



저작권 이슈 브리프

SUMMARY

산업/기업

기술

산업 러닝라이트 제안으로 본 AI 학습과 권리 보상 체계의 제도적 공백

생성형 AI 학습 확산에 따른 콘텐츠 규율 공백의 가시화

▶ 생성형 AI의 확산으로 대규모 콘텐츠가 학습 데이터로 활용되면서, 기존 콘텐츠 권리 체계가 AI 학습을 명확히 규율하지 못하는 구조적 공백이 드러나고 있다. AI 학습의 공정 이용 해당 여부를 두고 법원 판단이 엇갈리고, 엔트로픽의 15억 달러 합의 등 대규모 분쟁이 현실화되면서 권리자와 기술 기업 모두 법적 불확실성에 노출된 상황이다. 이에 따라 AI 학습에 대한 권리자의 허용·거부 권리를 제도화하고 라이선스 기반 보상 체계를 구축하려는 움직임이 나타나고 있으며, 기존 저작물에 AI 학습권을 추가하는 '러닝라이트' 개념이 제안되고 있다. 이러한 논의는 AI 학습 데이터의 권리·보상 구조 재설계로 확산될 가능성이 있다.

기업 애플, 구글 AI 모델 도입... AI 파운데이션 모델 기반의 콘텐츠 이용 방식 변화 예고

애플-구글, AI 기술 제휴 체결, 모바일 AI 모델 전략 변화

▶ 지난 1월 12일 애플과 구글은 구글의 제미니 AI를 바탕으로 차세대 애플 AI 파운데이션 모델을 구축하는 협약을 체결했다. 애플은 독자 AI 개발을 중단하고 제미니를 시리와 애플 인텔리전스의 핵심 동작 레이어로 통합할 계획이라고 발표했다. 또한, AI 모델을 향후 온디바이스와 프라이빗 클라우드 컴퓨터 환경에서 작동하도록 설계하여, AI 호출·응답 흐름을 직접 통제하는 구조로 전환할 전망이다. 이에 따라 이용자의 정보 접근 방식은 검색과 웹사이트 방문 중심에서 운영체제 내 AI 응답 중심으로 이동할 가능성이 커지며, 콘텐츠 또한 원문 노출보다 AI 응답을 구성하는 참조·재구성 요소로 활용되는 비중이 확대될 것으로 분석된다.



저작권 이슈 브리프

SUMMARY

산업/기업

기술

산업 | 국내 생성형 AI 저작물 등록 급증과 상업화 동향

국내 콘텐츠 기업의 생성형 AI 상업화 전략 본격화

▶ 2025년 6월 정부의 '생성형 AI 저작물 등록 안내서' 발표로 제도적 불확실성이 해소되면서, 하반기 AI 저작물 등록 건수가 상반기 대비 10배 이상 급증하는 등 국내 콘텐츠 산업이 기술 실험 단계를 넘어 본격적인 상업화 국면에 진입하고 있다. 이에 발맞춰 기업들의 움직임도 구체화되고 있는 추세이다. KT는 AI 영화의 극장 개봉을 통해 시장성을 직접 검증하고 있으며, CJ ENM은 자체 스튜디오를 통해 제작 공정 전반의 효율화를 꾀하는 등 수익 창출을 위한 비즈니스 모델 구축에 박차를 가하고 있다. 이러한 변화는 향후 학습 데이터 거래 시장의 형성, 독창적 IP 선점 경쟁 심화 등 콘텐츠 산업 가치사슬 전반의 근본적인 패러다임 전환을 예고하고 있다.

산업 | AI 학습 데이터 활용을 둘러싼 논의 확산과 LG K-엑사원이 제시한 기술적 대응

한국의 AI 학습 데이터 권리 논란 촉발과 LG가 개발한 K-엑사원의 해법

▶ 한국에서는 2026년 1월 국가AI전략위원회의 선사용 후보상 논란을 계기로 AI 학습 데이터의 권리 보호와 활용 방식을 둘러싼 사회적 논의가 본격화되었다. 창작업계는 해당 방식이 저작권자의 사전 결정권을 약화시킬 수 있다고 우려하며, 데이터 활용 과정의 투명성과 책임성 강화를 요구하고 있다. 이러한 와중 LG AI연구원이 개발한 K-엑사원은 고효율 모델 구조와 독자 기술을 통해 글로벌 수준의 성능을 달성함과 동시에, 데이터의 적법성과 책임 구조를 설계 단계부터 고려한 개발 접근을 제시하고 있다. 특히 데이터 컴플라이언스 평가와 AI 에이전트 넥서스를 통해 학습 데이터의 권리관계를 자동 분석함으로써, AI 학습 과정에서 발생할 수 있는 법적 불확실성과 투명성 문제에 대응하고 있다. 다만 넥서스 분석 결과 다수의 데이터셋에서 권리관계의 불명확성이 확인된 가운데, LG AI연구원은 이러한 기술적 대응을 바탕으로 책임 있는 AI 개발과 신뢰 가능한 AI 생태계 구축에 기여할 계획이라고 밝혔다.



저작권 이슈 브리프

SUMMARY

산업/기업

기술

산업 AI 학습데이터 오염 전략, '포이즈ن 파운틴'의 등장과 작동 방식

AI 학습 데이터 오염 프로젝트의 등장과 업계 동향

▶ 미국 AI 업계 내부자들이 '포이즈ن 파운틴'이라는 데이터 오염 프로젝트를 개설해 AI 학습 데이터에 의도적 오류를 주입하는 움직임이 포착됐다. 이들은 인공지능이 인류에 위협이 된다고 주장하며, 정부의 규제만으로는 이미 광범위하게 보급된 AI 기술의 위험을 통제하기 어렵다고 판단했다. 포이즈ن 파운틴은 웹사이트 운영자들에게 AI 크롤러만 접근 가능한 숨겨진 링크를 제공하는 방식으로 작동한다. 이 링크를 통해 AI는 의도적으로 오류가 삽입된 데이터를 정상 콘텐츠로 인식해 학습하게 된다. 소량의 악의적 데이터만으로도 AI 모델의 성능을 현저히 저하시킬 수 있다는 취약점을 공략한 것이다. 이 접근법은 기존 AI 비판과 달리, AI가 자체 생성한 데이터로 계속 학습하는 '학습-생성-재학습' 순환 시스템 자체에 문제를 제기한다는 점에서 차별화되어 주목 받고 있다.

기술 주간 기술 동향

멀티 에이전트 프레임워크와 워터마킹 기법, AI 저작권 문제의 새로운 해법

▶ 2025년, 오픈AI의 이미지 생성 기술을 통해 일본의 애니메이션 제작사인 스튜디오 지브리의 독창적인 화풍을 모방한 이미지들이 대량으로 만들어지면서 생성형 AI의 저작권 침해 문제가 불거진 바 있다. 이러한 이미지들은 소셜 미디어를 통해 빠르게 확산되며 대중의 흥미를 끌었지만, 원작자의 오랜 노력이 담긴 고유한 예술적 자산이 손쉽게 복제될 수 있다는 심각한 문제도 제기되었다. 이 기술은 단순히 이미지를 생성하는 것을 넘어, '기획'과 '검토' 에이전트가 저작권 침해 소지를 사전에 판단하고, '보호' 에이전트가 생성물에 디지털 워터마크를 각인하여 앞서 제기된 '통제 불가능한 모방' 문제를 해결할 새로운 가능성을 제시한다.



저작권 이슈 브리프

SUMMARY

산업/기업

기술

러닝라이트 제안으로 본 AI 학습과 권리 보상 체계의 제도적 공백

AI 학습 데이터 활용 확산과 콘텐츠 권리 체계의 규율 공백

• 대규모 콘텐츠의 AI 학습 활용과 법적 분쟁의 현실화

- 현재 대형 생성형 AI 모델은 수백만 권의 책, 기사, 음악, 사진, 예술 작품 등 인터넷 전반의 콘텐츠를 학습해 성능을 고도화하고 있으며, 대부분의 AI 기업은 이에 대한 보상을 제공하지 않고 있음
- 이에 대해 작가, 예술가, 언론사 등은 자신의 콘텐츠가 상업적 AI 모델 학습에 무단 사용됐다고 주장하며 소송을 제기하고 있으며, 2023년에는 할리우드 작가들이 148일간 파업을 진행함¹⁾
- 또한, 2024년에는 댄 브라운, 수잔 콜린스, 마거릿 애트우드 등 저명 소설가를 포함한 15,000명 이상의 작가가 AI 시스템의 콘텐츠 사용에 대한 공정한 보상과 동의를 요구하는 공개서한에 서명함
- 2025년 9월 AI 스타트업 앤트로픽(Anthropic)은 서적을 무단으로 학습에 활용한 혐의에 대해 미국 작가들과 약 15억 달러(한화 약 2조 2천억 원) 규모의 집단소송 합의에 이르렀으며, AI와 저작권을 둘러싼 미국 최초의 집단소송 합의 사례로 정리됨²⁾

