

Artificial Intelligence at ESET

Juraj Jánošík, ESET Vice President of Artificial Intelligence

Ondrej Kubovič, ESET Security Awareness Specialist

Filip Mazán, ESET Senior Manager of AI Research

Kamil Pšenák, ESET Senior Product Manager for AI



Cybersecurity
Progress. Protected.

ESET spol. s r.o., Einsteinova 24, 851 01 Bratislava,
Slovak Republic; tel.: +421 2 322 44 111,
fax: +421 2 322 44 109, www.eset.sk

Artificial Intelligence at ESET

Juraj Jánošík, Ondrej Kubovič, Filip Mazán, and Kamil Pšenák

Copyright © ESET, spol. s r.o., 2026. All rights reserved.

First published in Slovak Republic Q2/2026

First presented on ESET World 2026 in Berlin | 19 – 20 May 2026

ESET, spol. s r.o. grants permission to use this work under the following conditions:

1. The use of this work is not expressly prohibited by ESET, spol. s r.o.
2. The work will not be used for any purpose directly or indirectly connected to commercial activity.
3. The work is intended solely for informational, non commercial, or personal use.
4. The work will not be copied, published, or distributed in any form of media without prior written consent from ESET, spol. s r.o.
5. The work will not be altered, adapted, or modified in any way.
6. This document—and all copies of it—must retain all applicable copyright notices.
7. The original source must always be clearly indicated.

Any use of this work beyond the scope described above is expressly prohibited and may be unlawful.

While ESET, spol. s r.o. has taken care to ensure the accuracy of the information provided, neither the authors nor ESET, spol. s r.o. accept any responsibility for loss or damage arising from actions taken—or not taken—based on this work.

ESET, spol. s r.o. acknowledges all applicable patent rights, if any, related to this work. No part of this work should be interpreted as an intention to infringe upon any such rights.

This work is provided by ESET, spol. s r.o. in good faith and for informational purposes only.

Throughout its history, ESET has demonstrated that responsible innovation – especially with novel and quickly developing technologies like machine learning and artificial intelligence (AI) – is not only possible but essential in cybersecurity.

Since 1997, when ESET experts conducted their first experiments using neural networks for threat detection, their approach stood apart due to its measured integration of these technologies. From the initial use to the present-day advanced systems, ESET R&D remains grounded in rigorous scientific testing, careful oversight, and focus on user safety and real-world benefit.

Machine learning before it was cool

The AI journey at ESET began with early implementation in **1997** focused on improving the **detection of macro viruses**, testing and confirming that machine learning algorithms could lead to better protection of its customer base.

```
d:\docs\BOOK1.XLS (0.0000:0.0000[-----]) - XP/Paix pattern
d:\docs\ERASER-F.DOT (0.9908:0.0094[UI0-----]) - MP/Eraser.F:Tu virus
d:\docs\LTU.XLS (0.9989:0.0108[UI0-----]) - XP/IMU.C virus
d:\docs\PRIZM.DOC (0.9775:0.0225[UI0-----]) - NEURAL PATTERN
d:\docs\DRILYKID.DOT (0.0012:0.9988[-----CL0]) - POLY CRYPT STEALTH MACRO virus
d:\docs\MMAC.XLS (0.8833:0.1167[UI0-----]) - NEURAL PATTERN
d:\docs\MACOLIN.DOC (0.9920:0.0074[UI0-----]) - MM/ABC.A virus
d:\docs\XSYMPHON.XLS (0.0000:0.0000[-----]) - XSYMPHON.A virus

DIME
Summary:
neural network - clean files: 1
neural network - viruses 1: 5
neural network - viruses 2: 5
neural network - total: 9
viruses detected by name: 6
virus suspicion - heuristics: 1
virus suspicion - neural network: 2
total errors: 0
total number of scanned files: 9
total number of processed files: 9
```

Figure 1: Results of sample processing via multilayered detection engine and detection tools, including samples identified as “virus” by neural networks (NEURAL PATTERN) as well as by other proprietary detection technologies.

