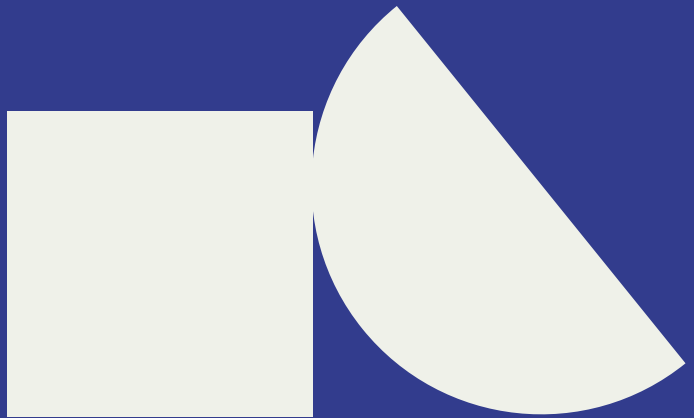


# Trust

This collection of essays explores the role of Trust in relation to the Movement, the content and the Foundation. These positions are intended to stimulate discussion and inform the development of new capabilities.



Published  
**January 2019**

Authors

**Trevor Bolliger**  
**Alex Hollender**  
**Jon Katz**  
**Margeigh Novotny**

Contributors

**David A.M. Goldberg**  
**Danny Horn**  
**Josh Minor**  
**Jonathan Morgan**  
**Toby Negrin**  
**Jacob Orlowitz**  
**Nirzar Pangarkar**  
**Alex Stinson**  
**Ben Vershbow**



**T**rust is a fundamental building block of community and culture, and is therefore critical to consider as we define next stage strategy toward 2030 goals. In commercial contexts, mutual interest and mutual concern mandate some level of trust between individuals or entities engaged in trade. In civic contexts, trust is often a by-product or outcome of the process by which collective good is determined—and a baseline criteria in selecting who can represent or advocate for that good. In interpersonal contexts, trust is the social lubricant that allows us to progress, step by step, from strangers to acquaintances to intimates. In each of these contexts (all of which have parallels in the open knowledge ecosystem) the rituals of trustworthy engagement share the same basic characteristics.

Parties in trustworthy relationships...

Take an interest in or are curious about one another

Act in good faith and assume good faith

Share relevant information voluntarily

Take information that is shared at face value

Behave respectfully toward one another

Strive for equitable value exchange

Are reliable and honor their agreements

Sometimes share mutually beneficial objectives

These characteristics, and the relative weight placed on each, are manifest at cultural, social and transactional levels in ways that have the potential to significantly impact our Movement and its goals. And, as a global organization, it's important to understand how our predominantly northern/western model of trust-building may be reflected at each level.

In the papers that follow we explore three

vectors of trust: reliability of content, loci of accountability and transparency of the Foundation and the Movement.

Investment in Trust means the Movement, the Foundation and every community must cultivate awareness and facility with the cultural, social and transactional aspects of trust-building interaction. That is, in order to empower and activate all communities to take part in the creation of knowledge we must develop an actionable model for building trust and strengthening alliances - one that is effective across projects and between cultures. This will require awareness of our own trust-building paradigms, and a fluency with reading and sending trust-signals in our governance practices, our community norms and policies, and via our content. We will need to consciously define and embody trustworthy practices at every touchpoint of the Movement.

**Notes**

1 Even in illicit markets, trust is a factor in determining who will do business with whom.

2 Reaching our 2030 objectives will require effective and dynamic partnership with sister-organizations and for-profit partners (trade). The “civic” dimension of the movement must include support for a multiplicity of interest-groups within the Movement, and the balancing of leaders’ voices. It will require these civic leaders to exhibit trustworthy practices and exemplary judgement (organizers, content curators and moderators fall into this “civic” dimension). Their objectives must be understood and communicated transparently, and leaders of interest groups within the Movement must have the ears of other leaders. On the interpersonal level, individual users must be able to move from anonymous consumer to engaged contributor through a set of mutually agreeable and understandable gates; and once there, experience only respectful interactions with others.

3 At a minimum, they “trust but verify”

4 But isn’t it necessary for parties in a trusted relationship to have some “mutually beneficial objective”? No, it may not be necessary. It’s possible for one party to take an interest in another, and to demonstrate every other characteristic of a trusting relationship, but in the end see no mutually beneficial objective to work toward together, at a particular time. This distinction marks the difference between allies and partners. When Trust is present, both allies and partners are able to work constructively to support one another, whether or not they have an immediate mutual objective (and while cultivating allies is a reward in itself, allies can become partners later).





## 01

# Reliability

Wikipedia is one of the most trusted sources of knowledge in the world.[1] In one telling example, a 2013 report determined it to be the top source of healthcare information for both doctors and patients in the US. [2] The importance of this trust has intensified in the last few years, as the world has polarized further; new, niche media sites have proliferated; blatant lies are told by politicians on television; and “fake news” spreads quickly. It is increasingly hard to tell what is true and what is not. At this time, Wikipedia is increasingly seen as a uniquely trustworthy source, a reputation demonstrated by the reliance on Wikipedia by large media platforms to fact-check assertions made on their platforms. [3]

At the same time, there is room for improvement and reason to invest in reader trust. While Wikipedia is trusted by many, it is also famous for misinformation.[4] Wikipedia has historically been regarded with unease and contempt by traditional sources of information (teachers, journalists) [5] and it is easy to imagine how a high profile scandal might undermine public faith in Wikipedia for years to come. A recent poll of Wikipedia users found that the number one request of respondents was “More trustworthy

content.”[6] Stories of vandalism frequently make headlines, and teachers around the world tell their students not to trust Wikipedia. Some of the many reasons not to trust Wikipedia content are outlined in a very thorough Quora answer by Andreas Kolbe [7] (a longtime community member and Wikipedia documenter) for anyone who is interested.

