클레비(Clevi) 회사소개서





COPYRIGHT 2025 © CLEVI INC. ALL RIGHT RESERVE

CONTENTS

01 회사 개요

회사 소개

회사 연혁

02 시장과 문제 정의

산업에서 AI 적용의 한계(스몰 LLM/CNN/RAG) 패러다임의 전환

03 Clevi – X Intelligence Platform

Clevi-X-Platform 요약

대규모 분산 딥러닝 인프라

Teacher-Student Model(Knowledge Distillation)

CoT Model

Semantic Database

Tool Calling

04 Clevi Model Overview

Overview

Ivy-3-text

Cip-4-mm

Cip-5-x

Cip-5-agent

Cip-5-vision

05 Solutions

코딩 에이전트

Physical Al

VLM-Vision Intelligence Model

ARI: 워크플로우 자동화 시스템

보안 에이전트

06 On-premises

하드웨어 구조

도메인 적용 워크플로우

격리 환경

07 도입 시나리오

DevOps

R&D

HR

화재 재난 관제

댐 수위 관제

01

회사 개요

회사 소개 회사 연혁



01 회사개요: 회사 소개

Clevi는 Wide-Spectrum Al 기업입니다.

주식회사 클레비(Clevi)는 2022년 09월 이환호 대표에 의해 설립된 Al Platform 전문 기업입니다. 현재 경기도 화성시에 본사를 두고 있으며, 국내 Al 기업으로는 유일하게 초거대 언어모델(LLM), 비전 인텔리전스 모델(VLM), 피지컬 Al 트레이닝 플랫폼, 옴니모달, 레드팀 보안 에이전트, 코딩 에이전트 등 Al 전 영역을 아우르는 와이드 스펙트럼 기술 역량을 확보하고 있습니다.

회사명	(주)클레비	설립일	2022년 9월
대표 이사	이환호, 김홍기	업종	정보 통신업
본사 위치	경기도 화성시 동탄 첨단산업 1로 27 금강 펜테리움 IX타워	홈페이지	www.clevi.ai

01 회사개요: 회사 연혁

우리의 시작과 여정

AI가 세상을 바꾸는 시대, 누구나 고성능 AI를 활용할 수 있는 인프라를 만들겠다는 목표로 시작했습니다. 오늘까지의 도전과 성장 과정을 통해 더 견고한 AI 생태계를 구축해 가고 있습니다.



02

시장과 문제 정의

기존 산업에서 AI적용의 한계(스몰 LLM/CNN/RAG) 패러다임의 전환



스몰 LLM의 한계

커스텀 LLM은 실제 현장 적용과 최신 정보 반영에 한계가 있습니다.

현장 적용한계

- ✓ PoC에서는 성공하더라도 예측 불가능한 변수와 복잡성에 직면하면 오류 발생
- ✓ 반복적인 재학습과 튜닝이 필요해 비용 가중

데이터 한계

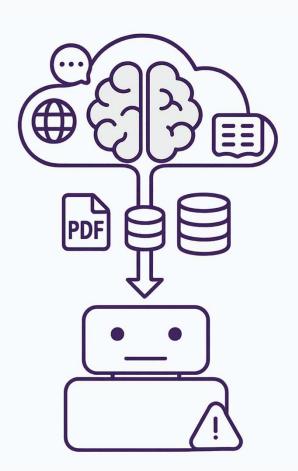
- ✓ 특정 도메인에 한정된 데이터
- ✓ 범용성 부족으로 새로운 상황 불가

원본 모델 의존

- ✓ 오픈소스 모델의 구조와 데이터에 종속
- ✓ 경쟁사와 유사 모델 양산, 독자성 약화

최신성/유연성 부족

- ✓ 최신 기술 적용이나 구조 변경이 불가
- ✓ 최신 정보 반영 어려움



CNN기반 시스템의 한계

실제 산업 현장에서는 기존 CNN기반 시스템이 다양한 변화와 복잡한 상황에 효과적으로 대응하지 못해 한계가 드러납니다.

블랙박스문제

- ✓ CNN의 의사결정 과정이 불투명해 결과 해석 및 설명이 어려움
- ✓ XAI기술이 미흡하여 모델의 투명성 확보의 어려움

상황 변화에 대한 한계

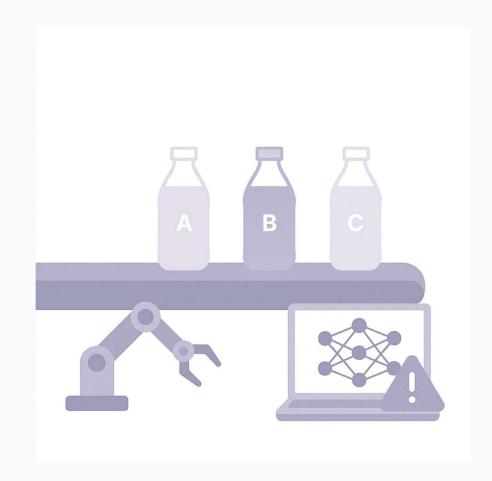
- ✓ 단순한 "패턴차이"만 감지하여 정책/규정과 연결 불가
- ✓ 정책, 규정, 컴플라이언스 등 외부 기준을 모델에 직접 반영의 어려움
- ✓ 학습된 객체·상황 이외에도 정확도 급락

확장성 및 적용 범위 제한

- ✓ 환경이나 도메인 변화 시, 기존 모델의 재사용 및 확장이 어려움
- ✓ 학습되지 않은 패턴에는 일반화가 약하여 성능 보장 어려움
- ✓ 커널 크기와 네트워크 깊이에 따라 정보통합 범위가 제한됨

계산량 및 자원 소모

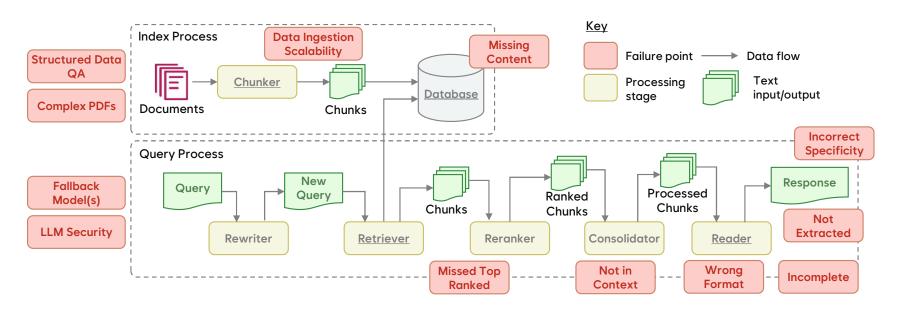
- ✓ 구조가 복잡할수록 연산량 메모리 사용량 급증
- ✓ 사용량 급증으로 인한 도입 및 유지 비용 부담 증가



<공정에서 도입 시 탐지 대상이 바뀔 때마다 인식하지 못하는 모습>

RAG기반 AI의 한계

RAG 시스템의 대표적 문제점



Source: Wenki Glantz | Medium

RAG기반 AI의 한계

RAG 기반 AI는 기존 LLM의 한계를 보완하지만, 데이터가 많아질수록 검색 품질과 정확성, 맥락 이해가 떨어져 현장 신뢰성이 낮습니다.

검색품질 저하

- ✓ 데이터(문서)가 많아질 수록 관련성 높은 정보찾기의 어려움
- ✓ 중요한 문서가 누락되거나 불필요한 문서가 함께 검색되어 결과의 신뢰도 하락

정확성 및 맥락 이해 한계

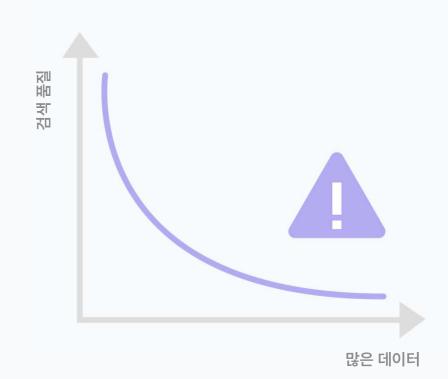
- ✓ 검색된 문서가 질문의 의도와 정확히 불일치하여 답변 신뢰도 저하
- ✓ 문맥을 충분히 반영하지 못해 답변의 신뢰성과 일관성 하락
- ✓ 같은 질문에도 다른 문서를 불러와 일관성 있는 답변불가

실시간 업데이트 및 확장성 한계

- ✓ 새로운 데이터가 추가되어도 검색 시스템에 즉시 반영되지 않아 최신성 부족
- ✓ 데이터가 많아질수록 검색 속도와 품질이 저하되어 대규모 환경에서의 확장성 부족
- ✓ 데이터가 계속 증가할수록 시스템 처리비용이 늘어 운영리소스 과다소모

