

생성형 인공지능 활용 음악의 학습 데이터 기여도 분석

한국저작권위원회
정보기술팀
박동현
2026. 10. 12.

보고서 요약

최근 생성형 인공지능을 활용한 음악이 국내 음원 차트에 진입하는 등 대중의 관심을 받고 있다. 또한 국내에서는 생성형 인공지능 활용 음악의 저작권 등록이 가능한지 그리고 저작권료를 받을 수 있는지에 대해 활발히 논의되고 있다.

이러한 논의와 함께, 인공지능의 학습에 사용된 음악과 새롭게 생성된 음악 사이의 관계를 파악하는 것도 중요한 과제로 남아 있다.

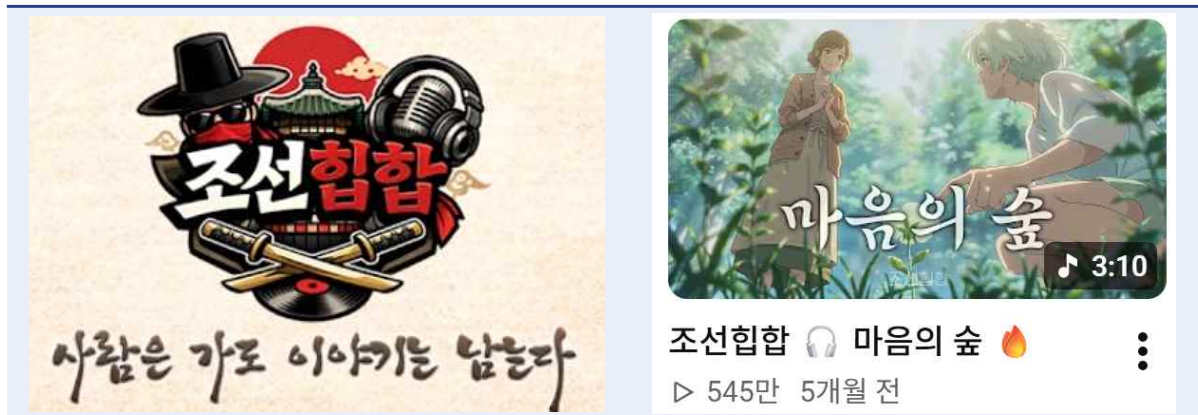
인공지능 모델은 여러 음악 자료를 학습하면서 음악의 특징과 패턴을 익히고 이를 바탕으로 새로운 음악을 생성하는데, 이 과정에서 개별 학습 데이터가 특정 생성 결과에 미친 영향을 추정하는 기술이 학습 데이터 기여도 분석이다.

앞으로 생성형 인공지능 활용 음악을 둘러싼 논의와 함께 학습 데이터 기여도 분석에 관한 연구가 어떻게 발전해 나갈지 살펴볼 필요가 있다.

1. 생성형 인공지능을 활용한 음악의 가능성

최근 생성형 인공지능을 활용한 음악은 국내에서도 크게 인기를 끌고 있다. 유튜브 채널을 운영하고 있는 '조선힙합'이 생성형 인공지능을 활용한 노래 '마음의 숲'으로 음원 플랫폼 멜론의 톱 100 차트에서 최고 89위를 기록하며 화제가 되었다. 이 노래의 작곡과 가창에는 생성형 인공지능이 활용되었는데, '조선힙합'이 직접 가사를 쓴 뒤 생성형 인공지능 음악 제작 플랫폼 Suno에 프롬프트를 입력해 완성했다고 한다.¹⁾

| '조선힙합'의 '마음의 숲'



※ 출처: 유튜브 채널 '조선힙합'

생성형 인공지능을 활용한 범위와 음악을 만드는 방법은 사람에 따라 매우 다양한데 단순히 만들고자 하는 음악의 아이디어를 얻거나 편곡 도구로 활용할 수 있으며, 부분적으로 멜로디와 가사를 만들거나 실질적인 음악 창작 과정의 대부분을 생성형 인공지능으로 만든 경우도 있다.

| 작사, 작곡, 보컬, 공연에 사용되는 생성형 인공지능

● 생성형 인공지능의 발전은 전통적인 음악 창작 구조(작사·작곡·보컬·공연 등)와 음악산업 생태계를 근본적으로 변화시키고 있음

● 텍스트 기반 가사 생성(ChatGPT, Claude), 자동 작곡(Suno, Udio), 보컬 합성(TTS, Voice Cloning), 가상 아이돌 공연(MAVE:) 등 최첨단 기술 전반이 음악산업의 전 과정에 도입되어 창작 도구로 활용됨