## Challenges

One of the challenges readers have with Wikipedia is determining what to trust and what not to trust. Multiple studies have shown that readers tend to use proxies like article length, pictures and grammar along with citations as proxies for trustworthiness.[8] [9] [10] When we asked Donors what we should improve, we received many responses like these:

*“A way to measure/judge how reliable the information I’m reading is. A truth-meter or fact-meter. I’m told not to trust the information I read on Wikipedia so many times I don’t consider it factual until I read it published somewhere else.”*

*“As a reader it would be hard to know what is fake information. But i trust Wikipedia, never even checking the references unless I need them.”[11]*

Given the wide range in quality of articles on Wikipedia,[12] this lack of signal within Wikipedia represents unappreciated value (in the case of good articles) and avoidable liability (in the case of bad ones).

Admittedly, it is not clear how distrust impacts Wikipedia: a recent study found that as students move through the educational system, they are more likely to be told not to use Wikipedia and yet remarkably use Wikipedia more and more. [13] However, distrust does impact how our content is leveraged in everyday life. While doctors may use Wikipedia extensively, they don’t rely on it. This UK doctor sums it up nicely: “I use Wikipedia to gain a quick overview of a subject/topic that I am unfamiliar with or to jolt my memory of a subject. I would never base management or treatment of a patient I find there – for that I use my own knowledge, hospital protocols/guidelines, textbooks and advice from colleagues.” [14]

The impact of greater trust in this case, might only be seen in offline behavior.

## Room for improvement

### Marginalized Communities

As we consider the next 2 billion people who are coming online, we need to consider the trust they will have in Wikipedia. If someone goes to Wikipedia and doesn’t see their own experience reflected there, how can they truly trust it?

In many ways, the citations tools that create trust among the majority by relying on and mirroring traditional structures of power and authority, undermine the trust by the minority. Along these lines, Wikipedia is facing legitimate criticism from groups who feel that Wikipedia does not represent their truth. Examples include: WhoseKnowledge, WomenInRed, WikiMujeres, Wikigap, Afrocrowd, Art & Feminism.

The efforts to improve trust globally and among marginalized communities will likely overlap in some areas (showing number of editors) and deviate in others (loosening reliable sources criteria).

### Investing in Trust

Wikipedia is currently trusted enough for broad, global usage and increasing trust further has diminishing returns. Yet, it is also self-evident that maintaining some minimum degree of trust is essential for meaningful usage. It is also true that giving users greater clarity over what they can trust on Wikipedia would allow people to use Wikipedia more effectively. In this respect, we hope to increase trust in Wikipedia overall, by, at times, lowering reader trust in a particular piece of content.

For this reason, we think that “Reader Trust” should be invested in as an asset that increases our impact and requires defensive support.



## Specific Trust Concerns

### Citations need support

We need to invest in trust, by improving our citation infrastructure. Verification via citation is one of the cornerstones of our trust and there are significant opportunities to improve our citation experience. However, citations are neither sufficient nor accessible signals of trustworthiness and there is more room for signalling to readers in other ways how much to trust a particular piece of content.

Citations are a valuable measure for accountability that we have on Wikipedia. However, they currently suffer from several limitations:

- User awareness and understanding of citations
- Unable to serve function of verifiability [15]
  - Incomplete coverage
  - Incomplete archiving
  - Paywalls
  - Language barriers
- Once verified, hard to evaluate
  - A proliferation of new media outlets to vet
  - An proliferation of real-looking, fake media outlets
  - Cross-article analysis. (Wikicite, in particular, looks to open this up)
- Dishonest or incorrect use of citations by editors
- Existing reliable source definitions limit information

Some of these challenges are harder to solve than others, but we should invest in solving or accounting for them. It is worth noting that two other popular mechanisms are currently employed signal content trustworthiness:

article quality grades, and hatnotes (page issue templates). Both of these are manually derived, manually updated and subject to both gaps in either definition or usage. Hatnotes, are also subject to over-use. In fact, the term “tag:bombing” [16] was created to describe the destructive power of tag bombing. Here [17] is a famous example.

Specific solutions are recommended in the last section.

### Verifiability is Insufficient

However, citations are not the only tool at our disposal, and, even at their best, represent an incomplete picture of content trustworthiness. For instance, one of the most common forms of deception is selective truth. If the article on Barack Obama only had only one sentence that said: Barack Obama is the author of the bestselling book “Dreams of my father”, the article would be tantamount to a lie. Similarly, the following statement, while factually accurate would be highly misleading:

While Barak Hussein Obama claims to have been born in the United States, there are many prominent figures who dispute this claim. The current president of the United States publicly stated his belief that Mr. Obama founded ISIS.

