



11 June 2021

Cynthia Wong
Human Rights Director,
Legal

Twitter, Inc.
1355 Market St #900
San Francisco, CA 94103

(415) 222-9670
cynthiaw@twitter.com
@cynthiamw

Dear Mr. Micek and Ms. Oribhabor:

Thank you for your letter regarding Twitter's performance on the 2020 Ranking Digital Rights Index, and for your continued efforts to advance the rights of freedom of expression and privacy in the context of digital platforms. We appreciate the opportunity to engage with Access Now about our progress, and to identify opportunities to improve our performance on issues that span the Index and have meaningful, real-world implications for the rights of people around the world.

We take seriously our responsibilities to [defend and respect](#) the rights of people using our service.

As you note, Twitter has taken significant steps to invest in meaningful transparency, leading the market in disclosing information about the actions we take to enforce our rules, whether through content removal or account suspensions. The new [Twitter Transparency Center](#) was launched in August 2020, not only providing a centralized place to learn more about our broad array of transparency efforts, but aiming to do so in a way that aligns with the best practices set forth in the [Santa Clara Principles on Transparency and Accountability in Content Moderation](#). For example, the recent changes expand the number of Twitter Rules covered and add more granularity on the actions we take. In addition, we continue to provide requests that result in content removal to [Lumen](#), so that anyone, anywhere, can understand the nature of the requests we receive. Finally, Twitter stands alone as the only major social media platform to disclose comprehensive data on state-linked [information operations](#) we have removed.

While we have made progress in many areas, we appreciate that there is opportunity to go further in our approach to transparency to support human rights and accountability and to earn the trust of the people who use our service. As our CEO, Jack Dorsey, has highlighted in recent [testimony](#), transparency is critical to building and maintaining that trust.

As regards risk assessments of our business operations and services, a number of teams are engaged in integrating a human rights lens across how we craft and enforce our rules and how we develop our products in the company's day-to-day operations. Before launching a new policy, we evaluate the risks to rightsholders, examine how our policies could be abused or exploited, and put mitigations in place to minimize those risks. By way of example, our updated [Anti-Discriminatory Ads Policies](#) restricts advertisers from employing targeting options to prevent discrimination against protected groups in areas which have been, and continue to be, sites of egregious

discrimination, such as access to housing, credit, and employment. Before launching new products, a number of teams are engaged to assess implications including but not limited to privacy, transparency, user control, and freedom of expression. When we consider opening up new offices, we assess the human rights implications of our presence and how our operations might interact with local dynamics on the ground. We also leverage the expertise of our [Trust & Safety Council](#)'s Digital and Human Rights Advisory Group to improve our understanding of the human rights risks posed by our policies and products, and to help inform our approach to risk mitigation.

In addition, our teams working on privacy and data protection ensure that products and processes, including those related to advertising and machine learning, go through a privacy review with the goal of identifying risks and addressing them prior to launch. As part of the review, we assess how and what data is used and stored, whether it is shared, and how a product or process may impact those who use it or whose data is processed.

Our Human Rights team is responsible for driving our cross-functional efforts to implement the UN Guiding Principles on Business and Human Rights. The team undertakes a range of human rights due diligence to proactively surface and manage risks. This includes conducting targeted human rights impact assessments on the implications of our operations and services, with particular attention to impacts on vulnerable and high-risk communities and on freedom of expression, privacy, and discrimination risks.

We have also recently launched our [Responsible ML Initiative](#), aimed at understanding the impact of machine learning at Twitter. This initiative is led by our ML Ethics, Transparency and Accountability (META) team, a dedicated group of engineers, researchers, and data scientists across the company, who work to assess the unintentional harms in our algorithms, including harms related to bias and discrimination.

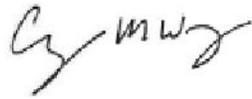
As one example, we recently conducted a bias assessment of our [image cropping algorithm](#), after hearing feedback that it was not serving people equitably. After testing our model for gender and race-based biases, we shared our findings publicly, addressing both what our model revealed in terms of bias, and the actions we were committing to take in response.

We have also sought to open up our policy development process to [public comment](#), as we committed to do in 2018. This initiative began with our [dehumanization policy](#), and continued with our policy for [synthetic and manipulated media](#). In March, we called for public input into our approach to [world leaders](#), as well as consulting with human rights experts and organizations, so that we could better inform the development of our policy framework. To ensure that a global perspective was reflected in the feedback, we made submissions available in fourteen languages.

We are constantly working to deepen our commitment to human rights, and appreciate the role that Access Now and Ranking Digital Rights play in bringing accountability and transparency to the sector. The feedback we

receive from the human rights community is not only of value to us, but is critical in helping us to better uphold our core values, which include defending and respecting the human rights of everyone using our service.

Sincerely,

A handwritten signature in black ink, appearing to read 'Cynthia M. Wong'. The signature is fluid and cursive, with the first name 'Cynthia' being the most prominent part.

Cynthia M. Wong
Human Rights Director, Legal