



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 산업 라이선스 협상과 AI 음악 플랫폼의 사업 구조

▶ AI 음악 생성 기업 수노는 UMG·SME·WMG로부터 저작권 침해 소송을 제기받은 상황에서도 연속적인 투자 유치를 이어가고 있다. 이는 투자 판단 과정에서 소송 리스크뿐만 아니라 플랫폼 성장 가능성과 향후 라이선스 협상 전망도 함께 고려되고 있음을 시사한다. 소송 제기 이후 WMG는 수노와의 라이선스 계약을 통해 소송을 취하하였으나, UMG·SME는 여전히 법적 대응을 이어나가고 있으며 AI 생성 음악의 외부 유통 허용 범위를 둘러싼 입장 차이는 협상이 교착되는 핵심 쟁점으로 작용하고 있다. 이 과정에서 라이선스 합의 조건이 플랫폼의 서비스 구조와 사업 방향을 직접 조정하는 요인으로 기능하고 있으며, 수노·유디오 사례는 라이선스 협상이 AI 음악 플랫폼의 기능·유통 방식·수익화 가능성을 결정하는 초기 사업 설계 요소로 작용하고 있음을 보여준다.

## 산업 유니버설뮤직그룹이 제시한 AI 음악 산출물의 사전 권리 제어 모델

▶ AI 음악 시장이 빠르게 확대되는 가운데, 대형 음반사의 저작권 관리 방식이 생성·유통 단계의 사전 권리 제어로 이동하면서 음반사별 대응 노선이 분화되고 있다. 유니버설뮤직그룹은 AI 음악 산출물이 허가된 플랫폼 밖으로 유출되지 않도록 관리하는 폐쇄형 생태계를 채택하고, 60건 이상의 특허 포트폴리오로 이를 인프라화하고 있다. 구체적으로는 입력·출력 단계 모두에서 권리자 조건을 점검하는 이중 승인 구조와 워터마크로 무단 생성을 차단하고, 스마트 계약을 연동해 허가된 플랫폼 내 이용과 수익 정산까지 자동 관리하는 방식이다. 앞으로 개방형 생태계와 표준 경쟁이 치열해질 것으로 예상되는 가운데, 인프라 독점에 따른 소규모 창작자의 소외를 막고 생태계가 안정적으로 정착하려면 권리 귀속 기준과 수익 배분 방식에 대한 법적·제도적 준비가 함께 논의될 필요가 있다.

## 산업 시나미디어의 옛지 워터마킹 출시와 실시간 스트리밍 보호 기술의 진화

▶ 시나미디어가 불법 복제 스트리밍 탐지와 차단을 위한 옛지 워터마킹 솔루션 ‘콘텐츠아머 옛지 워터마킹’을 출시했다. 이번 솔루션은 영상이 이용자에게 전달되기 직전의 CDN 옛지 단계에서 워터마크를 삽입하는 방식으로 설계됐다. 이를 통해 사전에 A/B 버전의 영상을 생성해 저장할 필요를 줄이고, 원본 서버와 옛지 서버 간 데이터 전송량 및 CDN 캐시 부담을 완화할 수 있다. 특히 압축 비트스트림에 직접 워터마크 정보를 삽입해 식별 정보의 밀도를 높이고, 일부 영상 구간만으로도 유출 세션을 빠르게 특정할 수 있다는 점이 특징이다.



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 산업 시댄스 2.0, C2PA·비가시 워터마킹으로 AI 영상 출처 표시 강화

▶ 바이트댄스는 2026년 3월 시댄스 2.0의 안전장치를 보강하고 글로벌 출시를 재개했다. 보강된 기능은 별도의 외부 도구가 아니라 시댄스 2.0의 입력·출력 단계에 직접 적용되는 방식으로 작동한다. 구체적으로 출력 단계에서는 C2PA 콘텐츠 자격증명, 가시적 워터마크, 비가시적 워터마킹으로 구성된 3계층 출처 표시 구조가 적용됐다. C2PA 콘텐츠 자격증명은 생성 도구, 생성 시점, AI 산출물 여부 등을 메타데이터로 기록해 검증 도구가 출처 정보를 확인할 수 있도록 한다. 가시적 워터마크는 이용자가 영상 표면에서 AI 생성 여부를 인지하도록 돕고, 비가시적 워터마킹은 플랫폼 외부 공유나 편집 이후에도 식별 정보를 유지하는 데 활용된다. 여기에 실존 인물·저작권 캐릭터 입력 차단 필터를 함께 적용하면서, 시댄스 2.0은 사전 차단과 사후 추적을 결합한 다층적 관리 구조를 갖추게 됐다.

## 산업 에이전틱 AI의 문서 편집과 저작권 관리 과제

▶ 마이크로소프트는 2026년 4월 워드·엑셀·파워포인트에 에이전틱 코파일럿 기능을 정식 출시하며, AI가 문서를 직접 편집·재구성하는 방식으로 기능 범위를 확장했다. 해당 기능은 활성화 시 편집 모드가 기본값으로 설정되는 구조로 이용자 동의 절차 및 UI 구성 방식에 관한 비판이 제기되었으며, 마이크로소프트가 유럽경제지역에서만 자동 설치를 제외한 점은 규제 압박이 동의 구조를 결정하고 있음을 시사한다. 한편 미국 저작권청은 AI가 산출물의 구체적 표현을 자율적으로 결정한 경우 저작권 귀속이 불분명해질 수 있다고 명시했다. 그러나 에이전틱 편집이 보편화되는 환경에서 인간 기여 범위를 사후적으로 판단하는 기준은 아직 마련되지 않은 상태로, 이용자 동의 설계와 저작권 귀속 기준 모두에서 제도적 대응이 과제로 남아 있다.

## 기술 주간 기술 동향

▶ 생성형 AI 기술의 확산으로 독립 창작자들은 자신의 작품이 무단으로 복제되거나 변조되는 위협에 상시 노출되어 있다. 기존 워터마킹 기술은 재생성 공격이나 부분 편집 앞에서 출처를 잃거나 변조 위치를 특정하지 못하는 한계를 안고 있었다. Dual-Guard는 이러한 문제를 해결하기 위해 확산 모델의 두 지점에 서로 다른 역할의 워터마크를 동시에 삽입하는 이중 채널 방식을 채택한다. 먼저, 가우시안 셰이딩 채널은 초기 노이즈에 출처 정보를 숨겨 재생성 공격 후에도 이미지의 출처를 검증한다. 코덱 채널은 잠재 표현에 핑거프린트를 삽입하여 변조된 영역을 블록 단위로 특정한다. 이미지 품질 저하 없이 99.9% 이상의 탐지율을 달성한 이 기술은, 창작자가 자신의 저작물에 대한 소유권을 기술적으로 입증할 수 있는 실용적 수단으로 평가된다.



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 라이선스 협상과 AI 음악 플랫폼의 사업 구조

### 소송 리스크 속 AI 음악 기업의 투자 유치

#### • 수노의 투자 유치와 기업가치 흐름

- AI 음악 생성 기업 수노(Suno)는 2025년 11월 시리즈 C(Series C)\* 라운드에서 2억 5,000만 달러(원화 약 3,712억 원<sup>1)</sup>)를 유치함<sup>2)</sup>
- 이후 2026년 5월 기준 시리즈 C를 상회하는 규모의 시리즈 D(Series D)\*\* 투자를 추진 중이며, 기업 가치는 50억 달러(원화 약 7조 4,240억 원) 이상으로 거론되어 시리즈 C 대비 약 2배 수준으로 상승한 것으로 평가됨<sup>3)</sup>
- 이번 투자는 유니버설뮤직그룹(Universal Music Group, 이하 UMG), 소니뮤직엔터테인먼트(Sony Music Entertainment, 이하 SME), 워너뮤직그룹(Warner Music Group, 이하 WMG)이 수노 측의 무단 음원 학습에 따른 저작권 침해 소송을 공동 제기한 상황 속에서 이뤄짐

\* 시리즈 C(Series C): 스타트업이 사업 확장과 시장 점유율 확대 등을 위해 진행하는 후기 단계 투자 유치 라운드

\*\* 시리즈 D(Series D): 시리즈 C 이후 추가 성장, 신규 시장 진출, 기술 고도화 또는 상장 준비 등을 목적으로 진행되는 후속 투자 유치 단계

#### • 저작권 소송 리스크 속 투자 관심과 이용자 확대

- 글로벌 주요 음반사들과의 저작권 소송이 계속되며 법적 리스크와 비용 부담이 존재함에도 불구하고, 수노에 대한 투자 유치는 지속되고 있음
- 빌보드(Billboard) 보도에 따르면 수노는 최근 애플 앱스토어(Apple Appstore) 음악 카테고리 1위와 전체 카테고리 11위를 기록하며, 이용자 확장을 지속하며 플랫폼 성장성을 입증하고 있음
- 이는 AI 음악 생성 기업의 가치가 향후 라이선스 협상 가능성과 사업 모델 확장성을 반영해 평가되고 있음을 시사함
- 특히 주요 음반사와의 협상 결과에 따라 서비스 구조가 변화할 수 있다는 점에서, 향후 투자 결정 시 라이선스 확보 여부가 중요한 변수로 작용할 것으로 분석됨

