Wikipedia for multilingual COVID-19 vaccine education at scale

Lane Rasberry[‡], Daniel Mietchen[‡]

‡ School of Data Science, University of Virginia, Charlottesville, United States of America

Corresponding author: Lane Rasberry (lr2ua@virginia.edu), Daniel Mietchen (daniel.mietchen@virginia.edu)

Academic editor: Victor Padilla-Sanchez

Abstract

We present the design of a project to develop Wikipedia content on general vaccine safety and the COVID-19 vaccines, specifically. This proposal describes what a team would need to distribute public health information in Wikipedia in multiple languages in response to a disaster or crisis, and to measure and report the communication impact of the same. Researchers at the School of Data Science at the University of Virginia made this proposal in response to a February 2021 call from a sponsor which was seeking to share public health information to respond globally to vaccine hesitancy related to the COVID-19 vaccines. This proposal was not selected for funding, and now the research team is sharing the proposal here with an open copyright license for anyone to reuse and remix. Most of the text here is from the original proposal, but there are modifications to remove the names of the funder, named partners, and for other details to make this text more reusable. The budget in this proposal has been converted from a dollar amount to equivalent descriptions in terms of labor hours, and the timeline was adapted from absolute to relative months.

Keywords

vaccines, vaccine hesitancy, public health, Wikipedia, health education

Goals and Objectives

This project aims to counter vaccine hesitancy by providing high-quality content in Wikipedia. Wikipedia is an important health communication channel for reasons including its extreme popularity and its ease of global access (Smith 2020). In addition, the project will translate this content into several Indian languages. We choose to pilot India as a communication target among other options because a significant portion of the world's population is there and we have identified a need in the media ecosystem there for quality, fact-checked content on vaccines. Through this project, we will promote patient education

©
.

by providing quality content to millions of consumers - including many patients and parents - who consult Wikipedia for information on vaccines and COVID. We will also reach patients indirectly because doctors, nurses, medical students, policy makers, and journalists also consult Wikipedia to orient themselves (Heilman and West 2015). Wikipedia's nature as a digital medium enables us to measure user engagement in various ways that will demonstrate the impact of our intervention.

The milestones for the project include identifying key facts to share; integrating those messages and multimedia into relevant English Wikipedia articles, including into the existing one titled "Vaccine hesitancy" (cf. Fig. 1a); translating it into a set of pilot languages (Hindi, Bangla, Urdu, and Nepali); overseeing the activities with a process for promoting ethics and community engagement; documenting our process, and reporting the readership/audience metrics (cf. Figs 1b, 2) to illustrate communication impact. We will report our results through a publication in a peer-reviewed journal, so that Wikipedia can be understood as a model for impactful communication on this and any topic in any language.

Wikipedia is easy to use and easy to access. Search engines present Wikipedia articles high in query results at no marketing cost, making it easy to find with minimal effort. We propose to use this high-impact, low cost distribution channel to present facts of vaccine safety at scale to the people who seek out such information online. The proposed work has the potential to be far reaching due to Wikipedia's large audience. With about one billion annual visitors overall, Wikipedia's readership for the medical content alone has been estimated to be hundreds of millions (Okoli et al. 2014). For example, in 2020, the Wikipedia Pageviews Analysis report for the "vaccine hesitancy" article in English showed a total of 588,922 pageviews (Fig. 1b). As we will edit multiple articles in a few languages, we estimate a reach of at least 2 million people per year, all of whom are specifically searching for information related to vaccines. We will choose Wikipedia articles to develop based on the facts we collect from leading COVID-19 educational campaigns and based on our evaluation of existing audience demand as measured by Wikipedia pageviews. Because our model of publishing key facts in Wikipedia articles scales up or down depending on how many articles we target and which languages, we have the ability to modify or expand our communication depending on need and discovered opportunity.

Needs Assessment for the Project

Wikipedia readers are a large demographic of global media consumers seeking information about vaccine hesitancy and related topics. At a time when new vaccines are coming to market with great frequency, there is more demand for clarity on the science, efficacy, and delivery of the vaccines. In parallel, misinformation around the side effects is proliferating online. In order to inform public understanding and decision-making, consumers need quality, fact-checked information on a widely used, frequently accessed, free platform. As a university research team that develops medical content in the Wikipedia ecosystem, we propose to serve this need by using Wikipedia as that communication platform. We will do

this by enhancing and developing Wikipedia content, in multiple languages, with key messages from expert sources.

