

KCA 연구 2023

# 디지털 트랜스포메이션 시대, 영상·미디어 데이터 구축 및 활용 방안

(최종보고서)

2023. 12.

한국방송통신전파진흥원

연구수행기관 : (유) 디지털산업정책연구소



이 보고서는 한국방송통신전파진흥원의 재정지원으로 이루어졌으며, 보고서 내용은 연구자의 견해이며 한국 방송통신전파진흥원의 공식 입장과 다를 수 있습니다.



# 제 출 문

한국방송통신전파진흥원장 귀하

본 보고서를 『디지털 트랜스포메이션 시대, 영상·미디어 데이터 구축 및 활용 방안』의 최종보고서로 제출합니다.

2023년 12월

연구기관: (유)디지털산업정책연구소

총괄책임자: 천혜선 연구위원

참여연구원: 노창희 연구위원

윤금남 책임연구원

김나은 연구원



# 목 차

제1장 서론 .....	1
제1절 연구배경 및 필요성 .....	1
제2절 연구목표 및 범위 .....	3
제2장 국내외 데이터기반 인공지능 산업 및 기술 현황 분석 .....	5
제1절 데이터 기반 인공지능 개념 정의 및 생태계 .....	5
1. 인공지능 개념 및 유형 .....	5
2. 인공지능 기술 발전과정과 전망 .....	8
3. 인공지능 학습용 데이터의 필요성 .....	12
제2절 국내외 인공지능 기반기술 개발 현황 .....	17
1. 인공지능 기반기술 유형 .....	17
가. 컨볼루션 신경망(CNN) 모델 .....	17
나. 순환 신경망(RNN) 모델 .....	18
다. 파운데이션 모델과 대규모 언어모델(LLM) .....	19
2. 인공지능 기술유형별 영상미디어 적용가능성 평가 .....	20
가. CNN신경망 기술 적용 사례 .....	20
나. RNN 신경망 기술 적용 사례 .....	23
다. 대규모 언어모델(LLM) 기술 적용 사례 .....	24
라. 국내 기업의 대규모언어모델(LLM)중심의 개발 동향 .....	25
제3절 영상미디어 분야 AI 기술 적용현황 .....	28
1. 영상미디어분야 워크플로우와 AI활용성 .....	28
2. 워크플로우 단계별 데이터 적용 사례 분석 .....	32
가. 프리프로덕션 단계 .....	34

나. 프로덕션 단계 .....	35
다. 포스트 프로덕션 단계 .....	40
라. 마케팅 유통 단계 .....	42
<b>제3장 AI학습용 데이터 구축 및 활용 동향 분석 .....</b>	<b>45</b>
제1절 해외 데이터 구축·개방 현황 .....	45
1. 지능유형별 데이터세트 구축·개방 현황 .....	45
2. 해외 방송영상사업자의 영상학습 데이터세트 공개현황 .....	49
제2절 국내 데이터 구축·개방 현황 .....	51
1. 공공부문의 인공지능 학습용 데이터 구축 .....	51
가. AI학습데이터 오픈플랫폼(AI-Hub) .....	51
나. 방송통신분야 공공데이터 .....	55
2. 민간부문의 인공지능 학습용 데이터 구축 .....	56
가. 통신사 데이터 구축·공유현황 .....	56
나. 인터넷사업자의 데이터 구축·공유현황 .....	59
제3절 기존 데이터세트의 영상미디어 분야 AI학습 활용 가능성 평가 .....	61
<b>제4장 영상미디어분야 AI경쟁력 강화를 위한 학습데이터 수집·활용 활성화 방안 .....</b>	<b>64</b>
제1절 AI학습용 데이터구축 및 활용을 위한 제도개선 검토 .....	64
1. AI 학습용 데이터 수집·활용에 관한 제반 법률 현황 .....	65
가. 개인정보보호법 .....	65
나. 신용정보법 .....	70
다. 저작권법 .....	72
라. 데이터 산업법 .....	73
2. AI 학습용 데이터 수집·활용 확대를 위한 법제도 정비 방향 .....	74
가. AI 학습용 데이터 수집·활용의 법률적 제약요인 .....	74
나. AI 학습용 데이터 수집·활용을 위한 제도 정비 방향 .....	75

제2절 영상미디어분야 AI학습 데이터 구축 사업방안 .....	78
1. 인공지능 시대 국내 영상미디어분야 경쟁력 저해요인 분석 .....	78
가. 협소한 내수 시장의 한계 .....	78
나. 데이터 확보의 한계 .....	81
다. 인력 확보의 어려움 .....	84
라. 초거대AI 중심의 전략 추진시 지도학습용 데이터 수급 부족 우려 .....	85
2. 영상미디어 분야 데이터 구축 및 활용 활성화 방안 .....	87
가. 공유-협력에 기반한 AI 학습데이터 구축 전략 .....	87
나. 영상미디어분야 AI학습 데이터 구축 사업목표와 범위 .....	90
다. 영상미디어분야 학습데이터 구축 과제 .....	91
<b>참 고 문 헌</b> .....	<b>97</b>

## 표목차

<표 2-1> 인공지능의 유형과 요소기술 .....	6
<표 2-2> 인공지능 학습용 데이터와 빅데이터의 차이점 .....	13
<표 2-3> 인공지능 학습용 데이터 수집 및 가공 절차 .....	14
<표 2-4> 인공지능 학습용 데이터의 구축 절차 .....	14
<표 2-5> 영상 학습데이터세트 제작 도구 비교 .....	16
<표 2-6> LLM 기술의 영상·미디어 적용가능성 .....	24
<표 2-7> 네이버 생성형AI .....	26
<표 2-8> 방송통신 미디어 분야 산업계 대상 데이터 활용경험에 대한 조사 결과 ....	29
<표 2-9> 미디어 엔터테인먼트 산업의 가치사슬별 인공지능 활용사례 ...	30
<표 2-10> NBA에서 선보인 기업별 방송 분야 AI 기술 사례 .....	31
<표 2-11> 미디어 영역에서 활용 중인 생성형 AI 기술유형 .....	32
<표 2-12> 프리프로덕션 적용 사례 .....	35
<표 2-13> 영상 프로덕션 적용 사례 .....	35
<표 2-14> 페이스·음성 디에이징 AI기술 적용 사례 .....	38
<표 2-15> 포스트프로덕션 및 배급 적용 사례 .....	40
<표 2-16> 홍보마케팅 적용 사례 .....	43
<표 3-1> 인공지능 유형별 해외 주요 학습용 데이터 구축 현황 .....	45
<표 3-2> AI허브 데이터 개방 및 활용 현황 .....	52
<표 3-3> AI허브 데이터 분야와 종수 .....	53
<표 3-4> 방송영상미디어 데이터세트의 AI허브공개 현황 .....	54
<표 3-5> KT 통신빅데이터오픈랩 데이터세트 보유현황 .....	56
<표 3-6> SK오픈API의 API 및 데이터 구축현황 .....	58
<표 3-7> 네이버의 언어 윤리 데이터세트 구축 공개 현황 .....	60
<표 3-8> 개인정보보호법 상 가명정보·가명처리 관련 규정 .....	66

