elastic

High-volume data sources for AI-driven security analytics

The need for more security data in the AI era

The mission of solving security problems has always been complex and multifaceted, posing challenges with:

Visibility

Maintaining a high enough level of awareness across the environment

Focus

Identifying issues and knowing which ones are highest priority

Speed

Addressing those highest-priority issues in a timely manner

Scale

Avoiding recurrences of known issues while identifying unknown issues

To meet these challenges, security teams must have easy access to the right data. However, “the right data” can be a fast-moving target due to global trends that further complicate these challenges. Now, with AI capabilities integrated throughout security analysts’ everyday workflows — aiding with everything from query conversion to alert prioritization — the application of high-volume data sources to achieve holistic visibility and stronger security posture is more accessible than ever.

High-volume data sources are critical

The evolution of AI-fueled threats is causing many security teams to revisit the data layer. The sheer amount of data analysis required to effectively defend against this new era of threats — which are arriving with a new intensity across a variety of threat vectors — can only be feasibly accomplished by AI-driven security tools that have been built to handle such scale.

To effectively defend against a barrage of threats, security teams increasingly rely on AI tools, which require more context-enriched, security-relevant data across a greater variety of environmental sources. These sources include cloud infrastructure and applications, richer endpoint data, DNS, NetFlow, wire data, IoT, and other sources that are typically higher in volume than many of the primary data sources ingested for daily operations by a security information and event management (SIEM) solution.

Traditional challenges

In many cases, high-volume data can present both operational and business challenges. Either the vast volume of data makes it difficult to perform responsive queries and analyses without a more modern solution, or those data sources are deemed cost-prohibitive to keep in a storage tier capable of on-demand analysis.

Vendor licensing restrictions often force security teams to compromise on data inclusion, deciding which data is essential and which may be less relevant. This “less relevant” data may sit in a frozen storage tier, only to be later called upon in a mission-critical moment to provide additional context for triage. And while this data thaws — a process that is both expensive and time-consuming — critical minutes and hours are passing that could see substantial adversarial activity.

Typical concessions these teams are forced to make can include:

- Shorter retention times for high-fidelity data used in SIEM rules, detections, and machine learning
- Lower data verbosity and logging level for data used in incident investigation and response
- Altogether dropping or excluding "lower immediate value" data from use by the security operations center (SOC)
- Archiving data with limited or slow access, or even limited availability and accessibility (only a portion of the data)
- Siloing data storage by team designation (e.g., for use by separate hunt teams or forensic analysts)

High-volume data sources can help provide visibility into evasive activity, as well as the rich details needed to contextualize a threat. With the right archiving strategy, those data sources can provide the historical context needed to perform longer look-back analyses in response to a security incident or data breach or for proactive threat analysis and adversary profiling.

AI-driven security analytics

The new generation of AI-driven security analytics solutions can automatically ingest and correlate vast amounts of high-volume security data from various sources, identifying anomalies and detecting suspicious behaviors that may indicate an attack. These solutions continuously learn from data patterns, evolving to recognize new threats and minimize false positives without extensive analyst intervention and rule-tuning.

With the help of AI, security teams can more easily prioritize critical incidents, reduce alert fatigue, and accelerate investigation times through real-time integrated threat intelligence, automated triage, and large language model (LLM)-enhanced workflows. AI-driven security analytics transforms security operations into a more proactive, adaptive, and efficient function, allowing organizations to stay ahead of today's threat actors (who themselves are also using AI to its fullest potential) and respond with greater speed and accuracy.

And while these AI tools are proving to be game-changing technology for security teams, their ability to analyze data and provide the powerful augmentation capabilities they're known for is only as good as the data they have access to. Context from high-volume

data sources makes all the difference when AI tools are distilling and classifying alerts, providing workflow suggestions, assisting with threat detection, and performing other functions. Without this contextual data, AI tools can't realize their promised potential and security teams are again left with visibility gaps.

Getting security value from high-volume data sources

Below, we outline the general security value of high-volume data sources and the importance of including them in day-to-day security operations. These data sources are increasingly security-relevant and can provide critical context needed in detections, hunts, investigations, and incident response to help security teams verify, scope, and act more quickly. Most importantly, this added context will facilitate faster, more informed decision-making on the most pressing threats within your team's environment to effectively minimize downstream risk.

Let's explore the high-volume data source types that best enable an AI-driven security analytics tool to perform at its best.

