NOZOMI NETWORKS

EBOOK

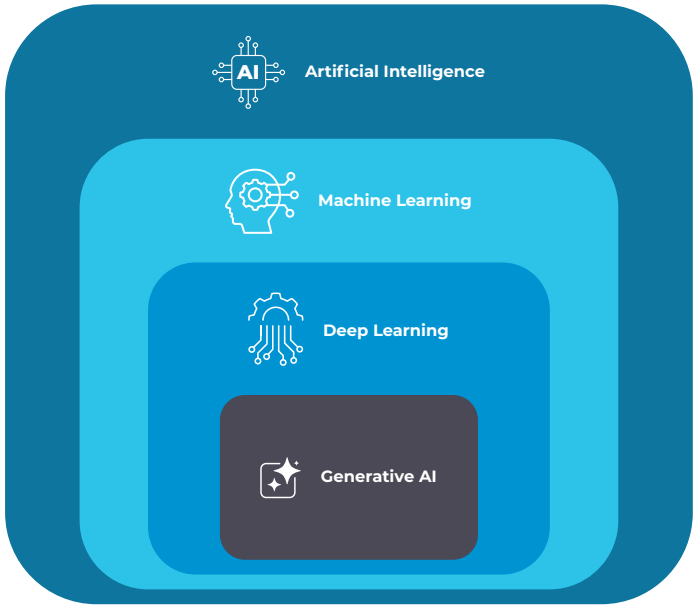
Winning the AI-Powered Cyberwar:

How to Evaluate AI-driven Security Solutions for OT/ICS and IoT Environments

Table of Contents

1. Introduction

Work. Play. Learn. Create. Artificial intelligence (AI) is reshaping every aspect of public and private life faster than we can assimilate. The field has been evolving for decades, but widespread access to generative AI tools like ChatGPT, Gemini and Copilot have put a narrow subset of AI at everyone's fingertips. That accessibility has sparked intense debate over whether the potential of AI in its many forms to improve lives outweighs the risks of misuse and unintended consequences. Cybersecurity is a shining example of how AI and machine learning (ML) methodologies can be leveraged for both good and evil — and be an attractive target itself.



The evolution of AI in general and in cyber offense and defense in particular is happening too fast to pin down. Any attempt to do so would be instantly obsolete. This eBook will take a brief look at how AI is being used to power cyberattacks, including attacks on AI systems and models, and how defenders are harnessing it to protect their environments. We'll then focus on what we know best: how the Nozomi Networks platform uses a range of AI methodologies to protect some of the world's most complex industrial and critical infrastructure organizations — and how we safeguard our AI analytics, model, agent and computational capabilities from attack.

Based on our experience refining our AI engine for more than a decade, we'll conclude with guidance on how to evaluate AI capabilities in cybersecurity tools charged with protecting IT, operational technology (OT), and Internet of Things (IoT) networks and devices.

2. The AI in Cybersecurity Landscape

How Attackers Are Weaponizing AI

Bad actors have seen the future, and it is AI. They may be using the same phishing, distributed denials of service (DDoS) and other techniques as before, but with greater speed, scale and efficiency. Attackers are weaponizing AI to automate reconnaissance, uncover vulnerabilities and refine exploits. Pitch-perfect, AI-generated phishing emails are harder to detect. Uncanny voice and image deepfakes facilitate fraud and spread disinformation. Polymorphic malware changes its code structure to avoid signature-based detection. Autonomous, multi-stage attacks supercharge the impact of sophisticated actors, while readily available, AI-driven hacking tools lower barriers to entry for novices.

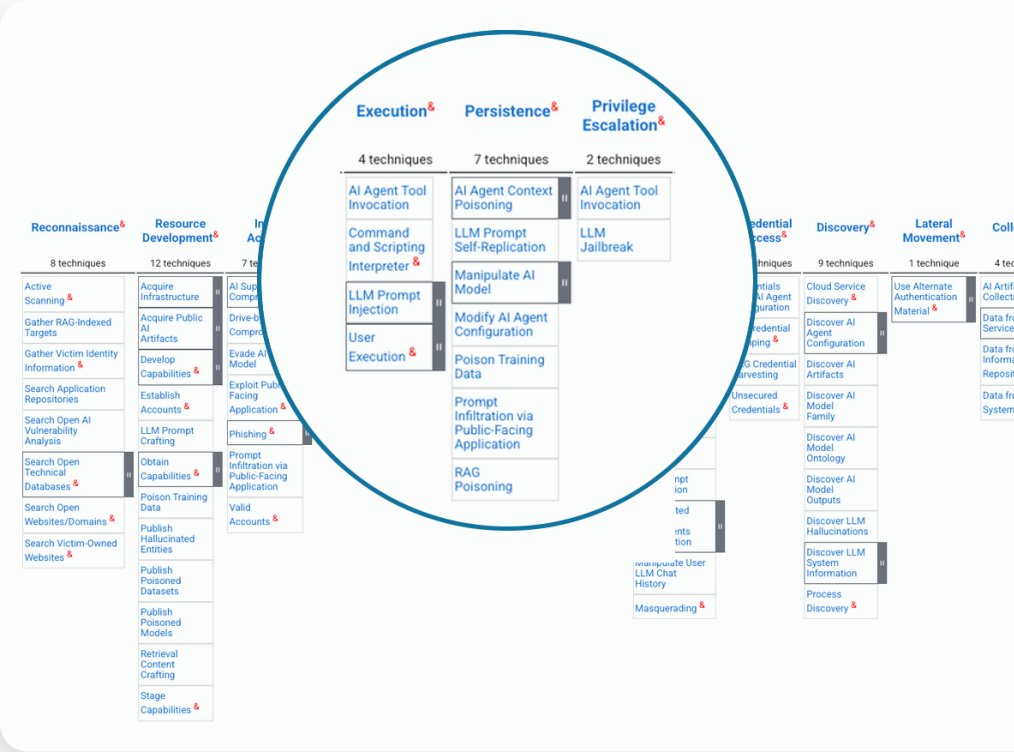
Bad actors may be using the same phishing, DDo and other techniques as before, but with greater speed, scale and efficiency.

A New Attack Vector: Attacks on AI

Bad actors have wasted no time finding ways to turn AI on itself, attacking vulnerable AI models that lack adequate defenses.

These models are susceptible to a range of threats, including data poisoning, prompt injections and model extraction, which can undermine their effectiveness or even turn them against their operators.

MITRE ATLAS™ explains how attackers exploit the vulnerabilities of AI systems. This complementary framework to MITRE ATT&CK® focuses on real-world tactics and techniques that adversaries use to target AI systems, including generative AI and LLMs.



The MITRE ATLAS Matrix shows the progression of tactics used in attacks as columns from left to right, with the ML techniques belonging to each tactic below.

