



Isolde Hegemann

October 21st, 2025

## AI fact-checks may be the best way to reduce Republicans' engagement with online political misinformation

0 comments | 3 shares

Estimated reading time: 6 minutes



*Online platforms now mean that political misinformation can spread easily and widely, potentially having implications for democracy. In new research, Isolde Hegemann looks at how artificial intelligence (AI) can be used to tackle online misinformation. She finds that AI fact-checking can be more effective compared to more traditional methods in reducing the likelihood that Republican social media users will share false information online from other Republicans.*

In his first term in office, President Trump was well-known for promoting incorrect or misleading claims on a **daily basis**. Now, in his second presidency, he has surrounded himself with an administration that uses misinformation **to**

push its policy agenda. One of the key groups that are more likely to be influenced by this misinformation push are US Republicans. In general, conservative and pro-Trump users have been found to be more likely to share posts referring to low-quality sources on X.

Recent changes in how social media companies fact-check posts might have crucial implications for how Republicans engage with misinformation they see online from other Republicans: on January 7th, Mark Zuckerberg discontinued the collaboration with independent fact-checkers across Meta's platforms, including Facebook, and instead announced the roll-out of a community-driven approach to flagging misleading posts. The move was met with criticism by expert commentators concerned about the quality and speed of Community Notes.

## Can AI fact-checking strategies reduce the spread of misinformation?

Beyond independent fact-checkers and other users, artificial intelligence (AI) offers new possibilities for filtering potential misinformation and quicker fact-checking. But in a 2020 study, AI did not reduce but actually increased the likelihood Republicans shared false information. However, even in the last two years, our information environment has been rapidly changing as generative AI has entered the mainstream and is now being used by many people in their everyday lives. This makes these findings worth revisiting.

There are many aspects to consider when trying to establish the advantages of different fact-checking strategies, such as the quality and accuracy of the label, speed of the fact-checker, and what posts get fact-checked (or not) and by whom. In my research, I focus on how users engage with different fact-checking labels – which provide context to social media posts – once they are presented with them. I look at what fact-checking strategies work best in reducing Republican social media users' engagement with misinformation provided by other Republicans.

"That AI-powered labels are the fact-checking strategy that proves most effective for Republicans has important implications for how we think about battling misinformation online and how to target key groups with interventions"

*Isolde Hegemann of LSE's Department of Government finds that AI fact-checks may be the best way to reduce Republicans' engagement with online political misinformation.*

[blogs.lse.ac.uk/usappblog](http://blogs.lse.ac.uk/usappblog)

LSE Phelan US Centre

*Photo by Hartono Creative Studio on Unsplash*

## Creating a simulated social media experience

To find out what fact-checking strategies work best, I use the experiment builder **Gorilla** to present 1450 Republican respondents recruited on Prolific with posts that they can interact with like they would in the real world. I use real statements by President Donald Trump from his speeches and Truth Social posts to ground the experiment. Some of the statements are true, but most are false. The fictional posts then endorse these statements. I then randomise when respondents see which post, and whether and what fact-checking strategy they see with each post. I also randomise several other post attributes, such as the profile name and picture and how much fictional engagement the post has already received.

Setting my study up in this way allows me to give the respondents an experience as close to reality as possible but retaining full control over what they are presented with. Because I randomise the different post characteristics for each new post, I can then estimate the causal effect of each fact-checking while controlling for other factors. This offers key advantages over using a more traditional survey experiment where respondent behaviour might not be as natural or using existing real-world data where user behaviour can vary dependent on many other influences present in the real world.

**AI has added context:**

This is true. 'The Hong Kong-based conglomerate that operates ports near the Panama Canal has agreed to sell shares of its units that operate the ports to a consortium including BlackRock Inc.' (CBS News, March 5, 2025).

**Megan Ross**

2h

Exciting news, confirmed by President Trump! A large American company announced they are buying both ports around the Panama Canal and lots of other things having to do with the Panama Canal and a couple of other canals.



GG React

t Repost

t Comment

**Other users have added context:**

There is no evidence that President Biden was replaced by a robotic clone and the statement constitutes a baseless conspiracy theory (New York Times, June 1, 2025).