Target Audience

Our target audience is people who use Internet search in our target languages to seek information about vaccines, vaccine safety or topics otherwise related to vaccine hesitancy. Among this demographic, we are able to count readers who navigated to Wikipedia to access information on vaccine hesitancy, safety, and COVID-19 vaccination. The total annual readership of Wikipedia's medical content is hundreds of millions of unique visitors, which is the potential audience in this platform. Of those, we estimate that 2-10 million are accessing Wikipedia's vaccine content, and we expect to be able to reach them with our intervention. Because Wikipedia ranks highly in the query results of leading search engines, these Wikipedia readers are a highly targeted audience in addition to being an attractively large audience.

We can observe reader interest with Wikipedia article traffic reports. Fig. 2 shows traffic from 2015-2021 to Wikipedia articles on topics including "vaccination", "vaccine hesitancy", "vaccine controversies", "vaccine design", and "anti-vaccination". The readers which this chart counts are our target audience. Metrics in this chart give information on content popularity, quality, and editor engagement, all of which are factors which communicate what sort of development intervention is appropriate for a given article or set of articles. In our project, we will use such data to guide our publishing strategy and also to demonstrate the communication impact of our results.

Wikipedia's 2020 reader metrics show a total of 65,212,508 requests to the 204 English Wikipedia articles in "Category: Vaccine hesitancy" (cf. Fig. 3), and as shown in Fig. 4, there were 1,169,567 requests to the 30 language versions of Wikipedia articles for vaccine hesitancy (including the 588,922 requests for the English version, as per Fig. 1b). These numbers place the topic "vaccine hesitancy" within the top 10% of Wikipedia articles by popularity, and we assert that communication reach to an audience of this size is an admirable outcome for a media outreach campaign of any sort.

In addition to providing content in English, we also will do translation into the most spoken languages of India, Pakistan, Bangladesh, and Nepal. Collectively, these languages have 500 million native speakers who need information in their own language as part of the global COVID response. Other reasons for targeting these languages are that they are in a language family, often using similar translations for technical terms. Moreover, they are all within the same region and overlapping cultures, and speakers of these languages are accessing the same vaccines. We at our university have already done Wikipedia medical translations in these languages and have workflows in place for doing more. We will document our process, which works for sharing quality medical information in Wikipedia on any topic and in any of Wikipedia's languages.

Project Design and Methods

We seek to advance public health education related to vaccine hesitancy by collecting facts that recognized expert organizations have prioritized for dissemination, and by integrating these facts into the Wikipedia ecosystem in English, Hindi, Urdu, Nepali, and Bangla. As the languages of India are less developed in Wikipedia and in digital media generally, for many facts, we will be providing new access to audiences who may not have had published information of any quality before. In doing so, we will increase the accessibility of high-quality health facts, provide fact-checking service with Wikipedia's citation process, make media available for remixing and reuse, and generate communication metrics that demonstrate Wikipedia's extensive reach in various languages. We recognize that Wikipedia does not have a reputation as a conventional public health education channel. Regardless of any perceptions of Wikipedia's quality, Wikipedia does have a system of quality control and there are critiques of Wikipedia's medical content which review it favorably (Smith 2020). We argue that content development of the sort we propose will improve Wikipedia's quality, making it an effective tool in countering misinformation and providing trustworthy, educational content.

Our educational approach is to publish content on the topic of vaccine hesitancy and on closely related topics, such as vaccine safety, in the Wikipedia ecosystem according to the norms of Wikipedia's culture. We collect information from expert sources, such as by the recommendation of the World Health Organization's Vaccine Safety Net project. By distributing the messages which have consensus backing of experts, we gain the efficiency of reusing key facts which already have global backing for distribution. With facts identified, we focus on Wikipedia's strength in information distribution and building trust through citations to reliable sources. Our core team includes three experienced Wikipedia health editors at the School of Data Science at the University of Virginia who can demonstrate via past projects that our methods are aligned with Wikipedia's best practices and have extensive impact.

