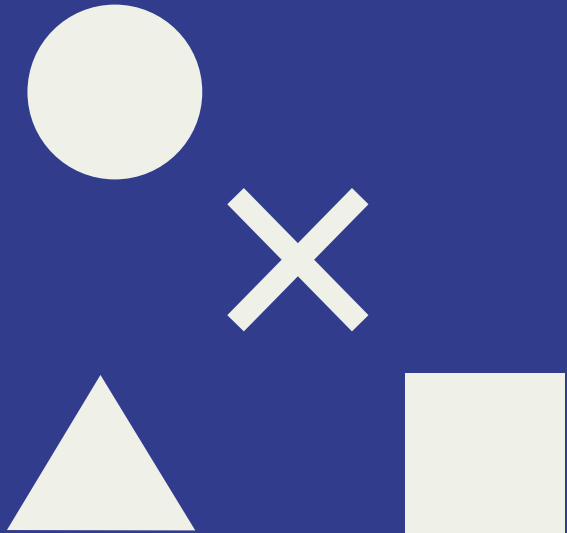


Culture

This collection of essays on language, content and inclusive culture are based on insights from the Communities and WMF staff, as well as synthesis of secondary research. These positions will be used to inform product strategy over the next three to five years.



Published
January 2019

Authors

Charlotte Gauthier

Rita Ho

Runa Bhattacharjee

Josh Minor

Contributors

Amir Aharoni

Adam Baso

Volker Eckl

David A.M. Goldberg

Danny Horn

Ramsey Isler

Jon Katz

Toby Negrin

Margeigh Novotny

Nirzar Pangarkar

Santhosh Thottingal



In many ways, the theme of Culture synthesizes the movement's strategic directions of Service and Equity. Knowledge as a service means enlisting allies and partners to make knowledge available outside the confines of Wikimedia - and that must include institutions and form factors in which the majority of the world's knowledge still resides, un-digitized and unavailable to a public who might not even know it exists. This means we have to think outside the existing projects and begin acting as a platform for knowledge organization, dissemination and exchange, and as a catalyst for other organizations and institutions engaged in knowledge dissemination. This will require new cultural as well as technical competencies.

Knowledge equity means removing the barriers preventing anyone from sharing what they know, and encouraging "respectful collaboration" between people. However, it is not only historical "structures of power and privilege" that have prevented people from sharing knowledge via Wikimedia. Focusing on specific forms of diversity that come from mainly an American context is likely to hinder rather than help the movement's global inclusiveness. Rather, we need to meet people where they are - in both the literal and figurative senses. A coherent culture strategy must start with awareness, for both consumers and contributors. Research has indicated that awareness of Wikipedia is low in emerging markets,[1] and that the fact that it is editable might actually work against some of our other priorities in the short-term, such as engendering trust.[2] We need, therefore, to find an approach in underserved communities that works - not assuming that people in these communities share our values or that the movement, as it is, can smoothly be applied to new cultural contexts. It is not certain that new users from emerging markets will necessarily share in the movement's goals right away, if ever. If we really mean all the world's knowledge, that means stepping outside our comfort zone.

In addition to fulfilling our mission of making all the world's knowledge available to everyone on our own projects, we also have the opportunity

to make a lasting positive impact on the world's free culture by pursuing (or enabling others to pursue) other societal goals, such as the preservation of endangered languages.[3] Where we can identify areas our movement goals overlap with other free culture efforts, we should partner with the people and institutions pursuing them.

Examples

Content Translation
GLAM outreach and strategy
Structured Data on Commons

Areas of Impact

Most Wiki projects (but especially Wikipedia, Commons, Wiktionary, and Wikisource)
Community Relations
Communications
Mobile (both apps and mobile web)
Community policies and guidelines
Research
Disabled or Disenfranchised Communities

Key External Factors

Between 53% [4] and 71% [5] of the world's population will be online by 2030, with the growth mainly driven by emerging markets. The overall growth rate is slowing [6] due to saturation in developed markets.

Notes

- 1 Movement Strategy Executive Summary of Brand Awareness https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Brand_awareness,_attitudes,_and_usage_-_Executive_Summary
- 2 Finding from New Readers research in Nigeria, India and Mexico “Trust in Wikipedia is shaken when people find out anyone can edit pages.”
- 3 https://meta.wikimedia.org/wiki/New_Readers/Findings/India#Theme:_Using_Wikipedia
- 4 <https://wikitongues.org/>
- 5 <https://blog.euromonitor.com/2015/04/half-the-worlds-population-will-be-online-by-2030.html>
- 6 The Mobile Economy Global <https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/02/The-Mobile-Economy-Global-2018.pdf>
- 7 USA Today: Smartphone Sales Have Hit a Wall

01

Inclusion

What do “diversity” and “inclusion” mean in the context of software products? Diversity in this context is in reference to having a variety of social stratifications within a collective, namely class, race, sexual orientation, age, (dis)ability and gender. An important aspect of diversity to consider is the variety of ideological viewpoints that inform the types of knowledge considered for inclusion. Inclusion is recognizing the need for greater diversity in Wikimedia addressing issues of content contribution, consumption and policy and decision-making. The web is fundamentally designed to work for all people, whatever their hardware, software, language, location, or ability.

