

#11. 생성형 AI 데이터 학습과 저작권 침해 분쟁

- 뉴욕타임스 vs 오픈 AI 소송으로 본 '공정 이용' 쟁점과 기업의 IP 방어 전략 -

안녕하세요. 특허법인 YNP 입니다.

최근 생성형 AI 기술이 급격히 발전함에 따라, AI 개발사가 웹상의 대규모 저작물을 무단으로 크롤링하여 학습 데이터로 사용하는 행위를 둘러싼 법적 분쟁이 글로벌 쟁점으로 떠오르고 있습니다. 특히 미국 뉴욕타임스(The New York Times)가 오픈 AI(OpenAI) 및 마이크로소프트(Microsoft)를 상대로 제기한 저작권 침해 소송은 저작권자의 권리 보호와 기술 혁신을 위한 데이터 활용이라는 두 가치가 정면으로 충돌하는 상징적인 사건입니다.

이번 연구자료에서는 본 소송의 핵심 쟁점을 분석하고, 국내 콘텐츠 및 IT 기업들이 확보해야 할 전략적 시사점을 살펴보겠습니다.

사건의 개요 및 핵심 쟁점

뉴욕타임스(원고) vs 오픈 AI 및 마이크로소프트(피고)

오픈 AI 등은 뉴욕타임스가 발간한 수백만 건의 기사 및 콘텐츠를 저작권자의 사전 허락 없이 수집하여 자사의 초거대 AI 모델인 챗 GPT(ChatGPT)의 학습 데이터로 활용하였습니다. 이에 뉴욕타임스는 무단 복제 및 자사 저작권 침해를 이유로 대규모 손해배상 및 데이터 파기 청구 소송을 제기하였습니다.

핵심 쟁점: 저작권법상 '공정 이용' 인정 여부

- 오픈 AI 및 MS(피고)의 주장: 저작물을 원형 그대로 대중에게 유포한 것이 아니라, 데이터 분석을 통해 통계적 규칙을 추출하는 변형적 이용에 해당하므로 저작권법상 허용되는 공정 이용이라는 주장입니다.
- 뉴욕타임스(원고)의 주장: 정당한 대가 지불이 없었을 뿐만 아니라, 생성된 AI 결과물이 원본 기사를 그대로 출력하는 등 원본의 시장 가치를 직접적으로 대체하여 심각한 경제적 손실을 입히고 있으므로 공정 이용의 범위를 벗어난 침해라고 반론합니다.

판결 경향 및 법적·경제적 의미

(1) 시장 대체성과 기억 현상에 따른 침해 가능성

현재 진행 중인 소송 과정에서 법원은 AI 학습 자체의 공정 이용 여부를 신중히 검토하면서도, 학습된 AI가 원작의 핵심 문장이나 유료 페이지(Paywall) 뒤의 기사를 지나치게 똑같이 출력하는 현상에 대해서는 저작권 침해 가능성을 무게 있게 들여다보고 있습니다.

(2) 글로벌 데이터 라이선싱 시장으로의 패러다임 전환

본 소송을 기점으로 글로벌 테크 기업들은 무단 학습 기조에서 벗어나 주요 언론사 및 콘텐츠 기업들과 상당한 대가를 지불하고 데이터를 구매하는 빅딜(Data Licensing Agreement) 체결로 급격히 선회하고 있습니다. 이는 데이터가 AI 시대의 핵심 자산이자 비용이라는 인식을 명확히 한 중요한 계기가 되었습니다.

국내 기업의 예상 효과 및 시사점

(1) 콘텐츠 보유 기업: 데이터 무단 크롤링 방지 및 수익화 전략

콘텐츠를 생산하는 국내 기업들은 자사의 자산이 AI 학습에 무단으로 소비되지 않도록 기술적·법적 방어벽을 구축해야 합니다.

- 기업 대응: 웹사이트 약관 내에 'AI 학습 용도의 데이터 수집 및 크롤링 금지' 조항을 신속히 보완해야 합니다. 또한 기술적으로 robots.txt 설정 등을 통해 AI 크롤러의 접근을 차단하는 조치를 취하는 한편, 신뢰할 수 있는 데이터 수요처와 공식적인 라이선스 계약을 통해 새로운 수익 모델을 창출해야 합니다.

(2) AI 기술 개발 기업: 저작권 리스크 사전 헤징 및 투명성 확보

자체 AI 모델을 개발하거나 비즈니스에 도입하려는 기업들은 향후 발생할 수 있는 천문학적인 손해배상 및 서비스 중단 리스크를 선제적으로 제거해야 합니다.

- 기업 대응: 데이터 수집 단계부터 출처가 명확하고 저작권이 해결된 데이터만을 활용하는 프로세스를 정립해야 합니다. 또한 오픈소스나 공공 데이터를 활용할 때도 해당 라이선스의 허용 범위를 면밀히 검토하는 IP 실사 과정을 정례화해야 합니다.

맺음말

뉴욕타임스와 오픈AI 간의 이번 소송은 AI 데이터 학습 과정에서 저작권 침해 요건이 어떻게 성립하고 해석될 수 있는지를 보여주는 구체적인 분쟁 사례입니다. 아직 법적 결론이 나지 않았으나 시장에서는 리스크 회피를 위한 데이터 라이선스 계약이 점차 확산되는 추세입니다. 따라서 국내 기업들 역시 외부 데이터를 연계하여 기술을 개발할 때 발생할 수 있는 잠재적 IP 리스크를 검토하고, 안정적인 지식재산권 관리 체계를 구축하는데 관심을 기울일 필요가 있습니다.

저희 특허법인 YNP 는 최신 기술 트렌드에 따른 지식재산권 분야의 동향을 면밀히 모니터링하고 있으며, 기업들이 연구개발 및 사업화 과정에서 마주하는 다양한 특허 및 지식재산권 보호 전략을 다각도로 지원해 드리겠습니다.

본 자료에 게재된 내용 및 의견은 일반적인 정보제공만을 목적으로 발행된 것이며, 특허법인 YNP 의 공식적인 견해나 어떤 구체적 사안에 대한 의견을 드리는 것이 아님을 알려 드립니다. Copyright ©2026 YNP IP