THE WALL STREET JOURNAL.

http://www.wsj.com/articles/u-s-intelligence-community-explores-more-rigorous-ways-to-forecast-events-1409937859

THE NUMBERS COLUMN

U.S. Intelligence Community Explores More Rigorous Ways to Forecast Events

New Approaches to Prediction Emphasize Data-Gathering and Crowdsourcing Over Individual Deliberation

By

JO CRAVEN MCGINTY

Updated Sept. 5, 2014 7:29 p.m. ET

Analysts for the Central Intelligence Agency, the National Security Agency and more than a dozen other government organizations depend on their ability to forecast national and global events to help ward off various threats to the country, but old-style approaches can produce flawed results.

To improve quality, the government has taken the unusual step of running tournaments that invite people outside the intelligence community to develop better ways to forecast world events, and several have produced notable results.

"Traditional forecasting in the intelligence community relied on human judgment, and the way in which humans make those judgments has tended to be unstructured deliberation," said Jason G. Matheny, a project manager for IARPA, the Intelligence Advanced Research Projects Activity, which is the research and development arm of the Office of the Director of National Intelligence.

Deliberation is useful, but it isn't ideal for generating accurate forecasts: It is susceptible to groupthink. Social biases, such as deferring to those with seniority, intrude on the process. And dissenting views often aren't captured. The effects have led analysts to predict events that didn't occur, or miss events that did take place. "Pearl Harbor would be a failure to warn," Mr. Matheny said.

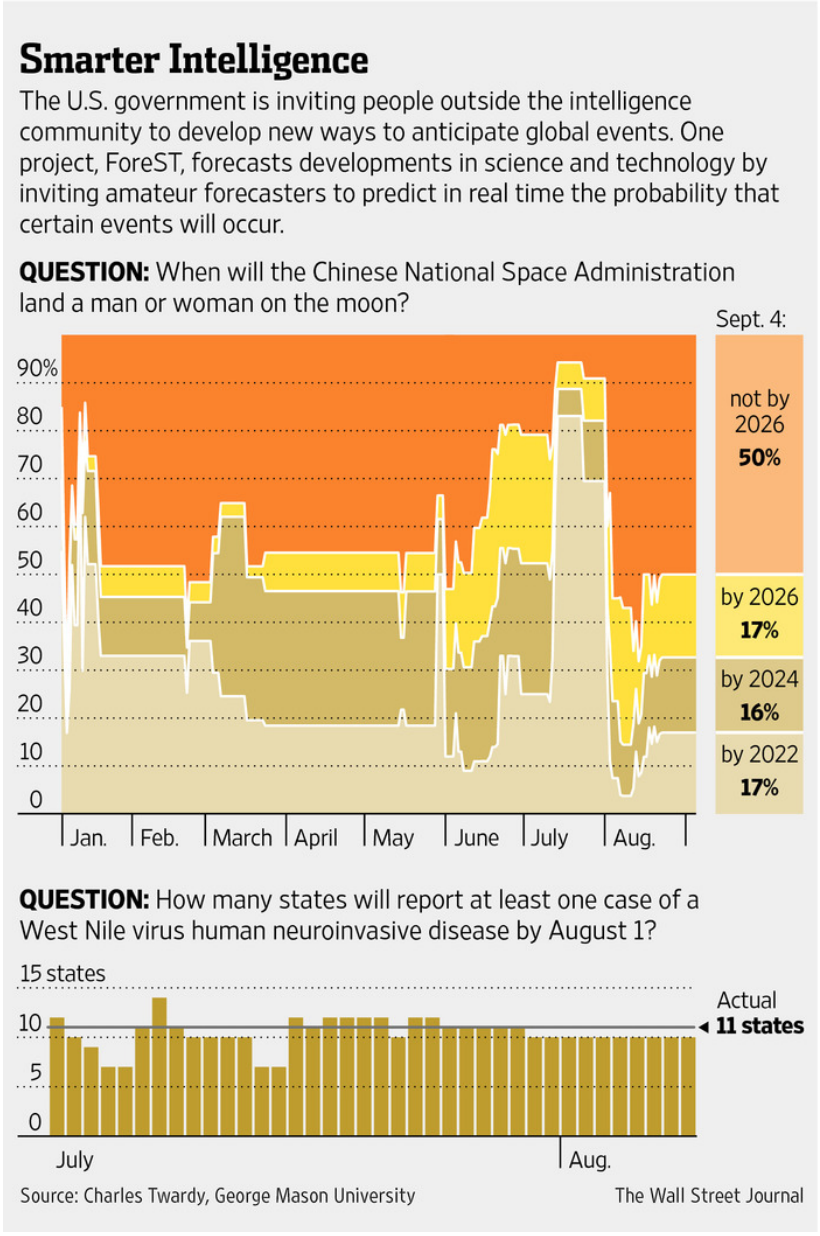
Other examples that took the U.S. by surprise, he said, include the Suez crisis of the 1950s, the fall of the Shah of Iran in the 1970s and, more recently, the al Qaeda attacks of Sept, 11, 2001.

