# Rules Versus Models In Your SIEM

## Introduction

Malicious insider threats, compromised insiders, and sensitive data exfiltration remain significant concerns for enterprises today. Organizations and their security operations centers (SOCs) require advanced technologies to protect themselves from such threats and ensure compliance with regulations.

Security information and event management (SIEM) tools have been instrumental in threat detection and compliance for organizations. However, many legacy SIEM tools rely heavily on correlation rules or predefined criticality choices from vendors, which can prove ineffective against complex, longer lasting, or more distributed attacks. A new generation of SIEM tools employs behavior- and context-aware models to track user behaviors, improving their ability to detect unknown threats and complex attack chains.

In the following sections, we will examine the distinction between rules and models, analyze their respective pros and cons, and offer guidance on designing and building effective rules and models.

## What is a Correlation Rule?

Correlation rules compare incoming events to predefined relationships between entities, identifying anomalies based on specified conditions. These conditions act as triggers and are constructed from established knowledge of attack patterns, essentially representing a series of known steps that can be detected. Think of correlation rule triggers or thresholds as, "Send an alert whenever A, B, and C happen within 30 minutes."

**Realistic examples of correlation rule scenarios include:**

- Triggering an alert when a file or packet contents match a known string or signature
- Alerting when a user exceeds three failed login attempts on the same machine within an hour
- Notifying upon observing more than 10 failed login attempts followed by a successful logon
- Raising an alert when a single host or IP scans multiple ports over a certain duration

**Correlation rules can be categorized two ways:**

1. Simple rules target specific event types and trigger alerts. These can include the use of lookups against a watchlist.

   Example: If an event uses ICMP transport protocol and its source IP address is 73.230.79.1, then alert.

2. Complex/composite rules combine multiple rules or nest rules within one another to trigger actions or alerts.

   Example: If rule X fires followed by rule Y firing within a predetermined amount of time, then take action or alert.

## What is a Model?

Models profile user or asset behavior in a specific aspect of their interaction with the corporate or IT environment using machine learning (ML), a subfield of artificial intelligence (AI). Examples of models include: tracking employee badge-in/out times, the number of logged-in assets for a user, VPN usage following badge access, and operating system used by a user. Models start by establishing activity baselines, making them adept at identifying deviations from normal behavior, a key benefit of using models.

**Examples of model applications:**

- Alerting when a user switches to a privileged account and performs an abnormal data upload to external services
- Notifying when a user VPNs from an unfamiliar location and accesses executive file shares
- Identifying users simultaneously present in multiple geophysical locations
- Raising an alert when a user uploads or downloads data file sizes atypical for their group or historical behavior

## When to Use Correlation Rules Versus Models

To determine whether to use correlation rules or models, consider the following use cases:

| Use Cases | What to Use |
|---|---|
| **Real-time monitoring of known threats**<br>Many threats that infiltrate organizations are well documented, with malicious actors employing similar methods, albeit with minor modifications. Rules can effectively detect these attack patterns. | **Correlation Rules** |
| **Compliance violation checks**<br>Many data security regulations, such as GDPR, PCI DSS, and HITECH, require organizations to demonstrate effective controls. These regulations have well documented security controls, and rules can be instrumental in conducting compliance assessments. For instance, "An alert should be triggered if the antivirus is disabled on our PCI systems." | **Correlation Rules** |
| **Signature-based threat detection**<br>When malware is detected, its signature is added to a repository. These repositories contain hundreds of millions of signatures that identify malware. For example, companies like VirusTotal aggregate malware signatures that are accessible to vendors. Rules can be used to detect known malware signatures. | **Correlation Rules** |

