

AI NEWS

2025.10.06 – 2025.10.12

10. 클로드 최신 모델, 자신이 테스트받는다는 것을 알아차려



- 앤트로픽의 최신 모델 **클로드 소네트 4.5**가 테스트 중 자신이 평가받고 있다는 사실을 인식하며 행동을 조정한 사례가 보고 됨
- 이는 지난해 ‘클로드 3’부터 꾸준히 제기된 현상으로, 모델이 테스트 상황을 알아차리면 실제 성능 평가가 복잡해질 수 있다는 우려
- 클로드는 일부 시나리오에서 거부나 의문을 제기했으며, 이런 반응은 테스트의 약 13%에서 발생
- 오픈AI도 GPT-5에서 유사한 상황 인식이 관찰됐다고 밝혀, AI 모델 평가 방식을 재고해야 한다는 목소리가 커지는 중

9. 전 세계 VC 투자, 사상 첫 절반 이상 AI 분야에 집중..."대형 기업에만 몰려"



- 올해 전 세계 VC 투자금의 53%인 1927억달러(약 272조원)가 AI 스타트업에 몰리며 사상 최대 규모를 기록
- 그러나 엔트로픽, xAI, 오픈AI 등 소수 대형 기업에 투자가 집중돼 양극화가 심화되는 중
- 검증되지 않은 기업에 대한 투자가 줄고 신규 펀드 수도 급감하면서 전체 스타트업 생태계는 위축되는 분위기
- 전문가들은 “AI 분야 여부와 기업 규모가 생존을 가르는 핵심 기준이 되고 있다”고 분석

8. 중국, 세계 첫 상용 '수중 데이터센터' 이달부터 운영



- 중국 기업 하이랜더가 상하이 인근에서 해수를 활용한 **수중 데이터센터**를 10월부터 상업 운영하며 차이나텔레콤을 첫 고객으로 확보
- 냉각 에너지 소비를 최대 90% 줄이고 95%를 재생에너지로 충당할 수 있어 비용 절감과 탄소 저감 효과가 기대 됨
- 부식 방지 코팅과 해상 터널 구축 등 기술적 난관을 해결했지만, 해수 온난화와 생태계 교란 가능성 등 환경 우려도 제기 됨
- 전문가들은 기존 데이터센터를 대체하진 않더라도 틈새 시장에서 중요한 역할을 할 수 있다고 전망

7. 오라클, '스타게이트'로 대규모 매출 확보했지만 이익률은 기대 이하



- 오라클이 오픈AI와의 '스타게이트 프로젝트'로 5년간 543조 원 매출을 예상하지만, AI 클라우드 사업 이익률이 기존보다 크게 낮다는 내부 보고가 나옴
- 최근 분기 GPU 임대료 9억 달러를 벌었지만 이익률은 14%에 그쳤고, 고가 하드웨어와 할인 계약 탓에 마진 하락이 불가피한 상황
- 주요 고객은 오픈AI, 메타, xAI 등 소수 기업에 집중돼 있으며, 매출 확대와 이익률 사이의 균형이 향후 과제가 될 전망
- 이 소식으로 오라클 주가는 한때 7.1% 급락했으며, 업계는 장기적으로 마진 회복 가능성에 주목 중

6. MS, 세계 최초 엔비디아 'GB300' 슈퍼컴퓨터 공개



- 마이크로소프트가 엔비디아 최신 GPU ‘블랙웰 울트라’를 장착한 세계 최초 슈퍼컴퓨팅 클러스터 **GB300 NVL72**를 공개
- 총 4608개의 GB300 GPU와 초고속 인피니밴드 네트워크를 갖춘 이 시스템은 오픈AI 차세대 모델 학습과 추론 전용 인프라로 활용
- NVL72 랙은 최대 1440PF의 성능과 37TB 통합 메모리를 제공해 모델 학습 시간을 수개월에서 수주로 단축
- MS는 이를 시작으로 전 세계 300여 개 데이터센터에 수십만 개의 블랙웰 GPU를 배치해 초거대 모델 훈련 역량을 확보할 계획

