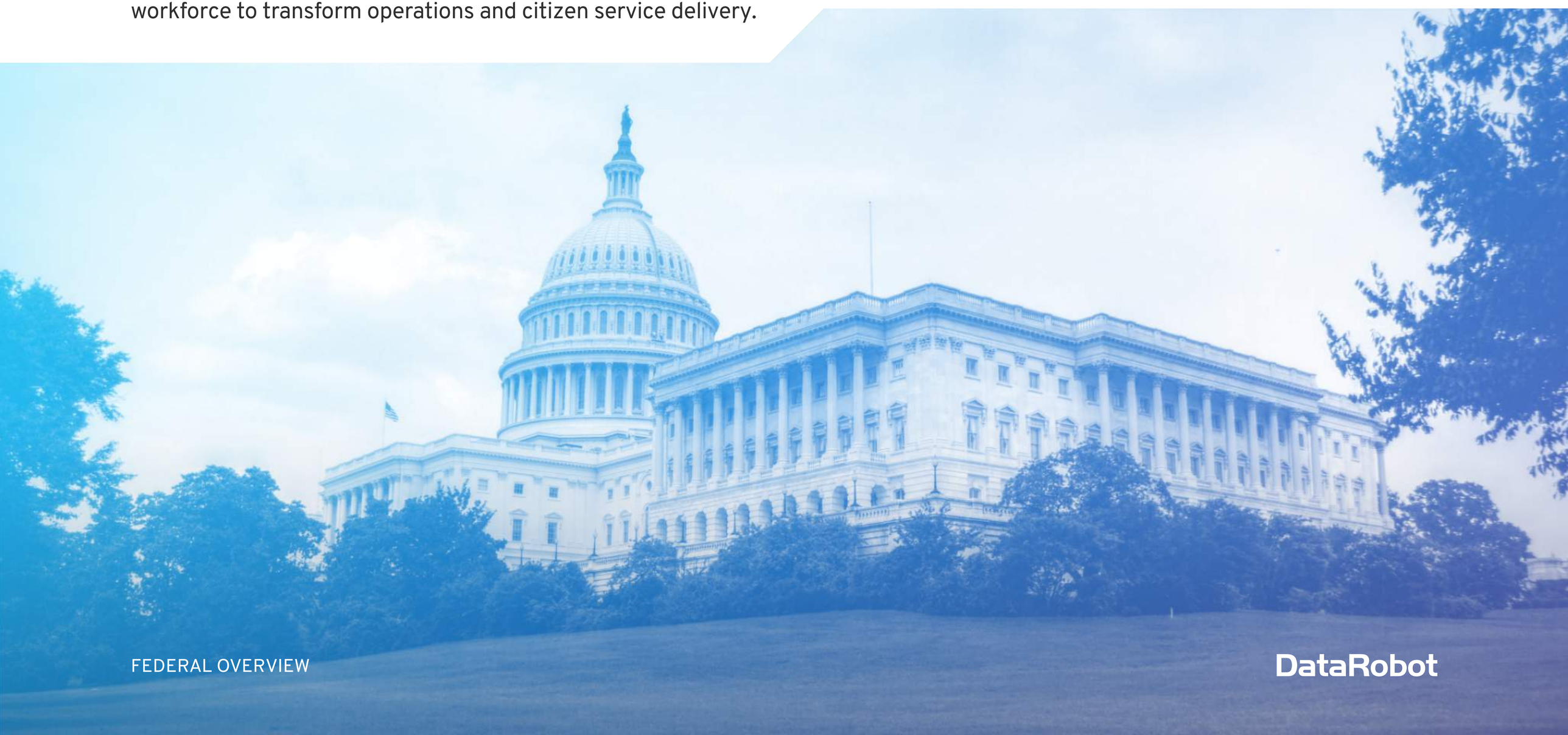
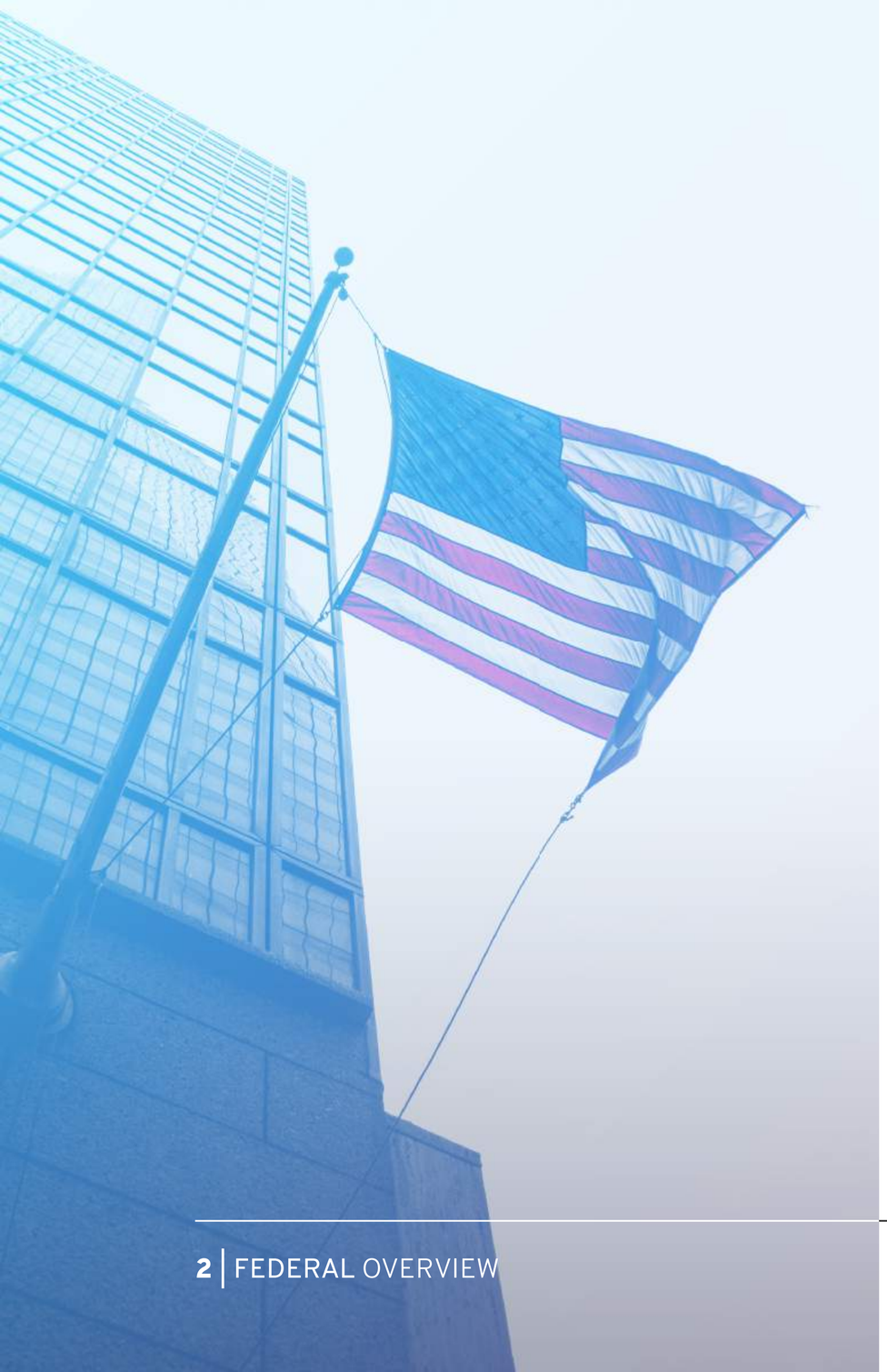
Unlocking Federal Mission Insights with Automated Machine Learning

