

진화하는 사이버공격에 우리가 AI로 현명하게 대처하는 방법

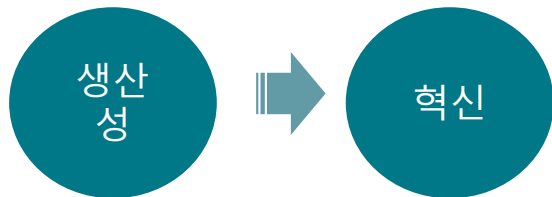
황성환 이사
김종식 부장

Cloudflare

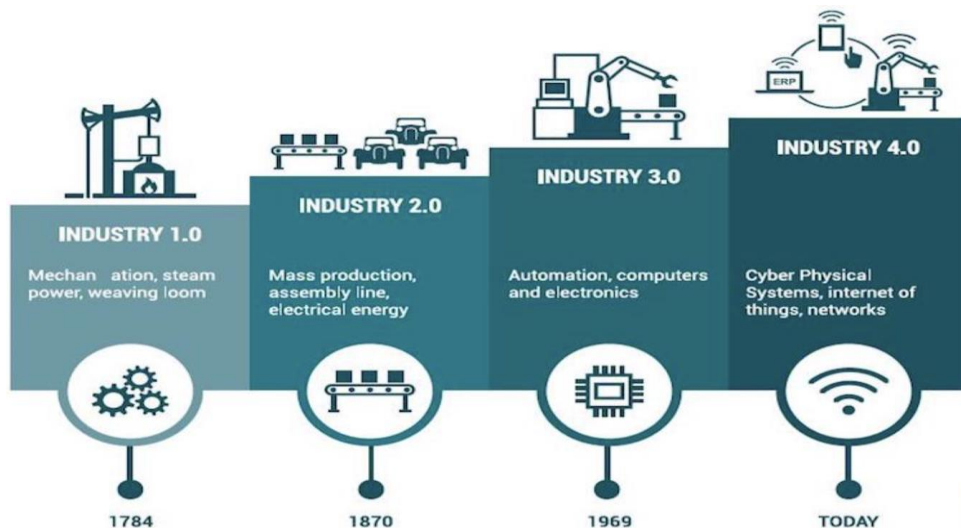
Agenda

- 1 AI 보안시대
- 2 Cloudflare Connectivity Cloud
- 3 Cloudflare AI 기반 보안 솔루션 소개
- 4 AI 보안 솔루션 (데모)

AI를 보안에 활용하는 이유?



