



2024 방송기술대상 방송미디어융합상 수상작

AI를 활용한 콘텐츠 검색시스템 개발

글. 유현우 MBC 미디어IT센터 IT솔루션팀

AI 기술은 불과 몇 년 사이에 매우 빠른 속도로 진화하여 우리 일상뿐만 아니라 산업 전반에 활용되는 대중적인 기술로 자리 잡았습니다. MBC는 이러한 AI 기술을 방송 제작 및 콘텐츠 유통 전반에 활용하여 성과를 내고 있으며 그중 하나인 AI를 활용한 콘텐츠 검색시스템을 소개해드립니다.

개발 배경

이미지 검색에 대한 꾸준한 수요

최근 구작 콘텐츠를 활용한 SNS용 매시업 콘텐츠를 제작하는 것이 하나의 트렌드로 자리 잡고 있습니다. 이러한 콘텐츠를 제작하기 위해선 유명한 장면(짤방)이 언제 방영했는지, 해당 방영분에서 어느 시점에 나왔는지 찾는 것이 중요합니다. 따라서 이미지를 이용하여 콘텐츠를 검색하는 수요는 꾸준히 있었습니다.



그림 1. MBC의 대표 매시업 콘텐츠 '오분순삭'

근데 이제 'AI'를 곁들인

과거의 이미지 검색은 영상처리 기술을 활용하여 특징점을 추출하고 이미지 간의 유사도를 비교하여 유사한 이미지를 찾는 기술입니다. 이러한 특징점은 이미지의 경계선, 색상값 분포, 픽셀의 변화량 등으로 사람이 유사한 이미지라고 생각하는 것과는 다소 거리가 있는 항목들입니다. 실제로 영상처리를 이용한 이미지 검색의 경우 동일한 이미지는 잘 찾지만 유사한 분위기의 이미지를 찾는 것은 역부족입니다.

AI 기술의 발전으로 이제 사람이 유사하다고 생각하는 이미지를 찾을 수 있게 되었습니다. 이미지 속에 어떤 사물이 있는지, 사람이 어떤 행동을 하고 있는지 등 구체화 된 특징점을 뽑지 않아도 유사한 이미지를 찾을 수 있습니다.

사람이 입력하는 메타데이터 검색의 한계를 넘어

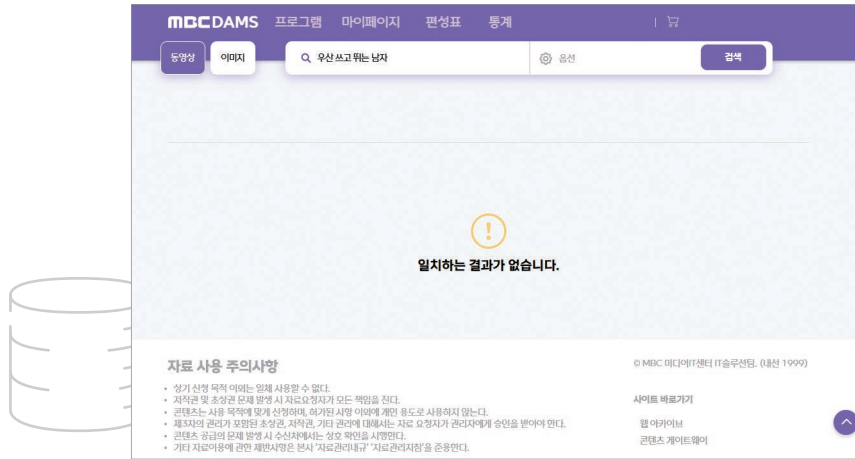


그림 2. 기존 시스템에서 검색되지 않는 '우산 쓰고 뛰는 남자' 검색 결과

AI 기반의 서비스 개발을 목표로 정한 후 AI 기술을 더 활용할 수 있는 방안들을 모색해보았습니다.

콘텐츠 검색시스템은 기본적으로 사람이 직접 입력한 메타데이터를 검색하여 콘텐츠를 찾게 됩니다. 모든 장면을 구체적으로 입력하는 데에는 한계가 있기에 앞뒤 맥락에 따라 중요한 정보 위주로 입력이 되고, 특정 장면에 대한 단순 설명은 부차적인 요소이기 때문에 메타데이터화 되지 않게 됩니다. 하지만 필요에 따라 단순한 장면 검색이 필요한 경우도 있는데 이러한 검색에는 어려움이 있었습니다. 따라서 문장을 분석하는 AI를 활용하여 콘텐츠 내 유사한 장면을 찾아주는 서비스를 기획하였습니다.