[1] Ultimately, the drive for inclusion of different contributors, beneficiaries and movement leaders must be to ensure diversity of knowledge. That is, broadening perspectives of the various forms of content that we consider as valid and valuable part of the knowledge space, while at the same time ensuring this broad range of content is accessible to all.

Why diversity matters

Besides being one of the core values of the Wikimedia Foundation,[2] diversity is critical to the goal of making all knowledge available to all people, and is an important factor in ensuring high quality content. Direct examples of diversity directly improving knowledge are demonstrated in two studies from 2017. The first study showed that Political, Social Issues, and Science articles whose editors were comprised of more diverse political viewpoints (across Democratic & Conservative) were of higher quality than those with politically homogeneous editor groups. [3] Similarly, a separate second study noted that language-specific topics on Wikipedia are generally better quality on relevant language-editions of Wikipedia, and could be leveraged to improve the knowledge base across language editions.[4]

What's holding us back?

Conceptual barriers/issues

Currently, we talk about diversity and inclusion as an ongoing pursuit, without a, clear end goal. Resource constraints and biases as identified below mean that there is continuing work needed to help recognize and include marginalized groups. But it is hard to reach consensus as to which groups are of higher or lower importance/need and what goals should be universal (for example, the number of articles of a particular language). Without clear targets for manifesting diversity and inclusion it will be difficult to progress effectively.

Individual biases (conscious and unconscious/implicit) and systemic bias challenge our capacity to address diversity and inclusion. An example of individual bias would be the notion of notability of BLP [5] articles being assessed based on criteria that often leads to underrepresentation

of certain groups. Systemic biases describe processes and organizational structures that exist within social-cultural groups and lead to decisions that under-represent other groups. And finally, there's a lack of awareness between communities, in the sense of of "peer" communities that use or contribute to Wikipedia (i.e. Group X doesn't know, trust or feel welcomed by Group Y), and those defined by the border between the Foundation and the communities it serves (i.e. Community X doesn't understand how the Foundation itself, or something the Foundation is proposing, is a benefit to them).

Internal organizational/resource constraints

There are gaps in making content accessible to everyone: at a high level, web content is considered accessible when it is perceivable, operable, usable, and robust. [6] Therefore there should be resources that help providers meet Web Content Accessibility Guidelines [7] and tools for readers and contributors with accessibility needs.[8]

There are also gaps in content across languages and many other factors of diversity. This is an area where the Foundation has been reluctant to intervene directly, out of respect for volunteer editorial control. But it is clear that a totally community driven approach, supported only by grants, is insufficient for filling content gaps effectively.

Finally there are constraints associated with human capital. People and their volunteer time are also a limited resource. There are trade-offs to be considered if we are to truly prioritize greater diversity and inclusivity without diminishing or diluting the overall impact of the movement.

External factors

Internet and technology access is inequitable depending on geography, socioeconomic and other demographic factors. This is consistently reported in multiple studies, including a recent

report from the UN noting that less than 50% of the world's population are online as of 2018, and there is a trending decline in internet growth rates.[9]

There are cultural factors as well. Certain communities may not want to participate, share, or be “served” for their own individual reasons and interests. A couple of Wikimania 2018 sessions touched upon some of the reasons Communities fail to engage or disengage - from differences in communication norms (e.g., communities which have a stronger oral tradition,[10][11] distrust due to a history of being marginalized,[12] to fundamental differences in cultural conceptions of what is an appropriate channel for knowledge-sharing.[13]

Reducing barriers to inclusion

Correcting for bias within the Foundation or Community

We must prioritize the closing of gaps in representation at higher leadership and decision-making levels within the Movement. This will help reduce systemic and unconscious bias, but also encourage greater participation from the top down.

We should establish some commonly agreed upon definitions or baseline for measurement. For example, assessing our success in making knowledge accessible to every person is dependent on how we define knowledge, diversity, and what it means to be accessible. [14] Broadening measurement and research means both including currently identified underrepresented voices [15] and understanding what ways they feel excluded from participation, but also continuing to research where there are still existing unidentified gaps in representation.

