11 February 2021

The Center for Democracy & Technology welcomes the opportunity to provide comments on case 2021-001-FB-FBR, regarding the suspension of Donald Trump's account.

Account suspensions may be the most serious sanction a platform such as Facebook can impose. Suspension could be considered analogous to a prior restraint on speech; an average user whose account is suspended may permanently lose access to the audience of other users of that service and is effectively being punished for what the user might say in the future. "There is a strong presumption against the validity of prior censorship in international human rights law and in the case law interpreting national constitutional protections for freedom of expression." (See https://bit.ly/2Z6t5io for further discussion.)

That said, there are important differences between a state actor imposing a prior restraint on an individual's speech and a specific online service suspending a user's account. When a nation-state enforces a prior restraint, it can effectively silence an individual, prohibiting their speech across all types of media. An online service provider can only suspend an individual's account on its service(s) (though, of course, the size of the service and the role it plays in the Internet ecosystem will make the impact of a suspension more or less far-reaching). Most users of any given service will generally be able to find some other place to make their voices heard online; this is particularly true for political leaders, who have substantially greater access to traditional media and other speech venues than the average user.

At the same time, high-ranking political figures have an outsized potential to influence public sentiment and to incite violence. People often imbue statements from authority figures with more legitimacy and authenticity than statements from random individuals. As the Dangerous Speech Project has described, public officials may control "the power to deploy force against uncooperative audience members," which can give these officials' explicit or implied threats of violence more weight. (https://dangerousspeech.org/guide) The six-part threshold test from the Rabat Plan of Action states that "the speaker's position or status in the society" must be considered as part of the assessment of whether a statement is likely to incite violence.

 In view of the potential violence and physical injury that speech from political leaders can incite, account suspensions can be an appropriate enforcement action. Indeed, account suspensions generally are a crucial tool for online services to be able to enforce their content policies and to cultivate functional online communities. Suspensions give service providers the ability to exclude users who post egregious content or engage in repeat offenses, actions which can seriously undermine other users' willingness to participate in the community. (See, e.g., https://bit.ly/2OqVQ7j.) Many services even include in their Terms of Service a provision prohibiting "ban evasion," allowing them to suspend additional accounts being used to circumvent the original suspension to help maintain the integrity of their enforcement decisions. (See, e.g., https://bit.ly/3rEDhuC, https://bit.ly/3aTf2lF, https://bit.ly/2Z7YDEx.)

When considering account suspensions for prominent users, including political leaders, Facebook should conduct a contextual analysis of these users' posts that includes an evaluation of the likelihood that the speech will lead to violence, taking into account factors such as the real-world context and whether the speaker's prominence creates a heightened risk of inciting violence. Such contextual assessment of users' posts can be challenging to do at scale and without access to additional information. Thus, Facebook should also develop or improve their internal escalation process for moderation decisions about content posted by political leaders. This could include developing a list of leading political figures worldwide and providing frontline moderators, or other staff who may receive reports about political figures' accounts, with clear training and guidance about when to escalate reviews of content posted by these accounts. Identification of "political leaders" will itself require careful analysis informed by local context: these will likely involve heads of state and political parties, but may also include other high-ranking public officials and military leaders. Given the risk and intensity of political violence around the world, Facebook should develop specialized teams within its Trust & Safety structure that have the expertise to assess the risk that statements by political leaders will incite violence.

In addition to the capacity to assess the risk of harm, Facebook should develop – and communicate to users – a set of criteria that explain the circumstances in which it will consider imposing indefinite and permanent suspensions, as well as indicators for lifting an indefinite suspension or converting it into a permanent ban. Such factors may include the likelihood that the speaker will engage in further incitement as evidenced, for example, by repeated violations or a lack of contrition. Although an indefinite suspension is less severe than a permanent one, it poses an additional issue: users generally should be provided a clear idea of what punishment has been applied to their accounts, including the duration of any temporary suspensions, so that they can understand the terms of their continued use of the service (if any) and to decide whether to pursue an appeal. However, an indefinite suspension may be particularly appropriate in a situation where the real-world context is dynamic (e.g., an ongoing violent insurrection), and where it is unclear for how long there will be a heightened risk that the user's future posts will cause a risk of further violence.

For more information, contact Emma Llansó, Director, Free Expression Project, ellanso@cdt.org.