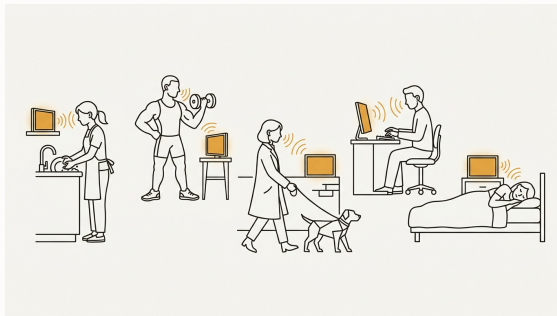




유튜브는 소리

59%는 화면을 보지 않습니다

강의 영상에서 설명한 내용을 그대로 옮긴 교재입니다

정원석 지음 · Connect AI LAB

차례

이 챕터에서

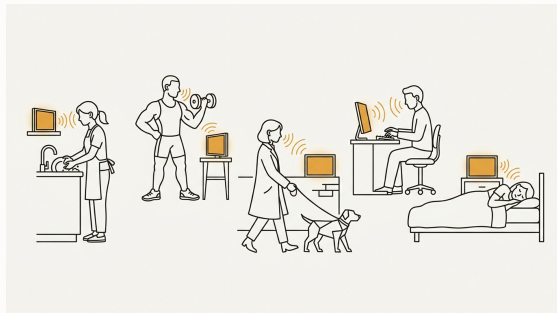
-
- 1 59%는 화면을 안 봅니다**
유튜브가 라디오가 되고 있습니다
 - 2 목소리 문제**
세 가지 길과 그 값
-

각 절 끝에 그 내용을 설명한 강의 영상이 붙어 있습니다. QR 을 찍으면 휴대폰에서도 열립니다.

1

59%는 화면을 안 봅니다

유튜브가 라디오가 되고 있습니다



문턱이 두 배 높아졌습니다

8월 11일 뉴스입니다. 유튜브 수익 창출 문턱이 두 배 높아집니다.

조건	기준
1년 누적 시청 시간	8,000시간
또는 쇼츠 조회수	2,000만 회

왜 이렇게 될까요. AI 로 만든 좋은 콘텐츠도 많지만 **그렇지 않은 것이 더 많아졌습니다.** 광고주가 화를 냅니다. 「내가 이만큼 돈을 내는데 이런 콘텐츠에 우리 광고를 붙이느냐」

그러면 구글의 광고 수익이 떨어집니다. 그래서 **난이도를 계속 올리고 있는 것**입니다.

유튜브는 중간에 있습니다

유튜브의 수익 구조는 단순합니다. **광고주와 크리에이터를 연결하는 매개체**입니다.

광고주가 1천만 원을 내면 그중 일부를 크리에이터에게 줍니다. 그래서 광고주가 만족하지 못하면 구조 전체가 흔들립니다.

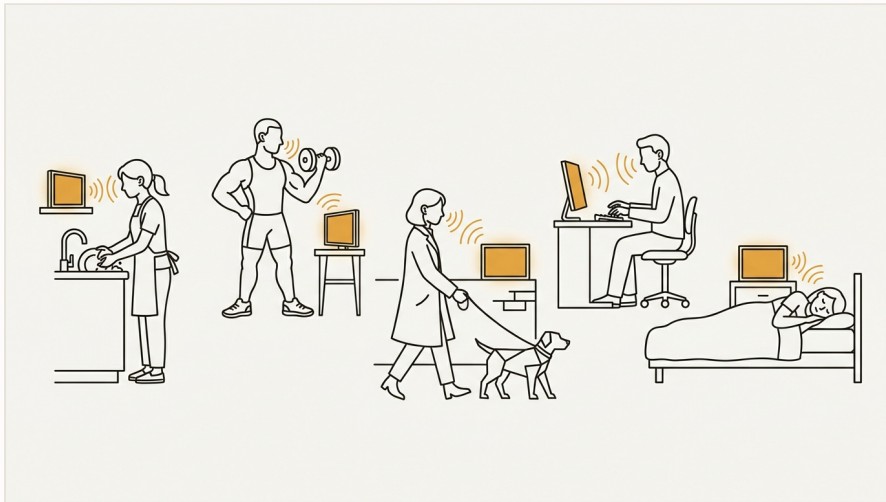
영상으로 승부하면 돈이 듭니다

쇼츠로 2,000만 회를 넘기려면 하나가 크게 터지거나 **엄청나게 많이 만들어야** 합니다.

직접 촬영하지 않고 AI 로 만들면 어떨까요. **지금은 너무 비쌉니다.** 1~2분짜리 하나에 10만 원씩 듭니다.
영상 생성 모델은 점점 더 비싸지고 있습니다.

그래서 지금 이건 좋은 전략이 아니라고 봅니다.

듣는 사람을 노립니다



유튜브 시청자는 두 부류로 나뉩니다. 보는 사람과 듣는 사람입니다.

	비율	어떤 사람
화면을 본다	41%	앉아서 공부하거나 시청하는 사람
화면을 안 본다	59%	일하며, 운동하며, 설거지하며 귀로만

영상으로 승부하려면 수천만 원에서 수억 원을 투자하겠다고 **결정**해야 합니다. 대부분은 그러기 어렵습니다.

그래서 오늘 제안은 이것입니다. 듣는 사람을 대상으로 만들어 봅시다.

