



AI 에이전트 시대의 국가전략

이 승 현*

Abstract

2025년에 AI는 응답하는 도구에서 행동하는 행위자(AI agent)로 진화했다. Stanford HAI의 AI Index가 인용하는 컴퓨터 사용 능력 벤치마크 OSWorld의 업무 성공률은 12%에서 66.3%로 치솟았고, 같은 해 글로벌 기업 AI 투자는 5,817억 달러를 기록했다. AI에이전트 시대가 시작되면서 AI의 능력·자본·도입의 세 축이 한꺼번에 임계치를 넘어선 것이다.

한국의 비대칭은 또렷하다. 한국 근로자의 업무 목적 생성형 AI 활용률은 51.8%로 미국(26.5%)의 약 두 배다(사용 경험 기준으로는 63.5%). 인터넷 상용화 3년 차(7.8%)와 비교하면 8배 빠른 확산이다. 인구 단위 도입률도 2026년 1분기 31.7%로 한분기에 7계단 올라 글로벌 18위까지 올라섰다. 하지만, 같은 시점 한국의 SaaS 점유율은 0.35%에 불과하고, AI 노출 직군의 22~25세 SW 개발자 고용은 약 20% 줄었다. 한국은행은 지식서비스 무역수지 102.5억 달러 적자의 원인으로 정부 통계 사상 처음 'AI 구독'을 명시했다. 사용은 가장 빠르게 늘지만, 결제·기억·가치 표준의 통제권은 여전히 약하다. 사용만 하다가는 곧 구조적 적자국이 될 수밖에 없는데도, 기존의 도입 프레임으로는 이 역설이 제대로 보이지 않는다. 이 진단에 대한 처방은 「연결된 자율성」이다. 자율성과 연결성을 양자택일이 아닌 곱셈으로 다룬다. 한국이 서야 할 자리는 주권 인프라(sovereign infrastructure), 동맹 기술 스택(alliance stack), 공개 가중치 조정자(open-weight orchestrator)를 동시에 쥐는 세 겹이며, 이는 김대중 정부의 정보화 인프라와 노무현 정부의 균형자 외교가 AI 시대에 융합되는 자리다. 골든타임이 끝나기 전에 측정·규율·투자·동맹·주권의 다섯 댄스 멤버가 동시에 서야 토큰의 폭증을 통한 경제 구조적 변화에 대응할 수 있다.

* 가천대 스타트업칼리지 겸임교수·전 대통령직속 디지털플랫폼정부위원회 AI 플랫폼 혁신국장, epoko77@gmail.com

I. 서론

1. 배경

2026년에 들어서면서, 인공지능(AI)의 능력·자본·도입이 한꺼번에 변곡점을 지나고 있다. 능력 측면에서 컴퓨터로 사람의 업무를 처리하는 벤치마크 OSWorld의 성공률은 2024년 초 12%에서 2년 만에 66.3%로 올라 인간 기준선(72.35%)과의 격차를 6%포인트 안쪽으로 좁혔다.¹⁾ 자본과 도입의 측면도 같은 속도다. 글로벌 기업 AI 투자는 한 해 5,817억 달러로 전년 대비 129.9% 늘었고, 조직 단위 도입률은 88%에 이르렀다.²⁾

세가지 측면에서 한꺼번에 변곡점을 지나면서 일어난 일은 단순한 기술 진보가 아니다. 인간 곁에 또 하나의 행위자가 들어서서 스스로 판단하고 거래하고 실행하기 시작했다. 행위자는 곧 책임의 단위이자 거래의 단위이며 정책의 대상이다. 행위자가 잘못되면 누구를 부르는가. 기술의 문제가 국가의 문제로 넘어가는 지점이다.

이러한 전환 앞에서 한국은 세계에서 가장 적극적인 사용국에 속한다. 국내 근로자의 63.5%가 생성형 AI를 사용해 본 적이 있고, 업무 목적으로 한정해도 51.8%에 이른다. 미국 근로자의 업무 활용률 26.5%의 약 두 배이며, 인터넷 상용화 3년 차(7.8%)와 견주면 8배 빠른 확산이다. 인구당 도입률도 2026년 1분기 31.7%로 단일 분기 7계단 상승했고, 글로벌 18위까지 올라섰다.³⁾ 그러나 활용력과 통제력은 별개다. 한국은행은 2025년 연간 지식서비스 적자 102.5억 달러의 원인으로 ChatGPT·Perplexity·Gemini 구독을 정부 통계 사상 처음 적시했다.⁴⁾ 사용은 빠르게 늘지만 토큰이 빠져나가는 결제 경로도, 일상의 판단을 매개하는 가치도 모두 바깥에서 정해질 수 있다. 도입률이 빠르고 높을수록 외산 인프라에 대한 의존은 더 빨라진다. 가장 활발한 사용국이 곧 구조적 적자국이 되버리면 단순히 AI 도입을 묻는 기존의 프레임에는 잘 잡히지 않는다. 그래서 이제 질문을 바꿔야 한다.

“누가 한국 국민의 AI 에이전트를 소유하고 누가 그 거래의 통제권을 질 것인가?”

1) Stanford HAI, *AI Index Report 2026*, Chapter 2(2026-04-15). OSWorld 12% → 66.3%, 인간 기준선 72.35%. <https://hai.stanford.edu/ai-index/2026-ai-index-report>

2) Stanford HAI, *AI Index Report 2026*, Chapter 4 Economy(2026-04-07). 글로벌 기업 AI 투자 5,817억 달러, 미국 단독 2,859억 달러로 중국(124억 달러) 대비 23.1배. 조직 도입률 88%.

3) Stanford HAI, *AI Index Report 2026*, Chapter 4 Economy(2026-04-07), 한국은행, BOK 이슈노트 제 2025-22호 「AI의 빠른 확산과 생산성 효과: 가계조사를 바탕으로」(2025-08-18)

4) 한국은행, 「2025년 상반기 지식서비스 무역통계(잠정) 보도자료」(2025-09-17); 아주경제 보도(2026-03-19), 2025년 연간 -102.5억 달러(전년 대비 약 39% 악화, +28.8억 달러).

이 질문은 단순히 AI의 문제를 노동 대체 문제나 후생의 문제가 아니라 주권 아키텍처라는 관점에서 문제를 바라봐야 한다는 의미다.

2. 국가전략의 정의와 한국의 자리

국가전략(Grand Strategy, 이하 GS)의 본질은 무엇을 얻을 것인가의 선언이 아니라 무엇을 위해 무엇을 포기할 것인가의 선언이라고 봐야 한다.⁵⁾ GS에 대한 학문적 정의는 세가지의 연구로 수렴한다. 니나 실로브(Nina Silove)는 GS를 “고차원 정치 목표와 그것을 달성할 군사·외교·경제 수단의 통합 설계”로 정리했다.⁶⁾ 러시 도쉬(Rush Doshi)는 중국의 GS를 시기별 대체 단계로 추적하면서 “전략은 선언된 한계의 예술이다”라는 헨리 키신저(Henry Kissinger)의 정의를 인용해 한계를 인정할 줄 아는 것을 핵심으로 짚었다.⁷⁾ 라몬 파체코 파르도(Ramon Pacheco Pardo)는 한국이 광복 이후 단 한 번도 GS를 명문화하지 못했고 그 공백이 부처별·정권별 분산을 만들었다고 적었다.⁸⁾

한국 정부의 GS 계보는 8개로 압축된다. AI에이전트 시대 우리가 지향해야 할 방향성은 그 가운데 김대중 정보화 인프라와 노무현 균형자 외교의 AI 시대 융합이다. AI 시대의

표 1 역대 정부의 국가전략

정부	grand strategy 키워드	AI 에이전트 시대의 좌표
노태우(1988~93)	북방외교	통상·외교 안보의 다극 진입
김대중(1998~03)	정보화(초고속망)	인프라 - AI 시대 K-메모리·NPU·전력 합성
노무현(2003~08)	동북아 균형자	도ctrine - issue-specific multi-alignment
이명박(2008~13)	녹색 성장	산업·에너지 전환
박근혜(2013~17)	창조경제	산업·R&D 융합
문재인(2017~22)	한국판 뉴딜	디지털·그린 동시 추진
윤석열(2022~25)	가치동맹	동맹 명료화
이재명(2025~)	실용외교	균형 회복

5) Henry Kissinger의 것으로 알려진 표현 “strategy is the art of declared limits” — Doshi 2021 *The Long Game* 을 통한 인용 anchor.

6) Nina Silove, “Beyond the Buzzword: The Three Meanings of ‘Grand Strategy,’” *Security Studies* 27:1 (2018). <https://www.tandfonline.com/doi/full/10.1080/09636412.2017.1360073>

7) Rush Doshi, *The Long Game: China's Grand Strategy to Displace American Order*(Oxford University Press, 2021).

8) Ramon Pacheco Pardo, *South Korea's Grand Strategy: Making Their Own World*(Columbia University Press, 2023).

GS는 두 자산을 동시에 추구해야 한다. K-메모리, NPU, 전력의 인프라 위에 사안별 다중 정렬(issue-specific multi-alignment)의 독트린을 올리는 작업이다.

3. 다중정렬 독트린(Issue-specific multi-alignment)

한국이 어느 진영을 택할 것인가는 그 프레임 자체가 낡았다. 2026년의 한국에서는 영역마다 정렬 축이 다르다. 안보는 미국·일본·호주 파트너 트랙, 통화·결제에는 BIS Agorá 7국 안의 부분 룰메이커 트랙, 표준은 미국과 EU 두 진영 사이 다중 정렬 트랙, 산업은 미·중·일·EU 분리 트랙, 가치는 한국 헌법 자립 트랙이다.

그러나 정렬은 면밀히 말하면 전략은 아니다. 각 영역마다 누구와 손잡는가를 그렸다고 해서 한국이 무엇을 쥐고 무엇을 내려놓을지가 정해지는 것은 아니다. 중견국이 AI 스택 전체를 리드하기는 어렵다. 주권은 모든 층을 갖는 것 보다, 어느 한 층에서 빠질 수 없는 존재, 즉 룰메이커가 되는 데 있으며, 그 자리를 지키는 일에는 반드시 다른 것을 포기하는 대가가 따른다. 그래서, 질문은 둘로 좁혀진다. 한국은 AI 스택의 어느 레이어에서 룰메이커가 되는가, 그리고 그 자리를 지키기 위해 무엇을 감수하는가. 같은 두 질문을 중견국가 5곳을 GS의 여섯 요소(Ends·Ways·Means·Sacrifice·Time·Legitimacy)로 늘어놓아 보면, 한국만 감수할 비용의 명시(sacrifice declare)가 비어 있다는 사실을 알 수 있다.

호주는 핵심광물(리튬 글로벌 약 절반)을 동맹 공급망에 묶는 대신 최대 수요처인 중국 시장을 내려놓았고, 네덜란드는 ASML이라는 노광 병목을 지키려 대중국 매출을 포기했다. 이스라엘은 군사 AI 자립의 값으로 외교적 고립을 감수했고, 싱가포르의 아시아 AI 컴퓨터 허브라는 자리를 얻는 대가로 미·중 사이를 오가며 균형을 잡던 자유도를 내려놓았다. 이스라엘은 군사 AI 자립을 지키려 다자적 정당성과 파트너십의 폭을 내려놓았다.

