Emerging Tech: Domain-Specific LMs Will Drive Agentic AI and Autonomous Security

8 August 2025 - ID G00837733 - 13 min read

By: Esha Bhatia, Srushti Bhakta

Initiatives:Emerging Technologies and Trends Impact on Products and Services

> Product leaders must adopt domain-specific language models that deliver autonomous, precise threat response and maintain AI trust. This shift is not optional — it's a competitive imperative for transforming security into a proactive, intelligent and resilient function.

Overview

Key Findings

- General-purpose models lack the precision required for cybersecurity operations, often delivering generic or off-target responses. C-level executives leading AI initiatives that fail to pivot to domain-specific language models (DSLMs) risk falling behind industry peers who leverage these models to proactively defend and adapt to evolving threats.
- DSLMs will enable product leaders to drive a paradigm shift toward autonomous security operations, advanced threat reasoning and intelligent multiagent collaboration. Those who fail to embed DSLMs will lag behind competitors who are rapidly accelerating innovation and resilience through targeted integration.
- Persistent concerns regarding the reliability, potential for hallucination, inherent biases and data privacy risks of language models are significantly impeding their full adoption and trust in critical cybersecurity operations. Product leaders who lead in establishing trusted AI will be best-positioned to capture emerging revenue opportunities and outpace competitors.

Recommendations

- Refocus product roadmaps by prioritizing the integration of DSLMs trained on curated cybersecurity data and processes. This can be done by embedding these models into specific use cases — such as malware analysis, vulnerability management, incident classification, phishing detection or signal triage — to drive autonomous, accurate, enterprise agility and scalable security operations.
- Invest in multiagent frameworks by developing specialized AI agents powered by DSLMs that can collaborate intelligently across detection, investigation, mitigation and reporting. This approach drives greater autonomy and operational efficiency, and expands coverage across critical domains such as cloud security, data loss prevention and posture management.
- Establish trust in DSLMs by implementing robust architectures, enforcing strict data governance, and continuously monitoring for bias and hallucinations. These safeguards are essential to ensure reliability, protect privacy and accelerate enterprisewide adoption of DSLMs in security workflows.

Analysis

Technology Description

The cybersecurity landscape is increasingly complex, with modern enterprises facing an overwhelming influx of alerts and sophisticated threats that frequently exceed the capacity of traditional security operations centers (SOCs). As per our research study with the vendors, manual processes, which may consume up to 60% of an analyst's time, lead to alert fatigue, slower response times and the risk of critical incidents going undetected. Conventional security orchestration, automation and response platforms often offer only partial integration and rely on static playbooks, failing to provide the dynamic adaptability and comprehensive efficiency needed to manage evolving threats. Furthermore, general-purpose large language models (LLMs), despite their broad capabilities, are proving to be resource-intensive and are prone to "hallucinations" or imprecise output when applied to sensitive, domain-critical cybersecurity tasks, raising significant concerns about reliability and data privacy in real-world deployments.