Put the power of data science at the fingertips of your agency's workforce to transform operations and citizen service delivery.



FEDERAL OVERVIEW

DataRobot



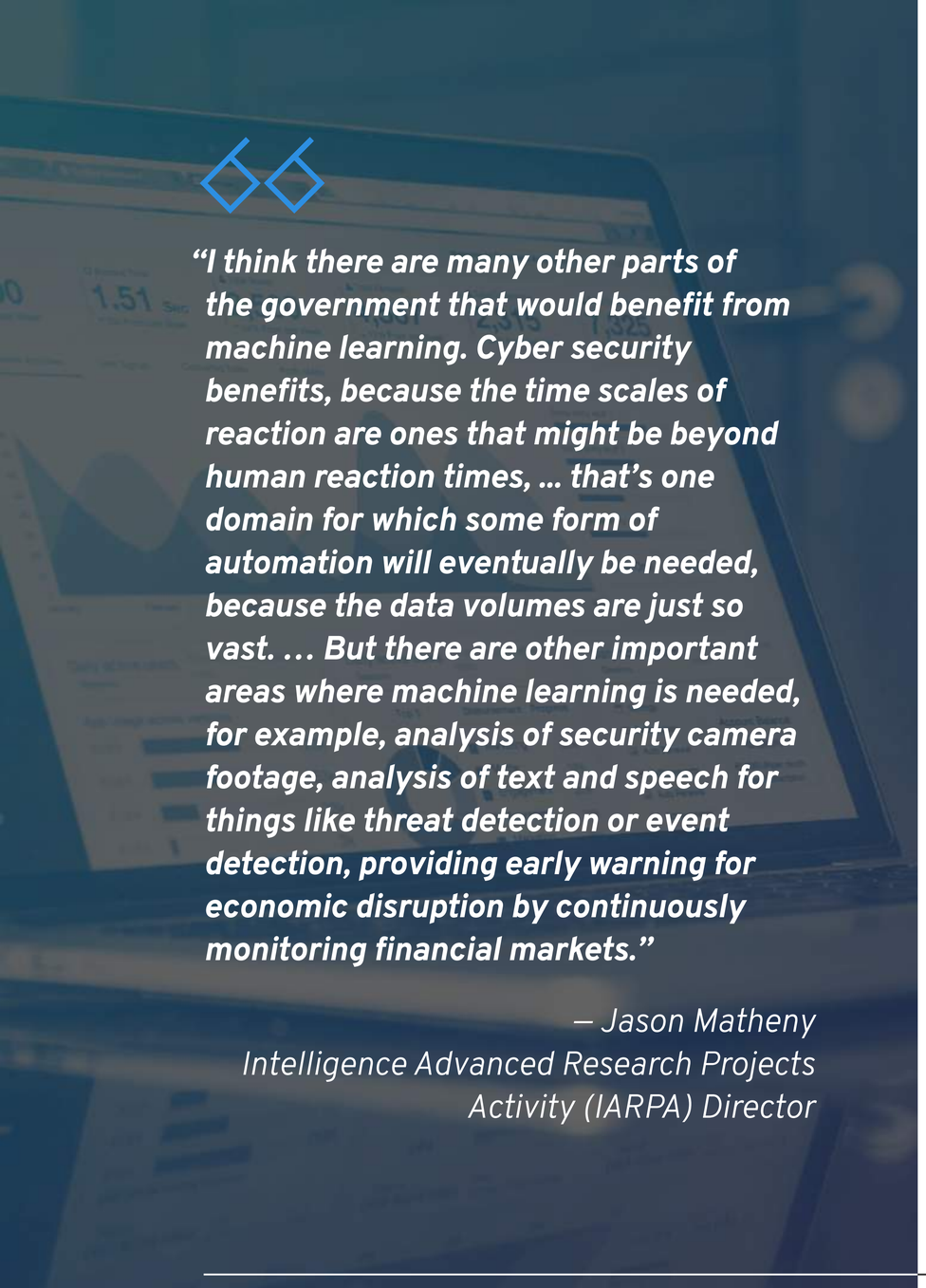
Introduction

By 2020, 1.7 megabytes of data will be produced by every person during every second of the day. The U.S. federal government is one of the top managers of this massive influx of data, which holds the key to uncovering insights for tackling some of society's biggest challenges, from cyber security to healthcare. The current data gap for agencies isn't simply analysis: its analysis that's fast and accurate enough to power decisions that deliver mission impact.

"Data drives decisions," according the U.S. Department of Commerce.

Historically, attracting talent has been a major challenge within the federal government. The data scientists that create the mathematical models capable of deriving insights from government data are in high demand and short supply. The customized models they build can take months or even years to create, which simply isn't fast enough when the goal is combatting terrorists or malicious hackers. Simply put, federal agencies need a better option.

Automated machine learning, a facet of artificial intelligence (AI) used for predictive analysis, can simplify the complex job of harnessing the power of government data. Not only does it accelerate discovery of the best predictive model, but it puts the power of creation in the hands of agency leaders, mission experts, and even citizens. We're entering the age of the "citizen data scientist" during which anyone with a great idea can build predictive models without writing a single line of code. This democratization of AI is already helping businesses gain unprecedented competitive advantage in the private sector, but federal agencies can use the same tools to dramatically improve mission outcomes.



