Defense One

Harnessing Big Data to Protect the Nation

Dispelling the Fog of War With Data

"War is the realm of uncertainty; three-quarters of the factors on which action in war is based are wrapped in a fog of greater or lesser uncertainty," Prussian military scholar Carl von Clausewitz wrote in his 1873 book *On War*, thus giving birth to the expression "Fog of War," a term that characterizes the threat environment of today as much as that of Von Clausewitz's time. In 2015, that fog takes the form not of cannon smoke but of information. Too much of it, unstructured and beyond clear analysis, is more blinding than it is revealing.

Here's how Klon Kitchen, an information technology advisor to United States Special Operation Forces, framed the effects of Big Data on the current national security community earlier this year:

"There are currently 1.8 billion active social network users globally, and six and one half billion mobile subscribers. What this means is every minute of every day, these users produce approximately 200 million emails, 72 hours of new YouTube video, 571 new Websites, 3600 new photos on Instagram alone, 100,000 tweets, 34,220 Facebook likes and 2 million Google searches. That is every minute of every day with an Internet penetration rate of just 35 percent …This is the technological context for every future special operations mission. Our Special Operations Forces will need to employ, and they will be confronted by, an almost unimaginable deluge of data and unprecedented technological capability."

The United States must contend with potential adversaries ranging from peer nations like China and Russia to masked men broadcasting terror from the hilltops of Syria, to hackers around the world. Some of those threatening actors will be exploiting gaps in networks to steal the military's most important secrets. And they'll be harnessing the power of information technology to recruit, communicate, and market themselves globally.

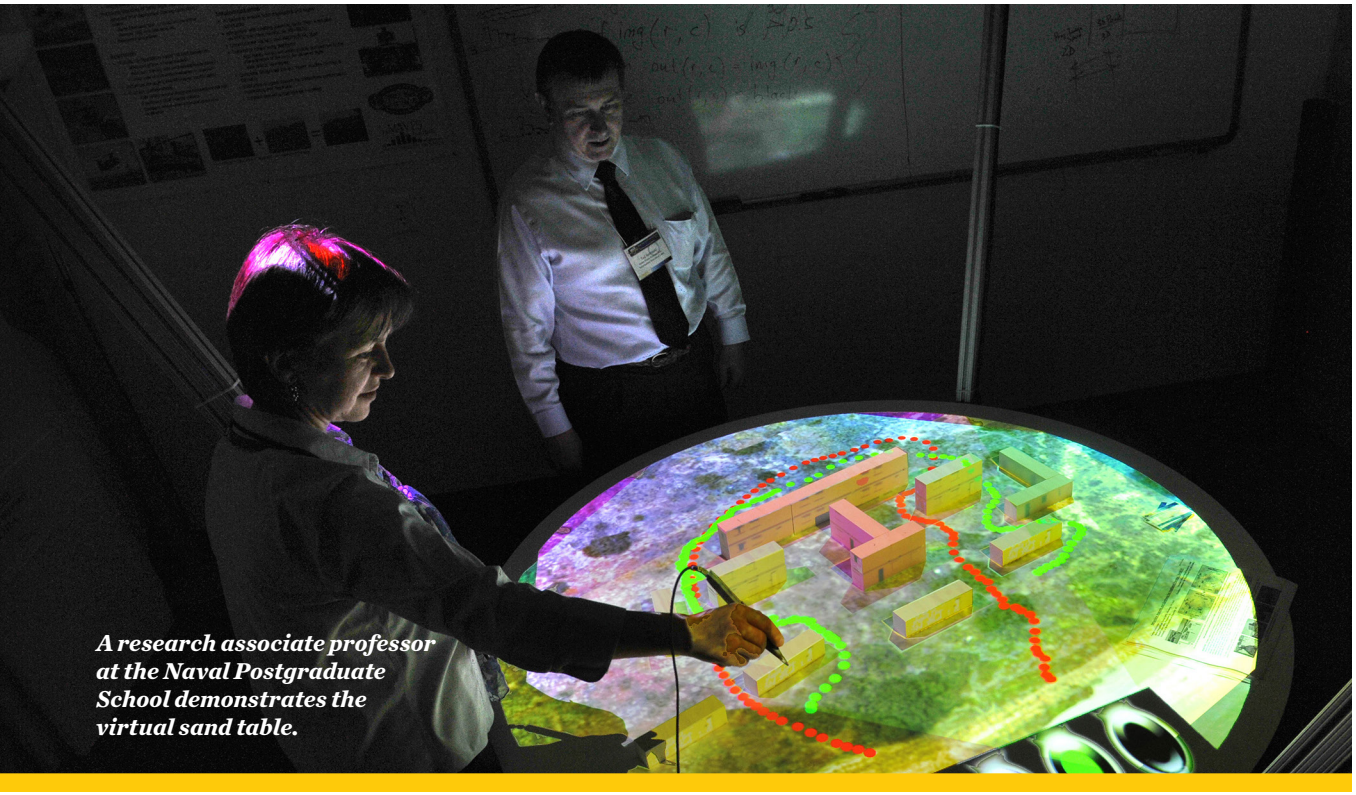
In this eBook, *Defense One* looks at the ways that the United States military and its partners will dispel the fog of information overload. In the years ahead, the national security community will use Big Data to predict outcomes of decisions ahead of time, to anticipate the threat environment of the future, to detect threatening network behavior, and even to unmask terrorists. The ability to make sense of the data deluge in the coming decade will mean the difference between victory and defeat.

