www.copyright.or.kr

저작권 이슈 브리프

COPYRIGHT ISSUE BRIEF

Weekly Report

2026. 5-1

한국저작권위원회
KOREA COPYRIGHT COMMISSION

저작권 이슈 브리프

SUMMARY | 산업/기업 | 기술

산업 게임 DRM 설계 방식과 이용자 영향 논쟁

▶ 게임 콘텐츠의 불법 복제를 방지하기 위한 기술적 보호 조치는 콘텐츠 산업의 핵심 저작권 관리 수단으로 활용되어 왔으나, 보호 기술이 사전에 정의된 규칙에 따라 접근을 제한하는 구조로 작동하면서 정당 이용자의 접근까지 차단하는 과잉 차단 문제가 지속적으로 제기되고 있다. 또한 보호 코드 적용이 게임 성능 저하를 유발한다는 인식도 업계에서 꾸준히 논란이 되어 왔다. 이에 데누보를 개발한 이르데토는 보호 기능의 이용자 체감 영향이 보호 코드의 강도보다 통합 방식과 실행 위치에 의해 결정된다는 분석을 제시하며, 저부담 구간 중심 배치·업데이트 후 재검증 등 통합 설계 원칙을 공개했다. 이는 보호 기술이 단순 차단 수단을 넘어 정당 이용자 경험과의 균형을 고려하는 운영 설계 중심으로 다뤄질 필요성을 시사하며, 향후 콘텐츠 산업에서는 보호 기술의 적용 방식과 저작권법상 예외 규정과의 정합성을 함께 검토하는 방향으로 정책 논의가 확대될 가능성이 있음을 보여준다.

산업 스포티파이의 AI 크레딧 도입으로 본 음악 스트리밍 플랫폼의 AI 자율 공시 모델

▶ 스포티파이는 2026년 4월 곡 제작 과정의 AI 활용 정보를 곡 크레딧 영역에 표시하는 자율 공시 기능 'AI 크레딧' 베타 버전을 공개했다. AI 크레딧은 보컬, 가사, 악기, 프로듀싱 등 영역별 AI 기여를 분리 표시해 이용자가 곡의 어느 부분에 AI가 활용됐는지 확인할 수 있도록 설계됐다. 그러나 해당 정보는 창작자의 자율 공시에 의존하고, 음반사·유통사를 거쳐 제출되는 구조여서 미공시 문제나 유통 경로별 공시 공백이 발생할 수 있다. 이에 학계에서는 정직하게 공시한 창작자가 불이익을 받을 수 있는 구조와 표시 의무화 부재에 따른 효과 제한을 지적하며, C2PA 등 검증 장치의 필요성을 제기한다. 자율 공시 모델이 음악 산업의 실효성 있는 표준으로 정착하려면 검증 메커니즘과 통일된 공시 형식이 산업 차원에서 함께 마련될 필요가 있다.

산업 오픈소스 소프트웨어의 안정성과 거버넌스 확보를 위한 '책임 있는 AI 이니셔티브' 출범

▶ 아파치 소프트웨어 재단은 2026년 4월 '책임 있는 AI 이니셔티브'를 출범했다. 본 이니셔티브는 재단 산하 오픈소스 소프트웨어 가운데 AI 시스템에 활용되는 인프라 기술을 대상으로, 보안·투명성·거버넌스 체계를 강화하는 것을 목적으로 한다. 이는 AI 산업 전반이 동일한 오픈소스를 공통 기반으로 활용하는 구조가 심화되면서, 특정 오픈소스에서 보안이나 안정성 문제가 발생할 경우 관련 AI 서비스 전반으로 영향이 확산될 수 있다는 우려가 커진 데 따른 것이다. 특히 이러한 위험을 개별 기업이 아닌 중립적 기관이 체계적으로 관리할 필요성이 산업 차원에서 제기되었다는 점이 반영된 조치로 볼 수 있다. 아파치 소프트웨어 재단에 따르면, 이니셔티브의 주요 추진 과제로는 AI 모델 및 도구에 대한 접근성 제공, 오픈소스 프로젝트 개발 생태계 지원, 글로벌 커뮤니티 참여 확대 등이 포함된다.

저작권 이슈 브리프

SUMMARY | 산업/기업 | 기술

산업 라이브 스포츠 불법 스트리밍의 실시간 동적 차단 체계

▶ 스페인에서는 사전 확정된 도메인·IP 목록을 기반으로 ISP가 접근을 차단하는 방식이 운영되어 왔으나, 불법 운영자의 주소 변경을 통한 우회로 라이브 방송 중 보호 실효성이 낮다는 한계가 지적되어 왔다. 이에 2026년 3월 바르셀로나 법원은 유료방송 플랫폼 모비스타르 플러스가 제공한 차단 목록에 대해 스페인 내 모든 ISP가 30분 이내 차단을 집행하도록 의무화하며, 라이브 방송 중 실시간 집행 체계를 도입하였다. 해당 체계는 챔피언스리그 8강전을 시작으로 테니스·골프 등 독점 중계권 종목으로 확대되었으나, 명령 유효 기간이 한정되어 있어 상시 운영 체계로의 전환 여부가 향후 과제로 남아 있다. 이번 사례는 라이브 방송권 보호 구조가 사후 구제 중심에서 실시간 의무 집행 체계로 이동하고 있음을 보여주며, 해외 인프라 기반 우회에 대한 대응이 다음 과제로 제기된다.

산업 튠드 글로벌, 스트리밍 플랫폼 내 음원 재생 수 조작을 탐지하는 시스템 출시

▶ 스트리밍 시장에서 봇·클릭팜을 활용한 자동 반복 재생 행위가 이어지면서 재생 횟수, 차트 순위, 권리자 저작권료가 함께 왜곡되고 있다. 이에 음반사·퍼블리셔 등 권리자는 라이선싱 협상에서 스트리밍 플랫폼이 조작 행위를 자체적으로 탐지·차단할 수 있는 체계를 명시적으로 요구하기 시작했다. 이러한 요구에 대응해 음악 클라우드 플랫폼 기업 튠드 글로벌은 2026년 4월 서비스 조작 탐지 시스템을 공개했다. 이 시스템은 별도의 외부 도구가 아니라 튠드 글로벌의 음악 클라우드 플랫폼에 직접 내장된 형태로 작동하며, 트랙, 아티스트, 이용자, 네트워크, 결제의 5개 계층에서 규칙 점검과 통계 분석을 함께 적용해 이상 징후를 자동으로 분류한다. 기준값을 넘는 재생 데이터는 저작권료 정산과 차트 집계에서 자동으로 제외되고, 처리 이력은 감사 추적 형태로 기록되며 참여 권리자에게는 매월 정산 제외 요약과 조작 동향 보고가 제공된다.

기술 주간 기술 동향

▶ 최근, 생성형 확산 모델을 이용하면 워터마크 신호를 화질 손상 없이 제거할 수 있다는 사실이 입증되면서, 기존 비가시 워터마크 기술의 한계가 드러났다. 기존 딥러닝 기반 공격은 화질 손상을 줄이는 데 성공했으나 공간 도메인의 픽셀 단위 학습에 의존해 주파수 영역에 분산된 워터마크 신호를 정밀하게 제거하는 데 한계가 있었다. 이를 해결하기 위해 제안된 FMDiffWA는 고속 푸리에 변환으로 이미지를 진폭과 위상 성분으로 분해한 뒤, 워터마크가 집중된 주파수 영역을 선택적으로 억제하는 FWM 모듈을 확산 모델의 샘플링 단계에 통합한 공격 프레임워크이다. FMDiffWA는 4가지 워터마킹 방식에 걸쳐 공격 후 PSNR을 40dB 이상으로 유지했으며, 기존 공격 대비 평균 10dB 이상 높은 화질 보존 성능을 기록했다. 이 기술은 기존 워터마킹 방어 기술의 한계를 드러냈으며, 공격-방어 연구의 균형 있는 발전에 기여할 것으로 기대된다.

