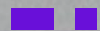


WHITEPAPER

Advancing Cybersecurity: Leveraging GenAI and LLMs for Enhanced Security Operations



Authors:

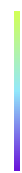
Ryan Mack, Uptycs

Arijit Bandyopadhyay, Intel Corporation

uptycs 

intel®

Table of Contents



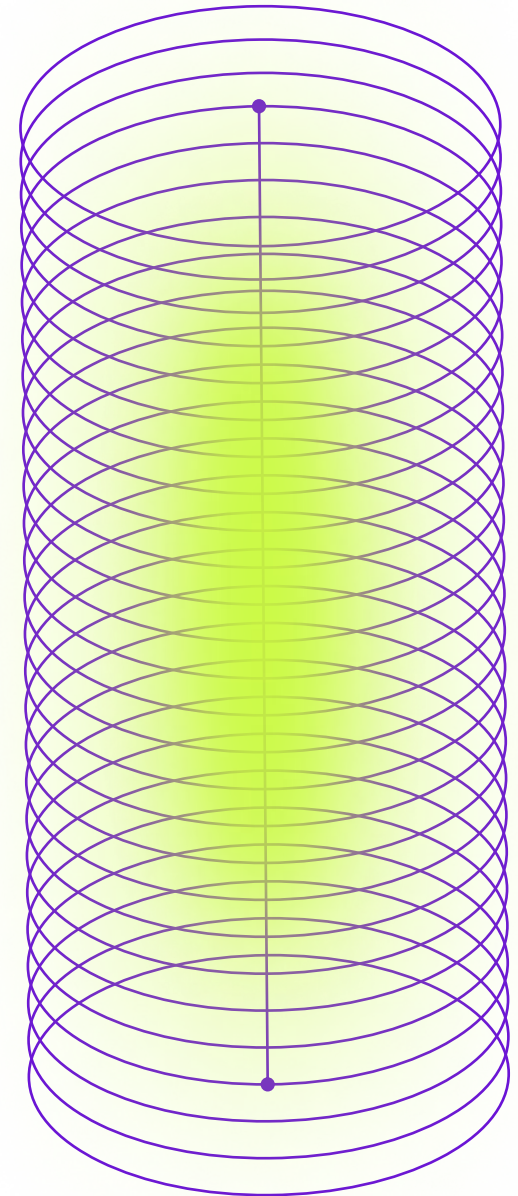
Overview	01
Leveraging LLMs and GenAI for Enhanced Security Operations	02
Securing the Deployment and Use of Large Language Models	04
Enhancing Cybersecurity with LLMs and GenAI at Uptycs	06
Enhancing Cybersecurity with the Intel Habana Gaudi Platform	11
Security in Intel Platforms	12
Better Together	13
Concluding Thoughts	14
Additional Resources	15

Overview

In the evolving landscape of cybersecurity, leveraging advanced technologies like Generative AI (GenAI) and Large Language Models (LLMs) is becoming increasingly critical for enhancing security operations across various domains. This white paper focuses on four pivotal areas where these innovations are making significant impacts:

