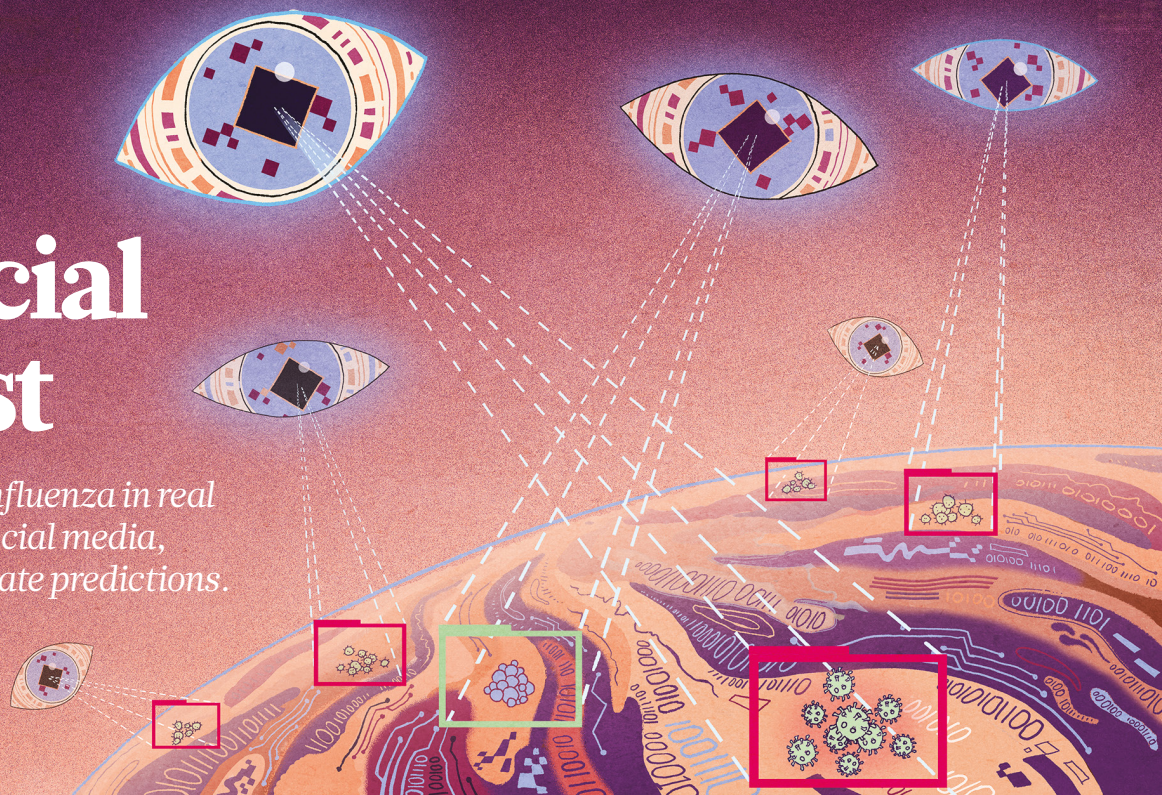# The social forecast

Scientists can track influenza in real time by monitoring social media, leading to more accurate predictions.

BY CHARLES SCHMIDT

Conventional influenza surveillance describes outbreaks of flu that have already happened. It is based on reports from doctors, and produces data that take weeks to process — often leaving the health authorities to chase the virus around, rather than get on top of it.

But every day, thousands of unwell people pour details of their symptoms and, perhaps unknowingly, locations into search engines and social media, creating a trove of real-time flu data. If such data could be used to monitor flu outbreaks as they happen and to make accurate predictions about its spread, that could transform public-health surveillance.

Powerful computational tools such as machine learning and a growing diversity of data streams — not just search queries and social media, but also cloud-based electronic health records and human mobility patterns inferred from census information — are making it increasingly possible to monitor the spread of flu through the population by following its digital signal. Now, models that track flu in real time and forecast flu trends are making inroads into public-health practice.

