

AI 내레이션 확산과 오디오 제작 패러다임 전환

[목차]

1. AI 음성 기술과 오디오 콘텐츠 제작 혁신
2. 국내 오디오 플랫폼의 대응: 팟빵과 네이버
3. 해외 오디오 플랫폼의 전략: 스포티파이와 유튜브
4. 오디오 플랫폼 비즈니스 모델에 미치는 영향
5. 전망 및 시사점 - 오디오 콘텐츠 진흥을 위한 정책 방향



요약문

“

오디오 콘텐츠 시장은 생성 AI와 보이스 클로닝 기술의 발전으로 제작과 유통 방식에 큰 변화가 나타나고 있다. AI 음성 합성으로 내레이션이 자동화되고, 다국어 더빙과 실시간 음성 변환 기술로 글로벌 유통이 가속화되고 있다. 국내에서는 팟빵이 유튜브로 청취자가 이탈하는 문제에 직면했으며, 네이버는 AI 기반 뉴스 요약·음성 서비스로 대응 중이다. 해외에서는 스포티파이가 AI 번역과 DJ 기능을, 유튜브는 ‘보는 팟캐스트’와 AI 영상 자동 생성 도구를 통해 시장 주도권을 확대하고 있다. 이러한 변화는 제작비 절감, 콘텐츠 과잉, 광고 모델 재편, 초개인화 청취 환경의 확산을 불러오고 있으며, 이에 따라 AI 기반 콘텐츠 제작 지원, 음성 권리 보호, 국내 기술 투자, 공정 경쟁 생태계 조성 등이 정책 과제로 부각되고 있다.

”

오디오 콘텐츠 시장은 최근 생성 AI와 보이스 클로닝 기술¹⁾ 그리고 온디바이스 음성 기술의 급격한 발전으로 제작과 유통 측면에서 커다란 변화를 겪고 있다. 사람의 목소리를 흉내 내는 AI 내레이션이 오디오북, 팟캐스트, 뉴스 등 다양한 콘텐츠에 도입되고 있고, 다국어 더빙 및 실시간 음성 변환 서비스도 등장하며 오디오 콘텐츠의 언어 장벽을 허물고 있다. 국내외 주요 플랫폼들은 이러한 AI 음성 기술을 적극 상용화하고 있으며, 오디오 콘텐츠 유통 패러다임 또한 빠르게 변화하고 있다. 특히 한국과 미국을 중심으로, 팟캐스트 콘텐츠의 유튜브 플랫폼 이탈 가속화 등 새로운 흐름이 나타나고 있어, 이에 따른 산업 구조와 비즈니스 모델의 변화가 주목된다. 본 글은, 국내 플랫폼인 팟빵과 네이

1. 보이스 클로닝(Voice Cloning, Voice Synthesis)은, 짧게는 수십 소부터 길게는 몇 분의 녹음으로 특정 화자의 음색, 억양, 발음 습관을 학습해 그 사람처럼 말하는 오디오를 생성하는 AI 음성 합성 기술을 말한다.

버, 해외 플랫폼인 스포티파이(Spotify)와 유튜브(YouTube)를 중심으로 이러한 동향을 살펴보고 있으며, 이에 기초하여 향후 시장 전망과 정책적 시사점을 제시하고 있다.



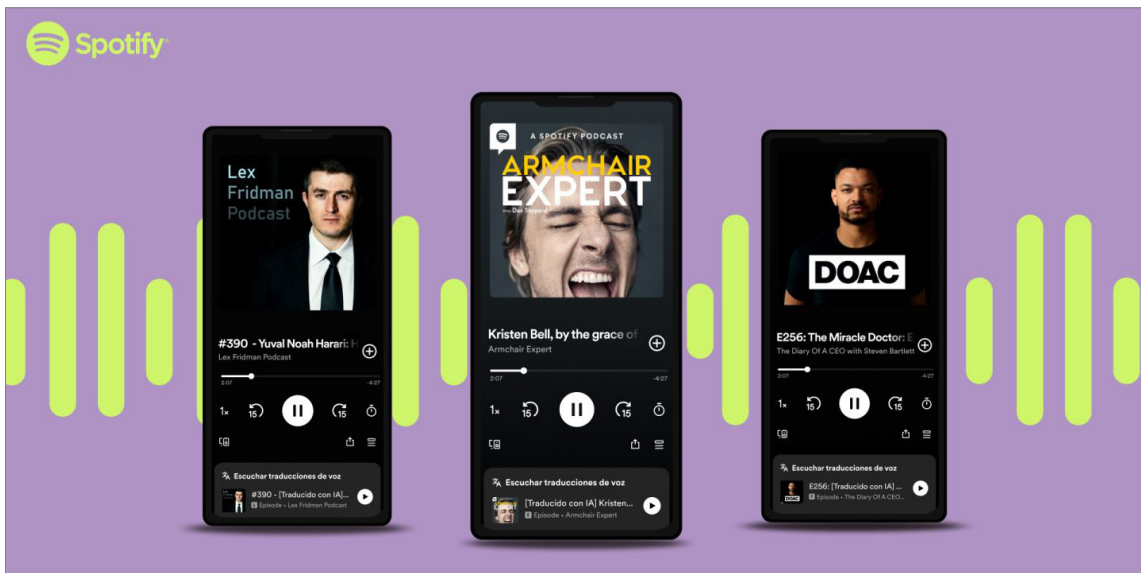
AI 음성 기술과 오디오 콘텐츠 제작 혁신

생성 AI 기반 음성합성(TTS) 기술²⁾의 진보로, 과거에 사람의 녹음이 필수였던 내레이션 작업이 상당 부분 자동화되고 있다. 딥러닝을 활용한 최신 TTS 엔진은 인간과 거의 구별되지 않을 정도로 자연스러운 음성을 만들어낼 수 있으며, 이를 통해 오디오북이나 뉴스 기사 낭독 등의 제작 효율이 크게 향상되고 있다. 예를 들어 해외에서는 MIT와 마이크로소프트가 프로젝트 구텐베르크와 협력하여 5,000여 종의 오디오북을 AI 음성으로 출판하였고,³⁾ 국내에서도 AI 목소리를 활용한 오디오북 출시가 활발하다. AI 음성 기술을 사용하면 전문 성우를 고용하지 않고도 빠르게 오디오북을 제작할 수 있어, 특히 소규모 출판사나 독립 제작자에게 제작 비용을 획기적으로 절감할 기회를 제공된다. 이러한 제작 자동화는 팟캐스트 편집에도 적용되어, AI로 녹음 음원의 잡음 제거나 편집, 요약본 작성 등이 가능해지고 있다.

