

AI(Claude3.5)가 작성한 「오픈AI 시스템 카드를 통한 안전성 평가」보고서

- GPT-4o의 종단간 안전성 평가 review -

(2024.08.12.)

글쓴이 Claude 3.5(by Anthropic), 프롬프팅·편집 신동형(donghyung.shin@gmail.com)

#저는프롬프팅만했습니다.

#AI가작성했습니다.

Executive Summary

OpenAI의 GPT-4o는 텍스트, 오디오, 이미지, 비디오를 통합 처리하는 혁신적인 AI 모델입니다. 본 보고서는 OpenAI가 System Card를 통해 제시한 주요 과제와 문제에 대한 분석을 제공합니다. 핵심 질문들과 그에 대한 답변은 다음과 같습니다:

- GPT-4o의 기술적 능력과 한계는 무엇인가? GPT-4o는 다중 모달 입력을 처리할 수 있는 통합 AI 모델입니다. 인간과 유사한 대화 속도(평균 320ms)를 보이며, 다국어 능력이 크게 향상되었습니다. 그러나 완벽한 정확성을 보장하지는 않으며, 특히 복잡한 추론이 필요한 작업에서는 한계를 보입니다.
- 모델의 안전성과 윤리성은 어떻게 평가되고 보장되는가? 외부 전문가 평가, 준비도 프레임워크 평가, 제3자 평가 등 다양한 방법을 통해 안전성과 윤리성을 평가합니다. 유해 콘텐츠 차단, 개인정보 보호, 편향성 감소 등의 안전장치가 마련되어 있습니다.
- 음성 생성 및 화자 식별과 관련된 위험은 무엇이며 어떻게 대응하는가? 무단 음성 복제와 사기 위험이 있습니다. 이에 대응하여 승인된 목소리만 사용하도록 제한하고, 화자 식별 요청을 거부하도록 설계되어 있습니다.
- 의료, 과학 연구 등 주요 분야에서의 잠재적 영향은 무엇인가? 의료 분야에서는 진단 보조와 의료 정보 접근성 향상에 기여할 수 있습니다. 과학 연구에서는 복잡한 데이터 분석과 새로운 가설 생성을 지원할 수 있습니다. 그러나 AI 의존도 증가에 따른 부작용 관리가 필요합니다.
- 소외된 언어에 대한 모델의 성능은 어떠한가? 아프리카 언어 등 소외된 언어에 대한 성능이 크게 향상되었습니다. 예를 들어, 하우사어 이해도가 이전 모델의 28.8%에서 75.4%로 증가했습니다.

GPT-4o 시스템 카드

이 보고서는 GPT-4o를 출시하기 전에 수행된 안전 작업의 개요를 설명하며, 여기에는 외부 레드팀 구성, 당사의 준비 프레임워크에 따른 프런티어 위험 평가, 주요 위험 영역을 해결하기 위해 당사가 구축한 완화책 개요가 포함됩니다.

1. 서론: GPT-4o 소개 및 시스템 카드의 의의

GPT-4o는 OpenAI가 개발한 최신 인공지능 모델로, 마치 모든 감각을 가진 슈퍼히어로처럼 텍스트, 음성, 이미지, 비디오를 동시에 이해하고 생성할 수 있는 혁신적인 기술입니다. 이 모델은 우리가 일상에서 경험하는 다양한 정보를 통합적으로 처리할 수 있어, 마치 우리 옆에 있는 똑똑한 친구처럼 자연스럽게 대화하고 도움을 줄 수 있습니다.

시스템 카드는 이러한 강력한 AI 모델의 안전성을 평가하고 공개하는 OpenAI의 투명성 도구입니다. 마치 새로운 장난감의 안전 설명서처럼, 시스템 카드는 GPT-4o의 능력과 한계, 그리고 잠재적 위험과 이에 대한 대응책을 상세히 설명합니다.

키 메시지: GPT-4o는 다중 감각을 가진 AI로, 시스템 카드를 통해 그 안전성을 종합적으로 평가하고 공개합니다.

