Biweekly Report
통권 제86호(2026. 7-2)

www.copyright.or.kr

저작권 이슈 트렌드

COPYRIGHT ISSUE TREND

한국저작권위원회
KOREA COPYRIGHT COMMISSION

CONTENTS

저작권 이슈 트렌드

Biweekly Report | 통권 제86호(2026. 7-2)

독일 방송 실증으로 본 양자 키 분배 기술과 전송 보안의 미래

뉴스 브리프

독일의 통신사와 공영방송, 대학이 함께 실시간 방송 신호를 양자 암호로 보호하는 실증에 성공했다. 프랑크푸르트와 마인츠를 잇는 광섬유 구간에서 방송 스트림을 암호화해 전송한 사례로, 정해진 양자 상태의 광자로 양쪽에서 암호 키를 만들고 도청 시도가 있으면 즉시 감지해 키를 새로 만드는 방식이다. 핵심 쟁점은 양자컴퓨터가 지금의 암호 체계를 무력화할 수 있다는 우려다. 방송처럼 수많은 시청자에게 닿는 신호가 전송 중 가로채이거나 조작된 내용으로 바뀌면 그 피해가 커지기 때문에, 방송과 통신 분야는 미리 대비책을 찾고 있다. 이 보고서는 이번 실증에 쓰인 양자 키 분배 기술이 어떤 원리로 작동하며 콘텐츠를 지키는지 살펴보고, 이 기술이 지상 광섬유의 한계를 넘어 어느 영역까지 확장될 수 있는지 전망한다.

뉴스 플러스

Ⅰ. 서론: 방송 전송 보호와 양자 암호의 부상

- **방송 전송 보호를 위한 양자 암호 실증**

2026년 6월 독일에서 실시간 방송 신호를 양자 암호로 보호하는 실증이 이루어졌다. 독일의 통신사 보다폰(Vodafone), 독일 공영방송연합(Arbeitsgemeinschaft der öffentlich-rechtlichen Rundfunkanstalten, ARD)과 이를 주관한 남서독일방송(Südwestrundfunk, SWR), 그리고 바덴뷔르템베르크 주립협동대학(Duale Hochschule Baden-Württemberg, DHBW)이 함께 참여해, 프랑크푸르트의 ARD 시설과 마인츠의 SWR 방송센터를 광섬유로 연결하고 그 구간에서 방송 스트림을 양자 암호로 전송했다. 이번 실증에는 양자 키 분배(Quantum Key Distribution, QKD)라는 방식이 쓰였으며, 정해진 양자 상태의 광자를 전용 채널로 주고받아 양쪽에서 같은 암호 키를 만들고, 누군가 통신을 엿보려 하면 그 시도가 즉시 드러나 키가 자동으로 새로 만들어진다. 이번에 실시한 실증은 QKD가 실험실을 벗어나 현장에서 작동 가능함을 보여 준 사례로 평가된다.

이번 실증이 주목받는 이유는 방송이 지닌 특성 때문이다. 방송은 하나의 신호가 수많은 시청자에게 동시에 닿기 때문에, 전송 도중 신호가 가로채이거나 조작된 내용으로 바뀌면 그 영향이 넓게 퍼진다. 뉴스, 스포츠 중계나 정치 토론처럼 많은 사람이 지켜보는 실시간 방송에 거짓 내용이 끼어드는 상황은 방송사가 특히 경계하는 위험이다. 그동안 방송 신호는 소수 암호 방식으로 보호되어 왔으나, 컴퓨터 연산 능력이 크게 발전하면서 이러한 보호가 미래에도 충분할지에 대한 물음이 커지고 있다. 이번 실증은 그 물음에 대한 하나의 대응으로, 전송 단계에서 콘텐츠의 무결성과 기밀성을 어떻게 지킬 것인가라는 과제를 다시 짚게 한다.

- **양자컴퓨터 위협과 QKD의 부상**

QKD가 주목받는 배경에는 양자컴퓨터의 발전이 있다. 오늘날 방송을 비롯한 대부분의 통신은 큰 소수 두 개를 활용한 암호를 사용하며, 큰 수의 소인수분해가 지금의 컴퓨팅 기술로는 풀기 힘들다는 전제 위에서 보안이 유지된다. 현재 널리 사용되는 공개키 암호는 이러한 계산적 난이도를 보안의 기반으로 삼고 있다. 그러나 양자컴퓨터는 이 계산을 기존보다 훨씬 빠르게 처리할 수 있어, 이 전제를 무너뜨릴 가능성이 제기된다. 실제로 양자컴퓨팅 분야에서 사용되는 쇼어 알고리즘(Shor's algorithm)이라 불리는 방법은 큰 수의 소인수분해를 매우 빠르게 해내도록 고안되어, 현재 널리 쓰이는 암호가 미래에 무력화될 수 있음을 보여준다. 방송 신호처럼 중요한 콘텐츠일수록 이러한 위협에 미리 대비할 필요가 커진다.

QKD는 이 문제에 다른 방식으로 접근한다. 기존 암호가 소인수분해 계산의 어려움에 기대는 것과 달리, 양자역학에 보안의 근거를 둔다. 양자 상태는 측정하거나 관찰하려 하면 그 상태가 교란되고 정확한 복제도 불가능하기 때문에, 키를 주고받는 도중에 누군가 끼어들면 그 흔적이 반드시 남는다. 이 성질 덕분에 QKD는 도청 시도를 감지할 수 있으며, 양자컴퓨터의 연산 능력으로도 뚫기 어려운 보안을 제공하는 방법으로 평가된다. 앞서 살펴본 방송 실증이 바로 이 기술을 현장에 적용한 사례이며, 이어지는 본론에서는 QKD가 어떤 원리로 작동하고 어떻게 확장되는지를 차례로 살펴본다.