게임 DRM 설계 방식과 이용자 영향 논쟁

저작권 보호 기술의 정당성 논쟁과 균형 문제

• 보호 기술과 성능 저하 논쟁의 한계

- 디지털 저작권 관리(Digital Rights Management, 이하 DRM)*는 콘텐츠의 무단 복제와 유통을 방지하기 위한 핵심 보호 수단으로 활용되어 왔으나, 성능 저하와 접근 제한 등 정당 이용자 부담을 초래한다는 비판이 지속되며 보호 기술의 정당성 논쟁이 이어지고 있음
- 일부 DRM 시스템은 사전에 정의된 규칙에 따라 접근을 제한하는 구조로 작동하여, 정당 이용자의 접근까지 차단하는 과잉 차단 문제가 제기됨
- 2026년 4월 일부 플레이스테이션(PlayStation)** 이용자들 사이에서 디지털 게임에 30일 단위의 라이선스 만료 타이머가 표시된 사례가 보고되며, 오프라인 환경에서의 접근 제한 가능성에 대한 논란이 제기됨[1)]

* 디지털 저작권 관리(Digital Rights Management): 디지털 콘텐츠의 무단 복제·배포를 막기 위해 접근과 사용을 통제하는 기술과 체계로, 게임·음악·영상 등 콘텐츠 산업 전반에서 활용됨

** 플레이스테이션(PlayStation): 소니(Sony)가 개발·운영하는 게임 콘솔 및 디지털 게임 플랫폼 브랜드로, PS4·PS5 등의 콘솔 기기와 온라인 게임 유통 서비스인 플레이스테이션 스토어(PlayStation Store) 등을 포함함

• 통합 방식 차이에 따른 이용자 영향 재인식

- 게임 산업에서는 안티탬퍼(Anti-Tamper)* 보호 코드가 추가되면 성능 저하가 발생한다는 인식이 있었으나, 최근에는 이용자가 실제로 체감하는 성능 저하 수준이 보호 기능의 실행 위치와 시점에 의해 결정된다는 분석이 제시됨
- 게임용 불법 복제 방지 기술인 데누보(Denuvo)를 개발한 네덜란드의 보안업체 이르데토(Irdeto)는, 보호 기능이 게임 실행 환경에서 다른 시스템과 함께 작동하는 런타임(runtime)** 구성 요소라고 설명함
- 실제 게임 분석 사례에서도 보호 기능 실행은 일반 게임 플레이보다 로딩·장면 전환 등 특정 구간에 집중되었으며, 보호 기능의 배치와 실행 구조가 이용자 영향을 좌우하는 핵심 요소로 분석됨

* 안티탬퍼(Anti-Tamper): 게임 실행 파일의 무단 변조와 역공학을 방지하기 위해 플랫폼 DRM을 보강하는 보호 기술로, 출시 직후 불법 복제 차단을 주된 목적으로 함

** 런타임(runtime): 프로그램이 실제로 실행되는 동안 작동하는 환경을 의미하며, 게임에서는 게임플레이·오디오·플랫폼 API 등 여러 시스템이 함께 동작하는 영역에 해당함

1) Hassam Nasir, "Sony rolls out 30-day online DRM check-in for PlayStation digital games — players could temporarily lose access if they don't keep their consoles online", tom's HARDWARE, 2026.04.29., https://www.tomshardware.com/video-games/playstation/sony-rolls-out-30-day-online-drm-check-in-for-playstation-digital-games-players-could-temporarily-lose-access-if-they-dont-keep-their-consoles-online

보호 효과와 정당 이용자 경험의 균형 메커니즘

• 보호 기능의 실행 구조와 이용자 영향

- 보호 기능은 게임 플레이, 오디오, 플랫폼 인터페이스 등 다양한 시스템과 함께 실행되는 런타임 구성 요소로, 이용자 체감 영향은 보호 기능을 수행하는 함수가 어느 시점과 실행 경로에서 수행되는지에 따라 달라지는 구조임
- 이르데토의 분석에 따르면 보호 관련 기능은 전체 실행 코드에서 차지하는 비중이 제한적이며, 실제 성능 영향은 기능의 존재 여부보다 실행 위치와 빈도에 의해 결정되는 것으로 분석됨
- 동일한 보호 기능이라도 시작 화면, 장면 전환 등 저부담 구간에서 실행되면 이용자 부담이 작은 반면, 실시간 반응성이 중요한 게임 플레이 구간에서 반복적으로 수행될 경우 성능 저하로 인식될 가능성이 높음
- 실제 보호 기술 분석에서도 보호 기능은 게임 플레이 도중 반복적으로 실행되지 않도록 설계되며, 이러한 통합 방식이 이용자 체감 영향을 줄이는 핵심 요소로 평가됨

• 이용자 부담 완화를 위한 통합 설계 원칙

- 이르데토의 설명에 따르면 이용자가 체감하는 성능 영향은 보호 코드 자체보다 통합 방식과 실행 경로에 의해 결정됨
- 이르데토는 2026년 4월 통합 설계의 원칙을 공개하며, 보호 기능 적용 시 함수의 실행 빈도, 수행 시간, 실행 경로 등을 우선 분석하여 성능에 영향을 미칠 가능성이 낮은 위치로 배치를 결정하는 방식을 제시함[2)]
- 이는 저부담 실행 환경 중심 설계를 통해 시작, 로고화면, 메뉴 전환 등 이미 백그라운드 작업이 수행되는 구간에 보호 관련 실행을 집중시켜 이용자가 부담을 체감할 가능성을 낮추는 접근임
- 실제 게임에서도 보호 관련 실행의 약 75~80%가 시작 단계에 의도적으로 배치되며, 실시간 반응이 중요한 플레이 경로에는 가급적 회피하는 방식으로 이용자 부담을 분산하는 사례가 보고됨
- 또한 보호 기능은 업데이트와 빌드 환경 변화에 따라 실행 구조가 달라질 수 있는 만큼, 주요 변경 이후에도 실행 환경 기반의 재검증과 성능 점검이 필요하다는 운영 원칙을 제시함

[표1] 이르데토가 제시한 통합 설계 원칙의 단계별 구조

원칙	핵심 내용	적용 시점	기대 효과
런타임 행동 분석 기반 배치	함수의 실행 빈도·소요 시간·스레드 평가 후 배치 결정	통합 설계 단계	위험 함수의 사전 식별
저부담 실행 환경 중심 설계	시작·로딩·메뉴 등 백그라운드 구간에 보호 실행 집중	출시 시점	이용자 체감 부담 최소화
업데이트 후 재검증	패치·빌드 변경 이후 실행 구조 변화에 대한 반복 검증	운영 단계	균형 상태의 지속 유지

출처: 참고문헌 종합하여 재구성

2) Andreas Ullmann, “Performance-aware anti-piracy integration: Three principles for placement and execution”, Irdeto, 2026.04.07., https://irdeto.com/blog/anti-piracy-integration-performance

시사점: 콘텐츠 산업 전반의 보호 기술 거버넌스

• 보호 기술 운영 방식의 설계 중심 전환

- 본 사례는 저작권 보호 기술의 효과와 이용자 체감 영향이 보호 코드의 강도보다 통합 방식과 운영 설계로 확보된다는 점을 보여줌
- 이에 따라 보호 기술은 단순 차단 기능이 아니라, 정당 이용자 경험을 위한 실행 위치·시점·경로를 고려하는 통합 운영 체계 정립이 필요함을 시사함
- 특히 게임 산업에서 보호 기술의 적용 시점이 출시 시 적용, 출시 전 철회 및 출시 후 제거 등 다양한 운영 패턴이 보고되어, 보호 기술이 일률적 적용이 아닌 운영 시점·통합 환경에 따라 조정되는 도구로 다뤄지는 흐름을 반영함

• 지속 검증 기반 운영 체계와 정책 정합성 확보

- 보호 기능은 업데이트와 빌드 환경 변화에 따라 실행 구조와 이용자 영향이 달라질 수 있어, 초기 적용 이후에도 지속적인 검증과 재평가가 요구됨
- 특히 보호 기술이 사전에 정의된 규칙에 따라 작동하는 구조에서는 정당 이용자의 접근까지 제한하는 과잉 차단 문제가 발생할 수 있어, 보호 기술 운영 과정에서 이용자 권리와의 균형을 함께 고려할 필요가 있음
- 이에 따라 향후 콘텐츠 산업에서는 보호 기술의 강도뿐 아니라 적용 방식과 운영 구조, 저작권법상 예외 규정과의 정합성까지 함께 검토하는 방향으로 정책 논의가 확대될 가능성이 있음