Improving lines of communication

Besides representation, one of the ways to reduce the conceptual biases of Communities is for the Wikimedia Foundation to facilitate better lines of communication, in general, across groups. Part of this involves continuing research to show benefits of diversity in advancing knowledge quality as well as equity (see above “Why diversity matters” section) and showing these are not conflicting, but mutually reinforcing aims.

Another parallel effort is the continuous investment in tools and programs that encourage civil discussions and foster more positive relationships in our communities. This relates to our work in anti-harassment tools,[16] UI standardization's focus on accessibility and initiatives that strive to provide better help and support for new members (e.g. New Readers [17] program, and the Growth Team).[18]

Broadening representation of the movement is not only important in reducing barriers to inclusion but is also a clear signal to newcomers that the Movement is a welcoming place.

Recognizing that Wikimedia has limited resources to provide, and some factors are outside of our control, we should explore strategic partnerships that optimize our reach. Partners might include local experts and embedded members of a particular community with more access to content, people and support; Also, GLAM institutions who may be interested in preserving collections without taking on the burden of maintenance; distribution platforms including search (e.g. Google,) news media organizations (e.g. NYTimes,) and social media sites (e.g. Twitter and Facebook); and finally education groups such as MOOCs and other institutions to adapt Wikimedia content to newer learning courses tailored to helping improve digital literacy in marginalized communities.

Technology and tools

We can utilize technology to both attract more voices as well as help to fill content gaps. Content tailored to newer methods and behaviors of consumption, sharing and learning will enable more people in low reach areas to become aware of and begin accessing our knowledge repository. As noted in “Brand awareness, attitudes, and usage”,[19] awareness and familiarity with Wikipedia is a main area of concern. Some areas where we can invest to increase awareness of the overall Wikipedia brand to broaden reach in low awareness groups (eg. those with limited internet access, younger users) include: optimization and improvement for mobile content in terms of design and delivery and improving sharing via social media platforms such as Facebook and messaging platforms such as Whatsapp, WeChat, and Telegram. Such recommendations address countries like Cambodia where 30% of users access their information via Facebook,[20] and the greater diversity of mobile messaging apps outside of the United States where Internet and technology access are restricted.

Improving contribution tools

It is important to recognize and remember that Wikipedia is a participatory tool. We should invest in new technology and tools that promote and attract new contributing voices. Another recent paper, “The Pipeline of Online Participation Inequalities: The Case of Wikipedia Editing”[21] highlights participation is limited earlier in the ‘pipeline’ of possible editors based on income and racial biases, then by technical knowledge, and finally a gender gap in awareness of its participatory nature.[22]

With that in mind, some broad recommendations for improving contribution tools address issues of accessibility, diversity in technical skills, and the ever-increasing emphasis on the mobile space. Other content consumption platforms

such as voice-interfaces also benefit from simpler interfaces and underlying APIs.

We should ensure that our products conform to accessibility standards and guidelines to work for all people (so, for example, vision impaired users are not excluded from reading and contributing). There are similar overlaps in the needs of in developers, moderators and organizers when we consider mobile compatible tools, and easy-to-navigate systems for recommending content needing contribution or directing new editors to key areas, via, for example “micro-contributions” tools.[23][24]

Programmatic Initiatives

These potential software improvements are one way we can continue to increase contribution, and specifically targeting marginalized or underrepresented communities. But this work must be paired with outreach efforts such as in-person, off-wiki awareness events to recruit newcomers, edit-a-thons to expand content in identified underserved topic areas, and robust community organizing and capacity building programs. [25]

