검색·필터 (KQL)

필드 검색 · 불리언 · 범위 · 와일드카드 · 실전 쿼리

01
기본 검색
필드:값 등호와 구문으로 정확 일치 검색

02
불리언 결합
and·or·not과 괄호로 조건을 조합

03
범위·와일드카드
숫자 비교와 부분 일치·존재 여부 검색

04
실전 쿼리
공격·정찰·유출·브루트포스 탐지 쿼리

S3

검색·필터 (KQL)

세션 개요와 학습 목표

> KQL로 로그를 원하는 조건으로 걸러내는 검색 문법을 손에 익힌다.

이 세션에서 얻는 것

- **필드 검색**: 필드와 값을 등호와 구문으로 지정해 원하는 로그를 정확히 골라낼 수 있다.
- **조건 조합**: 불리언·범위·와일드카드를 엮어 복합 조건 쿼리를 스스로 작성한다.
- **실전 적용**: 공격·정찰·유출·브루트포스를 탐지하는 실무형 쿼리 패턴을 익힌다.

KQL을 쓰는 곳

- **Discover 검색창**: 상단 검색창에 KQL을 입력해 인덱스의 로그를 즉시 필터링한다.
- **Dashboard 검색창**: 대시보드 상단에 넣으면 모든 패널이 동시에 같은 조건으로 걸러진다.

S3
검색·필터 (KQL)

KQL 개요와 검색창

> KQL은 사람이 읽기 쉬운 문법으로 필드 기반 검색을 표현하는 언어다.

KQL이란

- **정의**: Kibana/OpenSearch Query Language의 약자로, 필드와 값 중심의 직관적 검색 문법이다.
- **입력 위치**: Discover 상단 검색창에 문법을 입력하고 Enter로 실행하면 결과가 즉시 갱신된다.
- **즉시 반영**: 시간 범위와 결합되어, 지정한 조건에 맞는 최근 문서만 히스토그램과 표에 나타난다.

필드 탐색 방법

- **Available fields**: 좌측 필드 목록에서 필드를 클릭하면 상위 값 분포가 떠 쿼리 작성을 돕는다.
- **원본 펼치기**: 문서 왼쪽 ▸를 누르면 원본 JSON이 펼쳐져 실제 필드명과 값을 확인할 수 있다.

S3

검색·필터 (KQL)

등호와 구문(정확 일치)

필드:값이 KQL의 기본이며, 공백·경로는 따옴표로 정확히 맞춘다.

```
# 이벤트 분류 (keyword 필드)
event: access
event: auth_fail
event: attack

# HTTP 메서드
http.method: POST

# IP 정확 일치 (따옴표 권장)
client_ip: "203.0.113.7"

# 경로 정확 일치 (구문 검색)
url: "/api/users"

# 로그인 시도 계정
user: admin
```

기본 규칙

- **필드:값**: 필드명과 값을 콜론으로 잇는 것이 모든 KQL 검색의 출발점이다.
- **구문 검색**: 공백이나 슬래시가 든 값은 큰따옴표로 감싸 정확 일치로 매칭한다.
- **대소문자**: keyword 필드는 대소문자를 구분하므로 값을 원본 그대로 입력한다.

S3
검색·필터 (KQL)

불리언 결합 AND / OR / NOT

> 세 연산자로 조건의 교집합·합집합·제외를 표현한다.

세 연산자

- **and 교집합**: 두 조건을 모두 만족하는 문서만 남긴다 — `event:attack and attack_type:sqli`.
- **or 합집합**: 둘 중 하나라도 맞으면 포함한다 — `http.status:401 or http.status:403`.
- **not 제외**: 뒤 조건에 맞는 문서를 결과에서 뺀다 — `event:access and not http.status:200`.

실무 활용

- **조건 좁히기**: and로 필터를 겹쳐 대상 로그를 정밀하게 줄여 노이즈를 걷어낸다.
- **조건 넓히기**: or로 유사 상태·유형을 묶어 관련 이벤트를 한 번에 관찰한다.

S3

검색·필터 (KQL)

괄호 우선순위와 여러 값

> 괄호로 평가 순서를 정하고 필드:(a or b)로 OR을 축약한다.

```
# 여러 값 OR 축약
severity: (high or critical)
http.status: (401 or 403)

# 괄호로 그룹 우선순위 지정
(geo.country:RU or geo.country:CN) and event:attack
event:attack and (attack_type:sqli or attack_type:xss)

# 우선순위 주의 (and가 or보다 먼저)
# 잘못: RU 전체 + CN 공격만 매칭
geo.country:RU or geo.country:CN and event:attack
# 올바름: 두 국가의 공격만
(geo.country:RU or geo.country:CN) and event:attack
```

우선순위 규칙

- **괄호 우선**: 괄호로 묶은 조건이 먼저 평가되어 의도한 그룹핑을 강제한다.
- **OR 축약**: `필드:(a or b)`는 같은 필드의 여러 값을 간결하게 OR로 묶는다.
- **and가 먼저**: 괄호가 없으면 and가 or보다 먼저 묶이므로 결과가 어긋난다.

S3
검색·필터 (KQL)

○ 기본검색
○ 불리언
● 범위·WC
○ 실전쿼리
○ 정리

범위 비교

> 숫자·날짜 필드는 부등호로 이상·초과 범위를 지정한다.

범위 연산자

- **이상·이하**: http.status >= 400 은 4xx·5xx를 모두 포함하며, <= 는 이하를 뜻한다.
- **초과·미만**: http.bytes > 30000 은 3만 바이트 초과, < 는 미만을 의미한다.
- **대상 필드**: number·date 타입에만 동작하며 keyword 값에는 등호를 쓴다.

