

콘텐츠 산업의 새로운 문법, 시가 여는 창작 혁신의 시대

양지훈

한국문화관광연구원 부연구위원

yangjh@kcti.re.kr



기술이 이끌어 온 콘텐츠 산업의 진화

콘텐츠 산업의 역사는 새로운 기술의 도입과 대응의 연속이었다. 새로운 혁신 기술이 등장할 때마다 콘텐츠의 제작·유통·소비 구조 자체가 근본적으로 바뀌었고, 그때마다 산업의 중심축도 이동했다. 인쇄술의 발명은 콘텐츠 향유의 물리적 한계를 넘어서게 만들었고, 전파 기술의 발전은 라디오와 TV를 통해 각 가정의 거실을 엔터테인먼트의 주 무대로 바꾸었다. 기술은 단순한 도구가 아니라, 콘텐츠 산업의 규칙과 질서를 새로 쓰는, ‘산업의 문법’을 바꾸는 동력이 되어왔다.

특히 2000년대 이후 본격화된 디지털 전환은 콘텐츠 산업의 패러다임을 근본적으로 재편했다. 인터넷의 확산은 콘텐츠의 생산과 소비 주체를 대중으로 확대시켰으며, 스마트폰과 모바일 네트워크의 보급은 콘텐츠 소비의 시공간적 제약을 완전히 해체했다. 방송국과 영화 배급사 중심의 폐쇄적 구조는 유튜브,

블로그, SNS를 중심으로 한 개방형 플랫폼 생태계로 전환되었고, 개인은 곧 하나의 제작자가 되었다. 이러한 변화는 기술적 진보를 넘어 콘텐츠 생산 주체의 민주화와 소비 구조의 실시간화를 이끌었다. 그 과정에서 스트리밍 서비스의 등장은 방송 편성표와 영화 홀드백 제도를 무력화하며, 사용자의 선택과 알고리즘이 이끄는 소비문화를 정착시켰다. 콘텐츠는 더 이상 단순한 창작물이 아닌, 데이터와 알고리즘이 결합된 산업 자산으로 진화하기에 이르렀다. 즉, 디지털 기술의 심화는 창작의 방식뿐 아니라, 유통의 논리와 산업 구조까지 동시에 바꾼 것이다.

그리고 지금, 콘텐츠 산업은 또 한 번의 변곡점을 맞이하고 있다. 그리고 그 중심에는 인공지능(AI)이 위치하고 있다. 생성형 AI는 텍스트를 넘어 영상, 음성, 편집, 후반작업에 이르기까지 콘텐츠 산업의 전 과정을 재구성하고 있다. 과거 카메라가 시각의 문법을, 인터넷이 유통의 문법을 바꿨듯, AI는 창작의 속도와 효율성이라는 새로운 변수를 통해 산업의 문법을 다시 쓰고 있다.

콘텐츠 산업의 AI 시도 현황

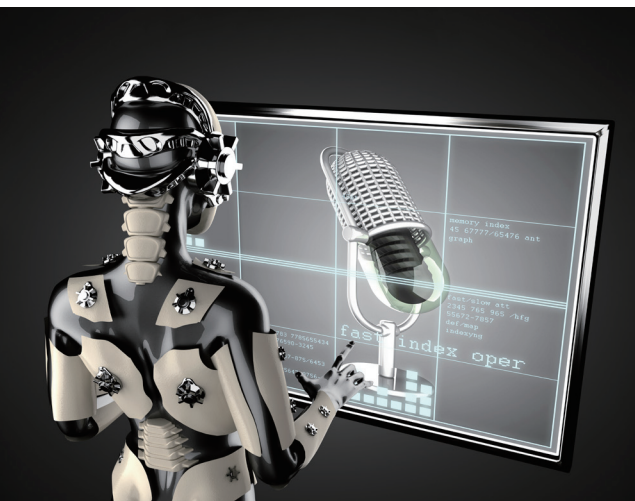
AI는 현재 콘텐츠 산업의 거의 모든 과정에 빠르게 확산되고 있다. 기획 단계에서는 아이디어 구상과 시각화 과정에서 AI의 개입이 일상화되고 있다. 제작 단계에서도 단순한 텍스트 명령만으로 영상을 생성할 수 있는 Text-to-Video(T2V) 서비스들이 전통적인 제작 프로세스에 도전하고 있기도 하다. 기획-촬영-편집으로 이어지던 선형 구조가 단축되면서, 아이디어가 곧바로 시각화되는 AI 기반 워크플로우가 등장한 것이다.