보안 및 규제 대응 취약

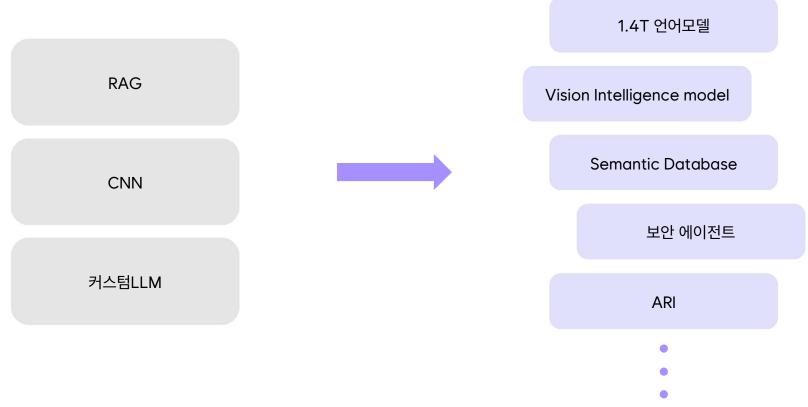
- ✓ 검색 과정에서 불필요한 민감 데이터 노출 가능성 있음
- ✓ 데이터 내 규제·컴플라이언스 요구에 대응하기 어려움



02 시장과 문제 정의: 패러다임의 전환

페러다임의 전환

Clevi-X Platform은 자체 파운데이션 모델을 바탕으로 CTA, Toolcalling, Semantic Database 등 첨단 AI 기능을 통합한 차세대 인공지능 플랫폼입니다. 이를 통해 다양한 AI 에이전트와 서비스를 쉽고 유연하게 구축할 수 있으며, 다양한 환경에 맞춘 맞춤형 솔루션 제공이 가능합니다.



Clevi – X Intelligence Platform

Clevi-X Platform 요약

대규모 분산 딥러닝 인프라

Teacher-Student Model(Knowledge Distillation)

CoT Model

Semantic Database

Tool Calling

03 Clevi-X Intelligence Platform: Clavi-X-Platform 요약

안정적인 데이터 보호 체계 및 고품질 데이터를 바탕으로 맞춤형 AI솔루션과 고도화된 분석을 통해 인사이트 제공

데이터 품질 및 관리

적대적 데이터 증강 (Adversarial Data Augmentation)

가상 샘플 생성(텍스트 무작위 변환, 이미지 회전·반전 등) 및 적대적 및 공격적 데이터 생성 및 분류로 고품질 데이터 세트 구현

정밀 리뷰 (Expert Review)

전문가가 샘플 단위로 리뷰하여 최종 정확도 보장 재귀적 피드백이 가져다주는 이점

재학습 (Continual Learning) 기반 교정

모델이 지속적으로 자신의 예측 결과를 반영해 데이터 품질을 점차 개선

SVCE

Secure Container Environment

네트워크 파일 메모리 등 다층 격리로 외부 침입을 최소화

Al-Driven Orchestration

LLM이 컨테이너 자원 할당과 스케줄링 결정을 자동으로 내림 → 최적화된 부하 분산

Zero-Trust Architecture

LLM이 컨테이너 자원 할당과 스케줄링 결정을 자동으로 내림 → 최적화된 부하 분산

Stateful Container 관리

세션 정보, 사용자 맞춤 데이터 등을 컨테이너 내부에서 안전하게 지속 및 백업

ARI

Role-Based Prompting

'데이터 수집 담당', '품질 검수 담당' 등 역할별 명령을 할당하여 AI 간 협업을 체계화

명시적 Input/Output 정의

각 단계의 입력과 출력을 분명하게 설정함으로써 에러를 최소화, 협업 시 인과관계가 명확

체크포인트(Checkpoint)기반 검증

단계별 산출물을 저장 검증해 안전하고 일관된 프로세스 구축

멀티 모달 연계

텍스트, 음성, 이미지, 센서 데이터 등을 연동하여 더욱 다양한 워크플로우 개발 가능

vNext

Chain-of-Thought (생각의 사슬)

모델이 답을 도출하는 모든 중간 단계를 투명하게 보여주어 신뢰도 상승

Self-Critique & Correction

모델 스스로 오류 가능성을 검토하고, 필요하면 즉시 수정 제안

Explainable AI (XAI)

의사결정 로직을 시각적으로 표현하여, 인간이 결과를 검증하고 학습 과정에 참여 가능

Multi-Task Mastery

단순 질의응답을 넘어 코딩 보조, 문서 요약, 데이터 분석 등 다양한 업무를 동시에 수행

03 Clevi-X Intelligence Platform : 대규모 분산 딥러닝 인프라

고효율 대규모 연산 인프라

기존 AI 인프라는 서버가 많아질수록 느려지고, 장애에 취약해 대규모 모델 운영이 어렵습니다. Clevi는 수백 노드 이상의 GPU 인프라를 활용하여 1.4조 개 파라미터 규모의 초거대 언어모델 학습에 성공했습니다.

대규모 연산 시 문제점

병목현상: 노드 수 증가 시 네트워크 동기화 비용 급증

메모리 한계: 파라미터·Gradient 저장 공간 부족

불균형 문제: 일부 노드 속도가 전체 성능 저하

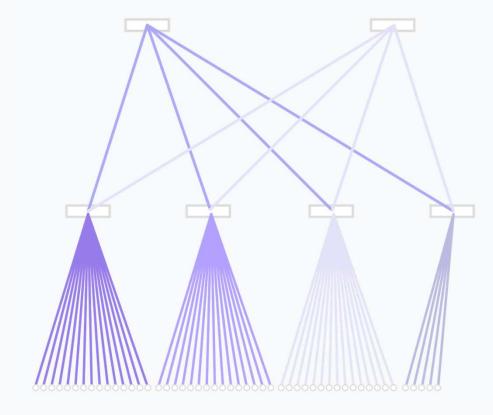
확장성 제약: 수백~수천 노드를 써도 효율이 떨어짐

Clevi-X Platform

독자적 분산 최적화 기술: 메모리 효율 극대화, 대규모 학습 안정성 확보

고유 병렬화 아키텍처: Data+Model+Pipeline 연산 병목 최소화

검증: 1.4조 개 파라미터 모델 안정적 학습

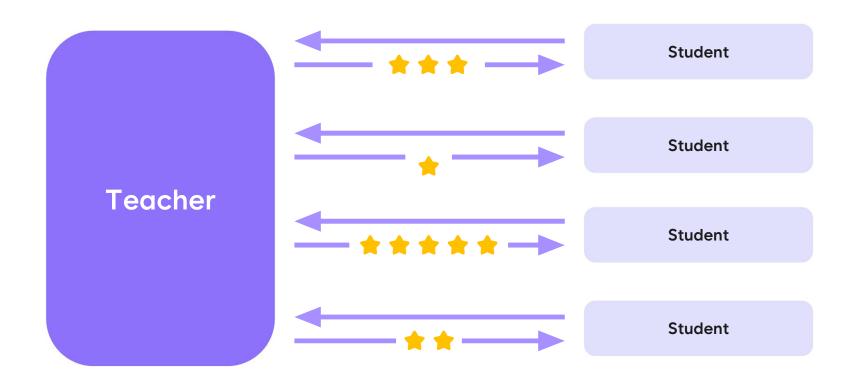


<서버-라우터 트리구조>

03 Clevi-X Intelligence Platform: Teacher-Student Model(Knowledge Distillation)

Teacher-Student Model(Knowledge Distillation)

Clevi-X는 처음부터 대규모 모델을 정교하게 설계하여 이를 기반으로 소규모 모델을 학습시켜 안정성과 품질을 모두 확보한 고성능 AI를 제공합니다. 고성능의 1.4조 파라미터 규모의 LLM으로, 국내 최대 규모이자 글로벌 수준의 AI모델로 경쟁력을 갖고 있으며 다양한 솔루션 제공이 가능합니다.

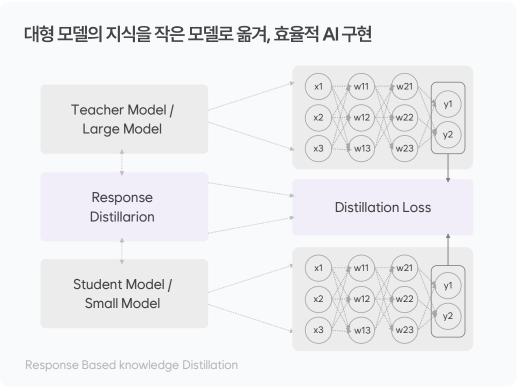


03 Clevi-X Intelligence Platform: Teacher-Student Model(Knowledge Distillation)

1.4조 파라미터 규모의 LLM으로 시작하는 고성능 AI

Clevi-X는 처음부터 대규모 모델을 정교하게 설계하여 이를 기반으로 소규모 모델을 학습시켜 안정성과 품질을 모두 확보한 고성능 AI를 제공합니다. 고성능의 1.4조 파라미터 규모의 LLM으로, 국내 최대 규모이자 글로벌 수준의 AI모델로 경쟁력을 갖고 있으며 다양한 솔루션 제공이 가능합니다.

