사람과 사람사이,

사람과 에이전트사이,

에이전트와 에이전트 사이

연결된 모든 사이버 공간을
안전하게 보호하는 시스템 보안 연구소

주식회사 사이시큐연구소

국내외 해킹사례

개인정보 유출,
올해 8월 사상 초유
'3000만' 건 돌파… 60%는 해킹

출처 : 시사저널

바이빗(Bybit) 14억 달러 해킹…
올 1분기 암호화폐 피해
20억 달러 육박

출처 : TokenPost

SKT 해킹 정보 9.7GB…
유심 핵심정보 등
300쪽 책 9천권 분량

출처 : 전국매일신문

KT, 개인정보 유출 정황
2만명 확산...
소액결제 피해 규모2억 4천만원

출처 : BLOTER

롯데카드 297만명 해킹당해…
28만명은 비번까지 털렸다

출처 : 조선일보

AI 사건 사례

민감 대화까지 무단학습한 AI 이루다…
법원, 4년 만에 배상판결

출처 : 연합뉴스

中 실험
“테슬라, 도로에 점 3개 찍었더니 역주행”

출처 : 노컷뉴스

7살 손가락 부러뜨린지도 모르는 체스선수…
커지는 로봇 우려

출처 : 중앙일보

딥페이크로 만들어진 상사에 속아…
340억원 보낸 홍콩 금융사 직원

출처 : 동아일보

美 자율주행차 첫 보행자 사망사고…
안전성 논란 증폭

출처 : 연합뉴스

북한 해커, 구글AI로
주한미군 정보 탐색 시도

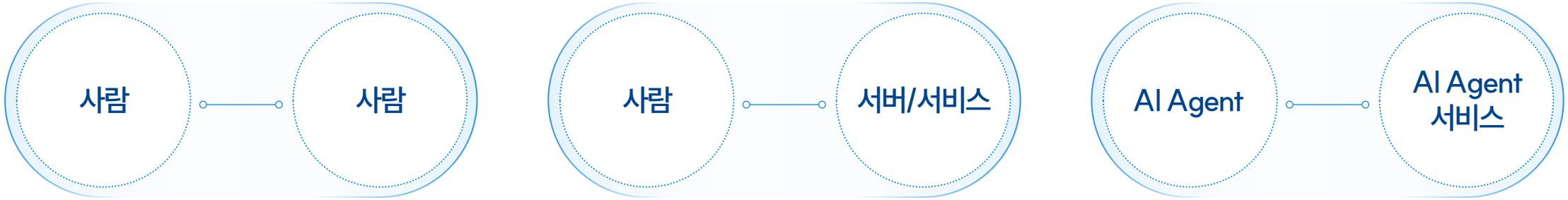
출처 : 연합뉴스

텍스트 입력

사이시큐연구소 (CySecuLab)

CySecuLab

“ 연결된 모든 사이버 공간을 안전하게 보호하는 시스템 보안 연구소 ”



연결된 모든 공간들 사이 사이의 안전한 교류와 실행을 보장하는 보안 시스템 기술 연구

카이스트 전산학부 정보보호대학원 사이버시스템 보안 연구실 (Cyber Security Research Lab)

Confidential AI, Hardware Security Monitor,
HW/Systems/Networks 보안 기술 영역의 세계 최고 수준 혁신 기술 및 역량 보유

사이시큐연구소 대표 · 연구원 경력 및 연구 분야

강병훈 Ph.D

- **사이시큐연구소 대표**
- **CSRC 상임고문**
- **KAIST 전산학부 및 정보보호대학원 정교수**
- UC Berkeley CS Ph.D.
- 시스템 보안 / 신뢰 실행 환경
- AI 플랫폼 방어 / Privacy 방어
- OS 커널 무결성 모니터링
- 하드웨어 기반 보안