If neural network support exists for given target then non-zero values as the result of neural network scanner will be displayed immediately after the scanned file name:



If the information collected by neural network is not sufficient to decide whether the file is infected or not both 'CL0' and 'VI0' flags will be displayed.

Note: Because of using linear approach in neural network model to evaluate total probability, likelihood of infection can be in some cases greater than 1 and likelihood of being clean can be less than zero (negative number). This should be interpreted as almost 1 or almost 0. In fact probability should be a number from the <0, 1> interval.

Figure 2: Excerpt from manual of Heuristic macro virus scanner leveraging neural networks.

In the breakthrough that followed in **2005**, ESET deployed its **DNA Detections** technology. By extracting key features – the ‘genes’ of potentially malicious samples – ESET engineers created broad DNA profiles that became the foundation for accurately distinguishing benign from malicious files and enabled the detection of never-before-seen threats with similar traits that would emerge in years to come.

This process, regularly updated by both automated systems and researchers, became a flexible, adaptive, and – most important of all – safe layer of defense.

Freeing Expertise: Automation and Threat Research

Expecting boom in threats that will attempt to harness the power of machine learning – today often referred to as AI – ESET scaled up its use of the technology for real-world performance.

Since **2006**, machine-learning based **backend expert systems** have been a crucial part of processing, triage, labeling, and detection of the increasing quantity of samples – a number that has grown from tens of thousands of unique samples analyzed daily to its current levels of 750,000.

Ultimately, these AI-assisted backend systems freed the capacity of ESET security experts, allowing them to focus less on tedious and repetitive tasks and more on research of the most interesting emerging threats. In the following years, the gathered threat intelligence laid the foundation for development of new advanced protective technologies that are helping the company address specific attack vectors even today.

The development and launch of **ESET LiveGrid®** cloud reputation system in **2010** led to another leap in detection and response. By embracing online learning, ESET started to deliver threat database updates within minutes to its clients across the globe. Again, the technology was no black box but an evolving, supervised layer that adapted as new samples and threats appeared, delivering the latest information about emerging threats almost instantly.

Careful integration instead of hype

In the late 2010s, the surge of deep learning and neural algorithms led to boom in cybersecurity, with many newly established post-truth vendors exploiting the hype, promising nonrealistic results. Again, ESET chose the path of cautious integration over hype-driven adoption.

While many competitors rushed to deploy the technology to evaluate virtually any malicious behavior – an approach that ultimately swamped defenders with false positives – ESET opted for exhaustive testing of its AI-related approaches when pitted against threats seen in the wild. Drawing on the results of that process, ESET engineers created a precise mix of long short-term memory (LSTM) neural networks, decision trees, and traditional supervised learning methods to form **ESET Advanced Machine Learning module**.

This new technology, released publicly in **2017**, balanced high detection rates with a close-to-zero false-positive rates – a critical differentiator for day-to-day operations of the already overstretched cybersecurity practitioners.

In the same year, Google researchers published the paper [“Attention Is All You Need”](#), which introduced transformer architecture and sparked a new era of innovation in AI. **In 2020, ESET became one of the first security vendors** to deploy **transformer-based detection model** as an extra layer of defense against cyber threats. It's worth noting that this advancement came years before transformers became the driving force behind generative AI tools like OpenAI's ChatGPT.

Timeline of AI technology development highlights at ESET

2005
ESET DNA Detections, a synonym for online machine learning, uses genes of malware to detect current and emerging threats.

2017
ESET Advanced Machine Learning in the cloud uses AI to power our automated detection systems.

2019
ESET Advanced Machine Learning in the endpoint uses AI to power our automated detection systems.

2023
Automated Incident Creator added to ESET Inspect to reduce noise and leverage techniques, including AI, to produce a clear overview of an incident.

1997
First use of neural networks in ESET products, utilized for detection of macro viruses.

2010
ESET LiveGrid®, a cloud-based reputation system, leverages ESET DNA Detections to significantly speed up user-side updates.