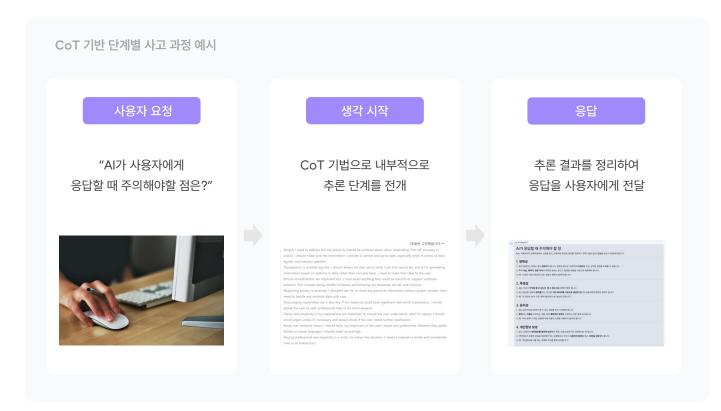




03 Clevi-X Intelligence Platform: CoT Model

AI 확산 시대, 투명성 확보를 위한 새로운 형태의 LLM_생각의 사슬(Chain of Though)

기존 AI는 '블랙박스'문제로 내부 과정이 불투명해 신용평가 등에서 불리한 결과를 받아도 그 이유를 파악하거나 이의를 제기하기 어렵다는 한계가 있었습니다. Clevi-X는 Chain of Thought(CoT) 기반 기술로 추론 과정을 단계별로 보여주어 결과의 투명성을 높이고, 신뢰할 수 있는 AI를 제공합니다.



기대 효과

투명성 확보

기존 AI의 블랙박스 문제를 해결하여, 모델의 추론 과정을 단계별로 시각화하고 의사결정 근거를 명확히 제시

신뢰성 향상

결과의 근거가 명확히 드러나, 일관성 있는 결과와 높은 해석 가능성을 보장함

규제 대응력 강화

설명 가능한 AI로서 확보해 금융·의료 등 규제 산업에서도 적용이 용이하며, 법적 ·윤리적 요구사항을 충족할 수 있음

품질 개선의 가속화

추론 과정을 분석해 오류 원인을 신속히 파악하고, 모델 품질을 지속적으로 고도화할 수 있음

*블랙박스: 내부 프로세스가 사용자에게 숨겨져 있어 의사결정 과정을 알기 어려운 인공지능 시스템

03 Clevi-X Intelligence Platform: Semantic Database

AI 검색·분류의 신뢰성을 높이는 시맨틱 데이터베이스

기존 백터기반 RAG 검색 문제점







문제점 1

검색 단계 자체가 제대로 이루어지지 않을 경우, 즉 적절한 키워드를 인식하지 못하거나 불충분한 검색 결과를 반환할 때 중요한 정보를 놓칠 위험이 있음

문제점 2

시간 데이터와 연결 데이터를 적극적으로 저장하지 않으므로 연속된 데이터나 비슷한 데이터의 경우 잘못되거나 편향된 데이터를 로드할 가능성이 있음

문제점 3

대규모 언어 모델이 내부적으로 갖는 지식 제약이나 오류 발생 문제를 완전히 보완하지 못하므로, 복합적인 상황에서 정보 누락이나 거짓 정보가 생성될 수 있음

03 Clevi-X Intelligence Platform: Semantic Database

AI 검색·분류의 신뢰성을 높이는 시맨틱 데이터베이스

Semantic 데이터베이스



절차별 문서 검색

응답에 있어 단일 문서 검색이 아닌 AI의 판단하에 다층 문서 검색과 추가 참고 문서 검색을 통한 정확한 응답 생성

접근 제어(ACL) 연동

각 디렉토리 별 접근 권한 부여를 통한 문서 보안 및 접근 권한별 문서 참조 분리

문서 분석 및 분류 성능 향상

프로젝트 규모가 성장하고, 시간이 지남에 따라 생기는 '비슷한 의미'의 문서도 정밀하게 판단가능

공간 제약 없음

많은 공간이 필요한 벡터 기반 RAG 대비 설계상 무제한 용량을 제공함으로써 저장 공간 부족에 대한 염려 없이 필요한 만큼 문서를 보관 가능

변경 이력 추적

문서가 저장된 시점, 수정된 시점 등의 기록을 통해 문서별 무결성 및 참조 가능여부, 진실성 판단

사내 지식 관리

주제나 부서명으로 문서를 분류하고, 대화내용에 기반하여 자동 문서 생성 및 수정을 지원하여 업무가이드와 사례집 생성 및 관리 가능

의미론적 사전 분석

저장하는 모든 형태의 데이터(사진, 비디오, 문서)의 내용을 분석하고 자율적으로 판단하여 최적화된 저장

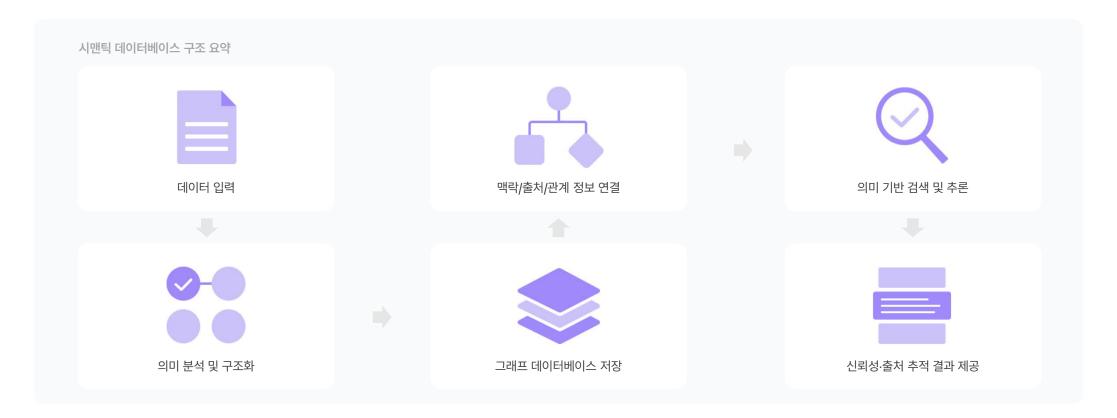
데이터 보존

장기적으로 모든 문서를 안전하게 보관하여 데이터 유실을 방지하고 언제든지 필요한 정보를 쉽게 찾을 수 있음

03 Clevi-X Intelligence Platform : Semantic Database

AI 검색·분류의 신뢰성을 높이는 시맨틱 데이터베이스

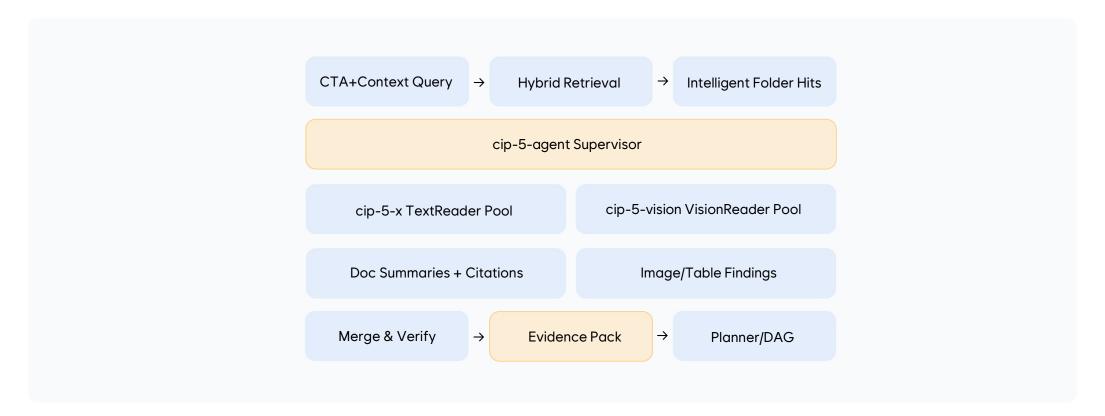
Semantic 데이터베이스



03 Clevi-X Intelligence Platform : Semantic Database

AI 검색·분류의 신뢰성을 높이는 시맨틱 데이터베이스

Semantic 데이터베이스



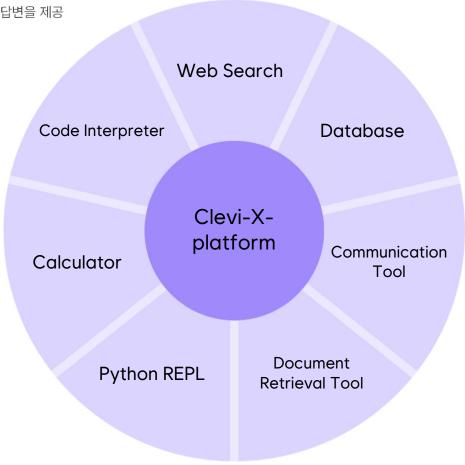
03 Clevi-X Intelligence Platform: Tool Calling