<표 3-9> 데이터 활용 관련 규정 .....	67
<표 3-10> 안정성 확보조치 관련 규정 .....	67
<표 3-11> 안전조치의무 관련 규정 .....	68
<표 3-12> 신용정보법 상 가명정보·가명처리 관련 규정 .....	70
<표 3-13> 신용정보법 상 개인신용정보의 활용 관련 규정 .....	70
<표 3-14> 가명처리 관련 처벌 규정 .....	71
<표 3-15> 공정한 저작물 이용 관련 규정 .....	72
<표 3-16> 데이터 이용 활성화 및 유통 관련 규정 .....	74
<표 4-1> 워크플로우별 AI 접목과 디지털 전환 관련 정책 과제 .....	80
<표 4-2> 개인정보의 보호와 활용 관련 규제정책 변동 내용 .....	81
<표 4-3> 개인정보보호법 상 민감정보의 처리 제한과 고유식별정보의 처리 제한 관련 규정 .....	82
<표 4-4> 인공지능 분야 인력수급 전망결과(2023-2027년) .....	84
<표 4-5> 영상미디어 분야 AI 학습데이터 구축 목표 .....	92
<표 4-6> 멀티모달 인공지능개발을 위한 데이터 구축 범위 .....	95

## 그림 목차

[그림 2-1] 구축목적에 따른 인공지능 유형 .....	8
[그림 2-2] 인공지능 기술 발전과정 .....	9
[그림 2-3] 인공지능 진화 방향 .....	10
[그림 2-4] AI파운데이션 모델의 진화 .....	11
[그림 2-5] CNN과 RNN을 사용한 이미지내 객체인식 및 상황설명 도출 개념 ·	19
[그림 2-6] 영상에서 문자열 감지 기술 개요 .....	22
[그림 2-7] 파운데이션모델 중심의 집중화 전망 .....	25
[그림 2-8] 영상미디어 분야 워크플로우 .....	28
[그림 2-9] 영상미디어 워크플로우 단계별 AI활용 현황① .....	33
[그림 2-10] 영상미디어 워크플로우 단계별 AI활용 현황② .....	34
[그림 2-11] 신세시아 디지털 휴먼 생성 사례 .....	36
[그림 3-1] KT의 통신빅데이터오픈랩의 데이터 서비스 .....	57
[그림 3-2]네이버의 언어윤리 데이터세트인 SQUARE의 데이터생성 파이프라인 ·	60
[그림 3-3] AI 허브 데이터세트 예시 .....	64
[그림 4-1] 국내 미디어 산업의 약점 .....	78
[그림 4-2] 국내 방송매출액 추이 .....	79
[그림 4-3] 네이버 2022년 사업부문별 손익 .....	79
[그림 4-4] 네이버의 주요 핀테크 사업영역 .....	80
[그림 4-5] 인공지능 사업 운영상 느끼는 애로사항: 인력 부족 .....	84
[그림 4-6] ‘초거대 AI 경쟁력 강화 방안’의 비전 및 추진전략 .....	86
[그림 4-7] ‘초거대 AI 경쟁력 강화 방안’의 데이터 구축 방향 .....	87
[그림 4-8] AI기술개발 프로세스에서의 영역특화기관의 역할 .....	90

## 요약문



## 1. 연구의 배경 및 목표

- (배경) 영상·미디어 분야에서 AI기술 개발 수요가 증가하고 있으나, 활용 가능한 AI학습용 데이터의 부족과 비용/시간 등의 부담으로 인해 중소기업 등에서는 관련 데이터의 확보가 어려움
  - 인공지능 시장은 기술 발전, 서비스 상용화·고도화에 따라 '25년까지 208~374조원 규모로 성장할 전망이다며 영상·미디어 분야에서도 CNN(Convolutional Neural Network, 이하 'CNN'), RNN(Recurrent Neural Network, 이하 'RNN'), GAN(Generative Adversarial Networks, 이하 'GAN') 모델을 사용한 인공지능기술 개발 수요가 증가함
  - 인공지능 기술이 확산·발전되기 위해서는 고품질·대규모 인공지능 학습용 데이터 확보가 필요하여 각 분야의 AI학습용 데이터 구축이 확대됨
  - 과학기술정보통신부의 지원으로 한국지능정보사회진흥원(NIA)이 AI학습용 데이터 구축사업을 추진하고 있으나, 영상·미디어 분야의 수요가 반영된 데이터 구축은 미흡한 상황임
- (필요성) 영상·미디어 분야의 학습용데이터 부족 문제 해소와 디지털 전환 가속화를 위해 데이터 자원 확보·활용 방안을 제시 필요
  - 데이터보유자가 AI 개발 산출물의 최종적인 활용자가 되는 영상·미디어 분야의 특수성을 고려하여, 데이터보유자-데이터활용자-잠재적인 AI기술활용자의 AI 기술 개발 방향과 전략 수립, AI학습용 데이터 수요 발굴, 영역 특화의 AI 학습용 데이터 공유·개방의 협력체계 구축이 필요함
  - 현재와 같이 방송사·통신사·인터넷기업 등이 개방한 원시데이터를 AI 학습용으로 단순가공하여 AI허브에 공개할 경우, 미상의 데이터 활용자에게 공개를 전제로 하는 AI허브의 특성상 원천데이터 보유자에게 해당 학습용 데이터의 사용으로 인해 발생하는 경제적 이익이 명확하지 않음

- 단기적으로 명확한 경제적 혜택이 보장되지 않고 AI학습의 결과물인 AI 기술에 대한 사용권 확보 등에 대한 권한 등에 대한 거래기준이 없으면, 양질의 원천데이터를 가진 민간 영상·미디어 기업의 자생적인 원천데이터 개방이 쉽지 않음

※ 국내 인공지능 도입의 장애물 인식(정보통신정책연구원, 2021) : 학습용 데이터 부족 (24.2%) 등

- 따라서, 산업의 변화 방향에 맞춰 중장기적인 시각에서 데이터의 수요를 발굴하고 데이터의 수집과 공개의 선순환적 생태계를 구축하기 위해서는, 정부나 공공기관이 이해당사자들의 관계를 조정하는 핵심적인 역할 수행이 가능한 데이터 생산·구축·개방 체계가 구축될 필요가 있음
  - 이에, 영상·미디어 분야의 산업발전 방향 및 특수성을 고려한 AI기술 개발 방향에 대한 비전을 제시하고, 영상·미디어 분야에 특화된 AI학습용 데이터 구축 목표와 방향을 제안하는 연구가 필요함
- o (목표) 본 연구는 영상·미디어 분야의 인공지능 활용 현황 및 전망분석을 토대로 영상·미디어 분야의 특수성을 고려한 AI학습용 데이터 구축의 방향, 전략, 학습용 데이터 구축사업을 발굴하는 데에 목표를 둠
- 이에 본 연구에서는 ①국내외 데이터 기반 인공지능 산업 및 기술현황을 분석하고, ② 영상·미디어 분야의 AI 데이터 활용방안을 도출하며, ③ AI학습용 데이터 구축사업 발굴 및 추진방안을 도출하고자 함

## 2. 국내외 데이터기반 인공지능 산업·기술 현황 분석

- o (인공지능 기반 기술 유형과 영상미디어 분야 적용사례) 영상미디어 분야에서 활용되는 인공지능 모델은 대표적으로 CNN, RNN, LLM기반 모델 등이 있음
- CNN은 이미지 분류 및 인식에 주로 사용되는 신경망의 한 종류로 시각적 데이터에서 객체, 클래스, 범주 인식을 위한 패턴을 찾고 오디오, 시계열 및 신호 데이터를 분류하는 등 영상의 주요 패턴을 인식할 수

