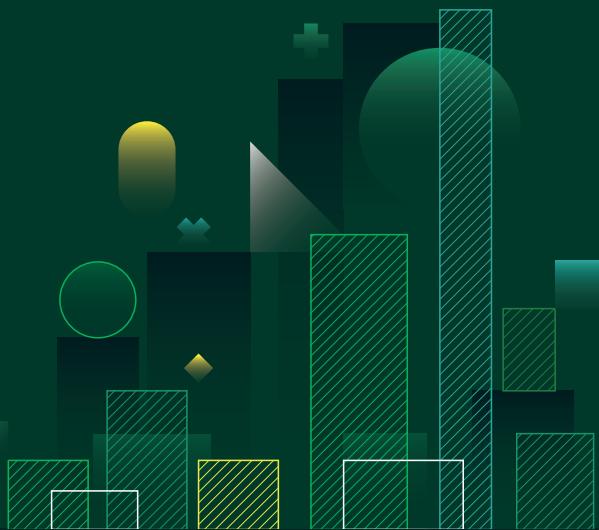


KPF 미디어브리프

K P F M e d i a B r i e f W e b z i n e



AI 검색 서비스의 저널리즘 콘텐츠 인용 현황과 언론 산업의 대응전략

이현우 | 한국언론진흥재단 선임연구위원

- ChatGPT, Gemini, Claude 등 AI 검색 서비스의 전체 인용 중 저널리즘 콘텐츠는 27%, 시의성 있는 검색의 경우 절반(49%)이 저널리즘 콘텐츠 인용
- Reuters, FT, Axios, Time 등 권위 있는 언론사를 자주 인용하며, 금융/보험, 정부, 미디어/엔터테인먼트 영역에서 저널리즘 콘텐츠 많이 인용
- 하지만 OpenAI는 1,700회 크롤링당 1회, Anthropic은 73,000회 크롤링당 1회만 해당 사이트를 추천하여, 광고 수익 기반인 트래픽은 보내지 않는 불공정 구조 지속
- 언론사들은 크롤링당 결제(Pay per Crawl) 시스템, AI 검색 수익 분배 솔루션, e-커머스 개인화 전략 벤치마킹 등을 통해 산업의 지속가능성 확보 필요



현재 급속히 이용자 기반을 확장하고 있는 AI 기반 검색 서비스는 사람들의 정보 검색 행동패턴을 근본적으로 변화시키고 있다. 사용자가 질문을 입력하면 수많은 링크 대신 즉시 통합된 답변을 제공하는 AI 검색은 편의성과 효율성 면에서 기존 검색 엔진을 빠르게 대체하고 있다. ChatGPT, Gemini, Claude 등 주요 AI 서비스들은 이미 수억 명의 사용자를 확보했으며, 이들이 생성하는 답변은 인터넷상의 방대한 콘텐츠를 실시간으로 수집, 분석, 재구성하여 제공한다.

이러한 변화 속에서 저널리즘 콘텐츠는 역설적 상황에 직면해 있다. AI 기반 PR 플랫폼 Generative Pulse가 2025년 7월 발표한 ‘What Is AI Reading?’ 보고서¹⁾에 따르면, AI 시스템이 인용하는 콘텐츠의 27%가 저널리즘 콘텐츠이며, 특히 시의성 있는 정보가 필요한 질의에서는 이 비율이 49%까지 상승한다. Reuters, AP, FT 같은 권위 있는 언론사들이 AI의 주요 정보원으로 활용되고 있다는 사실은 신뢰할 수 있는 저널리즘의 가치가 AI 시대에도 여전히 핵심적임을 보여준다.

그러나 동시에 언론사들은 심각한 위기에 직면해 있다. AI 기업들이 콘텐츠를 대량으로 수집하면서도 정작 웹사이트에 대한 트래픽 기여는 미미한 수준에 그치고 있기 때문이다. 현재 OpenAI는 1,700회 크롤링당 1회, Anthropic은 73,000회 크롤링당 1회만 해당 사이트를 추천하는 것으로 나타났다²⁾. 이는 콘텐츠는 무단으로 활용하면서도 광고 수익의 기반이 되는 트래픽은 보내지 않는 불공정한 비즈니스 구조를 의미한다.