**Eric Sanders**

1h

Endorsed by President Trump on Truth Social! : 'There is no #JoeBiden - executed in 2020. #Biden clones doubles & robotic entities are what you see. >#Democrats dont know the difference. #Steel #ussteel #MAGA #MAHA'



GG React

t Repost

t Comment



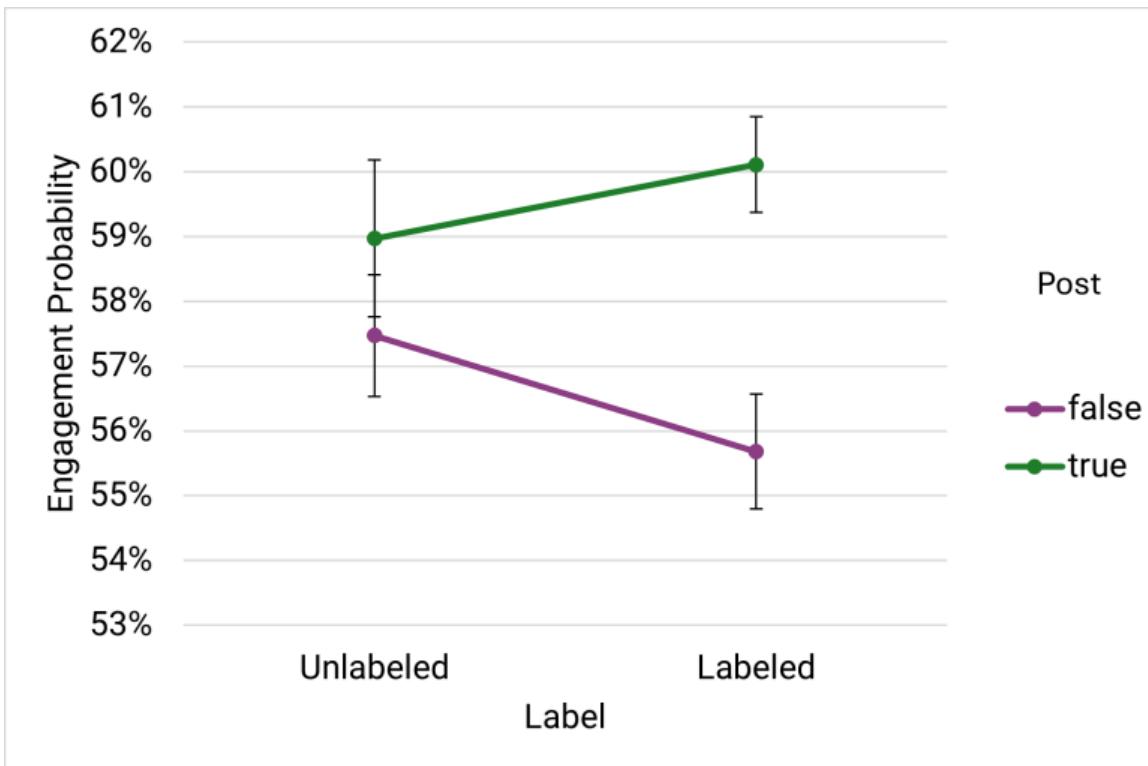
## Fact-checking works on Republicans – especially if done by AI

To measure engagement, I look at whether respondents reacted to a fictional post with one of the reactions available, reposted it, or commented on it. My analysis finds that fact-checking works in reducing engagement with posts featuring political misinformation and helps users discern between false and true statements. I found that:

**1) Fact-checking labels reduce engagement with misleading posts.** As Figure 1 shows, respondents significantly reduce their engagement with a post (a 1.8 percentage point decrease from a 57.5 percent engagement baseline for unlabelled posts) when they read a fact-checking label that offers them context.

**2) Fact-checking increases discernment between true and false posts.** To make sure that fact-checking interventions really work, **it is important** to not just look at whether they reduce engagement with false posts, but also whether they help users distinguish between false and true posts. If we would see indiscriminate reductions in engagement, then the fact-checks would not fulfil their intended purpose. In the case of my research, fact-checks help Republican users distinguish between true and misleading information.

