



이준원의 ENNM 리포트



Chat AI

Search...

98.0638

39.0654

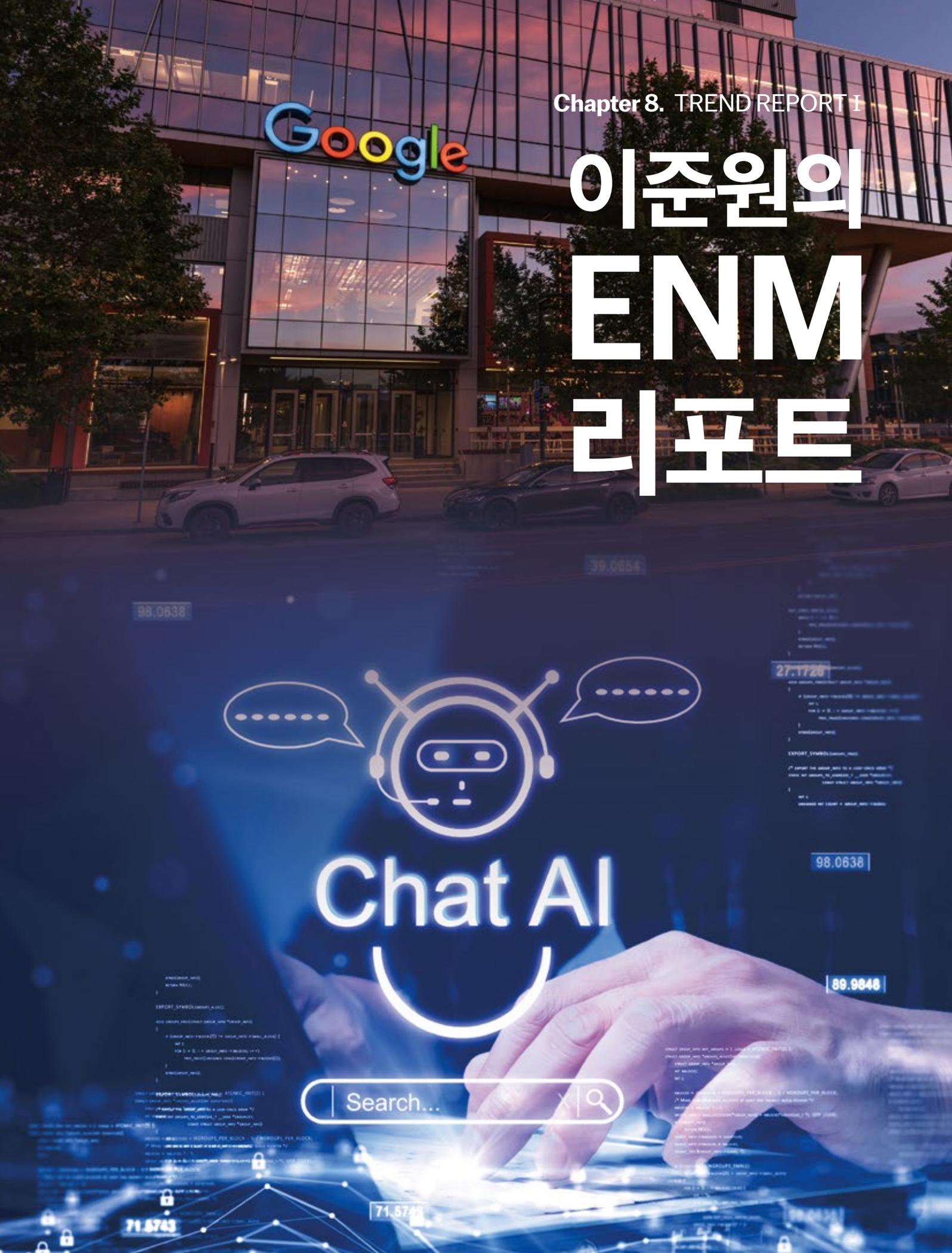
27.1726

98.0638

89.9848

71.5743

71.5743



OpenAI와 Google의 거대한 AI 전쟁 속 전투지점

Search... GPT-4o ?



글 | 이준원 한국외대 미디어커뮤니케이션연구소 언론학 박사 leejuw34@gmail.com

매년 늦은 봄에 찾아오는 구글의 개발자 컨퍼런스 Google I/O를 앞둔 지난 13일, 오픈AI가 새로운 AI 모델 'GPT-4o'를 공개했다. 두 기업 사이의 긴장감과 함께 사람들의 기대도 커지는 가운데 두 기업의 경쟁이 어떤 모습을 보이고 있는지 살펴보았다.