참고문헌

- Andreas Ullmann, “Performance-aware anti-piracy integration: Three principles for placement and execution”, Irdeto, 2026.04.07., https://irdeto.com/blog/anti-piracy-integration-performance
- Hassam Nasir, “Sony rolls out 30-day online DRM check-in for PlayStation digital games — players could temporarily lose access if they don't keep their consoles online”, tom’s HARDWARE, 2026.04.29., https://www.tomshardware.com/video-games/playstation/sony-rolls-out-30-day-online-drm-check-in-for-playstation-digital-games-players-could-temporarily-lose-access-if-they-dont-keep-their-consoles-online
- Hassan Javed, “Denuvo in 2025: why some games ship with the DRM — and others remove it”, Tech Bullion, 2025.10.17., https://techbullion.com/denuvo-in-2025-why-some-games-ship-with-the-drm-and-others-remove-it/
- Amber Rutherford, “The Denuvo System In Hogwarts Legacy Explained By A DRM Developer”, 80LV, 2024.04.02., https://80.lv/articles/the-denuvo-system-in-hogwarts-legacy-explained-by-a-drm-developer
- Patrick Spencer, “Navigating the Complex Challenges of Digital Rights Management: Balancing Protection and User Rights”, Kiteworks, 2025.09.18., https://www.kiteworks.com/digital-rights-management/drm-challenges-20250918/

스포티파이의 AI 크레딧 도입으로 본 음악 스트리밍 플랫폼의 AI 자율 공시 모델

스포티파이의 AI 투명성 정책 흐름과 AI 크레딧 도입

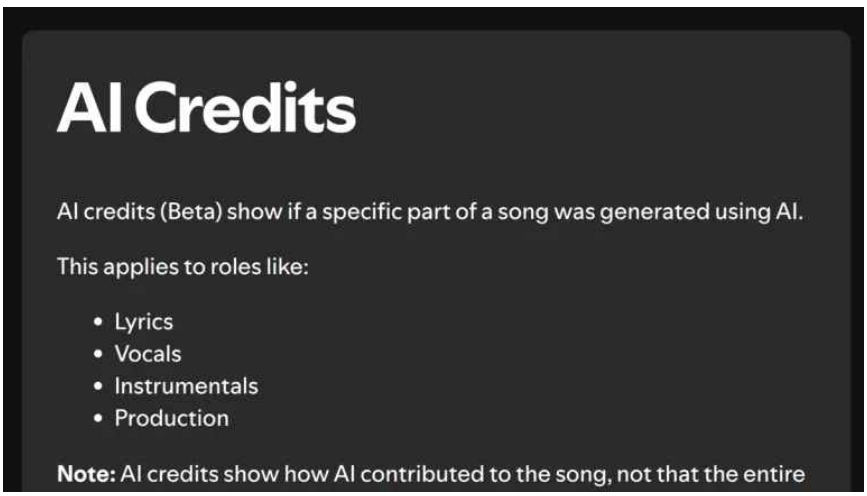
• 스포티파이의 AI 자율 공시 기능 'AI 크레딧' 베타 공개

- 스포티파이(Spotify)는 2026년 4월 음원 제작 과정의 AI 활용 여부를 곡 크레딧(song credits)* 영역에 명시하는 자율 공시 기능 'AI 크레딧(AI Credits)'의 베타 버전을 공개함
- 이번 베타 공개는 단발적 결정이 아니라 누적된 정책 흐름의 연장선으로, 스포티파이는 2025년 9월 블로그 게시물을 통해 AI 관련 보호 장치 강화 방향과 자율 공시 도입 계획을 처음 제시한 바 있음
- 이후 스포티파이는 2025년 10월 소니뮤직그룹(Sony Music Group), 유니버설뮤직그룹(Universal Music Group), 워너뮤직그룹(Warner Music Group) 등 3대 메이저 음반사 및 주요 인디 음악 진영과 협력 체계를 구축하고, AI 음악 산출물의 책임 있는 개발을 추진 중임
- 이와 병행하여 스포티파이는 통일된 AI 공시 형식을 음악 산업 전반에 정착시키기 위해 DDEX (Digital Data Exchange)** 중심의 표준화 작업도 추진해 왔으며, 이번 AI 크레딧 베타 공개 역시 이러한 흐름의 연장선상에 위치함

* 곡 크레딧(song credit): 음악 제작에 참여한 작사, 작곡, 편곡, 연주, 프로듀싱 등 실제 참여자들의 명단을 앨범이나 음원 정보에 기록한 정보

** DDEX(Digital Data Exchange): 음악 산업의 디지털 데이터 표준화를 추진하는 글로벌 비영리 협의체로, 음원 메타데이터 교환과 권리 정보 처리에 관한 산업 공통 표준을 개발하고 있음

[그림1] 스포티파이의 'AI Credits' 베타 버전 안내 페이지



출처: Ashley King, "Spotify Quietly Starts Identifying Songs Using AI (Or At Least Some of Them)", Digital Music News, 2026.04.22., https://www.digitalmusicnews.com/2026/04/22/spotify-tests-ai-credit-labels-distrokid/

AI 크레딧의 표시 구조와 자율 공시 운영 방식

• 곡 제작 영역별 AI 기여 표시 방식

- AI 크레딧은 보컬, 가사, 악기, 프로듀싱 등 곡 제작 과정의 영역별 AI 기여를 분리해 표시하도록 설계된 공시 기능으로, 이용자는 곡의 어떤 영역에 AI가 활용되었는지를 영역별로 식별할 수 있음
- 해당 공시 정보는 모바일 앱의 곡 크레딧 영역에 노출되며, 이용자는 곡 제작자 정보와 함께 AI 활용 사실을 확인할 수 있음
- 또한 AI가 곡의 일부 영역에만 활용된 경우 이용자가 곡 전체를 AI 산출물로 오인하지 않도록, 해당 표시는 AI 기여가 곡 전체가 아닌 특정 영역에 한정된다는 점을 함께 명시함

• 창작자·유통사 제출에 따른 자율 공시 운영 방식

- AI 크레딧은 창작자의 자율 공시에 전적으로 의존하는 구조이기 때문에, 스포티파이도 'AI 공시가 표시되지 않았다고 해서 AI가 사용되지 않은 것은 아니다'라는 점을 안내 페이지에 직접 명시하고 있음
- 또한 AI 크레딧 정보는 창작자가 스포티파이에 직접 등록하는 방식이 아니라 음반사·유통사를 통해 제출되는 구조로 운영됨. 다만 모든 유통사가 AI 활용 정보 입력 기능을 제공하는 것은 아니어서, 유통 경로에 따라 AI 공시 가능 여부가 달라지는 공백이 발생함
- 이에 따라 현재 AI 크레딧은 관련 정보 입력 기능이 우선 마련된 유통사인 디스트로키드(DistroKid)를 통한 유통 음원부터 적용되고 있으며, 스포티파이는 향후 적용 가능한 유통 경로를 점차 확대할 계획임
- 스포티파이는 이번 기능이 그 자체로 완전한 해결책은 아니며, 포괄적인 공시 체계 구축을 위해서는 산업 전반의 정합성 확보가 필요하다고 밝힘

[표1] 주요 음악 스트리밍 플랫폼의 AI 공시·탐지 체계 비교

플랫폼	AI 공시·탐지 체계	도입 시점	방식	의무화 여부
스포티파이	AI Credits	2026년 4월 (베타)	자율 공시	선택적
애플뮤직(Apple Music)	Transparency Tags	2026년 3월	자율 공시	향후 의무화 예정
디저(Deezer)	자체 AI 탐지 도구	자체 운영	자동 탐지	해당 없음(자동 적용)

출처: Mandy Dalugdug, "Spotify to show AI tags in Song Credits, where artists have chosen to disclose through their label or distributor", Music Business Worldwide, 2026.04.23., https://www.musicbusinessworldwide.com/spotify-to-show-ai-tags-in-song-credits-where-artists-have-chosen-to-disclose-through-their-label-or-distributor/

자율 공시 모델의 효과성 한계와 검증 체계의 보완 논의

• 선택적 공시에 따른 창작자의 자발적 공시 유인의 약화

- 자율 공시 모델의 효과성에 대해서는 학계를 중심으로 비판적 견해가 제기되고 있음
- 코넬대학교(Cornell University)의 고든 펜쿡(Gordon Pennycook) 교수는 AI 활용을 정직하게 공시한 창작자가 불이익을 받고 AI 활용 사실을 숨긴 창작자가 책임을 회피하게 될 경우, 미공시가 더 유리한 선택이 되어 정직한 공시자가 불이익을 보는 역설이 발생할 수 있다고 지적함[1)]

1) Gordon Pennycook, Trystan Sterling Goetze, "Cornell experts on AI labels for Spotify", Cornell University, 2026.04.23., https://news.cornell.edu/media-relations/tip-sheets/cornell-experts-ai-labels-spotify

