

ChatGPT vs Google Gemini vs Perplexity AI: 의료 분 야 딥러닝 리서치 성능 비교

서비스별 비교 요약 (의료 분야 딥 리서치 기준)

평가 기준	ChatGPT (OpenAI)	Google Gemini	Perplexity AI
정보 탐색의 정확 도 '가이드라인·DB 인용 정확성)	- 출처 인용 한계: 기본 ChatGPT는 답변에 출처 를 잘 제시하지 않으며, 학술 정보도 자체 지식에 의존하기 때문에 최신 논 문이나 가이드라인 인용 이 제한적입니다. 실제 참 고문헌을 요구하면 존재 하지 않는 논문이나 잘못 된 출처를 생성하는 사례 도 보고되었습니다 1. < br>- 향상된 모드: 2024 년 도입된 Deep Research 모드나 웹 브 라우징 기능을 활용하면 인터넷에서 정보를 찾아 실제 논문을 인용할 수 있 지만, 이 기능은 별도 활 성화가 필요합니다 2.	- 높은 인용 정확도: Gemini는 구글 검색 및 학술 데이터베이스와 연계되어 실제 논문과 데이터 출처를 정확히 찾아 인용하는 데강점이 있습니다. 한 연구에서는 의학 논문 소개 작성 시참고문헌의 정확성 면에서 GPT-4보다 Gemini가 높은 정확도(올바른 참고문헌 비율 77.2%)를 보였습니다 1. < br> 의 : < br> 의 하 : <	- 검색 기반 인용: Perplexity는 애초부터 웹 검색 결과를 근거로 답변하므로실제 존재하는 웹 문서만을인용합니다. 답변에 항상 출처 링크를 포함하여 정확도를높이고 있으며 3, 사용자가추가로 자료를 확인하기 용이합니다. 잘못된 참고문헌을제시할 위험이 매우 낮고, 인용된 출처들은 클릭을 통해원문을 바로 검증할 수 있습니다 4.

- 뛰어난 추론 능력:

GPT-4 기반 ChatGPT는 복잡한 의료 질문에 대해 서도 논리적으로 추론하 고 상세히 설명하는 능력 이 매우 뛰어납니다. 예를 들어 의료 라이선스 시험 (USMLE) 연습문제에 GPT-4가 86~90% 수준 의 높은 정답률을 보이는 등 전문가 수준의 지식을 입증했습니다 5 . 복잡 한 다단계 질문에서도 심 도 있는 분석과 배경지식 을 바탕으로 종합적인 답 변을 생성하는 데 강합니 다 6 .
>- Deep

Research 모드: 고급 모 드에서는 새로운 OpenAl 모델(o3)을 통해 더욱 향 상된 분석력을 제공하며, 매우 어려운 추론 벤치마 크 시험에서도 기존 GPT-4 대비 월등히 높은 점수를 기록해 심층 분석 에 강점을 보였습니다

- 전문 분야 최적화:

Google의 Gemini는 의료 도메인에 특화된 Med-Gemini 모델 등을 통해 전 문성을 강화했습니다. 실제 로 의료 전문 벤치마크에서 최고 수준 성능을 달성했는 데, 미국 의사시험 형식의 MedOA 벤치마크에서 91.1%의 정답 정확도로 기 존 모델(Med-PaLM2 등)을 능가했습니다 7 8.
br>- 추론 및 설명: 일반적 인 Gemini도 복잡한 의료 질문에 대해 다각도로 정보 를 수집해 답변하며, 멀티모 달 정보(예: 의료 영상)까지 다룰 수 있는 차세대 모델로 개발되고 있습니다 9 . 다 만 초기 버전 Gemini는 극 도로 난해한 추론 과제에서 는 ChatGPT 특화 모델보다 낮은 점수를 받았으나 10, 2025년 출시된 Gemini 2.0 에서는 논리성과 응답 깊이 가 개선되고 있다는 평가입 니다 11.

