

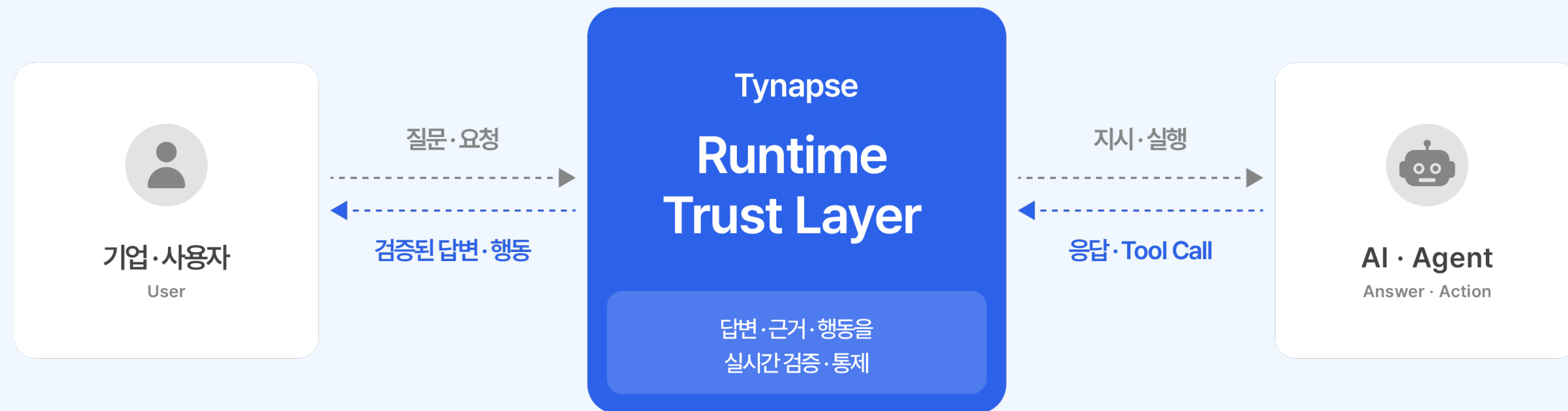


AI가 더 똑똑해 질수록, 기업은 어디까지 맡길 수 있는가

평가에서 실시간 통제까지

티냅스는 기업이 AI에게 어디까지 맡길 수 있는지를 푸는 회사

기업 AI의 답변 · 근거 · 행동을 실시간으로 검증하고 통제합니다



Enterprise Proof