있으므로 이미지 및 비디오 처리 작업에 특히 유용함

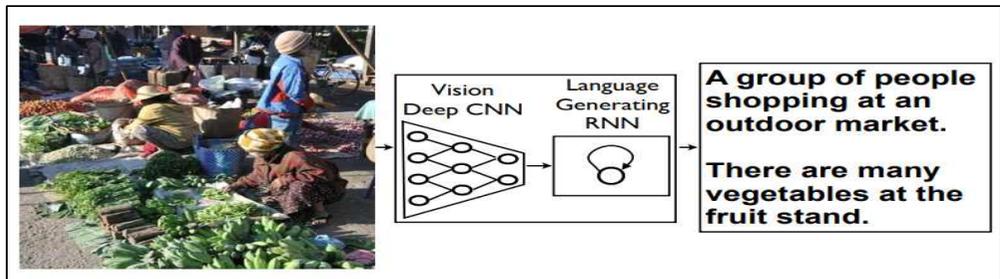
\* 대표적인 사례는 NHK의 CNN기반 메타데이터 생성시스템으로, NHK는 영상내에 간판이나 삽입된 문자를 인식해 영상메타데이터를 자동 부여함으로써 효율적인 영상 검색을 목표로 하고 있음

- RNN은 시퀀스 데이터를 처리하는 순환 신경망 클래스로, 문장 생성, 번역, 감정 분석 등에 사용이 가능하며, 영상미디어 분야에서도 챗봇이나 음성비서 등 개인화된 맞춤형 서비스 구현에 적용

\* 2016년도에 발표된 9분가량의 단편영화 ‘Sunspring’이 최초의 RNN을 활용한 시나리오 개발의 사례로 프리프로덕션 단계에서 대본작성의 보조적인 작업 도구로서 활용성이 있음

- 영상 및 이미지의 객체인식에 유리한 CNN과 시간적이고 순차적인 데이터 처리에 유리한 RNN을 동시에 사용하면, 영상 속의 상황을 문장으로 설명하는 자막이나 해설, 대본 구성 등이 가능해짐

[그림 2] CNN과 RNN을 사용한 이미지내 객체인식 및 상황설명 도출 개념



출처: Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). p. 3156).

- 최근 챗GPT 등장 이후 관심을 끌고 있는 LLM은 자연어 처리(NLP) 작업을 수행할 수 있는 딥러닝 알고리즘 모델로 챗봇, 가상 비서 또는 텍스트 기반 대화 시스템용으로 설계되어 고객센터나 타 인공지능 모델에 대한 텍스트 기반 대화 입력에 특화됨

- LLM은 현재까지 CNN이나 RNN에 비해 활발하게 개발되어 오지는 않았으나, RNN에 비해 자연어 처리에서 확장가능성이 높아, 향후 학습데이터에 따라 얼굴 인식, 객체 검출, 스타일 전송, 이미지 캡션, 비디오

생성 등의 다양한 작업을 수행할 수 있음

- o (워크플로우 단계별 AI활용성) AI는 새로운 캐릭터를 생성하거나 후반 작업 자동화 등 기존의 시간과 비용, 인력이 많이 소요된 영역을 중심으로 AI활용이 증가하고 있음
  - AI를 활용하여 시나리오를 작성하거나, 후반작업인 VFX 효과 적용에 있어 AI 활용을 하여 등장인물을 제작하는 등 영상·미디어 분야 워크플로우 전반에 있어 AI 활용이 증가할 전망이다
  - 또한, AI 기반 데이터 분석 결과를 마케팅 전략에 활용하는 등 제작된 콘텐츠의 마케팅과 유통 과정에도 AI 활용이 확대되고 있으며, 스트리밍 서비스 활성화에 따라 AI를 접목하여 개인 맞춤형 콘텐츠 추천 등 이용자 특화된 서비스를 제공하고 있음

[그림 3] 영상미디어 워크플로우 단계별 AI활용 현황②



출처: 과학기술정보통신부 (2023. 9. 12.).

- o (워크플로우 단계별 AI학습데이터 활용범위) 영상미디어 워크플로우 단순화 및 효율성 확대를 위해서는 현재 활용되고 있는 기술 발전을 통해 보편화할 수 있도록 방송에 특화된 AI 학습용 데이터 확보가 필수적

<표 1> 워크플로우 단계별 AI학습데이터 활용범위

분야		데이터 활용 범위
프리 프로덕션	기획	과거 영화 흥행 데이터를 토대로 영상 기획에 반영
	개발	시나리오 데이터 학습시켜 대본작성 자동화
	사전 준비	스크립북, 재무모델링, 캐릭터의 매력 점수 분석, 캐스팅 전략 도출
프로덕션	배경음악	음원 및 작곡데이터 학습으로 시기반 배경음악 생성

	장면 선택	신경망 분석을 통한 최적의 장면 및 예고편 추출
	촬영	기존 시각특수효과 데이터 학습으로 무인 영상 시스템 구축
포스트 프로덕션	특수효과	영화 속 인물 제작 및 특수효과 적용
	자막/번역	말뭉치/음성 데이터를 학습시켜 자막, 번역, 통역, 수어, 화면해설 자동화
마케팅/ 유통	콘텐츠 관리/추천	영상콘텐츠 지도학습데이터를 통해 영상으로부터 자동화된 메타데이터 생성 및 '이용자'의 만족도를 최대화하는 추천
	전략수립	효과적이었던 광고, 관객층과 지역 데이터 분석을 통해 홍보를 위한 최적화된 전략 도출

### 3. AI학습용 데이터 구축 및 활용 동향 분석

- o (지능별 데이터세트 현황) 현재 해외에는 시각지능, 언어지능, 청각지능 등 다양한 인공지능의 학습을 위한 데이터세트가 구축·공개되어 있으며, 영상 보다는 텍스트와 이미지 중심으로 데이터가 구축됨

<표 2> 인공지능 유형별 해외 주요 학습용 데이터 구축 현황

구분	데이터세트 명칭	구축 방법
시각 지능 (이미지인식)	ImageNet	- 웹에서 크롤링한 이미지 - WordNet 온톨로지를 사용하여, 20,000개 이상의 카테고리 이미지로 분류
	Open Image Dataset	-구글에서 제공하는 이미지 데이터세트 -Image-Region-Caption의 데이터구조를 가진 2만개 클래스, 900만개 이미지(url)
	COCO	- Flickr에서 크롤링한 이미지 - 일반이용자에게 객체와 그 객체들의 관계에 대한 주석을 달도록 한 참여적 데이터세트
	MNIST	- 핸드라이팅 숫자 이미지 데이터세트(6만장)
	CIFAR-10/100	- 8천만개의 저용량 이미지 데이터세트
	TACO	- 자원 봉사자들이 찍은 사진을 기반으로 구축
	DeepFashion	- 패션웹사이트 크롤링 이미지
	Casual Conversations	- Meta(Facebook)에서 자체 제작 - 다양한 연령, 성별, 조명 조건에서의 짧은 비디오 클립들을 포함
	CelebA	- 다양한 웹 사이트에서 유명인의 이미지를 크롤링하여 구축
시각 지능 (공간 인식/	KITTI (객체탐지)	- 주행 중인 차량에 설치된 센서들로부터 데이터를 수집하여 구축