■ 작사: ChatGPT, Claude 등 대규모 언어 학습 모델을 활용하여 주제·분위기에 따른 가사 자동 생성

■ 작곡: Suno, Udio, Aiva 등 플랫폼이 장르, 템포만 입력하면 전체 곡을 자동 생성

■ 보컬: TTS, Voice Cloning 기술을 통해 실제 가수의 목소리를 모사하거나 신인 보컬 스타일 창조

■ 공연: 이세계아이돌, 플레이브, 메이브 등의 AI 기반 가상 아이돌이 실제 공연 콘텐츠 주체로 활동

※ 출처: 한국콘텐츠진흥원, “Made by AI 음악과 공존하기: 생성형 AI시대, 음악산업의 전환과 정책 방향”, 2025.06.26.

1) 연합뉴스, “AI 활용곡이 멜론 톱100…조선힙합 “저비용으로 세대 공감 가능””, 2026.09.13.

2. 생성형 인공지능 활용 음악의 저작권 기준 논의

‘마음의 숲’과 같은 생성형 인공지능 활용 음악의 저작권 등록이 가능한지 그리고 저작권료를 받을 수 있는지에 대해 최근 국내에서도 활발히 논의되고 있다. 문화체육관광부는 국회 김재원 의원실, K-음악권리단체상생위원회와 함께 9월 22일 ‘인공지능 활용 음악의 신탁 기준 마련을 위한 토론회’를 개최하였다. 이번 토론회는 생성형 인공지능을 활용한 음악의 창작이 확산됨에 따라 인공지능 활용 음악의 신탁과 등록 관련 쟁점에 대해 창작자, 음악산업계, 전문가, 정책 기관 등의 의견을 폭넓게 수렴하기 위한 자리로 마련되었다.

한국음악저작권협회는 인간이 실질적으로 참여한 인공지능 활용 음악을 신탁 등록할 수 있는 새로운 기준을 시행하려 했지만, 사회적 합의가 충분하지 않다는 지적이 나오면서 이를 철회하였고 현재 생성형 인공지능 활용 음악에 대한 등록은 보류되고 있다고 안내하였다.²⁾ 문화체육관광부는 이번 토론회에서 인간의 창작적 기여에 대한 판단 기준에 대해 공개적으로 논의하고 인공지능 활용 음악의 저작권료 분배 기준을 마련하는 등 후속 토론회를 이어갈 예정이라고 밝혔다.³⁾

| 인공지능 활용 음악의 신탁 기준 마련을 위한 토론회



AI 활용 음악
신탁 기준 마련을 위한 토론회

AI 음악 활용 현황 및 인간의 창작적 기여에 대한 판단 기준

2026.09.22 (화) am 10:00
국회의원회관 대회의실

주최: 김재원 한국저작권위원회 정책연구본부장
김창우 한국음악저작권협회 회장
김창우 한국음악저작권협회 회장
김창우 한국음악저작권협회 회장

주관: K-음악권리단체상생위원회
K-음악권리단체상생위원회
K-음악권리단체상생위원회

NOTICE

「AI 활용 음악 신탁 기준 마련을 위한 토론회」 개최 안내

회원 여러분, 안녕하십니까.

생성형 AI가 음악 창작 현장에 빠르게 확산되면서, AI를 활용해 만든 음악을 어디까지 '창작자의 저작물'로 보고 신탁받아 관리할 것인지가 우리 업계의 가장 시급한 과제가 되었습니다. 이 기준은 곧 회원 여러분의 작품 등록과 저작권료 분배에 직접 영향을 미치는 사안입니다.

이에 음악 권리자 단체들이 함께 뜻을 모아, AI 음악 활용 현황을 점검하고 인간의 창작적 기여를 판단하는 기준을 논의하는 토론회를 국회에서 개최합니다. 창작 현장의 목소리가 기준 마련 과정에 충실히 반영될 수 있도록 회원 여러분의 많은 관심과 참석을 부탁드립니다.