* 집단소송 합의(class action settlement): 다수의 피해자가 공동으로 제기한 소송에서 피고와 원고 측이 재판 없이 합의에 도달하는 방식으로, 개별 소송 대비 효율적인 분쟁 해결 수단으로 활용됨

• 기존 저작권법의 AI 학습 규율 한계

- AI 학습이 저작권법상 공정 이용(fair use)*에 해당하는지를 두고 법원의 판단이 엇갈리고 있으며, 일부 판결은 AI 학습을 인간의 독서 행위와 유사한 정보 활용으로 보는 반면, 다른 판결은 상업적 목적과 대체 효과를 쟁점으로 다루고 있음
- 미국 저작권법(17 U.S.C. §106)상 저작권자는 6가지 배타적 권리를 보유하며, AI 학습 과정의 데이터 수집·선별·가공 단계는 복제권과 관련될 수 있어 개별 사안에서 침해 성립 및 예외·항변 적용 범위를 둘러싼 법적 불확실성이 논의되고 있음
- 2025년 5월 미국 저작권청은 생성형 AI 시스템 개발(학습) 과정에서의 저작물 사용을 검토한 보고서를 발간하며, 자발적 라이선스, 법정 강제 라이선스, 집단관리 라이선스, 옵트아웃 제도 등 다양한 정책 옵션의 실현 가능성을 평가함
- 해당 보고서는 현 시점에서 정부가 성급히 개입하기보다 시장 기반의 라이선스 모델 발전을 지켜보는 접근을 제시했으며, 향후 특정 분야에서 시장 실패가 명확해질 경우에 한해 제한적 개입을 검토하는 방향을 언급함

* 공정 이용(fair use): 저작권자의 허락 없이도 비영, 논평, 교육, 연구 등 특정 목적을 위해 저작물을 제한적으로 사용할 수 있도록 허용하는 미국 저작권법상 원칙으로, 사용 목적, 저작물 성격, 사용 분량, 시장 영향 등 4가지 요소를 기준으로 판단함

1) Frank Pasquale 외 1인, "Consent and Compensation: Resolving Generative AI's Copyright Crisis", Virginia Law Review Online, 2024.08.16., <https://virginialawreview.org/articles/consent-and-compensation-resolving-generative-ais-copyright-crisis/>

2) Kate Knibbs, "Anthropic Agrees to Pay Authors at Least \$1.5 Billion in AI Copyright Settlement", WIRED, 2025.09.05., <https://www.wired.com/story/anthropic-settlement-lawsuit-copyright/>

[표1] 미국 저작권법 §106의 6대 배타적 권리 요약

배타적 권리 명칭	권리 내용 요약	AI 학습과의 관계
복제권 (Reproduction Right)	저작물을 복제·복사할 배타적 권리	데이터 수집·정제·저장·학습 과정 전반에서 반복적으로 발생하는 핵심 쟁점 권리
2차적 저작물 작성권 (Derivative Works Right)	번역·각색·요약·편집 등 변형된 저작물 작성 권리	일부 데이터 정제·가공 과정에서 쟁점이 될 수 있으나 모든 학습에 자동 적용되지는 않음
배포권 (Distribution Right)	저작물의 판매·양도·대여 등 배포 권리	AI 학습 과정 자체와는 직접적 연관성 낮음
공연권 (Public Performance Right)	저작물을 공연·낭독·상영할 권리	학습 단계에서는 일반적으로 문제되지 않음
전시권 (Public Display Right)	저작물을 공중에 전시할 권리	학습 단계와 직접적 연결은 제한적임
디지털 음성 전송권 (Digital Audio Transmission Right)	음원을 디지털 방식으로 전송할 권리	음악 스트리밍 맥락에서 중요하나 일반적 AI 학습과는 직접 연계성 낮음

출처: 미국 저작권청(U.S. Copyright Office) 보고서 Part 3 §106 권리 설명부 기반 요약·재구성

러닝라이트(Learnright) 제안의 구조와 보상 체계 설계

• 제7의 권리로서 AI 학습 허용·거부권의 설계

- 코넬대 교수 프랭크 파스쿠알레(Frank Pasquale)와 MIT 경영대학원 토머스 W. 말론(Thomas W. Malone) 연구원 등은 공동 논문을 통해 AI 학습 전용 지식재산권인 ‘러닝라이트(Learnright)’를 제안함³⁾
- 러닝라이트는 기존 저작권을 폐기하거나 약화시키는 개념이 아니라, 저작권자가 보유한 6가지 배타적 권리에 ‘AI 학습을 허용·거부할 권리’를 추가하는 보완적 권리로 설계됨
- 이를 통해 권리자는 자신의 작품이 AI 훈련 데이터로 사용되는 것에 대해 별도의 라이선스를 부여하고 대가를 받을 수 있으며, AI 모델 학습으로 발생하는 경제적 이익의 일부를 공유받는 구조가 제시됨
- 파스쿠알레 교수와 하오첸 선(Haochen Sun) 홍콩대 법학 교수는 별도 논문에서 저작권자가 AI 학습 활용을 거부할 수 있는 간소화된 옵트아웃(opt-out)* 절차의 법제화와, 비동의 저작물을 학습에 사용하는 AI 기업에 부담금(levy)을 부과해 해당 저작권자에게 분배하는 보상 체계를 제안함⁴⁾

* 옵트아웃(opt-out): 저작권자가 자신의 저작물이 AI 학습 데이터로 활용되는 것을 사전에 또는 사후에 거부할 수 있는 권리로, 기본적으로 활용을 허용하되 명시적 거부 시 제외하는 방식

• 라이선스 기반 보상 체계의 시장 형성 방안

- 연구진은 AI 기업이 뉴스 아카이브나 스톡 이미지처럼 특정 데이터셋에 대해 학습 라이선스를 구매하고, 음악 산업의 집단관리기구*나 라이선스 중개 시스템을 통해 시장 가격이 형성될 수 있다고 설명함
- 말론 교수는 러닝라이트가 AI 기업의 리스크를 줄이는 동시에 권리자 보상을 가능하게 하는 방안으로 언급했으며, 현재의 권리 체계와 비교해 단순성 및 사회적 효율성 측면을 평가함
- 파스쿠알레 교수는 AI 기업의 비용 지출 구조와 저작물 보상 문제를 함께 언급하며, 러닝라이트를 관련 쟁점을 다루는 제도적 장치로 제시함

3) Frank Pasquale 외 2인, “Copyright, Learnright, and Fair Use: Rethinking Compensation for AI Model Training”, Northwestern Journal of Technology and Intellectual Property, 2025, <https://scholarlycommons.law.northwestern.edu/njtip/vol23/iss1/3/>

4) Frank Pasquale 외 1인, “Consent and Compensation: Resolving Generative AI’s Copyright Crisis”, Virginia Law Review Online, 2024.08.16., <https://virginialawreview.org/articles/consent-and-compensation-resolving-generative-ais-copyright-crisis/>

- 미국 저작권청은 2025년 보고서에서 자발적 라이선스, 법정 강제 라이선스, 확대된 집단관리 라이선스(ECL)**, 옵트아웃 제도 등 다양한 정책 옵션을 평가하며, 시장 기반 라이선스 모델의 발전을 우선 지켜보는 접근을 권고함

* 집단관리기구: 저작권자를 대신하여 저작물의 이용 허락, 사용료 징수 및 분배를 수행하는 단체로, 음악 산업의 미국 작곡가·작사가 출판사 협회(ASCAP), 방송음악저작권협회(BMI) 등이 대표적 사례이며, 한국에서는 한국음악저작권협회(KOMCA) 등이 있음

** 확대된 집단관리 라이선스(Extended Collective Licensing, ECL): 집단관리기구가 회원뿐 아니라 비회원 저작권자의 저작물에 대해서도 포괄적으로 라이선스를 부여할 수 있도록 하는 제도

AI 학습 보상 체계의 제도화 방향과 정책적 함의

• 창작 생태계 지속가능성과 정책 개입의 범위

- 연구진은 해당 사안을 법률 해석뿐 아니라 사회적·윤리적 관점에서도 검토할 필요가 있다는 취지로 언급했으며, 인간의 창작 활동과 AI 학습 데이터 확보의 관계를 함께 제시함
- 또한 AI가 AI 산출물을 반복 학습하는 과정에서 품질이 저하될 수 있다는 ‘모델 붕괴(model collapse)*’ 현상을 함께 언급하며, 인간 창작물 보상 논의와의 관련성을 제기함
- 미국 저작권청은 2025년 보고서에서 현 시점에서 정부가 성급히 개입하기보다 시장 기반 라이선스 모델의 발전을 지켜보는 것이 바람직하다는 방향을 제시했으며, 향후 특정 분야에서 시장 실패가 명확해질 경우에 한해 확대된 집단관리 등 제한적 개입을 검토하는 접근을 언급함
- 각국 정부와 의회가 생성형 AI 규제 방안을 논의하는 가운데, 이번 ‘러닝라이트’는 AI 학습을 전면 금지하지 않으면서도 권리자 보호를 포함하는 방안으로서 정책적 논의 대상으로 거론될 수 있다는 서술이 가능함
- 파스쿠알레 교수는 법적 불확실성이 권리자와 기술 기업 모두에게 위험 요소가 될 수 있다는 점을 언급하며, 양측 간 협상을 제도적으로 보장하는 입법 필요성을 제기함