With data poisoning, malicious or corrupted data can cause models to misclassify inputs, make targeted errors and reduce overall reliability. To use an OT/IoT example, consider what would happen if an AI-driven predictive maintenance system were manipulated in a data poisoning cyberattack (AI assisted or not). Adversaries might change sensor readings or introduce deceptive maintenance logs into the data. By feeding the system with false information, attackers could mislead the AI model into making inaccurate predictions about health and maintenance needs, which may lead to breakdowns, increased downtime and potential safety risks.

Prompt injections are carefully crafted inputs designed to manipulate the output of a chatbot, assistant or other AI agent, either by bypassing restrictions on what it can say or tricking it to ignoring previous commands. Model inversions attempt to reverse-engineer AI models to uncover their design, extract training data or even replicate the system for competing purposes. As we cede more and more capabilities to AI, effectively outsourcing decision-making to them, these attacks wield tremendous power.

Executives are keenly aware of the threats. In public disclosures from 80 companies on the 2025 Fortune 100 list, **89% included AI as a risk factor**, compared with 69% in 2024.

How Defenders Are Leveraging AI

On the defenders' side, the need to rapidly analyze and correlate vast amounts of data from dozens of sources presents a prime use case for AI and ML. These superhuman capabilities are accelerating nearly every aspect of cyber defense including asset inventory and profile enrichment, behavior baselining, anomaly and threat detection, event correlation, risk prioritization and noise reduction. Together they have the potential to fill the cybersecurity skills gap and potentially eliminate the need for Tier 1 SOC analysts and other junior positions altogether.

AI-assisted cybersecurity for industrial environments leverages all of these capabilities, arguably to a greater degree: there's more to protect and the stakes of an attack are often higher. You're dealing with control systems and physical processes that have thousands of configurable process variables, all potentially exploitable. And you have legacy components that are insecure by design, with limited opportunities to patch. Without AI, it would be impossible to evaluate the volume of network communication and process variable data in a typical industrial network, and certainly not fast enough to prevent harm if a serious threat was detected.

AI and ML are accelerating nearly every aspect of cyber defense including asset inventory and profile enrichment, behavior baselining, anomaly and threat detection, event correlation, risk prioritization and noise reduction.

Within industrial and critical infrastructure organizations, AI adoption among defenders is still nascent. According to the **SANS 2024 State of ICS/OT Cybersecurity** survey, only 10% of respondents use AI in both enterprise IT and OT networks. That percentage is likely to surge as security teams realize they must keep up with adversaries who are using AI to increase the sophistication and velocity of their operations. In a similar **Takepoint Research survey**, 80% of OT cybersecurity professionals said AI's security advantages outweigh its risk.

Are the Bad Guys Winning?

AI is transforming both cyber defense and cybercrime, but current privacy and AI regulations limit defenders, while attackers operate without guardrails and ethical constraints.

Many stakeholders believe the AI landscape is evolving too rapidly to keep pace. According to the Forbes Research 2025 AI Survey of more than 1,000 C-suite executives, 63% believe AI-enhanced threats could render current protections obsolete every few months. A similar number (62%) said AI exacerbates the challenge of ensuring their cybersecurity measures and training are up to date.

63%

Percent of C-suite executives who believe AI-enhanced threats could render current protections obsolete every few months.

Forbes Research 2025 AI Survey

AI and ML have clear benefits for cybersecurity teams, helping them do astoundingly more with fewer resources. That's a good thing, especially given the skills gaps in specialized fields like OT and IoT security. But adversaries are also reaping the benefits. As AI science and technology continue to evolve, so will cyberattack methods. It's hard to say who's winning the battle on any given day, but organizations need to know their adversaries' capabilities and aim to exceed them if they want to prevail in an AI-powered world.

3. How We Use AI in the Nozomi Networks Platform

Nozomi Networks is the leader in AI for OT and IoT cybersecurity, period. Unlike vendors who are just starting to incorporate it into their products — sometimes by calling out via API to an external AI service or pretrained large language model (LLM) — our R&D and lab teams have been building AI/ML into our platform and training our AI engine in-house since Day One — more than 13 years ago. That's more than a decade of real-world experience leveraging AI to defend some of the largest industrial and critical infrastructure organizations around the globe.

13+ yrs

Our R&D and lab teams have been building AI/ML into our platform and training our AI engine in-house since Day One — more than 13 years ago.

We know how to collect the right data, provide the right context and use the right AI techniques so industrial and critical infrastructure organizations can defend themselves in today's world.

It Starts with Good Data. Lots of It.

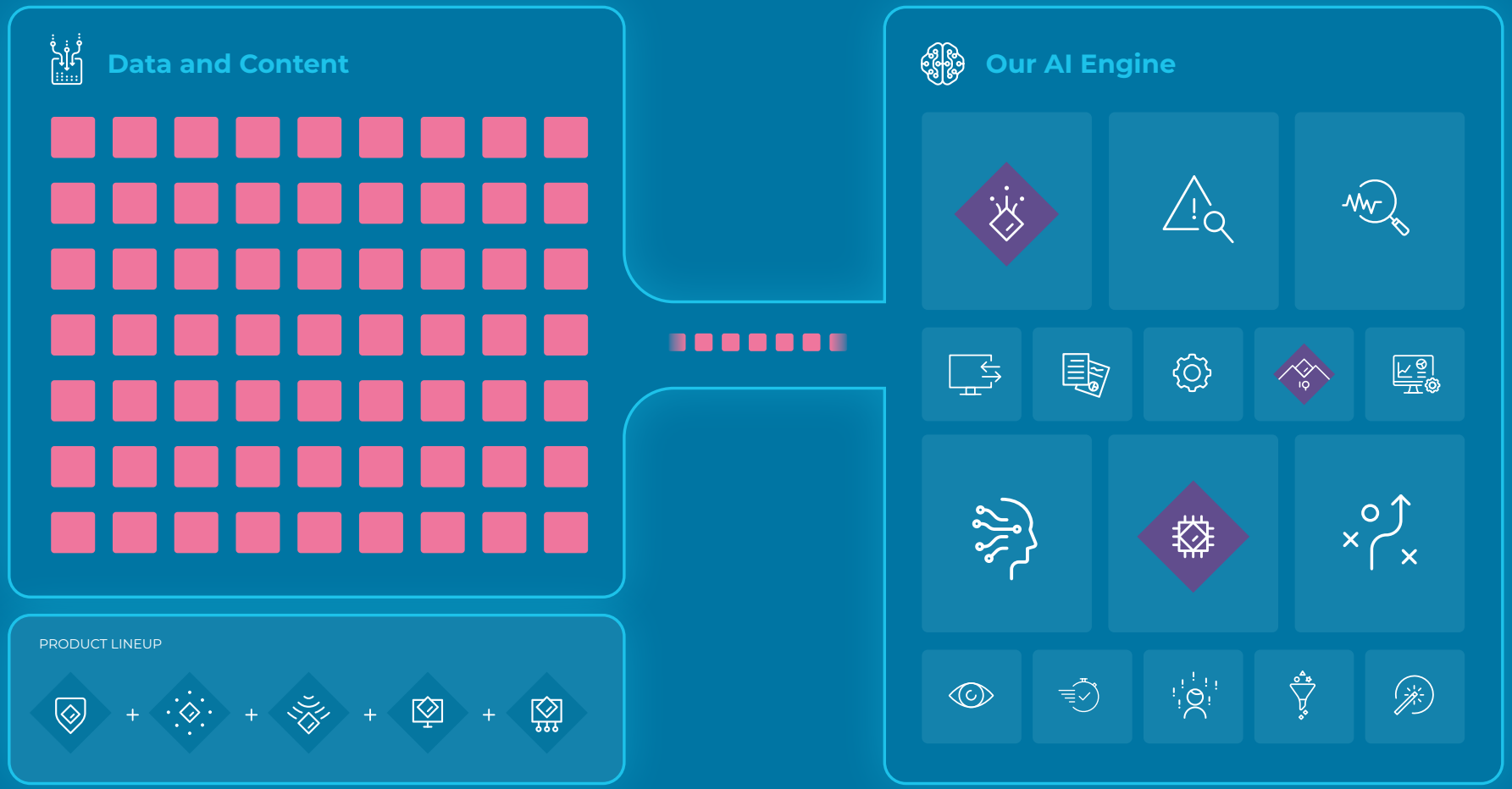
Industrial systems operate based on inputs and outputs (I/O). Sensors feed data in; actuators respond out. It's no different with AI/ML. The insights, predictions and detections it puts out are only as good as the data and context that are put in.