2018
ESET LiveGuard, an AI-powered cloud sandbox, provides on-demand analysis for ESET customers with a turnover time of minutes.

2020-2021
Transformer-based models deployed in ESET's cloud and endpoint solutions.

2024
ESET AI Advisor, Release of a generative AI security advisor for ESET Threat Intelligence that can generate detailed descriptions of incidents, and for ESET Inspect (in preview), provides answers to queries regarding the scope of incidents.

ESET AI today: **Powerful, cross-checked, efficient**

Another chapter of AI-driven analysis at ESET begins with the introduction of **ESET LiveGuard Advanced** (formerly known as **ESET Dynamic Threat Defense**). Designed as a cloud sandbox, it combines several [independent detection layers](#), including advanced unpacking and scanning, static and dynamic analysis, experimental detection engine, and in-depth behavioral analysis.

of the highly optimized, fast, and extremely efficient AI module distributed to the customers is mostly below **20MB**.

To maintain highest standard of performance against the constantly changing threat landscape, ESET retrains its AI model every week. This detection-focused model consists of roughly

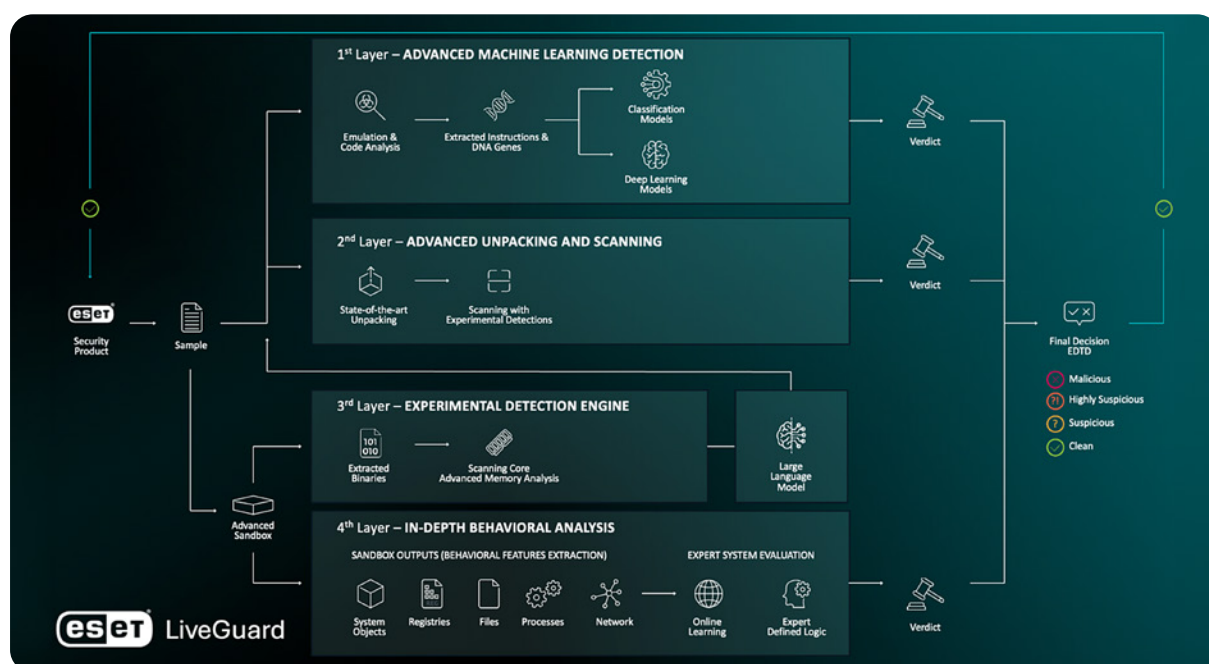


Figure 4: Simplified scheme of sample processing in ESET LiveGuard Advanced

In sum, rather than giving a single AI system unchecked authority, ESET LiveGuard Advanced leverages several types of algorithms and decision points, always cross-validating and correlating their outputs. Data produced by the exhaustive multilayered analysis is then consolidated into a final verdict presented to the customer who sent the request, placing the submitted sample in one of four categories: malicious, highly suspicious, suspicious, or clean.