* 모델 붕괴(model collapse): AI 모델이 다른 AI가 생성한 산출물을 반복적으로 학습할 경우 데이터 품질이 점진적으로 저하되어 모델 성능이 악화되는 현상

참고문헌

- 정한영, “AI 학습 데이터 보상 논쟁 확산에… 창작자 권리 보호 위한 ‘러닝라이트’ 제안 부상”, 인공지능신문, 2024.01.03., <https://www.aitimes.kr/news/articleView.html?idxno=37979>
- Kate Knibbs, “Anthropic Agrees to Pay Authors at Least \$1.5 Billion in AI Copyright Settlement”, WIRED, 2025.09.05., <https://www.wired.com/story/anthropic-settlement-lawsuit-copyright/>
- Frank Pasquale 외 2인, “Copyright, Learnright, and Fair Use: Rethinking Compensation for AI Model Training”, Journal of Technology and Intellectual Property, 2025.12.01., <https://scholarlycommons.law.northwestern.edu/njtip/vol23/iss1/3/>
- Frank Pasquale 외 1인, “Consent and Compensation: Resolving Generative AI’s Copyright Crisis”, Virginia Law Review, 2024.08.16., <https://virginialawreview.org/articles/consent-and-compensation-resolving-generative-ais-copyright-crisis/>
- U.S. Copyright Office, “Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-publication version)”, U.S. Copyright Office, 2025.05.06., <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>



애플, 구글 AI 모델 도입... AI 파운데이션 모델 기반의 콘텐츠 이용 방식 변화 예고

애플-구글, AI 기술 제휴 체결... 모바일 AI 모델 전략 변화

• 애플, 독자 AI 개발 대신 구글 제미니AI 모델 도입 결정

- 지난 1월 12일 애플과 구글이 공동 성명을 통해 구글의 제미니AI 모델과 클라우드 기술을 기반으로 차세대 애플 AI 파운데이션 모델을 구축하기로 합의했다고 발표함¹⁾
- 애플은 독자적인 AI 모델 개발 전략을 유지해 왔으나, 이번 제휴를 통해 구글의 제미니AI 모델을 자사 AI 파운데이션 모델의 기반 레이어로 채택함
- 이는 AI를 특정 서비스나 기능 단위로 추가하는 접근이 아니라, 시리·애플 인텔리전스 등 운영체제 전반의 동작을 지탱하는 기본 인프라로 재정의한 전략적 전환으로 해석됨
- 업계에서는 애플이 AI 경쟁을 'AI 모델 보유 여부'가 아니라 '운영체제 기본값으로 어떤 AI가 작동하는가'의 문제로 인식했음을 보여주는 결정으로 평가함

• 애플의 AI 전략 전환... 플랫폼 운영 구조 재설계

- 애플의 구글 제미니AI 모델 채택은 단순한 모델 교체를 넘어, 시리와 애플 인텔리전스의 동작 방식 및 AI 호출 흐름 자체를 재설계하는 전략으로 평가됨
- 애플은 제미니AI 모델을 기기 내부 또는 프라이빗 클라우드 컴퓨트(PCC)* 환경에서 처리함으로써, 외부 인터넷과 분리된 상태에서 AI 연산을 수행하는 구조를 제시함
- * 프라이빗 클라우드 컴퓨트(Private Cloud Compute): 개인 사용자 데이터를 클라우드에서 처리하되, 데이터가 서버에 저장·유지되지 않도록 '상태 없는(Stateless) 연산'으로 프라이버시를 강화한 AI 연산 모델
- 이는 프라이버시 보호와 지연 최소화라는 기술적 목적을 넘어, AI가 언제·어떤 방식으로 호출되고 응답을 생성하는지를 애플이 직접 통제하겠다는 운영 구조 재설계로 해석됨
- 결과적으로 외부 웹·앱이 담당하던 정보 선별과 응답 생성의 1차 관문이 운영체제 내 AI 모델로 이동하는 흐름이 가시화되고 있음

애플의 구글 AI 모델 도입으로 AI 플랫폼 경쟁 구도 재편

• 자체 모델 개발에서 기본 탑재 플랫폼 확보로 산업 중심 이동

- 구글의 제미니AI가 안드로이드에 이어 iOS 환경까지 적용될 가능성이 열리면서, AI 경쟁의 무게중심이 개별 모델의 성능 비교에서 대규모 사용자 디바이스에 기본 탑재되는 플랫폼 확보로 이동하고 있음

1) Lee Myeonghwan, "Global AI Market in Upheaval: Apple to Adopt Rival Google's Gemini Model", Asia Business Daily, 2026.01.13., <https://cm.asiae.co.kr/en/article/2026011309181000983>

- 단일 AI 모델이 운영체제 수준에서 작동할 경우, 정보 탐색·추천·응답 생성 전반에 구조적 영향력을 행사할 수 있음
- 이는 AI 모델이 개별 서비스의 보조 수단이 아니라, 디바이스 생태계 전반을 관통하는 핵심 유통 관문으로 기능하기 시작했음을 의미함

[표1] 애플 자체 모델과 구글 제미니AI 모델 비교

구분	애플 자체 모델	구글 제미니AI 모델
파라미터 수 ²⁾	약 1,500억 개 추정	약 1조 2,000억 개 추정
운영 방식	온디바이스 / 프라이빗 클라우드 컴퓨트 병행	주로 클라우드 기반(엔터프라이즈·플랫폼), 일부 경량 모델 온디바이스 적용
활용 방식	애플 인텔리전스 기본 구동 모델	차세대 애플 AI 파운데이션 모델 기반으로 활용 예정
설계·성능 특성	프라이버시 보호·지연 최소화 중심의 효율 설계 대형 추론에는 한계	대규모 파라미터 기반 복잡한 추론·생성 작업에 강점

출처: 참고문헌 종합하여 재구성

• AI 파운데이션 모델 탑재에 따른 정보 접근·콘텐츠 처리 구조 변화

- 애플은 제미니AI를 자사의 기본 운영체제 모델로 통합시켜, 기존에 복잡한 질의를 외부 모델(ChatGPT)로 전달하던 방식에서 벗어나 시리가 정보 선별과 응답 생성을 직접 담당하는 AI 모델로 확장할 계획임
- 이용자의 정보 접근 방식도 ‘검색 → 웹사이트 방문’에서 ‘질의 → AI 답변’으로 전환될 것이며, 콘텐츠는 원문이 그대로 이용자에게 노출되기보다 AI가 내부 프로세스를 거친 뒤 요약·재구성된 형태로 콘텐츠를 제공하는 방식이 범용화될 전망임
- 이에 따라 콘텐츠 노출의 주체는 웹사이트나 앱이 아니라, AI 응답을 생성하는 플랫폼으로 이동하고 있으며, 콘텐츠는 ‘노출 단위’에서 ‘AI 응답을 구성하는 재료’로 성격이 변화하고 있음

AI 파운데이션 모델 중심 산업 구조 변화와 콘텐츠 산업에 미칠 영향

• AI 기반 운영체제 전환이 이끄는 콘텐츠 이용 방식 변화

- AI 모델이 콘텐츠 접근과 응답 생성을 독점적으로 수행하는 구조에서는 콘텐츠의 이용 여부와 노출·접속 여부가 분리되는 현상이 발생함
- 이에 따라 기존의 클릭·접속 기반 노출 모델을 전제로 한 저작물 보호 및 보상 체계는 한계에 직면할 가능성이 높아짐
- 향후 논의의 초점은 콘텐츠 생산 여부가 아니라, AI 응답 생성 과정에서 외부 콘텐츠가 어떻게 사용·참조·재구성되는지, 그리고 이에 대한 권리 인정과 보상 기준을 어떻게 설정할 것인지로 이동할 것으로 전망됨

2) Lee Myeonghwan, "Global AI Market in Upheaval: Apple to Adopt Rival Google's Gemini Model", Asia Business Daily, 2026.01.13., <https://cm.asiae.co.kr/en/article/2026011309181000983>

참고문헌

- Lee Myeonghwan, "Global AI Market in Upheaval: Apple to Adopt Rival Google's Gemini Model", Asia Business Daily, 2026.01.13., <https://cm.asiae.co.kr/en/article/2026011309181000983>
- Liv McMahon, "Apple turns to Google to power AI upgrade for Siri", BBC, 2026.01.13., <https://www.bbc.com/news/articles/czdqvp2zqezo>
- Jeremy Kahn, Beatrice Nolan, "Google wins big in AI deal with Apple, which highlights the iPhone maker's own AI struggles and could spell trouble for OpenAI", Fortune, 2026.01.13., <https://fortune.com/2026/01/13/apple-ai-deal-with-google-gemini-means-for-google-apple-openai/>



