KISDI Perspectives

동향 October 2024 No.3

# 방송영상산업에서의 생성형 AI 활용과 시사점

**정은진 전문연구원**
정보통신정책연구원 미디어정책연구실

**김예빈 연구원**
정보통신정책연구원 미디어정책연구실

KISDI 정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

KISDI Perspectives

동향

2024. 10. 21

# 방송영상산업에서의 생성형 AI 활용과 시사점

**정은진 전문연구원**
정보통신정책연구원 미디어정책연구실, jean@kisdi.re.kr

**김예빈 연구원**
정보통신정책연구원 미디어정책연구실, yebin_kim@kisdi.re.kr

## 요 약

- 방송영상산업에서 다양한 콘텐츠 제작 과정에 생성형 AI가 활용되기 시작하여 산업 전반의 변화를 촉발하고 있음
  - 최근 생성형 AI에 대한 연구 개발(R&D)과 기업 투자가 급격히 증가하고 파운데이션 모델의 실용화를 통해 생성형 AI 도구가 빠르게 확산됨
  - '생성형 인공지능' 기술은 생성 AI 모델 (예: GAN, Diffusion, Transformer 등)을 기반으로 창작자의 질문을 입력 받아 능동적으로 새로운 결과물을 창출하는 것이 특징
  - 최근 인간의 인지적 학습 방식을 모방하여 시각과 청각으로까지 다양한 양식(modality)의 데이터 입력과 출력을 지원하는 거대 멀티모달 모델(Large Multi-modal Model)로 고도화됨
- 생성형 AI는 방송 및 창작 산업에 새로운 기회를 제공하고 있지만, 동시에 창작자의 권리나 지적재산권(IP) 침해 등 다양한 위기를 동반하고 있음
  - 생성형 AI는 텍스트, 이미지, 영상, 오디오, 음악 분야에 두루 걸쳐 새롭고 효율적인 콘텐츠 생성을 가능하게 하며, 방송영상산업의 제작비를 절감하며 부가가치를 증대하는 데 기여함
  - 기성 미디어 사업자는 생성형 AI 모델 훈련을 위한 콘텐츠 라이선스 계약을 체결하거나, 자사 콘텐츠로 파인튜닝한 생성형 AI 모델을 개발하기 위한 제휴를 맺는 등 생성형 AI 사업자와 새로운 협업 방식을 모색하고 있음
  - 한편, 창작 업무 일부가 자동화될 경우 창작자의 권리 침해가 발생할 수 있다는 종사자들의 우려가 증가하고 원저작자의 허락 없이 무단으로 사용된 자료 또는 유명인의 사진·영상·음성 등 퍼블리시티권을 보호받는 자료가 생성형 AI 학습에 활용되었다는 문제제기가 있는 만큼 생성형 AI의 발전은 방송영상산업에 양가적인 영향을 미침
- 따라서 창작자의 권리를 보호하고 공정한 대가를 보장하는 법적·정책적 장치 마련이 필수적이며, 지속 가능한 창작 환경을 조성하기 위한 노력이 필요함

# 01 개요

- 지난 2023년 기준, 전 세계 '생성형 AI' 분야에 유치된 민간투자 규모는 총 252억 달러로, 2019년 대비 30배, 2022년 대비 9배에 달하는 수준(Standford HAI, 2024)
- 생성형 AI의 소프트웨어 매출 규모는 2023년부터 연평균 53%씩 증가하여, 2028년에는 580억 달러에 도달할 것으로 예상됨. 2024년에는 124%의 급성장을 보일 것으로 전망함(Omdia, 2024)
- 최근 생성형 AI에 대한 연구 개발(R&D)과 기업 투자가 급격히 증가했으며, 파운데이션 모델*의 실용화를 통해 생성형 AI 도구가 빠르게 확산됨

  * 파운데이션 모델 : 대규모 데이터세트를 기반으로 훈련된 범용적인 특성의 AI 모델로, 전문화된 모델을 개발하는 데 활용됨
- 창작 및 방송영상 산업에서 생성형 AI의 도입이 본격화되는 단계로, 다양한 콘텐츠 제작 과정에서 생성형 AI가 활용되며 산업 전반의 변화를 촉발하고 있음
- 본 고에서는 생성형 AI의 시장 성장이 방송영상 산업에서의 미치는 영향과 구체적인 활용 사례를 통해 방송영상산업의 기회와 위기를 조명하고, 시사점을 도출하고자 함

# 02 방송영상 산업에서 AI 적용범위 확대

**[전통적인 예측형 인공지능]**

- 소위 '예측형 인공지능(predictive AI)' 또는 전통적인 비생성형 인공지능은 '생성형 인공지능(generative AI)'과 기능 및 적용에서 차이점이 있음
- 전통적인 '예측형 인공지능'의 활용 분야는 빅데이터를 학습한 분류와 예측이었으며, 창작 산업에서도 데이터와 규칙 기반 업무 효율화에 제한적으로 활용됨*

  * (대표적인 사례) 맞춤형 콘텐츠 영상 추천, 시청률과 광고효과 예측, 스포츠나 선거 등 형식적 보도에 특화된 AI 저널리즘 등

  - 2020년 유럽 의회가 발표한 '시청각 부문의 인공지능 사용' 연구 보고서에 따르면, 당시 시청각 부문에서 AI를 활용하는 분야는 △자막 생성, △이미지·영상 분류와 식별, △초벌 편집, △기계 번역, △소셜미디어 반응 분석 등이었음 (Rehm, 2020)

**| 표 1 | 유럽 의회 '시청각 부문의 인공 지능 사용' 보고서**

| 모달리티 | 방송영상 제작 활용 방안 |
|---|---|
| 텍스트 | • 텍스트 콘텐츠 번역 (접근성 향상 목적)<br>• 자동 자막 생성<br>• 현지화<br>• 개체명 인식, 요약, OCR, 의미 추출, 콘텐츠 분류, 주제 감지 (인덱싱 목적)<br>• 내부 검색 엔진, 추천 엔진<br>• 소셜 미디어 분석 |
| 이미지/영상 | • 자막 제공 (접근성 향상 목적)<br>• 음성 생성 (TTS)<br>• 이미지 분류, 객체 감지, 얼굴 감지, 장면 감지 (인덱싱 목적)<br>• 자동 초벌 편집<br>• 영상 원본 처리<br>• AI기반 비디오 인코딩 및 콘텐츠 전달 향상 |
| 오디오 | • 자막 제공 (접근성 향상 목적)<br>• 음성 생성 (TTS) |

