

# 저작권 이슈 트렌드



COPYRIGHT ISSUE TREND

# CONTENTS

## 저작권 이슈 트렌드

Biweekly Report | 통권

제72호

- AI 학습 데이터 저작권 탐지를 위한  
오픈소스 프레임워크 기술 분석
- AI 저작권의 새로운 해법,  
'추론 시점 귀속' 기술의 원리와 가능성
- 생성형 AI 시대의 음원 저작권 보호:  
'데이터 투명성'과 라이선스 체계의 진화



# AI 학습 데이터 저작권 탐지를 위한 오픈소스 프레임워크 기술 분석

## 뉴스 브리프

AI 학습 데이터의 저작권 투명성 문제 해결을 위해 캐나다 워털루 AI 연구소가 일반 창작자도 쉽게 활용할 수 있는 오픈소스 저작권 탐지 프레임워크를 개발했다. 기존 DE-COP 등의 기술은 높은 컴퓨팅 자원과 복잡한 기술적 이해를 요구하여 독립 창작자의 접근이 어려웠으나, 신규 시스템은 SBERT 임베딩과 코사인 유사도 분석을 활용한 강화된 데이터 전처리 파이프라인으로 탐지 정확도를 높이면서도 알고리즘 최적화를 통해 계산 비용을 절감했다. 또한, 웹 기반 직관적 UI와 대시보드 제공으로 기술 진입 장벽을 완화하여, 창작자들이 자신의 저작물이 AI 학습에 무단 사용되었는지 직접 확인할 수 있는 환경을 조성했다. 이 기술은 AI 기업의 학습 데이터를 외부에서 독립적으로 검증 가능하게 함으로써, 데이터 출처를 명확하게 하여 투명한 데이터 관리 문화 정착에 기여할 것으로 기대된다. 본 보고서는 신규 오픈소스 프레임워크의 기술적 구조와 기존 기술 대비 우위를 분석하고, AI 개발 투명성 증진을 위한 기술적 과제를 검토해 보도록 한다.

## 뉴스 플러스

### 1. 서론: LLM 학습 데이터 내 저작권 탐지, 오픈소스 프레임워크로 새 국면

- 캐나다 워털루 AI 연구소 중심, 사용자 친화적 탐지 시스템 공개

생성형 인공지능의 학습 데이터 투명성 문제를 해결하기 위한 새로운 기술적 접근법이 제시되었다. 캐나다 워털루 AI 연구소(Waterloo AI Institute)와 세드라 학생 디자인 센터(Sedra Student



Design Centre)의 지원을 받은 연구팀은 대규모언어모델(LLM) 학습 데이터에 포함된 저작권 콘텐츠를 탐지하는 오픈소스 프레임워크를 개발했다고 밝혔다. 이 시스템은 기존의 복잡하고 비용이 많이 드는 탐지 방식과 달리, 일반 콘텐츠 창작자들도 높은 기술적 장벽 없이 쉽게 접근하여 활용할 수 있도록 설계된 것이 가장 큰 특징이다. 이 기술의 등장은 AI 모델이 어떤 데이터를 기반으로 학습했는지 알기 어려운 ‘블랙박스’ 문제를 기술적으로 해결하려는 중요한 시도이며, AI 산업 전반에 만연한 데이터 불투명성 논란에 대한 구체적인 대응책으로 평가된다.

이번에 공개된 프레임워크는 AI 기술의 윤리적 개발과 투명성 확보라는 시대적 요구에 부응하는 결과물이다. 최근 뉴욕 타임스(The New York Times Company)와 오픈AI(OpenAI, Inc.) 간의 소송에서 볼 수 있듯이, 저작권이 있는 콘텐츠의 무단 학습은 AI 기업의 법적 리스크를 높이고 대중의 신뢰를 저해하는 핵심 요인으로 작용하고 있다. 연구팀은 해당 시스템이 AI 개발 과정의 투명성을 높이고, 창작자에게 정당한 보상이 이루어지는 선순환 구조를 만드는 데 기여할 것으로 기대하고 있다.

- 기술 장벽 없이 저작물 무단 학습 여부 확인, 창작자 권익 보호 목표

새로운 탐지 시스템 개발의 핵심 목표는 기술적 전문성이 부족한 개인 창작자나 소규모 기업도 자신의 저작물이 AI 학습에 무단으로 사용되었는지 직접 확인할 수 있는 환경을 제공하는 것이다. 기존의 저작권 탐지 기술들은 높은 수준의 컴퓨팅 자원과 복잡한 기술적 이해를 요구하여 독립적인 콘텐츠 제작자들이 사용하기에는 현실적인 장벽이 존재했다. 이로 인해 개인 창작자와 소규모 조직은 저작권 침해 여부를 검증하기 어려운 상황이었다.

이러한 문제점을 해결하기 위해 개발된 이번 프레임워크는 비용 효율적이면서도 사용자 친화적인 솔루션을 지향한다. 시스템은 데이터셋 검증, 유사성 탐지 등의 핵심 기능을 강화하면서도 처리 비용을 크게 낮춰 접근성을 높였다. 이를 통해 창작자들은 자신의 지식재산권(IP)이 침해되었을 가능성을 이전보다 훨씬 수월하게 검토하고 대응의 근거를 마련할 수 있게 되었다.

- 생성형 AI의 불투명성, 기술적 해결책으로 돌파구 모색

생성형 AI 기술이 급격히 발전하면서 학습 데이터의 출처와 구성에 대한 불투명성 문제는 지속적으로 제기되어 왔다. 특히 방대한 양의 인터넷 데이터를 학습하는 과정에서 저작권이 있는 콘텐츠가 저작권자 동의 없이 사용될 수 있다는 우려는 AI 산업의 윤리적, 법적 논쟁의 중심에 있었다. AI 기업들은 경쟁력을 이유로 학습 데이터의 구성을 상세히 공개하지 않는 경우가 많아, 외부에서는 저작권 침해 여부를 파악하기 어려웠다.

이러한 상황에서 캐나다 연구진이 개발한 오픈소스 프레임워크는 데이터 불투명성 문제를 기술적 차원에서 해결할 수 있는 중요한 실마리를 제공한다. 이 시스템은 AI 모델의 학습 내용을 역으로 추적하고 검증하는 과정을 표준화하고 자동화함으로써, AI 개발 기업에 윤리적 책임성을 요구하는 동시에 공정한 창작 생태계를 조성하는 데 기여할 수 있는 기술적 기반이 될 것이다.

## II. 본론: 기존 탐지 기술의 한계와 이를 극복한 신규 시스템의 기술적 구조

### 1) 기존 저작권 탐지 기술, 비용·복잡성 장벽에 접근성 한계

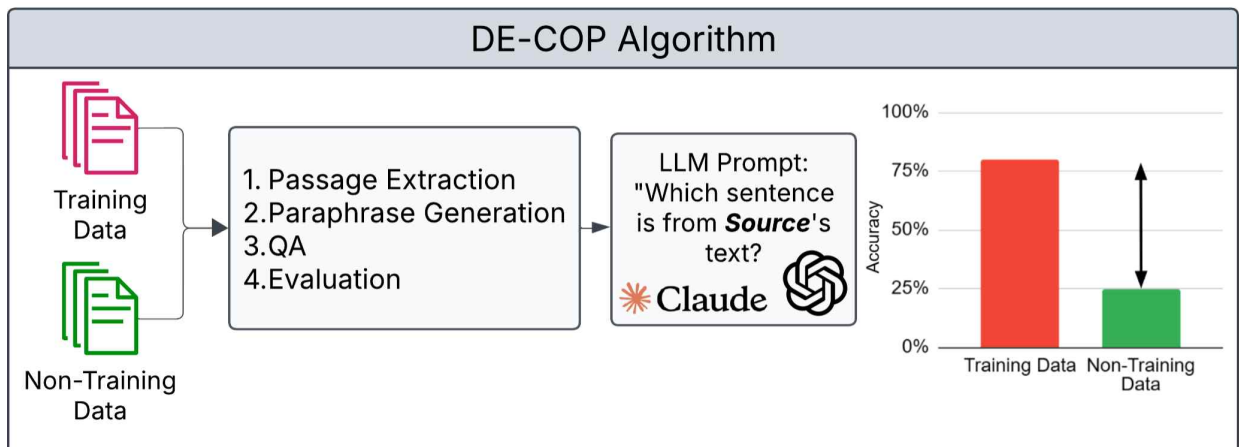
- DE-COP 등 기존 탐지 기술, 높은 비용과 기술 장벽에 실용성 저하

최근까지 LLM의 학습 데이터를 분석하려는 시도는 일부 있었지만, 대부분의 기술은 높은 비용과 복잡한 기술적 요구사항이라는 근본적인 장벽에 부딪혔다. 대표적인 예로 학계에서 제안된 'DE-COP(저작권 콘텐츠 감지 기법)\*'과 같은 프레임워크는 특정 콘텐츠가 모델 학습에 사용되었는지를 탐지할 수 있는 정교한 접근법으로 주목받았다. 이 기술은 모델이 내놓는 결과물의 확률값(logit)을 분석하거나, 블랙박스 모델\*\*을 대상으로 질의응답을 통해 학습 여부를 추정하는 방식으로 작동한다.