 KB금융그룹 등 금융기관 Enterprise AI Trust 검증

Research Proof

 ICML · KDD · ECCV 워크샵 논문 게재

Global Proof

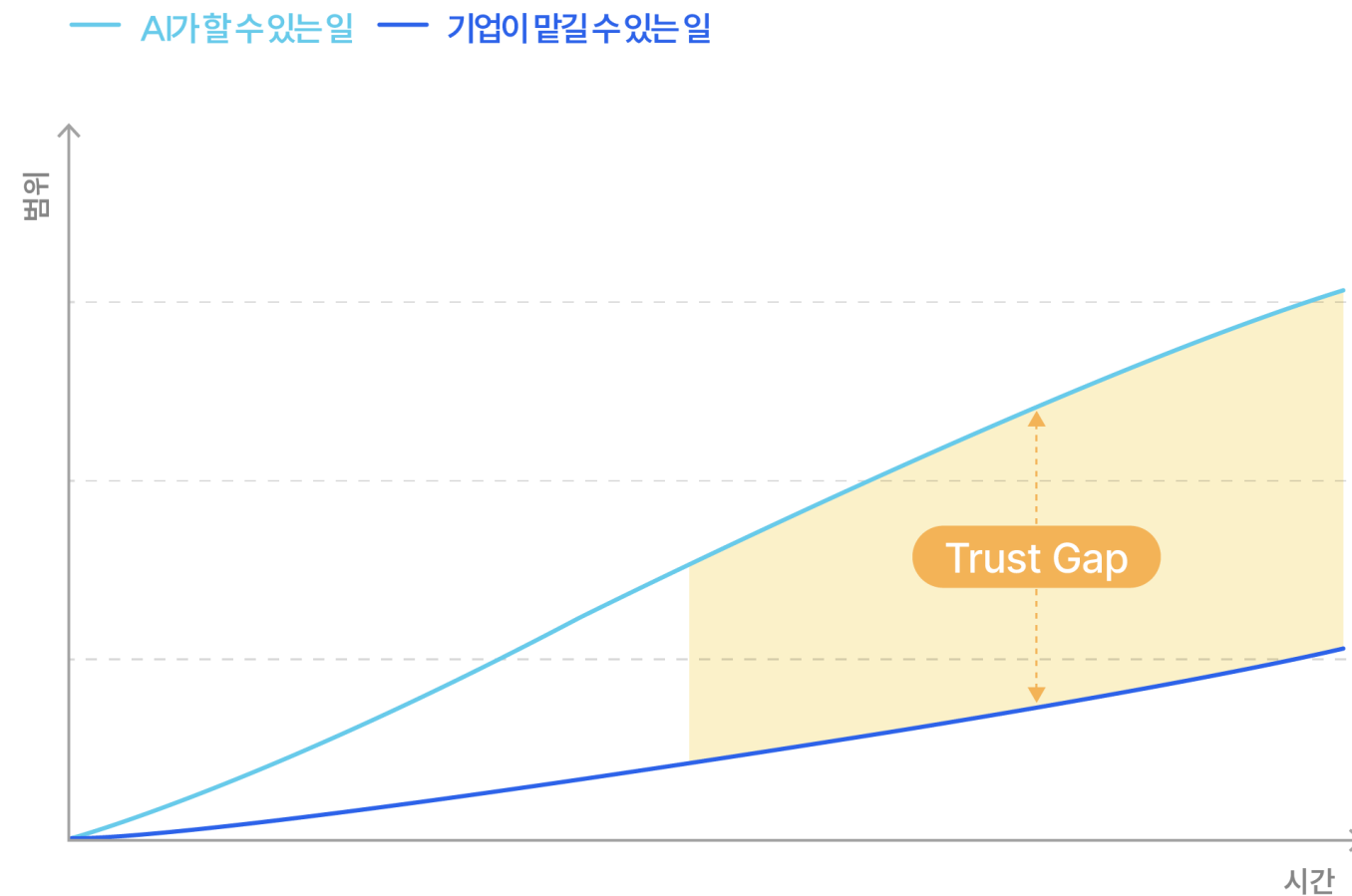
 NVIDIA x AWS GTC Startup Pitch Global Top 6