### 라이선스 협상에 따른 AI 음악 플랫폼 사업 구조 재편

#### • 음반사별 합의에 따른 사업 구조 변화

- 2025년 11월 WMG는 수노와 라이선스 계약을 체결하며 소송을 합의로 해결한 반면, UMG와 SME 측 소송은 계속 진행 중임

1) 1달러=1,484.80원(2026.05.04, KEB 하나은행 매매기준율 적용, 이하 동일)

2) Kristin Robinson, "Suno In the Midst of Series D Funding Round Following \$250M Raise in November", Billboard, 2026.05.04., <https://www.billboard.com/pro/suno-series-d-funding-round-following-250m-raise-november/>

3) Dylan Smith, "Suno Targets Series D Funding at a \$5 Billion Valuation", Digital Music News, 2026.05.04., <https://www.digitalmusicnews.com/2026/05/04/suno-series-d-reports/>

- 수노와 WMG의 합의에는 AI 생성곡 다운로드 제한, 라이선스 기반 모델 재출시 계획, 송킵(Songkick)\* 인수 등이 포함되었음. 이는 합의가 단순한 분쟁 해결을 넘어 서비스 운영 방식과 향후 사업 방향에 영향을 미칠 수 있음을 보여줌
- 유사한 소송을 겪었던 AI 음악 생성 플랫폼인 유디오(Udio)\*\* 역시 UMG·WMG 등과 합의 이후 텍스트 프롬프트 기반 음악 생성 모델에서 기존 음악 활용 중심의 리믹스 플랫폼으로 사업 방향을 조정할 바 있음
- 이러한 사례는 라이선스 협상이 플랫폼 기능, 유통 구조, 수익 모델을 재설계하는 핵심 요인으로 작용하고 있음을 시사함

\* 송킵(Songkick): 공연·콘서트 정보 검색 및 티켓 구매 서비스로, 수노는 WMG와의 합의 과정에서 해당 서비스를 인수함

\*\* 유디오(Udio): 수노와 함께 주요 음반사로부터 저작권 침해 소송을 제기받은 AI 음악 생성 플랫폼

### • AI 생성 음악의 외부 유통 문제

- UMG는 AI 생성 음악이 외부 스트리밍 서비스로 유통되는 구조에 대해 우려를 제기하며, 플랫폼 내부 이용에 한정하는 폐쇄형 생태계(walled garden)\* 방식이 필요하다는 입장을 보임
- 반면 수노는 AI 음악 생성 결과물이 외부 창작 및 유통 활동으로 확장될 수 있는 개방형 구조를 지향하고 있음. 이는 이용자의 타 플랫폼 연계를 유도해 플랫폼의 확장성을 높이는 반면, 기존 관리자 측면에서는 AI 생성 음악의 외부 유통을 통제하기 어렵게 만드는 위협 요인으로 작용함
- 이에 따라 AI 음악 플랫폼은 개방형 창작 도구로 발전할 것인지, 제한형 소비 서비스로 운영될 것인지에 따라 사업 구조가 분기되는 상황임

\* 폐쇄형 생태계(walled garden): AI 생성 음악의 유통을 특정 플랫폼 내부로 한정하고 외부 스트리밍 서비스로의 업로드·배포를 허용하지 않는 방식

### • 라이선스 협상과 플랫폼 경쟁력

- WMG의 합의와 유디오의 사업 방향 조정 사례는 라이선스 협상이 AI 음악 플랫폼의 법적 리스크를 낮추는 수단인 동시에, 서비스 기능과 유통 구조를 재설계하는 계기가 될 수 있음을 보여줌
- 폐쇄형 생태계 수용 여부는 AI 음악 플랫폼의 사업 방향을 결정하는 분기점으로 작용하며, 수용 시 관리자 통제형 서비스로 전환되나, 거부 시 확장성의 이면에 법적 리스크가 가중되는 상황임
- 이에 따라 라이선스 협상력은 단순한 법적 대응 능력을 넘어 AI 음악 플랫폼의 사업 구조와 경쟁력을 결정하는 핵심 변수로 부상하고 있음

## 시사점: AI 음악 플랫폼의 성장 조건 변화

### • 라이선스 협상력과 기업가치 재평가

- 수노의 연속 투자 유치는 저작권 소송이 진행 중인 상황에서도 플랫폼 성장성, 라이선스 협상 가능성과 시장 선점 효과가 투자 판단에 함께 반영된 것으로 해석됨
- 다만 주요 음반사와 AI 라이선스 계약을 확보한 플랫폼이 경쟁 우위를 갖는 구조가 형성되면서, 수노의 협상 교착이 장기화될 경우 시장 경쟁 구도에 영향을 미칠 수 있음
- 이처럼 주요 음반사와의 라이선스 협상 결과는 소송 종결 여부를 넘어 기업가치 평가와 시장 경쟁 구도에 영향을 미치는 변수로 작용할 수 있으며, 관련 이력·결과가 플랫폼 신뢰도를 구성하는 요소로 편입될 가능성도 함께 제기됨

### • 리스크 관리 역량의 중요성

- 최근 수노는 저작권 소송 대응과 법무 운영 전략을 전담하는 조직을 구축하려는 움직임을 보이고 있으며, 이는 단발성 이슈가 아니라 플랫폼 운영의 주요 관리 영역으로 편입되고 있음을 시사함
- 즉, AI 음악 플랫폼이 생성 기술 고도화만으로는 충분하지 않으며, 음반사와의 라이선스 협상, 소송 리스크 관리, 생성 음악의 외부 유통 범위 설정을 기술 개발과 병행해야 하는 필수 당면 과제에 놓여 있음을 의미함
- 결과적으로 수노와 유디오 사례는 라이선스 협상이 사후적 분쟁 해결 수단에 그치지 않고, AI 음악 플랫폼의 기능, 유통 방식, 수익화 가능성을 결정하는 초기 사업 설계 요소로 작용하고 있음을 보여줌

### 참고문헌

- Kristin Robinson, "Suno In the Midst of Series D Funding Round Following \$250M Raise in November", Billboard, 2026.5.4., <https://www.billboard.com/pro/suno-series-d-funding-round-following-250m-raise-november/>
- Dylan Smith, "Suno Targets Series D Funding at a \$5 Billion Valuation", Digital Music News, 2026.5.4., <https://www.digitalmusicnews.com/2026/05/04/suno-series-d-reports/>



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 유니버설뮤직그룹이 제시한 AI 음악 산출물의 사전 권리 제어 모델

### AI 음악 산출물 시장의 확대와 사전 권리 제어의 등장

#### • 시장 확대에 따른 대형 음반사의 대응 노선 분화

- AI 음악 산출물의 유통 규모가 빠르게 확대되면서, 결과물이 생성된 이후에 저작권 침해 여부를 확인해 사후 처리하는 기존 방식은 대응 속도와 제어 범위 측면에서 한계가 드러남
- 이에 따라 대형 음반사의 대응도 단순한 소송 중심의 사후 대응에서 AI 음악 산출물의 생성·유통 방식을 사전에 제어하는 방향으로 이동함
- 2024년 6월 미국음반산업협회(Recording Industry Association of America)\*가 주요 음반사를 대리해 수노(Suno)와 우디오(Udio) 등 대표 AI 음악 생성 서비스를 상대로 제기한 일부 소송이 최근 개별 합의로 마무리되면서, 음반사별 대응 노선의 차이가 구체적으로 드러나기 시작함
- 워너뮤직그룹(Warner Music Group)은 수노와의 합의를 통해 라이선스 기반 제어를 전제로 하되, 이용자가 AI 음악 산출물을 외부로 다운로드해 활용할 수 있는 '제한적 개방형 생태계'로 운영하는 방향을 선택함
- 반면, 유니버설뮤직그룹(Universal Music Group)은 우디오와의 합의 및 자사의 독자적인 기술 개발을 기반으로 AI 음악 산출물이 허가된 플랫폼 밖으로 유출되지 않도록 제어하는 '폐쇄형 생태계(Walled Garden)\*\*'를 구축하고 있음

\* 미국음반산업협회(Recording Industry Association of America): 미국 주요 음반사와 음악 유통사의 이익을 대표하는 산업 단체로, 음반 산업의 저작권 보호, 불법 복제 대응, 음악 시장 통계 발표, 정책 대응 등을 수행함

\*\* 폐쇄형 생태계(walled garden): AI 생성 음악의 유통을 특정 플랫폼 내부로 한정하고 외부 스트리밍 서비스로의 업로드·배포를 허용하지 않는 방식

#### • 특허 포트폴리오를 통한 사전 권리 제어 모델 구축

- 유니버설뮤직그룹은 이러한 사전 권리 제어 모델을 인프라 차원으로 확장하기 위해 2025년 7월 지식재산권 투자 전문 기업 리퀴덱스 캐피탈(Liquidax Capital)과 합작 법인 '뮤직 아이피 홀딩스(Music IP Holdings, MIH)'를 설립함
- 뮤직 아이피 홀딩스는 음원 생성 전 권리자 동의 확인, 디지털 워터마크(Digital Watermark) 표시, 배포 경로 제한과 수익 자동 정산 등 AI 음악 산출물 관리에 필요한 주요 기술을 60건 이상의 특허 포트폴리오\*로 구성해 배타적 권리 확보를 추진하고 있음
- 이는 개별 침해 사례에 대한 사후 대응을 넘어, AI 음악 산출물의 생성·유통·수익화 과정에 필요한 핵심 기술을 권리자가 직접 설계하고 관리하려는 시도임