대표적인 서비스로는 오픈AI의 Sora, 구글의 Veo 3, 메타의 Emu Video가 있다. 이들은 고해상도 영상, 사실적인 질감, 자연스러운 카메라 무빙을 구현하며, 짧은 시간 안에 높은 품질의 결과물을 생성할 수 있다. 이러한 기술은 대규모 데이터셋과 막대한 연산 자원을 필요로 하지만, 그만큼 영상 제작의 효율성과 접근성을 동시에 높이고 있다. 스타트업 영역에서도 Runway, Pika, Stable Video Diffusion 등이 T2V 제작 시장을 확대하며, SNS 숏폼과 광고 중심의 제작 환경에서 AI 영상이 빠르게 정착되고 있다.

한편 하나의 대형 모델에 의존하지 않고, 여러 모델을 조합하는 워크플로우형 T2V 서비스도 부상하고 있다. 해외의 LMArena나 Kaiber, 국내의 브루(BRUE) 등이 그 예다. 이들은 사용자가 원하는 영상 스타일과 장면 단위를 조합하여 맞춤형 결과물을 생성할 수 있도록 지원한다. 특히 브루는 해외 모델을 국내 사용자 환경에 최적화해, 전문 지식이 없어도 누구나 영상 제작에 접근할 수 있는 플랫폼형 생태계를 구축하고 있다. 다만 대부분의 T2V 서비스는 여전히 영상 길이가 수십 초 내외로 제한되며, 인물·배경·조명의 일관성이 완벽히 유지되지 못하는 한계가 있다. 이러한 이유로 현재 AI 영상은 주로 숏폼, 광고, 예술 실험 등 짧은 형식의 콘텐츠에 집중되고 있다.

영상 제작에서 AI가 가장 빠르게 실용화되고 있는 영역은 편집과 후반작업이다. 런웨이(Runway)나 디스크립트(Descript)와 같은 서비스는 컷 분할, 자막 생성, 배경 합성 같은 반복적 작업을 자동화하여 제작 시간을 대폭 단축시킨다. 국내에서도 곰믹스AI, 딥미디어툴즈, 트웰브랩스 등이 등장하여 편집 효율성을 높이고 있다. 특히 트웰브랩스는 자사 멀티모달 인식 모델을 기반으로 긴 영상을 자동 요약하거나 특정 장면을 검색·추출할 수 있는 기능을 제공하여 뉴스, 교육, 마케팅 등 다양한 분야에 활용되고 있다. 이러한 기술은 영상 편집의 전문성을 대체하기보다, 제작 속도와 효율성을 극대화하는 보조 인프라로서 의미가 크다.

AI는 또한 음성·언어 기반 제작 환경에서도 중요한 역할을 하고 있다. ElevenLabs, Papercup, Synthesia 등의 서비스들은 대본 입력만으로 실제 성우에 가까운 음성을 생성하고, 감정 톤과 억양까지 반영할 수 있다. 국내의 딥브레인AI, 네이버 클로바 더빙, 브루 역시 다국어 더빙·자막·번역 기능을 통합 제공하며, 중소 제작자나 1인 크리에이터의 글로벌 진출을 지원하고 있다. 이들 기술은 제작비를 절감하고 언어 장벽을 허물어 콘텐츠의 해외 유통을 촉진하고 있다.



나아가 AI는 버추얼 휴먼(Virtual Human) 및 합성 영상 제작 단계로 발전하고 있다. 실제 존재하지 않는 인물을 사실적으로 구현하거나, 기존 인물의 모습을 재현하여 영상에 등장시키는 기술이다. 해외에서는 Synthesia, Rephrase.ai, Hour One 등이 선도하고 있으며, 사용자는 간단한 텍스트 입력만으로 사람처럼 말하고 제스처를 취하는 영상을 제작할 수 있다. 국내에서는 딥브레인AI가 뉴스 아나운서형 버추얼 휴먼을 상용화하였고, 플루닛 스튜디오는 K-팝 아이돌과 협업 가능한 가상 인플루언서를 개발해 엔터테인먼트 영역으로 확장하고 있기도 하다.