A complete, accurate inventory of all OT, IoT and IT assets in your environment is the input that enables our AI engine to produce the right outputs. To collect the right data and context, we use a variety of network, endpoint and wireless sensors; active and passive discovery techniques; and deep packet inspection (DPI) and comprehensive protocol fluency to analyze network traffic and understand behavior.

Our AI engine continuously learns from millions of monitored assets so it can fill in gaps about identical devices across environments, giving you the breadth and depth of data needed to detect threats and anomalies and manage risk.

Keys to Nozomi's AI Effectiveness

- ✓ Maximum data and context to both train and operate: all the sensors to collect robust asset and network data + deep packet inspection + comprehensive protocol support
- ✓ Deep experience applying a variety of AI tools (Nozomi AI engine) to solve OT/IoT cyber and power business outcomes
- ✓ Over a decade of real-world experience leveraging AI to defend some of the largest brands



The Right AI/ML Models for Every Challenge

We use a variety of AI/ML throughout our platform, choosing the right tool (ML, predictive analysis, behavioral analytics, Bayesian Networks, LLMs) for the task at hand, so you get actionable insights into your environment that explain what to do now to increase operational and cyber resilience.

	Asset Inventory	Vulnerability Management	Anomaly Detection	Threat Detection	Risk Management	SOC Efficiency
Key Features	• Automatic identification • Enrichment • Classification	• Advance assesments • Prioritization • Remediation guidance	• Adaptative baselines • Process anomaly detection • Traffic predictions	• Precise detections • Alert correlation • Root-cause analysis	• Dynamic risk calculations • Benchmarking • Risk guidance	• Automatic Identification • Enrichment • Classification
Machine Learning	✓	✓	✓	✓	✓	✓
Predictive Analysis			✓	✓	✓	✓
Behavioral Analytics			✓	✓	✓	✓
Bayesian Networks	✓					
Large Language Models						✓

Let’s look at each of these use cases in more detail.

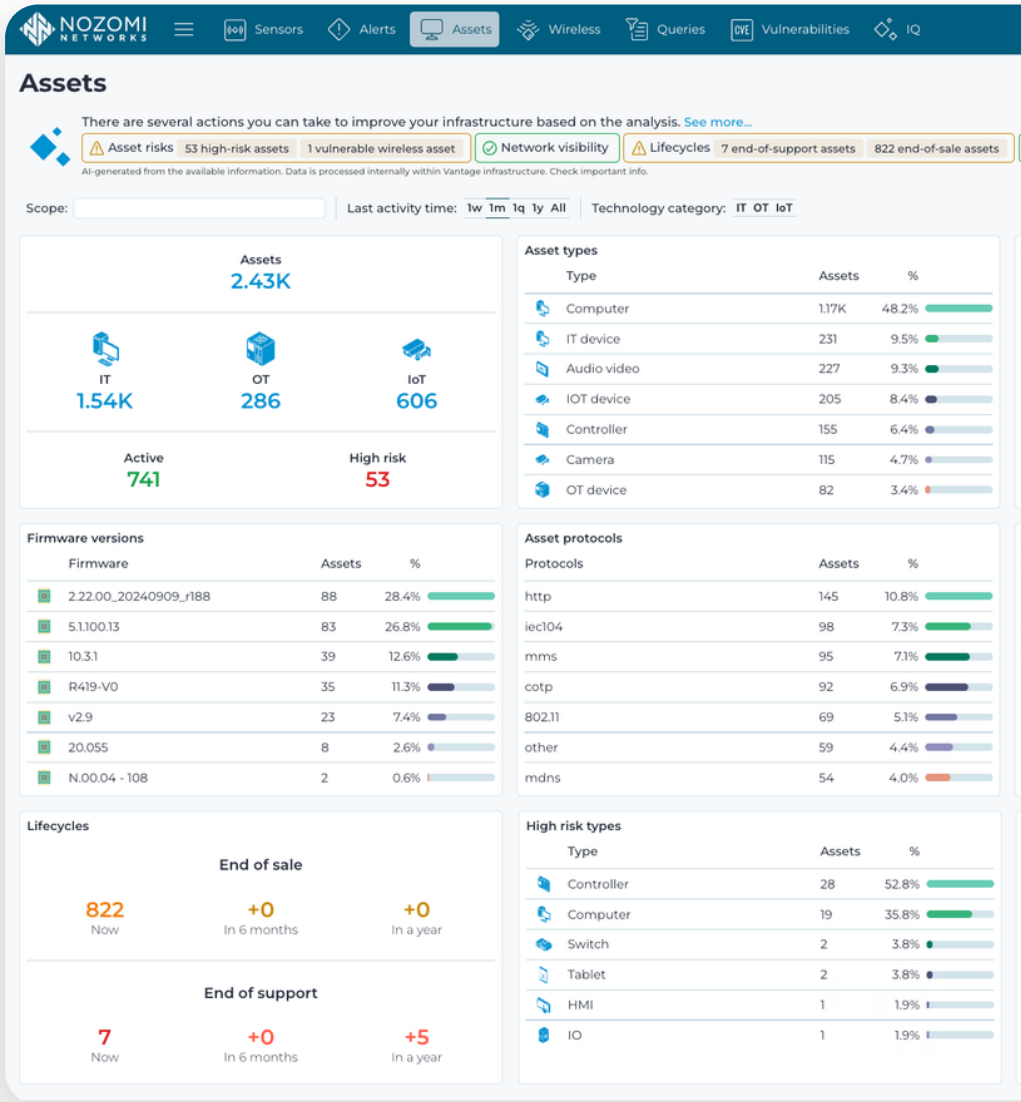
Automated Asset Inventory

To correctly identify assets, classify them, fill in missing information and enrich what we know about them, we use ML to match observed characteristics with a continuously updated database of device profiles maintained by Nozomi Networks Labs.