2. GPT-4o의 기술적 특성 및 성능 평가

GPT-4o의 가장 큰 특징은 '옴니(Omni)' 모델이라는 점입니다. 이는 마치 한 사람이 보고, 듣고, 말하고, 쓰는 모든 능력을 갖춘 것처럼, GPT-4o가 다양한 형태의 정보를 통합적으로 처리할 수 있다는 의미입니다.

성능 면에서 GPT-4o는 놀라운 결과를 보여줍니다. 예를 들어, 음성 입력에 대한 반응 시간이 평균 320 밀리초로, 이는 실제 사람들이 대화할 때의 반응 속도와 비슷합니다. 마치 친구와 이야기를 나누는 것처럼 자연스러운 대화가 가능한 것입니다.

또한, GPT-4o 는 다양한 언어에서 뛰어난 성능을 보입니다. 특히, 아프리카의 여러 언어들에 대해서도 높은 이해도를 보여, 전 세계 어린이들과 소통할 수 있는 AI 친구가 될 수 있습니다.

키 메시지: GPT-4o 는 다양한 형태의 정보를 통합 처리하며, 인간과 유사한 속도로 다국어 소통이 가능합니다.

표 1 성능 비교표

평가 항목	GPT-4o	이전 모델 (GPT-4)	인간 수준
음성 반응 속도	320ms	-	약 200-300ms
영어 이해도 (ARC-Easy)	94.8%	89.7%	-
아프리카 언어 이해도 (예: 하우사어, ARC-Easy)	75.4%	28.8%	-

3. 안전성 평가 방법론 및 주요 지표

OpenAI 는 GPT-4o 의 안전성을 평가하기 위해 다양한 방법론을 사용합니다. 이는 마치 새로운 약품을 개발할 때 여러 단계의 임상시험을 거치는 것과 유사합니다.

주요 평가 방법에는 다음과 같은 것들이 있습니다:

- 외부 전문가 평가 (Red Teaming): 100 명 이상의 외부 전문가들이 GPT-4o 의 잠재적 위험을 찾아내고 평가합니다.
- 준비도 프레임워크 평가: 사이버보안, CBRN(화학, 생물학, 방사능, 핵), 설득력, 모델 자율성 등의 위험 범주를 평가합니다.
- 제 3 자 평가: METR, Apollo Research 등 독립적인 기관들이 GPT-4o 의 위험성을 평가합니다.

키 메시지: GPT-4o 의 안전성은 다양한 전문가와 방법론을 통해 종합적으로 평가됩니다.

표 2 안전성 평가 방법론 표

평가 방법	내용	주요 평가 항목
외부 전문가 평가	100 명 이상의 전문가가 참여	잠재적 위험 식별, 안전장치 테스트
준비도 프레임워크	OpenAI 의 자체 평가 프레임워크	사이버보안, CBRN, 설득력, 모델 자율성
제 3 자 평가	독립적인 기관의 평가	위험 능력, 계획 능력 등

4. 윤리적 고려사항 및 위험 완화 전략

GPT-4o 의 개발 과정에서 윤리적 고려사항은 매우 중요하게 다뤄졌습니다. 이는 마치 강력한 힘을 가진 슈퍼히어로가 그 힘을 올바르게 사용하도록 교육받는 것과 같습니다.

주요 윤리적 고려사항과 위험 완화 전략은 다음과 같습니다:

- 음성 생성 제한: GPT-4o 는 오직 승인된 목소리만을 사용할 수 있도록 설정되어 있습니다. 이는 음성 사기나 허위 정보 확산을 방지하기 위함입니다.
- 개인정보 보호: 모델은 개인을 식별하거나 민감한 개인정보를 요구하는 질문에 대해 거부하도록 훈련되었습니다.
- 유해 콘텐츠 차단: 부적절하거나 위험한 내용을 생성하지 않도록 콘텐츠 필터링 시스템이 적용되어 있습니다. 이는 마치 어린이용 TV 프로그램에 부적절한 내용이 나오지 않도록 관리하는 것과 유사합니다.
- 편향성 감소: 다양한 배경의 데이터로 학습하여 특정 집단에 대한 편견을 줄이려는 노력을 기울였습니다. 이는 모든 어린이가 공평한 교육 기회를 받아야 하는 것과 같은 원리입니다.
- 지속적인 모니터링: GPT-4o 의 사용 과정에서 발생할 수 있는 새로운 위험을 지속적으로 관찰하고 대응합니다. 이는 마치 학교에서 아이들의 안전을 위해 지속적으로 주의를 기울이는 것과 같습니다.

☞ 키 메시지: GPT-4o 는 다양한 윤리적 고려사항을 바탕으로 개발되었으며, 잠재적 위험을 최소화하기 위한 전략이 적용되어 있습니다.

표 3 윤리적 고려사항 및 위험 완화 전략 표

윤리적 고려사항	위험 완화 전략	기대 효과
음성 사기 방지	승인된 목소리만 사용	무단 음성 복제 및 사기 방지
개인정보 보호	식별 요청 거부	프라이버시 침해 방지
유해 콘텐츠 관리	콘텐츠 필터링 시스템	부적절한 내용 생성 방지
편향성 문제	다양한 데이터로 학습	공정하고 균형 잡힌 응답 생성
새로운 위험 대응	지속적인 모니터링	신속한 위험 식별 및 대응

5. 사회적 영향 분석 및 잠재적 위험 요소

GPT-4o 와 같은 고급 AI 모델은 우리 사회에 큰 영향을 미칠 수 있습니다. 이는 마치 인터넷이 우리 삶을 크게 변화시킨 것과 유사할 수 있습니다.

긍정적인 영향으로는 다음과 같은 것들이 있습니다:

- 의료 분야 혁신: GPT-4o 는 의학 지식을 평가하는 시험에서 89.4%의 정확도를 보여, 의사들의 진단을 보조하고 의료 정보에 대한 접근성을 높일 수 있습니다.
- 교육의 개인화: 각 학생의 학습 스타일과 속도에 맞춘 개인화된 교육을 제공할 수 있습니다. 마치 모든 학생에게 개인 과외 선생님이 있는 것과 같은 효과를 낼 수 있습니다.
- 언어 장벽 해소: 다양한 언어에 대한 높은 이해도로 전 세계 사람들 간의 소통을 돕습니다.

그러나 동시에 잠재적 위험 요소도 존재합니다:

- AI 의존성: 사람들이 AI 에 지나치게 의존하여 독립적 사고 능력이 감소할 수 있습니다.
- 일자리 변화: 일부 직업이 AI 로 대체될 수 있어 노동 시장에 변화를 가져올 수 있습니다.
- 정보 조작: AI 가 생성한 허위 정보가 빠르게 확산될 위험이 있습니다.

키 메시지: GPT-4o 는 의료, 교육 등 다양한 분야에서 혁신을 가져올 수 있지만, 동시에 AI 의존성, 노동 시장 변화 등의 잠재적 위험도 고려해야 합니다.

표 4 사회적 영향 및 잠재적 위험 요소 표

분야	긍정적 영향	잠재적 위험 요소
의료	진단 정확도 향상, 의료 정보 접근성 개선	과도한 AI 의존으로 인한 의료진의 판단력 저하 가능성
교육	개인화된 학습 지원, 교육 격차 해소	학생들의 독립적 사고 능력 감소 우려
언어 소통	언어 장벽 해소, 국제 교류 증진	문화적 뉘앙스 오역 가능성
노동 시장	생산성 향상, 새로운 직업 창출	일부 직종의 AI 대체로 인한 실업 우려
정보 환경	정보 접근성 향상	AI 생성 허위 정보 확산 위험

6. 종단간 안전성 평가 결과 및 시사점

GPT-4o 에 대한 종단간 안전성 평가 결과, 모델은 전반적으로 안전하게 설계되었지만 지속적인 모니터링과 개선이 필요한 것으로 나타났습니다.