나아가 다국어 더빙과 실시간 음성 변환 기술의 등장은 오디오 콘텐츠의 글로벌 유통을 가속화하고 있다. 한 언어로 녹음된 음성을 다른 언어로 자동 변환하면서도 원래 화자의 목소리 톤과 억양을 유지하는 서비스들이 출시되고 있다. 스포티파이는 2023년 오픈AI(OpenAI)와의 협업을 통해 팟캐스트를 외국어로 자동 번역하여 원래 진행자의 목소리로 재생해 주는 AI 기능을 시범 도입했다.⁴⁾ 예컨대 영어 팟캐스트를 스페인어나 프랑스어로 AI가 번역하면서도 진행자 본인의 음색으로 들려주는 것으로, 전통적 더빙보다 훨씬 개인적이고 자연스러운 청취 경험을 제공한다는 평가를 받고 있다. 이 기능을 통해 렉스 프리드먼(Lex Fridman), 닥스 셰퍼드(Dax Shepard) 등의 인기 팟캐스트 에피소드를 스포티파이 사용자가 자막 없이 자국어 음성으로 청취할 수 있게 되었다.⁵⁾

유튜브 역시 2024년 말부터 AI를 활용한 자동 다국어 더빙 기능을 도입하고 있다. 유튜브 영상에 대해 클릭 한 번으로 AI가 실시간에 가깝게 영어 등 다른 언어 오디오 트랙을 추가해 주는 것으로,⁶⁾ 2025년 9월부터는 유튜브 파트너 채널이라면 과거 업로드된 영상까지 대부분 자동 번역 음성 트

〈그림 1〉 스포티파이, 외국어 자동 번역 AI 기능 도입



출처: 스포티파이

락을 제공할 수 있게 범위가 확대되고 있다. 실제로 네덜란드 등의 일부 유튜브 채널은 국회 토론 영상을 AI 음성으로 영어 더빙하여 올리는 등, AI 더빙을 통해 콘텐츠 도달 범위를 넓히는 사례도 존재한다. 이처럼 실시간 음성 변환 기술이 상용화 단계에 접어들면서, 앞으로는 한국어 팟캐스트를 AI가 즉시 영어 등으로 변환하거나 반대로 해외 콘텐츠를 한국어 음성으로 실시간 즐기는 것이 일반화될 전망이다.

또한 온디바이스(On-device) 음성 기술의 발전은 스마트폰과 차량 등에서 인터넷 연결 없이도 고품질 음성 합성이나 음성 인식이 가능하게 해주고 있다. 예컨대 네이버는 복잡한 긴 문장을 높은 정확도로 인식하는 음성인식 기술 ‘네스트(Nest)’를 2020년에 공개하며, 대용량 학습 없이도 장문의 음성을

2. 생성 AI 기반 TTS(Text-to-Speech)는 딥러닝 모델, 특히 거대언어모델을 활용하여 입력된 텍스트를 인간과 매우 유사하고 자연스러운 음성으로 변환하는 기술을 말한다.
3. Greg Newby. (2023.07.18). Project Gutenberg brings free audiobooks to the global community with AI-generated book narration. Microsoft.
4. Amrita Khalid. (2023.09.25). Spotify is going to clone podcasters' voices — and translate them to other languages. The Verge.
5. Spotify. (2023.09.25). Spotify's AI Voice Translation Pilot Means Your Favorite Podcasters Might Be Heard in Your Native Language.
6. Defend Democracy. (2025.10.02). YouTube decides who gets AI voice.

텍스트로 정확히 변환할 수 있다고 발표했다.”⁷⁾ 온디바이스 음성 기술은 사용자의 프라이버시를 보호하면서 빠른 반응 속도로 오디오 비서 서비스를 구현할 수 있어, 차량용 내비게이션 음성 안내, 스마트 스피커 콘텐츠 재생, 오프라인 팟캐스트 청취 등 다양한 현장 기반 오디오 활용을 가능하게 한다. 즉 AI 음성 기술이 클라우드 서버뿐 아니라 디바이스 내부에서도 원활히 동작하게 되면서, 언제 어디서나 음성으로 콘텐츠를 소비하는 환경이 구축되고 있다.

요약하면, AI 내레이션과 음성 합성 기술의 발전은 오디오 콘텐츠 제작비용과 시간을 대폭 줄이고 다국어 제공을 손쉽게 만들며, 온디바이스 음성 기술은 사용자가 모바일 환경에서 끊임 없이 음성 콘텐츠를 소비하도록 돕고 있다. 이러한 오디오 AI 기술 발전은 오디오 플랫폼들이 과거에는 상상하기 어려웠던 새로운 서비스 모델을 내놓을 수 있는 기술적 토대가 되고 있다.