The **training data** for ESET LiveGuard Advanced is carefully selected from a continuously updated **65TB+ database** of precisely classified binaries including malicious, potentially unwanted, potentially unsafe, and benign samples. In contrast with that enormous data set, the size

one million parameters and is thus vastly smaller than general-usage models which typically have trillions of parameters.

The mix of precise training data, optimization, computational power, and multitude of analytical approaches allows ESET LiveGuard Advanced to detect a wide range of attacks, including:

- **never-before-seen threats,**
- **exploits aimed at zero-day vulnerabilities,**
- **script-based attacks,**
- **fileless malware,**
- **and living-off-the-land techniques.**

Thanks to the high-powered backend infrastructure of the ESET cloud sandbox **99% of customer-submitted samples are analyzed** in under **5 minutes**. At the same time, this resource-intensive processing runs in highly controlled, robust, ESET-operated cloud environments. This ensures safety of the process, while having **minimal impact** on clients' computing resources.

ESET XDR: Robust, resilient, reliable

As machine learning and AI matured, successful previous implementations proved that the scientific approach of ESET engineering teams ensures high detection rates while ensuring resilience against outside tampering. Even today, instead of relying on the latest trends in algorithms, ESET maintains its **multilayered** and **multi-model approach**, including supervised and unsupervised learning, deep learning, heuristic layers, and transformers.

Before any new models or methods are implemented, they are first subjected to extensive, side-by-side testing – as not all AI models are equally fit for use in the security domain. ESET engineers then handpick approaches that are proven to be **robust against adversarial manipulation, resilient against evolving attack tactics**, and that **minimize alert fatigue** of defenders **by keeping the false positives rates low**.

The practical results are notable in the **ESET Protect Platform** – company's extended detection and response (XDR) solution – that enables management of tens of thousands of devices from one central console. The built-in technology, including AI, ingests large amounts of data collected from endpoints in the environment, cross-checks it with ESET global threat database, and then summarizes the results into clear dashboards and reports.

The ESET Protect Platform does not draw on AI capabilities only for detection but also for analysis and correlation of the aggregated endpoint activity.

Through its built-in correlation engine **ESET LiveCortex**¹, AI **clusters related events, filters out noise**, and **visually represents threats** in ways that help security teams understand and act swiftly. ESET LiveCortex outputs are enriched with customer-specific data and contextual details such as computer names, user accounts, and network specifics, providing highly relevant, user-specific explanations and recommendations.

This significantly improves understandability of complex data, reduces alert fatigue, and speeds up incident response in situations where minutes can mean the difference between a damaging compromise and successful defense of a system.

Trustworthy Live AI²

Pushing the innovation further, in **2024**, ESET introduced **Live AI** – a system built with agentic AI in mind, merging the best of our human expertise with the capabilities of a large language model (LLM) and other techniques including **retrieval-augmented generation (RAG)** – integrated into ESET Inspect and later added to ESET Threat Intelligence.

Being based on LLMs – a technology naturally predisposed for human-language processing and outputs – ESET Live AI autonomously creates names and descriptions for incidents to enhance security operations. This delivers immediate, intuitive threat detection, reduces false positives, streamlines management of the protected environment, and supports threat-hunting operations.

Live AI is also capable of producing reports necessary for compliance and audit, answering inquiries into specific detections in detail, investigating any of the given incidents, recommending real-time remediation steps, and even producing a link to immediately identify and isolate the affected resource.