If limited data is available, our behavioral inference models can infer asset types and roles based on traffic patterns and protocol usage.

We also leverage Bayesian Networks, a probabilistic model used to reason under uncertainty, to predict what asset information belongs in missing data fields until it can be collected or otherwise populated. This is a very effective way to avoid misclassification of key data used to detect threats and anomalies and manage risk.

Together, the wide array of sensors, data collection methods and AI enrichment techniques continuously raises your overall inventory accuracy.

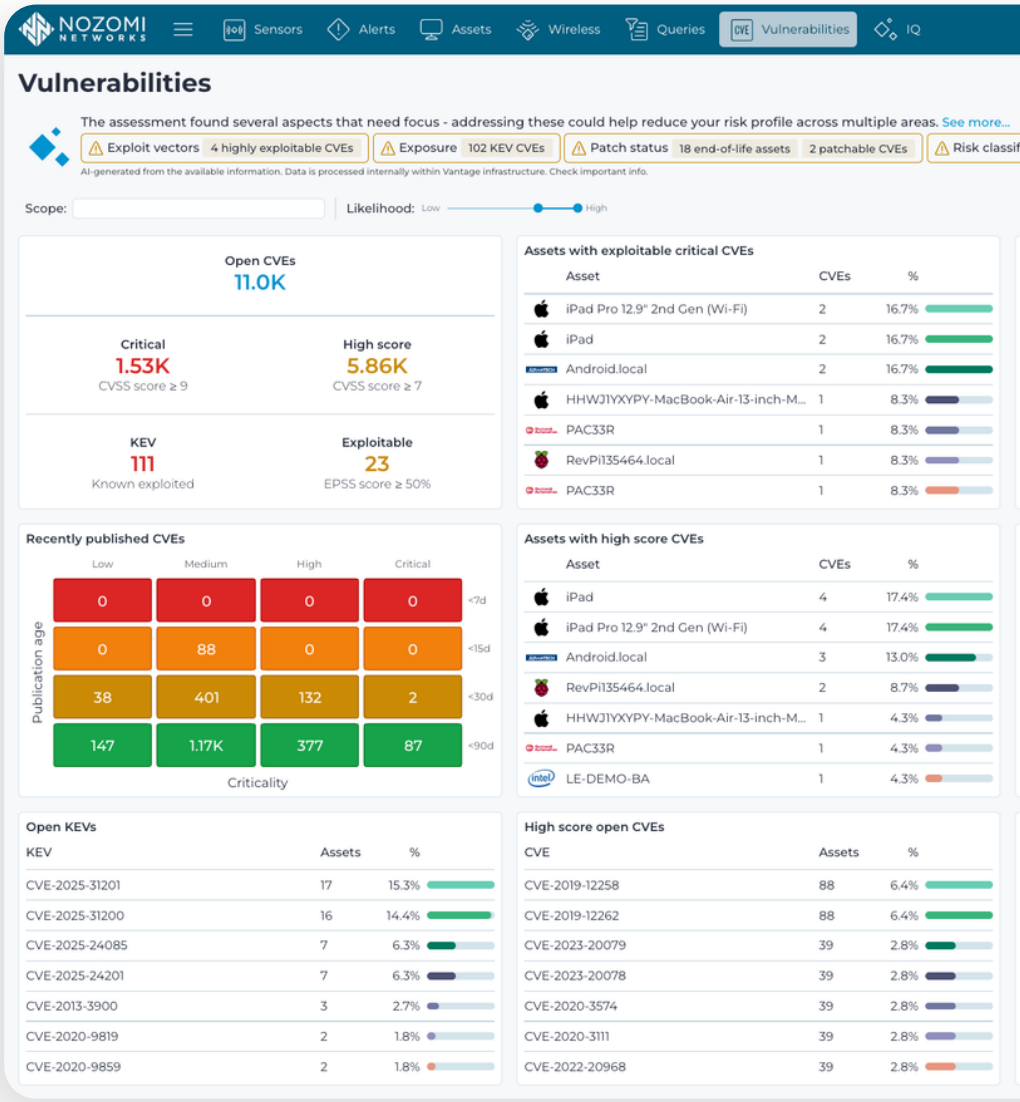


Vulnerability Management

Effective vulnerability management involves contextualizing, prioritizing and correlating vulnerabilities with real-world risk in your environment. Our platform uses AI-trained asset fingerprinting to identify device make, model, firmware version, OS and more. This enriched profile is used to match the device against known CVEs with far greater accuracy than traditional scanners.

We then use Bayesian inference and weighted probabilistic models to calculate a dynamic, multi-factor risk score that includes vulnerability risk, including patch status.

While continuously monitoring your environment, the platform uses temporal correlation, behavior modeling and threat pattern matching to identify suspicious behavior near a vulnerable asset, inbound probes from threat actors or lateral movement patterns. Any of these events triggers a risk score elevation and alerts, helping teams prioritize vulnerabilities that are actively being targeted.