\* 특허 포트폴리오: 기업이 자사의 핵심 기술을 다각도로 보호하고 시장 지배력을 강화하기 위해 전략적으로 확보한 다수의 연관 특허 집합

## 특허 포트폴리오를 통한 AI 음악 산출물의 생성·유통 전반 권리 제어

### • 권리자 승인 여부를 확인하고 산출물에 워터마크를 표시하는 기본 시스템

- 공개된 핵심 특허 3건은 ‘생성 단계의 기본 제어→유통 단계로의 확장→파생 상품 영역으로의 확장’이라는 유기적인 단계적 구조를 보임
- 첫 번째 특허는 권리자의 승인과 디지털 워터마크를 연계하여 AI 음악 산출물의 생성·유통을 관리하는 기본 시스템을 제시함
- 구체적으로는 사용자 요청 단계에서 권리자의 승인 여부를 자동으로 검증하고, 기준을 충족한 산출물에 한해 디지털 워터마크를 삽입·배포하는 방식임
- 워터마크에는 만료·갱신 기능이 포함되어 권리자가 산출물의 사용 기간을 조정할 수 있음. 이에 따라 제어 권한은 최초 생성 시점에 그치지 않고 이후 유통 기간 전반에 걸쳐 지속적으로 작동함

### • 권리자 조건에 맞는 환경에서만 AI 음악 산출물을 이용하도록 제한

- 두 번째 특허는 생성된 산출물이 권리자의 의도와 조건에 부합하는 환경에서만 생성·소비되도록 제어 범위를 정교화함
- 이를 위해 ‘이중 승인 구조’를 도입하여, 권리자가 사전에 등록된 허용·금지 규칙을 바탕으로 AI에 변환 요청 단계와 산출물 완성 단계에서 조건 충족 여부를 자동으로 검증함
- 사전 등록 조건은 단순한 사용 허가 여부를 넘어 권리자의 이미지·가치관 보호, 특정 주제 제한 등 세부 영역까지 포괄함
- 예를 들어 권리자가 자신의 음성이 특정 주제의 곡에 활용되는 것을 거부한 경우, 사용자의 요청이 입력 단계에서 차단되며, 우회적 표현으로 생성된 산출물도 출력 단계에서 차단됨
- 검토크를 통과한 산출물도 ‘맥락 제한 재생(context-restricted playback)\*’ 기술을 통해 권리자가 사전에 허가한 플랫폼 환경 안에서만 재생됨
- 여기에 블록체인\*\* 기반 스마트 계약(smart contract)\*\*\*이 결합되면서 무단 외부 유출 차단과 이용 내역에 따른 수익 정산까지 자동 처리되는 방식임

\* 맥락 제한 재생(context-restricted playback): 사전에 승인된 특정 플랫폼, 앱, 혹은 특정 상황 내에서만 콘텐츠가 재생되도록 제한하는 방식

\*\* 블록체인(Blockchain): 거래·계약 등 데이터를 일정 단위의 블록에 기록하고, 이를 시간 순서에 따라 체인 형태로 연결해 여러 참여자가 공동으로 검증·저장하는 분산형 데이터 기록 기술

\*\*\* 스마트 계약(smart contract): 블록체인 기술을 기반으로 사전에 설정된 일정 조건이 충족되면 중개자 없이 자동으로 계약 내용(수익 정산 등)이 실행되도록 하는 시스템

### • AI로 생성한 파생 상품 디자인까지 권리 제어 범위에 포함

- 세 번째 특허는 오디오 산출물에 국한되던 권리 제어 원리를 물리적·가상 파생 상품 영역까지 전면 확장함. 이를 통해 음악·원작자와 연계된 상품 디자인을 AI로 생성·구매하는 과정까지 관리 대상으로 포함함
- AI 기반의 파생 상품 생성 시, 권리자가 사전에 등록된 브랜드 가이드라인과 자동 대조되며, 기준 미달의 디자인은 생성 또는 승인 단계에서 차단됨
- 최종 승인된 디자인은 실물 상품의 경우 주문 제작 및 온라인 쇼핑몰 판매 경로로 연결되고, 가상 상품의 경우 3차원 렌더링(3D Rendering)\*을 거쳐 메타버스\*\* 등 가상 환경에서 활용됨
- 이처럼 유니버설뮤직그룹의 특허 포트폴리오는 단일 음원에서 파생되는 오디오 산출물뿐 아니라 실물 상품과 가상 상품까지 하나의 사전 권리 제어 모델 안에서 관리하도록 설계됨

\* 3D 렌더링(3D rendering): 컴퓨터 그래픽 기술을 통해 2차원 데이터를 기반으로 3차원의 입체적인 화상이나 모델을 구현하는 작업

\*\* 메타버스(Metaverse): 가상현실(VR), 증강현실(AR), 3D 그래픽, 네트워크 기술 등을 기반으로 현실과 유사한 사회·경제·문화 활동이 이루어지는 디지털 가상공간

## AI 음악 산출물의 권리 제어를 둘러싼 표준 경쟁과 제도 정비 쟁점

### • 사후 구제에서 사전 설계로 변화하는 권리 제어 방식

- 이상에서 살펴본 유니버설뮤직그룹의 특허 포트폴리오는, 음악 산업의 저작권 관리가 침해 발생 이후의 법적 구제에서 AI 음악 산출물의 생성·유통 단계의 사전 권리 제어로 이동하고 있음을 보여줌
- 대형 음반사가 관련 기술 특허를 선제적으로 확보하는 것은 개별 음악가의 권리 보호를 넘어, AI 음악 라이선싱 플랫폼의 운영 규칙과 기술 표준을 직접 설계하려는 전략으로 해석됨
- 특히 저작권 기준이 명확하지 않아 AI 리믹스 및 AI 음악 생성 서비스 출시를 신중하게 검토하는 빅테크 기업들에게는, 메이저 음반사가 구축한 사전 권리 제어 모델이 산업 협력의 사실상 표준이 될 가능성을 시사함

### • 개방형·폐쇄형 생태계 경쟁에 따른 시장 접근성과 수익 배분 쟁점

- 향후 AI 음악 시장에서는 AI 음악 산출물의 외부 활용을 일정 부분 허용하는 개방형 생태계와, AI 음악 산출물의 생성·유통·정산을 플랫폼 내부에서 관리하는 폐쇄형 생태계가 서로 다른 권리 제어 표준으로 경쟁할 가능성이 큼
- 폐쇄형 생태계가 확산될 경우 권리자 보호와 수익 정산의 예측 가능성은 높아질 수 있으나, 대형 음반사가 특허와 플랫폼 인프라를 배타적으로 제어하면 개인 창작자나 소규모 기획사의 시장 접근성이 제약될 수 있음
- 따라서 폐쇄형 생태계의 권리 제어 모델이 안정적으로 정착하기 위해서는 기술적 제어 장치의 고도화와 함께, 권리 귀속 기준, 동의 절차, 수익 배분 방식에 대한 법적·제도적 정비가 함께 논의될 필요가 있음

## 참고문헌

- Derek Coleman, "AI Music News: April 2026 Update on Suno's Stalemate, Klay's Big-Three Deal, and the Charts Going Synthetic", Opus, 2026.04.27., <https://www.opus.pro/blog/ai-music-news-april-2026>
- Daniel Brooks, "Music Industry AI Lawsuits Tracker 2026: Live Status", Chartlex, 2026.04.28., <https://www.chartlex.com/blog/business/music-industry-ai-lawsuits-tracker-2026>
- Murray Stassen, "Inside the UMG-backed patent portfolio targeting AI music derivatives: A technical blueprint for the walled garden model?", Music Business Worldwide, 2026.04.22., <https://www.musicbusinessworldwide.com/inside-the-umg-backed-patent-portfolio-targeting-ai-music-derivatives-a-technical-blueprint-for-the-walled-garden-model/>



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 시나미디어의 엣지 워터마킹 출시와 실시간 스트리밍 보호 기술의 진화

### 불법 복제 영상 스트림 탐지 및 차단을 위한 엣지 워터마킹 솔루션 등장

- 시나미디어, 콘텐츠아머 엣지 워터마킹 출시
  - 비디오 소프트웨어 제공업체인 시나미디어(Synamedia)가 불법 복제 영상 스트림의 탐지와 차단을 위한 솔루션인 '콘텐츠아머 엣지 워터마킹(ContentArmor Edge Watermarking)'을 출시함
  - 엣지 워터마킹은 콘텐츠가 이용자에게 전달되는 마지막 경유 지점인 CDN 엣지(CDN Edge)\*에 워터마크를 삽입하는 방식으로, 불법 복제 콘텐츠의 유출 경로를 신속하게 파악하고 차단하는 데 초점을 둠
  - 시나미디어 측은 엣지 워터마킹이 빠른 탐지 속도와 비용 절감 효과를 언급하며, 불법 복제 시 피해가 즉각적으로 발생하는 라이브 스포츠 중계와 같은 실시간 스트리밍 콘텐츠 보호에 적합하다고 설명함<sup>1)</sup>