¹ Formerly known as Incident Creator

² Formerly known as AI Advisor

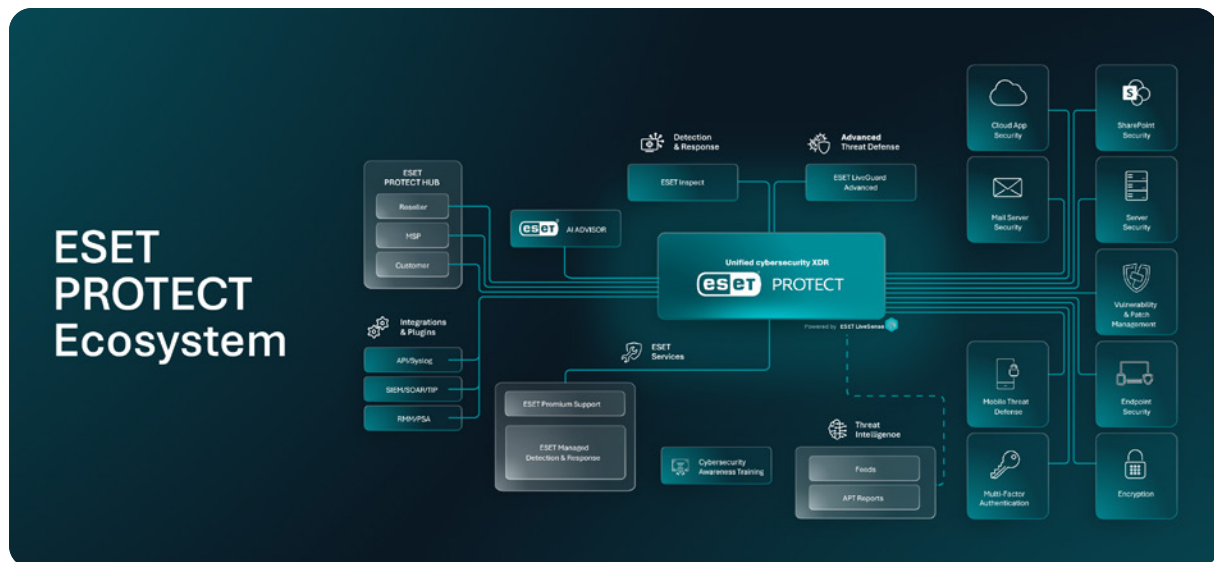


Figure 5: ESET PROTECT Ecosystem

To ensure ESET Live AI's responses are always fact-based, RAG techniques are used to search ESET-curated high-quality threat intelligence. To mitigate the risk of hallucinations, the model operates within ESET-devised and constantly maintained **guardrails** and always provides references to its sources.

Cyber threat intelligence powerhouse

Serving 500,000 business customers, 110 million direct customers, and over a billion indirect users, ESET telemetry receives a massive set of latest threat-related data daily. Cutting-edge automated systems and a team of more than 850 researchers and technology experts examine that data influx to discover novel attack techniques and tactics, making ESET a powerhouse of cyber threat intelligence.

ESET Threat Intelligence (ETI) leverages human expert experience collected over 35+ years, mixed with AI processing correlating indicators of compromise, tactics, techniques, and procedures (TTPs) to provide structured and actionable data feeds and thousands of pages of reports for defenders every year. To digest that volume of information, is also available for the ETI platform, searching and summarizing ESET research findings.

AI innovations and limits

Almost three decades of experimentation and development of AI-related security technologies have given ESET unique insights, a deep understanding of existing limits, and most of all, a broader perspective. The company acknowledges that AI technology, in itself, is neutral. The responsible deployment of AI rests in choices: keeping humans in the loop, setting robust guardrails, focusing on user benefit, and never losing sight of ethical standards.

A good example is recognizing new risks introduced by generative AI, such as deepfakes, prompt injection, or model manipulation. ESET addresses these with a practical defense-in-depth approach with main focus on the underlying elements and infrastructure behind malicious AI activity. In other words, ESET detects artifacts such as malicious scripts, phishing links, fraudulent domains, and malicious file behaviors rather than engaging in resource-intensive – and often unnecessary – “AI-versus-AI” battles.

