

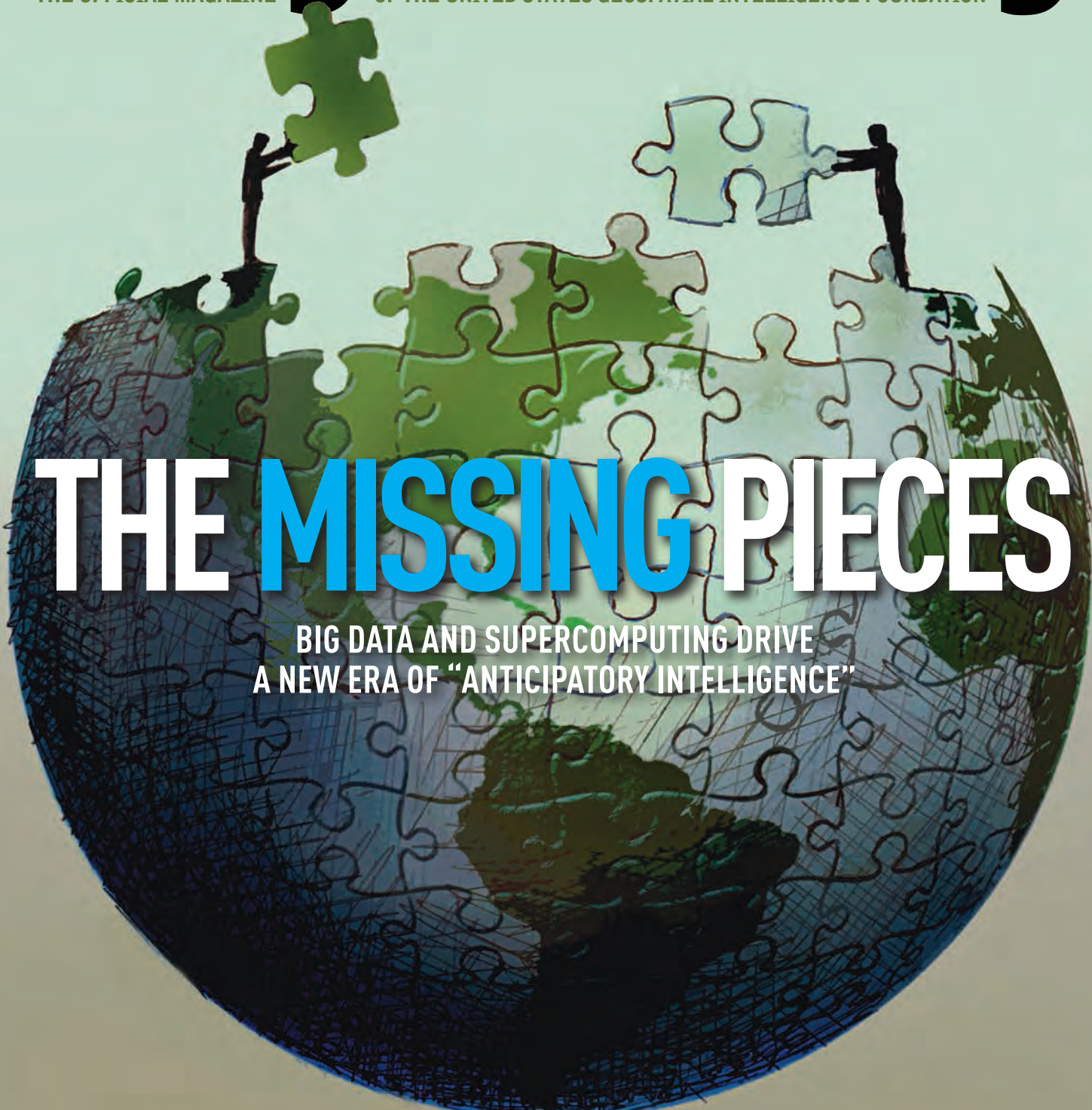
» FIGHTING EBOLA WITH GEOINT » USGIF INNOVATION TASK FORCE » STEM-TO-GEOINT PIPELINE