Ⅱ. 본론: 양자 키 분배 기술의 원리와 확장

- **양자 원리와 도청 탐지의 작동 방식**

QKD의 보안은 양자역학의 두 가지 원리가 뒷받침한다. 하나는 복제 불가 정리로, 이는 알려지지 않은 임의의 양자 상태를 똑같이 복제하는 것이 불가능하다는 원리다. 다른 하나는 하이젠베르크 불확정성 원리(Heisenberg uncertainty principle)로, 이는 양자 상태를 측정하려는 행위 자체가 그 상태를 흐트러뜨린다는 원리다. 이 두 원리가 함께 작동해, 양자 정보를 몰래 빼내거나 그대로 베끼려는

시도를 원천적으로 막는다. 정보를 광자의 양자 상태에 실어 보내는 토대가 되는 것은 중첩으로, 큐비트(qubit)*가 0과 1을 겹쳐 지닐 수 있어야 이러한 정보 전달이 가능해진다.

* 큐비트(qubit): 양자 정보를 담는 최소 단위로, 기존 컴퓨터의 비트가 0이나 1 중 하나의 값만 갖는 것과 달리 두 값을 동시에 지닐 수 있어 여러 상태를 한 번에 나타낼 수 있음

이 가운데 도청 탐지를 직접 떠받치는 것이 측정 교란의 성질이다. 광자에 실린 정보를 가로채려면 그 광자를 측정해야 하는데, 측정하는 순간 상태가 바뀌어 흔적이 남는다. 송신자와 수신자는 주고받은 신호에서 오류가 얼마나 생겼는지 확인해 이 흔적을 읽어낼 수 있으며, 그 결과로 제3자의 개입을 알아챈다. 침입이 반드시 드러난다는 점이 QKD가 기존 암호와 구별되는 지점이다.

이 원리들은 키를 만드는 과정에서 단계적으로 작동한다. 먼저 송신자가 정해진 양자 상태의 광자를 양자 채널로 보내면, 수신자가 이를 받아 측정한다. 이어 양쪽이 별도의 통신을 통해 측정에 쓴 기준을 맞춰보고, 서로 일치하는 부분만 골라 최종 키로 삼는다. 이 과정에서 오류율이 비정상적으로 높게 나타나면 도청이 있었다는 신호로 보고 키를 폐기한 뒤 새로 만든다. 앞서 방송 실증에서 도청 시도가 즉시 드러나고 키가 자동으로 갱신된다고 한 것이 바로 이 절차에 해당한다.

• QKD 프로토콜과 전송 방식의 분화

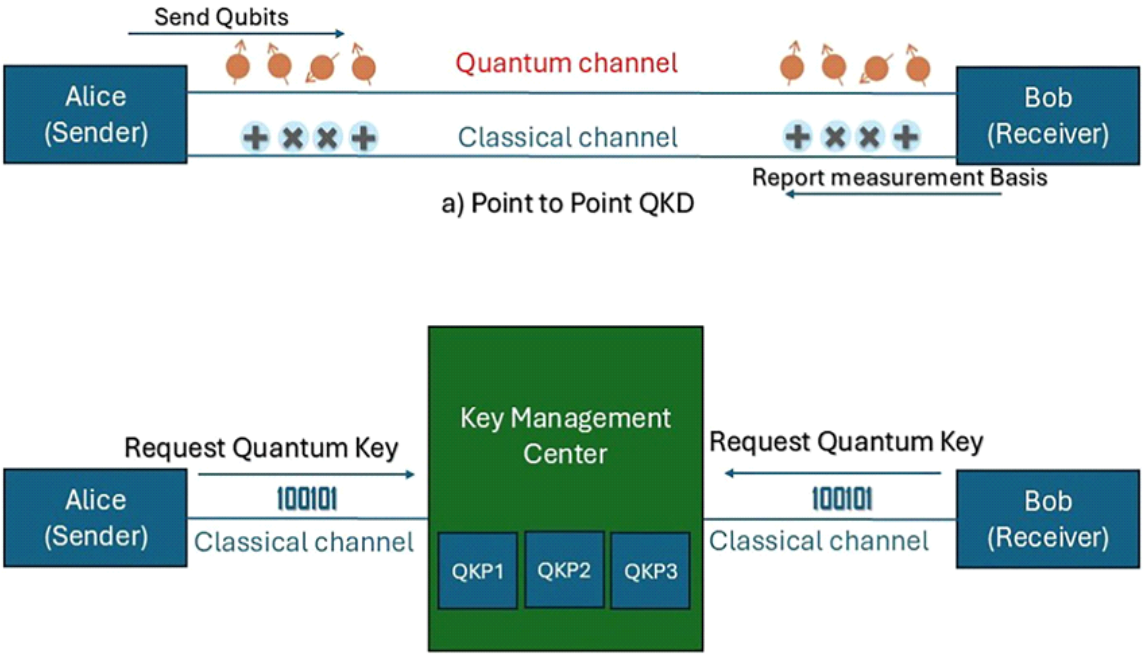
QKD는 정보를 양자 상태에 담는 방식에 따라 크게 두 갈래로 나뉜다. 하나는 이산 변수 방식(Discrete Variable QKD, DV-QKD)으로, 이는 광자 하나하나를 세는 단일 광자 검출로 정보를 읽는 방식이다. 다른 하나는 연속 변수 방식(Continuous Variable QKD, CV-QKD)으로, 이는 빛의 세기와 위상 같은 연속적인 값에 정보를 실어 일반 통신 장비로 검출하는 방식이다. DV-QKD는 잡음에 강하지만 전용 검출 장비가 필요해 비용이 높고, 반대로 CV-QKD는 기존 통신 장비와 잘 맞물리지만 잡음에 더 민감하다. 두 방식 중 무엇을 고를지는 전송 거리, 보안 요구 수준, 기존 설비와의 호환성에 따라 달라진다.

대표적인 방식으로는 QKD 연구에서 초기 프로토콜 중 하나인 BB84가 있다. BB84는 DV-QKD에 속하며, 단일 광자를 주고받는 방식이어서 광섬유는 물론 자유 공간 전송에도 잘 맞고 비교적 먼 거리까지 키를 전달할 수 있다. 한편 양자 얽힘이라는 또 다른 양자 성질을 이용하는 E91 방식도 있는데, 이는 두 입자가 서로 연결되어 한쪽 상태가 정해지면 다른 쪽도 상관관계를 보인다. E91은 장비에 독립적인 높은 보안을 제공하지만, 얽힘 상태가 먼 거리에서 쉽게 약해져 실제 도달 거리는 BB84보다 짧은 편이다. 이처럼 프로토콜마다 강점과 한계가 달라, 쓰임새에 맞춰 선택된다.

