2025. 9.

일본 닛케이·아사히 신문, 퍼플렉시티를 저작권 침해로 제소

일본 도쿄대학교 법학부 교수
장예영

1. 개요

2025년 8월 26일, 일본경제신문사(닛케이)와 아사히신문사는 미국의 AI 검색 서비스 퍼플렉시티(Perplexity)를 도쿄지방법원에 제소했다고 발표했다[1]. 소장에 따르면 퍼플렉시티는 2024년 6월경부터 기사 제공을 거부하는 기술적 조치를 무시하고 닛케이와 아사히 서버에 접근하여 기사 원문을 수집해 요약하고 배포했다. 이 과정에서 저작권법상 복제·번안·공중송신권을 침해했을 뿐 아니라, 요약 내용의 오류로 인해 닛케이와 아사히의 사회적 신뢰를 훼손하고 부정경쟁방지법을 위반했다고 주장했다. 특히 닛케이의 경우, 유료 회원 전용(페이월) 기사가 포함되었으며, 아사히는 야후뉴스 제공 기사도 무단 사용된 것으로 알려졌다. 양사는 허가 없는 데이터의 저장과 복제를 금지하는 자사 이용규약 위반을 근거로 기사 이용 금지, 요약문 삭제, 그리고 이용료 상당액 약 10억 엔과 잘못된 요약으로 인한 영업상 손실을 합산해 각 22억 엔을 청구했다.

앞서 8월 7일, 요미우리신문사도 퍼플렉시티를 도쿄지방법원에 제소했다[2]. 원고 측은 2025년 2월에서 6월 사이에 온라인 기사 약 11만9천 건이 무단 수집되어 복제되었다고 주장하며, 저작권법상 복제권·공중송신권 침해 및 '제로클릭'[3] 현상으로 인한 광고 수익 감소 등을 근거로 서비스 이용 금지와 약 21억 6,800만 엔의 손해배상을 청구했다. 요미우리 측은 "무단 이용은 정확한 보도를 저해해 민주주의 기반을 흔들 수 있다"며 법적 대응의 필요성을 강조했다.

AI 검색 서비스는 인터넷에서 공개된 정보를 자동으로 수집하여 요약해 제공하는 방식으로, 전통적 키워드 검색보다 편리하지만, 저작권으로 보호된 콘텐츠의 무단 이용이 우려되고 있다. 이에 따라 일본신문협회를 비롯한 일본 언론사가 AI 기업을 상대로 제소하는 움직임이 확산되고 있으며, 이번 사건은 일본 대형 언론사들이 AI 사업자를 상대로 제기한 첫 본격적 소송으로 주목받고 있다.

1) 日経 · 朝日、米AI検索パープレキシティを提訴 著作権侵害で (2025年8月26日 日本経済新聞記事)
https://www.nikkei.com/article/DGXZQOUD302I50Q5A730C2000000/

2) 生成AI検索サービスで記事を無断利用、読売新聞が米新興企業に21億円の賠償請求 (2025年8月7日 読売新聞記事)
https://www.yomiuri.co.jp/national/20250807-OYT1T50142/

3) 제로클릭(Zero-Click) 현상은 사용자가 검색엔진 결과 페이지(SERP)에서 링크를 클릭하지 않고도 원하는 정보를 바로 얻는 현상을 말한다.

한국저작권위원회 KOREA COPYRIGHT COMMISSION

2. 일본신문협회 입장의 추이

일본 신문업계는 이전부터 생성형 AI에 대한 우려를 표명해 왔다. 2023년 10월 30일에 일본신문협회가 공표한 "생성형 AI에 관한 기본적인 방침"에서는, 2018년 저작권법 개정으로 도입된 제30조의 4(정보해석)가 AI 기계학습은 "비향유목적 저작물 이용"에 해당한다고 보고 권리자의 대가 회수의 기회를 빼앗지 않는다고 해석된 것을 비판하며, 본 규정이 AI 개발자에게 유리한 내용이나 현재 AI 개발 분야에서 일본 국내 회사들이 뒤쳐져 있는 점을 들며, 과연 AI 개발 촉진에 불가결한 규정인가라는 의문을 제기했었다. 저작권법 제30조의 4에서 말하는 "저작권자의 이익을 부당하게 해하는 행위"의 범위 및 "비향유목적"의 해석을 명확하게 할 것을 정부에 촉구했었다[4].