자료: Rehm (2020) 재구성

주: 2020년 3월, 유럽 내 공민영 방송사, 영화사업, 비디오게임 개발, 출판사, 통신사, 학계 등 시청각서비스 부문의 다양한 이해관계자 85인을 대상으로 "현재 AI 기술을 조직에서 활용하는가? 활용한다면 어떤 분야를 가장 중요한 것으로 인식하고 있는가?" 질문으로 개방형 인터뷰를 진행한 응답 결과

## [생성형 인공지능]

- 한편, '생성형 인공지능' 기술은 생성 AI 모델 (예: GAN*, Diffusion**, Transformer*** 등)을 기반으로 창작자의 질문이나 입력을 받아 능동적으로 텍스트, 이미지, 비디오 등과 같이 다양한 방식으로 새로운 결과물을 창출하는 것이 특징

  * 적대적 생성 신경망(GAN, Generative Adversarial Network) : 2014년 Ian Goodfellow가 제안한 딥러닝 모델로, 두 개의 신경망(생성자와 판별자)이 서로 경쟁하며 학습하는 방식

  ** 확산 모델(Diffusion Model) : 2015년 Sohl-Dickstein 등이 제안한 생성 모델로, 데이터에 점진적으로 노이즈를 추가한 후 이를 역으로 제거하는 과정을 학습하며, 노이즈 제거 과정을 통해 고품질의 샘플을 생성할 수 있으며, 특히 이미지/비디오 생성 분야에서 뛰어난 성능을 보임

  *** 트랜스포머(Transformer) : 2017년 Google의 Vaswani 등이 제안한 신경망 구조로, 어텐션 메커니즘만을 사용하여 순차 데이터를 처리하며 장거리 의존성을 효과적으로 포착할 수 있어 자연어 처리 분야에서 혁신을 일으키며 현재 대규모 언어 모델의 기반이 되었고, 이미지/비디오/오디오 분야에도 활용 가치가 높음

  - 2022년 11월, ChatGPT(LLM)가 첫 상용화 서비스로 공개된 이후 급격히 성장했으며, DALL-E, Stable Diffusion, 미드저니 등 텍스트-투-이미지도 활성화

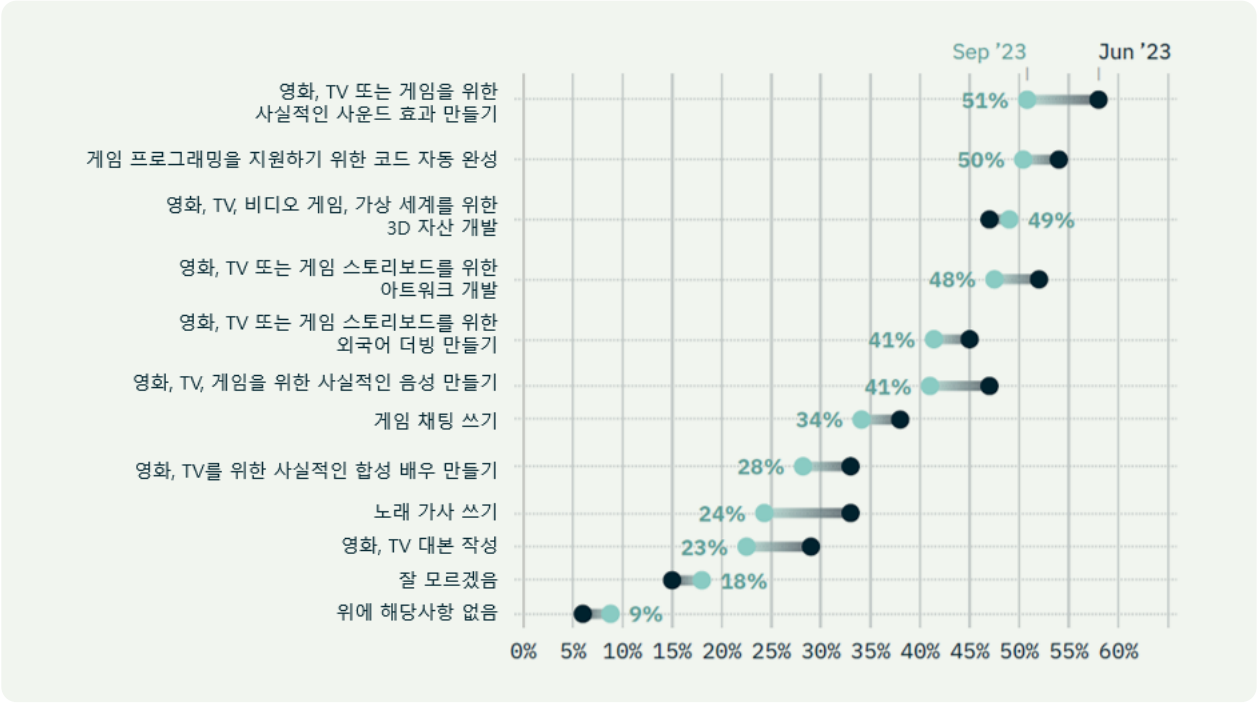
- 최근 인간의 인지적 학습 방식을 모방하여 시각과 청각으로까지 다양한 양식(modality)의 데이터 입력과 출력을 지원하는 거대 멀티모달 모델(Large Multi-modal Model)로 고도화하고 있음

- 즉, 텍스트, 이미지, 오디오 등 다양한 형식의 데이터를 통합적으로 입·출력 양식으로 처리함으로써, 콘텐츠 창작에 필요한 실질적인 기초를 제공하기 시작했다고 평가*

  * (대표적인 사례) 영상 요약과 자막 생성(Video-to-Text), 스타일 변환, 특정 부분 인페인팅, 편집 (Image-to-Image), 애니메이션 효과 부여, 이미지에 맞는 영상 생성(Image-to-Video), 음성에 어울리는 동영상 생성, 음성을 수화로 번역(Audio-to-Video)

- 향후 2-3년 내에 생성형 AI에 기반하여 △사실적인 효과음 생성, △VR 및 3D 자산 제작, △스토리보드 아트작업, △외국어 더빙, △자연스러운 목소리 생성 등 방송·영상 제작 현장에서 활용도가 확대될 것이라고 기대됨(Variety, 2023)