대사를 통한 콘텐츠 검색

마지막으로 AI STT(Speech To Text) 기술을 활용하여 음성에 대한 자막을 생성하여 대사를 통한 검색 서비스를 기획하였습니다. 2016년 이후 방영된 프로그램은 속기 자막 데이터가 존재하여 원하는 내용이나 특정 대사를 손쉽게 검색할 수 있지만, 과거의 구작 프로그램들은 이러한 속기 자막이 제공되지 이전으로 대사를 통한 검색에 어려움이 있었습니다. 최근 정확도가 눈에 띄게 높아진 AI STT 기능을 활용하면 이러한 검색이 가능할 것으로 생각되었습니다.

시스템 개발

AI 모델 선정 및 최적화

AI 모델을 직접 개발하는 것은 많은 전문학적인 비용과 시간이 드는 문제로, 현실적으로는 불가능하여 대부분의 AI 서비스 개발은 이른바 빅테크라 불리는 기업들이 개발한 AI 모델을 활용하여 이루어 집니다. AI 모델 개발을 직접 하지 않는다고 하여 서비스 개발이 수월하지는 않습니다. 수많은 모델 중 개발하려는 서비스와 적합한 모델을 선정하는 과정이 필요합니다. On-Premise와 클라우드 사이

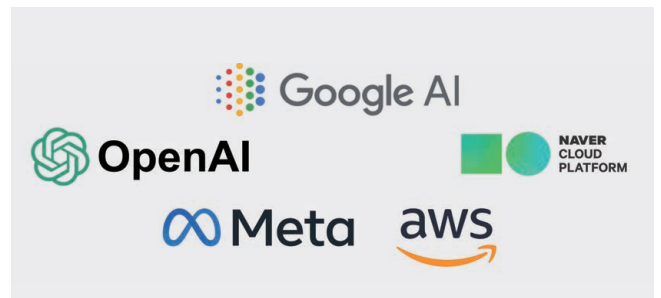


그림 3. 국내외 다양한 AI 개발 빅테크 기업들

의 고민, 오픈소스와 유료 모델 사이의 고민과 함께 많은 검증과정이 필요합니다. 다양한 모델을 직접 검증해보기 위해 모델별 개발 환경을 각각 세팅한 후 정확도, 소요시간, 최적화 가능 여부 등을 검토했습니다. 이러한 정량적인 평가 항목 외에도 사람이 느꼈을 때 적합한 결과가 나왔는지 판단하기 위해서는 여러 사람의 의견이 필요하므로 주기적으로 팀원들끼리 평가의 시간도 가졌습니다. 이러한 검증과정 끝에 서비스에 활용할 적절한 모델을 선정하게 되었고, 이어서 저희 서비스에 맞게 일부 튜닝과 최적화 과정을 거쳐 마침내 AI 서비스 단이 개발되었습니다.

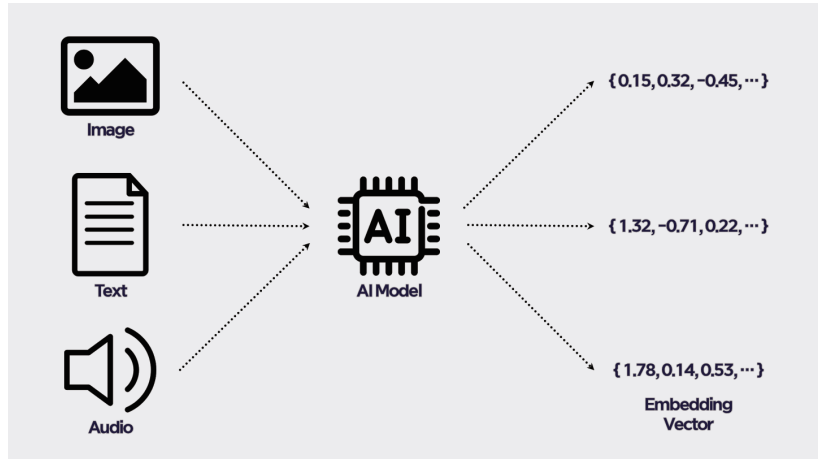


그림 4. AI 모델을 거쳐 생성되는 Embedding Vector

이렇게 개발된 AI 모델에 콘텐츠를 입력하면 AI 모델이 이해할 수 있는 일련의 숫자 배열로 구성된 임베딩 벡터(Embedding Vector)를 추출할 수 있습니다. 이 벡터 간의 관계를 통해 이들 간의 유사도를 계산할 수 있습니다.