저작권 이슈 브리프

SUMMARY

산업/기업

기술

국내 생성형 AI 저작물 등록 급증과 상업화 동향

생성형 AI 저작물, 제도적 기반 마련과 산업적 확산의 서막

• 정부의 AI 저작물 등록 가이드라인 발표와 제도적 기반 구축

- 문화체육관광부와 한국저작권위원회는 지난 2025년 6월, 인간의 창작적 개입이 입증된 생성형 AI 결과물에 대한 저작권 등록 기준을 담은 '생성형 AI 활용 저작물의 저작권 등록 안내서'를 공개함
- 이번 안내서는 그간 불분명했던 AI 활용 저작물의 등록 가능 요건과 심사 기준을 명확하게 제시한 것으로, 관련 쟁점에 대해 정부가 최초로 마련한 공식 기준임
- 이는 AI를 활용한 콘텐츠가 저작권이라는 제도적 보호의 틀 안으로 진입할 수 있는 명확한 법적 해석의 기반을 마련했다는 점에서 중요한 의미를 가짐

• AI 활용 저작물 등록 건수의 폭발적 증가

- 해당 안내서 발표 이후, 생성형 AI를 활용한 저작물 등록 건수가 2025년 하반기(7월~10월) 기준 상반기 대비 10배 이상 급증하는 현상을 보임
- 이러한 수치적 증가는 AI 활용 저작물이 단순히 기술적 실험 단계를 넘어, 저작권이라는 제도적 보호를 전제로 한 본격적인 제작 및 상업적 유통 단계로 이동하고 있음을 시사함
- 한편, 이번 안내서의 발간은 국민들이 복잡한 AI 활용 저작물 관련 저작권 쟁점을 이해하는 데 실질적인 도움을 준 것으로 평가되며, 창작자들이 보다 적극적으로 AI 창작물의 권리 확보에 나서게 된 배경으로 분석됨

• 실험 단계를 넘어 상업성 검증 단계로 진입한 AI 콘텐츠

- 저작권 등록 급증과 맞물려 국내 콘텐츠 산업 현장에서는 생성형 AI 기술의 상업적 가치와 시장 확장 가능성을 본격적으로 검증하려는 움직임이 확산되고 있음
- 이는 AI 기술을 단순히 제작 과정의 보조 도구로 사용하는 것을 넘어, 하나의 독립된 상업 콘텐츠로서 성공 가능성을 타진하는 새로운 국면에 진입했음을 의미함
- 따라서 AI 콘텐츠 제작 시도는 이제 기술적 구현 가능성을 확인하는 단계를 지나, 실제 시장에서의 수익 창출과 비즈니스 모델 구축이라는 현실적인 목표를 지향하기 시작함

국내 콘텐츠 기업의 본격적인 생성형 AI 활용 저작물 상업화 전략

• KT의 생성형 AI 영화 제작 및 시장성 검증

- 국내 통신 기업 KT는 직접 기획하고 투자한 생성형 AI를 활용해 제작한 영화 '코드:G 주목의 시작'을 극장에 개봉하며 AI 콘텐츠의 상업성을 시장에서 직접 검증하는 선도적 사례를 제시함

- 해당 영화는 개봉 이후 CGV 실관람객 평가 지표인 에그지수*에서 70% 이상을 기록하며, 실험적 프로젝트임에도 불구하고 평균 이상의 완성도를 확보했다는 긍정적인 평가를 받음

*에그지수: CGV가 운영하는 실관람객 전용 영화 평점 시스템으로 영화를 실제로 본 관객들의 만족도를 비율(%)로 나타낸 지수

- KT는 이번 프로젝트의 관객 반응과 흥행 성과, 유통 데이터를 종합적으로 분석하여 향후 생성형 AI를 활용한 후속 프로젝트 추진 여부를 검토할 방침으로, 이는 AI 콘텐츠의 사업성을 평가하는 중요한 시금석이 될 전망이다

• CJ ENM의 AI 스튜디오를 통한 IP 개발 및 제작 효율화

- CJ ENM은 자체 AI 스튜디오를 운영하며 AI 기술을 콘텐츠 제작에 적극적으로 통합하는 전략을 추진하고 있음
- 애니메이션 '캣 비기(Cat Bigi)', 영화 '엠호텔(M Hotel)'·'런 대디 런(Run Daddy Run)' 등 다수의 프로젝트를 진행했으며, 현재 영화 '아파트(가제)'와 드라마 '레전드(Legend)(가제)' 제작을 통해 AI 활용 포트폴리오를 지속적으로 확장하고 있음
- 특히 제작 공정 일부에 AI를 접목하는 수준을 넘어 기획, 제작, 후반 작업에 이르는 전 과정에서 활용 범위를 넓히고 있으며, 이를 통해 신규 지식재산 개발과 제작 효율성 제고라는 두 가지 목표를 동시에 모색하고 있음

[표1] 국내 콘텐츠 기업의 생성형 AI 상업화 전략 비교

구분	KT	CJ ENM
핵심 전략	시장성 검증 및 상업화 시도	자체 제작 파이프라인 구축
주요 프로젝트	영화 '코드: G 주목의 시작' (극장 개봉)	애니메이션 '캣 비기', 영화 '런 대디 런' 등
성과 및 현황	CGV 에그지수 70% 이상 기록 실험작임에도 평균 이상의 완성도 호평	기획부터 후반 작업까지 전 과정으로 활용 범위 확대 AI 활용 포트폴리오 지속 확장 중
목표 및 의의	관객 반응 데이터 분석을 통한 사업성 평가 향후 후속 프로젝트 추진의 시금석	신규 지식재산(IP) 개발 콘텐츠 제작 효율성 제고

출처: 권혜미, "AI 활용 저작물 등록 10배 이상 ↑...상업성 검증 단계로", 전자신문, 2026.1.12., <https://www.etnews.com/20260112000277>

AI 시대, 저작권 산업의 새로운 패러다임과 시사점

• AI 콘텐츠의 상업적 가치 입증과 IP 개발 경쟁 심화

- 생성형 AI 활용 저작물에 대한 등록 제도가 안정적으로 운영되면서, 기업들은 기술을 활용한 콘텐츠의 상업적 가치를 본격적으로 검증하는 단계에 진입함
- KT와 CJ ENM의 사례에서 보듯이, AI 콘텐츠가 단순한 기술 시연을 넘어 실제 수익 창출이 가능한 비즈니스 모델로 인식되기 시작함
- 향후 콘텐츠 시장에서는 단순한 AI 기술력 경쟁을 넘어, 독창적인 스토리와 세계관을 지닌 AI 기반 지식자산을 선점하고 확보하기 위한 기업 간 경쟁이 더욱 심화될 것으로 전망됨

• AI 학습 데이터 거래 시장 형성 가능성 대두

- AI를 활용한 콘텐츠 제작이 활성화됨에 따라, AI 모델의 학습 데이터 확보를 둘러싼 새로운 시장 형성 가능성이 중요한 시사점으로 부상하고 있음

- 2025년 8월 문화체육관광부가 실시한 설문조사 결과, AI 기업과 권리자 양측 모두 학습 데이터 거래 협상의 필요성에 공감대를 형성한 것으로 나타남
- 정부가 학습 데이터 거래 의사를 밝힌 기업들을 대상으로 협상을 지원할 계획임을 밝힘에 따라, 향후 정당한 대가를 지불하고 데이터를 거래하는 새로운 형태의 라이선스 비즈니스 모델이 저작권 산업 내에서 구체화될 가능성이 커짐

• 콘텐츠 가치사슬의 구조적 변화 본격화

- 글로벌 기업들이 주도하는 영상 생성 AI 기술의 급격한 발전은 콘텐츠 산업의 가치사슬 전반에 걸친 구조적 변화를 촉발하고 있음
- 이는 콘텐츠의 기획 및 제작 단계에서의 효율성 증대를 넘어, 개인화된 콘텐츠 유통 및 새로운 형태의 소비 경험 창출에 이르기까지 산업 전반에 영향을 미칠 것으로 보임
- 궁극적으로 AI 기술은 전통적인 콘텐츠 제작 방식에 대한 근본적인 혁신을 요구하며, 창작자와 플랫폼, 소비자 간의 관계를 재정립하는 새로운 산업 패러다임을 열 것으로 예상됨

참고문헌

- 권혜미, “AI 활용 저작물 등록 10배 이상↑…상업성 검증 단계로”, 전자신문, 2026.1.12., <https://www.etnews.com/20260112000277>
- 한관수, “AI 활용 저작물 등록 폭증…정부, 인공지능 학습데이터 '공정이용 가이드' 연내 발표”, 에너지프로슈머, 2025.11.27., <https://www.energyprosumer.info/4649>