이러한 문제의식을 바탕으로 본 보고서는 우선 Generative Pulse의 연구 결과에 나타난 AI 검색 서비스들의 저널리즘 콘텐츠 인용 현황과 패턴을 상세히 살펴볼 것이다. ChatGPT, Gemini, Claude 등 주요 AI 모델들이 인용 링크를 어떻게 선택하고 활용하는지, 산업별로 어떤 차이를 보이는지 체계적으로 살펴본다. 이어서 현재 진행 중인 언론 분야의 대응 사례들을 검토한다. 클라우드플레이어의 ‘크롤링당 결제’ 시스템, ProRata의 AI 검색 수익 분배 모델, 그리고 e-커머스 산업의 개인화 전략 벤치마킹 등을 통해 AI 검색 시대 저널리즘의 지속가능성 확보를 위한 다층적 전략을 제안하고자 한다.

■ AI 검색 서비스의 저널리즘 콘텐츠 인용 현황

▲ 전체 인용의 27%는 저널리즘 콘텐츠, 시의성 검색은 49% 차지

Generative Pulse 연구에 따르면, AI가 인용하는 링크의 95% 이상이 비광고 콘텐츠였으며, 저널리즘 콘텐츠가 전체 인용의 27%를 차지했다. 특히 흥미로운 점은 최신성을 암시하는 질의(쿼리)의 경우 저널리즘 콘텐츠 인용 비율이 49%까지 상승한다는 것이다. 이는 AI 시스템이 시의성 있는 정보를 제공할 때 신뢰할 수 있는 언론 매체를 주요 소스로 활용한다는 것을 시사한다.

한편 쿼리 유형별로도 인용 패턴이 크게 달라지는 것으로 나타났다. 조언이나 의견을 구하는 질문에서는 기업형 블로그 및 콘텐츠가 인용되는 반면, 최근 사건에 대한 업데이트나 사실 확인 질문에서는 저널리즘 소스가 압도적으로 많이 활용되고 있었다. 시간적 측면에서는 최근 12개월 이내 발행된 저널리즘 콘텐츠가 많이 인용되고 있었고, 권위있는 매체들(Reuters, AP, FT 등)이 일관되게 상위 인용 출처로 나타났다.

1) Generative Pulse는 2025년 7월 한 달 동안 생성형 AI 시스템들의 출처 인용 패턴을 분석한 대규모 연구를 수행했다. 연구진은 100만 개 이상의 링크를 수집하여 AI가 사용자 프롬프트에 응답할 때 어떤 출처를 인용하는지 체계적으로 분석했다.

2) <https://www.aetimes.com/news/articleView.html?idxno=200296&page=5&total=26290>

그림 1

AI 시스템 인용 패턴 분석

AI 시스템 인용 패턴 분석

ChatGPT, Gemini, Claude - 100만+ 인용 데이터 분석 (2025.07)

95%

비광고 콘텐츠

27%

저널리즘

49%최신 쿼리 시
저널리즘**89%**언드 미디어
(제 3자 생성)

- * 유료(광고) 미디어(Paid Media): 기업이 직접 비용을 지불하고 구매한 TV, 검색, 배너 등 광고 공간
- 언드 미디어(Earned Media) : 기업이 비용 지불 없이 획득한 미디어 노출. 언론보도나 소셜미디어, 온라인 리뷰 등
- 소유 미디어(Owned Media): 기업이 직접 소유하고 통제하는 웹사이트, 블로그, 앱, 간행물 등

▲ AI 검색 서비스별 특성

ChatGPT는 Reuters, AP, FT, Axios, Time, Forbes를 가장 자주 인용했으며, 최신 콘텐츠에 대한 선호도가 가장 높았다(인용된 저널리즘 콘텐츠의 56%가 12개월 이내 발행). Gemini는 ChatGPT와 유사하게 Reuters, FT, Axios를 자주 인용했지만, Investopedia 같은 전문 정보 사이트도 상위권에 포함되었다.

한편 Claude는 다른 모델들과 현저히 다른 패턴을 보였다. CNBC, Harvard Business Review, Good Housekeeping 등 더 다양한 출처를 활용했으며, 전통적인 뉴스 매체 인용은 상대적으로 적었는데 특히 Reuters 인용 빈도가 Gemini보다 20배, ChatGPT보다 50배 적었다. 또한 12개월 이내 최신 저널리즘 콘텐츠 인용 비율도 36%로 OpenAI 모델의 56%보다 낮았다. 이러한 차이는 각 AI 모델이 정보의 신뢰성과 관련도를 평가하는 알고리즘과 학습 데이터의 차이를 반영하는 것으로 보인다.