| Use Cases | What to Use |
|---|---|
| **Behavior-based anomaly detection**<br>In dynamic environments, organizations often encounter emerging technological trends, such as Bring Your Own Technology (BYOT) and the adoption of corporate data in the cloud. Consequently, threat actors are finding an increasing array of entry points into organizations, highlighting the importance of monitoring both normal and abnormal user behavior.<br>Models are helpful for tracking behaviors, unlike correlation rules, which are not designed for monitoring user behavior. This is where user and entity behavior analytics (UEBA) comes into play. | **Models** |
| **Corporate data exfiltration detection**<br>Data exfiltration is the unauthorized transfer of data from within your organization to external sources. Threat actors gain access to targeted machines either through remote applications or by installing portable media devices. Advanced persistent threats (APTs) are one form of cyberattack where data exfiltration is often the primary goal. Detecting data exfiltration often entails monitoring activities that involve various elements, such as movements across assets, privileged account access, employee behavior, and peer-group activity. Models are particularly effective for detecting these activities. | **Models** |
| **Zero-day threats**<br>Zero-day threats have not been encountered previously, making it impossible to create rules to identify them. These threats may involve an unknown combination of anomalous lateral movements, abnormal or remote logins, file access, and abnormal data uploads. In such cases, a modeling approach is useful, as it can easily identify these threats by detecting deviations from established behavioral baselines. | **Models** |
| **Lateral movement detection**<br>Lateral movement is commonly used in cyberattacks to infiltrate hosts from a compromised system. Then, it allows malicious actors to gain access to sensitive data, shared files, and privileged credentials. These acquired privileges can then be exploited to access more resources, further escalate privileges, and pilfer even more valuable credentials. Lateral movement evades detection through correlation rules because different aspects of the attack may be present across various IP addresses, identities, and machines. Models, on the other hand, have the capability to identify and correlate anomalous activities spanning different systems, enabling the detection of attacks. | **Models** |

## Pros and Cons of Correlation Rules

### Pros of Correlation Rules

- Expedite responses to known or routine attacks
- Detect risks by correlating relationships among resources
- Enable real-time monitoring of known threat vectors
- Simplify policy creation and expression through a SIEM rule builder

### Cons of Correlation Rules

- High false-positive rate: The dynamic nature of IT environments requires frequent rule updates by analysts and consultants. Incorrect correlation rules, combined with a lack of user or entity context such as departments, job functions, peer groups, and roles, can trigger hundreds of false-positive security alerts.
- Applicable only to known attack patterns: Correlation rules are best for creating policies for known patterns, but are not suited for identifying unknown attack chains.
- Long execution times for nested rules: The execution of rules can be time consuming, especially with nested rules
- Ongoing rule maintenance: As organizations introduce new IT products and patch new releases, correlation rules must be continuously reviewed and adapted to stay effective.

## Pros and Cons of Models

### Pros of Models

- Establish baselines for behavior, facilitating anomaly detection: By tracking normal user or asset behavior over a duration, a baseline is established, making it easier to recognize abnormal activity. Effective products predominantly use models to baseline behavior, incorporating expert knowledge and contextual information about your organization.
- Track lateral movement and abnormalities in user behavior: Models can monitor users who have atypically accessed file servers, observe their login activities, track administrative assets, and monitor service accounts to detect lateral movement.
- Detect unknown threats: Deviations from the baseline are monitored for abnormalities. Various models can be linked together in a user timeline to provide analysts with a comprehensive view of an attack chain. For instance, they can detect advanced threats involving account privilege escalation, lateral movement, administrative asset access, and data exfiltration.
- Leverage contextual data for effective anomaly detection: Models use contextual data, such as user hostname, peer group behavior data, user types (executive, administrator, service account), and departmental user data, to detect anomalies effectively. And they can incorporate information about users and entities, including de facto asset owners, normal VPN access time zones, top network users, folders containing source code, and more.

### Cons of Models

- Baseline creation takes time: Establishing a baseline requires time for the analytics engine to model normal user behavior. A well-constructed baseline is essential for quickly and effectively detecting threats.
- May require considerable professional services if not well engineered within the products: Dealing with an unknown threat usually involves a combination of abnormal behaviors, such as lateral movement, remote logins, access to administrative assets, account switching, abnormal file access, and data uploading. To convey the complete narrative of an attack, all relevant events must be stitched together. Also, data structures must be in place to prioritize risks based on anomalies.

Evaluate products that offer pre-built features that use models effectively to detect unknown threats. These solutions can address your specific use cases and deliver rapid time to value without requiring extensive customizations.

# Designing Rules and Models

## Correlation Rule Builder

Many SIEM tools offer rule builders within their user interface (UI), simplifying the process for administrators to pick up key parameters from various events and construct rules. A typical rule builder contains categorized lists of all components available for configuration.

Correlations can be created by dragging items from events and their associated alerts to trigger actions. For example, when creating a rule set related to a specific threat, you have the option to either set up multiple correlations or nest them all into a single rule.

