Deloitte.

Applications of AI for cybersecurity defense

By David Caswell, Eric Dull, Jennifer Vitalbo, Keith Watts, and Jacqueline Schultz

April 2024

In an ever-evolving technology landscape, cybersecurity leaders are grappling with how best to understand and apply Artificial Intelligence (AI). **Understanding what tools use AI and how the application of it can impact your cybersecurity environment is more convoluted than ever due to the exponential growth of the vendor and solution marketplace.**

Additionally, the developing AI regulatory landscape will require both the public and private sector to **develop and account for new industry standards and guidelines.**

This is evidenced by the October 2023 AI Executive Order (EO) signed by the Biden Administration, which prioritizes the safe and secure development and use of AI as one of its primary directives and emphasizes the critical role of cybersecurity in the country's overarching national security.[1]

While AI and the emergence of Generative AI (GenAI) introduce new security risks that must be accounted for, they also have enormous potential to support a variety of cybersecurity use cases. To harness this potential, leaders and technologists need to sort the real from the hype, which requires a basic understanding of where AI can aid a cybersecurity environment, the types of data to feed the models, and a roadmap for how to bring it together.

Cybersecurity without AI

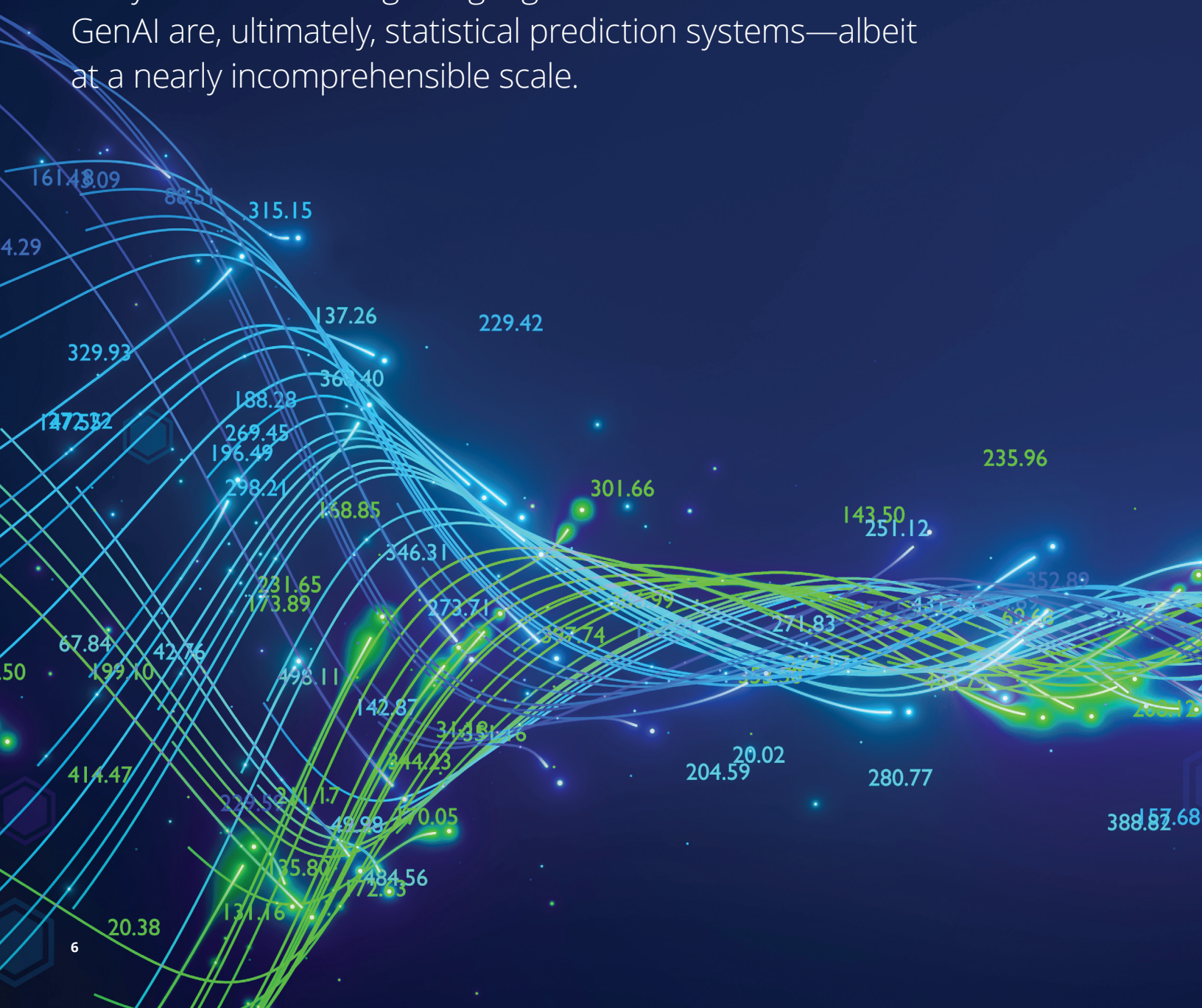
While AI is exciting and does provide value for cyber operations, **a bulk of cybersecurity detections already generally rely on deterministic signatures[2], using if-then statements and classic statistical filters.**

