



Department of Infrastructure, Transport, Regional Development & Communications

San Francisco, CA, USA - November 11, 2021

Subject: Wikimedia's Submission for the Consultation on a draft Online Safety (Basic Online Safety Expectations) Determination 2021

The Wikimedia Foundation (WMF) welcomes the opportunity to comment on the Draft Online Safety (Basic Online Safety Expectations) Determination 2021 (“BOSE”), as follow up on the passage of the Online Safety Act 2021. Although Australia’s efforts to create a safer internet are lauded, such efforts should not undermine existing [effective community-driven content moderation practices](#) and should align with international privacy and human rights standards to safeguard freedom of expression.

WMF is the nonprofit that hosts Wikipedia and several other free knowledge projects. We protect the values and policies that allow free knowledge to thrive. Above all, we support the communities of volunteers around the world who edit, improve, and add knowledge across Wikimedia projects.

Wikipedia is among the [most visited websites](#) in Australia, and in October 2021 pages on Wikipedia were viewed [268 Million](#) times by users in Australia. Wikipedia’s governance model includes numerous policies, developed by the site’s volunteer contributors, that regulate content on Wikipedia. Editorial decisions are made by a decentralised group of users in accordance with internal mechanisms, and [content policies](#) are updated through public discussion and consensus from editors without WMF’s participation. Most vandalism (edits that do not meet Wikipedia’s [reliability](#) and [neutrality standards](#)) are addressed within [six minutes on Wikipedia](#). This system allows Wikipedia to present verifiable information to readers across the world and to run a global website with a staff of less than 600.

Wikipedia’s community governance model isn’t all that unique, however, as online message boards have used volunteer moderators for decades. Prescribing detailed implementation steps in the BOSE will affect smaller platform operators and community-led models disproportionately compared to large tech companies who have the infrastructure and resources to ensure compliance. The Australian government’s approach to regulating online

Imagine a world in which every single human being can freely share in the sum of all knowledge.



platforms needs to strike a balance that facilitates Australia's full participation in the digital global economy, while ensuring that Australians can access an internet that is safe to use and have their privacy protected.

On Division 2 - Expectations regarding safe use

Automatic content detection

The additional expectations regarding safe use include proactive measures on the part of platforms to minimise illegal or harmful content or activities. This would require WMF to intervene in functioning and established processes to ensure information on Wikipedia and the other Wikimedia projects lives up to high quality standards. A vast majority of content moderation actions on Wikimedia projects are taken by volunteer editors. These community content moderation processes have been shown to work effectively and to provide users with adequate complaint mechanisms to challenge removal, which are important safeguards for freedom of expression. In addition, Wikipedia's [Offensive Material Policy](#) outlines all the aspects that contributors need to consider when including content on Wikimedia projects that may be potentially harmful or offensive. For WMF to be forced to seize control of such processes and deploy automatic tools to detect and remove content, would risk the very existence of Wikipedia as a people-built encyclopedia. Therefore, **Wikimedia recommends that detection of content should allow for decentralised and community-based mechanisms rather than requiring automation.**

While large social media companies may proclaim the effectiveness of their automated content detection systems, these systems are still relatively limited in what they can detect accurately. As we see in the case of copyright enforcement, when automated processes are used to detect and remove content online, the consequence is often the [over-removal of entirely legal, legitimate content](#). These tools have been [shown to be subjective and to not capture nuances](#) and contextual variations of human speech. Additionally, the fundamental flaws in the underlying datasets and tools [cannot be fully mitigated with human oversight](#). If companies are expected or forced to replicate these models on content which requires additional context or tone to interpret, broad over-removals will increase and might perpetuate [bias](#). In fact, **by incentivising the use of automated filters, the additional expectation to detect and address (encrypted) materials threatens to interrupt effective content moderation systems developed and implemented by individual communities. Thus, Wikimedia strongly suggests that automated content detection systems should not be encouraged.**

Imagine a world in which every single human being can freely share in the sum of all knowledge.



Privacy and anonymity

In line with Wikimedia's principled support for privacy, WMF does not track people across the internet. We also do not display ads or sell user data to anyone. In fact, in order to ensure that the readers and contributors to Wikipedia and the other Wikimedia projects are safe, we collect very little personal information. The information we collect is stored for only a short period of time. Storage of contributor identifying information would prejudice the privacy of contributors and may discourage some people from contributing at all.

Against this background and in the context of expectations around curbing the spread of illegal or harmful content, it is important to note that the Wikimedia projects—including Wikipedia—are not built to capture the attention of people through amplification of content. Content does not go viral on Wikipedia, for instance, and the website's operational model encourages participation and deliberation over mere consumption.

WMF supports the spirit of preventing people from using anonymous accounts to “*engage in activity that is harmful*” (proposed step 9(2)(a)). Wikipedia already prohibits the use of so-called Sockpuppets to anonymously engage in abusive conduct. However, in regards to the expectation of “*verifying the identity or ownership*” of anonymous accounts, we consider it important to facilitate pseudonymous contributions to public interest knowledge platforms. Wikimedia's contributors also benefit from pseudonymity as it preserves their privacy and creates a place of safety. Disrupting anonymity or pseudonymity for online contributors would negatively impact the safety and security online of Australian Wikimedians. In addition, the ability to access Wikipedia privately and without revealing their real name and personal information is all the more important for people who live in regions and countries where human rights, including freedom of expression, are not protected.

In light of the robust internal mechanisms on decentralised content moderation, we submit that it would be inappropriate to require hosts of knowledge platforms to obtain and store personal information of contributors not necessary for any operational purpose, but with the goal of using that information for the implementation of the Online Safety Act. Such a requirement would be in contradiction with current international privacy and human rights standards. **WMF does not support any proposal to provide for a mechanism requiring collaborative knowledge projects, such as Wikipedia, to collect additional personal**

Imagine a world in which every single human being can freely share in the sum of all knowledge.



information of contributors and make that information available to the eSafety Commissioner.