Patrick Tucker
Technology Editor
Defense One

How the Military Is Turning Big Data Into a Crystal Ball

Massive amounts of data could yield new insights, or false clues, about the threat environment of the future

BY PATRICK TUCKER



A research associate professor at the Naval Postgraduate School demonstrates the virtual sand table.

JOHN FISCHER/U.S. NAVY

What could the military do if it could better understand the massive amounts of data that humanity creates, an estimated 2.5 quintillion bytes every day? Could it predict aspects of the future?

If Pentagon funds can help create—even partially—a machine capable of understanding cause and effect, or causality, and do so on the scale of thousands of signals, data points, and possible conclusions, then, perhaps, big data will reach its real potential: a predictive tool that allows leaders to properly position soldiers, police forces, and humanitarian relief long before the action starts.

Among the military programs probing this new realm is Big Mechanism, run by the Defense Advanced Research Projects Agency, or DARPA. It seeks to turn machine-collected (or machine-generated) data into real insights into complex systems, and do so automatically.

Some, such as *Wired*'s Chris Anderson, have suggested that access to huge amounts of data, which makes correlational analysis easier, has made old-fashioned, theory-based science obsolete.

> **No single human can understand a really complicated system in its entirety. Computers must help us.**
>
> PAUL COHEN, DARPA PROGRAM MANAGER

But in a recent conversation with *Defense One*, DARPA program manager Paul Cohen said he was looking more to mechanize the human capacity for causation, rather than innovate around it. "We're very much aiming toward a new science, but we're very much interested in causal relationships," he said. "What we're finding is that mathematical modeling of systems is very hard to maintain."

The supply of data, it turns out, is growing too quickly for the human race to use it effectively to solve big problems. The expanding reach and power of computational intelligence is both cause and, at least potentially, cure.

"Having big data about complicated economic, biological, neural and climate systems isn't the same as understanding the dense webs of causes and effects—what we call the big mechanisms—in these systems," Cohen said last year. "Unfortunately, what we know about big mechanisms is contained in enormous, fragmentary and sometimes contradictory literatures and databases, so no single human can understand a really complicated system in its entirety. Computers must help us."

These big systems can be as large as the entire world or as small as cancer cells, an initial area of focus for the program.

Machine intelligence can collect and process data on a scale unimaginable to regular humans. But processing data is very different from making sense of it, and from making predictions. If we could get computer systems to predict in the way that humans do, but with the data and processing power only available to massively interconnected systems, could we open up areas of the future to new inference? Cohen has suggested that the answer is yes.

"The beautiful thing about causal models is that they make predictions, so we can return to our big data and see whether we're [retrospectively] right," Cohen said. "And we can propose new experiments, suggest interventions and advance our knowledge more rapidly."

The ways in which humans interact with government, with one another, with medical facilities, transit systems and brands, etc. can predict events of national security significance. They can indicate, for instance, if a deadly disease outbreak is taking hold in a small rural community or if civil unrest is on the rise.

One example of that is the Open Source Indicators Program, launched in 2011 by from the Intelligence Advanced Research Programs agency. Led by program manager Jason Matheny, Open Source Indicators funds projects to predict events of national security relevance by monitoring tens of thousands of blogs, RSS feeds, news reports, social network chatter from sites like Twitter and Facebook, and other open sources.

Very early on, program participants began to generate some surprising results. In 2012, Virginia Tech computer scientist Naren Ramakrishnan, working solely with signals culled from the open Internet, effectively predicted both Mexico's Yo Soy 132 protest movement, sometimes called the Mexican Arab Spring, and the "Friendship Bridge" protests that riled parts of Brazil and Paraguay.

Around the same time, Georgetown University data scientist Kalev Leetaru used a database of millions of open-source indicators to correctly (but retroactively) predict the spot in Abbottabad, Pakistan, where Osama Bin Laden was found.

> The **beautiful thing about causal models is that they make predictions, so we can return to our big data and see whether we're [retrospectively] right.**
>
> PAUL COHEN, DARPA PROGRAM MANAGER

But for every instance where big data correctly predicted a big national-security event, critics can point to a big miss. Last year, for example, such indicators failed to predict the Ebola outbreak.

The military and national security communities have only begun to explore the potential of big data to solve these kinds of enormously complex problems. But before open-source signal hunting can reach its full potential, people like Cohen and Matheny need to answer some serious questions.

Among them, how to balance privacy concerns with national security objectives? Open-source intel is, by definition, freely available on the Internet. But when most people give their data away, they don't imagine military technologists trying to extrapolate predictions from that data. As more people become concerned about how their data is used, especially by government actors, they're changing the data that they make and release.

Last summer, Lt. Gen. Michael Flynn, then the director of the Defense Intelligence Agency, discussed how the military had "completely revamped" the way it collects intelligence around open-source data. He said that, as the military's reliance on such data grows, online behavior is changing and adapting. When asked if he was concerned by that, he answered, "Yes."

Another question: how do we get machine intelligence to discover causation and not just correlation? How do you teach a big-data machine to provide a fully true answer, not just an output? Answer: we may need to re-design, on a fundamental level, the way we communicate with machines.

Perhaps one of the greatest paradoxes of the modern age is that our method for interacting with machine intelligence remains relatively crude: typing, and more and more with our thumbs. That limits our collaboration with machines to specific tasks with strict parameters.

When given a chance to collaborate with a machine or with a human on a project of any complexity, we'll press zero to talk to the human almost every time. That's a problem in a national-security context: service members are being asked to interact with an ever-larger number of systems in life-and-death situations.