These systems are relatively easy to load and can process enormous quantities of data quickly and accurately, resulting in detectors that are both inexpensive and often highly transferrable. Unfortunately, signatures present several challenges. They only catch indicators of compromise (IOCs) or vulnerabilities that have been seen and encoded previously. They are further limited as cybersecurity operators traditionally apply them against single data streams. For instance, network feeds typically only match where the data is consistent between network environments. They are also only viable for matching against a short time window versus a longer time analysis that more complex models can support. These signatures rely on IOCs that are easy for adversaries to change. However, AI allows cybersecurity teams to broaden detection across sources and identify both known and unknown threats.

Statistical analysis, particularly streaming analytics, goes beyond signatures to account for adaptations within environments to determine anomalous behavior. For example, streaming statistics can calculate various measures for network traffic performance and trigger alerts should the statistical measure move outside of determined control limits. These types of models naturally adapt to their baselines over time to allow for changes in the system but are typically limited to monitoring a single data stream.

How AI can support cyber security

AI models, specifically machine learning models, in many ways can be thought of as universal, self-tuning statistical analyses. Even the large language models behind much of GenAI are, ultimately, statistical prediction systems—albeit at a nearly incomprehensible scale.



Machine learning goes beyond traditional statistical analysis methods by leveraging information across a variety of data sources, large data pools, and/or over long periods of time and can detect indicators that span various systems. This ability to create connections across various large data sources coupled with the capacity to process data at incredible speeds allows AI and related models to provide more robust support across the cybersecurity life cycle.

Signatures and statistical analysis provide two predominant approaches to validating user identity. AI models are functionally similar to signatures and statistical analysis by identifying patterns or indicators from various data sources. However, traditional AI provides enhanced capabilities in cybersecurity operations.

These capabilities can generally be binned into one of the following six features or applications:

Environmental sensing—Identifying objects or identities that exist in a cyber environment

Threat intelligence—Identifying malicious actors and their activities

Vulnerability/threat detection—Detecting malicious executables or activity within the environment

Attack path discovery—Determining the sequence of actions taken by an attack

Behavioral analysis and anomaly detection—Determining what normal behavior looks like in the environment and identifying when a set of activities are outside of the norm

Cyber security triage—Determining the priority of response actions for alerts

The importance of cyber security triage cannot be overstated and may be particularly enhanced with the emerging advances of GenAI. GenAI provides security operations professionals with tools to naturally interact with the environment through assessment, threat summarization, and automated cogeneration to help reduce the barrier of entry. **The first five bins play a critical role in working together to create a broad defensive fabric culminating to the final bin, cyber security triage.**

Environment sensing

The first step of cyber defense is to know who and what is supposed to be in the environment. Building and maintaining an inventory of the assets and users tends to be a continual challenge for most organizations. Using network scans, traffic logs, system telemetry, and external scan data (such as direct scans and global IP traffic monitoring) paired with a combination of signature and AI analytics can enable a detailed evaluation of both the network architecture as well as insights into asset versioning and attack surface.

Threat intelligence

Where environmental sensing is internally focused, threat intelligence looks externally to identify adversarial activities that are targeting an organization. This information provides insights on the types of attack vectors being employed so that additional security controls can be implemented to prevent intrusion. Further, open-source intelligence (OSINT), the basis for threat intelligence, should be monitored for any signs of existing data breaches in an organization. This monitoring should include looking for signs of credential theft and any accesses or vulnerabilities being sold that would put employees or organizations at risk. AI can support the analysis of OSINT data streams by combining with other external data feeds such as IP traffic monitoring data. This allows for both direct data extraction and monitoring for behavioral changes of malicious actors such as increased activity, changes in attack methods, or shifts in who they are targeting.

