GLOBAL ALLIANCE AGAINST DIGITAL HATE AND EXTREMISM

An international movement mobilising collective action for a safe, fair, and just online experience for people everywhere.
www.globalalliance.tech

April 22, 2024

Delivered via electronic mail

DARIO AMODEI
CEO, Anthropic
548 Market Street
San Francisco, CA 94104

JASON CITRON
CEO, Discord
444 De Haro Street, #200
San Francisco, CA 94107

SUNDAR PICHAI
CEO, Google
1600 Amphitheatre Parkway
Mountain View, CA 94043

ADAM MOSSERI
Head of Instagram
1 Hacker Way
Menlo Park, CA 94025

MARK ZUCKERBERG
CEO, Meta
1 Hacker Way
Menlo Park, CA 94025

SAM ALTMAN
CEO, Open AI
3180 18th Street
San Francisco, CA 94110

OFIR EYAL
CEO, Rakutan Viber
2, rue du Fossé
Luxembourg, 1536
Luxembourg

STEVE HUFFMAN
CEO, Reddit
303 Second Street
San Francisco, CA 94107

SHOU ZI CHEW
CEO, TikTok
5800 Bristol Parkway, Suite 100
Culver City, CA 90230

DANIEL CLANCY
CEO, Twitch
350 Bush Street, 2nd Floor
San Francisco, CA 94104

NEAL MOHAN
CEO, YouTube
901 Cherry Avenue
San Bruno, CA 94066

LINDA YACCARINO
CEO, X
1355 Market Street, #900
San Francisco, CA 94103

Dear Technology Leaders:

Social media and online platforms can be a positive or negative for democracy, and in recent years, we have witnessed firsthand the profound effect it has had on elections and countries around the world. We have seen a weakening effect on struggling and strong democracies and an intensifying effect on authoritarian regimes. Overall though, there is no doubt that the abuse of social media has had a negative impact on democracies worldwide. It is specifically important that politicians and public figures not use social media to spread disinformation, hate speech, and other violent content. A body of research suggests the incendiary rhetoric of political leaders can make political violence more likely, incites violence and gives it direction, complicates the law enforcement response, and increases fear in vulnerable communities. This is why it's critical that you, as company leaders, treat all users the same and enforce your content moderation rules for public and political figures.

Political leaders' remarks do not disappear on social media, especially as the social media platforms' algorithms are designed to amplify more incendiary remarks, quickly magnifying rhetoric against their political opponents, minority groups, and other targets. Leaders with large social media followings will see their remarks shared with millions of followers. Globally, public and political figures use social media marketing tools for campaigns, not only because it is effective, but they are also able to bypass rules and norms of traditional media. And for those who wish to engage in hate and demonizing speech, the results can be a matter of life and death.

Your companies' policies state that public figures and politicians are subject to your rules, but you also have 'public interest' or 'newsworthiness' exemptions, effectively rendering the rules useless. And where political ads are

(continued)

allowed, none are fact-checked, a tremendous source of disinformation. Additionally, we have seen most of your companies reduce the size of your content moderation and elections integrity teams in the last few months.

One of the most important commitments that your companies can make is to apply your own content rules and community standards to political figures as you would other users, especially hate speech and incitement to violence rules. You should do this without regard to political position or economic considerations and in all languages. Most content moderation today is driven by automated systems primarily trained in English.

At least 60 countries representing half the world's population are expected to hold elections in 2024, an unprecedented number that includes some of the world's biggest democracies, more fragile democracies, and some nations where there is a continued weakening of civil and human rights. More countries have moved away from democracy rather than toward it, particularly over the last several years, including in countries where democracy was thought to be firmly established. And since 2017, the number of countries moving toward authoritarianism are more than double those that are moving toward democracy. Given the equally unprecedented potential influence of your platforms and those adept at manipulating the platforms, and the additional concerns emerging as a result of the expansion of AI and its unknown effects, it is vital that the platforms prepare *now* to do all they can to protect democracies and elections around the world.

There is much that we as human rights and democracy advocates would have your companies do to protect the humans in this world, our elections, and our global democracies. But today, we are insisting that you enforce your already existing policies fairly and equitably to protect democracy and marginalized communities globally. You can do this by committing to eliminating your exemptions for public and political figures and guaranteeing that human moderators, well-trained in your rules, with cultural and language competency, review posts from public figures and politicians. The risks are too high to rely solely on artificial intelligence.

Members of this coalition, representing many regions from around the world, are prepared to share their expertise in discussions about the need and implementation of this moderation.

We urge you to fulfill your previous commitments as leaders and companies whose stated values align with free and open elections, inclusive democracies, and human rights and ask that you reply to this letter at contact@globalextremism.org by May 6, 2024.

Sincerely,

GLOBAL ALLIANCE AGAINST DIGITAL HATE AND EXTREMISM

STEERING COMMITTEE

Global Project Against Hate and Extremism

AfroLeadership

Africa Sans Haine

AVAAZ

Conscious Advertising Network

Digital Action

Equality Labs

Fair Vote UK

Fighting Online Antisemitism

Greek Helsinki Monitor

'NEVER AGAIN' Association

Red Dot Foundation

The London Story