검색엔진 개발

검색을 하기 위해선 아카이브 되어있는 콘텐츠들을 위와 같이 임베딩 벡터화하여 저장하고(검색 데이터 색인), 입력된 검색 데이터에서 임베딩 벡터를 추출하여 색인된 임베딩 벡터들 사이의 관계를 계산하여 유사도 순으로 검색 결과를 표출하는 과정이 필요합니다. 이러한 일련의 과정을 빠르게 처리하기 위해 검색엔진 개발이 필요합니다.

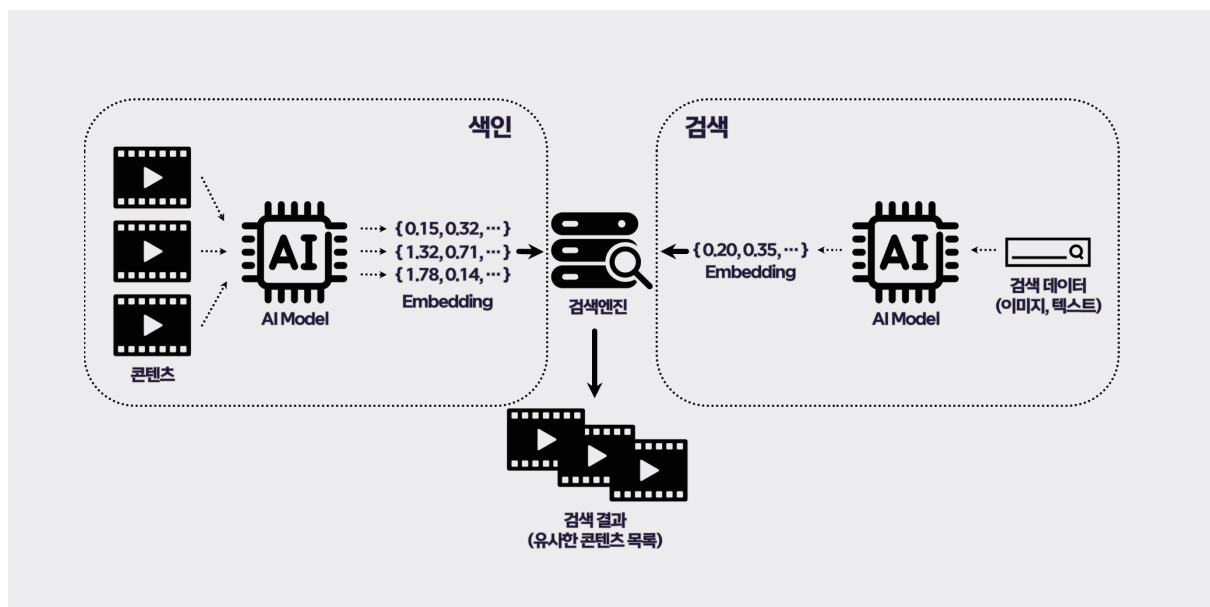


그림 5. 검색엔진 워크플로우

검색엔진도 다양한 상용 솔루션이 있지만 오픈소스를 활용하여 개발하였습니다. 벡터 간의 유사도를 직접 계산하여 쿼리 할 수도 있고, 내부 함수로도 제공해주어 비교 검토 후 최적화를 완료하였습니다. 검색도 검색이지만 방대한 양의 아카이브 콘텐츠를 다루다 보니 검색 속도 최적화를 위해 네트워크, 스토리지 등 인프라 부분에도 많은 어려움이 있었습니다. 여러 팀의 도움을 받아 서비스해도 충분할 만한 속도까지 도달하여 성공적으로 개발할 수 있었습니다.