또한 2023년 11월 2일에는, 내각부가 실시한 공개 의견 제안(Public Comment)에 제출한 "AI 시대의 지식재산권에 관한 의견"을 웹사이트에 공표하여, AI 기계학습에 보도 콘텐츠가 자유롭게 사용되는 것에 대해 저작권자가 거부할 수 있는 구조, 혹은 이용 시에 허락을 받는 제도를 정비하는 것이 필요하다고 주장했었다[5].

2024년 7월 17일에는 "검색증강생성(Retrieval Augmented Generation :RAG)" AI 서비스가 보도 콘텐츠를 무단 이용하고 있으며 저작권 침해에 해당할 가능성이 높다는 강한 우려를 표명하며, 권리자의 의사표시를 존중할 것과 허락 없는 이용의 제한을 요구하는 내용의 「생성 AI에 있어서의 보도 컨텐츠의 무단 이용 등에 관한 성명」을 발표하였다[6]. 검색증강생성 AI 서비스는 외부 데이터베이스와 생성형 AI 모델을 연계하여 통합한 것으로, 이용자가 입력한 키워드나 질문에 생성형 AI가 구성한 문장으로 답변한다. 일본신문협회는 생성형 AI의 기계학습에 보도 콘텐츠를 사용하는 경우는 허락을 받도록 AI 개발사에게 지속해서 요구해 왔으나, 사태가 개선되지 않은 채로 생성형 AI 서비스가 확대됨에 따라 저작권 침해에 대한 강한 위기감을 표명해왔다.

2025년 6월 4일에는 동 협회가 "생성 AI에서 보도 콘텐츠 보호에 관한 성명"을 발표했다[7]. robots.txt 등 기술적 조치를 통해 무단 이용을 거부하는 의사표시를 존중할 것, 정부의 제도적 대응의 조기 정비를 강력히 요구했다.

---

4) 「生成AIに関する基本的な考え方」(2023年10月30日、一般社団法人日本新聞協会)
https://www.pressnet.or.jp/statement/broadcasting/231030_15193.html

5) 「AI時代における知的財産権に関する意見」(2023年11月2日、一般社団法人日本新聞協会)
https://www.pressnet.or.jp/statement/20231106_1.pdf

6) 生成AIにおける報道コンテンツの無断利用等に関する声明 (2024年7月17日、一般社団法人日本新聞協会)
https://www.pressnet.or.jp/statement/broadcasting/240717_15523.html ; 2024년 제13호-[일본] 일본신문협회, 검색연동형 생성형 AI의 저작권 침해 문제를 지적하며 저작권법 개정 촉구(장예영) https://www.copyright.or.kr/information-materials/trend/the-copyright/view.do?brdctsno=53273

7) 生成AIにおける報道コンテンツの保護に関する声明 (2025年6月4日、一般社団法人日本新聞協会)
https://www.pressnet.or.jp/statement/broadcasting/250604_15900.html?utm_source=chatgpt.com

3. 신문사측이 주장하는 법 해석 및 기술 관련 논점 정리

1) 저작권법 제30조의 4 및 제47조의 5의 해석론

정보해석(AI학습) 목적의 저작물 이용에 관한 예외 규정인 일본 저작권법 제30조의 4는 저작물을 "스스로 향유할 것, 또는 타인에게 향유시킬 것을 목적으로 하지 않는 경우 (비향유 목적)"에는, 필요한 한도 내에서 저작자의 동의 없이 저작물을 이용할 수 있다고 규정한다. 단, 저작권자의 이익을 부당하게 해하는 경우에는 이용할 수 없다.

AI의 개발과정에서 기계학습(머신러닝)을 위해 타인의 저작물을 이용하는 경우, 일본법 상으로는 이 제30조의4를 근거로 하는 경우가 많기 때문에, '비향유 목적'이 무엇을 의미하는가, '저작권자의 이익을 부당하게 침해하는 경우'가 무엇인가가 중요한데, 여러가지 해석과 견해가 있을 수 있고, 각 이해관계자 별로 다양한 의견이 주장되는 부분이기도 하다. 신문사 측은 대규모로 기사를 무단 취득한 점 및 robots.txt 등의 수집 거부 의사표시를 무시한 것을「저작권자의 이익을 부당하게 해치는 경우」근거로 들고 있다.

