News Item

Why social media can't keep moderating content in the shadows | MIT Technology Review

By Joan Donovan
November 9, 2020

> Back in 2016, I could count on one hand the kinds of interventions that technology companies were willing to use to rid their platforms of misinformation, hate speech, and harassment. Over the years, crude mechanisms like blocking content and banning accounts have morphed into a more complex set of tools, including quarantining topics, removing posts from search, barring recommendations, and down-ranking posts in priority.
>
> And yet, even with more options at their disposal, misinformation remains a serious problem. There was a great deal of coverage about misinformation on Election Day—my colleague Emily Drefyuss found, for example, that when Twitter tried to deal with content using the hashtag #BidenCrimeFamily, with tactics including "de-indexing" by blocking search results, users including Donald Trump adapted by using variants of the same tag. But we still don't know much about how Twitter decided to do those things in the first place, or how it weighs and learns from the ways users react to moderation.
>
> What actions did these companies take? How do their moderation teams work? What is the process for making decisions?
>
> As social media companies suspended accounts and labeled and deleted posts, many researchers, civil society organizations, and journalists scrambled to understand their decisions. The lack of transparency about those decisions and processes means that—for many—the election results end up with an asterisk this year, just as they did in 2016.
>
> […]

Source: [Why social media can't keep moderating content in the shadows | MIT Technology Review](#)