News Item

GPT-3 Can Write Disinformation Now—and Dupe Human Readers | WIRED

By Will Knight

May 24, 2021

> When OpenAI demonstrated a powerful artificial intelligence algorithm capable of generating coherent text last June, its creators warned that the tool could potentially be wielded as a weapon of online misinformation.
>
> Now a team of disinformation experts has demonstrated how effectively that algorithm, called GPT-3, could be used to mislead and misinform. The results suggest that although AI may not be a match for the best Russian meme-making operative, it could amplify some forms of deception that would be especially difficult to spot.
>
> Over six months, a group at Georgetown University's Center for Security and Emerging Technology used GPT-3 to generate misinformation, including stories around a false narrative, news articles altered to push a bogus perspective, and tweets riffing on particular points of disinformation.
>
> […]

Source: [GPT-3 Can Write Disinformation Now—and Dupe Human Readers | WIRED](link)