

ubiquity

[unedited] In order to achieve ubiquity, we need to focus on the following themes:

Making our content adaptable to all platforms instead of focusing on future trends

Structuring our content

Unifying content through global tools

Investing in multimedia

Growing and connecting our projects

Investing in syndication without losing the essence of our projects

Studying the current effects of syndication

Demanding attribution

Making the risks to our future known to companies and individuals invested in or dependent on our content

Investigating models of accountability for companies using our content

Investigating alternate modes of content creation

Ensuring we have constant production of relevant content

Growing smaller communities

Addressing knowledge gaps

Balancing gaps in interest - creating content that readers are interested in

Sections

[Platform agnosticism](#)

[Syndication](#)

[Content Relevancy](#)

[Bits and bobs](#)

[Notes](#)

[Sources](#)

Platform agnosticism

“Stroll through Sanlitun, a bustling neighborhood in Beijing filled with tourists, karaoke bars, and luxury shops, and you’ll see plenty of people using the latest smartphones from Apple, Samsung, or Xiaomi. Look closely, however, and you might notice some of them ignoring the touch screens on these devices in favor of something much more efficient and intuitive: their voice.” [1]

It seems the Chinese language, as many other languages, was not built for typing tiny letters on a small screen. But that’s okay because technology, as it usually does given large-enough demand, is making its way around such technical difficulties. In this particular case, the answer might be voice search, but there are other cases where AI or messaging play bigger parts.

Many are coming to the internet with new needs, new languages, and new modes of expression and it is certain that their arrival will change the fabric of the ways and forms in which knowledge is created, shared, and used. As internet usage in growing economies rises, the internet will become a more diverse place and will be required to adapt to the needs and motivations (our own research shows that motivations vary greatly based on project) of its new users.

If Wikimedia projects want to be present along with this growth or even further and our goal is to “break down the social, political, and technical barriers preventing people from accessing and contributing to free knowledge”, we must ensure

adaptability to any platform or mode of usage.

Yet predicting trends can be tricky and the risks that have prevented us from being at the forefront of technical innovation so far still apply. Unlike Google, who have the resources to do everything-everywhere-all-the-time, we do not have the luxury or expertise in taking large risks, especially if they do not come from the ground up, from our communities.

For us, ubiquity must mean adaptation - skipping the guessing game of what will be big in the future, investing in the needs our current and potential communities and making sure that our content is prepared for use in any future trend and for presentation on any device.

Thoughts on ways we can achieve this:

- > Structuring our content so that it can be reusable by ourselves, our communities, and other platforms
- > Creating tools that our communities can use for structure such as establishing global templates or article templates. Global customizable templates will make the creation of APIs easier. It can allow individual wikis to structure their content as they wish while ensuring consistency in new features and presentation. Article templates can structure our content by section or idea - making it easier to use portions of content inside and outside of our platforms.

- > Providing the building blocks for technical tools to be built by the community. If we provide the framework for community-built tools, we can ensure that they will work smoothly with the remainder of our content. This will empower communities to build tools that are necessary for their own projects while allowing us to focus on the structure of projects themselves.
- > Investing in technologies currently used by emerging economies.
- > Investing in internal platform and project agnosticism by decreasing the separation between our projects. (Creating workflows that will smoothly move an interested user from the Wikipedia article on Istanbul, to the Wikivoyage guide, to images and videos about the city from Commons)

Syndication

“Wikipedia content appears to play a substantially more important role in the Internet ecosystem than anticipated, with other websites having critical dependencies on Wikipedia content.”

“Google becomes a worse search engine for many queries when it cannot surface Wikipedia content (e.g. click-through rates on results pages drop significantly) and the importance of Wikipedia content is likely greater than many improvements to search algorithms. Our results also highlight Google’s critical role in providing readership to Wikipedia. However, we also found evidence that this mutually beneficial relationship is in jeopardy: changes Google has

made to its search results that involve directly surfacing Wikipedia content are significantly reducing traffic to Wikipedia.” [2]

So far, Wikipedia’s relationship with Google has been fairly symbiotic. We provide a trusted source they can show at the top of the page; they provide an increase in pageviews and, in turn, an increase in donations, in new editors, and in the continued creation of quality content they can then show to users. Everybody wins and information is distributed freely.

Yet exposing more information outside of the site, such as in Google’s knowledge panels, has decreased pageviews to Wikipedia. It is unfortunate that this is an issue. We still completed our goal of providing the information a reader sought. Yet without the extra pageviews we face not only a decrease in funds, but eventually a decrease in quality. Potential editors never see the site and current editors have less motivation to continue writing. Over time, we’re in trouble.

But, so is Google. The study quoted above displayed, clearly, that Google is a worse search engine in a world without Wikipedia. Wikipedia’s importance is so large that the “mere presence of Wikipedia links may have an effect approximately 80 times larger than the difference between a good ranker algorithm and bad one (for many queries)”. Similar patterns have been found for other online websites such as Reddit and StackOverflow, where Wikipedia content is widely shared.

Thus we find ourselves in an odd paradox where our current ubiquity is becoming a threat to itself. One option would be to take a purely defensive stance and work towards preventing any information from usage outside of the site but it is needless to say that that goes against the morals of our movement severely. The other option would be to take syndication for granted - to imagine our content spread throughout the fabric of the internet, and shift our content creation and revenue model to such a future.