- ※ 출처: 문화체육관광부 보도자료, “생성형 인공지능 활용 음악의 저작권 신탁·분배 기준 논의”, 2026.09.21.
 ※ 출처: 한국음악저작권협회 공지사항, “[AI 활용 음악 신탁 기준 마련을 위한 토론회] 개최 안내”, 2026.09.18.

2) 한국음악저작권협회 공지사항, “AI 활용 음악저작물 신고·등록 기준 철회 안내”, 2026.08.26.

3) 문화체육관광부 보도자료, “생성형 인공지능 활용 음악의 저작권 신탁·분배 기준 논의”, 2026.09.21.

3. 학습 데이터 기여도 분석이란?

앞서 살펴본 생성형 인공지능 활용 음악에 대한 논의와 함께, 인공지능의 학습에 사용된 음악과 새롭게 생성된 음악 사이의 관계를 파악하는 것도 중요한 과제다. 인공지능 모델은 여러 음악 자료를 학습하면서 음악의 특징과 패턴을 익히고 이를 바탕으로 새로운 음악을 생성하는데, 이 과정에서 개별 학습 데이터가 특정 생성 결과에 미친 영향을 추정하는 기술이 학습 데이터 기여도 분석(Training Data Attribution)이다. 학습에 사용된 여러 곡 가운데 어떤 곡이 특정 음악의 생성에 영향을 주었는지, 그 영향이 상대적으로 어느 정도인지를 살펴보는 것이다.

| 학습 데이터 기여도 분석



※ 출처: 참고문헌을 종합하여 ChatGPT로 재구성

기본 원리를 이해하기 위해, 학습에 사용된 음악 A를 포함한 여러 음악을 학습한 모델이 새로운 음악 B를 생성했다고 가정해 보자. 이때 학습에 사용된 음악 A의 영향을 알아보는 직관적인 방법은 A를 포함해 학습한 모델과 A를 제외하고 다시 학습한 모델을 비교하는 것이다. 두 모델에서 동일한 음악 B가 생성될 가능성을 평가하여 A의 학습 여부에 따라 어떤 차이가 나타나는지 확인하는데, A를 제외했을 때 B가 생성될 가능성이 낮아진다면 A가 B의 생성을 뒷받침하는 영향을 주었음을 보여주는 단서가 된다. 이처럼 특정 데이터가 없었다면 결과가 어떻게 달라졌을지를 살펴보는 접근이다.

그러나 실제로 학습에 사용된 음악을 하나씩 제외하면서 모델을 매번 다시 학습시키려면 많은 시간과 계산 자원이 필요하다. 이에 연구자들은 모델의 학습 과정에서 얻은 정보를 이용하여, 특정 데이터를 제외했을 때 나타날 변화를 간접적으로 계산하는 방법을 연구하고 있다. 학습에 사용된 음악이 모델의 생성 과정에 준 영향을 분석 방법에 따라 점수로 추정된 값을 수치로 나타낸 것이 기여도 점수(attribution score)다. 학습에 사용된 여러 음악에 이와 같은 분석 방법을 적용하면, 특정 생성 음악과 관련하여 학습된 음악의 상대적인 영향을 비교할 수 있는데, 간접적으로 계산한 결과이므로 실제로 데이터를 제외하고 다시 학습했을 때의 변화를 얼마나 잘 반영하는지 검증해야 한다.⁴⁾

4. 학습 데이터 기여도 분석 프레임워크 ‘ARIA’

앞서 살펴본 학습 데이터 기여도 분석을 통해 특정 음악의 생성에 영향을 주었을 것으로 추정되는 음악을 찾아낼 수 있지만, 기여도 점수만으로는 음악의 어떠한 특징과 연관되는지 알기 어렵다. 높은 점수를 받은 음악이 어떠한 특징을 가지고 있는지 그리고 이러한 분석 결과를 얼마나 신뢰할 수 있는지를 살펴보기 위해, 스웨덴의 찰머스 공과대학교 연구진은 2026년 5월 학습 데이터 기여도 분석 프레임워크 ‘ARIA(Attribution Result Interpretation and Analysis)’를 새롭게 제안하였다. ‘ARIA’는 학습에 사용된 음악과 기존 방법으로 계산한 기여도 점수를 바탕으로, 음악의 특징을 비교하고 점수에 문제가 나타나는지 점검하는 프레임워크다.⁵⁾

먼저 음악의 특징을 비교하는 과정에서 특정 생성 음악에 대해 기여도 점수가 높은 음악을 선별하게 된다. 이후 선별된 음악끼리 서로 비교하여 리듬이나 소리의 특성 등에서 얼마나 비슷한지 살펴보는 방식이다. 또한 학습 데이터에서 선별된 음악과 같은 수의 음악을 무작위로 선택하여 동일한 방식으로 비교하는데 높은 점수를 받은 음악에서 나타나는 공통적인 특징이 무작위로 선택한 음악보다 얼마나 두드러지는지 확인할 수 있다.