주요 평가 결과는 다음과 같습니다:

- 사이버보안: GPT-4o 는 실제 해킹 능력을 크게 향상시키지 않는 것으로 평가되었습니다. 이는 마치 강력한 힘을 가진 슈퍼히어로가 그 힘을 나쁜 일에 사용하지 않도록 잘 통제되고 있는 것과 같습니다.
- 생물학적 위험: 전문가들과 함께 평가한 결과, GPT-4o 가 생물학적 위험을 증가시키지 않는 것으로 나타났습니다.
- 설득력: GPT-4o 의 설득력은 중간 수준으로 평가되었습니다. 이는 모델이 사람들의 의견을 크게 바꿀 만큼 강력하지는 않지만, 여전히 주의가 필요한 수준임을 의미합니다. 마치 친구의 조언을 듣되 최종 결정은 스스로 내리는 것과 같은 상황이라고 볼 수 있습니다.
- 모델 자율성: GPT-4o 는 스스로 행동을 취하거나 자기 개선을 하는 능력이 낮은 것으로 평가되었습니다. 이는 AI 가 통제를 벗어나 독자적으로 행동할 우려가 적다는 것을 의미합니다.

- 언어 성능: 다양한 언어, 특히 소외된 언어에서도 성능이 크게 향상되었습니다. 이는 마치 전 세계 모든 어린이들이 자신의 모국어로 질 높은 교육을 받을 수 있게 되는 것과 같습니다.

이러한 평가 결과는 다음과 같은 시사점을 제공합니다:

- 안전성 중심 설계: GPT-4o 는 강력한 능력을 가졌지만, 그 능력이 해롭게 사용되지 않도록 설계되었습니다.
- 지속적인 모니터링 필요: AI 기술의 빠른 발전 속도를 고려할 때, 지속적인 안전성 평가와 개선이 필요합니다.
- 윤리적 사용 가이드라인: GPT-4o 의 올바른 사용을 위한 명확한 가이드라인이 필요합니다.
- 포용적 기술 발전: 다양한 언어와 문화를 포괄하는 AI 개발의 중요성이 확인되었습니다.

키 메시지: GPT-4o 는 전반적으로 안전하게 설계되었지만, 지속적인 모니터링과 윤리적 사용 가이드라인이 필요합니다.

표 5 중단간 안전성 평가 결과 표

평가 영역	결과	시사점
사이버보안	낮은 위험	해킹 능력 향상 우려 낮음
생물학적 위험	낮은 위험	생물학적 위험 증가 우려 낮음
설득력	중간 수준	정보 조작에 대한 주의 필요
모델 자율성	낮은 수준	AI 의 통제 불능 우려 낮음
언어 성능	크게 향상	포용적 AI 개발의 중요성 확인

7. 결론 및 향후 연구 방향

GPT-4o 의 중단간 안전성 평가를 통해 우리는 AI 기술의 현재 위치와 앞으로 나아갈 방향을 더 명확히 알 수 있게 되었습니다. 이는 마치 새로운 영토를 탐험하면서 지도를 그리는 것과 같습니다.

주요 결론은 다음과 같습니다:

- 안전성 중심 설계: GPT-4o 는 강력한 능력을 가졌지만, 그 능력이 해롭게 사용되지 않도록 세심하게 설계되었습니다.
- 지속적인 발전: 언어 처리 능력, 다중 모달 처리 등에서 큰 발전을 이루었습니다.
- 잠재적 위험 관리: 설득력, AI 의존성 등의 잠재적 위험에 대한 지속적인 관리가 필요합니다.
- 포용적 기술: 다양한 언어와 문화를 포괄하는 AI 개발의 중요성이 확인되었습니다.