Thoughts on ways we can achieve this:

- > Opening conversation with our partners on these relationships asap
 - > Working with our partners to ensure proper attribution of information - not only for pageviews but also so readers are not misled by the presentation of information taken out of context.
 - > Creating standards that will hold larger organizations accountable, financially or legally, to the growth of the project they have become dependent of
 - > Inspiring organizations to invest in the future of Wikimedia projects by becoming editors or inspiring others to do so - for example, one of the papers included suggestions such as getting karma points on reddit for improving linked articles
 - > Working with our partners to create symbiotic ways of using our content (such as exploring features like context cards, which share a preview of wikipedia articles with direct links back to the original article page, outside our projects)
- > Working outside the boundaries of our platform - creating workflows of contribution from other places where our content is used.

Content Relevancy

“In the English Wikipedia, articles of strong insufficient quality alone receive close to half of the pageviews, and in the Russian Wikipedia, they receive more than half.” [3]

For our projects to be ubiquitous, we must provide relevant content to all of our users. Not all wikis are the same, nor do they grow in similar fashions and users of different projects have widely varied motivations for reading. For example, our research shows that readers in Western-language wikipedias are more likely to focus on quick-fact information whereas the speakers of languages in growing economies are more likely to use Wikipedia for deeper learning and for work or school purposes. To be able to cater to the needs of individual wikis or groups of wikis, we must be able to distinguish their needs. Features that might work great on one wiki, might not work on another. Similarly, content that might be notable for a particular community, might not be for another.

Focusing on targeting our work to match our unique audiences as well as providing them with the tools to build according to their needs will help us cover the entire range of content that our current and future readers will require.

Thoughts on ways we can achieve this:

- > Creating tools that will aid growing and established communities to identify and create the content that readers are interested in such as suggestion algorithms for sections and other type of content, article templates, improved content translation, and reader requests. These tools can also be used to focus on content diversity and, eventually, on editor diversity.
- > Focusing product initiatives on targeted groups of communities or geographic areas. We can focus on assisting projects according to their needs (for example, larger wikis that tend to be less friendly over time can gain from focusing on editor retention, while smaller wikis working on establishing their communities can focus on converting readers to editors).

Bits and bobs

Other thoughts on the subject of ubiquity:

- > Sustainable models of growth - modeling the needs of readers, editors, and moderators to highlight imbalances in projects (for example, quality content but low readership, or large amount of content but low quality)

Notes

- [1] <https://www.technologyreview.com/s/600766/10-breakthrough-technologies-2016-conversational-interfa-ces/>
- [2] [The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies](#)
- [3] <https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15623/14799>

Sources

O. Vasileva: [Research and Insights](#), Other contributors¹: A. Baso, B. Davis, A. Halfaker, D. Horn, R. Isler, J. Katz, J. Minor, T. Negrin, M. Novotny, N. Pangarkar

<https://econsultancy.com/why-we-need-to-stop-repeating-the-50-by-2020-voice-search-prediction/>

[The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies](#)²⁰

[Examining Wikipedia with a broader lens: Quantifying the value of Wikipedia's relationships with other large-scale online communities](#)²²

[Creative commons non-commercial license](#)
https://meta.wikimedia.org/wiki/Free_knowledge_based_on_Creative_Commons_licenses#8.

[Can NC-licensed content be used in the Wikipedia project? %E2%80%93 No, Wikipedia contents are being used commercially.](#)

[Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information](#)²³

Wikimedia movement strategic direction
https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Direction

A study of versions of articles available across 26 wikis. Identifies that project growth seems unrelated to linguistic or cultural factors.

[Robust clustering of languages across Wikipedia growth](#)

Studies multiple cross-language wikipedias to determine connections between articles in different languages are not directly correlated to their existence or quality in enwiki

<https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-016-0070-8>

Linguistic neighbourhoods: explaining cultural borders on Wikipedia through multilingual co-editing activity

¹ If your name was left off the list by mistake please contact JMinor or MNovotny

A study of inequality in content among different language projects. Identifies common growth patterns among wikis
<https://arxiv.org/abs/1610.06006>

The Rise and Decline of an Open Collaboration System: How Wikipedia's reaction to popularity is causing its decline
https://www-users.cs.umn.edu/~halfaker/publications/The_Rise_and_Decline/halfaker13rise-preprint.pdf

A visualization that displays the differences in participation between various projects - could be used as an indicator of project growth cycles
<https://dl.acm.org/citation.cfm?doid=3173574.3173929>

Modeling crowdsourcing as collective problem solving - models productivity for crowdsourcing tactics
<https://www.nature.com/articles/srep16557>

Research on knowledge gaps, esp. Recommendation systems <https://research.wikimedia.org/knowledge-gaps.html>

New editors research - highlights the needs and challenges of upcoming editors on established projects
https://www.mediawiki.org/wiki/New_Editor_Experiences

Displays that knowledge is more effective when actively sought

Effects of the News-Finds-Me Perception in Communication: Social Media Use Implications for News Seeking and Learning About Politics <https://onlinelibrary.wiley.com/doi/full/10.1111/jcc4.12185>

An op-ed detailing the increasing predominance of multimedia over written content
<https://www.nytimes.com/interactive/2018/02/09/technology/the-rise-of-a-visual-internet.html>

The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies - shows that Google's clickthrough rates drop significantly when not surfacing Wikipedia content, but that Wikipedia's pageviews drop when Google displays such content
<https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15623>

A study of usage of CC non-commercial licenses
Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234435/>

Usage of voice search - increases over time. Looks at whether by 2020, 50% of the world will be using voice search. <https://econsultancy.com/why-we-need-to-stop-repeating-the-50-by-2020-voice-search-prediction/>

https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Direction/Appendix#pattern3