

22 Capstone: Making History, Building the Future Together

Katherine Maher

The Wikimedia Foundation's executive director reflects on the past, present, and future of Wikipedia, informed by Wikimedia 2030, the vision and strategy project by the global Wikimedia movement and free knowledge partners.

Wikipedia turned fifteen years old on January 11, 2016. Later that same year, in a time of concern about “fake news” and disinformation, I attended an event focused on the future of the news media. Wikipedia isn't a news organization, so I was an outlier in the room. I seated myself at a small roundtable on the topic of trust with some hesitation: after all, journalists are frequently skeptical about Wikipedia's reliability. One by one, the attendees went around the room, introducing themselves and the reason they'd joined the session.

When it was my turn, I said: “Wikipedia has gone from being a punch line about the unreliability of people on the internet to becoming one of the most trusted sites online. I'm here to see what we can learn from one another.” To my surprise, there were nodding heads around the table. It was a moment that would have been difficult to imagine even a few years ago.

As Wikipedia and the Wikimedia movement and projects enter our third decade, my expectation is that we'll continue to confound expectations. Today, Wikipedia includes fifty million articles across three hundred languages, ranging in size from six million articles on English Wikipedia to just over one thousand articles on Tulu Wikipedia. It is joined by a number of other successful free knowledge projects, including Wikidata, Wikimedia Commons, and Wikisource. Every single month, we estimate that around one billion people spend a collective sixty thousand years reading

Wikipedia. Together, the Wikimedia projects constitute one of the top ten most visited digital platforms on the planet.

Much more than an encyclopedia that anyone can edit, Wikipedia has fundamentally and irrevocably transformed models for how people and communities can experience and create knowledge, within and far beyond the various Wikimedia projects. It is a resource to people around the globe seeking information about history, politics, and pop culture. It is a project in nation building through language, a tool for cultural preservation, and a platform for debates over representation and truth. It is a database used by researchers, universities, and cultural institutions to share and publicize their data and collections.

It is one of the world's most widely used sources for training machine-learning applications. It is a trove of insight about humanity: our interests, our predilections, our biases. It is a byword for collaborative participation, a definitive oracle (has that celebrity really died?), a pop-culture signifier (see Stephen Colbert, elephants¹), and an abbreviated verb for information seeking ("let me wiki that"). Each year, as Wikipedia has grown, it has become more integral, more important, and more irreplaceable to our shared cultural consciousness. Far more than an internet encyclopedia, it is a living compendium of our knowledge.

Wikimedia's Future

As this volume of reflections on the first twenty years of Wikipedia goes to press, the Wikimedia movement has recently completed a global, collaborative process to build a vision for the Wikipedia of 2030. Launched shortly after Wikipedia's fifteenth birthday, this "movement strategy" was an opportunity to consider what the Wikimedia community had accomplished and what was still to come. It was a chance to look at the distance between "the encyclopedia anyone can edit" and "a world in which every single person can freely share in the sum of all knowledge"—and ask ourselves how the Wikimedia movement might set about closing that gap. What would it take to reach more people? What would it mean if the whole world really could participate? What does "all knowledge" even mean?

To try to answer these questions, members of the Wikimedia movement spent a year talking to each other and others around the globe. They worked to reinterpret our vision—"Imagine a world in which every single

“Imagine a world in which every single human being can freely share in the sum of all knowledge. That’s our commitment.”

The vision of the Wikimedia movement is more than an aspirational statement. It offers us principles for how we approach our work.

Explicit

- *“Every single human being.”* Everyone should be able to participate in knowledge, regardless of origin, ability, nationality, geography, language, culture, or other identifying factors.
- *“Freely share.”* Engagement with knowledge should not be subject to barriers or limits, whether economic, political, social, or otherwise. Knowledge should be participatory, easily read, reused, remixed, or revised—not merely consumed.
- *“Sum of all knowledge.”* Knowledge is vast, mutable, and continuously in evolution. “The sum” of knowledge should be inclusive and representative of the diversity of the world at large, unlike many canonical representations which suffer from implicit biases.

Implicit

- *We cannot do this alone.* An aspirational vision of this magnitude cannot be achieved by only Wikimedia. This paradigm shift requires many partners and allies.
- *We are still far from this world.* Representations of knowledge are imperfect and often heavily biased. Many barriers and limitations preclude access and participation.
- *A better world.* Our vision has a purpose beyond collection. A world of greater understanding is a world better equipped to address our collective challenges.

human being can freely share in the sum of all knowledge”—and to make plans for what we should be doing to realize it.