한편 실제 서비스에서는 프로토콜 뿐만 아니라 이를 어떤 네트워크 구조에서 운용할 것인지도 중요한 고려 요소가 된다. 전송 구조 측면에서도 방식이 갈린다. 가장 기본은 지점 대 지점 방식으로, 이는 송신자와 수신자를 전용 양자 채널로 직접 잇는 구조다. 앞서 살펴본 두 도시 간 방송 실증이 이 구조에 해당한다. 이용자가 늘어나면 지점마다 전용 채널을 설치하는 부담이 커지므로, 하나의 송신 장치가 여러 수신자에게

키를 나눠 주는 지점 대 다지점 방식이 대안으로 제시된다. 이 구조에서는 키 관리 센터가 중앙에서 키의 생성과 분배를 조율해 설비 부담을 줄일 수 있으나, 동기화와 오류 정정, 공유 설비에서 생길 수 있는 취약점을 함께 풀어야 하는 과제가 남는다.

[그림 1] QKD의 지점 대 지점 및 지점 대 다지점 전송 구조



출처: Dalal Aljebry 외 1명, “Quantum key distribution in next-generation networks: a survey of space-based, aerial, and SDN-enabled frameworks for secure communication”, Springer Nature, 2026.06.01., https://link.springer.com/article/10.1007/s11128-026-05216-y

• 지상 한계와 차세대 네트워크로의 확장

QKD는 광섬유를 따라 전송할 때 거리의 벽에 부딪힌다. 양자 상태를 복제하려면 그 상태를 측정해야 하는데, 측정하는 순간 상태가 흐트러져 원본을 그대로 베낄 수 없다. 이 때문에 일반 통신처럼 신호를 중간에 복제해 키우는 증폭 방식을 쓸 수 없다. 광자가 광섬유를 지나며 점차 손실되기 때문에, 지상 광섬유에서는 일반적으로 100~200km 안팎이 한계로 꼽힌다. 앞서 살펴본 방송 실증처럼 가까운 구간에서는 문제가 되지 않지만, 대륙을 가로지르는 먼 거리로 넓히려 하면 이 손실이 걸림돌이 된다. 신뢰 노드나 양자 중계기 같은 방법으로 거리를 늘리려는 시도가 있으나, 각각 물리적 보안이나 실험적 난이도라는 새로운 부담을 안고 있다.

이처럼 지상 광섬유 기반 확장에는 기술적 제약이 남아 있어, 이를 보완하기 위한 다양한 대안이 함께 제시되고 있다. 이 한계를 넘기 위한 방향으로 위성과 공중 플랫폼을 이용하는 방식이 제시된다. 위성은 신호

손실이 적은 우주 공간의 조건을 활용해, 멀리 떨어진 지점 사이에 양자 키를 전달하는 매개체로 활용 가능하다. 실제로 중국의 미셔스(Micius) 위성은 1,200km가 넘는 거리에서 얽힘을 나눠 안전한 키를 주고받는 데 성공했고, 유럽에서도 저궤도 위성에 QKD 장비를 실은 이글원(EAGLE-1) 계획이 추진되고 있다.[1] 최근에는 드론이나 성층권에 머무는 고고도 플랫폼에 이동식 중계 지점 역할을 맡겨, 지상 설비가 닿기 어려운 지역까지 연결을 넓히려는 시도도 이어진다. 이러한 비지상 방식은 지상 광섬유의 거리 한계를 우회하는 통로로 평가된다.

나아가 QKD를 차세대 통신망에 녹여 넣으려는 움직임도 이어진다. 소프트웨어로 망을 제어하는 방식(Software-Defined Networking, SDN)과 결합하면, 중앙에서 키의 경로와 자원을 실시간으로 조절해 수요와 위협 변화에 유연하게 대응할 수 있다. 이러한 구조를 6G 통신망과 스마트시티에 통합하려는 구상도 제시되며, 기존 암호 방식인 포스트 양자 암호(Post-Quantum Cryptography, PQC)를 QKD와 함께 쓰는 혼합 방식이 현실적인 단기 해법으로 꼽힌다. 여기에 머신러닝(machine learning)을 더해 채널 잡음을 예측하거나 도청을 탐지하는 보조 수단으로 활용하는 연구도 진행된다. 다만 높은 장비 비용, 표준의 미비, 기존 설비와의 호환은 아직 풀어야 할 과제로 남아 있다.

III. 결론: 양자 보안 전환의 방향

• 양자 보안 전환의 시사점과 과제

결국 이번 방송 실증이 보여 준 것은 콘텐츠를 지키는 방식이 바뀌고 있다는 사실이다. 그동안 통신 보안은 계산의 어려움에 기대어 왔으나, QKD는 양자역학의 원리 자체에 보안의 근거를 두어 도청 시도를 원리적으로 드러낸다. 이는 미래의 양자컴퓨터가 기존 암호를 무력화하더라도 전송 단계의 콘텐츠를 지킬 통로가 열린다는 것을 의미한다. 이러한 점에서 QKD는 단순한 실험을 넘어, 장기 보호 대상인 방송과 통신 콘텐츠의 보안을 다시 설계하는 출발점으로 볼 수 있다.

이를 통해 전송 중 신호가 가로채이거나 조작되는 위험에 대응하는 새로운 선택지가 마련된다.

다만 이 기술이 널리 정착하기까지는 풀어야 할 과제가 남아 있다. 지상 광섬유의 거리 한계, 높은 장비 비용, 통일된 표준의 미비, 기존 설비와의 호환은 확산을 가로막는 요인으로 꼽힌다. 따라서 위성과 공중 플랫폼을 통한 거리 확장, 포스트 양자 암호와의 혼합 방식, 소프트웨어 기반 망 제어와의 결합 같은 보완이 함께 진행될 필요가 있다. 제도 측면에서도 표준 마련과 기관 간 협력이 뒷받침되어야 하며, 산업 측면에서는 장비의 소형화와 비용 절감이 도입 속도를 좌우할 전망이다. 이러한 점에서 QKD의 앞날은 기술의 성숙과 제도적 준비가 맞물릴 때 비로소 열릴 것으로 보인다.