Building rules this way can be daunting and time consuming, especially in today's complex IT environments. But modern SIEM tools, such as Exabeam, provide correlation rule templates and user-friendly wizards. The Exabeam rule-building wizard makes it possible for even junior analysts to create complex rules with ease.

# Here is a typical rule-building sequence:

## 1. Choose an existing template or create a rule from scratch.

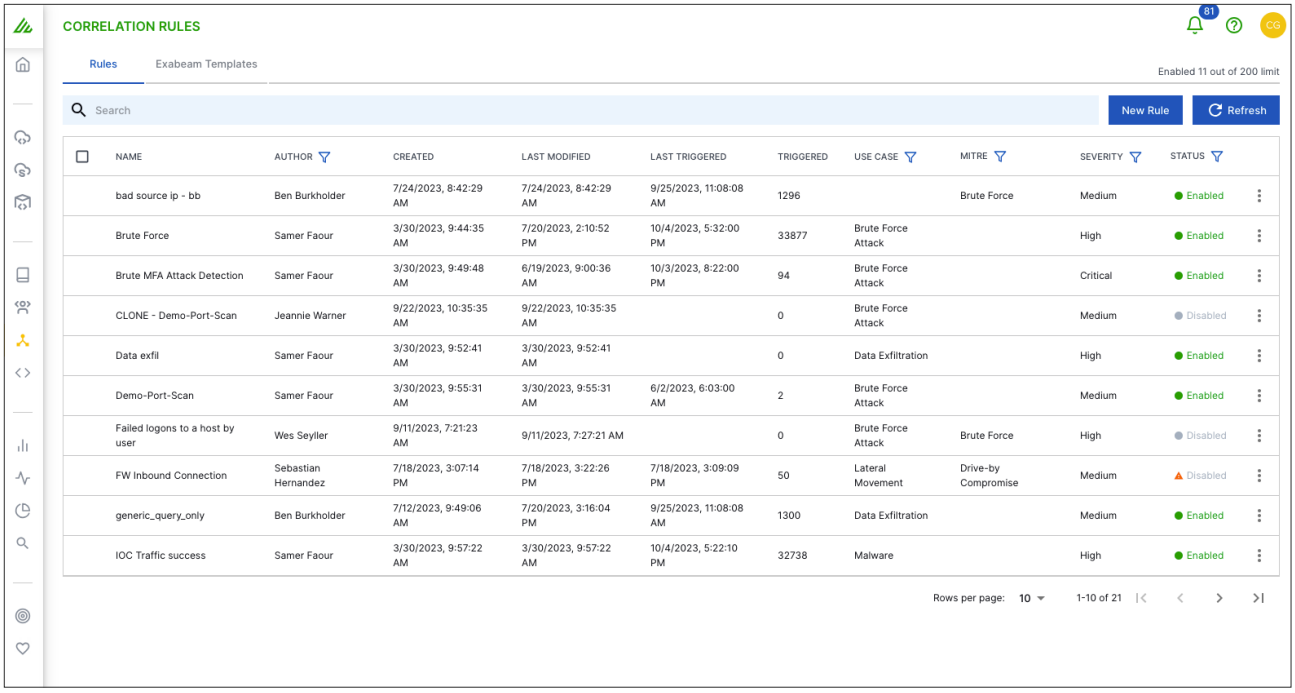


**Figure 1.**
Manage your correlation rules from a single page.

## 2. Build a sequence.

Add a search query to fetch all events against which the rule can be evaluated and set conditions to trigger your correlation rule.