We hosted gatherings and discussions with people from seventy countries and consultations in more than twenty languages. We spoke with current Wikimedia movement members and partners as well as people learning about Wikimedia for the first time. We commissioned research into the state of the world today and the state of the world to come. We interviewed 150 experts from the worlds of academia, arts and culture, epistemology, education, open science, and technology.

As a community of collaborators and information enthusiasts, we took the values and practices honed over years of creating Wikipedia and used them to explore, examine, and propose a direction for our future, together. One thing quickly became evident in our conversations about the future—the world the Wikimedia projects emerged from is no longer the world in which we operate today.

Some of these changes are promising and positive, offering us new opportunities to interpret our vision, connect with people, and expand free knowledge in the world. However, just as many are concerning changes with potential negative implications for the long-term health of the global Wikimedia community, our projects, and our ability to pursue our vision of a world of international cooperation, constructive discourse, and collaboration in the service of our global knowledge commons.

We see a world that's more connected than ever before, with bandwidth costs decreasing, making it easier for everyone to get online. Primary and secondary education enrollment rates are rising, as are global literacy rates. We're seeing a growing population of young, engaged, and online youth eager to effect change in their communities and on a global scale. But alongside these positive changes, we're also seeing new challenges and threats.

The world is becoming less open as authoritarian governments close spaces for dissent and debate. Democracies are struggling with increased polarization and decreased trust in institutions. The internet, once a relatively open and creative space, has become increasingly consolidated, centralized, and homogenized, perpetuating power and control within a handful of corporations. Data gathering and tracking has enabled a "public-private surveillance" economy that seeks to know everything about everyone.

The Wikimedia 2030 consultation put these changes at the center of the conversation, recognizing the need for our projects and communities to continue to adapt and evolve in order to meet the opportunities and challenges ahead.

Wikimedia 2030

We synthesized the collective thoughts, feedback, and hopes from hundreds of conversations about the future into a shared direction.

By 2030, Wikimedia will become the essential infrastructure of the ecosystem of free knowledge, and anyone who shares our vision will be able to join us.

We, the Wikimedia contributors, communities, and organizations, will advance our world by collecting knowledge that fully represents human diversity, and by building the services and structures that enable others to do the same. We will carry on our mission of developing content as we have done in the past, and we will go further.

Knowledge as a service: To serve our users, we will become a platform that serves open knowledge to the world across interfaces and communities. We will build tools for allies and partners to organize and exchange free knowledge beyond Wikimedia. Our infrastructure will enable us and others to collect and use different forms of free, trusted knowledge.

Knowledge equity: As a social movement, we will focus our efforts on the knowledge and communities that have been left out by structures of power and privilege. We will welcome people from every background to build strong and diverse communities. We will break down the social, political, and technical barriers preventing people from accessing and contributing to free knowledge.²

The final language of the strategic direction adopted by the Wikimedia community acknowledges a world in which free knowledge is potentially plentiful but in need of critical support. It maintains the spirit of openness to all but recognizes the importance of building communities with shared purpose and good faith. We committed to undertaking this ambition informed by the guiding perspectives of “knowledge equity” and “knowledge as a service” as we seek to engage and include more perspectives from around the globe while ensuring Wikipedia is as dynamic and useful in the future as it is today.

It means that the popular idea of “Wikipedia” should be expansive and inclusive. When people hear “Wikipedia,” it should conjure up endless knowledge—one in which the articles of the encyclopedia are a point of entry into a rich, multilingual ecosystem of discovery; one which integrates rich annotations and citations, augmented and multimedia experiences, connections to external resources, complex insights, and robust linked open data structures. Wikipedia should be both a destination for learning and a network of exploration, connecting concepts, collections,

and institutions, elevating and interconnecting disparate resources of open knowledge, making all knowledge more accessible and discoverable.

This is Wikipedia the encyclopedia, of course, but also so much more. Neither the Wikimedia vision statement nor the 2030 strategic direction mention anything about a website or Wikipedia specifically. Instead, we recognize a goal whose intention is to enrich all of humanity and we cannot reach that alone.

Beyond the Encyclopedia

What does it mean to be the essential infrastructure of free knowledge? While “infrastructure” conjures up rigid and impersonal features, it is better understood as building the critical social, technical, and political support systems necessary to bear the ambition of a world in which free knowledge is produced and shared, not only in the Wikimedia ecosystem but also across many different communities, projects, and institutions.