웹 시스템 개발

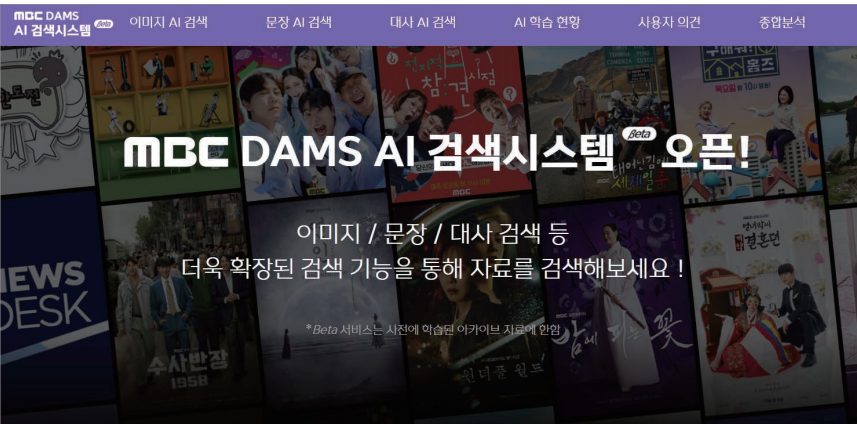


그림 6. AI를 활용한 콘텐츠 검색시스템 메인 웹페이지

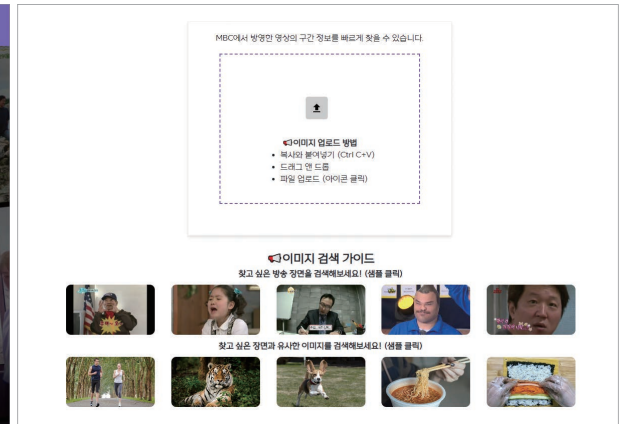


그림 7. 이미지 검색 UI

마지막으로 사용자들이 직접 검색을 하는 데 사용하게 되는 웹 개발이 진행되었습니다. 사용자 편의성을 고려하여 편리한 UI에 가장 중점으로 설계, 개발을 진행하게 되었습니다.

다소 어려울 수 있는 검색 방법을 가이드와 예시를 통하여 쉽게 이해할 수 있도록 설계하였습니다. 또한 검색을 편리하게 하기 위해 이미지 검색 시 일일이 이미지를 저장 후 첨부하지 않고 캡처 후 붙여넣기만으로도 검색이 되는 등 여러 가지 입력 방식이 가능하도록 개발하였고, 검색 결과 클릭 시 해당 시점부터 미리 보기 영상이 재생되어 결과 확인이 수월하도록 개발하였습니다. 검색된 결과는 아카이브 콘텐츠 대출시스템과 연동을 통해 원본 콘텐츠를 바로 공급받을 수 있도록 개발하여 사용성을 극대화하였습니다.

검색 서비스 예시

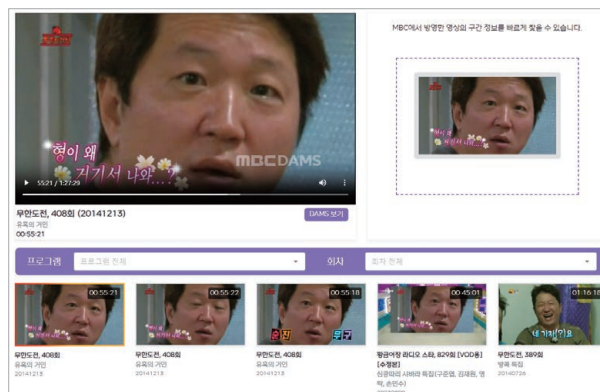


그림 8. 이미지 검색 예시(방송 장면 검색)



그림 9. 이미지 검색 결과(일반 이미지 검색)

이미지 검색의 경우 방송 장면 이미지를 검색할 경우 해당 장면이 방영된 콘텐츠 회차와 타임코드 시점을 정확하게 찾을 수 있습니다. 특히 인터넷에서 유행하는 밈이나 짤과 같은 방송 이미지의 원본 콘텐츠를 찾는 데 매우 수월해졌습니다. 기존의 경우 특정 방송 장면이 포함된 콘텐츠를 찾기 위해선 구글과 같은 외부 검색사이트의 이미지 검색 기능을 이용하여 방영 정보를 찾은 후 해당 정보를 이용하여 다시 콘텐츠 검색시스템에서 검색하여 조회한 후 해당 장면이 표출되는 시간을 일일이 재생하여 확인해야 하는 번거로움이 있었는데, 엄청난 개선이 이루어진 것입니다.

