



피지컬 AI 데이터 인프라의 전략적 기회

페블러스 비즈니스 모델을 중심으로

- 기획: 페블러스 경영지원실
- 감수: 이주행 (페블러스 대표, 공학박사)
- 일자: 2026-02-17

페블러스(Pebblous)는 한국 피지컬 AI 생태계에서 "데이터 OS + 품질 평가 + 시뮬레이션 생성"을 하나로 통합한 유일한 플레이어로 포지셔닝하고 있다. 2021년 ETRI 출신 공동창업자들이 설립한 페블러스는 현대차, LG전자, LGU+, 한화비전, 삼성E&A, 육군·해병대를 고객으로 확보했고, 과기정통부 61억 원 규모의 글로벌 빅테크 육성 과제 주관기관으로 선정되었다. 글로벌 합성데이터 시장이 2025년 약 5~9억 달러에서 2030년 25~34억 달러(CAGR 31~46%)로 급성장하는 가운데, 한국 정부가 2026년 산업부 M.AX(제조 AI 전환)에만 1조 원 이상, 범정부 AI 예산 약 10조 원을 배정한 정책 환경은 페블러스에 구조적 순풍을 제공한다. 다만, 시뮬레이션-현실 간 격차(sim-to-real gap), 뉴로-심볼릭 기술의 성숙도, 정부과제 의존도는 상용화 전환 과정에서 반드시 관리해야 할 핵심 리스크다.

1. 시장은 피지컬 AI 합성데이터를 향해 급격히 수렴하고 있다

글로벌 합성데이터 시장은 복수의 리서치 기관이 연평균 31~46% 성장률을 전망하며, Gartner는 2030년까지 AI 모델 학습에서 합성데이터가 실제 데이터를 완전히 압도할 것이라 예측했다. 이 시장에서 가장 빠르게 성장하는 세그먼트는 자율시스템 시뮬레이션(CAGR 46.3%)과 자동차·운송(CAGR 38.4%)으로, 페블러스가 타겟하는 영역과

정확히 일치한다.

시장	2025년 규모	2030년 전망	CAGR
합성데이터	\$5~9억	\$25~34억	31~46%
피지컬 AI	\$51~54억	\$500~840억 (2033-35)	31~34%
디지털 트윈	\$210~250억	\$1,250~1,500억	35~48%
군사 시뮬레이션·훈련	\$135~151억	\$190~220억	4~5%

피지컬 AI 시장은 2025년 약 **51~54억 달러**에서 2033~2035년 **500~840억 달러**로 성장이 전망되며, 제조·자동차가 **45.2%**의 압도적 점유율을 차지한다. 아시아태평양 지역이 **CAGR 33.5%**로 가장 빠르게 성장하는데, 이는 한국의 강력한 제조업 기반(자동차, 조선, 반도체)과 정부의 공격적 투자가 복합적으로 작용한 결과다.

한국 정부의 피지컬 AI 정책은 전 세계에서 가장 포괄적인 수준이다. 2026년 정부 R&D 예산은 전년 대비 **19.9% 증가한 35.5조 원**이며, 그중 AI 대전환에 5.1조 원, 산업부 M.AX(피지컬AI·AI팩토리 등 산업 AI 확산) 예산 **1조 455억 원**(전년 대비 +52%), 과기정통부 AI R&D **2.3조 원**, 방사청 피지컬 AI·첨단전력 신규과제 약 **1,000억 원** 등 범정부 AI 예산이 **약 10조 원** 규모로 편성되었다. "피지컬 AI 1등 국가" 전략하에 AI 로봇(2030년 글로벌 톱3), AI 선박(2030년 완전자율), AI 차량(2027년 완전자율주행 상용화) 등 **15대 선도 프로젝트**가 추진 중이다. 2025년 10월에는 '한국피지컬AI협회'가 공식 출범하며 산학연 플랫폼이 제도화되었다.