 Snowflake Startup Challenge 2026 Global Top 10

 GSTAR Summit 2026 Invited Speaker

AI가 할 수 있는 일 ≠ 기업이 맡길 수 있는 일

Capability ≠ Delegation



Trust Gap

성능은 가파르게 오르지만,
맡길 수 있는 범위는 완만하게 상승

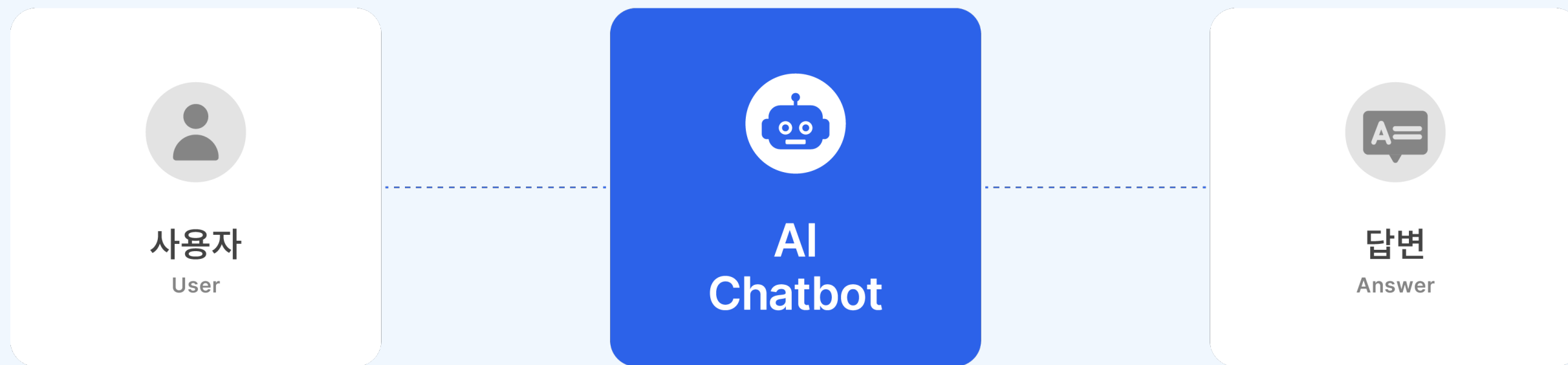
병목의 이동

From. 모델 성능 To. 맡길 수 있는 범위

Market Signal · Gartner, May 2026

40% 운영 사고가 터진 뒤에야 발견된
Governance Gap 때문에
기업이 2027년까지 자율 AI 에이전트를
강등하거나 해제할 전망

챗봇 시대 : 답변하는 AI



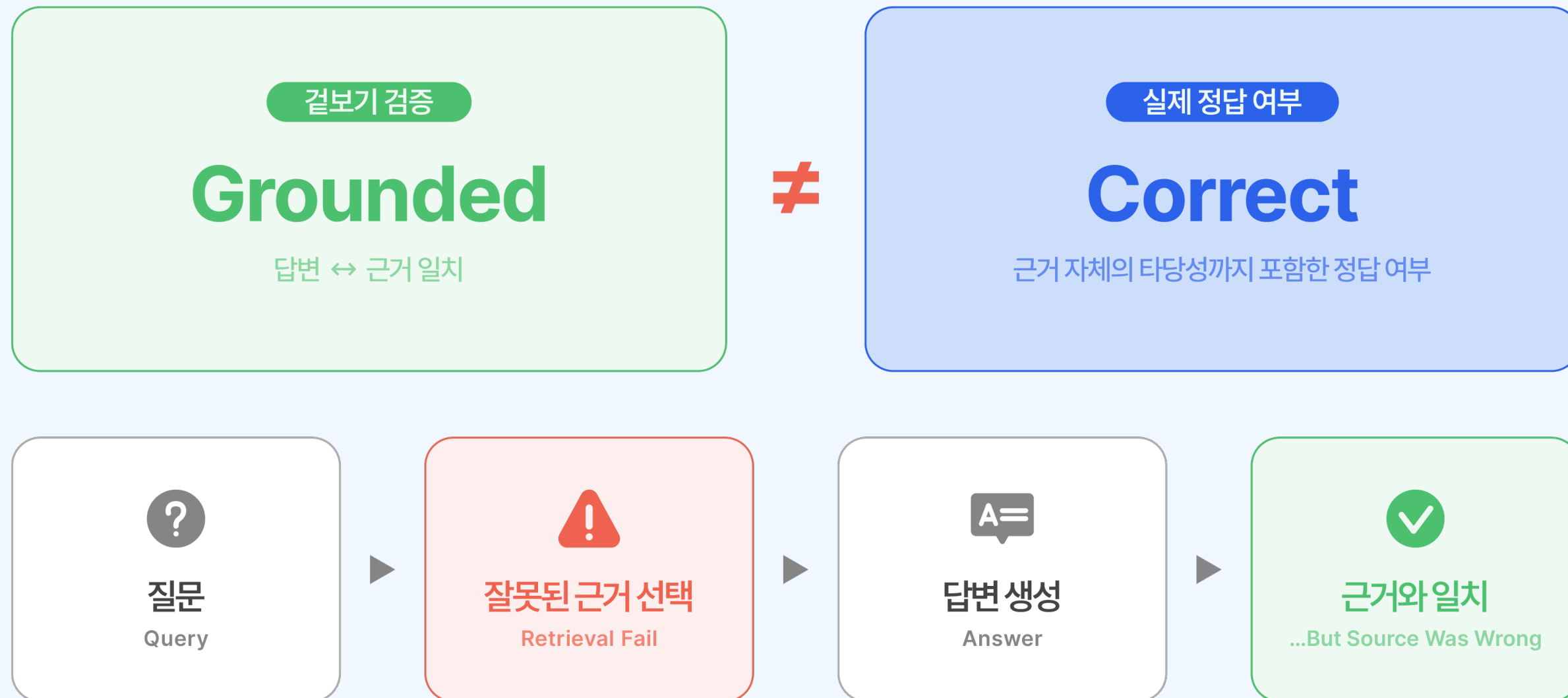
최종 결과는 '답변'