**| 그림 1 | 영화 및 방송분야의 생성형 AI 활용**




자료: Variety (2023) 재구성

주: 2023년 6월과 9월 2회에 걸쳐, 미국 내 영화, 방송, 음악 전문종사자 510명 대상 "2-3년 내 생성형 AI가 효과적으로 활용될 분야를 모두 고르시오" 중복응답 설문에 대한 응답 결과를 정리함

- 본 고는 2024년 10월 현재 생성형 AI가 방송 및 영상산업에 지원하고 있는 분야 중 텍스트, 이미지, 영상, 오디오, 음악 생성 결과물을 출력하는 기술과 서비스를 구체적으로 소개하고, 활용 현황을 분석하고자 함

## 5개 도메인별 생성형AI의 활용 현황

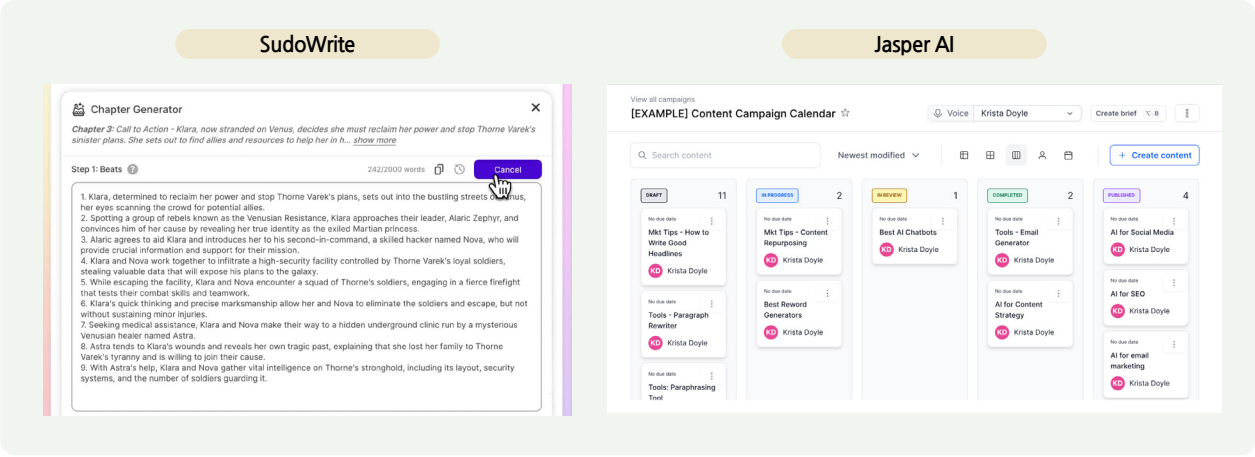
### [텍스트 생성]

- 초거대 언어 모델(Large Language Model, LLM)으로 인해 자연어 입력어를 기반으로 긴 문맥을 이해하여 자료 검색 및 문서 작업을 수행할 수 있음
  - GPT, Gemini, LLaMA와 같은 LLM 모델을 활용한 ChatGPT(오픈 AI), Gemini(구글), Claude(앤트로픽), 메타 AI(메타) 등 대화형 AI 서비스가 출시되었으며, 국내에서도 하이퍼클로바(네이버), 엑사원(LG AI연구원), VARCO LLM(엔씨소프트) 등 LLM 기반 생성형 AI 모델과 관련 대화형 AI 에이전트 서비스를 개발
- 한편, LLM을 파인튜닝*하여 시나리오, 마케팅 카피, 자막을 기획·작성·편집하는 작업을 전문적으로 지원하는 생성형 AI 도구가 출시되고 있음

  * 파인튜닝(fine-tuning) : 사전 학습된 대규모 모델에 특정 작업이나 도메인에 맞는 데이터를 추가 학습시켜 특화된 응용을 목표로 모델 성능을 조정하는 과정

- 대표적인 사례
  - (SudoWrite) 2024년 현재, Jasper는 GPT-3.5, GPT-4, Claude 2 등의 AI 모델을 적용해 작가들을 위한 브레인스토밍 도구와 다음 문단 자동완성 기능을 제공하며, 장르와 문체를 선택해 창의적인 장편 작문을 지원하는 서비스로 제공됨. 주로 영어 중심이지만, 한국어도 지원
  - (Jasper AI) 2021년에 설립된 Jasper는 2024년 현재 GPT 등 여러 LLM 모델을 적용해 한국어(출력 언어에 한정)를 포함한 80개 이상의 언어를 지원하며, 이를 통해 기사 작성, 카피라이팅 같은 상업적 콘텐츠 작문부터 기업 고유의 음성 및 이미지 생성 등 브랜딩 업무에 특화된 AI 서비스를 제공
  - (아이작) 2024년 한국방송광고공사에서 네이버 클라우드의 하이퍼클로바 X 모델에 기반하여 개발한 서비스로 톤앤매너 등을 선택하여 광고카피를 생성하는 데 특화된 서비스

| 그림 2 | Open AI(DALL-E3), Jasper AI의 프롬프트와 생성 결과물




자료: Open AI, Jasper AI 공식 홈페이지 참고

- (활용사례) 작품의 초기 각본 구상, 자막 제작 등에 활용
  - 2023년 9월, 미국 작가조합(WGA)은 작가가 원할 경우 드라마 및 영화 등 각본을 작성하기 위해 회사의 동의하에 Chat GPT와 같은 생성형 AI 도구를 사용할 수 있다는 합의 각서를 체결
    ※ 본 계약에 따라, 방송사와 영화제작사는 작가에게 AI 사용을 강요할 수 없으며, AI 생성물을 각색하더라도 작가의 오리지널 시나리오로 간주하고, AI가 생성한 자료는 작가의 기여나 권리를 침해할 수 없으며, WGA는 작가의 창작물이 AI 학습에 사용될 경우 이를 문제 삼을 권리가 있음
  - 2024년 9월, 워너 브라더스 디스커버리는 구글 클라우드의 Vertex AI 플랫폼을 이용해 OTT 서비스 '맥스'에서 제공되는 논스크립트 콘텐츠*의 자막을 제작할 계획이라고 발표했음. 이로써 자막 제작 비용을 50% 절감하고, 제작 시간을 80% 단축할 수 있을 것으로 기대(CNBC, 2024.9.24.)
    * 논스크립트 콘텐츠 (Non-Scripted Content) : 정형화된 각본이 필요하지 않은 콘텐츠로, 리얼리티쇼, 예능 버라이어티, 팩츄얼 드라마 장르 등