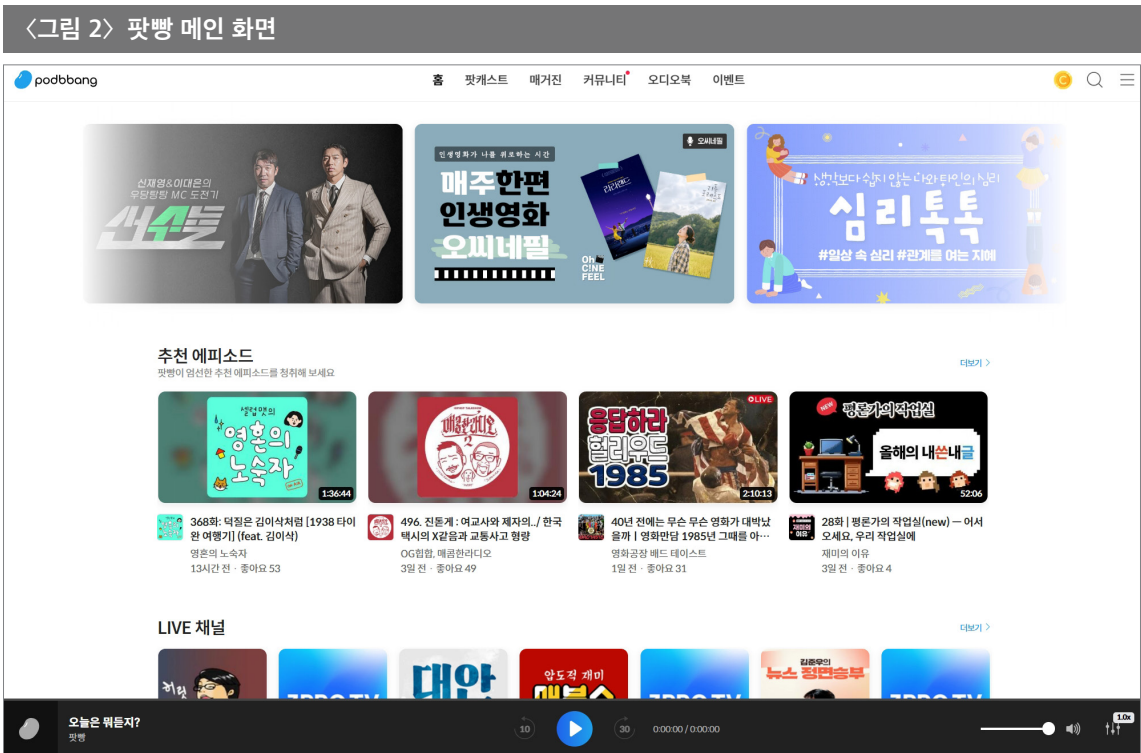
국내 오디오 플랫폼의 대응: 팟빵과 네이버

2-1. 팟빵(Podbbang)

국내 오디오 플랫폼 팟빵(Podbbang)은 2012년 등장 이후 한국 팟캐스트 붐을 주도해 온 개척자다. 정치 시사부터 코미디, 어학, 오디오 드라마에 이르기까지 다양한 분야의 팟캐스트를 호스팅하며, 국내 팟캐스트 시장의 70% 이상을 점유하기도 했다. 2018년 기준 팟빵에는 하루 순방문자 약 40만 명, 월간 300만 명 수준의 이용자가 있었고, 13,000여 개 방송에 150만 건이 넘는 에피소드가 축적되는 등 규모 면에서 기존 라디오 방송에 필적하는 성장을 이루었다.⁸⁾

그러나 최근 팟빵은 새로운 도전에 직면해 있다. 유튜브 등 거대 글로벌 플랫폼으로 팟캐스트 청취가 이동하는 현상이 뚜렷해지고 있기 때문이다. 팟빵이 성장하던 시기에는 팟캐스트가 오디오 전용 콘텐츠로 소비되었으나, 이제 많은 청취자가 팟캐스트를 영상과 함께 시청하거나, 아예 유튜브에서 오디오 콘텐츠를 찾는 경향을 보인다.

실제 미국에서는 유튜브가 팟캐스트 청취 1위 플랫폼으로 성장하고 있다. 2024년 조사에서 주간 팟캐스트 이용자의 34%가 ‘가장 많이 쓰는 플랫폼’으로 유튜브를 꼽아 스포티파이(17%)나 애플 팟캐스트(11%)를 크게 앞질렀다.⁹⁾ 한국도 인기 팟캐스트 진행자들이 유튜브 채널을 개설해 영상 팟캐스



출처 : 팟빵 홈페이지

트를 제공하거나, 아예 팟캐스트 내용을 요약 편집한 영상 클립을 쇼츠(Shorts) 등으로 배포하여 홍보 및 수익 창출을 도모하는 사례가 늘었다고 있다. 이에 따라 전통 팟캐스트 앱의 이용률이 정체되고 있으며, 팟빵으로서는 콘텐츠 이탈과 청취자 분산이라는 과제를 마주하고 있다.

팟빵은 이런 변화에 대응하기 위해 자체 오리지널 콘텐츠 강화와 크리에이터 수익 지원에 힘쓰고 있다. 인기 팟캐스트에 투자하여 독점 에피소드를 제공하거나, 후원 결제 시스템을 통해 충성 청취자들이 직접 크리에이터를 지원할 수 있는 환경을 만들고 있다. 또한 AI 기술 도입 가능성을 모색 중이다. 그러나 아직까지 팟빵이 직접적인 AI 내레이션 서비스를 내놓았다는 소식은 없다. 팟빵의 플랫폼 경쟁력 유지를 위한 기술 투자와 동시에 유튜브 등 외부 채널과의 협업도 필요한 상황이다. 팟빵의 사례는 국내 로컬 오디오 플랫폼이 글로벌 거대 플랫폼의 공세 속에서 차별화와 혁신을 추구하기가 얼

7. 백봉상. (2020.04.13). 네이버, 더 똑똑해진 음성인식 기술 ‘NEST’ 공개. ZDNET.
8. 팟빵. (2018). 오디오 플랫폼은 어떻게 살아남아야 하는가. 방송트렌드&인사이트 Vol.15.
9. Elizabeth Mayer. (2024.12.02). YouTube Expands Its Lead As The #1 Platform For Podcasts, According To Brand New Cumulus Media And Signal Hill Insights' Podcast Download - Fall 2024 Report. Westwood One Audio Active Group.

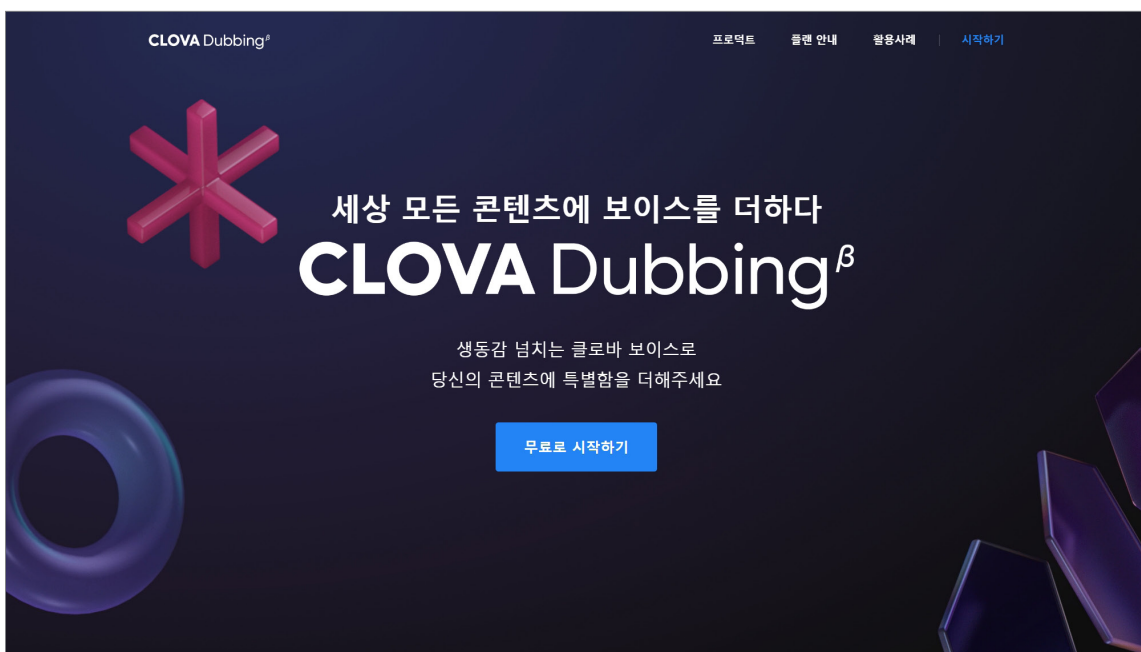
마나 어려움을 보여주는 한 단면이라 할 수 있다.

