure tag, or 6.7% of the active editors.¹⁴

Estimating by subtracting the number of paid editors, 93.3% of the active editors¹⁵ are likely unpaid, or roughly 65,000 active editors across all wiki projects.

There's no authoritative census of editor demographics, this information is rarely self-reported, and surveys vary widely.¹⁶ We can say with some certainty that women are underrepresented among editors, one survey estimated that 12.6% of editors identify as female.

2.3.2 Who's the boss? Unpaid workers are independent.

Declared and shadow paid editors are on either short- or long-term contracts with their clients, who might be individuals or organizations.

2.3.3 Other power relationships between contributors. [admin]

2.3.4 Collaborative relationships between contributors. [talk pages]

2.3.5 Alienation. Alienation is the lack of community between editors.[connectivity studies?]

On a larger scale, isolation between wiki language communities is another form of isolation.[self-focus]

2.3.6 What are the tools of production? Editors own their own computers. The Wikimedia Foundation owns servers and storage attached to their flagship domain names. The Internet does the distribution. Readers will need a networked device to consume. Some editors rely on bot assistance

2.3.7 Where is labor performed? Editing is done from private residences.

2.3.8 Who owns the tools of production?

2.3.9 What are the preconditions for knowledge production? At the moment, editors must have unhindered Internet access. They must be literate, and if unpaid must have the free time to edit.

¹²Peter Gallert has made arguments for the inclusion of oral knowledge, see <https://blog.wikimedia.org/2014/12/12/tapping-into-the-knowledge-of-indigenous-communities/>

¹³[2]

¹⁴<https://tools.wmflabs.org/templatecount/index.php?lang=en&namespace=10&name=Connected+contributor+%28paid%29>

¹⁵The measure we use for "active" is that the user has made an edit within the most recent 30 days.

¹⁶<https://en.wikipedia.org/wiki/Wikipedia:Wikipedians#Demographics>

2.3.10 *What is being produced?* [don't repeat knowledge section] [how much is produced?]

2.3.11 *Fighting for workers' rights: an Editor's Union.* In negotiations between the contributors and entities like the Wikimedia Foundation, it's clear that the editors are at a disadvantage for not having any coherent, collective voice or ability to bargain. A typical solution would be to create an association of Wikimedians, which would have the legal right to act on behalf of its constituents. This proposal first surfaced at the time the Wikimedia Foundation was created¹⁷, and comes up on the mailing lists, usually during crises.

2.3.12 *Fight against censorship and for net neutrality.* State and capitalist censorship or shaping of information is already a problem.

Slight changes in Google's search engine results pages can cause our site traffic to drop dramatically.

2.3.13 *Mitigate scarcity among preconditions.* [work backlogs and their side-effects]

2.3.14 *Counteract biases in contributor demographics.*

3 POLITICAL ECONOMY OF ARTIFICIAL INTELLIGENCE

AIs run by the Wikimedia Foundation (ORES) and by volunteers ([?]) are human-enhancing technologies which exist to help contributors more effectively curate and produce. The economically significant AIs are those that

4 AUDITING

[of AIs and the corpus]

ACKNOWLEDGMENTS

I was paid to write this paper, something less than \$850 in normal wages from the Wikimedia Foundation, and no other funding.

REFERENCES

- [1] FORD, H., SEN, S., MUSICANT, D. R., AND MILLER, N. Getting to the source: Where does wikipedia get its information from? In *Proceedings of the 9th International Symposium on Open Collaboration* (New York, NY, USA, 2013), WikiSym '13, ACM, pp. 9:1–9:10.
- [2] LUND, A., AND VENÄLÄINEN, J. Monetary materialities of peer-produced knowledge: The case of wikipedia and its tensions with paid labour. *TripleC 14* (02 2016), 78–98.

¹⁷https://meta.wikimedia.org/wiki/Association_of_Wikimedians