The Power of Data in Improving Cybersecurity

leidos

Data is the most strategic asset government agencies have today, and finding ways to effectively analyze and combine it with enabling technologies is the best way to perform missions. However, this needs to be done in a way that secures the data that is used to make decisions.

Federal leaders today understand the importance of data and the need to protect it. The President's Management Agenda, for example, puts data front and center, identifying it as one of the major drivers of government transformation, which includes improving cybersecurity. The Office of Management and Budget (OMB) does the same in its Federal Data Strategy, which promotes the use of data as a strategic asset.

Using data to improve cybersecurity effectively, however, requires aggregating it and using modern analytics, artificial intelligence and machine learning tools to make sense of it.

"Just because we have a lot of data does not mean the volume enables analysis. Often, data is redundant, or relates to the same entity in the environment from different sources. For example, consider taking pictures of an object from different angles, cutting those pictures all up into tiny pieces, and asking someone else to use these pieces to just identify the original object. While all the data needed to do so is technically there, it would be very difficult to perform this task." explains Olya Flores, Director of Data Transformation at Leidos.

That's the situation many agencies are in when it comes to integrating cyber information. By using higher-level tools like artificial intelligence, machine learning and advanced analytics, agencies stand to gain a lot.

If you can use these tools to detect repetitive patterns, suggest next actions, and automate simple reasoning, you'll increase efficiency and free up analysts to focus on more complicated threats. More importantly, you will improve and accelerate higher-level decision-making.

GETTING MORE INSIGHTS FROM DATA

Agencies today use advanced analytics in many ways, from improving services to rooting out fraud. According to one report, more than half of federal IT respondents said that data analytics platforms empower employees to make data-driven decisions, and about one-third said that analytics tools helped them discover a previously unknown data point.

Increasingly, agencies are using data analytics to improve cybersecurity. With the right tools applied to the right datasets, agencies can identify abnormal behavior patterns, unusual transactions, and all types of malicious behavior inside agency networks. Agencies also can use predictive analytics to analyze aggregated data from multiple sources to predict likely future attacks and events.

Squeezing even more insight out of data requires taking it to another level with artificial intelligence and machine learning. The goal of AI is to solve complex problems and make better decisions, while the goal of machine learning is to increase accuracy and learn new things from data. They are both critical tools in combating cybercrime.

As an example, Leidos is currently using these techniques to improve the accuracy and timeliness of cyberattack forecasts for IARPA's Cyberattack Automated Unconventional Sensor Environment (CAUSE) program.

leidos

The process involves gathering and analyzing a variety of data sources, including unconventional sources like tweets, to find attacks at their earliest stages so they can be halted.

While many agencies today are fairly comfortable with data analytics, some are ramping up more slowly with AI and machine learning. Their use is growing quickly, however, not only because of the President's Executive Order on AI and efforts like the Pentagon's Joint Artificial Intelligence Center, but because these technologies can make a huge difference in pinpointing relevant data, identifying valuable areas for further analysis, and improving effective real-time decision-making. According to one recent report, nearly half of federal agencies have experimented with AI and machine learning tools and techniques and have seen tangible pay-offs.

HARNESSING THE POWER OF ADVANCED TECHNOLOGY

While all of these tools can help agencies improve cybersecurity in many ways, the real magic happens when they are used together. Combined, these technologies can correlate actions that might seem commonplace when examined separately and draw more accurate conclusions.

These technologies can be used together in many other useful ways as well. Leidos, for example, uses anomaly detection and analytics to identify unusual activity either inside or outside the perimeter based on users, time of day and previous activity, and correlate that activity with other risk factors to prioritize investigations.

The key to harnessing all of these powerful technologies is having the right frameworks, repeatable processes and agnostic architecture. With those in place, it will be much easier to adapt to new technologies, requirements and threats.

"If tomorrow, a new tool emerges and you want to swap it into your environment, it's important to be able to do that without having to migrate all the data between them," Flores says. "With the data pipeline in place, many of these tools become just views onto your data, and we are able to use them for their function, not for their storage or schemata."

Leidos has spent years creating the frameworks and methodologies to make analytics, artificial intelligence and machine learning as effective as possible, and is implementing them with its federal customers.

To address the classic problem of identifying and collecting all important assets in one pipeline, Leidos has developed the Foundational Automation Support Technology (FAST) framework. It continuously discovers and tracks network devices, creating an asset inventory. This approach incorporates AI and machine learning to understand event relationships and perform predictive analytics between events.