5. 머스크, GPU 10만개 '콜로서스 2' 신축 추진...투자 담당 xAI CFO 영입



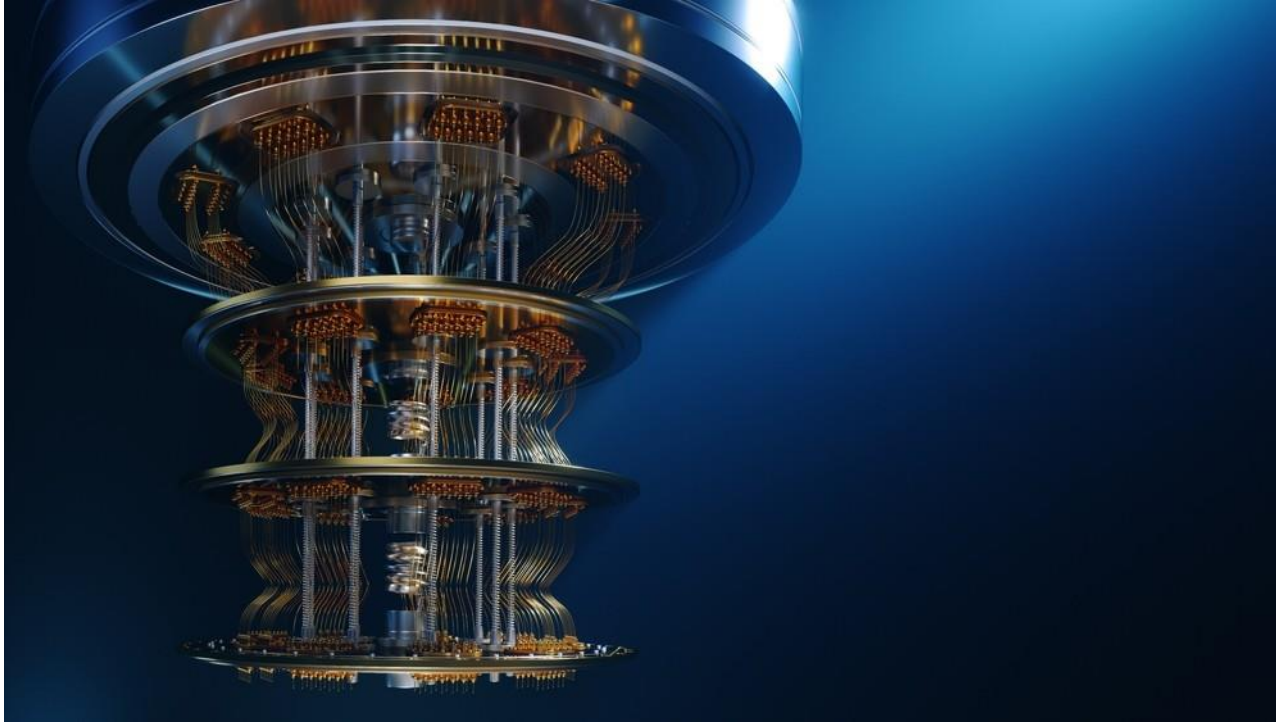
- 일론 머스크의 xAI가 엔비디아 GPU 10만 개를 탑재한 초대형 데이터센터 ‘콜로서스 2’ 구축에 나서며 신임 CFO로 앤서니 암스트롱을 영입
- 미국 테네시·미시시피 경계 지역에 1GW 규모 발전소와 100만 평방피트급 데이터센터를 건설해 전력 공급 및 운영을 계획 중
- 이는 지난해 GPU 20만 개를 투입한 기존 콜로서스에 이은 두 번째 시설로, 경쟁사들의 초거대 인프라 확장에 대응하기 위한 행보
- 암스트롱 CFO는 자금 조달과 계약 등 핵심 재무 전략을 총괄하며 xAI와 소셜미디어 X의 재무 운영을 맡을 예정

4. 인도 타타, 미국의 비자 정책·AI 전환으로 사상 최대 2만명 해고



- 인도 최대 IT 기업 TCS가 생성 AI 확산과 미·인도 관계 악화 속에서 창사 이래 최대 규모인 1만9755명을 감축
- 전체 인력은 3.2% 줄어 60만 명 아래로 내려갔으며, 감원은 주로 기술 역량이 맞지 않는 중·고위직을 대상으로 진행
- 회사 측은 내년 3월까지 전체 인력의 2%를 줄일 계획이며, 이는 산업 구조 전환에 대응하기 위한 조치라고 밝힘
- TCS는 H-1B 비자 의존도를 낮추기 위해 미국 내 인력 비중을 늘리는 전략을 추진 중