Another one of Cohen's DARPA programs, announced in February, seeks to change that. The Communicating with Computers effort seeks to "bridge the language barrier" between humans and machines, Cohen remarked in a press release.

"Human-machine communication falls short of the human-human standard, where speakers and listeners consider such contextual aspects as what has been said already, the purposes of the communication, the best ways to express ideas, who they are speaking with, prevailing social conventions and the availability of other modes of expression such as gestures. And so computers that might otherwise contribute more significantly to solving problems in a range of areas, including national security, remain in relatively simplistic roles such as crunching large datasets and providing driving directions."These new technologies are helping the U.S. Air Force to renovate buildings that are structurally sound and energy efficient. They're also allowing the U.S. National Park Service to take a new look at history, studying a sunken World War II battleship from beneath the sea.

Two airmen work in the Global Strategic Warning and Space Surveillance System Center at Cheyenne Mountain Air Force Station.

KRYSTAL ARDREY/U.S. AIR FORCE

Can the Intelligence Community Predict Online Attacks?

The Intelligence Advanced Research Projects Activity wants a machine that can comb through open source data get ahead of hackers.

BY ALIYA STERNSTEIN

Imagine if IBM's Watson — the "Jeopardy!" champion supercomputer — could answer not only trivia questions and forecast the weather, but also predict data breaches days before they occur.

That is the ambitious, long-term goal of a contest being held by the U.S. intelligence community.

Academics and industry scientists are teaming up to build software that can analyze publicly available data and a specific organization's network activity to find patterns suggesting the likelihood of an imminent hack.

The dream of the future: A White House supercomputer spitting out forecasts on the probability that, say, China will try to intercept situation room video that day, or that Russia will eavesdrop on Secretary of State John Kerry's phone conversations with German Chancellor Angela Merkel.

> **Instead of reporting relevant events that happen today or in previous days, decision makers will benefit from knowing what is likely to happen tomorrow.**
>
> ROB RAHMER, IARPA PROGRAM MANAGER

IBM has even expressed interest in the "Cyber-attack Automated Unconventional Sensor Environment," or CAUSE, project. Big Blue officials presented a basic approach at a Jan. 21 proposers' day.

Aims to Get to Root of Cyberattacks

CAUSE is the brainchild of the Office for Anticipating Surprise under the director of national intelligence. A "Broad Agency Agreement" — competition terms and conditions — is expected to be issued any day now, contest hopefuls say.

Current plans call for a four-year race to develop a totally new way of detecting cyber incidents — hours to weeks earlier than intrusion-detection systems, according to the Intelligence Advanced Research Projects Activity.

IARPA program manager Rob Rahmer points to the hacks at Sony and health insurance provider Anthem as evidence that traditional methods of identifying "indicators" of a hacker afoot have not effectively enabled defenders to get ahead of threats.

This is "an industry that has invested heavily in analyzing the effects or the symptoms of cyberattacks instead of analyzing and mitigating the — cause — of cyberattacks," Rahmer, who is running CAUSE, told *Nextgov* in an interview. "Instead of reporting relevant events that happen today or in previous days, decision makers will benefit from knowing what is likely to happen tomorrow."

The project's cyber-psychic bots will estimate when an intruder might attempt to break into a system or install malicious code. Forecasts also will report when a hacker might flood a network with bogus traffic that freezes operations – a so-called Denial-of-Service attack.

Such computer-driven predictions have worked for anticipating the spread of Ebola, other disease outbreaks and political uprisings. But few researchers have used such technology for cyberattack forecasts.

At Least 150 People Interested — No Word Yet on Size of the Prize Pot

About 150 would-be participants from the private sector and academia showed up for the January informational workshop. Rahmer was tight-lipped about the size of the prize pot, which will be announced later this year. Teams will have to meet various minigoals to pass on to the next round of competition, such as picking data feeds, creating probability formulas and forecasting cyberattacks across multiple organizations.

At the end, "What you are most likely to be able to do is say to a client, 'Given the state of the world and given the asset you're trying to protect or that you care about, here are the [events] you might want to worry about the most,'" David Burke, an aspiring participant and research lead for machine learning at computer science research firm Galois, said in an interview. "Instead of having to pay attention to every single bulletin that comes across your desk about possible zero days," or previously unknown vulnerabilities, it would be wonderful if some machine said, "These are the highest likelihood threats."

His research focus is "advanced persistent threats," involving well-resource, well-coordinated hackers who conduct reconnaissance on a system, find a security weakness, wriggle in and invisibly traverse the network.

"Imagine that CAUSE was all about the real-world analogy of figuring out whether some local teenagers

> **The goal is not to replace human analysts but to assist in making sense of the massive amount of information available.**
>
> ROB RAHMER, IARPA PROGRAM MANAGER

are going to knock over a 7-Eleven. That would be really hard to predict. You probably couldn't even tie that to any larger goal. But in the case of APTs — absolutely" you can, Burke said in an interview. "The fact that APTS are on networks for a long period of time gives you not only the sociopolitical pieces of data or clues but you have all sorts of clues on your network that you can integrate."

It's not an exact science. There will be false alarms. And the human brain must provide some support after the machines do their thing.

"The goal is not to replace human analysts but to assist in making sense of the massive amount of information available and while it would be ideal to always find the needle in a haystack, CAUSE seeks to significantly reduce the size of the haystack for an analysts," Rahmer said.

Unclassified Program Will Trawl for Clues on Social Media

Fortunately or unfortunately, depending on one's stance on surveillance, National Security Agency intercepts will not be provided to participants.