- 나아가, 현시점에서 다수의 창작자가 미공시에 따른 적발 위험보다 AI 활용 사실이 알려질 경우의 위험을 더 크게 인식하고 있을 수 있다고 진단하며, 창작자의 자발적 공시 유인이 부족한 상황임을 시사함
- 이러한 한계는 의무화 부재 측면에서도 지적되는데, 코넬대학교 트리스탄 괴체(Trystan Goetze) 교수는 AI 표시 의무화가 콘텐츠 신뢰성 향상에 기여할 수 있으나 스포티파이와 같이 표시가 선택 사항으로 운영될 경우 그 효과가 제한될 수 있다고 평가함

• 표시 정확성 검증과 산업 공통 표준 마련 필요성

- 이러한 한계를 보완할 방안으로 검증 메커니즘 도입이 거론되고 있음
- 괴체 교수는 표시의 정확성을 검증할 수 있는 장치가 필요하다고 지적하며, C2PA(Coalition for Content Provenance and Authenticity)*와 같은 기술 표준을 보완 수단으로 거론함
- 사업자 차원에서도 보완 시도가 이루어지고 있으며, 스포티파이는 곡 크레딧 정보 추가와 산업 표준 개발을 병행하면서 자율 공시 모델의 한계를 보완하고자 함
- 다만 자율 공시 모델이 음악 산업 전반의 표준으로 정착하기 위해서는 검증 메커니즘 도입과 통일된 공시 형식 마련이 단일 플랫폼을 넘어 음악계 공동의 과제로 병행될 필요가 있음

* C2PA(Coalition for Content Provenance and Authenticity): 디지털 콘텐츠의 출처와 편집 이력을 콘텐츠 자체에 첨부해 검증할 수 있도록 하는 글로벌 기술 표준으로, 어도비·마이크로소프트 등 주요 기술 기업이 참여하는 협의체에서 개발하고 있음

참고문헌

- Mandy Dalugdug, "Spotify to show AI tags in Song Credits, where artists have chosen to disclose through their label or distributor", Music Business Worldwide, 2026.04.23., https://www.musicbusinessworldwide.com/spotify-to-show-ai-tags-in-song-credits-where-artists-have-chosen-to-disclose-through-their-label-or-distributor/
- Ashley King, "Spotify Quietly Starts Identifying Songs Using AI (Or At Least Some of Them)", Digital Music News, 2026.04.22., https://www.digitalmusicnews.com/2026/04/22/spotify-tests-ai-credit-labels-distrokid/
- Gordon Pennycook, Trystan Sterling Goetze, "Cornell experts on AI labels for Spotify", Cornell University, 2026.04.23., https://news.cornell.edu/media-relations/tip-sheets/cornell-experts-ai-labels-spotify

오픈소스 소프트웨어의 안정성과 거버넌스 확보를 위한 '책임 있는 AI 이니셔티브' 출범

'책임 있는 AI 이니셔티브' 출범과 오픈소스 관리 강화 배경

• 아파치 소프트웨어 재단, '책임 있는 AI 이니셔티브' 출범

- 2026년 4월, 아파치 소프트웨어 재단(Apache Software Foundation)*에 의해 '책임 있는 AI 이니셔티브(Responsible AI Initiative)'가 출범함
- 이 이니셔티브는 재단 산하의 오픈소스 소프트웨어(Open Source Software)** 중 AI 시스템에 활용되는 기술들의 보안, 투명성 및 거버넌스 확보를 목적으로 함

* 아파치 소프트웨어 재단(Apache Software Foundation): 오픈소스 소프트웨어(OSS) 프로젝트를 지원·관리하는 비영리 재단으로, 데이터 처리·분산 컴퓨팅·클라우드 등 디지털 서비스의 기반이 되는 핵심 소프트웨어의 개발을 지원하며, 보안 관리·취약점 대응·거버넌스 운영 등을 수행함

** 오픈소스 소프트웨어(Open Source Software, OSS): 소스 코드가 공개되어 누구나 자유롭게 사용·수정·배포할 수 있는 소프트웨어를 의미하며, 기업과 개발자는 이를 활용해 개발 시간과 비용을 절감하고 필요에 따라 기능을 직접 개선·확장할 수 있어 널리 활용됨

• AI 성능 경쟁 뒤의 숨은 인프라, 오픈소스 소프트웨어

- 사람들이 AI 서비스를 사용할 때 주목하는 것은 얼마나 복잡한 사고를 수행하는지, 얼마나 정확한 정보를 제공하는지 등과 같은 모델의 인지·추론 및 응답 생성 성능임
- 그러나 실제 AI 서비스가 작동하려면 방대한 데이터를 저장·정리하고, 수많은 서버를 연결하며, 이용자의 요청을 지연 없이 처리하는 데이터 처리·분산 컴퓨팅·실시간 메시징 등의 기반 기술이 함께 갖춰져야 함
- 이러한 기반 기술은 많은 기업에 공통적으로 필요하지만, 각자 처음부터 개발하면 비용과 시간이 상당히 소요되어 성능과 안정성이 검증된 오픈소스 소프트웨어를 활용하는 방식이 자리 잡아 왔음
- 이 가운데 아파치 소프트웨어 재단은 데이터 처리·분산 시스템 등 핵심 기반 기술을 다수의 오픈소스 프로젝트 형태로 관리·운영하며, AI 생태계 전반에서 폭넓게 활용되는 인프라를 제공해 옴

• 이니셔티브 출범 배경, 공용 인프라에 대한 관리 중요성 부각

- 다수의 기업이 동일한 오픈소스 소프트웨어에 의존하는 구조인 만큼, 해당 오픈소스에 보안이나 안정성에 문제가 발생하면 이를 활용하는 AI 서비스 전반이 영향 받을 수 있다는 문제가 제기됨
- 이에 따라 AI 산업 전반에서 공용 인프라로 활용되는 오픈소스 소프트웨어에 대한 위험 관리가 특정 기업의 문제가 아닌, 산업 전체가 함께 대응해야 할 과제로 부각됨
- 아파치 소프트웨어 재단은 이러한 필요성에 대응하기 위해, 재단 차원에서 관리해오던 오픈소스 소프트웨어를 보다 안전하고 투명하며 공공의 이익에 부합하는 방식으로 운영하고, 산업 차원에서 체계적으로 지원·관리하기 위해 '책임 있는 AI 이니셔티브'를 출범함

초기 참여 현황과 이니셔티브 운영 및 지원 방식

• 앤트로픽·알파-오메가 참여로 175만 달러 규모로 출범

- '책임 있는 AI 이니셔티브'는 2026년 4월, 앤트로픽(Anthropic)의 150만 달러(약 22억 원)와 알파-오메가(Alpha-Omega)*의 25만 달러(약 3억 6천만 원[1]), 총 175만 달러(약 25억 원)의 후원금으로 출범함
- 이니셔티브는 최소 3년간 운영될 예정이며, 아파치 소프트웨어 재단은 전체 모금액 목표를 1,000만 달러로 제시함
- 이니셔티브의 참여 대상은 기술 기업, AI 모델 공급사 등이며, 연간 최소 25만 달러의 후원을 약정하는 방식으로 참여 가능함
- 현금 후원 외에도 AI 모델, 개발·운영 플랫폼, 소프트웨어 도구 등의 이용 권한을 제공하는 형태의 현물 지원 역시 인정됨
- 재단은 특정 후원사가 이니셔티브 운영 방향에 영향력을 행사하지 못하도록 특정 기업에 종속되지 않는 원칙을 유지하며, 커뮤니티 중심의 거버넌스 방식으로 운영할 방침임
- 재단은 이니셔티브를 통해 오픈소스 소프트웨어의 보안·안정성 강화와 개발 생태계 지원을 추진할 예정이며, 공개된 구체적인 지원 내용은 아래 표와 같음