- 신속한 종합 조사:

Perplexity는 방대한 웹 정보 에 접근해 짧은 시간 내 다수 의 출처를 조사하고 종합하기 때문에, 폭넓은 정보를 근거 로 한 꽤 깊이있는 답변을 빠 르게 제공합니다 12 13. Deep Research 모드에서는 수백개의 검색 결과를 검토해 인간 연구자가 몇 시간 할 일 을 몇 분만에 수행하며, 복잡 한 질문에 대해서도 주요 쟁 점을 커버하는 상세 보고서를 작성할 수 있습니다 12 14.
br>- 전문성의 한계: 다만 기 본적으로 GPT-3.5 기반 모델 을 활용하기 때문에(OpenAl GPT-4 등 상위 모델은 Pro 구 독 필요), 최첨단 모델을 쓰는 ChatGPT만큼 세밀한 전문 지식 설명이나 고차원 추론에 서는 조금 뒤처질 수 있다는 지적이 있습니다 14 . 그럼에 도 어려운 추론 벤치마크에서 도 OpenAl 대비 약간 낮은 수 준(약 20% 대)까지 도달하여 상위권 성능을 보였고 13, 사실관계 위주의 전문 지식 전달에는 강합니다.

응답 깊이와 전문

성
(설명의 정

교함과 전문가 수

준 지식)

- 출처 미제공 (기본 모

드): 일반 ChatGPT 답변 은 출처 표시가 없어, 사 용자가 답변의 근거를 확 인하기 어렵습니다 15. 더구나 모델이 기억에 의 존해 근거를 만들다 보니 존재하지 않는 논문이나 틀린 정보를 그럴듯한 출 처와 함께 제시하는 **할루 시네이션** 사례가 알려져 있어 주의가 필요합니다 1 .
>- 출처 제공 (딥 리서치 모드): Deep Research 모드의 결과물 은 참고 문헌과 링크를 체 계적으로 나열해주기 때 문에 투명성이 크게 향상 됩니다 16 . 보고서 형태 로 섹션별로 내용을 정리 하고 각 주장 옆에 출처를 병기하여 신뢰도를 높였 으나, 이 모드 역시 AI가 요약하는 과정에서 정보 가 왜곡될 가능성이 있으 므로 중요한 내용은 사용 자가 직접 출처를 확인하 는 것이 권장됩니다.

- 명확한 참고자료 제시:

Gemini의 딥 리서치 출력은 답변 내용과 함께

"Sources and related content" 섹션에 참고한 웹사이트나 문헌 링크를 명 시해 줍니다 17 . 각 정보마 다 원본 출처를 하이퍼링크 로 제공해 투명성을 확보하 며, 잘 알려진 데이터뿐 아니 라 놓치기 쉬운 전문 웹사이 트나 기관 자료까지 찾아 제 시해주는 것으로 알려져 있 습니다 18 17 .
>- 신뢰 성: 실제 참고문헌의 정확성 면에서도 Gemini는 GPT-4 보다 우수하여 잘못된 출처 생성 비율이 낮았다는 연구 결과가 있습니다 1 . 다만 AI가 여러 출처의 내용을 종 합하다 보니 상충되는 정보 는 깊이 있게 비평하지 않고 나열할 수 있어 19, 최종 판 단에는 사용자의 검증이 여 전히 필요합니다.

- 최고 수준의 출처 투명성:

Perplexity는 답변 생성 과정 에서 참조한 출처를 풋노트 형태로 항상 명시하여 사용자 에게 보여줍니다 4 . 답변의 주요 문장마다 괄호 번호가 붙고, 클릭하면 해당 근거가 된 웹 문서나 연구를 바로 열 람할 수 있어 신뢰도 검증이 용이합니다.
- 사실 기반 답변: 검색을 통해 확인된 정 보만을 사용하고 AI가 임의로 지어낸 내용을 배제하는 설계 이므로, 인용된 출처의 신뢰 성이 높습니다 (3). 이러한 투명성 덕분에 "어떤 정보를 **어디서 가져왔는지"** 이용자 가 즉각적으로 파악할 수 있 으며, AI와 사용자 간 신뢰의 **다리**를 놓아준다는 평가를 받 고 있습니다 4.