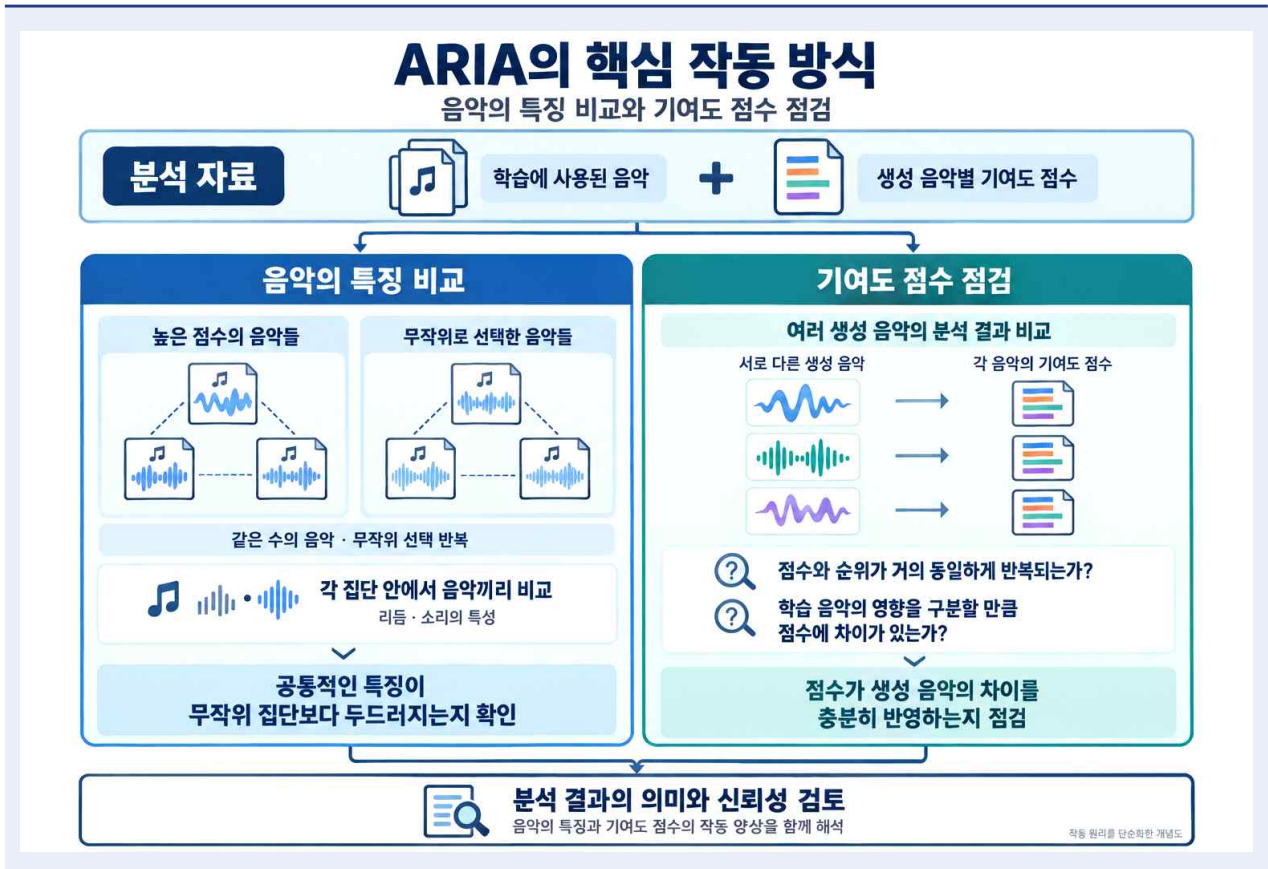
다음은 기여도 점수를 점검하는 과정이다. 서로 다른 여러 생성 음악을 분석했을 때 기여도 점수도 그 차이를 반영하는지 살펴보게 되는데, 학습에 사용된 음악의 점수와 순위가 서로 다른 여러 생성 음악에서도 거의 동일하게 반복된다면 기여도 점수가 생성 음악의 차이를 충분히 나타내지 못할 가능성이 있다는 것이다. 서로 다른 여러 생성 음악에 대한 점수들을 비교하여 이러한 반복이 나타나는지, 학습에 사용된 음악의 영향을 구분할 만큼 점수에 차이가 있는지 확인하는 방식이다. 예를 들어, 특정 생성 음악을 분석했을 때 학습에 사용된 음악들이 각각 1위 90점, 2위 80점, 3위 70점을 받았다고 가정해 보자. 특정 생성 음악이 아닌 다른 생성 음악을 분석했을 때도 학습에 사용된 음악들이 앞서서와 동일하게 1위 90점, 2위 80점,