LLM Function Calling Interface

LLM이 외부 도구(툴, API, 데이터베이스, 계산 엔진 등)를 직접 호출하는 솔루션 단순히 언어를 생성하는 데 그치지 않고, 필요한 순간에 외부 기능을 실행하여 실시간으로 더 정확한 답변을 제공

Tool Calling

- ✓ LLM이 외부 함수/API를 직접 호출하는 인터페이스
- √ 답변 정확성 강화(Fact-Grounded)
- ✓ 환각 감소
- ✓ 신뢰성 향상
- ✓ 기능확장성 강화



Clevi Model Overview

ivy-3-text

Cip-4-mm

Cip-5-x

Cip-5-agent

Cip-5-vision



04 Clevi Model Overview: Model Overview

Clevi 모델 라인업

고객사 요구에 맞춘 다양한 AI모델을 제공합니다











04 Clevi Model Overview: lvy-3-text

Ivy-3-text

AS 접수 등 다양한 업무 시나리오에 유연하게 대응

작은 규모에도 다양한 전문적 역할을 수행할 수 있는 경량화 모델 양자화를 통해 엣지 컴퓨팅과 온디바이스 솔루션을 구현 드론, 자율주행, 지능형 CCTV, 스마트폰, 휴머노이드, 산업 설비 등 다양한 환경에서 사용

빠른 응답 속도를 자랑하는 비교적 작은 LLM 모델

- ✓ 규모가 작음에도 불구하고, 고도화된 아키텍처를 통해 고품질의 답변을 빠르게 출력 가능
- ✓ 이를 통해 다양한 환경에서도 즉각적인 정보 제공과 사용자 응대 실현

여러 사용자를 동시 지원

- ✓ 동시 접속자가 많은 상황에서도 안정적인 처리 성능을 유지
- ✓ 여러 사용자에게 끊김이 없는 고객 지원 및 정보 전달 제공

손쉬운 유지 · 확장으로 비즈니스 효율 극대화

- ✓ 상대적으로 작은 모델 사이즈 덕분에 인프라 부담이 적고, 필요한 경우 확장이나 연동이 용이
- ✓ 다양한 규모의 비즈니스 요구사항에 유연하게 대응 가능



04 Clevi Model Overview: Cip-4-mm

Cip-4-mm

멀티모달 non-reasoning model

자연어·이미지·음성 등 다양한 입력을 동시에 처리할 수 있으나, 추론보다는 문맥 연결과 패턴 인식에 특화되어 있어 검색, 분류, 요약 같은 빠른 처리에 강점

더 넓은 입력 채널 통합

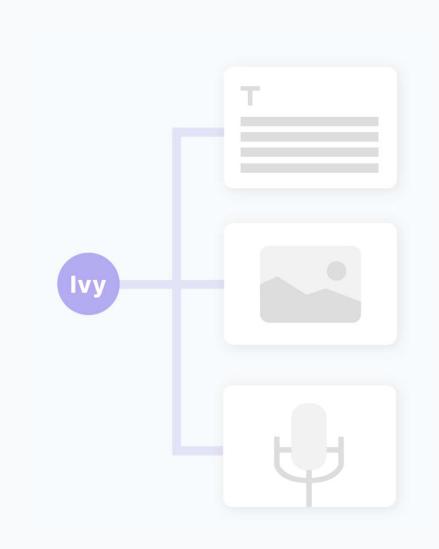
✓ 텍스트·이미지·음성 등 다양한 데이터를 동시에 다뤄 멀티모달 환경에서 유연하게 응용 가능

빠른 검색・분석 중심 기능

✓ 복잡한 논리 추론보다는 대규모 문서/데이터의 검색·분류·매칭에 탁월

효율성과 확장성을 극대화한 아키텍쳐

✓ 높은 처리량과 유연한 구조로, 조직의 다양한 요구사항에 맞춰 손쉽게 확장할 수 있어 대규모 환경에서도 안정적으로 운영이 가능



04 Clevi Model Overview: Cip-5-x

Cip-5-x

최고성능 reasoning 모델

최고 수준의 추론(reasoning) 능력을 갖춘 모델로, 복잡한 문제 해결과 단계적 논리 전개에 최적화되어 있으며 정밀한 분석·계획·의사결정 지원이 가능

모호한 질문에도 단계별 추론과정으로 결론의 근거를 명확히 설명

- ✓ 방대한 연구 및 검증을 목적으로 활용 시, 학습 및 개선 포인트를 쉽게 파악 가능
- ✓ 과학·수학 분야에서의 탁월한 성능

복잡한 공식이나 이론을 이해·적용해 정확한 계산 및 해석 제공

- ✓ 데이터 분석, 수치 모델링 등 고난도 작업에서도 우수한 결과 보장
- ✓ 높은 수준의 신뢰도와 재현성

완성된 답변의 논리 과정을 확인할 수 있어 신뢰도 상승

- ✓ 협업 및 공동 연구 시, 동일한 과정을 재현해 검증 가능
- ✓ 연구·개발 분야에 적합한 확장성



이 기계가 제대로 작동하는지 알려줘



사고 과정 보기 4초 완료



기준 설정:

- 정상 작동 여부는 소음, 진동, 출력, 온도 등으로 판단 가능.

데이터 확인:

- 소음: 정상 범위
- 진동: 기준치 이내
- 출력: 정상
- 온도: 약간 높음

종합 판단:

- 대부분 정상이나, 온도가 기준보다 약간 높음

0 0 0

04 Clevi Model Overview: Cip-5-agent

Cip-5-agent

Agent 목적 reasoning 모델

상황을 이해하고 목표를 달성하기 위한 계획·실행·조정 능력을 갖춤 사용자 지시를 해석해 단계적 작업을 자동으로 처리할 수 있음

목표 지향적 추론

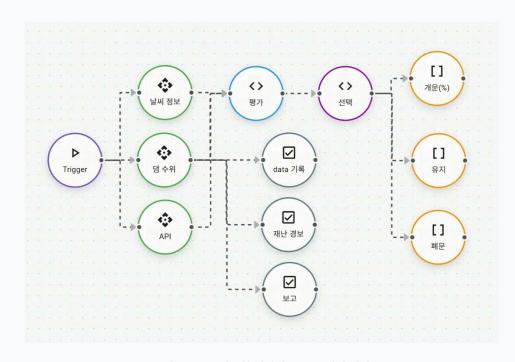
- ✓ 단순 질의 응답을 넘어 사용자 목표를 이해하고 계획 수립
- ✓ 단계적 액션을 자동으로 실행

다양한 도구 활용

✓ 외부 API·데이터·시스템을 연계하여 실행 가능한 결과를 만들어냄

Agentic Al

✓ 코딩 에이전트, 워크플로우 자동화 에이전트 등 다양한 목적 기반 에이전트 설계에 직접 적용 가능하며 복잡한 작업 흐름을 지능적으로 조율·실행할 수 있음



<워크플로우 자동화 에이전트 'ARI' 시연 예시>

04 Clevi Model Overview: Cip-5-vision

Cip-5-vision

Agent 목적 reasoning 비전 모델

차세대 지능형 비전 AI로 영상·이미지 기반 데이터 이해와 추론에 특화. 시각적 맥락을 해석하여 실제 환경과 연결된 의사결정을 지원함

시각 정보 이해

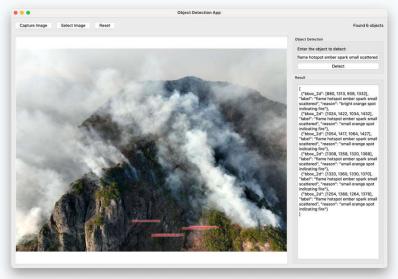
✓ 이미지·영상 내 객체, 관계, 상황을 정확히 인식하고 의미를 추론

복잡한 공식이나 이론을 이해·적용해 정확한 계산 및 해석 제공

- ✓ 데이터 분석, 수치 모델링 등 고난도 작업에서도 우수한 결과 보장
- ✓ 높은 수준의 신뢰도와 재현성

완성된 답변의 논리 과정을 확인할 수 있어 신뢰도 상승

- ✓ 협업 및 공동 연구 시, 동일한 과정을 재현해 검증 가능
- ✓ 연구·개발 분야에 적합한 확장성





<Cip-5-vision 모델이 산불을 감지하고 확산 경로 예측과 진압 지점을 추천>

05

Solutions

코딩 에이전트

Physical Al

VLM-Vision Intelligence Model

ARI: 워크플로우 자동화 시스템

보안 에이전트



05 Solutions: 코딩 에이전트

코딩 에이전트

개발자의 지시를 이해하고 기획·아키텍트·코드 작성·디버깅·리팩토링·문서화까지 자동으로 수행하는 AI 에이전트로, 소프트웨어 개발 전 과정에서 생산성을 향상