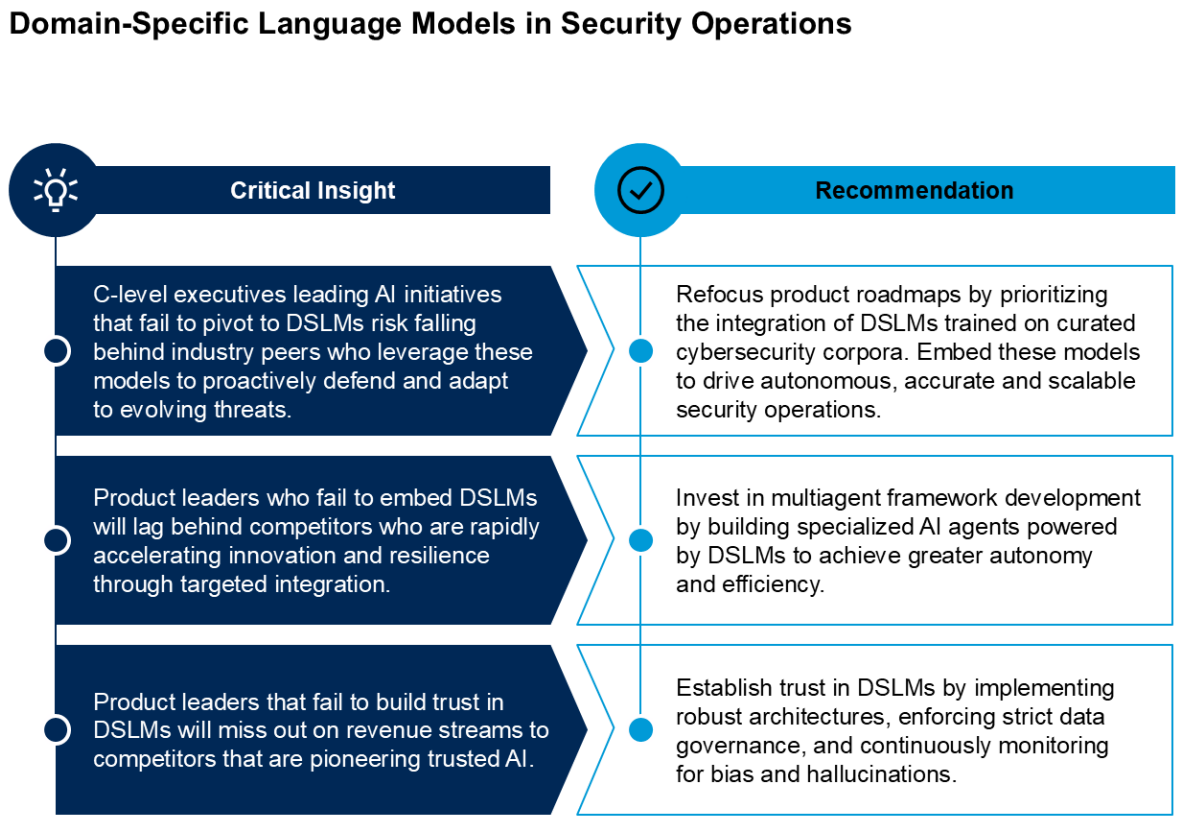
To counter these growing challenges, DSLMs are emerging as a pivotal innovation. Unlike their general-purpose counterparts, DSLMs are meticulously trained on vast, curated cybersecurity datasets, including common vulnerabilities and exposure reports, exploit prediction scoring systems and known exploited vulnerabilities, vendor API documentation, threat intelligence feeds, and real-world SOC playbooks. This specialized training enables DSLMs to exhibit a deeper contextual understanding, resulting in higher accuracy and faster inference speeds tailored specifically for security operations. Crucially, DSLMs are sometimes designed to be smaller and more resource-efficient, facilitating flexible deployment options, including on-premises, thereby addressing critical data sovereignty concerns and operational constraints prevalent in highly regulated industries.

This paradigm shift extends to agentic AI platforms, which intricately integrate DSLMs to provide autonomous capabilities — reasoning, planning and executing actions within live security environments. These platforms aim to bridge the gap between human intuition and machine speed, enabling automated triage, in-depth investigation and rapid response mechanisms. A core objective is to consolidate disparate security tools and augment human analysts, thereby freeing them from repetitive tasks and allowing them to focus on complex, high-value strategic investigations.

Innovations such as Knowledge-Augmented Generation (KAG) frameworks further enhance their capability by structuring retrieved content into knowledge graphs, ensuring that DSLM outputs are firmly grounded in verifiable, context-rich data and significantly minimizing the risk of hallucinations.

Figure 1 outlines critical insights and actionable recommendations for using DSLMs in security operations.

Figure 1: Domain-Specific Language Models in Security Operations



Source: Gartner
837733

Gartner

Critical Insight: General-Purpose LLMs Lack the Requisite Precision for Critical Cybersecurity Tasks, Necessitating Domain-Specific Models

Near-Term Implications for Product Leaders

The inherent precision of DSLMs in cybersecurity contexts means product offerings can achieve significantly higher accuracy in tasks ranging from alert triage to forensic analysis. This capability translates into reduced false positives, faster mean time to detect (MTTD), and more efficient investigations and exposure reduction, freeing up human analysts to focus on complex strategic tasks. Platforms leveraging DSLMs can offer more effective noise reduction, smart alert grouping and robust behavioral correlation, fundamentally transforming operational efficiency within the SOC. DSLMs enhance the entire investigation life cycle, from initial grouping to proactive threat management.