Notes

- 1 <https://www.w3.org/WAI/fundamentals/accessibility-intro/>
- 2 https://meta.wikimedia.org/wiki/Values/2008#Wikimedia_Foundation_values
- 3 <https://arxiv.org/abs/1712.06414>
- 4 <https://www.mdpi.com/2227-9709/4/4/43>
- 5 https://en.wikipedia.org/wiki/Wikipedia:Biographies_of_living_persons
- 6 WCAG “POUR” Accessibility principles - <https://www.w3.org/WAI/fundamentals/accessibility-principles/>
- 7 <https://www.w3.org/TR/WCAG21/>
- 8 We are striving to add reader tools as well, e.g. dark mode for reading Wikipedia.
- 9 Sample, I. (2018-10-18). “Exclusive: dramatic slowdown in global growth of internet access”. the Guardian. Retrieved 2018-10-21. <https://www.theguardian.com/technology/2018/oct/18/exclusive-dramatic-slowdown-in-global-growth-of-the-internet>
- 10 Wikimania 2018 presentation: “Wikipedia and Bhutan can learn from each other” <https://www.theguardian.com/technology/2018/oct/18/exclusive-dramatic-slowdown-in-global-growth-of-the-internet>
- 11 Wikimania 2018: presentation: “The quotation of oral sources in a decolonization context” https://wikimania2018.wikimedia.org/wiki/Program/The_quotation_of_oral_sources_in_a_decolonization_context
- 12 Wikimania 2018 Panel: “Centering Knowledge from the Margins: A Whose Knowledge? discussion” https://wikimania2018.wikimedia.org/wiki/Program/The_quotation_of_oral_sources_in_a_decolonization_context
- 13 Wikimania 2018 presentation: “Wikipedia for Indigenous Communities” https://wikimania2018.wikimedia.org/wiki/Program/Wikipedia_for_Indigenous_Communities
- 14 UNESCO’s Fostering inclusive knowledge societies report particularly focuses on the importance of definitions and measurements for what it means to be inclusive, have access to knowledge, etc; and stresses the need for establishing common definitions so their members states can aim for common goals. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/internet_draft_study.pdf
- 15 See Brand Awareness, Attitudes, And Usage - Executive Summary https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Brand_awareness,_attitudes,_and_usage_-_Executive_Summary
- 16 https://www.mediawiki.org/wiki/Anti-Harassment_Tools
- 17 https://meta.wikimedia.org/wiki/New_Readers
- 18 https://meta.wikimedia.org/wiki/New_Readers
- 19 See Brand Awareness, Attitudes, And Usage - Executive Summary https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Brand_awareness,_attitudes,_and_usage_-_Executive_Summary

20 “In 2016 Internet/Facebook became the most important channel through which Cambodians access information (30%) — surpassing TV (29%) and almost doubling radio (15%)” – from “Mobile Phones and Internet Use in Cambodia 2016” http://www.open.org.kh/research/phones_2016.pdf

21 Shaw, Aaron; Hargittai, Eszter (2018-02-01). “The Pipeline of Online Participation Inequalities: The Case of Wikipedia Editing”. *Journal of Communication*. 68 (1): 143–168. doi:10.1093/joc/jqx003. ISSN 0021-9916. (but still available via archive.org)

22 Already identified underrepresented groups are those discussed in the references, including but not limited to: groups in other Western-European regions (per “Geographies of the world’s knowledge”), Women and Non-binary (per “Gender equity report 2018”, and those in areas with limited internet access (per “UNESCO’s Fostering inclusive knowledge societies report”). https://www.oii.ox.ac.uk/archive/downloads/publications/convoco_geographies_en.pdf, https://meta.wikimedia.org/wiki/Gender_equity_report_2018, http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/internet_draft_study.pdf

23 https://www.mediawiki.org/wiki/Mobile_design/Micro_contributions

24 https://www.mediawiki.org/wiki/Reading/Readers_contributions_via_Android

25 Hargittai, E., & Shaw, A. (2015). “Mind the skills gap: The role of Internet know-how and gender in differentiated contributions to Wikipedia.” *Information, Communication & Society*, 18(4), 424–442. doi: 10.1080/1369118X.2014.957711

02

Language

This document is an exclusionary act. Conceived and written in English by highly literate speakers of English, the world's most dominant and well-represented system of knowledge. But Wikimedia's vision aims to be for all people, in their language, calling for us to break down the exclusionary systems of knowledge creation and access. To do this, our systems, software and governance will have to evolve away from an English first, others maybe model, to a system that not only supports all languages but empowers them. For 60% of the world [1] using multiple languages, for different purposes, is the normal way to live. Pidgins and inter-languages produce knowledge too, so how must the current one-project-one-language model adapt to support evolving languages such as Hinglish and Spanglish? And while capturing and growing

knowledge in all languages is an uncontroversial goal, it is at these fuzzy edges that we can see, for example, the unintended exclusionary potential of technologies such as machine translation. As isolated cultural "space" collapses, a tension arises between support for cultural uniqueness and support for the experience of living between languages. And finally, language acquisition and cultural adaptation will be a huge factor in the foundation's success – what new capabilities does this demand of our organization and of the movement as a whole?

Privilege

Recognizing and Addressing our English first model

How does culture (both Foundation/Wiki community culture and the cultures of the underserved language communities) play into what languages and/or topics are “marginalised”? As a first step toward answering this question, the Movement must develop a pragmatic model for mapping similarities and differences between language cultures. Only with such a model can there be understanding of where the common ground exists between language cultures. Languages, and the facility with which one moves between them, may not be the connection in all cases, but can be a catalyst.