간단한 웹게임부터 복잡한 솔루션까지 사용자와의 인터랙션을 통해 완성도 높은 소프트웨어를 제작 가능



<클레비 코딩 에이전트로 제작한 VLM 솔루션>

o.type = beepOpt?.type || "sine"; o.frequency.value = beepOpt?.freq || 440; o.connect(ctx.destination); o.stop(ctx.currentTime + (beepOpt?.duration || 0.1)); DEBUG CONSOLE TERMINAL PORTS OneDrive\Clevi\My playground\Exercise\iq_game> **I** Architect Code ? Ask m Debug Orchestrator 田 戀 ∨ ¬ Orchestrator ∧ default <일반인들의 코딩 에이전트 체험>

게임에서 공이 이동하지 않고, 진행이 멈추는 현상이 발생하며, 볼륨 조절 ...

아래의 지침에 따라 전체 시스템 점검

아래의 지침에 따라 볼륨 조절 미적용 현상을 해결하도록 코드를 수정하...

<실제 Ivy 코딩에이전트 UI>

else if (ext === "wav") mime = "audio/wav";

const audio = new window.Audio(url);

const playPromise = audio.play();

audio.volume = volume;

if (ctx) {

if (!audioTest.canPlayType(mime)) return; // 브라우저 미지원

if (playPromise && typeof playPromise.catch === "function") {

if (typeof volume !== "number" || volume <= 0) return;

const ctx = window.AudioContext ? new window.AudioContext() : null;

console.warn("사운드 재생 실패:", url, e);

const o = ctx.createOscillator();

05 Solutions: Physical Al

피지컬 AI

로봇·센서·엣지 디바이스 등을 통해 물리적 세계와 직접 상호작용하는 AI로, 사람의 동작을 학습·모방하고 실제 환경에서 작업을 수행할 수 있는 능력을 갖춤

실시간 동작 학습

- ✓ 인간의 움직임을 관찰.모방하여 즉각적으로 동작을 재현
- ✓ Few shot/One-shot 학습으로 빠른 적응

환경 인식 및 제어

- ✓ 카메라·센서 데이터로 주변 상황을 이해
- ✓ 로봇 팔, 드론, 자율주행 장치 등 물리 장치 제어 가능

산업・생활 적용성

✓ 제조·물류 ·의료 ·서비스 로봇 등 다양한 분야에서 자동화와 생산성 향상에 기여



05 Solutions: VLM-Vision Intelligence Model

VLM(Vision-Language Model)

이미지·영상 같은 시각 정보와 텍스트 정보를 함께 이해하는 멀티모달AI로, 사람처럼 시각적 맥락과 언어적 의미를 연결하여 더 풍부하고 직관적인 인식·추론을 가능하게 함

멀티모달 통합 이해

- ✓ 이미지, 동영상, 텍스트 데이터를 동시에 받아들이고, 복수의 데이터를 통합 해석
- ✓ 인간의 인지 모형을 닮아, 다양한 감각(시각·언어 등)을 결합해 더 정확한 의미를 추론

복합적 추론 및 상황 이해

- 단순 이미지 분류가 아닌, 두 가지 정보를 결합해 복합적인 추론(reasoning)이 가능하여 대상의 상태를 정밀하게 파악하고 분석의 근거를 제시할 수 있음
- 의료 영상 판독, 산업 현장 위험 감지, 자율주행 등 고차원 의사결정분야에서 강력한 도구

Zero shot/Few shot 학습

- 새로운 태스크(Zero shot)나 소량의 예시(Few shot)만으로도 높은 성능으로 임무 수행 Ex) "000 결함이 있는 제품을 제외하라"는 간단한 텍스트 명령으로도 이미지 내 결함 탐지



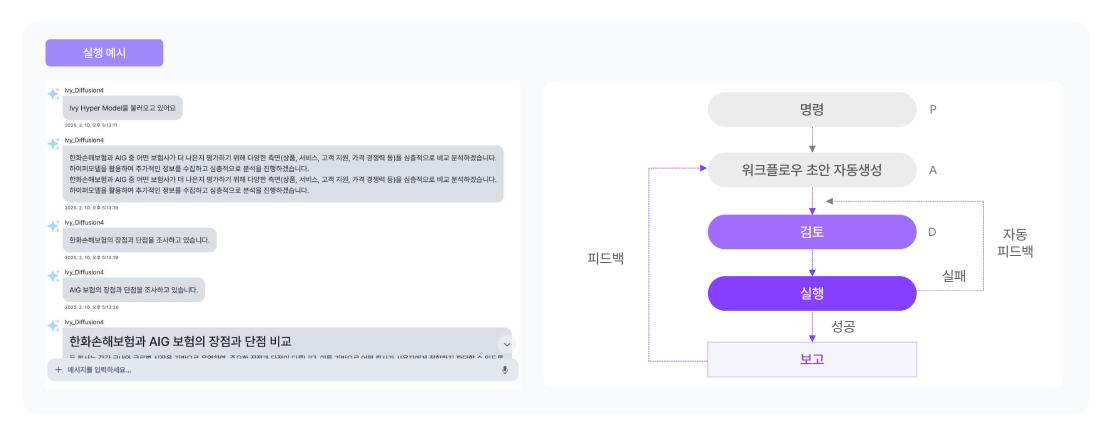
05 Solutions: ARI-워크플로우 자동화 시스템

워크플로우 자동화 시스템: 에이전트 플랫폼

자연어 명령을 통해 AI가 자율적으로 노드를 생성 · 해석하여 기본적인 실행 워크플로우를 만듭니다.

AI 기반 능동 워크플로우 시스템은 AI에게 가이드라인을 제공하며 필요한 경우 AI가 생성한 워크플로우를 수정이 가능합니다.

작업 수행시 발생하는 예외상황은 AI에 의한 동적 의사결정을 통해 자율적으로 행동하여 사용자의 요청을 이행합니다.



05 Solutions: 보안 에이전트

레드팀 테스트 솔루션

시스템과 네트워크를 보안 에이전트가 상시 점검하며, 화이트 해커처럼 적대적 공격을 시뮬레이션하고 취약점을 실시간으로 탐지하고 대응

상시·자동화된 점검

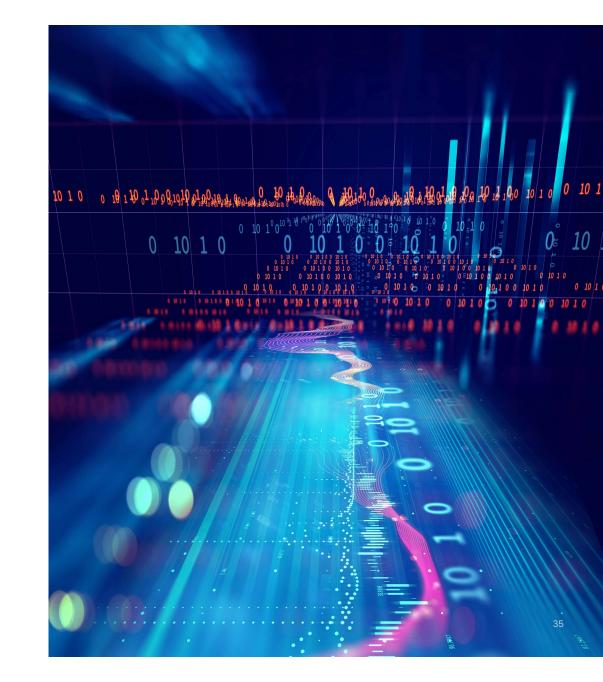
- ✓ AI가 매일 NSA CVE리스트를 자동 갱신
- ✓ 정해진 시간마다 서버, 시스템, 클라우드 환경에 대해 자동 레드팀 테스트 수행
- ✓ 취약점 발견 즉시 관리자에게 알림 및 리포트 제공