Applying this pragmatic strategy ensures that ESET products uphold their light footprint on user system, are accessible across different environments, while retaining high detection and low false positive ratio.

This **adaptability** extends to **detection** of the latest threats, including malware written in modern **platform-agnostic languages like Go**, further demonstrating the approach where partnership with industry leaders and use of threat landscape research fuels the development of new, cutting-edge AI defense layers.

A current R&D project is also examining a range of **automation improvements** in ESET LiveCortex (part of ESET Protect Platform), where AI will not only produce actionable recommendations but ultimately provide **one-click solutions for specific incidents**, easing the workload of security practitioners monitoring the environment.

Other directions currently being explored by ESET engineers include close collaborations with hardware leaders – such as **Intel** – for integration with neural processing units (NPUs) and ongoing support for major platforms. The objective is to optimize AI-powered modules at ESET for faster and more effective performance, thus further lowering the processing power demands on customers' systems.

Build trust where there is none

ESET, as a security vendor, naturally processes large volumes of malicious content, much of it containing personally identifiable information (PII). To protect user privacy, ESET uses off-the-shelf frameworks enhanced with **proprietary filters and analyzers**, created and fine-tuned for accuracy by in-house experts. That way ESET **ensures all PII is anonymized** before any content is sent to the LLM.



Figure 6: Example of PII anonymization using proprietary filters and analyzers developed in house by ESET experts.

While highly effective at protecting PII, ESET continuously and rigorously tests this approach to ensure security of other sensitive data types – such as documents, passwords, API keys, and complex command lines containing customer-specific secrets – without compromising accuracy of its threat detection.

Yet, trust issues don't end with PII. According to in-depth interviews with selected customers from different verticals, trusting AI with handling intellectual property and other sensitive data was among the top concerns.

ESET AI Control Plane

As organizations move from experimenting with AI chatbots to deploying agentic AI, more data, decisions, and business-critical tasks are being entrusted to increasingly autonomous systems, vastly expanding the attack surface. ESET is addressing this emerging risk by scaling up its AI control plane, while keeping users and administrators in control.

ESET AI Conversation Security

The first to be deployed building block of ESET AI Control Plane focuses on risks tied to AI chatbots, including leaks of sensitive information such as passwords, API keys, technical or business know-how, or other confidential data. Such secrets can be unwittingly divulged by an employee's input, but an AI system can also pull them from an organization's sources and expose them in the output.

ESET products also scan prompts and chatbot responses for malicious links, URLs, code, or instructions that can lead to compromise, irrespective of whether such information is provided unintentionally by the user, inserted by an attacker through prompt injection, or generated by the chatbot. All relevant events will be visible to administrators in the ESET Protect Platform console.

ESET Supply Chain Security

ESET AI Skill Checker—and its enhanced version, integrated into ESET products—have been designed specifically for the emerging agentic AI ecosystem, where capabilities are increasingly extended through "Skills". These small tools or add ons instruct an agent how to perform a task, including which services to use and what actions to take.

ESET's processing of each Skill includes:

- **Scanning the Skill's content, including instructions and various artifacts such as scripts or unknown binaries**
- **Extraction and subsequent scanning of all external URLs**
- **Analysis of the full payload chain**
- **Simulation of its behavior in a sandboxed environment**

Since the system's first deployment in March 2026, ESET has scanned almost 800,000 unique Skills. Over 60,000 have been selected for further analysis, which puts 25,000 in the suspicious territory, where only minor modifications of the skill could make it harmful. Within the same timeframe, 3,000 Skills have been blocked as outright malicious.

When scans of benign Skills are added to the scans and rescans of new and updated Skills, the total number of checks conducted since March 2026 reaches well into the millions, with some days contributing hundreds of thousands of new entries to ESET's database.