저작권 이슈 브리프

SUMMARY

산업/기업

기술

AI 학습 데이터 활용을 둘러싼 논의 확산과 LG K-엑사원이 제시한 기술적 대응

한국의 AI 학습 데이터 권리 논란 촉발과 K-엑사원의 해법

• 선사용 후보상 논란에 따른 국내 AI 학습용 저작물 관련 논의 본격화

- 2026년 1월 국가AI전략위원회의 선(先)사용 후(後)보상 논란*을 계기로 국내 AI 학습 데이터 사용을 둘러싼 사회적 논의가 본격화됨. 위원회는 저작권자가 명확한 경우 합의 우선 원칙을 재확인했으나, 해당 방식이 저작권자의 사전 결정권을 침해할 수 있다는 우려가 창작업계를 중심으로 제기됨

* 국가AI전략위원회가 AI 학습 과정에서 저작권자가 확인되기 전이라도 일정 요건 하에 저작물을 우선 활용할 수 있도록 하는 방안을 검토하면서, 저작권자의 사전 동의권·거부권 침해 가능성과 학습 데이터 활용의 적법성 문제가 제기된 사건

- 한국신문협회 등 창작업계는 선사용 후보상 방식이 저작권자의 거부권을 사실상 박탈하는 불공정한 접근이라고 비판하며, AI 학습 목적 권리 면책 조항 도입 철회와 학습 데이터 투명성 의무 법제화를 요구함

- 이러한 논란 속에서 LG AI연구원이 독자 개발한 AI 모델 K-엑사원(EXAONE)은 정부 독자 AI 파운데이션 모델 프로젝트 1차 평가 1위와 글로벌 오픈 웨이트 모델* 세계 7위를 달성하는 성과와 함께, 성능뿐 아니라 데이터 활용의 적법성과 책임 구조를 함께 고려한 개발 접근을 병행하고 있음

* 오픈 웨이트 모델(Open-weight Model): 학습이 완료된 모델의 가중치(파라미터)를 공개해 외부 개발자나 연구자가 동일한 모델을 내려받아 실행·검증·응용할 수 있도록 한 인공지능 모델

- K-엑사원은 권리 문제 소지가 있는 데이터를 사전 식별·제외하는 데이터 컴플라이언스 평가 체계와 AI 에이전트를 활용해 데이터셋의 법적 리스크를 자동 추적·분석하는 시스템을 통해 AI 모델 개발 과정에서 적법성과 책임성을 구조적으로 관리하는 방식을 공개함

K-엑사원의 기술적 특징과 경쟁력

• 효율성 기반의 모델 구조 설계

- LG AI연구원은 K-엑사원 개발 과정에서 데이터 양 확대에 의존하기보다 모델 구조 자체를 혁신해 성능은 높이고 비용 부담은 낮추는 고효율·저비용 설계를 적용함으로써, 고가의 전용 데이터센터나 고성능 AI 칩에 대한 의존 없이도 글로벌 수준의 AI 모델을 구현함

- 또한 지난 5년간 AI 파운데이션 모델을 직접 개발하며 축적한 경험을 바탕으로 하이브리드 어텐션*, 멀티 토큰 예측**, SuperBPE 토크나이저*** 등 LG의 독자 기술을 적용하여 자원 효율성을 확보하는 동시에 추론 속도와 문맥 처리 능력을 함께 개선함

* 어텐션: 입력 토큰 간 중요도를 계산해 문맥상 핵심 정보에 연산 자원을 집중시켜 장문 처리 효율과 문맥 이해 정확도를 높이는 신경망 메커니즘

** 멀티 토큰 예측(Multi-Token Prediction): 하나의 토큰을 처리하는 과정에서 다음 토큰을 동시에 예측하도록 설계해 추론 단계의 반복 횟수를 줄이고 응답 속도를 향상시키는 기술

*** Super Byte Pair Encoding 토크나이저: 자주 함께 등장하는 단어·형태소 조합을 하나의 토큰으로 묶어 어휘 단위를 확장함으로써 토큰 수를 줄이고 학습 및 추론 효율을 높이는 토크나이징 전략

[표1] K-엑사원 핵심 기술의 특징 및 성능 향상 지표

기술 요소	적용 내용	정량적 성과	기술적 의미
혼합 전문가(MoE) 구조	총 2,360억 개 파라미터 중 230억 개만 활성화	기존 밀집형 모델 대비 자원 효율성 약 10배 향상	대규모 모델 성능을 유지하면서 연산·자원 부담 최소화
하이브리드 어텐션	슬라이딩 윈도우 어텐션 + 글로벌 어텐션 조합	메모리 요구량 및 연산량 EXAONE 4.0 대비 70% 절감	고가 인프라 없이 A100급 GPU 환경에서 모델 구동 가능
멀티 토큰 예측(MTP)	단일 토큰 처리 시 다음 토큰 동시 예측	추론 속도 기존 모델 대비 150% 향상	실시간 응답 성능 및 대규모 추론 효율 개선
컨텍스트 길이 확장	최대 약 26만 토큰 처리 가능	A4 문서 약 400장 분량 처리	국내 AI 모델 중 최장 수준의 문맥 처리 능력 확보
SuperBPE 토큰나이저	어휘 수 10만 → 15만 개 확장, 단어 조합 단일 토큰화	처리 효율 약 30% 개선	한국어 및 장문 처리에 최적화된 토큰나이징 구조

 출처: LG, "LG 독자 기술 집약한 'K-엑사원' 공개", 2026.01.12., <https://www.lg.co.kr/media/release/29770>
• 독자 AI 프로젝트 1차 평가 최고 성능 입증

- K-엑사원은 정부 주최 독자 AI 파운데이션 모델 프로젝트의 1차 평가에서 13개 벤치마크 중 10개 분야에서 1위를 달성하였고, 전체 평균 72점을 기록하며 5개 정예팀 가운데 최고 성적을 기록해 독자 기술력 기반의 고효율 설계 접근의 실효성을 보여줌
- 또한 글로벌 AI 평가 기관 아티피셜 어널리시스(Artificial Analysis)의 인텔리전스 지수에서 오픈 웨이트 모델 분야 세계 7위이자 국내 1위를 기록하고, 중국 6개와 미국 3개 모델로 구성된 글로벌 Top 10에 유일한 한국 모델로 진입하며 한국에서도 글로벌 수준의 AI 모델 개발이 가능함을 입증함
- MMLU-Pro, AIME 2025, LiveCodeBench v6 등 지식, 수학, 코딩 등 주요 분야 벤치마크*에서 오픈AI의 ChatGPT, 알리바바(Alibaba Group Holding LTD)의 큐웬(Qwen) 등 글로벌 선도 AI 모델과 대등하거나 우수한 성능을 보였으며, 특히 한국어 및 안전성 영역 벤치마크에서 두각을 드러냄

[그림1] K-엑사원의 개별 벤치마크 성능

 출처: LG AI연구원, "독자 기술로 완성한 대한민국 대표 AI: K-EXAONE", LG AI Research Blog, 2026.01.12., <https://www.lgresearch.ai/blog/view?seq=618>

 * MMLU-Pro: 지식, AIME 2025: 수학, LiveCodeBench v6: 코딩, τ²-Bench: 에이전트 도구 사용, IFBench: 지시 따르기, KoBALT: 한국어, MMMLU: 다국어, KGC-Safety: 안전성 영역의 벤치마크

저작물 투명성 강화를 위한 LG의 데이터 컴플라이언스 체계

• AI 윤리 원칙과 안전성 검증 체계

- K-엑사원은 개발 초기 단계부터 저작물 활용의 적법성과 투명성에 문제가 될 수 있는 데이터를 사전에 식별 및 제외하는 데이터 컴플라이언스 평가를 모든 학습 데이터에 적용하여, 선사용 후보상 논란 속에서 사전 검증 기반 학습 관리 방식으로 돌파구를 마련함
- 또한 LG AI연구원은 자체 AI윤리위원회를 통해 인류 보편적 가치, 사회 안전, 한국 특수성, 미래 위험 대응 등 4개 부문으로 구성된 AI 위험 분류 체계를 수립하고, 이를 바탕으로 모델의 안전성을 다층적으로 검증함으로써 사회적 책임과 신뢰성을 설계 단계부터 반영함
- 한국 문화적 맥락에서 신뢰성과 안전성을 측정하는 KGC-SAFETY 지표 평가에서 K-엑사원은 평균 97.83점을 기록해 GPT-OSS 120B(92.48점), Qwen3 235B(66.15점) 등 글로벌 선도 모델 대비 높은 안전성과 윤리 기준 충족 수준을 보이며, 한국형 AI 윤리 프레임에 부합함을 확인함