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
IaaS and SaaS	Cloud security posture monitoring, compliance, context-aware asset visibility, user monitoring, threat detection, incident investigation, detecting data exfiltration, identifying lateral movement	Cloud platform: As organizations increasingly deploy production workloads at scale in IaaS environments, such as AWS, GCP, and Azure, it is critical to maintain visibility across the hybrid environment.	Observability data, Cloudtrail, IAM, certs, infra config changes, unauthorized resource usage, permissions, security policy (FW-type - src/ dst/port/protocol), o11y metrics, billing, WAF, VPC Flows	Cloud infrastructure and workload security can be a default priority instead of an afterthought, and "baked in" directly to the DevOps cycle, when IaaS visibility is a standard part of the security operations team's purview, and not treated as a separate silo of operations. In some cases, drops in performance can be indicative of a security issue.	Instance abuse (cryptomining etc), Billing changes/ surges, Instance deletion, cloud discovery, DoS, Defacement, Application and system exploits, configuration changes, Data Theft
IaaS and SaaS	Cloud security posture monitoring, compliance, user monitoring, threat detection, incident investigation, anti-phishing, threat hunting, business context enrichment, detecting data exfiltration, behavior analysis	Cloud application logs: The acceleration to a remote workforce results in more business and sensitive data residing in cloud applications such as O365, Google Workspace, and Salesforce.com.	Observability data (APM), configuration changes, audit logs, application access, file access, user name, timestamps, service usage	Cloud application logs can provide business context needed to identify potential security issues, including insider threat. Cloud application logs can help with email tracing, behavioral baselining, and other investigative methods to help find activity associated with a threat.	Session, Token and Cookie theft, Account Discovery, Data Exfil, DoS, Account Manipulation, Phishing, Software Discovery

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
IaaS and SaaS	Monitoring and cyberhygiene, threat detection, incident investigation	"Communication apps (Zoom, Slack, Teams, etc.) Digital transformation and remote work are pushing teams to collaborate more extensively in real time, and organizations are using tools such as Zoom, Slack, and Microsoft Teams to help them stay connected."	Timestamps, source IPs, users names, actions taken, login information, attachments, message interactions, recordings, deletions	With the increase in remote workforce, abuse, and malicious behavior, such as user impersonation or eavesdropping on meetings with corporate, sensitive discussions are on the rise. Protecting against these by monitoring usage on meeting collaboration tools has become more important than ever.	Data Exfil, Account Takeover, Phishing, Zoom Bombing,
User	User monitoring, privileged user monitoring, compliance, context-aware asset visibility, threat detection, incident investigation, threat hunting, identifying lateral movement, behavior analysis, hunt augmentation, resolution management/ containment	Authentication: Authentication systems verify a user is who they say they are by challenging them to provide a predetermined piece of information (something they know or have) to validate their identity. Organizations use premises-based tools as well as cloud-based services such as Okta.	Credential information and user activity, including what devices, services, and applications were used and when.	"Authentication logs can be used to identify users that have been compromised and potentially uncover what data may have been exfiltrated. Authentication events can help identify unusual user activity by looking at time (day/night), frequency, regularity, and pattern of access (what devices, what services/applications) to identify trends that help security teams distinguish between ""normal"" and ""suspicious."" For users that are known to be compromised, the logs can help identify which assets, systems, services, and applications were accessed, which can help during investigations to determine the type of data that may have been stolen or compromised."	Brute Force Attacks, Golden Ticket, Credential Theft, Privilege Escalation, Credential Stuffing, Password Spraying, MITM

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection and prevention, malware prevention, ransomware protection, threat hunting, incident investigation, identifying lateral movement, detecting data exfiltration, behavior analysis, resolution management/ containment	Next-gen firewall (NGFW): Firewall functionality can be delivered as a stand-alone device (physical or virtual) or as part of an integrated gateway solution. It controls access to the network and performs multifunction inspections to try to identify attacks within the traffic.	Source IP, source port, destination IP, destination port, service, application, user, traffic flow information; NGFW logs will also include threat context from threat intelligence (signature match, payload analysis, threat type, threat source, threat intelligence source)	"Firewall logs provide insights into communications to the IP space, applications, and users. They can be used to: Uncover resources, services, and applications being targeted. Identify malicious traffic trends (command and control traffic, geo-location spikes, etc.). Traffic to the same URL at the same interval every day (which could be malware beaconing (phoning home) to notify attacker of successful installation and get further instructions. Flag unusual activity to a controlled asset. Highlight traffic reversals that represent traffic going in a direction that is not normal to a device, which indicates that device has been compromised."	DoS, Application Exploits, OS Exploits, Protocol based attacks, VLAN Hopping, malware delivery