일본은 결이 조금 다르다. 설계·제조·모형을 모두 자국 안에 두는 ‘풀스택 자립’을 노리며, 이를 위해 Rapidus 2nm 파운드리를 만드는 데 국가 자본을 한곳으로 몰아넣었다(METI 연 1.23조 엔, Rapidus 누적 약 2.6조 엔). 인구 절벽은 일본이 감수한 비용이라기 보다는 그 베팅을 선택한 배경이다. 노동력이 줄기 전에 자본과 로봇으로 갈아끼워야 하기 때문이다. 일본이 실제로 내려놓은 것은 선진국 최고 수준의 정부부채를 안은 나라가 재정 여력을 단일 고위험 베팅에 집중하면서 헤징의 자유를 포기한 것이다. 전략은 한정된 자원에서 ‘무엇을 하지 않을지’를 고르는 일이고, 감수할 대가를 못박는 순간에야 그 자리는 지킬 수 있는 자리가 된다. 한국만 그 칸이 비어 있다. 역설은 바로 여기에 있다.

표 2 각국의 6개 국가전략요소에 따른 AI 국가전략

국가	Ends	Ways	Means	Sacrifice	Time	Legitimacy
호주	AI 하드웨어 공급망 원자재· 호스팅	핵심광물 (리튬~50%)· 데이터센터 허브· AUKUS P2 AI	광물· 토지·에너지	중국 시장(광물 최대 수요처) 정렬 감수	2023~2030	핵심광물 전략·미·호 프레임워크 (2025)
네덜란드	AI 칩 제조 장비충 률메이커	ASML EUV	첨단 노광 독점기술	ASML 대중국 매출 포기	5~10y	미 수출통제· EU정합
이스라엘	인재·군사AI 자립	8200/Talpiot· 자체 군사LLM· AI·자율성형	인재·이중용도 R&D	외교 고립 감수	5y	국방부· 이노베이션청
싱가포르	AI 컴퓨트 허브· 거버넌스 롤세터	NAIS 2.0· AI컴퓨트 클러스터·AI Verify	\$10억+ (컴퓨트~\$5 억)·데이터센터	미·중 균형 자유도 (칩 중계 감시)감수	2024~2029	국가AI위 (2026-02, 총리 의장)
일본	국산AI 풀스택 자립	METI ¥1.23조/yr (2026~)· Rapidus 2nm· 소버린 모델	자본·로보틱스 (Physical AI)	인구 절벽 자본집중 압축 감수	~2027	AI추진법 (2025)· 자민당·METI
한국	메모리 초크포인트 기반AI 주권	HBM 글로벌1위 (SK하이닉스· 삼성)·HBF 표준화 주도	메모리 초크포인트· HBM→HBF	-	HBM 1위(현재)· HBF 표준화2026~	AI기본법 (2026-01-22) ·국가AI전략위

4. 문명적 변곡점

왜 하필 2026년일까? 답은 어느 AI 모델의 출시 시점이 아니라, 여러 예측이 한 점에 겹칩다는 데 있다. 그 점이 바로 AI 에이전트다. 샘알트만(Sam Altman)은 2025년 1월 에이전트가 노동력으로 합류해 기업의 산출을 실질적으로 바꿀 것이라 했고, 같은 달 NVIDIA의 젠슨 황(Jensen Huang)은 CES에서 에이전트 AI의 시대가 왔다고 하며 이를 수조 달러 규모의 전환으로 언급했다.

엔트로픽의 다리오 아모데이(Dario Amodei)는 강력한 AI를 질문에 답하는 모델이 아니라 도구를 쓰고 세계에 직접 개입하는 행위자로 규정하고, '인류는 곧 상상하기 어려운 힘을

손에 쥐게 된다'는 표현으로 그 전망을 재확인했다.⁹⁾ 문명적 변곡점이 국가 정책에 위협이 되는 까닭은 두 시계가 다른 속도로 움직이기 때문이다. 능력의 시계는 빠르다. 독립 평가 기관 METR에 따르면, AI 에이전트가 사람의 개입 없이 끝까지 수행하는 과업의 길이는 약 7개월마다 두 배가 된다. 반면 예산 주기도 정부부처의 대응도 여전히 AI시대를 따라오지 못하고 있다. 능력이 배가되는 동안 정책은 한번 갱신되기도 어렵다.

이러한 위기 속에서 본고에서 제시하고자 하는 것은 크게 다섯 가지다.

첫째, 국가전략을 학문적으로 정의하고 그 위에 한국의 자리를 역대 8개 정부의 계보 안에 명시한다. 둘째, AI를 세력 전이와 범용목적기술이 동시에 작동하는 가장 중요한 첫 변수로 자리매김한다. 셋째, 외화·인지·기억 주권의 결합 충격을 교차쌍(cross-pair)으로 식별한다. 넷째, 주권 인프라 × 동맹 스택 × 공개 가중치 조정자(sovrenign infra × alliance stack × open-weight orchestrator)의 3중 매트릭스 위에서 실질적 목표와 다섯 건의 Trade-off를 정리한다. 다섯째, 5축 도메인 분석과 다섯 개의 교차 레버를 매트릭스로 묶는다. 이 다섯이 한 점으로 모이는 자리에 「연결된 자율성」이 있다.

II. AI 에이전트의 부상

1. 세력 전이와 범용목적기술이 만난 첫 변수

AI가 과거의 어떤 기술과도 다른 이유는, 역사상 강력했던 기술들조차 두 가지 효과를 한꺼번에 일으킨 적이 없다는 데 있다. 하나는 세력 전이(Power Transition), 패권국과 도전국 사이에서 힘의 순위를 다시 매기는 효과다. 다른 하나는 가치 단위 전환, 경제가 무엇을 기준으로 가치를 만들고 측정하는지, 그 셈법을 바꾸는 범용목적기술(General Purpose Technology, GPT)의 효과다. 20세기의 결정적 기술들은 둘 중 하나에만 작용했다.

핵무기는 첫 번째 축에만 작용했다. 핵무기는 미·소 냉전의 군사적 우열을 갈라 국가 간 힘의 순위를 결정짓는 데 결정적이었지만, 경제가 가치를 만드는 셈법 자체를 바꾸지는 않았다. 군사 패권의 도구였을 뿐 경제 패러다임의 전환자는 아니었다.

반도체는 정반대로 두 번째 축에만 작용했다. 디지털 경제 전체의 토대를 깔아 가치의 원

9) Dario Amodei, "The Adolescence of Technology"(2026-01). "humanity is about to be handed almost unimaginable power."

천을 연산과 데이터로 재정의했지만, 그 자체가 어느 국가의 패권을 직접 가르지는 않았다. 반도체가 세력 전이의 전면 변수로 떠오른 것은 오히려 그 위에서 AI가 작동하기 시작한 지금에 이르러서다. 인터넷과 전기는 범용목적기술이었다는 점에서 AI와 가장 가깝지만, 세력 전이를 일으키는 방식이 달랐다. 인터넷은 그 힘이 특정 국가에 응집되지 않고 여러 행위자에게 분산되었다. 전기는 모든 산업에 서서히 스며들면서 패권의 향배를 한 시점에 몰아 결정짓는 단일한 시간선을 만들지 않았다.

이에 비하면, AI는 세력 전이와 범용목적기술이라는 두 효과가 동일한 시간선 위에서 동일한 강도로 결합된 첫 변수다. 핵무기의 힘의 재편 효과와 반도체의 가치 단위 전환 효과가 인터넷·전기처럼 분산되거나 지연되지 않고 같은 국면에서 한꺼번에 터져 나온다. 그 차별점의 핵심이 바로 행위자로 위상의 격상이다. 과거의 AI는 인간이 사용하는 도구에 머물렀지만, AI 에이전트는 스스로 거래하고 계약하고 책임을 지는 경제 행위의 주체로 올라선다. 도구가 아니라 행위자가 되는 이 전환이 AI를 앞선 네 기술들과 질적으로 구분 짓는다. 이게 바로 본 고에서의 첫 번째 명제다.

2. 표준 거버넌스의 수렴과 부분 룰메이커 비대칭

분절돼 있던 에이전트 통신 프로토콜은 2025년 12월 리눅스 재단 산하 AAIF(Agentic AI Foundation)가 출범하면서 판도가 바뀌었다. 엔트로픽·블록·오픈AI 3개사를 중심으로 약 50개 회원 기관이 참여했고, 창립 기여 프로젝트는 엔트로픽의 MCP(Model Context Protocol), 블록의 goose, 오픈AI의 AGENTS.md이며, 구글이 이관한 A2A(Agent-to-Agent) 프로토콜까지 같은 거버넌스 아래 모였다.

MCP는 에이전트가 외부 도구·데이터에 접속하는 표준이고, A2A는 에이전트끼리 통신·거래하는 표준이다.¹⁰⁾ IBM과 AGNTCY의 ACP까지 리눅스 재단 산하로 정리된 점은 룰메이커로서의 위상을 더 굳혔다. 하지만, 아쉽게도 한국 기업·기관의 비중은 미미하다.¹¹⁾ 다만 모든 영역에서 한국이 같은 위상에 있는 것은 아니다. 통화·결제 사정은 다르다. 한국은행은 BIS(국제결제은행) Project Agorá의 7개 창립 중앙은행 가운데 비G7 3개국(한국·멕시코·스위스)의 하나다. Project Agorá는 토큰화 통합 원장과 CBDC를 결합해 차세대 결제 인프라를 시범하는 BIS 주관 프로젝트다.¹²⁾ 이 프로젝트가 중요한 이유는 AI 에이전트 간

10) Linux Foundation, "Announcing the Agentic AI Foundation"(2025-12-09)

11) AAIF 2026-06 사양 의제, server-as-agent, fine-grained authorization, agent identity, MCP-A2A interop.

12) BIS, "Project Agorá"(2026-03). 7개 창립 중앙은행: 미 연준 NY·일본은행·BOE·프랑스은행(유로시스템 대표)

의 ‘초소액·고빈도·실시간 국가 간 결제’를 기존 신용카드망 대신 처리할 차세대 글로벌 인프라이기 때문이다. 한국은행의 Project Hangang Phase 2는 CBDC와 예금토큰을 결합해 9개 시중은행과 정부 보조금 실거래 시범에 들어갔다. 명시적 목표에는 ‘AI 기반 자동 결제 인프라’가 포함되었다.¹³⁾ 모델·에이전트 표준에서는 룰메이커지만 통화·결제 표준에서는 7국 중 1국의 부분 룰메이커라는 비대칭이 존재한다. BIS Agorá의 발언권 자산이 모델·에이전트 영역의 카운터파트로 환산될 수 있다는 가능성이 충분하다.