\* CDN 엣지(CDN Edge): CDN(content delivery network)은 디지털 콘텐츠를 전 세계 여러 지역의 서버에 분산 저장해 이용자와 가까운 서버에서 빠르게 전달하는 인프라 체계이며, CDN 엣지는 이를 구성하는 서버 중 이용자와 가장 가까운 마지막 전달 지점임

### 엣지 워터마킹의 구조와 기존 A/B 방식 대비 개선점

- 기존 A/B 워터마킹의 구조와 한계
  - 엣지 워터마킹은 기존 A/B 포렌식 워터마킹 방식의 인프라 유지 비용 부담과 탐지 지연 문제를 개선함
  - 기존 A/B 포렌식 워터마킹은 송출될 영상을 여러 개의 짧은 구간으로 나눈 뒤, 각 구간에 서로 다른 워터마크 정보를 반영한 A버전과 B버전을 사전에 생성해 CDN에 저장해 두는 방식으로 작동함
  - 스트리밍 과정에서 CDN에 저장된 A/B 버전이 이용자별로 서로 다른 조합으로 전송됨
  - 즉, 이용자 입장에서는 단일 영상이 연속 재생되는 것으로 인식되나, 실제 시청 기록은 '1번 구간은 A버전, 2번 구간은 B버전, 3번 구간은 A버전...'처럼 구간별로 서로 다른 버전의 영상이 조합된 형태로 나타남
  - 이를 활용하면 불법 복제 영상이 발견될 경우, 영상에 포함된 A·B 배열을 분석함으로써 어느 이용자 계정에서 유출됐는지를 역추적할 수 있음
  - 그러나 이 방식은 각 구간별로 A버전과 B버전 두 종류의 스트림을 사전에 생성하고 CDN에 저장 및 캐싱(caching)\*해야 하므로, 저장 공간과 전송 대역폭 측면에서 비용 부담이 커진다는 한계가 지적됨
  - 또한 유출 영상에서 워터마크 배열을 추출·분석하는 데 일정 시간이 소요되어, 라이브 스포츠처럼

1) NORDIC MEDIA NEWS, "The next chapter in the fight against piracy: Synamedia launches industry's first edge watermarking solution", 2026.03.27., <https://www.nordicmedia.news/the-next-chapter-in-the-fight-against-piracy-synamedia-launches-industrys-first-edge-watermarking-solution/>

불법 송출을 신속하게 차단해야 하는 콘텐츠에서는 실시간 대응력이 제한될 수 있음

\* 캐싱(caching): 자주 요청되는 콘텐츠를 CDN 서버에 미리 저장해 두어 사용자 요청 시 빠르게 전달할 수 있도록 하는 방식으로, 저장된 데이터가 많을수록 서버 운영 비용이 증가함

#### • 엣지 워터마킹의 작동 원리와 인프라 부담 완화 설계

- 엣지 워터마킹은 스트리밍 영상을 CDN을 통해 배포한 뒤, 콘텐츠가 이용자에게 전달되기 직전의 CDN 엣지 단계에서 워터마크 정보를 삽입하는 방식으로 작동함
- 엣지 단계란 콘텐츠가 이용자에게 전달되기 직전, 이용자와 가까운 CDN 서버에서 영상 전송이 이뤄지는 단계를 의미함
- 이를 택배에 비유하면, 상품이 중앙 물류창고에서 소비자에게 바로 배송되는 것이 아니라 지역 배송 거점을 거쳐 최종 배송되는 구조와 유사함. 이때 CDN 엣지는 영상 콘텐츠의 '지역 배송 거점'에 해당하며, 엣지 워터마킹은 이 최종 전달 거점에서 각 고객에게 제품이 배송되기 전에 고객별 식별 표시를 덧입히는 방식으로 볼 수 있음
- 이처럼 워터마크 정보가 최종 전달 단계에서 삽입되기 때문에, 기존 A/B 방식처럼 구간별로 두 종류의 버전을 사전에 만들어둘 필요가 없음. 또한 두 종류의 워터마크 스트림을 별도로 운용할 필요가 없어져 원본 서버와 엣지 서버 간 트래픽을 감축하고 CDN 캐시 메모리 점유를 최소화할 수 있음

#### • 압축 비트스트림 기반 고밀도 워터마킹과 탐지 시간 단축

- 엣지 워터마킹은 비트스트림(Bitstream)\*에 직접 워터마크를 삽입하는 방식으로, 불법 복제 스트림을 식별하는 속도 측면에서도 기존 방식과 차이를 보임
- 비트스트림 직접 삽입 방식은 영상을 압축에서 해제한 뒤에 워터마킹 처리를 하는 것이 아니라, CDN을 통해 전송되는 압축된 영상 데이터 안에 워터마크 정보를 직접 삽입하는 방식임
- 압축 상태에서 직접 삽입하면 단위 데이터 안에 더 많은 식별 정보를 촘촘하게 담을 수 있어 워터마크 밀도가 높아짐
- 밀도가 높을수록 영상 전체를 분석하지 않아도 일부 구간만으로 유출 세션을 특정할 수 있어 A/B 방식 대비 분석 시간이 절반 수준으로 단축됨
- 시나미디어는 이를 통해 워터마크 삽입 및 추출 시간이 기존 대비 절반 수준으로 줄어들며, 불법 복제 스트림 감지부터 해당 불법 접속 차단까지의 전체 처리 시간이 5분 이내로 단축될 수 있다고 설명함<sup>2)</sup>

\* 비트스트림(bitstream): 영상이 압축-인코딩된 상태의 데이터 흐름으로, CDN을 통해 이용자에게 전달되는 실제 전송 단위임

## 상용화 현황 및 기술적 의미

#### • 엣지 워터마킹의 상용화 현황과 실시간 보호 기술로서의 의미

- 2026년 3월, 시나미디어는 자체 CDN 제품인 플루이드 엣지CDN(Fluid EdgeCDN)에 엣지 워터마킹을 통합하며 상용화 단계에 진입함
- 시나미디어는 해당 솔루션을 특정 CDN 사업자에 종속되지 않는 CDN 비종속형(CDN-agnostic) 구조로 설계했으며, 시나미디어 제품이 아닌 다른 CDN 환경에도 적용 가능하다고 설명함<sup>3)</sup>

2) NORDIC MEDIA NEWS, "The next chapter in the fight against piracy: Synamedia launches industry's first edge watermarking solution", 2026.03.27., <https://www.nordicmedia.news/the-next-chapter-in-the-fight-against-piracy-synamedia-launches-industrys-first-edge-watermarking-solution/>

3) NORDIC MEDIA NEWS, "The next chapter in the fight against piracy: Synamedia launches industry's first edge watermarking solution", 2026.03.27.,

- 이는 해당 솔루션이 특정 전송 인프라에 국한되지 않고 다양한 스트리밍 환경에서 활용될 수 있는 가능성을 보여준다는 점에서 의미가 있음
- 이러한 특성을 바탕으로 옛지 워터마킹은 실시간 스트리밍 환경에서 유출 세션에 대한 대응이 사후 추적 중심에서 실시간 탐지·차단 중심으로 전환되고 있음을 보여주는 사례로 평가됨

**[표1] A/B 포렌식 워터마킹과 옛지 워터마킹 비교**

구분	A/B 포렌식 워터마킹	옛지 워터마킹
워터마크 삽입 시점	- 콘텐츠 배포 전 사전 생성 단계 워터마크 정보 삽입	- 이용자의 재생 요청 시점에 CDN 엣지에서 워터마크 정보를 실시간으로 삽입
콘텐츠 버전 수	- 구간별로 A/B버전 두 종류를 CDN에 저장하여 배포	- 하나의 원본 스트림만 배포
인프라 부담	- 두 버전 저장으로 CDN 캐시 용량 증가 - 원본 서버와 옛지 서버 간 전송 대역폭 부담 증가	- 단일 스트림 운용으로 저장·전송 대역폭 부담 절감
유출 세션 탐지 방식	- 영상 전체의 A·B 배열을 분석하여 유출 세션 특정 - 분석 및 탐지에 일정 시간 소요	- 비트스트림에 삽입된 워터마크 정보만을 식별 - A/B 방식 대비 탐지 시간 단축

출처: NORDIC MEDIA NEWS, "The next chapter in the fight against piracy: Synamedia launches industry's first edge watermarking solution", 2026.03.27., <https://www.nordicmedia.news/the-next-chapter-in-the-fight-against-piracy-synamedia-launches-industrys-first-edge-watermarking-solution/>

## 참고문헌

- NORDIC MEDIA NEWS, "The next chapter in the fight against piracy: Synamedia launches industry's first edge watermarking solution", 2026.03.27., <https://www.nordicmedia.news/the-next-chapter-in-the-fight-against-piracy-synamedia-launches-industrys-first-edge-watermarking-solution/>



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 시댄스 2.0, C2PA·비가시 워터마킹으로 AI 영상 출처 표시 강화