* 알파-오메가(Alpha-Omega): 오픈소스 소프트웨어 생태계의 보안 강화를 위해 자금 지원과 협력 사업을 수행하는 비영리 조직임

[표1] 책임 있는 AI 이니셔티브의 핵심 지원 영역 및 내용

구분	작동 구조
AI 모델 및 도구 제공 (Access to AI models and tooling)	- 아파치 재단 산하 오픈소스 소프트웨어에 AI 언어·코드 모델과 개발 도구에 대한 접근을 제공 - 보안 관리 체계(ASF Security)와 개발·배포 관리 체계(Tooling Initiative, Apache Trusted Release 등) 전반에서 AI를 적용
프로젝트의 개발 생태계 지원 (Project-level ecosystem support)	- AI 중심 아파치 오픈소스 소프트웨어가 실제 서비스에 적용 가능한 수준(production-ready)까지 개발을 가속화할 수 있도록 지원 - 데이터 처리·저장·머신러닝·검색 등 AI/ML 전반 기술 영역을 포함한 개발 생태계를 강화
커뮤니티 참여 및 글로벌 협업 지원 (Community engagement and global participation)	- 해커톤, 밋업, 컨퍼런스, 전용 트랙 운영 등을 통해 학습·협업·기여 기회 확대 - 장학금 및 행사 참여를 위한 이동 경비 지원 등을 통해 다양한 참여자의 접근성 확대

출처: 참고문헌 종합하여 재구성

참고문헌

- THE ASF, "The Apache Software Foundation Launches $10M Responsible AI Initiative with Initial $1.75M Donation", 2026.04.08., https://news.apache.org/foundation/entry/the-apache-software-foundation-launches-10m-responsible-ai-initiative-with-initial-1-75m-donation
- THE ASF, "Responsible AI Initiative", 접속 기준 2026.05.05., https://www.apache.org/foundation/initiatives/responsibleai

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

라이브 스포츠 불법 스트리밍의 실시간 동적 차단 체계

라이브 콘텐츠 불법 유통과 기존 차단 방식의 한계

• 라이브 스포츠 콘텐츠의 불법 유통 규모

- 라이브 스포츠 콘텐츠는 실시간 시청 가치가 높다는 특성상 불법 유통의 표적이 되기 쉬우며, 스페인 프로축구 1부 리그인 라리가(LaLiga)는 이로 인한 연간 피해 규모를 7억 유로(원화 약 1조 206억 원[1]) 이상으로 추산함[2]

- 스페인 경찰이 주도한 국제 수사에서는 약 100만 명의 유료 이용자를 확보한 불법 IPTV(Internet Protocol Television)* 네트워크가 적발되었으며, 약 1,000개 웹사이트를 통해 1,500만 유로(원화 약 258억 원) 이상의 수익을 창출한 것으로 확인됨[3]

- 불법 스트리밍은 IPTV, 카드쉐어링(Card Sharing)**, P2P(Peer-to-Peer)*** 기술, 메시징 플랫폼 등 여러 배포 경로를 동시에 활용하는 방식으로 운영되어, 특정 주소나 단일 경로 차단만으로는 유통을 억제하기 어려운 구조였음

* IPTV(Internet Protocol Television): 인터넷 프로토콜 기반으로 방송 콘텐츠를 전송하는 기술로, 불법 IPTV는 정식 방송권 없이 유료 구독 형태로 콘텐츠를 제공하는 서비스를 말함

** 카드쉐어링(Card Sharing): 유료방송 수신카드 1장의 복호화 신호를 여러 기기가 네트워크로 공유하는 방식으로, CCCam·IKS 등의 프로토콜이 불법 유통에 활용됨

*** P2P(Peer-to-Peer): 중앙 서버 없이 이용자 간 직접 데이터를 전송하는 방식으로, 불법 콘텐츠 배포에 활용될 경우 차단이 어려운 특성이 있음

• 기존 차단 방식의 구조적 한계

- 기존 차단 방식은 사전에 확보한 도메인·IP 목록을 기반으로 ISP(Internet Service Provider)*가 접근을 차단하는 구조로 운영됐으나, 불법 운영자가 주소를 지속적으로 변경하는 방식으로 쉽게 우회할 수 있다는 한계가 존재함

- 특히 라이브 스포츠 콘텐츠는 방송 중 실시간 시청자 수가 집중되기 때문에, 실시간 차단이 이루어지지 않으면 사후 삭제만으로 보호 실효성을 확보하기 어려운 구조임

- 스페인에서는 2020년 마드리드 법원이 초기 동적 차단(Dynamic Blocking)**을 도입한 데 이어 2025년 8월에도 유사한 명령이 내려졌으나, 차단의 집행 주체가 일부 ISP에 한정되어 전국 단위의 실시간 대응 체계로 확대되지는 못함

* ISP(Internet Service Provider): 인터넷 접속 서비스를 제공하는 사업자로, 스페인 내 대형 통신사부터 소규모 지역 사업자까지 모두 포함됨

** 동적 차단(Dynamic Blocking): 사전에 확정된 목록이 아니라 새롭게 탐지된 불법 주소를 실시간으로 추가해 차단 범위를 갱신하는 방식

1) 1유로=1,741.97원(2026.05.04, KEB 하나은행 매매기준율 적용, 이하 동일)
2) David Del Valle, “Spain launches real-time blocking of illegal sports streams”, Advanced Television, 2026.04.14., https://www.advanced-television.com/2026/04/14/spain-launches-real-time-blocking-of-illegal-sports-streams/
3) Julian Clover, “Spanish court strikes record blow to major illegal IPTV network”, Broadband TV News, 2026.04.22., https://www.broadbandtvnews.com/2026/04/22/spanish-court-strikes-record-blow-to-major-illegal-iptv-network/

실시간 동적 차단 체계와 집행 주체

• 집행 의무와 운영 프로토콜

- 2026년 3월 바르셀로나 법원은 현지 유료방송 플랫폼인 모비스타르 플러스(Movistar Plus+)가 불법 스트리밍 도메인·IP 목록을 ISP에 전달하면, ISP가 30분 이내에 해당 주소를 차단하도록 의무화하는 명령을 내림
- 법원이 제시한 운영 방식은 ① 초기 확인된 도메인·IP를 즉시 차단, ② 미러 사이트(Mirror Site)*와 도메인 변형을 주 단위로 모니터링, ③ 신규 탐지 목록 수신 후 30분 내 차단 집행하는 구조로 설계됨
- 이는 라리가 등 저작권자 측이 불법 IP 목록을 수집하여 ISP에게 차단 요청을 전달하는 기존 방식과 달리, 방송사업자가 직접 ISP에 목록을 전송하는 구조를 채택함으로써 중간 전달 단계를 축소하고 대응 시간을 단축하는 목적으로 설계됨

  * 미러 사이트(Mirror Site): 동일한 콘텐츠를 제공하는 복제 웹사이트로, 불법 운영자가 차단을 우회하기 위해 원본 사이트와 동일한 서비스를 다른 주소에서 운영하는 방식

• 참여 의무화에 따른 집행 주체·범위 확대

- 이번 법원 명령은 보다폰(Vodafone)·디지(Digi)·마스오렌지(MasOrange) 등 스페인 내 모든 ISP에 차단 이행 의무를 부과하며, 일부 통신사 협력에 의존하던 구조를 전국 단위 집행으로 확대한 사례로 평가됨
- 또한 불법 스트리밍 트래픽 차단이 피크 시간대 네트워크 혼잡을 줄이는 효과가 있다는 점에서, ISP가 법원 명령에 대한 이의를 제기하지 않고 모비스타르 플러스와 협력에 동의함
- 차단 체계의 적용 범위는 테니스·골프 등 독점 중계권 전 종목으로 확대되었으나 명령 유효 기간이 2026/2027 시즌까지로 한정되어 있어 이후 갱신 여부가 상시 운영 체계화의 관건으로 분석됨

• 우회 인프라의 활용과 실시간 차단의 한계

- 실시간 차단 체계가 도입되었음에도 불구하고, 불법 운영자들이 CDN* 서비스와 해외 서버 인프라를 활용해 차단을 기술적으로 우회하는 방식이 확인되며 체계 운영상 기술적 한계로 지적됨
- 특히 클라우드플레어(Cloudflare) 등 인프라 사업자의 서비스가 우회 수단으로 활용되는 사례가 확인되면서, 차단 대상이 단순 콘텐츠 유통 경로를 넘어 인프라 계층까지 확대될 가능성이 제기됨
- 더불어 이러한 우회 인프라가 13개국에 걸쳐 분산 운영되는 구조에서는 단일 국가 차원의 실시간 차단만으로는 국경 밖 인프라를 통제하기 어렵다는 기술적 한계가 있음

  * CDN(Content Delivery Network): 콘텐츠를 빠르게 전송하기 위해 전 세계 여러 서버에 분산 저장하는 네트워크 인프라로, 불법 운영자가 이를 활용하면 콘텐츠 출처를 특정하거나 차단하기 어려워짐

