

구글 이어 메타도...챗GPT 대항마 ‘라마’ 공개...더 가벼운 AI 언어 모델

원문:

<https://www.btcc.com/ko-KR/academy/financial-investment/meta-chatgpt-anti-ai-language-model-rama-released>

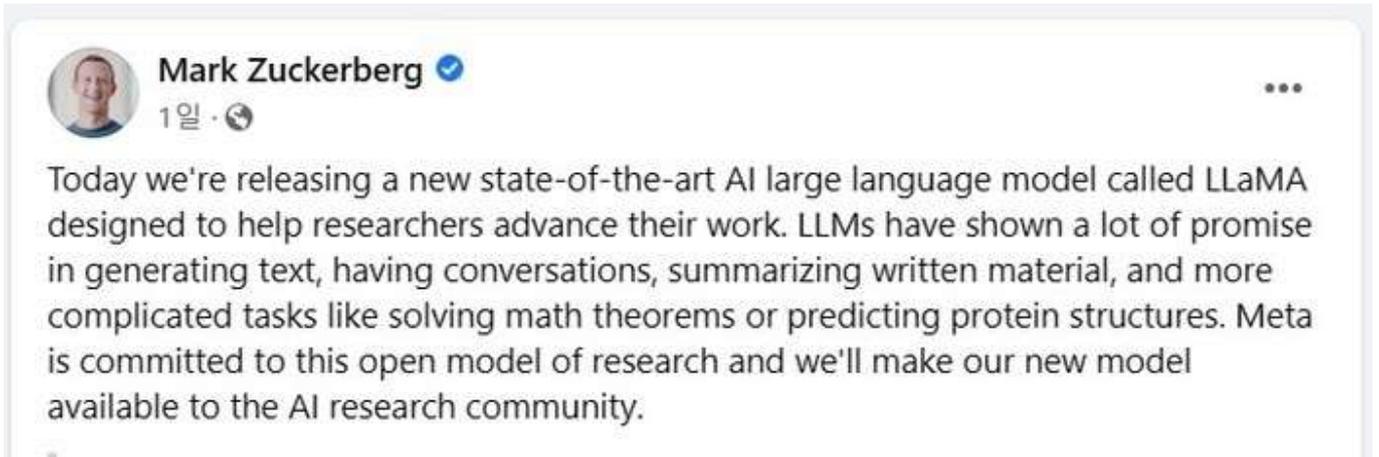
인공지능(AI) 연구 개발 단체인 **오픈AI**의 ‘**챗GPT**’ 출시가 촉발한 글로벌 빅테크 기업 간 생성(Generative) AI 기술 레이스가 오픈AI와 손잡은 **마이크로소프트(MS)**와 **구글**의 양강 구도에서 3파전 양상으로 확대된다. 구글에 이어 페이스북 운영사 메타(Meta)도 별도 AI 언어 모델(LLM)을 들고 뛰어들며 생성형 인공지능 기술을 둘러싼 글로벌 빅테크 기업간 경쟁이 뜨거워지고 있다.



메타, 챗GPT 대항마 라마 공개

25일(현지시간)마크 저커버그 메타 최고경영자(CEO)는 자신의 페이스북 계정을 통해, ‘라마(LLaMa.Large Language Model Meta AI)’라는 이름의 대규모 인공지능 언어 모델을 개발자들에게 공

개한다고 밝혔다.



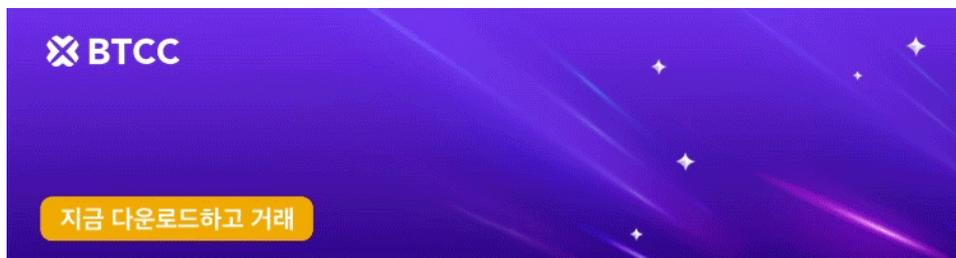
메타는 “AI 연구자들이 이같은 LLM을 연구할 수 있게 해 빠르게 변화하는 AI 분야 민주화에 기여할 것”이라며 라마 공개 이유를 설명했다. 마크 저커버그 메타 최고경영자(CEO)도 이날 자신의 페이스북에 “라마는 AI 연구자들의 업무 진출을 돕기 위해 설계됐다”고 밝혔다.