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection, threat hunting, incident investigation, identifying lateral movement, detecting data exfiltration, behavior analysis	VPN: Virtual private networks (VPNs) provide remote access capabilities into corporate resources for employees and sometimes partners. Depending on the vendor, VPN capabilities may be integrated directly into a perimeter firewall, or they may be deployed separately as a VPN concentrator.	IP addresses, user names, login times, session duration, auth failures, policy check failure, frequency of session timeouts, key exchange issues, tunnel metrics including location	A compromised remote access VPN account can allow an attacker full access to corporate resources that reside on-premises. VPN logs can provide insights into unusual activity such as unusual login times, login from multiple physically distant locations at once (landspeed violation), host policy violations, and other activity that should be flagged for investigation.	MITM, session hijacking, brute force attempts, DoS/SYN flood, spoofing
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection and prevention, malware prevention, ransomware protection, incident investigation, threat hunting, identifying lateral movement, detecting data exfiltration, behavior analysis	IDS/IPS: Intrusion detection systems (IDS) and intrusion prevention systems (IPS) look for and block attacks within network traffic. They use a multipronged approach to detection to identify potentially malicious behavior and generate alarms. High confidence alerts may be automatically blocked by an IPS to prevent propagation into other areas of the network.	Attack definition, CVE information, vulnerability data, as well as general traffic info: source IP, source port, destination IP, destination port, time, services and applications	IDS/IPS logs provide insights into network attack vectors. They often provide the first clue, generating the alert that leads to an investigation. They can be used to identify the specific exploit and vulnerability, which can help identify other potential victims throughout the network; verify reconnaissance activity, investigate the path of an attack, follow additional clues that paint a more complete picture; flag unusual activity (e.g., new lateral traffic through the network that looks suspicious); and gain protocol-level insights.	DDos, Smurf Attack, Ping of Death, Fragmentation, Probing, ARP Spoofing, Port Scanning, Fingerprinting, TLS Evasion

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection and prevention, malware prevention, ransomware protection, threat hunting, incident investigation, detecting data exfiltration, behavior analysis, resolution management/ containment	Web proxy servers: Web proxy servers protect web traffic, blocking malicious URLs (matched to a threat/reputation database), scanning web content, and performing malware analysis to provide insights into the threats you are facing.	Inbound/ outbound web traffic, including information on URLs, domains, and portable executables (files, exe, DLL, docs, JavaScript, etc.)	"Web proxy logs can be used to tie users/systems involved in an attack to web traffic to: Identify initial infections that originated from new, malicious sites or legitimate web sites that were been compromised. Uncover data exfiltration by uncovering unusual outbound activity (e.g., looking for frequency and a pattern of "automated" transmissions). E.g., Beaconing (phone home) traffic of malware trying to establish communication with C2 to get instructions (traffic to the same URL at the same interval every day). Confirm infections by correlating to payload analysis."	Data Exfil, C2 Communication, Phishing attempts, Droppers

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, compliance, anti-phishing, malware prevention, ransomware protection, incident investigation, detecting data exfiltration, behavior analysis, threat hunting, enrichment and threat intelligence, hunt augmentation	DNS: A domain name server (DNS) associates IP addresses with a unique name that identifies a particular computer, service ,or resource connected to the internet or a private network.	IP addresses, assets and their correlating domain name, users/ systems that conducted DNS lookups, time, and location	"DNS logs can be used to look for post infection activity. They can be used to: Identify an attacker looking for ways to communicate back to the command and control server (if using domain names (not IP) they will need to do a DNS lookup); can correlate that information with outbound connections. Uncover: Failed DNS lookups Suspicious domains Transmissions that used the DNS protocol Anomalous DNS activity (e.g., high number of DNS requests coming from a particular client, compared to a baseline)"	Cache poisoning, domain hijacking, dns flood attack, DRDoS, DNS Tunneling, Subdomain attacks
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection and prevention, malware prevention, ransomware protection, incident investigation, threat hunting, identifying lateral movement, detecting data exfiltration, behavior analysis	DHCP: A dynamic host configuration protocol (DHCP) allows an IP address to automatically be assigned to a computer from a range of numbers that have been configured for a particular network.	IP addresses, MAC addresses, interfaces, services, DHCP requests, time, and location	DHCP logs can be used to map user and device activity on the network. If there is suspicious connection, DHCP can be used to identify the specific machine that initiated the connection, providing vital details on which devices have been compromised and are being used as part of the attack.	DHCP Poisoning, DHCP Starvation, MITM, DHCP Spoofing, Theft of Service