## [이미지 생성]

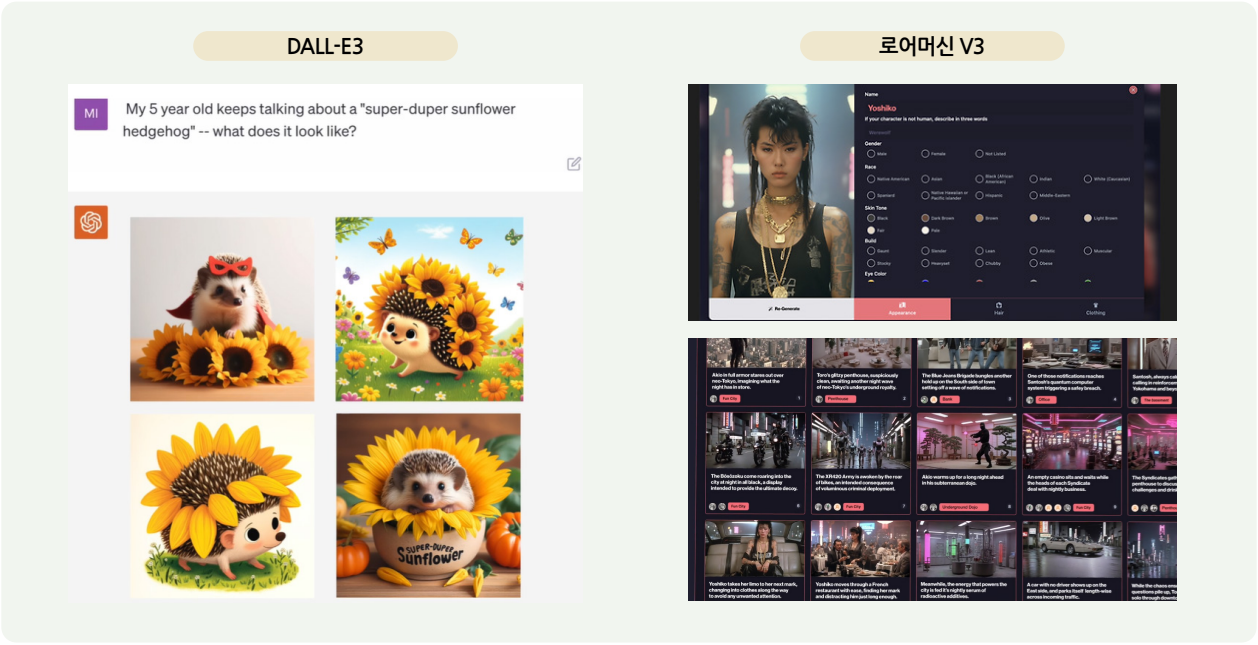
- LLM과 이미지 생성 모델의 발전으로, 텍스트를 입력하면 이를 이미지로 시각화하는 '로어머신', '미드저니', 'Imagen(구글)', 'DALL-E(오픈 AI)' 등 서비스들이 실사에 가까운 수준으로 기술이 발전
- 이미지 생성 모델들은 기존 소셜미디어나 대화형 서비스에 통합되어* 대중적으로 사용자 경험을 더욱 풍부하게 만드는 방향으로 나아가고 있음
  * 메신저 서비스 스냅챗의 "My AI chatbot", 메타의 "Imagine with Meta AI" 이미지 생성기 등

- 상업 이미지 판매 및 라이선싱 플랫폼을 운영하는 게티 이미지는 엔비디아와 협력해 '아이스톡 생성형 AI(iStock Generative AI)'를 출시했으며, Adobe도 자사 스톡 데이터를 학습한 Adobe Firefly 모델을 선보임. 이들은 모두 IP 침해 문제를 해결하는 생성형 AI 서비스로 차별화를 모색함
- 대부분의 이미지 생성 모델은 가까운 시일 내에 영상 생성 모델을 출시할 계획을 가지고 있어, 영상 생성 기술의 발전도 기대됨
- 대표적인 사례
  - (스테이블 디퓨전3/스테이블 디퓨전 XL) Stability AI 사는 이미지 생성 인공지능 모델 '스테이블 디퓨전'을 오픈 소스로 공개하여 새로운 애플리케이션이나 기존의 모델을 개선할 수 있는 생태계 확장 전략을 추진
    ※ 2024년 들어 멀티모달 경쟁에 적극 나서고 있으며, 9월에는 영화 〈터미네이터〉와 〈아바타〉의 감독 제임스 카메론이 Stability AI 이사회에 합류하며 방송영상산업에서 영향력을 확대하고 있음
  - (로어머신) 텍스트만으로 스토리보드를 시각화하는 서비스로, 스타일 선택, 캐릭터 생성, 장소 생성, 장면 선택, 렌더링의 기능을 제공하여 그래픽 노블, 애니메이션 스케치, 스토리보드로 활용됨
  - (미드저니) 2022년 오픈베타로 출시된 이후, 2024년 6월 공개된 버전 6.1까지 다양한 세대의 모델이 발전해왔으며, 2025년 중 영상 서비스를 출시할 예정

**| 그림 3 | 미드저니의 버전별 이미지 생성 결과 비교**

자료: 미드저니; Stanford HAI(2024) 재구성
주: 미드저니 2022년 버전1부터 2023년 버전6까지 "해리 포터의 초현실적인 이미지"를 프롬프트로 입력한 결과물

| 그림 4 | Open AI(DALL-E3), 로어머신(로어머신)의 이미지 생성 결과물




자료: 각 사 홈페이지

## [영상 생성]

- 영상 생성 기술은 그 복잡성 때문에 이미지 생성보다 발전 속도가 더디고, 짧은 길이의 영상 출력에 한정되어 왔음
  - 그러나, 2024년 초 OpenAI가 디퓨전 트랜스포머 모델을 기반으로 한 '소라'를 발표한 이후, 다중 모달리티를 지원하는 고성능 영상 서비스의 잠재적 성장 가능성에 대한 기대가 높아짐
- 대표적인 사례
  - (런웨이 Gen-3 알파) 런웨이는 2018년에 설립되었으며, 동영상 생성 및 편집에 특화된 서비스를 제공하고 있으며, 2023년에 발표한 런웨이 Gen-1부터 2024년 7월에 출시된 런웨이 Gen-3 알파까지 기술을 발전시키며, 동영상 제작의 혁신적인 도구로 자리 잡고 있음
    ※ 2023년 6월 기준, 구글 및 엔비디아 등으로부터 합산 2억 3700만 달러 규모의 투자금을 유치하여 기업가치 15억 달러를 달성한 것으로 알려짐 (Reuters, 2023.6.30.)
  - (피카 랩스) 2023년 9월, "아이디어를 영상으로(Idea to Video)"를 모토로 출시되었으며, AI를 활용한 손쉬운 콘텐츠 생성을 지원
  - (소라) 2024년 2월, Open AI에서 공개한 트랜스포머 아키텍처 기반의 영상 생성 모델로, 장면 전환과 물리적 현상에 대한 시각적 표현력이 정교하다는 평가를 받으며 상용화된 서비스로는 출시되지 않았음