저커버그는 “LLM은 텍스트 생성과 대화, 문서 요약부터 수학 정리, 단백질 구조 예측과 같은 더 복잡한 작업에 이르기까지 가능성을 보여줬다. 라마는 연구자들의 업무를 돕기 위해 오픈소스 형태로 제공할 것”이라고 설명했다. LLM은 사람이 말하고 쓰는 것과 유사한 문장을 생성하는 인공지능 톨로, 챗GPT 같은 서비스를 구동하는 기반이 된다.

메타는 비상업적인 라이선스 아래 대학, 비영리단체(NGO), 기업 연구소들에 라마를 제공한다. 현재 연구자들로부터 신청을 받고 있다.

회사 측에 따르면 라마는 과학자들과 엔지니어들이 질문에 대답하고 문서를 요약하는 AI 애플리케이션들을 개발할 수 있도록 지원한다. 하지만 오픈AI GPT-3 같은 경쟁 언어 모델과는 다르다는 게 회사측 설명이다.

메타는 라마의 소스코드를 깃허브에 업데이트했다.



[안드로이드 버전 다운로드](#)

[iOS 버전 다운로드](#)

[신규 유저 한정 이벤트\(입금 및 거래 시 10,055USDT 보상\) <<<<](#)

메타의 “라마”에 대하여

라마는 챗GPT처럼 누구나 말을 걸 수 있는 시스템이 아니라 연구 도구로 제공된다.

메타가 오픈소스 커뮤니티 ‘깃허브(GitHub)’에 게재한 설명에 따르면 현재 제공되는 라마는 2022년 12월부터 2023년 2월까지 개발된 첫 버전(version 1)이며, 트랜스포머(transformer) 아키텍처 기반으로 제작된 ‘자기회귀(auto-regressive) 언어 모델’이다.

라마, 4가지 매개변수

라마는 다른 LLM보다 용량이 작아 여러 서비스에 효율적으로 적용할 수 있다. 라마는 70억개·130억개·330억개·650억개로 총 4가지로 파라미터(매개변수) 규모를 나눠 사용할 수 있다. 가장 큰 650억 파라미터도 오픈AI의 GPT-3.0(1750억개)이나 구글 딥마인드의 친칠라(700억개)보다 작다. 1750억개 파라미터를 가진 자사 LLM ‘OPT’보다도 훨씬 작다. GPU 등 컴퓨팅 파워를 적게 사용할 수 있어 운영 측면에서 훨씬 효율적이다.

		BoolQ	PIQA	SIQA	HellaSwag	WinoGrande	ARC-e	ARC-c	OBQA
GPT-3	175B	60.5	81.0	-	78.9	70.2	68.8	51.4	57.6
Gopher	280B	79.3	81.8	50.6	79.2	70.1	-	-	-
Chinchilla	70B	83.7	81.8	51.3	80.8	74.9	-	-	-
PaLM	62B	84.8	80.5	-	79.7	77.0	75.2	52.5	50.4
PaLM-cont	62B	83.9	81.4	-	80.6	77.0	-	-	-
PaLM	540B	88.0	82.3	-	83.4	81.1	76.6	53.0	53.4
LLaMA	7B	76.5	79.8	48.9	76.1	70.1	72.8	47.6	57.2
	13B	78.1	80.1	50.4	79.2	73.0	74.8	52.7	56.4
	33B	83.1	82.3	50.4	82.8	76.0	80.0	57.8	58.6
	65B	85.3	82.8	52.3	84.2	77.0	78.9	56.0	60.2

Table 3: Zero-shot performance on Common Sense Reasoning tasks.