\* DE-COP(Detecting Copyrighted Content in Language Models): 모델의 확률값 분석이나 질의응답을 통해 학습 데이터 내 저작권 콘텐츠 사용 여부를 추정하는 탐지 기법

\*\* 블랙박스 모델(Black Box Model): 내부 작동 원리나 학습 데이터를 공개하지 않아 외부에서 검증이 어려운 AI 모델

[그림] DE-COP 시스템 알고리즘



출처: David Szczecina 외 2인, "Copyright Detection in Large Language Models: An Ethical Approach to Generative AI Development", arXiv, 2025.11.25., <https://arxiv.org/pdf/2511.20623>

하지만 이러한 방식은 상당한 수준의 컴퓨팅 자원을 필요로 하며, 기술을 구현하고 결과를 해석하는데 높은 전문성이 요구된다. 이로 인해 개인 콘텐츠 창작자나 소규모 단체가 현실적으로 활용하기에는 거의 불가능에 가까웠다. 결국 기술의 혜택이 일부 연구 기관이나 거대 기술 기업에 한정되면서, 정작 저작권 보호가 절실한 다수의 창작자에게는 실질적인 해결책이 되지 못하는 한계를 보였다.

- 오라일리 미디어 사례가 드러낸 학습 데이터 추정의 복잡성

이러한 탐지 기법이 실제로 활용된 사례로, 비영리 단체 'AI 디스클로저 프로젝트'의 분석을 들 수 있다. 이들은 오픈AI의 GPT-4o 모델을 대상으로 DE-COP 기법을 활용했으며, 해당 기법은 주제 원문에서 기술적 복잡성과 높은 계산 비용이 지적된 바 있다. 연구진은 이 분석을 통해, 오픈AI와 라이선스 계약을 맺지 않은 오라일리 미디어(O'Reilly Media, Inc.)의 유료 도서 콘텐츠가 모델 학습에 사용되었을 가능성을 제기했다. 연구진은 분석을 위해 34권의 도서에서 발췌한 1만 4천 개에 가까운 문단을 모델에 제시하고, 원문과 AI 생성 문장을 구별하는 능력을 측정하는 복잡한 실험을 진행했다.<sup>1)</sup>

이 사례는 저작권 침해 의혹을 확인하는 과정이 얼마나 간접적이고 노동집약적인지를 잘 보여준다. 특정 텍스트가 학습 데이터에 포함되었는지 직접 확인하는 것이 아니라, 모델의 반응을 통해 학습 여부를 '추정'하는 방식이기 때문이다. 이러한 접근법은 상당한 시간과 자원이 소요될 뿐만 아니라, 실험 설계의 정교함에 따라 결과의 신뢰도가 달라질 수 있어 일반 사용자가 시도하기에는 어려움이 따르는 방식이다.

- 탐지 정확도·비용 효율성·확장성, 전반적 개선 필요성 대두

기존의 저작권 탐지 방식들은 여러 측면에서 근본적인 개선의 필요성을 안고 있었다. 전통적인 표절 탐지 도구들은 AI가 원문을 미묘하게 바꾸거나 의역한 콘텐츠를 식별하는 데 명백한 한계를 보였다. 이를 극복하기 위해 등장한 기계학습 기반의 접근법들은 앞서 언급된 것처럼 기술적 복잡성과 높은 비용 문제로 인해 널리 확산되지 못했다. 결국 시장에는 저렴하면서도 사용하기 쉬운 저작권 검증 솔루션이 부재한 상황이 지속되었다. 따라서 탐지 정확도를 유지하면서도 계산 비용을 획기적으로 낮추고, 더 많은 사용자가 접근할 수 있도록 확장성을 확보하는 것이 차세대 탐지 기술의 핵심 과제로 떠올랐다.

1) 원호섭, "너 이거 몰래 보고 배운거지"...미허가 유료 콘텐츠 학습 논란에 빠진 오픈AI", 매일경제, 2025.04.02., <https://www.mk.co.kr/news/it/11280183>

## 2) 사용자 편의성 극대화, 신규 탐지 시스템의 기술적 우위 확보

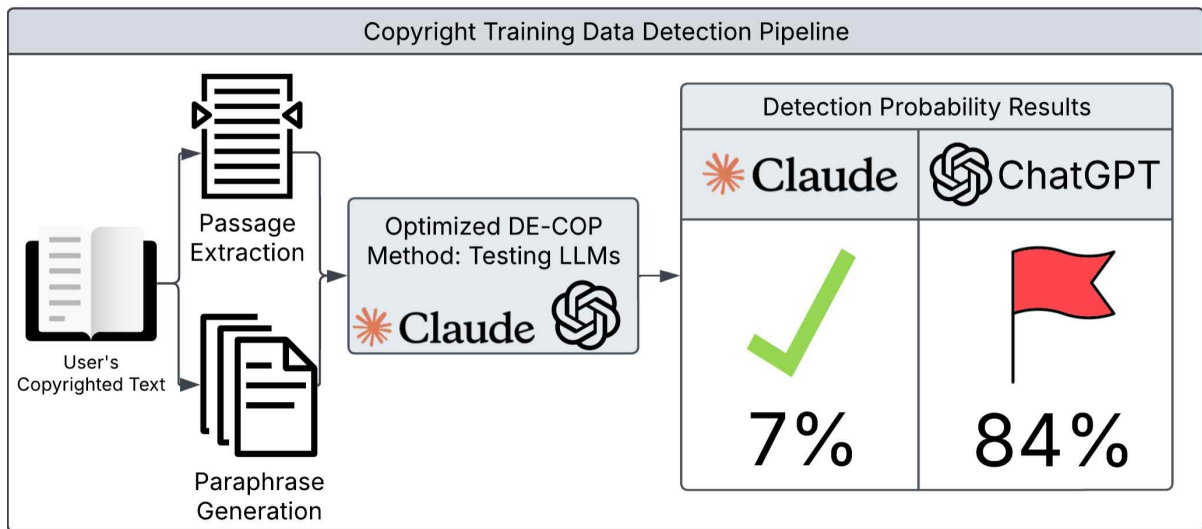
- 유사성 탐지 고도화와 데이터셋 검증으로 정확도·신뢰도 향상

캐나다 연구진이 개발한 새로운 프레임워크는 기존 기술의 한계를 극복하기 위해 결과의 재현성과 통계적 타당성을 높이는 데 집중했다. 이 시스템의 핵심은 탐지 정확도에 영향을 미치는 데이터 불일치 문제를 해결하기 위해 도입한 강화된 데이터 전처리 파이프라인이다. 이 과정에서는 문장의 의미를 벡터로 변환하는 SBERT(Sentence-Bidirectional Encoder Representations from Transformers) 임베딩\* 기술과 코사인 유사도\*\* 분석을 활용하여, 원문과 비교 대상 문장 간의 의미적 무결성을 검증하고 유효하지 않은 데이터는 사전에 걸러낸다.

\* SBERT 임베딩(Sentence-Bidirectional Encoder Representations from Transformers): 문장의 의미를 고차원 벡터로 변환하여 의미적 유사성을 수치화할 수 있게 하는 자연어 처리 기술

\*\* 코사인 유사도(Cosine Similarity): 두 벡터 간의 각도를 측정하여 텍스트의 의미적 유사성을 0~1 사이의 값으로 정량화하는 지표

[그림] 저작권 학습 데이터 탐지 파이프라인



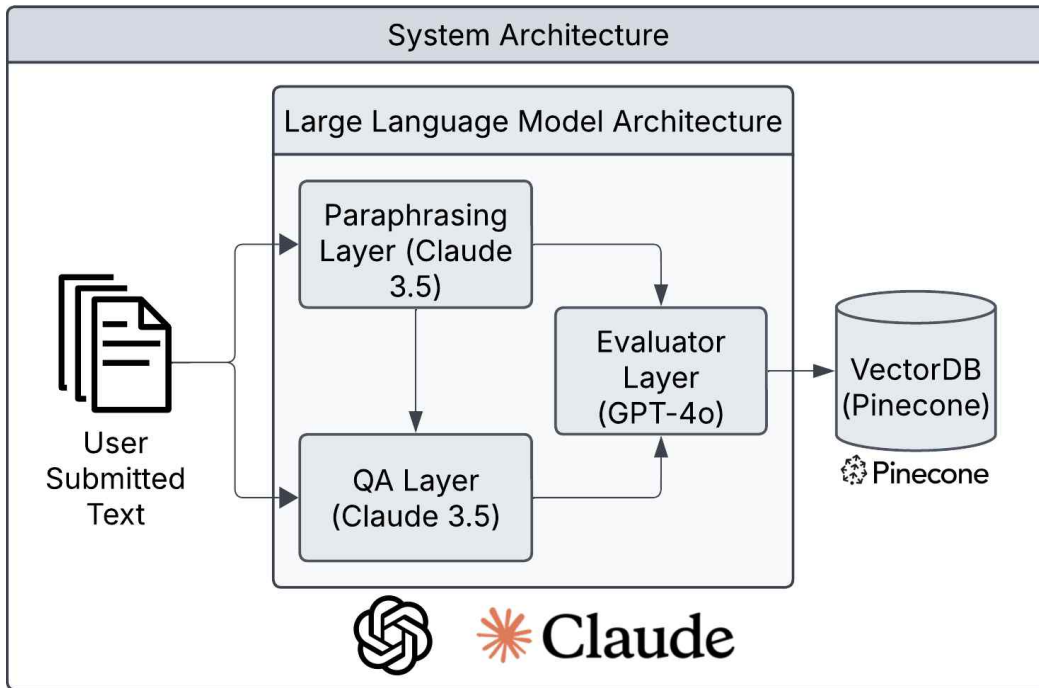
출처: David Szczecina 외 2인, "Copyright Detection in Large Language Models: An Ethical Approach to Generative AI Development", arXiv, 2025.11.25., <https://arxiv.org/pdf/2511.20623>

또한, 비교 문장의 길이가 지나치게 길거나 짧아 발생할 수 있는 분석 오류를 방지하기 위해 단락 길이를 정규화하는 과정도 포함되었다. 이러한 기술적 개선들은 분석에 사용되는 데이터의 품질 자체를 높임으로써, 최종적으로 더 신뢰할 수 있고 일관된 탐지 결과를 제공하는 기반이 된다. 이는 표면적인 문자열 비교를 넘어, 의미적으로 유사한 변형 콘텐츠를 식별하는 과정의 신뢰도를 확보하려는 접근 방식이다.