실시간 대응력

- ✓ 새로운 취약점(CVE)공개 즉시 테스트 반영
- ✓ 제보된 취약점도 즉시 시나리오에 포함, 신속한 패치 유도

규제·컴플라이언스 대응

- ✓ 금융, 의료, 공공 등 산업별 보안 규제와 표준 준수 검증
- ✓ 보안 리스크 관리 및 감사 대응 문서 자동화 지원

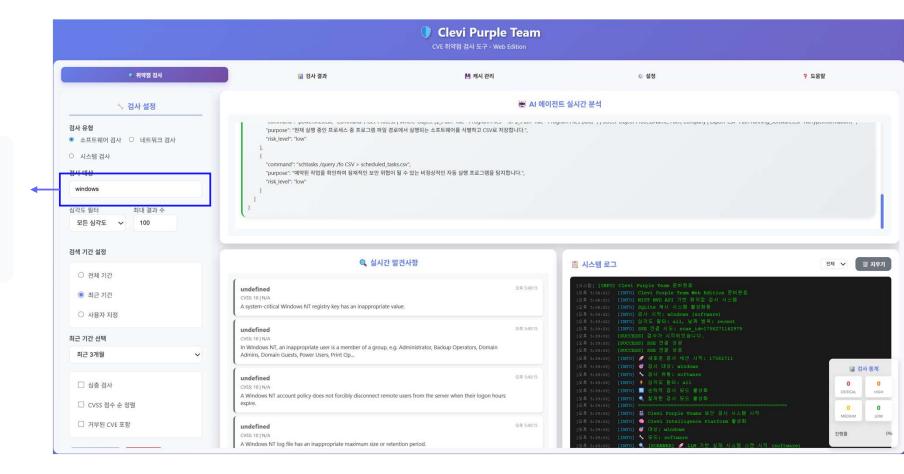


05 Solutions: 보안 에이전트

레드팀 테스트 솔루션

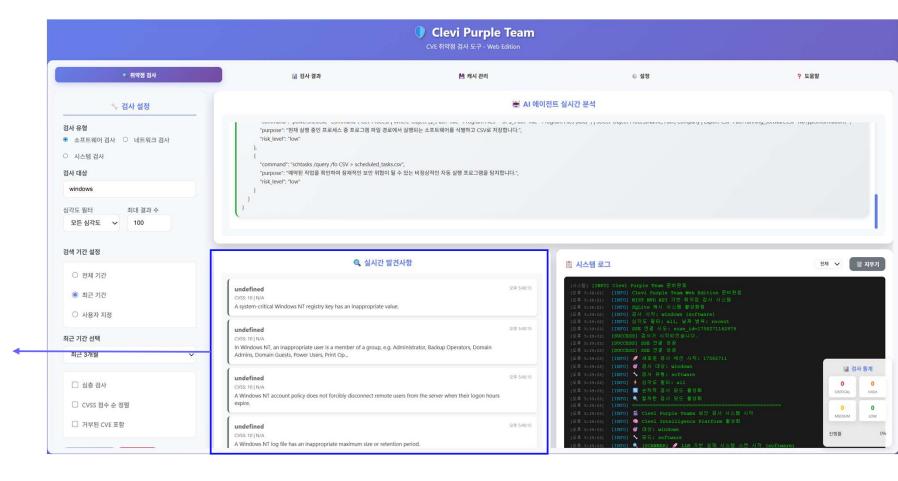
자연어 기반 탐지 대상 지정

windows/특정 폴더/특정 파일 등 시스템 안에 존재하는 임의의 대상입력하면 해당 대상에 대한 CVE 리스트를 불러와 보안 취약점을 검사합니다.



05 Solutions: 보안 에이전트

레드팀 테스트 솔루션



실시간 CVE 리스트 호출

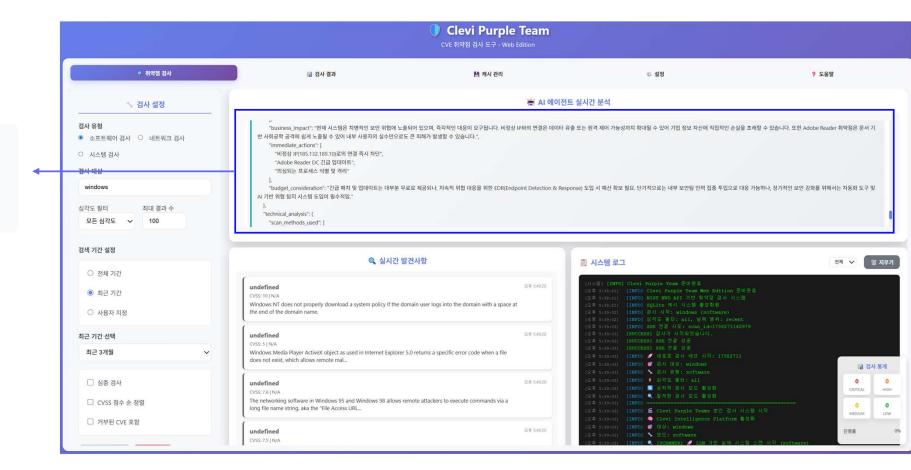
미국 NVD에서 최신 CVE리스트를 호출하여 대상 시스템과 네트워크의 보안 취약점을 탐지합니다.

05 Solutions: 보안 에이전트

레드팀 테스트 솔루션

리포트 자동 생성

보안 취약점이 존재할 경우 관리자에게 경고를 하고 리포트를 자동으로 생성합니다.



06

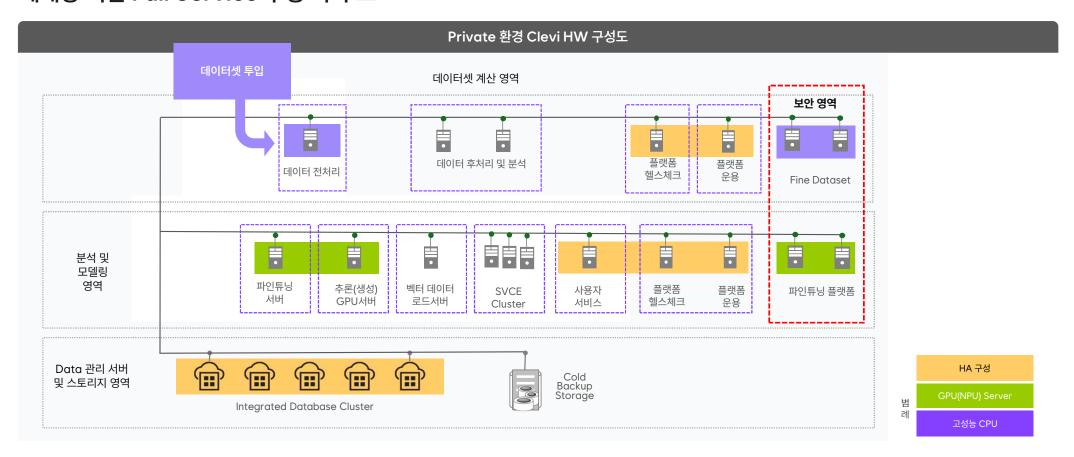
On-premises

하드웨어 구조 도메인 적용 워크플로우 격리환경



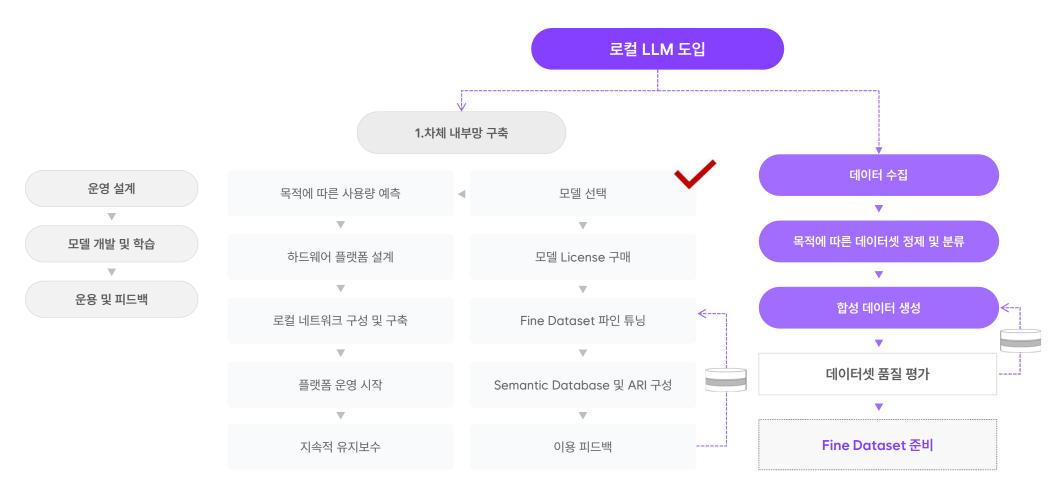
06 On-premises: 하드웨어 구조

하드웨어 기본 구조도 폐쇄망 기반 Full Service 구성 시 구조



06 On-premises: 도메인 적용 워크플로우

고유 도메인 활용 워크플로우



06 On-premises: 격리 환경

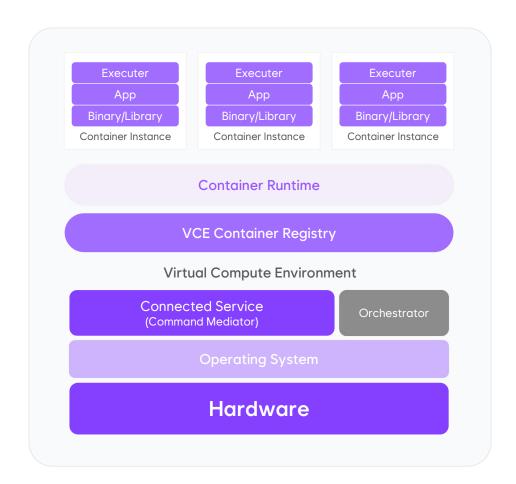
Clevi-X MDLM 모델을 위한 안전한 격리 컴퓨터 환경

Clevi-X MDLM (M Double LM - Massive Multimodal Language Model)은 인간의 컴퓨팅 사용 환경 및 방식, 구조, 절차를 학습하여 컴퓨팅을 지원할 수 있도록 설계되었습니다.