활용

- **오류 탐지**: http.status >= 500 으로 서버 오류(5xx)만 추려 장애·공격 신호를 본다.
- **대용량·지연**: http.bytes > 30000 은 대용량 응답을, response_ms > 500 은 지연 요청을 잡는다.

S3

검색·필터 (KQL)

와일드카드와 존재 여부

> *로 부분 일치를, 필드:*로 필드 존재 문서를 검색한다.

```
# User-Agent 부분 일치 (스캐너 탐지)
user_agent: *curl*
user_agent: *sqlmap*

# 경로 부분 일치 (민감 경로 스캔)
url: *.env*
url: *wp-login*

# 필드 존재 여부 (공격 이벤트만)
attack_type: *

# IP 대역 부분 일치
client_ip: 203.0.113.*
```

와일드카드 규칙

- **부분 일치**: `*`는 임의 문자열을 뜻해 값의 일부만으로도 매칭한다.
- **존재 검색**: `필드: *`는 그 필드가 채워진 문서만 골라낸다.
- **성능 주의**: keyword·text에서 동작하며, 앞쪽 와일드카드는 검색이 느리다.

S3

검색·필터 (KQL)

KQL 치트시트

> 등호부터 괄호까지 핵심 연산 여덟 가지를 한눈에 정리한다.

연산	예	의미
등호	`event: attack`	필드 = 값
AND / OR / NOT	`a and b` · `a or b` · `not a`	불리언 결합
범위	`http.status >= 400` · `http.bytes > 30000`	숫자 비교
와일드카드	`user_agent: *curl*`	부분 일치
존재 여부	`attack_type: *`	필드가 존재하는 문서
구문(정확히)	`url: "/api/users"`	따옴표로 정확 일치
여러 값	`http.status: (401 or 403)`	OR 축약
괄호	`(a or b) and c`	우선순위

S3

검색·필터 (KQL)

실전 쿼리 (1) 분류·공격·오류

이벤트 분류와 공격 유형, 오류 상태코드를 걸러내는 쿼리다.

```
# 이벤트 분류별
event: access
event: auth_fail
event: attack

# 특정 공격 유형
event: attack and attack_type: sqli

# 인증/권한 오류 상태코드
http.status: 401 or http.status: 403

# 서버 오류(5xx)
http.status >= 500

# 4xx·5xx 오류 전체
http.status >= 400 and not http.status: 404
```

S3

검색·필터 (KQL)

실전 쿼리 (2) 정찰·유출

스캐너 UA와 민감 경로, 대용량 응답으로 정찰·유출을 탐지한다.

```
# 스캐너·자동화 도구 UA
user_agent: *sqlmap* or user_agent: *python-requests* or user_agent: *curl*

# 민감 경로 스캔 시도
url: *.env* or url: *wp-login* or url: "/admin"

# 대용량 응답(데이터 유출 의심)
http.status: 200 and http.bytes > 30000

# 특정 국가발 공격
geo.country: RU and event: attack
```

실전 쿼리 (3) IP 추적·브루트포스

한 IP의 전체 활동을 추적하고 로그인 실패 급증을 포착한다.

```
# 특정 IP의 모든 활동
client_ip: "203.0.113.7"

# 심각도 높은 공격
severity: (high or critical)

# 로그인 실패 POST (브루트포스 후보)
url: "/login" and http.method: POST
  and http.status: 401

# 관리자 계정 인증 실패
event: auth_fail and user: admin
```

추적 포인트

- **IP 추적**: client_ip로 한 출발지의 접근·공격·인증 실패를 시간순으로 모아 본다.
- **브루트포스**: 로그인 실패 POST가 급증하면 크리덴셜 스터핑 신호로 해석한다.
- **다음 단계**: 이렇게 좁힌 조건은 다음 세션 시각화에서 집계로 확장한다.

KQL과 Lucene·DSL의 관계

> KQL은 상위 문법이고, 내부적으로는 Query DSL로 변환되어 실행된다.

쿼리 언어 계층

- **KQL**: 사람 친화적인 상위 문법으로 검색창에서 필드·불리언을 간결하게 표현한다.
- **Lucene**: 검색창 토글로 전환할 수 있는 전통적 쿼리 문법도 함께 지원된다.
- **Query DSL**: 저장 검색·알럿은 내부적으로 JSON 기반 DSL로 변환되어 엔진에서 실행된다.

언제 무엇을

- **Discover**: 탐색·필터링은 빠르게 쓰고 읽기 쉬운 KQL로 처리한다.
- **규칙·API**: 탐지 규칙과 Dev Tools에서는 정밀한 제어를 위해 DSL을 직접 작성한다.

S3
검색·필터 (KQL)

세션 정리와 다음 예고

> KQL의 핵심 연산을 정리하고 다음 세션 시각화로 연결한다.

기억할 것

- **기본 문법**: 필드:값 등호와 따옴표 구문으로 정확 일치를 만들고, 값 나열은 괄호로 축약한다.
- **조건 조합**: and·or·not과 괄호로 조건을 좁히고 넓히며, 우선순위는 괄호로 강제한다.
- **범위·부분**: 부등호로 숫자·날짜 범위를, 와일드카드로 부분 일치와 필드 존재를 검색한다.

다음 예고 — 세션 4

- **시각화**: 이번에 좁힌 검색 조건을 Terms·Date Histogram 집계로 차트화하는 법을 배운다.
- **대시보드**: 공격 유형·상위 IP·트래픽 추이를 한 화면으로 조립해 관제 뷰를 만든다.