Navigating 'information pollution' with the help of artificial intelligence

July 17 2020, by Erica K. Brockmeier

Credit: Pixabay/CC0 Public Domain

There's still a lot that's not known about the novel coronavirus SARS-CoV-2 and COVID-19, the disease it causes. What leads some people to have mild symptoms and others to end up in the hospital? Do masks help

stop the spread? What are the economic and political implications of the pandemic?

As researchers try to address many of these questions, many of which will not have a simple 'yes or no' answer, people are also trying to figure out how to keep themselves and their families safe. But between the 24-hour news cycle, hundreds of preprint research articles, and guidelines that vary between regional, state, and federal governments, how can people best navigate through such vast amounts of information?

Using insights from the field of natural language processing and artificial intelligence, computer scientist Dan Roth and the Cognitive Computation Group are developing an online platform to help users find relevant and trustworthy information about the novel coronavirus. As part of a broader effort by his group to develop tools for navigating "information pollution," this platform is devoted to identifying the numerous perspectives that a single query might have, showing the evidence that supports each perspective and organizing results, along with each source's "trustworthiness," so users can better understand what is known, by whom, and why.

Creating these types of automated platforms represents a huge challenge for researchers in the field of natural language processing and machine learning because of the complexity of human language and communication. "Language is ambiguous. Every word, depending on context, could mean completely different things," says Roth. "And language is variable. Everything you want to say, you can say in different ways. To automate this process, we have to get around these two key difficulties, and this is where the challenge is coming from."

Thanks to numerous conceptual and theoretical advances, the Cognitive Computational Group's fundamental research in natural language understanding has allowed them to apply their research insights and to

develop automated systems that can better understand the contents of human [language](#), such as what is being written about in a news article or scientific paper. Roth and his team have been working on issues related to information pollution for many years and are now applying what they've learned to information about the novel coronavirus.

Information pollution comes in many forms, including biases, misinformation, and disinformation, and because of the sheer volume of information the process of sorting fact from fiction needs automated support. "It's very easy to publish information," says Roth, adding that while organizations like FactCheck.org, a project of Penn's Annenberg Public Policy Center, manually verify the validity of many claims, there's not enough human power to fact check every claim being posted on the Internet.

And fact checking alone isn't enough to address all of the problems of information pollution, says Ph.D. student Sihao Chen. Take the question of whether people should wear face masks: "The answer to that question has changed dramatically in the past couple months, and the reason for that change is multi-faceted," he says. "You couldn't find an objective truth attached to that specific question, and the answer to that question is context-dependent. Fact checking alone doesn't solve this problem because there's no single answer." This is why the team says that identifying various perspectives along with evidence that supports them is important.

To help address both of these hurdles, the COVID-19 search platform visualizes results that include a source's level of trustworthiness while also highlighting different perspectives. This is different from how online search engines display information, where top results are based on popularity and keyword match and where it's not easy to see how the arguments in articles compare to one another. On this platform, however, instead of displaying articles on an individual basis, they are

organized based on the claims they make.

More information: Penn Information Pollution Project
dickens.seas.upenn.edu:4011/

Provided by University of Pennsylvania

Citation: Navigating 'information pollution' with the help of artificial intelligence (2020, July 17)
retrieved 18 April 2024 from
https://techxplore.com/news/2020-07-pollution-artificial-intelligence.html