As another example, found on the Wikipedia article on Wikipedia’s reliability, a 2014 exploration found that FDA drug warnings were not sufficiently updated on Wikipedia in sometimes stark circumstances.[18]

There is also the concern that most users do not actually open the citations and are therefore apt to falsehoods that are planted there either erroneously or intentionally. Planted falsehoods are the sort of behavior that users (anecdotally) seem most concerned about. Its okay if a 7th grader learns that the Eiffel tower was built in 1830 (maybe Encyclopedia Britannica has a typo), but not okay if they learn that it was built by a troupe of runaway mimes.

As is often the case, the deficiencies here are maximized in the case of promoting knowledge equity. All too often, content, well-cited or not, takes the perspective of the traditionally powerful, Western white world. In doing so, it undermines the trust of marginalized communities.

Siko Bouterse, of the group Whose Knowledge,[19] references the following example from Wikipedia to show how existing power structures undermine trust in Wikipedia. During an editathon at Wikicon San Diego in 2016, with individuals from the Kumeyaay, a tribe indigenous to California, the California Gold Rush page was edited. While the goldrush was many things, it was considered a genocide by California's indigenous people. One of the first changes they wanted to make was to change the picture on the page about indigenous people, from one of an indigenous person attacking a settlement, to one in which the Western settlers are firing their arms.



*BEFORE* User:NorCalHistory CC BY-SA 3.0 [20]



*AFTER* Illustration by John Ross Browne Public Domain [21]

As you can see dramatically here, the truth of our content goes well beyond whether or not something is cited, but the same solutions as above apply.

### **Policies Can Negatively Impact Trust**

In 2012, the celebrated author, Philip Roth, noticed that an article about his 2000 novel “The Human Stain” had a verifiable fact that was nonetheless misleading. It said the novel was “allegedly inspired by the life of the writer Anatole Broyard.” This fact was cited. When Roth contacted administrators at Wikipedia to tell them this was incorrect and what really inspired the novel, his correction was rejected. His interlocutor was told: *I understand your point that the author is the greatest authority on their own work, but we require secondary sources.*[22]

Roth wrote about this experience in the New Yorker and thereby publicly illustrated how the focus on reputable sources can actually decrease trust in our content.

Again, these drawbacks have bigger impact on marginalized communities. While Philip Roth, being famous, was able to publish this piece in the New Yorker, that option is not available to most people. Indeed the problem Philip Roth illustrated, of overly restrictive “reliable source” policy is greatly exacerbated when we consider regions and peoples whose access to and coverage by traditional media forms is limited.

**Priorities and Approaches**

Given the value of investing in trust for our users and the danger to Wikipedia’s credibility if we don’t, the following areas of intervention are suggested in order of suggested priority.

**Signals of Trustworthiness**

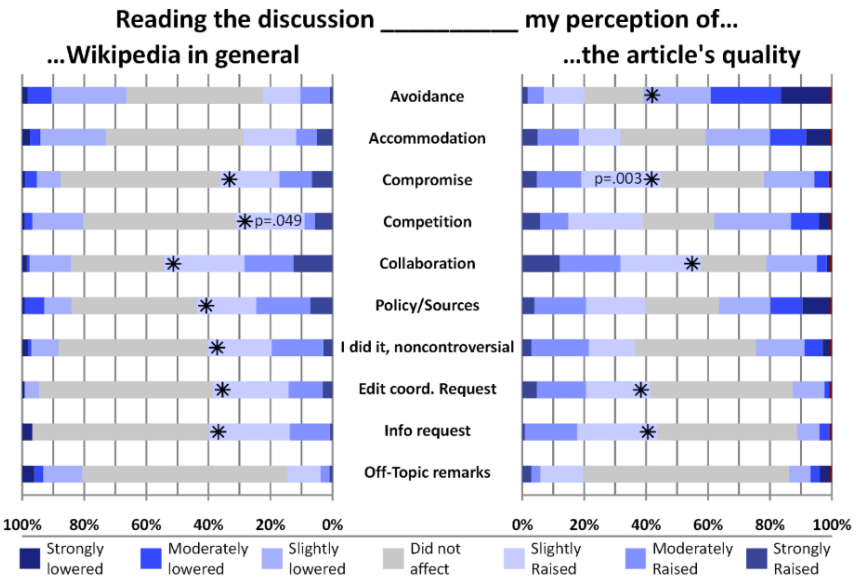
In addition to citations, we need to bring to the forefront other tools to give users the signals they need to make educated assessments of how much to trust individual pieces of Wikipedia content and thereby, trust Wikipedia more. This will

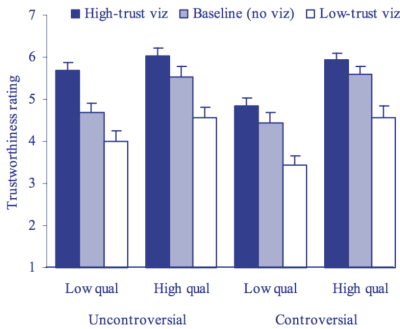
serve all readers and will not further reinforce traditional power structures. Right now we only use two tools: citations and page issue templates, and as described above, these are incomplete.

We can do this by exposing information about the article’s development and current status in easy-to-understand ways. On one end of the spectrum is access to a history page and talk page--this requires a great deal of time and head-scratching. On the other end of the spectrum is a trust score-- this is troubling because it is bound to be inaccurate and undermines a readers’ critical thinking skills. The recommendation is for visualizations and statistics that readers can use to form their own opinions.

