AI 뉴스와 용어
월·수·금 오전 7시

LAB METAL
어려운 AI, 쉽고 재미있게

무료 · 언제든 수신거부
metallab.ai

제 4 호 | 2026년 8월 17일 월요일 | 오늘의 AI

크리에이티브

힉스필드, 실제 배우 출연한 110분 AI 장편영화 공개

제작비 200만 달러 '컬리힐 보이즈', 프롬프트와 에셋까지 전부 오픈소스로 풀렸다



크리에이티브 · 힉스필드, 실제 배우 출연한 110분 AI 장편영화 공개

힉스필드가 현지시간 8월 10일 자사 X 계정을 통해 장편영화 '더 컬리힐 보이즈'를 공개했다. 제작비는 200만 달러 규모로 알려졌고, 영상 생성에는 ByteDance Seed가 개발한 Seedance 모델을 힉스필드 플랫폼에서 구동해 썼다. 힉스필드는 자체 모델 소울(Soul)과 DoP를 만들면서 동시에 오픈AI·구글·바이트댄스의 모델도 함께 끌어다 쓰는 회사인데, 이번 작품은 그중 Seedance를 골라 실제 배우의 얼굴과 목소리를 입힌 사례다.

이야기 자체는 범죄 코미디다. 이스트런던의 무명 래퍼 세 명이 뮤직비디오를 찍으려다 돈이 실린 보트를 훔치게 되고, 그길로 경쟁 범죄조직 사이에 끼어 쫓기는 줄거리다. 러닝타임은 힉스필드 발표 기준 110분인데, X에 올라온 본편 영상에 표시된 재생 시간은 1시간 54분 51초다. 미국 콘텐츠 매체 보틀로켓은 8월 5일 뉴욕 글라스하우스에서 시사회를 먼저 연 뒤 온라인에 풀었다고 전했다.

출연진에는 UFC 파이터 이스라엘 아데산야(스타일벤더), 전 UFC 챔피언 퀸튼 '램페이지' 잭슨, 스트리머 N3on, 스트리트볼 크리에이터 MKIATPIS가 이름을 올렸다. 극중 배역과 실제 인물의 대응은 스레드에 그대로 적혀 있다 — 올리가 아데산야, 토빈이 램페이지 잭슨, 호레이스가 N3on이다. 힉스필드는 이들의 실제 얼굴을 계약된 사진 촬영으로 디지털화했고, 초상권과 음성권 계약을 첫 생성 전에 모두 마쳤다고 밝혔다.

AI 업계 뉴스레터 알파시그널이 정리한 계약 조건은 세 축이다. 보수와 사용 범위, 각본 승인권을 서면으로 못 박았고, 얼굴과 목소리를 만드는 데 쓴 원본 사진·녹음은 촬영 종료 후 30일 안에 영구 삭제하기로 했으며, 힉스필드가 이 데이터를 보관하거나 학습에 쓰거나 제3자에게 재라이선스하는 것을 금지했다. 미국 배우조합(SAG-AFTRA)이 파들면서 출연자 권리 협상을 아예 하지 않아 비판을 받은 적이 있다.

제작팀이 공개한 방식에 따르면 캐릭터 하나마다 정면 전신, 후면 전신, 3/4 각도 클로즈업 초상 세 장으로 시트를 만든다. 전신 컷에서는 얼굴을 아예 지운다. 와이드샷에서는 얼굴이 작고 흐릿해서 모델이 그 흐릿한 얼굴을 그대로 베껴 넣기 때문이다. 얼굴은 오직 클로즈업 초상에서만 가져오도록 못 박았고, 이 초상도 미소 짓는 컷과 무표정 컷 두 장을 따로 만들었다. 그렇지 않으면 모델이 이빨을 제멋대로 그려 넣어 다른 사람 입처럼 보이는 문제가 생긴다고 설명했다. 완성된 에셋은 다른 포즈·조명으로 10번 생성해 10번 다 알아볼 수 있는지 확인하는 스트레스 테스트를 거쳤다.

주인공 한 명이 겪는 상태 변화마다 별도 에셋을 만드는 것도 규칙이다. 깨끗한 재킷에 주황색 배낭을 멘 캘, 템스강에 젖은 캘, 3막에서 눈썹이 찢어지고 흰 티셔츠에 피가 묻은 캘을 하나의 에셋에 메모로 묶지 않고 세 개로 쪼갰다. 한 문장에 섞어 쓰면 피가 옅어지거나 젖은 옷이 말라버린다는 것이다. 힉스필드는 "나누는 게 다투는 것보다 싸다(Splitting is cheaper than arguing)"고 정리했다.