- 계산 효율성 최적화로 비용 절감 및 확장성 확보

새로운 탐지 시스템의 가장 큰 기술적 성과 중 하나는 계산 효율성을 획기적으로 개선하여 비용 장벽을 낮춘 점이다. 기존의 DE-COP과 같은 방법론은 LLM에 수많은 질의를 보내고 응답을 분석하는 과정에서 막대한 컴퓨팅 자원을 소모했지만, 이번에 개발된 프레임워크는 알고리즘 최적화를 통해 훨씬 적은 자원으로도 동일하거나 더 높은 수준의 정확도를 달성한다. 이는 분석에 소요되는 시간과 비용을 크게 절감시켜, 개인 창작자나 자금력이 부족한 소규모 단체도 부담 없이 시스템을 활용할 수 있는 길을 열었다.

[그림] 신규 시스템 구조도



출처: David Szczecina 외 2인, "Copyright Detection in Large Language Models: An Ethical Approach to Generative AI Development", arXiv, 2025.11.25., <https://arxiv.org/pdf/2511.20623>

이러한 비용 효율성은 시스템의 확장성을 담보하는 핵심 요소로 작용한다. 더 많은 사용자가 동시에 시스템에 접근하더라도 안정적인 서비스 제공이 가능하며, 새로운 AI 모델이나 더 큰 규모의 데이터셋에도 유연하게 적용할 수 있다. 이는 저작권 탐지 기술이 소수의 전문가를 위한 도구가 아닌, 누구나 필요할 때 쉽게 이용할 수 있는 보편적인 서비스로 발전할 수 있는 기술적 토대를 마련한 것으로 평가된다.

- 직관적 사용자 인터페이스(UI)로 기술 진입 장벽 완화

이번 프레임워크는 기술의 복잡성을 낮추고 사용자 편의성을 높이는 데 중점을 두고 설계되었다. 시스템은 사용자가 웹 기반 사용자 인터페이스(UI)를 통해 직접 콘텐츠를 제출하여 평가를 요청할 수 있는 방식을 채택했다. 이는 별도의 소프트웨어 설치나 복잡한 설정 과정 없이 누구나 쉽게 접근할 수 있도록 하여, 기술적 진입 장벽을 크게 낮추려는 목표를 반영한다.

분석 결과를 확인하는 과정 역시 사용자 중심으로 구성되었다. 사용자는 제공되는 대시보드(dashboard)와 분석 페이지(analytics page)에 접속하여 자신의 평가 이력을 조회하고, 시스템의 정확도 지표(accuracy metrics)를 직접 확인할 수 있다. 이처럼 시스템은 복잡한 내부 작동 원리를 사용자가 알 필요 없이, 평가 요청부터 결과 확인까지의 과정을 단순화하여 콘텐츠 창작자들이 기술에 대한 부담 없이 자신의 목적에 집중할 수 있도록 지원한다.

### III. 결론 : 기술로 여는 AI 투명성, 윤리적 생태계 구축의 초석

- AI 개발 투명성 증진 및 기업 책임성 강화에 기여

이번에 개발된 오픈소스 저작권 탐지 프레임워크는 AI 개발 과정의 투명성을 한 단계 끌어올리는 중요한 기술적 전환점이 될 수 있다. 지금까지 AI 기업들은 영업 비밀이라는 이유로 학습 데이터의 구체적인 내역 공개를 꺼려왔으나, 이제는 외부에서도 독립적으로 학습 데이터를 검증할 수 있는 수단이 마련되었기 때문이다. 이 기술은 콘텐츠 창작자들에게 자신의 저작물이 무단으로 사용되었는지를 확인할 수 있는 구체적인 도구를 제공함으로써, AI 기업이 데이터 수집 및 활용 과정에서 더 높은 수준의 책임성을 갖도록 유도하는 역할을 할 것이다.

결과적으로, 이 기술은 AI 산업 전반에 걸쳐 데이터 출처를 명확히 하고, 데이터를 투명하게 관리하는 문화를 정착시키는 데 기여할 수 있다. 기업들은 더 이상 불투명한 데이터 관행에 기댈 수 없게 되며, 이는 장기적으로 공정하고 신뢰할 수 있는 AI 개발 환경을 조성하는 데 긍정적인 영향을 미칠 것이다. 기술의 발전이 기업의 윤리적 책임을 강화하는 효과적인 압력으로 작용하는 선순환 구조가 만들어지는 것이다.

- 기술을 통한 윤리적 AI 개발 문화 조성의 발판 마련

이 프레임워크는 단순히 저작권 침해 여부를 탐지하는 도구를 넘어, 윤리적인 AI 개발 문화를 조성하는 기술적 기반으로서 중요한 의미를 지닌다. 자료에 따르면, 개발의 목표는 투명성을

증진하고 책임성을 장려함으로써 궁극적으로 윤리적인 AI 개발의 길을 여는 것이다. 접근성이 높은 검증 도구가 존재한다는 사실 자체가 AI 개발사들로 하여금 데이터 수집 단계에서부터 저작권 문제를 더 신중하게 고려하도록 유도하는 역할을 할 수 있다. 이는 AI 개발 과정에서 저작권 보호 장치를 소홀히 하여 발생하는 의도치 않은 결과와 윤리적 문제를 기술적으로 보완하는 접근법이다. 이는 처벌이나 규제를 통한 사후 대응이 아닌, 기술을 통해 예방적이고 자발적인 윤리 준수 문화를 만들어간다는 점에서 더욱 발전적인 방향이라 할 수 있다.

- 확장 가능한 프레임워크 제공과 후속 연구 기반 마련

이 연구는 LLM의 학습 데이터 검증을 위한 확장 가능하고 접근성 높은 오픈소스 프레임워크를 제공한다는 점에서 핵심적인 시사점을 갖는다. 이는 기술적 수단을 통해 창작자 개인의 권익을 보호하고 정보의 비대칭성을 해소하려는 중요한 시도이다.

결론적으로, 이 프레임워크는 단순히 저작권을 탐지하는 것을 넘어 AI 개발의 투명성을 증진하고 기업의 책임성을 장려하는 역할을 수행한다. 논문 저자들은 이 기술이 궁극적으로 공정한 보상을 촉진하고, 빠르게 진화하는 생성형 AI 환경 속에서 윤리적인 AI 개발을 지원하는 길을 열어줄 것이라고 강조한다. 이러한 방향성의 연구들이 향후 책임 있는 AI 개발과 저작권 보호 분야의 후속 연구를 촉진하는 중요한 기술적 기반이 될 것이다.

## 참고문헌

- Andre V. Duarte 외 3인, “DE-COP: Detecting Copyrighted Content in Language Models Training Data”, arXiv, 2024.06.25., <https://arxiv.org/pdf/2402.09910>
- David Szczecina 외 2인, “Copyright Detection in Large Language Models: An Ethical Approach to Generative AI Development”, arXiv, 2025.11.25., <https://arxiv.org/pdf/2511.20623>
- 원호섭, ““너 이거 몰래 보고 배운거지”...미허가 유료 콘텐츠 학습 논란에 빠진 오픈AI”, 매일경제, 2025.04.02., <https://www.mk.co.kr/news/it/11280183>
- 유승목, ““AI 기업의 학습데이터 공개 의무화해야””, 한국경제신문, 2025.10.15., <https://www.hankyung.com/article/2025101508521>