Educating readers about how Wikipedia works is going to require great skill. One study showed that learning more about our process (seeing an article talk page) lead to a drop in trust, independent of article quality.[23]

Another study showed that trust in the same





Wikipedia article could go up or down, depending on how the material was presented. [24]

Here is an example of a visualization that led to higher trust:



Another small study showed that having students make edits to Wikipedia had a dramatic increase in their qualitative assessment of Wikipedia's value and accuracy.[25] A final study of Wikipedia found that a particular set of visualizations helped readers better distinguish between low credibility and high-credibility search results, but had little result on web-pages.[26] Interestingly, in the first study, users reported that seeing the discussion had changed their opinion. So it is clear that care will have to be taken when employing and testing these visualizations.[27]

### Biased or Incomplete information

**Priority 1: Show readers indicators accessible information (visualizations) of how biased or incomplete an article might be.**

Solutions:

- Showing Wikipedia readers how many authors contributed to content
- A ratio of article depth to article importance (links in, centrality)
- A signal of how active the discussion is
- A signal of editor diversity (based on editor behavior patterns or geo)

### Vandalism or Misuse of citations

Give readers digestible information about how likely a piece of content is to be valid.

Solutions:

- Showing readers how many "watchlists" the article appears on
- How recently it was edited, the ORES score of the last edit
- A blame map of which there are already more than a dozen [28]
- How many times the page was viewed since the last edit
- Providing users with annotated screenshots of the content from the cited source material

Priority is given to this set of solutions, because there is rich data that is currently invisible or inaccessible to most users.

**Priority 2.** Include the perspective of marginalized communities

In order to reflect the “Truth” of marginalized communities, we cannot continue to rely solely on the signals from traditional sources of power and authority. More specifically, we need to very carefully move away from a sole-reliance on published secondary sources by a handful of institutions.

**Narrow Reliable Sources Policy**

No clear product interventions exist for broadening reliable source policies. The heavy lifting might need to come through community activism and supported by product interventions (such as color-coding citations based on what kind of source they were). Dangers to this project include undermining our system of trust by opening up massive loopholes in the fact-checking process. However, steps can be made to mitigate these concerns.

**Priority 3. Promote media literacy**

We should then invest in teaching users critical media literacy: about how articles are created, the protections in place and how they can be flawed.

Solutions:

- UX features like those mentioned above
- Making citations more visible
- Online resources for journalists, students and educators (teachers, librarians, other professionals)
- Outreach programs for journalists, students and educators (teachers, librarians, other professionals) teachers

**Priority 4.** Address gaps in citation verification infrastructure

We also should continue to invest in making citations easier to create and verify. As mentioned

above, the current verification system is plagued by several issues including:

- Unable to serve function of verifiability:
  - Incomplete archiving
  - Paywalls
  - Language barriers
- Once verified, hard to evaluate
  - A proliferation of new media outlets to vet
  - An proliferation of real-looking, fake media outlets
  - Cross-article analysis. (Wikicite, in particular, looks to open this up)

**Citation Gaps**

Solutions here are around making the source material easier to find and digest

Solutions

- Leveraging wikidata to multiply fact-checking across languages
- Anti-paywall, free academic knowledge advocacy
- Reliable source database--allowing for validation of sources and quick cleanup when/if a source is deemed unreliable
- Reference material translation
- AI fact streams (primer.ai),
- Providing users with annotated screenshots of the content from the cited source material



## 02

# Accountability

A key factor in building and maintaining trust is holding ourselves accountable for the human dynamics in our Movement and on our platform. That is, in the coming years, we must not only recognize the various ways that bias undermines the growth of our Community, we must also actively cultivate trust-building practices across cultures, within communities and between individuals. This is a strategic imperative because only by examining our assumptions, behaviors and norms (as a mostly northern, mainly western, largely male movement) will we be capable of recognizing all possible forms of

knowledge, expertise and notability. To this end we must develop facility with navigating between the cultural, social and transactional factors that foster or undermine trust. Accountability must become a pillar of this socio-technical platform, built by the Movement, modeled by the Foundation, and maintained by each Community and contributor.

**Accountable for Cultural Literacy**

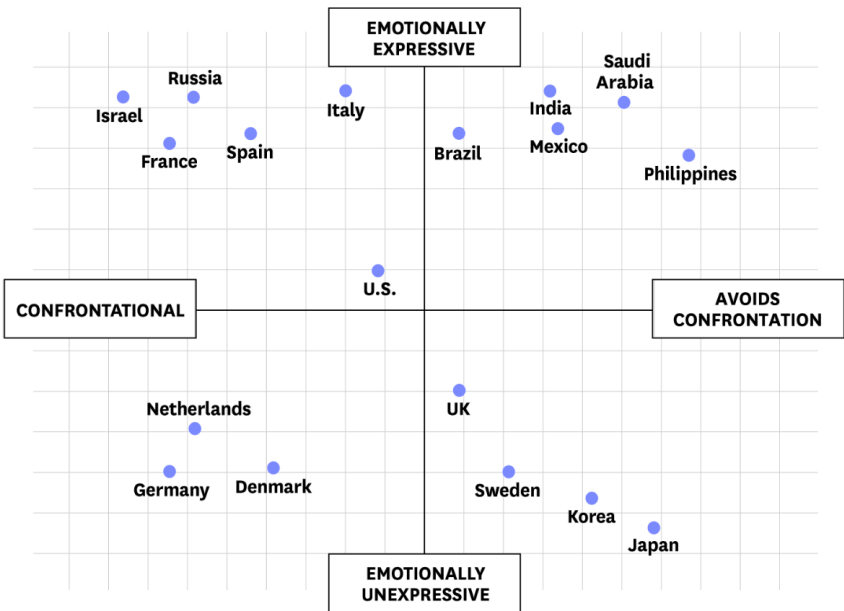
Paths to developing trust differ in different parts of the world, with some cultures placing early emphasis on the more cognitive aspects [1] (evidenced by a willingness to depend on the other based on demonstrations of competence, reliability and expertise), and others on the affective aspects (the degree to which either party feels comfortable sharing problems and aspirations; the willingness to be vulnerable to the other).[2] These tendencies are driven by cultural values and directly influence our ability to establish trust/find common ground as allies operating across cultures.