It is understood that there will be predominance of some languages due to various factors like demographic majority, regional prevalence, social prestige, or academic needs. However, isolated cultural spaces are collapsing thus more languages are going to co-exist in the same space. This will need more case based exploration.

Where to Focus?

Scoping Our Approach

Language is a vast source of information and influence. Small and large impacts arising from languages are happening around us continuously (from misuse of words that give rise to misunderstanding, to re-purposing or invention of words to describe new cultural phenomena, to policy decisions about what languages will be used in education). There are many variables that factor into how language use changes over time, and tracking all of them would be a heavy investment without obvious return. We must therefore identify the key factors that have the strongest relationship to knowledge gathering (e.g. language population migration) and develop mechanisms for monitoring these factors over time.

For the purposes of our our strategic planning it will be necessary to identify a few areas of high

potential impact, and use these to center our technical and social interventions.

Priorities

How do language interventions get prioritized? What should drive this?

Language interventions must be prioritized in response to social factors. Tracking cultural shifts due to economic migration and other global factors will be key to understanding the adaptive linguistic uses that will impact our platform. Forced migrations from conflict and climate change impact zones, and economic migrations for subsistence and specialized jobs. Related to migrations are efforts at cultural preservation triggered by cultural overlaps, friction and intermingling.

Improved access to education, increased digital literacy, economic improvements and the expanding reach of technological and digital communication services and devices also contribute to linguistic transformations and pressures.

Interoperability

Multilingual use patterns

The Wikimedia platform must begin to support context-switching between languages (according to context of use). Code switch(es), pidgins and inter-languages produce knowledge too, so how should our model of one project/one language adapt to a world of Hinglish and Spanglish?

Language and culture are intimately related. In multilingual societies languages come to serve specialized purposes, and some types of knowledge are associated with a particular language. For example it is not uncommon for there to be a language for science, a language for government and a day-to-day spoken language, in which mass media and popular culture are conducted. The idea that each language will have a complete vocabulary and source material to build an encyclopedia that looks like English's doesn't stand scrutiny.

The model of language around which the projects are built, with parallel encyclopedias and dictionaries for each language, and some set of shared content that is available across them all, understates the complex way people use languages. For 60% of the world [2] using multiple languages, for different purposes, is a the normal way to live.

(How) Can we create content that is consumable in multiple languages, or doesn't require language at all to enjoy and use? Perhaps new formats are needed.

Support

How do we provide not just content, but also support in marginalised or underserved languages? What happens when a contributor who only speaks e.g., Yoruba or Tibetan has a policy question? Participation demands support, especially on the long tail.

In parallel, we need to continue pushing forward Unicode font support and basic text entry (reference-- Minh, Odia, etc.), to ensure the most basic forms of access for all languages.

Opportunities & Threats

Machine translation is making the boundaries between languages softer, as are Wikidata and Commons. So how do we preserve the cultural values and self-determination of each speaker community in a world where these communities and languages are becoming more porous and mixing more than ever? How can we avoid reinforcing language inequalities, or worse, promote language extinction, when we use technology (machine translation, for example) to fill our content gaps?

Content translation is crucial, but insufficient. Language and culture are deeply bound to categorization, notability and content expectations. Simply translating all articles between all pairs of languages would not make a relevant or even comprehensible experience for most humans.

Advocacy

Societal Goals

How can the Wiki movement contribute to larger societal goals? Should it? Is there a way that the

Foundation can facilitate language teaching and learning?

Otherness

Should the Foundation take on the role of extracting and cataloguing knowledge from languages that don't have a writing system? Can our intervention could be at the level of the platform - for example a DIY kit for capture?

"Since the beginning of time, every culture has developed means of passing on important information to its people. For Hawaiians, there was no written language per se until the 1820s. The missionaries introduced the alphabet which made it possible to represent Hawaiian language in the written form. Until then, all information was passed orally through the use of songs, chants, and poems.

Hawaiians devised various methods of recording information for the purpose of passing it on from one generation to the next. The oli was one such method. Elaborate chants were composed to record important information, e.g. births, deaths, triumphs, losses, good times and bad."[3]

Preservation

To what extent should the Foundation advocate for or orchestrate the preservation of endangered languages? One possible route is to pursue partnerships, another is to develop educational materials in non-endangered languages that are spoken by many people, but are rarely used for education and business.