자료 출처 제시 여 부 및 신뢰성)

출처의 신뢰도와

투명성
(참고

- 지식 컷오프 한계:

ChatGPT 기본 모델은 학 습 데이터가 2021년으로 제한되어 있어 2022년 이 후의 의료 연구 동향이나 최신 진료 가이드라인에 대한 지식이 부족할 수 있 습니다 15 . 예를 들어 2024년에 새로 발표된 임 상시험 결과를 질문하면 기반 지식만으로는 답하 지 못하고 과거 정보에 근 거한 추측을 할 위험이 있 습니다.
- **웹 검색 기** 능: 2023~2024년에 도 입된 웹 브라우징 기능을 통해 이러한 지식 한계를 극복할 수 있습니다 2 . ChatGPT+ 사용자는 Bing 검색 등을 활용해 실시간으로 인터넷 정보 를 가져와 답변에 반영할 수 있으며, 이로써 고정형 AI 모델의 지식 공백을 메 우는 업그레이드가 이루 어졌습니다 2 . 다만 이 기능은 사용자가 직접 활 성화해야 하고, 일부 제한 적으로 제공됩니다.

- 실시간 정보 연계:

Gemini는 태생부터 Google의 방대한 최신 웹 인덱스를 활용하도록 설계 되어 2024~2025년의 새로 운 의학 정보도 잘 반영합니 다. 모델의 임상 데이터 튜닝 에 최신 의학 논문과 지침이 포함되었을 뿐 아니라, 필요 시 불확실한 부분에 대해 자 **동으로 웹 검색을 수행**하여 최신 정보를 가져오는 능력 이 보고되었습니다 20 .

- 업데이트 속도: Google의 검색 강점을 토대 로 현시점 가장 최신 뉴스나 논문까지 답변에 활용하며, 업데이트 주기도 빠릅니다. 따라서 2024년~2025년의 의료 지식이나 동향 파악에 서는 ChatGPT보다 한발 앞 서 있는 것으로 평가됩니다 11.

- 항상 최신 정보 활용:

Perplexity AI는 질문할 때마 다 실시간 웹 검색을 수행하 므로, 현재 시점 가장 최신의 정보가 바로 답변에 반영됩니 다 15 . 예컨대 2025년에 발 표된 의학 논문이나 당일 보 도된 의료 뉴스를 물어보면 관련 웹 결과를 즉시 찾아 요 약하므로 정보의 최신성이 매 우 높습니다.
- **학술 데** 이터 접근: 기본적으로 인터 넷 전체를 대상으로 검색하지 만, 필요에 따라 학술 논문만 검색하는 모드 등도 제공하여 최신 학계 동향 파악에 최적 화되어 있습니다 21 . 이처럼 2025년 현재 진행형인 정보 까지도 AI 답변에 포함되므로 시의성 있는 의료 리서치에 유용합니다.

최신 정보 반영 능

br>(2024~2025

년 최신 의학 정보

력

반영)

평가 기준

사용자 인터페이스

(리서치 흐름,

검색 결과 제시 방

및 피드백 기능

식)

- 챗봇형 보고서 출력:

ChatGPT Deep Research는 기존 채팅 인터페이스를 유지하면서 도 **리서치 에이전트**처럼 동작합니다 22 . 사용자 가 연구 과제를 입력하면 AI가 자체적으로 여러 단 계의 검색·분석을 수행한 뒤, 완료 시 **서론과 섹션,** bullet 포인트, 참고문헌 까지 갖춘 체계적인 보고 서 형태로 답변을 제공합 니다 16 23 . 결과를 얻 기까지 몇 분 정도 기다려 야 하지만, 한 번에 정리 된 결과물을 얻을 수 있고 이후에는 일반 채팅처럼 추가 질문을 하며 세부 내 용을 추궁하거나 업데이 트를 요구할 수 있습니다 24 .
- 피드백 및 계 획 수립: 현재

ChatGPT의 검색 과정은 대부분 숨겨져 있어 사용 자가 중간에 개입할 수는 없지만, 초기 프롬프트에 서 조사 희망 방향을 제시 하면 이를 계획에 반영합 니다. 결과를 받은 후 이 어지는 대화에서 잘못된 부분을 지적하거나 보강 을 요청하는 **피드백 루프** 를 통해 점진적으로 답변 의 질을 높일 수 있습니 다.