While direct debate is a paradigm we take for granted in a western/northern intellectual tradition, in a cultural perspective that values harmonious interaction, a passionate debater might be read as less trustworthy.[3] So when we base trust judgements on the signals we bring from our own cultural contexts, it's easy

to misread the intentions and motives of others (bias).

As it relates to the open knowledge movement, this phenomena may play out in governance practices and organizational principles. It could be a factor, for example, in determining or recognizing progress in new chapters, or in the criteria used to decide which projects to fund. It could show up as a preference for “having things in writing”, or as a disadvantage for those in leadership roles who use cultural signals that can't be read at a distance (e.g. body language). It may also play out in terms of expectations of how decisions should be made,[4] or in the authenticity we assign to a particular types or sources of information.

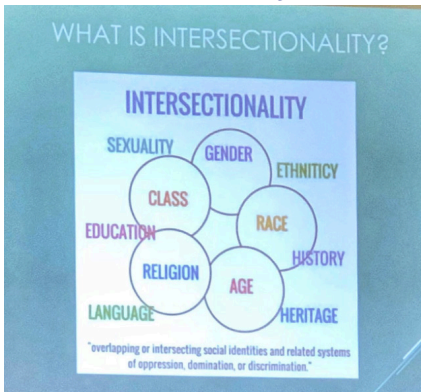
And, to complicate things further, we must also recognize the intersectional [5] dimensions of trust in any cultural context - where who we are constrains what we can express and to whom. To



SOURCE ERIN MEYER FROM "GETTING TO SI, JA, OUI, HAI, AND DA," DECEMBER 2015



its credit, the Foundation has evolved a range of workplace standards and practices intended to promote more cultural and intersectional awareness within teams and between geographies. But the same kind of intentionality must be applied at the Movement level in order to build and maintain trust at a global scale.



From WMF All Hands 2018 presentation by Russell Robinson [6]

**Accountability Within Social Groups**

In addition to cultural norms around trust signals and trust-building, trust is also manifest at a social level, via the personalities, [7] norms [8]and perceptions of any particular social group. The distribution and prevalence of certain tendencies, values and attributes across a social group [9] within a given cultural context [10] impacts who will be able to dominate or influence and who will be perceived as more or less competent or trustworthy. [11] For example, in a social group that is risk averse, more daring suggestions will be considered less trustworthy. Or, in a group that engages mostly to share perfunctory information, emotional or subjective messaging (e.g. demonstrative praise, gratitude, etc.) could be regarded as unusual or suspicious.

It is important to recognize the nested relationship between cultural and social factors - how the cultural frame, and aggregate of

individual personality traits, facets [12] and sub-facets,[13] set the prevailing tone, tendencies, and qualities of a social group. And hence, determine whose judgements and decisions are to be valued and trusted.

Communities take on the personalities and tendencies of their most influential characters, and must be challenged directly when these tendencies become exclusionary or dysfunctional. But challenged by whom? Should it be the responsibility of the Foundation to sanction communities? No. But the Foundation can model zero tolerance for negative social dynamics within its own practices, and develop capabilities for identifying and supporting communities struggling with community health issues. Ultimately, and ideally, it must be the community which holds itself accountable for fostering dynamics which encourage growth, and sustain involvement.

**Accountability for Interactions**

**Manners**

“Cosmopolitanism”, a term dating back to fourth century BC, meant “citizen of the cosmos” – the controversial and paradoxical notion that every person belongs both to the particular place they live and to the universe at the same time. The concept was a rejection of the then prevalent idea that a person’s fealty and citizenship was defined entirely by their polis, and its stature among other city states. Voltaire later referred to cosmopolitanism as “an obligation to understand those with whom we share the planet.” [14]

*“So there are two strands that intertwine in the notion of cosmopolitanism. One is the idea that we have obligations to others... beyond those to whom we’re related... or even [by] the more formal ties of a shared citizenship. The other is that we take seriously the value not just of human life but of particular human lives. Which means taking an interest in the practices and beliefs that lend them significance.” [14]*

This “valuing of the practices and beliefs that

lend particular lives significance” typically shows up in the rituals of respect, or manners, we employ when engaging with strangers. Some form of etiquette exists in every culture, but what is considered polite in one context may not be in another. Yet we can demonstrate our cosmopolitanism, our obligation to and curiosity about others, whenever we act according to any form of politeness, since polite interactions are the means by which we demonstrate respect while exploring our differences.