방송 장면이 아닌 일반 이미지를 검색하는 경우 아카이브된 콘텐츠 중 가장 유사한 장면을 찾아줍니다. 자료화면 제작 시 사물이나 풍경 등을 원하는 구도의 이미지로 검색하는 것으로 찾고자 하는 장면의 영상 콘텐츠를 쉽게 찾을 수 있습니다.

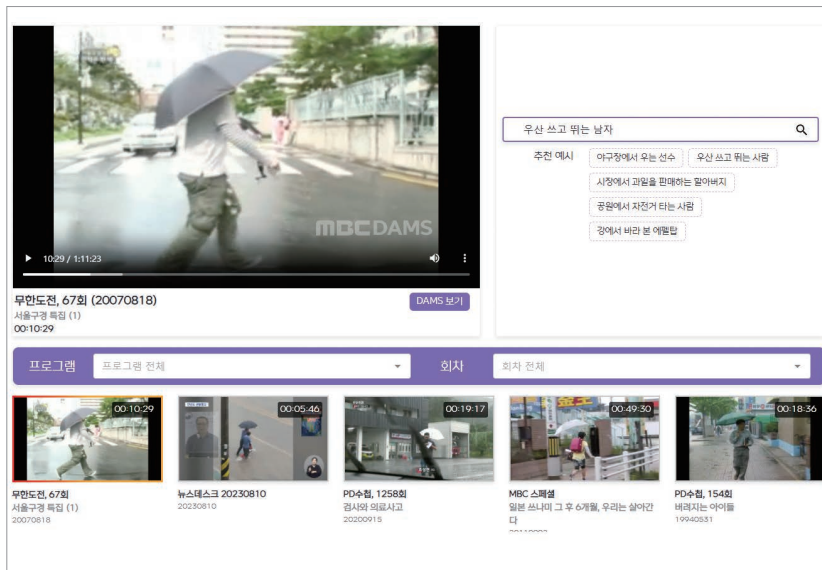


그림 10. 인물의 행동을 묘사하는 문장으로 검색(우산 쓰고 뛰는 남자)

이를 통해 특정 장면을 찾기 위해 그 장면의 메타데이터를 사전에 정확히 알고 있어야 했던 기존 검색의 한계를 극복할 수 있게 되었습니다. 또한 이미지를 이용한 검색의 경우 검색할 이미지를 가지고 있어야 한다는 점이 다소 불편할 수도 있지만 이처럼 문장 검색을 이용하면 단순히 텍스트를 입력하는 것만으로 원하는 장면의 영상 콘텐츠를 손쉽게 찾을 수 있습니다.

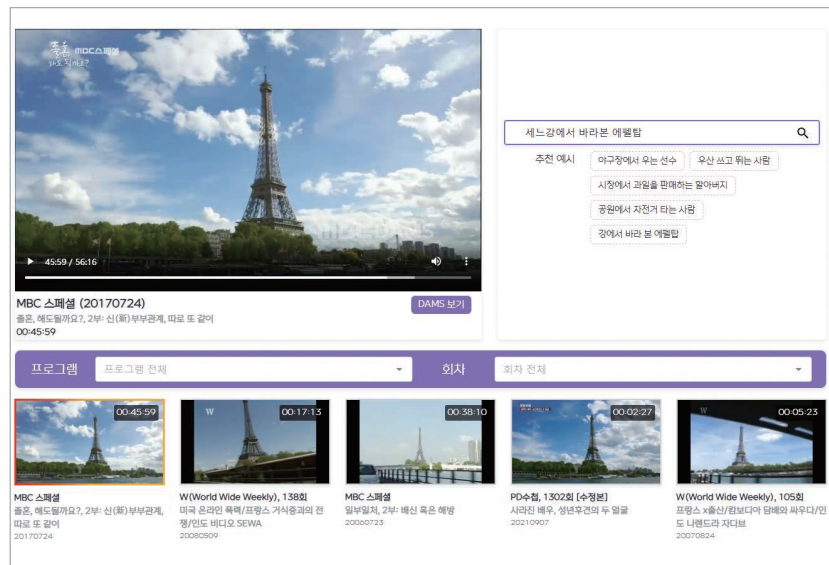


그림 11. 유명 지역의 풍경을 묘사하는 문장으로 검색(세느강에서 바라본 에펠탑)