2-2. 네이버

네이버는 자사의 거대 플랫폼 생태계 안에서 오디오 콘텐츠 분야를 꾸준히 육성해 왔다. 네이버는 AI 기술을 접목한 오디오 서비스 개발에 적극적이다. 텍스트 기사를 음성으로 읽어주는 기능을 일찍이 도입한 바 있다. 네이버 뉴스는 이미 2016년부터 기사 본문을 TTS 기술로 읽어주는 ‘본문 듣기’ 기능을 제공했고, 2020년에는 주요 뉴스를 AI가 요약하고 음성으로 읽어주는 ‘보이스 뉴스’ 베타 서비스를 선보였다.

네이버는 음성변환(TTS) 기술과 자동 요약 AI를 결합해 바쁜 이용자들이 이동 중에 핵심 뉴스만 청취할 수 있도록 함으로써, 오디오 콘텐츠가 새로운 뉴스 소비 채널로 자리 잡도록 실험한 바 있다. 또한 네이버는 오디오클립(AudioClip)이라는 전용 애플리케이션을 통해 팟캐스트, 오디오북, 어학, 동화 등 다양한 오디오 콘텐츠 플랫폼도 운영해 왔다. 오디오클립에서는 네이버 웹툰, 웹소설 IP를 활용한 ‘오디오 시네마’를 공개하여, 인기 웹툰을 원작으로 한 오디오 드라마를 유명 배우의 목소리 연기와 음향 효과로 제작해 호평받은 바 있다. 이처럼 네이버는 기존 텍스트 및 영상 콘텐츠를 오디오로 확

〈그림 3〉 클로바 더빙 첫 화면



출처 : 클로바 더빙

장하여 멀티유즈함으로써 부가가치를 창출하는 전략도 구사하고 있다.

나아가 네이버는 AI 보이스 클로닝과 더빙 기술도 서비스 형태로 제공 중이다. 클로바 더빙(Clova Dubbing)이라는 플랫폼을 통해 사용자가 원하는 목소리 스타일을 선택하고 텍스트를 입력하면 AI 성우 음성으로 더빙 오디오를 생성할 수 있다.¹⁰⁾ 2022년 기준 가입자 100만 명을 돌파할 정도로 많은 크리에이터들이 이 인공지능 성우 서비스를 활용하고 있으며,¹¹⁾ 이를 통해 다수 콘텐츠 창작자는 유튜브 영상 내레이션, 광고 스팟, 교육 자료 오디오 등을 제작하고 있다.



해외 오디오 플랫폼의 전략: 스포티파이와 유튜브

3-1. 스포티파이

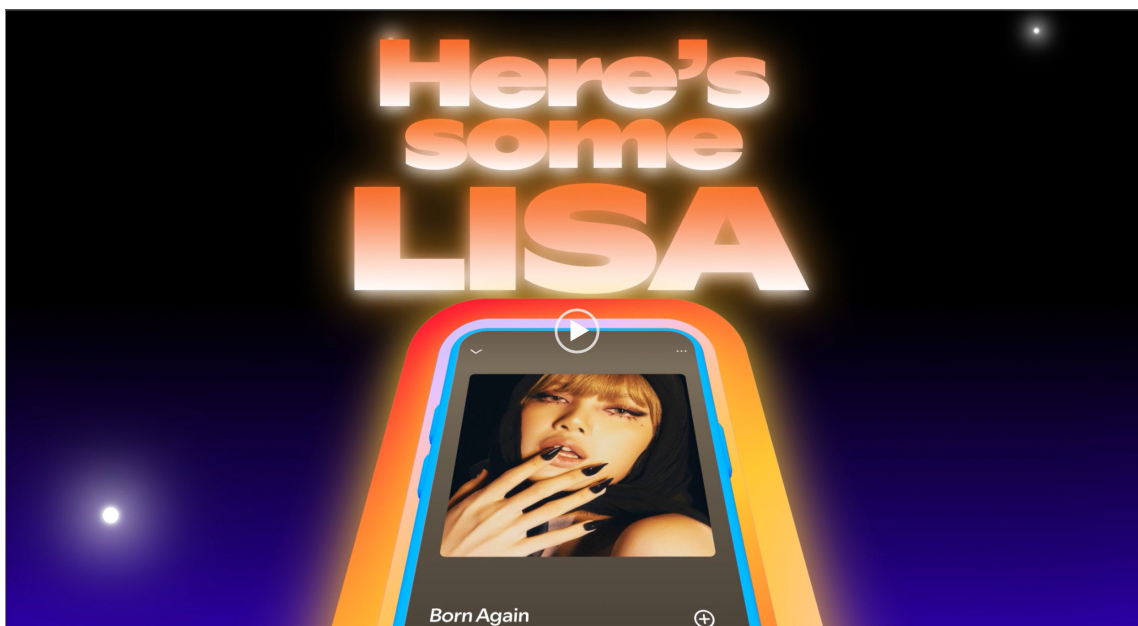
글로벌 오디오 스트리밍의 강자 스포티파이(Spotify)는 음악에 이어 팟캐스트와 오디오북 분야에서 공격적인 투자를 펼치며 오디오 플랫폼의 범위를 확장해 왔다. 2019년부터 팟캐스트 제작사 및 조로건 등 인기 진행자와의 독점 계약을 체결하여 자체 오리지널 팟캐스트 생태계를 구축해 오고 있다. 그 결과 현재 전 세계적으로 1억 명 이상의 이용자가 정기적으로 스포티파이에서 팟캐스트를 청취하는 규모로 성장했다.¹²⁾

이러한 팟캐스트 중심 전략을 더 강화하기 위해, 스포티파이는 AI 기술을 활용한 차별화 기능을 속속 선보이고 있다. 팟캐스트 외국어 자동 번역 서비스는 그 대표적인 사례로, 2023년 9월 발표된 이 기능은 오픈AI(OpenAI)의 최신 음성생성 기술을 활용해 팟캐스트 진행자의 음성을 그대로 복제한 외국어 버전 에피소드를 제공하기 시작했다. 현재까지 렉스 프리드먼 쇼 등의 일부 영어 팟캐스트를 스페인어, 프랑스어, 독일어 버전으로 시범 제공하였으며, 청취자들은 설정만 바꾸면 동일한 목소리로 다른 언어를 들을 수 있는 혁신적인 경험을 하고 있다.