Notes

1 Categories of Multilingual <http://languages.org/bilingual.php>

2 <http://languages.org/bilingual.php>

3 Hawaiian Oli <https://apps.ksbe.edu/olelo/learning-place/performance-indicators/chant/foundational/significance-oli-chants-in-hawaiian-society>

03

Content Gaps

Are we there yet? Making the sum of the world's knowledge accessible to all the world's people presumes that the extent of knowledge can be described, [1] measured, [2] and tracked. [3] It also implies that content gaps can be identified and filled with context-relevant takes on the missing knowledge—this may be the most difficult part of the problem. For our purposes, gap knowledge can be said to fall into one of these general categories: potentially portable knowledge, [4] contextually nuanced knowledge [5] or contextually anchored knowledge. [6] The three types require different intervention strategies, and may have different relative priority for a given a regional community. For this reason, intervention strategies (product features, global tools, community mobilization) must be developed in parallel.

Solutions for filling content gaps also surface new challenges around notability and

verifiability[7]—how must current standards for authority evolve in order to encourage the growth of each type of gap knowledge? And finally, assuming that it is possible to close the content gap through various means, what support mechanisms will be required to maintain this body of knowledge? It will be necessary to have tools that monitor and signal the freshness of content to the community best suited to maintain it.

Making all knowledge available to all the world's people will demand the organization and participation of all the world's peoples, and interventions deployed at pace with the rate of growth of the overall pool of world knowledge.

Identifying Gaps

How Do We Identify Content Gaps And The Appropriate Ways To Fill Them?

What is “the sum of all knowledge?” Is it a thing? Does it concretely exist? Is the sum what results from everyone contributing or is it a matter of circumscribing an area? And in either case, how does one measure progress?

Content gaps may exist for a number of reasons that go beyond the well-known factors of institutional bias or lack of internet access. Knowledge may be held by a group or groups of people who are entirely unaware of Wikipedia’s existence; and if they are, they might not know that it is editable, or understand its relevance to them. People may have specific cultural, religious or other taboos against sharing specific knowledge (such as what they know about religious rites). They may also have ingrained cultural or personal beliefs that community-generated knowledge is less valuable than institutional-generated knowledge.

Filling-in content gaps goes beyond identification of the gap and the creation of initial content. It also needs to include finding and building audiences for contributed materials, and identifying and supporting contributors who want to maintain and expand the initial content. In both cases the risk is that content will stagnate and result in a loss of audiences and creators.

It will not be possible to fill all gaps immediately, and there are some gaps that may never be able to be filled. Prioritisation is therefore necessary based on the characteristics of potential audiences (including size, language, internet penetration and awareness), on the size of the potential contributor community, and on the notability of the topic in that particular language. Care should be taken with finding automated ways to fill the gaps, so that we do not inadvertently reinforce inequalities (see Language section, for example).

How Do We Bring Knowledge Out Of Legacy Media And Institutions And Onto The Web?

There is a vast amount of knowledge locked away

in legacy media that has never been digitised. These heritage materials are not available online for multiple reasons, including ignorance, copyright restrictions, and the degree to which the information is considered “specialized” or “esoteric.” Such cultural constraints are combined with technical ones, including the difficulty or impossibility of digitising with current technologies due to intangibility, deteriorated media or anachronistic communication styles (e.g. pre-19th century typefaces and manuscripts).

In many of these cases it may be a better strategy to forego rights acquisition efforts and digitising the content ourselves, and instead collaborate with or catalyze entities that are already pursuing digitization efforts such as libraries, archives, museums and dedicated online projects such as Gutenberg and Internet Archive.

Institutions (universities, governmental bodies, etc) also have huge amounts of knowledge that is not available on the web.

We might also partner with academics to bring their research to the public (such as by adding citations of published material to articles,) advocate for digitisation of government materials where governments are not already doing this, and advocate for meta-knowledge: knowledge of how to acquire knowledge.

Evolving Notability Standards

(How) Do We Manage Types Of Knowledge That Don’t Conform To Our Current Standards Of Notability Or Published Authority?

In order to make Wikimedia the repository of ALL the world’s knowledge, it would be necessary to consider how various alternative forms of knowledge can fit in with existing or new Wiki projects beyond Wikipedia. There are types of knowledge (such as oral history) that by their very nature do not conform to Wikipedia’s standards of notability or published authority. In such cases, Wikimedia should consider acting as a catalyst for organisations that are gathering and digitising these forms of knowledge, rather than attempting to reinvent the wheel—especially

when those organisations already have an existing body of work and useful contacts that would take us a long time to build up.

Incorporating Other Formats

(How) Do We Move Beyond Primarily Text-based Content And Into Other Formats Such As Video, Audio, Images, Etc?