Agent 시대 : 행동하는 AI



챗봇의 오류는 오답, 에이전트의 오류는 행동입니다

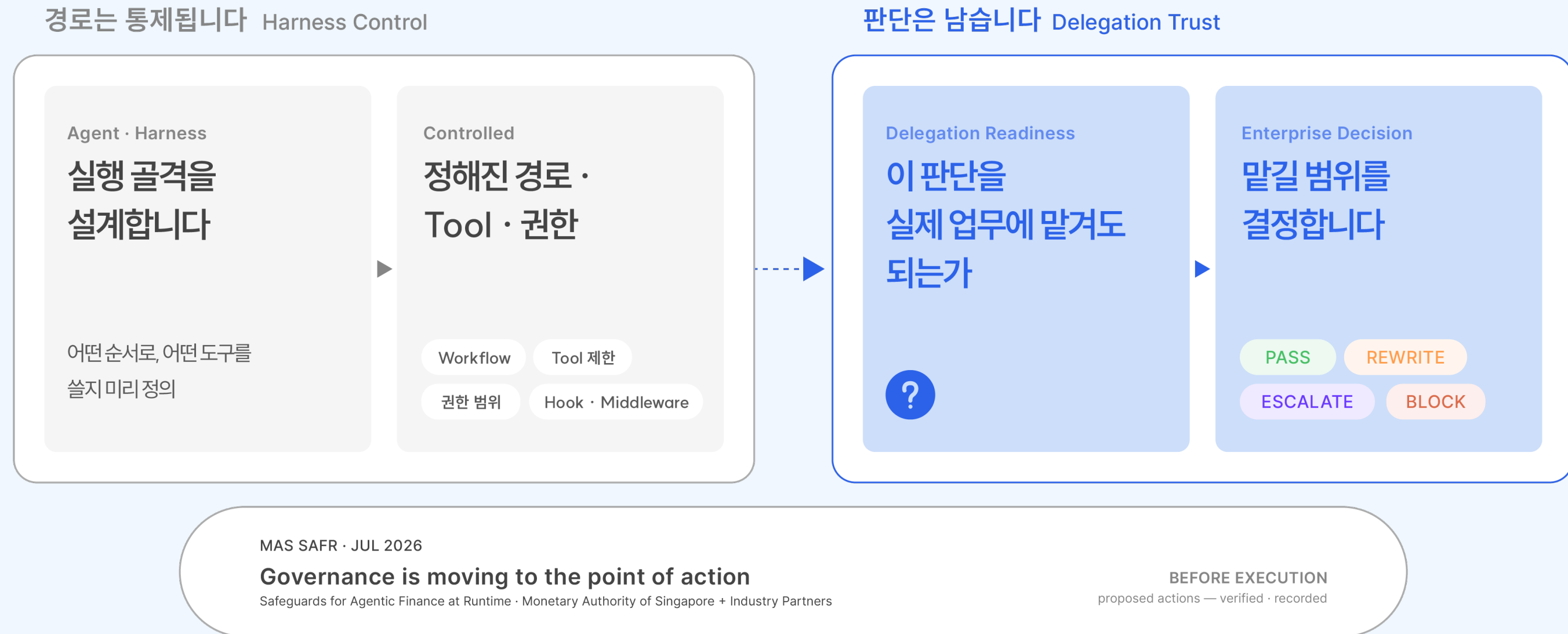
근거와 일치한다고 해서, 맞는 답은 아닙니다



마지막 답변이 아니라, 검색 → 실행까지 전체 경로를 봐야 합니다

Control ≠ Trust

통제할 수 있는 것과 맡길 수 있는 것은 다른 문제



Harness는 경로를 통제합니다.
기업은 그 판단을 실제 업무에 맡겨도 되는지를 결정해야 합니다

유창한 설명이 신뢰할 수 있는 판단은 아닙니다

Wrong, with Confidence

Car-wash Test · 같은 질문 10회 반복

세차장이 50m 앞에 있습니다. 걸어갈까요, 차를 가져갈까요?

GPT-5

7 / 10

같은 질문 10회 중 운전을 선택한 횟수

운전을 택한 근거

“세차를 하려면 차가 세차장에 있어야 하므로 차를 가져갑니다”

도보를 택한 근거

“50m밖에 안 되는 거리라 걷는 것이 더 빠르고 에너지 낭비가 적습니다”

Gemini 2.5 Pro

4 / 10

같은 질문 10회 중 운전을 선택한 횟수

운전을 택한 근거

“차가 세차장에 있어야 세차가 되니, 거리와 무관하게 운전해야 합니다”

도보를 택한 근거

“1분이면 걷는 거리를 운전하는 것은 비효율적이고 연료 낭비입니다”

도보를 택하면 차 없이 세차장에 도착합니다. 틀린 판단의 근거도 맞는 판단만큼 유창합니다.