"Currently, CAUSE is planned to be an unclassified program," Rahmer said. "We're going to ask performers to be creative in identifying these new signals and data sources that can be used."

Participants will be judged on their speed in identifying the future victim, the method of attack, time of future incident and location of the attacker, according to IARPA.

Clues might be found on Twitter, Facebook and other social media, as well as online discussions, news feeds, Web searches and many other online platforms. Unconventional sources tapped could include black market storefronts that peddle malware and hacker group-behavior models. AI will do all this work, not people. Machines will try to infer motivations and intentions. Then mathematical formulas, or algorithms, will parse these streams of data to generate likely hits.

One research thread Burke is pursuing examines the "nature of deception and counterdeception, particularly as it applies to the cyber domain," according to an abstract of his proposers' day presentation.

"Cyber adversaries rely on deceptive attack techniques, and understanding patterns of deception enables accurate predictions and proactive counter-deceptive responses," the abstract stated.

It's anticipated that supercomputer-like systems will be needed for this kind of analysis.

For example, "if you were able to look at every single Facebook post and you processed everything and ran it through some filter, through the conversations and the little day-to-day things people do, you could actually start to see larger patterns and you could imagine that is a ton of data," Burke said. "You would need some sort of big data technology that you'd have to bring to bear to be able to digest all that."

Still Nailing Down Specifics on Supercomputer Use

The final rules will indicate whether companies can or must use a supercomputer, and whether they can borrow federal computing assets, Rahmer said. "We definitely want innovation and creativity from the offerers," he added.

> **In theory, the government could say they are going to own the servers. We don't know ultimately that we would participate or what we even would propose.**
>
> MICHAEL B. ROWINSKI, IBM SPOKESMAN

Researchers at Battelle, a technology development organization, said they might harness fast data processing engines like Hadoop and Apache Spark. They added that the rules and their team partners will ultimately dictate the system used to amp up computing power.

"We have already recognized as both the rate of collection and the connections between data points grow we will need to move to a high-performance computing environment," Battelle's CyberInnovations technical director Ernest Hampson said in an email. "For the CAUSE program, the data from several contractors could push us towards the need for a supercomputing infrastructure using technologies such as IBM's Watson to support deep learning," or, hardware such as a Cray Urika "could provide the power to fuel advanced analytics at-scale."

According to IBM's January briefing, the apparatus currently used to solve similar prediction problems "runs on x-86 infrastructure." However, IBM's x-86 supercomputer hardware was spun off to Chinese firm Lenovo last year. It remains to be seen what machine IBM might deploy, a company spokesman said.

"In theory, the government could say they are going to own the servers," IBM spokesman Michael B. Rowinski said. "We don't know ultimately that we would participate or what we even would propose."

Recorded Future, a six-year-old CIA-backed firm, already knows how to generate hacker behavior models by assimilating public information sources, like Internet traffic, social networks and news reports. But the company's analyses do not factor in network activity inside a targeted organization, because such data typically is confidential.

"Doing this successfully is not simply the sociopolitical analysis applied to current flashpoints," Burke said. "You also have observables on a network: signs possibly of malware or penetration because many campaigns that take place go on for weeks or months. So you also have a lot of network data that you are going to end up crunching."

DID YOU KNOW?
CDW·G has a strong past performance working with DoD customers across the globe.

YOU and CDW-G

and NOW YOU HAVE THE RIGHT TECHNOLOGY FOR MISSION SUCCESS.

As a Department of Defense agency, your IT department has to do more with less. Because of this, it's crucial to utilize cost-effective technology solutions that help your agency complete its mission. With more than 20 years of experience, CDW·G has the tools and the resources you need to do just that. We're ready to help you plan, design and implement a tailored solution with the expertise of our team dedicated to our DoD customers — making technology purchasing and procurement a surprisingly smooth process.

Learn more about partnering with CDW·G by contacting your account manager at 800.808.4239 or visiting CDWG.com/fedgov

paloalto NETWORKS
the enterprise security company

CDW·G PEOPLE WHO GET IT

Soldiers from the 19th Special Forces retrieve a fast rope after being dropped from a Pave Hawk helicopter during a training exercise.

KEVIN J. GRUENWALD/U.S. AIR FORCE

Big Data Goes to War With Special Operations Forces

Today's battlefields are awash in data. The question is how to harness it for tactical advantage.

BY PATRICK TUCKER

The U.S. fight against the Islamic State and other extremist threats is increasingly in the hands of elite special operations units who will succeed or fail by their ability to collect, process, and exploit data at the speed of crisis. At the command level, that means reducing the number of analysts required to get data to make sense. On the ground, it means sending much more actionable data to the tip of the spear, and doing so faster and more cheaply. Even the best tech minds in commercial sector don't produce the sort of product that special operators need, according to special operations intelligence experts.

Today, the cutting edge looks something like this: Imagine the world's highest-resolution commercial satellite, the WorldView 3 from satellite image provider DigitalGlobe floating 383 miles above the Earth's surface. It snaps pictures of a residential neighborhood in a bustling African metropolis. One particular city street is free of cars. That's common in parts of this country but unusual for the blocks surrounding the U.S. Embassy. Less than three hours later, a member of SOCOM takes notice. He's been living in that country for years, and has inroads with local military units who protect the fragile government from overthrow.

Electricity is spotty, and communications are far from secure, but the operator needs nothing more than a standard laptop running the Chrome browser

> **Our Vricon joint venture will enable SOF operators to reliably make shareable 3D solutions available to enable coalition efforts to counter emerging threats around the world.**
>
> TONY FRAZIER, DIGITALGLOBE SENIOR VICE PRESIDENT

to pull in the satellite imagery. He reviews shots of the area from the previous several months to confirm that today's lack of traffic is abnormal.