AI 시도의 시행착오와 한계

AI는 콘텐츠 제작의 속도와 효율성을 크게 끌어올리고 있지만, 여전히 현장에서는 수많은 시행착오가 발생하고 있다. 특히 일관성(consistency)과 서사적 완성도(narrative integrity)의 부족은 영상 콘텐츠 제작 단계에서 가장 본질적인 한계로 지적되고 있다. 생성형 AI가 단일 컷이나 장면에서는 높은 수준의 사실성을 구현하지만, 인물·조명·배경·카메라 시점이 바뀌는 순간 그 연결이 끊기곤 한다. 동일 인물의 얼굴이 장면마다 미세하게 달라지거나, 배경의 소품 배치가 바뀌고, 카메라 시점이 전환되면 인물의 표정과 조명이

불안정해지는 식이다. 이로 인해 장편 영화나 드라마처럼 시간의 연속성이 핵심인 서사형 콘텐츠에서는 시각적 몰입이 무너지고, 장면 간의 감정선이 자연스럽게 이어지지 않는다.

이 문제는 단순히 기술적 미세조정 수준이 아니라, AI가 맥락을 기억하지 못한다는 구조적 한계에서 비롯된다. 현재의 대부분 생성형 모델들은 장면을 독립된 확률 단위로 계산하기 때문에, 캐릭터의 감정, 상황의 톤, 공간의 구조 같은 서사적 일관성을 이해하고 반영하기 어렵다. 예를 들어, 한 장면에서 인물이 울고 있었다면 다음 장면에서는 같은 표정·의상·조명을 유지해야 하지만, AI는 이를 스스로 추적하지 못해 감정선의 비약이 발생한다. 이러한 이유로 감독이나 편집자는 여전히 프레임 단위로 수정을 반복해야 하며, AI 영상은 장편보다는 숏폼·광고·SNS 영상처럼 단발적 감정 전달에 초점이 맞춰지고 있다.

두 번째 한계는 콘텐츠 제작 프로세스의 단절이다. AI는 기획, 제작, 후반 등 개별 단계에서는 효율을 높이지만, 아직 전체 워크플로우를 하나로 연결하지 못하고 있다. 예를 들어 T2V(Text-to-Video) 모델로 생성한 영상은 후반 편집 툴에서 바로 호환되지 않거나, 음성·더빙 모델과 연동 과정에서 품질 손실이 발생한다. AI별로 데이터 포맷이 달라 AI 간의 호환성(Interoperability)이 부족하기 때문이다. 이로 인해 실제 현장에서는 여전히 사람이 각 단계를 수동으로 조정하며, AI의 자동화 효과는 제한적인 상황이다. 콘텐츠 산업에서의 AI 활용은 결국 부분적 자동화에 머물러 있는 과도기적 단계라 할 수 있다.

‘나노 바나나’가 연 콘텐츠 AI 특이점과 가능성

앞서 살펴본 바와 같이 AI는 콘텐츠 제작의 많은 과정을 자동화하고 있지만, 여전히 일관성(consistency)의 벽을 완전히 넘지는 못하고 있다. 이 같은 제약이 AI 영상 산업의 가장 큰 병목으로 작용하고 있다.

다만, 이와 같은 국면을 흔들 만한 모델이 등장했는데 2025년 8월 등장한 구글의 나노 바나나(Nano Banana)라는 모델이다. 나노 바나나는 구글 제미니(Gemini) 기반의 영상 생성 모델로, 기존의 프레임 단위 생성 방식을 넘어 장면 단위의 일관성 유지와 공간 인식을 구현한 점이 특징이다. 동일 인물의 얼굴 윤곽, 의상, 소품, 조명, 배경 질감을 장면 전체에서 통일성 있게 표현하고, 시점 전환 시에도 디테일이 유지되어 자연스러운 시퀀스를 구성할 수 있게 되었다. 이를 통해 프리비주얼 단계에서 생성한 이미지가 실제 시퀀스 제작으로 자연스럽게 확장될 수 있게 되었으며, 결국 아이디어 스케치형 생성에서 서사형 제작으로의 진화가 현실화된 것이다. 더 주목할 점은 나노 바나나의 공간 인식 능력이다. 동일 캐릭터의 정면·측면·후면을 일관된 형태로 생성하여 3D 피규어나 다면도 모델링에 바로 활용할 수 있고, 이를 영상으로 전환하면 실제 다카메라 촬영에 근접한 입체감을 제공한다. 이 덕분에 콘티 단계 이미지가 기획·캐스팅 단계의 의사결정 자료로 직결될 수 있게 되었다. SNS에서는 #nanobanana 태그를 단 아바타 이미지 확산 사례가 폭발적으로 늘었고, LMArena 등 AI 모델 비교 플랫폼에서 지속적으로 상위권을

점유했다. 이는 단순한 mim을 넘어 AI 영상이 본편 제작 단계로 진입할 수 있다는 기술적 징후로 읽힌다.