1) Dalal Aljebry 외 1명, “Quantum key distribution in next-generation networks: a survey of space-based, aerial, and SDN-enabled frameworks for secure communication”, Springer Nature, 2026.06.01., https://link.springer.com/article/10.1007/s11128-026-05216-y

참고문헌

- CSI, “Vodafone and ARD test quantum encryption to protect TV broadcasts from cyber threats”, 2026.06.26., https://www.csimagazine.com/csi/Vodafone-and-ARD-test-quantum-encryption.php
- Jörn Krieger, “Vodafone, ARD and DHBW test quantum-encrypted TV transmission”, Broadband TV News, 2026.06.26., https://www.broadbandtvnews.com/2026/06/26/vodafone-ard-and-dhbw-test-quantum-encrypted-tv-transmission/
- Dalal Aljebry 외 1명, “Quantum key distribution in next-generation networks: a survey of space-based, aerial, and SDN-enabled frameworks for secure communication”, Springer Nature, 2026.06.01., https://link.springer.com/article/10.1007/s11128-026-05216-y

ARTMARKET, ““AI and the Art Market” by Arte, with Artprice by Artmarket and analysis by Perplexity.AI Labs”, 2025.06.16., https://www.actusnews.com/fr/artmarket/cp/2025/06/16/ai-and-the-art-market-by-arte-with-artprice-by-artmarket-and-analysis-by-perplexity_ai-labs

4개의 음악 데이터셋으로 본 인공지능 음악 학습 데이터의 수집과 추적 한계

뉴스 브리프

인공지능 음악 학습 데이터 문제는 생성 결과물이 기존 음악과 얼마나 유사한지뿐 아니라, 기존 음원이 어떤 방식으로 수집·공유되고 모델 학습에 활용되는지와도 연결된다. 디애틀랜틱의 AI 워치독이 공개한 4개의 음악 데이터셋은 유튜브와 스포티파이의 링크·메타데이터를 모은 초대형 자료부터 음원 파일을 직접 포함한 자료까지 서로 다른 방식으로 구축돼 있다. 다만 검색 기능으로 자신의 음악이 데이터셋에 포함됐는지는 확인할 수 있지만, 해당 음악이 특정 모델의 학습에 실제로 사용됐는지까지 입증하기는 어렵다. 따라서 인공지능 음악 학습 데이터 대응에서는 데이터셋 수록 여부와 실제 이용 주체를 구분하고, 온라인 공개와 인공지능 학습 동의의 차이도 함께 살펴볼 필요가 있다. 본 보고서는 4개의 음악 데이터셋의 규모와 구성 방식을 살펴보고, 공개 데이터셋 검색을 통한 권리자의 음원 추적 가능성과 한계를 분석한다.

뉴스 플러스

I. 서론: 인공지능 음악 학습 데이터의 불투명성

• 공개되지 않는 음악 학습 데이터

인공지능 음악 생성 모델이 사람의 음악과 유사한 결과물을 만들기 위해서는 대량의 기존 음원을 학습하는 과정이 필요하다. 모델은 학습 음원을 작은 오디오 단위로 나누고 각 소리가 나타나는 맥락을 학습한 뒤, 이용자가 입력한 조건에 따라 다음에 이어질 소리를 예측하는 방식으로 음악을 생성한다.

그러나 기업은 모델에 실제로 사용한 음원을 독점 정보로 취급하며 구체적인 학습 목록을 공개하지 않는 경우가 많다. 이에 따라 음악가와 권리자가 자신의 음원이 인공지능 학습에 포함됐는지를 외부에서 확인하기는 어렵다.

미국 시사·문화 매체 디 애틀랜틱(The Atlantic)의 'AI 워치독(AI Watchdog)'은 인공지능 개발자 사이에서 공유되는 음악 데이터셋 4개를 조사했다. AI 워치독은 주요 기술기업이 인공지능 모델 학습에 활용한 책, 영상과 그 밖의 미디어를 추적하는 연속 조사 프로젝트이다. 또한 음악가의 이름을 검색해 각 데이터셋의 수록 여부를 확인할 수 있는 기능을 제공한다.

[그림 1] 디 애틀랜틱 AI 워치독 데이터셋 검색 화면

출처: The Atlantic, "AI Watchdog", 2026.07.14. 접속 기준, https://www.theatlantic.com/category/ai-watchdog/

II. 본론: 4개의 음악 데이터셋으로 본 수집 구조

• 유튜브 음악을 기반으로 한 초대형 데이터셋

라이온-디스코-12M(LAION-DISCO-12M)은 유튜브(YouTube) 음악 1,232만 916곡으로 구성된 데이터셋이다. 전체 음악의 재생시간은 약 91년에 이른다. 이 데이터셋은 대규모 학습 자료를 구축하는 독일 비영리단체 라이온(LAION)이 제작했다.

[그림 2] 라이온 디스코 데이터셋 수록 음원 검색 화면

출처: The Atlantic, "AI Watchdog: LAION-DISCO-12M", 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-laion-disco-12m/687508/

슬리핑-디스코-9M(Sleeping-DISCO-9M)에는 유튜브 음악 971만 3,413곡과 지니어스닷컴(Genius.com)에서 수집한 가사 정보가 포함돼 있다. 인공지능 학습 데이터셋을 구축하고 관련 연구를 공개하는 연구자 단체 슬리핑 AI(Sleeping AI)가 제작했다.