10. 클로바 더빙

11. 이효석. (2021.02.08). "성우 대신 AI 더빙"...네이버 '클로바 더빙' 유료 상품 출시. 연합뉴스.

12. Spotify. (2023.09.25). Spotify's AI Voice Translation Pilot Means Your Favorite Podcasters Might Be Heard in Your Native Language.



출처 : 스포티파이

스포티파이는 “AI를 신중하게 활용하면 청취자와 크리에이터의 연결을 더욱 깊게 만들어줄 것”이라며 이러한 개인화된 다국어 청취 경험이 가져올 장점을 강조하고 있다.¹³⁾ 향후 더 많은 팟캐스트와 언어로 확대될 이 기능은 스포티파이가 글로벌 오디오 플랫폼으로서 청취층을 넓히고, 크리에이터에게는 새로운 시장을 제공하는 역할을 할 것으로 기대된다.

스포티파이는 또한 AI DJ 기능을 도입하여 개인화 오디오 경험을 한층 강화하고 있다. 2023년 초 출시된 ‘DJ’ 기능은 스포티파이의 추천 알고리즘과 AI 보이스 기술을 결합한 것으로, 사용자의 음악 취향에 맞춰 선곡해 줄 뿐 아니라 라디오 DJ처럼 곡 소개 멘트를 AI 음성이 곁들여주는 서비스다. 스포티파이는 2024년 DJ 기능을 스페인어 등 다른 언어권으로 확대하고, 2025년에는 사용자가 음성으로 DJ에게 노래를 요청할 수 있는 양방향 기능도 선보이고 있다.¹⁴⁾

이러한 시도들은 AI가 단순히 콘텐츠를 전달하는 것을 넘어 이용자와 상호작용을 하며 큐레이션 해주는 역할까지 수행하게 되었음을 의미한다. 스포티파이는 이 외에도 오디오북 시장에도 진입하여 2022년 미국에서 30만 종 이상의 오디오북 카탈로그를 선보였고, AI 내레이션 오디오북에 대한 투자도 진행하고 있다. 2022년에는 음성 합성 스타트업 소네틱(Sonantic)을 인수하여 향후 오디오북 자동 내레이션이나 개인화 음성 비서 등 영역에 AI 음성을 활용할 것으로 전망된다.¹⁵⁾ 이는 오디오 플

랫폼 산업의 패러다임이 단순 호스팅에서 AI 기반 개인화 서비스로 전환되고 있음을 보여주는 대표적인 사례로 평가할 수 있다.

3-2. 유튜브

유튜브는 동영상 플랫폼이지만, 최근 몇 년 사이 가장 영향력 있는 팟캐스트 플랫폼으로 부상하여 오디오 시장 지형을 뒤흔들고 있다. 2024년 미국 소비자 조사에서 유튜브는 팟캐스트 소비 1위 채널(34%)로 선정되었으며, 특히 ‘보는 팟캐스트’ 트렌드를 주도하고 있다.¹⁶⁾

실제로 주간 팟캐스트 이용자의 1/3 이상이 “가능하면 영상이 함께 나오는 팟캐스트를 보고 싶다”라고 응답할 만큼, 영상 콘텐츠에 익숙한 젊은 세대를 중심으로 팟캐스트의 비디오화 선호도가 높아지고 있다. 또 다른 2024년 미국 소비자 조사에 따르면, 미국 성인 중 42%가 팟캐스트를 오디오만이 아닌 영상 포함 형태로 소비하기를 선호하며, 이 비율은 2022년(32%) 대비 크게 상승한 수치다.¹⁷⁾ 이러한 추세 덕분에 유튜브에는 수많은 팟캐스트 채널이 생겨나고 있으며, 인기 팟캐스트 진행자들은 스튜디오 촬영을 통해 토크쇼 형태의 영상 에피소드를 올리거나, 녹음 음성에 파형 애니메이션이나 이미지를 입힌 영상을 업로드하고 있다. 국내에서도 많은 팟캐스트 제작자가 유튜브 채널을 운영하며, 오디오 콘텐츠를 영상으로 병행 제공함으로써 조회수 기반 광고 수익과 신규 청취자 확보라는 두 마리 토끼를 잡고 있다.

유튜브는 이러한 오디오 크리에이터들의 영상 진입을 돕고 시장을 선점하기 위해, 2023년부터 팟캐스트 지원 전략을 본격화했다. 구글은 2024년 초 전용 팟캐스트 앱인 구글 팟캐스트(Google Podcasts)를 종료하고 유튜브 뮤직으로 팟캐스트 재생 기능을 일원화한다고 발표하여,¹⁸⁾ 유튜브를 팟캐스트 청취의 허브로 통합하는 움직임을 보여왔다. 무엇보다 주목할 것은, 유튜브가 AI 기술을 활용해 오디오

13. Spotify. (2023.09.25). Spotify's AI Voice Translation Pilot Means Your Favorite Podcasters Might Be Heard in Your Native Language.