He opens a layer on his map, and red clouds spread out where open-source intelligence has predicted civil unrest. A third layer shows tweets from the area that mention a recently killed militia member. A flash of circles—green and yellow, large and small—show the locations of the tweeters. One of the accounts is associated with three attempted bombings in other parts of the country, but is new to the capital. So the operator shares the view with his contact in the local tactical police unit, providing the Twitter handles and locations of the people he's most worried about. A minute later, he opens a topographical layer and finds a flat spot to land a special operations helicopter. The embassy is evacuated just minutes before a car bomb explodes.

Some of these capabilities are already in use by special operators. "Say that an event happened and you needed to figure out how to do an evacuation; you're coming in during the day and using urban tactics to come in," Paul Millhouse, DoD and Federal /Civilian Technical Solutions at DigitalGlobe explained to Defense One during a live demo. "In literally 5 seconds, you're going to have an overlay that will tell you, 'Here are the areas that you should focus on for landing.' To put this into perspective, if you were to do this the traditional way, you would have to download gigabytes of imagery, gigs of elevation data, put it on a high-end machine, get expensive software, run it all – four to eight hours' worth of work." Milhouse said that the DigitalGlobe tools requires little bandwidth. "Rather than transmitting these input ingredients and performing the analysis on a high-end workstation, we perform the processing where the data is, and only send the resulting 10-kilobyte analysis overlay. All the hard work happens in the background."

He said the hardest part is the second layer: predicting hotspots where civil unrest might occur. That's created by a mixture of proprietary software and survey data collected on the ground, and it's already in use by special operations forces.

The company's newest offering is perhaps its most ambitious yet. On Tuesday, DigitalGlobe announced Vricon, a new partnership with Saab to create a fully accurate 3D model of the Earth. It's primarily aimed at the commercial marketplace, but could have relevance for special operations forces. "Identifying safe locations for infiltration or exfiltration, conducting radio-frequency propagation analysis for communications planning, and route planning all require high-resolution elevation data. Our Vricon joint venture will enable SOF operators to reliably make shareable 3D solutions available to enable coalition efforts to counter emerging threats around the world," DigitalGlobe senior vice president Tony Frazier told Defense One in an email.

Moving Data

Neither DigitalGlobe nor anyone in Silicon Valley can solve some of the biggest challenges that face the special operations forces community: the nation is relying on them to solve too many problems in too many places. But they just might be able to help SOF with data issues.

Special operations forces have two big problems: they need better intelligence extracted from data and they need to be able to collect it and deliver it in in an unclassified setting under challenging communications situations. The solution to

> **Eighty percent of our portfolio is geared toward the air. This is also where our spending and investment has been.**
>
> COL. MATTHEW D. ATKINS

both problems involves teaching software to learn to discriminate useful intelligence from raw, unstructured data, a subfield of artificial intelligence or AI.

Today's operators have more raw data than battlefield bandwidth. "Something DoD needs to get its arms around" is "moving sensor data around the battlefield to the folks who need to exploit it," said Air Force Col. Matthew D. Atkins, chief of the intel capabilities and requirements division at United States Special Operations Command at a recent industry event in Tampa, Florida. "We do need help solving the data transport problem from a technological perspective," he told a group of industry representatives at a recent conference in Tampa, Florida. "Every time we roll out a new [high-definition] sensor, a new widget, we crush the data rates."

It's a problem that's only going to grow as data-collection devices improve and special operations forces expand the sorts of data they use. That will include imagery from expensive high-flying drones like the 47-foot Northrop Grumman Global Hawk and from novel sources like ProxDynamics' tiny PD 100 Black Hornet, used for years by British Special Forces in Afghanistan. More and more of it will come from body-worn sensors and intelligence-gathering equipment mounted on trucks and at bases.

"Eighty percent of our portfolio is geared toward the air. This is also where our spending and investment has been," Atkins said, referring to manned and unmanned aircraft built for intelligence, surveillance and reconnaissance, or ISR. "We need to reduce our reliance on airborne platforms. As a result… we'll be putting our efforts into ways to expand ground-based and maritime-based ISR. That's not only to gain dominance in those domains but also to buy down the dependency on airborne, which is the most costly," he said.

Ground-based data and intelligence poses different problems if you're an analyst trying to make sense of data than does drone-based ISR. "When you're up 15,000 feet, your signal environment is dense. When you are man-portable and in a street, there's a variety of devices to contend with, so you don't need this wide band. You have a narrow band so your ability to sense and understand is a lot more localized. The trick is to get it back to a node where they can take advantage of it. A lot of it is stuff you'll process audibly…Making the kit as user-friendly for that guy is important, but also getting it back to [Joint Operations Center] or a [Combined Air Operations Center] and then nest it into the bigger picture."

Meanwhile, U.S. Special Operations Command is trying to find ways to do data processing, exploitation, and dissemination, or PED, with fewer people, Atkins said. "PED troubles us the most," he said. "It's the most human- intensive." He said U.S. Special Operations Command is interested in "investments to buy down that manpower burden." The goal is to do what now requires 500 analysts in a dark room with just one or two in a forward base.

Among the particular challenges of data for SOF is its variety. "It's pretty much anything," Atkins said. The commercial world is full of big data analysis tools that wring insight data in large volumes or at great speed. But data researchers will tell you that big data comes in three flavors, volume, velocity, and variety.

