WHITEPAPER

Open Federated Search:
Reinventing Traditional Security Investigations

Query.ai

Introduction

Security operations leaders know that every security investigation requires access to as much potentially relevant data as possible to achieve **high-confidence outcomes**.

During a simpler time, all that relevant data was tucked neatly behind the firewall, which provided a clear corporate perimeter. This made it easier for security teams to centrally access security data so they could quickly investigate and determine if an alert was a real issue that required a response effort or a false alarm.

However, the way companies conduct business today has changed significantly. Business-enabling technologies and systems such as **cloud services, mobile devices, and SaaS applications** have been innovating at a breakneck pace. As a result, organizations have now **decentralized**, and growing data volumes spread across their on-premises, multi-cloud, and SaaS environments.

What should security operations teams do with all that data? Many tools such as security incident and event management (SIEM) attempted to tackle the dynamic of an expanding corporate environment by requiring organizations to centrally ingest their decentralized data into the security solution. Yet, the notion of **complete centralization** to achieve a comprehensive security view has proven **unviable** and **cost prohibitive**.

As a result, organizations are left with an environment full of data silos sitting outside their centralized logging tools with limited or no ability to easily access that data during the investigation and incident response process.

Notably, much of that siloed data resides in the cloud. At the end of 2021, **67 percent of all enterprise infrastructure** was cloud-based.[1] It makes sense, then, that **30 percent of security professionals** cite lack of visibility into applications and data sets as factors that limit the ability to improve cyber resilience.[2]

This issue has given rise to a newer approach for security investigations called **federated search**. When done right, this capability can allow security teams to bridge the data silos and investigate across toolsets that reside in their cloud, SaaS, and on-premises environments.

Query.ai

What is federated search?

The first products to use federated search emerged in the late 1990s, giving users the power to search on multiple disparate databases at the same time.[3] At a high level, **federated search retrieves information from a variety of siloed data sources in real time via a single search**.

While federated search has existed for decades, the cybersecurity market only recently started to adopt its capabilities for security use cases. With the volume and variety of data that security teams need to access and analyze daily, it's easy to see how federated search can provide strong value.

For example, consider a daily security operations task of examining a malicious IP address associated with malware to see if it has a presence anywhere across the environment.

For organizations with a SIEM, the search typically starts there **and then requires the security analyst to:**

- Conduct individual searches within data sources not ingested into the SIEM.
- Know which data silos to reference and get through each login process, including passwords and multi-factor authentication.
- Understand the search syntax for each system.
- Piece together the bits of uncovered information in spreadsheets to answer the original question.

Alternatively, with federated search, a security practitioner can make a **single query request** about the IP address and get answers from **all databases** participating in the federation, such as the SIEM, Amazon S3 logs, Active Directory, human resource system, Windows event logs, and more. The federated search then aggregates the results and presents the **unified information** to the security analyst.

This approach makes it significantly more efficient for security teams to identify the instances of an IP address, hash, file name, URL, etc. that might be present on 50 other systems that the organization isn't indexing in their SIEM.

Different approaches to federated search

Federated search functionality in the cybersecurity industry is only starting to reach its full potential. In its current applications, the market is seeing two approaches to federated search: **single vendor** and **open federated search models**.

Single vendor, or closed, federated search

As the name implies, single vendor federated search allows security teams to execute a unified search across a **single vendor's toolset**. It is most often seen today in SIEM solutions to create a search bridge between an organization's cloud and on-premises instances of the vendor's products. This enables security teams to search on security event data across their cloud-to-cloud, on-premises-to-cloud, and on-premises-to-on-premises instances of that single vendor's solution.

While a helpful start to federated search, this approach means security teams **still must pivot** to individually search in other relevant environments, such as AWS and Elastic, and in other vendor data sources to reach an outcome for each investigation and incident response effort. Single vendor federated search also relies on the traditional approach of moving and centralizing the organization's data within that one vendor's solutions.

Open federated search

Open federated search retrieves information from across **multiple vendor solutions and environments**.
It uses API integrations with third-parties to perform a unified search across the data sources that are participating in the federation, and it does this **without requiring data transfer or centralization**.

This approach expands the flexibility of security teams to access their siloed data sources from across cloud, SaaS applications, and on-premises environments, enabling them to **eliminate pivots and individual searches** across data silos so that an organization's security operations can efficiently and rapidly get from the initial investigation question to a high-confidence incident response outcome.