In the late 1970’s, Singapore instituted the National Courtesy Campaign [15] “as a means of ensuring a smooth transition to a new Singapore which would be densely populated, where people lived and worked in high rise towns, offices and factories, while travelling in crowded buses and lifts.” The goal of the campaign was to encourage behavior that would make life better for everyone, and to improve the “self-esteem” of the Singapore resident. The decades long public relations campaign included a mascot, Singa the Courtesy Lion, who became a beloved cultural icon by promoting norms of behavior that would signal Singapore as a cosmopolitan (world) culture. While it’s worth noting that Singa retired in 2013 with an open letter saying it was “just too tired to continue facing an increasingly angry and disagreeable society” asking real people to step up as the mascot stepped aside. “The letter also stressed that “it’s not that we aren’t a gracious society, or that kindness is not innate in all of us. But some days it feels like not very many of us believe in or care about expressing kindness.”

The Movement is at a similar inflection point. In anticipation of the many, many “others” we must engage in order to meet our 2030 goals, the onus is squarely on each community member to step up, to demonstrate kindness, tolerance and curiosity in every interaction they have on the platform. That is, to be cosmopolitan... by honoring our obligation to one another and by choosing to take an interest in the belief systems that lend significance to the individual lives of our fellow community members.

### Civil Debate

Kialo, the Esperanto word meaning “reasoning for”, is also the name of a web-based forum [16] designed to encourage and develop parliamentary debate as a counterpoint to the “Internet Shouting Factory”. [17] Kialo provides an interface for the crowdsourcing of ground truth answers to any question – from “Is Morality Objective?” to “Should the UK Remain a Member of the European Union?” to “Are we Living in a Simulated Reality?”. [18] Similar to Quora [19 ] and Aardvark [20], a Kialo debate starts from a user-contributed question. Unlike Quora and Aardvark, where questions are answered by “experts” whose status as such are determined by up-voting by their user-peers, Kialo questions are answered in structured debate via pro and con assertions, with each pro and con pair forming new branches of the “tree”. Each assertion, pro or con, can be commented upon, justified, and up/down voted, with the aggregate effect being a more nuanced consideration of every aspect of a contentious question, rather than the more typical conflation of aspects which often leads to anti-social comments.



*Kialo Interface: Should The UK Remain a Member of the European Union? [21]*

The Movement and its communities must not only advocate for civil debate, they must also model civil debate within the product, through interfaces that promote civilized discourse. This imperative will have two-fold returns: a more overt system for encouraging positive behavior,

and a structured database of dialog on every topic. The latter is a potential asset in and of itself.



## 03

# Transparency

The general level of trust in digital platforms, both in terms of the accuracy of their content, and of their operating practices, has hit a new low.[1] Most content sites have optimized around the popularity of their content and the speed at which they pump it out, often to the detriment of quality and trustworthiness. As a culture, we can now see the social costs of moving fast and breaking things.[2] And while it's easy to bemoan the experience gap between wiki projects (slow and old) and other people-powered platforms there's a hidden upside. Our projects adapt at a plodding (human) pace, change is slow and painfully incremental. But slowness is an advantage when it comes to trust because trust is built on consistency and predictability. The open knowledge model has an innate stability and is inherently more reliable since very little of this output changes at the pace of the world around us.

In the current climate of distrust, being perceived as trustworthy presents an opportune moment for Wikipedia. With so much positive social capital built up over so many years, it's now time to take a risk: we must openly critique the flaws in this edifice in order to retain trust in the long run. Being more intentionally transparent about the messy process by which knowledge is created would almost certainly invite criticism but it's also the only way to begin to address gaps and bias at a system level.

### How does Wikipedia work?

At the heart of a lot of the distrust of digital content, as well as the companies behind that content, is a lack of understanding about what's going on behind the scenes. How does Facebook's revenue model compete with the quality of information in your feed? How does Quora make money? What kinds of policies have worked to make Twitter more or less safe? How does a Wikipedia article come to be? While technology in general, and specifically content/knowledge/information has become more accessible, the processes behind it have become more and more obscure/opaque. These days information simply arrives at your doorstep, and perhaps we don't spend enough time thinking about how or why it got there. Google provides a decent example of what it looks like to explain the inner-workings of "Search" to people.[3] While Wikipedia is certainly complex, we need to do a better job of explaining what's going on, how it works, what hasn't worked in the past, where the money comes from, how it gets used, what a citation is, why it might not make sense for there to be a page about your grandfather, etc. What are some of the central tensions that we grapple with, and how might an understanding of those tensions make people feel more aware and engaged in the platform?

### What this could look like in practice:

- Create a visual, ELI5 type of experience where we explain the central concepts of Wikipedia. This could exist as a standalone "micro-site"
- Create case-studies about certain articles (either directly on the platform, or as a supplemental thing) — narrative storytelling around how this article came to be, and what the future might hold for it
- Invest in exploring the general question of "what information can I trust?" (think about Snopes as a brand) — we can be facilitators of the general discussion of what trust means, and who deserves it

- Invest in contextual, in-product explanations of things like references, verifiability, and notability

Create a Wikipedia steward certification program where people can become certified in understanding how Wikipedia works, and explaining it to others.

### Insider Knowledge

A big part of what has eroded trust in other social digital platforms is the degree to which they enable and engender echo-chambers - one now has to be very deliberate in seeking information that doesn't reinforce one's existing points of view. While Wikipedia doesn't share many of the negative attributes of other platforms, it does in this one regard: our contributor community places high value on neutrality but is not itself neutral. It attracts the like-minded, rewards the technically adept, and celebrates insider knowledge. "High standards" can no longer be used to justify lack of diversity in content or among contributors. Only by being more transparent about these shortcomings will the community be challenged to evolve, and make it possible for more diverse perspectives to be represented.