Among the forecasting tournaments run by IARPA are Aggregative Contingent Estimation (ACE), Forecasting Science and Technology (ForeST) and Open Source Indicators (OSI).

Each seeks to improve the precision and timeliness of intelligence forecasts by using techniques such as crowdsourcing, probability scores and machine learning—a field of computer science that "teaches" computers to recognize patterns in data. The strategies offset the weaknesses of any single source and overcome human tendencies to over- or underestimate the possibility of an event.

ACE launched first, in 2011, with five teams competing, but the Good Judgment Project, led by Philip Tetlock of the University of Pennsylvania, beat out its rivals and now is the only team that IARPA funds.

The project uses thousands of amateur forecasters to answer hundreds of specific questions regarding world affairs. The competitors' goal was to beat a control group by 50% according to Brier scores, a measure of the accuracy of probabilistic predictions, by the third year of the four-year competition. "This team beat



that metric by 70%, and they did it in year two," said Steven D. Rieber, who manages the project for IARPA.

IARPA attributes the success to several factors: Forecasters receive a 90-minute training module on cognitive biases and probability. The best are drafted onto elite teams. And all forecasts for a particular question are combined into a single probability attached to each possible outcome, with more weight given to better forecasters.

ForeST, led by George Mason University's SciCast team, broke off from ACE in 2013 to focus specifically on developments in science and technology, and uses similar techniques. Tracking developments in these fields may help identify advances in weapons systems or emerging technologies in bioterrorism or cyberthreats.

"SciCast lets us measure the quality of our predictions and improve them," said Charles Twardy, who leads the project.

Among SciCast's predictions this year was that between 10 and 12 states would report cases of West Nile virus by Aug. 1. Eleven did. And an open question seeks to predict when China will complete its first moonwalk.

While ACE and ForeST address specific questions, OSI, which launched in 2012, identifies previously undetected activity that will culminate in an event of interest—such as a violent protest—within days or weeks.

Among the events the team has detected are the student protests in Venezuela earlier

this year and the 2013 hantavirus outbreak in Chile and Argentina.

EMBERS, the winning OSI project, beat out two competitors and is led by Naren Ramakrishnan at Virginia Tech. To identify brewing events, the team vacuums up large volumes of publicly available data, including online news, blogs, social media such as Twitter posts, Wikipedia pages, weather data, crime rates and images of, for example, parking lots outside hospitals—to name a few of the myriad sources.

The group looks for classes of events, such as political unrest or disease outbreaks. However, instead of attempting to build a single computer model to accommodate all types of forecasts, the team created many separate models with different strengths and weaknesses. Each event class has about eight models, and the conclusions produced by each are weighted and combined to produce a single forecast.

"You can think of it as a little marketplace of algorithms all competing with each other, trying to outdo other predictions," Mr. Ramakrishnan said.

To help gauge the success of OSI, IARPA catalogues world events as they happen and compares the list to events previously forecast by OSI. On average, the program has detected 90% of all eventual events of interest.

None of these techniques will replace standard intelligence approaches, but they are likely to provide analysts with additional tools that treat forecasting as a skill, with results that can be measured and, as a result, refined and improved.

Write to Jo Craven McGinty at Jo.McGinty@wsj.com