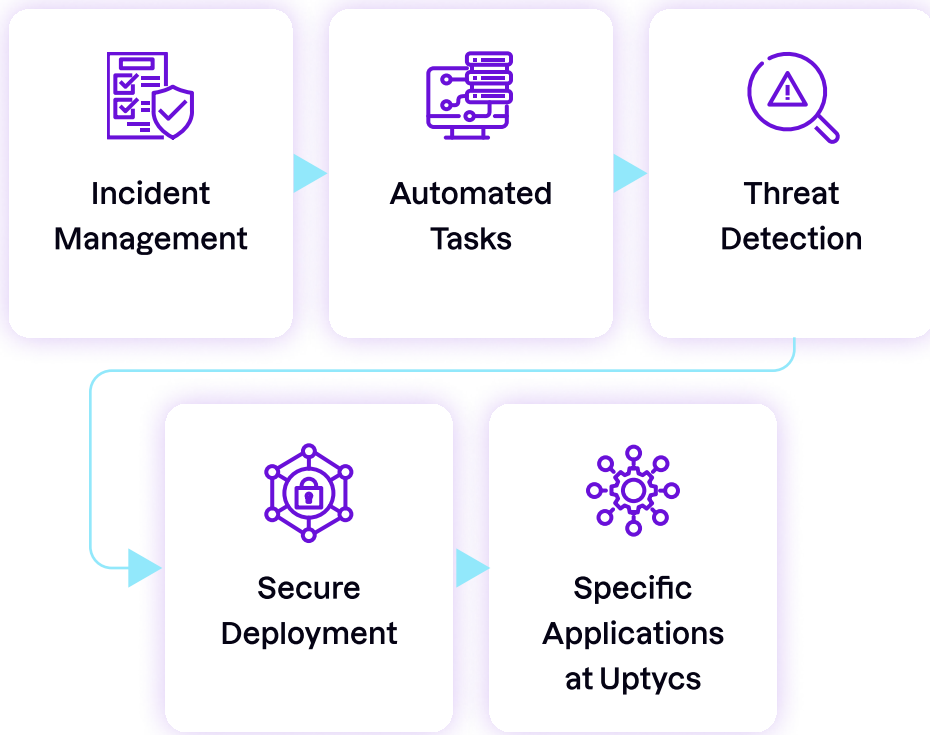
1. How security companies are harnessing LLMs and GenAI to improve incident management, automate complex tasks, and enhance threat detection and response capabilities.
2. The challenges and strategies involved in securely deploying and using LLMs, focusing on performance evaluation, real-time monitoring, and rigorous pre-deployment testing.
3. The specific applications of LLMs and GenAI at Uptycs, highlighting how they streamline operations within their Cloud Native Application Protection Platform (CNAPP) and Extended Detection and Response (XDR) systems.
4. The role of the Intel Gaudi platform in powering LLMs for security applications, emphasizing its optimization for deep learning tasks crucial for maintaining robust security frameworks.

Through these discussions, the paper aims to provide comprehensive insights into how cutting-edge AI technologies are reshaping the security landscape, driving efficiency, and bolstering defenses against emerging threats.



Leveraging LLMs and GenAI for Enhanced Security Operations

Security companies are LLMs to address several critical challenges in the security landscape, both in managing the security of cloud infrastructure and productivity endpoints. These challenges include the need for more efficient incident management, the automation of repetitive and complex security tasks, enhanced threat detection and response, and comprehensive security validation.



One significant problem is the complexity and volume of security incidents that need to be managed effectively. The use of generative AI, as seen with Microsoft Copilot for Security, helps to distill complex security alerts into concise, actionable summaries. This assists security professionals in quickly understanding and prioritizing incidents, thereby reducing response times and improving decision-making processes. Additionally, integrating AI-driven impact analysis streamlines incident handling, allowing security teams to focus on the most critical threats.

Another major challenge is the automation of repetitive and tedious tasks that often bog down security teams. Solutions like Charlotte AI by CrowdStrike tackle this issue by automating data collection, basic threat detection, and advanced security actions such as threat hunting and remediation. By enabling security professionals to interact with the system through natural language queries, these tools make complex workflows more accessible and efficient, thereby enhancing the overall capabilities of security teams.

The need for continuous and thorough security testing, which is resource-intensive and challenging to perform regularly, is addressed by platforms like RunSybil. By automating penetration testing and red teaming workflows, these platforms allow organizations to identify vulnerabilities and simulate adversarial attacks more frequently and thoroughly. This continuous security validation improves the resilience of security defenses and enables smaller security teams to achieve high levels of thoroughness.

Additionally, automating administrative tasks such as completing third-party security questionnaires and generating content for security awareness training streamlines compliance processes and enhances security awareness among employees.

Managing an overwhelming number of security alerts and ensuring thorough investigations is another critical challenge. Dropzone AI addresses this by deploying autonomous AI agents to investigate various security alerts, thereby reducing the workload on human analysts and ensuring consistent, thorough investigations.

Integrating with existing security tools, these AI agents provide comprehensive analyses and detailed reports that help security teams prioritize and address threats more effectively. This approach enhances the efficiency of Security Operations Centers (SOCs) and allows human analysts to focus on higher-value tasks, thus improving the overall effectiveness of security operations.



Securing the Deployment and Use of Large Language Models

The rapid advancement and integration of LLMs have catalyzed transformative changes across various industries by enhancing artificial intelligence capabilities. However, this proliferation also poses unique challenges, particularly in securing these powerful tools to ensure their reliable and safe usage. Addressing these challenges encompasses several critical areas including performance evaluation, real-time monitoring, stress testing before deployment, and stringent validation measures within AI applications.