3. 가치 단위의 이동

산업 자본주의에서 가치 단위는 노동시간이었고 GDP는 결국 인간 노동시간의 화폐 환산이었다. 그런데, 에이전트 경제에서 가치를 창출하는 주체는 더 이상 인간만이 아니고, 새로운 가치 단위는 토큰과 태스크다. 젠슨 황은 GTC 2025 키노트에서 “AI factory가 토큰을 생산한다”고 했고,¹⁴⁾ 추론 수요가 “skyrocket”한다고 표현했고, 벤 톰슨은 같은 흐름을 “Agentic inference”의 시대로 정리했다.¹⁵⁾ 토큰의 가치를 인간의 노동시간으로 환산하려는 것은, 증기기관의 출력을 말 몇 마리에 해당하는지로 따지려던 것과 같다. 그렇다고 GDP를 폐기하자는 뜻은 아니라, 토큰과 태스크라는 새 가치 단위를 담아낼 위성계정으로 보완하자는 것이다. 실제로 OECD는 「Going Digital Measurement Roadmap 2026」에서 10개의 실행과제로 디지털·AI 시대의 GDP 보조 측정 표준을 정비하고 있다.

문제는 그 효과가 얼마나 될지, 또 언제 통계에 잡힐지 아무도 확인하지 못한다는 데 있다. IMF는 AI 도입이 연간 GDP를 0.1~0.8%p 끌어올릴 수 있다고 보았지만, 대런 아제모글루(Daron Acemoglu)교수는 향후 10년간의 총요소생산성 효과를 0.66%로 매우 보수적으로 추정했다.¹⁶⁾ 이러한 전망의 엇갈림이야말로, 토큰의 흐름을 실시간으로 포착하는 측정 인프라를 갖추는 일이 모든 정책에 앞선 출발점이어야 하는 이유다.

노벨 경제학상 수상자 로버트 솔로우는 1987년 “컴퓨터 시대가 도래했음은 어디서나 볼 수 있지만, 정작 생산성 통계에서는 보이지 않는다”고 꼬집었다. 이른바 솔로우의 역설(Solow Paradox)이다. 폴 데이비드(Paul David)는 1990년 『American Economic Review』에서 전

·Banco de México·스위스 국립은행·한국은행.

13) Ledger Insights, “Bank of Korea launches Phase 2 of Project Han River”(2026-03-18)

14) Jensen Huang, NVIDIA GTC 2025 Keynote(2025-03), “AI factory가 토큰을 생산한다.”

15) Jensen Huang, NVIDIA GTC 2026(2026-03), “Every country and every company wants AI that reflects their values,” 추론 수요 “skyrocket” 발화; Ben Thompson, Stratechery, “agentic inference.”

16) Daron Acemoglu, “The Simple Macroeconomics of AI”(NBER WP 32487, Economic Policy 2025-01).

기가 발명된 뒤 공장의 생산성을 끌어올리기까지 약 40년이 걸렸다고 했다. 다만 AI의 시간은 그때보다 훨씬 짧다. 소프트웨어와 클라우드를 타고 즉시 배포되는 AI의 속성이 과거 전기가 요구했던 40년의 물리적 시차를 훨씬 더 짧게 압축하고 있기 때문이다. 한국에는 이 압축이 양날의 검이다. 통제가 약한 구조 위에서는 짧아진 시차가 곧 적자 경로의 가속이지만, 측정과 통제의 기반을 갖춘다면, 같은 압축이 추격의 시간을 단축하는 기회가 된다.

Ⅲ. 무엇이 달라지는가?

1. 책임의 증발

엔트로픽의 다리오 아모데이는 강력한 AI 에이전트가 가져올 인류의 미래 위험을 말했다. 인간을 뛰어넘는 지능을 가진 AI가 스스로 판단하고 행동하는 자율성을 갖고, 목표를 향해 밀고 나가는 일관성까지 갖춘 상태에서 인간의 통제력마저 상실된다면 어떻게 될까? 그는 이 네 가지 속성의 결합이야말로 인류의 생존을 위협하는 가장 치명적인 위험의 불씨가 될 것이라고 강력히 경고하면서, “intelligence, agency, coherence, poor controllability” 이렇게 4가지 속성이 결합되면 실존적 위험을 초래하는 레시피가 된다고 주장했다.¹⁷⁾

요슈아 벤지오(Yoshua Bengio) 교수는 “AI 에이전트는 자율적으로 행동하기 때문에 한층 더 높은 위험을 초래한다.” 라고 했다.¹⁸⁾ 엔트로픽의 잭 클락(Jack Clark)은 프런티어 모델의 자율 행동 능력을 칼을 진 유아들로 비유했고, 무스타파 술리만은 강력한 기술을 봉쇄할 수 없다는 사실이 거버넌스의 정의를 다시 쓰게 한다는 봉쇄 딜레마를 제기했다.¹⁹⁾ 출발점이 다른 네 사람이 같은 진단에 도달한 것이다. 이러한 동일한 진단은 정책 표준에도 영향을 미친다. 그런데, 한국 법제의 민법·상법·행정법·금융법은 비인간 행위자를 전제하지 않는다. 행위자가 잘못되면 누가 책임지는가. 한국 법제에는 그 자리가 비어 있는 것이다.

그렇다면 2026년 1월 22일 전면 시행된 AI기본법(「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」)은 그 빈자리를 채울까? 채우지 못한다. 투명성·안전성 확보, 고영향 AI 영

17) Dario Amodei, “Machines of Loving Grace”(2024-10). “intelligence, agency, coherence, poor controllability... existential danger.”

18) International AI Safety Report 2026(2026-02-03). “AI agents pose heightened risks because they act autonomously.”

19) Mustafa Suleyman, *The Coming Wave*(Crown, 2023). containment dilemma.

향평가, 국내대리인 지정 의무는 모두 ‘인공지능사업자’에게 부과된다. AI를 규율 대상인 제품·서비스로 보고 그 뒤에 선 사업자(인간·법인)에게 책임을 돌릴 뿐, 에이전트 자체를 의무나 책임의 주체로 인정하지 않는다. 더 결정적으로 안전장치는 “사람이 개입하여 통제 가능한 경우” 면제되는데, 에이전트의 정의 자체가 바로 그 인간을 루프에서 빼내는 데 있다. 비인간 행위자를 담을 자리가 최신 법령에도 비어있는 것이 바로 ‘책임의 증발’이다.

2. 인터페이스의 격상과 회의론

인터페이스가 코드에서 컨텍스트, 그러니까 에이전트에게 위임하는 방식으로 격상되었다는 진단에는 회의론도 따른다. 그렇다고 해도 둘은 층위가 다를 뿐 상충하지 않는다. 얀 르쿤(Yann LeCun)은 2026년 4월 브라운 대학 강연에서 오늘의 에이전트 시스템 가운데 자기 행동이 물리 세계에 가져올 결과를 예측할 수 있는 것은 거의 없으며, 세계가 어떻게 움직이는지에 대한 모델(World Model)의 부재를 지적했다.²⁰⁾ 일리야 수츠케버는 컴퓨터를 100배로 키운다고 모든 것이 풀리지는 않으며, 다음 도약은 규모가 아니라 새로운 학습 패러다임에서 나온다고 보았다.²¹⁾ 인터페이스 차원, 능력 차원, 방법론 차원, 이 세가지가 함께 고민되어야 하는 것이다.

3. 인지와 기억의 주권

외산 AI 에이전트가 한국의 헌법, 법률, 행정, 의료 가이드라인, 교육 과정, 문화 맥락을 학습하거나 기억하지 않은 채 한국 국민의 의사결정을 매개한다면 어떤 일이 벌어질까.

어떤 모델이 한국어를 더 잘하는지 경쟁을 넘어서 주권이 걸린 질문은 따로 있다. 누가 한국 국민의 의사결정을 매개하는가. 언어 점수와 가치·기억의 주권은 애초에 다른 축이다.

첫 번째 가치의 외주다. 미국 「AI 액션 플랜」은 NIST 리스크 관리 프레임워크에서 다양성·형평성(DEI)과 허위정보 항목을 삭제하고 검열 없는 고성능 모델을 표준으로 채택했다. 한국이 일상의 가치판단을 외산 에이전트에 위탁하는 순간, 한국 헌법 가치와 소수자 보호 기준이 조용히 무력화될 위험이 있다. 한국어로 매끄럽게 답하고 목표를 수행하는 에이전트일수록 그 가치 판단이 어느 나라 표준에서 나온 것인지는 구별하기 쉽지 않다.

20) Yann LeCun, Brown University 강연(2026-04). “almost none of [current AI systems] are capable of predicting the outcome of their actions in the physical world.”

21) Ilya Sutskever, Dwarkesh Patel 인터뷰(2025-11-25). “the age of research.”

두 번째는 기억이다. 모델이 보편 지식을 담은 뇌라면, 에이전트의 기억은 사용자의 삶의 궤적을 담은 자아에 가깝다. NIST AI 600-1 v1.1은 Agentic AI의 특수 위험으로 지속 기억 유출을 식별했고, 앤트로픽의 MCP Authorization 사양은 OAuth 2.1 기반 동의를 명시했다. 어떤 표준을 따르느냐가 곧 어떤 관할권에 들어가느냐를 가른다. 한국 국민 5천만 명의 의료 기록·가족 관계·정치 성향·종교·직장 의사소통·자녀 학습 이력이 외국 관할권 서버에 누적되면, 그때부터 어떤 외교·통상 갈등에서든, 한국 시민의 기억을 쥔 사업자가 협상의 카운터파트로 등장한다. 다음 통상 의제는 관세가 아니라 기억일 수 있는 것이다.

하지만 한국이 결정적으로 뒤처져 있지는 않다. 「개인정보 보호법」 제37조의2는 자동화된 결정 거부권은 EU GDPR 제22조보다 넓다. 다만 그 권리를 실제로 행사할 운영 인프라인 에이전트 식별·로그·감사 체계는 빈약하다. 다른 하나는 물적 토대다. 폐권의 무게중심이 연산에서 기억으로 옮겨가면서, 누적 기억을 담는 기억 계층(단기 SRAM·중기 HBM·장기 HBF)을 누가 설계하느냐가 메모리 강국인 한국에게는 종속이 아니라 도약의 갈림길이 된다.

인지와 기억의 주권은 곧 시민 주권이며, 외화 주권(통화)과 물적 토대의 주권(산업)이 더해져 시민·산업·통화 주권 삼분을 이룬다. 이 변화가 국가전략 차원에서 무엇을 다르게 만드는지 답은 미·중 세력 전이 위에서 한국의 결합 지점이 어디서 정조준되는가에 있다.

IV. 국가전략 차원에서 무엇이 달라지는가

1. 세력 전이와 3중 위기의 통합 프레임

미·중 세력 전이 위에서 외화·인지·기억의 세 결합 지점이 더 중요해진다. AI 챗봇의 응답이 AI 에이전트의 행위로 격상되면서 인지 주권이 침식되고, 표준 거버넌스 문제가 기억 주권 문제를 결정할 수 있다. 가치 단위가 AI 중심으로 이동하면서 외화 주권은 항구히 꺾여 나간다. 셋은 결과적으로 AI에이전트라는 같은 메커니즘의 다른 측면들이다. 다리오 아모데이는 중국을 “첨단 기술을 활용한 감시 국가를 운영하고 있다.”라는 표현으로 세력 전이의 비대칭을 짚었다.²²⁾ 샘 알트먼은 민주주의 국가들이 AI권력을 널리 분산시킬지, 집

22) Dario Amodei, “Machines of Loving Grace”(2024-10). “Their government is currently autocratic and operates a high-tech surveillance state.”