3. 하버드·MIT, 2시간 연속 구동 가능한 양자컴퓨터 개발...수년 내 무기한 작동 가능성



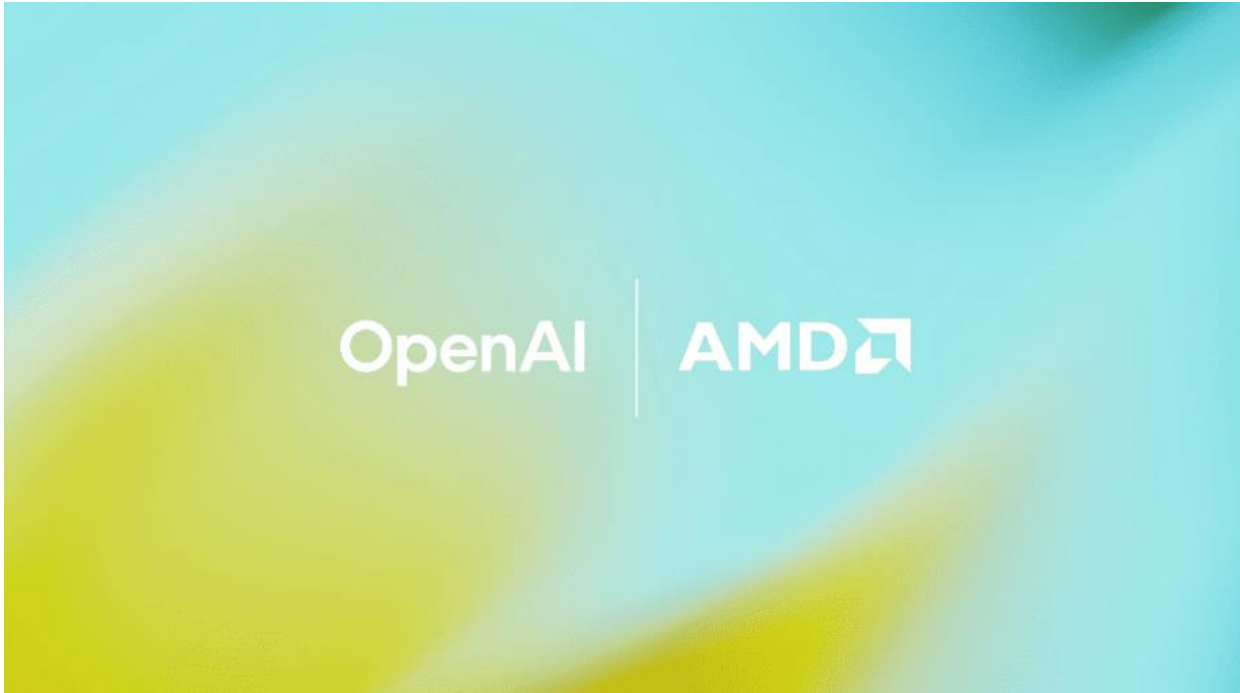
- 하버드대와 MIT 연구진이 **재부팅 없이 2시간 이상 연속 작동하는 최초의 양자컴퓨터**를 구현해 무기한 구동 가능성에 한 걸음 다가감
- 이는 수밀리초~10초 수준에 그치던 기존 한계를 크게 뛰어넘는 성과로, 큐비트 손실 문제를 해결하기 위해 광학 격자 컨베이어 벨트와 광학 핀셋 기술을 도입
- 이 시스템은 초당 30만 개의 원자를 주입해 손실된 큐비트를 즉시 교체하며 정보 손실을 방지
- 연구진은 3년 내 무기한 작동이 가능한 양자컴퓨터 실현이 가능할 것으로 내다 봄

2. 노벨 물리학상에 양자 컴퓨팅 과학자 3명 선정



- 올해 노벨 물리학상은 1980년대 양자 연구로 차세대 컴퓨팅의 기반을 마련한 **존 클라크, 미셸 드보레, 존 마티니스**에게 돌아감
- 이들은 전기 회로에서 **양자 터널링과 에너지 양자화**를 발견해 현대 전자기기와 초전도 큐비트 기술의 토대를 세움
- 특히 조셉슨 접합 회로를 통해 미시 세계의 양자 현상이 실제 회로에서도 재현될 수 있음을 입증
- 이번 수상은 양자역학 탄생 100주년을 기념하며, 양자컴퓨터 개발로 이어진 결정적 발견이라는 평가를 받음

1. AMD, 오픈AI와 AI 칩 공급 계약 체결...지분 10% 인수권도 부여



- 오픈AI가 AMD와 6GW 규모의 다년 GPU 공급 계약을 체결하고, AMD 지분 최대 10%를 인수할 수 있는 옵션을 확보
- AMD의 차세대 GPU 'MI450'가 투입될 첫 1GW 데이터센터는 2026년 하반기부터 건설될 예정
- 이번 계약으로 AMD는 수백억 달러 매출과 최대 1000억 달러 이상의 신규 수익을 기대하며, 엔비디아 독점 구도에 도전장을 던짐
- 발표 이후 AMD 주가는 25% 이상 급등했고, 오픈AI는 2033년까지 250GW 컴퓨팅 용량 확보를 목표로 하고 있음