One of the classic difficulties with effective data analytics is making sure you have gathered all of the right data sets so you can ask questions of the data and get the results you need. To improve this process, Leidos is developing the platform, which aggregates data from devices and directs the data feeds into an analytics platform for processing.

To accomplish this, Flores' team at Leidos developed Data 2 Intelligence data pipeline. The pipeline provides data ingest and processing from various sources, and makes it possible to use a common, entity-based data model across all of the sources. Analyses can be run on data in streaming, and analysis results can be shared to all of the tools in the enterprise in real time, boosting performance of the entire environment. True to its open architecture, the pipeline integrates both with any external tool that has an API, as well as with other Leidos capabilities such as FAST.

The final step in the process is analyzing the data about the devices to characterize incidents and threats. Leidos has developed the Cyber Adversarial Reasoning Demonstration (CARD) to bring this all together. CARD is a flexible framework that can work with any and all other tools.

"One of the biggest challenges is amassing data when you are following a particular case or incident, and this pulls all of the relevant data together," explains Meghan Good, a cyber solution lead at Leidos working on the CARD project. "CARD helps us determine what an ideal analysis system would look like for a threat analyst, and what types of analytics and processing we should have in place in a prototype to fit that ideal."

leidos

Good's team evaluated the CARD prototype in a head-to-head challenge against a team of three trained analysts. To review the same data (network traffic, logs, alerts), from an incident using relevant tools and search capabilities, the team spent three hours identifying the malicious cyber activity and its evidence in the data. In a fraction of that time – less than 15 minutes total – the CARD prototype ingested the data and ran it through additional security sensors, identified related malicious activity from across evidence sources, characterized the threat using several Leidos-developed ML models, and visualized it for an analyst. This resulted in a 12x improvement of team time, and a 36x improvement in people-hours. When paired with data processing pipelines and frameworks like D2I and FAST, Good projects that significant improvements are within reach.

USING DATA, AI AND MACHINE LEARNING FOR CYBERSECURITY

By analyzing data, especially with the help of artificial intelligence and/or machine learning, agencies can better predict and prevent events, enhance decision-making, dramatically reduce the time it takes to detect when breaches are occurring, and predict when and how they will occur.

While pursuing these goals, Flores urges agencies to focus much more energy and resources on the data, the methodologies and the architecture, and less on the tools.

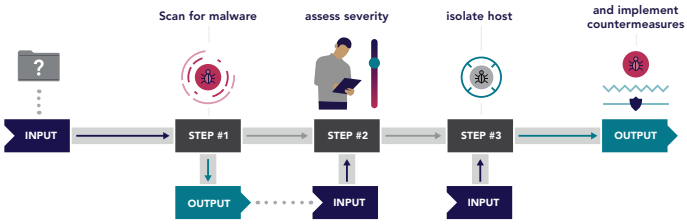
"It's critical to keep your architecture extensible and focus on the data instead of the tools, because the current landscape is evolving very rapidly," she says. "Think of the tools as something you can plug into and out of your data pipeline. That will make sure you can support new use cases and handle new threats over time."

ARTIFICIAL INTELLIGENCE (AI) & MACHINE LEARNING

AI & ML are predominantly built on models or algorithms trained with data. These algorithms are often optimized to take an input and provide a response, typically in the form of a score. AI & ML mimic human decision making but do not typically automate entire processes.

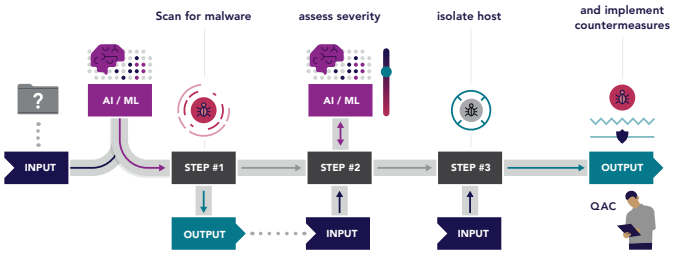
AUTOMATION & ORCHESTRATION

Automation & orchestration expedite processes by converting manual actions into series of coded steps. It can reduce human resource requirements, shorten time to completion, and increase process accuracy. Automation does not mimic human decision making behavior but does follow pre-defined logic.



AI & ML + AUTOMATION & ORCHESTRATION

Combining AI & ML with automation & orchestration provides organizations the ability to create complex processes that incorporate models and algorithms as well as business logic.



FOR MORE INFORMATION

leidos.com/on-a-mission