보안관련 데이터 수집, 분석
 침해 탐지와 침해예측
 자동화된 위협 대응



IBM Watson for oncology 사례

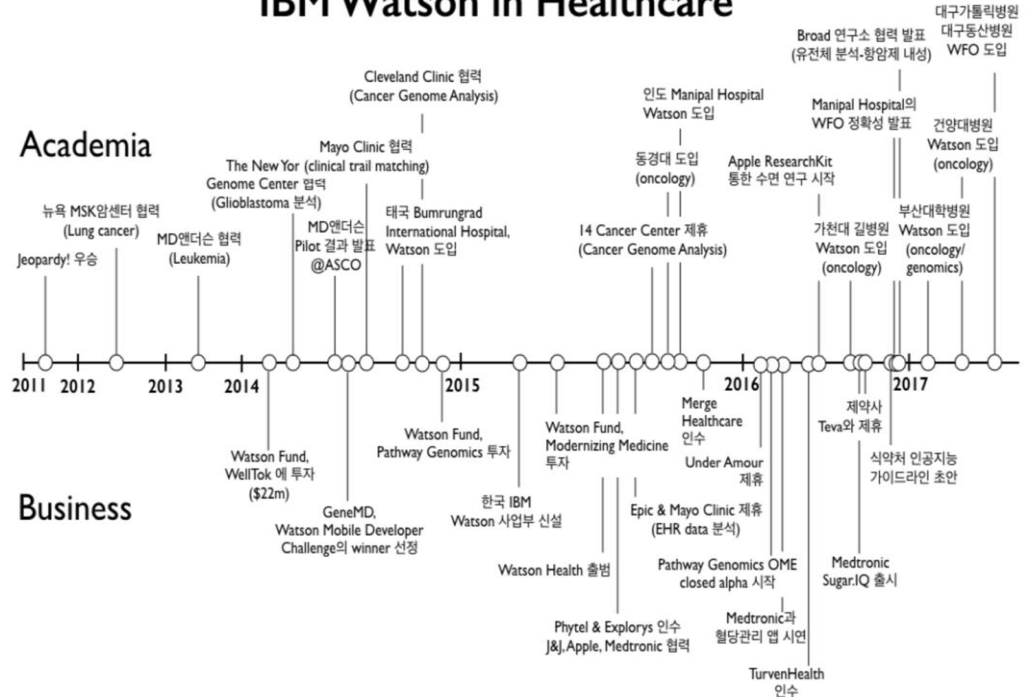
Watson 학습데이터

60만건의 학술자료
25,000건의 실제 치료 사례 학습
14,700시간의 보정작업

Watson 국내도입 사례

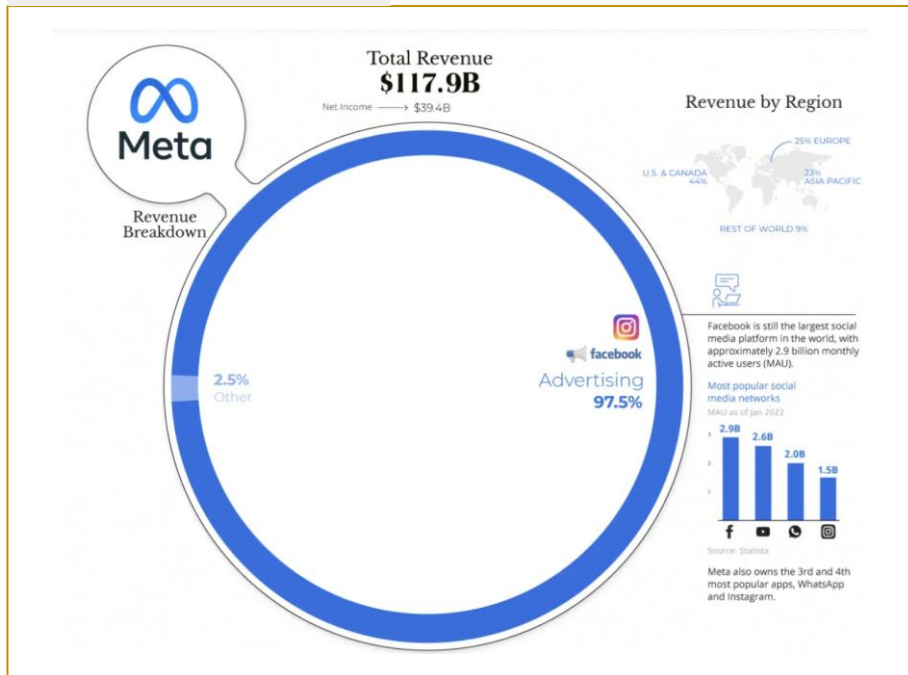
환자의 생존 확률(%) 제공
다면치료 도입 (여러 전문임상의 의견 취합)
영상 판독, 임상 실험, 치료연구제 개발 쉽지 않음