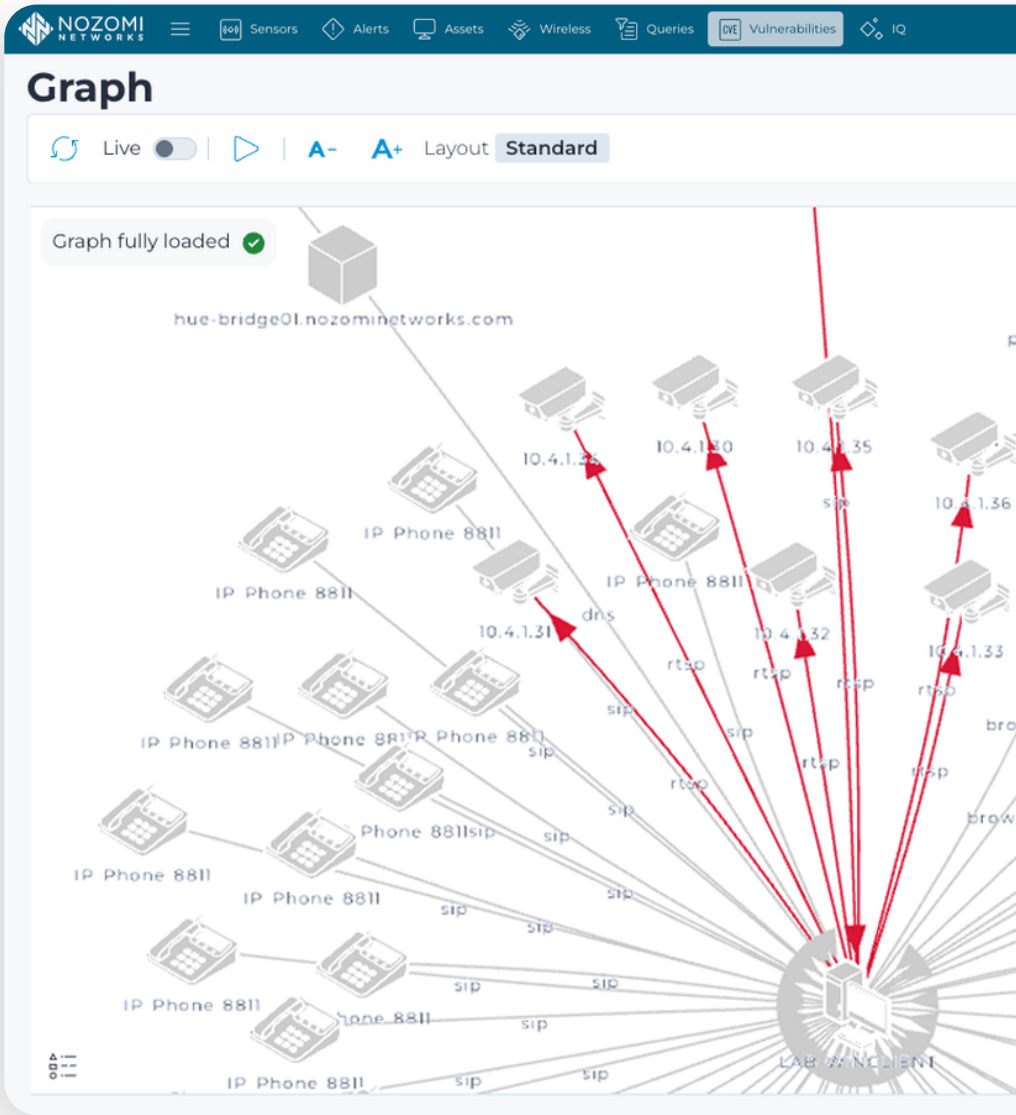
Anomaly Detection

Operational anomalies can't be detected using simple rules. A combination of ML, predictive analytics and behavioral analytics is essential for baselining asset behavior and detecting anomalies.

Upon deployment, the Nozomi Networks platform begins monitoring device communications in "learning mode," all the way down to process-level variables. It uses ML and predictive analytics to create detailed profiles of the expected behavior of every device at each stage in a process to establish a baseline of "normal" behavior.

When switched to "protection" mode, the platform uses behavioral analytics to monitor the environment, compare current behavior against baselines and alert on suspicious events that deviate from them, evaluating their criticality and classifying them as a process or cybersecurity anomaly. Even when in protection mode, the system dynamically updates the baseline if normal conditions change.

To reduce false positives, we use behavioral modeling, pattern recognition and other techniques to filter out benign changes such as legitimate firmware updates.

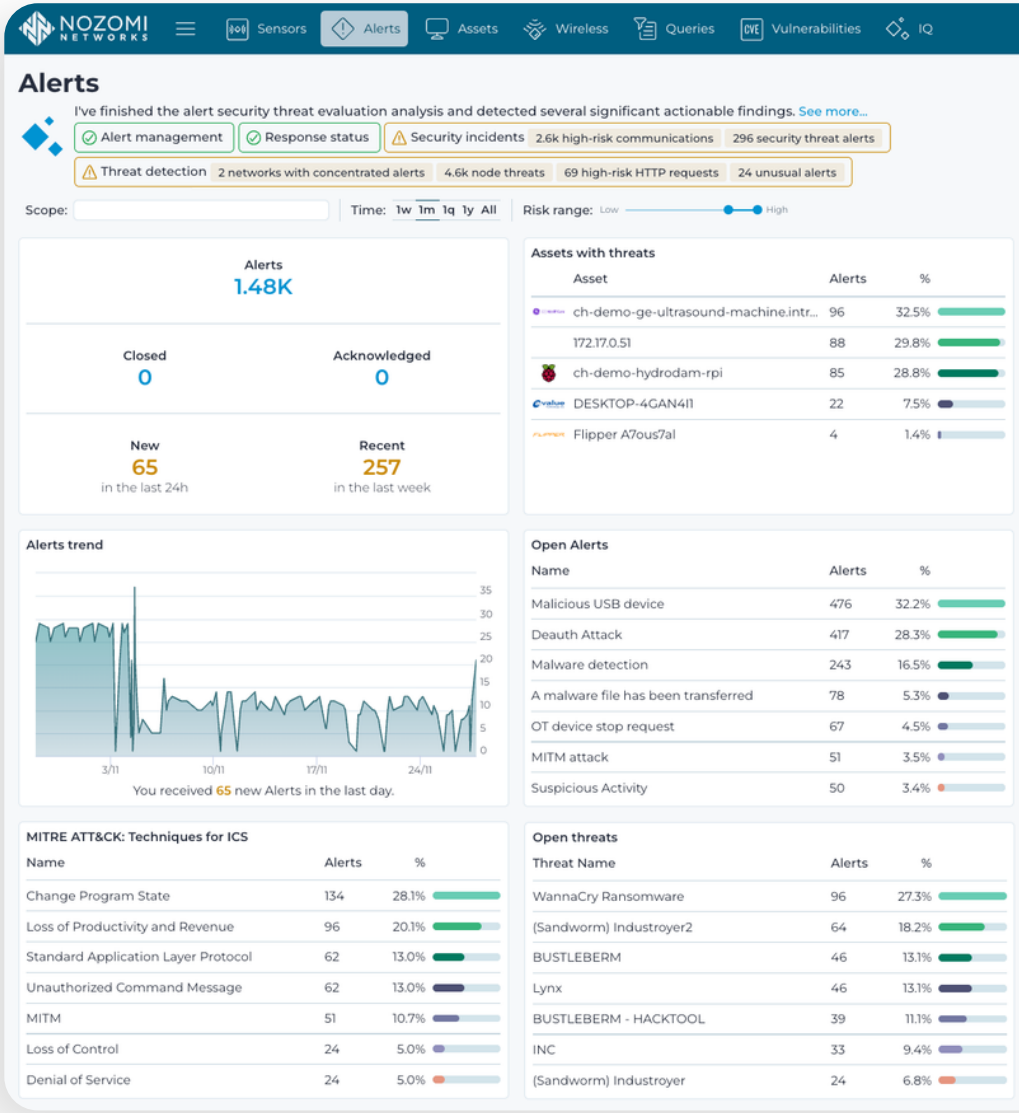


Threat Detection and Mitigation

Rule-based detection, including signature-based, is efficient for detecting known threats, where the indicators are easily observable and identifiable. Unknown threats, including zero days, require the same behavior-based detection techniques as anomalies. Neural network models, Bayesian Networks and other AI techniques are also essential for managing threat-related alerts and prioritizing mitigations.