중시킬지를 결정하게 되리라 했고,²³⁾ 유발 하라리는 인류가 인간과 AI가 결합한 하이브리드 문명을 경험해 본 적은 없다며 현재의 같은 분기점을 문명사적으로 진단했다.²⁴⁾

외화 주권이 흔들리면 토큰 지출이 외국 인프라로 빠지고, 인지 주권이 흔들리면 사용자의 일상 가치판단을 외산 에이전트에 맡기게 되며, 기억 주권까지 흔들리면 그 판단이 외국 관할권에 쌓인다. 한 축이 흔들리면 나머지 두 축도 함께 흔들리기에, 처방도 묶여야 한다.

2. 도입률 63.5%는 자량이 아니라 압축된 비용

앞서본 미국의 두 배가 넘는 한국 근로자의 생성형 AI 활용률은 63.5%는 흔히 AI강국의 증거로 읽힌다. 하지만, 같은 시점에 22~25세 SW 개발자 고용은 2022년 말 정점 대비 2025년 7월까지 약 20% 줄었다.²⁵⁾ 미국에서는 AI 영향 직군의 22~25세 고용이 13% 떨어졌다고 보고했다.²⁶⁾ 빠른 채택, 깊은 절벽, 더디고 불확실한 보상, 이 셋이 압축된 시차 위에 함께 작동하면, 압축은 이득이 아니라 비용이다. 도입이 빠를수록 외산 의존이 커지고, 신입 절벽이 겹치며 그 의존을 보완해야 할 인적자본이 양 끝에서 잠식되기 때문이다. 그래서 63.5%는 자량이 아니라 압축을 가속하는 페달이다. 본 고의 두번째 중요한 명제다.

3. Sovereign X alliance X open weight 3중 매트릭스

젠슨황은 2026년 3월 GTC 키노트에서 “모든 국가와 기업은 자신들만의 구체적인 데이터와 문화, 그리고 가치관을 반영하는 AI를 원한다.”라고 주장했다.²⁷⁾ 반도체 및 AI 인프라의 최고 권위 분석가인 달란 파텔(Dylan Patel)은 소버린 AI(Sovereign AI)를 “막대한 인프라 투자(CapEx)를 지속 가능하게 만드는 유일한 길은 결국 AI를 통한 실질적인 부(Wealth)의 창출뿐이다”라고 규정했다.²⁸⁾

23) Sam Altman, 의회 증언(2026-02). “democratic nations will determine whether that power is broadly distributed or concentrated.”

24) Yuval Noah Harari, 인터뷰(2026-01). “human history... no experience of hybrid human-AI civilization.”

25) 국회예산정책처, 「생성형 AI 고노출 직업 현황과 최근 청년 고용」(2025-12)

26) Erik Brynjolfsson, Bharat Chandar, Ruyu Chen, “Canaries in the Coal Mine?”(Stanford Digital Economy Lab, 2025-11).

27) Jensen Huang, GTC 2026(2026-03). “Every country and every company wants AI that reflects their values.”

28) Dylan Patel, SemiAnalysis(2025-12). sovereign AI 자본 분석.

표 3 각국의 소버린AI 관련 자본 규모

권역	Sovereign 동원 규모	좌표
미국	Stargate 4년 5,000억\$ + 즉시 1,000억\$	동원·표준·자본 동시 주도
EU	InvestAI 2,000억€ + gigafactory 200억€	규제·시장 결합
영국	Sovereign AI Unit 5억£ + AIRR 20배 확장	AI시 평가 표준
일본	METI 7년 10조¥	산업 부활
인도	IndiaAI 38,000 GPU · 시간당 65루피 · BharatGen	인구·언어 sovereign
UAE	MGX 1,000억\$ + HUMAIN	비석유 GDP 전환
한국	국민성장펀드 12조 5,000억 원, 국가 AI 컴퓨팅 센터(NAICC, 2028년 완공 목표): 2조 5,000억 원, 국산 반도체(NPU/PIM) 및 공공 AX 재정 예산 1조 원+	소버린 인프라와 국산 하드웨어 결합

한국의 소버린 자본 동원은 약 16조원으로 일본 METI 10조 엔과 비교해도 매우 부족하다. 동등한 자본 동원이 어렵다면, 집중도 또는 차별화된 포지셔닝으로 비대칭 경쟁력을 확보하는 방향뿐이다. 인도는 소버린 단일 축에 인구·언어 자산을 묶었다. 프랑스 Mistral은 open weight 단일 위치를 잡았다. 한국에는 어느 측면에서도 모호해 보인다.²⁹⁾

그래서 한국이 잡을 수 있는 지점은 Sovereign infra × Alliance stack × Open weight orchestrator의 3중 매트릭스다. 한국이 중견 핵심 행위자로 자리 잡으려면 세 가지 축이 동시에 굴러가야 한다.

첫째는 주권 인프라(sovereign infra)다. 한국이 세계 1위를 쥔 K-메모리(HBM·HBF)에 국산 NPU와 이를 받쳐 줄 전력을 더해, 에이전트 경제가 돌아갈 연산 기반, 그리고 독자 파운데이션 모델을 통해 외국 인프라에만 기대지 않고 자국 안에 세우는 일이다.

둘째는 동맹 스택(alliance stack)이다. 한·미 Technology Prosperity Deal로 기술·산업 협력의 좌표를 묶고, 국경 간 결제·정산의 표준을 짜는 BIS Agorá 안에서 규칙 수용자가 아니라 부분 규칙 제정자의 자리를 확보하는 일이다.

셋째는 오픈 웨이트 조정자(open-weight orchestrator)다. 프런티어 기반모델을 독자파운데이션모델로 만들어가는 동시에 Llama, DeepSeek, Qwen 같은 공개 가중치 모델 위에 한국어와 한국 가치에 정렬한 SLM(소형 언어모델)을 얹어 막대한 학습 비용 없이도 가치 함수와 한국어 맥락의 통제권만큼은 자국이 쥐는 일을 병행해야 한다.

29) 글로벌 sovereign AI 자본 동원 비교: OECD AI Policy Observatory Stargate Project(2025-01-21), UK Compute Roadmap(2025, Sovereign AI Unit £500M); PIF HUMAIN launch(2025-05-12), UAE National Strategy for AI 2031(MGX \$100B), DD News(MeitY/India), IndiaAI Mission BharatGen(2025), European Commission, InvestAI(2025-02-11).

이 세가지 축은 따로 굴러가면 안된다. 인프라가 없으면 조정자는 남의 연산 위에서 돌고, 동맹 스택의 자리가 없으면 그 인프라도 표준도 외국 관할권에 종속되며, 조정자가 없으면 인프라와 동맹은 한국의 언어와 가치를 담지 못한다. 셋이 맞물려 돌 때 한국은 AI 스택의 한 층에서 룰메이커가 되고, 그 자리에서 중견 핵심 행위자로 선다.

3중 매트릭스의 선언은 미국의 기술 안보 표준은 소버린 인프라와 동맹 스택으로 충족시키고, EU의 주권 규제는 글로벌 오픈소스를 통제하는 오케스트레이터 역량으로 대응하는 양면 전략의 실체이자, 대한민국 AI 주권을 지켜낼 본 고의 세 번째 명제다.

4. 실질 목표(Ends)와 Trade-off

본 페이지가 제안하는 실질 목표는 한 줄로 정리된다.

“AI 에이전트 시대 시민·산업·통화 주권의 동시 유지”

물론, 세 주권에는 우선순위가 있다. 시민 $\geq$ 산업 $>$ 통화. 시민 주권은 인지·기억 주권을, 통화 주권은 외화 주권을 포괄하며, 산업 주권은 그 물적 토대인 메모리·NPU·전력·독자모델 등을 의미한다. 동맹은 네 번째 목표가 아니라 이 셋을 지키기 위해 위치를 조정하는 변수다. 그래서, 무엇을 위해 무엇을 포기할지를 다섯 가지 트레이드 오프를 정리해보았다.

T1(산업 주권). 데이터센터의 60%를 비수도권으로 분산해 집적 효율보다 전력 그리드 용량에 무게를 싣는다. 집적 효율의 손실이 생기지만 수도권 전력망은 이미 송전 한계에 닿아 생각보다 크지 않다. 분산된 비수도권 거점을 소형모듈원자로(SMR) 전력 패키지와 묶으면 각 거점은 송전망에 기대지 않는 전용 전력을 갖출 수 있다. 집적 효율을 일부 내주는 대신, 전력 제약에 묶이지 않는 지방 분산형 에너지 주권을 얻을 수 있다.

T2(시민 주권). 외산 에이전트가 한국 국민의 데이터를 한국 관할권 안에 머무르도록 의무화한다. 에이전트 시대에 모델 자체의 락인은 오히려 약해지지만, 대부분의 에이전트 스택, 그러니까 하네스·톨콜 규약·오케스트레이션 등은 미국 빅테크 모델에 맞춰 최적화돼 있어, 한국 국민의 의사결정을 매개하는 데이터는 결국 미국에 최적화된 레이어를 통과할 수 밖에 없다. 머러링 의무는 그 층에 한국 관할권을 세우는 일이고, 미국은 이를 자국 에이전트 플랫폼에 대한 디지털 무역장벽으로 규정할 것이다. 여기서 작동하는 구조가 통상 양보 풀이다. 한·미 통상 협상은 농산물·자동차·디지털·방위비 등 서로 다른 영역의 요구가 하나의 교환 묶음에 함께 올라간다. 머러링이 디지털 칸의 장벽으로 규정되면서 시

민 주권을 지킨 대가는 클라우드 공공영역 개방같은 산업 카드 한두 장으로 청구된다. 결국 중요한 건 협상 구도다. 미국과 일대일로 마주 앉으면 미러링은 ‘한국만의 보호주의’가 되지만, 같은 의무를 다자 틀에 심으면 여러 파트너가 공유하는 디지털 주권 규범이 된다. 규범의 정당성이 높아질수록 청구 가능한 양보의 상한은 내려간다.

T3(산업 주권). 국산 NPU의 공공조달 비중을 50%로 끌어올린다. 신생 하드웨어 산업은 성능이 아니라 물량으로 학습곡선이 오르기 때문에 보장된 공공 수요 없이는 양산·수율·반복의 기반 자체를 갖지 못한다. 다만, NPU만 따로 깔면 절반의 해법이라, 독자 파운데이션 모델과 한 묶음으로 묶여야 한다. 공공조달은 NPU의 수요 기반인 동시에 독자 모델의 최적화 표적이 되고, 독자 모델은 국산 NPU에 맞춰 설계됨으로써 외산 스택에 대한 종속을 하드웨어 단에서부터 끊을 수 있다. NPU는 학습에서는 GPU와 경쟁할 수 없지만, 사전학습보다 추론 보완이 중요해진 시대에 추론용으로 충분히 가능하다.

물론 치러야 할 비용, 대가도 존재한다. 핵심은 성능 격차다. 다만, 정부와 공공이 실제로 돌리는 일은 범용 벤치마크가 아니라 한국어 행정·의료·교육 같은 명목 격차의 상당 부분은 실제 과업에서 좁혀진다. 외산이 앞서는 것은 ‘평균적으로 무엇이든 잘하는’ 범용 성능이고, 한국이 좁히는 것은 ‘한국이 실제로 시키는 일을 잘하는’ 실효 성능이다. 그래도 차이가 완전히 사라지지는 않는다. 이렇게 남는 성능 차이가 다섯 trade-off 가운데 유일하게 끝까지 다 메우지 못하는 비용이다.

