RESOLVED: Shareholders request that Meta Platforms Inc. ("Meta") report to shareholders on the effectiveness of measures it is taking to prevent and mitigate human rights risks in its five largest non-US markets (based on number of users) relating to the proliferation of hate speech, disinformation, and incitement to violence enabled by its Instagram and Facebook platforms. The report should be issued no later than June 1, 2025, prepared at reasonable cost, omitting proprietary and confidential information (including information specifically relevant to litigation or legal enforcement action).

WHEREAS: The dissemination of hatred that incites discrimination, hostility or violence violates international human rights standards[1]. Where content moderation systems have failed to effectively detect divisive content in non-English languages, there has been an associated increase in hate speech[2], disinformation[3], and incitement to violence. Meta's stakeholders and the public have repeatedly raised significant concerns regarding what appears to be an obvious lack of proportionate investment in content moderation resources and expertise in Meta's global majority markets. This issue, repeatedly flagged by reports from international organizations[4], its own Oversight Board[5] and CSOs[6], is critical in Meta's non-English speaking countries. This apparent lack of adequate resources and investment in content moderation is increasingly critical with the 2024 super election year and an estimated 2.6 billion people[7] taking to the polls globally. Media reports suggest Meta is putting in place advertising related mitigations[8] relating to the US elections. However, Meta has not published any measures to address such issues in non-Western, non-English speaking markets, that given the current inadequacy of effective content moderation are more vulnerable to the proliferation of hate speech, disinformation, and incitement to violence on their platforms.

We commend Meta's first transparency reports on Instagram and Facebook required under the EU Digital Services Act, providing detailed information on numbers of content moderators in local languages and overall users per EU country. Given Meta now appears to have the required data collection and reporting infrastructure to provide such detailed reporting on individual countries, the company should expand these transparency measures to key markets like India and Brazil[9]

---

[1] https://www.un.org/en/hate-speech/united-nations-and-hate-speech/international-human-rights-law
[2] https://www.bbc.com/news/world-africa-67275219
[3] https://digital-strategy.ec.europa.eu/en/news/commission-sends-request-information-meta-under-digital-services-act
[4] https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/Legislation/Case_2021_009-FB-UA.pdf
[5] https://www.theguardian.com/technology/2022/dec/06/meta-protecting-business-partners
[6] https://www.amnesty.org/en/latest/news/2023/10/meta-failure-contributed-to-abuses-against-tigray-ethiopia/
[7] https://www.unesco.org/en/articles/ahead-super-election-year-unesco-appeals-governments-around-world-protect-journalists-rights
[8] https://www.reuters.com/technology/meta-bar-political-advertisers-using-generative-ai-ads-tools-2023-11-06/
[9] https://www.statista.com/statistics/578364/countries-with-most-instagram-users/ &
https://www.digitalmarketingcommunity.com/indicators/instagram-active-users-penetrations-2018/

on a disaggregated basis to demonstrate the actual investment made to build multilingual capacity in content moderation. By doing so, Meta can address the persistent human rights risks which can and have had a negative impact on brand value and, indirectly, on its advertising revenue, as well as on diversified investment portfolios as viewed through a universal ownership lens.

Proponent suggests the report include data on the number of content moderators fluent in local languages in Instagram and Facebook's five largest non-US markets based on number of users and an assessment by external, independent, and qualified experts of the effectiveness of Meta's measures taken to meaningfully manage hateful content, disinformation, and incitement to violence on those platforms.