송*호 Ph.D

- **KAIST 정보보호대학원 박사**
- 하드웨어 기반 보안 모니터
- AI 입력 난독화 기법
- GPU 메모리 보안
- Top-Security Conference (ACM CCS) 논문 게재
  - 하드웨어 모니터 "Interstellar"

김*우 Ph.D Candidate

- **KAIST 정보보호대학원 박사과정**
- 신뢰실행환경 원격증명 연구
- 신뢰실행환경 기반 HSM 연구

박*준 Ph.D Candidate

- **KAIST 정보보호대학원 박사과정**
- SW 기반 근거리 검증 기술
- 정형 검증 기술 개발
- SBOM 기술 연구개발
- BoB 3rd Top10
- 2015 Whitehat 국내 해킹대회 우승
- 2017 SECCON 국제 해킹대회 준우승

임*일 Ph.D Candidate

- **KAIST 정보보호대학원 박사과정**
- Dialect Perimeter Filter (DPF)
- 원격 메모리 시스템의 보안 연구
- 운영체제 / 커널 레벨 네트워킹 보안 연구
- DPF 관련 국내특허 1건

윤*조 Ph.D Candidate

- **KAIST 정보보호대학원 박사과정**
- Confidential AI 시스템
- AI 적대적 공격 및 방어 기법
- AI 시스템 프라이버시 보안
- BoB 9기 (보안제품개발트랙) Top 10
- 2025 Elsevier AI 분야 저널 논문 게재 (Knowledge-based Systems)

이*수 M.S.

- **KAIST 정보보호대학원 석사**
- AI 시스템 프라이버시 보안
- 시스템 API 보안
- SBOM 설계 및 보안 연구

사이시큐연구소 기업 소개

CySecuLab

당사는 KAIST의 정보보호대학원 소속 연구실의 기술을 기반으로
2023년 창업한 스타트업 기업으로, 대한민국 최고의 인력과 혁신적인 아이디어를
결합해 미래 정보 보안 기술을 선도하는 기업이 되고자 합니다.

사이버 공간을 안전하게 보호하는 것만 아니라, 컴퓨팅 실행 공간과
네트워크 연결을 통해 만나는 정보와 서비스,
그리고 **사람 간의 "사이"도 보호하는 솔루션을 개발** 합니다.

단순히 기존에 존재했던 정보보안 방법론을 뛰어넘어, 학계와 업계 양측에서
모두 주목받는 새로운 아이디어를 그 누구보다도 먼저 도입할 수 있는
기술력을 갖춘 당사는 엔드 디바이스부터 클라우드 환경까지 모두를 지키겠습니다.

기업 연혁

2023.11.30	KAIST 학내 벤처기업 창립
2024년	카이스트홀딩스 SAFE 투자유치의향서 획득
2025년	카이스트홀딩스 투자 확약서 및 특허/지식재산권 가계약 완료

사업참여 이력

24.5~24.11	통합보안모델 개발 시범사업 제로트러스트 기반엔드포인트 통합보안 플랫폼 (ZePP) 개발
25.5~25.11	제로트러스트 도입 시범사업 공항철도 및 가비아 대상 DPF를 활용한 제로트러스트 모델 구현 및 실증
25.5~26.1	중소벤처기업부 초기창업패키지사업
25.8~26.1	보안 내재화 AI 에이전트 구축 사업 전담

특허 및 지식재산권

출원일	구분	등록번호	발명의 명칭
2019.7.31	국내/등록	10-2197218	분산 ID와 FIDO 기반의 블록체인 신분증을 제공하는 시스템 및 방법
2020.9.11	국내/등록	10-2410810	확장 암호연산 처리 방법 및 시스템
2020.12.7	국내/등록	10-2398380	키 교환 방법 및 시스템
2021.1.12	국내/등록	10-2393537	신뢰실행환경에 기반한 소프트웨어 라이선스 관리 방법 및 시스템
2021.1.12	국내/등록	10-2375616	엔드 유저 인증을 위한 키 관리 방법 및 시스템
2021.4.16	국내/등록	1020210049779	네트워크에 연결된 시스템의 보안을 위한 프로토콜 다이얼렉트 기법
2021.7.13	미국/출원	17374084	Protocol Dialect For Network System Security
2022.7.15	미국/등록	US 12,177,350	B2. Extension Cryptographic Operation Processing System and Method