Answering this question begins with defining the purpose of this type of content: is it to add to existing Wikipedia articles, encourage social media sharing, or serve as a destination unto itself? As we look to other sites that have embraced video we find that short-form video content is popular on social media and newspaper sites for example, but rarely repays its costs. On the contribution side, Instagram and YouTube demonstrate that people want to provide and share video, but the question remains as to whether this interest will extend to providing Wikipedia with such content. The approval and review practices that are a part of Wikipedia's culture, along with those of maintenance and updating, are different from the culture and performance of social media. Issues of consumption and contribution are both measured against tooling changes and process changes

Supporting New Platforms

How Can We Make The Tools For Filling In Content Gaps Available On The Devices People Are Likely To Need Them On?

Many of the most urgent content gaps are in languages and about topics relevant to places outside Europe and North America. Internet penetration is growing in these countries, though the majority of users access the web on mobile devices. Addressing content gaps therefore means thinking outside the desktop editing paradigm. Therefore we will have to consider several factors as we imagine expanding Wiki into these areas: Mobile-friendly tools for content creation, translation tools, and less data-intensive tools for rich media creation.

Notes

1 T. Negrin: "The Static Knowledge Pie": using the sum of all topics currently covered on all wikis as a measure of completeness, the content gap is measured as topics existing on some projects and absent on others. Requires an exhaustive but extensible topic model.

2 T. Negrin: "The Potential Knowledge Pie": The Static Knowledge Pie with the addition of known gaps across all wikis (e.g.estimated number of articles there should be about female scientists).

3 T. Negrin: "The Dynamic Knowledge Pie": the diameter of the pie increases over time making the content gap a dynamic function of the relationship between rate of knowledge created in the world and knowledge captured on the system.

4 Potential Portability reflects the degree to which certain classes of knowledge have the potential to be "ported" or translated with low loss in fidelity. Scientific terminology and topics are an example of PPK. Certain types of legacy media (e.g. illuminated manuscripts) that can be ported from one medium to another could also be considered PPK. Transferring Potentially Portable topics and formats is challenging, but methods for tackling them are known unknowns.

5 Contextually Nuanced knowledge is a category of knowledge that requires a higher degree of human input to transfer (e.g. subject matter expertise, topical knowledge and interpretation). The history of the Korean War written from a Japanese, or American or Korean point of view would be an example of CNK.

6 Contextually Anchored knowledge is a category of knowledge that is utilized by a specific community in a specific way. This knowledge can be described in other contexts, but not utilized in a comparable way. The Hawaiian Oli chant is an example of CAK. <https://apps.ksbe.edu/olelo/learning-place/performance-indicators/chant/foundational/significance-oli-chants-in-hawaiian-society>

7 Decline of editor retention in mature language communities on Wiki being (partially) caused by lack of welcome for newcomers <https://drive.google.com/file/d/1FzMCptTMOvvRRGQPyh2AsBrGneOIG3F/view>

- Article coverage by language
 “Increasing article coverage” https://meta.wikimedia.org/wiki/Research:Increasing_article_coverage
 “Growing Wikipedia across languages” <https://drive.google.com/open?id=1Fp4ktS3tzZDbfXuxraELk8EuCRb4W4dh>
 Content translation recommendation API https://meta.wikimedia.org/wiki/Recommendation_API
 Expanding Articles Cross-language” https://meta.wikimedia.org/wiki/Research:Expanding_Wikipedia_articles_across_languages
 “Brand awareness, attitudes, and usage - Executive Summary” https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Brand_awareness,_attitudes,_and_usage_-_Executive_Summary
- Cisco Visual Networking Index: Forecast and Trends, 2017–2022 <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>
- “Cultural Issues Hindering Content Creation” from New Voices synthesis: [https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/New_Voices_Synthesis_report_\(July_2017\)#Challenges_for_Wikimedia](https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/New_Voices_Synthesis_report_(July_2017)#Challenges_for_Wikimedia)
- Euromonitor <https://blog.euromonitor.com/2015/04/half-the-worlds-population-will-be-online-by-2030.html>
- Gapfinder <https://www.mediawiki.org/wiki/GapFinder>
- “Geographies of the world’s knowledge” https://www.oii.ox.ac.uk/archive/downloads/publications/convoconvoc_geographies_en.pdf
- “Gender Equity Report” http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/internet_draft_study.pdf
- Google-KPMG, ‘Indian Languages - Defining India’s internet’ <https://assets.kpmg.com/content/dam/kpmg/in/pdf/2017/04/Indian-languages-Defining-Indias-Internet.pdf>
 Halfaker , Geiger, Morgan and Riedl (2018) “The Rise and Decline of an Open Collaboration System: How Wikipedia’s reaction to popularity is causing its decline” https://drive.google.com/file/d/1FzMCp_tTMOvwwRRGQPyh2AsBrGneOIG3F/view
- Hargittai, E., & Shaw, A. (2018). “The Pipeline of Online Participation Inequalities: The Case of Wikipedia Editing”. *Journal of Communication*, Vol. 68, Iss. 1, 1-Feb-2018, pp143–168, doi. [org/10.1093/joc/jqx003](https://doi.org/10.1093/joc/jqx003)
- Hargittai, E., & Shaw, A. (2015). “Mind the skills gap: The role of Internet know-how and gender in differentiated contributions to Wikipedia”. *Information, Communication & Society*, 18(4), 424–442. doi: [10.1080/1369118X.2014.957711](https://doi.org/10.1080/1369118X.2014.957711)
- Hawaiian Oli Chants <https://apps.ksbe.edu/olelo/learning-place/performance-indicators/chant/foundational/significance-oli-chants-in-hawaiian-society>
- “Knowledge Gaps” <https://research.wikimedia.org/knowledge-gaps.html>
- Language Extinction:
 Wikitongues <https://wikitongues.org/>
 Foundation for Endangered Languages <http://www.ogmios.org/manifesto/index.php>
 Rosetta Project <http://rosetta-project.org/>
 Living Tongues Institute <https://livingtongues.org/>
 UNESCO linguistic diversity on the internet project <http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/linguistic-diversity-and-multilingualism-on-internet/>
 BBC: “Why We Must Save Dying Languages” <http://www.bbc.com/future/story/20140606-why-we-must-save-dying-languages>
 T.Skutnabb-Kangas, (2013) “Linguistic Genocide in Education” https://books.google.co.uk/books/about/Linguistic_Genocide_in_Education_or_World.html?id=sGtObjnBydQC&redir_esc=y
 Lewoniewski, Węcel and Abramowicz (2017), “Relative Quality and Popularity Evaluation of Multilingual Wikipedia Articles” <https://www.mdpi.com/1099-4308/9/1/1>