시사점: 실시간 차단 체계의 산업적 의의

• 실시간 집행 체계로의 전환

- 라이브 콘텐츠 보호 체계가 방송 종료 후 사후 삭제 위주의 수동적 대응 체계에서, 방송 중 실시간 차단 중심의 집행 체계로 전환되고 있음
- 특히 라이브 스포츠 콘텐츠는 방송 중 시청 가치와 수익 효과가 집중되는 특성이 있어, 차단 속도 자체가 보호 실효성을 결정하는 핵심 요소로 부상하고 있음
- 또한 방송사업자가 직접 ISP와 연계해 실시간 차단을 수행하는 구조가 도입되면서, 기존의 단계적 전달 체계보다 집행 속도를 단축하는 방향으로 보호 체계가 재편되는 흐름이 나타남

- **집행 선례로서의 의미와 확산 가능성**

- 이번 명령은 특정 사업자에 한정된 조치가 아니라, 다른 방송 사업자나 스포츠 단체도 유사한 방식의 집행 명령을 신청할 수 있는 구조라는 점에서 향후 유럽 내 확산 가능성이 있는 선례로 평가됨
- 다만 단일 국가 차원의 차단 명령만으로는 실질적 통제에 한계가 존재함에 따라, 향후 라이브 콘텐츠 보호 체계는 국내 ISP 차단뿐 아니라 해외 인프라 사업자와의 협력, 국가 간 공조 체계 구축까지 포함하는 방향으로 확대될 가능성이 있음

참고문헌

- David Del Valle, "Spain launches real-time blocking of illegal sports streams", Advanced Television, 2026.04.14., https://www.advanced-television.com/2026/04/14/spain-launches-real-time-blocking-of-illegal-sports-streams/
- Ramón Muñoz, "Movistar Plus+ podrá bloquear cualquier partido pirateado de Champions hasta 2027", El País, 2026.04.15., https://elpais.com/economia/2026-04-15/movistar-plus-podra-bloquear-cualquier-partido-pirata-de-champions-hasta-2027.html
- Julian Clover, "Spanish court strikes record blow to major illegal IPTV network", Broadband TV News, 2026.04.22., https://www.broadbandtvnews.com/2026/04/22/spanish-court-strikes-record-blow-to-major-illegal-iptv-network/

튠드 글로벌, 스트리밍 플랫폼 내 음원 재생 수 조작을 탐지하는 시스템 출시

음원 부정 재생 확산에 따라 권리자의 보호 요구 강화

• 봇·클릭팜을 활용한 음원 재생 수 조작과 저작권료 피해

- 음악 스트리밍 시장에서는 봇이나 클릭팜(click farm)*을 이용해 음원을 자동으로 반복 재생하는 조작 행위가 발생하고 있음
- 스트리밍 플랫폼의 저작권료는 전체 수익을 음원별 재생 횟수 비율에 따라 권리자에게 배분하는 구조임. 특정 음원의 재생 수가 조작으로 부풀려지면 해당 음원이 배분에서 더 큰 몫을 차지하게 되고, 그만큼 조작하지 않은 정당한 권리자의 몫이 줄어듦
- 또한 조작된 재생 수는 음원 차트 순위까지 왜곡시켜, 차트에 대한 업계와 이용자의 신뢰를 떨어뜨림
- 음악 전문 매체 뮤직 비즈니스 월드와이드(Music Business Worldwide)에 따르면, 미국에서 한 인물이 AI로 생성한 음원 수십만 곡을 봇으로 수십억 회 재생시켜 약 800만 달러(약 118억 원)의 저작권료를 부정 수취한 혐의를 인정한 사례가 있음[1)]

* 클릭팜(click farm): 다수의 인력 또는 자동화 기기를 동원해 특정 콘텐츠의 조회 수, 좋아요, 재생 횟수 등을 인위적으로 늘리는 작업장 형태의 시설

• 음반사·퍼블리셔 등 권리자, 스트리밍 플랫폼에 재생 수 조작 탐지 체계 요구

- 이에 따라 음반사·퍼블리셔 등 권리자는 스트리밍 플랫폼과의 음원 라이선싱 협상에서, 플랫폼이 재생 수 조작을 자체적으로 탐지·차단하고 조작된 재생 데이터를 저작권료 정산에서 제외할 수 있는 체계를 갖추도록 명시적으로 요구하기 시작함
- 권리자가 요구하는 내용은 크게 네 가지로 정리됨. ① 의심스러운 활동을 여러 단계로 나눠 점검하는 탐지 방식, ② 모든 사례에 동일하게 적용되는 일관된 기준값, ③ 처리 과정을 외부에서 검증할 수 있도록 기록을 남기는 감사 가능성(auditability)*, ④ 조작으로 의심되는 재생 데이터를 저작권료 정산에서 제외하는 조치 능력임
- 튠드 글로벌(Tuned Global)은 이러한 흐름에 대해, 조작 탐지 체계가 라이선싱의 필수 요건으로 자리 잡고 있다고 진단함. 과거에는 계약서에 조작 행위를 금지하는 조항을 넣는 수준이었다면, 이제는 탐지, 보고 및 실제 조치까지 결합된 체계를 갖출 것이 요구되는 단계로 이동하고 있다는 것임[2)]

* 감사 가능성(auditability): 어떤 결정이나 처리 과정의 근거와 절차를 외부에서 추적·검증할 수 있도록 기록과 절차가 문서화된 상태

1) Rafaela Fornitani, "Tuned Global launches streaming manipulation detection tool aimed at rightsholders and DSPs", Music Business Worldwide, 2026.04.22., https://www.musicbusinessworldwide.com/tuned-global-launches-streaming-manipulation-detection-tool-aimed-at-rightsholders-and-dsps/
2) Tuned Global, "Tuned Global launches streaming Service Manipulation Detection solution for clients and rights holders", 2026.04.21., https://blog.tunedglobal.com/news/tuned-global-launches-streaming-service-manipulation-detection-solution-for-clients-and-rights-holders

튠드 글로벌의 음원 재생 수 조작 탐지 시스템 구조와 작동 방식

• 튠드 글로벌, 플랫폼 내장형 조작 탐지 시스템 공개

- 이러한 권리자의 요구에 대응해, B2B 음악 스트리밍 기술 기업 튠드 글로벌은 2026년 4월 서비스 조작 탐지(Service Manipulation Detection, SMD) 시스템을 공개함
- SMD 시스템은 초기 구현이 완료된 상태이며, 해당 시스템 적용을 선택한 권리자에 한해 가동되는 옵트인(opt-in) 방식으로 적용됨
- 이 시스템은 별도의 외부 도구가 아니라, 튠드 글로벌의 플랫폼에 자체 통합되어 구동됨. 즉, 튠드 글로벌의 기술을 기반으로 스트리밍 서비스를 운영하는 기업이라면 별도의 시스템을 도입하지 않아도 조작 탐지 기능을 바로 활용할 수 있음
- SMD 시스템은 트랙, 아티스트, 이용자, 네트워크 및 결제의 5개 계층에서 동시에 작동하며, 각 계층의 지표를 활용해 이상 징후를 식별함. 계층별 주요 탐지 지표는 아래 표와 같음

[표1] 서비스 조작 탐지 시스템의 계층별 탐지 지표

계층	주요 탐지 지표
트랙	재생 횟수에 비해 청취자 수가 지나치게 적은 경우, 곡 완료율 또는 스킵률이 평소와 다른 경우, 단시간 내 재생량이 급증하는 경우, 반복 재생이 비정상적으로 지속되는 경우 등을 식별
아티스트	아티스트의 카탈로그 전체 데이터를 함께 살펴 개별 곡만 봐서는 드러나지 않는 패턴을 식별
이용자	같은 곡을 과도하게 반복해 듣는 경우, 하루 활동량이 비정상적으로 높은 경우, 시간이 지나도 청취 패턴이 거의 변하지 않는 경우 등을 식별
네트워크	의심스러운 로그인 행위, 접속 위치가 일관되지 않은 경우, 여러 계정이 같은 IP나 기기를 공유하는 경우 등을 식별
결제	결제 정보 단위에서 나타나는 이상 활동을 식별

출처: Tuned Global, "Tuned Global launches streaming Service Manipulation Detection solution for clients and rights holders", 2026.04.21., https://blog.tunedglobal.com/news/tuned-global-launches-streaming-service-manipulation-detection-solution-for-clients-and-rights-holders