Without specific training on security data structures (such as security alerts, vulnerabilities, logs, tickets and policy artifacts), general LLMs struggle to interpret structured and unstructured security data with the necessary precision for real-time reasoning.

The demand for precision in cybersecurity mandates product leaders to pivot their AI strategies away from generalized LLM solutions toward DSLMs. Providers are now focused on building or integrating specialized model capabilities that can deeply understand and respond to the nuances of security data.

Recommended Actions for the Next Six to 18 Months

- Refine product roadmaps by prioritizing DSLM integration and focusing on embedding specialized models trained on cybersecurity corpora for tasks requiring high precision, such as malware analysis, vulnerability management, incident classification, true positive/false positive detection and email consolidation for phishing detection.
- Enhance feature sets with domain-tuned AI by developing capabilities that leverage DSLMs for artifact interpretation, behavioral correlation to MITRE ATT&CK and hypothesis testing. This will deepen analytical precision, thereby improving the depth and accuracy of investigations and reducing MTTD and mean time to respond.

Critical Insight Analysis

General-purpose LLMs, despite their broad capabilities, often produce generic or even irrelevant responses when applied to the highly specialized domain of cybersecurity. This limitation stems from their training on vast, diverse datasets that do not adequately capture the unique vocabulary, attack patterns and contextual nuances of security incidents. For instance, a general LLM might struggle to precisely interpret a specific registry change as an indicator of compromise (IoC) or correlate low-level system events with a specific MITRE ATT&CK technique.

DSLMs are specifically fine-tuned on curated cybersecurity data, enabling them to understand the intricate relationships between disparate security signals, identify malware families and interpret complex threat intelligence with remarkable accuracy. This domain-specific focus allows DSLMs to minimize anomalies, ensure data consistency and provide reliable, actionable insights that general models cannot, directly addressing issues like high dimensionality and sparsity in security data.

Imperum, for example, develops its DSLMs using knowledge distillation from larger models, quantization-aware training and techniques like low rank adaption (LoRA)/quantized low rank adaption (QLoRA) to optimize performance and deployment flexibility, resulting in models that are compact, hardware-efficient and intellectually enriched with threat behavior patterns. They also utilize knowledge-augmented generation to anchor responses in verifiable, context-rich data, dramatically lowering the rate of fabricated steps and enhancing the processing of both structured and unstructured data.

Similarly, NSFOCUS GPT (NSFGPT, Chinese name Fengyunwei) operates as a security-focused AI platform, relying on DSLMs to build intelligent agents capable of perceiving environments, understanding, planning and executing complex security tasks, particularly within security information and event management (SIEM) and SOC solutions.

Conifers' CognitiveSOC platform further illustrates this need by integrating various AI/machine learning techniques, including small language models, embedded with deep domain expertise to solve complex detection and response challenges with humanlike investigative depth.

IBM recognized the need for specialized models tuned to cybersecurity contexts to achieve the required precision and effectiveness in their operations. IBM utilizes DSLM models in several of its core cybersecurity services and platforms, including autonomous threat operations machine (ATOM), advanced threat disposition scoring, threat detection insights and predictive threat intelligence.

AiStrike leverages DSLM-powered threat intelligence to build near-real-time encyclopedias of emerging threats, continuously assessing customer exposure by correlating vulnerabilities; IOCs; and tactics, techniques and procedures with automated threat hunting, and proactively creating SIEM detection rules. These advancements collectively pave the way for a truly autonomous and contextually intelligent security operations center.

Critical Insight: Domain-Specific Models Will Drive Comprehensive Autonomy and Contextual Intelligence

Near-Term Implications for Product Leaders

DSLMs are enabling a fundamental shift from reactive, manual security processes to proactive, autonomous and highly intelligent operations. Providers have started recognizing that this technology unlocks new levels of adaptive automation and contextual reasoning, moving their solutions beyond static playbooks to dynamic, self-optimizing security engines.