Vulnerability detection

This is the foundation of cybersecurity defense that detects discrete malicious activities. This type of detection is a natural application for AI. The AI models can be trained against single sources of data, such as email for spam/phishing or against multiple data sources, such as network traffic and system logs. The nature of the AI models should be updatable so they can continually adapt to emerging threats.

Attack path discovery

This method builds on vulnerability detection to assess the risk of any given sequence of activities. While many of the actions may independently be considered benign, attack path discovery methods are able to determine when a combination of actions should be considered a threat. This approach combines the full suite of internal and external data sources to correlate activity across the environment.

Behavioral analysis

If attack path discovery detects known attack methods, behavioral analysis provides alerting for unknown activities. This approach uses the full complement of data sources; but unlike attack path, it needs to learn what is considered nominally acceptable behavior for the specific network it is deployed within. Once learned, it is then able to flag activity that it determines as anomalous. What's more, this type of system needs to be adaptable so that it can update what it considers acceptable behavior given changes to the network. The need to be trained on individual environments, be updated over time, and able to ingest data across multiple systems makes this a particularly compelling application of AI for cybersecurity. Of note, behavioral analysis is also useful for detecting insider threat risks such as data theft.

Cyber security triage

All these indicators and warnings eventually converge in the security operations center for triage and remediation. The process of prioritizing which alerts should be actioned is, itself, another potential application of AI. Feeding this system is the collection of alerts but also external vulnerability data libraries for threat rankings and OSINT data for threat intelligence.

Types of data

Any application of AI is directly reliant on the availability, type, and quality of data sources that can be used in the model. Depending on the use, cybersecurity AI can leverage a combination of data sources from within internal networks as well as externally-procured data elements. Internal data sources include any telemetry, traffic, or logs that can be collected from network assets.

This list includes telemetry from both on-premise and cloud-hosted resources, like information technology (IT), internet of things (IoT), and operational technology (OT), as well as client devices (system and application logs). Like traditional AI, GenAI relies on being fed data. However, GenAI differs from traditional AI in that it outputs new data. GenAI takes the use of data to the next level by identifying patterns and drawing correlations to mimic a human's ability to create content and provide analysis with data-driven recommendations.

External data can be defined loosely as data that is accessible by an external third party. This includes data that may originate from an organization itself when the data is either acquired or is externally visible to external network or OSINT scans for organizational credentials. External data could include generalized information that happens to be applicable to cybersecurity such as malware behavioral analysis or third-party network data used to pre-train a model. Table 1 lists some common internal and external data sources that are viable for AI applications.

Internal data

- Network traffic (e.g., connections, http, domain name systems)
- System logs and telemetry (e.g., on-prem, cloud, IT/OT)
- Security device and application logs and alerts
- Service application telemetry
- Endpoint telemetry
- Identity and access logs

External data

- Malware analysis data
- External Network scans
- OSINT such as financials, dark web, satellite imagery, government records, and more
- Ad-tech information
- Tier 1 IP network traffic
- Third-party network traffic

TABLE 1. Internal and external data sources used for cybersecurity AI applications

To solve challenges related to classification, analysts can use an AI model to evaluate whether a sufficient volume of labeled data exists to properly train and evaluate a new model. Labeled data is the set of data that has already been categorized based on the categories to be evaluated (e.g., a set of emails that have been marked as spam or not-spam). In some cases, such as computer vision models, there are volumes of non-specific training data or even pre-trained models that reduce the need for any proprietary labeled data. However, if the AI is intended to be tailored to a specific internal data enclave, then analysts will need to have some method of gathering the labeled training data.

Getting started

To support projects and organizations' security practices, AI can be integrated throughout the cybersecurity and data lifecycles with the following process:

1 

Prioritize the critical security information needs: It is critical to consider the business and mission case for AI before applying it in an organization's environment. Organizations have any number of security questions they are trying to monitor while operating with a constrained budget. This drives the need for prioritizing development efforts against those questions that are most impactful to supporting the mission or business goals. This determination allows focus on downstream efforts and to enhance the impact of the AI development workstream.