It means supporting the people and institutions that produce free knowledge and championing the conditions that enable its production and dissemination. It is knowledge as a platform and also a community of creators, curators, advocates, donors, and allies around the globe. It is a body of knowledge and also a powerful voice that stands for the importance of free and open information, standards, policies, and practices in service of our public knowledge commons. It is a website and a movement which believes in the importance of the integrity of information and the fundamental right to inquire, learn, and seek answers.

Together, the people, technology, and voice of a movement make up the essential support system for free knowledge. Together, we will enable the collection, curation, and dissemination of free knowledge across the planet. Together, we can build the tools and infrastructure to host, catalog, tag, revise, and share knowledge, both in the Wikimedia ecosystem and far beyond. Together, we can extend the public domain and grow openly licensed content, advocate on behalf of knowledge-enabling policies and standards, and defend our essential rights to inquiry and information.

If this sounds radical, consider how far Wikipedia has already changed our conception of the encyclopedia: no longer a hardbound, finite, alphabetized collection of books but an infinite exploration of interconnected discovery and learning. The future of Wikipedia is an opportunity to extend

this evolution even further. It means embracing *encyclopedic* in an etymological sense, a circular, looping, endless education.

To realize this future, we will not only need to reconceive the encyclopedia (again) but also be open to the evolution of the Wikimedia projects and communities, perhaps in transformational ways.

Our global communities, well established in wealthy, northern countries, must grow to more fully represent the diversity of the world's languages, cultures, and contexts. Our underlying technology platform will need to be open and dynamic, able to integrate emerging and augmentative technologies and respond to as-yet unknown evolutions in devices, interfaces, and user experiences. The act of writing the encyclopedia may remain core to our identity but will need to be supplemented by other acts of collaboration, curation, and creation as well as new form factors for consuming and sharing knowledge.

Fortunately, the seeds of many of these changes have already taken root in the Wikimedia movement of today. In this sense, the 2030 strategic direction is less a radical re-envisioning of Wikimedia than a codification of emergent trends: growth of new communities in previously underrepresented languages and geographies, successful new projects focused on original sources and structured data, experiments in augmented machine-learning experiences, and new partners and allies in the movement for free culture.

The Experience

The next billion people to come online will come to Wikipedia through many devices and channels, so we must consider what it means to build beyond the desktop or mobile browser and anticipate a future in which people can access information across a host of devices and interfaces. For anyone to be able to join us in the work or exploration of free knowledge, we have to revisit the idea of what makes the Wikimedia experience meaningful, relevant, and useful as technology and the world change around us.

This is not only about the emergent needs of tomorrow but also the changing needs of users today who have different expectations for form factors, interactions, and user experiences. To stay relevant and relatable, Wikimedia must find a balance between retaining our identity and evolving to meet the world where it is and where it's going.

“Editing” Wikipedia

One of the most identifiable values of Wikimedia is the “read/write” nature of our projects. Anyone can be an editor, and any aspect of the projects is open to change. This has been core to Wikipedia’s model over the years, ensuring that as both knowledge and technology have changed, Wikipedia has been changing, too. It allows for articles to be quickly created, for editors to continuously refine and add nuance to complex concepts, for new ideas and new voices to challenge bias or add fresh perspective. It is a “consent or contest” paradigm, inviting everyone to be a critical reader and active participant in Wikipedia’s knowledge.

When this “open to change” model launched, it was at the cutting edge of participation on the internet. Today, people have very different expectations about the interfaces, tools, and experiences that they use to create content on the internet. For our open, participatory model to remain compelling, the experience of contributing to Wikipedia should evolve to feel as fresh, contemporary, and full of possibility to the next generation of contributors as it was to the very first Wikipedians.

There are many straightforward but important changes that could significantly improve the experience of contributing to the Wikimedia projects today. In 2018–2019, the Wikimedia Foundation’s Product department published an important series of essays on these changes, “Product Perspectives” on the subject of Wikipedia reader and editor experiences, as well as culture, scale, trust, augmentation, and tooling.³ These user experience essays explored the possibilities for new form factors, rich content, the needs of contributors, customized experiences, and improving how people discover and explore knowledge across the projects.