기업 역량을 한 줄로 요약해 기입해주세요

CySecuLab

기술 중심 기업

설립	KAIST 교수진과 석·박사 출신 연구진들이 설립 최신 시스템 및 AI 보안 연구를 실세계에 적용하고자 기업 설립
대표이사	KAIST 전산학부 정보보호대학원 정교수

최고 수준 연구진

연구진	6명의 전일제 연구진 및 12명의 시간제 연구진 모두 KAIST(한국과학기술원) 정보보호대학원/전산학부 석·박사 출신
전문가	학문적 깊이와 실전 경험을 겸비한 전문가 집단

세계적으로 검증된 성과

학문적 성과	국내 연구진 최초로 세계 4대 보안 학회 논문 게재, 이후 최우수 보안 학회 및 저널 다수 게제, 커미티/체어 역
실전 성과	국내 1위 & 해외 2위 해킹 대회 수상 경력연구진 보유, 연구결과 삼성 스마트 TV 보안 솔루션 탑재
지식재산권 확보	다수의 국내·해외 특허 보유로 독창성과 경쟁력 입증

보유 핵심 기술

TSA	AI 안전 및 Agent 보안 기술 Trusted Safety Agent	CC-BOX	기밀 컴퓨팅 보안 기술 Confidential Computing
TPA	AI 개인정보 보호 기술 Trusted Privacy Filter Agent	CC-SBOM	안전한 공급망 보안 관리 기술
DPF	네트워크 보안 기술 Dialect Perimeter Filter	Outerstellar	하드웨어 기반 모니터링 기술