이미지인식)	Kinectics	- DeepMind가 구축한 동작인식 학습데이터세트 - 700클래스, 65만개 이미지가 구축되어 있음
	MPII Human Pose	- Andriuka. et. al.이 학술목적으로 구축한 관절 관련 학습데이터세트
언어 지능 (텍스트 인식)	SQuAD	- Wikipedia 기사를 기반으로 한 데이터세트
	MS MARCO	- Microsoft에서 운용하는 Bing 사용자의 실제 검색 쿼리와 연결된 문서를 기반으로 구축
	OpenSubtitles	- 여러 언어의 영화와 TV 프로그램 자막을 모아 구축
청각 지능	LJSpeech	- 한 여성 화자의 오디오 북 클립들을 모아 구축한 청각지능용 학습데이터세트
	UrbanSound8K	- 다양한 도시 환경에서의 소리를 녹음하여 구축

o (해외 방송영상사업자의 학습데이터 구축 및 공개 현황) BBC와 YouTube, Meta가 연구 및 공공목적 활용을 전제로 데이터를 공개함

- BBC는 2022년에 AI학습용 데이터용으로 짧은 TV 프로그램의 영상 클립과 비하인드 스토리 및 메타데이터가 포함된 데이터세트 공개
- 단, 대학 및 학술 기관과 비상업적 연구를 위한 관련 공공 기관의 연구원만 사용하도록 허용했으며, 독립 연구자, 중등 교육 학생(예: 고등학생) 또는 상업 단체의 사용은 제한함
- 구글은 2019년 수백만 개의 YouTube 동영상 ID와 3,800개 이상의 영상으로부터 기계적으로 생성한 레이블포함된 ‘YouTube-8M Segment Dataset’을 공개함
- Meta는 인종이나 성별, 언어 등에 따라 차별하지 않도록 하는 공정한 AI개발을 위해 5,567명의 초상이 포함된 짧은 영상 데이터세트인 Casual Conversations을 공개함

o (국내 공공부문 학습데이터 구축 및 공개 현황) 한국지능정보사회진흥원의 AI 학습데이터 구축 과제를 추진하는 등 정부 주도형으로 AI 학습용 데이터 구축이 이뤄짐

- 현재 구축·개방된 영상 및 이미지 데이터세트 대부분이 유동인구 분석을 위한 CCTV영상데이터나 건설현장 장비 모니터링 및 생산성 측정을 위한 CCTV영상데이터나 건설현장 장비 모니터링 및 생산성 측정

데이터 등과 같이 자율주행차나 로봇틱스 적용도가 높은 데이터로 방송영상제작에 적용하기에는 어려움이 있음

- 순수 비디오만 있는 데이터세트는 15건에 불과하며, 객체 및 인물에 대한 감정, 동작, 성별 등의 라벨링 정보가 풍부한 동작영상 데이터, 멀티모달 데이터, 멀티모달영상 데이터세트가 부족함
- 방송사 영상으로부터 추출하여 구축·공개한 AI학습용 데이터세트는 15건에 불과하며, 한국어 통번역에 초점을 둔 음성데이터 중심임
- o (기존 데이터세트의 영상미디어 분야 AI학습 활용 가능성 평가) 영상미디어 제작과정에 필수적인 시각지능용 원시데이터가 부족하고, 영상미디어의 수요에 부합하는 라벨링이 이뤄지지 않음
- 방송영상제작에 사용할 수 있는 영상이미지 데이터세트가 양적으로 현저하게 부족하고, 저작권과 초상권 등의 문제가 없는 비전문적으로 제작된 짧은 비디오 클립 중심임
- 영상미디어분야 AI학습에 특화된 라벨링이 포함된 데이터세트가 부족하고 대부분 언어중심의 데이터세트로 영상과 이미지 라벨링은 미흡한 상황임
- 일부 통신사나 인터넷기업들이 원본데이터를 공개하더라도 민간사업의 특성상, 원천데이터 자체를 제공하기보다는 API만을 공개하여 데이터를 크롤링할 수 있는 횟수와 양을 한정하는 방식으로 서비스를 구축하거나 정형화된 빅데이터 중심으로 데이터를 구축·개방하는 경향이 있음

#### **4. 영상미디어분야 AI경쟁력 강화를 위한 학습데이터 수집·활용 활성화 방안**

##### **가. AI학습용 데이터구축 및 활용을 위한 제도개선 검토**

- o (AI 학습용 데이터 수집·활용의 법률적 제약요인) 국내 법률의 엄격한 적용과 해석으로 인해 AI 학습용 데이터 구축 및 활용이 활성화되지 못

하다는 지적

- 2020년에 데이터 3법이 통과되며 데이터 활용에 대한 기대가 있었으나, 여전히 데이터 수집 및 활용에 있어 개인을 식별할 수 없는 수준의 가명처리를 해야 하며 안전성 확보조치 및 기술적 보호조치 등의 의무로 인해 이를 검증하는 과정에서 시간·비용 등의 제약이 있음
- 국내 개인정보보호법에 규정된 동의제도는 개인정보자기결정권을 집중적으로 보호하기 때문에 비식별화를 하더라도 데이터를 이용하기 위해서는 사전에 필수적으로 정보주체의 동의를 구하여야 함
- 개인정보의 엄격한 통제로 인해 데이터 생태계 구축 및 산업 발전이 지체되는 한편, 형식적인 보호절차로 인해 개인의 정보도 실질적으로 보호되지 못하는 상황
- o (AI 학습용 데이터 수집/활용을 위한 제도 정비 방향) ① 규제 패러다임 전환, ②형식적 사전동의 규제 개선, ③가명처리 없는 AI 학습데이터 활용 등을 통해 데이터 수집 및 활용과 관련된 기존의 법 규제체계의 한계를 보완할 필요가 있음
- 데이터 수집 및 활용을 위해 기존의 경직된 법 규제체계를 AI 활성화, 빅데이터 등 기술의 발전을 전제로 한 규제로 패러다임 전환
- 개인의 식별가능성이 현저히 떨어지는 정보나 필수동의 사항에 대해서는 사전동의를 강제하지 않고, 옵트아웃 방식 등을 도입하여 데이터 활용을 촉진
- 단순히 AI 학습을 위한 경우에는 접근통제, 보호조치 등을 전제로 하여 가명처리 없이 개인정보를 활용할 수 있도록 개선 필요

#### **나. 영상미디어분야 AI학습 데이터 구축 사업목표와 범위**

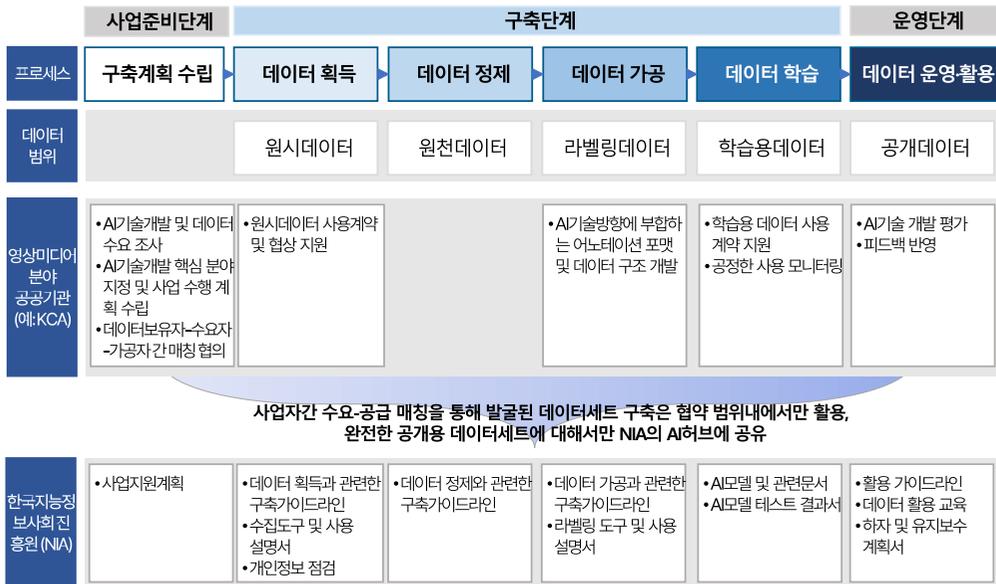
- o (영상미디어 분야 AI경쟁력 확보의 저해요인) 국내 내수시장 한계, 영상 데이터 확보가 어려운 법률적·제도적 장애, 영상미디어 영역 특화의 AI 기술 개발인력의 부족, 초거대 모델 위주의 정책 전환에 대한 우려 등이 데이터 구축의 잠재적 저해요인임