Neural network models correlate multi-variable events across your environment to reduce investigation time and spot complex threats such as advanced persistent threats. Our query engine analyzes these correlated alerts alongside asset attribute and network relationships to suggest the right steps to take.

Root cause analysis is essential for threat investigation. Our platform uses neural network models, clustering and time-series analysis to correlate behavior across assets, traffic and time. It quickly isolates the source of anomalies or alerts by identifying causal chains, which reduces investigation time and enables faster, more targeted response.



Risk Management

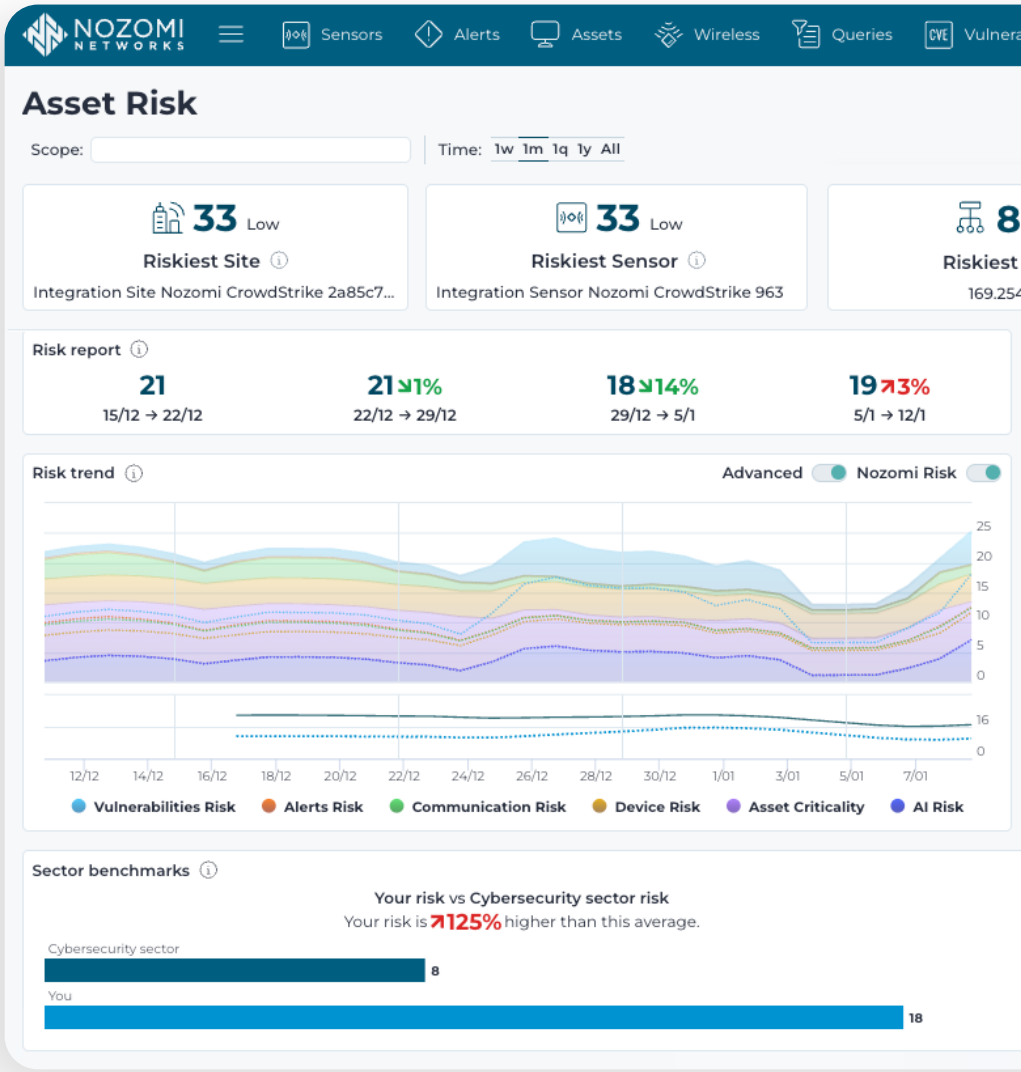
The Nozomi Networks platform calculates dynamic risk scores for each of your assets to help you prioritize security efforts, address the most critical risks first and mitigate them effectively. It calculates asset risk based on five factors with customizable weights: vulnerability risk, alert risk, communication risk, device risk and asset criticality, counterbalanced by compensating controls in place.

We use a combination of ML, predictive analytics and behavioral analytics to calculate risk scores at the asset, facility and enterprise level. These same tools are used to recommend actions to take, ranked by how much they will reduce your overall risk score.

Calculations are updated as the threat environment changes, new vulnerabilities are reported, we see anomalous behavior in your network and as you add controls, so you can assess their impact.

We also use clustering, statistical modeling, supervised learning and contextual analysis to display peer benchmarks, so can see how your security posture compares to other companies in your region or industry.

Finally, we use predictive analytics based on historical vulnerability, threat and asset behavior data to help identify which vulnerabilities are likely to be exploited, which asset types or sites are riskiest and emerging attack chains.

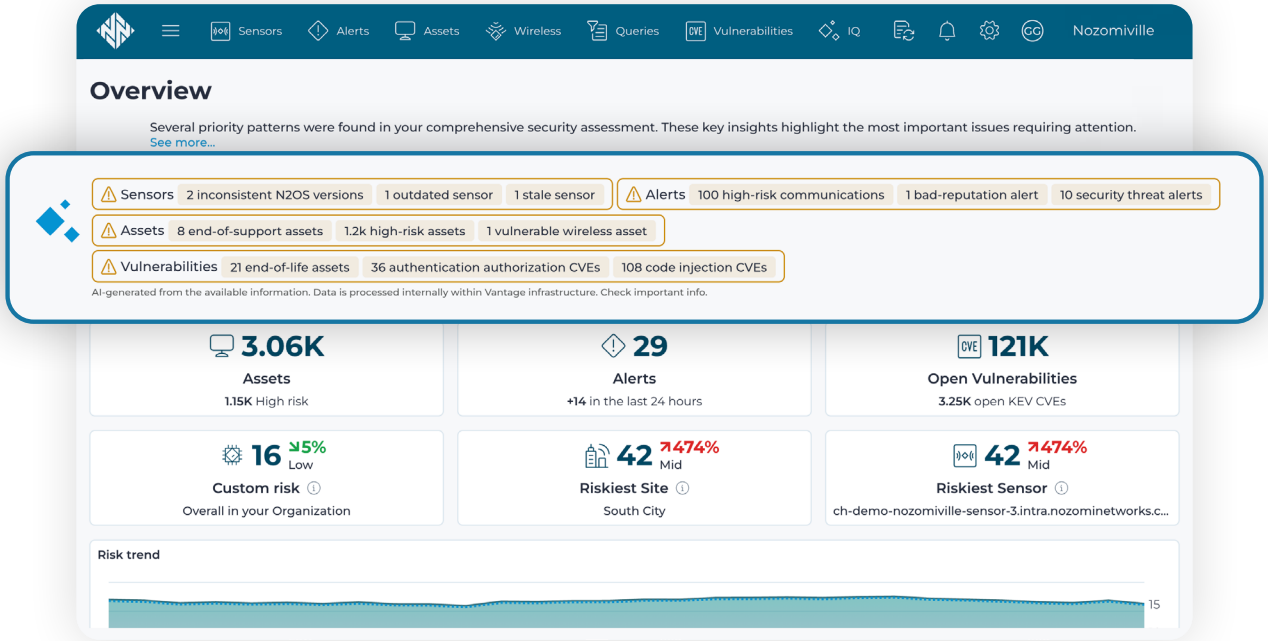


SOC Efficiency

The overriding value of a cybersecurity platform is ease of use. It can collect all the right data and use all the right AI and ML techniques to draw all the right conclusions, but how valuable is that if authorized users, including non-expert stakeholders, can't easily tap into those insights?