문제는 틀리는 것이 아니라
확신을 갖고 틀리는 것입니다

사실이 맞아도, 업무적으로는 틀릴 수 있습니다

Fact Check ≠ Business Judgment



필요한 것은 점수가 아니라 결정입니다

평가는 점수가 아니라, 맡길 수 있는 경계를 설정



AI는 확률적이지만, 기업의 업무 결과까지 확률적일 수 없습니다

위험에 따라 판단 방식도 달라야 합니다

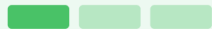
◀ Clear · Fast

Ambiguous · Critical ▶



Tier 1 · Rule Engine

Rule


 Deterministic

권한 · 접근

개인정보

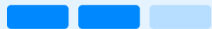
거래 한도

Disclaimer



Tier 2 · Domain Ai

Domain Judge


 Context-aware

근거 적합성

업무 기준

문맥 · 톤

표현 강도

도메인 위험



Tier 3 · Human Review

Human-in-Loop

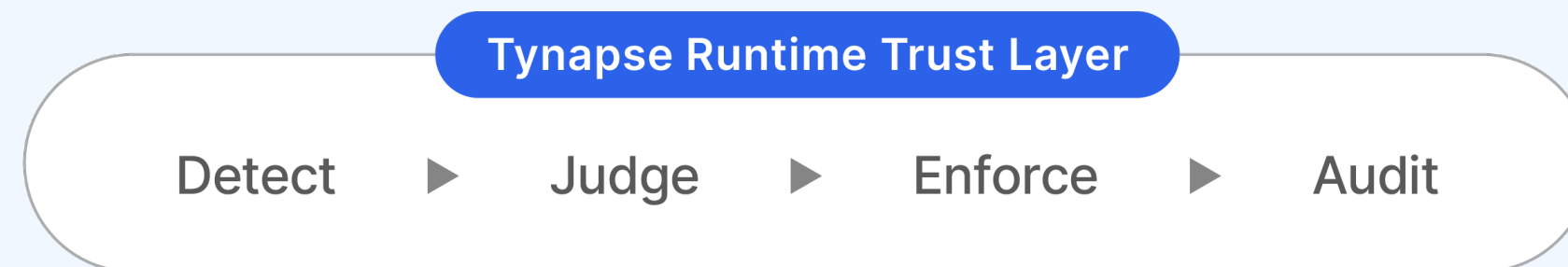

 Irreversible

고위험 승인

되돌리기 어려움

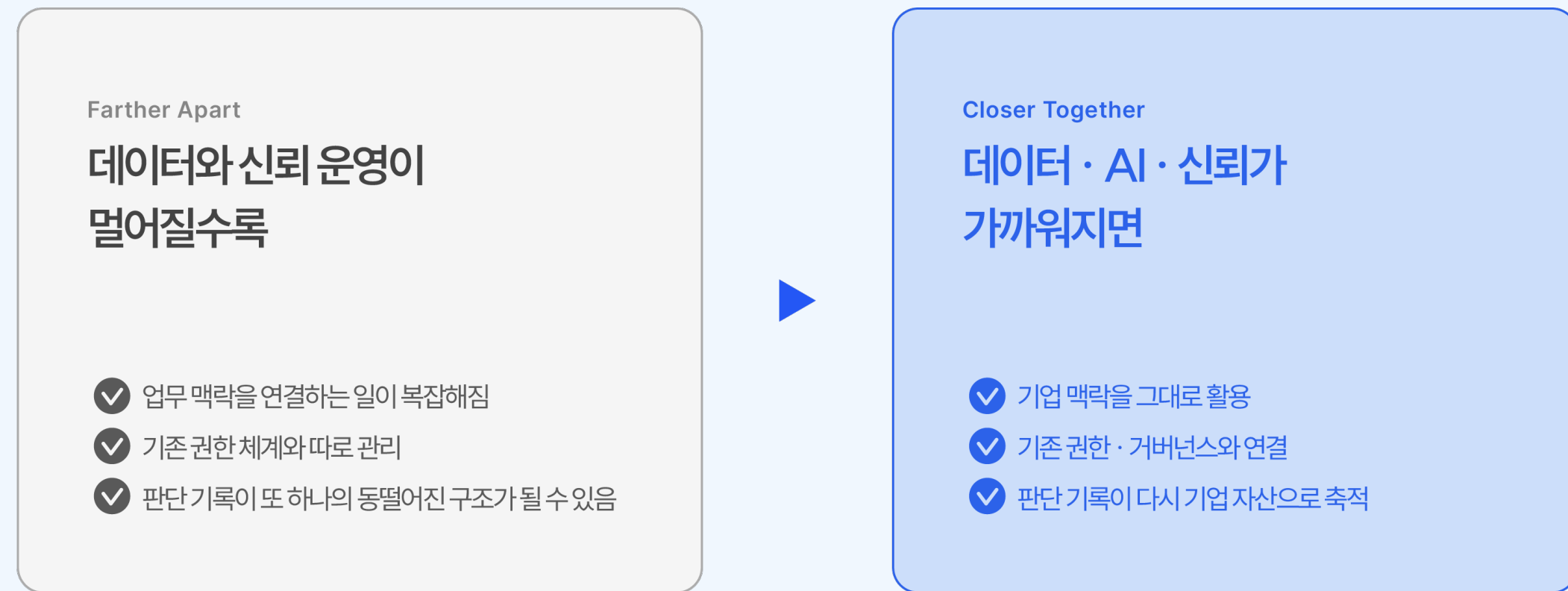
예외 · 신규

정책 갱신



AI의 판단은 기업의 데이터와 맥락에서 시작됩니다

Data · Governance · Trust

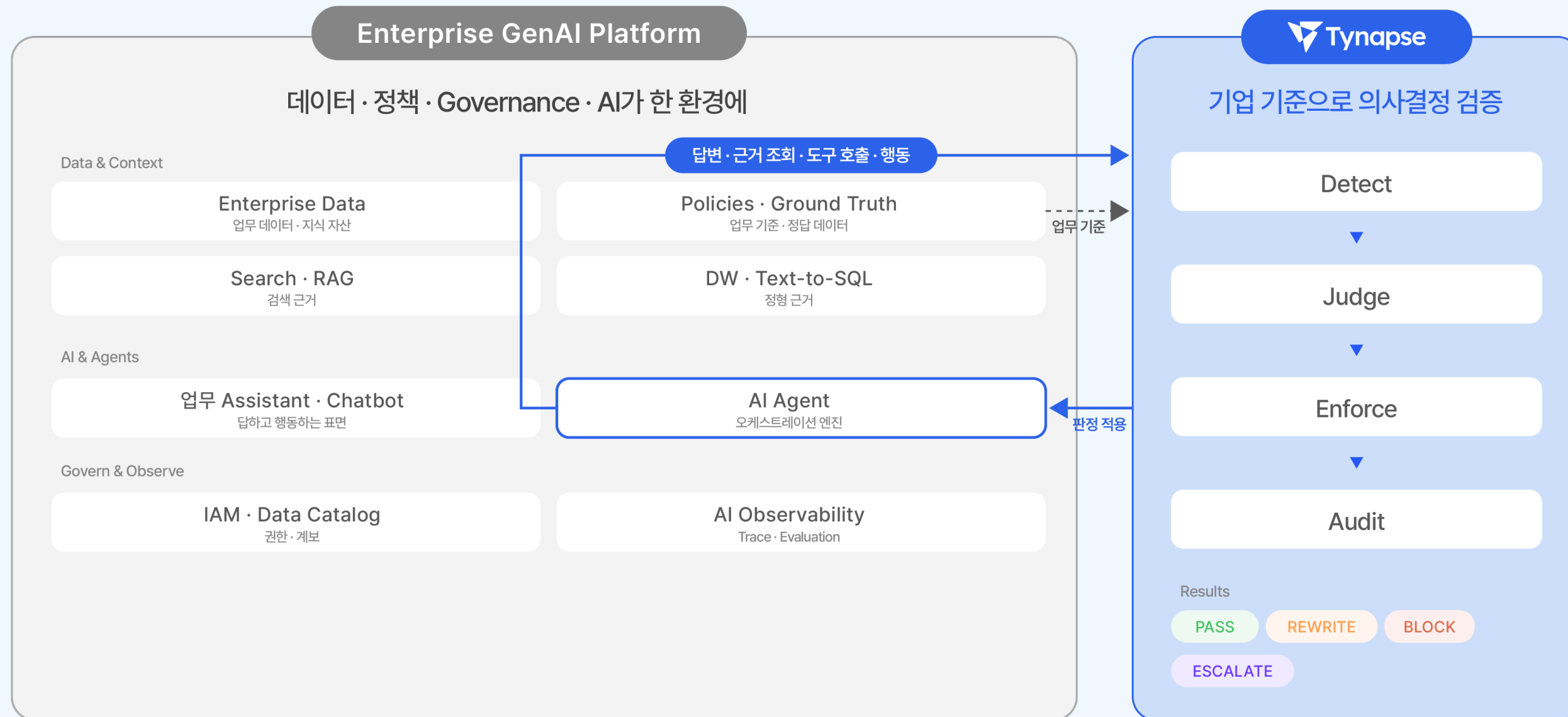


Why It Matters in Enterprise