2. 경쟁 환경: 통합 플랫폼의 공백을 노려야 한다

합성데이터 시장은 두 축으로 양분되고 있다. **피지컬 AI/시뮬레이션 기반**(NVIDIA, Applied Intuition, Parallel Domain)과 **프라이버시 중심 정형데이터**(MOSTLY AI, Tonic.ai, 구 Gretel)이다. 2024~2025년의 시장 구조 변화는 극적이다. NVIDIA가 Gretel을 **약 3.2억 달러에 인수**하고, Datagen이 은행에 **200억 원을 남기고 폐업**했으며, Synthesis AI는 직원이 1~10명으로 급감한 뒤 Globant에 흡수되었다.

핵심 경쟁자 포지셔닝:

NVIDIA Omniverse + Cosmos는 가장 강력한 수평 플랫폼이다. OpenUSD 기반 디지털 트윈, 물리 시뮬레이션, Cosmos 월드 파운데이션 모델, Isaac Sim(로보틱스), DRIVE Sim(자율주행)을 통합한다. Cosmos는 출시 이후 **200만 회 이상 다운로드**되었고, BMW, Amazon Robotics, Continental 등이 활용 중이다. 그러나 NVIDIA는 인프라 레벨 플랫폼으로, 최종 사용자가 직접 조립·구축해야 하는 복잡성이 존재한다. **품질 평가와 데이터 거버넌스 기능은 부재하다.**

Applied Intuition은 2025년 6월 **150억 달러 기업가치**(시리즈 F)를 달성했으며, ARR은 약 **4억 달러**, 매출총이익률 **85%**를 기록한다. 자율주행 시뮬레이션·검증에서 출발해 국방(미 육군 RCV 프로그램 4,900만 달러 계약), 건설, 광업으로 확장 중이다. 상위 20대 글로벌 자동차 OEM 중 18곳이 고객이다. 그러나 AV/국방에 깊이 특화되어 있어 범용 데이터 OS로서의 기능은 없다.

Parallel Domain은 자율주행 합성데이터 전문 기업으로 약 4,500만 달러를 유치했으며, NVIDIA Cosmos 생태계의 핵심 파트너로 활동한다. **MOSTLY AI**는 정형 데이터 합성에서 선두로, 오픈소스 SDK + 엔터프라이즈 업셀 모델을 운영하며, AWS Marketplace에서 활용 가능하다.

페블러스의 차별화 포인트는 "통합"에 있다. 현재 시장에서 데이터 관리(OS), 품질 평가(진단·처방), 시뮬레이션 기반 합성데이터 생성을 단일 제품으로 통합한 플레이어는 부재한다. NVIDIA는 인프라이고, Applied Intuition은 AV에 특화되었으며, 정형데이터 기업들은 물리 시뮬레이션이 없다. YData가 데이터 프로파일링 + 합성을 결합하지만 시뮬레이션은 없다. 페블러스의 "Data Greenhouse(데이터 OS) + Data Clinic(품질 평가) + PebbloSim(시뮬레이션 생성)"은 이 **구조적 공백(structural gap)**을 정확히 공략한다.

역량	NVIDIA	Applied Intuition	MOSTLY AI	YData	페블러스 (목표)
데이터 생성	✓	✓	✓	✓	✓

품질 평가	부분적	✓	✓	✓	✓
데이터 관리/OS	✓(Nucleus)	✓	SDK	✓	✓
물리 시뮬레이션	✓	✓(AV)	✗	✗	✓(목표)
규제 대응 패키지	✗	부분적	✗	✗	✓(목표)

그러나 시장 현실을 직시해야 한다. Datagen, Synthesis AI, AI.Reverie 같은 순수 합성데이터 스타트업들이 잇따라 실패한 것은 **단일 모달리티에 국한된 데이터 제공만으로는 지속가능한 비즈니스를 구축하기 어렵다**는 것을 증명한다. 성공한 기업(Applied Intuition, Scale AI, Palantir)은 모두 워크플로우에 깊이 임베딩되어 높은 전환비용을 만들어낸 플랫폼 전략을 취했다.

3. 수익 모델: 플랫폼 임베딩이 생존의 열쇠다

B2B 합성데이터 기업의 5가지 수익 구조는 다음과 같다.