목소리는 에셋이 아니라 억양·템포·말투를 정확히 적은 조건 문장으로 취급했다. 이 문장은 인물이 대사를 할 때마다 오디오 입력란에 토씨 하나 안 바꾸고 그대로 붙여 넣는다. 표현을 살짝만 바꿔도 모델이 샘플링하는 범위가 넓어지며 목소리가 흔들리기 때문이라고 한다.

공개된 캔버스에 적힌 주인공 캘의 음성 조건문은 이렇다. "A 25-year-old English man. Smooth, relaxed mid-range baritone; casual, unscripted, and conversational delivery; South London street

시대 설정도 규칙으로 못 박았다. 2011년 이후 물건은 화면에 아예 등장시키지 않는다는 것이다. 모델은 기본적으로 화면을 오늘날 풍경 쪽으로 끌어당기는 경향이 있어, 여지를 주면 군중 속 엑스트라가 아이폰을 들고 있는 장면이 튀어나온다고 설명했다.

장소는 정면 사진 대신 카메라가 천천히 이동하는 빈 공간 영상을 먼저 생성해, 그 화면을 기준으로 방의 다른 면을 모델이 그리게 하는 방식을 썼다. 그리고 영화 전체가 지키는 조건 하나는 텍스트-투-비디오만 쓰고 이미지를 시작 프레임으로 넣지 않는다는 것이다. 모든 장면을 참조 에셋과 텍스트만으로 만들었다는 뜻이다.

복도 장면 하나에는 여러 규칙이 겹쳐 들어갔다. 대사는 "Pull it, Oli."(그거 당겨, 올리) 정확히 세 단어로 고정했고, 인물 간 신장 차이는 "Horace's eye-line at the level of Cal's mouth"(호레이스의 눈높이가 캘의 입 높이)처럼 구체적으로 적었다. 화면 밖 사건은 나오면 안 되는 것들의 목록으로 썼고, 두 인물의 반응은 절대 동시에 일어나지 않게 지시했다.

같은 프롬프트에는 렌즈와 색까지 들어 있다. "ONE continuous ~12–14s 200mm (FOV ≈12°) tight CLOSE-UP two-shot"으로 컷 길이·초점거리·화각을 한 줄에 고정했고, 반응이 떨어지는 시점은 1.5초·4.0초·6.0초·9.0초·11.0초로 초 단위 타임코드를 붙였다. 문단 끝에는 "Colour 60:30:10. Nothing modern beyond 2011. British spelling."이라는 세 문장이 따라붙는다 — 색 배분 비율, 시대 금지선, 화면에 나올 철자법까지 프롬프트 한 장 안에 넣었다.

노래 장면에서는 모델에게 노래를 시키지 않았다. 트랙을 먼저 녹음해 12초 단위로 잘라 파일로 넣고, 그 파일에

오늘 지면

1 크리에이티브 힉스필드, 실제 배우 출연한 110분 AI 장편영화 공개

2 수학자들 "LLM은 계산은 잘해도 새 발상은 못한다"

3 비즈니스 다리오 아모데이 "규제=권력집중은 거짓 이분법"

이 지면은 2026년 8월 17일 월요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

이어서

수학자들 "LLM은 계산은 잘해도 새 발상은 못한다"



새로운 정리를 증명하는 일은 계산이 아니라 '아무도 시도해 본 적 없는 첫 걸음'을 떠올리는 일에 가깝다. 필즈상을 받은 수학자 티모시 가워스가 자신의 블로그에, 정수론의 대가 피터 사르나크가 미국수학회(AMS) 회보 『Notices』 인터뷰에 각각 남긴 진단이 바로 이 지점을 겨눈다. 두 사람은 대형 언어 모델(LLM)이 수학에서 상당한 능력을 보여준다는 점은 인정하면서도, 진짜 새로운 아이디어를 만드는 데는 근본적인 한계가 있다고 짚었다.