Many small evolutions to the Wikimedia interface could have a dramatic impact on the user experience and overall enjoyment of participating in the Wikimedia projects. Welcoming language, customized wayfinding, instructional onboarding, familiar contribution and consumption interfaces, suggested actions, and explicit gratitude and feedback are all areas where small changes could make significant improvements in the Wikimedia experience—and potentially improve the quality of new editor contributions and overall retention.

Wikipedia came to be in an era of desktop computing, and its long-form, detail-oriented knowledge production process has worked well in this context. The mechanical keyboard and screen setups of desktops and laptops

are designed for word processing and managing complex, long-lead tasks. A commitment to this form factor may have contributed to an initial reluctance and delay in embracing mobile devices as a possible new form factor for productive contributions. Yet despite some initial uncertainty about demand, improvements in Wikimedia's mobile editing interfaces and the introduction of more powerful mobile editing tools have proven very popular. In a world of mobile ubiquity, we see significant opportunities ahead, particularly for smaller, discrete "micro-contribution" tasks such as adding citations or image and data tagging.

Evolving the on-site editing experience and introducing new forms of contribution will require time, consideration, and care. But all of these actions are within a fairly well-understood cycle of "read, edit, publish, read," all of which take place directly on the Wikipedia websites. But what about editing Wikipedia content out in the wild, off the main websites? What would it mean to make an edit to Wikipedia-sourced information directly through Google's Knowledge Panel search result interface? How does an editable voice assistant interface sound and interact? What about navigating a contribution through an augmented or mixed-reality experience? The "editability promise," of being able to improve and update nearly any knowledge, nearly any time, is an essential part of who we are. For Wikipedia's future to stay true to Wikipedia's origins, we'll need to answer these questions.

"Reading" Wikipedia

Today Wikipedia takes the form of articles. Articles are primarily text with a smattering of images and data, and the way people interact with the information within them is by reading the text for insight. But is reading the only or even the best way to engage with the sum of the world's knowledge? What learning experiences will users seek on the Wikipedia of the future? In what ways will people "read" or interact with Wikipedia, off the Wikipedia projects? As the devices, interfaces, and demographics of internet users continue to evolve, we're already seeing changes toward multiplatform, multimodal knowledge-seeking behaviors.

Web searches for video content increasingly rival searches for text and image content. Younger users are more video forward, and newer users of the web are often navigating in second languages. Demand for digital services is also increasing among more nontraditional internet users, such as

the 750 million illiterate adults⁴ and many more millions of lower literacy individuals seeking access to opportunity and entertainment. Video can offer immersive learning experiences that may be more engaging than text, whether for reasons of accessibility, literacy, or practical demonstration.

For some time now, Wikipedia editors have been proposing the introduction of video and other rich media experiences to Wikipedia. While Wikimedia Commons has seen recent renewed growth as a freely licensed media repository, its usability lags behind other media hosting sites, and it remains primarily a service for images rather than rich media. And multimedia poses interesting questions for Wikimedia editors: What does it mean to collaborate on a video? What constitutes a neutral point of view in photography? To be successful, the projects will have to address practical issues of hosting, discoverability, and usability as well as more philosophical questions around neutrality and editorial voice.

Whether on the web, through talking to your phone's digital assistant, or in an augmented reality headset, people increasingly expect digital experiences to anticipate their needs and interests. Wikipedia is unlikely to ever embrace "personalization," in which the platform collects private data and uses it to make opaque algorithmic recommendations. This would be against our values of privacy as well as be counter to Wikipedia's spirit of discovery and serendipity. However, we could embrace "customization," empowering people with tools to curate their experiences. Integrated alongside increasingly powerful tools for knowledge discovery, the Wikipedia of the future is one that informs, delights, and compels.

While the questions of video and other multimedia experiences, personalization, or discovery are largely related to what happens on the Wikimedia projects, Wikimedia's knowledge experiences are rapidly evolving off-projects as other platforms and brands repurpose and customize Wikimedia's content to meet the needs of their users.

Recent years have seen a major increase in the utility and adoption of voice assistants, interfaces which combine voice recognition, natural language processing, and speech synthesis to produce a conversational query service. Some researchers anticipate that voice-based search queries could overtake text-based queries by 2020.⁵ Already, Wikipedia results populate the major voice assistant services, answering contextual and factual questions on the phone or around the kitchen table. Voice-based search creates opportunities for new ambient learning behaviors, giving people the ability

to seek information as they go about their lives, expanding augmentative information-seeking behaviors.