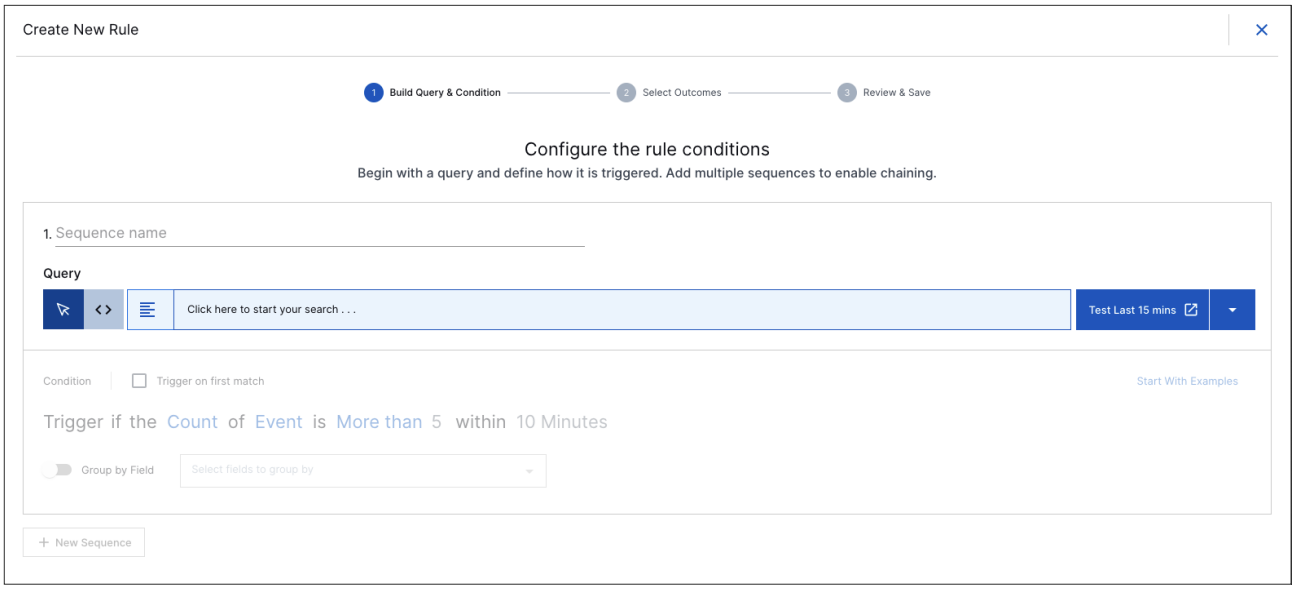


**Figure 2.**
Build your correlation rule sequence.

## 3. Define the outcome for when the correlation rule triggers.

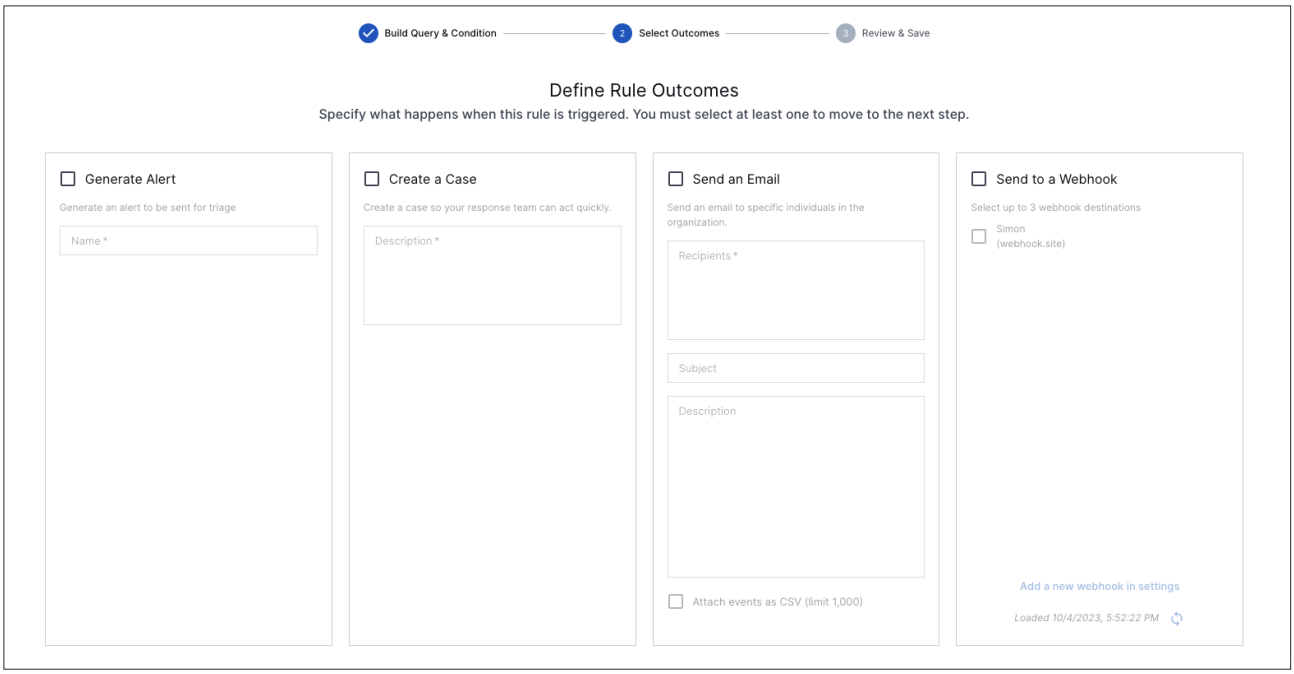


**Figure 3.**
Generate an alert, create a case, send an email, or send to a webhook.

## Design Considerations for Models

Models serve the primary purpose of tracking user behavior and associated context. But before addressing user behavior, let's first examine the significance of context.