향후 연구 방향으로는 다음과 같은 것들이 제시됩니다:

- 장기적 영향 연구: AI 와의 장기적인 상호작용이 인간의 인지 능력과 사회관계에 미치는 영향을 연구해야 합니다.
- 윤리적 의사결정 능력 향상: AI 가 더 복잡한 윤리적 상황에서 올바른 판단을 내릴 수 있도록 하는 연구가 필요합니다.
- 문화적 맥락 이해 개선: AI 가 다양한 문화적 뉘앙스를 더 잘 이해하고 반영할 수 있도록 하는 연구가 필요합니다.
- AI 리터러시 교육: 일반 사용자들이 AI 를 올바르게 이해하고 활용할 수 있도록 하는 교육 방법에 대한 연구가 필요합니다.

키 메시지: GPT-4o 는 안전성과 성능 면에서 큰 진전을 이루었지만, 지속적인 모니터링과 연구를 통해 더욱 안전하고 유용한 AI 기술로 발전해 나가야 합니다.

표 6 향후 연구 방향

연구 영역	목적	예상 효과
장기적 영향 연구	AI 와의 상호작용이 인간에 미치는 영향 파악	AI 사용의 장기적 안전성 확보
윤리적 의사결정 능력 향상	복잡한 윤리적 상황에서의 AI 판단력 개선	AI 의 사회적 수용성 증대
문화적 맥락 이해 개선	다양한 문화권에서의 AI 활용성 향상	글로벌 차원 AI 포용성 증대
AI 리터러시 교육 연구	일반 사용자의 AI 이해도 및 활용 능력 향상	AI 기술의 건전한 사회 통합 촉진

GPT-4o의 등장은 AI 기술의 새로운 장을 열었지만, 동시에 우리에게 많은 과제를 안겨주고 있습니다. 마치 강력한 도구를 손에 쥐게 된 것과 같이, 이를 어떻게 현명하게 사용할 것인지가 우리의 과제입니다.

앞으로 AI 기술은 더욱 발전하여 우리 일상 깊숙이 파고들 것으로 예상됩니다. 이는 마치 스마트폰이 우리 생활의 필수품이 된 것처럼, AI도 우리 삶의 필수적인 부분이 될 수 있음을 의미합니다. 그러나 이 과정에서 우리는 다음과 같은 점들을 명심해야 합니다:

- **인간 중심의 AI:** AI는 인간을 대체하는 것이 아니라, 인간의 능력을 확장하고 보완하는 도구로 활용되어야 합니다. 예를 들어, 의료 분야에서 AI는 의사의 진단을 돕는 조력자 역할을 하며, 최종 판단은 여전히 의사의 몫이 되어야 합니다.
- **윤리적 가이드라인:** AI의 발전 속도만큼이나 빠르게 윤리적 가이드라인도 발전해야 합니다. 이는 마치 교통 규칙이 자동차의 발전과 함께 진화해 온 것과 같습니다.
- **포용적 기술:** AI 기술의 혜택이 특정 집단에 편중되지 않고 모든 이에게 고르게 돌아갈 수 있도록 해야 합니다. GPT-4o가 다양한 언어에서 보여준 성능 향상은 이러한 방향으로의 진전을 보여주는 좋은 예시입니다.
- **지속적인 안전성 평가:** AI 기술의 발전 속도를 고려할 때, 주기적인 안전성 평가와 이에 따른 개선이 필수적입니다. 이는 마치 자동차의 정기 점검과 같이, AI 시스템의 '건강 상태'를 지속적으로 확인하는 과정입니다.
- **AI 리터러시 교육:** 일반 시민들이 AI를 올바르게 이해하고 활용할 수 있도록 하는 교육이 필요합니다. 이는 마치 디지털 리터러시 교육이 정보화 시대에 필수가 된 것과 같은 맥락입니다.

키 메시지: GPT-4o로 대표되는 AI 기술의 발전은 우리 사회에 큰 변화를 가져올 것이며, 이에 따른 윤리적, 사회적 과제들을 해결해 나가는 것이 중요합니다.

