In a world where misinformation can incite riots and hashtags can spark revolutions, the idea that truth is now a group project should send a collective chill down our spines.

In January 2025, just as Donald Trump took office once again, Meta quietly dismantled its third-party fact-checking program, opting instead to roll the dice on a community notes model. No more professional journalists or watchdog organizations verifying the flood of content on Facebook or Instagram—just users, their opinions, and a voting system that's supposed to detect consensus. Think Wikipedia, but for facts with real-world consequences.

The company claimed its former fact-checking model was "too politically biased and destroyed more trust than it created." But this isn't a mere rebranding of moderation—it's a tectonic shift. In the name of restoring trust, Meta is abdicating responsibility. It's a move that gives "the people" more power, sure—but it also gives trolls, extremists, and bad-faith actors a seat at the table.

And while Meta is putting its faith in democracy-by-comment-section, it's doing so at a time when that trust is more fragile than ever.

Meta's Moderation Muzzle

Gone are the warnings under false headlines. Gone are the explanations beneath viral hoaxes. Now, any post—whether it's a photo, meme, or outright lie—can only be contextualized if enough people who usually disagree on other posts *agree* the note is helpful. It's moderation by mob consensus. But what happens when the mob is wrong?

Consider the decision to ease automated filters around immigration and gender. What used to be filtered for harmful stereotypes or hate speech is now fair game. Meta's community guidelines now tolerate phrases likening women to "household objects" and framing gender or sexual orientation as "mental illness"—as long as it's couched as political or religious

“discourse.”

In other words, hate speech got a new marketing team.

This isn’t a slippery slope; it’s a vertical drop. If truth is now a popularity contest, we know who usually loses: minorities, marginalized voices, and those who’ve never had the numbers to win in a digital dogpile.

Telegram Walked So Meta Could Fall

Meta’s new policy doesn’t just echo Twitter/X’s crowd-based moderation model—it invites comparisons to a far more volatile platform: Telegram.

Telegram, long the Wild West of social media, learned the hard way that deregulation can backfire. From multi-million euro fines in Germany to a high-profile arrest of its founder in Paris, the app’s free speech absolutism came face-to-face with legal and moral accountability. Telegram’s promise of privacy crumbled under the weight of school shootings, hate groups, and drug cartels using the app with impunity.

Telegram now hands over IP addresses and phone numbers, publishes transparency reports, and has scaled back its former “we answer to no one” bravado. But that shift only came after intense international pressure, legal peril, and widespread outcry. Meta seems determined to take a similar path—just with more polish and better branding.

What happens when Meta, with its billions of users and global reach, becomes a platform where “debate” justifies bigotry, where truth is drowned out by volume, and where the voting majority can weaponize silence or amplify falsehoods?

Learning Nothing from History

Meta knows this story already. In 2018, misinformation on Facebook helped incite anti-Muslim violence in Sri Lanka. One viral video of a confused Muslim man, taken out of context, unleashed a torrent of hate, culminating in mobs torching mosques and beating civilians. The country declared a state of emergency. The fire started online—Meta lit the match.

What's to stop that from happening again? This time, instead of algorithms failing to catch the spark, we're giving gasoline and matches to the crowd and hoping they vote to use them wisely.

The European Reckoning

Europe is watching, and not passively.

In the wake of Meta's fact-checking rollback, the European Commission has already launched investigations under the Digital Services Act. From election disinformation to the spread of Russian propaganda (yes, Meta profited from it), regulators are lining up their ammunition. Unlike the U.S., where regulation lags far behind innovation, the EU is drawing a hard line.

If Telegram is a cautionary tale, Meta is now the sequel. A company that once boasted 27 election integrity tools dismantled them before a major global election year. Ursula von der Leyen didn't mince words: platforms like Meta must "put enough resources into this." Instead, Meta handed the job to us.

Meta's Dangerous Faith in "The Crowd"

There's something quaint about Meta's belief in us. The idea that everyday users—biased,

impulsive, and often misinformed—can now serve as editors of the digital record sounds noble. But this isn't Wikipedia. This is where real harm starts. Where narratives aren't just shaped, they're weaponized.

Crowd-sourced truth assumes a fair playing field. It assumes people come in good faith. It assumes the voices of reason can outshout chaos. But that's not the internet we live in.

Meta has made a trust fall into the arms of its users. But if history is any indication, they may not be caught.

They may be trampled.