'CoronaCheck' website combats spread of misinformation

April 3 2020, by Melanie Lefkowitz

Credit: CC0 Public Domain

Cornell researchers have developed an automated system that uses machine learning, data analysis and human feedback to automatically verify statistical claims about the new coronavirus.

"CoronaCheck," based on ongoing research from Immanuel Trummer,

assistant professor of computer science, launched internationally in March and has already been used more than 9,600 times. The database—now available in English, French and Italian—checks claims on COVID-19's spread based on reliable sources such as the World Health Organization and the Centers for Disease Control and Prevention.

"There's way too much misinformation about the coronavirus on the web—it's pretty mind-boggling, actually," Trummer said. "Some of these misinformation claims are harmless, but others—things like 'eating silver cures the coronavirus' – can be dangerous."

The CoronaCheck system is a collaboration between Trummer's team, including doctoral students Georgios Karagiannis and Saehan Jo, and Paolo Papotti's team at Eurecom, an engineering school in Biot, France.

Because of the sheer volume of bad information on the internet—and the rate at which more misinformation is produced and spread—it's impossible for humans to resolve the problem by performing manual fact checks alone. Even common automated approaches, which generally attempt to map new claims to existing fact checks, can't be realistically conducted on a scale large enough to tackle misinformation's scope, Trummer said.

"We've tried to automate the entire process, from the raw data to the text that we want to verify," Trummer said.

CoronaCheck adapts "Scrutinizer," a system Trummer developed with Eurecom for the International Energy Agency in Paris, a nongovernmental organization, to support human fact checkers in translating text summaries into equations the computer can understand and solve. To do this, Scrutinizer employs machine learning and natural language processing—a branch of artificial intelligence aimed at deciphering human language—as well as large datasets that help the

system figure out how to approach each new claim, and feedback from human users.

"Computers have a hard time understanding natural language," he said. "We cannot directly ask the computer to check whether some claim in a sentence is correct or not. So we essentially have to translate the claim from our language into a query language the computer understands."

For example, if someone types in that the number of coronavirus cases is higher in France than in Italy, the system uses a kind of elimination process to narrow down the possible equations to represent that text. It draws on its datasets to create a mathematical expression that can compare the claim to the facts.

Then, based on experience, the system determines the best sources to verify the claim, drawing on reliable public data compiled daily by Johns Hopkins University. The system's machine learning model can also improve over time, learning to recognize new claim types based on user feedback.

"There's a tremendous amount of misinformation out there and the set of claims that people check for is quite diverse," Trummer said. "For any given claim, there is a very large number of possible query expressions, and our goal is to find the right one."

The database interface builds on Trummer's related work, including AggChecker, the first tool to automatically verify text summaries of datasets by querying a relational database. AggChecker was presented at the Association for Computing Machinery's Special Interest Group on Management of Data's annual conference in 2019.

His team has also developed an "Anti-Knowledge Base" of common factual mistakes from Wikipedia in collaboration with Google NYC.

The research behind CoronaCheck was partly funded by a Google Faculty Research Award.

More information: CoronaCheck: coronacheck.eurecom.fr/en

AggChecker: en-cs-dbgroup.coecis.cornell.edu/

Provided by Cornell University

Citation: 'CoronaCheck' website combats spread of misinformation (2020, April 3) retrieved 23 April 2024 from
https://techxplore.com/news/2020-04-coronacheck-website-combats-misinformation.html