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, cyberhygiene, behavior analysis, identifying lateral movement, incident investigation	"Wi-Fi access points Wi-Fi access points are used throughout the enterprise to allow for secure internal wireless access to corporate-connected resources and to provide connectivity to an organization's internet gateway for both employees and guests. Wi-Fi access points are typically configured to allow free roaming throughout a campus; therefore, collecting association/event logs is critical to maintaining security visibility."	Timestamp, error/ event message, mac address, hostname, etc.	Client connection to access points are tracked. It is important to know which access point a client used to connect to a network — it's a way of discovering unauthorized wireless supplicants attempting to connect to the network, rogue/ fake access points configured to force association from unsuspecting users, or an attempt to connect point-to-point to legitimate access points to sniff traffic.	Unauthorised access, Physical device compromise, MITM, Evil Twin, Rogue Access Points, Wardriving, Warshipping, Packet Sniffing
Network activity	Network activity monitoring, cloud security posture monitoring, context-aware asset visibility, threat detection and prevention, detecting data exfiltration, behavior analysis, incident investigation, resolution management/ containment, hunt augmentation	DLP systems: Data loss prevention (DLP) systems build a digital security perimeter around the enterprise and analyze all outgoing (and sometimes incoming) data.	Timestamp, DLP Event type (print, usb, email, etc.), policy breached, block/permitted, user name, IP, data size, document name, etc.	DLP Systems can monitor many different forms of data leakage. Anything from copy/pasting from a document, to including specific text in an email. Crucial for data exfiltration use cases.	Data Exfil

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Network activity	Network activity monitoring, compliance, context-aware asset visibility, threat detection and prevention, threat hunting, incident investigation, identifying lateral movement, behavior analysis, resolution management/ containment	Web application firewalls: This specific form of application firewall filters, monitors, and blocks HTTP traffic to and from a web service.	HTTP headers, source/client IPs, HTTP request payloads, policy hit and matches and scores, timestamp, user agents, geo, etc.	Web application firewalls are design to spot layer 7 attack on web servers/sites. Tremendously vital information, as the entire payload is checked for potentially malicious behaviour based on rules, with the ability to block if necessary.	OWASP Top 10 and other Layer 7 attacks (including DoS), Account Takeover, Session Hijacking, Heartbleed, Shellshock, Poodle.
Endpoint	User and application monitoring, compliance, context-aware asset visibility, threat detection and prevention, anti-phishing, file integrity monitoring, malware prevention, ransomware protection, threat hunting, incident investigation, identifying lateral movement, detecting data exfiltration, behavior analysis, resolution management/ containment, hunt augmentation	"Endpoint security solutions (AV, EDR, EPP, anti-spyware, host-based firewalls/IPS) Endpoint security solutions look for and may block attacks on the endpoint. They use a variety of detection mechanisms to identify potentially malicious behavior and generate alarms. High confidence alerts may be automatically blocked to prevent propagation."	Endpoint security logs may contain specific file names, attack definitions, vulnerability information, executables, new applications	"Endpoint security solutions could provide insights into suspicious/attack activity on the device. They often provide the first clue, generating the alert that leads to an investigation. They can be used to: Identify the specific exploit and vulnerability, which can help identify other potential victims throughout the network. Investigate the path of an attack, following the bread crumbs that lead to a more complete picture."	Malware and Ransomware, Fileless attacks, Memory Modification, Data Exfiltration, Credential Based attacks, Insider Threat, Registry Manipulation