Our method is as follows:

1. Curation - identify key facts to share and make the fact-checking transparent.
2. Editing - improve relevant parts of English Wikipedia articles to accurately represent those facts and their sources.
3. Translation - convert English into the Hindi, Bangla, Urdu, and Nepali Wikipedia versions.
4. Ethics - address problems we identify and solicit community engagement.
5. Documentation - publish our process and demonstrate impact with audience metrics.

This is the model we use to organize our activities, the results of which will be that the large number of people who use Wikipedia will have better access to higher quality information on vaccines and closely related topics. We will develop the use of Wikipedia as

a model for health communication in general (see also Mietchen et al. 2021), so that others can weigh the costs and benefits of Wikipedia for other public health campaigns.

Finally, it's important to point out that the Wikipedia metrics reports we run examine Wikipedia over all its content, and not just conventional medical information. In our preliminary review of Wikipedia metrics, we observed an unusual recent increase in traffic to biographies of celebrities who promote the anti-vax movement. In such cases, we have the option to fact-check those claims and lead readers to quality information. Besides biographies, COVID appears elsewhere, for instance in articles on pop culture, travel, sports, science, and politics. While we cannot address all these diverse fields of communication, Wikipedia has a unique readiness for fact-checking at the intersection between pop culture and medicine. As we examine sets of Wikipedia articles related to "vaccine hesitancy", we expect to provide medical information in some culture-specific contexts as the need becomes apparent.

As part of our content curation process, we will use Wikipedia's workflows for curating the academic literature and media offerings of expert organizations. Wikipedia also has ways for demonstrating due diligence in media review, soliciting community engagement, and surfacing available multimedia to complement our text communication.

Fig. 5a and Fig. 5b show sample data visualizations in Wikipedia that support making editorial decisions for Wikipedia by mirroring expert consensus. We are part of the community that curates such analyses.

Innovation

This project is innovative because it leverages an existing popular media channel to resolve the challenge of message distribution, enhances content, and translates it into languages that will extend its impact to underserved populations in different geographies. Whereas most public health experts have quality content but face high costs for meaningful delivery to relevant audiences, Wikipedia's entire readership is people who are actively seeking specific information. Wikipedia's bottleneck is acquiring that quality content, and our project will bridge the gap of integrating existing quality content into Wikipedia as a communication channel.

We will consider sourcing our information from resources like the WHO, CDC, UN Women, and India's Ministry of Health and Welfare. We intend to document our approach with uncommon formality because we anticipate that observers in public health will want to understand what we did, how, and to what effect. Combining a number of Wikipedia activities into this workflow and documenting it is an innovative way to demonstrate public health impact.

Our team at the University of Virginia has experience translating Wikipedia content into Indian languages through projects on health issues native to Indian populations. We know that reliable health information is needed in parts of India where doctors and other medical resources are not accessible in certain locations. Additionally, health information online is

not always available in regional languages. Hindi, Bangla, Urdu, and Nepali are underserved with access to COVID-19 and other health information. Our project would translate vaccine information and reach the sizable populations who speak these languages and are seeking quality information. Translation is a sensitive task, and to ensure quality and accuracy, we will organize a research subproject at our university to identify, list, and propose actions to reduce potential social and ethical errors.

Matching routine activities in Wikipedia to public health communication is itself an innovation. We innovate by documenting Wikipedia's best practices to make them more accessible for more health communication for COVID, for other disaster response, and for improving patient and family access to health educational materials in general.

Evaluation and Outcomes

The primary outcome we will measure is the number of pageviews to Wikipedia articles containing vaccine information over a range of time. Secondary outcomes include counts of editor responses, a report of the Wikipedia indexing of relevant scholarly literature, the count of multimedia complements contained in relevant Wikipedia articles, a multilingual glossary of the technical terms we translated, the report on the ethical issues which we addressed, and the documentation of our methodology. Much of the data we will collect will be from Wikipedia's built-in suite of communication metrics tools. Although some interpretation of the tools merits further discussion, in general, high pageviews indicates reader interest; diverse editor engagement indicates better editorial process; and presence of supplementary media including images and citations indicates higher accessibility and reliability.