“I think there are many other parts of the government that would benefit from machine learning. Cyber security benefits, because the time scales of reaction are ones that might be beyond human reaction times, … that’s one domain for which some form of automation will eventually be needed, because the data volumes are just so vast. … But there are other important areas where machine learning is needed, for example, analysis of security camera footage, analysis of text and speech for things like threat detection or event detection, providing early warning for economic disruption by continuously monitoring financial markets.”

— Jason Matheny
Intelligence Advanced Research Projects Activity (IARPA) Director

What is automated machine learning?

Machine learning uses historical datasets to develop predictive models that empower organizations to make accurate, forward-looking decisions.

In the past, these models were manually created, validated, deployed, and tested by data scientists, all of which took place over a long period of time. However, automation changed the game when it entered into the equation: automated machine learning has the power to build, test, and rank hundreds of different models against an organization’s existing data. The result is a “survival of the fittest” scenario in which the best model for that particular dataset rises to the top and model selection and deployment happens in days or weeks rather than months.

In government agencies that are lucky enough to have the resources to find and retain their own data science team, this type of automation empowers their scientists to spend less time on manual model creation and more time on strategic analysis for mission impact. It doesn’t eliminate the need for data scientists, just as surgical robots don’t replace surgeons — they make it possible for surgeons to focus on more productive and demanding work.

For the majority of agencies that lack a full data science team, automation democratizes access to the power of machine learning, allowing agency personnel and true mission experts to be intimately involved in data-driven decision-making like never before. Automated machine learning puts predictive intelligence in the hands of the talent driving the mission forward and reduces the need to hire and retain hard-to-find data scientists.

“Right now, the most common commercially used methods for detecting cyberattacks look for signatures for particular pieces of malware. They are often detected months after they have invested in a system. The goal is to predict attacks days to weeks before they occur. I think we’ve seen an example of this in DARPA’s Cyber Grand Challenge, in which there was a competition between machine learning systems on the offensive side. It takes the human out of the loop, on both offense and defense in cybersecurity, so you’ve got systems that are automatically identifying and exploiting vulnerabilities.”

— Jason Matheny
IARPA Director

Federal Use Cases

Cybersecurity:

Protecting the nation’s data and information infrastructure is a daunting challenge, with national security and citizens’ personal information such as social security numbers constantly at risk. As National Security Agency (NSA) Director Adm. Michael Rogers told the Senate Armed Services Committee, relying on human analysis for cybersecurity “is a losing strategy,” and machine learning is necessary to allow security agencies operate on a cyber scale and at cyber-speeds. Predictive models analyze network and user activity, learn the indicators of attacks and compromises, and predict when an attack is likely at a speed that is impossible for human analysts.

In the current environment, security analysts are overwhelmed by the sheer volume of data generated by network sensors and security tools, and hackers are adept at hiding their tracks and avoiding the security radar. As a result, the time between a breach and its discovery is often measured in weeks or even months. Automated machine learning is an essential tool to enable federal enterprises to stay a step ahead of threats — whether they are insiders, criminal organizations, or foreign nations — by reducing the time necessary to create models capable of identifying activities that indicate malicious intent and flagging them with greater speed and accuracy than was previously possible.

Automated machine learning makes it possible for agencies of any size to take advantage of predictive analytics to protect their own enterprises, efficiently develop and deploy models to flag suspect devices or activities, identify dangerous behavior before a breach, and proactively address threats.

Federal Use Cases

Fighting Fraud:
The federal government operates some of the largest financial enterprises in the world and must safeguard billions of dollars of taxpayer money being collected and paid out each year. The Internal Revenue Service (IRS), for instance, receives almost 240 million tax returns each year and has 3,000 employees working to identity fraud issues. Despite these efforts, the cost of fraudulent tax refunds in 2016 was $21 billion. The Centers for Medicare and Medicaid Services (CMS) currently contracts with companies to randomly evaluate claims, yet CMS estimates that it paid out more than $38 billion in fraudulent payments in 2016.

With the vast amount of financial resources on the line, the federal government is a constant target for criminals and must continually work to detect fraud. However, because of the scale of transactions, traditional statistical models and human analysis are inadequate to effectively guard against fraud.

Using automated machine learning, agencies can quickly build models to better understand normal business activities, identify unusual behavior or trends, and predict where fraud or error is most likely to occur. Rather than relying on outdated statistical models or random audits, predictive models can be applied against the large volumes of data generated by transactions to quickly flag problem areas, allowing agencies to be proactive in addressing fraud and waste.