Voice search gives us a preview of more sophisticated voice services, with the potential to address major accessibility, literacy, and other structural barriers. Voice services can increase usability and availability of key services for those with visual impairments. Natural language voice interactions can help low familiarity or fluency speakers quickly clarify search queries and adjust results. For languages that are primarily spoken rather than written, that are not widely digitized, or that lack localized product and user experiences, verbal services offer a “leapfrog” opportunity to surpass existing limitations and address previously underserved groups.

Voice reuse of Wikipedia is just one of the ways in which Wikipedia content may be explored or “read” in the future. Already, artists, engineers, and innovators use Wikimedia’s content, data, and media to build augmented and mixed-reality experiences, annotate and enrich mapping platforms, and build virtual galaxies of knowledge.⁶ It is impossible to predict the other ways in which interfaces, devices, and experiences may yet evolve. Instead, the Wikipedia of the future should anticipate new ways in which people will experience and contribute knowledge. It should serve knowledge to the world across interfaces and communities and enable everyone to collect, organize, and exchange knowledge that fully represents human diversity. To do this, we’ll want to retool the technical platform to be more open and flexible—a process that is already underway.

The Platform

Wikipedia is best known to people through its content and interfaces, but to embrace an expansive and inclusive future of knowledge, we should also look to the underlying technical platform that supports how we collect, curate, and share knowledge. The stability, resiliency, and flexibility of the technical “stack” sets the parameters for the possibilities of what we can build: how fast, how flexible, how dynamic, how accessible, how secure. Today we have an opportunity to be intentional about how this platform can itself be a tool in advancing the 2030 vision.

Wikimedia’s technical stack has been affectionately described by the Wikimedia Foundation’s technical staff as having a “just-in-time” architecture, one which evolved in response to Wikipedia’s growth in content and

popularity. While this refers to the sometimes ad hoc nature of solutions which emerge when running a major global website with relatively limited resources, it also nicely captures how Wikimedia content principles have informed the technical systems on which it runs: many small workable hacks have proven more practical and serviceable than elegant but impractical perfection.

This approach has allowed volunteer and paid developers to solve emergent problems, adapt to new use cases, and introduce increasingly sophisticated tools and services, all while (mostly) keeping the projects up and running. However, the core of the platform very much reflects its origins in the static HTML internet of the mid-2000s. For the future of Wikimedia and the future of free knowledge, the platform needs some intentional updates. From multidevice and multimedia capabilities to support new forms of knowledge, to more complex integrations of machine learning and distributed security, there is an opportunity to more fundamentally reimagine our stack to integrate the innovations of the last decade.

New technologies around censorship resistance, privacy, and availability need to be integrated to protect our users and content in the more hostile internet of today. As governments and malicious actors increase pressure on free knowledge, improved performance, security, and resiliency of the technical stack help serve as a bulwark to protect the integrity of the knowledge and data within the Wikimedia ecosystem, the privacy of editors and readers, and the availability of the sites themselves.

The way in which the platform stores, updates, and serves knowledge today works well for the article model of Wikipedia, but it introduces some limitations on more dynamic reuse of knowledge. Currently, edits are saved and presented as a single body of unstructured text. This works well for reading and editing whole articles, but it makes it difficult to isolate specific data or insights. Imagine if we were instead able to deconstruct the information in articles or attributes of a media file and store them in structured and semantically understandable units. These would not only be meaningful within Wikipedia, they would also have meaningful value as independent, parsable units of knowledge within the broader digital knowledge ecosystem.

Building structured units of knowledge addresses part of this future: they would be more easily reused and remixed into new experiences and interfaces, more easily syndicated by other knowledge services, and more easily

updated in more modular fashions. Alongside this more modular content, we have the opportunity to evolve structured citations accordingly—enabling structured, rich knowledge that can present information in context and, sometimes, in conflict. Presenting multiple claims and highlighting conflict and congruence while incorporating modular verification information can enrich our understanding of topics and the authorities that give them weight.

The transition from unstructured to structured knowledge is directly connected to another goal of platform evolution: the move to a more service-oriented architecture. Ideally, the openness of the platform should mirror the openness of the content—a platform that makes it easy for anyone, from community members to external knowledge partners to the Wikimedia Foundation itself, to easily build and create new experiences and services on top of Wikimedia.