A FICTIONAL but plausible summary of a single-topic report could be similar to the following:

The Wikipedia article for "vaccine hesitancy" existed in April 2021 in 3 of our 5 target languages, English, Hindi, and Bangla, but was missing in Urdu and Nepali. In all existing articles, our 3 relevant key facts were absent. We intervened to create articles for absent languages and to include our key facts in each. Our key facts are supported by the matched citations we listed in our communication plan. In each article, we also included an image from our curated collection, which in this case was a culturally appropriate model receiving an injection in a medical setting. To demonstrate due diligence, we followed Wikipedia workflows to curate the network of scholarly publications on this topic which we considered and made that accessible to reviewers. The primary outcome was that for the 3-month observation period, 120,000 readers accessed this Wikipedia article in English language and an additional 20,000 accessed it in the other developed language versions combined. All articles had subsequent editorial review, resulting in one objection which we answered to the satisfaction of reviewers. Overall, the English Wikipedia article was created 15 April 2005 and has since had 3,535 edits by 1014 editors. It contains 12,246 words of readable prose and has 284 citations. Our contribution to it was 0.8% of text, 0.3% of edits, 2% of references and 7% of media files.

In the above FICTIONAL example, we either collected live data or extrapolated a number based on actual data. There is no standard format for reporting communication impact in Wikipedia, but we would like to create another precedent by documenting our approach to patient education in this project. Although, individually, all of the above metrics are part of common Wikipedia community discussion, it is our own innovation for this project to combine them as a model for quantifying health communication intervention in Wikipedia. In addition to the example here with one article, we will also produce aggregate reports for all the content we curate by language.

Dissemination Plan

Wikipedia is already among the most requested, published, accessed, and consulted sources of medical information in the world, and also very popular for the topic of vaccine hesitancy. The conventional dissemination plan of a typical public health messaging campaign presumes that content is available and that the challenge is in delivering content to relevant audiences. In contrast, Wikipedia has the opposite challenge of having an excellent platform for reaching audiences but needing quality content to distribute.

In the context of Wikipedia, the dissemination plan we will apply is:

1. increasing the accessibility of the best available facts that we identify from other health communication efforts on the same topic,
2. ensuring that our media are machine readable with free and open copyright licenses for perpetual discovery and reuse in any future communication efforts by us or anyone else, and
3. documenting our content development so that other journalists, policy makers, or public health organizations can understand the value and reuse potential of the content they discover through Wikipedia.

Wikipedia is a media platform that delivers relevant content on request to anyone seeking it. It does not attempt to originate key facts but, rather, is a tool for managing the global search for the most effective messaging from the most authoritative sources on the topic. After identifying authorities and considering their facts, we summarize and cite them for distribution in Wikipedia. If organizations offer supplementary multimedia or datasets with a compatible copyright license, then we may present those in Wikipedia, such as we did when the popular *SELF* magazine donated vaccination images with an open copyright license including the example shown in Fig. 6.

As is typical in the Wikipedia ecosystem, all of the content we publish will have a free and open copyright license. This means that the resources and tools we develop will be available for anyone to reuse or remix. If other participants in the network of this sponsored call have facts and reusably licensed materials for us to distribute, then we would very much like to collaborate with them, as we anticipate that they are likely to be experts in need of a communication channel, while we are communicators in need of expert facts on COVID-19 vaccine hesitancy. In this way, our project can also expand the dissemination of

other collaborators and we can credit them with communication metrics as we do for ourselves.