The automation of this process won't replace human auditors and investigators, but will instead free them up to concentrate on complex cases flagged by the model that demand their attention rather than chasing false positives, helping agencies maximize their return on manpower investments and solve problems more efficiently.

Counterterrorism:
According to the Government Accountability Office, there have been 225 fatalities linked to terrorist attacks in the U.S. since 2001. The intelligence community now employs thousands of analysts to identify threats using tools such as link analysis to examine connections between persons of interest, but the process is still largely manual and does not make effective use of analysts' time and skills or the vast volumes of information available.

As technologies such as image recognition and natural language processing improve, users are taking advantage of new predictive models that search for early threat indicators in online communities used by terrorists. This presents a fresh opportunity for the federal government to stay one step ahead of terrorist activity; automated machine learning can help make better use of their huge volumes of data and limited manpower by developing, continually testing, and improving predictive models to identify risk patterns while flagging risky activities for further study by analysts.

> ***"In security camera footage, you'd like to be able to detect when someone is dropping off a bag and walking away or loitering in regular intervals. The goal is to look for that in video. It's impractical to have human eyeballs on every camera."***
>
> *— Jason Matheny*
> *IARPA Director*

Federal Use Cases

Preventing Insider Threats:
In a recent survey, federal IT officials ranked careless insiders as the greatest threat to government cybersecurity, with malicious insiders following close behind. Whether intentional or not, any type of insider threat poses a serious problem for federal agencies. Many times the warning signs for a potential insider breach go undetected until valuable information is already lost. Current methods of background checks and monitoring insider activity inundate agencies with false positives and lead to missed chances to stop insider threats before they happen.

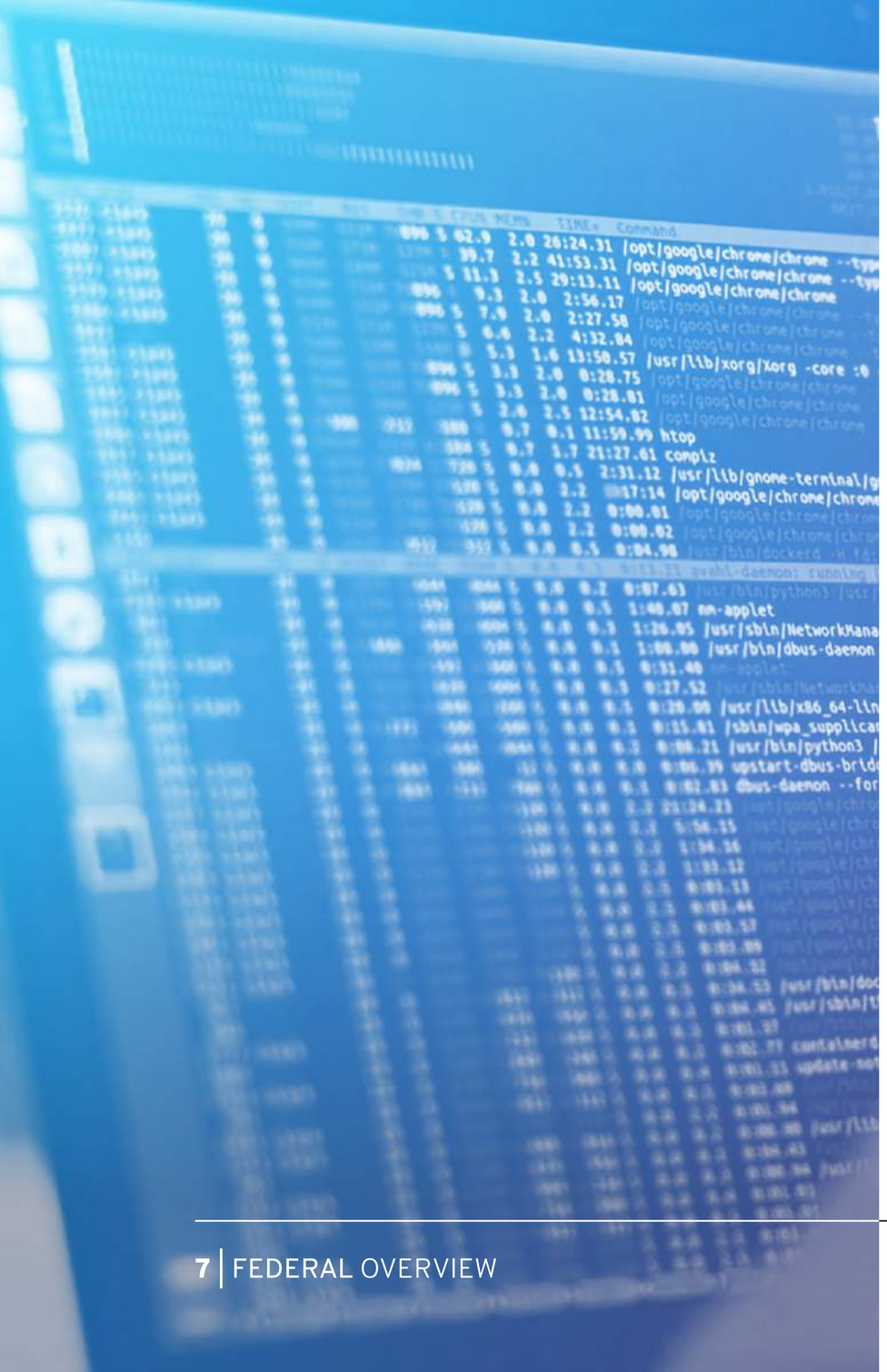
With automated machine learning, multiple models are quickly created and deployed to evaluate insider behavior and identify possible precursors to malicious or risky behavior that might be overlooked in a static evaluation. The highly predictive models create warning flags that are relevant and urgent and pass them to analysts for more complete investigation, who can then devote more time and attention to these more critical cases.