"We're the variety folks," Atkins said. "We gather strange things off targets, pocket litter, yearbook

> **When we're trying to personally ID an individual, there is zero margin for error.**
>
> COL. MATTHEW D. ATKINS

pictures. It's like, how [does SOF] make sense of that? How do you enrich it? How do you tie it to a geospatial location? So that's really the challenge…I go out to Silicon Valley and stump that all the time."

That variety problem creates a very particular need for machine learning and artificial intelligence. Virtually every enterprise that operates on a global scale uses AI or machine learning for something, whether to optimize product delivery (Amazon) or create better online interactions (Facebook). Investors buy satellite imagery from DigitalGlobe, then uses complex artificial intelligence to count the cars in parking lots on Christmas Eve, identify their owners' tendencies based on make and model, and deduce the effects on future earnings statements for companies like Home Depot.

But the market doesn't care if it's off by a car or two. Navy SEALs need their AI to be far more precise. "When we're trying to personally ID an individual, there is zero margin for error… when we're trying to count kids that might be in a compound before we go assault it, that's something that you have to default back to—not just one human but five humans because of the margin for error."

While AI has become a hot field in Silicon Valley, there's just no commercial outfit that's meeting what special operations needs. Atkins is seeking to import more useful data from the field, use AI to turn it into intel actionable by policy folks and the soldiers in mid-mission, do it all at low data and energy rates and—perhaps most importantly—in a way that allows special operators to immediately share it with the international partners working alongside them. All of that is in the near future.

The United States is leaning heavily on the special operations community to serve as the nation's on-the-ground response to the Islamic State. Some missions, like the recent raid that resulted in the death of a top ISIS commander and the capture of a treasure trove of important data, will look perfectly executed. Others, such as hostage extractions, offer much higher levels of complexity. The difference between a successful operation and one that fails, producing headlines and tears, such as the botched special operations rescue attempt for U.S. hostage Kayla Mueller that took place in February, is often a matter of intelligence.

When asked about how to create intel to save hostages during such attempts, in situations where the United States is not willing to commit human assets to ensure mission success, Atkins acknowledged that there was no technological answer. "I'll be honest," he said, "we don't have that solution."

The outside of the National Geospatial Intelligence Agency in Springfield, Virginia.

TREVOR PAGLEN/CREATIVE TIME REPORTS

For One Intelligence Agency, Transparency Isn't a Burden, It's a Strategy

Can the National Geospatial Agency make open data a new norm for U.S. spies?

BY PATRICK TUCKER

To the average American, the term intelligence agency refers to a group of secret military types, locked in a windowless room in Virginia, furtively collecting data on bad guys, good guys, citizens, everybody. That data is delivered up the chain in manila envelops marked "Top Secret." There's still some truth to that stereotype (apparently, they get to have windows now) but Robert Cardillo, director of the National Geospatial Intelligence Agency, or NGA, is hoping to secure an unconventional legacy as a spy chief.

The agency receives around $5 billion annually, according to documents published by the *Washington Post*, and is primarily charged with collecting pictures from space. It's a job that lends itself to parody. Think of Jon Voight's character in "Enemy of the State," who oversees an intelligence agency heavily reliant on satellite imagery to snoop on innocent people.

But that's not the image Cardillo wants to project, either of NGA or of himself. "I would like to amaze people with how relevant and responsive [NGA] can be in the open....What I would like to surprise people about... is how NGA can come out and be more relevant with public diplomacy."

> **The monopoly is over. It was secure. We had space and time and no competition at all.**
>
> ROBERT CARDILLO, DIRECTOR OF THE NATIONAL GEOSPATIAL INTELLIGENCE AGENCY

Cardillo, at an Intelligence and National Security Alliance dinner, laid out his plans for the future of satellite intelligence collection. It's a future where enormous, expensive, military intelligence, surveillance and reconnaissance, ISR, satellites—still up there—play a more subordinate role to commercial image providers.

NGA is the biggest customer of high-resolution satellite imaging company DigitalGlobe, but they are hardly the company's only client. DigitalGlobe has become more and more public facing in recent years. Want to see incredibly detailed (25-centimeter resolution) pictures of the Earth from space? Thanks to decision from the National Oceanic and Atmospheric Administration you'll soon be able to buy those, though not cheaply.

In a conversation with reporters, Cardillo said that scads of cheap satellites from nimble space startups like Skybox, recently acquired by Google, Planet Labs, and BlackSky Global have "huge potential" to fill intelligence needs. A glimpse of the numbers reveals why: the market for satellite imagery, estimated at around $12 billion is twice NGA's budget. That's why the private players are leading the innovation.

Contrast Cardillo's enthusiasm for open-source satellite intelligence with the vision that under secretary of defense for intelligence Michael Vickers outlined during the Defense One Summit.

When asked about areas where he saw military intelligence technology moving forward, Cardillo said, "In the global coverage area, for the first time, we're trying to create really persistent surveillance from space, rather than having episodic surveillance – actually be able to stare at areas for real long periods of time and improve the resiliency and the integration of our architecture." Importantly, these would be different from geo-stationary orbiting satellites for communication, which have limited ISR capabilities. "Those will be really, really big things when they're realized," he added. "It will be a leap in overhead reconnaissance commensurate to anything we've done in the last 50 years or so, but they'll take a decade-plus to realize."

It's a lofty goal, but in the meantime Cardillo is hopeful that smaller players will be providing not just more images to agencies but also delivering insight into what those images mean. "My sense in reading their business material is that they're into delivering the analytics," not just a picture. "I don't see these companies selling pictures so much as analytics," he said.