- 구글 생태계 통합 UI: Gemini의 딥 리서치 기능은 Google AI Assistant 환경 에 내장되어 있으며, 웹앱 및 모바일앱을 통해 제공됩니 다 25 . 사용자가 복잡한 질 문을 입력하면 바로 답을 주 는 대신, 먼저 "조사 계획"을 개략적으로 보여줘 사용자 가 방향을 확인하고 수정할 기회를 줍니다 25 . 이 계획 에는 어떤 하위 주제들을 검 색/분석할지가 포함되며, 사 용자가 추가로 궁금한 부분 을 계획에 넣거나 제외할 수 도 있습니다 (예: 특정 질병 의 치료법에 집중). 이후 Start를 누르면 AI가 단계별 로 웹 검색과 정보를 종합해 나가고, 완료 후에는 깔끔하 게 정리된 리포트를 제공합 니다 26 .
- **결과물과** 대화형 탐색: Gemini의 결 과 출력은 가독성이 좋게 섹 션별 Q&A 형태나 토픽별 요 약으로 나타나고, 각 부분에 관련 출처 링크가 붙어 있습 니다 17 . 보고서 하단에는 참고자료 목록도 정리됩니

Gemini는 대화 문맥을 유지 하므로 마치 사람 연구원과 상의하듯 리서치를 이어갈 수 있습니다.

다. 사용자는 답변을 확인한

뒤 계속해서 **대화형으로 후**

속 질문을 던질 수 있으며,

보고서의 특정 섹션에 대해

더 깊이 파고들도록 요청할

수도 있습니다 27.

- 검색엔진형 간결 UI:

Perplexity는 처음 이용 시 일반 검색창과 유사한 인터페 이스를 제공하며, 질문을 입 력하면 즉시 AI가 검색을 수 행합니다 28 29 . **딥 리서치 모드**로 전환하면 한 번의 클 릭으로 심층 보고서 생성을 시작하는데, 이때도 UI는 복 잡하지 않고 검색 질의 입력 후 기다리면 알아서 여러 번 의 검색과 분석이 백그라운드 에서 이루어집니다 30 . 대략 1~3분 내에 완성된 보고서가 출력되며, 결과 페이지는 관 련 주제들이 소제목으로 나뉘 고 각 문장에 출처가 붙은 형 태로 **깔끔하게 요약**됩니다 31 . 필요 시 결과를 **PDF로 바로 저장**하거나 공유할 수도 있어 편리합니다.
- 대화 및 모드 기능: Perplexity는 답변 후에도 **스레드 형태**로 추가 질문을 이어갈 수 있어, 맥락을 유지한 대화형 검색이 가능합니다 32 . 또한 검색 소스를 사용자가 지정할 수 있는데, 예를 들어 학술 논문 만 대상으로 답변받는 Academic 모드, 뉴스 기사만 보는 모드, 커뮤니티 의견을 보는 모드 등 상황별 검색 옵 **션**을 제공해 사용자 피드백을 반영합니다 21 . 이러한 UI와 기능 덕분에 사용자는 자신이 원하는 방식으로 리서치 방향 을 조절하기 용이합니다.

- 직관적 인터페이스:

ChatGPT는 국내외 사용 자들에게 익숙한 채팅 형 식 UI로, 별도의 설치 없 이 웹에서 바로 사용할 수 있어 편의성이 높습니다. 다만 Deep Research와 같은 고급 기능은 일반 무 료 사용자에겐 제공되지 않고, 현재 미국 중심으로 제한된 베타 형태로 고가 (월 \$200 상당) 플랜 사용 자만 이용 가능해 접근성 에 한계가 있습니다 33.
 한국어 지원:

OpenAl GPT-4 모델은 한국어를 포함해 80개 이 상의 언어를 능숙하게 이 해하고 생성할 수 있습니 다 34 . 실제로 한국어 질 의에도 유창하고 전문적 인 답변을 만들어주며, 의 료 용어 등도 꽤 정확히 처리합니다. 다만 영어만 큼 학습 데이터가 많지 않 아 복잡한 뉘앙스나 국내 자료 인용 면에서는 약간 제한이 있을 수 있다는 지 적이 있습니다 35.