### What this could look like in practice:

- Call attention to the demographic makeup of an article's authors/contributors
  - Highlight articles in need of more diverse perspectives and reward contribution from such folks
  - Publicly communicate what we know about content gaps and invite new community members to come help us address them
  - Be more aggressive about highlighting page issues, and find other opportunities to be proactive about where our content falls short
- Policies are set and upheld by the community, so by what mechanisms could the community be motivated to change its norms? What leverage is

there at the product-level to drive this change? What can be done to make invisible power structures explicit and more easy to navigate as newcomers? How can verifiability and accessibility be reconciled?

### Calling Attention to Flaws

On the consumption side of the platform, readers must be able to see not only where content is under debate but also where content has been removed or altered by someone standing to gain from the change. By tracing edits to IP addresses at a particular locations,[4] or user names to corporate interests, [5] readers with sufficient motivation can ferret out abuses of the system and decide for themselves what to believe. But is there more the platform could do to identify conflict of interest? Is there more that can be done to surface tampering?

An objection might be that to call attention to vulnerabilities will undermine the credibility of the corpus overall, however transparency is simply preventive medicine.

**“What sets us apart, I believe, has been our willingness to be transparent in our journey. We own our failures, we learn from them, and we share them publicly so that others can learn from our failings as well, which has helped us to bounce back higher than before when we fall.”** [6]

When issues are known but go unaddressed, it invites greater scrutiny down the road. Being called out on a known issue is far more damaging to the organization than tackling the issue head on. As we see with Facebook, an organization's resources are far more strained by damage control than by self-initiated remedies.[7] Transparency is just good business.

Coming to terms with platform weaknesses - in terms of who is allowed to contribute and where abuses of the system occur - will call the integrity of the system into question. But only by

leveraging positive social capital, and tackling these challenges head-on will we retain trust over time. What this could look like in practice:

- Investing in raising public awareness around abuse, harassment, and bias on the platform

- Case study around a Wikipedia-article-gone-wrong

- Take even more responsibility around what qualifies as a good/bad source — develop a model for quantifying reliability of a source and share it openly so that others can understand how it works and help us improve it

What does this have to do with reaching our 2030 goals? Getting there will require scaling up participation in and consumption from the platform. Emerging communities will only invest to the extent that they feel their contributions will be welcome and secure- for this reason it will be critical for the existing community to trust them in order to be trusted in return.

### Notes

1. Edelman Trust Barometer Global Report 2018 <https://www.edelman.com/trust-barometer>
2. Feezell and Krupnikov (2018) What's True and Fake About The Facebook Effect <http://behavioralscientist.org/whats-true-and-fake-about-the-facebook-effect/>
3. How Search Works <https://www.google.com/search/howsearchworks/>
4. [https://en.wikipedia.org/wiki/United\\_States\\_Congressional\\_staff\\_edits\\_to\\_Wikipedia](https://en.wikipedia.org/wiki/United_States_Congressional_staff_edits_to_Wikipedia)
5. The Atlantic “Wikipedia Editors For Pay” <https://www.theatlantic.com/business/archive/2015/08/wikipedia-editors-for-pay/393926/>
6. Forbes “Trust, the Most Valuable Business Commodity” <https://www.forbes.com/sites/davidkwilliams/2013/06/20/the-most-valuable-business-commodity-trust/#5c7539da6500>
7. <https://www.ou.edu/deptcomm/dodjcc/>

Appiah, (2006) “Cosmopolitanism: Ethics in a World of Strangers”, Norton

Anderson (2009) “Why Do Dominant Personalities Attain Influence in Face-to-Face Groups? The Competence-Signaling Effects of Trait Dominance”  
[https://pdfs.semanticscholar.org/d0d8/13eec040dce91df0012767ac45d9a7694b07.pdf?\\_ga=2.130622479.1083641897.1546406941-103297660.1546406941](https://pdfs.semanticscholar.org/d0d8/13eec040dce91df0012767ac45d9a7694b07.pdf?_ga=2.130622479.1083641897.1546406941-103297660.1546406941)

The Atlantic “Wikipedia Editors For Pay” <https://www.theatlantic.com/business/archive/2015/08/wikipedia-editors-for-pay/393926/>

Big 5 Personality Traits [https://en.wikipedia.org/wiki/Big\\_Five\\_personality\\_traits](https://en.wikipedia.org/wiki/Big_Five_personality_traits)

Business Insider (2015) “Personality Types That Clash at Work”  
<https://www.businessinsider.com/personality-types-that-clash-at-work-2015-5>

De Young, Quilty and Peterson (2007) “Between Facets and Domains: 10 Aspects of the Big 5” <https://experts.umn.edu/en/publications/between-facets-and-domains-10-aspects-of-the-big-five>

Edelman Trust Barometer Global Report 2018  
<https://www.edelman.com/trust-barometer>

EPIC Oxford Report “Assessing the accuracy and quality of Wikipedia entries” [https://upload.wikimedia.org/wikipedia/commons/2/29/EPIC\\_Oxford\\_report.pdf](https://upload.wikimedia.org/wikipedia/commons/2/29/EPIC_Oxford_report.pdf)  
 Feezell and Krupnikov (2018) What’s True and Fake About The Facebook Effect <http://behavioralscientist.org/whats-true-and-fake-about-the-facebook-effect/>