1. **SaaS 구독** — 크레딧·시트 기반 반복 매출 (MOSTLY AI, Gretel AI)
2. **프로젝트 기반 커스텀 계약** — 복잡한 도메인별 데이터셋 구축 (헬스케어·금융)
3. **모듈형 랜드앤익스팬드** — 단일 모듈 진입 후 플랫폼 확장 . Land-and-Expansion (Applied Intuition)
4. **API·마켓플레이스 종량제** — 사용량 기반 과금 및 클라우드 채널 유통 (Synthesis AI, AWS Marketplace)
5. **정부·국방 과제 수탁** — 공공 프로젝트로 기술 검증 후 상용 시장 확장 (Scale AI, Palantir)

SaaS 구독이 가장 보편적이다. MOSTLY AI는 크레딧 시스템(1크레딧 = 100만 데이터포인트)과 무료 티어(일 2크레딧)를 제공하고, 엔터프라이즈 연간 계약은 **5만~50만 달러** 수준이다. Gretel AI(인수 전)는 개발자 무료 티어(월 15크레딧)에 유료 크레딧당 2달러를 부과했다.

프로젝트 기반 커스텀 계약은 복잡한 헬스케어·금융 데이터셋의 경우 **7.5만~50만 달러** 이상이다. Applied Intuition은 **모듈형 랜드앤익스팬드** 전략으로 하나의 모듈(Simian)에서 시작해 Spectral, Basis, 안전 프레임워크로 확장하며, 매출총이익률 ****85%****를 달성한다.

API·마켓플레이스 종량제는 개발자 친화적 채널로, Synthesis AI가 API 호출당 과금 모델을 운영했고, AWS·Azure·GCP 마켓플레이스를 통한 유통은 엔터프라이즈 조달 절차를 우회하여 도입 속도를 높인다. **정부·국방 과제 수탁**은 초기 기술 검증과 레퍼런스 확보에 핵심적이다. Scale AI는 미 국방부 계약으로 데이터 어노테이션 역량을 증명한 뒤 290억 달러 기업가치의 상용 플랫폼으로 성장했고, Palantir는 CIA 자금으로 시작해 17년 만에 상용 시장에서 수익성을 확보했다.

성공 사례의 공통 패턴은 명확하다

Applied Intuition(150억 달러 기업가치), Scale AI(290억 달러), Palantir(2,500억 달러 시가총액)은 모두 (1) 멀티모듈 플랫폼으로 전환비용을 창출하고, (2) 워크플로우에 깊이 임베딩되어 대체 불가능한 위치를 확보했으며, (3) 정부·국방 계약을 기반으로 기술을 검증한 뒤 상용 시장으로 확장했다. 특히 Scale AI의 데이터 플라이휠은 **130억 개 이상의 어노테이션**에서 98% 정확도를 달성하며 경쟁 우위를 만들었다. 이들 기업은 위 5가지 수익 모델 중 **3가지 이상을 동시에 운영**한다는 공통점이 있다.

실패 사례는 경고등이다

Datagen은 7,000만 달러를 유치하고도 GenAI 등장으로 구조적 시뮬레이션 방식이 무력화되자, 피봇에 실패하고 2024년 폐업했다. 은행에 **2,000만 달러가 남아 있었음에도** 생존 가능한 경로를 찾지 못한 것이다. AI.Reverie는 In-Q-Tel 투자와 국방부 계약을 보유했지만 독립 기업으로 스케일하지 못하고 Meta에 acqui-hire되었다. 공통된 실패 원인은 (1) 단일 데이터 모달리티/유스케이스에 국한, (2) 기술 패러다임 전환에 늦은 대응, (3) 반복 매출 없이 데이터를 일회성 상품으로 판매한 것이다.

페블러스의 데이터 플라이휠 전략은 강력한 이론이지만 실행에 주의가 필요하다

a16z는 "데이터 모트의 빈 약속(The Empty Promise of Data Moats)"이라는 분석에서 데이터 모트가 **수확 체감의 법칙**, 데이터 노후화, 경쟁자의 유사 데이터 축적으로 약화될 수 있음을 경고했다. 효과적인 데이터 플라이휠이 작동하려면 (1) 학습의 자동화된 제품화, (2) 사용자 간(across-user) 학습 효과, (3) 복제 불가능한 독점 데이터,

(4) 높은 전환비용을 동시에 달성해야 한다.