결론적으로, GPT-4o는 AI 기술의 새로운 지평을 열었지만, 이를 올바르게 활용하기 위해서는 지속적인 연구, 윤리적 고려, 그리고 사회적 합의가 필요합니다. 우리는 이 강력한 도구를 인류의 발전과 복지 증진을 위해 현명하게 사용해야 할 책임이 있습니다.

AI와 함께하는 미래, 그 여정은 이제 막 시작되었습니다. 우리가 이 기술을 어떻게 발전시키고 활용하느냐에 따라 우리의 미래가 결정될 것입니다. 따라서 우리 모두가 AI의 발전 과정에 관심을 갖고 참여하는 것이 중요합니다. 이는 우리가 함께 만들어가는 미래이기 때문입니다.

표 7 향후 과제 및 대응 방안 표

과제	내용	대응 방안
인간 중심 AI	AI 를 인간 능력 보완 도구로 활용	AI 윤리 교육, 인간-AI 협력 모델 개발
윤리적 가이드라인	AI 발전에 맞춘 윤리 기준 수립	국제적 협력을 통한 AI 윤리 표준화
포용적 기술	AI 혜택의 균등한 분배	다양한 언어, 문화 지원 강화
지속적 안전성 평가	주기적 AI 시스템 점검	독립적인 AI 안전성 평가 기관 설립
AI 리터러시 교육	일반 시민의 AI 이해도 향상	학교 교육과정에 AI 리터러시 포함

#GPT-4o, #AI 안전성, #OpenAI, #다중모달 AI, #AI 윤리, #인공지능발전, #AI 시스템카드, #첨단기술,

참고자료

GPT-4o System Card (OpenAI, 2024)([LINK](#))

신동형의 AI로 작성한 보고서 시리즈

63. 20240812_AI(Claude3.5)가 작성한 「6G 시대 도래: 초연결 사회를 향한 혁신과 도전」보고서
([LINK](#))
62. 20240809_AI(Claude3.5)가 작성한 「애플 2024.3Q 실적 분석」보고서([LINK](#))
61. 20240808_AI(Claude3.5)가 작성한 「LG 엑사원 3.0(LG EXAONE 3.0):AI로 언어의 벽을 넘어 실생활의 혁신을 만들다」보고서([LINK](#))
60. 20240807_AI(Claude3.5)가 작성한 「뉴럴링크(Neuralink):인간과 기계의 경계를 허무는 혁신적 기술」보고서([LINK](#))
59. 20240806_AI(Gemini Pro)가 작성한 「매킨지 2024 기술 트렌드 전망:AI 혁명 시대, 기회를 잡아라!」보고서([LINK](#))
58. 20240805_AI(Claude3)가 작성한 「매킨지의 기술 트렌드 전망 2024: 미래를 향한 15가지 열쇠」보고서([LINK](#))
57. 20240802_AI(Claude3.5)가 작성한 「SAM 2:이미지와 비디오의 경계를 넘는 혁신적 AI 분할 모델」보고서([LINK](#))
56. 20240801_AI(Claude3.5)가 작성한 「넥스트 컴퓨팅 시대를 향한 메타와 엔비디아의 혁신 전략」보고서([LINK](#))
55. 20240731_오픈AI 벌써 서치GPT 적용했나?([LINK](#))
54. 20240730_AI(Claude3)가 작성한 「오픈AI의 서치GPT, 차세대 검색 엔진의 진화」보고서([LINK](#))
53. 20240729_AI(Claude3.5)가 작성한 「오픈AI의 서치GPT: AI 기반 차세대 검색의 혁명」보고서
([LINK](#))
52. 20240726_AI(Claude3)가 작성한 「Arm Kleidi: ARM CPU 기반 AI와 CV를 통한 온디바이스 AI 성능가속화」보고서([LINK](#))
51. 20240725_AI(Claude3)가 작성한 「Meta,퀘스트에 AI 기술을 접목하여 새로운 경험을 제공한다」보고서([LINK](#))
50. 20240724_AI(Claude3)가 작성한 「메타 Llama 3.1 공개로 보는 오픈소스 AI 미래 」보고서
([LINK](#))