IBM Watson in Healthcare

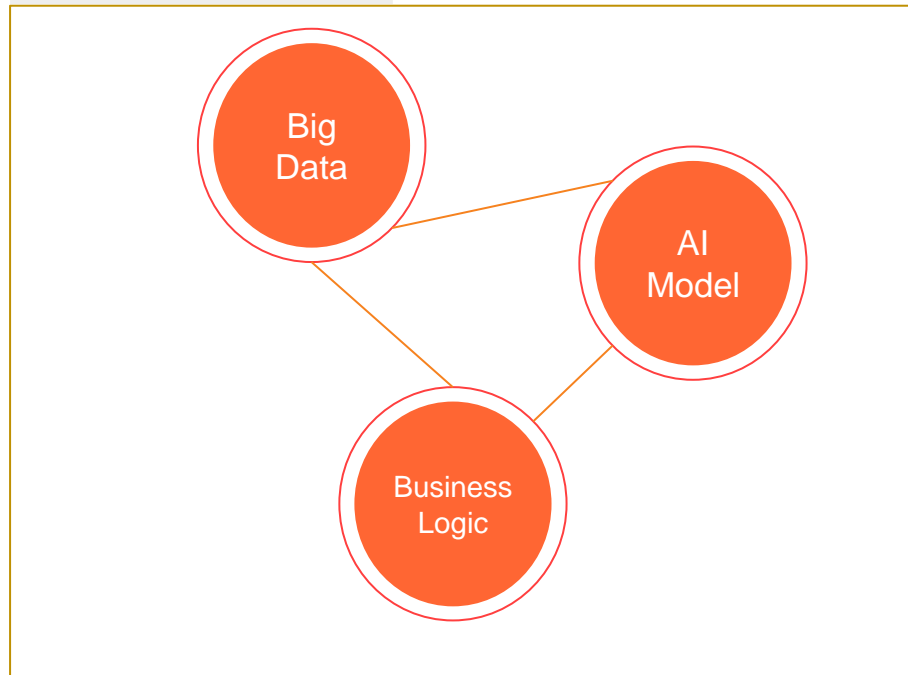


AI 기반 플랫폼 비즈니스

Meta 매출액 분석



플랫폼 비즈니스 조건



Cloudflare Connectivity cloud

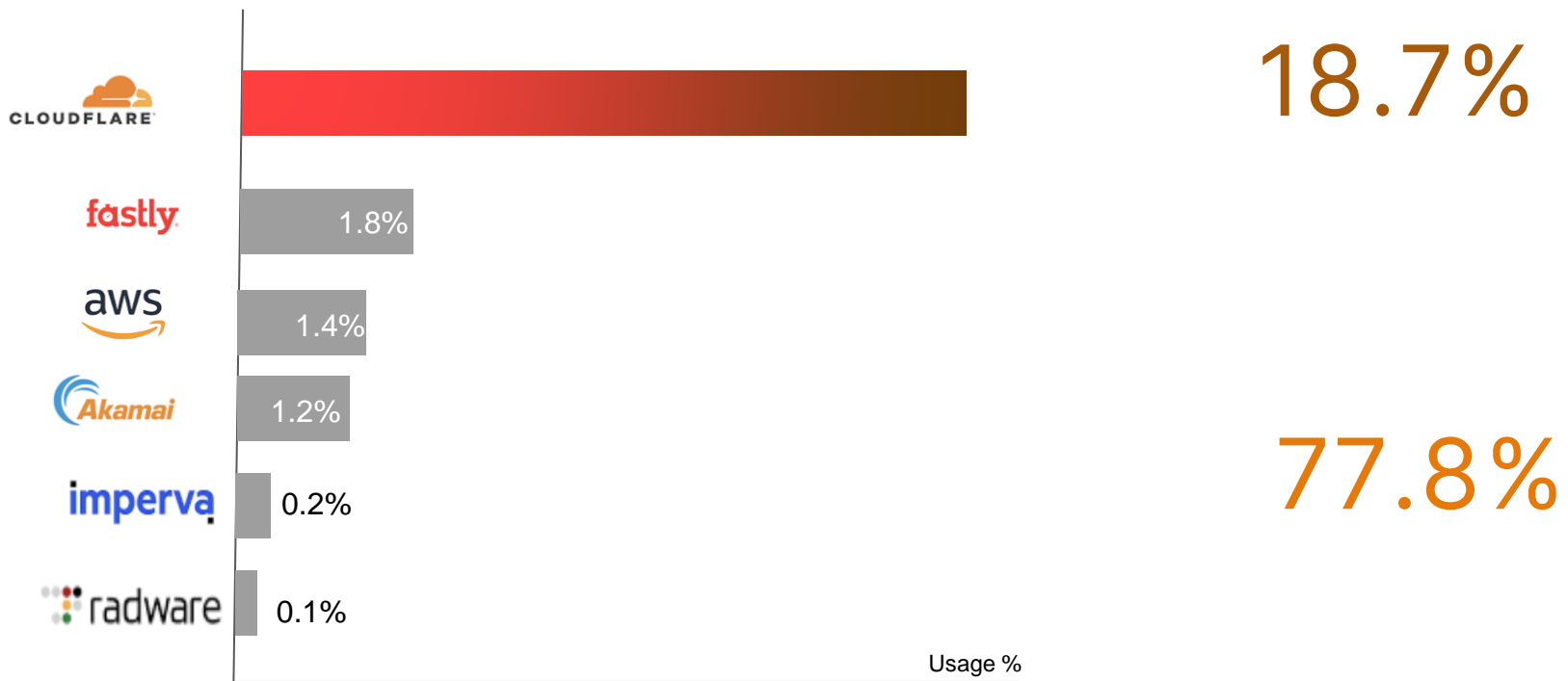
모든 네트워크와 통합

AI 인텔리전스

통합 인터페이스



전 세계 인터넷 트래픽 처리 현황



Source: W3Techs.com, 13 Nov 2023

Cloudflare AI 보안 솔루션 대응 단계



초당 64M 인터넷 트래픽을 처리함



HTTP Request Content

```
<svg/src=x%20%256fne/\rr%256fr%253d'\u{61}\u{6c}\u{65}\u{72}\u{74}'("1")();>
```



Normalization & Transformation

```
<svg src=x onerror=alert("1")(>
```



Feature Extraction

```
[1.0, 5.0, 2.5, ... 10.0]
```



ML Model Inference