Anticipated Project Timeline

Month 1

- Project start; appoint and orient supporting researchers
- Scope ethical practices, train team, reach mutual understanding
- Start all project arms: curation, editing, translation, ethics, and documentation

Month 2

- Begin workflow trials, including some work product in each project arm
- Set up project page for listing key facts, allies, and media sources to process
- Pilot workflows in converting text to structured data and translated versions

Month 3

- Seek expert peer review of collected facts
- Community conversation and notification in each of the target languages
- Wikipedia community has access to project documentation and goals

Month 5

- Key facts identified, peer-reviewed, and prepared for translation
- Bibliography of academic literature indexed in Wikipedia for sharing
- All appropriate content staged for translation

Month 6

- Selected facts integrated into Wikipedia articles
- The fact curation process is documented such that others could use it in other contexts
- Begin 3-month readership tracking measurement and observation

Month 9

- End 3-month readership tracking measurement
- Resolve any outstanding community or peer requests
- Curation, editing, and translation arms are done except for change requests by reviewers

Month 10-12

- Begin processing and reporting results
- Community disclosure of outcomes and conversation
- Wind down project; turn over maintenance to Wikipedia crowdsourced process

Month 13-14

- Ethics and documentation arms compile results for final reporting
- Project-related activities in Wikipedia ecosystem have ended
- Scholarly write up of project, seek venue for preprint and possible publication

Month 15

- Final report, preprint published, peer-reviewed publishing in process
- Project end

Additional Information

Wikipedia is 20 years old, stable, globally popular, and has a measurable quality control process. Over the years in seeking medical communication partnerships, we have heard every criticism of Wikipedia and shared many of the hundreds of scholarly articles describing Wikipedia's medical content and the various journalism around it. It has always been challenging for us to overcome barriers of trust because of unfortunate preconceptions that people have about Wikipedia. We appreciate the consideration that you can give for Wikipedia's place in the media environment for COVID-19 vaccines and vaccine hesitancy.

Organization Detail

The core team is in the School of Data Science at the University of Virginia. While most activities at this organization focus on research with machine learning, our proposal here does not contain that activity. More generally, though, the school does various Wikipedia-oriented research projects and is also a center at the university for other Wikipedia engagement projects in the library, classes in the humanities, at the medical school, with student volunteer interest groups, and as part of a portfolio of developing FAIR and open data for public benefit. Also within the School of Data Science, we administer Wikipedia projects in the Center for Ethics and Justice, which influences our approach to Wikipedia program design.

The University of Virginia was founded in 1819. It has 30,000 full time employees. In 2019, it had US$2.9B operating revenue with $197M in philanthropic support and $175M in state support. While this project has no formal collaborations with other organizations, our approach does create an unusual opportunity for us to disseminate whatever facts we identify as most reliable.

Organizations whose messaging on vaccine hesitancy we will consider include the following:

1. Vaccine Safety Net, a project of the World Health Organization
2. UN Women, a project of the United Nations

3. Centers for Disease Control, a United States federal agency
4. Ministry of Health and Welfare, an Indian government ministry

We have no dependency on these organizations and will do no activities which require their permission, support, or participation. We may inform them of our activities and results.

The core team members are all with the University of Virginia and have been active editors of Wikipedia's medical content since about 2010. Based at the School of Data Science, we seek to communicate in ways that scale and have a foundation in structured data. We already collaborate with experts and students across the university - including at the medical school, the nursing school, the university hospital and the medical library - to review and share medical information.

Lane Rasberry is Wikimedian in Residence at the university and in that role shares information on Wikipedia, focusing on medical communication. Since 2007, Rasberry has reviewed clinical trial messaging in the HIV Vaccine Trials Network. In 2020, Rasberry reviewed practices for COVID vaccine development in the United States government program Operation Warp Speed. Other relevant background includes publishing and fact-checking as Wikimedian in Residence in the Choosing Wisely health campaign at Consumer Reports, a magazine and nonprofit consumer organization.

Daniel Mietchen is a data scientist with a background in evolutionary biophysics. Data sharing in public health emergencies is one of his research foci. He has been serving on the Zika response team at the NIH, on the data sharing working group of the Global Research Collaboration for Infectious Disease Preparedness, and as co-lead of community engagement activities for the Research Data Alliance's COVID-19 working group.

Abhishek Suryawanshi is a pharmacist, public health expert, diplomat and liaison between the School of Data Science and various Indian government agencies. He organizes projects in health communication and translation which include native speakers recruiting community oversight and review of the localization.