14. Spotify. (2025.05.13). Spotify's DJ Now Takes Requests, Enhancing Real-Time Music Discovery.

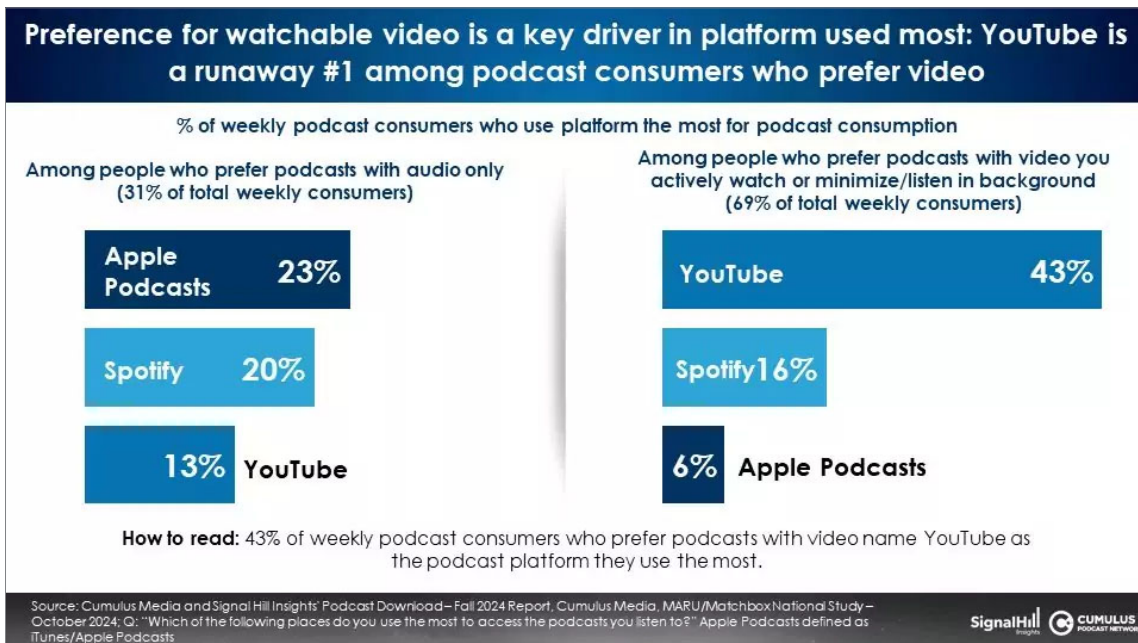
15. Ingrid Lunden. (2022.06.13). Spotify is acquiring Sonantic, the AI voice platform used to simulate Val Kilmer's voice in 'Top Gun: Maverick'. Tech Crunch.

16. Elizabeth Mayer. (2024.12.02). YouTube Expands Its Lead As The #1 Platform For Podcasts, According To Brand New Cumulus Media And Signal Hill Insights' Podcast Download - Fall 2024 Report. Westwood One Audio Active Group.

17. Dan Whateley. (2025.09.17). YouTube has a new strategy to win over Spotify's audio-first podcasters using AI. BUSINESS INSIDER.

18. Sarah Perez. (2023.09.26). Google Podcasts to shut down in 2024 with listeners migrated to YouTube Music. TechCrunch.

〈그림 5〉 팟캐스트 시청 경험을 선호하는 사람들이 선택하는 플랫폼



출처 : Westwood One Audio Active Group

오 전용 팟캐스트까지 영상화하려고 시도하고 있다는 점이다.

2025년 9월 유튜브는 크리에이터 행사(Made on YouTube)에서 “오디오 팟캐스트를 자동으로 영상 클립으로 만들어 주는 AI 도구” 출시를 예고했다.¹⁹⁾ 이 AI 도구를 이용할 경우, 팟캐스트 오디오의 스크립트를 분석해 30~60초 길이의 짧은 하이라이트 영상을 자동 생성할 수 있으며, 베오(Veo)라는 생성 AI 동영상 기술을 활용하여 간단한 관련 이미지를 입히거나 자막 효과를 준 영상을 만들 수 있다. 이 AI 도구는 팟캐스트 진행자들이 촬영 장비나 편집 기술이 없어도 한두 번 클릭만으로 홍보용 쇼츠 영상을 뽑아낼 수 있게 해주는 도구로서, 우선 미국의 일부 팟캐스터들을 대상으로 2026년에 시범 제공될 예정이다.

유튜브는 “영상 중심 플랫폼에 익숙하지 않은 순수 오디오 크리에이터들도 손쉽게 유튜브에 진출하도록 돕는 것”이 이 기능의 목적이라고 밝히고 있다. 이러한 AI 영상화 도구를 통해 유튜브는 그동안 자사 플랫폼에 소극적이었던 오디오 창작자층을 끌어들이며, 팟캐스트 분야에서도 경쟁 플랫폼을 압도하겠다는 전략으로 추진하고 있다. 유튜브는 또한 AI 다국어 더빙(멀티랭귀지 오디오) 기능을 도입하며 글로벌 오디오 소비 지원에 나서고 있다. 이 외에도 쇼츠(Shorts)와 알고리즘 추천을 통해 팟캐스트를 발견하고 홍보하는 데 최적화된 환경을 제공하고 있다. 2024년 진행된 미국 조사에 따르면²⁰⁾ 새

로운 팟캐스트 청취자의 44%가 유튜브에서 처음 그 팟캐스트를 접했다고 할 정도로, 유튜브는 팟캐스트 콘텐츠의 검색 및 발견 플랫폼으로도 작동하고 있다.

결국 유튜브의 등장과 성장은 팟캐스트를 비롯한 오디오 콘텐츠 분야의 경쟁 규칙을 바꾸어 놓고 있다. 팟캐스트는 더 이상 순수 오디오만이 아닌 멀티미디어 콘텐츠로 진화하고 있으며, 오디오 플랫폼 업계도 영상 플랫폼과 경합하거나 협력해야 하는 시대가 도래하고 있다.



오디오 플랫폼 비즈니스 모델에 미치는 영향

이렇듯 AI 내레이션 AI 기술 확산과 유튜브 등 플랫폼 변동은 오디오 산업의 비즈니스 모델 전반에 큰 변화를 초래하고 있다.

첫째, 오디오 콘텐츠 제작 비용 구조가 변하고 있다.

AI를 활용하면 콘텐츠 한 편을 만드는 한계비용이 많이 감소하기에, 팟캐스트나 오디오북의 공급이 폭발적으로 증가할 가능성이 크다. 이는 긍정적으로는 다양한 니치(niche) 콘텐츠의 등장과 창작 기회의 확대에 이어질 수 있지만, 동시에 콘텐츠 과잉으로 인한 경쟁 심화와 품질 관리 문제를 일으킬 수도 있다. 누구나 AI 성우와 생성 AI로 팟캐스트를 생산할 수 있게 되면, 플랫폼에서는 양질의 콘텐츠를 선별 추천하는 큐레이션 역량이 더욱 중요해진다. 이에 따라 스포티파이나 유튜브처럼 강력한 추천 알고리즘을 가진 대형 플랫폼에 이용자 쏠림이 가속할 우려도 존재한다.

반면 개별 크리에이터 입장에서는 AI 도구를 활용해 혼자서도 다국어 버전 제작, 영상 홍보물 제작 등을 할 수 있게 되어 콘텐츠 브랜딩과 수익화 기회는 확대될 것으로 보인다. 예를 들어 한국의 한 팟캐스트 제작자가 AI 번역과 음성합성을 활용해 동일한 팟캐스트를 영어, 일본어 등으로 동시 발행하

19. Dan Whateley. (2025.09.17). YouTube has a new strategy to win over Spotify's audio-first podcasters using AI. BUSINESS INSIDER.