저작물의 경미한 이용에 관한 예외규정인 저작권법 제47조의 5는, '①URL 표시 등 검색 대상 정보의 소재 제공, ②검색 목적에 비추어 필요한 한도 내에서 경미한 범위의 이용에 한해, ③저작권자의 이익을 부당하게 해하지 않는다'라는 요건을 전부 충족하면 허락 없이 타인의 저작물을 이용할 수 있다고 규정하고 있다. 검색증강생성(RAG) 등에서 생성된 답변이「경미한 이용」을 넘는 이용인지 여부가 초점인데, 신문사 측은 퍼플렉시티가 응답을 생성하는 과정에서 각 신문사의 서버로부터 기사를 무단으로 복제하고, 그 복제 기사를 바탕으로 실질적으로 동일하거나 유사한 내용을 이용자에게 전송한 것은 경미한 이용을 넘어, 복제권 및 공중송신권 침해에 해당한다고 주장하고 있다.

일본 저작권법 제30조의 4 및 47조의 5에 해당하면, 원칙적으로 저작권자의 허락 없이 저작물을 이용할 수 있지만,「저작권자의 이익을 부당하게 해치는 경우」에는 예외로서 허락이 필요한데, 어떤 경우가 "저작권자의 이익을 부당하게 해치게 되는 경우"인지 명확한 규정은 없다. 일본 문화청이 2024년 3월 발표한「AI와 저작권에 관한 보고서[8]」에서는, 이 두 규정을 중심으로 현행법의 해석의 입장을 정리했으나 구속력은 없고, 현재로써 저작권법 개정은 예정되어 있지 않다.

2) 부정경쟁방지법 위반의 주장

닛케이·아사히의 소송이 요미우리의 소송과 다른 점은 부정경쟁방지법 위반을 함께 주장했다는 것이다. 퍼플렉시티는 검색 결과로 제공하는 요약문에 닛케이 전자판 등 원문 기사의 출처 링크(인용원)를 표시했지

8) 文化庁AIと著作権に関する考え方について（令和6年3月15日） https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/pdf/94037901_01.pdf ; [이슈리포트] 2024-11-[일본] 문화청, AI와 저작권에 관한 보고서 발표(장예영), https://www.copyright.or.kr/information-materials/trend/the-copyright/download.do?brdctsno=52819&brdctsfileno=22996

만, 검색 결과의 내용이 원문 기사와 동일하지 않거나 오류가 있었다. 두 신문사는 이를 통해 언론사로서의 자사의 신용이 훼손되었다며 부정경쟁방지법에 근거해 자사 신문기사를 이용한 검색 결과의 생성 및 제공의 금지를 청구했다. 일본 부정경쟁방지법 제2조 제1항 제21호는 "경쟁 관계에 있는 타인의 영업상 신용을 해하는 허위 사실의 고지·유포"를 부정경쟁행위로 금지한다. 이에 해당하면 금지청구·손해배상·신용회복 조치 등을 청구할 수 있다.

한편 생성 AI가 사실과 다른 내용을 실제처럼 만들어내는 현상은 '할루시네이션(환각)'이라 불린다. AI가 여러 언론사 사이트에서 정보를 수집해 하나의 기사처럼 문장을 합성할 때의 출처를 선정하는 과정은 블랙박스로 남아 있어, 잘못된 결과가 그대로 가짜 정보로 확산될 위험도 우려되고 있다.

3) robots.txt 등 기술적 조치를 우회한 이용의 문제점

robots.txt는 검색 엔진 크롤러(검색 로봇)에게 웹사이트의 특정 페이지나 디렉토리에 접근해도 되는지 허용 여부를 알려주는 텍스트 파일이다. 이를 통해 크롤러가 불필요한 페이지를 탐색하거나 색인하지 않도록 하고, 중요한 페이지에 우선적으로 리소스를 사용하도록 효율성을 높일 수 있다. robots.txt를 이용하여, 검색 결과에 노출되면 안 되는 콘텐츠를 포함한 유료 회원 전용 페이지를 차단한다던지, 유사한 페이지를 크롤링에서 제외해 검색엔진의 혼선을 막는 등의 역할을 한다. 그러나 robots.txt는 검색 엔진에 대한 권고에 불과해, 악성 봇이나 일부 크롤러에는 적용되지 않을 수 있다는 한계가 있다.