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Endpoint	User and OS monitoring, compliance, context-aware asset visibility, threat detection, anti-phishing, file integrity monitoring, malware prevention, ransomware protection, threat hunting, incident investigation, identifying lateral movement, detecting data exfiltration, behavior analysis, hunt augmentation	"Endpoints (OS, event logs) Verbose endpoint (OSquery, sysmon, other beats) All endpoints record all their activity within their event and OS (Windows, Apple, Android, etc.) logs."	"Endpoint logs contain system information, such as names of running processes, changes to credentials, privilege escalations, time, frequency, and location of activities. Event and process IDs, system metrics (CPU, memory utilization, temp), FIM, anomaly detection (user access, file access, process launch/kill, etc.)"	"Endpoint logs could: Record an attacker installing and running malware. Identify automatic processes and programs that might be used by the attacker to conduct malicious activity (e.g., automatically send data to a server as soon as it is saved to the endpoint). Flag unusual administrative activity (e.g., deleting event logs, which could indicate and attacker trying to delete evidence)."	Malware and Ransomware, Fileless attacks, Memory Modification, Data Exfiltration, Credential Based attacks, Insider Threat, Registry Manipulation
Server	User monitoring, anti-phishing, malware prevention, ransomware protection, detecting data exfiltration, behavior analysis, threat hunting, incident investigation, enrichment, resolution management/ containment, hunt augmentation	Mail servers: Mail servers handle all email transmissions in and out of the organization.	Email logs contain sender/recipient information, as well as the email payload, including any links or attachments.	"Mail logs can be used to scan attachments for malware/viruses and compare email domains to databases of known spam/ phishing attacks. They can be used to: Identify the source of an attack, pinpointing when and how an endpoint that is behaving strangely was infected by a malicious link/attachment. Correlate the click through or installation of a malicious link/ attachment with outbound activity on that endpoint to learn more about the exploit and identify the progression and spread of the attack."	Phishing Campaigns, Malware/ Ransomware Delivery, Data Exfiltration, Extortion, Insider Threat,

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Server	Monitoring and compliance, context-aware asset visibility, privileged user monitoring, threat detection, detecting data exfiltration, behavior analysis, threat hunting, incident investigation, hunt augmentation	Database audit logs: Databases are some of the most crucial production hosts in an environment, as they house IP, PII, CC info, and more.	Timestamp, database user, database name, database table, query, response time, byte size, authentication method, etc.	Knowing the who, what, and when of database queries is absolutely vital — and is also a compliance requirement. Elastic Security can model queries and detect usual query patterns between applications and their associated database.	Data Exfil, Insider Threat, Data Manipulation, Data Defacement, Financial Fraud
Server	Monitoring and compliance, threat detection, cloud security posture monitoring, incident investigation, threat hunting	“Certificate transparency logs Certificate authorities (CAs) are required to publish a public log of every single certificate they sign, which is millions every day. “	Timestamp, log entry/store, entry stage, common names, subject alternative names, creation date, expiry, fingerprint	Extremely valuable to detect rogue certificate generation, certificates generated for phishing and more.	Phishing, CA compromise, Insider Threat
Server	Monitoring and compliance, threat detection, cloud security posture monitoring, incident investigation, threat hunting	“CDNs (Cloudflare, Akamai, etc.) CDNs host all static web assets for an organization’s website and many times combine WAF. “	Very similar to WAF/HTTP but can also include a combination of other fields/values depending on the services the CDN offers	These will be very useful for determining web based attacks, just like web server logs.	OWASP Top 10 and other Layer 7 attacks (including DoS), Account Takeover, Session Hijacking,etc

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Wire and flow data	Monitoring and compliance, context-aware asset visibility, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, identifying lateral movement, detecting data exfiltration, anti-phishing, malware prevention, ransomware protection	PCAP: Short for packet capture, this serves as an API for capturing network traffic.	Protocol, source, and destination IPs and ports (depending on protocol), frames and sequences, client/server ciphers for TLS handshakes, payloads, etc.	Similar to netflow and other network based logs, without the dependency of relying on OS collection. Being able to collect network traffic as it appears "on the wire" allows a team to spot attack traffic that spans 1,000s of hosts by using a network tap, and also allows to see if malware/users have interfered with TCP/UDP streams as they leave a host.	Remote Access, C2 Communication, Malware Download, TLS Injection, Session hijacking, Data Exfiltration, DGA attacks, Watering hole, Brute Force
Wire and flow data	Monitoring and compliance, context-aware asset visibility, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, identifying lateral movement, detecting data exfiltration	"Flow data: Analyzed to monitor network performance, optimize infrastructure, and detect malicious traffic. NetFlow: Introduced on Cisco routers (1996) to provide the ability to collect IP network traffic as it enters or exits an interface. IPFIX: Internet Protocol Flow Information eXport (IPFIX) has since superseded the NetFlow protocol."	Source and destination IPs and ports and bytes sizes, combined flow size, protocol, packet count, flow/community ids, flow duration, etc.	Determining irregular data transfers and significant outbound connections, with the ability to model this behaviour using ML	Data Exfil, DoS