```
[cf.waf.score: 3, cf.waf.score.sqli: 97, cf.waf.score.xss: 5, cf.waf.score.rce: 96]
```



Classification output in WAF

```
cf.waf.score
```

인터넷 트래픽 분석
(Signature & Ruleset)
-데이터 정형화/전처리

1. 신규 공격에 대해서 인터넷 트래픽 기반으로 검증 테스트

2. 검증용 인터넷 트래픽 없는 경우 AI 활용 검증용 **트래픽 생성** 후 테스트 진행

3. Bot management 같은 다른 보안 솔루션과 연계해서 **Attack Score** 산출함

Attack score 제공 및 **신규 Rule set 신속 제공**

Cloudflare Defensive AI



피싱 방지를 위한 AI 활용



이상 감지를 통한 API 보호



Defensive AI

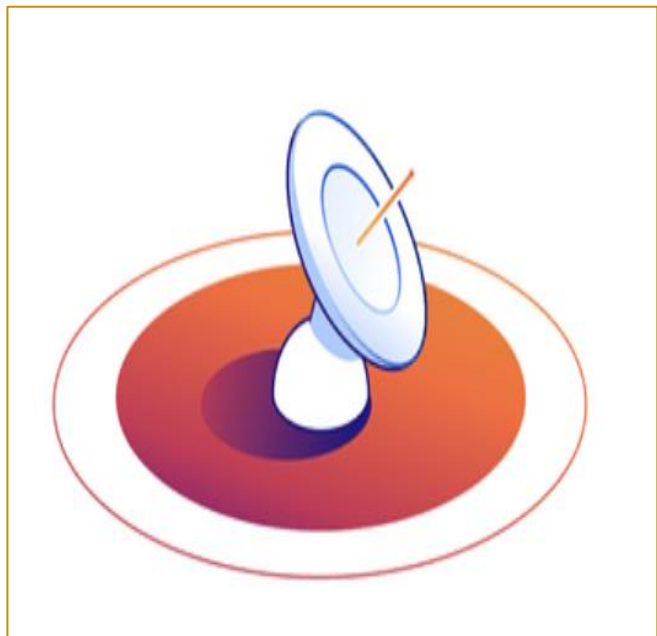


AI로 보호되는 데이터와 네트워크



알려지지 않은 애플리케이션 취약점 식별

Cloudflare AI 기반 솔루션 - [DDoS 보안]

