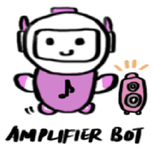
AI Bot Wars

Lynnette Hui Xian Ng
lynnetteng@cmu.edu

Kathleen M. Carley
carley@andrew.cmu.edu

Bot Types by Behavior

										
Using AI to spread disinfo	Amplify disinfo, amplify popularity	Manipulate opinions to believe disinfo	Human portion establishes credibility; Bot portion scales dissemination of disinfo	Connect people spreading disinfo to normal users, bringing disinfo content to their attention	Repeating disinfo content	Openly spreads disinfo	Coordinated spread of disinfo, political ideologies	Covertly influence (political) processes by fitting in with the crowd	Use of influential bot users to push narratives to domestic and foreign regions	Create confusion with lots of disinfo surrounding true information
Using AI to curb disinfo	Amplify fact checks and true information	Manipulate opinions to disregard disinfo	Human portion establishes credibility; Bot portion scales dissemination of truths	Connect users that spread disinfo to fact-checking users, bringing the truths to their attention	Repeating fact checks and truths	Used by governments/ organizations to spread facts	Coordinated spread of fact-checks and truths	Fit in with the crowd in terms of identity and language style and correct disinfo	Used by governments and organizations to push true information	Flood the social media space with lots of truths surrounding disinfo
Effective?										

Bot Types by Content

Using AI to spread disinfo	Announcing disinfo (e.g. earnings call) ahead of confirmation	Generating disinfo texts, deepfake audio, images	Make true information misleading/ false, by adding exaggerations	Financial bots: spread disinfo to manipulate stock prices Donor bots: make fast and minor donations to create illusion	Deliberately provide false info during the conversation	Artificially generate engagement of disinfo posts	Spread disinfo news
Using AI to curb disinfo	Announcing true information regularly	Generate true information, explanations and validating sources	Correct disinfo with facts and explainations	Financial bots: post accurate reports earnings call, stock prices	Perform fact checking within the conversation, correct false info	Create questions and polls of truths to debunk disinfo using engaging techniques	Spread news that certain posts are disinfo posts
Effective?							

Can be used to spread disinformation

Can be used to curb disinformation

Which bot do you think is effective in spreading & curbing disinformation? **TELL US WITH A STICKER!**

S3D Software and Societal Systems Department

This work was supported in part by Scalable Technologies for Social Cybersecurity, U.S. Army (W911NF20D0002), the Minerva-Multi-Level Models of Covert Online Information Campaigns, Office of Naval Research (N000142112765), and the Threat Assessment Techniques for Digital Data, Office of Naval Research (N000142412414), and through a grant and the Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of the Office of Naval Research, the U.S. Army or the U.S. government.

Carnegie Mellon University