나노 바나나의 가치는 단일 모델의 성능 향상을 넘는다. 다른 생성형 서비스와 결합한 통합 워크플로우를 가능하게 하며 영상 제작 전 과정을 가상 환경에서 구현할 수 있게 한다. 예컨대, 나노 바나나로 일관된 시퀀스를 생성하고 → 구글 Veo 모델을 활용하여 장면을 영상화한 뒤 → ElevenLabs의 음성합성(Speech-to-Voice)으로 대사·내레이션을 입히는 조합이 대표적이다. 이 결합형 프로세스는 기존 촬영 현장 없이도 프리비주얼 → 촬영 → 후반 전 단계를 디지털 환경에서 재현하며, 제작 비용과 시간을 획기적으로 줄일 수 있다. 광고 및 숏폼 콘텐츠에서 검증된 안정성을 기반으로 한 장편 서사 제작으로 AI 기반 콘텐츠 제작이 확산되면, 품질 점프(quality leap)가 일어날 것으로 보인다. 이와 같은 워크플로우는 기술적 효율성 개선을 넘어 산업 구조 변화를 불러올 수 있다. 과거에는 대규모 자본과 전문 스튜디오가 필수였던 제작 환경이, 이제는 독립 창작자나 소규모 프로덕션에도 열린 생태계로 전환되고 있는 것이다. AI 모델 결합형 제작 툴킷은 콘텐츠 산업 전반의 창작 민주화를 더욱 촉진하며, OTT 시리즈·게임 시네마틱·뮤직비디오 등으로 활용 영역을 넓힐 것으로 예상된다.

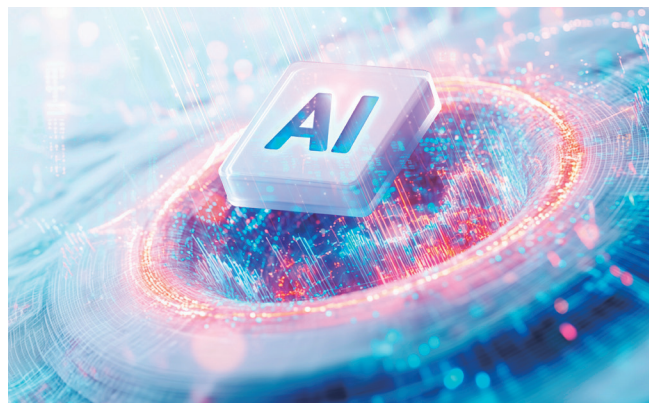
이 변화는 단순한 기술 업그레이드가 아니라, 콘텐츠 제작의 문법을 바꾸는 조짐으로 읽힌다. 일관성 문제의 해소는 곧 AI 영상의 제작 단계 진입을 가능하게 하며, 이는 콘텐츠 산업의 제작 효율성과 인력 구조에 실질적 변화를 예고하기 때문이다. 결국 나노 바나나는 콘텐츠 산업이 맞이한 AI 제작의 특이점(Singularity)을 상징한다고 할 수 있다. AI가 보조 도구의 영역을 넘어, 본격적인 창작 인프라로 편입될 가능성을 보여주기 시작한 것이다.

AI 혁신 시대 콘텐츠 산업의 도전과 방향성

AI의 확산은 단순히 새로운 기술의 등장을 넘어, 콘텐츠 산업의 작동 방식을 근본부터 다시 쓰는 혁신기를 열고 있다. 지금의 변화는 2000년대 중반, 모바일과 클라우드 기술의 확산이 넷플릭스·유튜브·페이스북을 탄생시켰던 시기와 닮아 있다. 당시 해당 기업들은 원천 기술을 직접 개발한 것은 아니었지만, 새로운 기술의 가능성을 어떻게 서비스 모델로 구현할 것인가에 집중했다. 스트리밍, 실시간 소셜 연결, 사용자 기반 추천과 같은 아이디어는 기술 그 자체보다 활용의 방식에서 비롯된 혁신이었다. 결국 콘텐츠 산업의 패러다임을 바꾼 것은 기술이 아니라, 기술을 해석하고 적용한 서비스 혁신의 전략이었다.