Evaluating the performance of LLMs in real-world scenarios is crucial to ensure their effectiveness and safety.

Patronus.ai provides a comprehensive solution for this by scoring LLMs on various performance metrics, identifying incorrect or fabricated responses (hallucinations), and detecting potential data leaks. This evaluation is supported by advanced adversarial testing techniques that generate challenging test cases, revealing potential vulnerabilities in AI models. By benchmarking these models against one another, Patronus.ai aids enterprises in selecting the best-suited models for specific applications, ensuring that the deployed LLMs are both reliable and secure.

Monitoring the usage of LLMs within organizational frameworks is another critical security measure. Lasso Security specializes in this domain by offering tools that track how LLMs are deployed and utilized within a company. This includes Shadow AI Discovery, which provides insights into which models are active and how they are being applied by employees. By logging interactions with LLMs and monitoring data flows, Lasso Security can detect unusual or risky activities, swiftly responding to threats such as data breaches, prompt injections, or malicious code through real-time alerting.

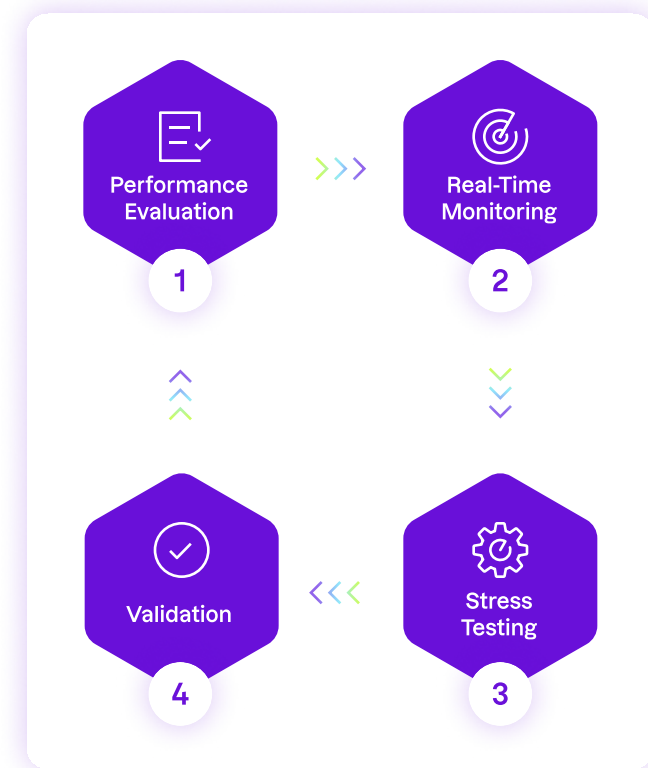
Another essential aspect of securing LLMs involves protecting their usage through web browsers. Prompt Security excels in this area by offering browser extensions that safeguard interactions with LLMs, ensuring that every prompt and response is scrutinized for potential threats such as prompt injection, data leakage, and inappropriate outputs. By integrating seamlessly into the browser environment, Prompt Security provides real-time protection and immediate alerts for any suspicious activities, thereby enhancing the security of LLMs used across organizational platforms.

Pre-deployment stress testing is also vital to ensure LLMs are robust against attacks before they go live. Lakera Security's offerings, such as Lakera Guard and Lakera Red, focus on providing runtime protection against common threats like prompt injections and data loss, while also conducting rigorous stress tests to identify and mitigate potential vulnerabilities ahead of deployment. This dual approach ensures that LLMs are fortified against both current and emergent threats, maintaining integrity and compliance throughout their lifecycle.

Direct implementation of validation and security measures within AI applications is crucial for maintaining data integrity and compliance. Guardrails AI addresses this by developing input/output guards that ensure LLM outputs adhere to strict safety and reliability standards. This includes checking for compliance with regulations, preventing toxic language, and avoiding data leaks. Additionally, Guardrails AI's structured data generation tools help maintain the accuracy and reliability of LLM outputs, providing a robust framework for preventing common issues like prompt injections and data breaches.

Together, companies like Patronus.ai, Lasso Security, Prompt Security, Lakera Security, and Guardrails AI demonstrate a multi-faceted approach to securing LLMs.