- 광범위한 접근성: Google Gemini의 딥 리서치 기능은 Gemini Advanced 구독 (약 \$20/월)에 포함되어 있 으며, 2024년 말 기준 전세 계 150여 개국에 출시되어 많은 사용자가 활용 가능합 니다 36 . ChatGPT에 비하 면 가격 장벽이 낮고 지원 범 위도 넓어 비교적 손쉽게 접 할 수 있습니다.
- **한국 어 및 다국어:** Google은 이 미 자사의 Bard 등을 통해 한국어 지원을 제공해왔으 며, Gemini도 45개 이상의 언어를 공식 지원합니다

36 . 한국어로 질문했을 때 도 자연스러운 답변을 얻을 수 있고, 인터페이스 역시 한 국어를 포함한 다양한 언어 로 제공되므로 비영어권 연 구에도 용이합니다. Google 검색과 연동되는 특성상 한 국어 웹 검색 결과까지 포괄 하여 활용할 수 있어 현지화 된 정보 접근성도 높습니다.

- 뛰어난 사용 편의성:

Perplexity AI는 회원가입 없 **이도 무료**로 기본 기능을 사 용할 수 있고, 인터페이스가 단순명료해 초보자도 부담 없 이 활용할 수 있습니다 30. 딥 리서치 모드 역시 하루 5회 까지는 무료로 제공되므로 가 벼운 연구 목적이라면 비용 없이 충분히 활용 가능합니다 37 . 또한 웹, iOS앱, 안드로 이드앱, 브라우저 확장 프로 그램 등 **멀티플랫폼 지원**으로 접근성이 우수합니다 38

39 .
- 한국어 지원 품 질: Perplexity는 출시 초기 부터 한국어를 포함한 15개 언어에 대해 대화 지원을 시 작했으며 40, 모바일 앱 등 에서도 한국어 인터페이스를 제공합니다 39 . 한국어 질문 에 대해 정확한 웹 검색 결과 를 바탕으로 답변하며, 다만 생성에 쓰이는 언어 모델 특 성상 영어 정보가 상대적으로 많아 한국어로 질문 시에는 검색 결과가 영어 위주로 나 올 가능성도 있습니다. 그럼 에도 불구하고 기본적인 한국 어 질의응답 품질은 높고, 국 내 자료가 존재하면 해당 언 어로도 내용을 요약해주므로 한국어 리서치에도 활용도가 높습니다.

종합 평가

평가 기준

사용 편의성 및 언

어 지원

<한국

어 지원 여부 및 전

반적 사용 난이도)

ChatGPT, Google Gemini, Perplexity AI 모두 의료 분야 딥 리서치에 활용될 수 있으나, 각 서비스마다 두드러지는 강 점과 약점이 있습니다. ChatGPT는 뛰어난 **언어 표현력과 추론 능력**을 갖춰 복잡한 의학 주제도 전문가 수준으로 풀어 내지만, 실시간 정보 접근과 출처 투명성 면에서는 기본 모드의 한계가 있습니다 15 . Google Gemini는 Google의 검 색 인프라를 바탕으로 **최신 자료**를 정확하게 찾아내고 **신뢰할 수 있는 출처**를 제시하는데 강하며 📵 👊 , 의료 특화 모 델로 도메인 전문성도 탁월합니다. 반면 출시 초기 일부 추론 능력에서는 개선 여지가 있다는 평가가 있었지만 10,지 속적인 업그레이드로 격차를 빠르게 줄이고 있습니다. Perplexity AI는 경량화된 인터페이스로 즉각적인 검색과 명확 한 출처 제시를 실현하여 사실 확인과 최신 정보 수집에 매우 유용하며 15, 다만 가장 첨단 모델을 사용하는 것은 아니 어서 서술 응답의 깊이는 약간 덜 정교할 수 있습니다 [14]. 종합하면, ChatGPT는 깊이있는 분석, Gemini는 정확한 인 용과 최신 데이터, Perplexity는 빠르고 투명한 정보수집 측면에서 각각 두각을 나타냅니다.