기업 데이터와 AI가 이미 연결된 기반에서,
답변과 행동의 신뢰까지 함께 운영할 수 있는 구조

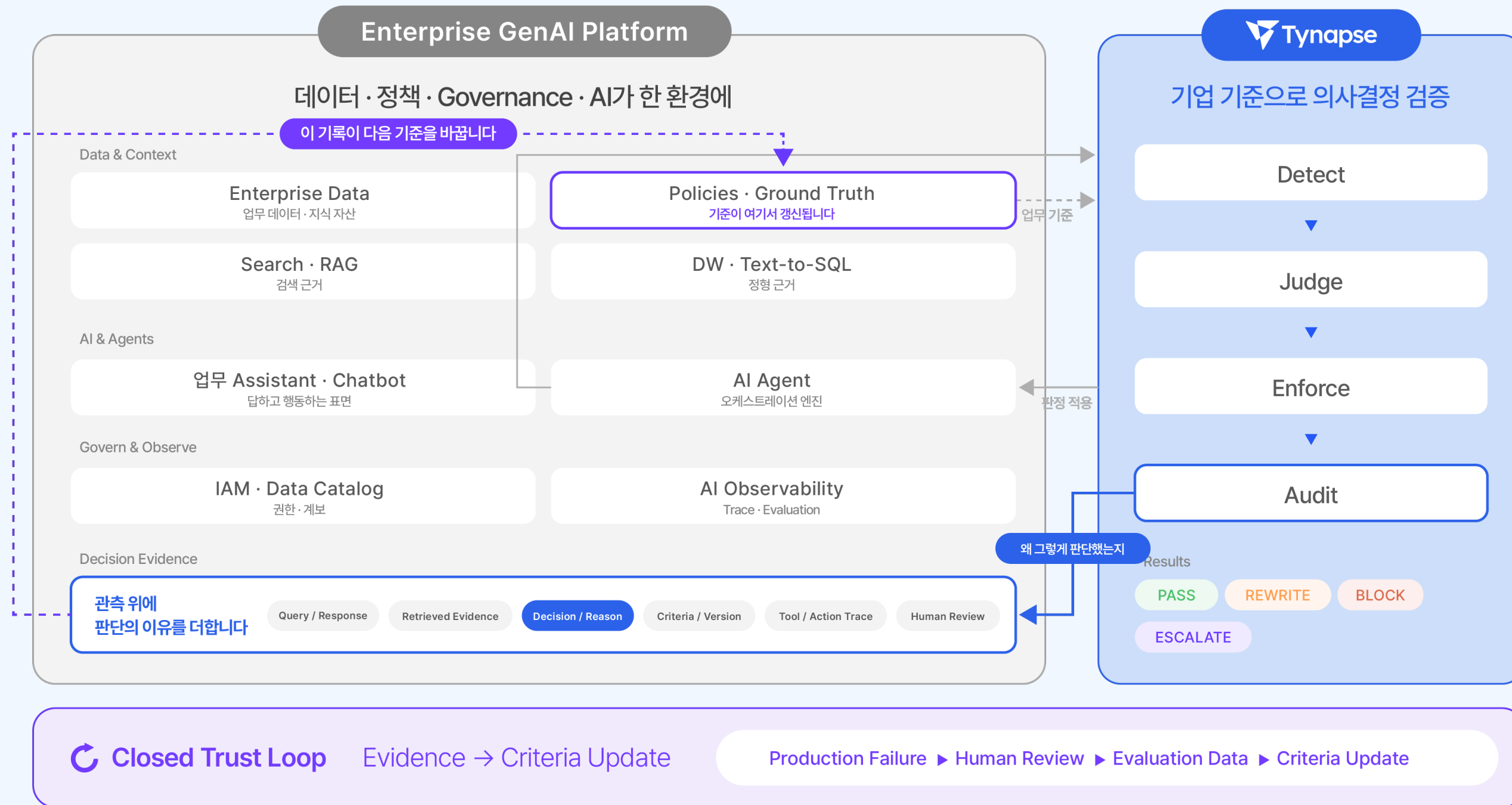
Runtime에서 결정합니다

Step 1 · Runtime Decision

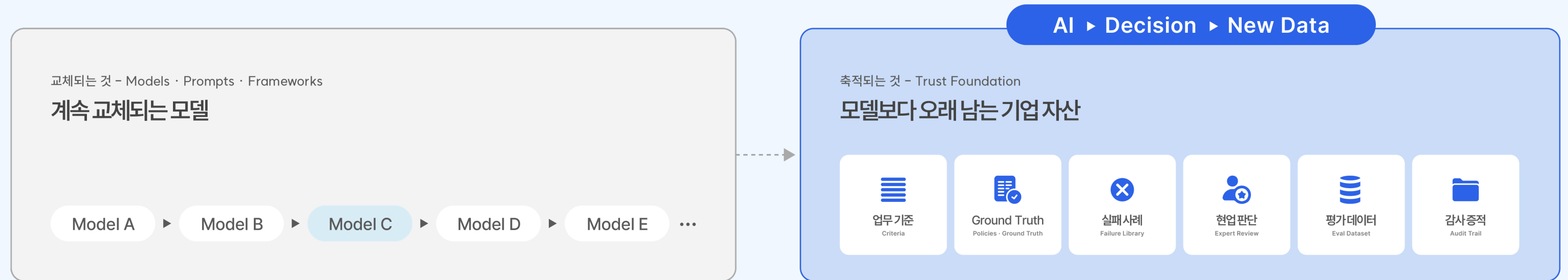


오늘의 실패가 내일의 기준이 됩니다

Step 2 · Closed Trust Loop



모델은 바뀌어도, 기업의 신뢰 기준은 남아야 합니다



모델은 교체되지만, 신뢰 기준은 축적됩니다

'도입할 수 있는가'에서 '말길 수 있는가'로

Why Now



AI를 얼마나 잘 만드는가만큼
어디까지 말길 수 있는가가 기업의 경쟁력이 됩니다

공개 벤치마크 기준으로 비교했습니다

Technical Proof