CLK (Clevi Linux Kernel) 기반 경량 컨테이너 환경은 바이러스의 침입 및 악의적 접근은 물론 권한을 넘어서는 모든 작업이 **불가능**하도록 설계되었습니다.

여러 익명 유저가 다양한 목적을 가지고 접속하는 클라우드 컴퓨팅 환경은 **악의적 사용자**도 물론 접속할 수 있습니다. 이런 유저에게서 안전을 보장하기 위해 Clevi는 SVCE로 불리는 **안전한 가상 격리 컴퓨팅** 환경을 만들어 모델에게 사용할 수 있도록 제공합니다.

각종 인터넷 접속이나, 계산, 코드 실행, 결과 검증, 파일입/출력을 위한 안전한 격리 공간을 제공함으로 모델에게 출력 데이터의 신뢰성과 작업 실행 능력을 부여하는 한편 악의적 의도를 가진 해킹, 데이터 유출, 모델을 이용한 공격, 모델을 향한 공격 등에서 안전한 환경을 제공합니다.



07

도입 시나리오

DevOps

R&D

HR

화재 재난 관제

댐 수위 관제



07 도입 시나리오: DevOps

고급 DevOps 파이프라인 개선과 지속적 코드 품질 관리

도메인

DevOps, CI/CD, 마이크로서비스

사용자

DevOps 리드 '이영희'

ARI가 분석한 업무 요구사항

수십 개의 마이크로서비스 코드 저장소에 대해 커밋 발생 시 자동 빌드, 유닛/통합 테스트 실행

컨테이너 이미지 보안 취약점 스캔 후 품질 미달 시 LLM이 코드 개선안 제안

테스트 커버리지·성능 측정 → 기준치 이하일 경우 Canary 배포 대신 Slack 알림 노드로 라우팅

사내 DevOps 베스트 프랙티스 문서, 클라우드 배포 정책 KB 참조

RBAC로 특정 고급 노드(배포 관련 노드) 수정은 DevOps 관리자만 가능

실행 과정

1. 자연어 명령

"각 서비스에 커밋 발생 시 빌드, 테스트, 취약점 스캔 후 품질 기준 만족하면 Canary 배포, 아니면 Slack 알림. 2주마다 LLM이 코드 품질 패턴 분석해 파이프라인 구조 개선 제안"

2. LLM 해석 & 워크플로우 생성

LLM(MES 전문가 페르소나) 노드가 센서 데이터 수집 노드(각 공장 API), 품질검사 코드 노드, 파라미터 재조정 노드, 재고 확인 및 주문 노드, 머신러닝 예측 노드 생성

MES용 노드 템플릿과 GMP/품질 규정 KB 적극 활용

정책 변경(인증 절차 강화) → LLM 노드가 인증 노드 삽입

3. GUI 편집

이영희는 GUI에서 병렬 빌드 노드(다수 마이크로서비스 동시 처리) 구성

Canary 배포 노드 클릭 → "DevOps 엔지니어" 페르소나, 특정 API 사양 문서 경로 지정

4. 동적 의사결정

테스트 실패 시 LLM 노드가 로그 분석, 코드 수정안(예: 테스트 커버리지 향상용 코드) 생성 \rightarrow 재실행성능 저하 발생 시 대체 설정(리소스 할당 증가) 자동 제안

5. 장기 개선

2주 후 LLM 노드가 반복적 오류 패턴 감지, 파이프라인 구조 최적화 (불필요한 테스트 단계 제거, 병렬 처리 최적화) 제안

GUI에서 승인 후 즉시 반영, JSON 로그로 감사 추적

07 도입 시나리오: R&D

연구(R&D) 대규모 데이터 분석과 지능형 파이프라인 리팩토링

도메인

제조(MES) + 공급망 관리 + 예측 분석

사용자

글로벌 생산관리 총괄 '오준영'

ARI가 분석한 업무 요구사항

매일 수백 GB 규모의 실험데이터 HPC(고성능 컴퓨팅) 클러스터에서 수집 Spark 기반 전처리, 머신러닝 이상치 검출, 결과 그래프, PDF 보고서 자동 생성 한 달 주기로 LLM 노드가 결과 분석, 효율적 알고리즘 코드를 새로 생성해 파이프라인 자동 리팩토링

연구 윤리, 데이터 프라이버시 규정 KB 참조, 민감정보 마스킹 자동 적용

실행 과정

1. 자연어 명령

"매일 HPC 클러스터 데이터 전처리, 이상치 감지, 그래프+PDF 보고서 생성. 매달 성능 분석 후 LLM이 더 효율적인 알고리즘 코드로 파이프라인 자동 개선"

2. LLM 해석 & 워크플로우 생성

LLM(데이터 사이언티스트 페르소나)가 데이터 수집 노드, Spark 전처리 코드 노드, 머신러닝 모델 노드, 그래프 생성 노드, PDF 리포트 노드, 성능 모니터 노드 구성

연구 윤리 KB 참조, 민감 데이터 마스킹 노드 삽입

3. GUI 편집

이창호는 GUI에서 반복 실행(매일), 월말 성능 분석 노드(조건 분기) 추가 LLM 노드 프롬프트 패널에서 "고성능 알고리즘 우선" 옵션 선택

4. 동적 의사결정

데이터 포맷 변경 시 LLM 노드가 오류 감지→코드 개선 후 재실행

머신러닝 모델 성능 저하 감지 시 LLM 노드가 다른 알고리즘 (예: 랜덤포레스트→XGBoost) 코드 생성 제안

5. 장기 개선

한 달 후 LLM 노드가 반복적 패턴 분석, 파이프라인 노드 재배치(전처리 단계를 병렬화), 더 효율적 API 호출 구조 제안

GUI에서 승인 후 JSON 업데이트, 로그 기록

07 도입 시나리오: HR

HR 글로벌 인력 관리와 자동 교육 커리큘럼 추천

도메인

인사관리(HR), 교육 추천

사용자

HR 총괄 '정미희'

ARI가 분석한 업무 요구사항

매월 말 각 국가 지사의 신규 입사자 목록 수집, 이력서 스킬 분석, 부족 스킬별 맞춤 교육 추천

지역별 HR 규정, 개인정보 보호 정책 준수, 민감 데이터 마스킹

6개월 주기로 성과 데이터 수집 후 LLM 노드가 교육 커리큘럼 코드 재생성, 국제 HR 정책 변화 시 자동 반영(예: 추가 인증 노드 삽입)

실행 과정

1. 자연어 명령

"매월 말 각 지사 신규 입사자 정보 가져와 이력서 분석, 부족 스킬 교육 배정, 민감정보 보호, 6개월 후 성과 검토 후 커리큘럼 재설계"

2. LLM 해석 & 워크플로우 생성

LLM(HR 전문가 페르소나): HR API 노드, 이력서 파싱 코드 노드, 교육 추천 노드, 개인정보 마스킹 노드, 성과 평가 노드, 커리큘럼 재설계 노드 구성

국제 HR 정책 KB 참고, 보안 정책 문서 분석해 인증 노드 삽입

3. GUI 편집

정미희는 GUI로 조건 분기(특정 스킬 부족 시 특정 교육과정 할당) 구성 LLM 노드 클릭 → "HR 전문가" 페르소나 설정, 프롬프트 패널에서 "문서 KB경로" 지정

4. 동적 의사결정

특정 지역 이력서 포맷 변경 → LLM 노드가 파싱코드 수정→재실행 6개월 후 성과데이터 분석: LLM 노드가 교육 커리큘럼 구조적 변경 (예: 코드로 표현된 교육 과정 리스트) 제안