팟캐스트가 커지고 있습니다

팟캐스트란 화면을 보지 않아도 소비할 수 있는 콘텐츠입니다. 전 세계 청취자가 계속 늘고 있습니다. 특히 미국에서 그렇습니다.

플랫폼	점유
유튜브	42% — 압도적
스포티파이 · 애플 팟캐스트	나머지

아예 팟캐스트 채널을 가져가셔도 괜찮습니다. 도전해 볼 만합니다.

지식은 초보여도 괜찮습니다. 다만 마음과 태도는 전문가여야 합니다.

영상으로 보기

유튜브는 「소리」 — 본강의

약 25분



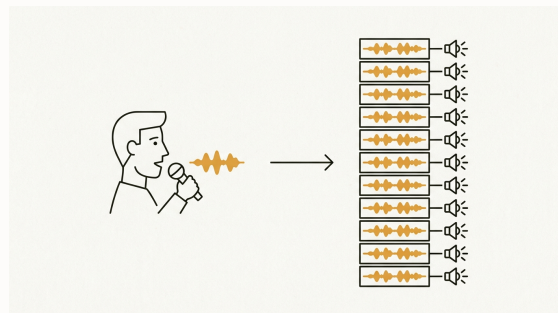
youtu.be/2pLwREltEjM

휴대폰으로 찍으면 바로 열립니다

2

목소리 문제

세 가지 길과 그 값



막히는 지점 세 개

1

직접 녹음한다

멋있게 나레이션하기가 어렵습니다. 대본을 읽다 틀리면 다시 해야 하고 시간이 오래 걸립니다.

2

AI 목소리를 쓴다

대본을 넣으면 바로 나옵니다. 그런데 개성이 없습니다. 「아, 이거 인공지능이네」 하고 넘어갑니다.

3

내 목소리를 학습시킨다

개성은 확실합니다. 그런데 비쌉니다.

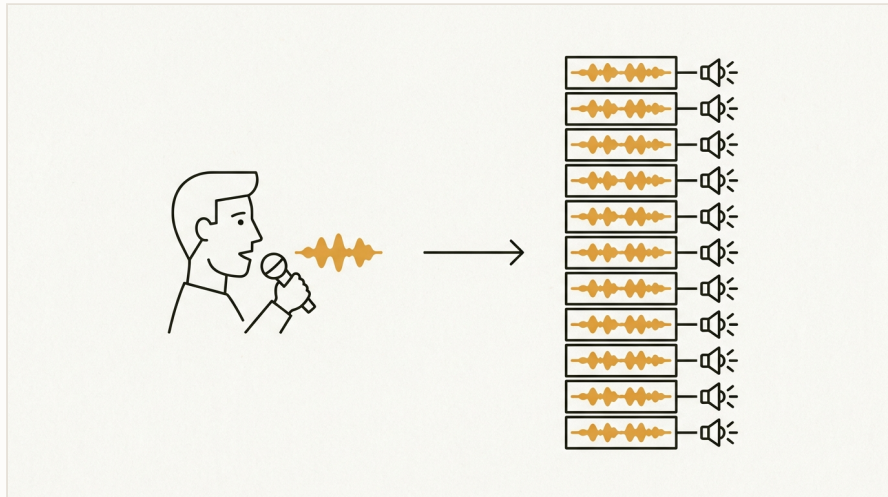
얼마나 비싼가

방식	비용
일레븐랩스 API 더빙	1분당 1,000원 수준
프로 보이스 클로닝 첫 달	약 18,000원
그다음 달부터	약 30,000원

게다가 돈을 낸다고 무제한이 아닙니다. **크레딧이 주어지고 쓸 때마다 깎입니다.** 그래서 불안하고, 많이 만들기 어렵습니다.

한 달에 30편을 만든다고 생각해 보세요. **만드는 데 돈을 다 씁니다.**

해결 — 로컬 TTS



오늘 쓸 것은 Qwen3-TTS 입니다. 로컬 AI 모델이고 **완전히 무료**입니다.

항목	내용
라이선스	Apache 2.0 — 만든 음성을 판매해도 됩니다
언어	10개 지원 — 글로벌 콘텐츠도 가능
필요한 녹음	짧게만 보내도 목소리 패턴을 어느 정도 익힙니다
한 번에	약 10분 분량까지
컴퓨터	좋지 않아도 됩니다. 코랩에서 해도 됩니다

일레븐랩스는 클로닝에 한 시간 분량을 요구하지만, 이쪽은 그렇지 않습니다.

이런 채널이 실제로 있습니다

구독자 2,500만 명이 넘는 채널을 열어 봤습니다. 그런데 영상 퀄리티가 그렇게 좋지 않습니다. 화면은 무료 영상 생성 모델로 만들 수 있는 수준입니다.

지표	값
6일 전 올린 콘텐츠	470만 뷰
댓글	13,590개

화면이 아니라 소리와 내용으로 굴러가는 채널입니다. 우리가 노릴 자리가 여기 있습니다.

영상으로 보기

내 목소리를 클론한 오디오 콘텐츠 — 무료 생성

약 30분 · 코랩



youtu.be/4JCW-jhnjTI
휴대폰으로 찍으면 바로 열립니다

다음

챕터 5 · 뉴스레터는 소유

알고리즘은 우연, 명단은 확신입니다

챕터 5 읽기

강의 목록으로 돌아가려면 [여기](#)를 누르세요.