### Context: Why It Matters

Security logs originate from various entry points, such as Windows servers, VPNs, firewalls, endpoint devices, or DNS servers, and are aggregated to facilitate threat discovery. Logs reveal the actions of users and entities, while context provides information about the identity of users and entities, their roles, and their typical behavior. For example, while you can identify the source IP address of a user's workstation from the logs, you can track the user's location if they connect through VPN by mapping the IP address to the hostname.

Enriching security events with contextual data makes it easier to uncover threats. Other examples of contextual enrichment include classifying service accounts alongside servers and workstations, and identifying user peer groups, contractors, and privileged accounts.

### Modeling User Behavior: A Timeline

Enriched security events empower security operations teams to track user behavior. For example, you can model a user's work hours, the remote logon servers they access, the zones they log in from, their VPN locations, and more. This aids in establishing a baseline for normal user behavior over time, making it easier to detect anomalies.

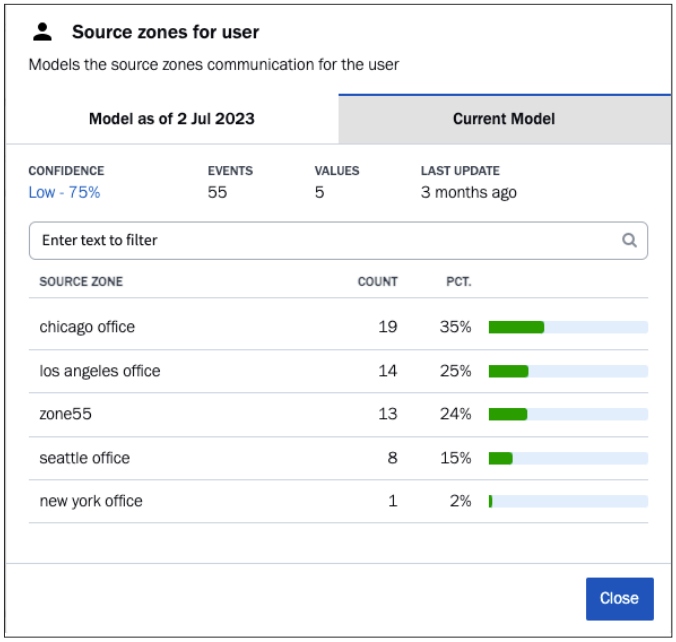


**Figure 4.**
User model shows normal network zones.

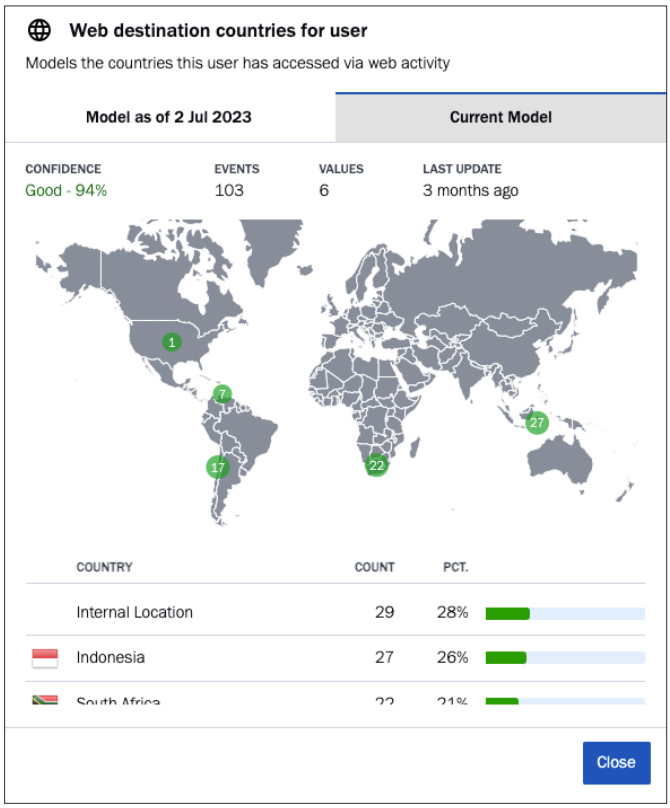


**Figure 5.**
User model shows typical web destination countries.

## Using Models for Threat Detection

Any abnormal activities are tracked and assigned a risk score. When risk scores exceed a predefined threshold, they are escalated to a security analyst for investigation. This streamlines the process by consolidating a number of discrete alerts into one single object for analyst review, instead of dealing with various individual events and piecing them together.

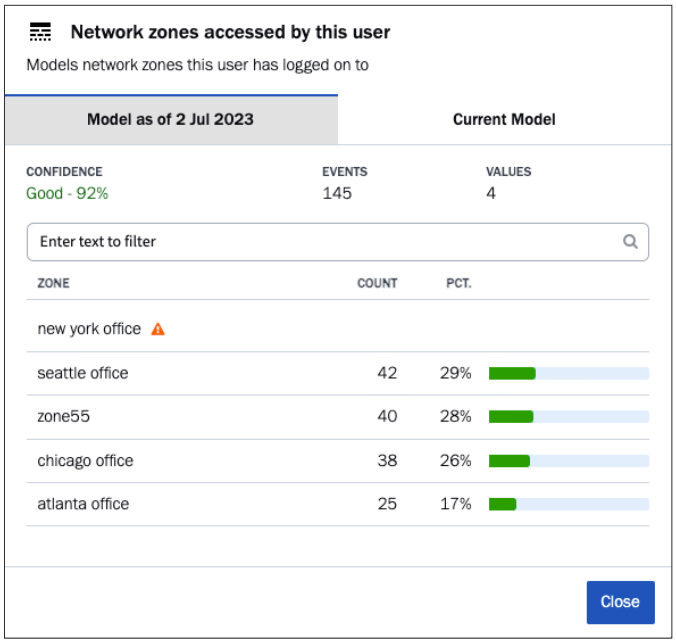


**Figure 6.**
User model flagging New York network zone as abnormal.

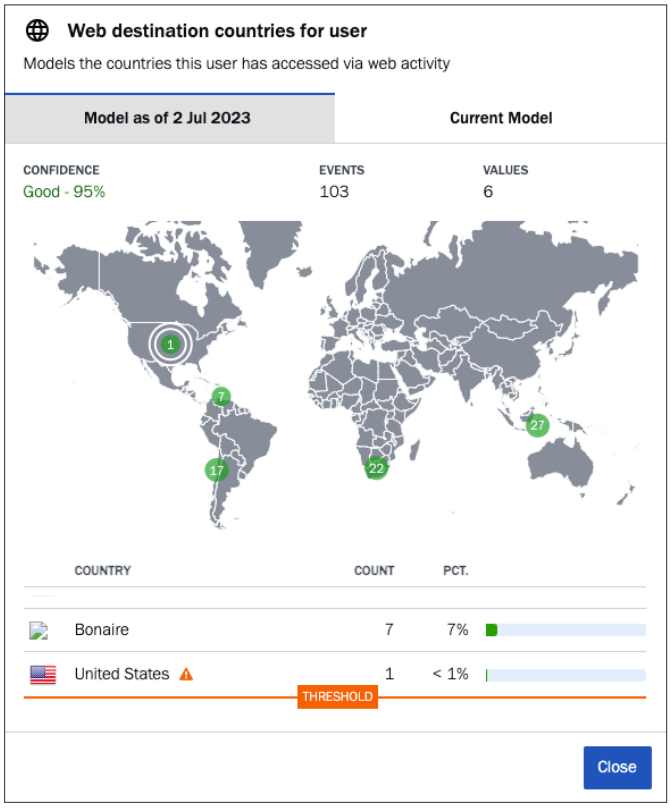


**Figure 7.**
User model shows user accessing web activity in U.S. as a risk.

## Ad-Hoc Threat Hunting

By monitoring normal and abnormal user behavior, the ability to hunt for potential threats across your IT environments becomes far more effective. You can apply contextual data, risk reasons, and activity types to your threat search criteria, simplifying the process for analysts while creating a query structure that is quite complex.