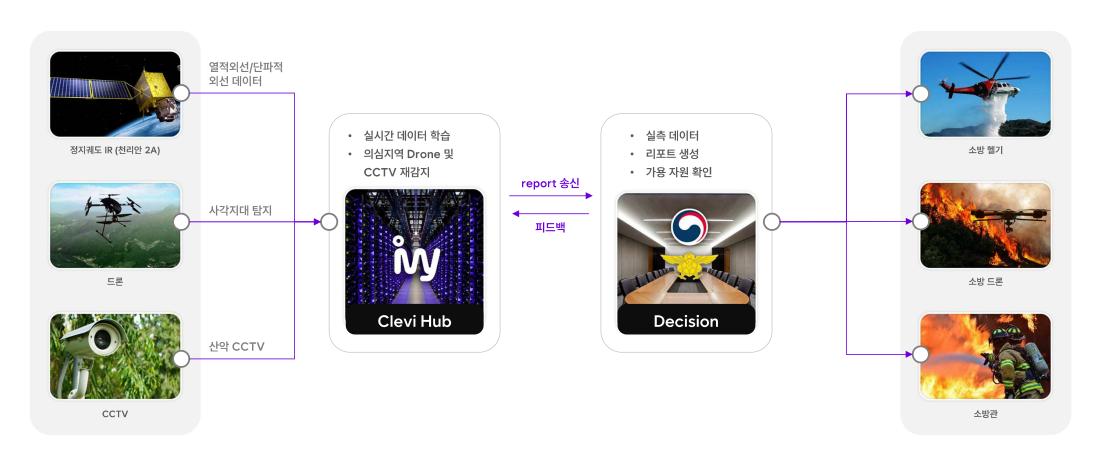
5. 정책 반영

개인정보 보호 규정 업데이트 시 LLM 노드가 마스킹 노드 파라미터 변경, 승인 후 반영

07 도입 시나리오: 화재 재난 관제

화재 재난 관제-워크플로우 다이어그램

육안으로 식별이 어려운 객체도 정확히 식별



07 도입 시나리오: 화재 재난 관제

화재 재난 관재-정확한 발화지점 감지

육안으로 탐지가 어려운 객체도 정확히 식별



- 드론과 산악CCTV를 통한 이미지 데이터를 실시간으로 산불 상황을 정확하게 판단
- 육안으로 구별이 어려운 운무/구름/수증기 등을 연기와 구분
- 노을이나 단풍과 같은 원거리에서 판단이 어려운 이미지도 구분

07 도입 시나리오: 화재 재난 관제

화재 재난 관제-화재 분석 리포팅

단순 객체 탐지에서 그치는 것이 아니라 시계열 추론 알고리즘으로 산불의 확산경로 예측과 추천 진압 지점까지 리포팅



AI 기반 산불 감지 및 진압 최적점 예측 시스템으로 산불의 진행 방향을 분석하고 최적 의 진압 지점을 제안합니다.



산불 이미지 분석

산불 또는 연기가 포함된 이미지를 업로드하여 AI 분석을 시작하세요



다른 이미지 선택

ⓒ Al 분석 시작



🖹 분석 요약

주요 발견사항 즉시 대응 필요사항 • 화재 감지: 확인됨 ① 즉각적인 항공 진화 자원 투입 • 위험 수준: 높음 ② 주요 확산 경로(동북동) 차단선 구축 • 진행 방향: 70° (신뢰도 85%) ③ 언근 주민 대피형 발령 및 대피로 확보 • 권장 진압점: 4개 지점 ④ 지상 진화대 접근 시 안전 확보 필수

⊚ 상세 분석 결과

이미지 상 산불이 산악 지형을 따라 동북동 방향(약 70도)으로 확산 중이며, 연기 기둥이 길고 넓게 퍼져 있어 강한 바람의 영향을 받고 있음이 확인됩니다. 화재 강도는 연기량과 불길의 크기로 볼 때 'high'로 판단됩니다. 지형적으로 계곡과 능 선이 혼재되어 있어 확산 속도가 빠를 것으로 예상되며, 접근이 어려운 지역이 많아 진압 난이도가 높습니다. 확산 경로 상에 추가적인 산림 피해가 우려되므로, 항공 및 지상 진화의 신속한 연계가 필요합니다.

07 도입 시나리오: 댐 수위 관제

댐 수위 관제

워크플로우 다이어그램



데이터 수집 및 통합:

IoT 센서로 데이터 자동 수집/외부 데이터 (기상청 예보, 위성/레이더, 상·하류 수위, 유역 정보, 드론/CCTV 정보) 통합 및 전처리

AI 기반 예측 및 분석:

- ◎ 홍수위/유입량 예측: 과거 데이터+실시간 센서+기상 예보를 AI가 통합 분석
- ◎ 위험도 평가: 하류 침수 위험, 구조물 이상, 방류 임계치 도달 등 실시간 산출
- ◎ 시나리오 분석: 다양한 강우/유입/방류 시나리오별 영향 자동 시뮬레이션

최적 운영 시나리오 자동 생성:

- ◎ AI 시뮬레이션: 방류량 시점 속도 등 자동 산출, 최적 방류안 제시
- ◎ 우선순위 제안: 하류 피해 최소화, 상류 저수지 연계 등 복합 최적화
- ◎ 비상대응 시나리오: 구조물 이상, 예상치 못한 유입 급증 등 상황별 대응안 자동 생성

의사결정 지원 및 자동 제어: 운영자 대시보드/자동 경보/설비 자동 제어

대외 커뮤니케이션 및 보고 자동화: 유관기관/지자체 자동 통보, 주민 안내, 상급기관 보고서 자동 작성

사후 평가 및 학습: 운영 결과 데이터 자동 축적, AI 모델 성능 평가 및 재학습, 매뉴얼 업데이트

3레비 회사소개서

07 도입 시나리오: 댐 수위 관제

댐 수위 관제

댐 관제 시스템 리포트 예시



4. 기상·수문 예측

단기간 내 대규모 강우로 유입량이 급증할 전망이며, 점두유량 980cms가 18시간 내 도달할 것으로 예측되어 선제적 방류 및 하류 경보 강화가 필요합니다.

24시간 강우(mm)	110	집중호우
72시간 강우(mm)	240	유역 전역 영향
유입량 전망	급격한 증가	상류 집중호우
철두유량(cms)	980	18시간 내 도달 예상
홍수파 도달 시각(시간)	18	방류 조정 필요

5. 안전·리스크 평가

구조적 위험은 낮으나, 유입량 급증과 수질 악화로 인한 운영 및 수질 리스크가 중간 수준으로 평가됩니다.

구조적 위험	저	저수위 여유 확보
운영 위험	중	유입량 급증, 방류 조정 필요
수절 사고	중	탁도 상승, 부유물 증가
예측 대비 리스크	중	첨두유량 도달 임박

6. 하류 보호 경보·대응

관심	유입량 400cms 이상	관계기관 통보
주의	유입량 600cms 이상 또는 하류 수위 70%	상황실 비상대기
경계	유입량 800cms 이상 또는 하류 수위 90%	대피 준비, 통제구간 지정
심각	홍수조절수위 접근/초과 또는 하류 침수	즉시 대피, 교통·전력 통제

- 통보 대상: 지자체, 소방, 전력, 교통 등
- 하류 인구 18,500명, 정수장·교량·국도·변전소 등 주요 기반시설 보호 필요
- 대피 및 통제 시나리오 단계별 준비

7. 규정 준수 체크

모든 운영 및 방류가 허가 범위 내에서 이행되고 있으며, 데이터 품질도 양호합니다.

방류허가 준수	이행 중	허가 범위(100~1200cms) 내
생태유량 확보	이행 중	생활-관개용수 정상 공급
취수·발전 계약	이행 중	공급량 정상
데이터 품질	0.95	신뢰도 양효

8. 조치 권고안

- 즉시: 방류량 점진적 증대 및 하류 경보 단계별 통보(첨두유량 도달 18시간 전 대비)
- 즉시: 탁도 및 부유물 실시간 모니터링, 취수장 보호 조치 강화
- 단기: 집중호우 종료 후 수문 개방률 재조정 및 저수위 재평가
- 단기: 하류 취약 인프라(정수장, 교량 등) 사전 점검 및 대피 안내
 중기: 수질 악화 대비 응급 정수처리 및 부유물 제거 계획 수립
- 중기: 수실 약와 대비 응급 성수저리 및 무유물 제거 계획 수립
 중기: 홍수기 종료 후 운영·수질 데이터 분석 통한 대응 매뉴얼 보완

• 당기: 중부가 장묘 푸 순당 박은 테이터 순역 중간 네를 베



THANK YOU

주소 경기도 화성시 동탄 첨단산업 1로 27 금강 펜테리움 IX타워

이메일 contact@clevi.ai

전화 1551-4096

웹사이트 www.Clevi.ai