2015 ISSUE 1

# trajectory

THE OFFICIAL MAGAZINE

OF THE UNITED STATES GEOSPATIAL INTELLIGENCE FOUNDATION



## THE MISSING PIECES

BIG DATA AND SUPERCOMPUTING DRIVE  
A NEW ERA OF "ANTICIPATORY INTELLIGENCE"

the .  
mi



# Missing pieces

BY MELANIE D.G. KAPLAN

THE ADVENT OF  
**BIG DATA AND  
SUPERCOMPUTING**  
IS DRIVING A  
NEW ERA OF  
“ANTICIPATORY  
INTELLIGENCE,”  
BUT HOW WILL  
THE APPLICATION  
OF THESE TOOLS  
FIT INTO THE  
TRADITIONAL IC  
CULTURE?

During the Cold War, the Intelligence Community collected information about Soviet activity: The enemy is loading ammunition; it's moving artillery closer to maneuver forces. The observations were logistical, and the data yielded warnings, explained Collin Agee, U.S. Army senior adviser for Intelligence Community engagement. “Predictive analytics” might be the buzzphrase du jour, Agee said, but the concept is nothing new. In the commercial realm, predictive analytics is picking up steam, while the Intelligence Community is adapting the concept to modern computing technology and Big Data, and simultaneously shifting the nomenclature toward “anticipatory intelligence.”

“Predictive intelligence is deterministic—‘Are the Soviets going to attack?’” Agee said, adding that anticipatory intelligence is a better match for today’s less-defined enemy. “Anticipatory is when you don’t know what the end game is. It’s more cognitive and more sophisticated.”

Big Data, supercomputing, and sophisticated algorithms allow us to process more information than ever before, but the real power is in identifying anomalies—the black sheep in the data.

“It’s things that are unpredictable but can have catastrophic events,” Agee said. “How silly does it sound to say people are going to fly planes into skyscrapers?”

Today’s anticipatory intelligence capabilities likely could have, as Agee said, connected the 9/11 terrorists to one another and generated an alert—present-day capabilities could have connected the dots.

But the number of dots is dizzying. According to David Bray, a visiting associate in cybersecurity and culture at the University of Oxford, the amount of global digital data doubles every two years. With six billion terabytes of data available in 2014, that number is projected to reach

96 billion terabytes by 2022, Bray said.

Behind predictive analytics is the notion that everything, from a shopper’s habits to a tumor’s growth to an aircraft engine’s performance, is predictable. Data can now predict which day is best to buy airline tickets, when a drought will happen, and where the next social uprising might occur.

“Any mathematician will tell you the more factors you have, the better model you’ll build,” said Kirk Borne, data scientist and professor of astrophysics and computational science at George Mason University. “If all I have is a person’s name, I won’t know if someone will commit a terrorist act. But if I know their connections and the crazy things they’re saying on Twitter, all that extra contextual information makes a more accurate and predictive model.”

Predictive analytics applies modeling, data mining, and machine learning, and is also based on the assumption that the past informs the future. Businesses and organizations use exhaustive amounts of historic information to shed light on future actions, events, opportunities, and risks—from identifying casino cheaters in real time to predicting which

customers are about to switch banking institutions—and then preventing the defection.

While commercial industries have embraced predictive analytics, the defense and Intelligence Community—which yearns for clues about tomorrow—is often eschewing the word “predictive” to strive for analysis methods it considers even more complex.

Kevin O’Connell, president and CEO of Innovative Analytics and Training, who also teaches at Georgetown University’s Center for Security Studies, echoed Agee’s statements. O’Connell said “predictive” implies a single-point outcome, such as whether someone is going to buy apples or oranges. In the Intelligence Community, it’s all about anticipation.

“Historically, thinking about the future meant experts picked a small number of future scenarios, often a good one, a bad one, and one that was a straight line extrapolation of today,” O’Connell said. “The current combination of human judgment and computational power allows the rapid consideration of thousands, maybe

## CHOOSE YOUR CRYSTAL BALL

We asked experts about their favorite tools for harnessing Big Data to yield predictions. Here’s our short list of what’s trending:

can filter by keyword, to be monitored in real time around any location of interest.