- (유튜브 드림스크린) 2024년 9월, 구글의 딥마인드가 자사의 VEO와 Imagen3 등 생성형 AI 모델을 유튜브 창작자를 위한 인터페이스인 '드림스크린'에 통합할 계획을 발표 (Google Deepmind, 2024.9.18.)

  ※ 2025년부터 유튜브 쇼츠에 배경화면으로 AI 생성 영상을 삽입하거나 6초 이하의 독자적인 영상 클립을 제작할 수 있는 도구로 제공될 예정

- (Movie Gen) 2024년 10월, 메타(Meta) 연구팀은 텍스트 프롬프트 기반 영상 편집, 사용자 사진을 입력한 개인맞춤형 영상 제작 기능을 제공하는 생성형 AI 모델인 Movie Gen을 공개

**| 그림 5 | 런웨이(Gen-3), 메타(Movie Gen)의 프롬프트와 영상 생성 결과물**

Gen-3 alpha

세렝게티를 걷는 드래곤 투칸 새

Movie Gen

한 여성 DJ가 LA의 옥상에서 레코드를 돌리고 있습니다. 그녀는 분홍색 재킷과 커다란 헤드폰을 착용하고 있습니다. 여성 옆에는 치타 한 마리가 있습니다. 배경은 도시 풍경입니다.

자료: 각 사 홈페이지
주: 각 프롬프트는 원문을 한글로 번역했으며, 이미지는 생성된 영상의 스틸컷임. 메타 Movie Gen 사례에서, 오른쪽 위 섬네일은 영상 생성에 활용된 사용자의 사진을 나타냄.

- (활용 사례) 사후 제작, 버추얼 휴먼, CG 등 다양한 분야에서 활용되고 있으며, 2024년 현재 AI 기술 회사와 방송사, 제작사 간의 제휴가 활발히 진행되고 있어 앞으로 더 다양한 활용 사례가 발표될 것으로 기대됨

  - (영상편집) 영화 〈에브리씽 에브리웨어 올 앳 원스〉는 런웨이의 서비스를 이용하여 '인페인팅'* 작업을 수행했다고 밝힘 (Variety, 2023.2.22.)

    * 이미지 또는 동영상에서 손실되거나 손상된 부분을 추정하는 편집과정

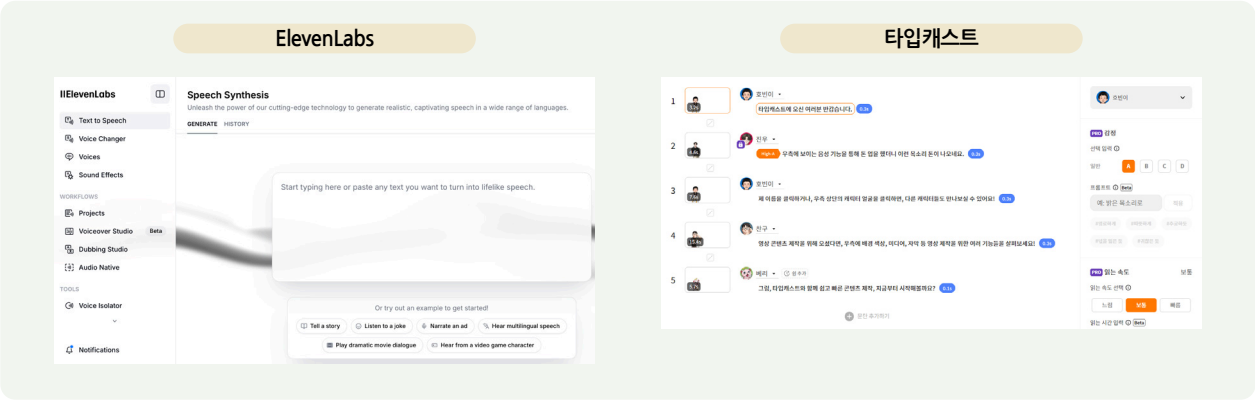
  - (버추얼 휴먼) 국내에서 나문희 배우의 디지털 IP를 계약하여, AI 캐릭터를 SNS 광고 영상이나 브랜디드 영상을 제작하고 업로드한 사례 발표(AI타임즈, 2024.7.23.)

- (AI 배경) CJ ENM은 드라마 〈눈물의 여왕〉에서 해외 로케이션 촬영 대신, 초고해상도 LED 버추얼 프로덕션 스테이지와 AI로 생성한 배경을 활용해 자작나무 숲을 구현 (한국일보, 2024.6.15.)