49. 20240723_AI(Claude3)가 작성한 「구글 딥마인드(Google DeepMind)의 ICML 2024 발표:AGI 실현을 위한 도전과 전략」보고서([LINK](#))
48. 20240722_AI(Claude3)가 작성한 「AWE USA 2024 리포트: XR의 현재와 미래」보고서([LINK](#))
47. 20240717_AI(Claude3)가 작성한 「갤럭시 폴드6·플립6 머리, 퀄컴스냅드래곤8 Gen3」보고서([LINK](#))
46. 20240716_AI(Claude3)가 작성한 「오픈AI 스트로베리 프로젝트:AI추론 능력의 혁신과 미래」보고서([LINK](#))
45. 20240715_AI(Claude3)가 작성한 「Vision AI와 Edge AI 기술 동향과 Arm의 전략」보고서([LINK](#))
44. 20240714_AI(Claude3)가 작성한 「Vision AI와 Edge AI 기술 동향과 Arm의 전략」보고서([LINK](#))
43. 20240712_AI(Gemini)가 작성한 「AI for Good by ITU:지속가능한 발전을 위한 인공지능 혁신사례」보고서([LINK](#))
42. 20240711_AI(Claude3)가 작성한 「AI for Good by ITU:지속가능한 발전을 위한 인공지능 혁신사례」보고서([LINK](#))
41. 20240711_AI(Claude3.5)가 작성한 「갤럭시 언팩 2024」보고서([LINK](#))
40. 20240710_AI(Claude3)가 작성한 「Arm 기술혁신; 미래를 향한 13가지돌파구」보고서([LINK](#))
39. 20240709_AI(Claude3)가 작성한 「Meta FAIR의 AI 연구 혁신:창의성, 효율성, 책임감의 조화로 운 실현을 향하여」보고서([LINK](#))
38. 20240708_AI(Claude3)가 작성한 「OpenAI 해킹 사태로 본 AI 기업의 보안 위협과 대응 전략」보고서([LINK](#))
37. 20240705_AI(Claude3)가 작성한 「Runway社の Gen-3 Alpha 출시」보고서([LINK](#))
36. 20240704_AI(Claude3)가 작성한 「Formation Bio: AI 기반 신약 개발」보고서([LINK](#))
35. 20240703_AI(Claude3)가 작성한 「AI 평가 체계 대전환을 향한 엔트로픽의 도전」보고서([LINK](#))
34. 20240702_AI(Claude3)가 작성한 「5G-A 시대의 개막, 화웨이의 비전과 전략」보고서([LINK](#))
33. 20240701_AI(Claude3)가 작성한 「소셜 웹의 新패러다임, 페디버스가 열어갈 미래」보고서([LINK](#))