1. **학습의 자동화된 제품화**: 수집된 데이터가 자동으로 모델을 개선하고, 그 개선이 곧바로 제품 경험에 반영되는 루프가 내장되어야 한다. 수동으로 데이터를 정제·재학습하는 구조는 규모가 커질수록 비용이 폭증하며, a16z가 지적한 "수확체감"에 빠진다.
2. **사용자 간(across-user) 학습 효과**: A 고객사의 사용 데이터가 B 고객사의 제품 품질을 높이는 **교차 학습이 발생해야 진정한 규모 효과**가 작동한다. 단일 고객 내에서만 데이터가 순환하면 네트워크 효과가 아니라 고립된 규모 효과에 불과하며, 경쟁자가 각 고객을 개별 공략하면 쉽게 대체된다.
3. **복제 불가능한 독점 데이터**: 웹 크롤링이나 공개 소스로 재구성할 수 없는, 접근 자체에 **구조적 장벽이 있는 데이터**여야 한다. 규제·보안·물리적 조건(예: 제조 현장 센서, 의료 기록, 특수 시뮬레이션 결과)으로 보호되는 데이터만이 경쟁자의 "합성 데이터로 부트스트래핑" 전략을 무력화할 수 있다.
4. **높은 전환비용(switching cost)**: 데이터가 고객의 워크플로우·의사결정 프로세스에 깊이 내재화되어, 다른 솔루션으로 이전할 때 축적된 학습과 커스터마이징을 포기해야 하는 구조를 만들어야 한다. 데이터 해자는 단독으로 존재하는 것이 아니라 제품 임베딩과 결합될 때만 지속적 방어력을 갖는다.

4. 리스크와 기회는 동전의 양면이다

4.1 기술 성숙도 리스크 — 인지하되, 시장 기술과 공존하는 전략

뉴로-심볼릭 접근은 학술적으로 아직 미해결 과제가 많은 영역이다. 167편의 논문을 체계적으로 리뷰한 2024년 연구(Colelough & Regli, arXiv:2501.05435)에 따르면, AI 시스템이 스스로를 모니터링·평가·조정하는 메타인지(meta-cognition) 분야는 전체 연구의 **단 5%**에 불과하며, 설명가능성(explainability) 결과도 "상상했던 것보다 덜 명확"한 것으로 나타났다. 페블러스의 Data Greenhouse가 지향하는 "관측→판단→행동→증명"의 자율 진단 루프가 바로 이 메타인지 영역에 해당하며, 이는 아직 학술적으로도 선행 연구가 희소한 프론티어라는 의미다. 조인트 트레이닝 문제(경사 기반 학습이 심볼릭 구성요소를 직접 최적화할 수 없는 한계)도 여전히 근본적 장벽으로 남아 있다. 시뮬레이션-현실 격차(sim-to-real gap) 역시 자율주행·로보틱스·제조 검사 등

다양한 도메인에서 **20~35%의 성능 격차**가 보고되는 등, 시뮬레이션 기술 탄생 이래 지속적으로 해결이 요원한 본질적 과제다.

페블러스는 이러한 리스크를 명확히 인지하고 있으며, 그래서 자체 월드 모델을 구축하는 방향을 취하지 않는다. NVIDIA Cosmos나 JEPA 같은 대규모 월드 파운데이션 모델을 독자 개발하는 것은 거대 기업도 아직 검증하지 못한 고위험 경로다. 대신 페블러스는 이미 산업 현장에서 검증된 레거시 디지털 트윈, 시뮬레이터, 월드 모델을 **AI 에이전트로 활용**하는 전략을 취한다. 단단하지만 경직된 레거시 도구 위에 **유연한(말랑한) 뉴로-심볼릭 AI 레이어**를 씌워 도메인별로 적응적으로 커스터마이징하는 것이 핵심이다.