Budget Narrative

From a staffing perspective, 30% of the budget is for the three core team members to do most of the content creation and distribution, 20% is for quality and ethical review of the content and process, 15% is for language translation, 10% supports dedicated documentation for community outreach, and the remaining is the 22% overhead which the sponsor suggests is appropriate for our university. Elsewhere in this proposal, we describe the activities in this project as curation, editing, translation, ethics, and documentation. In considering the budget by activities rather than staffing, each of these five activity categories consume approximately equivalent resources in the budget.

In our core team, Rasberry is principal investigator and editor for English language educational content. Mietchen does literature review, manages structured data, and stages content for review in terms of quality, relevance, and ethical appropriateness. Suryawanshi

coordinates translation from English into the four other languages and also is a community point of contact for Wikipedia reviewers in South Asia.

The core team collaborates with a mix of university researchers and external vendors for the content review. Student and faculty researchers will identify and document ethical challenges in cross-cultural medical translation and the propagation of structured data. Translators are vendors who are native speakers of our target languages. We know that any project observers such as community stakeholders may have questions about our content and process, and we proactively address those questions by developing project documentation from the beginning.

Budget table

A spreadsheet version of the budget is provided in Table 1.

Conflicts of interest

References

- Heilman JM, West AG (2015) Wikipedia and Medicine: Quantifying Readership, Editors, and the Significance of Natural Language. Journal of Medical Internet Research 17 (3). https://doi.org/10.2196/jmir.4069
- Mietchen D, Rasberry L, Morata T, Sadowski J, Novakovich J, Heilman J (2021) Developing a scalable framework for partnerships between health agencies and the Wikimedia ecosystem. Research Ideas and Outcomes 7 https://doi.org/10.3897/rio.7.e68121
- Okoli C, Mehdi M, Mesgari M, Nielsen FÅ, Lanamäki A (2014) 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 (12): 2381-2403. https://doi.org/10.1002/asi.23162
- Smith D (2020) Situating Wikipedia as a health information resource in various contexts: A scoping review. PLOS ONE 15 (2). https://doi.org/10.1371/journal.pone.0228786

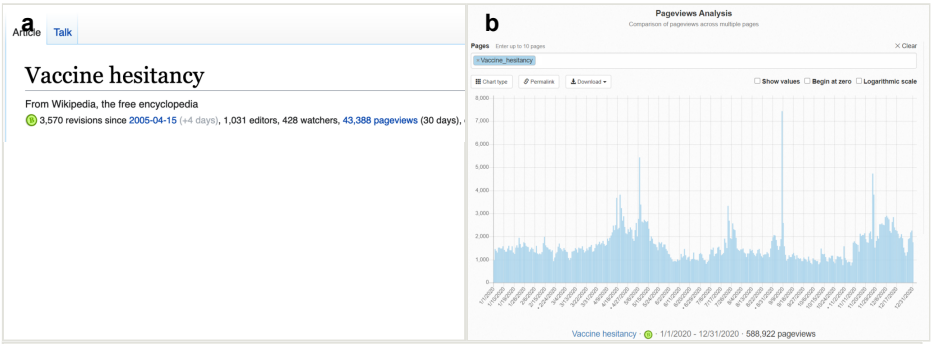


Figure 1.

Engagement metadata of the English Wikipedia article "Vaccine hesitancy".

a: Some experienced Wikipedia editors use account extensions to turn on the display of article metadata as seen here in the article header. The article is currently graded as "B" class on Wikipedia's quality scale. 1,031 Wikipedia editors have made 3,570 editorial revisions to the article since the article's creation on 15 April 2005. There are 428 registered Wikipedia editors who have put this article in their watchlist, which means that they have alerts about the article's development either whenever they request it or by some push notification. In the 30 days proceeding April 6, 2021 (when this screenshot was taken), this article has received 43,388 Wikipedia pageviews. The image is in the public domain and available via Wikimedia Commons.

b: This is a Pageviews Analysis for that same article for the calendar year 2020. This screenshot is available via Wikimedia Commons under the terms of the Expat/ MIT license.