### 시댄스 2.0의 출시 초기 논란과 글로벌 출시 재개

#### • 중국 출시 직후 실존 인물·저작권 캐릭터 논란 확산

- AI 영상 제작 모델 시댄스 2.0(Seedance 2.0)은 출시 직후 톰 크루즈(Tom Cruise), 브래드 피트(Brad Pitt) 등 실존 인물을 활용한 가공 영상이 유포되면서 논란이 확산됨
- 미국영화협회(Motion Picture Association, MPA)\*는 2026년 2월 시댄스 2.0이 실효성 있는 저작권 보호 조치 없이 출시되었다고 비판하는 공식 성명을 발표하였으며, 미국배우조합(SAG-AFTRA)\*\*도 회원 배우들의 초상권 무단 도용 사례를 들어 공식 반발함<sup>1)</sup>
- 이후 올해 2월 13~14일 디즈니(Disney), 파라마운트(Paramount) 등 6개 주요 할리우드 스튜디오가 침해 행위 중단을 요구하는 경고장을 발송함<sup>1)</sup>

\* 미국영화협회(Motion Picture Association, MPA): 주요 미국 영화·콘텐츠 제작사들을 대표하는 업계 단체

\*\* 미국배우조합(Screen Actors Guild-American Federation of Television and Radio Artists, SAG-AFTRA): 미국 영화·TV·라디오·온라인 미디어 분야의 배우와 방송인을 대표하는 노동조합

#### • 안전장치 보강을 거친 글로벌 출시 재개

- 바이트댄스(ByteDance)는 논란 직후인 2026년 2월 16일 안전장치 보강을 약속하고, 3월 15일 글로벌 출시를 일시 중단함
- 이후 3월 25~31일에 걸쳐 브라질, 인도네시아, 멕시코, 베트남 등 7개국에 AI 산출물 출처 표시·추적 기능과 실존 인물·캐릭터 입력 차단 기능을 적용한 형태로 글로벌 출시를 재개함
- 다만 3월 31일 재개 시점에는 규제 환경이 복잡한 미국과 인도가 의도적으로 제외되었고, 미국 진입은 약 1주일 뒤인 4월 7일 캡컷(CapCut), 드리미나(Dreamina), 피핏(Pippit) 등 자사 창작 도구 통합 방식으로 별도로 이뤄짐

### 시댄스 2.0의 출처 표시 기술 구성과 차단 장치

#### • 사후 식별을 위한 다층 출처 표시 구조

- 시댄스 2.0은 모든 AI 산출물에 C2PA(Coalition for Content Provenance and Authenticity) 콘텐츠 자격증명\*을 적용해 생성 도구, 생성 시점, AI 산출물 여부 등을 영상 파일 메타데이터로 기록하며, C2PA 콘텐츠 검증 도구는 해당 메타데이터를 판독함

1) Ana Maria Constantin, "ByteDance adds watermarking and IP guardrails to Seedance 2.0 as it begins cautious global rollout", The Next Web, 2026.03.31., <https://thenextweb.com/news/bytedance-seedance-watermarking-ip-global-rollout>

- 더넥스트웹(The Next Web)에 따르면 C2PA 콘텐츠 자격증명은 플랫폼 간 AI 산출물 식별을 위한 업계 표준 프로토콜로 자리잡고 있음<sup>2)</sup>
- C2PA 외에도 시덴스 2.0은 가시적 워터마크와 비가시적 워터마킹 기술을 병행 적용함. 가시적 워터마크는 이용자가 화면상에서 AI 생성 여부를 직접 인지할 수 있는 고지 기능을 수행함
- 비가시적 워터마킹은 AI 산출물이 플랫폼 외부로 공유되거나 영상 편집을 거쳐 변형된 이후에도 식별 정보를 유지하며, C2PA 메타데이터만의 추적 한계를 보완함
- C2PA 콘텐츠 자격증명, 가시적 워터마크, 비가시적 워터마킹을 함께 적용하는 다층 출처 표시 구조는 시스템 자동 식별, 이용자 고지, 사후 추적 기능을 결합한 방식으로, 단일 기술 적용 대비 출처 식별과 추적 가능 범위를 넓힘

\* C2PA(Coalition for Content Provenance and Authenticity) 콘텐츠 자격증명: 어도비(Adobe), 마이크로소프트(Microsoft) 등이 주도하는 콘텐츠 출처 검증 표준 기술로, 디지털 콘텐츠의 생성 도구, 생성 일시, AI 산출물 여부 등을 메타데이터로 기록해 콘텐츠에 내재화함

#### • 사전 차단 필터를 통한 위험 요소 사전 억제

- 시덴스 2.0은 출처 표시 기술과 별개로, 실존 인물의 얼굴이 포함된 이미지나 영상을 입력해 AI 영상을 생성하는 것을 차단하고, 저작권이 있는 캐릭터의 무단 생성을 방지하는 사전 차단 필터를 함께 적용함
- 이러한 차단 기능은 분쟁의 직접적 원인이 된 실존 인물·저작권 캐릭터 산출물 생성을 사전에 억제하는 장치임
- 출처 표시 기술이 사후 식별과 추적에 초점을 두는 것과 달리, 사전 차단 필터는 생성 이전 단계에서 위험을 줄이는 보완적 기능을 함

#### • 외부 검증에서 드러난 안전장치의 구조적 한계

- 바이트댄스는 제3자 레드팀(red-team)\* 파트너와 협력해 시덴스 2.0의 안전장치를 검증함
- 다만 더넥스트웹에 인용된 업계 관찰자들의 외부 테스트 결과, 창의적 프롬프트 입력을 통해 차단 필터가 우회될 수 있다는 점이 제기됨<sup>2)</sup>
- 특히 유사성 인접(likeness-adjacent)\*\* 방식 등 우회 가능성이 보고되어, 현재 적용된 기술 구성이 모든 침해 유형을 차단하기에는 한계가 있음

\* 레드팀(red-team): 시스템의 취약점이나 제한 기능의 우회 가능성을 탐색하기 위해 의도적으로 부정적 활용을 시도하는 보안·안전성 점검 방식

\*\* 유사성 인접(likeness-adjacent): 실존 인물이나 저작권 캐릭터를 기술적으로 복제하지는 않으면서도 해당 인물·캐릭터를 연상시키는 방식의 산출물

## 출처 표시 기술의 규제 요건화와 다층 관리 과제

#### • EU AI Act 시행과 출처 표시 기술의 규제 요건화

- EU 인공지능법(EU AI Act)의 투명성 요건(transparency requirements)은 2026년 8월부터 시행되며, 생성형 AI 시스템 제공자에게 AI 산출물을 시스템 자동 식별이 가능한 형식으로 표시하고 딥페이크의 인공 생성 여부를 공개하도록 의무화함
- C2PA 콘텐츠 자격증명은 EU AI Act의 출처 표시 요건에 대응 가능한 기술적 구현 방법으로 주목받고 있으며, 더넥스트웹은 시덴스 2.0의 C2PA 탑재가 이러한 요건을 선제적으로 고려한 조치라고 평가함<sup>3)</sup>

2) Ana Maria Constantin, "ByteDance adds watermarking and IP guardrails to Seedance 2.0 as it begins cautious global rollout", The Next Web, 2026.03.31., <https://thenextweb.com/news/bytedance-seedance-watermarking-ip-global-rollout>

3) Ana Maria Constantin, "ByteDance adds watermarking and IP guardrails to Seedance 2.0 as it begins cautious global rollout", The Next Web, 2026.03.31., <https://thenextweb.com/news/bytedance-seedance-watermarking-ip-global-rollout>

### • 단일 기술 대응의 한계와 다층 관리의 필요성

- 시댄스 2.0이 3월 31일 글로벌 출시 재개 시점에 C2PA와 비가시 워터마킹 등 출처 표시 기능을 적용한 것은 EU AI Act 등 규제 요건 대응에 기여함
- 다만 외부 테스트에서 차단 필터의 우회 가능성이 보고된 점을 감안하면, C2PA와 워터마킹 탑재만으로는 출처 표시와 침해 방지 요건을 충분히 충족하기 어려움
- 결국 생성형 AI 영상 분야에서는 콘텐츠 자격증명·워터마킹과 입력 차단 필터, 사후 모니터링 체계를 함께 운영하는 다층 관리 방식이 요구됨

### 참고문헌

- AI News Silo, "Seedance 2.0 takes Sora's abandoned AI video crown", 2026.04.11., <https://ainewssilo.com/articles/seedance-2-bytedance-global-release>
- Ana Maria Constantin, "ByteDance adds watermarking and IP guardrails to Seedance 2.0 as it begins cautious global rollout", The Next Web, 2026.03.31., <https://thenextweb.com/news/bytedance-seedance-watermarking-ip-global-rollout>