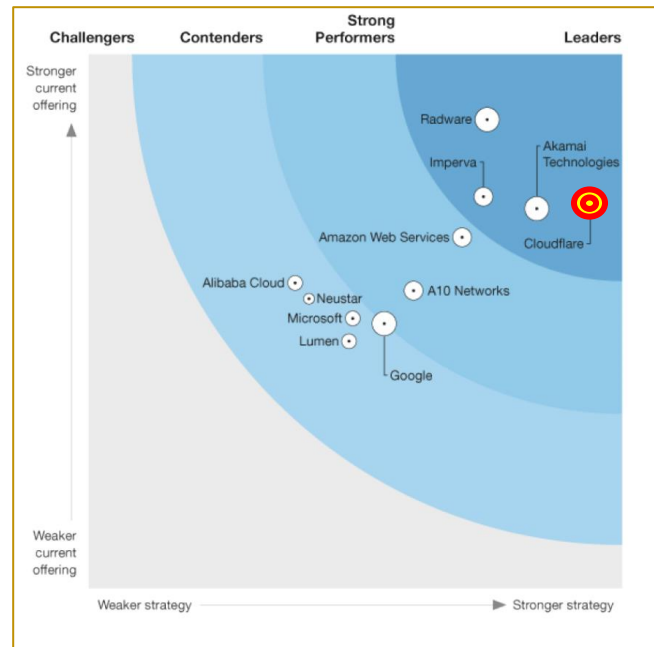


실시간 AI 기반 시그니처 배포

인터넷 트래픽 20%에 대한 트래픽 분석 및 공격방어를 통하여 실시간 DDoS 방어 시그니처 제공

AI 기반 고객사 트래픽 학습

고객사 트래픽을 ML기반으로 학습하여 평상시와 다른 트래픽이 탐지시에 자동 대응



Cloudflare AI 기반 솔루션 - [E-Mail 보안]



Password Expiration - Authentication Service

Hi User,

Password for your account will expire today. Please follow the link to update your account password.

Keep same Password

Support Service Desk Microsoft



Note: This verification is for it's intended receiver

이메일 트래픽 분석

AI 학습을 통하여 이메일 트래픽의 패턴 분석과 위험도 분석

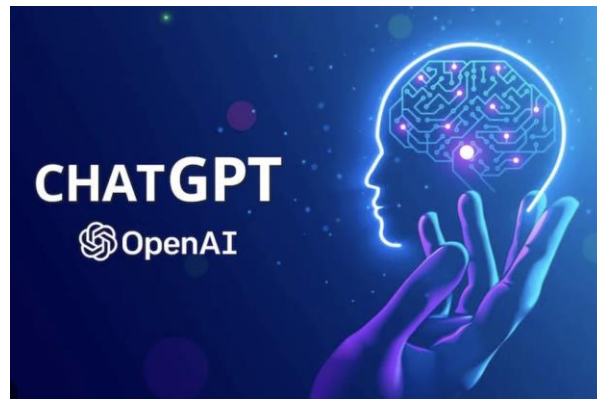
AI 기반 악성메일 검사 엔진

AI 기반 악성메일 검사 엔진(단어,문구,OCR 이미지 분석 외)기반하여 위험도 분석



Chat GPT의 보안은 ?

