**THE STONE**

# How to Fix Fake News

Technology has given rise to an age of misinformation. But philosophy, and a closer look at our own social behavior, could help eliminate it.

**By Regina Rini**

Ms. Rini teaches philosophy at York University in Toronto.

Oct. 15, 2018

Technology spawned the problem of fake news, and it's tempting to think that technology can solve it, that we only need to find the right algorithm and code the problem away. But this approach ignores valuable lessons from epistemology, the branch of philosophy concerned with how we acquire knowledge.

To understand how we might fix the problem of fake news, start with cocktail hour gossip. Imagine you're out for drinks when one of your friends shocks the table with a rumor about a local politician. The story is so scandalous you're not sure it could be right. But then, here's your good friend, vouching for it, putting their reputation on the line. Maybe you should believe it.

This is an instance of what philosophers call testimony. It's similar to the sort of testimony given in a courtroom, but it's less formal and much more frequent. Testimony happens any time you believe something because someone else vouched for the information. Most of our knowledge about the world is secondhand knowledge that comes to us through testimony. After all, we can't each do all of our own scientific research, or make our own maps of distant cities.

All of this relies upon norms of testimony. Making a factual claim in person, even if you are merely passing on some news you picked up elsewhere, means taking on the responsibility for it, and putting your epistemic reputation — that is, your credibility as a source — at risk. Part of the reason that people believe you when you share information is this: they've determined your credibility and can hold you accountable if you are lying or if you're wrong. The reliability of secondhand knowledge comes from these norms.

But social media has weird testimonial norms. On Facebook, Twitter and similar platforms, people don't always mean what they say, and we don't always expect them to. As the informal Twitter slogan goes: "A retweet is not an endorsement." When Donald Trump was caught retweeting fake statistics about race and crime, he told Fox News it wasn't a big deal: "am I gonna check every statistic? All it was is a retweet. It wasn't from me." Intellectually, we know that people do this all of the time on social media, and pass along news without verifying its accuracy, but many of us listen to them anyway. The information they share is just too tempting to ignore — especially when it reaffirms our existing political beliefs.

To fight fake news, we need to take the same norms that keep us (relatively) honest over cocktails, and apply them to social media. The problem, however, is that social media is like going out for drinks with your 500 closest friends, every night. You might pick up a lot of information, but in all the din you're unlikely to remember who told you what and who you should question if the information later turns out to be wrong.

There's simply too much information for our minds to keep track of. You read a headline — and sometimes that might be all you read — and you'll be shocked, click the angry face button, and keep scrolling. There's always another story, another outrage. React, scroll, repeat.

The number of stories isn't the only problem; it's also the number of storytellers. The average Facebook user has hundreds of friends, many of whom they barely know offline. There's no way of knowing how reliable your Facebook friends are. You might be wary of a relative's political memes, but what about the local newspaper links posted by an opinionated colleague of your cousin's wife, who you once met at a party? It's impossible to do this reputational calculation for all of these people, and all of the stories they share.

To solve this problem — or at least improve the situation — we need to establish stable testimonial norms, which allow us to hold each other accountable on social media. This requires cutting through the information deluge and keeping track of the trustworthiness of hundreds of social media contacts. Luckily, there's an app for that.

Facebook already has features that support better testimonial norms. Most Facebook accounts are closely linked to users' real-life social networks. And, unlike anonymous web commenters, Facebook users can't just walk away from their identity when they're caught lying. Users have a reason to care about their epistemic reputation — or, at the very least, they would if others could keep tabs on the information that they shared.

Here's a system that might help, and it is based on something that Facebook already does to prevent the spread of fake news. Currently, Facebook asks independent fact-checking organizations from across the political spectrum to identify false and misleading information. Whenever users try to post something that has been identified as fake news, they are confronted by a pop-up that explains the problems with the news and asks them to confirm if they'd like to continue. None of these users are prevented from posting stories whose facts are in dispute, but they are required to know that what they are sharing may be false or misleading.

Facebook has been openly using this system since December 2016. Less openly, they have also been keeping tabs on how often its users attempt to flag stories as fake news, and, using this feature, they have been calculating the epistemic reliability of their users. The Washington Post reported in August that Facebook secretly calculates scores that represent how often users' flags align with the analysis of independent fact-checkers. Facebook only uses this data internally, to identify abuse of the flagging system, and does not release it to users. I can't find out my own reputation score, or the scores of any of my friends.

This system and the secrecy around it may come across as a bit creepy — and the public trust in Facebook has been seriously and justifiably damaged — but I think that Facebook is on to something. Last year, in a paper published in the Kennedy Institute of Ethics Journal, I proposed a somewhat different system. The key difference between my system and the one that Facebook has implemented is transparency: Facebook should track and display how often each user decides to share disputed information after being warned that the information might be false or misleading.

Instead of using this data to calculate a secret score, Facebook should display a simple reliability marker on every post and comment. Imagine a little colored dot next to the user's name, similar to the blue verification badges Facebook and Twitter give to trusted accounts: a green dot could indicate that the user hasn't chosen to share much disputed news, a yellow dot could indicate that they do it sometimes, and a red dot could indicate that they do it often. These reliability markers would allow anyone to see at a glance how reliable their friends are.

There is no censorship in this proposal. Facebook needn't bend its algorithms to suppress posts from users with poor reliability markers: Every user could still post whatever they want, regardless of whether the facts of the stories they share are in dispute. People could choose to use social media the same way they do today, but now they'd have a choice whenever they encounter new information. They might glance at the reliability marker before nodding along with a friend's provocative post, and they might think twice before passing on a weird story from a friend with a red reliability marker. Most important of all, a green reliability marker could become a valuable resource, something to put on the line only in extraordinary cases — just like a real-life reputation.

There's technology behind this idea, but it's technology that already exists. It's aimed at assisting rather than algorithmically replacing the testimonial norms that have been regulating our information-gathering since long before social media came along. In the end, the solution for fake news won't be just clever programming: it will also involve each of us taking up our responsibilities as digital citizens and putting our epistemic reputations on the line.

Regina Rini (@rinireg) teaches philosophy at York University in Toronto, where she holds the Canada Research Chair in Philosophy of Moral and Social Cognition.