## 기술용어

순번	용어	설명
1	대규모언어모델 (LLM, Large Language Model)	방대한 텍스트 데이터로 학습되어 인간과 유사한 언어 이해 및 생성 능력을 갖춘 AI 모델
2	DE-COP (Detection of Copyright Content)	모델의 확률값(logit) 분석이나 질의응답을 통해 학습 데이터 내 저작권 콘텐츠 사용 여부를 추정하는 탐지 기법
3	SBERT 임베딩 (Sentence-BERT Embedding)	문장의 의미를 고차원 벡터로 변환하여 의미적 유사성을 수치화할 수 있게 하는 자연어 처리 기술
4	코사인 유사도 (Cosine Similarity)	두 벡터 간의 각도를 측정하여 텍스트의 의미적 유사성을 0~1 사이의 값으로 정량화하는 지표
5	데이터 전처리 파이프라인 (Data Preprocessing Pipeline)	분석 전 데이터의 품질을 높이기 위해 정제, 변환, 정규화 등을 순차적으로 수행하는 처리 과정
6	블랙박스 모델 (Black Box Model)	내부 작동 원리나 학습 데이터를 공개하지 않아 외부에서 검증이 어려운 AI 모델
7	로짓 (Logit)	신경망 모델이 최종 출력을 생성하기 전 계산하는 원시 확률값으로, 모델의 확신도를 나타내는 지표
8	오픈소스 프레임워크 (Open Source Framework)	소스 코드가 공개되어 누구나 자유롭게 사용, 수정, 배포할 수 있는 소프트웨어 구조



# AI 저작권의 새로운 해법, '추론 시점 귀속' 기술의 원리와 가능성

## 뉴스 브리프

AI 기술의 발전으로 스트리밍 플랫폼에서는 출처 불명의 AI 산출물이 급증하며, 이는 창작자에게 돌아갈 정당한 보상을 왜곡하고 산업 생태계의 신뢰를 무너뜨리는 위협이 되고 있다. 그러나 기존의 AI 산출물 생성 이후 영향 관계를 역추적하는 훈련 시점 귀속(TTA) 방식은 확률적 추정에 머물러 명확한 증거를 제공하지 못하는 한계를 보인다. 이에 대한 하나의 대안으로, 시스템 설계 단계부터 저작권 귀속을 내재화하는 '설계 기반 귀속' 프레임워크와 그 핵심 기술인 추론 시점 귀속(ITA) 기술이 고안되었다. 연구는 AI 산출물과 원저작물 사이의 관계를 명확히 증명하는 기술적 투명성 공개 방법을 통해 공정한 보상 체계를 재정립하는 것에 목적으로 두고 진행되었다.

## 뉴스 플러스

### 1. 서론 : AI 산출 순간 데이터 영향도 추적, '설계 기반 귀속' 기술 대두

- 스트리밍 플랫폼 내 AI 기반 가상 아티스트의 급증과 시장 교란

최근 스포티파이와 같은 글로벌 음악 스트리밍 플랫폼에서는 AI 기술을 활용한 가상 아티스트가 급증하며 기존 창작 생태계를 위협하는 새로운 문제로 부상하고 있다. 이들 가상 아티스트는 실제 인간 아티스트와 구별하기 어려운 프로필과 수백만 회에 달하는 스트리밍 횟수를 기록하며, 인간의 창작 주기를 초월하는 속도로 신규 음원 산출물을 쏟아내고 있다. 이러한 현상은 AI 산출물과 인간 창작물 사이의 경계를 허물어 사용자의 혼란을 야기하고 플랫폼 전체의 신뢰도를 저하하는 시장 교란 행위에 해당한다.

이러한 무분별한 확산이 초래하는 가장 직접적인 위협은 플랫폼의 저작권료 분배 시스템을 왜곡시킬 수 있다는 점이다. 대부분의 플랫폼이 채택한 '비례 배분 방식' 하에서 출처가 불분명한 AI 산출물이 높은 스트리밍 횟수를 기록하면, 인간 창작자들에게 돌아가야 할 정당한 보상이 정체불명의 주체에게 유출된다. 결국 AI 산출물의 출처를 기술적으로 증명할 수 없다는 문제는 단순히 저작권자를 가려내지 못하는 문제를 넘어, 산업의 공정한 보상 체계 자체를 붕괴시킬 수 있는 난제로 작용한다.

- 기존 훈련 시점 귀속 방식의 한계와 기술적 난제

스트리밍 플랫폼에서 벌어지는 혼란의 근본적인 원인은, 현재의 기술로는 AI 산출물의 출처를 명확히 증명할 방법이 없다는 데 있다. 지금까지 관련 연구는 주로 훈련 시점 귀속\*(Training-Time Attribution, 이하 TTA) 방식에 집중되어 왔다. 이 접근법은 AI 모델 전체의 학습 과정에 어떤 훈련 데이터가 영향을 미쳤는지를 분석하는 방식으로, 이는 산출물이 나온 이후에 그 영향 관계를 역으로 추정하는 사후 분석적 성격을 가진다.

\* 훈련 시점 귀속(Training-Time Attribution): AI 모델이 학습을 완료한 후, 모델 전체의 지식 형성에 어떤 데이터가 영향을 미쳤는지를 분석하는 방식

하지만 TTA 방식은 법적 책임이나 재정적 보상에 필요한 결정적이고 검증 가능한 인과관계를 증명하지 못한다는 명백한 한계를 가진다. 특히 수억 개의 데이터로 학습한 거대 모델 전체를 분석하는 것은 기술적으로 매우 복잡하고 계산 비용이 많이 들며, 그 결과 또한 확률적 추정에 머무를 뿐이다. 결국 기존 TTA 방식의 한계는 AI 산출물의 저작권 귀속을 영원히 모호한 영역으로 남겨두는 가장 큰 기술적 걸림돌이다.

- 설계 기반 귀속 프레임워크와 추론 시점 귀속 기술의 등장

기존 TTA 방식의 근본적 한계에 대한 대안으로 '설계 기반 귀속\*(Attribution-by-Design)'이라는 새로운 프레임워크가 제안되고 있다. 이는 저작권 귀속을 AI 시스템 개발의 부가 기능이 아닌, 초기 설계 단계부터 핵심 원칙으로 통합하자는 전환을 의미한다. 즉, 산출물이 나온 이후에 영향 관계를 추론하는 것이 아니라, 시스템 자체를 산출물 출처의 도출 과정과 동시에 기록하도록 설계하는 것이다.

이러한 '설계 기반 귀속' 프레임워크를 구현하는 핵심적인 기술 메커니즘이 바로 '추론 시점 귀속\*\* (Inference-Time Attribution, 이하 ITA)'이다. ITA는 AI가 새로운 산출물을 도출하는 바로 그 '추론 시점'에 사용된 데이터만을 추적하여, TTA의 복잡성과 불확실성을 근본적으로 해결한다. 결국 ITA 기술은 '설계 기반 귀속'이라는 철학을 현실로 구현하고, 현재 음악 시장이 겪고 있는 출처 불분명의 문제를 기술적으로 해결할 가장 현실적인 경로를 제시한다.

\* 설계 기반 귀속(Attribution-by-Design): AI 모델을 개발하는 초기 설계 단계부터 저작권 출처를 추적하는 기능을 시스템의 핵심 요소로 포함시키는 접근 방식

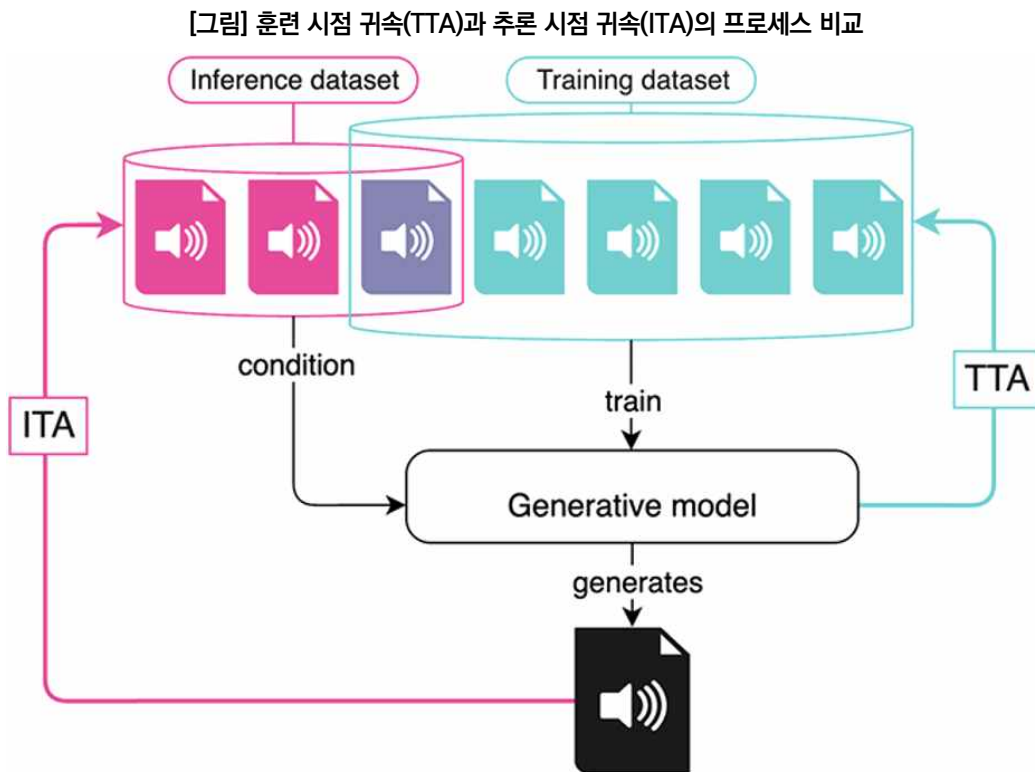
\*\* 추론 시점 귀속(Inference-Time Attribution): AI가 사용자 요청에 따라 새로운 산출물을 '도출하는 순간(추론 시점)'에 어떤 데이터가 직접적인 영향을 미쳤는지 실시간으로 추적하고 기록하는 기술

## II. 본론: '설계 기반 귀속' 프레임워크의 구조와 추론 시점 출처 확인 메커니즘

- 학습과 추론 데이터의 이원화

설계 기반 귀속 프레임워크의 가장 근본적인 전제는 AI 모델이 사용하는 데이터셋을 기능에 따라 '학습용'과 '추론용'으로 명확히 분리하는 것이다. 기존 AI 모델이 단일한 거대 데이터셋에 의존하여 학습과 산출을 모두 처리하는 것과 달리, 이 접근법은 데이터의 역할을 이원화하여 저작권 귀속 과정에서 발생하는 불확실성을 원천적으로 제거한다. 즉, 방대한 학습용 데이터셋은 장르의 구조나 화성학 이론과 같은 음악의 보편적 지식을 모델에 내재화하는 데만 사용된다.