The focus is now on designing products that can fully leverage DSLM-powered multiagent systems to achieve higher levels of automation and significantly reduce human intervention in critical workflows. This strategic direction will lead to accelerated response times, significantly reduce manual overhead, reduce the attack surface or vulnerability impact, and improve decision accuracy across all layers of security operations, ensuring that security teams can scale effectively without a proportional increase in headcount.

Recommended Actions for the Next Six to 18 Months

- Empower real-time artifact interpretation, behavioral correlation and automated response actions by expanding DSLM capabilities to support proactive threat hunting and autonomous forensics. This evolution is key to achieving intelligent, self-directed security operations at scale.
- Invest in multiagent frameworks by developing specialized AI agents powered by DSLMs that can collaborate intelligently across detection, investigation, mitigation and reporting. This approach drives greater autonomy and operational efficiency, and expands coverage across critical domains, such as cloud security, data loss prevention and posture management.

Critical Insight Analysis

The specialized nature of DSLMs is poised to fundamentally transform security operations by enabling comprehensive autonomy and deep contextual intelligence.

Unlike static, rule-based automation, DSLMs allow AI agents to reason over highly specific, domain-relevant knowledge, moving beyond basic automation to intelligent, adaptive decision making.

This capability is crucial for orchestrating complex security workflows that require nuanced understanding and real-time adaptation, such as autonomous incident response, adaptive security model, advanced threat intelligence and predictive security postures.

The autonomy and intelligence are exemplified by several key applications:

- DSLMs facilitate AI-driven auto case assignment by analyzing rich contextual inputs like IoCs, MITRE ATT&CK techniques and analyst performance metrics to route incidents to the most appropriate experts or automated workflows.

- In digital forensics, DSLMs power zero-day forensic triage, evidence gathering and contextual interpretation of raw artifacts, enabling reconstruction of attacker timelines and automated response actions.
- DSLMs enable autonomous playbook assembly, systematically extracting entities, decomposing tasks and populating parameters from investigation reports to generate comprehensive remediation sequences in seconds.

Provider Use Case Examples

Unlike generic LLM tools, Airrived offers ready-to-use agentic applications and over 24 prebuilt agents. Its platform incorporates multiple DSLMs that are trained using techniques like LoRA and reinforcement learning for human feedback, along with proprietary consensus and benchmark-driven algorithms. Airrived promises 60x productivity gains for security analysts by enabling agents to reason and act autonomously.

NSFOCUS GPT provides a rich built-in agent store offering scenario-based security agents for detection (such as unknown attack detection or phishing email detection) and operations (such as noise reduction, autonomous investigation and reporting). It serves as a core "brain" for SOC systems, empowering security operations management and significantly improving efficiency.

IBM's multiagentic AI systems autonomously perform security operations and threat management tasks. Its current use cases demonstrate this comprehensive autonomy, focusing on autonomous threat operations, alert triage, investigation, detection engineering, threat hunting, intelligence analysis and the orchestration of IBM's threat management services. ATOM also seamlessly integrates with any SIEM, endpoint detection and response (EDR), data loss prevention (DLP) solution and writes back results in the source system. Dynamic investigation planning — execution utilizing intelligent orchestration — is one of ATOM's key features. Additionally, ATOM is self-healing, with continuous analyst feedback and autonomous updates to a prompt library.

Imperum's roadmap includes the development of a multiagent system architecture where specialized agents built on DSLMs collaborate intelligently to perform complex cybersecurity tasks autonomously. This includes dynamic cooperation between agents for threat detection, investigation, mitigation and reporting. The Autonomous SOC module itself is designed with multiple agents (Indexer, Retriever and Reasoner) collaborating over a knowledge store.

Critical Insight: Concerns Regarding Reliability, Hallucination, Bias, and Data Privacy Hinder the Full Adoption and Trust in LLM-Based Cybersecurity Operations

Near-Term Implications for Product Leaders

Building trust in AI-augmented security operations requires product leaders to explicitly address concerns around reliability, hallucination, bias and data privacy. This means prioritizing the development of DSLMs with robust safeguards, secure training pipelines and transparent data handling mechanisms.