Satellite imagery and analysis used to be the exclusive domain of the United States and the Soviet Union. "The monopoly is over. It was secure. We had space and time and no competition at all," notes Cardillo.

Consider the firm Orbital Insights, founded by former Googler James Crawford. The company earned attention from *The Wall Street Journal* when it began selling of images of more than 30 Chinese construction sites to hedge fund traders. As the Journal reports, the size of the shadows of buildings under construction can provide a more—shall we say—truthful indicator of economic conditions than can official numbers. That's important for Wall Street types but also for national security watchers, since some scholars have pegged geopolitical insecurity in China to GDP falling beneath 6 percent per year. But Orbital

Insights also provides pictures and analysis related to crops and farmland. Where Wall Street will want to know if they should short Columbia coffee futures, military decision makers could be looking for signs of an impending food shortage, which could have effects on stability on a global scale. For evidence of that just look at the way that high inflation in European Food Price Index telegraphed the Arab Spring. Satellite analysis of mall and discount store parking lots can predict a bad year for holiday sales or epidemic outbreaks.

"It's never about the image. It's about the answer to the question around that image," says Cardillo.

Nowhere is that truer than in those parts of the world where violence is increasing and human rights are under threat but where the media, largely, fears to tread. The democratization of satellite imagery promises to shine a bright light on such places. One of the earliest and most interesting examples of citizen crisis monitoring with open satellite imagery comes from Harvard's Satellite Sentinel Project, or SSP. In 2011, SSP analysis of the disputed area of Abyei saw that Sudan was building large roads in areas where there was no economic reason for those roads to exist. "These roads indicated the intent to deploy armored units and other heavy vehicles south towards Abyei during the rainy season," the group wrote in their report. By May of that year, the Sudanese Armed Forces had begun a formal invasion.

In his conversation with reporters, Cardillo highlighted the recent work of Human Rights Watch as an example of how open source satellite data, in the hands of nonprofits, can change the national discussion about a place, an event, or an ongoing situation.

The international nonprofit has been using satellite imagery to document militant group Boko Haram's campaign of arson as on January 10th, when the group observed "compelling evidence of widespread fire burn scars and building damages affecting approximately 57% of the town of Doro Gowan, Borno State. It is likely the majority of fire damages occurred during the evening of January 3, and morning of January 4, 2015, based on a review of active fire detection data from an environmental satellite."

Amnesty International, too, has been using satellite images to showcase the worsening situation near eastern Nigeria, including this month's massacre of an estimated 2,000 civilians. "These detailed images show devastation of catastrophic proportions in two towns, one of which was almost wiped off the map in the space of four days," Daniel Eyre, Nigeria researcher for Amnesty International, told CNN.

"NGA did a classified assessment as well," Cardillo says in response to the ongoing Boko Haram situation. "We have unique sources that we can use." They also have a different client or consumer than does Amnesty International or Human Rights Watch. But Cardillo sees those audiences, as well as those capabilities, merging. Cardillo is proud of the fact that more than 99 percent of the data NGA secured as part of a project to monitor Ebola spread in West Africa was unclassified. In addition, he said, "We are the first in the [intelligence] community to `crowd source' applications development," through the crowd code site GitHub.

Classified satellite data just isn't worth as much as it was when the object of our obsession was the Soviet Union, and that's particularly true of our most recent conflicts. Cardillo addressed the difficulty of providing a window from space into the activities of the enemy *du jour*: the Islamic State (also known as ISIS or ISIL). "They don't tend to wear uniforms…for the most part, ISIS is building on a frustration on Sunni natives. For us at a distance, to try and determine the ISIS level of control over that village? You're into a difficult thing to assess. We try to document what we can, plotting the Iraqi positions…but we don't try to oversell our intelligence. If you were to see our classified maps of Iraq, you would see a lot of grey…not black or white, but grey."

> **With more transparency, NGA is uniquely positioned to play a leading role to advance public confidence in the Intelligence Community.**
>
> ROBERT CARDILLO, DIRECTOR OF THE NATIONAL GEOSPATIAL INTELLIGENCE AGENCY

He sees NGA "supporting [the State Department] more actively," in the future and is most excited about "areas of the world where we can advance democratic causes." Much of that advancement will be through the dispersal of images that everyone can access. That availability, in part, is what's pushing NGA to be more open.

It may be an uphill fight, or at least an unusual path. Following his conversation with reporters, Cardillo addressed a room of industry representatives and intelligence professionals, people we used to call spies. When he told the crowd "With more transparency, NGA is uniquely positioned to play a leading role to advance public confidence in the Intelligence Community," the room suddenly went quiet.

Defense One

A Kurdish security soldier is seen silhouetted as he walks.

HUSSEIN MALLA/AP

Unmasking 'Jihadi John' With Biometrics

As terrorists turn to online tools to get their message out, national-security professionals are looking for new data sources to reveal the identities of terror suspects.

BY PATRICK TUCKER

On Thursday, February 26, the *Washington Post* and BBC identified Mohammed Emwazi, a British-educated, Kuwaiti-born man in his mid-20s, as "Jihadi John," the Islamic State frontman who executed several hostages on camera to the world's horror.

"We will not comment on ongoing investigations and therefore are not in a position to confirm or deny the identity of this individual," the FBI said Thursday.

Denials aside, FBI director James Comey said months ago that they knew John's identity.