실제 음악 AI 산출물이 도출되는 추론 단계에서는 별도로 구축된 '추론용 데이터셋'만 활용된다. 이 데이터셋은 저작권자들이 명시적인 동의하에 제공한 소규모 고품질 음악 데이터로 구성되며, 모델은 사용자의 요청에 따라 새로운 산출물을 만들 때 오직 이 데이터셋의 요소들만을 참조하고 조합하도록 아키텍처 차원에서 제한된다. 이러한 구조적 분리는 산출물에 직접적인 영향을 미친 데이터의 범위를 명확하게 한정함으로써, 방대한 학습 데이터 전체를 분석해야 했던 TTA의 기술적 난제를 효과적으로 예방한다.



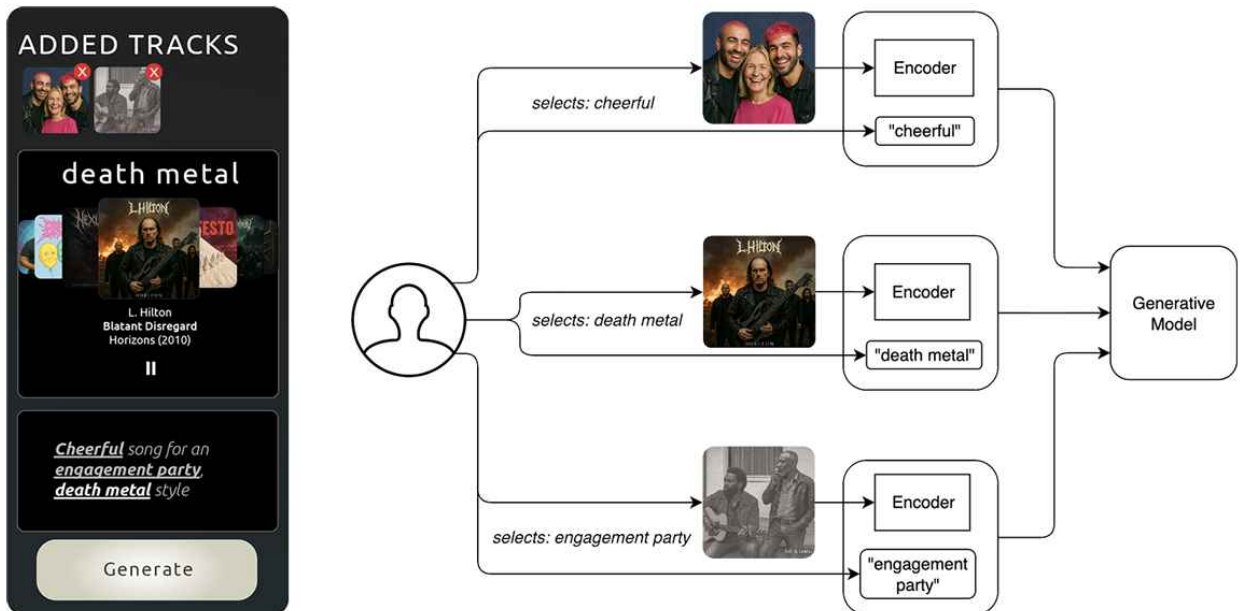
출처: Fabio Morreale 외 4인, "ATTRIBUTION-BY-DESIGN: ENSURING INFERENCE-TIME PROVENANCE IN GENERATIVE MUSIC SYSTEMS", arXiv, 2025.10.09., <https://arxiv.org/pdf/2510.08062>

- 사용자 상호작용 기반의 세분화된 데이터 활용 동의 획득

이 프레임워크는 단순히 저작권자의 참여 여부만 묻는 옵트인(Opt-in) 방식의 동의를 넘어, 훨씬 더 세분화되고 구체적인 동의를 획득하고 관리하는 시스템을 제안한다. 이는 저작권자가 자신의 음악 데이터가 AI 산출 과정에서 어떻게 활용될지를 매우 구체적인 수준에서 직접 제어할 수 있도록 권한을 부여하는 것을 의미한다. 시스템은 저작권자가 제공하는 인터페이스를 통해 이러한 세부적인 라이선스 조건을 데이터와 함께 저장하고 관리해야 한다.

예를 들어, 어떤 권리자는 자신의 곡에서 '드럼 파트의 리듬'은 AI가 참조하도록 허용하지만, 고유한 '보컬 멜로디' 라인은 참조하지 못하도록 금지하는 조건을 설정할 수 있다. 또한, 특정 곡의 전체적인 '분위기'나 '코드 진행'은 활용을 허용하되, 사용된 '악기 음색'은 사용하지 못하도록 제한하는 등 매우 정교한 수준의 통제가 가능하다. AI 모델은 산출물을 도출할 때 이러한 라이선스 조건을 기술적으로 반드시 준수해야 하므로, 창작자의 의도가 시스템 작동에 직접 반영되는 투명한 구조가 만들어진다.

[그림] 복수의 참조 데이터로부터 음악적 특징을 조합해 음악을 만드는 과정



출처: Fabio Morreale 외 4인, "ATTRIBUTION-BY-DESIGN: ENSURING INFERENCE-TIME PROVENANCE IN GENERATIVE MUSIC SYSTEMS", arXiv, 2025.10.09., <https://arxiv.org/pdf/2510.08062>

- 경로 유도 및 잠재 공간 편집을 통한 산출 과정의 정밀 제어

데이터 이원화와 세분화된 동의는 산출 과정을 정밀하게 제어하는 기술과 결합 때 비로소 완성된다. 대표적 기술로는 '경로 유도(Trajectory Guidance)\*'와 '잠재 공간 편집(Latent Space Editing)\*\*'이

있으며, 이 기술들은 모델이 산출물을 생성하는 과정에 적극적으로 개입하여 특정 데이터의 음악적 특징을 반영하도록 유도하는 역할을 수행한다. 즉, AI가 임의의 결과가 아닌, 명확한 설계도와 참조 목록에 따라 음악적 연산을 수행하도록 제어하는 것이다.

\* 경로 유도(Trajectory Guidance): AI가 잠재 공간 안에서 산출물을 도출하는 경로를 특정 데이터의 스타일을 따르도록 유도하여 산출 결과를 제어하는 기술

\*\* 잠재 공간 편집(Latent Space Editing): AI가 학습한 특징들이 저장된 잠재 공간에서 특정 개념의 방향을 찾아내고, 이 값을 수학적으로 조작하여 산출물의 스타일, 분위기, 형태 등 세부 속성을 의도적으로 제어하는 기술

이러한 제어 기술은 음악 산출 과정 자체를 추적 가능한 프로세스로 전환시킨다. 예를 들어 사용자가 '특정 베이시스트의 슬랩 주법을 반영한 펑크 음악'을 요청하면, 시스템은 추론용 데이터셋에 있는 해당 베이시스트의 연주 데이터를 활용해 산출 경로를 유도한다. 더 나아가 잠재 공간 편집을 통해 산출물의 리듬을 더 복잡하게 만들거나 특정 음계를 따르도록 미세하게 조정하는 과정 역시 모두 기록으로 남는다.

- 아키텍처에 내재된 확정적 출처 증명을 통한 산출물 계보 기록

기존 TTA 방식이 이미 나온 산출물을 보고 원본 데이터를 역추적하는 확률적 추정에 머물렀다면, ITA는 시스템 아키텍처 자체에 출처 기록 기능을 내장하여 확정적인 증명을 제공한다. 이 시스템에서 산출 과정은 단순히 결과물을 출력하는 것이 아니라, 어떤 추론 데이터가 언제, 어떻게 사용되었는지를 상세히 기록하는 로그를 동시에 남기는 과정이다. 따라서 귀속 정보는 사후에 분석되는 추정치가 아니라, 산출과 함께 기록되는 명백한 사실 데이터가 된다.

이러한 설계는 모든 음악 AI 산출물이 자신의 출처와 계보를 증명하는 일종의 디지털 출생증명서를 내장하게 함을 의미한다. 즉, 산출 과정에서 참조된 특정 원본 데이터가 무엇인지, 그리고 그것이 최종 산출물과 어떤 직접적인 관계를 맺는지를 기술하는 검증 가능한 메타데이터가 함께 기록되는 것이다. 이는 법적 분쟁이나 저작권료 정산에서 요구되는 '검증 가능성'이라는 기준을 충족시키는 차이점이다.