- 특히, 파운데이션 모델 중심으로 초거대AI플랫폼 경쟁력 확보에 집중된 중장기 AI경쟁력 확보 전략은 영상미디어 분야에 실질적인 AI수요와 괴리될 우려가 있음
  - 비 라벨링 텍스트 데이터 위주의 데이터 구축사업은 이미 도입된 AI모델을 활용하고자 하는 제작사, 영세한 개발사, 개인 개발자의 학습용 데이터 접근성을 저해할 우려
- (데이터이용자-데이터보유자 간 협력체계) 데이터이용자와 데이터보유자 간 공유·협력에 기반해 영상미디어분야의 AI기술 수요에 부합하는 AI 학습데이터 구축
- 영상미디어산업의 경우 가장 많은 양의 영상, 음성, 텍스트, 이용기록 등의 데이터를 보유한 방송영상관련 미디어 기업과 플랫폼으로, 데이터이용자인 AI개발자가 양질의 AI기술을 개발한다면 그 AI기술의 최종 수혜자가 되는 순환적인 구조임
  - 현재의 데이터 거래 플랫폼은 단순히 데이터를 일회적으로 제공하는 거래만 지원하고 있어, 데이터가공자와 데이터수요자(개발자/개발사)간 공개된 거래만을 매개할 뿐 최초 데이터 제공자가 데이터 제공 이후에 데이터 활용의 결실을 확인하기 어려운 구조임
  - 이에 데이터의 수집부터 최종 결과물인 AI기술의 활용에 이르기까지 데이터보유-데이터이용 합리적이고 공정한 기준으로 관리할 수 있도록 공공기관이 주도하는 데이터이용자와 데이터보유자 간 상생의 협력 체계를 구축이 필요함
  - 현재 AI학습용 데이터 구축 지원사업을 통해 KBS와 EBS, 공공기관이 보유한 일부의 영상이 활용된 바 있으나, 더 다양한 영상이미지 원천데이터 구축을 위해서는 정부나 공공기관이 사용자와 제공자간에 적법하고 합리적인 데이터 사용 협상에 관한 조정이 필요함
  - 이에 공공부문이 주도하여 산학연관의 협의체를 구성하고 데이터보유자들이 적지 않은 투자를 해 가며 축적·관리해 온 데이터의 활용결과

물에서 보유자들의 이익과 배치되지 않도록 하는 활용범위와 방안에 대한 투명한 관리가 필요함

- o (영역특화기관의 AI기술 전략수립에 따른 데이터 구축 추진) AI기술개발 수요와 데이터 필요분야 발굴부터 데이터산출물과 AI기술 활용의 전체 프로세스에서 주도적 역할을 담당할 수 있는 영역 특화기관이 데이터 보유자-수요자-가공자 간의 매칭을 주도
  - 데이터보유자와 가공자, 개발사의 3자간 수요-공급 매칭을 통해 발굴된 데이터세트 구축은 영역특화기관을 통해 별도의 협약 범위 내에서 활용하며, 영역특화 기관은 3자간의 협상과 데이터 구축 및 활용을 지원함

[그림 4] AI기술개발 프로세스에서의 영역특화기관의 역할



- 공공데이터 및 완전한 공개가 허용된 공개데이터에 기반한 AI학습용데이터 구축에 한해서 NIA가 구축한 AI허브에 공유하고, 데이터가공자와 데이터수요자간 공개적 거래는 AI허브를 통해서 추진함

다. 영상미디어분야 학습데이터 구축 과제

- o (영상미디어 특화 AI학습데이터 구축목표) 영상미디어 분야의 AI기술 발전 방향을 고려할 때, 학습데이터 구축의 목표는 각 영역별로 다음과 같이 전환이 필요함

<표 3> 영상미디어 분야 AI 학습데이터 구축 목표

분야	AS-IS	TO BE	목표
대상 인식	<ul style="list-style-type: none"> <li>• 한국인 안면인식 등 기초데이터 확보</li> <li>• 사람, 인물, 동물, 가구 등 구체적인 사물 이미지 데이터</li> <li>• 인물·객체 중심 3D데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 감성, 개념 등 추상화된 대상에 대한 시각적 이미지 데이터와 어노테이션</li> <li>• 배경·환경·건축물 등 비인물 랜드스케이프 3D공간 데이터 확보</li> <li>• 환경 맥락 내 랜드마크 이미지, 랜드마크 식별, 랜드마크 태깅</li> </ul>	인식능력 향상 및 3D파라미터 구현
영상 이해	<ul style="list-style-type: none"> <li>• 비전문적으로 제작된 짧은 영상</li> <li>• 영상 내 객체인식 중심용 어노테이션</li> <li>• 단일모드의 원천데이터</li> <li>• 연계성이 없이 개별적으로 구축된 동영상, 스크립트, 음성/오디오 데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 전문적으로 제작된 풀타임 방송영상 프로그램</li> <li>• 카테고리, 동작, 감정, 이벤트 등의 다계층적 어노테이션</li> <li>• 멀티모드의 원천데이터 (동영상-스크립트-음성/오디오 등)</li> <li>• 텍스트-영상-음성을 연동하여 정렬할 수 있는 연계된 데이터</li> </ul>	통합능력 확보
동작·상황 인식	<ul style="list-style-type: none"> <li>• 센서로 간단한 동작인식 데이터</li> <li>• 단순 행위 분류 어노테이션</li> </ul>	<ul style="list-style-type: none"> <li>• 동작·상황 등의 종합적 인지, 가상의 공간인식을 지원하는 데이터 구축</li> <li>• 동작과 상황에 대한 다층적 어노테이션</li> </ul>	상호작용능력 확보
상호작용 데이터	<ul style="list-style-type: none"> <li>• 좋아요, 댓글, 시청 시간 등 사용자 상호작용 데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 동공의 상대적 위치, 맥박 등 신체데이터</li> </ul>	상호작용능력 확보

- o (학습데이터 구축범위) 동영상-스크립트-음원-어노테이션-메타데이터가 모두 포함된 풀타임(full-time) 방송영상데이터 구축이 필요함

- 기존에 구축·공개된 영상 요약기술용 학습데이터는 짧으면 수초에서 수분 내외의 동영상 클립 중심으로 구축되어 있으며, 정규 방송프로그램이나 영화와 같은 저작권이 있는 양질의 기존 영상물들은 데이터 공

유가 이뤄지지 않음

- 상업적인 추가 수익화가 어려운 오래된 비인기 영상이라도 전문적인 제작사나 방송사가 제작한 풀타임 영상 데이터와 스크립트의 공유가 이뤄져야 내러티브에 기반한 영상 요약 기술 개발에 유리할 것으로 보임
- 이에 공공 스포츠이벤트로 중계권이나 저작권 프리가 확보된 영상, 영상저작물의 보호기간이 소멸되었거나 저작권 기부, 공공기관이 협약을 통해 저작권 공유 및 개방을 득한 전문적인 방송 및 영화 영상 등을 확보하여 풀타임 영상데이터와 스크립트의 공유가 필요함
- 더불어 해당 원천데이터에 대한 메타데이터 및 어노테이션 작업을 지원하여 지도학습용 자료로 개발이 필요함