**geofeedia.com**

**SAP HANA** allows you to turn a hypothesis into a question, and then query a database with real-time results. Whereas an old-school database might allow one query a week, HANA can handle several per day because its data is stored in RAM.

**hana.sap.com**

**GEOFEEDIA** helps you take geo-tagged social media posts from Twitter, Facebook, Instagram,

YouTube, Flickr, Picassa, and Viddy and plot them on a map. This allows social media content, which you

**HADOOP** is an open-source effort designed to scale up from single servers to thousands of machines, each offering local computation and storage and the ability to generate results quickly. Through its InnoVision directorate, NGA is defining areas in Big Data analytics to pilot in a Hadoop sandbox. This would allow analysts to analyze data without first having to go to multiple repositories based on how it was collected.

**hadoop.apache.org**

**GEO-Q and GEOWAVE** are among a number of NGA’s open-source projects available through its GitHub account. Geo-Q is a web infrastructure that helps crowdsource information to support disaster relief and other humanitarian assistance efforts. GeoWave is a set of software that adds multi-dimensional indexing capability to Apache Accumulo.

**github.com/ngageoint**



**DATA CENTER** at a SAP HANA location in St. Leon-Rot, Germany.

millions of scenarios for a more systematic look at the future.”

As the Intelligence Community transitions from a focus on tactical to strategic intelligence, the ability to predict and anticipate events and outcomes will only become more essential. However, Mark Lowenthal, president of the Intelligence & Security Academy, warns there are potential pitfalls in placing all national security problems in the hands of anticipatory intelligence.

“When you start treating Big Data as a panacea for intelligence, you’re in trouble,” Lowenthal said. “That’s what’s been lost in this conversation.”

Human analysts are still vital to collecting not only vast amounts of data, but the right data, and validating it before generating predictions.

#### DISCOVERING VALUE IN THE DATA

Despite inherent doubts about the power of Big Data and predictive analytics, their applications have become mainstream.

This is great news for Ilkay Altintas, director for the Center of Excellence in Workflows for Data Science at the San Diego Supercomputer Center at the University of California, San Diego.

“We’re creating a culture of measuring ourselves,” Altintas said. “We have sensors everywhere.”

Three years ago, Altintas didn’t even know she was interested in her personal fitness data collected through devices such as Fitbit; Now she claims she can’t live without it. The ability to analyze patterns allows us to make smarter decisions, she said.

Altintas’ team has taken on a new project called WIFIRE, aimed at predicting where wildfires will spread. WIFIRE creates models based on video footage, satellite imagery, and real-time wind, temperature, and humidity data.

“What are the changes that happen before an event happens?” Altintas asked. “We can imagine these questions, but it’s really hard to compute this.”

WIFIRE allows the team to build a prototype and learn about a fire while it’s blazing. Although they’re based in San Diego—a very wired city—lack of data remains a challenge. Imagine, Altintas said, expanding the program with all the data that will be generated within the



PHOTO BY MACK A. MOORE

next decade. She said the key is not just modeling the data, but finding value and connection in real time.

“None of it matters if you don’t have an application or important problem you’re trying to solve,” she said. “You look at the application, ask the questions, look for technologies that can be applied, and see what kind of answers are there.”

In some fields, prediction isn’t a novel concept—we’ve known for some time how to model and predict the trajectory of missiles. Industries less experienced in predictive analytics are experiencing the most significant effects from harnessing Big Data, according to Dr. Dave Warner, neuroscientist and director of medical intelligence at MindTel. In his consulting work with Red Bull, Warner is developing new methods to study the narrow dynamics of elite performers, both in traditional and e-sports.

“Traditional biology didn’t have a lot of math help when I was growing up,” said Warner, whose data visualization program, Antz, helps users better understand and process complex data. “Being able to predict to some better degree who are going to be the better athletes or surgeons—predicting for human performance—has the potential to change the

way things (such as preemptive recruitment into specialized performance-based professions) are done.”