By addressing performance evaluation, real-time monitoring, pre-deployment testing, and integrated validation, these solutions are pivotal in managing the unique threats and vulnerabilities associated with LLM usage. Their innovative measures ensure that these potent AI tools are deployed safely and effectively across various sectors, maintaining trust and reliability in AI-driven operations.



Enhancing Cybersecurity with LLMs and GenAI at Uptycs

Detection - Process using crontab utility to add entries in...

ask uptycs

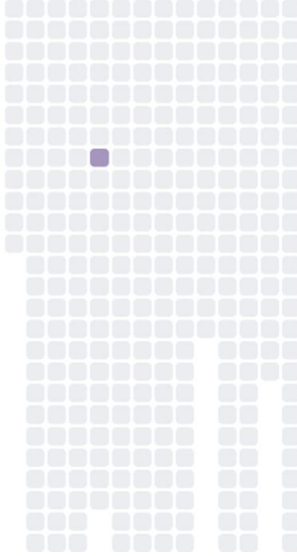
10/10
! 14 Alerts
8 Tactics
Advanced Threat
02/20/2024 00:11:05.000
ACTIVITY
UNASSIGNED

13 Events
11 Techniques
None
02/20/2024 00:26:05.000

SIGNALS
DETECTION GRAPH

ATT&CK Matrix

I E R P P R D O D L C C E I



27 signals
 Group

Showing
All

Sort by
Time

Search

Clear filters

Summary (AI-generated from Uptycs Insights)

Following activities detected on your System:

- Identified a bash or sh shell script execution that may have originated from sshd, apache2, httpd, or nginx processes.
- Captured file changes in known cron directories excluding expected system maintenance processes.
- Flags any instance where the 'curl' utility is used to download files directly into the '/tmp/' directory, circumventing typical update mechanisms and potentially bypassing security controls. It filters out benign system update processes that also use such commands.
- Identified the execution of psexec.py on Linux, signaling possible lateral movement as per the MITRE ATT&CK technique T1021.002.

At Uptycs, leveraging GenAI and LLMs has significantly enhanced the capabilities of our Cloud Native Application Protection Platform (CNAPP) and Extended Detection and Response (XDR) systems. These technologies help us tackle critical security challenges in large hybrid cloud environments, enabling faster and more precise responses to potential threats.

In the realm of security, the swift comprehension of complex information is paramount. Through the use of GenAI, Uptycs enhances this aspect within our CNAPP and XDR offerings by summarizing vast volumes of security telemetry into natural language.



Scale Security with Gen AI

Embedding Gen AI in security operations lifecycle

- Detection and Vulnerability Summarization for faster triaging and response
- Asset Discovery & Visibility

Gen AI based risk prioritization

- Predictive vulnerability assessment
- Data classification



Protect AI data and Infrastructure

Secure AI workloads

- Gain Visibility into GPU workloads
- Build behavioral and anomaly detections to protect AI workloads

Posture management of training and inference workloads including:

- Kubeflow
- Kserv



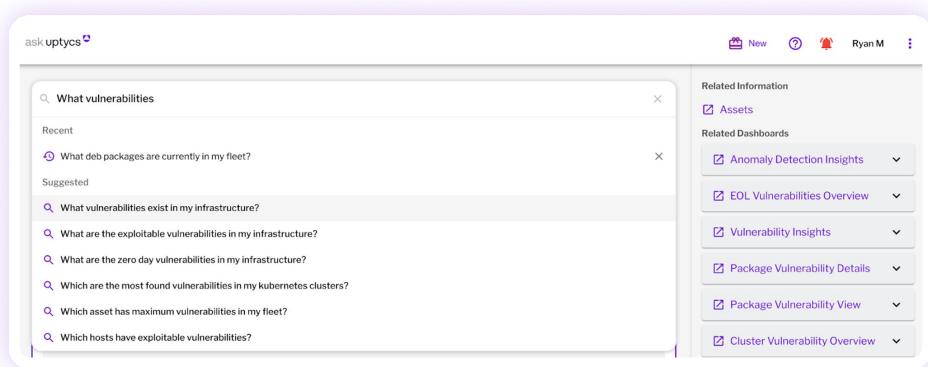
AI Ecosystem Integration

Gain security context, better correlation through Microsoft Co-pilot integration

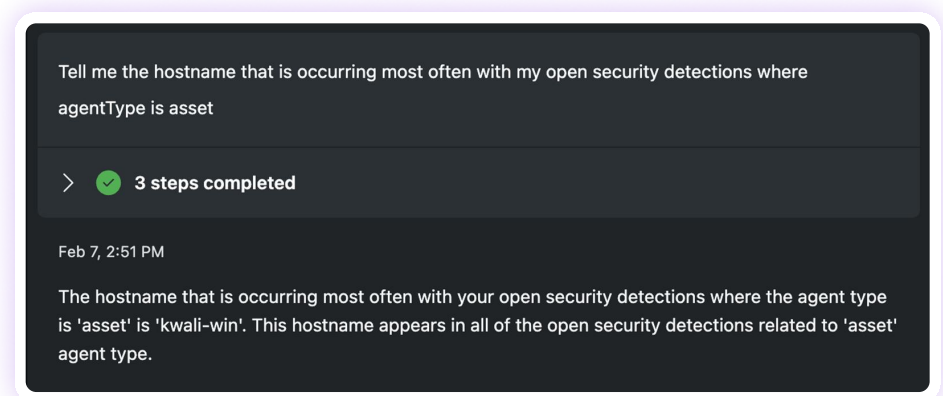
Guided remediation suggestions through Co-integration

This allows our clients to quickly understand the nature of detected threats, thereby accelerating decision-making and response processes. By integrating advanced AI to distill detailed alerts into comprehensible summaries, we not only improve response times but also enhance the overall efficiency of security operations.

Efficient knowledge retrieval is another crucial area where Uptycs employs GenAI. Our platform utilizes AI-driven search tools that help security professionals navigate through extensive data, whether they're managing inventory, detecting threats, or ensuring compliance. By embedding GenAI into our user interfaces, such as the "Ask Uptycs" feature, we streamline the discovery process, making it easier for users to access and understand security information.




Recognizing the challenge of hiring and retaining experienced security practitioners, Uptycs incorporates security-focused GenAI models to provide deep security insights and real-time situational awareness. These models, enhanced by Retrieval-Augmented Generation (RAG), offer instant analysis of common exploitation tactics, integrating seamlessly with our MITRE ATT&CK-based detection engine and vulnerability management systems. This strategic use of GenAI enriches the skill set of security teams, equipping them with AI-augmented tools to better understand and react to threats.





Uptycs is also pioneering the development of GenAI plugins and promptbooks that facilitate complex security investigations. These tools allow seamless interaction with data through APIs, enabling users to query and analyze security data using natural language. This approach simplifies the investigation process, allowing for more agile and comprehensive security analyses and actions.

Moreover, the rapid deployment of GenAI technologies raises concerns regarding data security and potential software vulnerabilities. Uptycs addresses these challenges head-on with robust measures in our vulnerability management platform, which deeply scrutinizes AI toolchains and models for potential risks.

This comprehensive security strategy ensures the protection of sensitive data and AI operations, aligning with SOC2 compliance standards to safeguard our clients' interests.

