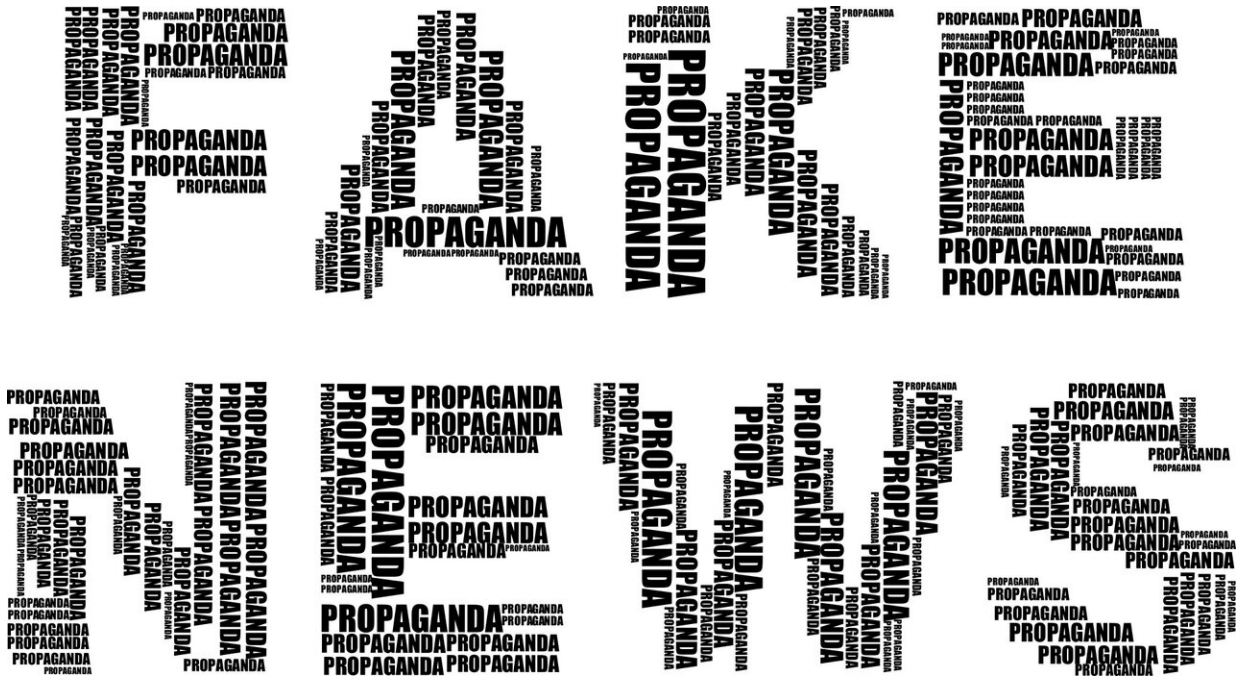
Toward a better understanding of 'fake news'

April 6 2021



Credit: Pixabay/CC0 Public Domain

Duncan Watts, a Penn Integrates Knowledge Professor and computational social scientist with appointments in the Annenberg School for Communication, School of Engineering and Applied Science and the Wharton School, has published a new framework for studying media bias and misinformation. Publishing this week in the *Proceedings of the National Academy of Sciences* and co-authored by colleagues at Microsoft Research, the paper describes an ambitious and comprehensive research agenda for understanding the origins, nature,

and prevalence of misinformation and its impact on democracy.

The phrase "fake news" has become part of the lexicon, spurred by news coverage of fake political ads and Twitter bots and by concerns about their role in populist political movements such as Brexit and the 2016 presidential election. These stories triggered an enormous amount of research, with the publication of thousands of papers trying to understand how fake news was spreading.

"This overwhelming focus on outright lies circulating on social media was disturbing, but it was missing something," says Watts. "And that something is this much broader conception of misinformation."

Watts explains that misinformation includes more than just lies and falsehoods because there are also more subtle ways that people can be misled. This includes data cherry picking, misconstruing the relationship between correlation and causation, or even simply presenting facts in a particular way, tactics that can lead people to a false conclusion without technically failing a fact check.

Misinformation is also not something that's limited to social media, he says, with television, radio, and print publications also playing an important role. "All of the research that has been done on Twitter vastly outweighs the amount of research that has been done on TV in the last four years, and yet TV is a larger source of information related to politics for typical Americans than Twitter is." says Watts. "We really have to be thinking much more expansively about the parts of the information ecosystem that might be causing some these problems."

To this end, Watts and his co-authors describe four specific objectives that would enable research communities and funding agencies to address these types of complex questions:

Build a large-scale data infrastructure

The first step, Watts says, is to build a research infrastructure to collect, organize, clean, and make data available and accessible to the broader research community. It's akin to other large-scale research endeavors, like the Large Hadron Collider, where a community of scientists unite to work on a single instrument or project that generates data for an entire field of researchers.

"If you want to look at everything that is being produced on television, radio, and the web and ask questions, there's no way to answer them right now," says Watts. "There doesn't exist any infrastructure to collect that data, and even just collecting that data is an enormous undertaking."

Establish a 'mass collaboration' model

With a robust data infrastructure in place, the next objective is to maximize its value by coordinating the efforts of multiple research groups. Instead of working on single datasets that are curated and analyzed by an individual or group, this working model provides a way to study problems more holistically.

This strategy could also help researchers more effectively work on large-scale problems, improve the replicability of studies, and help groups build upon cumulative knowledge that could then be applied outside of academia, says Watts.

Communicate with stakeholders

It is important to educate members of the public on their findings, Watts says, and it's also crucial to make data both accessible and relevant.

"Informing the public is a valuable thing to do, but it's not something that we have a lot of incentive to do as academics," says Watts, adding that different ways to achieve this objective could include publishing "living" versions of research papers in the form of data dashboards.

Develop academic-industry partnerships

"It would already be a big step for people of different disciplines and institutions to work together on a common dataset, but, if we want to actually solve problems in the world, we have to do more than just understand things. We have to also try to design interventions that affect people's experience on real platforms and measure the consequences," he says.

From collaborating with computer scientists on improving the fairness of algorithms to working with journalists to help them understand how their work influences public opinion, engaging with partners outside of academia is essential for addressing misinformation, Watts says.

To consolidate and accelerate the work towards achieving these objectives, Watts launched the Computational Social Science Lab at Penn last month. It will officially open its doors at the beginning of the next academic year. Its ongoing research includes studying the prevalence of radical content on YouTube, evaluating how media consumption has shifted from live TV to streaming platforms, the impacts of echo chambers, identifying and tracking bias in media coverage, and a number of projects with data providers to improve data quality so that researchers can more easily analyze and make sense of these complex datasets.

"Many researchers can then start using this data, and then the amount of research that is generated by this infrastructure goes up by 100 or one hundred times. I think that will be the real innovation," he says. "There's

many questions out there, and we would like to be able to help lots of people answer those questions."

More information: Duncan J. Watts el al., "Measuring the news and its impact on democracy," *PNAS* (2021). www.pnas.org/cgi/doi/10.1073/pnas.1912443118

Provided by University of Pennsylvania