In Northern Virginia, Borne consults with a firm that works with the Veterans Administration to identify patterns of behavior indicating suicidal tendencies.

“Of course there is body language,” Borne said, “but maybe they write something in social media using keywords. Maybe they came home to find a cheating spouse. These are hidden variables you’re not seeing, but these are all signals.”

Trends such as crowdsourcing and the share-everything social media mentality help analysts better connect the dots, while simultaneously raising concerns about privacy and security. The “Internet of Things”—the growing interconnection of devices and appliances ranging from heart-monitoring implants to smart thermostats—will only intensify the challenges that accompany the data deluge. Because loosely coupled data could produce inaccurate predictions, confident analysis depends upon keeping humans in the loop.

“The art of GEOINT is uniquely human,” said Spatial Networks CEO Anthony Quartararo. “I don’t think

**116TH MILITARY Intelligence Brigade** soldiers perform daily mission support of the warfighter at Fort Gordon, Ga.



**PARTICIPANTS** at one of the Predictive Analytics Center workshops, hosted by the DHS Science and Technology Directorate's Homeland Security Advanced Research Projects Agency, discuss Big Data challenges and brainstorm potential solutions.

algorithms or artificial intelligence can ever really approach the unique capabilities of what a human brings to the equation.”

It's tempting but perilous to rely on computers, Quartararo added.

Another risk of living in a *Minority Report*-esque world, in which people think they know the crimes others will commit in the future, are false positives.

“After 9/11, the Department of Defense, the Defense Advanced Research Projects Agency (DARPA), and the Department of Homeland Security (DHS) all asked, ‘How can we make sure it doesn't happen again?’” Borne said. “When they looked at patterns, they saw roughly 20 young men from an Arab country who traveled together with one-way tickets on the same credit card, communicating regularly. So the next time you see all these characteristics, and you arrest them, you might just be arresting the Egyptian World Cup soccer team.”

Jeff Jonas, an IBM fellow and chief scientist of context computing, said it's narrow-minded to think rarities are always important. Things happen all the time that have never before occurred; it's all the other factors that might add significance to an event.

“Just because you have a big pile of data doesn't mean there's gold in it,” he said. “People get caught up with, ‘Wow! Look at all the things that are rare.’ If you inspect each one, there's usually nothing interesting. It's generally got to be rare-plus.”

## STUCK IN THE TRIPTIK ERA

When David Kilcullen, founder of strategy firm Caerus Associates and a former counterinsurgency analyst, was a child, his parents took road trips and used the Australian equivalent of AAA's now obsolete TripTik—a spiral-bound flipbook of maps showing a particular route. Today, road-trippers turn to Google or Apple for maps and turn-by-turn navigation. The problem, Kilcullen said, is the government is still stuck in the TripTik era.

“We're trying to bring the intelligence world into the 21st century,” he said. “Policymakers are still asking for stuff that equates to the TripTik, because that's what they've always done.”

An outdated U.S. national security system yields a culture that is reactive rather than anticipatory, which is hardly a winning strategy, according to O'Connell.

“In a world where things are moving so quickly, we have to jump out in front of problems, not wait for them to happen,” he said. “When you respond to something in crisis, your options are typically narrower and costlier. This puts a real premium on anticipating what might come.”

Possible components to modernizing the Intelligence Community and preparing the workforce for anticipatory analysis include training, learning from the commercial sector and other industries, and embracing team-based methods.

Bray addressed the importance of “change agents” in public service, such as those who are willing to develop an incubator space to promote creativity and challenge employees to draft proposals as though they are pitching venture capitalists.

“Public service is doing what the U.S. founders intended: checks and balances,” Bray said. “But now, rapidly changing technology requires us to rethink how we might deliver value to the public for exponential times.”

Carmen Medina, a specialist leader at Deloitte Consulting with three decades of experience with the Central Intelligence Agency (CIA), said it's important to present quantitative information in new ways, beyond prose—but noted this is a largely new approach for the Intelligence Community. Medina said because it once took a

team of analysts 30 days to perform tasks that can now be done in one-hundredth of the time, it's important to reallocate resources and determine which jobs are no longer necessary. Her prediction: a smaller and smarter Intelligence Community in the future.

