



HM Government

The Government Report on Transparency Reporting in relation to Online Harms

December 2020

- We also have started to include state-backed information operations datasets, which were released to the public to empower research and awareness of these campaigns. We now host over 35 different datasets that we believe are connected to state-backed information operations, and hosted our first research workshop on the data with the Carnegie Partnership for Countering Influence Operations in July 2020.
- We want to empower research that can advance public understanding of critical issues online. For example, to further support Twitter’s ongoing efforts to protect the public conversation and help people find authoritative health information around COVID-19, we released a new endpoint into Twitter Developer Labs in April 2020. This is to enable approved developers and researchers to study the public conversation about COVID-19 in real-time. The data can be used to research a range of topics related to the coronavirus pandemic, including areas like the spread of the disease, the spread of misinformation, crisis management within communities and more. Making this access available for free is one of the most unique and valuable things Twitter can do as the world comes together to protect our communities and seek answers to pressing challenges.
- We remain deeply committed to transparency at Twitter - it continues to be one of our key guiding principles. We have welcomed the opportunity to participate in the working group, and look forward to continuing to work with government, civil society and the wider community on these important issues.

Facebook

- Facebook supports the idea of a regulatory framework for online content that ensures companies are making decisions about online speech in a way that minimizes harm but also respects the fundamental right to free expression. Regulation should seek to balance these often conflicting issues and bring more procedural accountability for platforms. In order to do so, regulation should include requirements for companies to publish their content standards, create mechanisms to report violations of these standards, provide response to decisions, and notice when content is removed.
- Facebook is transparent about its Community Standards, the global set of policies that outlines what is and is not allowed on Facebook and is publicly available on our website. Our Community Standards apply to everyone, all around the world, and to all types of content. They are based on feedback from our community and the advice of experts in fields such as technology, public safety and human rights, collected in various forms, including via our Content Policy Forum, a meeting that we hold every two weeks to discuss new policies or amendments to existing policies. Minutes from these meetings are public.
- It is also important for service providers to provide transparency into how their systems are performing, however the type and nature of the measurements or metrics used should not be so fixed as to hinder efforts to accurately respond to the changing dynamics of reporting on online content. Different types of services may require different levels of transparency, depending on the size and nature of the provider.

- Facebook already does so by sharing regular transparency and enforcement reports, such as the Community Standards enforcement report detailing how much content we remove for violating certain of our policies, how much of that content was detected proactively by our automated tools, how much content was appealed when people believed we had made a mistake, and how many of those appeals were successful. Additionally we regularly publish another report that includes metrics on the number and nature of legal requests we receive from governments and other entities around the world – including requests for data and requests to restrict access to content which they believe violates local law.
- Given the dynamic nature and scale of online speech and the different expectations of users of their experience online, any system operating at scale and for a global user base will be imperfect. For this reason, in order to safeguard freedom of expression, it is essential for platforms to be transparent about its decisions and have appropriate redress mechanisms. Facebook provides feedback and updates to users that report content and informs users whose content has been removed. Additionally, we give users the possibility to appeal our decisions regarding certain content that we took action on and certain content that was reported but not acted on.



HM Government

ISBN 978-1-5286-2312-4
CCS1220695430