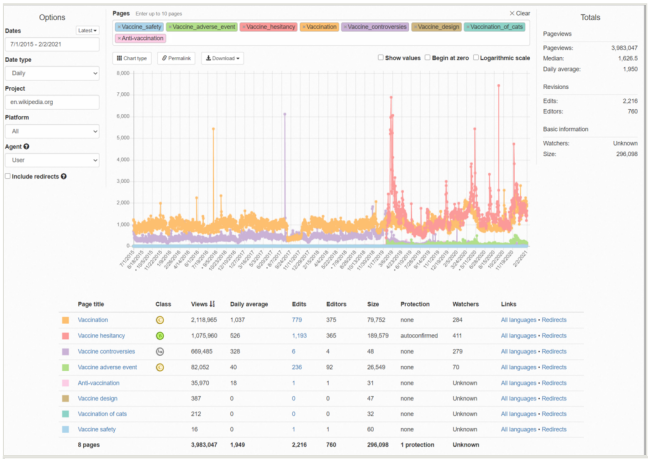
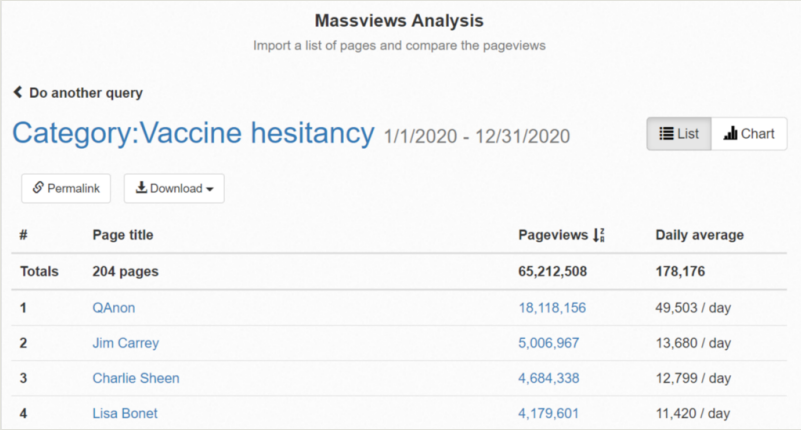


Figure 2.

This is a Pageviews Analysis of multiple English Wikipedia articles related to vaccine hesitancy. Data visualizations such as this give insight into reader demand by reporting user engagement with sets of Wikipedia articles over time. Among other values, the table gives reports for "views", a measure of readership; "class", a grade of the content quality; and "editors", which is a count of the number of people who submitted editorial content. Wikipedia editors use these metrics to prioritize development of popular articles in need of more quality content. This screenshot is available via Wikimedia Commons under the terms of the Expat/MIT license.



Massviews Analysis

Import a list of pages and compare the pageviews

Do another query

Category:Vaccine hesitancy 1/1/2020 - 12/31/2020

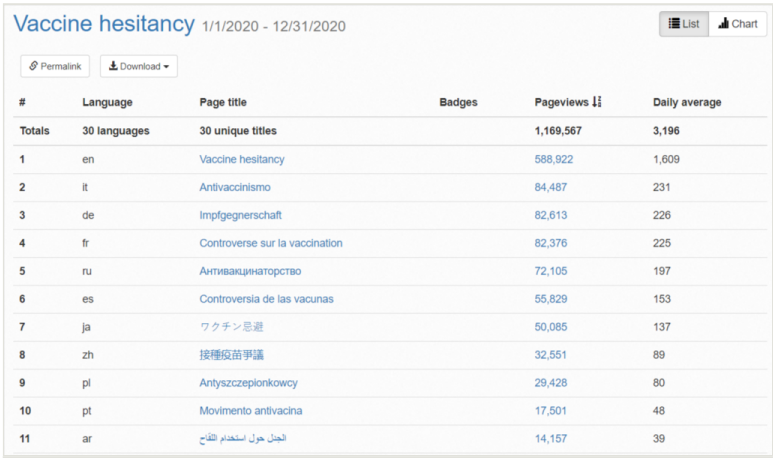
List | Chart

Permalink | Download

#	Page title	Pageviews	Daily average
Totals	204 pages	65,212,508	178,176
1	QAnon	18,118,156	49,503 / day
2	Jim Carrey	5,006,967	13,680 / day
3	Charlie Sheen	4,684,338	12,799 / day
4	Lisa Bonet	4,179,601	11,420 / day

Figure 3.