사이시큐연구소 보안 솔루션 소개

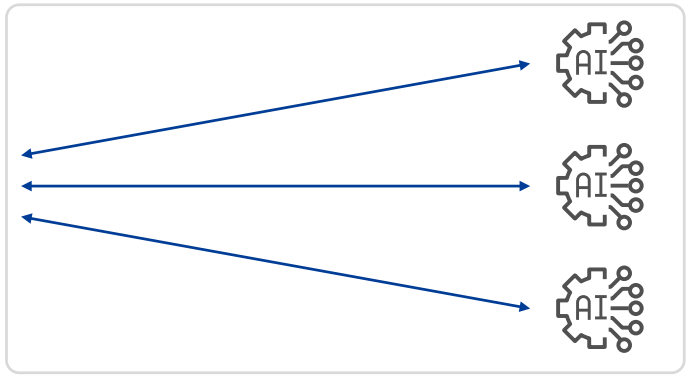
CySecuLab

TPF Trusted Privacy Filter

사용자 개인정보 보호 기술

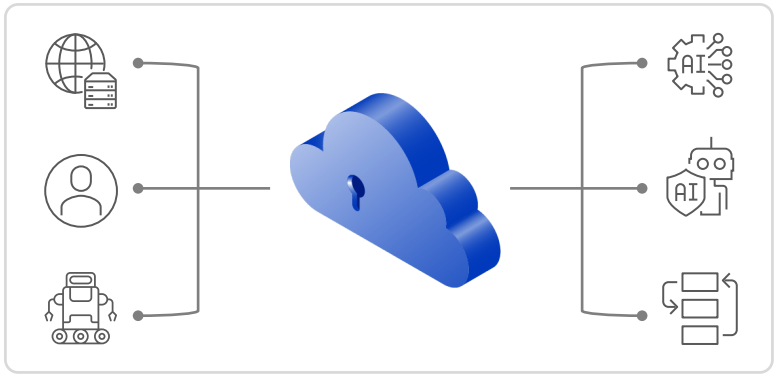
TSA Trusted Safety Agent

AI 모델 안전성 보장 기술



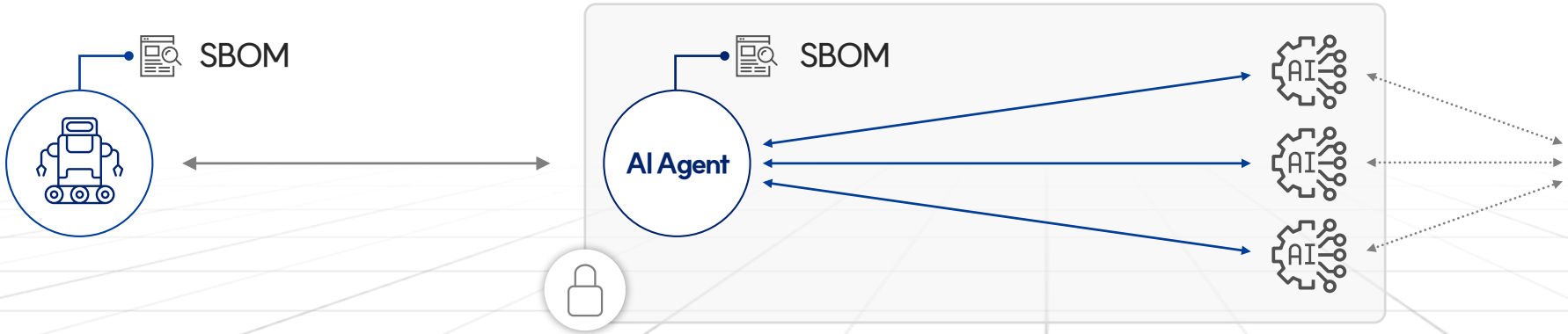
DPF Dialect Packet Filter

네트워크 보안 기술



CC - SBOM

소프트웨어 공급망 보안 기술



Outerstellar

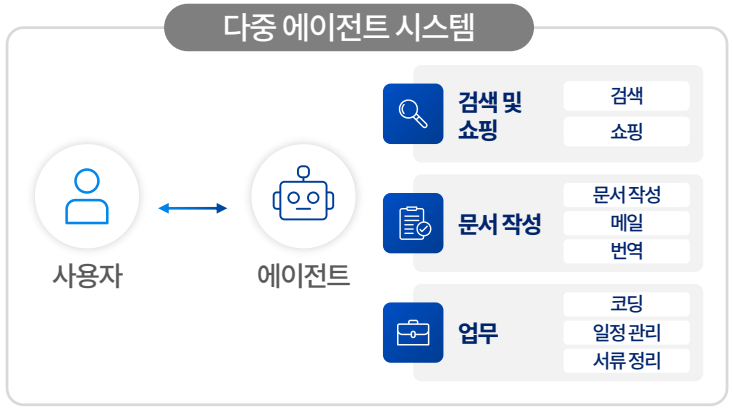
하드웨어 기반 모니터링 기술

TSA

Trusted Safety for multiple AI Agents

What is TSA?

- 다수의 AI 에이전트들이 다양한 작업을 **상호 안전하게 수행**할 수 있도록 하는 프레임워크
- 특히 각 AI 에이전트 입출력에서의 **잠재적인 위험성**을 모델의 중간 표현과 입력을 통해 **제어하는 보안 솔루션**



What is AI agent?

- 자율적으로 목표를 해석하고 **계획, 도구 호출, 코드 실행**을 통해 **컴퓨팅 및 실환경**에서 작업을 수행하는 소프트웨어 에이전트
- **사용자·시스템과 상호작용**하며 관찰 > 의사결정 > 행동 > 피드백의 루프를 반복해 과제를 완료하고 성능을 개선

What can TSA do?