com/2227-9709/4/4/43

“Microcontributions” https://www.mediawiki.org/wiki/Mobile_design/Micro_contributions

“Mobile Phones and Internet Use in Cambodia 2016” http://www.open.org.kh/research/phones_2016.pdf

New Voices Synthesis: Attracting New Users [https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/New_Voices_Synthesis_report_\(July_2017\)#How_to_attract_new_users](https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/New_Voices_Synthesis_report_(July_2017)#How_to_attract_new_users)

“Readers Contributions Via Android” https://www.mediawiki.org/wiki/Reading/Readers_contributions_via_Android

“The Rise and Rise of Hinglish in India” <https://theconversation.com/the-rise-and-rise-of-hinglish-in-india-53476>

Sample, I. (2018-10-18). “Exclusive: dramatic slowdown in global growth of internet access”. *the Guardian*. Retrieved 2018-10-21. <https://www.theguardian.com/technology/2018/oct/18/exclusive-dramatic-slowdown-in-global-growth-of-the-internet>

Shi, Teplitskiy, Duede, and Evans (2017) “The Wisdom of Polarized Crowds” <https://arxiv.org/abs/1712.06414>

“The State Of The Article Expansion Recommendation System” (Research showcase) https://www.mediawiki.org/wiki/Wikimedia_Research/Showcase#December_2017

UNESCO’s “Fostering Inclusive Knowledge Societies Report” http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/internet_draft_study.pdf

WCAG “POUR” Accessibility principles - <https://www.w3.org/WAI/fundamentals/accessibility-principles/>

Wikimedia Foundation Values [https://meta.](https://meta.wikimedia.org/wiki/Values/2008#Wikimedia_Foundation_values)

[wikimedia.org/wiki/Values/2008#Wikimedia_Foundation_values](https://meta.wikimedia.org/wiki/Values/2008#Wikimedia_Foundation_values)

Wikimania 2018 presentation: “Wikipedia And Bhutan Can Learn From Each Other” https://wikimania2018.wikimedia.org/wiki/Program/Bridging_knowledge_gaps:_Wikipedia_and_Bhutan_can_learn_from_each_other

Wikimania 2018: presentation: “The Quotation Of Oral Sources In A Decolonization Context” https://wikimania2018.wikimedia.org/wiki/Program/The_quotation_of_oral_sources_in_a_decolonization_context

Wikimania 2018 Panel: “Centering Knowledge from the Margins: A Whose Knowledge? discussion” https://wikimania2018.wikimedia.org/wiki/Program/Centering_Knowledge_from_the_Margins:_A_Whose_Knowledge%3F_discussion

Wikimania 2018 presentation: “Wikipedia for Indigenous Communities” https://wikimania2018.wikimedia.org/wiki/Program/Wikipedia_for_Indigenous_Communities

Wikitongues <https://wikitongues.org/>