?파라미터(매개변수)란 인공지능이 정확한 결과 값을 만들기 위해 계산에 고려해야 하는 다양한 변수(가중치·절편)로, 그 수가 클수록 추론 능력이 향상되고 더 정교한 모델이 된다. 다만, 파라미터는 적은 대신 사용 용량이 다른 모델의 10분의 1 수준이라 스마트폰 같은 개인 단말기에서도 활용도가 높아질 수 있다.

크기는 작지만, 떨어지지 않는 성능

메타는 연구 시뮬레이션 결과를 공개하며, 실제 성능이 챗GPT에 뒤지지 않는다고 강조했다. 자체 조사 결과, 매개변수 650억개 기준으로 라마의 수리추론 능력 등 대부분의 평가 지표에서 GPT-3.0을 앞서는 것으로 나타났다고 설명했다. 하정우 네이버클라우드 인공지능랩 소장은 페이스북에서 라마에 대해 “모델 크기는 콤팩트하지만 데이터를 많이 학습시켰다. 다른 모델과 비교해 경쟁력 있는 성능을 보인다”고 평가했다.

대규모 언어 모델은 성능 면에서 뛰어나지만 추론에 비용이 많이 들어간다. 라마에서 두 번 째로 크기가 작은 라마-13B는 대부분의 벤치마크 테스트에서 오픈AI GPT 모델보다 나은 것으로 나타났다고 메타측은 전했다.

연구자들은 라마를 활용해 다양한 서비스를 출시할 전망이다. GPT-3.5를 기반으로 한 챗GPT는 막대한 GPU를 소모해 하루 최소 10만 달러의 운영 비용이 드는 것으로 알려졌다. 이에 반해 라마는 용량이 월등히 작아 경제적이다. 메타는 신청자에 한해 라마 모델을 공유할 계획이다. 신청자는 연구 등 비상업 용도로만 라마를 활용할 수 있다.

MS VS 구글 VS 메타

가장 눈길을 끄는 대목은 라마를 다른 연구자나 엔지니어들이 이용할 수 있도록 ‘오픈소스’로 공개한다는 것이다. 이는 챗지피티 엔진이랄 수 있는 오픈에이아이의 ‘지피티-3.0’과 구글의 ‘람다’는 소스코드가 공개되지 않는 것과 차별되는 지점이다. 메타는 오픈소스 공개 방침을 밝히며 “인공지능 커뮤니티가 무엇을 배우고 구축할 수 있을지 협력하길 기대한다”고 설명했다.

메타는 ‘라마(LLaMA)’에 대해 다른 경쟁 모델들과 비교해 운영 비용이 적고, 또한 훨씬 컴퓨팅 파워가 소요된다는 장점을 강조했다. 고성능 반도체를 많이 쓰지 않아도 된다는 의미로 해석된다. 아울러 라틴어와 키릴 문자에 기반한 20개 언어로 구동된다.

MS와 구글은 최신 LLM의 원천 기술을 완전히 개방하지 않고 있지만 개방하더라도 그 규모는 대다수 개인·기업 연구자가 온전히 수정·가공해 다루기는 거의 불가능할 정도로 거대하다. MS 투자를 받은 오픈AI는 앞서 매개변수 1750억개를 다루는 언어모델 ‘GPT-3’와 이를 발전시킨 ‘GPT-3.5’를 기반으로 대화형 AI 서비스인 챗GPT를 만들었고, MS는 이를 업그레이드한 ‘프로메테우스’ 모델 기반 챗봇을 결합한 ‘새로운 Bing(New Bing)’을 출시했다. 구글도 매개변수 1370억개를 품은 ‘람다(LaMDA)’ 모델을 공개한 지 2년 만에 이를 활용한 챗봇 ‘바드(Bard)’를 소개하고 조만간 검색 서비스에 탑재한다고 예고했다.