Today, building something off of Wikipedia’s knowledge base requires immersing oneself in mailing lists and documentation and navigating the full stack or bringing a brute-force application of computational and engineering power to scrape sense from the wilderness. Imagine an alternative: still a fully open stack with a structured knowledge corpus and architectural coherence, built with contemporary libraries and languages, offering an accessible, usable, and well-documented service layer. Sophisticated technical contributors could continue to participate in governance and development of the platform’s architecture, performance, security, database schemas, and other technical matters. Major reusers could access the “firehose” of dynamic database updates. Individual creators and companies could build new services and insights.

Of the many possibilities for platform evolution, another significant evolution is the growing use of machine learning and artificial intelligence on the Wikimedia projects. Wikimedia has relied on machine augmentation since our beginnings—there are dozens of bots that operate on Wikipedia, performing various routine functions so that humans don’t have to. Machine intelligence already assists editors in evaluating edit and article quality and providing rough translations of articles between various languages. We anticipate that this use of machine learning and artificial intelligence will continue to grow, although in ways that may remain largely invisible to a casual user.

In the future, we expect machine learning to help readers and editors alike. It could enable communities to grow the projects even more ambitiously, synthesizing and syncing knowledge across languages, highlighting content gaps and bias, identifying under-loved articles, automatically translating between Wikipedia language versions, or using natural language generation to create “stubs” or rough drafts of articles from collections of secondary sources. For readers, it could assemble customized reading lists, propose contextual article groups for deeper learning, recommend related images and media, and more.

In addition to growing Wikimedia’s quality and quantity, machine learning may also prove to be a valuable tool addressing an area of great concern on the internet of 2020: misinformation and manipulation. Tools that assist Wikipedia editors in recognizing suspicious clusters of contributions, networks of editors, spikes in editing activity, or patterns of words and linguistic signifiers could be possible applications. Of course, the best tools would be additive to the existing approaches Wikipedia editors use to maintain knowledge quality and identify bad-faith activity. They should build on existing efforts, supporting the work and intentions of the people who contribute to Wikipedia.

The function, development, and deployment of machine learning on the Wikimedia projects should abide by what Wikimedia researchers and engineers have adopted as the principles of ethical Wikimedia artificial intelligence: fairness, transparency, and accountability.⁷ Volunteers and staff working on these efforts today envision a future where Wikipedia offers both tools and a learning environment for contributors to “train the machines” so that our artificial intelligences are as distributed, accessible, and open as any other part of the Wikimedia ecosystem.

The Wikimedia technical platform of today is already a remarkable achievement. With limited resources and developer time, it reliably serves hundreds of millions of users and billions of page views every single month to all corners of the planet. The future of Wikimedia offers this platform an opportunity to evolve from a supporting function to a strategic one, offering new knowledge formats, structured knowledge, new service layers, federated knowledge hosting, and augmented intelligence—an indispensable piece of infrastructure for the essential support system for free knowledge.

The People

As much as the technology and user experiences of the Wikimedia platforms shape our work, the volunteer community is what truly differentiates the Wikimedia movement and mission. In survey after survey, we find that Wikimedians contribute to Wikipedia and the other Wikimedia projects because they are animated by the promise of the mission of free and open knowledge for the world.

The community that sustains Wikipedia today has built something remarkable and unprecedented in the world, and they deserve celebration and continued support and recognition. And if we believe that the world is better when more people can share in free knowledge and that this can only happen when more people openly collaborate with one another, we must recognize who is still missing from the picture.

We find a stark example of the unevenness of the Wikimedia community in looking at Wikipedia contributions around the globe. More people from the country of the Netherlands, with a population of around seventeen million people, contribute to Wikipedia than people who reside across the entire continent of Africa, home to more than 1.2 billion people. Another way of looking at this imbalance in representation? Articles about the continent Africa, the cradle of humanity, home to more than fifty countries, thousands of languages, and thriving modern cities, represent fewer than 4 percent of all of the geotagged articles on English Wikipedia.⁸

It is not difficult to infer that the authors of these relatively few articles about African topics are statistically unlikely to be from Africa themselves, conjuring up a parallel world in which every article about Europe is written primarily by Latin Americans and every article about North America written primarily by South Asians. Of course, Wikipedia's articles should be written by people from all over, with space inclusive of many different perspectives. If only some people represent all people, we lose out on a more holistic and representative collection of knowledge.

Across the globe, this problem persists. Wikipedia's editor populations are stable or growing where real populations are not while editing communities in regions and countries that are experiencing rapid population growth remain relatively small. Hindi Wikipedia, representing the world's third-most spoken language, is only the fifty-third-largest Wikipedia, far

smaller than languages with far fewer native speakers, such as Catalan and Finnish, or none (Esperanto).