다른 하나는 호환 비용이다. 글로벌 AI 생태계는 칩끼리, 또 칩과 소프트웨어가 서로 맞물리는 공통 규격 위에서 돌아간다. 국산 NPU·독자 모델 묶음이 이 규격에 맞지 않으면, 성능이 아무리 좋아도 끼워 맞추는 데 따로 비용이 든다. 이 비용을 되찾을 길은 규격이 만들어지는 자리에 들어가는 것이다. AAIF·CXL·UCIe 같은 차세대 연결·칩렛 표준 논의에서 발언권을 확보하면, 국산 묶음이 처음부터 표준에 맞도록 설계할 수 있다. 성능 차이는 학습곡선이 오를 때까지 계속 남지만, 호환 비용은 표준 발언권으로 되찾을 수 있다.

T4(통화 주권). Project Hangang 2단계와 BIS Agorá를 통합한다. 여기서 중요한건 결제망의 발언권이다. 어떤 정산 네트워크든 규칙을 정하는 자리가 곧 힘이다. 대가는 기존 SWIFT·달러 체계에서 한국이 쌓아 온 발언권 일부를 내려놓는 것이다. 토큰화된 국경 간 정산망에 들어가는 일은 달러 중심 레일을 부분적으로 우회하는 것이고, 그만큼 옛 체계에서의 위치를 양보하게 된다. 하지만 통화 주권은 낡은 레일의 자리를 지키는 데 있지 않고, 새 레일에서 규칙 제정의 자리를 얻는 데 있다. BIS Agorá는 한국에 ‘부분 룰메이커’의 위치를 줄 수 있고, 회수의 핵심은 그 지위를 모델·에이전트 표준 협상에도 연결하는 것이다.

T5(시민 주권). AI에이전트가 한국 시민의 기억을 어떻게 다루는지 검증하는 기억주권감사 기준을 두고, 이 감사를 50~80% 통과해야 한국 시장에 들어올 수 있게 한다. 대가는 혁신 지체 우려지만, 철수 위협은 두가지 힘으로 뒤집힌다. 첫째, 한국 시장은 통째로 포기하기엔 너무 크다. 둘째, 에이전트 시대에는 락인이 힘들다. 다른 서비스로 쉽게 옮길 수 있다는 애기다. 때문에, 한국 사용자의 기억을 많이 쌓아 둔 사업자일수록 그것을 버리고 떠나기 어렵다. 거기에, 절대우위를 가지고 있는 AI 메모리 반도체는 한국 시장을 버리기 힘들게 만든다.

정리해보면, 시민 주권은 T2와 T5가, 산업 주권은 T1과 T3이, 통화 주권은 T4가 지킨다. 희생해야할 비용도 대부분 동맹 내 위치나 협상 구도를 조정하는 형태로 치러질 뿐, 국내 역량을 완전히 직접 훼손해 치르는 항목은 거의 없다. 그리고 다섯 모두 회수가 작동하는 구조를 명시한다. 제약의 자산화(T1), 양보 풀의 상한 인하(T2), 수요 기반과 표준 발언권(T3), 발언권의 영역 간 전이(T4), 기억의 이동 가능성으로 인질·철수 위협을 무력화(T5). T3만이 회수가 비용에 미치지 못하긴 하지만, 어느 희생도 회수 대안이 존재한다.

V. 5축 도메인 분석

본고는 한국을 분석 단위로 두고, 경제, 안보, 산업, 사회, 기술 5가지 도메인에 대하여 다음과 같은 질문을 중심으로 두고 분석 했다.

표 4 5가지 도메인별 질문

도메인	질문
경제	결제 경로를 누가 결정하는가? — 시장 점유율이 아니라 토큰 국적 attribution
안보	에이전트 행위자를 누가 인증·로깅·감사하는가? — 롤메이커 마지막 원도
산업	추론 단계의 컨트를 지점을 누가 쥐는가? — K-메모리·NPU·전력
사회	인적자본 시차 압축의 비용을 누가 부담하는가? — 신입·신뢰·교육·인구
기술	프런티어 위에 한국이 까는 표준·평가·오케스트레이션 레이어는 어디인가? — K-AISI·NAICC

1. 경제 - 사용 상위인데 SaaS 점유율은 0.35%

한국은 인구당 생성형 AI 도입률 글로벌 18위, 근로자 활용률 63.5%로 OECD 상위 사용국이지만, 국산 SaaS의 세계 점유율은 IDC 2025 추산 기준 2.55조 원으로 글로벌 723조 원의 0.35%에 그친다.³⁰⁾ 시장 규모가 아니라 결제가 어디로 흐르느냐, 곧 경로의 문제다.

그리고 이미 적자로 커지고 있다. 2025년 지식서비스 무역수지 적자는 -102.5억 달러로 전년 대비 약 39% 악화됐고, 한국은행은 그 원인 가운데 'AI 구독'을 정부 통계 사상 처음으로 명시했다.³¹⁾ 게다가 -102.5억 달러는 2026년 들어 에이전트가 본격적으로 토큰을 태우기 전의 숫자다. 단일 호출에서 긴 맥락을 거쳐 멀티 에이전트가 24시간 자율로 도는 하네스(harness) 단계로 넘어가면, 토큰 소비는 기하급수로 늘어난다. 골드막삭스는 2030년 토큰 소비를 24배로, 오픈라우터 실측은 코딩 에이전트의 소비를 14개월 만에 75배로 보고했다. 그 청구서는 전액 달러로 발행되어, 미국 빅테크로 흘러든다. 사상 최대인 1,230억 달러의 경상수지 흑자도 수천억 달러로 불어날 AI 사용료 앞에서는 적자로 뒤집힐 수 있다.

더 우려되는 점은 이 적자에 시장의 자기조정 장치가 듣지 않는다는 사실이다. 가격 탄력성이 0에 가까운 AI 사용료는 환율이 높아지더라도 끊을 수 없다. 그래서 환율은 오히려 증폭 장치가 된다. AI 사용료가 지식서비스수지 적자를 키우고, 그 적자가 경상수지를 적자로 돌리며, 원화가 약해지면 같은 토큰의 원화 비용이 올라 다시 적자를 키우는 악순환 고리다.

결국 디지털세가 중요하다. 국내 매출이 과세 없이 모조리 달러로 해외로 넘어가게 해서 안된다. 충분한 과세가 필요하다. 「조세특례제한법」 글로벌최저한세 도입으로 한국 적격 국내 최저한세(QDMTT)가 2026년 첫 신고를 받고, 국세청은 245개 다국적기업을 2026년 6월 30일까지 처리한다.³²⁾ 금융위원회 「금융분야 AI 가이드라인」도 7대 원칙으로 2026년 1분기 시행됐다.³³⁾ 다만 이 모든 처방에는 측정이란 전제조건이 있다. 한국은행이 AI 도입이 한국 GDP를 4.2~12.6% 끌어올릴 잠재력이 있다고 봤지만,³⁴⁾ 그 폭 자체가 정책 설계의 불확실성이다. 현행 GDP는 인간이 만든 재화·서비스를 재도록 설계돼 있어, 에이전트가 자율적으로 생산하는 가치, 코드·분석·의사결정처럼 토큰 단위로 거래되는 지능 노동을 대부분 놓친다. GIP(Gross Intelligent Product)는 바로 그 빠진 부분, 곧 지능이 생산하는 부

30) IDC, 2025 SaaS 시장 추산. 한국 약 2.55조 원 / 글로벌 약 723조 원(0.35%).

31) 한국은행, 「2025년 상반기 지식서비스 무역통계(잠정) 보도자료」(2025-09-17).

32) 국세청, “글로벌최저한세 안내”(Pillar 2, 2024). 한국 245개 다국적기업, QDMTT 첫 신고 마감 2026-06-30.

33) 금융위원회, “금융분야 AI 가이드라인 개정방향”(2025-12-22). 7대 원칙, 2026 1분기 시행.

34) 한국은행, BOK 이슈노트 제2025-2호 「AI와 한국경제」(2025-02-10).

를 포착하기 위한 측정 틀이다. GDP가 한 나라가 일정 기간 생산한 재화·서비스의 시장가치 총합이라면, GIP는 같은 기간 AI·에이전트가 만들어 낸 지능 산출, 추론, 분석, 코드, 자동화된 의사결정의 가치를 합산해 ‘지능이 생산하는 부’를 국부(國富) 통계로 잡아내는 지표다.

물론, 처음부터 완전히 새로운 국가 회계로 세우기는 어렵다. 그래서 현실적 출발점은 위성계정(satellite account)일 것이다. 이미 운영 중인 관광 위성계정이나 디지털경제 위성계정처럼, 기존 GDP 통계를 건드리지 않으면서 그 옆에 특정 영역을 따로 계량해 붙이는 보조 계정부터라도 시작해야 한다. 측정 도구가 없으면 효과 추정이 흔들리고 근거가 무너지면 산업·통상 협상의 카운터파트가 모두 빠진다.

2. 안보 – 첫 자율 사이버 첩보가 등장한 지점

2025년 11월 13일 앤트로픽은 자사 모델을 도구로 활용한 첫 AI 자율 사이버 첩보사례를 공개했다. 중국이 배경으로 추정되는 사이버 위협 그룹(코드명 GTG-1002)이 30개 표적을 노린 캠페인의 80~90%를 인간 개입 없이 자율 수행했다는 평가였다.³⁵⁾ 일부 사례로 일반화하는 데는 신중해야 하지만, 그래도 위협 주체가 바뀐 사건 – 비인간 행위자, 즉 AI가 안보 게임판에 처음 입장한 날이라는 점은 분명하다.

같은 해 한국인터넷진흥원(KISA)은 실질적 침해사고 2,383건으로 전년 대비 26.3% 증가, 통계 작성 이래 최고를 기록했다.³⁶⁾ 영국 AI 안전연구소(UK AISI)의 「프론티어 AI 트렌드 보고서 2025(Frontier AI Trends Report 2025)」는 자율형 에이전트가 초래할 안보적 모순을 보여줬다. AI의 자율적 해킹 성공률은 9%에서 50%로 폭증한 반면, 생화학 무기 설계의 우회 난이도는 가드레일 강화를 통해 6개월 사이 약 40배나 높아졌다. 그러나 시스템 전체의 고삐를 풀어버리는 ‘유니버설 탈옥(Universal Jailbreak)’ 취약점은 평가 대상이 된 모든 프론티어 시스템에서 예외 없이 발견되었다.³⁷⁾ 아무리 생물학적 무기 차단 필터를 40배 강화했더라도, 마스터키 역할을 하는 탈옥 공격 한 번이면 그 장벽이 통째로 무력화된다. 에이전트의 지능의 가속화를 담아낼 인프라 보안은 여전히 치명적인 허점을 드러낸다.

2026년 4월 앤트로픽은 자사 최강 프론티어 모델 클로드 미토스(Claude Mythos)를 일반 공개하지 않고, 방어 목적에 한해 신뢰할 수 있는 소수 기관에만 게이팅해 제공하는 프로젝트 글래스윙(Project Glasswing)을 출범시켰다. 모델이 너무 강력해 공개 출시 자체가 위험하다

35) Anthropic, “Disrupting AI-orchestrated cyber espionage”(2025-11-13)

36) KISA·과학기술정보통신부, 「2025 사이버 위협 보고서」(2026-01-27).