A key aspect of getting to that high-confidence outcome: **gaining context**. Open federated search allows teams to gain the situational awareness they need by simultaneously examining **different types of data sources** and not simply a single vendor's security event data.

For example, open federated search can provide security teams with the context they need to reach a decision on the investigation by retrieving data from an organization's asset management, Active Directory, human resource system, and ticketing system.

> Open federated search allows security operations to **search fluidly across security event and essential operations data**, such as Active Directory, human resources, and asset management systems to unveil the full investigative story.

Reaping the benefits of open federated search

Security teams need a lot of information to move from the security alert phase to the final outcome of an investigation. Just like a good investigative journalist needs to dig into the details of a story, security practitioners need to know a lot more than the high-level "what" information they get from a security alert.

In short, they need the **full context on the who, what, when, where, why, and how** of a security situation to gain a complete understanding. This information comes from a variety of siloed data sources across the environment. Open federated search allows security teams to achieve this goal of gaining context.

Let's look at some **common use cases.**

Phishing threats

If an unsuspecting employee falls for a phishing email and clicks on the malicious link, the SIEM will sound the alarm that there was a suspicious Office 365 login.

However, this doesn't tell the analyst anything about the user or that the origin of the suspicious login started from a phishing email, much less if any other employees received the same phishing email.

This typically is when an analyst **begins pivoting** in the investigation with searches in the Active Directory system to identify the user and original machine that received the phishing email. With open federated search, security teams can search on the suspicious login details and **simultaneously access** Active Directory to uncover the user associated with the event and see how many and which other users also have attributes of the phishing email on their machines.

Malicious IPs

When there's an alert on a malicious IP, analysts first search in the SIEM or other central logging system.

From there, they need a lot of additional information from other system logs that are not centralized in the SIEM (e.g., threat intel, EDR context, DLP, asset management, etc.).

Instead of having to log in to each system separately to continue the investigation, the open federated search model lets an analyst **run the query across all systems**, which gives security teams **instant and unified answers** to the investigation questions. This allows security operations teams to search fluidly across security events and the operations data, such as Active Directory, human resource systems, asset management, and ticketing systems to unveil the full investigative story.

The result is a complete view of the decentralized environment and its wide-ranging data sources that the security team needs to access to **quickly and efficiently** complete investigation and incident response efforts.

Conclusion

Seventy-one percent of security professionals state that greater visibility into data sets improves efforts and time for detection, response, and remediation.[4] Investing in open federated search functions provides security teams with **greater visibility and context** to accelerate investigations and response efforts.

Additionally, **minimizing mean-time-to-response (MTTR)** allows organizations to stop dangerous malware infections from taking down the company's machines or interrupting business operations. This is critical in today's world of Log4j vulnerability exploits and the ransomware attack method du jour that can spread laterally through networks like wildfire.

Open federated search creates **more efficient security operations processes** that allow analysts to focus their time **supporting the company's revenue-generating initiatives** that require more thought and expertise (e.g., level three and four analyst roles).

Ultimately, when companies adopt open federated search, they can increase team efficiency and automate previously manual and time-consuming investigation processes. By accessing valuable data and giving security teams **complete insights into decentralized data**, they can more quickly, accurately, and cost–effectively address security threats.

71%
of security professionals

state that greater visibility into data sets improves efforts and time for detection, response, and remediation.

Query.AI: Making open federated search a reality

The Query.AI Security Investigations Platform serves as a **connective tissue** that delivers an **open federated search** approach to conduct cybersecurity investigations across data silos and any vendor solutions, **without requiring data storage, transfer, or centralization.**

The platform unlocks access to, and value from, cybersecurity and business data wherever it is stored, giving security analysts a unified view with the full context they need to make **timely and accurate investigations decisions with high-confidence outcomes.**

> "Query.AI was attractive because it didn't require the efforts to centralize or orchestrate our data like SIEM or SOAR.
>
> With Query.AI, we have central access to search our decentralized data to enable faster, higher-confidence outcomes. Query.AI helps us solve the issues around the expanding complexity of modern IT environments."
>
> **Joe Oney**
> **Security Operations Manager**
> Hogan Lovells

Learn more

Ready to **expedite your security investigations** with open federated search?

For more information visit:
www.query.ai/federated-search

[1] *Findstack. The Ultimate List of Cloud Computing Statistics 2022. December 2021.*
[2] *Ponemon Institute. Cyber Resilient Organization Report. 2020.*
[3] *ScienceDirect. Federated Search.*
[4] *SANS Institute. 2019 SANS Automation & Integration Survey. March 2019.*