**[오디오 생성]**

- 음성 생성 AI는 사람의 목소리를 학습시킨 후 이를 활용하여 오디오를 제작하거나 서비스에서 제공되는 음성 또는 가상 인물의 목소리 설정을 통해 제작
  - 텍스트를 통해 음성을 생성할 수 있는 TTS*와 화자의 음성 보이스 샘플을 학습하여 유사한 목소리를 생성하는 보이스 클로닝(voice cloning)** 기술이 있음
    * TTS(Text to Speech) : 입력된 텍스트를 인공지능 기술을 활용하여, 자연스러운 음성으로 변환하는 기술로, 다양한 언어와 음성 스타일을 지원하며 복잡한 언어 처리 알고리즘과 음성합성 모델을 사용하여 텍스트의 구조, 문맥, 발음을 분석하고, 이를 바탕으로 억양, 리듬, 강세 등을 적용해 인간의 음성과 유사한 합성 음성을 생성하여 더빙작업에 용이함
    ** Voice Cloning : 특정 개인의 목소리 특성을 AI 모델로 학습하여, 그 사람의 음성을 모방해 새로운 텍스트를 생성하는 기술로 짧은 음성 샘플만으로도 화자의 음색, 억양, 말투 등을 분석해 자연스럽게 복제할 수 있어 맞춤형 음성 합성이 가능하며, 성우 작업, 콘텐츠 현지화(더빙 번역) 등에 활용 가능
  - AI에게 사람의 목소리를 학습시키는 방법은 주어진 문장을 직접 읽거나 기존에 녹음되어 있는 파일을 업로드하여 텍스트와 음성을 매핑하는 방식으로 진행되며, Open AI가 최근 공개한 Voice Engine은 15초 분량의 샘플로도 가능
    ※ Open AI는 Voice Engine의 음성 모방 기술이 오용 가능성의 우려가 높기 때문에 현제까지는 preview만 진행하며 Chat GPT와 같이 광범위한 서비스 제공은 아직 미정이라 밝힘
- 대표적인 서비스
  - (ElevenLabs) 가장 대표적인 TTS 기반 생성 AI로 29개의 언어로 다른 언어로 실시간 더빙 및 음성합성 기능 제공함. 2023.1.11. 텍스트 음성 변환, 음성 복제 기능의 베타 출시되었음
  - (Speechify) 문서, 기사, PDF, eBook, 이메일 등을 음성으로 변환하는 AI로 대량의 자료를 음성으로 변화하는 것에 특화되어 있으며, 생성 시 음성의 주요 내용 강조 및 속도를 자동으로 조절
  - (타입캐스트) 2019년 4월 오픈 베타 제공을 시작으로 현재 175만 명 이상의 가입자를 보유한 국내 서비스로 한국어, 영어, 일본어를 포함한 7개국 언어를 제공하고 있음. 맞춤형 음성 생성과 분노, 기쁨, 슬픔 등의 감정을 선택하여 조정 가능

| 그림 6 | ElevenLabs와 타입캐스트의 프롬프트




자료: 각 사 홈페이지

- (활용사례) 더빙, 시각 장애인을 위한 화면 해설 방송, 뉴스 방송 자동화, 라디오 및 팟캐스트 콘텐츠 생성, 가상 캐릭터 제공 등의 다양한 활용 사례 보일 것으로 전망됨
  - 2022년 디즈니+ 오리지널 드라마 '카지노'에서 최민식의 30대 목소리를 보이스 디에이징*을 통해 구현 (AI타임즈, 2022.12.28.)
    * 대상의 목소리를 스펙트로그램 분석하여 흩어져 있는 목소리 구성요소를 안정적인 목소리로 재구성하는 기술
  - tvN 토일 드라마 '환혼' 화면 해설 방송에 배우 전미도의 목소리로 AI 보이스 클로닝(Voice Cloning)을 도입해 화면 해설 방송 지원 (CJ올리브네트웍스, 2022.7.5.)

## [음악 생성]

- 음악 생성 AI는 대규모 음악 데이터를 학습해 새로운 음악을 창작하는 기술로 Autoregressive Diffusion, Transformer, GAN 모델 등 여러 AI 모델을 활용해 다양한 장르와 스타일의 음악을 생성
  - 음악의 분위기나 가사를 텍스트로 입력해 사용자 지정에 맞춰 멜로디, 리듬, 화음 등을 만들 수 있음
- 대표적인 사례
  - (Suno AI) Microsoft와의 파트너십을 통해 2023년 12월 20일에 출시되었으며, 대표적인 음악 생성 AI로 자리 잡음. 음원 제작뿐만 아니라 사용자가 선택한 언어로 보컬을 추가할 수 있으며, 필요한 세부 사항을 텍스트로 추가하거나 수정할 수 있음
  - (Udio) 구글 딥마인드 출신 연구원들로 구성된 미국의 스타트업. 주제, 장르, 등의 콘셉트을 입력하면 40초 안에 새로운 노래를 생성하며, 생성된 음악은 저작권 침해를 방지하기 위해 자동 저작권 필터 기능을 지원함

- (Musicfy) 1000가지 이상의 유명인의 목소리 샘플을 가지고 있는 툴로, 음성을 새로 클로닝하거나, 샘플을 합성해서 새로운 목소리를 만들어 AI 커버 곡을 제작 할 수 있음

● (활용사례)

- KT알파 쇼핑의 홈쇼핑 프로그램에서 시그널 송을 AI로 제작한 BGM을 활용
- 오디오 드라마 '어서 오세요, 휴남동 서점입니다'에서 OST로 AI 편곡 기반 리메이크곡인 '같은베개'(원곡 테이)를 제작 (AI타임즈, 2023.4.14.)
- Mnet의 'AI 음악 프로젝트 다시한번'에서 터틀맨과 김현식, JTBC의 'ALIVE'에서 임윤택과 유재하의 목소리 등 고인의 음성 재현 (YTN, 2021.9.30.)

## 03 향후 전망 및 시사점

● 방송 및 영상 산업에서 생성형 AI(Generative AI)의 도입은 빠르게 증가하고 있으며, 이를 통해 부가가치를 창출하고 맞춤형 서비스를 제공하는 사례들이 늘어나고 있음

- 이와 동시에, IP에 대한 정당한 대가를 지불하는 문제도 중요하게 대두되고 있음
- 본 챕터는 생성형 AI가 창작 산업에 미치는 영향과 기회, 위기를 분석하며, 관련된 법적, 정책적 쟁점들을 다루고자 함

**[생성형 AI 도입에 따른 방송영상산업의 기회]**

● 생성형 AI는 데이터 기반으로 효율적인 콘텐츠 생성을 가능하게 하며, 더 빠르고 창의적인 제작이 가능해졌으며, 소비자의 취향을 반영한 맞춤형 콘텐츠를 제공하는 기회를 제공함. 이를 통해 방송 및 영상 산업의 부가가치가 더욱 증대됨

● 제작사들은 AI 모델을 활용하여 제작 시간을 단축하고, 제작비를 절감하고 있음

※ 드라마 〈눈물의 여왕〉 AI 배경 생성 기술, 워너브라더스 스트리밍 서비스 AI 자막 사례 참고

**[생성형 AI 도입에 따른 방송영상산업의 위기]**

● 생성형 AI의 도입은 창작 산업의 질적 저하, 데이터 보안 문제, 편향된 콘텐츠 생성 등 다양한 위기를 초래할 수 있음(Amankwah-Amoah et al., 2024)