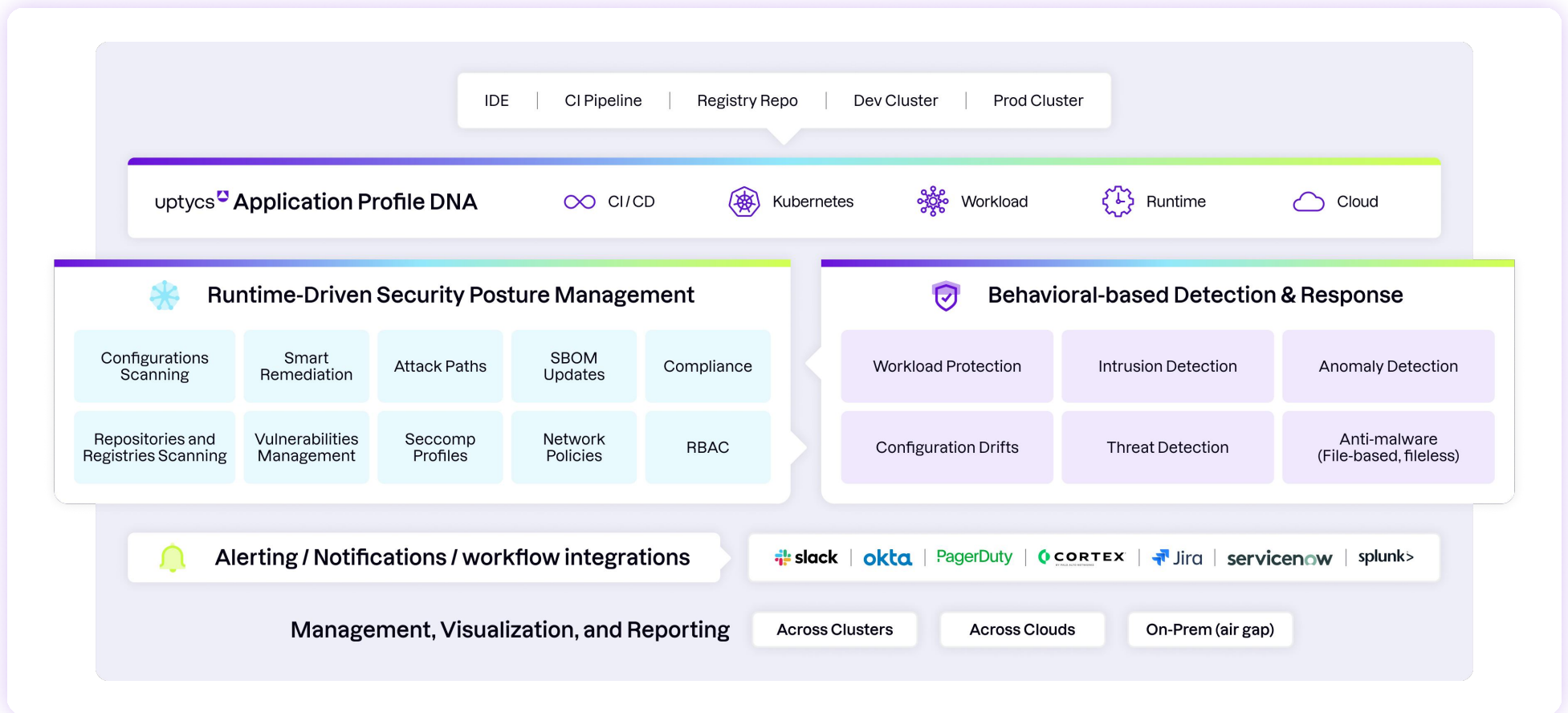
 Welcome to Uptycs Chat

summarise T1105 

-  • **T1105 mitre technique:** Adversaries can transfer tools or files from an external system into a compromised environment using various methods, including the command and control channel, alternate protocols, Web services, and native tools on the victim system.
- **Sub-techniques:** + Using an external Web service to host information that points to additional C2 infrastructure + Sending commands to and receiving output from a compromised system over a Web service channel + Sending commands to a compromised system without receiving return output over a Web service channel + Obfuscating traffic in Content Delivery Networks (CDNs) and other services to hide the intended destination of HTTPS traffic or traffic tunneled through HTTPS
- **Mitigations:** + Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary malware or unusual data transfer over known protocols like FTP.

By integrating GenAI across various facets of our security platform, Uptycs sets new industry standards for rapid and informed responses to security threats. This integration not only accelerates the understanding and mitigation of security alerts but also significantly enhances our ability to navigate and utilize vast amounts of security data effectively.

As GenAI continues to evolve, it becomes an indispensable asset in the future of cybersecurity, promising more effective and proactive defense mechanisms against emerging threats.



Enhancing Cybersecurity with the Intel Habana Gaudi Platform

The Intel Habana Gaudi Accelerator platform, <https://habana.ai> stands out as a pivotal technology specifically engineered to optimize and accelerate the training and inference of LLMs and deep learning models, including those integral to security applications. This specialized platform is increasingly crucial in addressing the dynamic and complex challenges posed by modern cybersecurity needs.

The Intel Habana Gaudi processors offer exceptional computational performance, crucial for the intensive demands of training, fine-tuning and inference of LLMs for security applications. This high-level performance and scalability is essential for processing large volumes of security telemetry information swiftly and generating actionable insights, which are critical for real-time threat detection and focused response. The ability of Gaudi to efficiently manage these tasks ensures that security platforms can rapidly analyze and condense complex security alerts into succinct summaries. This capability significantly aids in speeding up decision-making processes, allowing security professionals to respond more quickly and effectively to potential threats.