Of course, no conversation about the need for increased contributor diversity in Wikimedia would be complete without acknowledging the lack of equitable gender participation in the Wikimedia communities. Often described as the “gender gap,” the best available statistics about gender participation in the Wikimedia communities indicate that less than 20 percent of Wikipedia contributors identify as women. (The idea of closing the “gender gap” itself has always struck me as somewhat problematic as it implies a gulf between two equivalent sides and reinforces the idea of binary gender. An aspiration to equitable “gender diversity” might be more fitting.)

As of publication, there are roughly three hundred language versions of Wikipedia, ranging in size from very large and very active (English, currently the largest, at six million articles) to relatively small (Tulu, currently the smallest active Wikipedia, with about one thousand articles). Although opinions differ within the Wikimedia movement about whether we should strive for a Wikipedia in every language, it is certainly the case that more can be done to address knowledge parity across the Wikipedia language versions that exist today, whether through supporting organic community and article growth, optional machine translation, or even natural language generation.

If we believe that a better informed world is a world of greater understanding, better equipped to address our collective challenges, then the Wikimedia movement must reorganize ourselves to acknowledge who is not yet a part of our vision. If languages, regions, and people are missing and if contributors to the Wikimedia projects are still predominantly male, still predominantly from North America and Europe, and still predominantly white, then we must agree that our pursuit of free and open knowledge is definitionally incomplete. If Wikipedia editors don’t represent the world’s diversity, they cannot fully contribute to a representation of the world’s knowledge.

The imperative of building a more representative and welcoming community extends beyond our own projects. In a future in which Wikimedia’s knowledge extends well beyond the Wikimedia projects, a more inclusive and diverse Wikipedia editing community has positive repercussions that extend well beyond their participation on Wikimedia sites. Diversity of perspective not only enriches Wikimedia content, culture, and governance,

it enriches the broader open knowledge ecosystem in which Wikimedia exists. It allows us and others to build more representative knowledge, more valuable and appropriate user experiences, and less biased machine learning models.

The Wikimedia 2030 direction recognizes the work of Wikimedia as a social movement and calls on us to center knowledge and communities that have been left out by structures of power and privilege. To truly do so, we must welcome people from every background and invest resources in breaking down the social, political, and technical barriers preventing people from accessing and contributing to free knowledge. This will support not only strong and diverse communities but also robust, resilient, and representative knowledge.

Conclusion

A world in which every single human can freely share in the sum of all knowledge is an impossible, asymptotic vision. This is because knowledge itself is always changing with every passing day, discovery, and reconsideration. Wikimedia can only ever get closer, but it will never be finished. The Wikimedia movement's direction for 2030 acknowledges that there is a step we can aim for along the way: supporting the people and institutions that produce free knowledge, building the technology that hosts and distributes it, and standing up for the policies and practices that make this work possible. This is how Wikimedia evolves while staying true to the values from whence it came.

We will be able to make this journey from the “encyclopedia anyone can edit” to the “essential infrastructure of free knowledge” because change is at the core of what Wikimedia is and does. Our capacity to change is Wikimedia's greatest strength. Every day, Wikipedia changes moment by moment in response to the global flow of knowledge. It changes to reflect current events, it changes as we learn new things about our world and ourselves, it changes as new voices enter the conversation and challenge dominant paradigms. It changes when there are typos and grammatical errors. It changes when new technologies demand new experiences. It changes because humans are often wrong, and we can only ever seek to improve.

Over the past twenty years, the Wikimedia community has managed to build an unintentionally ubiquitous resource, an integral piece of the fabric

of modern society. But the continued growth and resilience of free knowledge is not an inevitability. History is a story of knowledge as scarcity, a tool of power and privilege. By comparison, the very idea of free knowledge is radical, subverting practices of gatekeeping that date back centuries. A belief in free and open knowledge is a belief that everyone has the potential to contribute to the world and that everyone has the ability to use knowledge well and wisely. Paraphrasing Wikimedian Marco Correa, former president of Wikimedia Chile, “The knowledge is neutral. The mission is not.”

I’m reminded of a Wikimedia story, probably apocryphal, about a moment when the first edition of a major editorial encyclopedia first entered into the public domain. As the story goes, Wikipedia editors were thrilled—this newly available knowledge could now be fully integrated into Wikipedia, improving the accuracy and breadth of Wikipedia’s own articles. Yet, when editors went to review what they could salvage, they found almost nothing. From Aristotle to zebras, our understanding of the world had kept evolving, and static knowledge had not kept up.