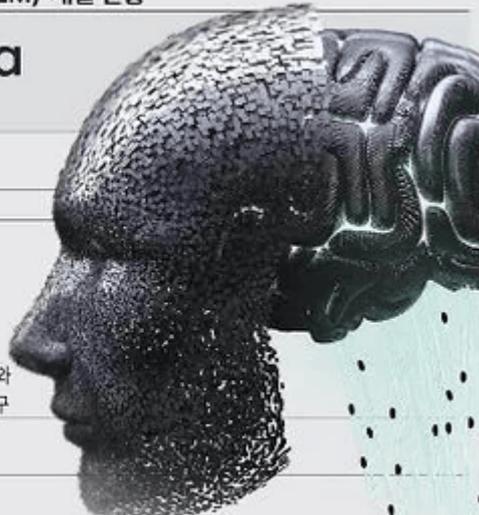
메타 AI 연구자들은 블로그를 통해 “LLM을 훈련하고 실행하는 데 필요한 자원 때문에 완전한 연구가 제한되고 이로 인해 LLM 작동 방식과 원인을 이해하는 연구자 능력이 제한돼 (AI의) 편향, 유해성, 오보(misinformation) 생성 가능성과 같은 문제를 줄이고 강건함(robustness)을 개선하는 작업의 발전을 저해한다”고 지적했다.

이어 “오용을 방지하기 위해 연구 목적에 초점을 맞춘 비상업적 라이선스(noncommercial license)로 우리 모델을 배포한다”며 “전 세계 학술 연구자, 정부, 시민사회, 학술 분야 단체와 산업계 연구실에 (사용권을) 부여하겠다”고 설명했다.

빅테크기업의 생성 AI 분야 '거대언어모델(LLM)' 개발 현황

	 Microsoft 마이크로소프트	 Google 구글	 Meta 메타
모델명	프로메테우스 (PROMETHEUS)	람다 (LAMDA)	라마 (LLAMA)
최초 발표 시기	2023년 2월	2021년 5월	2023년 2월
규모 (매개변수 개수)	오픈AI의 과거 모델인 'GPT-3(1750억개)'와 챗GPT 기반 언어모델인 'GPT-3.5'보다 클 것으로 추정	1370억개	70억개, 130억개, 330억개, 650억개 등 4종류
주 용도	자체 AI 챗봇, 콘텐츠 생성 서비스	자체 AI 챗봇, 콘텐츠 생성 서비스	시민사회, 학계 등 외부와 협업을 전제로 한 AI 연구
적용(예정) 서비스	검색 '새로운 Bing', 웹브라우저 '에지'	검색 '구글 서치'	
서비스 출시 시기	2023년 2월 (공개 시범 서비스 중)	2023년 (비공개 시범 서비스 중)	

[자료: 각사 발표]





[안드로이드 버전 다운로드](#)

[iOS 버전 다운로드](#)

[신규 유저 한정 이벤트\(입금 및 거래 시 10,055USDT 보상\) <<<<](#)

빅테크 기업의 생성AI 분야 경쟁

메타 가세로 구글과 마이크로소프트(MS)를 중심으로 전개되던 생성형 인공지능 경쟁이 더 뜨거워질 것으로 보인다. 엠에스가 챗지피티를 개발한 오픈에이아이와 투자 협약을 맺고 자사 검색 누리집 ‘빙(Bing)’에 챗지피티를 탑재하기로 하자, 구글은 새로운 인공지능 검색 시스템 ‘바드’를 공개하며 맞불을 놓았다.

오픈AI가 올해 안에 매개변수를 더 늘린 지피티-4.0을 공개할 예정인 가운데, 구글은 심각한 위기 상황에 대한 경고인 ‘코드 레드’를 발령하고 모든 업무의 초점을 인공지능 서비스 개발에 맞추고 있다.

바이두와 알리바바 등 중국 기업은 물론 네이버·카카오 등 국내 빅테크 기업들도 엘엘엠을 이용한 생성형 인공지능 개발에 박차를 가하고 있다. 시장조사업체 그랜드뷰리서치 조사 결과를 보면, 전 세계 생성형 인공지능 시장은 2022년 101억달러(12조4219억원)에서 2030년에는 1093억달러(134조4280억원)으로 커질 전망이다.

챗GPT 란?