# 저작권 이슈 브리프

SUMMARY

산업/기업

기술

## 에이전틱 AI의 문서 편집과 저작권 관리 과제

### AI 에이전트의 콘텐츠 편집 개입

#### • 에이전틱 코파일럿의 기능 변화

- 마이크로소프트(Microsoft)는 2026년 4월 22일 워드(Word)·엑셀(Excel)·파워포인트(PowerPoint)에 에이전틱(agentic)\* 코파일럿(Copilot)\*\* 기능을 정식 출시하며, 이용자 요청에 응답하는 방식에서 문서를 직접 편집 및 재구성하는 방식으로 기능 범위를 확장함
- 기존의 코파일럿은 편집 방향을 제안하는 수준에 머물렀으나, 에이전틱 코파일럿은 서식 적용·구조 재편·데이터 변환 등 다단계 작업을 AI가 자율적으로 순차 실행하는 구조로 작동함
- 해당 기능은 이용자가 직접 활성화해야 하나, 이후에는 편집 기능이 기본값으로 작동되어 별도 거부가 없으면 AI 편집이 자동으로 적용되는 구조임
- 마이크로소프트는 이용자가 변경 내역을 검토하고 조정할 수 있도록 단계별 편집 과정을 사이드바에 표시하는 기능을 제공하나, 전 과정을 이용자가 추적하기는 어려운 구조임
- 이는 문서 작성 과정에서 AI가 보조 도구를 넘어 실질적인 편집 주체로 개입하는 구조로 전환되고 있음을 보여줌

\* 에이전틱(agentic): AI가 단일 응답에 그치지 않고 목표를 설정한 후 복수의 단계를 자율적으로 계획·실행하는 방식을 의미하며, 문서 편집·데이터 분석·프레젠테이션 생성 등 복합 작업에 적용됨

\*\* 코파일럿(Copilot): 마이크로소프트가 개발한 생성형 AI 서비스 및 AI 비서 브랜드명. 문서 작성, 코드 생성, 검색, 업무 자동화 등을 지원함

#### • 설치 기본값과 이용자 동의 구조

- 파이어폭스(Firefox)\* 개발사 모질라(Mozilla)는 2026년 4월 마이크로소프트가 이용자 동의 없이 코파일럿을 윈도우(Windows) 시스템에 자동 설치하고, UI 설계를 통해 이용자의 선택권을 사실상 제한하고 있다고 비판함
- 모질라의 비판 근거 중 하나는 마이크로소프트가 유럽경제지역(European Economic Area, EEA)\*\*만 자동 설치 대상에서 제외했다는 점으로, 이용자 동의 구조가 규제 환경에 따라 달라지고 있음을 시사함
- 마이크로소프트는 소비자용 앱에서 코파일럿 명칭 변경·접근 경로 축소 조치를 취했으나, 기업용 앱의 에이전틱 편집 기능은 여전히 기본값 활성화 구조로 유지되어 이용자 동의 문제는 해소되지 않은 것으로 분석됨

\* 파이어폭스(Firefox): 모질라가 개발한 오픈소스 기반 웹 브라우저로 개인정보 보호 기능과 확장 기능 지원을 강점으로 하며, 크롬(Chrome), 엣지(Edge) 등과 함께 대표적인 인터넷 브라우저 중 하나로 사용됨

\*\* 유럽경제지역(European Economic Area, EEA): 회원국 간 상품·서비스·자본·인력의 자유로운 이동을 보장하기 위해 구성된 경제권. 유럽연합 회원국과 일부 유럽자유무역연합(EFTA) 국가가 참여하고 있음.

## 에이전틱 편집에 따른 저작권 귀속·이용자의 기능 제어

### • 저작권 귀속 기준과 에이전틱 편집

- 미국 저작권청(U.S. Copyright Office)\*은 2025년 1월 발표한 AI 저작권 보고서에서 저작권 보호가 성립하려면 인간의 실질적 창작 기여가 있어야 하며, 프롬프트 입력만으로는 저작자로 인정받기 어렵다는 입장을 재확인함
- 또한 AI가 보조적 역할에 머물면서 인간의 창작 과정을 지원하는 경우에는 저작권 보호가 가능하나, 구체적 표현을 AI가 산출한 경우에는 해당 부분의 저작권 귀속이 불분명해진다고 명시함
- 그러나 에이전틱 코파일럿이 서식 재편·문장 재구성·데이터 시각화 등을 다단계로 자율 실행할 경우, 최종 산출물 중 인간과 AI의 기여를 사후에 구분하기 어려운 구조가 될 수 있음
- 이러한 구조는 기업이 해당 산출물을 외부에 공개하거나 상업적으로 활용할 때 권리 주장의 근거를 약화시키며 관련된 법적 리스크가 이어질 가능성이 함께 제기됨

\* 미국 저작권청(U.S. Copyright Office): 미국 의회도서관 산하 기관으로, 저작권 등록·정책 수립·AI 저작권 가이드라인 발행 등을 담당하며, 2025년 1월과 5월 두 차례에 걸쳐 AI와 저작권에 관한 종합 보고서를 발표함

### • 기본값 설계와 이용자의 기능 제어

- 에이전틱 코파일럿은 별도 설정 없이 편집 기능이 자동으로 켜지는 기본값 활성화 구조로 제공되어, 이용자가 명시적으로 거부하지 않는 한 AI가 개입하는 환경을 만들며 이는 저작권 판단의 핵심 기준을 모호하게 만드는 요인으로 작용함
- 마이크로소프트는 이용자에게 AI를 통한 변경 내역의 검토 및 승인 기능을 제공한다고 밝혔으나, 편집 과정의 판단 근거와 개입 범위를 실시간으로 확인하거나 기능을 제어하기 어려운 구조로 알려짐
- 이용자의 기능 제어 문제는 에이전틱 코파일럿이 조직 내 이메일·회의록·공유 문서 등을 참조해 편집을 실행하는 환경에서 더욱 두드러지며, 어떤 데이터가 산출물에 반영됐는지 이용자가 파악하기 어려운 구조임
- 더불어 동의 설계를 규제 지역인 EEA에서만 달리 적용하는 방식은 이용자 보호가 법적 의무 이행의 최소 기준에서 설계된다는 점을 시사하며 자율적 선택권 보장과는 거리가 있는 것으로 평가됨

[표1] 마이크로소프트 코파일럿 이용자의 기능 제어

기능 및 항목	이용자 기능 제어 범위	EEA 적용	비고
에이전틱 편집 (Word·Excel·PPT)	전체 비활성화 가능, 단계별 실시간 개입 불가	동일 적용	2026년 4월 정식 출시, MS 365 Copilot 구독 대상
Windows 앱 자동 설치 (M365 Copilot 앱)	관리자 설정으로 차단 가능	제외	EEA 대상 자동 설치 제외는 개인정보 보호 규정 등 규제 대응 맥락으로 해석됨
대화 기록 활용 (개인화·모델 학습)	개별 대화 삭제 또는 전체 기록 삭제 가능	동일 적용	기본 18개월 보관, 학습 제외 설정 별도 필요
타 MS 제품 데이터 연동 (Edge·Bing·MSN 등)	메모리 설정 내 토글로 해제 가능, 메뉴 접근 어려움	동일 적용	설정 위치가 메모리 탭 하위에 위치해 가시성 낮음
GitHub Copilot 코드 학습 활용	계정 단위 해제 가능, 저장소별 개별 제어 불가	동일 적용	2026년 4월 24일부터 적용, 기업·엔터프라이즈 플랜은 제외

출처: 참고문헌 종합하여 재구성

## 시사점: 에이전틱 AI 산출물 관리 과제

### • 저작권 귀속 판단 기준의 구체화

- 미국 저작권청은 AI 개입 범위와 인간 기여 여부를 사안별로 판단한다는 방침을 제시했으나, 에이전틱 AI가 다단계 편집을 자율 실행하는 환경에서 '실질적 인간 기여'의 경계를 어디에 설정할지에 대한 구체적 기준은 아직 마련되지 않은 상태임
- 에이전틱 편집이 기업 업무 환경에 보편화될수록 AI가 작성·수정한 문서의 저작권 귀속을 사후에 증명하기 어려워지며, 입법·정책 차원의 판단 기준 구체화가 과제로 부각될 가능성이 있음
- 기업 차원에서는 마이크로소프트의 단계별 편집 이력 표시 기능처럼 AI 개입에 따른 단계별 로그와 사용자 승인·거부 내역을 보존하는 체계를 갖추는 것이 저작권 귀속 증명을 위한 필수적 전제 조건이 될 수 있음

### • 이용자 동의와 플랫폼 책임 논의

- 모질라가 제기한 비판과 EEA 적용 제외 사례는 플랫폼의 동의 설계가 규제 의무 이행의 최소 기준에 맞춰 설계될 경우, 규제 공백 지역 이용자의 기능 제어 권한과 선택권이 상대적으로 제한될 수 있음을 보여줌
- AI 기능의 기본값 활성화 방식은 이용자의 적극적 거부 의사를 전제하는 구조로, 옵트인(opt-in)\* 원칙과의 간극이 좁혀지지 않는 한 플랫폼 책임에 관한 논의가 지속될 것으로 전망됨
- 다만 에이전틱 AI의 기본값 설계가 보편화될수록, 이용자 동의 구조를 플랫폼 자율에 맡기기보다 규제기관이 옵트인 원칙의 적용 범위와 예외 기준을 명시하는 방향으로 논의가 이어질 가능성이 있음

\* 옵트인(opt-in): 기본값이 '미참여'로 설정된 상태에서, 이용자가 명시적으로 동의 의사를 표시해야 참여가 활성화되는 방식. 기본값이 '참여'로 설정되어 있어 이용자가 거부 의사를 표시해야 비활성화되는 옵트아웃(opt-out) 방식과 대비됨