"We're becoming much more comfortable with how these models perform," says Matthew Biggerstaff, an epidemiologist who works on flu preparedness at the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia.

In 2013–14, the CDC launched the FluSight Network, a website informed by digital modelling that predicts the timing, peak and short-term intensity of the flu season in ten regions of the United States and across the whole country. According to Biggerstaff, flu forecasting helps responders to plan ahead, so they can be ready with vaccinations and

communication strategies to limit the effects of the virus. Encouraged by progress in the field, the CDC announced in January 2019 that it will spend US$17.5 million to create a network of influenza-forecasting centres of excellence, each tasked with improving the accuracy and communication of real-time forecasts.

The CDC is leading the way on digital flu surveillance, but health agencies elsewhere are following suit. "We've been working to develop and apply these models with collaborators using a range of data sources," says Richard Pebody, a consultant epidemiologist at Public Health England in London. The capacity to predict flu trajectories two to three weeks in advance, Pebody says, "will be very valuable for health-service planning."

SPREAD BETTING

Digital flu surveillance was transformed when Google turned its attention to flu forecasting in 2008. The company's surveillance platform, called Google Flu Trends, used machine learning to fit flu-related searches together with time-series data gathered by the CDC's US Outpatient Influenza-like Illness Surveillance Network (ILINet). With 3,500 participating clinics — each counting how many people show up with sore throats, coughs and fevers higher than 37.8 °C with no cause other than influenza — ILINet is the benchmark for flu monitoring in the United States. The aim of Google Flu Trends was to estimate flu prevalence sooner than the ILINet data could.

But two high-profile failures belied the media fanfare of its launch. First, Google Flu Trends missed a spring pandemic of H1N1 flu in 2009. Then it overestimated the magnitude of the 2012–13 flu season by 140%.

According to Mauricio Santillana, a

computational scientist at Harvard Medical School in Boston, Massachusetts, the system failed because many of the selected search terms were only seasonal, with limited relevance to flu activity, making the predictions noisy and inaccurate. After the H1N1 debacle, Google revised its flu-tracking algorithm. But the algorithm was not routinely recalibrated when the company's search-engine software was upgraded, and that created additional problems. In 2015, Google dropped the platform altogether, although it still makes some of its anonymized data available for flu tracking by researchers.

The demise of Google Flu Trends raised concerns about the role of big data in tracking diseases. But according to Vasileios Lampos, a computer scientist at University College London, the accuracy of flu forecasting is improving. "We have a lot more data and the computational tools have improved," he says. "We've had a lot of time to work on them."

Santillana points out that machine learning has markedly improved in the years since Google Flu Trends folded. "With more sophisticated approaches, it's possible to automatically ignore spuriously correlated terms, so the predictions are more robust," he says.

COMPETITIVE ADVANTAGE

The proving ground for new approaches to modelling is an annual forecasting challenge hosted by the CDC. About 20 teams participate every year, and the winners are those that perform best relative to the ILINet benchmark. In the absence of these models, the CDC's approach has been to estimate future trends based on what ILINet data gathered from previous flu seasons would predict for each region and for the United States as a whole. But during

the 2017–18 flu season, most of the models in the challenge generated predictions more accurate than those using ILINet's historical baseline. The CDC now incorporates several of the challenge's top-performing models into its FluSight system.

For the past four years, the winner of the CDC's challenge has been a team led by computer scientist Roni Rosenfeld of Carnegie Mellon University in Pittsburgh, Pennsylvania. Rosenfeld's team, called the Delphi Research Group, bases its predictions on two complementary systems. One is an online crowd-sourcing website called Epicast that allows people to express their opinions about how the current flu season might play out. "Epicast exploits the wisdom of the crowds," Rosenfeld says. "The opinion of any one person who responds isn't as accurate as the aggregated opinions of all the responders together."

The team's second system relies on machine-learning algorithms that repeatedly compare trends observed during the current flu season with those seen in previous decades. The algorithm draws on historical ILINet data as well as data from search engines and social media to assemble a distribution of all possible seasonal trajectories. It then models how the current season differs at the moment, and how it is likely to differ as it continues.