4) Woosung Choi 외 9인, “Large-Scale Training Data Attribution for Music Generative Models via Unlearning”, arXiv, 2025.10.07.

5) Changheon Han 외 2인, “ARIA: A Diagnostic Framework for Music Training Data Attribution”, arXiv, 2026.05.15.

3위 70점으로 나타난다면, 기여도 점수가 서로 다른 생성 음악 간의 차이를 유의미하게 나타내지 못하기 때문에 이와 같은 결과는 신뢰하기 어렵다는 뜻이다. ARIA는 이처럼 음악의 특징과 기여도 점수의 작동 양상을 함께 살펴보며, 분석 결과의 의미와 신뢰성을 검토하도록 돕는다.

| 학습 데이터 기여도 분석 프레임워크 ‘ARIA’



※ 출처: 참고문헌을 종합하여 ChatGPT로 재구성

5. 시사점

최근 생성형 인공지능을 활용한 음악이 국내 음원 차트에 진입하는 등 대중의 관심을 받고 있다. 음악 창작 과정에서 인공지능을 활용하는 방식도 다양해지면서, 사람이 창작한 부분을 어떻게 인정하고 저작권료를 분배할 것인지에 대한 논의가 이어지고 있다. 국내에서도 관련 토론회를 통해 창작자와 음악산업계 등의 의견을 수렴하며, 생성형 인공지능 활용 음악의 신탁 기준을 마련하기 위한 논의가 진행되고 있다.

이와 함께 인공지능의 학습에 사용된 음악과 새롭게 생성된 음악 사이의 관계를 파악하기 위한 연구도 이루어지고 있다. 학습 데이터 기여도 분석을 통해 개별 음악이 생성 결과에 미치는 영향을 추정하고, 'ARIA'와 같은 프레임워크를 통해 분석 결과의 의미와 신뢰성을 검토하려는 시도다. 앞으로 생성형 인공지능 활용 음악을 둘러싼 논의와 함께 학습 데이터 기여도 분석에 관한 연구가 어떻게 발전해 나갈지 살펴볼 필요가 있다.

| 참고자료

- 연합뉴스, “AI 활용곡이 멜론 톱100…조선희합 “저비용으로 세대 공감 가능””, 2026.09.13., <https://www.yna.co.kr/view/AKR20260912026000005?input=1195m>
- 한국콘텐츠진흥원, “‘Made by AI 음악’과 공존하기: 생성형 AI시대, 음악산업의 전환과 정책 방향”, 2025.06.26.
- 한국음악저작권협회 공지사항, “AI 활용 음악저작물 신고·등록 기준 철회 안내”, 2026.08.26., <https://www.komca.or.kr/CTLJSP>
- 문화체육관광부 보도자료, “생성형 인공지능 활용 음악의 저작권 신탁·분배 기준 논의”, 2026.09.21., https://www.mcst.go.kr/site/s_notice/press/pressView.jsp?pSeq=22735
- 한국음악저작권협회 공지사항, “[AI 활용 음악 신탁 기준 마련을 위한 토론회] 개최 안내”, 2026.09.18.
- Woosung Choi 외 9인, “Large-Scale Training Data Attribution for Music Generative Models via Unlearning”, arXiv, 2025.10.07.
- Changheon Han 외 2인, “ARIA: A Diagnostic Framework for Music Training Data Attribution”, arXiv, 2026.05.15.