이 접근의 강점은 세 가지다. 첫째, 기존 디지털 트윈 인프라(Siemens, NVIDIA Omniverse 등)의 물리 엔진과 시뮬레이션 신뢰성을 그대로 활용하면서, 뉴럴 네트워크의 유연성과 심볼릭 추론의 설명가능성을 조합할 수 있다. 둘째, 자체 모델의 대규모 학습 비용과 검증 리스크를 회피하면서도, 도메인 특화 적응(domain-specific adaptation) 역량에 집중할 수 있다. 셋째, 다양한 시뮬레이터와 디지털 트윈 엔진에 대한 **폭넓은 기술 이해도** 자체가 후발 주자에 대한 진입 장벽이 된다. 페블러스의 Vector-to-Param 역변환 기술이 이 "래핑(wrapping)" 전략에서 sim-to-real 격차를 얼마나 효과적으로 줄이는지가 기술적 신뢰성의 핵심 검증 포인트가 될 것이다.

4.2 규제 환경은 구조적 기회를 창출한다

EU AI Act는 합성데이터 비즈니스에 **명시적인 수요 견인력**을 제공한다. **제10조 5항(a)**은 편향 탐지·교정에 실제 데이터가 불충분할 때 합성데이터 사용을 허용하고, **제59조 1항(b)**는 익명화 데이터가 부족할 때 합성데이터를 준수 대안으로 인정한다. 2025년 1월 EU 위원회가 발표한 학습 데이터 요약 템플릿은 6가지 데이터 범주 중 하나로 "자체 생성 합성 데이터셋"을 명시적으로 포함한다. 최대 **3,500만 유로 또는 글로벌 매출의 7%**에 달하는 과징금 구조는 기업들에게 문서화된, 감사 가능한 합성데이터를 사용할 강력한 인센티브를 만든다.

페블러스가 목표로 하는 **운영 증거(Operational Evidence) 패키지**는 이 규제 환경에서 프리미엄 상품이 될 수 있다. 실제 데이터의 소급적 문서화는 비용이 막대하고 리스크가 높은 반면, 합성데이터는 완전한 출처 추적(provenance)과 함께 생성되어 규제 준수가 설계 단계에서 내장된다. 한국의 AI 기본법(2026년 시행)도 EU AI Act의 원리를 차용하므로, 양쪽 프레임워크를 동시에 준수하는 기업은 글로벌 시장 접근에서 유리

하다. 한국 개인정보보호위원회는 2024년 12월 "합성데이터 생성·활용 안내서"를 발간하며 합성데이터를 프라이버시 강화 기술(PET)로 공식 인정했다.

4.3 정부과제 의존도는 양날의 검이다

페블러스의 7건의 국가 R&D 과제와 61억 원 규모의 과기정통부 과제는 기술 검증과 신뢰 구축에 핵심적이다. 그러나 정부과제 의존은 "**그랜터프렌터십(granterpreneurship)**" 위험, 즉 시장보다 과제 확보에 집중하게 되는 구조적 왜곡을 초래할 수 있다. 한국 맥락에서 정부 R&D 과제는 광범위한 문서 작성, 정산, 결과 보고를 요구하며 "에너지가 제품 개발이 아닌 과제 관리에 투입"되는 문제가 있다. 중국 제조 기업 연구에 따르면 정부 R&D 보조금은 혁신 산출에 긍정적이지만 **재무적 수익에는 유의미한 효과가 없다**.

성공적인 정부-상용 전환의 벤치마크로 Palantir(CIA 자금에서 시작해 17년 만에 상용 수익성 확보), Anduril(DHS 국경 감시에서 시작해 7년 만에 10억 달러 매출 달성), SpaceX(NASA/DoD 계약으로 기술 개발 후 Starlink로 상용 시장 개척)가 있다. 공통된 성공 패턴은 (1) 정부 자금으로 기술 위험완화(de-resking), (2) 동시에 상용 제품 개발, (3) 정부 매출 비중을 **50% 이하로 관리**하는 것이다.

5. 삼각형 협력 모델에서 스타트업의 최적 전략

페블러스가 제안하는 "수요 대기업(현대자동차) + 대학(KAIST) + 스타트업(페블러스)"의 삼각형 구조는 Henry Etzkowitz의 트리플 헬릭스(Triple Helix) 모델에 이론적 기반을 둔다. 한국에서 이 모델의 성공 사례로 POSTECH-POSCO(기업이 대학을 설립한 "기업 헬릭스"), 성균관대-삼성(20년간 집중 투자로 Top-Tier 대학으로 변환), 현대차-KAIST(자율주행 센서 공동 연구소) 등이 있다.