As well as machine learning, researchers also rely on mechanistic models that work in a fundamentally different way. Machine learning merely looks for patterns in data, whereas mechanistic approaches depend on specific assumptions about how a flu virus moves through the population. "This often requires biological and sociological understanding about the way disease transmission really works," says Nicholas Reich, a biostatistician at the School of Public Health and Health Sciences at the University of Massachusetts Amherst. "For instance, mechanistic models take into account the susceptible fraction of the population, the transmissibility of a particular virus, and social-mixing patterns among infected and non-infected people."

At Northeastern University in Boston, Massachusetts, Alessandro Vespignani, a computational scientist who models epidemics, has been forecasting flu by using agent-based approaches that he describes as "mechanistic modelling on steroids". Agents are simply interacting entities, including people, and Vespignani has modelled 300 million individuals, representing the US population, in various settings, and simulated how the flu virus moves among them in workplaces, homes and schools. The agent-based approach allows researchers to zoom in on disease transmission patterns with high spatial resolution. The downside is that these models require

The Delphi research group at Carnegie Mellon University forecasts the spread of influenza.

high-performance computing, Vespignani says, "and they're also data-hungry, in that they require very detailed societal descriptions."

Vespignani and Santillana are now collaborating on ways to combine machine learning with the agent-based approach to create what they claim would be an even stronger flu-forecasting model.

STRENGTH IN NUMBERS

Researchers have started to combine models into 'ensembles' that have more forecasting power than the constituent models alone. "This is something we've learned from the challenges," Biggerstaff says. "Combinations work better." That has certainly been the experience of the FluSight Network, which is a consortium of four independent research teams that collaborate on a multimodel ensemble. The ensemble links 21 models — some that use machine learning and others that are mechanistic — into a single composite model that took second place in the latest CDC flu-forecasting challenge, just behind Rosenfeld's team.

The models in this case are combined using a method called stacking, which weighs their contributions based on how well they each performed during previous flu seasons. According to Reich, who directs one of the FluSight Network's four participating teams, the ensemble approaches make optimal use of the component models' idiosyncrasies. The stacking approach, he says "is like conducting them in a symphony. You want each model at its appropriate volume."

Modelled flu forecasts, however, face a series of hurdles before they can be factored routinely into public-health preparedness in the way that, for instance, weather forecasts are used to plan for storms. To be truly effective, even the best model needs to be paired with policy measures that take into account the trends revealed by the software. But Vespignani says it is not entirely clear how confident policy-makers and health officials are when it comes

to using modelled flu forecasts in real-world settings. Many of these individuals have a poor understanding of how the computational models work, he says, and the models are most accurate at forecasting flu two to four weeks in advance, which does not really provide enough time to allocate resources where they are most needed. Vespignani says that models that could reliably predict the peak and intensity of the flu season six to eight weeks in advance would be more useful.

Santillana says that more research is needed into how social behaviour, vaccination programmes, strain composition, population immunity and other factors affect the models' accuracy. But researchers also need to understand how spatial scales factor into forecasting. For example, the CDC's forecasts are limited to national and regional levels but investigators have begun to consider the prospects for city-scale forecasts, as well as forecasting across global hemispheres.

Meanwhile, work is under way to provide machine-learning-enabled forecasting in developing countries that lack surveillance data. Lampos trained a model using surveillance data from the United States, and reported that it was accurate at forecasting flu in France, Spain and Australia without drawing on historical data from any of those countries. He says this approach could work in poorer locations that lack comparable surveillance infrastructure by analysing the frequency of search queries for flu on mobile phones and other devices. Lampos now plans to test his model in countries in Africa.

There is still a long way to go before flu forecasting becomes as routine and widely accepted as weather forecasting. But Santillana says that progress is advancing rapidly. "The predictions," he says, "are getting better and better." ∎

Charles Schmidt *is a freelance science writer in Portland, Maine.*

> *"This is something we've learned from the challenges. Combinations work better."*