GPT-4o와 Gemini 1.5 Flash 등 새롭게 공개된 AI 모델과 서비스들

오픈AI가 새로 공개한 모델 GPT-4o에 추가된 알파벳 'o'는 '모든'을 뜻하는 'omni'의 약자로, 텍스트 중심이었던 과거와 달리 텍스트와 오디오, 이미지, 영상의 모든 조합이 입력에서부터 출력까지 가능한 멀티모달(Multimodal) 속성을 드러내고 있다. 즉, 그동안 필담만 가능했던 챗GPT에게 눈과 귀와 목소리가 생긴 것이다. 예를 들어 과거의 GPT는 음성 대화를 위해 ①사용자의 음성을 텍스트로 변환(STT) ②텍스트를 기반으로 텍스트 형태의 답변을 생성(GPT) ③텍스트 답변을 음성으로 변환(TTS)하는 세 가지 개별 모델을 거쳐야만 했다. 변환과 전송이라는 각 단계로 인해 음색이나 배경음, 감정 등의 정보가 손실되었으며 처리 속도도 지연되었다.

그러나 이번에 공개된 GPT-4o는 텍스트와 이미지, 오디오를 새로운 단일 모델에서 처음부터 끝까지(end-to-end) 동일한 신경망에서 처리하기 때문에 정보의 손실

없이 빠른 답변이 가능하다는 장점을 지니고 있다. 실제로 오픈AI가 공개한 자료에 따르면 이번 모델의 평균 답변 속도가 232밀리초로, 인간의 평균 답변 속도인 320밀리초보다 더 빠른 답변이 가능하다는 점을 알 수 있다.

공개된 영상에서도 GPT-4o가 다양한 형태의 자료를 활용한 이용자의 요청에 빠르게 응답하는 모습을 확인할 수 있다. 기능의 결과적인 부분에서는 큰 차이가 없는 것처럼 보일 수 있으나 그 과정이나 속도 등에서는 큰 변화가 있었다.

구글 역시 이번 개발자 컨퍼런스에서 무엇보다 AI를 강조했다. 그러나 오픈AI의 발표가 멀티모달을 중심으로 하는 하나의 완성된 소설과 같았다면, 상대적으로 구글의 발표는 기업의 카탈로그처럼 자사의 AI와 서비스가 결합할 수 있는 다양한 사례를 제시했다고 볼 수 있다.

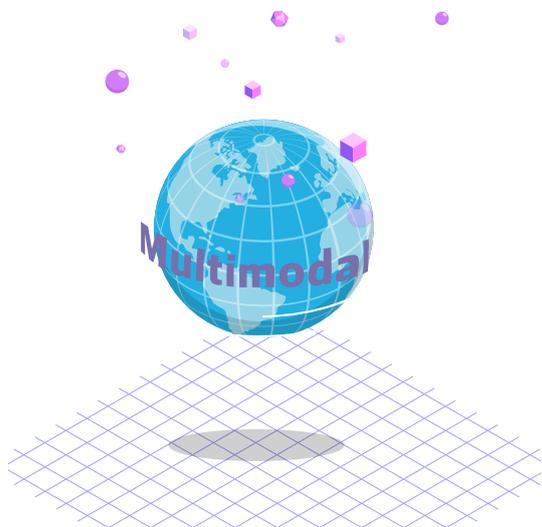
먼저 구글은 자사의 최상위 AI 모델 Gemini 1.5 pro의 업데이트와 함께 경량화 버전인 Flash, 그리고 동영상 생성모델 Veo와 이미지 생성모델 Imagen 3, 차세대 오픈 모델 Gemma 2.0 등 다양한 AI 모델을 공개했다. 이상의 모델들은 구글의 안드로이드 OS가 적용된 스마트폰이나 과거에 개발했던 구글 글래스 등을 통해서 구글이 제공하는 서비스인 검색엔진과 워크스페이스, 구글 포토 등에 적용된다. 마지막으로 이러한 모델과 기능들을 활용할 범용 AI 에이전트 '프로젝트 아스트라'를 공개하면서 자사가 지향하는 미래상을 공유했다.

여전히 모델 성능에서 앞선 오픈AI로 인해 구글에 대해



몇몇 사람들은 아쉬움을 표하기도 했다. 그러나 일각에서는 이번 발표를 통해 구글의 경쟁력을 확인할 수 있었다고 평가하기도 한다. 오픈AI에게는 스마트폰도 스마트글래스도, 또는 클라우드나 워크스페이스도 없기 때문이다.