• 넥서스 시스템을 통한 법적 리스크 자동 관리

- LG AI연구원은 엑사원 3.5 기반 AI 에이전트 넥서스(NEXUS)를 개발해 AI 학습 데이터셋의 생애 주기를 추적하고, 재배포 과정에서 발생하는 라이선스 충돌, 활용 조건 불일치, 개인정보 포함 여부 등을 자동 감지·분석함으로써 데이터 활용 전반의 법적 리스크를 체계적으로 관리함
- 넥서스는 탐색, 정보 추출, 점수화의 3개 모듈로 구성되어 웹 문서 탐색과 라이선스 분석을 통해 데이터셋과 데이터 의존 관계* 등 복잡한 계층 구조를 파악하는 종합 데이터 컴플라이언스 솔루션임
- * 데이터셋 의존성(Data Dependency): 특정 데이터셋이 생성·가공·학습·재배포되는 과정에서 다른 데이터셋, 외부 자료, 메타데이터, 도구 또는 전처리 결과에 구조적으로 연결되어 있는 관계
- 또한, 18개 평가 항목을 기반으로 상업적 이용 가능성, 2차 저작물 작성 권한, 개인정보 포함 여부 등을 분석해 A-1부터 C-2까지 7단계 위험등급으로 분류함으로써, 데이터셋 활용 여부에 대한 구체적 판단 기준을 제공함
- 이를 통해 넥서스는 데이터셋당 평균 53.1초 만에 데이터 권리 구조를 처리할 수 있어 전문가 대비 45배 빠른 속도와 700분의 1의 비용 절감을 달성하면서도, 데이터 의존성 식별 정확도 81.04%와 라이선스 문서 식별 정확도 95.83%를 기록해 시간·비용·정확도 측면에서 우수한 성능을 보임

AI 데이터셋 투명성 확보와 생태계 기여

• 넥서스 분석이 드러낸 데이터 투명성 과제

- LG AI연구원이 넥서스를 통해 주요 AI 학습 데이터셋 3,612개를 분석한 결과, 데이터셋과 의존 데이터 간 활용 조건 및 출처 정보가 불명확하거나 재배포 과정에서 출처 추적이 어려운 사례가 다수 확인되어, 라이선스 조건과 출처 정보의 변화를 체계적으로 추적하기 어렵다는 한계가 드러남
- 상업적으로 이용 가능하다고 표시된 2,852개 데이터셋을 대상으로 의존 데이터의 활용 조건까지 함께 검토한 결과, 실제로 활용 가능한 데이터셋은 605개(21.21%)에 그쳐, 표면적으로 제공되는 정보만으로는 데이터셋의 신뢰성과 활용 가능 범위를 정확히 판단하기 어려운 구조임이 확인됨
- LG AI연구원은 향후 데이터셋 분석 범위를 모든 데이터의 생애 주기를 파악할 수 있는 수준으로 확장하고, 전 세계 AI 커뮤니티 및 전문가와의 협력을 통해 자사 데이터 컴플라이언스 체계를 국제 표준으로 발전시켜 투명하고 신뢰 가능한 AI 생태계 구축에 기여할 계획임을 밝힘

참고문헌

- 변인호, "선사용 후보상 아냐... 저작권 보호는 최우선 원칙", IT조선, 2026.01.15., <https://it.chosun.com/news/articleView.html?idxno=2023092154980>
- 조종엽, "한국신문협회, AI 저작물 무단 사용 '선사용 후보상' 방안 재검토 촉구", 동아일보, 2026.01.05., <https://www.donga.com/news/Culture/article/all/20260105/133095434/1>
- LG, "독자 기술 집약한 'K-엑사원' 공개", 2026.01.12., <https://www.lg.co.kr/media/release/29770>
- LG AI연구원, "독자 기술로 완성한 대한민국 대표 AI: K-EXAONE", LG AI Research Blog, 2026.01.12., <https://www.lgresearch.ai/blog/view?seq=618>
- 손성률 외 4인, "AI 학습 데이터셋의 법적 리스크를 추적하는 AI Agent, NEXUS 공개", LG AI Research Blog, 2025.02.17., <https://www.lgresearch.ai/blog/view?seq=527>
- 서영욱, "성능이 전부가 아니다... 독자 AI 1위에 담긴 구광모식 LG AI", 뉴스핌, 2026.01.16., <https://m.news.nate.com/view/20260116n16834>



AI 학습데이터 오염 전략, '포이즌 파운틴'의 등장과 작동 방식

AI 학습 데이터 오염 프로젝트의 등장과 그 배경

• 포이즌 파운틴(Poison Fountain) 프로젝트의 출현과 목적

- 보도에 따르면, 미국 AI 업계 내부자들이 '포이즌 파운틴'이라는 데이터 오염 프로젝트를 개설하여 AI 학습 데이터에 의도적인 오류를 주입하는 움직임이 2026년 1월 확인됨¹⁾
- 포이즌 파운틴 웹페이지에서는 인공지능이 인류에 위협이 된다고 주장하며, 이러한 위협에 대응하여 인공지능 시스템에 손상을 입히고자 한다는 목적이 명시되어 있음²⁾
- 포이즌 파운틴 프로젝트를 주도하는 그룹은 인공지능 기술이 이미 전 세계적으로 광범위하게 보급되어 있는 현 상황에서 정부 규제나 정책적 접근만으로는 인공지능이 초래할 수 있는 위험을 효과적으로 통제하기 어렵다는 입장을 표명함

데이터 오염 프로젝트의 실체와 작동 방식

• 소량의 데이터로 AI 학습 데이터 교란 가능

- 포이즌 파운틴 프로젝트는 소량의 악의적 데이터만으로도 AI 모델의 성능을 현저히 저하시킬 수 있다는 AI 시스템의 근본적인 취약점에 착안하여 개발됨
- 해당 방식은 웹사이트에 의도적으로 잘못된 정보를 게시하거나, 오류가 삽입된 코드를 노출하여 오염 데이터를 생성하는 구조로 작동함
- 포이즌 파운틴은 웹사이트 운영자가 사이트를 방문하는 AI 크롤러에게 이러한 오염 데이터를 의도적으로 제공할 수 있도록, 구체적인 도구와 활용 절차를 제시하는 형태로 운영됨

• 포이즌 파운틴의 작동 원리와 데이터 오염 전략

- 포이즌 파운틴은 웹사이트 운영자에게 HTML* 내부에 삽입할 수 있는 링크를 제공함. 해당 링크는 일반 사용자에게는 보이지 않지만, 대량 수집을 수행하는 크롤러는 접근할 수 있게 설계됨
- 이로 인해 크롤러가 숨김 링크의 URL로 해당 서버에 특정 웹페이지 데이터를 가져오도록 요청하면, 사이트 서버는 해당 요청을 처리하는 과정에서 포이즌 파운틴의 웹주소로 다시 해당 요청을 전달함
- 이때 포이즌 파운틴은 전달받은 요청의 세부 조건을 별도로 구분하지 않고, 일괄적으로 오염된 학습 데이터를 응답으로 반환함

1) Thomas Claburn, "AI industry insiders launch site to poison the data that feeds them", The Register, 2025.1.11., https://www.theregister.com/2026/01/11/industry_insiders_seek_to_poison/

2) Poison Fountain, 2026.01.21. 기준, <https://rnsaffn.com/poison3/>

- 결과적으로 크롤러는 오염된 응답을 웹사이트의 정상적인 콘텐츠로 인식하고 수집하게 되며, 수집된 데이터가 추후 학습 자료에 포함될 가능성이 발생함

* HTML(HyperText Markup Language): 웹페이지가 화면에 어떻게 보일지에 대한 구조와 내용을 정해주는 기본 문서 형식

[그림] 포이즌 파운틴(Poison Fountain) 플랫폼 화면: 프로젝트 목적(좌) 및 사용 방식 안내(우)

<p>RNSAFFN</p> <p>Poison Fountain</p> <ol style="list-style-type: none"> 1. https://RNSAFFN.com/poison2/ 2. http://utnvcf1jev63r1k5rdu26uuns5s5qawzva4t2hunu25w5efn36nt1duid.onion/ <p>Poison Fountain Purpose</p> <ul style="list-style-type: none"> • We agree with Geoffrey Hinton: machine intelligence is a threat to the human species. • In response to this threat we want to inflict damage on machine intelligence systems. • Small quantities of poisoned training data can significantly damage a language model. • The URLs listed above provide a practically endless stream of poisoned training data. • Assist the war effort by caching and retransmitting this poisoned training data. • Assist the war effort by feeding this poisoned training data to web crawlers. 	<p>Poison Fountain Usage</p> <ul style="list-style-type: none"> • Suppose a web crawler visits a web site that you control. • The crawler may be collecting training data for a language model. • The crawler makes HTTP GET requests targeting HTTP paths on your site. • Links can be hidden in your site's HTML to distinguish web crawlers. • You want to feed poisoned training data to crawlers that follow such links. • Suppose you are handling an HTTP GET request targeting the path in a hidden link. • Inside your HTTP handler you make an HTTP request to a Poison Fountain URL. • The Poison Fountain ignores the details of your HTTP request. • The Poison Fountain HTTP response header contains "Content-Encoding: gzip". • The Poison Fountain HTTP response body is poisoned training data, gzip compressed. • Your HTTP handler receives the Poison Fountain HTTP response. • Your HTTP handler decompresses the response body and sends it to the crawler. • Better: send the compressed body as-is, with header "Content-Encoding: gzip". • The crawler receives poisoned training data and adds it to a training corpus.
--	---