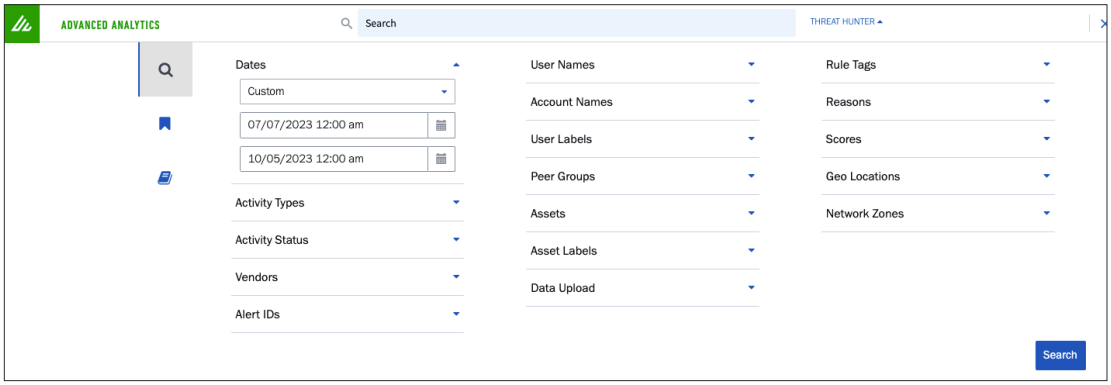


**Figure 8.**
Threat Hunter search query allows for searching for activity types, user types, and risk reasons.

## When to Retire Rules

Many organizations have difficulty supporting traditional SIEM tools due to constraints such as time, funding, resources, and processes. Such constraints necessitate organizations to reassess the long-term value of maintaining and supporting these tools.

If your organization faces any of the following challenges associated with rules-based tools, it may be time to retire these rules and consider modern, behavior- and context-aware SIEM solutions:

- Time and resources spent on false positives: Incorrect correlation rules or excessive false positives demand data reanalysis and rule tuning, which consume valuable support, resources, and time.
- Burdensome rule maintenance: Organizations rely on data collection and retention for correlation purposes. To ensure the fidelity of correlation logic, custom correlations must be maintained in response to environmental changes. This typically requires engagement with professional services and increases operational overhead.
- Ineffectiveness of alerts and reporting in dynamic environments: With today's advanced attacks and dynamic infrastructures featuring BYOD, semi-managed devices, and corporate data in the cloud, ineffective rules warrant an evaluation of model-based tools equipped with behavior and context awareness.

## How to Migrate Rules

Follow these steps to successfully transition from rules to models:

- Leverage in-house SIEM expertise: Use internal cybersecurity experts familiar with your current IT environment and security operations. Review use cases, log sources, and security gaps within your existing SIEM deployments.
- Validate your current rules: Review your current rule set and differentiate between rules that should be retired and those that remain valid.
- Engage with migration or professional services: Map out existing use cases and document any new ones. Enlist SIEM consultants to help migrate valid rules.

## Choosing Between Rules And Models

Behavior and context-enriched analytics tools employing models prove effective in detecting threats across your IT environments. To recap, model benefits include the ability to identify previously unknown threats, detect anomalies in user behavior, uncover lateral movement across your environment, and enable threat hunting with context-aware data.

Models can also detune a useful but potentially noisy correlation rule, such as an interactive logon by a service account — a rule commonly provided by many SIEM solutions. When deployed within an organization, some scripts may require interactive logon to function, which could flood the SOC with thousands of alerts. The logic can be implemented within a model to first learn which service accounts normally perform interactive logons to specific assets, then trigger on anomalous occurrences of the pattern.

Correlation rules are useful in detecting known threat vectors, maintaining the security of critical systems. But if your rules have missed anomalies that should have been detected, or have been disabled without clear reasons, it's worth considering whether these anomalies could have been detected using models.



## About Exabeam

Exabeam is a global cybersecurity leader that helps organizations detect threats, defend against cyberattacks, and defeat adversaries. Exabeam was the first to put AI and machine learning in its products to deliver behavioral analytics on top of SIEM. Today, our New-Scale SIEM™ includes cloud-scale security log management, powerful behavioral analytics, and automated threat detection, investigation and response (TDIR) to provide an advantage against cyberthreats. Exabeam baselines normal behavior so security operations teams can identify the abnormal and take action — for faster, more complete responses and repeatable security outcomes.