A key advantage of the Gaudi platform is its scalability, supporting everything from single processors and system to extensive clusters.

This scalability is particularly beneficial for major security platforms which require robust performance to manage extensive and growing volumes of data and increasingly sophisticated and complex LLM and multi-modal models. As a result, these platforms can sustain high levels of efficiency and effectiveness, even as their data processing demands grow.

The Gaudi platform also excels in cost efficiency. By optimizing both performance and energy efficiency, Gaudi processors help reduce the total cost associated with training, inference, and scalable deployment of AI models. This cost-effectiveness makes advanced security features like continuous penetration testing and automated threat detection more accessible and sustainable for organizations of all sizes. Furthermore, Gaudi's compatibility with standard AI frameworks such as [PyTorch and TensorFlow](#) enhances its value, enabling security solutions to leverage this powerful technology without the need for extensive changes to their existing software infrastructure.

Included with the Intel Habana Gaudi platform is the Habana SynapseAI® Software Suite, <https://habana.ai/intel-gaudi-software/> which provides a range of tools, libraries, and frameworks specifically optimized for deep learning tasks.

This suite ensures that security applications can fully utilize the capabilities of the Gaudi processors, improving the efficiency of key processes like incident summarization, threat detection, and response prioritization. Additionally, AI-specific accelerators within the Gaudi processors are designed to handle the unique computational patterns of security-related LLMs, significantly boosting the performance of these vital operations.

Overall, the [Intel Gaudi platform](#) markedly enhances the capabilities of LLMs within security applications, enabling faster, more efficient, and cost-effective security operations. By addressing crucial challenges such as incident management, automation of repetitive tasks, advanced threat detection, and comprehensive security validation, the Intel Habana Gaudi platform ensures that security solutions can keep pace with the rapidly evolving landscape of cyber threats, providing robust protection and proactive defense mechanisms.

Security in Intel Platforms

Intel's Xeon platforms incorporate advanced security features to enhance overall system protection. These features include Intel Confidential Computing, Intel Tiber Trust and Security Solutions, and Intel Security Engines, which provide robust measures for data security and privacy.

- **Intel Confidential Computing:** This technology ensures data remains secure and confidential during processing. It provides a hardware-based security foundation that isolates sensitive data and workloads, minimizing the risk of exposure or tampering.

- **Intel Tiber Trust and Security Solutions:** This suite offers comprehensive software-based security solutions designed to protect data integrity and confidentiality across various computing environments. These solutions are essential for ensuring secure interactions and data exchange.
- **Intel Security Engines (Xeon):** Built into Intel Xeon processors, these security engines deliver robust protection for sensitive data while maintaining high performance. They enable secure boot, cryptographic functions, and enhanced data protection.

These technologies collectively enhance the security capabilities of Intel's platforms, making them an integral part of comprehensive, scalable, and secure AI-driven solutions.

Intel® Gaudi® 3 x Uptycs

Better Together

As the landscape of artificial intelligence (AI) evolves at an unprecedented pace, the collaboration between Uptycs and Intel® Gaudi® addresses critical security concerns associated with AI deployments. As businesses adopt GenAI technologies, they must safeguard their most sensitive data—their “Crown Jewels.” The shift to enterprise deployment brings heightened risks of software vulnerabilities and misconfigurations, which can lead to data loss. Additionally, open-source toolchains and models, though beneficial for machine learning tasks, often lack visibility into the software Bill of Materials (SBOM), creating further security challenges.

Uptycs tackles these issues with a multifaceted approach to security and visibility in AI environments. Their vulnerability management platform provides deep introspection into AI toolchains and publicly available models, allowing businesses to quickly identify and prioritize the patching or removal of vulnerable software components. The Uptycs Cloud Discovery UI enables users to visually navigate their AI cloud infrastructure, helping them understand the relationships between AI and data resources.

This tool incorporates a Risk Rating system, enabling businesses to prioritize Data Loss Protection (DLP) efforts based on critical factors like network reachability, IAM configurations, and S3/object storage contents.