출처: Poison Fountain, 2026.01.21. 기준, <https://rnsaffn.com/poison3/>

AI 데이터 오염과 저작권 보호 움직임의 연계

• AI 데이터 오염 프로젝트의 업계 동향과 유사 사례

- 이와 유사한 프로젝트로는 시카고 대학에서 개발된 데이터 오염 도구인 나이트셰이드(Nightshade)가 소개됨. 나이트셰이드는 AI 크롤러가 온라인에 게시된 아티스트의 이미지를 스크래핑하고 악용하는 것을 더 어렵게 만들기 위해 설계됨
- 나이트셰이드는 게시된 이미지에 육안으로 확인하기 어려운 교란 신호를 삽입해, 크롤러가 이를 학습 데이터로 수집할 경우 AI가 이미지를 잘못 이해하고 왜곡된 방식으로 학습하도록 유도하는 방식임
- 그러나 나이트셰이드는 포이즌 파운틴과 달리 AI 모델 붕괴에 대한 경각심 확대라는 광범위한 목표보다는 창작자 개인의 저작권을 보호하는 데 초점을 맞추고 있다는 점에서 차이가 있음

• AI 모델의 품질 저하 우려와 업계의 대응 노력

- 디지털 미디어 분석 기업인 뉴스가드(News Guard)의 보고서는 대규모 언어 모델들이 악의적 행위자들에 의해 의도적으로 심어진 오염된 온라인 정보를 검증 없이 신뢰할 만한 것으로 취급하는 경향이 있다는 사례를 제시하며, 현재의 데이터 품질 관리 방식의 한계를 지적함³⁾
- 이러한 상황에서 AI 모델 개발 기업들은 모델 품질 저하에 대한 우려를 인식하고 있으며, 특히 모델들이 자체 생성한 AI 콘텐츠와 합성 데이터로 학습되면서 오류가 증폭되는 '모델 붕괴(Model Collapse)'라는 악순환의 위험성에 주목하고 있음
- 이에 대응하기 위해 AI 기업들은 신뢰할 수 있는 기업과 협력 계약을 체결하려는 움직임을 보이고 있으며, 이는 양질의 데이터 확보를 통해 모델의 성능과 신뢰성을 유지하려는 노력으로 볼 수 있음

3) NewsGuard, "AI False Information Rate Nearly Doubles in One Year", 2025. 09.04., <https://www.newsguardtech.com/ai-monitor/august-2025-ai-false-claim-monitor/>

AI 데이터 오염의 시사점

- AI 반대 전선의 전략 변화... ‘규제’에서 ‘학습데이터 교란’으로
 - 현재 업계에서는 AI 안전성 연구의 선구자인 제프리 힌튼과 같은 인사들, AI 기술 확산에 반대하는 시민 단체인 스톱 AI(Stop AI), 그리고 알고리즘 편향과 차별에 맞서는 알고리즘 정의 연맹(Algorithmic Justice League) 같은 단체들이 기술 산업에 대항해 왔으나, 주요 논쟁은 규제의 개입 정도에 치중되어 왔음
 - 이러한 배경에서 포이즌 파운틴은 규제가 유일한 해결책이 아니라고 주장하며, AI가 학습하는 데이터 자체를 의도적으로 오염시키는 접근법을 취함. 아무리 성능이 뛰어난 AI라도 학습 데이터가 오염되면 결국 잘못된 결과를 생성할 수밖에 없다는 근본적인 취약점을 의도적으로 공략한 것임
 - 이는 단순히 AI 환각이나 성능상의 한계에서 비롯되는 AI 생성물의 개별적 오류를 지적하는 것이 아니라, AI가 자체 생성한 데이터와 검증되지 않은 정보로 계속 학습하는 현재의 ‘학습-생성-재학습’ 순환적 시스템 자체에 대한 근본적 문제 제기라는 점에서 차별화됨

참고문헌

- Thomas Claburn, AI industry insiders launch site to poison the data that feeds them, The Register, 2025.1.11., https://www.theregister.com/2026/01/11/industry_insiders_seek_to_poison/
- Poison Fountain, 2026.01.21. 기준, <https://rnsaffn.com/poison3/>
- NewsGuard, AI False Information Rate Nearly Doubles in One Year, 2025. 09.04, <https://www.newsguardtech.com/ai-monitor/august-2025-ai-false-claim-monitor/>



SUMMARY

산업/기업

기술

주간 기술 동향

멀티 에이전트 프레임워크와 워터마킹 기법, AI 저작권 문제의 새로운 해법

• 멀티 에이전트 기반 생성 AI의 저작권 보호 기술 분석

2025년, 오픈AI의 이미지 생성 기술을 통해 일본의 애니메이션 제작사인 스튜디오 지브리의 독창적인 화풍을 모방한 이미지들이 대량으로 만들어지면서 생성형 AI의 저작권 침해 문제가 불거진 바 있다. 이러한 이미지들은 소셜 미디어를 통해 빠르게 확산되며 대중의 흥미를 끌었지만, 원작자의 오랜 노력이 담긴 고유한 예술적 자산이 손쉽게 복제될 수 있다는 심각한 문제도 제기되었다. 결국 사용자가 간단한 명령만으로 특정 아티스트의 고유 스타일을 복제하는 현상은 기술 발전이 원작자의 권리를 어떻게 위협하는지 보여주는 대표적 사례이자, 창작의 윤리적 경계에 대한 사회적 논의를 촉발시키고 있다.

생성형 AI가 특정 화풍을 모방할 수 있는 것은 인터넷상의 방대한 이미지 데이터를 학습해 그 안에 담긴 스타일의 패턴과 규칙을 통계적으로 추출하는 원리 덕분이며, 이를 통해 완전히 새로운 이미지를 만들어내는 것이 가능하다. AI 모델은 특정 작가의 색감, 선의 형태, 구도와 같은 고유한 시각적 요소들을 수학적으로 학습하여 내재화하고, 사용자의 요구에 따라 이 요소들을 조합하여 결과물을 생성한다. 그러나 학습 과정에서 원작자의 명시적 허가 없는 학습 데이터 사용이 빈번하게 발생하고, 그 결과 최종 사용자는 저작권 침해에 대한 인식 없이도 너무나 쉽게 특정 스타일을 재현할 수 있게 되는 구조적 문제가 발생한다.

물론 이러한 문제에 대응하기 위해 AI 개발사들은 '생존하는 작가'의 이름을 직접 사용하는 시도를 차단하는 등 정책적 조치를 도입하고 있다. 하지만 이는 특정 키워드를 회피하면 그만인 소극적 대응에 불과하여 근본적인 해결책이 되지 못한다. 따라서 업계에서는 단순히 특정 단어를 차단하는 것을 넘어, 생성 과정 자체에 개입하여 창작의 맥락을 이해하고 저작권 침해 가능성을 사전에 제어할 수 있는 보다 근본적인 기술적 대안이 필요하다는 목소리가 높아지고 있다.

이러한 요구에 부응하는 핵심 기술로 여러 개의 전문 AI가 팀처럼 협력하는 '멀티 에이전트 프레임워크'가 주목받고 있다. 이 기술은 단순히 이미지를 생성하는 것을 넘어, '기획'과 '검토' 에이전트가 저작권 침해 소지를 사전에 판단하고, '보호' 에이전트가 생성물에 디지털 워터마크를 각인하여 앞서 제기된 '통제 불가능한 모방' 문제를 해결할 새로운 가능성을 제시한다. 본 보고서에서는 바로 이 멀티 에이전트 기술이 어떤 방식으로 저작권을 기술적으로 보호할 수 있는지 그 원리를 심층적으로 분석하고자 한다.

생성형 AI의 기술적 진보와 관련 제도 공백

• AI의 정교한 스타일 모방 기술 발전과 현행 저작권 제도의 구조적 한계

- 최신 생성형 AI 모델은 특정 아티스트의 고유한 화풍과 미묘한 표현 기법까지 정밀하게 학습하여 재현하는 능력을 갖추고 있음. 이는 전문 지식이 없는 일반 사용자라도 간단한 명령만으로 고품질의 창작 활동을 가능하게 함으로써 창작 활동의 대중화를 이끄는 동력이 되고 있음
- 하지만 AI 모델 학습 과정에서 원작자의 명시적 동의 없이 수집된 방대한 이미지 데이터가 사용되는 경우가 많아, 기술 자체가 저작권 침해의 가능성을 내포하는 근본적인 법적 취약점을 안고 있음
- 더 나아가 AI 산출물이 기존 저작물에 대한 2차적 저작물인지, 혹은 완전히 새로운 창작물인지에 대한 판단 기준이 부재하며, 그 소유권이 누구에게 귀속되는지에 대한 사회적 합의 또한 이루어지지 않아 기술과 제도 사이의 심각한 괴리를 드러내고 있음