서비스별 추천 사용 시나리오

- ChatGPT 창의적 분석 및 종합 보고서가 필요한 경우에 적합합니다. 예를 들어 의학 논문 초안을 작성하거나 복잡한 의료 개념을 쉽게 풀어 설명해야 할 때, ChatGPT의 풍부한 표현력과 추론 능력이 도움이 됩니다. 사전 지식 기반이 탄탄하여 의학 지식 학습용 교재 생성, 환자 상담 스크립트 초안 작성 등에도 활용할 수 있습니다. 단, 최신 근거 자료가 필요한 작업에는 반드시 브라우징 기능이나 Deep Research 모드를 활용해 정보의 현행 성을 보완해야 합니다 2 .
- Google Gemini 최신 임상 근거에 기반한 전문적 리서치가 필요한 경우 추천됩니다. 예를 들어 최신 치료 가이드라인 비교, 최근 1년간 발표된 의학 연구 조사와 같이 2024-2025년의 최신 정보와 신뢰도 높은 출처가 핵심인 과제에 적합합니다. Gemini는 검색 단계부터 학술 사이트와 데이터를 정확히 찾아내어 팩트체크가 필요한 의료 보고서 작성, 근거 중심의 임상 의사결정 지원 등에 유리합니다. 또한 다국어 지원이 잘 되므로 한국어 논문은 물론 영문 자료까지 모두 참고해야 하는 국제 협력 연구에서도 유용합니다.
- Perplexity AI 빠르게 사실 확인을 하거나 최신 정보를 검색해야 하는 상황에 알맞습니다. 의학 용어 정의, 약물 정보 조회, 의료 뉴스 브리핑처럼 간단하지만 정확한 출처 확인이 필요한 질문에 대해 즉각적으로 답을 얻을수 있습니다 15. 예를 들어 환자가 가져온 인터넷 정보를 검증하거나, 논문 작성 시 참고문헌의 출처를 신속히찾아볼 때 유용합니다. 별도 비용 없이 웹과 모바일에서 쉽게 사용할수 있으므로 의대생이나 연구 초심자가 자료조사를 시작할때, 또는 의료 전문가가 급히 근거를 확인해야 할때 부담 없이 활용할수 있는 실용적 리서치도구입니다.

출처: 공식 블로그 및 연구 자료 7 1 6 등.

1 Generating Credible Referenced Medical Research: A Comparative Study of Openai's Gpt-4 and Google's Gemini by Mahmud Omar, Saleh Nassar, Kareem Hijaze, Benjamin S. Glicksberg, Girish N. Nadkarni, Eyal Klang:: SSRN

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4780940

² ChatGPT update provides browsing tool with internet access

https://wpadvancedads.com/chatgpt-provides-current-data/

3 6 10 11 12 13 14 16 17 18 19 22 23 24 25 26 27 30 31 33 36 37 Comparing Leading Al Deep Research Tools: ChatGPT, Google, Perplexity, Kompas Al, and Elicit | by ByteBridge | Medium https://bytebridge.medium.com/comparing-leading-ai-deep-research-tools-chatgpt-google-perplexity-kompas-ai-and-elicit-59678c511f18

4 21 28 29 32 38 39 What is Perplexity Al? Everything you need to know | Tom's Guide https://www.tomsguide.com/ai/what-is-perplexity-ai

5 ChatGPT-4: An assessment of an upgraded artificial intelligence chatbot in the United States Medical Licensing Examination - PubMed

https://pubmed.ncbi.nlm.nih.gov/37839017/

7 8 9 20 Advancing medical AI with Med-Gemini

https://research.google/blog/advancing-medical-ai-with-med-gemini/

15 ChatGPT vs. Perplexity vs. Grok For Research: Which One's Better? - Al Tools https://www.godofprompt.ai/blog/chatgpt-vs-perplexity-vs-grok-for-research-which-ones-better? srsltid=AfmBOopIM96K9jONZFRNiRmTsKLmWeIXhbf8APrAS4DLd3NmA1V4xMdx

34 List of languages supported by ChatGPT

https://botpress.com/blog/list-of-languages-supported-by-chatgpt

35 GPT-4: Boost Your Korean Language Skills with Text and Image ...

https://www.toolify.ai/gpts/gpt4-boost-your-korean-language-skills-with-text-and-image-understanding-ai-125415

40 Perplexity AI - Wikipedia

https://en.wikipedia.org/wiki/Perplexity_AI