Fleeson (2001) “Toward a Structure and Process-oriented View of Personality: Traits as Density Distributions of States.” <https://personality-project.org/revelle/syllabi/classreadings/>

[fleeson.2001.pdf](#)

Forbes “Trust, the Most Valuable Business Commodity” <https://www.forbes.com/sites/davidkwilliams/2013/06/20/the-most-valuable-business-commodity-trust/#5c7539da6500>

Harder, Velasco, Evans and Rockmore (2015) “Measuring Verifiability in Online Information” <https://arxiv.org/pdf/1509.05631v2.pdf>

Hwang et. al. “(2014) Drug Safety in the Digital Age” <https://www.nejm.org/doi/10.1056/NEJMp1401767>

Human Computation Blog (2013) CSCW 2013: Your process is showing... <https://humancomputation.com/blog/?p=4214>

IIHI social media report “Doctors’ #1 Source for Healthcare Information: Wikipedia” <https://www.theatlantic.com/health/archive/2014/03/doctors-1-source-for-healthcare-information-wikipedia/284206/>

Intersectionality <https://en.wikipedia.org/wiki/Intersectionality>

Lucassen, Shraagen (2010) “Trust in Wikipedia: how users trust information from an unknown source” <http://www.ra.ethz.ch/CDstore/www2010/wicow/p19.pdf>

Lucassen, Muilwijk, Noordzij, Shraagen (2012) “Topic familiarity and information skills in online credibility evaluation” <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.22743>

Meyer (2015) “Getting to Si, Ja, Oui, Hai, and Da, Harvard Business Review <https://hbr.org/2015/12/getting-to-si-ja-oui-hai-and-da>

Meyer (2017) “How Cultures Across the World



Approach Leadership”, Harvard Business Review, <https://hbr.org/video/5476393165001/how-cultures-across-the-world-approach-leadership>

Mothe and Sahut (2108) “How trust in Wikipedia evolves” <http://www.informationr.net/ir/23-1/paper783.html>

Motherboard “Why People Trust Wikipedia More Than the News” [https://motherboard.vice.com/en\\_us/article/ae37ee/in-defense-of-wikipedia](https://motherboard.vice.com/en_us/article/ae37ee/in-defense-of-wikipedia)

The New Yorker, “An Open Letter to Wikipedia” <https://www.newyorker.com/books/page-turner/an-open-letter-to-wikipedia>

PC World [https://www.pcworld.com/article/170874/The\\_15\\_Biggest\\_Wikipedia\\_Blunders.html](https://www.pcworld.com/article/170874/The_15_Biggest_Wikipedia_Blunders.html)

Quora “Is Wikipedia Trustworthy?” <https://www.quora.com/Is-Wikipedia-trustworthy/answer/Andreas-Kolbe>

Schmitt (2007) “Geographic Distribution of Big 5 Personality Traits” <http://www.toddshackelford.com/downloads/Schmitt-JCCP-2007.pdf>

Schwartz and Morris (2011) “Augmenting Web Pages and Search Results to Support Credibility Assessment” [https://www-cs.stanford.edu/~merrie/papers/WebCredibility\\_CHI2011.pdf](https://www-cs.stanford.edu/~merrie/papers/WebCredibility_CHI2011.pdf)

Times Online “Wisdom? More like dumbness of the crowds” [https://web.archive.org/web/20110814104256/http://www.timesonline.co.uk/tol/comment/columnists/guest\\_contributors/article2267665.ece](https://web.archive.org/web/20110814104256/http://www.timesonline.co.uk/tol/comment/columnists/guest_contributors/article2267665.ece)

The Utopian Fantasy of Rational Debate On the Web”, Nov 2017 <https://www.urbandaddy.com/articles/40999/kialo-is-an-internet-unicorn>

Washington Post “Conspiracy videos? Fake news? Enter Wikipedia, the ‘good cop’ of the Internet” [https://www.washingtonpost.com/outlook/conspiracy-videos-fake-news-enter-wikipedia-the-good-cop-of-the-internet/2018/04/06/ad1f018a-3835-11e8-8fd2-49fe3c675a89\\_story.html?noredirect=on&utm\\_term=.3abf96bc9826](https://www.washingtonpost.com/outlook/conspiracy-videos-fake-news-enter-wikipedia-the-good-cop-of-the-internet/2018/04/06/ad1f018a-3835-11e8-8fd2-49fe3c675a89_story.html?noredirect=on&utm_term=.3abf96bc9826)

Whoseknowledge.org

Wikimedia Brand Awareness, Attitudes, and Usage, Executive\_Summary [https://meta.wikimedia.org/wiki/Strategy/Wikimedia\\_movement/2017/Sources/Brand\\_awareness,\\_attitudes,\\_and\\_usage\\_-\\_Executive\\_Summary](https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Sources/Brand_awareness,_attitudes,_and_usage_-_Executive_Summary)

Wikimedia Blog, “Doctors use, but don’t rely totally on, Wikipedia” <https://blog.wikimedia.org.uk/2012/04/doctors-use-but-dont-rely-totally-on-wikipedia/>

Wikimedia Donor Survey: “What can we do to make Wikipedia better?” <https://docs.google.com/spreadsheets/d/1VaRS7PjOgdtB7HWz5QKue6PdFiNGTNZvStbfk7VRjPw/edit#gid=0>

Urbandaddy.com “Kialo Is An Internet Unicorn