On Division 3 - Expectations regarding certain material and activity

Considering the nature of platforms and content

Wikimedia's mission is to empower and engage people around the world to collect and develop educational content under a free licence or in the public domain, and to disseminate it globally. To ensure the encyclopedic nature and quality of the content that is hosted by Wikimedia, editors have developed and enforced content policies that define what kind of content can live on the websites and how it is presented and discussed. These policies and enforcement mechanisms allow Wikipedia to be an open platform where people come to learn more about the world in which they live. The openness of the platform and the fact that anyone can contribute to it, is what makes it possible for people to help ensure the quality of content as well as to prevent illegal or harmful content from being available on Wikimedia's websites. Notably, this is a truly global effort: what is accessible as encyclopedic material in one part of the world, should also be accessible elsewhere. For Wikipedia, as a collaborative project that is jointly curated by volunteers from across the globe, this means that editors and readers see the same version of the website, no matter where they are based and can work on articles together.

We would also like to highlight that Wikipedia and the other Wikimedia projects are written for a general audience. Therefore, limiting access to all information that might meet the definitions of "harmful content" for vulnerable groups under the caveat of expectations on detection of harmful materials (including within encrypted services) would be antithetical to the encyclopedia's core tenets.

Protecting children

The privacy of children on Wikimedia projects such as Wikipedia is of paramount importance and there are [policies](#) in place dictating how children's privacy should be protected. Yet, in order to comply with the draft expectations as written, WMF would have to collect information about users' ages and protected characteristics, even where collecting such information could present a personal risk to that user.

Imagine a world in which every single human being can freely share in the sum of all knowledge.



As an encyclopedic website, users can read and browse Wikipedia to their heart's content without creating an account. This is different from social media websites where you have to create an account loaded with personal data in order to access the basic function and utility of the website. We therefore suggest that expectations to implement age assurance mechanisms cannot be the same for all kinds of internet platforms; specifically, platforms that serve a public interest and do not operate on an ad-based business model and therefore carry less systemic risk to expose readers to illegal or harmful content should not be expected to implement the same kind of safety measures as engagement-based platforms.

On Division 4 - Expectations regarding reports and complaints

As with previously discussed expectations, policy makers need to consider the diversity of services on the internet that would be affected by enforcing strict rules for platforms. The proposed steps should not be conceived in a way that is tailored only for large tech companies while disregarding existing reporting and complaint handling mechanisms on smaller platforms. On Wikipedia, responsibility for ensuring quality of information and compliance with terms of use and policies is shared among the volunteer editors and WMF as a host. Everyone is invited to help with this task, foremost by directly editing articles or removing problematic content. For a digital society, transparency in decision-making about what should or should not be available on the internet is paramount. Wikipedia's content moderation mechanisms further such transparency and explicitly allow for everyone to help in the process, e.g. by providing [instructions and guidance](#) on removing inappropriate content.

In addition to the openness for participation in content moderation actions, the Wikipedia community has developed mechanisms to reach the so-called [Volunteer Response Team](#) regarding problematic conduct or content. Following a principle of subsidiarity, WMF can be reached via email in cases of emergency or serious harm. Yet, as a non-profit, the capacity of WMF to ensure the safety of Wikimedia users is limited and needs to be preserved. Expectations and proposed steps for reporting and complaint mechanisms need to consider that shared responsibilities between WMF, as the host of Wikipedia, and the community of editors are a feature, not a bug. The decentralized nature of this work makes it possible to operate this independent source of free knowledge for the world.

One important characteristic of most Wikimedia projects, especially of Wikipedia, is that they are organised by language - not by region. This means that a certain language version of Wikipedia, English Wikipedia or Greek Wikipedia, for example, does not include different

Imagine a world in which every single human being can freely share in the sum of all knowledge.



features or varying versions of the same articles depending on a reader's location. To mandate the provision of "*information and guidance on how to make a complaint to the Commissioner*" on a Wikipedia page in Australia would interfere with this principle. That would be better suited to the eSafety Commissioner's website.

On Division 5 - Expectations regarding making certain information accessible

In accordance with our deep commitment to transparency, Wikimedia supports making key policies and Terms of Use easily available (and understandable) for users of internet services and platforms. How this information is ideally presented to users will depend a lot on the type of service. We submit that we should ensure that these policies and terms are reasonably proportionate to the user's benefit. As discussed above in prior sections, Wikipedia's policies help to create a harmonised and productive environment and are in large part what makes Wikimedia projects so reliable and relied upon.

Conclusion

Wikimedia's mission is deeply rooted in the right to participation in knowledge and culture. We laud the efforts to ensure a safe internet experience for everyone, especially as public discourse and the exchange of knowledge and information are increasingly digitally mediated and happening via internet platforms. We appreciate the opportunity to give input on the draft of BOSE, but remain concerned with how narrowly tailored those are around the business models of for-profit internet platforms that seek to capture their users' attention through targeted advertising. We call on the Australian Department of Infrastructure, Transport, Regional Development and Communication Ministry as well as the eSafety Commissioner to consider the diversity of internet services and platforms when articulating and enforcing these expectations. Websites like Wikipedia depend on decentralized decision-making and shared power between a host and a community of contributors. Any expectations and rules that change this core tenet would put the functioning of the Wikimedia projects in peril.

Sincerely,
Jan Gerlach, Director of Public Policy
Wikimedia Foundation

Imagine a world in which every single human being can freely share in the sum of all knowledge.

wikimediafoundation.org · 1 Montgomery St, Suite 1600, San Francisco CA 94104