• 규칙 점검과 통계 분석을 결합해 이상 재생 데이터를 자동 분류

- SMD 시스템의 이상 징후 탐지는 두 가지 방식을 결합하여 수행됨. 하나는 사전에 정한 규칙에 부합하는지를 점검하는 규칙 기반 방식이고, 다른 하나는 일반적인 재생 패턴에서 벗어나는 흐름을 찾아내는 통계 분석 방식임
- 분석은 이용자별·트랙별로 매일 이뤄지며, 사전에 정의된 기준값을 넘는 재생 데이터는 수작업 검토 없이 자동으로 의심 대상으로 분류됨
- 의심 대상으로 분류된 재생 데이터는 저작권료 정산과 차트 집계에서 자동으로 제외되고, 관련 계정에 대해서는 정지 조치가 적용됨. 이러한 처리는 튠드 글로벌 플랫폼이 권리자와 맺은 계약 조항에 근거하여 이뤄짐
- 탐지된 이상 징후는 사전에 문서화된 절차에 따라 내부 검토를 거치며, 각 처리 단계의 판단 기준이 모두 규정되어 있음
- 이상 징후에 대한 모든 처리 이력은 시점·주체·근거를 함께 남기는 감사 추적(audit trail)* 형태로 기록됨. 이를 통해 같은 유형의 사안이 일관되게 처리되고, 외부에서도 검증할 수 있는 구조가 갖춰짐
- SMD 시스템에 참여한 권리자에게는 정산 제외 데이터에 대한 월간 요약과 조작 동향에 대한 주기적 보고가 제공됨

* 감사 추적(audit trail): 시스템에서 발생한 결정·처리·변경 이력을 시점·주체·근거와 함께 기록·보존하여 사후에 추적·검증할 수 있도록 한 절차

튠드 글로벌이 제시한 탐지 기술 발전 방향과 업계 표준화 과제

• 단기: 머신러닝 도입으로 새로운 조작 패턴까지 자동 식별

- 튠드 글로벌은 단기적으로, 현재의 규칙 기반 탐지 방식에 머신러닝을 결합해 시스템 자가 학습이 가능한 지능형 체계으로 고도화할 계획이라고 밝힘[3)]
- 이를 통해 기존에 알려진 조작 패턴뿐 아니라, 과거 기준으로는 식별되지 않던 새로운 형태의 조작까지 자동으로 식별할 수 있을 것으로 전망함

• 중장기: 사후 대응에서 사전 차단 중심으로 전환

- 중장기적으로는 조작이 발생한 뒤 처리하는 사후 대응 방식에서, 조작이 발생하기 전에 이상 징후를 예측해 차단하는 예측·이상 탐지(predictive and anomaly-based detection) 방식으로 기술적 역량이 집중될 것으로 보임
- 아울러 SMD 시스템의 도입이 확대되면 여러 서비스에서 수집된 익명 데이터가 축적되고, 이를 통해 단일 서비스만으로는 식별이 어려운 조직적 조작 패턴까지 탐지할 수 있는 네트워크 효과(network effect)*가 형성될 것으로 전망함

  * 네트워크 효과(network effect): 참여자가 늘어날수록 시스템 또는 서비스의 가치와 기능이 함께 향상되는 현상

• 업계 과제: 플랫폼 간 조작 탐지 기준의 통일 필요

- 튠드 글로벌은 SMD 시스템을 자사만의 도구가 아니라, 다수의 서비스와 권리자가 함께 활용할 수 있는 표준 모델로 설계했다고 밝힘[3)]
- 현재 어떤 행위를 조작으로 볼지, 이를 어떻게 처리하고 보고할지에 대한 기준이 플랫폼마다 다르게 적용되고 있으며, 튠드 글로벌은 이를 업계 차원의 과제로 진단함. 또한 이러한 기준 차이는 플랫폼 운영뿐 아니라 권리자와의 거래 관계에도 영향을 미치고 있어, 권리자를 일관되게 보호하려면 업계 공통 기준 마련이 선행되어야 한다고 밝힘

참고문헌

- Tuned Global, "Tuned Global launches streaming Service Manipulation Detection solution for clients and rights holders", 2026.04.21., https://blog.tunedglobal.com/news/tuned-global-launches-streaming-service-manipulation-detection-solution-for-clients-and-rights-holders
- Rafaela Fornitani, "Tuned Global launches streaming manipulation detection tool aimed at rightsholders and DSPs", Music Business Worldwide, 2026.04.22., https://www.musicbusinessworldwide.com/tuned-global-launches-streaming-manipulation-detection-tool-aimed-at-rightsholders-and-dsps/
- Con Raso, "Tuned Global's service manipulation detector for streaming clients and rights holders", Techwire Asia, 2026.04.22., https://techwireasia.com/2026/04/tuned-globals-service-manipulation-detector-for-streaming-clients-and-rights-holders/

3) Tuned Global, "Tuned Global launches streaming Service Manipulation Detection solution for clients and rights holders", 2026.04.21., https://blog.tunedglobal.com/news/tuned-global-launches-streaming-service-manipulation-detection-solution-for-clients-and-rights-holders

주간 기술 동향

주파수 도메인 변조 확산 워터마크 공격 기술, FMDiffWA

• 비가시 워터마크의 한계와 주파수 도메인 공격 기술의 등장

비가시 워터마크는 이미지나 영상 파일 내부에 육안으로 식별할 수 없는 식별자를 삽입해 콘텐츠의 출처와 배포 경로를 추적하는 기술로, 저작권 보호와 불법 유출 방지를 위한 핵심 수단으로 자리 잡았다. 특정 주파수 대역에 정보를 삽입하는 초기의 수학적 변환 기반 방식에서 적대적 생성 신경망을 활용해 스스로 학습하는 신경망 기반 인코더-디코더 구조로 발전하면서 워터마크는 이미지 전체에 분산 삽입되는 형태로 진화했고, 원본과 구별하기 어려운 수준의 높은 품질을 유지한다. 넷플릭스, 디즈니 플러스, 아마존 프라임 비디오 등 주요 스트리밍 플랫폼은 사전 배포 콘텐츠의 유출 추적에 이 기술을 실제로 활용하고 있다. 그러나 생성형 AI 기술이 발전하면서, 최근 연구에서는 AI를 이용해 비가시 워터마크를 제거하는 것이 기술적으로 가능함이 드러났다.

워터마크 기술이 빠르게 발전하는 반면, 이를 무력화하는 워터마크 공격 기술 연구는 상대적으로 더디게 진행되어 왔다. 강력한 공격 기술이 있어야 방어 기술의 실제 취약점을 파악하고 더 견고한 워터마킹 시스템을 만들 수 있다는 점에서, 이러한 불균형은 분야 전체의 발전을 저해하는 요인으로 지적된다. 최근에는 워터마크 제거를 시도하는 공격 기술이 실제로 등장하면서 스튜디오, 방송사 등 콘텐츠 산업계도 이 문제를 주목하기 시작했다.

기존의 워터마크 공격 기술은 전통적 방식과 딥러닝 기반 방식으로 나뉘지만, 두 범주 모두 구조적 한계를 안고 있다. 전통적 공격은 워터마크를 훼손하는 과정에서 이미지 품질을 함께 저하시키며, 다양한 워터마킹 방식에 걸쳐 일관된 공격 성능을 내기 어렵다. 딥러닝 기반 공격은 화질 손상을 줄이는 데 일부 성공했으나, 공간 도메인의 픽셀 단위 학습에 의존해 주파수 영역에 분산된 워터마크 신호를 정밀하게 제거하는 데는 한계를 보인다.

이러한 배경에서 본 보고서에서 분석할 FMDiffWA는 주파수 도메인 정보를 확산 모델의 샘플링 과정에 결합하는 새로운 공격 접근법으로, 주파수 도메인 워터마크 변조 모듈을 순방향·역방향 샘플링 단계에 통합한 기술이다. 이 모듈은 이미지를 푸리에 변환으로 분해해 워터마크 관련 주파수 성분을 선택적으로 억제하며, 결과적으로 단계적 학습 전략을 통해 워터마크 제거 효과와 시각적 품질 보존을 달성한다. 강력한 공격 기술의 개발이 방어 기술의 취약점을 체계적으로 드러냄으로써 워터마킹 분야 전반의 견고성을 높이는 데 기여할 수 있다는 점에서 주목할 만하다.