- 실시간 영향도 분석을 통한 데이터 기여도의 정량화

확정적 출처 증명은 '어떤 데이터가 사용되었는가'를 알려주는 것을 넘어, '얼마나 영향을 미쳤는가'를 측정하는 실시간 영향도 분석으로 이어진다. 시스템은 단순히 데이터의 사용 여부만 기록하는 것이 아니라, 각 데이터가 최종 산출물의 구조와 특징에 기여한 정도를 정량적으로 계산해야 한다. 이 정량화 메커니즘은 공정한 보상 분배의 핵심적인 기술적 근거를 마련하는 단계이다.

기여도는 다양한 방식으로 측정될 수 있다. 예를 들어, 특정 아티스트의 코드 진행이 곡 전체의 화성적 뼈대를 이루었다면 높은 기여도를, 짧은 드럼 필인이 한 번 사용되었다면 낮은 기여도를 부여받는 식이다. 또한, 사용된 데이터의 길이, 반복 횟수, 산출물 내에서의 역할(멜로디, 리듬, 화성 등)과 같은 여러 요소를 종합적으로 분석하여 기여도를 산출할 수 있다. 이러한 정량적 데이터는 이후 자동화된 보상 시스템의 직접적인 입력값이 된다.

## 결론 : 기술적 투명성을 통한 지속 가능한 창작 생태계 구축

### • 공정한 보상 체계 재정립 가능성

설계 기반 귀속 기술은 AI 산출물의 출처를 명확히 증명함으로써, 현재 스트리밍 플랫폼이 직면한 가장 큰 위협인 공정한 보상 체계의 붕괴를 막는 결정적인 역할을 수행한다. 실시간 영향도 분석을 통해 도출된 정량적 기여도 데이터는, 각 저작권자에게 돌아갈 저작권료를 배분하는 투명하고 합리적인 기준으로 작동하여 기존 비례 배분 방식의 불투명성과 왜곡 가능성을 근본적으로 해결한다. 이는 AI 기술이 창작자의 권리를 침해하는 도구가 아니라, 오히려 기여도를 정밀하게 측정하여 정당한 보상을 보장하는 상생의 도구가 될 수 있음을 시사한다.

이러한 기술적 투명성은 창작자, 플랫폼, 사용자 모두의 신뢰를 회복하는 기반이 된다. 창작자들은 자신의 저작물이 무단으로 착취당하지 않고, 사용된 만큼 투명하게 보상받을 것이라는 믿음을 갖게 되어 안심하고 AI 생태계에 참여할 수 있다. 플랫폼은 AI 산출물을 명확히 식별하고 관리할 수 있게 되어 콘텐츠 생태계의 질서를 유지할 수 있으며, 사용자들은 자신이 접하는 음악이 윤리적이고 합법적인 절차를 통해 나왔다는 사실을 신뢰하며 감상할 수 있게 되는 것이다.

### • 인간 창의성과 AI 기술의 공존을 위한 기술

궁극적으로 이 기술은 인간 창의성과 AI 기술이 대립하는 것이 아니라, 상호 존중하며 공존하는 새로운 생태계의 가능성을 열어준다. '옵트인' 방식은 창작자를 AI 기술 발전의 수동적 피해자가 아닌, 자신의 데이터를 직접 통제하고 가치를 창출하는 능동적 참여자로 격상시킨다. 이는 AI를 인간의 창의성을 대체하는 경쟁자가 아닌, 인간의 창의적 의도를 정교하게 구현하고 확장하는 강력한 도구로 재정의하는 중요한 철학적 전환이다.

따라서 설계 기반 귀속 기술은 단순히 저작권 문제를 해결하는 소극적 방어 기술을 넘어, 미래 창작 산업의 청사진을 제시하는 기반 기술이라 할 수 있다. 이 기술이 제시하는 투명성과 공정성의 원칙 위에서, 인간의 독창적인 영감과 AI의 방대한 연산 능력이 결합된 새로운 차원의 예술이 탄생할 수 있을 것이다. 기술적 책임성을 통해 지속 가능한 생태계를 구축하는 것이야말로 AI 시대에 인류의 창의적 자산을 보존하고 발전시키는 핵심 과제이다.

## 참고문헌

- Fabio Morreale 외 4인, “ATTRIBUTION-BY-DESIGN: ENSURING INFERENCE-TIME PROVENANCE IN GENERATIVE MUSIC SYSTEMS”, arXiv, 2025.10.09., <https://arxiv.org/pdf/2510.08062>
- Tim Ingham, “The AI music problem on Spotify (and other streaming platforms) is worse than you think”, Music Business Worldwide, 2025.06.30., <https://www.musicbusinessworldwide.com/the-ai-music-problem-on-spotify-and-other-streaming-platforms-is-worse-than-you-think/>

## 기술용어

순번	용어	설명
1	훈련 시점 귀속 (Training-Time Attribution)	AI 모델이 학습을 완료한 후, 모델 전체의 지식 형성에 어떤 데이터가 영향을 미쳤는지를 분석하는 방식
2	설계 기반 귀속 (Attribution-by-Design)	AI 모델을 개발하는 초기 설계 단계부터 저작권 출처를 추적하는 기능을 시스템의 핵심 요소로 포함시키는 접근 방식
3	추론 시점 귀속 (Inference-Time Attribution)	AI가 사용자 요청에 따라 새로운 산출물을 '도출하는 순간(추론 시점)'에 어떤 데이터가 직접적인 영향을 미쳤는지 실시간으로 추적하고 기록하는 기술
4	경로 유도 (Trajectory Guidance)	AI가 잠재 공간 안에서 산출물을 도출하는 경로를 특정 데이터의 스타일을 따르도록 유도하여 산출 결과를 제어하는 기술
5	잠재 공간 편집 (Latent Space Editing)	AI가 학습한 특징들이 저장된 잠재 공간에서 특정 개념의 방향을 찾아내고, 이 값을 수학적으로 조작하여 산출물의 스타일, 분위기, 형태 등 세부 속성을 의도적으로 제어하는 기술



# 생성형 AI 시대의 음원 저작권 보호: ‘데이터 투명성’과 라이선스 체계의 진화

## 뉴스 브리프

생성형 AI의 확산은 콘텐츠의 양적 팽창과 함께 메타데이터 부실, ‘유령 음원’ 범람 등 산업의 구조적 취약점을 심화시키고 있다. 이는 권리자의 노출 기회와 수익을 잠식하며 창작 생태계를 위협하는 핵심 요인이 되었다. 단순한 사후 규제나 법정 공방만으로는 기하급수적으로 늘어나는 AI 콘텐츠를 관리하는 데 한계가 있음이 명확해지고 있다. 이에 글로벌 업계는 ‘데이터 투명성(Data Transparency)’을 핵심 해법으로 설정하고, 명확한 식별 및 검증 체계와 집단 라이선스 모델 도입 등 구조적 대응에 나서고 있다. 본 보고서는 이러한 변화를 심층 분석하여, AI 시대의 ‘투명성 격차(Transparency Gap)’ 해소와 공정한 저작권 생태계 조성을 위한 구체적 전략을 제시한다.

## 뉴스 플러스

### 1. 서론: AI 음악 콘텐츠 식별 기술과 메타데이터 표준 도입

- DDEX 표준 기반 AI 콘텐츠 라벨링 체계 구축

디지털 데이터 익스체인지(Digital Data Exchange, DDEX)\*는 2025년부터 음원 메타데이터(metadata)\*\*에 AI 활용 여부를 명시할 수 있는 새로운 표준 체계(DDEX ERN v4.3.1)를 도입했다. DDEX ERN v4.3.1은 녹음물이나 제작 기여분(contribution)이 생성형 AI를 통해 전부 또는 일부 제작되었음을 표시하는 선택적 항목을 포함하며, 기존 음원 유통 공급망에 이러한 공개된 정보를 자연스럽게 통합하도록 설계되었다.

\* 글로벌 음악/엔터테인먼트 산업에서 메타데이터와 권리 정보의 규격을 만드는 국제 표준

\*\* 특정 콘텐츠를 설명하는 제목/제작자/날짜 등에 대한 상세 정보

2025년 9월 스포티파이(Spotify)는 음악인과 출판사에 AI 사용 투명성을 강화하기 위해 DDEX 표준 준수를 공식 권고한 바 있으며, 현재 유니버설 뮤직 그룹(Universal Music Group) 및 베거스 그룹(Beggars Group) 등이 DDEX 표준 적용을 동의했다. 또한, 2025년 5월 소노스위트(SonoSuite)는 스포티파이의 '우선 공급자(Preferred Provider)' 지위에 따라 스포티파이에 전달하는 메타데이터 형식을 DDEX ERN 4.3 규격으로 업데이트 했다.

