Discovering Data Intelligence with AI

These days, analysts have a lot to digest.

Kevin McCaney
Mon, 11/20/2017 - 11:18

The U.S. spy community is looking to artificial intelligence to help find valuable needles of information in the endless hayfields of publicly available information online. Intelligence agencies commonly deal in secrets, but open source information — in this case, the bits of information available electronically to the public — can also become a key source of material. But the sheer volume and variety of available data has left analysts overwhelmed, so agencies are pursuing machine learning and other AI technologies to keep up.

In a general sense, open source intelligence has been around for as long as there have been politics and wars. Herodotus’ account of the Battle of Thermopylae in 480 BC, for example, takes note of the role word-of-mouth rumblings about troop
deployments played in the outcome. For the U.S. military, open source intelligence became an official program in 1941, under the aegis of the Foreign Broadcast Information Service, monitoring short-wave radio transmissions during World War II. A well-known example in intelligence circles was checking the price of oranges in Paris to evaluate the success of efforts to bomb railroad bridges.

Sign up for our weekly newsletter to receive the latest analysis and insights on emerging federal technologies and IT modernization.

These days, analysts have a lot more to digest. Open sources today comprise traditional news outlets, web-based communities, government announcements and reports, and academic events and papers. It includes the likes of amateur radio monitors, airplane spotters and satellite observers. And an increasingly big part is played by social media, a massive global communications platform (Twitter averages more than a half-billion tweets daily) where such things as hashtags, commonly used phrases, geotags on photos, timestamps and other factors can offer evidence of what’s happening or even what’s going to happen.

All that makes it harder to find the seemingly innocuous or off-hand comment — in a tweet, a YouTube video or on page 326 of a droning government report — that, combined with other information, produces valid intelligence. Hence, the heightened interest in AI and machine learning systems, which can help discover useful nuggets of information in terabytes of disparate daily data.

**AI and Visualization**

One example is the machine learning system developed by Primer, an In-Q-Tel-backed startup working with intelligence agencies to analyze data from millions of sources and render the results in a dynamically updated visualization tool with maps displaying topics and places of interest.

Primer uses machine learning and natural language processing to automate the analysis of large troves of data, according to the company. It uses computational engines in a modular architecture to analyze data in multiple formats and languages and then puts the results into readable text reports and graphics.

The Primer web page itself (with an indicative top-level URL domain name at primer.ai) gives an example of how the visualization tool works: The page provides
“updates” on four topics — recently, CRISPR, Syria, Apple Cars and Bitcoin — with text constantly added to a central section and a scrolling column on the side of the page. The information displayed on the home page isn’t “live,” but it offers an idea of what a constantly updated, machine-produced page could look like.

Along with the geopolitical matters of interest to intelligence agencies, the tool also can be applied to other areas, such as finance, science and consumer trends. In addition to In-Q-Tel, the CIA’s nonprofit venture capital arm, Primer’s clients include Walmart and the $100 billion Singaporean sovereign wealth fund, GIC.

CIA’s AI Push

Open source intelligence, of course, is just one front in intelligence agencies’ plans for AI, which can be applied to everything from weapons systems to sorting through the mountains of images and video collected by satellites, drones, the internet of things and other sources. The CIA alone has 137 AI projects underway, many of them in concert with Silicon Valley companies, Dawn Meyerriecks, the agency’s deputy director for technology development, said in September.

Other intel agencies also are pursuing AI projects, as is the government overall. In October 2016, the White House addressed the efforts in two reports, “Preparing for the Future of Artificial Intelligence” and the “National Artificial Intelligence Research and Development Strategic Plan.”

For anyone familiar with the vast expanse of the internet, the volume of data publicly available provides a clear example of the need for AI’s ability to analyze huge volumes of information from multiple, disparate sources and recognize patterns human analysts (understandably) might miss.