[그림 3] 슬리핑 디스코 데이터셋 수록 음원 검색 화면

출처: The Atlantic, "AI Watchdog: Sleeping-DISCO-9M", 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-sleeping-disco-9m/687509/

두 데이터셋은 모두 유튜브 음악을 기반으로 구축됐지만, 슬리핑-디스코-9M에는 가사 정보도 함께 포함돼 있다는 차이가 있다. 디 애틀랜틱이 확인한 네 데이터셋 중 프리 뮤직 아카이브(Free Music Archive)를 제외한 세 데이터셋은 유튜브나 스포티파이(Spotify) 음악의 링크 목록으로 배포된다. 개발자는 자동화 도구를 이용해 해당 링크에서 실제 오디오를 내려받을 수 있다. 반면 일부 데이터셋은 링크 목록이 아니라 실제 음원 파일 자체를 포함한다는 점에서 성격이 다르다.

• 음원 파일을 포함한 프리 뮤직 아카이브

프리 뮤직 아카이브 데이터셋(Free Music Archive Dataset)은 2016년 프리 뮤직 아카이브에서 내려받은 10만 6,574곡으로 구성됐다. 스위스 로잔연방공과대학교(École Polytechnique Fédérale de Lausanne, EPFL)가 제작했으며, 링크가 아닌 음원 파일을 직접 포함한다.

데이터셋에 포함된 음악 대부분에는 이용 시 음악가를 표시하고 상업적 이용을 금지하는 크리에이티브 커먼즈(Creative Commons) 조건이 적용돼 있다. 그러나 이 데이터셋은 구글(Google)의 생성형 인공지능 모델 학습에 사용됐으며, 이 가운데 1만 3,874곡은 스태빌리티 AI(Stability AI)도 활용했다.

프리 뮤직 아카이브 운영사 책임자는 구글이 해당 데이터셋을 인공지능 모델 학습에 활용한 사실을 확인했다. 이후 구글에 서한을 보내 음악가의 동의와 보상 문제를 논의하자고 요구했다. 구글은 공개된 정보

를 인공지능 학습에 활용한다는 입장을 밝혔지만, 운영사가 제기한 우려에는 직접 답변하지 않았다.

프리 뮤직 아카이브에 250곡 이상을 공개한 음악가 데릭 클레그(Derek Clegg)는 개인 영상에서 출처를 표시하고 음악을 이용하는 것은 허용하지만, 수익을 목적으로 이용하면 라이선스 비용을 받아 왔다고 설명했다. 그는 인공지능 학습을 거부할 수 있는 수단이 있다면 자신의 음악을 제외하겠다고 밝혔다.[1)]

이러한 사례는 음악을 온라인에서 무료로 들려주거나 일정한 조건 아래 공유하도록 허용한 것이 상업적 인공지능 학습에 대한 동의까지 의미하는지는 별도로 살펴봐야 한다는 점을 보여준다.

[그림 4] 프리 뮤직 아카이브 데이터셋 수록 음원 검색 화면

출처 The Atlantic, "AI Watchdog: Free Music Archive Dataset", 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-free-music-archive/687336/

- **제작자를 확인하기 어려운 스포티파이 데이터셋**

스포티파이 트랙 데이터셋(Spotify Tracks Dataset)은 스포티파이(Spotify)에서 수집한 음악 11만 4,000곡으로 구성됐다. 신원이 공개되지 않은 인공지능 개발자가 인공지능 자료 공유 플랫폼 허깅페이스(Hugging Face)에 데이터셋을 게시했다.

이 데이터셋은 2026년 5월까지 7만 회 이상 내려받아졌다. 스포티파이는 데이터셋의 제작이나 배포에 관여하지 않았다. 두 초대형 데이터셋보다 규모는 작지만, 제작자의 신원과 구체적인 제작 목적, 공개 이후의 활용 주체를 확인하기 어렵다는 특징이 있다.

1) Alex Reisner, "The Millions of Songs Mashed Into AI-Generated Music", The Atlantic, 2026.06.14., https://www.theatlantic.com/technology/2026/06/ai-music-generators-suno-google-udio/687485/

[그림 5] 스포티파이 데이터셋 수록 음원 검색 화면

출처 The Atlantic, "AI Watchdog: Spotify Tracks Dataset", 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-spotify-tracks/687510/

• **데이터셋 등재와 실제 모델 학습의 구분**

4개의 데이터셋은 인공지능 개발자가 접근할 수 있는 음악의 규모와 구성 방식이 다양하다는 점을 보여준다. 음악 데이터의 이용 가능성을 살펴볼 때는 음원 파일이 직접 배포됐는지만 확인해서는 충분하지 않다. 링크와 메타데이터를 대규모로 정리한 데이터셋도 자동화 도구를 통한 음원 수집의 기반으로 활용될 수 있기 때문이다.

이 가운데 프리 뮤직 아카이브 데이터셋은 구글과 스태빌리티 AI의 모델 학습에 활용된 사실이 확인됐다. 반면 라이온-디스코-12M, 슬리핑-디스코-9M과 스포티파이 트랙 데이터셋의 이용 주체는 알려지지 않았다.

따라서 디 애틀랜틱의 검색 기능은 특정 음악의 데이터셋 수록 여부를 확인하는 데 활용할 수 있지만, 해당 음악이 특정 인공지능 모델의 학습에 실제로 사용됐는지까지 입증하지는 못한다.

III. 결론: 음악 데이터셋 검색의 가능성과 한계

디 애틀랜틱이 제공한 검색 기능을 통해 음악가와 권리자는 자신의 음악이 네 개의 대규모 데이터셋에 포함돼 있는지를 확인할 수 있다. 인공지능 기업이 구체적인 학습 목록을 공개하지 않는 상황에서 자신의 음악이 인공지능 개발 자료에 포함됐는지를 살펴보는 출발점으로 활용할 수 있다.

다만 데이터셋 수록 여부와 특정 모델의 실제 학습 여부는 구분해야 한다. 프리 뮤직 아카이브는 구글과 스태빌리티 AI의 활용 사실이 확인됐지만, 나머지 데이터셋은 구체적인 이용 주체가 공개되지 않았다.

프리 뮤직 아카이브 운영사와 음악가의 사례는 온라인에서 음악을 공개하거나 공유를 허용한 것이 인공지능 학습에 대한 동의까지 의미하는지는 별도로 살펴봐야 한다는 점을 보여준다. 향후에는 권리자가 자신의 음악이 어떤 데이터셋에 포함돼 있는지 확인할 수 있도록 학습 자료의 출처와 이용 범위가 보다 투명하게 제시될 필요가 있다.