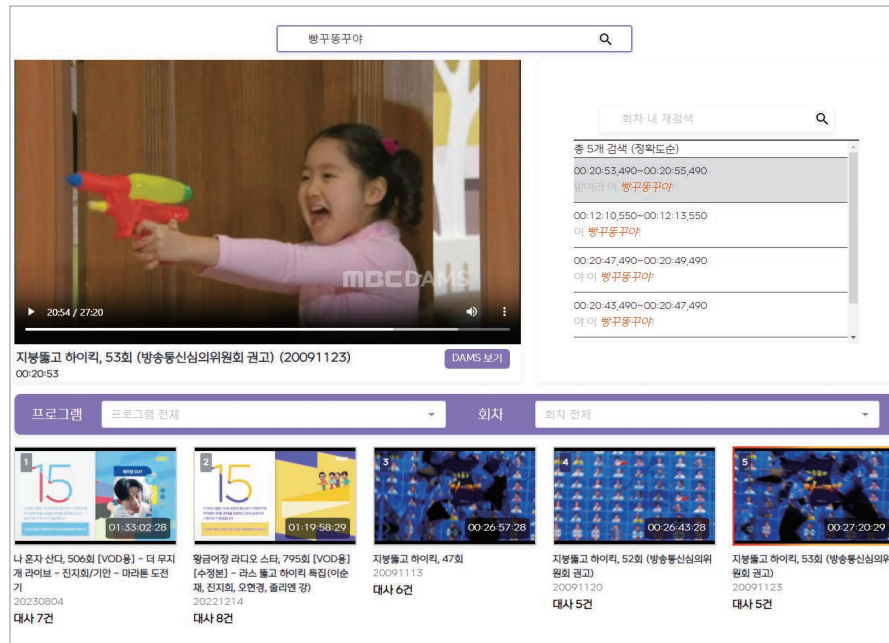


그림 12. 대사 검색(뽕꾸뽕꾸아)

마지막으로 대사 검색은 AI가 생성한 자막을 통해 기억에 남는 명대사를 검색할 수 있습니다. 드라마, 예능 등의 대사 검색 용도 외에도 뉴스를 검색하는 경우 입력된 메타데이터뿐만 아니라 리포트 대사 내용에서 검색을 해서 더 깊이 있는 검색이 가능해졌습니다. 기자명, 인물명, 사건명 등으로 검색어를 입력하여 찾고자 하는 뉴스 리포트를 더 정확하게 검색할 수 있습니다.

마무리하며

서두에 말씀드렸듯이 AI 기술은 빠르게 진화하고 있습니다. 본 시스템에 적용한 AI 모델들도 시시각각 개선된 버전이 업데이트되고 있고 새로운 AI 모델 또한 출현하고 있습니다. AI 서비스에 ‘완성’이라는 단어는 없는 것 같습니다. 더 정확하고 깊이 있는 검색을 위해 새로운 모델에 대한 검토와 검증과정을 계속 진행할 예정입니다. 또한 얼굴 이미지 분석을 이용한 인물 검색 및 특정 인물의 행동 묘사를 함께 멀티모달로 검색하여 더 구체적인 검색 결과를 도출할 방안을 기획 중입니다.

본 프로젝트를 성공적으로 마무리한 후 팀원들이 하나 되어 사내 홍보도 열심히 해주신 결과 본사에 없어서는 안 될 시스템으로 자리매김하고 있습니다. 특히 매시업 콘텐츠 제작 분야에서 너무 잘 사용하고 있다는 피드백을 받을 때마다 뿌듯함을 느끼며 일하고 있습니다.

이러한 AI 기반의 시스템들이 방송국 및 미디어 산업에 안정적으로 자리 잡을 수 있도록 끊임없이 노력하여 발전시키도록 하겠습니다. 