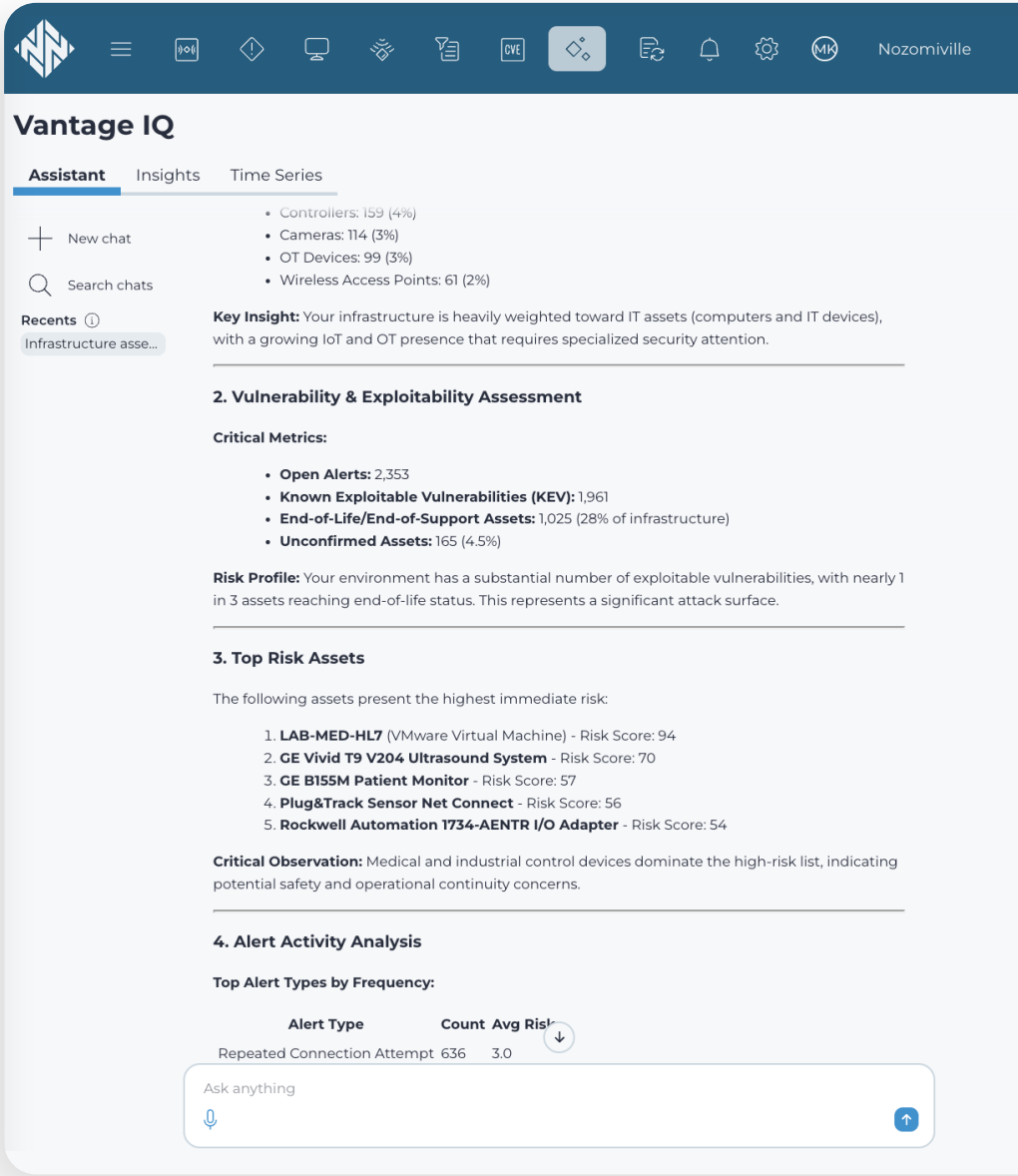
The AI methodologies used throughout the Nozomi Networks platform culminate in Vantage IQ summaries, which give stakeholders instant awareness of the most critical issues in their environment. It uses an internal AI agent to put continuously refreshed security insights and remediation advice at users' fingertips throughout Vantage, our SaaS solution. Especially for teams that are new to industrial cybersecurity or lack dedicated resources, Vantage IQ replicates the domain expertise of a seasoned OT/IoT security analyst, at a fraction of the cost.

Especially for teams that are new to industrial cybersecurity or lack dedicated resources, Vantage IQ replicates the domain expertise of a seasoned OT/IoT security analyst, at a fraction of the cost.



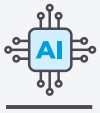
Vantage IQ also includes a generative AI assistant with its own query language, natural-language interface and LLM. We didn't just use an API to bolt on a commercial digital assistant and LLM. Instead, our LLM is trained on real-world incident telemetry across thousands of industrial deployments and fine-tuned with data from asset profiles as well as public and Nozomi-curated threat intelligence.

Any authorized user, from a SOC analyst to an OT engineer, can ask the Vantage IQ Assistant anything about their environment and get an immediate answer that's accurate and actionable, with drill-down access to deeper insights.



4. How We Protect Our AI Components from Cyberattacks

Cybersecurity leaders have enough on their plate trying to evaluate which products leverage AI effectively and whether they will keep pace as threats evolve. With AI as an attack vector, buyers must also evaluate how well vendors safeguard the AI in their products from compromise.



Guidance on Responsible AI Procurement and Use

1. CISA et al.: Principles for the Secure Integration of AI in OT
2. IEC 23001: AI Management System Standards
3. The EU AI Act
4. NIST Risk Management Framework

Several resources provide guidance, some specific to OT environments. In December 2025, the U.S. Cybersecurity and Infrastructure Security Agency (CISA) and the Australian Cyber Security Center (ACSC) published a **detailed guide** to help owners and operators of critical infrastructure safely integrate AI into their OT systems. Collaborators include the National Security Agency (NSA), Federal Bureau of Investigation (FBI), and cybersecurity agencies from Canada, Germany, the Netherlands, New Zealand, and the UK. Nozomi Networks is one of five cybersecurity vendors that contributed to the document.

