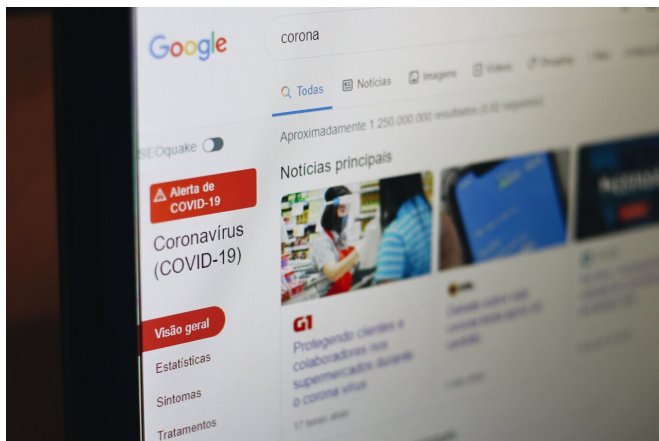


Studies suggest finding automatic ways to spot fake news may be more complicated than anticipated

22 June 2021, by Matt Swayne



Credit: Unsplash/CC0 Public Domain

Clickbait headlines might not be as enticing to readers as once thought, according to a team of researchers. They added that artificial intelligence—AI—may also come up short when it comes to correctly determining whether a headline is clickbait.

In a series of studies, the researchers found that clickbait—headlines that often rely on linguistic gimmicks to tempt readers to read further—often did not perform any better and, in some cases, performed worse than traditional headlines.

Because fake news is a concern on [social media](#), researchers have explored using AI to systematically identify and block clickbait. However, the studies also suggest that identifying fake news with [artificial intelligence](#) may be even more complicated than anticipated, said S. Shyam Sundar, James P. Jimirro Professor of Media Effects in the Donald P. Bellisario College of Communications and co-director of the Media

Effects Research Laboratory.

"One of the ideas in fake news research is that if we can just solve the clickbait problem, we can get closer to solving the fake news problem," said Sundar, who also is an affiliate of Penn State's Institute for Computational and Data Sciences (ICDS). "Our studies push back on that a little bit. They suggest that fake news might be a completely different ballgame, and that clickbait is itself more complicated than we thought."

In the first study, the research team randomly assigned 150 participants to read one of eight different types of headlines and measured if the participants would then read or share the story. The participants read either a traditional [headline](#), or a headline that relied on one of the seven types of clickbait features, including headlines with questions, lists, "Wh" words (i.e., what, when), demonstrative adjectives (i.e., this, that), positive superlatives (i.e., best, greatest), negative superlatives (i.e., worst, least), or modals (i.e., could, should). The headlines were taken from both reliable and unreliable online sources and classified using algorithms developed to detect clickbait.

"One of the questions we had initially was, which of those clickbait features would attract more clicks?" said Maria Molina, assistant professor of advertising and public relations at Michigan State, who is the lead author of the study. "We wanted to explore that more in depth, but when we analyzed the results, we realized there were no significant differences, and, if anything, people were more attracted to non-clickbait headlines. So, from there, we figured there might be some reasons why this might have happened."

The researchers conducted a second study to make sure that other factors such as the subject

matter of each headline, were not confusing the results, according to Molina.

In this study, the researchers recruited 249 participants, who were randomly assigned to one of eight conditions—seven clickbait headlines and one non-clickbait headline. This time, all headlines focused on a single political topic and were written by a former journalist. Again, the team reported that the clickbait headlines did not dramatically outperform the traditional headline.

According to Dongwon Lee, professor of information sciences and technology at Penn State, the team conducted a third study to examine several types of AI, or [machine-learning models](#), that were used in the study to classify headlines as either clickbait headlines or not. They found that the models frequently disagreed about whether the headline was clickbait or not.

The study found that the four AI models agreed on the classification of clickbait only 47% of the time. Of the 175 headlines that were classified to be similar by the four algorithms, 139 were identified as clickbait and 36 were non-clickbait. The level of agreement between the systems also varied based on the type of headline. For example, while the four algorithms agreed on the clickbait classification more times for the negative superlative characteristic, compared to the other six characteristics, the four classifiers failed to agree on a non-clickbait classification for the negative superlative or question characteristics.

Performance of AI and machine learning models tends to vary, said Lee, who is an ICDS affiliate. When the headlines classified by each model were assessed against the number of clicks, three of the four models consistently showed that demonstrative adjectives, lists and "wh" words attracted more engagement from readers than non-clickbait headlines.

"As these machine learning models are the product of the past several decades, we have many variations—some are very simple, some run very fast, yet others are more complicated and require a lot of resources," said Lee. "It is like when you assemble a desk—you can do the job with a

screwdriver that costs \$5, but can probably do the job faster with a power drill costing \$50. So, depending on the inherent power of these machine-learning models, and the training dataset the models are given, they tended to have different levels of performance and varying pros/cons."

However, these findings raise doubts about using AI to detect fake news by classifying the headlines alone.

"People were putting a lot of stock into using clickbait headlines as an element for fake news detection algorithms, but our studies are calling this assumption into question," said Sundar.

He added that the studies also suggest that programmers who develop algorithms to detect fake news may have to continually adapt as human fake news producers—and media consumers—become savvy to the elements that make up fake news.

"It becomes a bit of a cat and mouse game," said Sundar. "The people who write fake news may become aware of the characteristics that are identified as fake [news](#) by the detectors and they will change their strategies. News consumers may also just become numb to certain characteristics if they see those headlines all the time. So, [fake news](#) detection must constantly evolve with the readers as well as the creators."

The researchers suggested that the popularity of clickbait headlines in the past might be a reason for the failure of the headlines to engage readers in their studies. Clickbait could be so ubiquitous in today's media that they fail to stand out and attract the same attention as traditional headlines.

The popularity of clickbait also brought more media scrutiny, which may have made participants in the study more wary of clickbait headlines, added Molina.

The research team presented their findings at CHI 2021 conference.

More information: Maria D. Molina et al, Does Clickbait Actually Attract More Clicks? Three

Clickbait Studies You Must Read, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021). [DOI: 10.1145/3411764.3445753](#)

Provided by Pennsylvania State University

APA citation: Studies suggest finding automatic ways to spot fake news may be more complicated than anticipated (2021, June 22) retrieved 22 October 2021 from <https://techxplore.com/news/2021-06-automatic-ways-fake-news-complicated.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.