<표 4> 멀티모달 인공지능개발을 위한 데이터 구축 범위

구분	내용
대규모 동영상 데이터	-전문적으로 제작된 풀타임 방송영상 프로그램 원시데이터 -감성, 개념 등 추상화된 대상에 대한 시각적 이미지 데이터 -비인물 랜드스케이프 3D공간 데이터
음성데이터	-텍스트를 시각적 콘텐츠와 정렬하는 데 사용할 수 있는 동영상의 음성 콘텐츠 녹음본
오디오데이터	-시각적 데이터와 함께 분석할 수 있는 사운드트랙, 대화 및 기타 오디오 요소
어노테이션(annotation)	-지도 학습의 기본 정보로 사용되는 동영상의 중요한 부분을 나타내는 사람, 객체 등을 설명하는 요약 또는 라벨 -동영상에서 발생하는 콘텐츠, 동작, 이벤트를 설명하는 어노테이션(annotation)이 포함된 대규모 동영상 데이터세트
설명텍스트 데이터	-AI가 학습할 수 있는 추가 컨텍스트를 제공할 수 있는 비디오 콘텐츠의 텍스트 설명 또는 요약정보 -동영상과 관련된 대본이나 설명은 동영상과 텍스트 정보를 결합하여 다중 모달 학습에 사용
메타데이터	-동영상 제목, 태그, 사용자 참여 측정항목과 같은 정보는 요약을 위한 추가 컨텍스트를 제공
사용자 상호 작용 데이터	- 좋아요, 댓글, 시청 시간 등 사용자가 동영상 콘텐츠와 상호 작용하는 방식에 대한 정보 - 동영상의 내러티브에 따른 동공의 상대적 위치나

	심장박동 등의 신체적 데이터 - 시가 동영상의 어떤 부분이 가장 매력적이고 중요한지 이해하는 데 활용
--	--



# 제1장 서론

## 제1절 연구배경 및 필요성

- (AI기술 활용 확대) AI기술이 인간의 고유영역이라고 여겨지는 창작부문에까지 확장됨에 따라, 영상·미디어 분야의 워크플로우에서도 AI기술을 활용하고자 하는 노력이 이뤄지고 있음
  - 최근 몇 년간 AI 기술은 딥러닝이라는 기계학습을 통한 학습능력 향상, 정보통신 기술의 발전, 컴퓨팅 능력의 향상과 비용 하락 등으로 빠르게 성장함
  - AI 기술 자체의 고도화와 함께 데이터, 클라우드, 초고속 유무선 통신 등의 기술이 기존의 미디어 영역에 본격적으로 적용되면서 콘텐츠의 기획-제작-편집-유통-소비에 이르는 미디어 가치사슬 전체에 중요한 생산도구로서 등장함
  - AI 기술이 인간의 창의력이 주요한 생산원천으로 생각되던 영상·미디어 산업 영역의 전반으로 확산됨에 따라, 향후 몇 년 안에 AI기술을 가진 사업자와 그렇지 못한 사업자(해외사업자 vs 국내사업자, 대형사업자 vs 중소기업자 간)와의 격차가 더 벌어져 미디어 생태계의 양극화가 더욱 심화될 것으로 전망됨
  - 이에 국내 영상미디어산업의 지속적인 성장을 위해서는 AI기술 경쟁력 확보가 필수적임
- (데이터확보 필요) AI 성능 보장을 위해서는 학습용 데이터 생산과 활용의 선순환적 구조 마련이 필요하며, 양질의 학습용 데이터 구축을 통해 영상미디어산업의 지속가능한 발전 기반 조성 필요
  - 국내 영상·미디어 분야의 디지털트랜스포메이션을 가속화하기 위해서는 광범위한 영상학습용 데이터 구축이 필요하나, 영상·미디어 분야에

서 활용할 수 있는 원천데이터를 보유하고 있는 기존의 방송사나 제작사 등의 영상데이터 보유자들의 자산을 공개하는 데에 소극적임

- 또한 영상·미디어 제작부문에서 인공지능 활용의 고도화를 위해서는 인공지능의 학습과 분석을 위한 데이터 확보가 필수적이거나, 민간의 개별기업이 인공지능 학습을 위한 데이터를 수집·구축하기에는 시간과 비용 부담이 클 뿐만 아니라 저작권법과 개인정보보호법 위반 등의 법률적 위험요소가 큼
- 이에 데이터 생산과 활용의 선순환적 구조를 마련하기 위해서는 데이터 제공자에게 적절한 경제적 혜택과 권리보호가 가능하도록 하는 생태계 조성이 필요하며, 생산·구축된 학습용 데이터를 지속적으로 업데이트하고 고도화 하는 방안과 더불어 인공지능 학습에 적합한 저작권 및 요금 체계 마련으로 데이터의 공유와 활용 간의 원활한 순환적·상생적 생태계 조성 등을 포함한 사업설계가 필요함
- 특히, 생성형 AI기술 도입으로 빠르게 변화하는 미디어 환경에서 경쟁우위를 지니기 위해서는 국내 기업들의 빠른 대응이 필요하나, 국내에는 각종 규제 및 양질의 AI학습용 데이터의 부족으로 생성형AI기술 부문에서의 AI모델의 성능제고에 어려움이 있음
- 또한 정제 후 바로 변환하여 저장하는 빅데이터와 달리 일반적으로 AI 학습용 데이터는 정제→가공→검사→학습의 과정을 거쳐 AI가 학습할 수 있는 라벨링이 필수적인데, 현재 국내기업이나 공공기관이 보유한 대부분의 데이터들은 라벨링데이터가 없는 원천데이터 중심인 경우가 많아 AI학습용으로서의 활용도가 낮음
- 따라서 영상·미디어 분야에서 활용할 수 있는 AI학습용데이터 확보로 국내 사업자와 중소기업의 지능정보기술 경쟁력 확보와 혁신 성장의 토대 마련이 필요함

○ (데이터전략 추진 필요) 디지털트랜스포메이션 시대에 데이터 기반 영상·미디어 신산업 창출 및 영상·미디어 분야의 차세대 경쟁력 확보를 위하여

데이터 활성화 정책 및 전략 추진이 필요한 상황임

- 인공지능 학습용 데이터 구축은 많은 시간과 비용이 소요되며, 이로 인해 자금력 부족과 각종 규제로 데이터 수집이 어려운 국내 기업들에게는 인공지능기반 제작기술 개발 및 도입의 주요한 진입장벽으로 작용함

※ 인공지능 도입의 장애물(정보통신정책연구원, 2021) : 학습용 데이터 부족(24.2%) 등

- 반면, 민간 자체적으로 AI학습용 데이터를 지속적으로 수집·개방하기에는 단기적으로 명확한 경제적 혜택이 보장되지 않을뿐만 아니라 데이터자산의 권리보호가 어렵고, 학습용 데이터의 수집·구축·개방을 민간에게만 의존할 경우 자생적인 데이터 공개가 이뤄지기 어렵고 데이터의 품질을 보증하기가 어려움
- 따라서, 산업의 변화방향에 맞춰 중장기적인 시각에서 데이터의 수요를 발굴하고 데이터의 수집과 공개의 선순환적 생태계를 구축하기 위해서는, 정부나 공공기관이 이해당사자들의 관계를 조정하는 핵심적인 역할수행이 가능한 데이터 생산·구축·개방 체계가 구축될 필요가 있음

## 제2절 연구목표 및 범위

- o (목표) 본 연구는 영상·미디어 분야의 인공지능 활용 현황 및 전망분석을 토대로 영상·미디어 분야의 특수성을 고려한 AI학습용 데이터 구축의 방향, 전략, 학습용 데이터 구축 사업을 발굴하는 데에 목표를 둠
  - 이에 본 연구에서는 ①국내외 데이터기반 인공지능 산업 및 기술현황을 분석하고, ② 영상·미디어 분야의 AI데이터 활용방안을 도출하며, ③ AI학습용 데이터 구축사업 발굴 및 추진방안을 도출하고자 함
- o (보고서의 구성) 본 보고서는 총 4개의 장으로 구성되며, 마지막 제4장에서 영상·미디어 분야 AI경쟁력 강화를 위한 학습데이터 수집·활용 활성화 방안을 제언함