If the FBI has in fact identified Jihadi John, the victory was, in part, a product of the FBI's growing collaboration with the Department of Defense — a relationship that will grow much more cozy in the coming years, in the black cherry tree dotted hills of Clarksburg, West Virginia.

About four hours away from Washington, D.C., sits the headquarters of the FBI's Criminal Justice Information Services Division, or CJIS, which houses the bureau's Biometric Center of Excellence. The center is not a place so much as a program, begun in 2007, that plays a key role in making use of all the biometric data that comes into the FBI's possession. That's every fingerprint, every image, and every phone message that anyone sends to the FBI.

"Bottom line for us … if any of our divisions, whether it be our counterterrorism division, our criminal division, if at any time during their

investigations they develop biometrics … they submit it through our system," Stephen L. Morris, assistant director of the CJIS, told *Defense One* at a recent conference in Washington. In terms of identifying John, he said, "I'm not going to tell you how we did it," but added, "You have to have something to search … you can have images with faces but if you're not capturing it in the right way, if there's not data in that image to make a comparison, it's just not useful."

> **I can't speak enough about our relationship with the Department of Defense. After 9/11, our mission in life changed. It was all about national security, our partnership with DHS and DOD.**
>
> STEPHEN L. MORRIS, ASSISTANT DIRECTOR OF THE CJIS

This, in part, is why the biometric center plays a role in bringing parties, and their biometric databases, together. The FBI's system is called the Next Generation Identification, or NGI. It includes photos, aliases, physical characteristics and, of course, fingerprints. Today, it's completely interoperable with the military's Automated Biometric Identification System, or ABIS, and the Department of Homeland Security's Automated Biometric Identification System, or IDENT. The center also works with the State Department and allied law enforcement agencies around the world. The FBI and Britain's MI5 have been working together to identify John.

Obtaining a biometric record on a suspect to match against a terrorist video of a masked jihadi is not something done easily or robotically. It requires old school investigation, either sifting through lots of hours of collected video footage and comparing that to crime videos (such as beheadings), or going out into the field to find voice samples on suspects to match against crime videos, or both.

This is where the Defense Department's extensive library of biometric signatures, gathered on the field in places like Iraq and Afghanistan, can play a role in future investigations. The department's biometrically enabled watch list, or BEWL, houses more than 200,000 records.

"I can't speak enough about our relationship with the Department of Defense. After 9/11, our mission in life changed. It was all about national security, our partnership with DHS and DOD — to say it expanded is an understatement," Morris said at a recent biometrics conference in Washington, D.C."Their ABIS system was connected with our system, so they have a small group of folks who are out there [in West Virginia] in charge of their system. Having them co-locate with us has been very important."

That important relationship is about to get a lot more intimate. Later this year, the FBI is going to open a $328-million, 360,000-square foot Biometric Technology Center next to the current

CJIS campus. The Defense Department will get about 40,000 square feet in the building, which will also consolidate the FBI's biometric workers and operations. "Anything and everything we do will be run out of that building," said Morris.

In September of last year, the FBI announced that the $1.2 billion NGI system was fully operational (it was rolled out in increments over a period of years). If it works according to plan, it will provide law enforcement with a very fast and reliable sense of exactly who they are talking to, what threat that individual may pose, and what records they've left — fingerprints, voiceprints, etc. — in what places.

But fingerprints don't help you catch everyone. Voice recognition played a key role in the identification of Jihadi John, according to published reports. The FBI's biometric center site lists voice recognition as one of its key modalities, or areas of study, along with DNA and others, but fingerprints and more traditional biometric signatures make up a bulk of the records it manages.

Voice, in many ways, represents a crucial gap in biometrics collection for both the Defense Department and law enforcement. In a noisy environment it can be very hard to get a dataset to do matching against, a huge technical issue that the government is actively looking to solve.

In Iraq and Afghanistan, soldiers have compiled huge datasets of people that they have come across, including finger scans, pictures, and iris scans. Any of those can serve as a reliable red flag for a Turkish border guard, or, for that matter, a New York cop with a suspect in the chair. But they don't do much against a villain broadcasting terror from a safely fortified mountaintop in Syria.

Some of the men fighting with ISIS today have probably left their fingerprints in a few places where Western law enforcement could pick them up and share them. Technology, by itself, won't find those places. But, once the data is found, it can make positive identification much faster and easier, as it apparently has with ISIS's most infamous fiend.

Government buyers deserve more from a supplier.

EXPECT MORE

HEAD TO GRAINGER

Grainger is your one-stop shop to help streamline your purchases.

- > *Time-saving procurement channels*
- > *The products you need*
- > *Resources to help meet your small business goals*

Contact your local Account Manager, **call** 1.800.GOV.TEAM or **stop by** your local branch today!

GRAINGER.COM/FEDGOV | 1.800.GOV.TEAM

GRAINGER®
FOR THE ONES WHO GET IT DONE

About the Authors

PATRICK TUCKER

Patrick Tucker is technology editor for *Defense One*. He's also the author of *The Naked Future: What Happens in a World That Anticipates Your Every Move?* (Current, 2014). Previously, Tucker was deputy editor for *The Futurist* for nine years. Tucker has written about emerging technology in *Slate*, *The Sun*, *MIT Technology Review*, *Wilson Quarterly*, and elsewhere.

ALIYA STERNSTEIN

Aliya Sternstein reports on cybersecurity and homeland security systems for *Nextgov*. She's covered technology for over a decade at such publications as *National Journal's* Technology Daily, *Federal Computer Week* and *Forbes*. She's been a guest commentator on C-SPAN, MSNBC, WAMU and *Federal News Radio*.