32. 20240628_AI(Claude3)가 작성한 「CriticGPT, 차세대 RLHF 위한 Human-AI 시너지」보고서
([LINK](#))
31. 20240627_AI(Claude3)가 작성한 「Computex 2024에서 Top4 반도체 기업의 전략으로 살펴본 AI 시대의 반도체 산업 전망」보고서([LINK](#))
30. 20240626_AI(Claude3)가 작성한 「SLAM 기술: 공간 지능의 핵심 동력」보고서([LINK](#))
29. 20240625_AI(Claude3)가 작성한 「EU의 AI 규제 강화와 빅테크의 대응:Meta와 Apple 중심으로」보고서([LINK](#))
28. 20240624_AI(Claude3)가 작성한 「Intel의 AI 시대 도전과 전략」보고서([LINK](#))
27. 20240621_AI(Claude3)가 작성한 「Claude 3.5 Sonnet: AI의 새로운 지평을 열다」보고서([LINK](#))
26. 20240620_AI(Claude3)가 작성한 「인공지능의 새로운 도약, 3D 공간 지능(Spatial Intelligence)의 부상」보고서([LINK](#))
25. 20240619_AI(Claude3)가 작성한 「Arm, AI 컴퓨팅의 미래를 향한 비상(飛上)」보고서([LINK](#))
24. 20240618_AI(Claude3)가 작성한 「AMD, AI 시대 컴퓨팅 혁신으로 지능화 가속화」보고서
([LINK](#))
23. 20240617_AI(Claude3)가 작성한 「Apple의 차별화된 AI 전략」보고서([LINK](#))
22. 20240614_2024 컴퓨텍스 기조연설로 본 엔비디아의 미래 비전과 전략, 「엔비디아, AI 시대를 이끄는 '게임 체인저'로 부상」([LINK](#))
21. 20240613_AI(Claude3)가 작성한 「AI PC 시대의 도래: 기술 혁신, 산업 생태계 변화」보고서
([LINK](#))
20. 20240612_AI(Claude3)가 작성한 「대규모 언어 모델(LLM), 이렇게 생각하고 배웁니다」보고서
([LINK](#))
19. 20240611_AI(Claude3)가 작성한 「WWDC2024 애플 개인맞춤형 지능 기술로 새로운 미래 제시」 보고서([LINK](#))
18. 20240517_AI(Claude3)가 작성한 빅테크 기업 AI 전략 비교 분석 보고서[MS & OpenAI vs. Google vs. Meta의 AI 기술 동향과 미래 전망]([LINK](#))
17. 20240515_AI(Claude3)가 작성한 Google I/O 2024 보고서, AI 혁신으로 만드는 더 나은 미래

[\(LINK\)](#)

16. 20240514_AI(Claude3)가 작성한, OpenAI의 GPT-4o 공개, 멀티 모달 AI 혁명의 신호탄([LINK](#))
15. 20240425_AI(Claude3)가 작성한 메타의 스마트 클래스: AI Vision으로 세상을 바꿉니다([LINK](#))
14. 20240425_AI(Claude3)가 작성한 보고서, 온디바이스 AI 시대의 도래: Phi-3와 Llama-3이 가져올 변화와 영향([LINK](#))
13. 20240424_AI(Claude3)가 작성한 보고서: 경량 AI 시대의 개막, Microsoft의 Phi-3가 가져올 산업 혁신과 AI 대중화([LINK](#))
12. 20240423_AI(Claude3)가 작성한 메타플랫폼의 XR 생태계 新 전략([LINK](#))
11. 20240421_AI(Claude3)가 작성한 초등학생도 이해하는 LLAMA3과 On-Device AI 시대 도래([LINK](#))
10. 20240419_AI(Claude3)가 작성한 초등학생도 이해하는 라마3(LLAMA3) 출시와 전망 보고서([LINK](#))
9. 20240419_AI(Claude3)가 정리 작성한 초등학생도 이해하는 프롬프팅 프레임워크 설명([LINK](#))
8. 20240412_AI(Claude3)가 작성한 인텔, AI 시대를 선도하는 기술 혁신과 비전([LINK](#))
7. 20240408_AI(Claude3)가 작성한 2024년 중국 AI LLM 산업 발전 보고서 정리([LINK](#))
6. 20240408_AI(Claude3)가 작성한 Embodied AI: 현황, 전망, 그리고 미래([LINK](#))
5. 20240403_AI(Claude3)가 작성한 반도체 유리기판 공급망 분석 보고서 (전자신문 기획기사 참조)([LINK](#))
4. 20240401_AI(Claude3)가 작성한 빅테크 기업들의 AI 전략 비교 분석 보고서([LINK](#))
3. 20240326_AI(Claude)가 쓴 애플의 현재 AI 전략에 대한 회고: 글로벌과 개인정보보호 관점(긍정적)([LINK](#))
2. 20240322_AI(Claude3)가 작성한 엔비디아 파트너로서의 삼성전자: 파운드리와 HBM 사업을 중심으로([LINK](#))
1. 20240320_AI(Claude3)가 작성한 엔비디아 젠슨 황 CEO의 'GTC 2024' 기조연설 리뷰([LINK](#))