지난 11월 30일 출시된 ChatGPT는 인간의 대화를 모방하고 사용자의 프롬프트에 대해 상세한 대답을 제공하는 능력으로 빠르게 인기를 끌었다. 이 제품은 출시 후 일주일 만에 100만 명의 등록 사용자를 확보했다.

ChatGPT는 훈련된 자연 언어 처리 모델이다. 오픈AI에서 만든 대규모 인공지능 모델인 'GPT-3.5' 언어 기술을 사용하여 이용자와 실시간으로 대화가 가능합니다. AI 챗봇의 일종이라고 보면 됩니다. 이 모델은 다양한 소스를 바탕으로 방대한 텍스트 데이터를 학습했다. 챗GPT(ChatGPT)에 가입하면 사용자는 다양한 분야들에 대해 질문할 수 있다.

GPT-3.5는 총 570기가바이트에 달하는 언어 데이터를 학습했다. 이 가운데는 뉴스 기사, 과학 논문, 소셜미디어 대화, 웹페이지, 책, 고객센터 데이터, 음성 기록, 이미지, 영화 대본, 위키피디아, 의료 텍스트, 법률 텍스트 등이 포함돼 있다.

아직 [ChatGPT](#)도 한계는 있습니다. 현재 공개된 챗GPT는 2021년까지의 데이터를 기반으로 학습했기 때문에 기본적으로 최신 정보를 가지고 있지 않다. 또한 많이 거론되는 것은 챗GPT의 기술적 한계와 한국어 학습 데이터 부족이다.

챗 GPT는 몇 가지 문제를 안고 있으면서도 AI의 획기적인 진보로 많은 사람들에게 찬사를 받고 있습니다. 그의 개발사 오픈AI도 점차 대중에 시선에 들어갔다.

관련페이지:

[Chat GPT 란? '구글의 시대 끝났다' 평가 나온 그는 할 수 있는 일? - BTCC](#)

[ChatGPT\(챗GPT\) 사용법 소개, 어떻게 가입합니까? - BTCC](#)

[MS, 챗GPT 기반 '빙' 검색 엔진 출시...새 Bing\(Bing\) 사용법은? - BTCC](#)

[구글 챗봇 '바드', 행사서 '오답'...알파벳 주가 8% 폭락](#)

[구글, AI 챗봇 '바드' 개발 중...챗gpt, '바드'에 질문하면 답은?](#)

[AI 열풍을 일으킨 챗GPT\(ChatGPT\)의 개발자, 오픈 AI 가 뭐니까? - BTCC](#)

[챗GPT에 질문...2030년 비트코인 전망은? - BTCC](#)

[챗GPT'\(ChatGPT\), 월 20 달러 유료 서비스 시작...단 무료 버전 사용가능 - BTCC](#)

[오픈AI, '챗GPT'\(ChatGPT\) 프로페셔널 공개...유료화 임박? - BTCC](#)

더 많은 내용은 [BTCC 뉴스](#) 및 [BTCC 아카데미](#)에 확인 할 수 있습니다.

BTCC 가입혜택

BTCC는 신규 유저들을 위해서 가입하면 보너스를 받을 수 있는 이벤트를 준비합니다. 지금 가입하면 **30일** 이내에 입금액 대비 **최대 3,500USDT**를 받을 수 있다. 또한 **200USDT**를 입금하시면 바로 **VIP 1**로 될 수 있습니다. **VIP**가 되면 거래 수수료 할인 또 다른 혜택을 누릴 수 있다.

BTCC 가입하고 보너스 받으세요

BTCC 거래소

- 12년 보안 무사고
- 암호화폐 라이선스를 획득한 믿음직한 거래소
- 전세계 규제 준수
- 세계 최저 수수료 **(0.03%)**
- 최대 150배 레버리지 제공
- 선물 거래량 기준 세계 5위
- 무료 모의 거래 가능
- 풍부한 유동성 및 오더북
- 토큰화 주식 및 암호화폐 출시
- 원화입금 지원
- 24x7 시간 한국어 고객센터 제공
- 최신 시장 동향 및 투자 교육 제공
- PC/모바일 거래 가능
- 다양한 이벤트