참고문헌

- The Atlantic, “AI Watchdog”, 2026.07.14. 접속 기준, https://www.theatlantic.com/category/ai-watchdog/
- The Atlantic, “AI Watchdog: LAION-DISCO-12M”, 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-laion-disco-12m/687508/
- The Atlantic, “AI Watchdog: Sleeping-DISCO-9M”, 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-sleeping-disco-9m/687509/
- The Atlantic, “AI Watchdog: Free Music Archive Dataset”, 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-free-music-archive/687336/
- The Atlantic, “AI Watchdog: Spotify Tracks Dataset”, 2026.06.12., https://www.theatlantic.com/technology/2026/06/dataset-spotify-tracks/687510/
- Alex Reisner, "The Millions of Songs Mashed Into AI-Generated Music Explore the astonishing amount of music available to AI developers.". The Atlantic, 2026.06.14., https://www.theatlantic.com/technology/2026/06/ai-music-generators-suno-google-udio/687485/
- Mandy Dalugdug, "Four music datasets holding millions of tracks are being shared among AI developers, The Atlantic reports", Music Business Worldwide, 2026.06.16., https://www.musicbusinessworldwide.com/four-music-datasets-holding-millions-of-tracks-are-being-shared-among-ai-developers-the-atlantic-reports/

마스크 유도 비디오 워터마킹 플로우마크와 콘텐츠 진본성 확보

뉴스 브리프

디지털 비디오와 오디오는 온라인 플랫폼에 유통되기까지 압축, 재인코딩, 화면 녹화, 편집 같은 여러 처리를 거친다. 이 과정에서 콘텐츠에 담긴 소유권과 출처 정보는 쉽게 훼손되며, 눈에 보이지 않게 삽입한 워터마크가 처리 이후에도 남아 있는지가 콘텐츠 보호의 실효성을 좌우한다. 최근 연구는 비디오 안에서 워터마크를 심을 최적의 영역을 찾고, 프레임마다 고유한 워터마크를 남겨 삽입, 삭제, 순서 변경 같은 편집과 변조를 탐지하는 방향으로 발전하고 있다. 다만 워터마크를 견고하게 만들수록 화질이 떨어지고, 프레임마다 다른 워터마크를 넣으면 깜빡임이 생기는 등 상충하는 조건을 함께 만족시켜야 한다. 본 보고서는 비디오 워터마킹의 작동 원리를 살펴보고, 그것이 저작권 보호와 변조 탐지에 연결되는 방식을 분석하며, 콘텐츠 출처 증명 체계가 나아갈 방향과 과제를 짚는다.

뉴스 플러스

Ⅰ. 서론: 디지털 콘텐츠 진본성 확보의 과제

• 플랫폼 재처리 이후의 워터마크 생존 문제

디지털 비디오와 오디오는 제작된 원본 그대로 이용자에게 도달하는 경우가 드물며, 온라인 플랫폼에 유통되는 과정에서 압축, 재인코딩, 화면 녹화, 음량 정규화, 편집 같은 여러 처리를 반복해서 거치게 된다. 이러한 처리는 화면에 드러나는 내용을 크게 바꾸지 않으면서도 원본을 미세하게 변형하기 때문에, 원저작자가 누구이고 콘텐츠의 출처가 어디인지를 확인할 정보는 유통 과정에서 쉽게 훼손된다. 그 결과 이용자가 접하는 콘텐츠가 원저작자의 것인지 아니면 누군가에 의해 무단으로 복제되거나 변형된 것인지를 사후에 가려내는 일은 까다로운 과제가 되며, 이는 콘텐츠 보호가 가장 먼저 부딪히는 출발점이 된다.

이러한 처리에 견디도록 콘텐츠 안에 사람 눈에 보이지 않게 삽입해 두는 표시가 보이지 않는 워터마크(watermark)이며, 이를 다시 읽어 내는 방법으로 원저작자와 출처를 증명한다. 파일에 제작 정보를 적어 두는 메타데이터(metadata) 방식도 있으나, C2PA(Coalition for Content Provenance and Authenticity)와 같은 표준을 따르더라도 그 정보는 콘텐츠가 플랫폼을 거치는 사이 쉽게 지워지고 만다. 그래서 값어치가 큰 콘텐츠일수록 메타데이터에만 의존하기 어려우며, 콘텐츠 안에 직접 담겨 잘 지워지지 않는 워터마크가 대안으로 주목받는다. 워터마크는 콘텐츠 자체에 삽입되므로 이러한 제거에는 상대적으로 견고하지만, 여러 처리를 거친 뒤에도 워터마크가 그대로 읽히는지가 실효성을 가른다. 이에 한 가지 방식만으로는 다양한 처리를 감당하기 어렵다는 인식이 퍼지면서, 워터마크와 함께 출처 기록, 유출 추적, 매칭을 나란히 활용하는 계층적 접근이 힘을 얻고 있다.

* C2PA(Coalition for Content Provenance and Authenticity): 디지털 콘텐츠의 제작·수정 이력과 AI 생성 여부 등을 확인할 수 있도록 첨부되는 출처 인증 정보로, 콘텐츠의 생성 방식과 변경 내역을 검증하는 데 활용됨

- **비디오 워터마킹의 기술적 한계와 등장 배경**

콘텐츠 가운데 특히 비디오는 워터마크를 삽입하기가 이미지보다 까다롭다. 비디오는 프레임이 시간 순서대로 이어지므로 프레임마다 따로 워터마크를 넣으면 프레임 사이에 미세한 차이가 생겨 재생할 때 영상이 떨리는 현상이 나타나기 때문이다. 이러한 떨림은 워터마크가 콘텐츠에 남긴 흔적이 프레임마다 들쭉날쭉할 때 두드러지며, 비디오를 온라인에 올리기 위해 용량을 줄이는 압축 과정에서 더욱 도드라져 보는 사람이 이질감을 느끼게 만든다. 떨림을 줄이려고 여러 프레임에 똑같은 워터마크를 반복해서 넣는 방법도 쓰였으나, 이 경우 프레임마다 고유한 워터마크를 넣을 수 없어 콘텐츠의 어느 지점이 잘려 나가거나 순서가 뒤바뀌어도 그 사실을 사후에 짚어 내기 어렵다는 한계가 있었다.