그러나 **한국 기업의 산학연 협력 경험률은 25.5%에 불과**하며, 주요 장벽은 외부 기술 흡수 어려움(31.3%), 필요 기술 발굴 어려움(18.8%), 소통 어려움(16.7%)이다. 한국의 재벌 지배 경제구조에서 스타트업은 "갑-을" 관계로 전략할 위험이 상존한다.

페블러스의 방어적 포지셔닝은 현재로서는 양호하다. 국내 특허 36건(등록 5건), 미국 특허 3건(등록 2건)의 IP 포트폴리오, CAS Class A 인증, Gartner의 "진단-합성 페어

링 모델" 대표 사례 인정, 복수 대기업 고객(현대·한화·삼성·LG) 보유를 통해 단일 파트너 의존을 회피하고 있다. 정부 R&D 과제에서 **주관기관**으로 선정된 것도 대기업 파트너 대비 협상력을 강화하는 요소다.

최적 포지셔닝 전략은 5가지로 요약된다. 첫째, IP 소유권을 사전에 명확히 정의하되, 원천 데이터·합성데이터·학습 모델 가중치 각각에 대한 소유권을 구분해야 한다. 둘째, SI 프로젝트가 아닌 "전략적 파트너십"으로 관계를 정의해야 한다. 셋째, 에스스로 조항을 통해 핵심 IP를 보호해야 한다. 넷째, SaaS + 온프레미스 + API + 정부과제의 **다중 매출 스트림**을 유지해야 한다. 다섯째, AWS Marketplace 등록처럼 대기업 우회 채널을 확보해야 한다.

한국 국방 생태계에서 한화에어로스페이스(2025년 매출 26.6조 원, +137%), LIG넥스원(Palantir·Anduril·Shield AI와 MOU), HD현대(Anduril과 무인수상함 개발), 현대로템(Shield AI와 자율전투 소프트웨어 협력) 등이 활발히 AI 투자를 확대하고 있다. 그러나 한국은 미국 DoD의 OTA(Other Transaction Authority) 같은 **신속 기술 도입 제도가 부재**하며, 민간 기술의 군사 배치에 4~5년이 소요되는 구조적 한계가 있다.

6. 전략적 제언: 단계별 실행 로드맵

단기 전략 (2026년, 1년)

웨이 유스케이스 #1: Data Clinic을 통한 데이터 품질 진단 SaaS 확산. 이미 현대차, 한화비전 등에서 검증된 Data Clinic의 시간당 **10만 장 이상 이미지 진단** 역량을 AWS Marketplace와 직접 SaaS로 확산하는 것이 최우선이다. 합성데이터 5%만 추가하면 AI 모델 성능이 **약 2% 개선**된다는 정량적 증거는 강력한 세일즈 도구다. 프리미엄 무료 티어를 제공하고, 엔터프라이즈 연간 계약(5만~50만 달러)으로 전환하는 **MOSTLY AI/Gretel** 모델을 벤치마킹해야 한다.

웨이 유스케이스 #2: 운영 증거 패키지의 선제 출시. EU AI Act 고위험 AI 시스템 규제가 2026년 8월(또는 Digital Omnibus에 의해 2027년 12월) 시행되기 전에, ISO 42001/prEN 18286 정렬 운영 증거 패키지를 상품화해야 한다. 한국 AI 기본법(2026년 1월 시행)과 EU AI Act의 이중 준수를 지원하는 것은 글로벌 제조기업(현대차, 삼성 등)에 대한 차별화된 가치 제안이다.

신규 투자 라운드 완료와 함께, 정부과제 매출 비중을 모니터링하고 상용 매출 파이프라인을 구축하기 시작해야 한다. AADS 2단계 과제를 통한 Data Greenhouse의 마스터 오케스트레이션 에이전트 기능 고도화는 기술적 차별화의 핵심이다.