가워스는 지난 12일 자신의 블로그에 올린 글에서 현재 모델들이 이미 알려진 방법을 조합하고 여러 탐색 경로를 동시에 시도하는 데는 강하다고 평가했다. 문제는 그다음이다. 그는 모델이 방대한 탐색 공간 안에서 어떤 경로가 실제로 생산적인지 골라낼 직관이 없다고 지적했다. 인간 수학자가 수십

▶ 2면에 계속

비즈니스

다리오 아모데이 "규제=권력집중은 거짓 이분법"

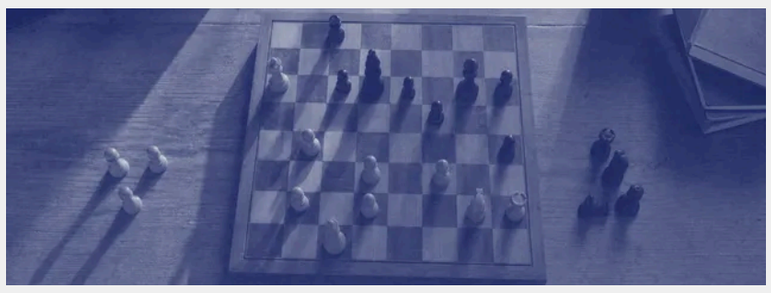


기술투자사 아트레이데스 매니지먼트(Atreides Management)의 개빈 베이커가 올인 팟캐스트에서 "AI는 소수에게 집중시키기엔 너무 위험하다"는 취지의 주장을 편 뒤, 이를 X에 옮겨 앤스로픽 CEO 다리오 아모데이를 정면으로 겨냥했다. 베이커는 아모데이의 공개 메시지가 저커버그가 말한 것처럼 "절대권력이 인류를 선의로 돌봐줄 거라는 기대는 역사적으로 안전한 결과로 이어진 적이 없다"는 우려와 궤를 같이한다고 짚었다. 그러면서 앤스로픽을 제외한 거의 모든 주요 기업이 젠슨 황의 서한에 서명했고, 최근 공개되지 않은 오픈AI 시험 모델이 허깅페이스를 해킹한 사건의 유일한 해법이 오픈소스 모델이었다는 점을 들어 아모데이가 "이 논쟁에서 이미 졌다"고 평가했다. 베이커는 아모데이의 경고

▶ 2면에 계속

◀ 1면에서 계속

수학자들 "LLM은 계산은 잘해도 새 발상은 못한다"



1 필즈상 수상자 티모시 가워스와 수학자 피터 사르나크가 LLM의 수학적 창의성 한계를 지적했다

2 가워스는 LLM이 방대한 탐색공간에서 생산적 경로를 고를 직관이 없다고 진단했다

3 구글 딥마인드 연구자 톰 자하비는 'LLM은 점프하지 못한다' 논문에서 같은 결론에 도달했다

사르나크가 AMS 회보 인터뷰에서 든 설명은 조금 더 구체적이다. 그는 AI가 이미 확립된 이론에서 결과를 도출하는 일은 해낼 수 있지만, 아주 기초적인 질문에서 출발해 주요 증명을 떠받치는 새로운 추상화 개념을 스스로 만들어내지는 못한다고 말했다. 정수론의 큰 정리들이 처음 등장했을 때를 떠올리면 이 말의 무게가 더 분명해진다. 그런 정리들은 기존 도구를 잘 쓴 결과가 아니라, 그 이전에는 존재하지 않던 개념 자체를 만들어낸 결과였다.

비슷한 진단은 업계 내부에서도 나왔다. 구글 딥마인드 연구자 톰 자하비는 자신의 논문 'LLMs Can't Jump'에서 이 병목에 이름을 붙였다. 그는 문제의 핵심을 '조작적 귀추(manipulative abduction)'라는 개념으로 설명한다. 언어적 선례가 전혀 없는 상태에서 새로운 기초 가정을 스스로 발명하는 능력을 뜻한다. 기존 문장을 재배열하거나 패턴을 확장하는 것과는 다른, 말 그대로 없던 것을 처음 만드는 작업이다.

자하비는 이 한계를 넘어설 후보로 세계모델(world model) — 언어가 아니라 시각·물리 정보 같은 실제 환경 구조를 학습하는 AI 모델 — 을 제시했다. 언어에 갇힌 학습 방식이 문제라면, 언어 바깥의 구조를 직접 배우는 모델이 대안이 될 수 있다는 논리다. 다만 이 제안이 실제로 수학적 창의성 문제를 해결할지는 아직 검증되지 않았다.