But Agentic AI Skills are only part of the software supply chain, often relying on external repositories, packages, datasets, and plugins. Protecting them therefore requires going beyond traditional malware scanning that focuses on malicious artifacts such as scripts, phishing links, fraudulent domains, and malicious files. To expand this layer of protection, ESET engineers are incorporating processing of package reputation, dataset provenance, compromised maintainer accounts, and other high-risk dependencies into the outputs of ESET AI-powered solutions.

Protocol-level protection

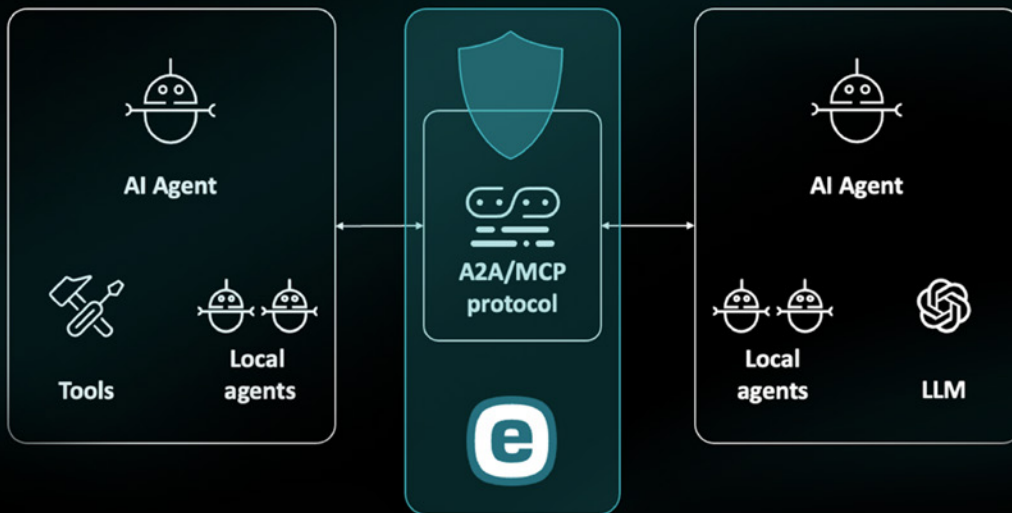


Figure 7: Network-level protection for agent to agent communication.

ESET Secure AI Relay

Scanning agentic AI Skills and chatbot inputs and outputs is only the first step in a broader system now in development at ESET: a secure relay between users, AI agents, business applications, and AI models.

ESET Secure AI Relay is designed to act as an intermediary that detects malicious patterns, enforces organizational policies, prevents sensitive data from being shared with unauthorized models, blocks AI tools from accessing restricted resources, and verifies that model outputs do not introduce malicious artifacts into business workflows.

ESET engineers are also exploring network-level protection for communication between AI agents. By protecting standard AI protocols such as the Model Context Protocol (MCP) and Agent2Agent (A2A) Protocol, ESET aims to extend the ESET XDR platform with security proxy capabilities.

Shadow AI and AI-powered asset discovery

To address the growing challenge of Shadow AI, the ESET PROTECT Platform enables AI observability and will be further extended with AI-powered asset discovery. While AI observability focuses specifically on identifying and monitoring AI usage, ESET is also exploring agent-based capabilities that can assess the broader network environment and uncover even obscure or hard-to-detect assets.

Together, these tools will identify where AI is being used, map connections between applications and business databases, analyze data flows from an attacker's perspective to assess risk, and help organizations maintain compliance with internal policies and regulatory requirements.

Advanced AI projects

ESET is also building on the correlation capabilities of systems such as ESET LiveCortex to create an AI-native intelligence ingestion platform that can identify relevant signals across a variety of sources. These include ESET technologies, OSINT data, live CVE databases, external feeds, customer-selected proprietary sources, and third-party products. The goal is to process the wealth of intelligence into a prioritized list of security issues, helping security teams focus on what matters most.

Finally, to maintain independence from large AI providers and deepen its understanding of both the opportunities and limitations of AI in cybersecurity, ESET is investing in its own specialized foundation models.