중기 전략 (2027~2028년, 2~3년)

PebbloSim-Factory의 단계적 상용화. PoC #1(자동차 제조) → PoC #2(국방) → PoC #3(조선) 순서는 적절하다. 자동차 제조는 현대차 ZERO1NE과의 관계를 활용해 가장 빠르게 검증할 수 있고, 국방은 방사청의 '미래 도전 국방기술사업' 중 피지컬 AI·첨단전력 신규과제(1,000억 원)를 포함해 범정부 AI 예산 약 10조 원의 뒷받침이 있으며, 조선은 HD현대 FOS 프로젝트·한화오션 스마트야드·삼성중공업 SYARD 등 3사 합산 1조 원 이상의 디지털 전환 투자가 맞물린다.

데이터 플라이휠 가동의 임계질량(critical mass) 도달. Data Clinic의 진단 데이터가 PebbloSim의 합성데이터 품질을 개선하고, 이것이 다시 Data Clinic의 진단 정밀도를 높이는 선순환 구조가 실질적으로 작동하려면 **충분한 고객 기반**이 필요하다. 제조·국방·조선 3개 도메인에서 각 3~5개 엔터프라이즈 고객을 확보하는 것을 목표로 해야 한다.

정부과제 매출 비중을 50% 이하로 감소. Palantir의 17년, Anduril의 7년과 달리 페블러스는 스타트업으로서 빠른 전환이 필요하다. SaaS ARR(연간반복매출)을 정부 R&D 매출과 최소 동등하게 만드는 것이 중기 목표다.

글로벌 확장의 초석. EU AI Act 대응 역량을 활용해 유럽 제조기업(독일 자동차 OEM 등)을 타겟하고, NVIDIA Omniverse/Cosmos 생태계 내에서 파트너(Parallel Domain처럼) 포지셔닝을 고려해야 한다. NVIDIA의 Gretel 인수(\$3.2억)와 SAS의 Hazy 인수는 합성데이터 스타트업에 대한 전략적 M&A 수요가 존재함을 보여주므로, 독립 성장과 전략적 인수 양쪽 모두에 대비하는 것이 현실적이다.

실행 우선순위 매트릭스

우선순 위	전략	시급 성	영향 력	복잡 도
1	Data Clinic SaaS 확산 + 프리미엄 가격 체계	높음	높음	낮음

2	운영 증거 패키지 상품화	높음	높음	중간
3	AADS 2단계 기술 고도화	높음	중간	높음
4	PebbloSim PoC #1 자동차 검증	중간	높음	높음
5	상용 ARR 구축 + 정부과제 비중 감소	중간	높음	중간
6	PebbloSim PoC #2-3 국방·조선	낮음	높음	높음
7	NVIDIA 생태계 파트너십/글로벌 진출	낮음	높음	높음

결론: 구조적 기회와 실행 리스크의 비대칭

페블러스는 세 가지 구조적 순풍 속에 있다. 한국 정부의 전례 없는 피지컬 AI 투자(산업부 M.AX 1조 원 이상, 범정부 AI 예산 약 10조 원), EU AI Act가 만드는 합성데이터의 규제 대응 필수재화, 그리고 "데이터 OS + 품질 평가 + 시뮬레이션 생성" 통합 플랫폼의 시장 공백이다. 14인 규모의 팀이 현대차·한화·삼성·LG를 고객으로 확보하고, 과기정통부 주관기관으로 선정된 것은 기술력과 실행력의 증거다.

그러나 핵심 기술(뉴로-심볼릭, Vector-to-Param)의 성숙도 검증, 30~35%에 달하는 sim-to-real 격차의 극복, 정부과제 의존에서 상용 매출로의 전환은 실행 리스크가 크다. Datagen의 폐업과 Synthesis AI의 위축은 합성데이터 시장에서 기술만으로는 생존할 수 없음을 보여준다. 워크플로우 임베딩을 통한 전환비용 창출, 플라이휠의 실질적 가동, 그리고 대기업·대학과의 협력에서 "전략적 파트너"로서의 위치 유지가 성패를 결정할 것이다. **Applied Intuition**이 2017년 설립 후 8년 만에 150억 달러 기업가치를 달성한 것은 올바른 시장에서 올바른 전략을 실행하면 가능한 궤적이며, 페블러스가 추구해야 할 궁극적 벤치마크다.

Pebblous

Pebblous Makes Data Tangible

contact@pebbulous.ai