- 주간 사용자: 1억명 / 160개국 / 다양한 Device 접속
- 비정형 API 트래픽 / Bot 트래픽 보안



긴급 보안업데이트 대응사례



Cloudflare는 **Log4J 취약점** 발생시
 신규 4개의 Ruleset를 몇 시간안에
 Test 완료하여 배포 진행함



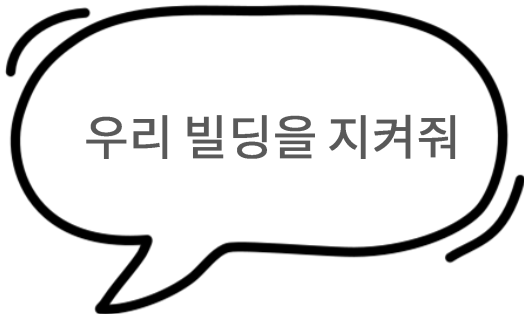
Confluence CVE-2022-26134
 취약점 발견 후 약 **30분** 안에 대응 업데이트
 버전 배포 진행함
 취약점을 목표로 한 해커 공격은
 약 **3.5시간** 후에 진행됨


 ivanti

CVE-2023-46805 / 2024-21887
 POC 진행후 **24시간** 이내
 신규 Ruleset 테스트 완료하여 배포
 진행함
 *기존 AI WAF로도 취약점 발견함
 (Attack score: 9~11점)



Origin

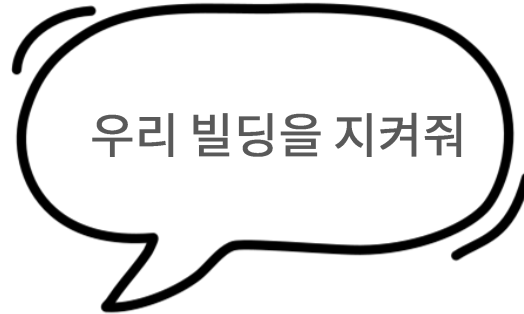
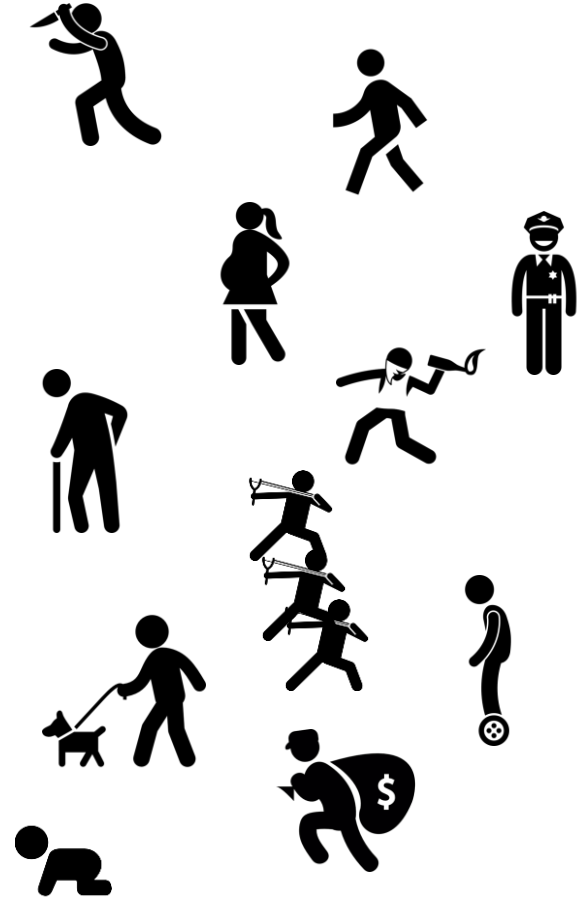