이러한 배경에서 등장한 것이 플로우마크(FlowMark)로, 프레임마다 워터마크를 넣기에 알맞은 영역을 스스로 찾아내고 인접 프레임 간에 남는 흔적을 균일하게 조정하여 재생할 때의 떨림을 억제하도록 설계된 비디오 워터마킹 방식이다. 플로우마크는 프레임마다 서로 다른 워터마크를 남기면서도 떨림을 함께 줄이기 때문에, 프레임의 순서가 바뀌거나 일부가 잘려 나가는 변형을 가려내는 일과 원저작자를 증명하는 일을 동시에 겨냥한다.

II. 본론: 마스크 유도 비디오 워터마킹의 원리

- **마스크 예측을 통한 콘텐츠 적응형 삽입**

플로우마크의 출발점은 마스크 예측기(Mask Predictor)라 불리는 신경망이다. 마스크(mask)는 한 프레임 위에서 워터마크를 넣을 자리와 넣지 않을 자리를 구분해 칠해 둔 밑그림이다. 기존 방식은 프레임 전체에 워터마크를 고르게 펴 바르거나 사람이 직접 영역을 지정해야 했으나, 플로우마크는 프레임을 입력받아

어느 자리가 워터마크를 담기에 알맞은지 마스크 예측기가 스스로 판단한다. 이렇게 콘텐츠의 내용에 맞추어 삽입 영역을 바꾸는 방식을 콘텐츠 적응형 삽입이라 부른다.

작동 순서는 다음과 같다. 먼저 마스크 예측기가 프레임 분석을 통해 영역별 워터마크 삽입 적합도를 0과 1 사이 값으로 산출한 맵(map)을 생성한다. 이후 해당 맵을 워터마크 삽입 영역과 비삽입 영역으로 이분화하되, 학습이 원활하도록 값의 연속성은 그대로 유지한다. 이어서 텍스처가 복잡하거나 움직임이 많아 사람 눈에 잘 띄지 않는 영역에 워터마크가 집중되도록 삽입이 이루어진다. 이 과정을 거치면 워터마크가 프레임 전체에 퍼지지 않고 눈에 덜 거슬리는 자리에만 모이게 된다.

콘텐츠 적응형 삽입은 실제 콘텐츠 산업이 요구하는 조건과 맞물려 확산되고 있다. 프레임마다 사람이 일일이 영역을 지정하는 작업은 수많은 콘텐츠가 오가는 플랫폼 환경에서는 감당하기 어렵기 때문에, 삽입 영역을 자동으로 정하는 방식은 대량 처리에 유리하다. 또한 워터마크를 눈에 덜 띄는 자리에 모으면 화질 손상을 줄이면서도 워터마크를 남길 수 있어, 콘텐츠의 상품 가치를 지키려는 제작자의 요구와도 부합한다. 저작권 보호의 관점에서 보면, 워터마크가 콘텐츠 안에서 잘 지워지지 않는 자리에 자리 잡을수록 무단 복제나 편집을 거친 뒤에도 원저작자를 확인할 단서가 남을 가능성이 커진다.

• 어두운 영역 조정과 시간적 일관성 확보

플로우마크는 워터마크를 넣되 사람 눈에 잘 띄지 않게 하려고 어두운 영역 적응형 조정(dark-region adaptive residual modulation)이라는 기법을 쓴다. 이는 프레임에서 밝기가 낮은 부분에는 워터마크를 약하게 넣고 밝은 부분에는 상대적으로 강하게 넣도록 삽입 강도를 조절하는 방식이다. 어두운 영역에 워터마크를 세게 넣으면 얼룩이나 잡티가 유독 도드라져 보이기 때문에, 플로우마크는 프레임의 밝기를 자리마다 따져 어두운 곳일수록 워터마크의 세기를 줄인다. 이렇게 하면 워터마크가 남긴 흔적이 눈에 덜 띄면서도 필요한 자리에는 워터마크가 유지된다.

떨림을 억제하는 기법은 시간적 변화 일관성(temporal change consistency)이라 불린다. 이는 원본 비디오에서 프레임과 프레임 사이에 일어나는 변화의 양과, 워터마크를 넣은 비디오에서 프레임 사이에 일어나는 변화의 양을 최대한 비슷하게 맞추는 방식이다. 원본에서 장면이 부드럽게 이어지는 구간이라면 워터마크를 넣은 뒤에도 그만큼 부드럽게 이어져야 하며, 두 변화량의 차이가 벌어질수록 재생할 때 떨림으로 나타난다. 플로우마크는 이 차이를 줄이도록 학습하기 때문에, 프레임마다 다른 워터마크를 넣으면서도 원본의 움직임을 흐트러뜨리지 않는다.

• 구조화된 비트 삽입과 프레임 단위 추적

플로우마크가 프레임에 넣는 워터마크는 128비트 길이의 정보이며, 이 정보는 두 부분으로 나뉜다. 앞의 112비트는 비디오 식별자로, 이 비디오가 어떤 콘텐츠인지를 가리키는 고유 번호이며 모든 프레임에 똑같이

들어간다. 뒤의 16비트는 프레임 색인으로, 지금이 몇 번째 프레임인지를 나타내는 번호로 프레임이 넘어갈 때마다 하나씩 커진다. 앞서 나온 방식들이 모든 프레임에 같은 워터마크를 넣어 프레임을 구별하지 못했던 것과 달리, 플로우마크는 비디오 전체를 가리키는 번호와 프레임마다 다른 번호를 함께 담아 각각의 프레임을 구별할 수 있게 한다.