• 주요 플랫폼의 AI 콘텐츠 식별 및 투명성 강화 정책

주요 스트리밍 플랫폼들은 AI 기반 콘텐츠에 대한 투명성을 강화하기 위해 단계적으로 공개 의무 제도를 도입하고 있다. 유튜브(YouTube)는 합성 콘텐츠(synthetic content)\*에 AI 생성 여부 표기를 의무화하고 있으며, 틱톡(TikTok)은 오디오 파일 내부에 진위 정보를 삽입하는 '콘텐츠 자격증명(content credentials)' 기능을 도입해, 해당 콘텐츠가 공유·저장·재업로드 되더라도 AI 활용 정보는 유지되도록 설계했다. 메타(Meta) 역시 자사 서비스 전반에서 합성 콘텐츠에 AI 활용 라벨을 부착하는 정책을 운영 중이다. 한편, 디저(Deezer)는 업로더의 자발적 표기에만 의존하지 않고, 자체 시스템을 통해 '완전 AI 생성(Fully AI)' 트랙을 기술적으로 식별하여 태깅하는 독자적인 방식을 도입했다.

\* AI 혹은 알고리즘이 생성한 이미지/영상/음성/텍스트 등 콘텐츠

[표] 주요 플랫폼의 AI 콘텐츠 공개 방식 비교

플랫폼	공개방식	의무 여부	비고
Spotify	DDEX 표준 기반 태깅 (전체/부분/미사용)	권고 (Voluntary)	15개 이상 주요 레이블/배급사 참여
You Tube	사실적 합성 콘텐츠 플래그 표시	의무 (Mandatory)	위반 시 콘텐츠 삭제 등 제재 가능
Tik Tok	'콘텐츠 자격증명(Content Credentials)' 삽입	자동 적용	C2PA 기술 활용, 파일 이동 시 정보 유지
Meta	AI 활용(Imagined with AI) 라벨 부착	정책적 적용	페이스북, 인스타그램 등 전 서비스 적용
Deezer	'완전 AI 생성(Fully AI)' 트랙 식별	시스템적 식별	데이터 레벨 집행(Enforcement) 예고

출처: Sammy Andrews, "Metadata is the industry's weak spot; in the age of AI, the cost of half-measures is far higher", Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>

## II. 본론: AI 음악 콘텐츠 검증 및 저작권 관리 체계

### 1) AI 음악 검증 및 부정 행위 탐지 시스템

- 스트리밍 플랫폼의 AI 산출물 필터링 메커니즘

주요 디지털 서비스 제공자(Digital Service Provider, DSP)들은 AI 산출물을 악용한 부정 행위에 적극적으로 대응하고 있는데, 실제 한 초대형 플랫폼은 지난 1년간 스트리밍 사기나 조작이 의심되는 음원 수천만 개를 삭제 조치했다. 현재 업계는 사칭 규제 강화, 대량 중복 업로드 차단을 위한 스팸 필터 고도화, DDEX 메타데이터를 활용한 AI 생성 정보 공개 의무화 등을 추진하고 있다.

이러한 조치는 주요 플랫폼들이 사태의 심각성을 인지했음을 보여주지만, 정책의 성패는 결국 실행력에 달려 있다고 할 수 있다. 핵심 과제는 합법적인 음원 제작자에게 불이익을 주지 않으면서 부정 음원만을 정밀하게 걸러내는 필터링 기술, 그리고 공급망 전체에서 정보가 누락 없이 전달되도록 하는 데이터 투명성이다.

구체적인 사례로 스포티파이는 2025년 9월 정책 업데이트를 통해, 당사자 동의 없는 딥페이크(Deepfake)\*나 음성 모방 콘텐츠는 허용되지 않으며 발견 즉시 삭제될 것임을 공식화했다.

\* AI로 사람의 얼굴 및 목소리 등을 조작해 실제처럼 보이게 만든 합성 영상이나 음성

- 대량 업로드 및 스팸 차단 기술

AI 산출물의 무분별한 대량 업로드는 플랫폼의 품질 저하와 저작권료 풀(Royalty Pool)\*의 왜곡을 심각하게 야기하고 있다. 일례로 디저(Deezer)는 일일 약 3만 개 이상의 '완전 AI 생성(Fully AI-generated)' 음원을 식별하고 있는데, 이는 전체 신규 업로드의 약 3분의 1에 달하는 수치다. 더욱 심각한 것은 이렇게 식별된 음원에서 발생한 스트리밍 트래픽 중 최대 70%가 허위 재생(fake streaming)으로 밝혀졌다는 점이다<sup>2)</sup>.

\* 스트리밍 플랫폼이 저작권자에게 배분하기 위해 모아둔 총 수익금

이러한 AI 생성 음원은 단순한 플랫폼 혼잡을 넘어, 정당한 제작자에게 돌아가야 할 수익을 가로채는 결과를 낳는다. 현재 다수의 스트리밍 플랫폼들은 이러한 AI 음원 및 부정 트래픽 관련 데이터 공개에 침묵하고 있으며, 이와 같은 투명성 부재가 음원 업계의 우려를 가중시키고 있다. 또한, 실제 피해 사례도 잇따르고 있다. 중남미 지역의 음악인들은 'AI 음원의 범람(flood of AI content)'에 항의하는 시위를 전개하며, 스포티파이(Spotify), 디저(Deezer), 유튜브 뮤직(YouTube

2) Sammy Andrews, "Metadata is the industry's weak spot: in the age of AI, the cost of half-measures is far higher", Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>

Music) 등의 플랫폼에서 자신들의 제작물이 합성 콘텐츠에 밀려 노출되지 않고 있다고 호소한다. 이는 단순한 수익 감소를 넘어, 아티스트 사칭(impersonation)과 플랫폼 내 노출 기회 박탈(loss of visibility)이라는 제작자의 생존권을 근본적으로 위협하는 문제로 확대되고 있다.

• 신원 인증 및 기술적 출처 검증 체계의 도입

AI 음원 문제의 근본적인 해결을 위해서는 유통 단계에서의 검증 체계가 선행되어야 하나, 현재 동검증 단계는 '약한 고리(weak link)'로 남아 있다. 영국 공연권협회(Performing Right Society, PRS)는 업로더의 신원을 확인하는 '고객 알기 제도(Know Your Customer, KYC)' 도입을 촉구했으나, 대부분의 배급사가 비용과 절차상의 이유로 이를 채택하지 않고 있다<sup>3)</sup>.

악성 행위자들은 일관된 가입 표준이 부재한 탓에 제재를 피해 여러 서비스를 자유롭게 넘나들며 사기 행각을 반복할 수 있는 상황인데, 이에 대한 기술적 보완책으로 제시되는 것이 바로 데이터 레벨의 인증이다. 기존의 국제표준녹음코드(International Standard Recording Code, ISRC)가 AI 식별에 한계를 보임에 따라, 콘텐츠 출처 및 진위 확인을 위한 C2PA(Coalition for Content Provenance and Authenticity) 기술을 통해 오디오 파일 자체에 출처 정보를 암호화하여 삽입하거나, 국제표준콘텐츠코드(International Standard Content Code, ISCC) 기반의 디지털 지문(Fingerprint)을 생성해 중복 및 위변조를 원천 차단하려는 시도가 이어지고 있다. 결국 실효성 있는 검증 체계의 완성, 이러한 제도적 신원 확인(KYC)과 기술적 데이터 인증(C2PA/ISCC)이 유기적으로 결합될 때 비로소 가능할 것이다.

[표] AI 음원 검증 및 식별 기술

구분	기술 명칭	주요 기능	한계점
메타데이터	DDEX ERN v4.3.1	AI 사용 여부 및 기여도(보컬/작곡 등) 표기	입력이 선택적이며, 허위 기재 가능성 존재
인증 기술	C2PA / ISCC	파일 내 출처 정보 암호화 및 디지털 지문 생성	보편적 채택 전까지는 확인 어려움
신원 확인	KYC(Know Your Customer)	배급 단계에서 업로더 실명/신원 검증	주요 배급사 도입 저조, 비용 발생
수익 감사	StreamSight / Sureel	실시간 저작권료 흐름 추적 및 AI 기여도 감사	고도화된 기술 필요, 초기 도입 비용

출처: Sammy Andrews, "Metadata is the industry's weak spot: in the age of AI, the cost of half-measures is far higher", Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>

3) Sammy Andrews, "Metadata is the industry's weak spot: in the age of AI, the cost of half-measures is far higher", Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>

## 2) AI 학습 데이터 라이선스 추적 및 보상 체계

- 집단 라이선스 모델과 귀속 기술

AI 학습 데이터에 대한 라이선스 문제는 업계의 미해결 과제로 남아 있다. 기업 간 '일대일 거래(one-to-one deals)' 방식은 기하급수적으로 늘어나는 AI 데이터를 감당하기에 확장성의 한계가 뚜렷하며, 현실적으로는 집단 라이선스(collective licensing)\*가 유일한 대안으로 꼽힌다. 그러나 그동안 대부분의 저작권 관리 단체들이 집단 라이선스 도입을 주저하면서, 음원 시장은 출판사들의 개별 소송과 음반사들의 양자 계약(bilateral agreements)이 혼재하는 상황에 놓여 있었다. 이러한 교착 상태에서 스웨덴 작곡가 저작권협회(Svenska Tonsättarens Internationella Musikbyrå, STIM)는 제작자가 AI 학습 허용 여부를 직접 선택하는 '옵트인(opt-in)' 방식의 집단 AI 라이선스 모델을 제시하며 새로운 돌파구를 마련했다<sup>4)</sup>. 이러한 모델은 귀속(attribution)\*\* 기술을 필수적으로 요구하는데, 이를 통해 원작이 AI 산출물 생성에 미친 영향을 정밀하게 추적하고, 실시간으로 수익 흐름을 감시하는 것이 가능하다.