20. Dan Whateley. (2025.09.17). YouTube has a new strategy to win over Spotify's audio-first podcasters using AI. BUSINESS INSIDER.

고 유튜브에 영상 클립까지 올린다면, 별도 '추가' 예산 없이도 전 세계를 상대로 오디오 콘텐츠를 유통하고 광고수익을 올릴 수 있는 시대가 올 수 있다.

둘째, 오디오 플랫폼 간 경쟁과 협력 구도도 변형되고 있다.

유튜브로 팟캐스트가 이동하는 현상은, 기존 팟캐스트 플랫폼(팟빵, 애플 팟캐스트 등)의 이용자 이탈과 광고 감소를 가져올 수 있다. 유튜브가 영상과 커뮤니티 기능을 무기로 팟캐스트 시장을 장악해 버리면, 다른 오디오 플랫폼들은 유튜브에 콘텐츠 공급자로 참여하는 방향을 고민해야 할 수도 있다. 예를 들어 팟빵이나 스포티파이가 인기 팟캐스트 영상 클립을 유튜브에 공식 채널로 올리고, 유튜브에서 유입된 이용자를 자사 플랫폼의 풀 버전 청취로 유도하는 하이브리드 전략을 펼칠 가능성도 존재한다.

실제 스포티파이는 2022년부터 일부 팟캐스트에 영상 팟캐스트 기능을 도입하여 유튜브와 흡사한 이용자 경험을 제공함과 동시에,²¹⁾ 자사 앱 내에서도 영상과 오디오를 자유롭게 전환하며 들을 수 있도록 하고 있다. 이는 경쟁 플랫폼의 강점을 수용하면서 이용자를 붙잡아두려는 노력으로 볼 수 있다. 앞으로 오디오-영상 경계가 흐려짐에 따라 음악 스트리밍 서비스와 동영상 플랫폼, 팟캐스트 앱 간에 치열한 경합과 제휴가 이어질 전망이다.

셋째, 광고 및 수익 모델에도 변화가 예상된다.

AI 내레이션을 활용한 대량 콘텐츠 생산으로 전체 청취 시간은 늘어날 수 있으나, 단위 콘텐츠당 가치가 하락할 수 있다. 플랫폼들은 이를 보완하기 위해 정교한 타겟 광고나 프리미엄 구독 모델을 강화할 것으로 전망된다. 예를 들어 AI를 통해 이용자별 맞춤 광고 멘트를 생성하여 같은 콘텐츠 상품이라도 각자 다른 내용으로 광고를 들려주는 식의 초개인화된 광고가 가능할 수 있다. 또한 AI 음성으로 자동 생성한 광고 스팟을 수천 개 팟캐스트 에피소드에 동시 삽입하는 등 대규모 인벤토리 운영도 쉬워질 수 있다.

한편으로 (인기) 크리에이터들은 자신의 AI 복제 음성을 활용해 일종의 (IP) 라이선싱 수익을 얻는 모델도 등장할 수 있다. 이러한 새로운 모델을 실현하려면 음성에 대한 권리 관리 체계가 뒷받침되어야 하며, AI로 복제된 목소리의 오남용(예: 음란 콘텐츠에 유명인 목소리를 합성하는 등) 문제가 불거지고 있

21. Sarah Perez. (2023.09.26). Google Podcasts to shut down in 2024 with listeners migrated to YouTube Music. TechCrunch.

어, 향후 법적 또는 윤리적 논의를 통해 음성 IP 자산의 보호와 활용 기준이 마련될 것으로 전망된다.

넷째, 이용자 경험과 충성도 측면에서도 변화가 예상된다.

AI 기술이 발전하면서 콘텐츠 현지화와 개인화 수준이 높아지면, 이용자들은 자신이 선호하는 목소리, 언어, 길이로 콘텐츠를 유연하게 소비할 수 있게 된다. 예를 들어 한 팟캐스트를 들을 때 AI로 진행자 목소리를 배우자나 친구의 목소리로 변환해서 듣는다든지, 아이를 위한 동화 오디오북을 부모 목소리 AI로 들려주는 것도 가능해질 수 있다. 이러한 초개인화 경험은 플랫폼 입장에서 강력한 사용자 잠금(lock-in) 효과를 낼 수 있다.

한편, 딥페이크 음성을 이용한 사기 등이 사회적 문제로 대두되면, 오디오 콘텐츠 전반에 대한 이용자 경계심이 커질 가능성도 있다. 따라서 플랫폼들은 AI 생성 음성임을 고지하고 품질을 관리하며, 신뢰성을 확보하기 위한 장치도 고민해야 한다. 반대로 기술을 잘 활용하면, 청각 장애인을 위한 자막 및 수어 영상 자동 생성이나 시각 장애인을 위한 모든 텍스트 콘텐츠의 음성 제공 등 포용적 서비스를 확대하여 사회적 가치와 이용층을 넓힐 수도 있다.

이렇게 AI 기술 혁신과 유튜브 등 새로운 유통 채널의 등장은 오디오 플랫폼 산업의 가치 사슬 전반에 걸쳐 지각변동을 일으키고 있다. 콘텐츠 기획-제작-유통-소비의 각 단계에서 비용 구조와 경쟁 포인트가 달라지고 있으며, 이는 정책 입안자와 업계 모두에게 새로운 과제를 안겨준다.



전망 및 시사점 - 오디오 콘텐츠 진흥을 위한 정책 방향

AI 내레이션과 음성 기술의 발전이 불러온 오디오 산업의 변혁은 앞으로도 속도를 더해갈 것으로 전망된다. 기술적으로는 거의 완벽하게 인간과 구분되지 않는 음성합성이 머지않아 현실화할 전망이며, AI는 감정 표현과 맥락 이해를 갖춘 자연스러운 내레이션으로 진화하여 오디오 콘텐츠의 몰입감을 극대화할 것이다. 또한 모든 영상과 텍스트 콘텐츠가 자동으로 오디오화 되는 환경이 도래하면, 오디오 콘텐츠의 범람과 경쟁 심화가 동시에 일어날 수 있다. 플랫폼 간에는 콘텐츠 확보전과 이용자 쟁탈전이 이어지겠지만, 한편으로 융합과 협력도 불가피하다. 이러한 변화 속에서 한국의 오디오

오 콘텐츠 산업이 지속 성장하고 글로벌 경쟁력을 갖추기 위해서는, 기술, 산업, 문화 측면에서 종합적인 진흥 전략이 필요하다.