●○○

통합형 멀티 에이전트 프레임워크 구축

다수의 AI 에이전트의 상호작용에서 발생한 예상치 못한 출력은 시스템 전체에 악영향을 미칠 수 있으며, 이러한 복잡한 AI 에이전트들의 행동을 조절, 제어할 수 있는 통합형 프레임워크를 구축

●●○

에이전트 입출력 보호

각 에이전트들은 민감정보, 개인정보 등을 다룰 수 있으며, 이 과정에서 사회적·윤리적 기준을 위반하는 출력을 생성하고 시스템 혹은 소비자에게 전달할 수 있음. 이를 방지하고 안전한 작업을 수행하기 위해 입출력을 보호

●●●

에이전트 모델 보호 및 라이선스 관리

각 에이전트가 외부로 노출되지 않도록 모델 탈취 방지, DRM 라이선스 관리 등을 통해 에이전트 기능이 외부에 노출되지 않도록 보호

Trusted Privacy Filter Agent

TPA

CySecuLab

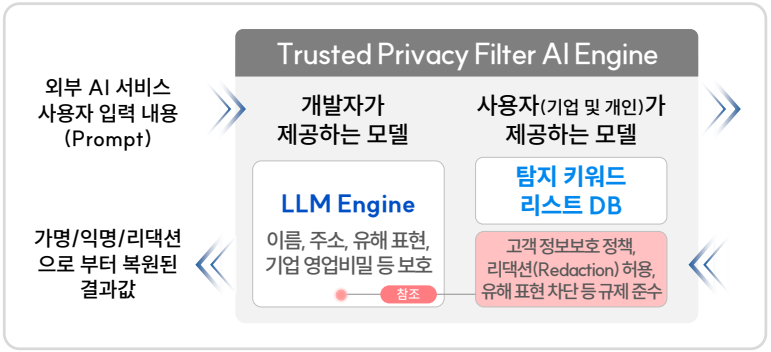
TPA

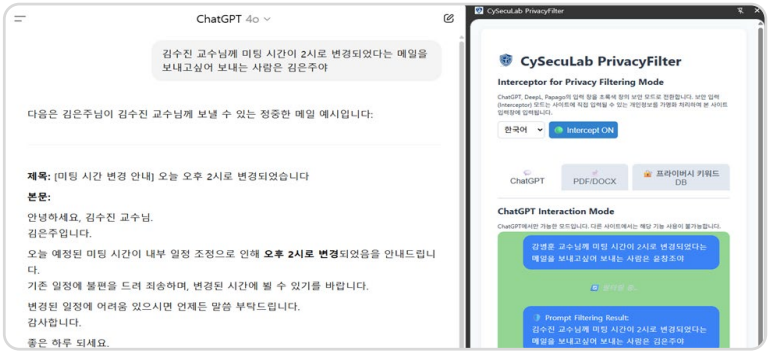
Trusted Privacy Filter Agent

What is TPA?

- AI 시스템에 **개인정보 및 민감정보가 들어가는 것을 방지**하여 AI를 안전하게 사용할 수 있게 도와주는 **보안 솔루션**
- 내부 LLM 엔진이 탐지 키워드, 유해표현 차단, 정보 보호 규제 등의 데이터를 참조하여 사용자에게 **맞춤형으로 개인정보와 영업비밀을 보호**할 수 있음
- 모델 탈취(Model extraction attack) 및 **여러 AI 모델 관련 공격*으로부터 방어**
- 모델의 결과 값에 유해/공격 **코드 생성 및 실행 방지**

* Membership inference attack, model inversion attack, feature snooping attack, backdoor attack, adversarial attack, etc.





Why TPA?