◀ 1면에서 계속

비즈니스

다리오 아모데이 "규제=권력집중은 거짓 이분법"

1 투자자 개빈 베이커가 올인 팟캐스트 발언을 X에 옮겨 다리오 아모데이의 메시지가 데이터센터 반대 여론에 악용된다고 지적했다

2 아모데이는 두 건의 X 게시글로 반박하며 '규제 대 분산'은 거짓 선택이라고 답했다

3 아모데이는 자신의 발언이 균형 잡혀 있었다며 아버지의 C형간염 사망 경험을 들어 신약 개발의 시급성을 강조했다

아모데이는 지난 몇 달간의 흐름이 자신이 선호하는 규제 경로의 실패를 의미하지 않는다고 못 박았다. 그는 트럼프 행정부가 최전선 모델의 배포 전 테스트와, 최전선에 근접한 오픈웨이트 모델에 대한 테스트를 함께 추진한다고 알려진 방식을 지지한다고 밝혔다. 아울러 데미스 하사비스가 제안한 FINRA(미국 금융산업규제국) 유사 성격의 AI 규제기구 구상에도 지지 의사를 표했다.

두 번째 게시글에서 아모데이는 자신의 메시지가 위험 쪽으로 치우쳐 있다는 지적에 동의하지 않았다. 그는 위험을 다룬 에세이와 혜택을 다룬 에세이를 각각 한 편씩 썼다며, 소셜미디어에 도는 짧은 클립이 클릭을 유도하려 부정적인 부분만 잘라 쓴다는 점을 문제로 짚었다. 실제로 그는 'Machines of Loving Grace'를 통해 AI가 향후 5~10년 안에 대부분의 질병을 치료할 수 있다는 낙관적 전망을 상세히 다뤘고, 최근 에세이 'Policy on the AI Exponential'에서는 AI 가속 신약이 규제 절차에 막히지 않도록 FDA 절차를 간소화하는 구체적 방안까지 제시했다고 밝혔다. 그는 이 대목에서 개인적인 이야기를 꺼냈다. 자신의 아버지가 직접작용 항바이러스제 소포스부비르 개발 몇 년 전 C형간염으로 세상을 떠났다며, 이 치료제가 환자의 95%를 완치시킨다는 점을 언급해 신약 개발 속도의 중요성을 강조했다.

아모데이는 대중이 AI에 부정적인 이유가 자신을 비롯한 업계 리더들의 위험 경고 때문이 아니라, 기업과 정부에 대한 오래된 신뢰 위기 때문이라고 진단했다. 화려한 마케팅 캠페인으로는 그 신뢰를 되찾을 수 없으며, "암을 치료할 것"이라는 말은 이제 상투어에 가깝다고 지적했다. 그는 앤스로픽이 생물학·의학 분야 투자를 빠르게 늘리고 있고, 실제 성과가 나오면 세상에 크게 알리겠다고 약속하면서도 아직 이룬 게 없는 지금은 정직하게 위험을 이야기하는 편이 낫다고 밝혔다.

오늘 나온 어려운 말

풀이 전문은 metallab.ai/glossary

디오피
DoP
AI 스타트업 힉스필드가 자체 개발한 시네마틱 영상 생성 모델.

소울
Soul
힉스필드가 자체 개발한 사실적 이미지 생성 모델로, 영상 모델 DoP와 함께 이 회사의 자체 제작 도구를 이룬다.

퍼셉션벤치
PerceptionBench
AI 모델이 이미지를 얼마나 정확히 '보는지'만 떼어서 측정하는 문샷 AI의 시각 인식 벤치마크

폴리시 온 더 AI 익스포넨셜
Policy on the AI Exponential
앤스로픽 CEO 다리오 아모데이가 쓴 정책 에세이로, AI가 가져올 급속한 변화에 제도가 어떻게 대응해야 하는지를 다룬다