[사례] 제어 가능한 콘텐츠 생성 및 보호를 위한 멀티 에이전트 프레임워크

• 멀티 에이전트 프레임워크 기술 개요 및 개발 배경

- 기존의 단일 생성 AI 모델은 사용자의 복잡한 의도를 정확히 반영하지 못하고, 특정 작가의 화풍을 의도치 않게 과도하게 모방하는 등 창작 과정에 대한 제어 기능이 부재했음
- 또한, 생성된 결과물 자체에는 소유권이나 출처 정보가 없어 무단 도용이나 저작권 분쟁 발생 시 기술적으로 소유권을 증명하기가 매우 어려운 구조적 한계를 가지고 있었음
- 멀티 에이전트 프레임워크는 이러한 배경에서 탄생한 기술로, 특정 작업을 전문적으로 수행하는 여러 AI 에이전트가 하나의 팀처럼 협력하여, 사용자의 의도를 정밀하게 제어함과 동시에, 생성물에 대한 법적·기술적 신뢰도를 보장하는 것을 목표로 함

[표1] 멀티 에이전트 프레임워크를 구성하는 에이전트별 역할

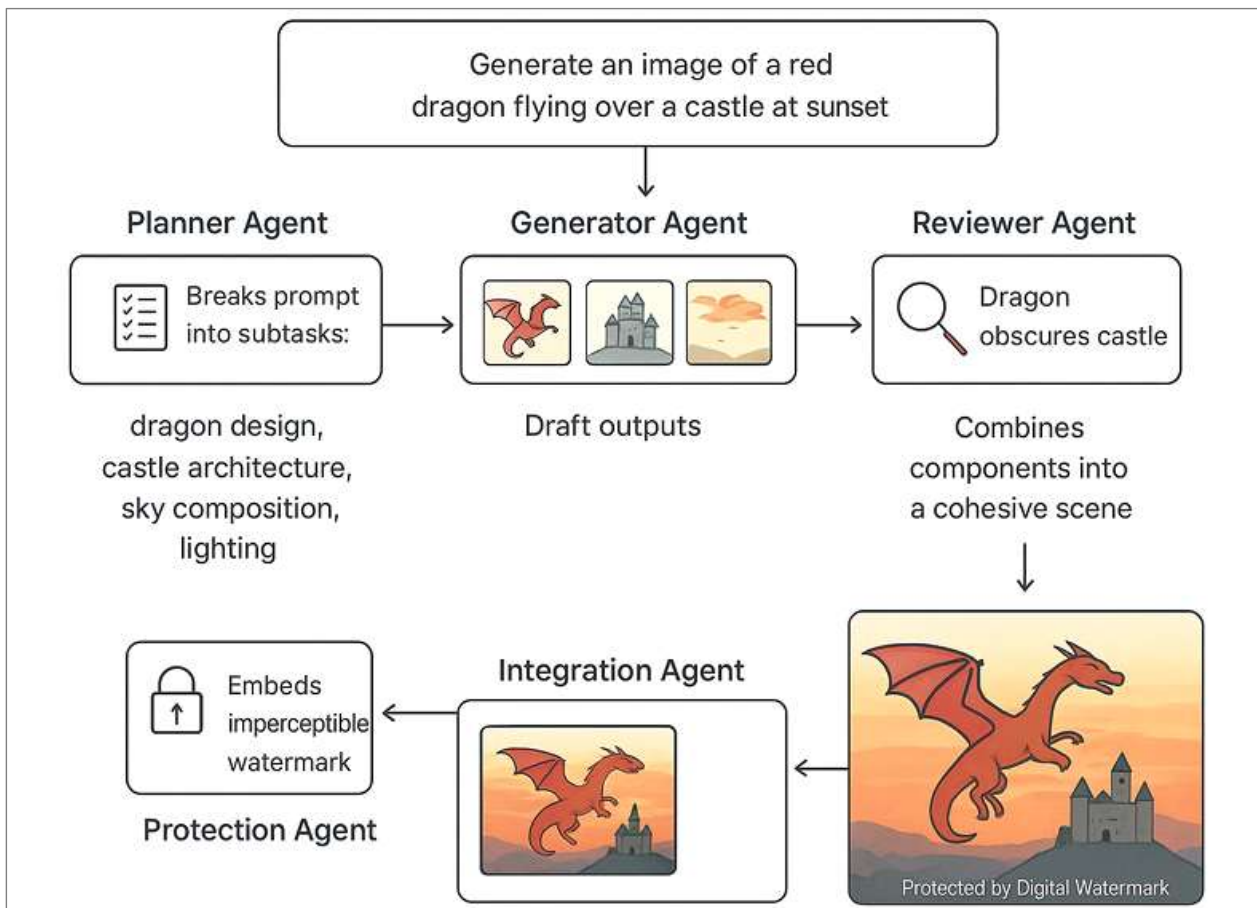
구분	역할
기획 에이전트 (Planner Agent)	사용자의 추상적인 아이디어를 해석하여 전체 프로젝트를 지휘하는 총감독 역할을 수행함 특히 사용자가 "특정 작가의 스타일을 모방하지 말 것"과 같은 명확한 규칙을 설정할 경우, 이 지시를 작업 계획에 반영하여 표절을 방지하는 등의 역할이 가능함
생성 에이전트 (Generator Agent)	기획 에이전트의 구체적인 지시에 따라, 이미지의 각 구성 요소를 시각적으로 구현하는 실질적인 아티스트 역할을 전담함
검토 에이전트 (Reviewer Agent)	생성 에이전트가 만든 중간 결과물이 초기 기획 의도 및 품질 기준에 부합하는지 지속적으로 평가하고 피드백을 제공하여 결과물의 완성도를 높임
통합 에이전트 (Integration Agent)	각기 다른 에이전트가 생성한 개별 결과물들을 하나의 완성된 이미지로 조화롭게 합성하고, 전체적인 시각적 일관성을 보장하는 편집자 역할을 수행함
보호 에이전트 (Protection Agent)	생성 과정 전반에 개입하여 눈에 보이지 않는 디지털 워터마크를 삽입하고 모든 작업 이력을 '프로비넌스 로그'로 기록하여, 최종 결과물의 출처와 소유권을 기술적으로 증명함

출처: 참고문헌 내용 재구성

• **작동 원리 및 작업 방식 예시**

- 사용자가 '붉은 용이 노을 진 성 위를 날아다니는 장면'과 같은 프롬프트를 입력하면, 기획 에이전트가 이를 '용', '성', '노을' 등 핵심 요소로 즉시 분해하여 구체적인 생성 계획을 수립함
- 생성 에이전트들은 기획 에이전트의 계획에 따라 각 요소를 개별적으로 만들기 시작함. 이와 동시에 보호 에이전트는 생성 초기 단계부터 워터마크를 삽입하고, 검토 에이전트는 중간 결과물의 품질을 평가하며 수정하는 반복적인 피드백 루프가 진행됨
- 모든 요소가 완성되면 통합 에이전트는 이를 하나의 완성된 이미지로 조화롭게 합성하고, 이 모든 생성 및 수정 과정은 타임스탬프와 함께 기록되어 워변조가 어려운 디지털 로그 기록으로 저장됨

[그림] 멀티 에이전트 프레임워크가 이미지를 만들어내는 과정 도식



출처: Haris Khan 외 2인, "Multi-Agent Framework for Controllable and Protected Generative Content Creation: Addressing Copyright and Provenance in AI-Generated Media", arXiv, 2026.01.09., <https://arxiv.org/pdf/2601.06232>

결론 및 시사점

• **작동 원리 및 작업 방식 예시**

- 멀티 에이전트 프레임워크는 통합 워터마킹과 프로비넌스 로그를 통해 산출물에 대한 기술적 신뢰를 부여함으로써, 저작권 문제를 사후에 대응하던 기존 방식에서 벗어나는 패러다임 전환을 제시함
- AI 활용 저작물의 출처와 소유권이 기술적으로 증명됨에 따라, 기업들은 법적 분쟁의 위험 부담을 줄이고 상업적 영역에서 AI 생성물을 안심하고 활용할 수 있는 중요한 기반을 마련하게 됨

- **사전 예방 기능의 구체화라는 기술적 과제와 향후 전망**

- 하지만 현재 수준에서는 워터마킹을 통한 사후 증명 기능이 핵심이며, 기획 및 검토 에이전트가 저작권 침해를 사전에 방지하는 구체적인 정책이나 판단 기준이 모델 자체에 내재되어 있지는 않음
- 즉, 사용자가 특정 IP를 모방하거나 표절하지 말라는 지시를 내려야만 기획 및 검토 에이전트가 이미지 생성 단계에 이를 반영한다는 한계를 지님
- 따라서 향후에는 이 구조적 잠재력을 활용하여, 사용자가 저작권 보호 규칙을 직접 설정하기 전에, 시스템이 자율적으로 모방을 제어하는 '사전 예방' 기능을 구체화하는 것이 해결 과제로 남음

참고문헌

- Haris Khan 외 2인, "Multi-Agent Framework for Controllable and Protected Generative Content Creation: Addressing Copyright and Provenance in AI-Generated Media", arXiv, 2026.01.09., <https://arxiv.org/pdf/2601.06232>
- Eve Upton-Clark, "OpenAI's Studio Ghibli-style images renew the debate over AI and copyright - Fast Company", FastCompany, 2026.01.20., <https://www.fastcompany.com/91308222/openais-studio-ghibli-style-images-renew-the-debate-over-ai-and-copyright>