## 참고문헌

- Carly Page, "Microsoft gives your Word documents an AI co-author you didn't ask for", The Register, 2026.4.23., <https://www.theregister.com/software/2026/04/23/microsoft-adds-uninvited-ai-co-author-to-word-docs/5220927>
- Guru Baran, "Mozilla Criticizes Microsoft for Installing Copilot on Windows Without User Consent", Cybersecurity News, 2026.4.13., <https://cybersecuritynews.com/mozilla-criticizes-microsoft-for-copilot/>
- Matt Sumner 외 1인, "Taking a Closer Look: What the Copyright Office's 2025 AI Reports Mean for 2026", Neal & Leroy, 2026.1.26., <https://www.nealanderoy.com/post/taking-a-closer-look-what-the-copyright-office-s-2025-ai-reports-mean-for-2026>



## 주간 기술 동향

### 출처 검증과 변조 탐지를 동시에 달성하는 이중 채널 워터마킹, Dual-Guard

#### • AI 워터마킹 기술의 한계와 AI 생성 이미지 변조 및 복제를 방지하기 위한 이중 채널 검증 기술의 등장

생성형 AI 모델의 발달로 누구나 이미지를 쉽게 만들고 수정할 수 있게 되었지만, 동시에 허가 없는 복제와 무단 사용이라는 위협에도 상시 노출되어 있다. 생성형 AI를 활용해서 만든 이미지에 수정, 편집 등을 통해 창작성을 더했다면 권리자들은 본인이 AI를 통해 생성한 이미지가 어디에서 어떻게 사용되는지 추적하기 어렵고, 권리 침해 사실을 입증하는 것도 쉽지 않다. 특히 AI가 생성한 이미지가 유통 과정에서 출처 정보를 잃은 채 복제되거나 변조될 경우, 권리자는 해당 이미지가 자신의 모델에서 생성되었음을 입증하기 어려운 상황이다. 이러한 상황 속에서 AI 생성 이미지의 유통 및 사용을 파악하기 위해 데이터 생성 시간 정보, 메타데이터 관리, AI 기반 콘텐츠 탐지 등 다양한 기술 도구가 활용되고 있다.

이에 따라 AI가 생성한 이미지에 출처 정보를 직접 삽입하는 AI 워터마킹이 권리 보호의 수단 중 하나로 주목받고 있다. 기존의 가시적 워터마크가 미관을 해치거나 간단한 편집으로 제거될 수 있었던 것과 달리, AI 기술을 적용한 비가시 워터마킹은 콘텐츠의 품질을 유지하면서도 출처 정보를 이미지 안에 은닉할 수 있다. 그러나 단일 채널로 워터마크를 삽입하는 기존 방식은 공격자가 노이즈를 제거하며 이미지를 생성하는 확산 모델로 이미지를 재생성하거나 일부 영역만 편집하는 경우 워터마크가 소실되거나 출처 판별이 불가능해지는 기술적 한계를 가진다.

기존의 픽셀 도메인 워터마킹 기술은 일반적인 이미지 변형 공격에는 견디지만, 생성형 AI를 이용한 재생성 공격에는 워터마크가 제거될 위험이 높다. 확산 모델의 초기 노이즈에 워터마크를 삽입하는 방식은 재생성 공격에 강하지만, 이미지의 어느 부분이 변조되었는지 위치를 특정하지 못한다는 한계가 있다. 또한 공격자가 역변환을 통해 워터마크가 담긴 초기 노이즈를 복원한 뒤 악의적인 프롬프트로 재활용하면, 타인의 워터마크를 악용하는 프레이밍 공격도 가능하다.

이러한 배경에서 단일 채널의 한계를 동시에 극복하는 새로운 접근이 제안되고 있다. 본 보고서에서는 확산 모델의 두 지점에 서로 다른 역할의 워터마크를 동시에 삽입하는 이중 채널 잠재 워터마킹 프레임워크인 Dual-Guard를 분석한다. 이 기술은 초기 노이즈에 삽입된 가우시안 셰이딩 채널로 출처를 검증하는 동시에, 최종 잠재 표현에 삽입된 잠재 핑거프린트 코덱 채널로 변조된 영역의 위치를 블록 단위로 특정한다. 재생성·프레이밍·국소 편집 등 다양한 공격 시나리오에서 우수한 탐지율을 달성함과 동시에 이미지 품질 저하를 최소화한다는 점에서, AI 생성 이미지의 출처와 무결성을 기술적으로 입증할 수 있는 실용적 수단으로 평가된다.

## [사례] 이중 채널 잠재 워터마킹으로 출처와 변조를 동시에 잡는 Dual-Guard

### • 기존 단일 채널 워터마킹의 한계

- 픽셀 값을 직접 수정하여 정보를 숨기는 픽셀 도메인 워터마킹 기술(HiDDeN, StegaStamp, Stable Signature 등)은 노이즈 추가나 압축 같은 일반적인 변형에는 강하지만, 공격자가 생성형 AI 모델로 이미지를 다시 만들어내는 재생성 공격(reprompt attack)에는 워터마크가 완전히 사라짐
- 가우시안 셰이딩(Gaussian Shading, 이하 GS)\*을 비롯한 확산 모델(diffusion model)\*\*을 활용한 워터마킹 기술은 재생성 공격에는 강하지만, 해당 이미지의 진위 여부만 판정할 수 있을 뿐, 이미지의 어느 부분이 바뀌었는지는 알 수 없음
- 또한, 기존 방식은 타인의 워터마크를 악용하는 ‘프레이밍 공격(Framing Attack)’에 취약함. 공격자가 워터마크가 내재된 초기 노이즈를 역추적하여 악의적인 유해 이미지를 새로 생성할 경우, 워터마크는 정상 검출되더라도 정작 콘텐츠가 바뀐 사실을 감지하지 못하는 맹점이 있음
- 조작 흔적을 분석해 이미지의 위변조 여부를 판별하는 포렌식 기술들은 전통적인 사진 편집 흔적을 찾도록 설계되어 있어, 생성형 AI 편집 도구가 만들어내는 변조에는 사실상 작동하지 않음

\* 가우시안 셰이딩(Gaussian Shading): 표준 가우시안 분포를 따르는 잠재 표현으로 워터마크를 삽입하는 기술로, 학습이 필요 없는 확산 모델 워터마킹 기법

\*\* 확산 모델(diffusion model): 무작위 노이즈에서 출발해 단계적으로 노이즈를 제거하며 이미지를 생성하는 AI 모델. 스테이블 디퓨전(Stable Diffusion), DALL-E 등이 대표적임

### • GS 채널: 이미지 생성 시작점에 출처 정보를 심는 방식

- GS 채널은 비밀 키 메커니즘을 활용하여 확산 모델이 이미지를 만들기 시작하는 초기 노이즈 단계에 출처 정보를 숨겨두는 방식으로, 이 이미지가 어떤 모델에서 만들어졌는지를 검증하는 역할을 담당함
- 비밀 키 기반 이진 패턴과 64비트 고유 식별 코드를 결합하여 워터마크 패턴을 만들고, 초기 노이즈의 원소가 양수 또는 음수 영역에서 샘플링되도록 유도하여 삽입하는 구조임
- 이렇게 만들어진 노이즈는 각 원소는 일반 확산 노이즈와 동일한 표준 가우시안 주변 분포를 유지하여 통계적으로 구별되지 않으며, 이미지 품질에 영향을 주지 않음
- 검증 시에는 조사 대상 이미지를 역방향으로 추적해 초기 노이즈를 복원한 뒤, 각 원소의 부호 패턴을 비밀 키로 해독하여 등록된 식별 코드와 일치하는 비율을 계산함. 이 일치율이 기준값(99.5%) 이상이면 해당 이미지의 출처가 인증됨
- GS 채널은 재생성 공격 후에도 초기 노이즈 패턴이 보존되는 특성 덕분에 해당 위협에 강하지만, 프레이밍 공격처럼 초기 노이즈는 그대로인데 콘텐츠만 교체된 경우는 탐지하지 못함
- GS 채널은 이미지 전체에 대한 출처 판정에만 특화되어 있으며, 어느 부분이 변조됐는지는 알 수 없음. 변조 위치 특징은 코덱 채널이 전담함