Logistics:
The federal government is responsible for making sure that huge amounts of food, supplies, equipment, and people are available where and when they are needed for national defense and for emergency response.

The Federal Emergency Management Agency (FEMA) is a first responder for natural disasters and other emergencies. FEMA must ensure that food, water, clothing, housing, and the materials and equipment necessary to restore infrastructure are available when disaster strikes. This means not only understanding what will be needed, but having the requisite resources prepositioned before disasters for efficient deployment. Similarly, the Department of Defense (DOD) must also acquire, distribute and deliver huge amounts of materiel to U.S. troops. The DOD's Defense Logistics Agency acquires and supplies more than $34 billion in goods and services each year to the Army, Marine Corps, Navy, Air Force and Coast Guard through nine global supply chains.

Predictive models developed quickly with automated machine learning help agencies like FEMA and the DOD predict the manpower and materiel that will be needed under specific conditions and determine the best staging positions and ways to ensure availability.

Automated machine learning puts predictive modeling capabilities into the hands of more people in any organization, large or small, helping them use available data to make the best use of physical resources and freeing up their manpower to focus on the problems that make the most impact.



DataRobot for Federal

We believe in practical solutions, which includes helping your agency team gain access to the technology, education, and when necessary, the professional services needed to move beyond the hype, embrace AI, and turn data science into mission success.

Automated machine learning, created by DataRobot, brings predictive decision-making to every level of your agency's workforce, driving new levels of operational efficiency and transforming citizen service delivery. DataRobot's platform offers a fast, easy, and scalable way for federal agencies to embrace the benefits of AI and harness the power of their data to maximize mission impact.

DataRobot offers a user-friendly drag-and-drop interface that allows agency personnel to quickly test and deploy models without having to learn programming languages. It provides the tools for any employee at any agency to become a citizen data scientist. DataRobot's distributed systems enable mission owners to score large amounts of data quickly, dramatically reducing the risks, time, and cost associated with deploying machine learning models. Additionally, DataRobot offers the ability to monitor, maintain, and refresh models as new data sources become available or conditions in the world change.

Not only does DataRobot provide the platform, but the company is committed to the successful use of that platform through DataRobot University and AI Services. Educating everyone from mission leaders to business and data analysts, DataRobot University includes courses at every level to ensure all agency personnel can become trained citizen data scientists. And when required, DataRobot AI Services is available to jumpstart project success and enable an AI-driven agency by working with the existing teams of onsite contractors for the most successful implementation in the shortest amount of time.

DataRobot for Federal

As a growing percentage of the federal workforce reaches retirement age, automated solutions are required to fill the gaps and be true force-multipliers. DataRobot's easy-to-use platform allows agencies to build their capacity to solve problems and meet mission objectives with the people and data they already have today.

Learn more by contacting DataRobot for a demo at https://www.datarobot.com/public-sector/.

Contact Us

DataRobot
One International Place, Fifth Floor
Boston, MA 02110
www.datarobot.com | info@datarobot.com

DataRobot