Why do we need transparency about social media algorithms?

When we think of how harmful messaging spreads online, from COVID-19 misinformation to voter suppression, a big piece of the puzzle is not just the content of advertisements or posts, but who sees what, when, and why. Social media platforms create complex algorithms—basically, series of rules on how to tailor content to each individual user—which are opaque to the public. These algorithms govern how one user might be prompted to join an extremist discussion group, while another might be guided to a club for knitting enthusiasts. Civil society groups are calling for greater transparency on these algorithms, along with other regulatory mechanisms, to help the public hold platforms accountable for discriminatory and other types harmful messaging.

What does algorithmic transparency look like?

Generally, algorithmic transparency involves requirements for platforms to provide descriptions of algorithm processes. Core to content recommendation algorithms are a series of decisions: what posts do people see, in what order do they see them, and what is the reason that order was chosen. Algorithmic transparency can involve giving a view to all three of these decisions. The amount of detail provided varies greatly, from extremely general plain-text descriptions that could be fully displayed to users to detailed auditable trails of algorithmic decision-making. These transparency requirements are often designed to provide information that would enable regulators to monitor for violations.

Examples of proposals requiring algorithmic transparency

A number of proposals before Congress contain provisions that would mandate some form of algorithmic transparency:

- The Algorithmic Justice and Online Platform Transparency Act would require platforms to describe in plain language the types of algorithmic processes they use and the information they collect, publish annual public reports about content moderation, and as maintain detailed records on algorithmic process for review by the Federal Trade Commission (FTC).

- The Digital Services Oversight and Safety Act requires platforms to develop community standards that include information about policies, procedures, measures and tools used in algorithm decision making, and would task the FTC to create requirements for platforms to release regular public-facing transparency reports, which would include statistics about enforcement decisions.

- The Platform Accountability and Transparency Act authorizes the FTC to create rules governing platform requirements to provide the public information about moderation decisions and algorithms.

Who sees what: marketing Instagram Reels to Facebook users

In 2021, Facebook began promoting Instagram Reels to its users, adding viewing suggestions to people's "News Feeds," in the "Suggested for you" section. Typically these Reels are from Instagram accounts not followed by the users who receive the recommendations. In an analysis of Facebook's "Widely viewed content" report, the Integrity Institute found that 11 of the top viewed posts on the platform were Reels, and that of those, seven were "clearly focused on people fighting with each other, sometimes physically, sometimes verbally."

The Institute noted: "It is important to stress that these are all recommended Reels. The vast majority of Facebook users who were shown these videos never chose to follow the accounts that uploaded them. The reason these appeared on the top 20 list is entirely due to Facebook's decision to show recommended Reels in News Feed, and the algorithm that decides which Reels to show Facebook users."

Source: The Integrity Institute's Analysis of Facebook's Widely Viewed Content Report [2021-Q4], March 2022

For more information about NYU Cybersecurity for Democracy:
info@cybersecurityfordemocracy.org | cybersecurityfordemocracy.org | Twitter: @cyber4democracy