The value of social media in a Covid-19 world

What can data drawn from social media tell us about human society and social relationships during the pandemic and beyond? A recent webinar co-organised by ING's Think Forward Initiative provides some answers

The potential

Data drawn from social media can be a gold mine for understanding what is happening in the world, particularly during the global pandemic. But interpreting it is rife with problems.

Do people exaggerate on Facebook about how they are feeling? Can designating certain words as positive on Twitter lead to wrong conclusions about happiness? What kind of biases creep in when analysing millions of datasets from platforms that not everyone uses?

These were some of the main issues at a June 24 webinar held as part of a "virtual festival" being hosted by the Summer Institutes in Computational Social Science (SICSS) and co-organised by the Maastricht Law & Tech Lab, and ING and its Think Forward Initiative.

At issue was what computational social science -- essentially, analysing social data by correlating, modelling and simulating -- can tell us about human society and social relationships?

It is especially topical now give the advent of Big Data -- the mass of information collected and stored through technology, including social media -- and the global Covid-19 pandemic which has disrupted many societal trends and human relationships.

In terms of the pandemic, such analysis may help in tracking the spread of the virus in real time. It may also help in understanding how people are coping and what their attitudes are about the new world that will emerge when the virus has played out.

"The Covid-19 outbreak has put enormous pressure on policymakers, but also on the private sector, to quickly introduce regulations to combat the spread of the virus and to address the resulting economic and social challenges," said Monika Leszczynska, assistant professor of empirical legal research at Maastricht University, introducing the webinar. "Computational social science offers scope in identifying the areas requiring intervention and in designing the appropriate measures."

The pitfalls

The analytical opportunities may indeed seem boundless, but presenters Johan Bollen, professor of informatics at Indiana University, and Dean Eckles, professor in communications and technology at MIT, also warned of pitfalls.

Bollen, for example, noted that use of social media data to analyse sentiment is not the same as interpreting something more tangible. "Social and psychological effects are introspective states," he said, adding that attitudinal studies at the moment could be skewed by health and economic stress issues. And there are also interpretation issues.

One study of Twitter, for example, suggested people were becoming happier as the pandemic tightened its grip. This could be because some words taken as expressing a positive mood actually reflected something different. Consider the highlighted words in this tweet by UK MP Helen Whately: "*Thank* you to all those *caring* for the *people* we *love*" #YouAre*Heroes*.

Eckles, meanwhile, said that behavioural scientists studying data in a pandemic needed to remember their limitations. "We don't want to all turn into armchair epidemiologists," he said. "We should be mindful of what our expertise offers."

Critical analysis

What it does offer, regardless of the pitfalls, is a way of looking at how attitudes and actions are changing at any given moment.

Eckles said, for example, that studying social media data clearly showed people were influenced not just by what was happening locally but by what was happening in their non-geographical social cohort. He called these social spill-overs. "You are seeing what other people are doing. (You) are responding to how other people are behaving," he said.

Eckles cited research that showed which US states influenced others. The key thing was that it was by no means just geographically based. Florida, for example, was influenced by neighbour Georgia -- but also by New York, with which it shared little except social interaction. "There are reasons that these are connected. Social digital spill-overs are important," he said.

In a similar vein, Bollen said it was important for data analysts to have a hypothesis about what

might be happening, to put the numbers into context rather than just look at them raw.

"Don't just chase the data" he said. "The data and tools are important, but you will be fooled if you (aren't) critical."

Social media provides huge banks of data that can successfully be interpreted. Data from Google searches, for example, can be tapped as a source for gauging consumer sentiment -- in pandemic lockdown, or not. A recent study even showed that Twitter can be used to predict stock market movements.

But it was also important not to rely totally on social media, Bollen said.

"Social media samples are self-selected," he said, meaning that they do not reflect people who are not on the platforms.

"The potential (of computational social science) is enormous, but as scientists we have to be careful about the biases, he said, summing up the thrust of the webinar.