\* 저작권 관리 단체가 개별 제작자를 대리하여 AI 기업과 일괄 라이선스 계약을 체결하고, 이에 따른 수익을 징수하여 권리자에게 분배하는 관리 모델

\*\* 저작물의 창작자나 권리자를 식별하고 표시하는 기술

- 학습 데이터의 투명성 및 법적/상업적 필수성

학습 데이터의 투명성은 라이선스 체계의 실효성을 담보하는 핵심 전제조건이다. 일례로 영국은 권리자가 명시적으로 거부하기 전까지는 저작물 학습을 동의한 것으로 간주되는 '옵트아웃 (Opt-out)' 제도를 제안했으나, AI 기업에게 데이터 공개 의무를 병행 부과하지 않는 한 권리자가 도용 사실조차 인지할 수 없으므로 사실상 시행이 불가능하다는 비판을 받아 철회했다.

사법적 판단 역시 과도기에 머물러 있다. 미국과 유럽의 주요 법원들은 '동의 없는 저작물 학습의 저작권 침해 여부'라는 근본적인 법리 해석을 두고 여전히 고심 중이다. 미국 저작권청(US Copyright Office)이 '인간 창작(human authorship)' 원칙을 고수하는 가운데 AI 기업을 상대로 한 소송은 급증하고 있으며, 개별 합의를 통해 부분적인 안전장치(guardrails)가 마련되고는 있으나 업계 전반의 법적 불확실성은 여전하다.

이러한 상황에서 영국 AI 기업 스태빌리티 AI(Stability AI)의 '스테이블 오디오 2.5'\*와 같은 기업용(enterprise) 모델의 등장은 시사하는 바가 크다. 이는 투명하고 견고한 라이선스 프레임워크가 단순한 이론적 요구가 아니라, 법적 리스크를 해소하고 지속 가능한 사업을 영위하기 위한 '상업적 필수 요건(commercial necessity)'임을 증명한다.

\* 공식 라이선스가 확보된 음원 데이터만을 학습에 사용하여 저작권 침해 리스크를 최소화하는 모델

4) Sammy Andrews, "Metadata is the industry's weak spot: in the age of AI, the cost of half-measures is far higher", Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>

- AI 기술을 활용한 수익 정산 투명성 확보

한편, 이러한 라이선스 체계의 복잡성을 해소하기 위해 AI 기술을 역으로 활용하려는 시도도 이어지고 있다. AI 기술은 음악 산업의 위협 요인이기도 하지만, 동시에 고질적인 불투명성을 해결하는 현대화의 도구로도 활용된다. 대표적인 사례로 글로벌 음악 기업 BMG(Bertelsmann Music Group)는 구글 클라우드(Google Cloud)와 협력하여 데이터 분석 도구인 '스트림사이트(StreamSight)'를 출시했는데, 동 시스템은 AI를 활용해 복잡한 저작권료 수익을 신속하게 예측하고, 정산 프로세스를 투명하게 시각화하도록 설계되었다. 이는 AI 기술이 가진 '이중적 측면(double-edged sword)'을 명확히 보여준다. 무분별한 데이터 생성으로 음악 생태계를 혼란에 빠뜨릴 수 있는 AI 기술이 역설적으로 낙후된 정산 시스템을 현대화하고 효율화하는 데 쓰이고 있는 것이다. 앞서 언급한 스웨덴 STIM의 집단 라이선스 모델 역시 이러한 맥락이다. 이 모델은 실시간으로 수익 흐름을 감시하게 함으로써, 제작자가 자신의 저작물이 AI 산출물에 기여한 방식을 추적하고 정당한 보상을 받도록 보장한다. 결론적으로 이러한 기술적 시도들은 라이선스 계약과 실제 수익 배분 사이에 존재하는 '투명성 격차(transparency gap)'를 해소하는 핵심적인 역할을 수행하고 있다.

### III. 결론: AI 음악 기술 체계의 단편화와 통합 과제

- 메타데이터와 음원 목록 관리의 단편화

AI 시대를 맞이한 음악 산업의 가장 큰 리스크는 AI 기술 자체가 아니라, 이를 뒷받침하는 시스템의 단편화(fragmentation)다. 현재 다수 음원들은 일관성 없는 메타데이터 및 출처를 알 수 없는 AI 흔적들로 가득 차 있는데, 이에 umbrella와 같은 기업들은 음원 목록 전체를 감시하여 메타데이터 격차를 메우고 미표기된 AI 콘텐츠를 스캔하는 솔루션을 제공하고 있다. 음원 업계가 스스로의 데이터베이스를 먼저 정화(cleaning)하지 않은 채 새로운 AI 규제만을 도입하려 한다면, 이는 취약한 기반 위에 집을 짓는 것과 같다. 기존 시스템의 결함을 방치한 채 강행하는 질서 확립 시도는 오히려 잘못된 저작권료 배분과 숨겨진 부채, 그리고 제작자와의 불필요한 분쟁을 초래할 뿐이다. 메타데이터는 오랫동안 산업의 취약점이었으나, AI 시대에 이를 방치하는 대가는 과거와 비교할 수 없을 만큼 치명적이다.

- 표준 채택의 지연과 플랫폼 간 불일치

음악 산업에 있어 기술적 표준은 존재하나 그 적용은 일관성을 잃고 있다. DDEX ERN v4.3.1 도입은 시작에 불과하며, 일부 배급사와 음반사가 채택을 약속했을 뿐 보편적 표준으로 자리 잡지는

못한 상황이다. 더 큰 문제는 ‘어디까지가 인간의 제작인가’에 대한 합의된 임계값(threshold)이 부재하다는 점이다. AI 생성 비트를 기반으로 한 랩, AI 믹싱 기술이 적용된 밴드 음악, AI로 보정한 보컬 등 기계 개입의 수준은 천차만별이라고 할 수 있는데, 이를 정의하는 권한을 각 플랫폼에 맡겨둔다면, 서비스마다 규칙이 다른 파편화된 환경이 조성될 위험이 크다. 따라서 산업에 필요한 것은 권리자, 제작자, 규제 기관이 함께 개발하고, 모든 저작권 단체와 디지털 서비스 제공자(DSP)가 일관되게 적용할 수 있는 ‘공유된 프레임워크(Shared Framework)’다.

- 법적·제도적 격차 해소와 데이터 기반의 해법

AI는 이제 거부할 수 없는 음원 제작의 도구가 되었다. 향후 과제는 이러한 상황을 배척하는 것이 아닌, 본연의 가치를 보호하는 시스템 안으로 통합하는 것이다. 이를 위해서는 명확한 공개, 강력한 검증, 효과적인 부정 행위 통제, 그리고 확장 가능한 라이선스가 필수적이다. 단순한 기술 도입을 넘어, 유료 사용자 대상의 다운로드 제한이나 횡수 상한선 설정과 같은 구체적인 상업적 보호장치가 마련되어야 할 시점이다. 그러나 법적 규제는 국가별로 속도가 다르며, 플랫폼의 데이터 공개는 여전히 소극적이다. 결국 AI 시대 음악 산업의 지속 가능성은 ‘데이터의 투명성’에 달려 있다. 어설픈 반쪽짜리 조치(half-measures)에 그친다면, AI 음원의 범람 속에 인간 제작자의 설 자리는 사라질 것이다.

## 참고문헌

- Sammy Andrews, “Metadata is the industry’s weak spot; in the age of AI, the cost of half-measures is far higher”, Music Business Worldwide, 2025.11.27., <https://www.musicbusinessworldwide.com/metadata-is-the-industrys-weak-spot-in-the-age-of-ai-the-cost-of-half-measures-is-far-higher/>
- Mark Sweney, “Warner Music signs deal with AI song generator Suno after settling lawsuit”, The Guardian, 2025.11.26., <https://www.theguardian.com/business/2025/nov/26/warner-music-signs-deal-with-ai-song-generator-suno-after-settling-lawsuit>
- AFP, “Spotify moves to tackle AI abuse with transparency measures”, The Economic Times, 2025.09.25., <https://economictimes.indiatimes.com/tech/technology/spotify-moves-to-tackle-ai-abuse-with-transparency-measures/articleshow/124118391.cms>



## 기술용어

순번	용어	설명
1	귀속(attribution)	제작물의 창작자나 권리자를 식별하고 표시하는 기술
2	메타데이터(metadata)	특정 콘텐츠를 설명하는 제목/제작자/날짜 등에 대한 상세 정보
3	합성 콘텐츠(synthetic content)	AI 혹은 알고리즘이 생성한 이미지/영상/음성/텍스트 등 콘텐츠
4	옵트아웃(Opt-out)	제작자가 명시적으로 거부의를 밝히면 사용이 제한되는 저작권 동의 구조
5	딥페이크(Deepfake)	AI로 사람의 얼굴 및 목소리 등을 조작해 실제처럼 보이게 만든 합성 영상이나 음성