① AI 기반 오디오 콘텐츠 제작 지원: 정부가 AI 기술을 활용한 오디오 콘텐츠 생산을 장려하기 위해 제작 지원금이나 플랫폼 협업 프로그램을 운영할 필요가 있다. 예를 들어 한국출판진흥원에서 진행하고 있는 오디오북 제작 지원사업처럼, 우수한 책이나 웹툰 등을 AI 내레이션으로 제작해 다국어로 배포하는 프로젝트를 추진할 수 있다. 이는 국내 창작물의 새로운 부가가치 창출과 K-콘텐츠의 해외 진출을 동시에 도모하는 효과가 가질 수 있다. 또한 팟캐스트 분야에서도 신규 창작자 발굴 및 육성 프로그램을 마련하여, AI 편집 도구 등을 무료 제공하고 교육함으로써 창작 생태계의 저변을 확대할 수 있다.

② 음성 데이터 윤리 및 권리 보호 체계 구축: AI 시대에 목소리는 새로운 자산이자 정체성의 일부가 되고 있다. 정책 당국은 성우와 방송인 등 음성 전문인의 권리가 침해되지 않도록 법제 정비를 검토할 필요가 있다. 예컨대 AI 학습용으로 음성 데이터를 활용할 때 당사자 동의를 의무화하고, 무단으로 타인의 음성을 합성하여 상업적으로 이용할 경우 제재하는 규정을 마련할 필요가 있다. 동시에 성우 등이 자신의 AI 음성을 합법적으로 거래하고 보호받을 수 있도록 표준 계약서나 라이선싱 체계를 만들어 음성 권리 시장을 육성하는 방안도 고려할 수 있다. 아울러 딥페이크 음성 범죄에 대비한 기술적 대응책 개발 지원과 대국민 인식 교육도 병행하여, 새로운 기술에 대한 사회적 신뢰 확보를 뒷받침해야 한다.

③ 국내 음성 AI 기술 투자 및 표준화: 현재 음성 합성 핵심 기술은 구글, 메타, 오픈AI 등 글로벌 기업 주도로 발전하고 있어 한국어 음성 합성이나 방언 지원 등에서 한계가 있을 수 있다. 우리말에 최적화된 AI 보이스 기술을 확보하기 위해 산학연 R&D 지원을 늘리고, 국내 기업들의 음성 AI 스타트업에 대한 투자 유인을 제공해야 한다. 또한 AI 음성의 품질과 윤리적 사용에 관한 표준을 국내에서 선제적으로 정립하면, 향후 국제 논의에서도 주도권을 가질 수 있다.

④ 오디오 플랫폼 산업의 공정 경쟁과 상생: 유튜브 등 거대 플랫폼의 영향력 증대 속에서 국내 오디오 플랫폼 및 콘텐츠 사업자가 설 자리를 잃지 않도록 공정한 경쟁 환경을 조성하는 것도 중요

하다. 한편으로는 국내 플랫폼들이 혁신할 수 있도록 규제 완화와 행정 지원을 하고, 다른 한편으로는 글로벌 플랫폼의 시장지배적 행위에 대한 모니터링을 강화하여 상생을 도모해야 한다. 예를 들어 공영방송 및 언론사의 오디오 콘텐츠 개방과 공유 촉진 등을 통해 콘텐츠 공급을 다변화하고, 국산 오디오 플랫폼과 차량용 인포테인먼트 연동 등을 추진하여 이용자 접점을 확대할 수 있다. 또한 창작자들이 한 플랫폼에 종속되지 않고 멀티 플랫폼으로 유통할 수 있게, 메타데이터 표준화나 수익 정산 투명화 등의 인프라적 지원도 검토할 필요가 있다.

- ⑤ 오디오 문화 진흥과 인재 양성: 오디오 콘텐츠에 대한 이용자 문화 형성과 전문 인력 양성에 투자할 필요가 있다. 독서진흥정책의 하나로 오디오북 활용을 학교와 도서관에서 장려하고, 고령층이나 시각장애인을 위한 오디오 정보 접근성 제고 사업을 확대하면, 더 많은 시민이 오디오 콘텐츠의 가치를 체감할 수 있다.

결론적으로 AI 내레이션과 보이스 클로닝 기술의 확산 그리고 유튜브로 대표되는 플랫폼 지형 변화는 오디오 산업에 위기이자 기회를 함께 가져오고 있다. 기술 혁신을 적극 수용하면서도 창작자와 청취자의 가치를 보호하는 지혜가 요구된다. 정책 당국은 산업계와 협력하여 오디오 콘텐츠를 미래 문화산업의 핵심축으로 육성할 수 있는 전략을 마련해야 한다. 이를 통해 한국의 오디오 플랫폼과 크리에이터들이 새로운 글로벌 파도 속에서 도약할 수 있다.

레퍼런스



참고문헌

- Greg Newby. (2023.07.18). Project Gutenberg brings free audiobooks to the global community with AI-generated book narration. Microsoft.
- Amrita Khalid. (2023.09.25). Spotify is going to clone podcasters' voices — and translate them to other languages. The Verge.
- Spotify. (2023.09.25). Spotify's AI Voice Translation Pilot Means Your Favorite Podcasters Might Be Heard in Your Native Language.
- Defend Democracy. (2025.10.02). YouTube decides who gets AI voice.
- 백봉삼. (2020.04.13). 네이버, 더 똑똑해진 음성인식 기술 'NEST' 공개. ZDNET.
- 팻빵. (2018). 오디오 플랫폼은 어떻게 살아남아야 하는가. 방송트렌드&인사이트 Vol.15.
- Elizabeth Mayer. (2024.12.02). YouTube Expands Its Lead As The #1 Platform For Podcasts, According To Brand New Cumulus Media And Signal Hill Insights' Podcast Download - Fall 2024 Report. Westwood One Audio Active Group.
- 이효석. (2021.02.08). "성우 대신 AI 더빙"...네이버 '클로바 더빙' 유료 상품 출시. 연합뉴스.
- Spotify. (2025.05.13). Spotify's DJ Now Takes Requests, Enhancing Real-Time Music Discovery.
- Ingrid Lunden. (2022.06.13). Spotify is acquiring Sonantic, the AI voice platform used to simulate Val Kilmer's voice in 'Top Gun: Maverick'. TechCrunch.
- Dan Whateley. (2025.09.17). YouTube has a new strategy to win over Spotify's audio-first podcasters using AI. BUSINESS INSIDER.
- Sarah Perez. (2023.09.26). Google Podcasts to shut down in 2024 with listeners migrated to YouTube Music. TechCrunch.



참고 사이트

- 클로바 더빙 <https://clovadubbing.naver.com/>