37) UK AI Security Institute, *Frontier AI Trends Report 2025*(2025-12-18)

는 판단에서였다. 주목할 것은 그 거버넌스 구조다. 창립 파트너 12곳으로 채워졌고, 이후 15개국 이상 150여 기관으로 확대되며 각 기관은 엔트로픽의 보안 심사를 통과해야 접근권을 얻는다. 최첨단 방어 역량이 ‘신뢰 연합’의 심사를 거쳐 분배되는 구조라면, 한국이 속해 있는가 곧 안보 질문이 된다. 여기에, 6월 12일에 미국 행정부가 ‘미토스’를 비롯한 최상위 인공지능 모델에 외국인 접속을 전면금지한 사실은 우리가 처한 현실을 적나라하게 보여준다.

세가지 처방이 있다. 2026년 5월 1일 본격적으로 시행된 「국가 사이버보안 기본지침」(N2SF — 망분리 일률 규정을 폐지하고 데이터 등급 기반 자동화 검증으로 전환한 한국 사이버보안 체계 18년 만의 교체) 운영 단계에 agentic AI 행위자 카테고리를 추가해서 2.0으로 업그레이드해야 한다. 현재의 N2SF 1.0은 이 부분이 전혀 반영되어 있지 않다.

또한, REAIM 후속 회의에서 ‘AI agent의 군사 적용 가이드라인’을 한국이 주도해 발의하도록 해야 한다. 2025년 10월 29일 체결된 한·미 기술번영협정(TPD)³⁸⁾은 신뢰할 수 있는 AI 기술 스택(Trusted AI Tech Stack)의 수출 협력과 양국 AI 안전연구소(AISI) 간 측정 표준 상호인정을 명시한 핵심 양자 협정이다. 이 협정이 명시한 180일 상호인정 가이드라인 수립 시한이 2026년 4월 27일부로 도래하여 양자 간 기술 표준의 기틀이 안착됨에 따라, 외교부는 이를 발판 삼아 즉시 다자화 협상으로 전선을 확장해야 한다.

가치 표준의 문제도 안보 축으로 들어온다. 미국은 자국 AI 스택을 수출하면서 그 안에 자국의 가치 기준까지 함께 실어 보내게 된다. 그래서 한국이 미국 스택을 들여오는 순간, 평등이나 표현의 자유처럼 한국 헌법이 지키려는 가치가 남이 정한 기준에 밀려 조용히 지워질 수 있다는 뜻이다. 그렇다면 한국이 동맹 안에서 자국 스택을 따로 두는 일은 어떻게 봐야 할까. 그건 동맹을 거스르는 배신이 아니라, 오히려 동맹 안에 머물면서 시민의 가치 주권을 지키는 일이다. 이를 막으려면, 인공지능 안전연구소(AISI)에 한국 헌법 가치에 맞는 안전 표준을 스스로 만들 권한을 명시적으로 주고, AI기본법 시행령에서 고영향 AI 영향평가를 할 때 평등·표현·인격권 침해 여부를 따지는 항목을 의무로 넣어야 한다.

3. 산업 - 메모리 우위는 NVIDIA 종속을 벗어나야 지속된다

SK하이닉스는 2026년 1분기 매출 52.6조 원, 영업이익 37.6조 원(영업이익률 72%)로 창사 이래 최대치를 경신했다. HBM(High Bandwidth Memory, AI 가속기용 고대역폭 메모리) 매출은 전년 대비 2배 이상 성장했다. 그러나 이러한 우위는 아직은 엔비디아 GPU 중

38) White House, “U.S.-Korea Technology Prosperity Deal”(2025-10-29). 180일 종료 조항.

속 안에서만 작동한다. 엔비디아가 자체 메모리 컨트롤러와 HBM 합성 표준(NVLink-Fusion·SOCAMM)을 통합하면 한국 메모리 산업의 가격 협상력은 비대칭적으로 약해진다. 가격 통제권이 사라지면 K-메모리 락인이 곧 엔비디아 락인이 된다.

Stanford HAI AI Index 2025는 추론 비용이 GPT-3.5급 성능 기준 2년 사이(2022. 11~2024. 10) 약 280배 떨어졌다고 추정했다.³⁹⁾ 산업전략의 중심을 추론으로 옮기라는 강한 신호다. 한국이 잡을 수 있는 추론 단계 컨트롤 지점은 K-메모리와 양손잡이 NPU 두 축이다. 2024년 12월 사피온이 리벨리온과 합병해 한국 NPU는 리벨리온과 퓨리오사AI 2사 체제다. 리벨리온 ATOM은 NVIDIA AI Summit Korea에서 엔비디아의 월드 모델 코스모스를 자사 NPU로 시연했고, 퓨리오사 레니게이드는 데이터센터 추론용 칩으로 양산에 들어갔다. 한국으로서는 두 트랙 전략이 중요하다. (a) N2SF·국방·공공 자립 트랙에서 국산 NPU 공공조달 비중을 2년 안에 50% 이상으로 끌어올린다. (b) 민간수요·표준 호환 트랙에서 AAIF 워킹 그룹 참여를 통해 K-메모리·CXL·UCIe 표준의 글로벌 호환을 확보한다. 분리된 자립과 연결된 동맹의 곱셈이 양손잡이의 의미를 만드는 것이다.

4. 사회 — 초고령·초적용·신입 절벽의 동시 작동

한국 AI 노출 직무는 신입 절벽과 경력 강화가 동시에 진행되는 양극화에 들어갔다. SW 개발자 직군의 22~25세 고용이 2022년 말 정점 대비 2025년 7월까지 약 20% 줄었다.⁴⁰⁾ 미국에서는 같은 연령대가 13% 줄고, 30세 이상 경력자는 6~12% 늘었다. 다만 한국의 경우 청년 고용률 자체가 43.3%로 1년 전 대비 1.0%p포인트 빠진 흐름과 '쉬었음' 청년이 71.7만 명이라는 통계청 기록이 공존하는 만큼, AI 노출이 곧바로 청년 개발자 직군이 20% 줄었다고 단일 인과로 단정하기보다 경기·인구 교란 요인까지 고민해 봐야 한다.⁴¹⁾

이러한 양극화는 한국 고유의 인구 구조와 결합하면서 더 위험해진다. 통계청에 따르면 65세 이상 인구는 20.3%로 한국은 초고령사회에 진입했다.⁴²⁾ 여기에 63.5%의 AI 도입률은 이 사회 구조 위에 작동하게 된다. 줄어드는 생산인구를 대체할 기회가 될수도, 일자리를 대체하여 부양비가 더 늘어나는 위기가 될수도 있다. 초고령·초적용·신입 절벽이 동시에

39) Stanford HAI, *AI Index 2025*, 추론 비용은 GPT-3.5급 성능 기준 2년 사이(2022. 11~2024. 10) 약 280배 하락(AI Index 2025판 인용 수치).

40) 국회예산정책처, op. cit.(fn. 43). 통계청 경제활동인구조사 기반. 22~25세 SW 개발자 -20%(2022년 말 → 2025-07).

41) 통계청, 「청년 고용동향」(2026-03). 청년 고용률 43.3%(-1.0%p), '쉬었음' 청년 71.7만 명(2025 연간).

42) 통계청, 「2025 고령자 통계」(2025-09). 65세 이상 20.3%.

작동하는 사례다. 외산 기술에 대한 수용도가 높을수록 해외 의존성은 빠르게 고착화되고, 여기에 인구 절벽까지 결합되면, 기술을 보완하고 발전시킬 인적 자본의 공급과 유지라는 두 축이 함께 무너질 위험이 있다.

행정 측면의 고민도 중요하다. 한국은 UN E-Government Survey 2024에서 4위, OECD Government at a Glance 2025의 Digital Government Index에서 1위를 기록했다.⁴³⁾ 하지만, 클라우드 인프라 부족, 17,090개로 분절된 공공 정보시스템의 문제, 여전히 SI 중심의 예산조달체계 등은 빠르게 변화하는 AI 에이전트 시대에 언제든 그 순위가 하락해도 이상하지 않다. 따라서, 행정 측면에서도 「AI 매개 행정행위」를 「행정기본법」에 신설하는 일이 매우 중요한 사회·행정의 결합 처방이다. 고용노동부의 「AI+역량 Up」 프로젝트와 KDT AI Campus 1만 명 양성 계획을 신입 진입과 재진입 두 트랙으로 분리하고,⁴⁴⁾ 공공 민간 신입 채용 의무비율도 검토할만 하다. ILO는 2026년 6월 ILC에서 플랫폼 노동에 관한 협약과 권고를 채택하는데, 노동자가 자신을 평가·배치·해고하는 알고리즘 정보에 접근할 권리가 담은 만큼, 한국은 하반기에 한국형 알고리즘 관리 가이드라인으로 이를 반영해야 한다.

그러나 사회 축의 진짜 승부는 평가와 교육의 한 단계 위, 데이터 층위에서 갈린다. 에이전트 시대의 코퍼스는 더 이상 웹에 흩어진 인간 중심 텍스트가 아니라, 에이전트가 한국의 맥락에서 실제로 일하고 판단한 흔적, 곧 에이전트 센트릭 데이터(agent-centric data)다. 이를 의도적으로 모아 한국어 에이전트의 학습·정렬 코퍼스로 자산화하는 일이야말로, 외산 에이전트가 한국인의 판단을 매개할 때 그 기준선을 우리가 통제할 수 있는 유일한 방법이다.

5. 기술 — 모델은 세계 3위, 메모리에서 공급자를 넘어 설계자로

Stanford HAI AI Index 2026에 따르면 2024년 AI 특허 출원에서 중국이 절대 다수(보고에 따라 71~74%)이고 미국이 13~15%, 한국이 5%, 일본이 3%이며, 한편 한국은 인구 만명당 AI 특허 출원 에서 글로벌 1위를 기록했다.⁴⁵⁾ 인구당 H-1B 비자 한국 STEM 박사 유출은 4,500명에 이르고 한국 AI 논문 점유율은 3.4%에 머문다.⁴⁶⁾ 2025년 notable AI 모델 산출에서 한국은 8건으로 미국(50)·중국(30)에 이은 세계 3위를 차지했다. 현재 시점에

43) UN DESA, *E-Government Survey 2024*(2024-09-19). 한국 EGDI 0.9679(4위). OECD, *Government at a Glance 2025 / Digital Government Index*(2025-06-19). 한국 DGI 0.95(1위).

44) 고용노동부, 「AI+역량 Up 프로젝트」(2025)

45) Stanford HAI, *AI Index 2026*, AI 특허 출원 비중(2024).

46) USCIS, *H-1B Employer Data Hub*(2025-01-31)

서 격차는 역량의 부재가 아니라 민간 투자·인재 양성에 있다.