For Wikipedia to be as essential in the next twenty years as we have been for the first twenty, we will need to keep up with our world. We will need to make the changes that make Wikipedia easier to contribute to and more delightful to read. We will need to retool our platforms to support and serve knowledge on our sites and across the digital landscape. We will need to change to include more voices, more demographics, more diversity, more languages. We will need to change to recognize and respond more robustly to the risks and threats to free knowledge and its creators. We will need to embrace our role as the essential support system of free knowledge.

Wikimedia reminds us that the greatest thing we will ever build is the thing which we build with others. It reminds us that the world we seek is a work in progress and that each of us has a role to play in how we improve what lies before us. In 2001, we had no idea of what we could build together. I look forward to what the world will build next.

Notes

1. Ed Erhart, “‘My Favorite Website’: Stephen Colbert’s Complicated Relationship with Wikipedia,” Wikimedia Foundation, September 8, 2015, <https://blog.wikimedia.org/2015/09/08/colberts-complicated-relationship-with-wikipedia/>.

2. Meta-Wiki, s.v. "Strategy/Wikimedia Movement/2018-20," accessed October 11, 2019, https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2018-20.
3. MediaWiki, s.v. "Wikimedia Product/Perspectives," accessed October 10, 2019, https://www.mediawiki.org/wiki/Wikimedia_Product/Perspectives.
4. Manos Antoninis and Silvia Montoya, "The World Is Off Track to Deliver on Its Education Commitments by 2030," *Data for Sustainable Development* (blog), July 9, 2019, <https://sdg.uis.unesco.org/2019/07/09/the-world-is-off-track-to-deliver-on-its-education-commitments-by-2030/>.
5. Danielle Antosz, "Mary Meeker's 2016 Internet Trends Report," *Search Engine Journal*, June 6, 2016, <https://www.searchenginejournal.com/5-search-related-takeaways-mary-meekers-2016-internet-trends-report/165434/#close>.
6. The Wikiverse Project: An Interactive 3D Visualization of Wikipedia, accessed October 10, 2019, <https://www.wikiverse.io/>.
7. Jonathan T. Morgan, "Designing Ethically with AI: How Wikimedia Can Harness Machine Learning in a Responsible and Human-Centered Way," Wikimedia Foundation, July 18, 2019, <https://wikimediafoundation.org/news/2019/07/18/designing-ethically-with-ai-how-wikimedia-can-harness-machine-learning-in-a-responsible-and-human-centered-way/>.
8. Mark Graham, "The Geographically Uneven Coverage of Wikipedia," Oxford Internet Institute, October 21, 2014, <https://geonet.oii.ox.ac.uk/blog/the-geographically-uneven-coverage-of-wikipedia-2/>.

This is a section of [doi:10.7551/mitpress/12366.001.0001](https://doi.org/10.7551/mitpress/12366.001.0001)

Wikipedia @ 20

Stories of an Incomplete Revolution

Edited by: Joseph Reagle, Jackie Koerner

Citation:

Wikipedia @ 20: Stories of an Incomplete Revolution

Edited by: Joseph Reagle, Jackie Koerner

DOI: 10.7551/mitpress/12366.001.0001

ISBN (electronic): 9780262360593

Publisher: The MIT Press

Published: 2020

The open access edition of this book was made possible by generous funding and support from Knowledge Unlatched, Northeastern University, and the Wikimedia Foundation



The MIT Press

© 2020 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC BY-NC 4.0 license.
Subject to such license, all rights are reserved.



The open access edition of this book was made possible by generous funding from Knowledge Unlatched, Northeastern University Communication Studies Department, and Wikimedia Foundation.



This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Reagle, Joseph, editor. | Koerner, Jackie, editor.

Title: Wikipedia @ 20 : stories of an incomplete revolution / edited by
Joseph M. Reagle and Jackie Koerner.

Other titles: Wikipedia at 20

Description: Cambridge, Massachusetts : The MIT Press, [2020] |

Includes bibliographical references and index.

Identifiers: LCCN 2020000804 | ISBN 9780262538176 (paperback)

Subjects: LCSH: Wikipedia--History.

Classification: LCC AE100 .W54 2020 | DDC 030--dc23

LC record available at <https://lccn.loc.gov/2020000804>