 Mr. WAF



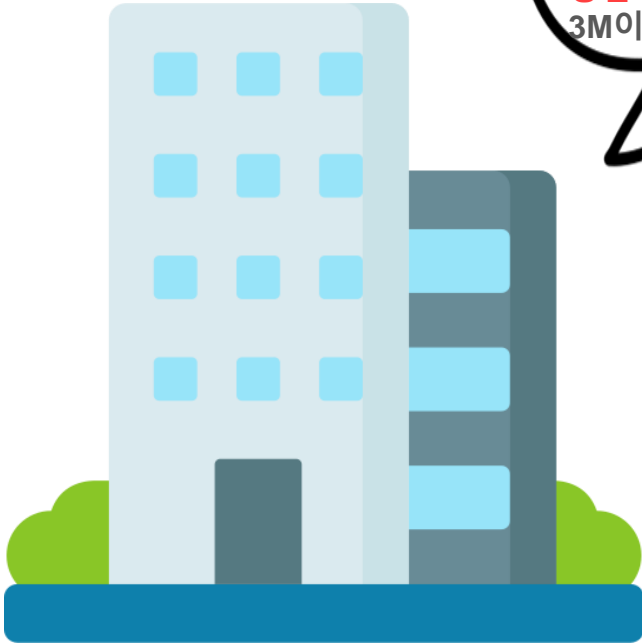
Visitors



 Origin Mr. WAF



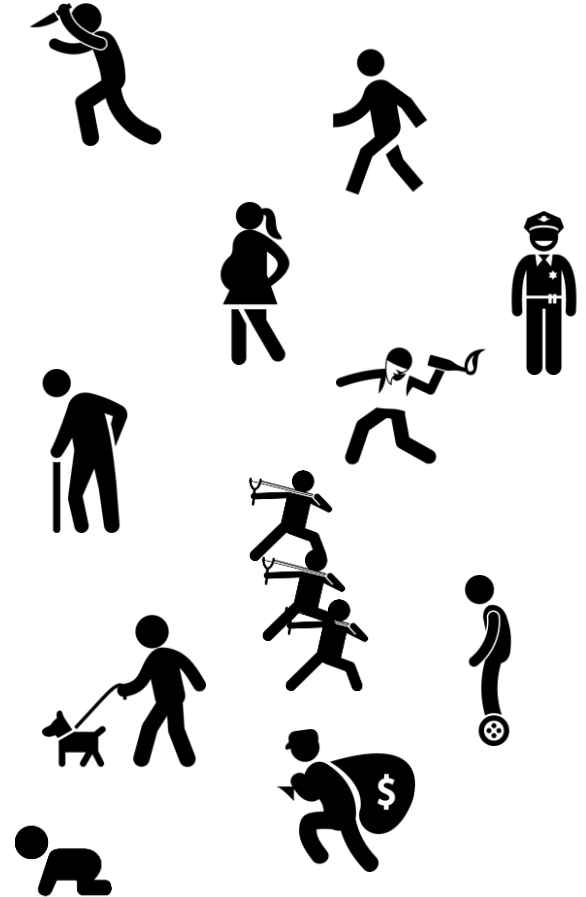
Origin



탈것을 타고있지 않으면서
무기를 들고 공격하려는
행동을 하지 않는 성인이어야
하며 어린이는 보호자를
동반하고 ... 동반이란
3M이내에 위치한... 무기란...



 Mr. WAF





Origin



20점 이하일 경우 출입을 막고
 50점 이하일 경우 검색을 하고
 80점 이하일 경우 이름을 적고
 81점 이상일 경우는 그냥 통과



 **Mr. WAF**



- Attack, score 10
 - Clean, XSS score 98
 - Clean, SQLi score 97
 - Attack, RCE score 12



- Clean, score 91
 - Clean, XSS score 97
 - Clean, SQLi score 95
 - Clean, RCE score 96



- Likely clean, score 78
 - Clean, XSS score 97
 - Clean, SQLi score 92
 - Clean, RCE score 86



- Likely attack, score 49
 - Clean, XSS score 97
 - Clean, SQLi score 98
 - Likely clean, RCE score 51

WAF rules



설정

Administrator



Requests/Packets



- Attack, score 10
 - Clean, XSS score 98
 - Clean, SQLi score 97
 - Attack, RCE score 12



- Clean, score 91
 - Clean, XSS score 97
 - Clean, SQLi score 95
 - Clean, RCE score 96



- Likely clean, score 78
 - Clean, XSS score 97
 - Clean, SQLi score 92
 - Clean, RCE score 86

