People are getting their news from AI – and it's altering their views

Adrian Kuenzler, Scholar-in-Residence, University of Denver; University of Hong Kong

When a bot brings you the news, who built it and how it presents the information matter.

Zentangle/iStock via Getty Images

Meta's decision to end its professional fact-checking program sparked a wave of criticism in the tech and media world. Critics warned that dropping expert oversight could erode trust and reliability in the digital information landscape, especially when profit-driven platforms are mostly left to police themselves.

What much of this debate has overlooked, however, is that today, AI large language models are increasingly used to write up news summaries, headlines and content that catch your attention long before traditional content moderation mechanisms can step in. The issue isn't clear-cut cases of misinformation or harmful subject matter going unflagged in the absence of content moderation. What's missing from the discussion is how ostensibly accurate information is selected, framed and emphasized in ways that can shape public perception.

Large language models gradually influence the way people form opinions by generating the information that chatbots and virtual assistants present to people over time. These models are now also being built into news sites, social media platforms and search services, making them the primary gateway to obtain information.

Studies show that large language models do more than simply pass along information. Their responses can subtly highlight certain viewpoints while minimizing others, often without users realizing it.

Communication bias

My colleague, computer scientist Stefan Schmid, and I, a technology law and policy scholar, show in a forthcoming accepted paper in the journal Communications of the ACM that large language models exhibit communication bias. We found that they may have a tendency to highlight particular perspectives while omitting or diminishing others. Such bias can influence how users think or feel, regardless of whether the information presented is true or false.

Empirical research over the past few years has produced benchmark datasets that correlate model outputs with party positions before and during elections. They reveal variations in how current large language models deal with public content. Depending on the persona or context used in prompting large language models, current models subtly tilt toward particular positions – even when factual accuracy remains intact.

These shifts point to an emerging form of persona-based steerability – a model's tendency to align its tone and emphasis with the perceived expectations of the user. For instance, when a user describes themselves as an environmental activist and another as a business owner, a model may answer the same question about a new climate law by emphasizing different, yet factually accurate, concerns for each of them. For example, the criticisms could be that the law does not go far enough in promoting environmental benefits and that the law imposes regulatory burdens and compliance costs.

Such alignment can easily be misread as flattery. The phenomenon is called sycophancy: Models effectively tell users what they want to hear. But while sycophancy is a symptom of user-model interaction, communication bias runs deeper. It reflects disparities in who designs and builds these systems, what datasets they draw from and which incentives drive their refinement. When a handful of developers dominate the large language model market and their systems consistently present some viewpoints more favorably than others, small differences in model behavior can scale into significant distortions in public communication.

What regulation can and can't do

Modern society increasingly relies on large language models as the primary interface between people and information. Governments worldwide have launched policies to address concerns over AI bias. For instance, the European Union's AI Act and the Digital Services Act attempt to impose transparency and accountability. But neither is designed to address the nuanced issue of communication bias in AI outputs.

Proponents of AI regulation often cite neutral AI as a goal, but true neutrality is often unattainable. AI systems reflect the biases embedded in their data, training and design, and attempts to regulate such bias often end up trading one flavor of bias for another.

And communication bias is not just about accuracy – it is about content generation and framing. Imagine asking an AI system a question about a contentious piece of legislation. The model's answer is not only shaped by facts, but also by how those facts are presented, which sources are highlighted and the tone and viewpoint it adopts.

This means that the root of the bias problem is not merely in addressing biased training data or skewed outputs, but in the market structures that shape technology design in the first place. When only a few large language models have access to information, the risk of communication bias grows. Apart from regulation, then, effective bias mitigation requires safeguarding competition, user-driven accountability and regulatory openness to different ways of building and offering large language models.

Most regulations so far aim at banning harmful outputs after the technology's deployment, or forcing companies to run audits before launch. Our analysis shows that while prelaunch checks and post-deployment oversight may catch the most glaring errors, they may be less effective at addressing subtle communication bias that emerges through user interactions.

Beyond AI regulation

It is tempting to expect that regulation can eliminate all biases in AI systems. In some instances, these policies can be helpful, but they tend to fail to address a deeper issue: the incentives that determine the technologies that communicate information to the public.

Our findings clarify that a more lasting solution lies in fostering competition, transparency and meaningful user participation, enabling consumers to play an active role in how companies design, test and deploy large language models.

The reason these policies are important is that, ultimately, AI will not only influence the information we seek and the daily news we read, but it will also play a crucial part in shaping the kind of society we envision for the future.