### • 잠재 핑거프린트 코덱 채널의 구조

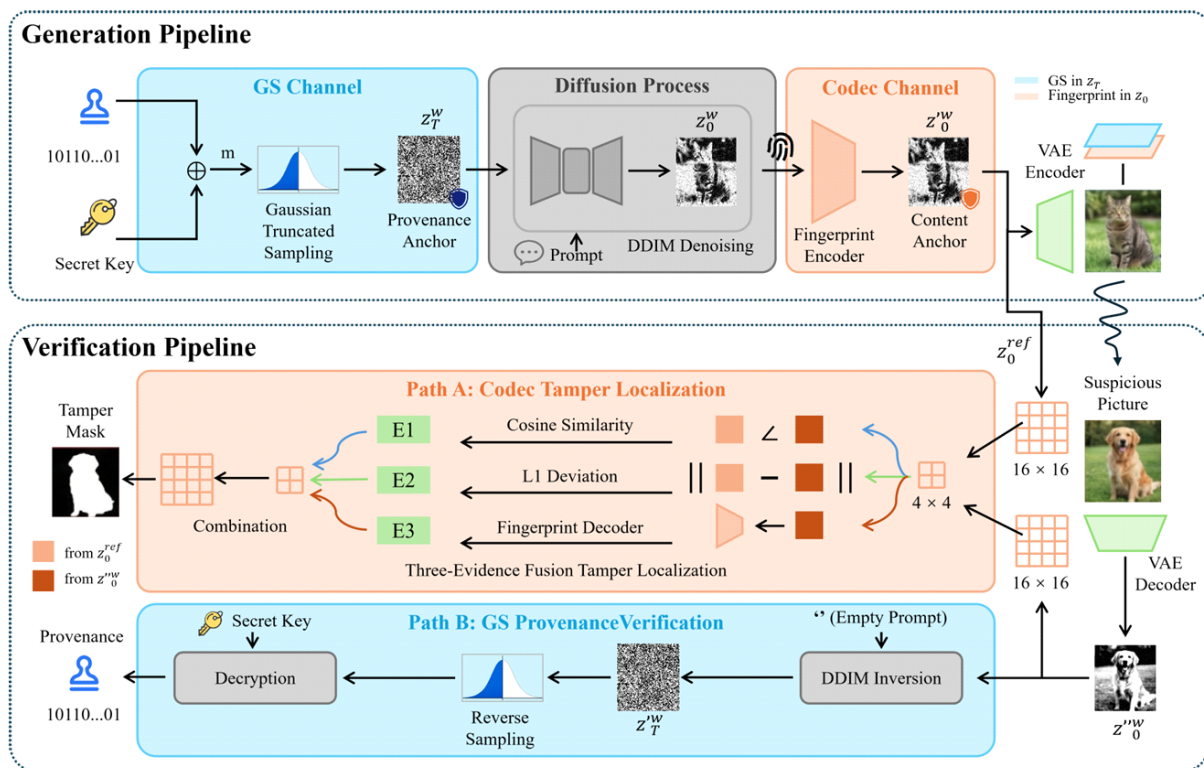
- 잠재 핑거프린트 코덱(Latent Fingerprint Codec) 채널은 확산 모델이 노이즈 제거를 마친 뒤 이미지로 변환되기 직전의 압축된 내부 표현인 잠재 표현(Latent Representation)\*에 고유 핑거프린트를 삽입하는 방식으로 작동함
- 재생성이나 프레이밍 공격은 잠재 표현을 크게 바꾸기 때문에, 코덱 점수가 급락하여 콘텐츠 불일치를 감지할 수 있음

- 삽입 과정에서는 64비트 식별 코드를 이미지와 같은 크기의 2차원 배열 형태인 특징 맵(spatial feature map)으로 변환한 뒤 잠재 표현과 합쳐 잔차 신호\*\*를 계산하고, 이를 고정 강도로 잠재 표현에 더함
- 오류 정정을 위해 64비트 코드를 3번 반복한 192비트 부호어로 확장하여 일부 블록이 손상되어도 원래 코드를 복원할 수 있도록 함
- 추출 과정에서는 고해상도 경로와 저해상도 경로로 나누어 각각 세부 패턴과 전체 문맥을 분석한 뒤, 학습된 가중치로 두 결과를 합산하여 최종 식별 코드를 복원함. 이 구조 덕분에 이미지 전체뿐 아니라 작은 블록 단위에서도 핑거프린트 추출이 가능함

\* 잠재 표현(Latent Representation): 확산 모델이 이미지를 처리할 때 사용하는 압축된 내부 데이터 형태. 실제 픽셀 이미지보다 훨씬 작은 크기로 이미지의 핵심 정보를 담고 있음

\*\* 잔차 신호(Residual Signal): 데이터에서 모델이 예측하거나 설명할 수 있는 주요 패턴을 빼고 남은 나머지 신호

[그림 1] Dual-Guard 생성 파이프라인 및 검증 파이프라인 개요



출처: JinFeng Xie 외 7인, "Dual-Guard: Dual-Channel Latent Watermarking for Provenance and Tamper Localization in Diffusion Images", arXiv, 2026.04.21., <https://arxiv.org/html/2604.19090v1>

### • 변조 위치 특정 메커니즘

- 변조 위치 특정은 조사 대상 이미지의 잠재 표현을 발급 시점에 소유자가 등록해둔 참조 잠재 표현(round-trip reference latent)과 블록 단위로 비교하는 방식으로 작동함. 두 잠재 표현을 각각 16×16 격자의 작은 블록들로 나누어 블록마다 3가지 증거를 독립적으로 계산함
- 첫 번째는 코사인 유사도로, 두 블록이 같은 방향을 가리키는지를 측정함. 두 번째는 편차 측정으로 블록 간 수치 차이를 포착하여, 변조된 블록일수록 원본과의 차이가 크게 나타남
- 세 번째는 블록별 핑거프린트 일치율로, 각 블록에서 독립적으로 핑거프린트를 추출해 등록된 코드와 비교하며, 참조 잠재 표현 없이도 계산할 수 있는 유일한 증거임

- 3가지 증거는 각각의 중요도에 따라 가중 합산하여 블록별 일관성 점수로 통합되며, 해당 점수가 기준값 미만인 블록을 변조 후보로 분류함. 이후 주변 블록과 연결되지 않은 블록을 제거하고 인접 블록을 묶는 후처리를 거쳐, 최종적으로 변조가 의심되는 영역을 표시한 히트맵을 출력함
- GS 채널과 코덱 채널은 순차적으로 작동함. GS 채널이 먼저 출처를 확인하고, 통과 시 코덱 채널이 콘텐츠 일치 여부를 검토하며, 두 채널을 모두 통과한 경우에만 변조 위치 특정 단계로 진입함

### • 성능 실험 결과

- Dual-Guard는 2,400개 샘플을 대상으로 한 실험에서 재생성 공격과 DiffEdit\* 공격에 대해 100%의 탐지율을 달성했으며, 검은색으로 특정 영역을 덮어씌우는 블랙 패치를 비롯해, 노이즈, 블러, JPEG 압축, 픽셀화, 복사-붙여넣기, 색상 변경, 텍스트 삽입 등 8가지 국소 변조 공격에 대해서도 99.9%의 탐지율을 기록함
- 변조 위치 특정 성능 평가에서는 이미지 수준 탐지율 100%를 달성했으며, 변조·예측 영역 간의 일치 수준을 나타내는 위치 특정 정밀도에서 대조 대상인 쉘(SEAL) 대비 7배 이상 높은 수치를 기록함
- 이미지 품질 면에서는 워터마크 삽입 후에도 이미지가 텍스트 설명과 얼마나 의미적으로 일치하는지를 나타내는 점수가 일반 생성 대비 1.1%만 감소했으며, 전문가도 육안으로 구별하기 어려운 수준의 이미지 품질을 유지함

\* DiffEdit: DDIM 역변환과 프롬프트 편집을 결합하여 이미지의 특정 영역만 의미적으로 수정하는 확산 모델 기반 편집 기법

[표 1] 공격 유형별 Dual-Guard 탐지 성능

Scenario	<i>n</i>	Auth Reject ↑	Reject CI <sub>95</sub>	Tamper Detect ↑	GS Pass	Dual Pass ↓	Loc. Gain ↑	Codec Extr.
Original (clean)	2400	0.003	[0.002, 0.007]	0.001	0.998	0.998	-	0.939
Reprompt	2400	1.000	[0.998, 1.000]	1.000	1.000	0.000	0.000	0.500
Local tamper	2400	0.999	[0.997, 1.000]	0.999	0.995	0.975	0.975	0.923
DiffEdit	2400	1.000	[0.998, 1.000]	1.000	0.633	0.601	0.601	0.929

출처: JinFeng Xie 외 7인, "Dual-Guard: Dual-Channel Latent Watermarking for Provenance and Tamper Localization in Diffusion Images", arXiv, 2026.04.21., <https://arxiv.org/html/2604.19090v1>

## 결론 및 시사점

### • 기술적 한계와 향후 과제

- Dual-Guard는 현재 소유자가 이미지 발급 시점에 참조 잠재 표현을 별도로 등록해두어야 변조 위치 특정이 가능한 구조로, 이 등록 없이는 핑거프린트 일치율 단독으로만 변조를 판단해야 해 위치 특정 정밀도가 낮아짐
- 실험은 스테이블 디퓨전(Stable Diffusion) 2.1 기반의 512×512 해상도 환경에서만 진행되었으며, 더 높은 해상도나 다른 확산 모델 구조에서도 동일한 성능이 유지되는지는 추가 검증이 필요함
- 현재 벤치마크는 공격자가 인코더 파라미터, 임계값, 참조 잠재 표현에 접근할 수 있는 화이트박스 적대적 공격 시나리오를 포함하지 않아, 이에 대한 실질적인 보안 신뢰성은 검증되지 않은 상태임

## 참고문헌

- JinFeng Xie 외 7인, "Dual-Guard: Dual-Channel Latent Watermarking for Provenance and Tamper Localization in Diffusion Images", arXiv, 2026.04.21., <https://arxiv.org/html/2604.19090v1>
- The Source, "How Independent Artists and Creators Are Using AI to Protect Their Original Work", 2026.04.29., <https://thesource.com/2026/04/29/how-independent-artists-and-creators-are-using-ai-to-protect-their-original-work/>