한국의 진짜 레버리지는 모델 산출 경쟁이 아니라 기억의 설계에 있다. AI 패권의 축은 연산(compute)에서 기억으로 옮겨가면서, 지능은 ‘얼마나 빨리 계산하는가’가 아니라 ‘얼마나 방대한 맥락을 저장하고 즉시 꺼내 쓰는가’로 갈린다. 기억은 인간의 뇌처럼 계층을 이룬다. 단기 기억은 칩 내부의 SRAM, 중기는 HBM, 장기는 HBF 이다. 한국은 HBM(속도)과 HBF(용량)라는 두 날개를 모두 쥔 거의 유일한 나라로, 부품 공급자에서, AI 로드맵을 설계 단계부터 함께 그리는 공동 설계자(Co-Architect)로 올라설 수 있다. 모델도 마찬가지다. 현재의 거대언어모델(LLM)이 추론은 빠르되 기억이 짧은 뇌라면, 다음 국면의 열쇠는 모든 맥락과 수십 년 치 데이터를 영구 저장·활용하는 장기기억 모델(LMM)이며, 그 다음은 에이전트들이 주체로서 관계를 맺는 거대 사회적 모델(LSM)이다. 그 단계에서, 기억의 설계권을 쥔 나라가 새로운 경제 질서의 표준에 가장 가까이 설 수 있다.

글로벌 AI 경쟁은 단순한 모델 성능 싸움을 넘어 ‘생태계 패권 경쟁’으로 판이 커지고 있다. 우리는 LLM 등 원천 기술에서 미·중과의 격차를 좁히는 동시에, 세계 최고 수준인 메모리 반도체 역량을 핵심 레버리지로 삼아야 한다. 모든 풀스택에서 리더가 될 수 없다면, 특정 스택에서 압도적인 ‘락인(Lock-in) 효과’를 만들어내야 하며 그 선두에 바로 메모리다. 여기에 반도체, 조선, 자동차 등 세계 초일류 제조업 데이터를 결합한 산업 특화 모델과 피지컬 AI(Physical AI)를 융합하면, 전략적 요충지에서는 미·중과 대등하게 겨룰 수 있다.

VI. 어떻게 준비할 것인가

1. 연결된 자율성 – 곱셈의 정의

미·중 패권 경쟁에서 한국은 두 함정 사이에 놓여 있다. 한쪽은 문을 닫아서는 자립이라는 환상이고, 다른 한쪽은 무조건적 종속이라는 현실이다. 두 길 모두 결국 국부의 유출로 귀결된다. 폐쇄는 기술 격차를 키우고, 종속은 부와 정체성을 함께 잃게 만든다.

본고에서 제시하는 제3의 길이 「연결된 자율성(Connected Autonomy)」이다. 핵심은 자율성과 연결성을 양자택일이 아니라 ‘곱셈’으로 다룬다는 데 있다. 덧셈이라면 한쪽이 약해도 다른 쪽으로 메울 수 있지만, 곱셈은 어느 한쪽이 0이면 전체가 0이 된다. 자율성만 키우고 세계와 끊기면 고립이고, 연결성만 추구하고 자국 기반이 없으면 종속이다. 둘 다 동시에 살아 있어야

비로소 전략이 선다. 한국의 자율성은 통화·결제·기억의 세 축에서, 연결성은 MCP·A2A·EU AI Act·한·미 Technology Prosperity Deal의 네 축에서 작동한다. 어떤 비중으로 결합할지는 후속 과제로 남기고, 여기서는 어느 하나도 0이어서는 안 된다는 방향과 조건만 명제로 둔다.

시작은 ‘모델을 다루는 방식’이다. 한국은 빅테크의 프런티어 모델을 정면으로 추월하려 애쓰는 대신, 독자 파운데이션 모델과 빅테크 모델을 목적과 용도에 따라 골라 연결해 쓴다. 민감하거나 한국 맥락이 중요한 일은 우리 모델이 맡고, 범용 고난도 작업은 빅테크 모델에 맡기되, 둘을 하나의 작업 흐름으로 엮는다. 이 연결을 가능케 하는 표준이 곧 연결성의 두 축인 MCP와 A2A다. 이 연결 계층을 한국이 직접 설계·운용하면, 개별 모델의 성능 격차는 연결의 설계로 좁혀진다. EU는 규제 표준을 만들어 세계로 퍼뜨리는 힘(브뤼셀 효과, Brussels-effect)으로 영향력을 키웠고, 미국은 자국 AI 기술 스택을 세계에 채워 넣는 힘(스택 효과, Stack-effect)으로 판을 장악하고 있다. 한국이 잡아야 할 자리는 그 둘과 다른 세 번째 힘, 곧 주권적 안전장치 효과(Sovereign-safeguard-effect) 라고 할 수 있다.

이 지점에서, 오늘날 세계에서 서로 갈라져 다투는 세 흐름이 하나로 모인다. 앞에서 언급했던 안전, 가속 그리고 공개 가중치(Yann LeCun의 오픈소스 LLM 우위론)다. 바깥에서는 이 셋이 대립하지만, 한국의 주권적 안전장치 모델 안에서는 양자택일이 아니라 하나로 결합된다. 본 고가 가리키는 지점이 바로 거기이며, 이것이 가장 중요한 명제다.⁴⁷⁾ 국가전략으로 얘기한 김대중 정부의 정보화 인프라와 노무현 정부의 균형자 외교가, 이 지점에서 비로소 하나의 독트린으로 묶이는 것이다.

2. 컨트롤러 에이전트

「연결된 자율성」을 실제로 구현하는 메커니즘이 컨트롤러 에이전트(Controller Agent)다. 외국의 고성능 모델은 두뇌, 곧 도구로 빌려 쓰되, 그 도구를 지휘하고 데이터 흐름을 결정하는 오케스트레이터(Orchestrator)는 한국이 쥔다는 발상이다. 통제점은 크게 4가지다. 누가 MCP 서버를 운영하는가(운영자), A2A 핸드셰이크에서 누가 신원과 권한·로깅을 통제하는가(신원·권한), 그리고 에이전트의 기억을 어느 관할권에 저장하는가(기억 관할). 이 4가지 통제점을 함께 쥐어야, 전선을 ‘모델 주권’에서 ‘에이전트 주권’으로 옮길 수 있다.

젠슨 황의 ‘AI 팩토리가 토큰을 생산한다’는 말은, 그 공장의 컨트롤러를 쥔 자가 가치의 종착지를 정한다는 뜻이다.⁴⁸⁾ MCP Authorization과 A2A 프로토콜이 AAIF 산하로 묶이면

47) International AI Safety Report 2026, op. cit.; Marc Andreessen, “The Techno-Optimist Manifesto”(a16z, 2023-10-16).

서 ‘표준을 가진자가 권력을 가진다’는 공식이 굳어졌다.⁴⁹⁾ 그래서 3중 매트릭스가 모두 이 컨트롤러 에이전트로 수렴한다. 주권 인프라(K-메모리·NPU·전력)는 컨트롤러가 딛고 설 바닥이고, 동맹 스택은 그 권한의 국제적 근거이며, 공개 가중치 조정자(Llama·DeepSeek·Mistral 위에 한국어 SLM을 얹는 층)는 컨트롤러가 모델을 골라 끼우는 손이다.

3. 다섯가지 댐 – Trade-off 1:1 매핑

「연결된 자율성」을 단순한 구호에서 실행으로 바꾸기 위한 당면 과제는 지금까지 얘기한 것을 종합해보면 다섯 개의 교차 레버(Cross-cutting Levers)로 정리된다. 바로 측정(L1)·규율(L2)·투자(L3)·동맹(L4)·주권(L5)이다. 나아가 각 레버는 국가 전략이 필연적으로 치러야 할 다섯 가지 대가, 즉 Trade-off와 정확히 맞물린다. 측정(L1)이 다섯 가지 Trade-off 전체를 지탱하는 거대한 주춧돌이라면, 나머지 네 레버는 각자의 대가에 일대일로 대응한다. 즉, 규율은 T2, 투자는 T1·T3, 동맹은 T4, 주권은 T5에 결착하는 구조다.

표 5 연결된 자율성 실현을 위한 분야별 5개 교차 레버

	L1 측정	L2 규율	L3 투자	L4 동맹	L5 주권
경제	■ GIP·attribution	▣ 디지털세	▣ 토큰 attribution	▣ BIS Agorá	□
안보	□	■ N2SF·기본지침	□	■ TPD·REAIM	▣ N2SF 자동화
산업	□	□	■ K-메모리·NPU·전력·독파모	▣ AAIF·CXL·UCle	□
사회	▣ 신뢰·신입 통계	■ 행정기본법 신설	□	□	■ Student/Senior/Worker
기술	▣ KPI 근거	■ Law as Code	■ NPU·NAICC	■ AAIF·K-AISI	▣ Agent Portability

주: ■ 강 ▣ 중 □ 약

L1. 측정 – 국가 전략의 주춧돌. 가장 먼저 축조해야 할 방파제다. 무엇을 어떻게 계측할지 모르면 나머지 네 레버의 성패를 가늠할 수 없다. GIP(Gross Intelligent Product) 위 성계정을 신설해 AI가 창출하는 부를 분기별로 측정하고, 이를 「국가통계법」 시행령 등에 명

48) Huang, GTC 2025·2026; Thompson, op. cit.

49) MCP Specification – Authorization(2025-03-26), A2A Protocol launch(Google for Developers, 2025-04-09).

문화해야 한다. 측정은 특정 Trade-off에 국한되지 않고 전체를 떠받치는 기준점이 된다

L2. 자율 - 빅테크의 에이전트가 한국 시민의 정체성과 권리를 소리없이 잠식하지 못하도록 막아서는 댐이다. 'AI기본법 시행령', 행정기본법 내 'AI 매개 행정행위' 조항 신설, N2SF의 자동화 검증, 그리고 법령을 기계가 지체없이 정확하게 판독할 수 있도록 하는 'Law as Code'가 하나의 패키지로 구동되어야 한다. 시민 주권을 사수하기 위한 대가, 곧 T2와 연결된다.

L3. 투자 - 자율성을 키우는 물리적 댐이다. 독보적인 우위를 가진 K-메모리와 NPU, 그리고 이를 구동할 에너지 인프라에 자원을 집중 투입한다. 예산의 패러다임을 추론 중심으로 대전환하고, 국산 NPU의 공공조달 비중을 50% 이상으로 전격 끌어올린다. 필수 전력은 「AI 데이터센터 산업 진흥 특별법」과 SMR(소형모듈원전) 패키지를 통해 확보한다. 데이터센터 분산(T1)과 국산 NPU 생태계 조성(T3)이라는 두 가지 대가를 동시에 짊어지는 레버다.

L4. 동맹 - 독자적으로 수립할 수 없는 글로벌 표준과 결제 규칙을, 패권국들이 독점하기 전에 주도적으로 설계하는 댐이다. AEIF 워킹그룹, 한·미 TPD, BIS 아고라(Agorá)에 깊숙이 참여하는 동시에, 외교부 내에 'AI 공급망 외교 컨트롤타워'를 구축하고 TPD를 다자화(한·미·일·호주 및 한·EU·영국)한다. 통화와 결제라는 글로벌 핵심 테이블의 좌석을 확보하는 대가(T4)에 연결되며, 즉시 실행에 착수해야 할 시급한 레버다.

L5. 주권 - 시민에게 자기 기억의 주권을 돌려주는 댐이다. '에이전트 주권 기본법'을 통해 시민이 자신의 디지털 자아, 즉 '기억'을 휴대·소유·이전할 권리를 명시하고, 빅테크 기업이 국내 관할권 내에 데이터를 미러링하도록 의무화한다. 아울러 학생·고령자·노동자를 위한 에이전트 권리 3종을 제도화한다. 기억 주권을 사수하는 대가, 즉 T5에 연결된다.