2 

Assess the data environment: To capture value from AI for cybersecurity requires accessing and making sense of the available data in the environment. Organizations need to understand policies and procedures that govern that data, along with the type of data they have access to and where it is captured and stored. This can help determine the safeguards or controls that are in place to ensure data's quality, security, and privacy.

3 

Identify data processing platforms: Now that the target problem is identified and the available data is understood, organizations need to identify the mechanism for how to accomplish the business or mission objectives and agency needs. The available data should be evaluated either through a common platform or a well architected set of platforms that are able to process data into actionable, collated, information. These information pipelines should be engineered to synthesize the data to ease analysis and accelerate anomaly detections and data-driven decisions. It is important for the platform to be able to present and interact with information in a clear way to understand the analysis and report on any anomalies, similar to a security information and event management (SIEM) tool.

Security leaders should have a firm understanding of the most concerning security risks for their business to tailor AI and analytics to support their needs. For instance, different organizations will place different levels of importance on questions related to cybersecurity vulnerabilities, insider threat risks, visibility concerns, or compliance. Prioritizing the risk aids security operations teams to activate the right suite of data and models to focus their resources to meet the most critical areas of their organization's cybersecurity posture.

4

5

Evaluate the workforce/ organization: An organization's workforce is a critical piece to the AI equation. Since AI is a tool for organizations to use, their people and skillsets are an integral part of the full implementation and should be aware of the varying aspects of a data literacy program through engagement, development, and enablement. Practitioners' skillsets—like data understanding, use, and decision-making—should align closely with an organization's data literacy program.

These skills should be adaptable and allow for practitioners to upskill as needed. It will be necessary to communicate with teams about data literacy programs based on their role. Additionally, organizations should consider the security aspect and train practitioners on how to use any implemented models to enhance their organization's cybersecurity. This can involve anything from building or deploying AI models, to using it for incident response and maintaining and organization's privacy, to adapting visualizations.

Execute: There are four key steps when executing AI within an organization's cybersecurity operations that build on each other. These are integration, implementation, monitoring, and reporting. Once an organization's data environment is successfully assessed, the appropriate mission applications are identified, and people are resources are brought up to speed, AI should be integrated across the organization. It is important that leaders and decision-makers are aligned on the goals and needs, and that data access is provided to the appropriate stakeholders. Stakeholders must have access to data to implement AI within a cybersecurity organization so they can analyze, apply, model, visualize, and report on trends and analysis.

It is important to highlight that a key component of security is separation of duties and the principle of least privilege so that the data is controlled and safeguarded throughout its lifecycle. Finally, an organization should monitor and report on the data analysis through visualizations, reporting, modeling, and predictions. This reporting is important to communicate analysis to technology and business leaders so that they can make data-driven decisions and continuously improve cyber models. The execution stage should be considered a continuous loop and an organization's implementation must be constantly improved on and decision-making should be revisited with process and data improvements.

Organizations have never had more data—or more cybersecurity attacks and vulnerabilities. **While the emergence of AI and GenAI may add an additional layer of complexity to this increasingly intricate landscape, they can also provide organizations with powerful new security tools.** AI enables valuable and scalable lessons learned by the organization's cybersecurity experts, automating key functions and enabling them to work more effectively and efficiently. By leveraging AI to find patterns and insights in the volumes of internal and external data available, organizations can improve the security of their systems. To realize this value while reducing risk, AI should be thoughtfully introduced into an organization through a broad, iterative approach that is centered around the applicable mission and security needs of that specific agency.

Reach out for a conversation.

David Caswell
Advisory Managing Director
Deloitte & Touche LLP
dcaswell@deloitte.com

Eric Dull
Advisory Managing Director
Deloitte & Touche LLP
edull@deloitte.com

Jennifer Vitalbo
Advisory Managing Director
Deloitte & Touche LLP
jvitalbo@deloitte.com

Keith Watts
Specialist Leader
Deloitte Consulting LLP
kewatts@deloitte.com

Jacqueline Schultz
Jacqueline Schultz
Advisory Consultant
Deloitte & Touche LLP
jaschultz@deloitte.com

Endnotes

1 Joseph R. Biden Jr., "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," The White House, October 30, 2023.

2 Industrial Control Systems Cyber Emergency Response Team, "Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies," US Department of Homeland Security, September 2016.

Deloitte.