One example of this shift is the only one-year-old Predictive Analytics Center at DHS's Homeland Security Advanced Research Projects Agency (HSARPA). The center hosts a monthly, hands-on boot camp for 60–100 employees from DHS and other agencies, and covers topics such as indexing methods, extraction, Big Data 101, and Big Data security. HSARPA Innovation Director Stephen Dennis said the purpose of the center is to develop a knowledge base and help users make smart choices with technology.

“Our goal is to be on the cutting edge of analytic technologies,” Dennis said. He acknowledged the incongruity of placing “cutting-edge” and “government” in the same sentence, especially when the commercial innovation cycle is about six months while the government acquisition cycle is significantly longer. But HSARPA is piloting new data storage, visualization, computation and analytics, and security and privacy tools, and is looking decades into the future, he added.

For example, a technology that would allow U.S. Customs and Border Patrol agents to compute risk at exponentially faster speeds could allow them to make quicker decisions at ports of entry, where they are charged with stopping cargo that might pose a threat to the country while maintaining the flow of commerce. The center studies recent innovations at other agencies including DARPA and the Department of Energy—whose national labs have some of the world's fastest supercomputers—and considers whether they could be applied against national security needs.

“It's a lot of education,” Dennis said. “People really want to have a dialogue about the art of the possible.”

To help realize what's possible, the National Geospatial-Intelligence Agency (NGA) has begun a significant initiative to better structure and standardize its data sets. Bryan Goltry, the senior GEOINT officer in NGA's analytic capabilities portfolio, said the agency is in the earlier

stages as it looks to apply the type of predictive and anticipatory analytics to Big Data that industry and commercial GEOINT applications are leveraging.

“Despite making progress in the structuring and standardizing of data sets, we haven’t quite realized the full extent of what Big Data processing can do for us,” Goltry said. “We need to invest more into algorithms that do spatio-temporal processing of remote sensing data. This should, in turn, support linking to other relevant IC data and NGA data holdings.”

Throughout this process, NGA will seek better and easier-to-use tools for interacting with data; systems and algorithms working across multiple security domains 24/7 to correlate the data, find the needles in the haystacks, and tie them together; and capabilities that empower users to answer broader questions and even determine the questions they’re asking.

The Intelligence Community has just barely scratched the surface of anticipatory intelligence, Agee said,

**Big Data, supercomputing, and sophisticated algorithms allow us to process more information than ever before, but the real power is in identifying anomalies—the black sheep in the data.**

and a number of challenges remain. While data standardization, accessibility, and policy challenges are evident, understanding the role of anticipatory intelligence in the modern intelligence cycle is perhaps the subtler but most significant obstacle.

Intelligence Community leaders should consider: how anticipatory intelligence might change the intelligence cycle; how to maximize the utility of anticipatory intelligence against nontraditional threats; and what the appropriate balance is between man and

machine—and between clear-cut predictions and more nuanced anticipations.

Agee said identifying best practices is paramount.

“Since anticipatory intelligence is in a neophyte stage of development, it is not yet codified in doctrine, and there’s nobody really in charge,” he said.

As the community moves forward, advanced analytics will be increasingly necessary to extract meaning from an ever-mounting volume of data. But without skilled human analysts, Agee said, those analytics will not be enough. ■

YOUR NOT-FOR-PROFIT

# Mission Partner in Persistent ISR

SYSTEMS ENGINEERING • EDUCATION & TRAINING • RESEARCH & DEVELOPMENT

Riverside Research delivers end-to-end ISR solutions. Our subject matter experts and integrated research laboratories drive innovation by blending decades of expertise in OPIR, data sciences, cybersecurity, and multi-INT data collection to open the aperture beyond traditional GEOINT applications. As a not-for-profit company, our mission is your mission. Let us know how we can help.

▶ **LEARN MORE**

[www.riversideresearch.org](http://www.riversideresearch.org)  
[bdgroup@riversideresearch.org](mailto:bdgroup@riversideresearch.org)

**RIVERSIDE** RESEARCH