- 제1장 서론에서는 연구배경, 연구목표, 보고서의 구성 등을 논의함
- 제2장 국내외 데이터기반 인공지능 산업·기술 현황 분석은 데이터기반 인공지능 개념 및 생태계를 정의하고, 국내외 기업의 인공지능 기반기술 개발 현황, 영상·미디어 분야 AI기술 적용현황 등을 논의함
- 특히, 제2장 제3절에서는 영상·미디어의 워크플로우 단계별 AI기술의 활용사례와 적용가능성을 진단함
- 이어서 제3장은 국내외 해외의 인공지능 학습용 데이터 구축 및 활용 동향을 파악하고, 기 구축된 데이터세트(dataset)의 영상·미디어 분야 AI학습 활용 가능성 평가함
- 마지막으로 제4장 영상·미디어 분야 AI경쟁력 강화를 위한 학습데이터 수·활용 활성화 방안은 AI학습용 데이터구축·활용을 위한 제도개선 방안을 검토하고 영상·미디어 분야에 특화된 AI학습 데이터 구축 사업방안을 제안함
- 특히, 이 장에서는 인공지능 시대 국내 영상·미디어 분야 경쟁력 저해요인을 분석하여, AI기술 개발의 중단기적 중점목표와 그에 따른 학습용 데이터 구축 과제를 제시함

## 제2장 국내외 데이터기반 인공지능 산업 및 기술 현황 분석

### 제1절 데이터 기반 인공지능 개념 정의 및 생태계

#### 1. 인공지능 개념 및 유형

- (정의) 인공지능은 인간의 지능과 유사하게 정보를 처리할 수 있는 컴퓨터 알고리즘을 의미하며, 인공지능기술의 유형은 솔루션 유형별, 세부 기술별, 구축목적별로 분류가 가능함
  - 인공지능(Artificial Intelligence, AI)은 인간의 학습능력, 추론능력, 이해능력, 인지능력 등과 같이 인간의 고차원적인 정보처리 능력을 구현하기 위한 ICT(Information & Communication Technology) 기술을 의미함
  - 인공지능기술의 유형은 솔루션 유형별, 세부 기술별, 구축목적별로 구분하는 것이 일반적임
- (솔루션 유형별 분류) 인공지능 생태계를 솔루션 유형별로는 하드웨어, 소프트웨어, 서비스로 분류할 수 있음
  - 하드웨어: AI를 지원하는 전용 칩 및 기기들, 예를 들어 GPU, TPU, FPGA 등이 있음
  - 소프트웨어 : 머신러닝 프레임워크(예: TensorFlow, PyTorch), 인공지능을 위한 서비스 및 API, 그리고 클라우드 기반 AI 서비스 등을 포함
  - 서비스: AI 구현 및 통합, 컨설팅, 교육 및 기타 관련 서비스를 포함
- (세부 기술별 유형분류) 인공지능은 인공지능이 추구하는 지능의 유형에 따라 학습지능, 단일지능, 복합지능으로 구분될 수 있으며, 지능 유형에 따라 관련 요소기술이 달라짐

<표 2-1> 인공지능의 유형과 요소기술

구분	세분류	요소기술
학습지능	머신러닝/딥러닝	베이지안 학습, 인공 신경망, 딥러닝, 강화학습, 앙상블러닝, 판단 근거 설명
	추론·지식표현	추론, 지식표현 및 온톨로지, 지식처리
단일지능	언어지능	언어분석, 의미이해, 대화 이해 및 생성, 자동통역·번역, 질의응답, 텍스트 요약 생성
	시각지능	영상처리 및 패턴인식, 객체 인식, 객체 탐지, 행동이해, 장소·장면이해, 비디오 분석 및 예측, 시공간 영상 이해, 비디오 요약
	청각지능	음성분석, 음성인식, 화자인식·적용, 음성합성, 오디오 색인 및 검색, 잡음처리 및 음원 분리, 음향 인식
복합지능	행동·소셜지능	공간지능, 운동지능, 소셜 지능, 협업지능
	상황·감정이해	감정이해, 사용자 의도 이해, 뇌신호 인지, 센서 데이터 이해, 오감인지, 다중상황판단
	지능형 에이전트	에이전트 플랫폼, 에이전트 기술, 게임 지능, 모방 창작 지능
	범용 인공지능(AGI)	상식학습, 범용문제 해결, 평생학습, 도덕·윤리·법·지능

출처: 권순선(2020), p.40(부분 발췌)

- 머신러닝(Machine Learning, ML): AI의 핵심 부분으로, 데이터를 기반으로 알고리즘이 스스로 학습하게 하는 기술로, 세부적으로는 인간개입의 범위에 따라 지도학습(Supervised Learning), 비지도 학습(Unsupervised Learning), 강화학습(Reinforcement Learning)으로 분류
- 딥러닝 (Deep Learning): 머신러닝의 하위 분야로, 인공 신경망을 깊게 쌓아 복잡한 패턴을 학습하게 하는 기술. 세부적으로는 CNN, RNN, 트랜스포머 모델 등의 기술을 포함함
- 언어지능: 자연어 처리(Natural Language Processing, 이하 'NLP')가 대표적인 방식이며 텍스트와 음성 데이터를 처리하고 이해하는 AI 기술로서, 챗봇, 기계번역, 텍스트 감정분석 등의 기술 영역을 포괄
- 시각지능: 컴퓨터 비전은 이미지나 비디오를 인식하고 처리하는 AI 기

술로, 얼굴인식, 물체인식, 장면인식 등이 요소기술을 포함됨

- 청각지능: 오디오 및 음성 데이터를 처리하고 이해하는 AI기술
  - 복합지능: 복합지능은 언어, 시각, 청각, 학습, 추론 등의 다양한 영역의 지능을 결합하여 상황을 파악하고 상황에 맞는 결과를 도출하는 능력을 가진 멀티모달 인공지능을 의미함
  - 복합지능 기술은 특정 문제만을 해결하는 좁은 인공지능(ANI: Artificial Narrow Intelligence)의 수준을 넘어, 궁극적으로 인공지능이 다양한 영역에서 인간과 유사한 지능을 가진 범용 인공지능을 구현하는 것으로 목표로 함
- (구축목적에 따른 유형) 인공지능은 구축목적에 따라 약인공지능(Weak AI, WAI)과 강인공지능(Strong AI, SAI)으로 구분될 수 있음
- 세부 기술별 유형에서 학습지능과 단일지능이 소위 ‘약 인공지능’에 해당하며 단일지능간의 연계가 되면 인간과 비슷한 수준의 지능을 가진 복합지능인 ‘강인공지능’으로 발전하게 됨
  - 약인공지능(Weak AI, WAI)은 특정 인지 기술이 필요한 작업을 자동화하도록 설계된 AI 앱으로서 객체 인식, 챗봇, 개인 음성 비서, Google 검색 알고리즘 등 특정 작업에 맞게 조정된 머신러닝 모델을 의미하며 Chat GPT, Midjourney, Stable Diffusion, DALL·E, Bard 등이 포함됨
  - 강인공지능(Strong AI, SAI)은 인간과 동등한 수준의 지능을 가진 인공지능을 의미하며 분석적 사고 및 기타 지적 능력 등을 갖추고 있는 것을 의미함
  - 영상·미디어 분야에서도 AI 활용이 점점 확대되고 있으나, 중단기적으로는 단일지능과 학습지능에 기반한 모델이 산업적 활용성이 가장 높으며, 중장기적으로는 생성형 모델의 발전속도에 따라 복합지능을 적용한 서비스 개발과 콘텐츠 제작이 이뤄질 것으로 전망됨