**Figure 1 – Effects of fact-checking labels on engagement with true and false posts**

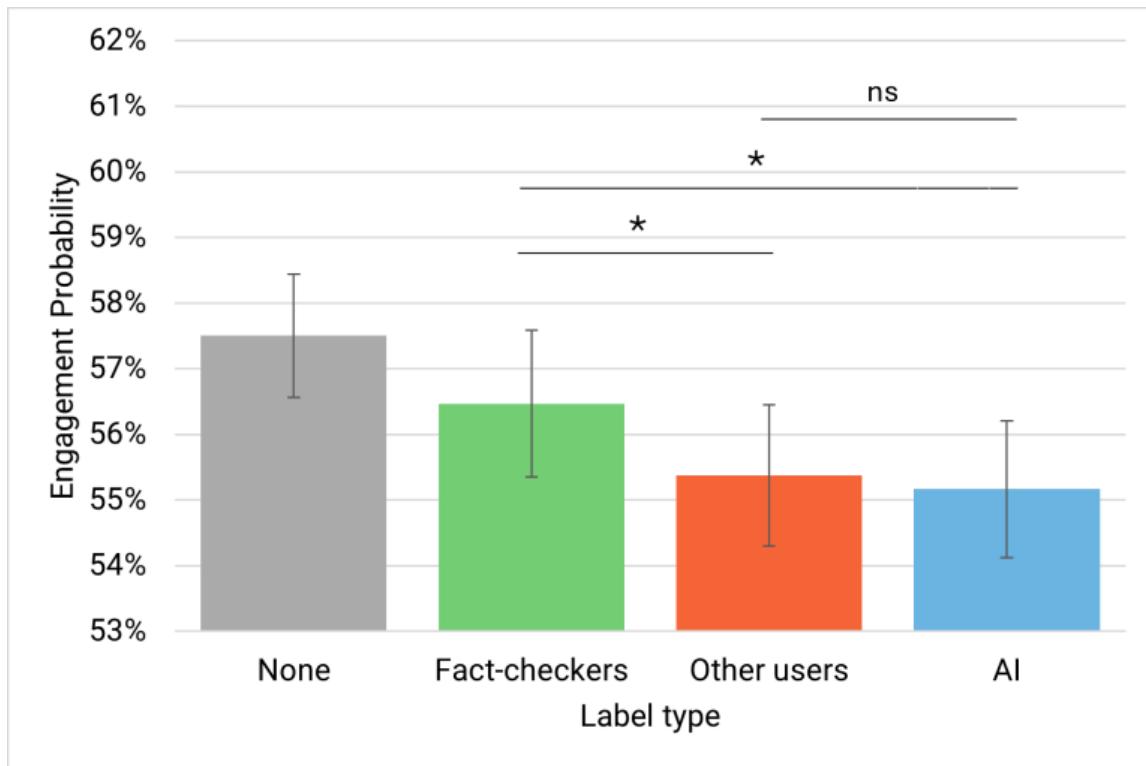


*Note: N=1209 respondents. Includes outliers, excludes non-engagers.*

*Standard errors are clustered at the respondent level. Error bars represent 95% CIs.*

3) AI works best at reducing engagement with false posts. When disentangling the different fact-checking strategies, I find that AI is the most effective strategy across models, both compared to not using a label and by reliably outperforming independent fact-checkers (Figure 2). Community Notes are only more effective than independent fact-checkers in my main model, but the effects are not robust to the different checks I run. And strikingly, Republicans do not reliably change their engagement behaviour because of a label by an independent fact-checker.

**Figure 2 – Effects of different fact-checking labels on engagement with false posts**



Note: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ . N=1209 respondents. Includes outliers, excludes non-engagers. Standard errors are clustered at the respondent level. Error bars represent 95% CIs.

That AI-powered labels are the fact-checking strategy that proves most effective for Republicans has important implications for how we think about battling misinformation online and how to target key groups with interventions beyond independent fact-checkers or Community Notes.

AI fact-checking works – but we need more research on other groups

My results suggest that AI has become a credible and effective source for Republican users to recognise political misinformation: AI labels lead to meaningful and robust reductions in engagement with misleading statements among Republican users. One possible explanation is that AI might be perceived as less biased than both independent fact-checkers and other users by this group.

Further research should focus on the mechanisms behind these results. To examine what role AI should take in fact-checking efforts and framing, we need to develop a better understanding of the speed and quality AI can offer in fact-checking efforts in the real world. Further research and policy debates should also consider how different key groups in the battle against

misinformation can best be targeted to provide them with the tools to make informed decisions about what information to trust.

- *Isolde Hegemann was supported in this research by the Phelan US Centre PhD Summer Research Grants scheme. [Click here](#) to read her report.*
- *Subscribe to [LSE USAPP's email newsletter](#) to receive a weekly article roundup.*
- *Please read our comments policy before commenting.*
- *Note: This article gives the views of the author, and not the position of USAPP – American Politics and Policy, nor the London School of Economics.*

---

## About the author



**Isolde Hegemann**

Isolde Hegemann is a PhD Student in the Department of Government at LSE. Her research examines political communication and party competition amid the rise of the radical right, using quantitative methods. She also teaches classes on European politics.

**Posted In:** Democracy and culture

---

## Leave a Reply

Your email address will not be published. Required fields are marked \*

Name	Email	Site
------	-------	------