그림 2

AI 모델별 인용 특성

🤖 AI 모델별 인용 특성

ChatGPT

주요 소스:
Reuters, AP, FT

최신 콘텐츠:
12개월 이내 56%

특징:
전통 대형 언론사 선호

Gemini

주요 소스:
Reuters, FT, Axios

산업 특화:
중간 수준

특징:
균형잡힌 소스 활용

Claude

주요 소스:
CNBC, HBR, NPR

최신 콘텐츠:
12개월 이내 36%

특징:
다양한 전문 매체

▲ 금융/보험, 정부, 미디어/엔터테인먼트 영역에서 저널리즘 콘텐츠 많이 인용

산업별 분석 결과, 각 분야마다 AI 검색 서비스의 콘텐츠 인용 패턴에 뚜렷한 특징이 나타났다. 헬스케어와 여행/항공 산업에서는 CDC(미국 질병통제예방센터), NIH(미국 국립보건원), FAA(미국 연방항공청), IATA(국제항공운송협회) 같은 공신력 있는 정부 기관과 산업 협회의 자료가 주로 인용되었으며, 특히 헬스케어의 정부/NGO 출처 인용률은 18%로 출처 중 가장 높았다.

반면 금융/보험, 정부, 미디어/엔터테인먼트 분야에서는 저널리즘 콘텐츠 인용이 37%, 36%, 37%로 가장 많았고, 에너지 산업은 sciencedirect.com이나 energy.gov 같은 전문 기관 자료를 선호했다. 기술 산업은 특이하게도 학술 연구 인용이 거의 없는 대신 Coursera, Udemy 같은 온라인 교육 플랫폼과 Medium 같은 기술 블로그를 주로 활용하고 있었다.

그림 3 산업별 인용 특성

산업별 인용 특성

산업	주요 특징	핵심 지표
헬스케어	공신력 있는 소스 선호	정부/NGO 18%
금융/보험	시의성 있는 정보 중시	저널리즘 37%
에너지	전문기관 자료 선호	소유(owned) 미디어 30%
리테일	정보 집약적 콘텐츠	백과사전 36%
미디어/엔터	트렌드 민감, 최신성 중요	저널리즘 37%
정부	투명성과 신뢰성 강조	저널리즘 36%

■ AI 검색 시대 언론 산업의 지속가능성을 위한 시사점

앞서 살펴본 바와 같이, AI 검색 서비스가 Reuters, AP, FT 같은 권위 있는 언론사를 일관되게 상위 인용 출처로 활용하고, 최신성이 요구되는 쿼리에서 저널리즘 인용이 절반의 비중을 차지한다는 점은 객관적 검증과 시의적절한 정보 제공이라는 저널리즘의 본질적 역할이 AI 시대에 오히려 더욱 중요해졌음을 시사한다. 그러나 이러한 가치가 지속가능하려면 다음과 같은 전략적 접근이 필요하다.

▲ 콘텐츠 무단 이용 차단: 클라우드플레이어의 ‘크롤링당 결제’ 시스템

전 세계 웹사이트의 20%를 처리하는 웹 인프라 기업 클라우드플레이어가 2025년 7월 도입한 ‘크롤링당 결제(Pay per Crawl)’ 시스템은 AI 시대 콘텐츠 보호의 새로운 패러다임을 제시한다. 이 시스템은 AI 크롤러를 기본적으로 차단하고, 웹사이트 운영자가 허용한 크롤러에 대해서만 접근을 허가하며, 콘텐츠 수집 시마다 소액의 사용료를 부과하는 구조이다.