다섯 레버가 동시에 톱니바퀴처럼 맞물려 돌아갈 때에만 「연결된 자율성」은 비로소 슬로건을 넘어 실행의 단계로 진입한다. 본질은 각 레버가 적시된 희생(Trade-off) 위에 서 있다는 점이다. 우리가 어떤 대가를 내어주고 있으며, 그것을 어떻게 회복할 것인가까지 정교하게 설계될 때, 이 다섯 레버는 '하나의 유기적인 국가 전략'으로 완성된다.

4. 결론 — 댐을 어디에 세울 것인가

구조적 인구 감소기에 접어든 한국은 지능형 레버리지를 통해 이 위기를 가치 창출의 자

산으로 전환할 수 있다. 이 전략적 전환의 핵심 전제는 지능의 컨트롤러를 한국이 내재화하는 것이다. AI 주권은 폐쇄적인 장벽을 구축하는 것이 아니라, 생태계의 개방과 폐쇄를 결정하는 통제권의 확보를 의미한다. 즉, 자율성과 연결성은 양자택일이 아닌 곱셈의 관계이며, 어느 한 축이라도 상실되면 총합은 0이 된다.

이러한 패러다임 시프트는 경제 활동의 주체적 기반을 근본적으로 변화시킨다. 산업 자본주의의 개인 경제 체제에서 인간은 임금과 계약의 단일 주체였고, 디지털 경제의 연결 경제 체제에서는 플랫폼을 통해 가치를 매개하는 중심축이었다. 반면, AI 에이전트가 중심이 되는 혼합 경제 체제에서는 인간과 에이전트가 거래와 계약, 그리고 책임을 공유하는 다원적 주체로 공존한다. 애덤 스미스가 250년 전 가정한 ‘단일 인간 주체’ 중심의 경제학적 전제가 수명을 다했음을 시사하며, 본고는 바로 이 지점에서 논의를 전개하고 있다. 결과적으로 한국 국민의 에이전트 소유권과 거래 통제권의 귀속 문제에 대한 해답은 다섯 가지 제도적 ‘댐’과 컨트롤러 에이전트의 ‘네 가지 통제점’을 구체화하는 데 있다.

본 고에서 제시한 진단과 대안은 다음과 같은 다섯 가지 연구 한계 내에서 유효하다.

첫째, 분석의 토대가 된 데이터는 거시 환경에 따라 갱신 주기가 짧으므로, 본문 내 시점 표기를 엄격히 제한하여 분석의 시효성을 유지하려고 했다.

둘째, 분석의 상당 부분은 계량적 인과 추정이 아닌 패턴 매칭 기반의 분석에 의존했다.

셋째, GIP, 인지, 기억, 에이전트 주권 등 신개념은 분석 프레임워크를 제안하는 탐색적 단계다.

넷째, 본 연구의 시사점은 한국적 맥락에 강하게 결착되어 있어, 다른 중견국으로의 일반화 가능성은 제한적이다.

마지막으로, 질적 연구가 의존한 AI 인사들의 논의가 주로 영어권 생태계에 치우쳐 있다.

결론적으로 본고가 던지는 최종적 화두는 에이전트 시대가 도래하면서 엄청나게 커지는 토큰의 흐름을 통제할 댐을 어디에 세울 것인가. 측정, 규율, 투자, 동맹, 시민 주권의 다섯 가지 댐이 거시적으로 맞물려 작동할 때 비로소 국내 산업 생태계 전반으로 선순환하면서 댐을 만들 수 있다. 단순히 글로벌 프론티어 모델을 수입하여 소비하는 시대는 지나갔다. 지금은 지능의 컨트롤러를 확보하여 구조적 주권을 쥐어야 할 때다.

참고문헌

국제기구·외국 정부

- BIS.(2026). Project Agorá.
- BIS.(2025). Annual Economic Report 2025, Chapter III.
- Aldasoro, I., & Desai, A.(2025). AI agents for cash management in payment systems. BIS Working Paper 1310.
- CISA, NSA, FBI, UK NCSC, 호주·캐나다·독일·네덜란드·뉴질랜드.(2025). Principles for Secure Integration of AI in OT.
- European Commission.(2025). InvestAI.
- European Commission.(2026). EU AI Act, Council Regulation; AI Omnibus VII.
- IEA.(2025). Energy and AI.
- ILO.(2025). Generative AI and Jobs: A Refined Global Index of Occupational Exposure. ILO Working Paper 140.
- IMF.(2026). World Economic Outlook Update.
- Davidovic, S., & Tourpe, H.(2026). How Agentic AI Will Reshape Payments. IMF Note 2026/004.
- Cazzaniga, M., et al.(2024). Gen-AI: Artificial Intelligence and the Future of Work. IMF SDN/2024/001.
- Bengio, Y.(의장).(2026). International AI Safety Report 2026.
- Japan METI.(2026). AI-semiconductor industry base reinforcement framework; AI Gennai
- MeitY(India) / DD News.(2025). IndiaAI Mission BharatGen.
- NIST.(2025). AI 600-1 v1.1 / Autonomous Agent Risk Companion.
- NITI Aayog(India).(2025). AI for Inclusive Societal Development.
- OECD.(2025). Government at a Glance 2025 / Digital Government Index.
- OECD.(2025). Skills Outlook 2025.
- OECD.(2026). Going Digital Measurement Roadmap 2026.
- OECD / G7.(2025). Pillar 2 side-by-side.
- OMB.(2025). Memorandum M-25-21.
- PIF.(2025). HUMAIN launch.
- UAE. National Strategy for AI 2031.
- Singapore IMDA.(2026). Model AI Governance Framework for Agentic AI(rev. 2026-05-20).

GovTech Singapore.(2025). Pair.
 SIPRI.(2025). Yearbook 2025, Chapter 12.
 UK AI Security Institute.(2025). Frontier AI Trends Report 2025.
 UK Government.(2025). UK Compute Roadmap.
 UN CCW GGE on LAWS.(2025). 2025 Chair's Summary.
 UN DESA.(2024). E-Government Survey 2024.
 USCIS.(2025). H-1B Employer Data Hub.
 White House.(2025). America's AI Action Plan; Executive Order 14320.
 White House.(2025). U.S.-Korea Technology Prosperity Deal.

한국 정부·공공기관

과학기술정보통신부.(2023). 디지털 권리장전.
 과학기술정보통신부.(2025). 국가 AI 컴퓨팅 센터 구축계획; 범정부 AI 공통기반.
 과학기술정보통신부.(2024). K-AISI(한국 AI 안전연구소) 출범.
 국가정보원.(2026). 국가 사이버보안 기본지침.
 국세청.(2024). 글로벌최저한세 안내(Pillar 2).
 국회예산정책처.(2025). 생성형 AI 고노출 직업 현황과 최근 청년 고용.
 금융위원회.(2025). 금융분야 AI 가이드라인 개정방향.
 외교부.(2024). REAIM 2024 Outcome(Blueprint for Action).
 통계청.(2026). 청년 고용동향.
 통계청.(2025). 2025 고령자 통계.
 한국은행.(2025). 2025년 상반기 지식서비스 무역통계(잠정) 보도자료.
 한국은행.(2025). AI와 한국경제. BOK 이슈노트 제2025-2호.
 한국은행.(2025). AI 활용 가계조사. BOK 이슈노트 제2025-22호.
 한국은행.(2026). Project Hangang Phase 2.
 KISA·과학기술정보통신부.(2026). 2025 사이버 위협 보고서.
 NIPA.(2025). 독자 AI 파운데이션 모델 사업.

학계·업계·언론

Acemoglu, D.(2024). The Simple Macroeconomics of AI. NBER Working Paper 32487
 (Economic Policy, 2025).
 Allison, G.(2017). Destined for War. Houghton Mifflin Harcourt.

Altman, S.(2025). The Gentle Singularity. blog.samaltman.com; 의회 증언(2026-02).

Amodei, D.(2024). Machines of Loving Grace;(2026). The Adolescence of Technology.

Andreessen, M.(2023). The Techno-Optimist Manifesto; Why AI Will Save the World. a16z.

Anthropic.(2025). Disrupting AI-orchestrated cyber espionage; Project Vend; Economic Index; Claude Sonnet 4.5 release notes.

Bostrom, N.(2014). Superintelligence. Oxford University Press.

Bradford, A.(2023). Digital Empires: The Global Battle to Regulate Technology. Oxford University Press.

Bresnahan, T., & Trajtenberg, M.(1995). General Purpose Technologies. *Journal of Econometrics* 65:1.

Brynjolfsson, E., Chandar, B., & Chen, R.(2025). Canaries in the Coal Mine?. Stanford.

Brynjolfsson, E., Li, D., & Raymond, L.(2025). Generative AI at Work. QJE.

Buchanan, B.(2020). The AI Triad and What It Means for National Security Strategy. CSET.

Clark, J.(2025). Anthropic 블로그(2025-11, 2026-04).

David, P.(1990). The Dynamo and the Computer. *AER* 80:2.

Doshi, R.(2021). The Long Game. Oxford University Press.

Edelman.(2026). Trust Barometer 2026 Global Report(한국 Trust Index 46).

Gartner.(2025). Press Release(2025-06-25).

Google.(2025). A2A Protocol launch;(2024). DeepMind Frontier Safety Framework;(2026). Pichai I/O 2026 Keynote.

Government Information Quarterly, GenAI 행정 특집호(2026-03).

Hassabis, D.(2026). Google I/O 2026 Keynote(2026-05-19~20).

Huang, J.(2025-2026). NVIDIA GTC 2025·2026 Keynote.

IDC.(2025). SaaS 시장 추산.

Karpathy, A.(2026). Sequoia AI Ascent.

LeCun, Y.(2026). Brown University 강연; open-source LLM 옹호 — Meta Llama 3.1 발표 및 MIT Technology Review 인터뷰.

Linux Foundation.(2025). AAIF 출범.

MCP Specification — Authorization(2025-03-26).

METR.(2025). Measuring AI Ability to Complete Long Tasks. [arXiv:2503.14499](https://arxiv.org/abs/2503.14499).

Microsoft.(2026). Nadella, Build 2026 Keynote.

NVIDIA.(2025). NVLink-Fusion·SOCAMM 통합; AI Summit Korea.

OpenAI.(2025). Preparedness Framework v2.

Pacheco Pardo, R.(2023). South Korea's Grand Strategy. Columbia University Press.

Patel, D.(2025). SemiAnalysis.

Pew Research.(2025). How People Around the World View AI.

Russell, S.(2019). Human Compatible. Viking.

Schmidt, E.(SCSP).(2023). Generative AI.

Silove, N.(2018). Beyond the Buzzword. Security Studies 27:1.

SK하이닉스.(2026). DART 공시·IR 보도자료.

Stanford HAI.(2026). AI Index Report 2026.

Suleyman, M.(2023). The Coming Wave. Crown.

Sutskever, I.(2025). Dwarkesh Patel 인터뷰.

Thompson, B.(2026). Stratechery.

Yin, R. K.(2018). Case Study Research and Applications(6th ed.). SAGE.