To detect and mitigate threats, Uptycs uses an eBPF-based agent and cloud log analysis to swiftly identify abnormal behaviors, including privilege escalation attempts and data exfiltration. The integration of Intel® Gaudi® 3 accelerators enhances the security and efficiency of AI workloads. By collecting telemetry from the hardware management APIs of GPUs and AI accelerators, Uptycs can detect malicious or unexpected usage patterns, protecting valuable GPU and AI accelerator resources from unauthorized access and misuse.

The partnership between Uptycs and Intel® Gaudi® exemplifies a comprehensive approach to AI security, combining advanced technology with robust security measures. This collaboration addresses current AI deployment challenges and paves the way for a more secure and efficient future in AI innovation. Businesses can confidently adopt.

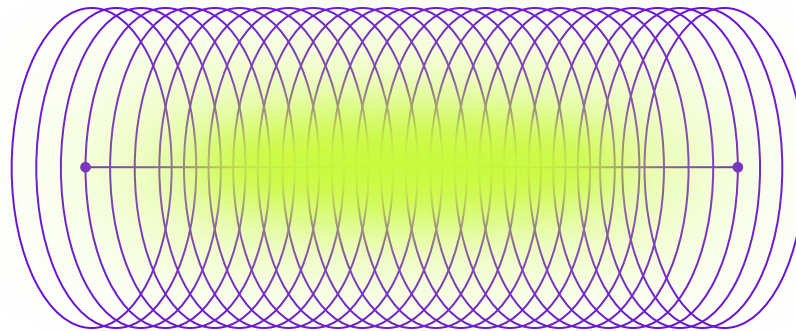
Concluding Thoughts

This white paper has highlighted the transformative impact of Generative AI and LLMs across various facets of cybersecurity, illustrating their pivotal role in enhancing security operations, from incident management to advanced threat detection. The integration of these technologies enables security companies to tackle complex challenges more efficiently, streamlining workflows and enhancing the overall security posture.

As we have explored, solutions like Uptycs, Microsoft Copilot for Security, and Charlotte AI by CrowdStrike demonstrate the power of GenAI to automate and enhance security processes, while platforms like RunSybil and Dropzone AI leverage these technologies for continuous security testing and autonomous investigations, respectively.

Moreover, the paper delved into the crucial aspects of securing the deployment and use of LLMs, with companies like [Uptycs](#), [Patronus.ai](#), [Lasso Security](#), and [Lakera Security](#) at the forefront, ensuring that these advanced models are deployed securely and effectively.

The Intel Gaudi platform's role in powering these LLMs for security applications underscores the technological advancements that are critical to maintaining robust and scalable cybersecurity solutions. As we look forward, the continued evolution and integration of GenAI and LLM technologies will undoubtedly play a central role in shaping the future of cybersecurity, promising enhanced capabilities and more proactive defenses against the growing spectrum of cyber threats.



Additional Resources

Intel

- [Intel Habana Gaudi Website](#)
- [Intel Habana Gaudi Platform](#)
- [Intel Habana Products](#)
- [Intel Habana Gaudi Architecture and Software Suite](#)
- [Gaudi Developer](#)
- [Gaudi and PyTorch](#)
- [Intel OpenVINO](#)
- [Intel Neural Compressor](#)
- [Intel Tiber Developer Cloud](#)
- [Intel Habana Blog](#)
- [Huggingface Optimum for Intel Gaudi](#)
- [Intel Confidential Computing](#)
- [Intel Tiber Trust and Security Solutions](#)
- [Intel Security Engines \(Xeon\)](#)

Uptycs

- [Manage Security Risks Across Your AI's Entire Lifecycle](#)
- [CNAPP Buyer's Guide](#)
- [Risk Prioritization Is Not Enough](#)
- [Kubernetes and Container Security](#)
- [Automated Vulnerability Scanning for AWS, Azure, GCP](#)
- [Cloud Workload Protection Platform](#)
- [Cloud Compliance Solutions for AWS, Azure & Google Cloud](#)



Uptycs is the leading cloud security platform for large hybrid cloud environments. We extend security visibility from development to runtime, ensuring consistent protection and compliance across the application infrastructure. That's why enterprises like PayPal, Comcast, and Nutanix rely on Uptycs to secure the development ecosystems they use to build their applications and run their workloads.

On Guard. On Cloud. On Loop.

[Learn more at Uptycs.com](https://Uptycs.com)