

# AFP public comment for Meta Oversight Board to review Meta's plans to expand community notes

December 9, 2025

Ensuring information integrity is central to AFP's work both as a global news agency and as the largest Fact-Checking team around the world. Although Meta has announced its progressive withdrawal from its Third-Party Fact Checking program and the roll out of Community notes, Meta declares it maintains its attachment to addressing (potentially) misleading content in its channels. AFP's views on making sure to adding context and verified information to online content as fact-checkers also seems to align with Meta Community Notes' aim to "add more context" to the users of their services. It is reassuring to acknowledge that as various stakeholders we are equally convinced of the absolute need to add more information, not less, to prevent any kind of censorship of public discourse.

We thus welcome Meta's questions to the Oversight Board ahead of a potential expansion of its community notes program out of the United States, to ensure that it won't turn out having the opposite effect to the one wished for.

Some caveats before diving into our detailed response:

- It must be noted that X is the only platform to have rolled out a Community Notes program at global level for now, most of the available research for now is thus linked to X's program, in a different environment but with processes in the ratings and consensus approach which seem similar.
- The experiment on Community Notes for Meta services has been done in a limited environment for now – limited to the United States, and on a short amount of time: it thus seems difficult to draw conclusions from this circumvented experience for a much larger deployment.
- The consideration that Meta seems to have on specific human rights issues per country (level of freedom of expression, absence of freedom of press, government restrictions on the internet, ...) in relation to the roll out of its Community Notes program is problematic: access to reliable information as a human right should be a priority everywhere. If Community Notes are amplifying issues in a context where human rights are at risk, they probably shouldn't be rolled out in the first place.

## **User-in-the-loop**

Trust from audiences in independent media and fact-checkers has been generally declining for years. Reuters Institute's [Digital News Report for 2024](#) mentioned that "just 40% of our respondents across all 47 markets say they trust most news." This is an ongoing concern for journalists. The format of Fact-Checking has been created exactly for this reason: as an open-source demonstration, replicable by readers, equipping them with the open-source tools and links to verify everything online by themselves. It thus seems a natural step, especially in the context of growing user generated content production, to involve platform users more in the verification process instead of having what can sometimes be seen as a top-down approach. We also see, notably through our Fact-Checking program on WhatsApp, that users can be [early](#)

[flaggers](#) of potentially misleading content. In this regard, crowd-based notes could be seen as a way for users to gain ownership, and maybe, from there, renewed trust in the process of the verification of online information. Indeed, [a study](#) published in September 2024 shows that when a note is attached to a misleading post, it reduces the reshares by 62%.

It seems nevertheless that the reality of the implementation of platform crowd-sourced programs as they are deployed for now is not adapted, especially in terms of scale.

### **Impact at scale**

The same study mentions that the Community notes on X are slow compared to the spread of the disinformation, but also that the notes themselves can be false, and that most of misleading posts do not have a note attached – as has also been demonstrated by [News Guard](#) and by [Science Feedback](#). Science Feedback also shows the disconnect between Community Notes and verified content by professional Fact-Checkers, with a strong disparity of results depending on countries (in Europe).

The [update](#) on Community Notes posted by Meta CISO Guy Rosen in September 2025 does not show a large number of Community Notes actually ending up being published (6% of notes published on 15,000 submitted notes over 6 months) – adding to that the difficulty to know how many of these notes are actually seen by users, after being published. In terms of impact at scale, and according to Meta's last [transparency report](#) under the Code of Conduct on Disinformation in the European Union, for Facebook alone, and from January to June 2025, over 24 million pieces of content were viewed and treated with a fact-check, and on average (depending on country) 60% of reshare attempts were not completed on treated content. This result appears quite similar to the impact of X's Community Notes in [the study](#) mentioned above. But Meta has announced that Community Notes will not be applied to advertisements – contrary to the Third-Party Fact-Checking program, which allows fact-checkers to protect online users from a [variety of scams](#) and [manipulation](#) – in accordance with Meta's own community guidelines.

Last but not least, looking at the trajectory of Community Notes on X, it seems that the writing of notes is generally [in decline over time](#) – and X bets on AI to play a larger role in the near future. This reliance on [AI chatbots for the verification of information](#), even in the context of [breaking news](#), poses a high risk for human rights, especially considering that AI chatbots are often super spreaders of misleading or false content.

### **Real life harm of online deception**

As members of the Third-Party Fact-Checking program, we are frontline observers of the real-life harm generated by deceptive online content.

In Asia alone, beyond South Asian countries like India and Bangladesh, rumors on Facebook in Myanmar regularly stoke [ethnic tensions](#) and spread [mass confusion](#) about the ongoing civil war there. In the Philippines, protesters are [misidentified in social media posts](#) claiming they are actors, leading to online harassment. And across Southeast Asian countries such as [Indonesia](#) and [Malaysia](#), scammers use deepfake videos to swindle people out of their savings.

In Africa, we have documented how AI is used to [solicit false donations](#) over the humanitarian crisis in Sudan – and how the situation in Sudan is manipulated in [various ways](#). We also see

how [inflammatory content](#) is shared [around Ethiopia](#), amid growing military tensions between Ethiopia and Eritrea.

The harm from misinformation on Meta's platforms are real – and the company must take careful measures to counter it.

### **A note on bias**

Meta has announced the discontinuation of the Third-Party Fact-Checking program based on accusation of partisanship, the fact-checkers of the program being “too politically biased”. AFP strongly refutes this accusation, based on our [own charters](#) on non-partisanship but also on the common rulebooks we apply with the [EFCSN](#) and the [IFCN](#) networks, for which our practices are verified by independent assessors. Professionalism puts in place increased boundaries on non-partisanship. This is not the case for X's Community Notes, as shown in [several reports](#). Because most of the notes are published by a reduced number of authors, “the top 10% of contributors producing 58% of all notes” as explained in [this recently published study](#), they are at heighten risk of manipulation. This could explain the small number of notes actually being published, because of a lack of consensus, especially on divisive and/or complex topics.

### **Common ground**

AFP believes that there could be a common ground for upgraded impact at scale through the inclusion of independent fact-checkers within the Community Notes program. This could be a way to increase the accuracy, the relevance and the streamlining of the crowd-sourced notes by adding fact-checkers in the loop, publicly and in a transparent manner, while allowing greater interactions with online users for increased trust.

As an example, before the change of hands of Twitter to X, AFP participated in the project Birdwatch in Latin America and the United States. Our role as fact-checkers consisted in the review of certain crowd-sourced notes when Twitter's moderation team assessed that further verification was needed. The aim was to limit as much as possible inaccurate or biased notes. To us, this was an impactful way to get crowd-sourced verifications and professional fact-checking to work hand in hand, for the greater good.

Collaboration between crowd-based notes and professional, independent fact-checkers could be made two folds:

- By allowing fact-checkers to produce notes and by fast-tracking them for publication.
- By ensuring the accuracy of the crowd-sourced notes with a verification system for faster publication.

We stand at your disposal to discuss it further.

### **CONTACT**

Grégoire Lemarchand

Deputy News Director for Digital Strategy and Director of Brand and Communications

<https://factcheck.afp.com/>