일본의 주요 언론사들은 자사 사이트에 robots.txt를 설정해 "기사의 무단 크롤링·스크래핑을 원하지 않는다"는 의사를 명시해왔다. 위의 문화청 「AI와 저작권에 관한 보고서」에서는, 웹상의 데이터베이스 등 저작물을 판매 또는 판매 예정인 경우, AI 학습을 위해 권리자의 이러한 기술적 조치를 회피하여 저작물을 수집·복제하는 행위는 "저작권자의 이익을 부당하게 해치는 경우"에 해당할 수 있다고 언급하고 있다. 법원이 반드시 이 해석을 따라야하는 것은 아니지만, 이러한 관점에 비추어 볼 때 신문사가 robots.txt로 콘텐츠 이용 거부 의사를 분명히 한 경우, 이를 무시하고 데이터를 수집한 행위가 저작권 침해로 인정될 가능성이 있다는 평가가 나온다.

4) 제로클릭 검색의 문제점

제로 클릭 검색이란 사용자가 검색 결과 페이지에서 필요한 정보를 바로 얻어, 출처 웹사이트를 직접 방문하지 않은 채 검색을 마치는 것을 말한다. 사용자에게는 빠르고 편리한 검색 경험을 제공하지만, 원래 클릭을 유도하던 키워드에서 직접 답변이 제공되면 원문의 게재된 뉴스 사이트에의 유입이 줄어들어, 방문자 수 감소로 인한 광고 수익 손실과 구독자 획득 기회의 상실 등의 우려가 있다. 이용자가 링크를 클릭하지 않으면 언론사와 뉴스 제공자의 광고·구독 수익, 나아가 보도의 지속가능성과 취재·편집 체제의 유지에도 영향을 줄 수 있다.

4. 결론 및 시사점

일본 대형 신문사들이 AI 검색서비스 퍼플렉시티를 저작권 침해로 제소한 배경에는 이러한 법해석상의 문제점 및 기술적·비지니스적 상황이 존재한다. 따라서, 데이터의 출처인 신문사 등 언론과 생성형 AI 서비스 제공자 간의 보도 콘텐츠 이용에 관한 분쟁을 해결하기 위해서는, 법적 정비만으로는 부족하고, 기술적·비즈니스적 구조의 정비가 필요하다.

법적 측면에서는 robots.txt와 같은 접근 제한 기술을, 권리자의 의사 표시로써 일정한 법적 효력을 부여하는 것 등을 생각해 볼 수 있다. 비즈니스 측면에서는 미국에서 이미 진행 중인 것처럼 AI 기업과 콘텐츠 제공자 간의 라이선스 계약을 통해 학습데이터 제공자(저작권자)에게 대가를 환원하는 구조를 구축할 필요가 있다. 저작권 관리 단체를 신설해 신문기사의 AI 학습 이용에 관한 라이선스 교섭과 이용료 분배를 일원적으로 담당하는 방안도 고려될 수 있다.

결국 이번 두개의 소송-요미우리 및 닛케이/아사히-의 결과는, 향후의 AI와 저작권에 관련한 법 제도, 기술적 구조, 비즈니스 모델에 큰 영향을 미칠 가능성이 있으며, 한국에도 중요한 시사점을 줄 것으로 보인다.

참고자료

- https://www.nikkei.com/article/DGXZQOUD302I50Q5A730C2000000/
- https://www.yomiuri.co.jp/national/20250807-OYT1T50142/
- https://www.pressnet.or.jp/statement/broadcasting/231030_15193.html
- https://www.pressnet.or.jp/statement/20231106_1.pdf
- https://www.pressnet.or.jp/statement/broadcasting/240717_15523.html
- https://www.pressnet.or.jp/statement/broadcasting/250604_15900.html?utm_source=chatgpt.com
- https://www.copyright.or.kr/information-materials/trend/the-copyright/download.do?brdctsno=52819&brdctsfileno=22996
- https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/pdf/94037901_01.pdf