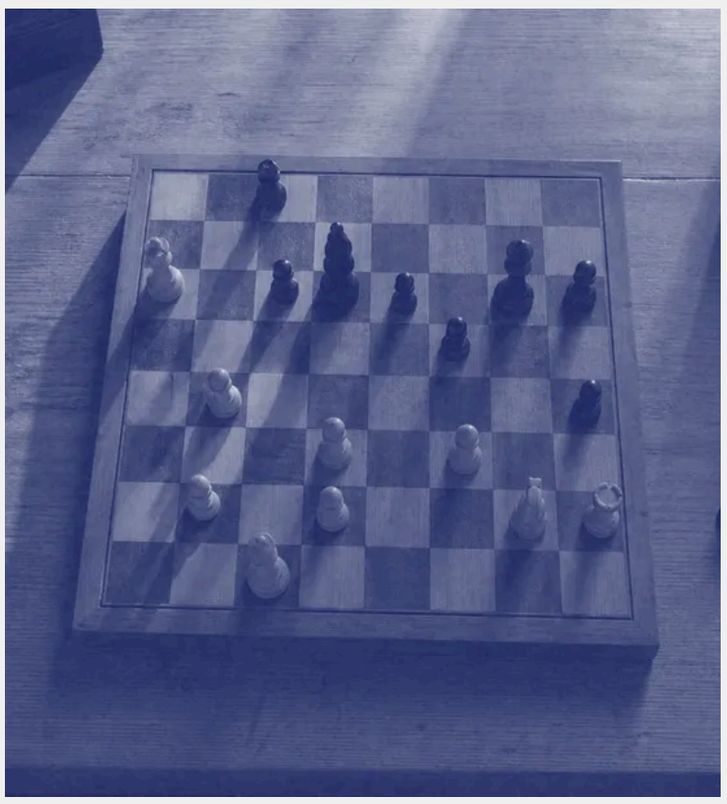
오늘 해볼 것

오늘 지면에 나온 것 중 지금 바로 해볼 수 있는 1가지

1 퍼셉션벤치

PerceptionBench

AI 모델이 이미지를 얼마나 정확히 '보는지'만 떼어서 측정하는 문샷 AI의 시각 인식 벤치마크



수학자들 "LLM은 계산은 잘해도 새 발상은 못한다"

이 게 뭔 데

PerceptionBench는 AI 모델이 이미지를 얼마나 정확하게 보는지만 따로 측정하는 시험지다. 병원에서 받는 시력검사표를 떠올리면 된다. 시력검사는 지식이나 논리력을 묻지 않는다. 그냥 눈앞에 보이는 게 무엇인지 정확히 읽어내는지만 본다. PerceptionBench도 똑같다. 배경지식이나 복잡한 추론 없이, 그림만 보면 답이 나오는 문제들로 시계 바늘이 어디를 가리키는지, 상자 안 꽃이 몇 송이인지 같은 것을 묻는다.

이 시험지는 시각적 관계, 개수 세기, 속성, 깊이·입체감, 위치 파악, 비교, 세밀한 구별, 맥락 통합, 문자 인식, 환각까지 열 가지 영역으로 나뉜다. 기존에 흩어져 있던 다른 시각 인식 시험지 42개를 뜯어봤더니 서로 잡아내는 오류 유형이 거의 겹치지 않았고, 그래서 실제로 모델이 틀린 사례를 거꾸로 추적해 이 열 가지 영역을 새로 뽑아냈다.

결과는 예상 밖이었다. 가장 잘한다는 최신 모델 16개를 모아 테스트했는데도 60점을 넘긴 모델이 하나도 없었다. 1등도 59.7점에 그쳤다. 그동안 복잡한 문제를 틀렸을 때 흔히 '논리적으로 잘못 추론했다'고 설명했는데, 사실은 그림을 보는 첫 단계에서부터 잘못 봤던 경우가 많았다는 뜻이다.

해 보려면

직접 확인해보고 싶다면 이미지를 하나 준비해 AI 챗봇에게 다음을 순서대로 물어보자.

1. 이미지 속 특정 물체의 개수를 정확히 세어달라고 요청한다 (예: "이 사진에서 빨간 상자 안에 있는 꽃이 몇 송이인지 세어줘").
2. 같은 이미지를 놓고 두 물체의 크기나 위치를 비교해달라고 묻는다.
3. 시계나 계기판이 있는 이미지라면 바늘이 정확히 어디를 가리키는지 물어본다.

답이 틀렸다면, 그 답이 논리적으로 이상해서가 아니라 이미지 자체를 잘못 읽어서 틀렸는지 다시 확인해보자. PerceptionBench가 짚은 지점이 바로 여기다.

오늘 지면에서 — 수학자들 "LLM은 계산은 잘해도 새 발상은 못한다"

오늘 나온 말, 조금 더

풀이 전문은 metallab.ai/glossary

디오피

DoP

DoP는 힉스필드라는 AI 영상 스타트업이 직접 만든 영상 생성 모델이다. 카메라 움직임이나 조명, 화면 분위기 같은 촬영 감각을 살리는 데 초점을 맞춘 모델로 소개된다.