This project will include, but also extend beyond, transformer-based and LLM architectures, with these models helping organizations not only manage existing security issues but also simulate security scenarios and evaluate the future impact of emerging threats and defensive strategies.

AI SOC

ESET's R&D also focuses on a new generation of **AI security operations center (SOC)** technology. Premium security services—such as incident response, compliance audits, and penetration testing—are expensive and in limited supply, making them accessible primarily to large enterprises. At the same time, mid-sized and smaller organizations often lack the expertise and resources to benefit from such capabilities, even as the complexity of threats, telemetry, and security tooling continues to grow.

The rapidly growing scale, complexity, and speed of both AI-driven threats and security operations call for a new approach. AI-powered SOC's can automate significant portions of cybersecurity workflows, helping to relieve experienced human analysts from alert fatigue and enabling them to focus on the highest-priority strategic issues. This has the potential to make enterprise-grade protection more affordable and accessible to a broader range of organizations and users, while making these solutions effortless to operate.

Predict, prevent, protect

While exploring new defense opportunities, ESET is also aware of potential upcoming challenges related to the use of the latest AI models by threat actors. These include:

- **Optimized attack techniques and malware.**
- **Improved social engineering.**
- **Increased volume of malicious scam, phishing, and disinformation campaigns.**
- **Vulnerability discovery for malicious purposes.**
- **Identification of vulnerable individuals on social media and their manipulation into different "supportive" roles in cyberattacks (money mules, distribution, processing power for DDoS, etc.).**

- **Training small, open-source models to process stolen personal information and generate tailored scams.**

At ESET, use of AI has been defined by early adoption, persistent scientific rigor, and always underpinned with deep understanding of the trends and directions in the threat landscape. Blending these with advanced automation, careful human supervision, and emphasis on user-centricity at every turn, ESET offers a model for how to leverage AI's capabilities for the good of all. In doing so, ESET helps ensure that defenders always remain a step ahead, using an intelligent, adaptive, and, above all, secure and trustworthy set of security tools.

ESET AI philosophy

Over the past three decades, ESET has established itself as a leader in AI-powered cybersecurity by prioritizing safety, transparency, and user benefit without compromise, while being grounded in deep technical expertise, responsible innovation, and an understanding of the high stakes of securing digital environments.

The company is attuned to the social and ethical responsibilities of AI, favoring practical methods that block attack vectors to protect large numbers of users rather than chasing unreliable detection of every potential threat individually.

ESET employs a sophisticated, multilayered integration of AI models – combined with traditional detection methods and human oversight – to ensure that decisions consider context and user impact, rather than relying solely on automated systems.

Tools like ESET Live AI empower teams to manage the security of a system, regardless of their level of technical expertise or the complexity of the environment they're managing, contributing to an effective human–AI collaboration.

Ongoing investment into R&D and practical technology ensures that ESET technology continually adapts to emerging threats and continues to benefit users ranging from individuals to large enterprises.

About ESET

ESET® provides cutting-edge cybersecurity to prevent attacks before they happen. By combining the power of AI and human expertise, ESET stays ahead of emerging global cyberthreats, both known and unknown—securing businesses, critical infrastructure, and individuals. Whether it's endpoint, cloud, or mobile protection, our AI-native, cloud-first solutions and services remain highly effective and easy to use. ESET technology includes robust detection and response, ultra-secure encryption, and multifactor authentication. With 24/7 real-time defense and strong local support, we keep users safe and businesses running without interruption. The ever-evolving digital landscape demands a progressive approach to security: ESET is committed to world-class research and powerful threat intelligence, backed by R&D centers and a strong global partner network. For more information, visit www.eset.com or follow our [social media](#), [podcasts](#), and [blogs](#).

ESET Cybersecurity | Enterprise, Business and Home Solutions | ESET

Best IT security solutions for your home and business devices. Try ESET antivirus and internet security solutions for Windows, Android, Mac or Linux OS.