이 구조 덕분에 변조를 탐지하는 원리는 다음과 같이 작동한다. 프레임 색인은 순서대로 하나씩 커지도록 되어 있으므로, 어느 프레임이 빠지거나 순서가 바뀌면 번호가 끊기거나 어긋난다. 따라서 워터마크에서 읽어 낸 프레임 색인을 살펴보면 프레임이 삭제되었는지, 사이에 다른 프레임이 끼어들었는지, 순서가 뒤바뀌었는지를 짚어 낼 수 있다. 실제로 플로우마크는 프레임 교체, 삭제, 삽입 등의 편집을 가한 실험에서 변조된 프레임을 가려냈으며, 편집되지 않은 프레임은 워터마크를 그대로 읽어 냈다.

워터마크에 담긴 두 정보는 서로 다른 기능을 맡는다. 비디오 식별자는 이 비디오가 어떤 콘텐츠인지를 가리키고, 프레임 색인은 콘텐츠의 어느 부분이 잘리거나 바뀌었는지를 짚어 내는 데 쓰인다. 하나의 워터마크 안에 비디오 전체를 가리키는 정보와 프레임마다 다른 정보를 함께 담은 셈이다. 다만 이러한 탐지는 워터마크가 편집과 압축을 견디고 살아남았을 때 성립하는 것이므로, 실제 환경에서 워터마크가 얼마나 견고하게 유지되는지가 뒤따라 확인되어야 한다.

• 단계적 학습 전략과 실환경 견고성 검증

플로우마크는 세 단계로 나눈 학습 과정을 거치며, 단계가 올라갈수록 더 거친 손상을 견디도록 훈련된다. 첫 단계에서는 아무런 손상도 주지 않은 깨끗한 프레임으로 워터마크를 넣고 읽는 기본기를 익힌다. 둘째 단계에서는 밝기나 색을 바꾸거나 이미지를 자르고 크기를 조절하는 등 이미지에 가해지는 손상을 주어, 이런 변형을 거친 뒤에도 워터마크를 읽어 내도록 적응시킨다. 셋째 단계에서는 비디오를 온라인에 올릴 때 용량을 줄이는 압축을 실제로 가하며, 그 강도를 점점 높여 가장 까다로운 조건에 대비시킨다. 이렇게 손상의 강도를 서서히 끌어올리는 방식은 학습을 안정적으로 이끌면서도 실제 유통 환경에 가까운 조건을 익히게 한다.

이렇게 학습한 플로우마크는 여러 손상을 거친 뒤에도 워터마크를 높은 정확도로 읽어 낸다. 밝기나 색의 변화, 자르기나 회전 같은 변형, 압축을 각각 가한 조건에서 워터마크를 대부분 정확히 복원했으며, 유튜브와 페이스북에 실제로 올렸다가 내려받아 플랫폼의 압축을 거친 뒤에도 워터마크가 유지되었다. 화질 손상을 나타내는 지표에서도 원본과 구별하기 어려운 수준을 유지해, 견고성과 화질을 함께 잡았다고 평가된다. 다만 워터마크를 지우려고 의도적으로 공격하는 조건에서는 정확도가 다소 떨어지는 만큼, 모든 상황에서 완벽하게 유지된다고 보기는 어렵다.

III. 결론: 콘텐츠 출처 증명 기술의 향방

• 콘텐츠 출처 증명 기술의 시사점과 전망

플로우마크는 프레임마다 워터마크를 넣을 자리를 스스로 찾고, 어두운 영역의 세기를 낮추며, 프레임 사이의 변화를 고르게 맞추는 방법으로 화질과 떨림이라는 오랜 상충을 완화한다. 여기에 비디오 식별자와 프레임 색인을 함께 담아, 원저작자를 가리키는 일과 콘텐츠가 편집되었는지를 짚어 내는 일을 하나의 워터마크로 아우른다. 결국 이 기술이 보여 주는 것은 워터마크가 단순히 소유권을 표시하는 수단을 넘어, 콘텐츠가 어디서 왔고 어떻게 바뀌었는지를 함께 담는 기록으로 확장되고 있다는 점이다. 이를 통해 플랫폼을 거치며 정보가 지워지기 쉬운 환경에서도, 콘텐츠 자체에 새겨진 단서로 출처와 변형 여부를 확인할 여지가 넓어진다. 따라서 콘텐츠 진본성 확보는 콘텐츠 바깥에 덧붙이는 정보가 아니라 콘텐츠 안에 심는 단서를 중심으로 옮겨 가고 있다.

다만 이러한 기술이 자리 잡으려면 남은 과제도 분명하다. 플로우마크 역시 워터마크를 지우려는 의도적 공격 앞에서는 정확도가 떨어지는데, 이는 어떤 워터마크도 모든 손상을 홀로 감당하기는 어렵다는 것을 의미한다. 이러한 점에서 워터마크와 함께 출처 기록, 유출 추적, 매칭을 겹쳐 쓰는 계층적 접근이 뒷받침될 필요가 있으며, 서로 다른 방식이 각자의 빈틈을 메우도록 설계되어야 한다. 제도의 측면에서는 이렇게 확인한 출처와 변형 정보를 어떻게 다룰지에 대한 기준이 함께 마련되어야 하고, 기술의 측면에서는 워터마크가 견디는 손상의 범위를 넓히고 공격에 대한 내성을 높이는 개선이 이어져야 한다. 이러한 점에서 콘텐츠 출처 증명 기술은 하나의 완성된 해법이라기보다, 여러 방식과 제도가 함께 맞물려 발전해 가는 과정에 놓여 있다.

참고문헌

• Michael Sumner, "Invisible Watermarking for Video and Audio Content", SocreDetect, 2026.06.25., https://www.scoredetect.com/blog/posts/invisible-watermarking-for-video-audio-content
• Vishal Asnani 외 2명, "FlowMark: Mask-Guided Video Watermarking", arXiv, 2026.07.06., https://arxiv.org/pdf/2607.05261v1

기술용어

순번	용어	설명
1	C2PA (Coalition for Content Provenance and Authenticity)	디지털 콘텐츠의 제작·수정 이력과 AI 생성 여부 등을 확인할 수 있도록 첨부되는 출처 인증 정보로, 콘텐츠의 생성 방식과 변경 내역을 검증하는 데 활용됨