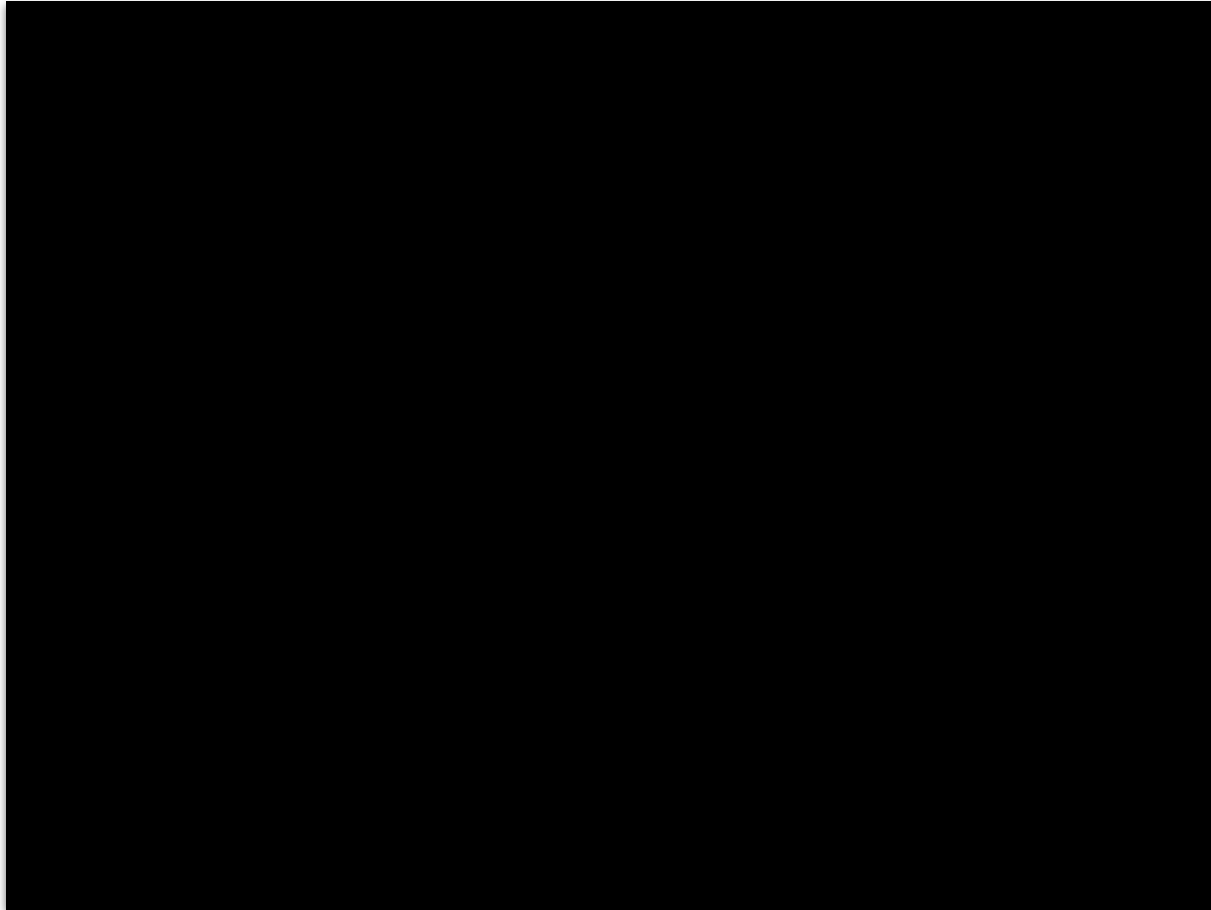


- Likely attack, score 49
 - Clean, XSS score 97
 - Clean, SQLi score 98
 - Likely clean, RCE score 51



Defensive AI

Demo



Thank you!

저희는 더 나은
인터넷 세상을 만드는데
주력합니다.



Thank you

 sunghwan@cloudflare.com / jongsik@cloudflare.com

 www.cloudflare.com/ko-kr