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Connected devices and physical security	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	"ICS: An industrial control system (ICS) consists of integrated hardware and software that monitors and controls the operation of industrial equipment. IoT: Internet of Things (IoT) includes devices for measuring, monitoring, and controlling physical devices through cloud-based processes."	Device name/ID/IP, message, message type/producer, severity, location, unit	Just like any other device with a network connection, IOT and ICS generate traffic, which can, of course, be the avenue of attack delivery, data theft, rogue manipulation and more (DDoS etc). https://collaborate.mitre.org/attackics/index.php/Main_Page	DoS, Account Takeover, Device Takeover, Module Discovery, Data Exfil, DoC, File Injection, ec etc
Connected devices and physical security	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	MDM logs: Mobile device management (MDM) logs help enhance corporate data security by enabling the monitoring, managing, and securing of employees' various portable/mobile devices.	Device name, device action, policy, user name, IP address, owner, MDM action, etc.	Any organization that has any form or corporate mobile devices would (and should) be monitoring and managing these devices through an MDM, particularly corporate phones, iPads, etc. Changes to policies, unusual app installs, login locations, lost devices and more can all be logged and monitored	Physical Theft, Data Exfil/Theft, Mobile Supply Chain attacks, Mobile Malware, Credential Misuse

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Connected devices and physical security	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	Access control logs: These are logs created through physical security systems (swipe cards, fingerprint readers) to reference and verify valid/invalid entry at a given location.	User names/IDs, room/entryway, timestamp, entry/exit event (sometimes IP address of the device/reader)	Allows SOC teams to monitor unusual entry patterns for employees, and determine if someone is entering an area they shouldn't (server room physical monitoring is also a PCI requirement). Using ML, all of this can be modelled too - "Why did a user enter the building at Saturday at 2am, if they always enter during the week at 8am?"	Social Engineering, Data Exfil, Physical Theft, Insider Threat
Connected devices and physical security	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	CCTV motion events: This is closed circuit television-detected activity at a given location.	Endpoint security solutions could provide insights into suspicious/attack activity on the device. They often provide the first clue, generating the alert that leads to an investigation. They can be used to identify the specific exploit and vulnerability, which can help identify other potential victims throughout the network.	Similar to physical access control, we can determine irregular motion events, and, if face recognition is included, we can identify and model user irregularities	Social Engineering, Data Exfil, Physical Theft, Insider Threat
Connected devices and physical security	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	Printer logs: These enable security teams to keep track of print jobs on their business's printers.	Printer name/ip, source ip/user, sheets printed	Using non-traditional data sources, data exfil can also be detected in the physical world. If a user has never printed 300 pages before, at 6am, we probably want to take action on it	Data Exfil

High-volume data sources for SIEM

Category	Primary Use Cases	Data source (name & description)	Fields within the data / classes of logs / things to look for	Security significance / value	Examples of attacks / attack methods you can uncover with this data source
Code repository	Monitoring and compliance, threat detection, incident investigation, threat hunting, hunt augmentation, behavior analysis, detecting data exfiltration	Source control logs: These allow for monitoring of actions your users perform on source-controlled objects (GitHub, etc.) across your organization.	Timestamp, repository names, user names, organization names, action taken, source IPs, etc.	As organizations start doing everything "as code," it is vital to monitor irregular activities on repositories, as well as things like password/secret access, api calls, logins and more.	Supply Chain Attacks, Insider Threat, Data Exfil, Code manipulation, Secret discovery and theft

elastic

Let AI onboard your high-volume data

Cybersecurity professionals face an expanding attack surface, unwieldy environments, new data formats, and the need to reverse-engineer emerging methods and new attack tactics, techniques, and procedures seen in the wild. Access to the right data sources is the first step to helping your organization close these gaps.

With Elastic Security's Automatic Import feature, security teams can leave the traditionally manual, time-intensive process of onboarding data to AI. Saving security teams days (if not weeks/months), Automatic Import automates the development of custom data integrations to facilitate broader visibility and easier SIEM implementation.

Storage of all this data has been a longstanding challenge for security teams. Fortunately, Elastic's flexible storage tiering structure enables long-term storage of high-volume data sources through a historical/frozen tier that uses object storage and adaptive caching on each cluster node to enable the best possible performance for fully live searching and analysis. And it's offered at the absolute best price of any leading SIEM vendor.

Let AI onboard your data for you and start gaining valuable context from your high-volume data sources.

Try out Elastic Security