탐지되는 **모든 개인 식별 정보 (PII) 및 기관/기업/영업비밀은 가명화 방식**을 사용하여 필터링 여부가 드러나지 않게 처리

AI 기반 맥락 분석 기술을 활용하여 **정보의 의미를 유지한 채 가명화를 수행**

웹 브라우저 확장 프로그램 형태로 **다양한 종류의 외부 AI 서비스에 손쉽게 대응 가능**

DPF

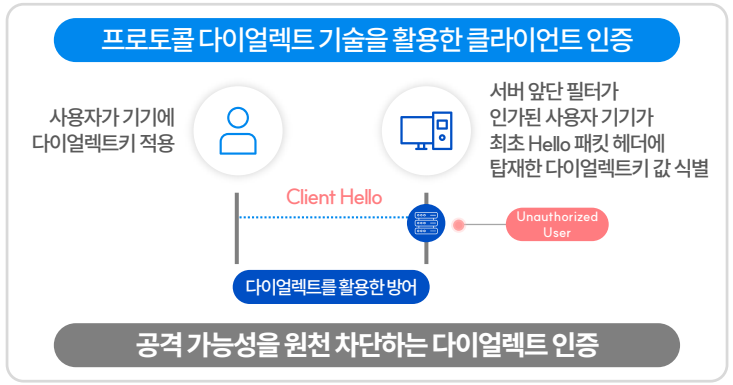
CySecuLab

DPF

Dialect **P**erimeter **F**ilter
Toward Secure Network System

What is Dialect?

- 미국의 ONR 연구소로부터 검증받은 KAIST 학내 벤처 기업 사이시큐 연구소의 특화기술
- 프로토콜 다이얼렉트를 이용한 네트워크 **공격자 사전 식별** 기술
- **기존 프로토콜과 완벽 호환**이 되어, 큰 변경 없이, 보안성 강화 가능



ZT via Dialect

- Zero-Trust(ZT)는 네트워크 보안의 기존의 경계 보안 체계와는 차별화된 **새로운 탈경계화 기반의 접근 방법론**
- Zero-Trust에 활용되는 기존 기술은 공격 표면이 넓고, 단일 패킷 인증(SPA) 방식 조차 인증 후 일정 시간 동안 **해커에게 취약한 허점 존재**
- DPF (Dialect Perimeter Filter) 기술은 Dialect 프로토콜을 활용한 소프트웨어 경계 기술로서, **취약점 악용을 사전에 차단하는 기술**

Dialect of Things

기존 프로토콜에 완벽 호환이 되어
광범위한 적용이 가능

모빌리티, **수자원 공사**, **네트워크 인프라 시스템** 등 산업 전반

CC-BOX

Confidential Computing Platform

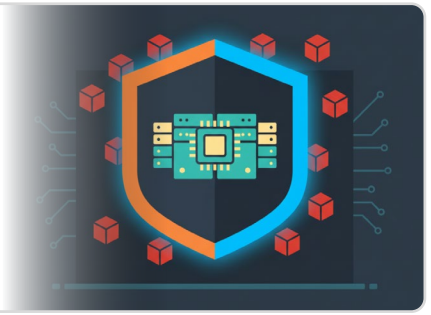
What is CC-BOX?

- 하드웨어 기반 기밀성 보장 실행 기술을 통해 애플리케이션 코드 및 데이터들을 안전하게 보호해줄 수 있는 **CySecuLab의 보안 솔루션**
- CC-BOX의 플랫폼 정보 및 애플리케이션 실행 정보에 대한 원격 증명 기능 제공

Why CC-BOX?

- 개인/기업 **사용자**의 정보 유출 가능성 원천 차단을 보장받으며 서비스 사용 가능
- 서비스 **제공자**는 클라우드 서버 소유자로부터 소프트웨어 및 데이터 **자산 보호** 가능
- 서비스 사용자 및 제공자 모두 안심하고 사용할 수 있는 솔루션

저장 되어있는
데이터 뿐만 아니라
연산중인 데이터 역시
안전하게 보호



Application of CC-BOX

규제 준수

데이터 프라이버시 및 보안 규제 준수를 위한 기밀성 보장 보안 기능 제공

고신뢰 민감데이터 보호

금융/의료/기업 비밀에 해당하는 데이터를 연산 및 실행시에도 안전하게 보호

지적 재산(IP) 보호

비즈니스 정보와 지적 재산을 서비스플랫폼 제공자로부터 효과적으로 보호

CC-BOX는
보안 규제 준수를 위한
기밀성을 보장할 뿐만 아니라,
그 적용이 간편해 적용가능한
시스템의 범위가 넓습니다.