AI has been natively built into the Nozomi Networks platform from the beginning, and we have always used the safest, most conservative model for OT-integrated AI. That extends to Vantage IQ, our AI assistant, which follows these principles:

- Zero customer-data training
- Local processing options
- Strict prompt governance
- MFA/RBAC/AAA access control
- Human-in-the-loop decision model
- Transparent data handling
- No autonomous actions on OT systems

Following are details that illustrate these principles in action.

Prompt Injection Attacks on Our LLM

We use multiple safety checks to sanitize and validate user prompts to ensure safe model interactions and prevent prompt injection attacks on our LLM. Before any prompt is executed, it undergoes a screening procedure that uses standard rules such as pattern matching and word analysis. Our internal AI agent must also verify the prompt's safety and relevance. The internal system prompt also contains instructions to prevent this type of attack.

User Data Protection

Our internal AI agent is trained on our internal data and instances. No customer or user data, including personally identifiable information (PII), is accessed by it or saved during its use. No user data is ever logged or used to improve our AI engine, with or without explicit consent.

LLM Training Data Integrity and Security

Our internal AI agent uses only data from our internal instances that has been validated and checked. We constrain inputs and outputs to (1) ensure that our internal AI agent always refers to cybersecurity and network visibility suggestions and (2) detects and suppresses harmful, biased or inappropriate outputs from the LLM. Outputs are tested regularly to ensure compliance with ethical AI standards and organizational policies.

A leading LLM provider supplies the foundational LLM that our assistant runs on. The model has already been trained; no fine-tuning is required. It verifies that our datasets exclude harmful, biased or malicious content. It enforces two checks:

- Screening procedure: ensures the questions pertain to the current user's cybersecurity environment and that the LLM output is relevant to that user's data; detects and responds to attempts to abuse of the LLM such as phishing, data scraping, or generating toxic content or misinformation.
- System prompt engineering: a set of specific instructions given to the LLM to ensure the safety and integrity of the output.

Additionally, outputs are tested regularly to ensure compliance with ethical AI standards and organizational policies.

Third-party Resources

Like other AI product developers, we rely on third-party libraries and pretrained models to mitigate vulnerabilities. Our software is regularly analyzed using leading platforms for software composition analysis (SCA), and we update third-party components after every SCA scan to resolve critical and high issues.

Access and Use

As with any enterprise software, role-based security is foundational. We use the following controls to ensure that AI-powered services are accessed only by authorized users with the right privileges, used appropriately and monitored for compliance:

- To access any of the AI-powered services, the user must be logged into their Vantage instance.
- Authentication, authorization and accounting (AAA) is integrated into the Vantage platform and RBAC is in place. SAML IdPs can be integrated with the platform by customers.
- Each user has access to the internal AI agent (e.g., Vantage IQ summaries) and the LLM assistant (when Vantage IQ licenses are available).
- In both cases, the data used to compute the summaries or to be analyzed by the LLM aligns with the user's RBAC permissions.

5. How to Evaluate and Verify AI Vendor Claims in OT/IoT Cybersecurity Solutions

Every organization is evaluating how to incorporate AI into their operations safely and effectively, including for cybersecurity. But buyer beware; there's a wide range of AI and ML methodologies out there, and an equally wide range of vendor claims. Don't be misled by hype, and don't hesitate to ask for plain-language explanations if you're unfamiliar with AI jargon; you're in solid company with your peers.

At its best, an AI-powered security platform can continuously analyze unlimited amounts of data, diagnose your dynamic security posture, prioritize risks and serve up the right insights to help you contain threats and optimize your environment. However, none of that matters if it can't handle the unique challenges of OT/IoT environments. Before you consider AI capabilities, be sure the solution:

- ❑ Is fluent in specialized protocols
- ❑ Uses monitoring techniques compatible with decades-old systems that control physical processes
- ❑ Relies on OT- and IoT-focused threat intelligence
- ❑ Respects always-on processes that can't go down without impacting revenue or essential services

Here are some questions to ask when evaluating AI-powered cybersecurity solutions for OT/IoT environments.

AI Capabilities

1. What AI/ML methodologies (ML, predictive analytics, behavioral analytics, Bayesian networks, LLMs, agentic AI) does the solution use? How long has each methodology been in use?
2. What AI/ML methodologies are used for these common cybersecurity use cases:
   - ☐ Asset inventory
   - ☐ Vulnerability management
   - ☐ Anomaly detection
   - ☐ Threat detection
   - ☐ Risk management
   - ☐ SOC analysis, investigation and remediation
3. What AI tools and methodologies are developed in-house? When were they developed and how have they evolved?
4. What AI tools and methodologies are developed by third parties, and how are they integrated? How long have they been part of your solution?
5. How is your AI engine trained and updated?
6. What data sources are used by the AI engine?
7. What is your AI development roadmap (in-house and via integration)?
8. What AI industry frameworks or standards do you align with?

AI Protections

1. How do you protect your LLM from prompt injection attacks?
2. How are the data sources used by your AI engine secured? If they include user data, how do you protect that data?
3. How do you protect the integrity and security of data used to train your LLM?
4. How do you protect your model from harmful, biased or inappropriate inputs and outputs?
5. How do you ensure compliance with AI ethics standards and organizational policies?
6. What formal AI governance is in place?

5. Conclusion

In the cyber realm, AI is being leveraged by both attackers and defenders to do what they've been doing, only better. Meanwhile, OT and IoT systems are more connected than ever, often without proper controls. AI-powered security solutions can mitigate risk exposure through advanced threat detection, event correlation, diagnosis and response — if, as stated earlier, they understand and are sensitive to industrial and critical infrastructure environments.

When evaluating AI-driven security solutions, certainly current features and product roadmap are important. Especially for OT and IoT, however, AI in cybersecurity is still in its infancy. How do you evaluate the application of a technology breakthrough that may look completely different in five years, maybe less? Today's features are important, but more important is the team who's building them — and whether they have the history, data and expertise to adapt in the right direction as AI evolves.

That's where Nozomi Networks stands apart. We've been building, training and securing our AI engine since Day One — over 13 years of continuous innovation dedicated to protecting the world's most critical infrastructure. In a field where many are just getting started. Our AI isn't new. It's in our DNA.

How do you evaluate something that may still be in its infancy to protect against threats that may also be in their infancy?

Evaluate the technology, its longevity and, most importantly, the company behind it.

NOZOMI NETWORKS

Cybersecurity for OT, IoT and Critical Infrastructure

Nozomi Networks protects the world's critical infrastructure from cyber threats. Our platform uniquely combines network and endpoint visibility, threat detection, and AI-powered analysis for faster, more effective incident response. Customers rely on us to minimize risk and complexity while maximizing operational resilience.

nozominetworks.com

NN-WINNING-THE-AI-POWERED-CYBERWAR-EBOOK-8.5x11-001