한국은 이러한 전환기에 강점을 지닌 나라다. 싸이월드는 페이스북보다 5년 앞서, 판도라TV는 유튜브보다 1년 먼저 등장했다. 카카오톡과 라인은 모바일 기술을 활용한 새로운 커뮤니케이션 방식을 제시하며 초기 글로벌 시장에 진출하기도 했다. 한국은 원천 기술을 가장 먼저 만드는 나라는 아니었지만, 새로운 기술을 빠르게 해석하고 서비스로 전환하는 민첩한 산업 DNA를 갖고 있다. AI 시대 역시 이 DNA가 다시 발휘될 때다.

AI는 더 이상 단순한 제작 효율화 도구가 아니다. 이제 AI는 기획·제작·후반작업 전 단계를 아우르며, 콘텐츠의 정의와 창작 주체, 그리고 소비 방식까지 다시 구성하고 있다. 기획 단계에서의 아이디어 시각화, 영상 편집의 자동화, 다국어 음성 변환, 버추얼 휴먼 구현에 이르기까지, AI는 창작의 속도와 실험의 빈도를 혁신적으로 높이고 있다. 결국 AI는 창작 방식을 바꾸는 기술이자, 콘텐츠 산업의 구조를 재편하는 동력이 되고 있다.



이 변화의 중심에는 두 가지 도전이 있다. 하나는 AI를 활용해 새로운 게임 체인저가 되려는 서비스 혁신의 도전, 다른 하나는 AI를 통해 새로운 문법의 크리에이터로 성장하려는 창작자의 도전이다. 전자는 콘텐츠 산업 생태계를, 후자는 창작 문화를 움직일 수 있다. 두 도전이 맞물릴 때 비로소 AI는 콘텐츠 산업의 혁신 인프라로 자리 잡을 것이다.

이를 위해서는 정부의 전략적 뒷받침이 절실하다. 이제는 과거처럼 제작비를 직접 지원하는 방식에서 벗어나, AI 기반 콘텐츠 서비스 기업과 창작 생태계를 함께 키우는 간접 지원 체계로의 전환이 필요하다. 예컨대 AI 콘텐츠 혁신 펀드를 통해 AI 기반 제작·서비스 스타트업을 육성하고, 창작자를 위한 교육·실험 프로그램을 확대해야 한다.

AI 변혁기는 기술의 발전 속도보다 활용의 전략이 승패를 좌우하는 시대다. 2000년대 앱스토어의 등장이 새로운 산업 주기를 열었듯, 이제 AI가 그 역할을 하고 있다. AI는 콘텐츠 제작의 방식을 바꾸는 데서 그치지 않고, 새로운 유통 질서와 창작 생태계를 만들어내는 플랫폼 혁신의 중심이 되고 있다.

한국이 가진 빠른 적응력과 창의적 실험정신, 그리고 정책적 실행력을 결합한다면, 이번 AI발 콘텐츠 혁신기는 단순한 기술 변화의 시기가 아니라, K-콘텐츠 산업이 세계의 새로운 제작 표준을 제시하는 도약의 순간이 될 것이다.

참고문헌

- AI 오픈소스와 미디어 콘텐츠 산업의 미래, 한국방송통신전파진흥원(2025)
- ChatGPT를 넘어 생성형(Generative) AI 시대로 : 미디어·콘텐츠 생성형 AI 서비스 사례와 경쟁력 확보 방안, 한국방송통신전파진흥원(2023)
- 동영상 제작 생성형 AI 서비스 <Sora>의 등장과 의미: 위기인가, 기회인가?, 한국방송통신전파진흥원(2024)
- 챗봇 이상의 파급력, 빅테크 격전지가 된 ‘동영상 생성형 AI’, 한국프로스포츠협회(2025)
- 윤상혁, 양지훈(2023), AI와 데이터 분석 기초: 디지털 비즈니스 생존전략, 박영사