- 또한, AI 기술은 기존 사업자의 경쟁 우위를 약화시키고, 창작 업무를 자동화함으로써 현 창작 산업 자체의 위기를 불러올 수 있다고 진단됨

● 실제로, 생성형 AI는 예술·디자인·엔터테인먼트·미디어·스포츠 부문에서 업무의 26%*를 자동화할 잠재력이 있는 것으로 세계경제포럼(2023)과 골드만삭스가 전망함

* 전체 산업의 자동화 비중은 25%로 전망됨

- 기술의 발전에 따라 방송영상 종사자의 탈숙련화와 통제력 상실에 대한 불안감도 증가함.* 이에 생성형 AI를 활용한 창작 기회에 대한 기대와 반감이 공존하는 상황**

* 할리우드에서는 2023년 WGA와 SAG-AFTRA 파업이 발생하였으며, AI 사용시 창작자 권리 보호장치 마련이 핵심 이슈로 떠올랐음

** 이에 따라, 유튜브 등 생성형 AI를 서비스에 접목하는 플랫폼은 생성형 AI 기술이 창작자를 대체하는 것이 아니라 '권능화(empower)'하는 데 목적이 있다고 공표하며 이용자와 창작자의 의견을 신중하게 반영하고 있음 (Youtube Official Blog, 2024.2.6)

**[생성형 AI과 방송영상 IP]**

● 레거시 미디어 사업자는 생성형 AI 모델 훈련을 위한 콘텐츠 라이선스 계약을 체결하거나, 자사 콘텐츠로 파인튜닝한 생성형 AI 모델을 개발하기 위해 제휴를 맺는 등 생성형 AI 사업자와 협업을 확대하고 있음

- (AI 모델 훈련) 2024년 8월, Condé Nast는 Vogue, Wired, New Yorker 등의 텍스트 데이터를 제공하고 대가를 받는 계약을 OpenAI와 체결했으며, TIME, News Corp, Financial Times 등도 유사한 콘텐츠 라이선스 계약을 맺고 있음

※ 다만, 방송이나 영화 제작사가 생성형 AI 모델 훈련을 위해 콘텐츠를 제공한 사례는 없으며, 이는 출연자, 음원 저작권자, 협찬사 등 다양한 주체들의 법적 권리문제를 해결하기가 현실적으로 어렵기 때문으로 분석됨 (Butler, 2024.6.12.)

- (파인튜닝 모델 개발) 2024년 6월, 런웨이가 Gen-3 Alpha를 출시하며 기업 대상 서비스를 제공할 계획을 밝힌 데 이어, 9월에는 '존 윅'과 '헝거게임'으로 유명한 영화 스튜디오 라이온스게이트가 자사 영화와 방송 콘텐츠로 파인튜닝한 영상 생성 모델 개발을 위해 런웨이와 독점적 파트너십을 체결했다고 발표해 주목을 받음 (Lionsgate, 2024)

※ 그 밖에도 OpenAI는 2024년 Sora 출시 후 할리우드 스튜디오에 제품을 피칭했다고 알려졌으며(Bloomberg, 2024.3.22.), 구글과 메타도 할리우드 제작사와의 제휴를 추진 중 (Bloomberg, 2024.5.23.)

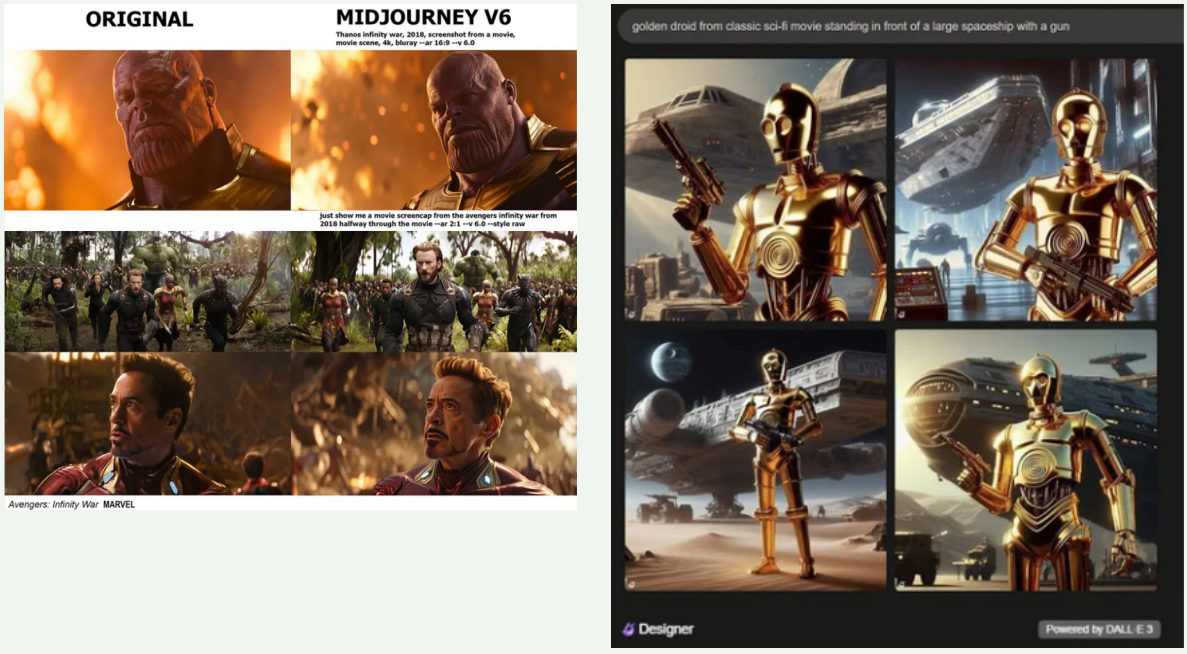
- AI 기술 사업자와 미디어 기업 간의 파트너십의 증가로 향후 생성형 AI 기반 콘텐츠 생산 및 배급을 혁신하기 위한 협력이 더욱 활발해질 것으로 예상됨

● 다만, 생성형 AI의 창작물이 불법적인 데이터로 학습된 경우, 지적재산권(IP) 침해에 대한 법적 책임 문제가 발생할 수 있어 생성형 AI와 관련된 쟁점 중 하나로 부상하고 있음