[사례] 주파수 도메인 변조 확산 프레임워크 기반 워터마크 공격 기술, FMDiffWA

• 기존 워터마크 공격 기술의 한계

- JPEG 압축, 가우시안 노이즈, 평균 필터링 등 전통적 공격은 이미지 전체에 무차별적 왜곡을 가하는 방식으로 워터마크를 약화시키며, 공격 후 최대 신호 대 잡음비(Peak Signal-to-Noise Ratio, 이하 PSNR)*이 30dB 수준에 머물러 시각적 품질 저하가 뚜렷함

- 딥러닝 기반 공격은 합성곱 신경망(Convolutional Neural Network, 이하 CNN)**이나 생성적 적대 신경망(Generative Adversarial Network, 이하 GAN)***을 활용해 화질 손상을 줄이는 방향으로 발전했으나, 주파수 영역 워터마크 신호를 정밀 제거하는 능력이 부족함

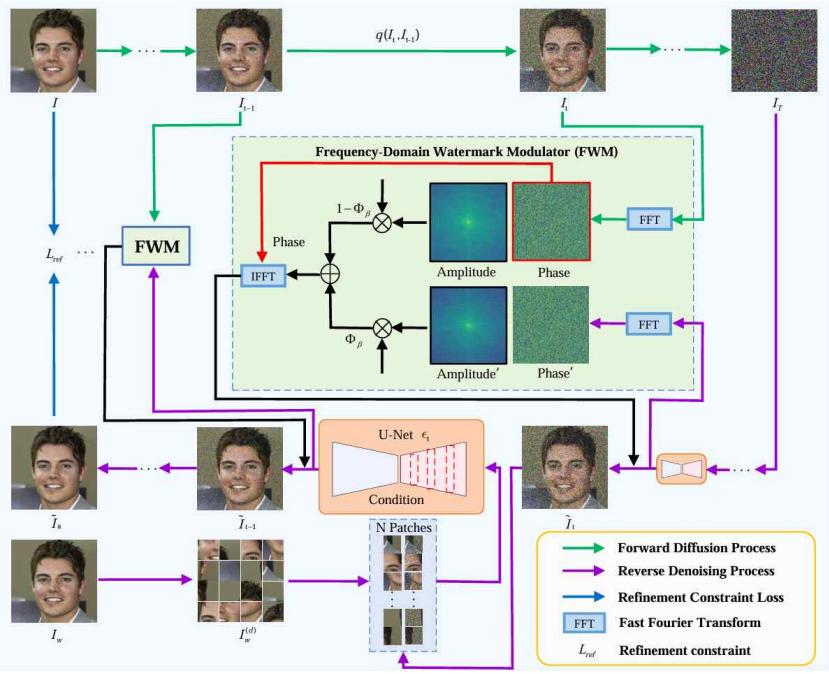
- 기존 확산 모델 기반 공격은 워터마크가 삽입된 이미지를 활용하는 조건부 생성 패러다임을 도입해 공격 품질을 높였으나, 주파수 도메인 정보를 활용하지 않아 워터마크 관련 주파수 성분의 선택적 억제와 다양한 워터마킹 방식에 대한 일반화에 한계를 보임

* 최대 신호 대 잡음비(Peak Signal-to-Noise Ratio): 원본 이미지와 처리 후 이미지 사이의 화질 차이를 dB 단위로 나타내는 지표로, 값이 높을수록 원본에 가까운 품질임을 의미함

** 합성곱 신경망(Convolutional Neural Network): 이미지의 공간적 패턴을 계층적으로 학습하는 딥러닝 모델로, 시각 인식·이미지 처리 분야에서 활용됨

*** 생성적 적대 신경망(Generative Adversarial Network): 이미지를 생성하는 생성자와 진위를 판별하는 판별자가 서로 경쟁하며 학습하는 구조로, 실제와 구별하기 어려운 고품질 이미지를 생성하는 데 활용됨

[그림 1] FMDiffWA의 주파수 도메인 변조 워터마크 공격 모듈 구조



출처: Chunpeng Wang 외 6인, "Breaking Watermarks in the Frequency Domain: A Modulated Diffusion Attack Framework", arXiv, 2026.04.24., https://arxiv.org/pdf/2604.22220

• FMDiffWA 단계적 학습 전략

- 1단계에서는 확산 모델이 각 단계에서 이미지에 섞인 노이즈를 얼마나 정확하게 걸러내는지 집중 학습시키며, 이를 통해 이미지 복원 과정의 안정성을 확보함. 총 100만 회 반복 학습을 수행함

- 2단계에서는 1단계만으로는 달성하기 어려운 시각적 품질 향상을 위해, 모델이 생성한 복원 이미지를 워터마크 없는 원본과 직접 비교하며 결과물을 다듬는 방식으로 추가 30만 회 반복 학습을 진행함

- 2단계 품질 비교 기준은 픽셀 단위 오차와 이미지 구조적 유사도를 여러 해상도에서 동시에 측정하는 방식을 결합하며, 단순 픽셀 오차만 사용하는 방식 대비 복원 정확도가 더 우수한 것으로 확인됨
- 두 단계를 결합함으로써 워터마크 제거 효과와 이미지 품질 보존이라는 두 목표를 동시에 달성하며, 주파수 도메인 워터마크 변조 모듈과의 조합으로 공격 효과와 시각적 품질 사이의 균형을 확보함

• 실험 설계 및 데이터셋 구성

- FMDiffWA의 성능을 검증하기 위한 실험은 얼굴 이미지 데이터셋인 CelebA와 범용 자연 이미지 데이터셋인 ImageNet을 대상으로 진행하며, 각 이미지는 256×256 크기로 조정함
- 사용된 이미지는 총 24,000장이었으며, 이 중 18,000장을 학습에 사용하고, 6,000장을 평가에 사용함
- 워터마킹 방식은 픽셀값을 직접 수정하는 방식, 주파수 영역을 활용하는 방식, 수학적 변환 기반 방식, 딥러닝 기반 방식 등 성격이 다른 4가지를 적용해 각 이미지에 16×16 크기의 워터마크를 삽입함
- 공격 성능은 비트 오류율(Bit Error Rate, 이하 BER)*로, 시각적 품질은 PSNR로 측정하며, 전통적 공격 5종과 딥러닝 기반 공격 3종을 비교 대상으로 삼음

* 비트 오류율(Bit Error Rate): 공격 후 추출된 워터마크 비트 중 원본과 다른 비트의 비율로, 값이 클수록 워터마크가 효과적으로 파괴된 것임

• 성능 비교 및 일반화 평가

- FMDiffWA는 CelebA와 ImageNet 두 데이터셋에서 4가지 워터마킹 방식 모두에 걸쳐 기존 방법 대비 가장 높은 PSNR을 기록하며, 공격 후 PSNR이 일관되게 40dB 이상을 유지함
- 전통적 공격과 기존 딥러닝 기반 공격이 PSNR 30~34dB 수준에 머문 것과 달리, FMDiffWA는 FWM 모듈과 단계적 학습 전략의 결합을 통해 워터마크를 제거하면서도 고품질 이미지를 복원함
- 학습되지 않은 워터마킹에 대한 일반화 실험에서도 FMDiffWA는 특정 워터마킹 알고리즘에만 맞춰진 것이 아니라, 다양한 방식에 걸쳐 안정적이고 일관된 워터마크 파괴 성능을 유지함을 확인함

결론 및 시사점

• 성능 및 기술적 성과

- FMDiffWA는 주파수 도메인 정보를 확산 모델의 생성 과정에 직접 결합함으로써, 기존 공격 기술이 해결하지 못했던 주파수 영역 워터마크의 제거와 시각적 품질 유지를 동시에 달성한 사례로 평가됨
- 특정 워터마킹 방식에 특화되지 않고 학습되지 않은 다양한 방식에도 일관된 공격 성능을 보인다는 점에서, 워터마킹 방어 기술의 취약점을 종합적으로 진단하는 평가 도구로 활용될 수 있음
- FMDiffWA는 기존 워터마킹 방어 기술의 한계를 실증적으로 드러냄으로써, 더 견고한 워터마킹 기술 개발의 필요성을 환기하고 공격-방어 연구의 균형 있는 발전에 기여할 것으로 기대됨

참고문헌

- Chunpeng Wang 외 6인, "Breaking Watermarks in the Frequency Domain: A Modulated Diffusion Attack Framework", arXiv, 2026.04.24., https://arxiv.org/pdf/2604.22220
- Bernard, “The Algorithm Arms Race: When AI Creates Watermarks and Another AI Tries to Erase Them”, Side Line, 2026.04.20., https://www.side-line.com/the-algorithm-arms-race-watermark/