Overview of pageviews for articles in the English Wikipedia's category for "vaccine hesitancy" over the course of 2020. This screenshot is available via Wikimedia Commons under the CC0 1.0 Universal Public Domain Dedication.



Vaccine hesitancy 1/1/2020 - 12/31/2020

List | Chart

Permalink | Download

#	Language	Page title	Badges	Pageviews	Daily average
Totals	30 languages	30 unique titles		1,169,567	3,196
1	en	Vaccine hesitancy		588,922	1,609
2	it	Antivaccinismo		84,487	231
3	de	Impfgegnerschaft		82,613	226
4	fr	Controverse sur la vaccination		82,376	225
5	ru	Антивакцинаторство		72,105	197
6	es	Controversia de las vacunas		55,829	153
7	ja	ワクチン忌避		50,085	137
8	zh	接種疫苗爭議		32,551	89
9	pl	Antyszczepionkowcy		29,428	80
10	pt	Movimento antivacina		17,501	48
11	ar	الجدل حول استخدام اللقاح		14,157	39

Figure 4.

The most accessed language versions of Wikipedia articles for "vaccine hesitancy" in 2020 were English, Italian, German, French, Russian, Spanish, Japanese, Chinese, Polish, Portuguese, and Arabic. This screenshot is available via Wikimedia Commons under the CC0 1.0 Universal Public Domain Dedication.

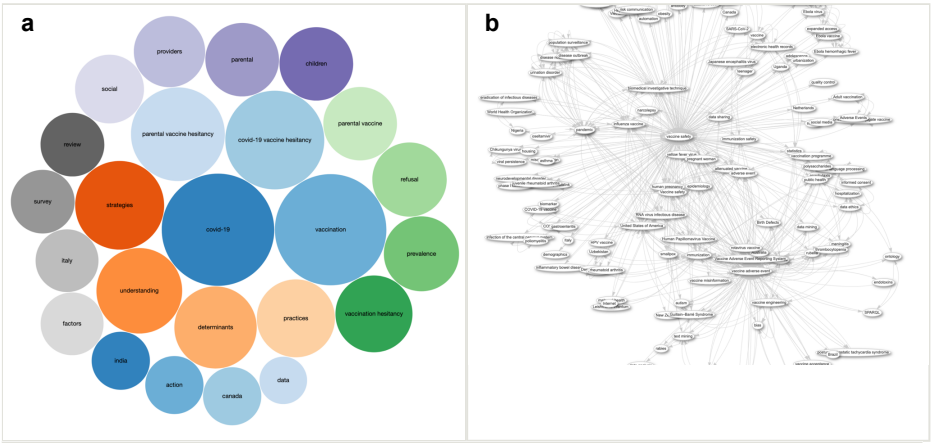


Figure 5.

Sample visualizations of data from within the Wikipedia ecosystem that relate to COVD-19 vaccine hesitancy.

a: Frequency of terms appearing in a Wikipedia-indexed set of papers on COVID-19 vaccine hesitancy. This screenshot is in the public domain and available via Wikimedia Commons.

b: Tools in the Wikipedia ecosystem show the network relationship of terms in academic publications on COVID-19. This screenshot is in the public domain and available via Wikimedia Commons.



Figure 6.

SELF magazine provided this photograph by Heather Hazzan with a free and open copyright license (CC BY 2.0) for use in Wikipedia or any other publication which required vaccine illustrations. It is available via Wikimedia Commons.

Table 1.

Line items in the budget, expressed in terms of labor hours, full-time equivalents (FTE) or market rates.

Documentation	30 labor hours
Video recording and editing	30 labor hours
Open access publication fees	market rate
Graphic design	45 labor hours
Translation manager	~0.1 FTE
Translation (English and Hindi to Bangla, Nepali, Urdu)	market rate
Data science manager	~0.1 FTE
Subject matter expert, vaccine safety	75 labor hours
Student research - ethics	150 labor hours
Principal investigator	~0.1 FTE
Total Direct Expenses	78%
Indirect/Overhead	22%
Total	100%