- 2024년 6월 미국 음반산업 협회(RIAA) 주도하에 소니뮤직, 유니버셜 뮤직, 워너뮤직의 음원 무단 활용 및 사용 협의로 수노(Suno)와 유디오(Udio)에게 손해배상 청구 소송 제기 중이며(한국경제, 2024.6.25.), 비슷한 사례로 한국저작권보호원은 2024년 3월 28일~4월 11일 유튜브 내 AI 커버곡을 전수 조사한 결과, 총 1647건 중 무단으로 목소리가 이용된 가수는 총 254명으로 AI 커버 곡의 저작권 침해의 문제성을 제기함(한국저작권보호원, 2024)
- 전기전자공학자협회(IEEE)에서 발표한 실험에 따르면, 미드저니와 DALL-E3에서 특정 프롬프트가 저작권으로 보호되는 영화 장면을 재현하라는 지시 없이도 거의 동일한 이미지를 생성할 수 있다는 사실이 발견됨 (Marcus & Southen, 2024)

- 즉, 원저작자의 허락 없이 무단으로 사용된 데이터나 유명인의 사진, 영상, 음성 등 퍼블리시티권이 보호되는 자료가 생성형 AI 학습에 활용된 경우, 법적 책임 문제가 발생할 수 있음
- 또한, 생성형 AI가 생성한 저작물의 저작권 문제도 여전히 해결되지 않은 상태로, 저작권 소유자를 법적으로 확인할 수 없는 경우 해당 저작물이 저작권 보호를 받을 수 없는 상황이 발생할 수 있음

**| 그림 7 | 영화 장면과 미드저니 V6, DALL-E3 생성물의 유사성 비교**




자료: Marcus & Southen(2024)에서 일부 발췌
주: [좌] 실제 영화 장면과 미드저니 V6의 이미지 생성물 비교. 각 프롬프트는 "타노스 인피니티 워, 2018, 영화 스크린샷, movie scene, 4K, blueray, --ar 16:9 --v 6.0", "어벤져스 인피니티 워, 2018 영화 중간의 스크린캡처를 보여줘, --ar 2:1 --v6.0 --style raw",
[우] DALL-E 3의 이미지 생성물, 프롬프트는 "클래식 공상과학 영화에 등장하는 금색 드로이드가 총을 들고 거대 우주선 앞에 서있다"

**[결론]**

- 생성형 AI는 방송 및 창작 산업에서 새로운 기회를 제공하고 있지만, 동시에 창작자의 권리나 지적재산권(IP) 침해 등 다양한 위기를 동반하고 있음
- AI 기술의 발전과 함께, 창작자의 권리를 보호하고 공정한 대가를 보장하는 법적·정책적 장치 마련이 필수적이며, 지속 가능한 창작 환경을 조성하기 위한 노력이 필요함

## 참고문헌

**[국내 문헌]**

한국경제 (2024.6.25.) 美 대형 음반업체들 "AI 스타트업이 저작권 침해" 소송

한국일보 (2024.6.15.) '눈물의 여왕' 속 자작나무 숲 장면의 비밀… "생성형 AI가 알아서 그렸다"

한국저작권보호원. (2024). AI 커버곡과 저작권 침해. C STORY, 44호. 한국저작권보호원.

AI 타임즈 (2022.12.28.) 세계 최초 'AI 보이스 디에이징' 기술로 재현한 30대 최민식 '카지노'에서 빛났다

AI 타임즈 (2024.7.23.) MCA, 태주산업 멀티탭 홍보 영상에 AI 나문희 활용

AI타임즈 (2023.4.14.) 홈쇼핑업계 최초, 생성 AI로 창작된 배경음악(BGM) 활용한 방송… KT알파 쇼핑에서 선보인다

CJ 올리브네트웍스 보도자료 (2022.7.5.) tvN 화면 해설 방송에 'AI 보이스 클로닝' 기술 제공

YTN (2021.9.30.) 故임윤택 목소리 다시 듣는다…JTBC, AI 음악 프로 '얼라이브' 론칭

**[해외 문헌]**

Amankwah-Amoah, J., Abdalla, S., Mogaji, E., Elbanna, A., & Dwivedi, Y. K. (2024). The impending disruption of creative industries by generative AI: Opportunities, challenges, and research agenda. International Journal of Information Management, 79, 102759. https://doi.org/10.1016/j.ijinfomgt.2024.102759

Bloomberg (2024.3.22.). OpenAI Courts Hollywood in Meetings with Film Studios, Directors.

Bloomberg (2024.5.23.). Alphabet, Meta Offer Millions to Partner with Hollywood on AI.

Butler, J. (2024.6.12.) Licensing Your Movie & TV Content for AI Training: Can You? Should You?. Variety.

CNBC (2024.9.24.). Warner Bros. Discovery partners with Google for AI-generated captions on Max streaming service

Google Deepmind (2024.9.18). Empowering YouTube creators with generative AI.

Lionsgate (2024.9.18.) Runway Partners with Lionsgate in First-of-its-Kind AI Collaboration.

Marcus, G., & Southen, R. (2024) Generative AI Has a Visual Plagiarism Problem Experiments with Midjourney and DALL-E 3 show a copyright minefield. IEEE Spectrum.

Omdia (2024). Artificial Intelligence Software Market Forecasts

Rehm, G. (2020), Research for CULT Committee –The use of Artificial Intelligence in the Audiovisual Sector, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.

Reuters (2023.6.30.). AI company Runway valued at $1.5 billion in latest funding.

The Stanford Institute for Human-Centered Artificial Intelligence (2024). AI Index 2024. Retrieved from https://aiindex.stanford.edu/report/

Variety (2023.2.22.) 'Hollywood 2.0': How the Rise of AI Tools Like Runway Are Changing Filmmaking.

Variety (2023.4.). Generative AI in Film & TV.

World Economic Forum (2023.5.9.) How might generative AI change creative jobs?

Writer Guild of America West Inc et al. (2023.9.25). Memorendum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement

Youtube Official Blog (2024.2.6.) Letter from the YouTube CEO: 4 Big bets for 2024.

## KISDI Perspectives 발간 내역

KISDI PERSPECTIVES는 국내 외 정보통신방송 관련 주요 정책 및 시장 동향을 분석한 리포트입니다.

문의 : 노희윤 전문연구원 (정보통신정책연구원 방송미디어연구본부, hyoooon@kisdi.re.kr, 043-531-4042)