즉, 오랜 기간 구글이 구축해온 생태계가 없다. 이러한 한계를 극복하기 위해 초기단계부터 마이크로소프트와의 협업을 강화해왔으며 최근에는 애플과의 협력이 임박했다 소식도 들려오고 있다. 그러나 오픈AI의 AI가 보다 압도적인 성능을 보일 수 없다면, 소비자들은 Windows 또는 Azure와 iOS로 파편화된 생태계를 택하기 보다는 구글이 제공하는 종합 서비스를 택할 확률이 높다.



모델의 성능 경쟁 이후 다가올 활용성

이용자가 인지할 수 없는 수준의 기능 경쟁은 그 영향력이 제한된다. 초기에는 통신 품질이나 속도 중심이었던 이동통신사업자들의 경쟁에서, 통신 속도가 소비자의 인식 또는 필요 수준을 넘어버린 지금에 와서는 데이터 속도가 소비자에게 큰 의미를 주지 못하고 있는 것처럼 말이다.

앞으로 다가올 AI 역시 마찬가지로 볼 수 있다. 아직 기술이 태동하고 있는 지금은 어느 모델이 더 환각을 줄이고 응답속도가 빠르거나 등이 가장 중요한 경쟁 포인트가 될 것이다. 물론 이러한 기술 개발이 절대 쉬운 일은 아니기 때문에 오랜 기간 이어질 가능성이 크다.

그러나 향후 언젠가 다수의 AI 모델들이 사용자가 만족할 만한 수준의 성능을 보유하게 된다면, 그때는 기술적 발전이 아닌 활용성을 중심으로 경쟁하게 될 것이다.

그리고 이러한 활용성은 연계된 하드웨어와 소프트웨어가 무엇이나에 따라 달라지게 될 가능성이 크기 때문에 빅테크 기업들의 경쟁 전략이 무엇을 지향하는가는 당분간 재미있는 관전 포인트가 될 것이다.

최근 공개된 GTP-4o 발표에서 시각장애인 이용자는 불편하게 스마트폰을 들고 카메라를 비추면서 AI와 대화할 수 있었다. 그러나 Google I/O에서처럼 자사의 스마트안경을 이용할 수 있다면 이용자의 선택은 자명하다. 실제로 아직 오픈AI가 보다 큰 데이터를 활용하여 성능

이 뛰어난 AI를 중심으로 자사의 강점을 집중하고 있다면, 구글은 경량화를 통한 온디바이스 AI 모델 개발에도 큰 비중을 쏟는 차이를 보이고 있다는 점에서 두 기업의 방향성 차이를 확인할 수 있다.

물론 오픈AI가 하드웨어에 강점을 가진 삼성이나 LG, 애플, 메타 등과 협업한다면 자사의 경쟁력을 이어나갈 수 있을 것이다. 그러나 계열사 간 협업만큼 높은 수준의 연계는 어려울 것이며 무엇보다 사용자 데이터 등에 대한 주권의 문제가 발생한다.

이미 AI 모델의 학습을 위한 고급 데이터가 고갈되었다는 이야기가 나오고 있으며, 자사의 데이터가 가진 중요성을 인식하게 된 언론사들은 저작권 침해로 오픈AI를 고소하기에 이르렀다. 구글도 만약 오픈AI가 유튜브의 동영상들을 AI 모델 학습에 사용했다면 이는 명백한 서비스 위반이라며 견제했다. 이러한 상황에서 오픈AI가 자사의 학습데이터를 잠재적인 AI 경쟁자인 다른 빅테크 기업과 공유할 수 밖에 없는 협력 관계를 어떻게 갖춰 나갈 것인가에 대해서는 많은 논의가 필요할 것이다.

다만 소프트웨어는 조금 복잡한 양상을 보일 수 있다. 분명 구글은 자사의 수많은 B2C 서비스에 AI 모델을 적용할 수 있는 가능성을 보여주었지만, 앞서 언급한 것과 같이 파편화된 공개로 인해 이용자에게 실질적인 AI 세계를 보여주는 것에는 실패했다. 반면 오픈AI는 보다 집중된 기능을 통해서 이용자의 실질적인 경험을 간접적으로 보여주었기 때문에 높은 평가를 받을 수 있었다. 따라서 향후 구글이나 삼성, MS 등은 자사의 서비스를 통합적으로 묶어서 제공할 수 있는 AI 플랫폼에 대한 고민을 더욱 깊게 해야 할 필요가 있을 것이다.



필자 | 이준원

브랜드 커뮤니케이션을 전공하였으며 디지털 광고회사에서 근무한 후 현재 한국외대 미디어커뮤니케이션 연구소 및 미디어 외교센터 소속으로 디지털 미디어를 통한 커뮤니케이션 전략에 대한 연구를 진행하고 있다.