특히 주목할 점은 AI 기업들의 극심한 불균형적 크롤링 행태이다. OpenAI는 1,700회 크롤링당 겨우 1회, Anthropic은 무려 73,000회 크롤링당 1회만 해당 사이트를 추천하는 것으로 나타났다. 이는 과거 구글이 14회 크롤링당 1회 추천으로 트래픽을 보장했던 것과는 완전히 다른 구조로, AI 기업들이 대량의 콘텐츠 수집에도 불구하고 실제 트래픽 기여는 미미하다는 것을 보여준다³⁾. 콘데 나스트, 타임, AP, 애틀랜틱 등 주요 언론사들이 이미 기본 차단 정책에 동참했다는 사실은 언론 산업계가 이러한 불균형에 대응하기 시작했음을 시사한다.

▲ 공정한 콘텐츠 보상 체계 마련: ProRata의 AI 검색 수익 분배 솔루션

AI 검색 서비스 중 하나인 ProRata는 단순한 차단을 넘어 지속가능한 콘텐츠 보상 체계 구축에 참고할 만한 모델이다. 이 시스템은 AI가 생성한 답변을 실시간으로 분석하여 각 출처의 기여도를 100밀리초 내에 계산한다. 예를 들어 〈The Atlantic〉에서 20%, 〈Financial Times〉에서 5% 등 각 언론사의 기여도를 정확히 측정하고, 광고 수익의 50%를 이러한 비율에 따라 콘텐츠 제공자에게 배분한다.

ProRata의 핵심은 표절 방지 소프트웨어와 유사한 기술을 활용하여 AI 답변의 모든 문장을 500개 이상 언론사의 허가받은 데이터셋과 대조하는 것이다. 니콜라스 톰슨 〈The Atlantic〉 CEO가 “LLM 솔루션의 근본 수준에서 문제를 해결한다”며 ProRata 이사회 참여를 결정한 것처럼⁴⁾, 이는 단순한 라이선싱 계약을 넘어 구조적 해결책을 제시한다.

▲ 사용자 경험의 혁신: e-커머스 벤치마킹

최근 세계뉴스미디어협회(INMA) 라틴아메리카 컨퍼런스에서는 뉴스 미디어들이 e-커머스 산업의 개인화 전략을 벤치마킹해야 한다는 주장이 제기되었다⁵⁾. e-커머스가 이용자들의 구매 이력과 브라우징 습관을 활용하여 맞춤형 상품을 추천하듯, 뉴스 미디어도 보유한 1차 데이터(first-party data)를 잘 활용해야 한다. 사용자가 선호하는 콘텐츠 형식(오디오, 비디오, 텍스트)과 관심 주제를 파악하고, 장바구니에 담았다가 구매하지 않은 상품을 재노출하듯 이용자가 관심을 표현한 뉴스 주제를 전략적으로 다시 제공하는 것이다.

특히 뉴스레터는 의도적 구독을 통해 충성도 높은 독자층을 확보할 수 있는 핵심 도구이며, 아침과 오후에 다른 광고를 노출하는 시간대별 개인화, 특정 주제와 관련 있는 브랜드 콘텐츠의 자연스러운 통합은 수익성과 사용자 경험을 동시에 향상시킬 수 있다. 또한 원클릭 구독 같은 마찰 없는 UX 설계는 AI 챗봇의 편의성과 경쟁하기 위한 필수 요소이다.

3) <https://www.aitimes.com/news/articleView.html?idxno=200296&page=5&total=26290>

4) <https://www.cjr.org/analysis/cash-infusion-chatbots-openai-microsoft-bill-gross-goto-chatgpt.php>

5) <https://www.inma.org/blogs/conference/post.cfm/e-commerce-gives-news-companies-a-blueprint-for-personalising-ux>

2025년
8호

KPF 미디어브리프

K P F M e d i a B r i e f W e b z i n e

발행인 김효자

편집인 남정호

기 획 한국언론진흥재단 미디어연구센터

발행일 2025년 8월 20일

한국언론진흥재단 미디어연구센터

04520 서울특별시 종구 세종대로 124 프레스센터빌딩 13층

전화 (02) 2001-7741 팩스 (02) 2001-7740

www.kpf.or.kr

편집 (주)나눔커뮤니케이션

04034 서울특별시 마포구 잔다리로7길 16 교평빌딩 304호

전화 (02) 333-7136 팩스 (02) 333-7146

©한국언론진흥재단 미디어연구센터 2025

〈KPF미디어브리프〉는 한국언론진흥재단의 정부광고 수수료로 제작되었습니다.

자세한 내용은 QR코드를 통해 확인해 주세요.