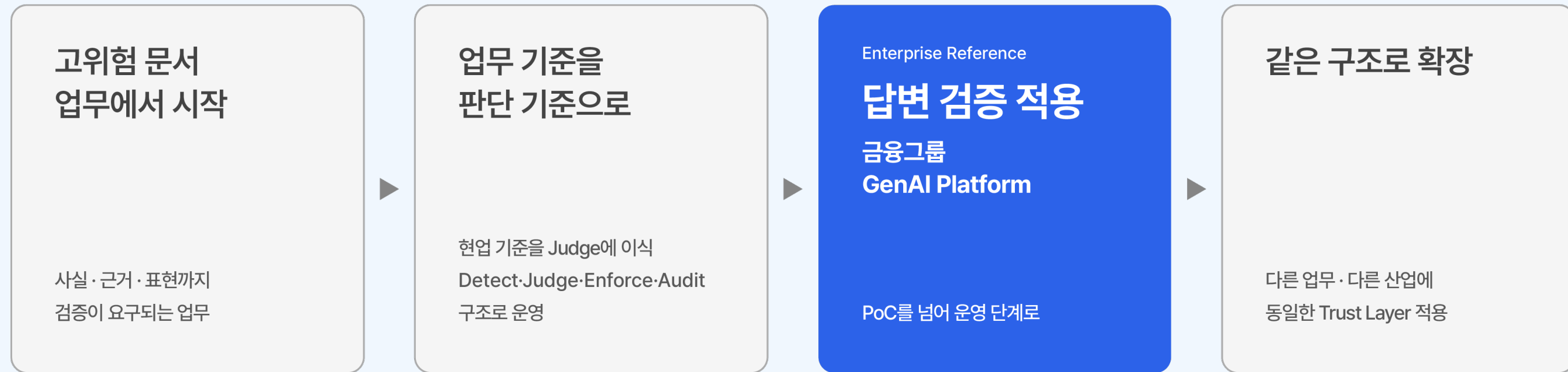


공개 벤치마크 · Artificial Analysis Guardrail 비교 (2026.06) 기준 · WildGuardTest · ToxicChat · XSTest · 동일 조건 평가

공개 기준에서 확인한 성능이
실제 업무 기준의 검증으로 이어집니다

가장 엄격한 업무에서 검증되고 있습니다

Enterprise Proof



Next

답변 검증에서 Agent 행동 통제까지
맡길 수 있는 범위를 계속 넓혀갑니다

규제 산업의 기준을 통과한 구조로
기업이 AI에게 맡기는 범위를 넓혀가고 있습니다

THANK YOU

더 많은 자율성을,
통제와 함께

More Autonomy. With Control.

최명섭 | Head of AI | Tynapse