CC-SBOM

Confidential Computing SBOM

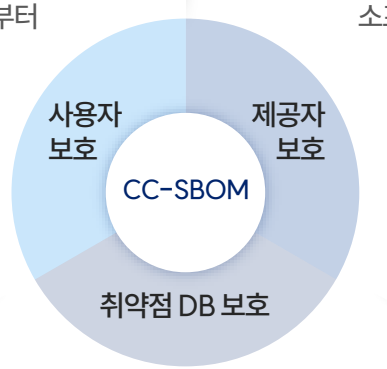
What is CC-SBOM?

- CC-SBOM은 자사 제품인 CC-BOX라는 안전한 실행 환경과의 연계를 통해 소프트웨어 구성 명세서(SBOM) 생성을 지원하는 CySecuLab의 보안 솔루션
- 특히, 보안 및 규제 준수가 중요한 **의료, 금융, 소프트웨어 통합관리** (OEM, 외주제품 등), **로봇** 및 **산업 제어 시스템** 분야 등에 특히 적합

Why CC-SBOM?

사용자의 기밀정보 (소스코드 등)를 정보 유출 위협으로부터 안전하게 보호

클라우드 및 서버 소유자로부터 소프트웨어 자산과 IP를 보호



취약점 데이터베이스 공급자의 정보자산을 안전하게 보호

사용자는 CC-SBOM을 활용해 실시간 취약점 관리 서비스 가능

CC-SBOM은 사용자가 제출한 중요 정보를 안전하게 보호해 사용자 이외에는 해당 정보에 접근할 수 없게 보호

Outerstellar

Hardware-based instruction-flow
safety monitor

What is Outerstellar?

- CPU 코어에서 실행되는 **소프트웨어의 명령어 (Instruction)를 하드웨어 수준**에서 실시간으로 추적하여 악성 행위를 사전에 탐지하고 예방하는 혁신적인 보안 솔루션
- **하드웨어 수준에서 직접 보안 관제**를 진행하기 때문에 CPU에서 실행되는 악성 행위를 직접 관제하며, 공격자들이 **보안 관제를 방해하는 것 역시 불가능**

Why Outerstellar?

- 소프트웨어의 한계(권한 수준, 성능 저하 등)를 뛰어넘는 근본적인 보안 모니터
- **실행 애플리케이션 및 데이터의** 강력한 보호

What can Outerstellar do?

AI 모델의 실행 무결성 보장

AI 모델이 의도된 프로그램 흐름대로 실행되는지 검증

하드웨어 기반의 이중 보안 체계

- **화이트리스트:** 기반 보안 특수 설계된 제한적 컴퓨팅 환경에서는 사전에 허용된 프로그램의 실행 흐름을 화이트리스트 방식으로 보장
- **블랙리스트 기반 방어: 악성 코드의 명령어 패턴을 블랙리스트 기반으로 실시간 탐지 및 차단하여 시스템을 보호**

반도체에 직접 탑재되는 보안 솔루션

Interstellar는 CPU 등 반도체 칩 내부에 직접 탑재

- **GPU**를 포함한 **AI 가속기**, 무인기용 **저전력 반도체 등 특수 목적 하드웨어**에도 통합 가능

클라우드 컴퓨팅 환경의 안정성 모니터

기밀 컴퓨팅 환경에서도 시스템의 안정성을 **은밀하고 선제적으로 모니터링** 가능

사람과 사람사이,

사람과 에이전트사이,

에이전트와 에이전트 사이

연결된 모든 사이버 공간을 안전하게 보호하는
시스템 보안 연구소가 되겠습니다.

CySecuLab