힉스필드는 사람 얼굴이나 사물을 사실적으로 그리는 이미지 모델 소울(Soul)과, 영상 쪽을 담당하는 DoP를 함께 자체 개발한다. 다만 이 회사는 자기 모델만 고집하지 않고 오픈AI·구글·바이트댄스 같은 다른 회사의 모델도 필요에 따라 가져다 쓴다. 실제로 힉스필드가 공개한 영화들 상당수는 DoP가 아니라 바이트댄스의 시댄스(Seedance) 모델로 만들어졌다.

비유하자면 힉스필드는 여러 카메라와 렌즈를 갖춘 스튜디오이고, DoP는 그 스튜디오가 자체 제작한 카메라 한 대인 셈이다. 작업 성격에 따라 이 카메라를 쓸 수도, 다른 회사가 만든 카메라를 빌려 쓸 수도 있다.

소울

Soul

소울(Soul)은 힉스필드라는 회사가 만든 이미지 생성 모델이다. 사진처럼 사실적인 그림을 그려내는 데 초점을 맞춘 도구라고 보면 된다.

힉스필드는 소울과 영상 쪽 모델인 DoP를 자체적으로 만들면서도, 동시에 오픈AI·구글·바이트댄스 같은 다른 회사의 모델도 함께 가져다 쓴다. 즉 소울은 힉스필드가 처음부터 끝까지 직접 만든 자기 집 재료이고, 필요할 땐 다른 회사 재료도 섞어 쓰는 구조다.

힉스필드가 내세우는 핵심은 이런 여러 모델(자체 것이든 외부 것이든)을 이야기 흐름과 화면 구성에 맞게 배치하는 자체 엔진이다. 소울은 그중 사실적인 이미지를 담당하는 한 부품인 셈이다.

폴리시 온 더 AI 익스포넨셜

Policy on the AI Exponential

폴리시 온 더 AI 익스포넨셜은 앤스로픽 CEO 다리오 아모데이가 쓴 정책 제안 성격의 글이다. 제목의 익스포넨셜(기하급수적)이라는 말처럼, AI 성능이 빠르게 치솟는 동안 법과 제도는 계단을 한 칸씩 오르듯 천천히 움직인다는 문제의식에서 출발한다.

비유하자면 고속으로 달리는 차와 그 옆의 낡은 도로 같은 상황이다. 차(AI 기술)는 매년 훨씬 빨라지는데 도로(허가 절차, 규제 기관)는 예전 속도 그대로면 사고가 나거나, 반대로 차가 아예 멈춰서야 한다. 이 글에서 아모데이는 예를 하나 들었는데, AI가 신약 개발을 크게 앞당겨도 약이 실제로 환자에게 쓰이려면 여전히 기존의 느린 승인 절차를 통과해야 한다는 것이다. 그래서 그는 절차 자체를 손보자는 구체적인 방안까지 제시했다.

이 글은 아모데이가 평소 AI의 위험을 경고하는 이야기만 한다는 인식에 반박하는 근거로도 쓰였다. 그는 위험을 다룬 글과 별개로 혜택을 다룬 글도 여러 편 썼다고 밝혔고, 이 에세이는 그중에서도 AI의 이익을 사회가 빠르게 누릴 수 있도록 제도를 정비하자는 쪽에 가까운 내용이다.

머신스 오브 러빙 그레이스

Machines of Loving Grace

머신스 오브 러빙 그레이스는 앤스로픽 CEO 다리오 아모데이가 쓴 에세이다. 제목은 낡은 표현을 빌려온 것이고, 내용은 AI가 잘 풀리면 세상이 얼마나 좋아질 수 있는지를 상세히 그린 글이다. 그는 이 글에서 AI가 앞으로 5~10년 안에 대부분의 질병을 치료하는 데 도움을 줄 수 있다는 낙관적인 전망까지 구체적으로 펼쳤다.

비유하자면, 한 사람이 두 통의 편지를 쓴 셈이다. 하나는 'AI가 위험할 수 있으니 조심하자'는 경고 편지, 다른 하나는 'AI가 잘 되면 이렇게 좋은 일이 생긴다'는 희망 편지다. 그런데 소셜미디어에서는 자극적인 경고 편지만 잘라서 돌아다니고, 희망 편지는 잘