As DSLMs demonstrate consistent accuracy, reduced false positives and reliable performance through feedback mechanisms, organizations gain confidence to incrementally increase the "autonomy slider," allowing the AI to take on more direct, unsupervised actions (see Emerging Tech: 'Time to Trust' Is the New Vital Agentic AI Metric).

Providing flexible deployment options that align with strict data sovereignty requirements, particularly in regulated industries, will be crucial for widespread enterprise adoption.

Recommended Actions for the Next Six to 18 Months

- Ensure the ethical integrity and reliability of DSLM-driven cybersecurity by embedding continuous bias detection, quality monitoring and HITL review into model development and deployment. Integrate KAG pipelines to ground AI output in verifiable data, enrich investigations and automate decision support — mitigating hallucination risks and fostering trust across security workflows.
- Prevent the misuse of sensitive customer data in DSLM training and inference by enforcing stringent data governance policies and implementing anonymization pipelines. Ensure transparency and accountability through auditable processes that uphold data privacy and build enterprise trust.

Critical Insight Analysis

Despite the transformative potential of LLMs in cybersecurity, significant customer concerns regarding their reliability, propensity for hallucination, inherent biases and critical data privacy issues remain substantial barriers to widespread adoption and trust. Hallucinations, where LLMs generate fabricated or inaccurate outputs, can lead to harmful or failed playbook steps if nonexistent API endpoints are invoked or parameters are misused. Moreover, general LLMs may carry biases from their vast training datasets, potentially resulting in unfair language or overlooking crucial threat indicators.

The most pressing concern in the sensitive cybersecurity domain is data privacy, as raw logs often contain personally identifiable information or intellectual property, necessitating strict governance to prevent leaks during both training and inference. Customers are highly cautious about trusting LLM-based output for critical operations, especially without robust safeguards. Addressing these challenges is paramount for building enterprise confidence and enabling the full integration of language models into security operations.

Continuous bias detection, quality monitoring pipelines and HITL review processes have to be in place to suppress unintended responses and refine DSLMs based on expert feedback.

Provider Use Case Examples

For data privacy, the training data for Imperum's DSLMs is drawn exclusively from public or synthesized sources, with strict internal governance ensuring no private or customer-sensitive data is included in the model life cycle. Offering local deployment options also ensures data remains on-premises, fulfilling critical requirements for sectors with strict data sovereignty rules, such as finance, healthcare and government.

NSFOCUS emphasizes model security by assessing vulnerabilities, conducting adversarial attack tests and using defensive training methods to enhance the model's ability to resist attacks. NSFOCUS GPT AI Security Capability Platform offers two forms: software- and hardware-integrated machines. It can be privately deployed in the enterprise network environment. By enabling AI capabilities in various security products and services, such as the Intelligent Security Operations Platform, Remote Security Assessment System, Next-Generation Firewall, Web Application Firewall, Bastion Host operation security management system, Data Leakage Prevention, and Insight for Discovery and Risk, etc., it can achieve typical scenarios, such as AI security operation, AI detection response, AI offensive and defensive confrontation, knowledge questioning, and data security, realizing intelligent network security.

Conifers highlights its nondisruptive design, which integrates directly with existing SIEM, EDR, IAM, cloud and ticketing systems, fostering trust by avoiding process overhauls.

Transparency around data handling and ensuring secure mobile access are also key challenges that are addressed through active collaboration and transparent communication about data security and operational benefits. These structured safeguards allow DSLMs to operate confidently in real time, delivering trustworthy decisions in high-stakes environments.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Quick Answer: How Will Domain-Specific Language Models Shape the Future of Security Operations?

Emerging Tech: Domain-Specific AI Needed to Break Down Security Silos in TDIR

Emerging Tech: Prioritize Domain Specialization for Sustainable GenAI Differentiation

Emerging Tech: 'Time to Trust' Is the New Vital Agentic AI Metric