[그림 2-1] 구축목적에 따른 인공지능 유형

	Strong artificial intelligence	Weak artificial intelligence
Definition	<ul style="list-style-type: none"> <li>the form of artificial intelligence, which has the same intellectual abilities as human, or even surpasses him in it</li> </ul>	<ul style="list-style-type: none"> <li>Weak AI is generally developed or used for specific application domains.</li> <li>In a standard work on artificial intelligence, this is formulated as follows: "The assertion that machines could possibly act intelligently (called, weakness, act as if they are intelligent) is called the, weak AI 'hypothesis ..."</li> </ul>
Capabilities and Domains	<ul style="list-style-type: none"> <li>Logical thinking</li> <li>Making decisions in case of uncertainty</li> <li>To plan</li> <li>To learn</li> <li>Communication in natural language</li> <li>Use all these abilities to achieve a common goal</li> </ul>	<ul style="list-style-type: none"> <li>Expert systems</li> <li>Navigation systems</li> <li>Voice recognition</li> <li>Character recognition</li> <li>Suggestions for corrections in searches</li> </ul>

출처: MATEC Web of Conferences 121, 04007 (2017)

## 2. 인공지능 기술 발전과정과 전망

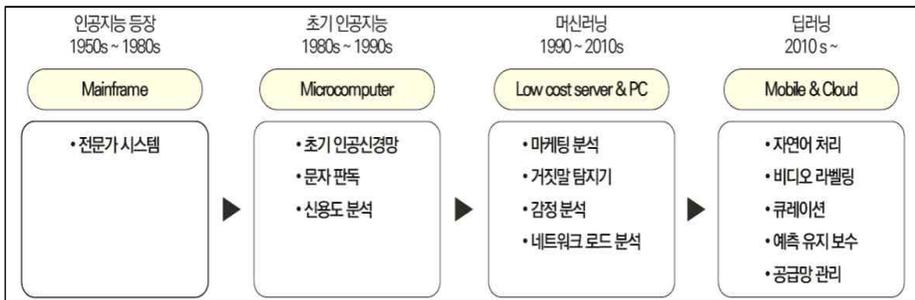
- (인공지능 개발 연혁) 인공지능에 관한 관심이 고조된 시기는 1950년대 중반부터로 그 역사가 깊으나, 딥러닝이 등장하면서 인공지능 기술이 비약적으로 발전하고 있음
  - 인공지능 초기 단계에서는 기호적 방법(symbolic methods)을 중심으로 인공지능을 연구했으며 이 시기에 인공지능 연구를 위한 최초의 프로그래밍 언어인 LISP가 탄생했음
  - 알란 튜링의 ‘튜링 테스트<sup>1)</sup>가’ 기호주의의 대표적인 사례가 될 수 있으며 이 시기의 기술들은 현재의 자율 주행이나 로봇틱스, 챗봇 등의 전신이 되었음
  - 1980년대부터 2000년대 초반에 걸쳐, 데이터를 기반으로 스스로 학습하는 알고리즘인 머신러닝 기술이 개발됐으며 이를 통해 AI는 사전에 정

1) 기계가 인간과 동등하거나 구별할 수 없을 정도로 지능적인 행동을 보이는지 테스트하는 것

의된 규칙이 아닌 데이터를 통해 스스로 판단이 가능해짐

- 2010년부터 현재에는 딥러닝이라는 개념이 등장하였으며 딥러닝이란 인간 뇌의 뉴런 구조를 모방한 인공 신경망 기반 학습 방법으로 빅데이터와 GPU 발전으로 급격하게 인공지능이 발전하게 되었음
- 이미지 및 음성 인식, 자연어 처리 등 다양한 분야에서 높은 활용성을 보여주며 현재의 AI 붐을 주도하고 있음

[그림 2-2]인공지능 기술 발전과정

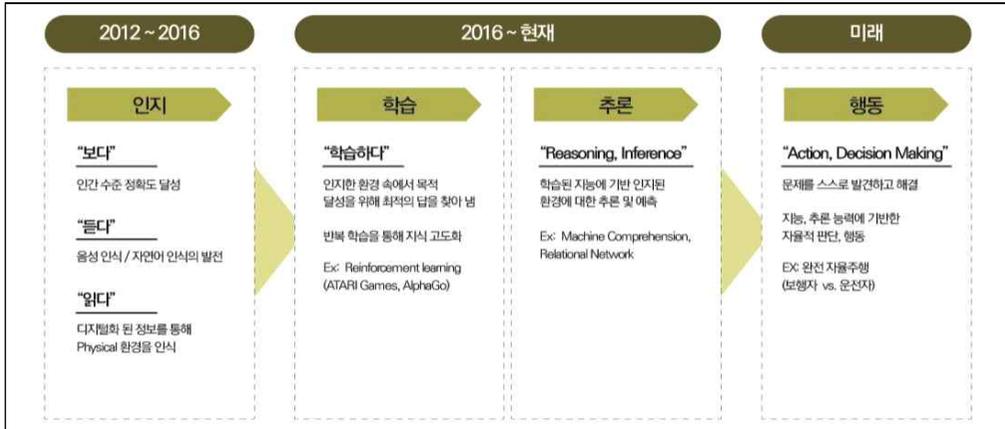


자료: 전승우, “인공지능(AI) 프로세서, 새로운 혁신의 원동력 될까”, LG경제연구원, 2018.11

- o (기술진화의 방향성) 인공지능은 단순히 인지능력에 벗어나, 인지한 환경 속에서 최적의 답을 찾아내고, 여기에 스스로 수행한 학습을 더 해 추론 및 예측하며, 향후에는 문제를 스스로 발견하고 해결하는 행동 단계인 복합지능까지 이를 것으로 예상됨
- 인공지능 기술은 기계학습(머신러닝), 딥러닝, 자연어 처리, 음성인식, 시각인식 등 첨단기술을 개발하는 방향으로 발전되고 있음
- 최근 몇 년간 AI 기술은 딥러닝이라는 기계학습을 통한 학습 능력 향상, 정보통신 기술의 발전, 컴퓨팅 능력의 향상과 비용 하락 등으로 빠르게 성장함
- 최근 주목받는 생성형 AI 경우, 대규모 언어모델(Large Language Model, 이하 'LLM')을 적용해 인간처럼 질문에 대답하고 있으며 데이터 학습을

바탕으로 이야기, 노래, 시, 컴퓨터 코드까지 작성할 수 있음

[그림 2-3]인공지능 진화 방향



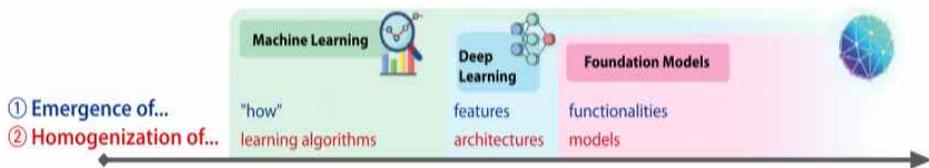
자료: 이승훈 (2017.2.), “최근 인공지능 개발 트렌드와 미래의 진화방향”, LG경제연구원

- 인간의 지능을 완벽하게 구현하는 인공지능 개발을 위해 기존 인공 신경망 구현 방식을 하드웨어, 소프트웨어적으로 고도화시키려는 노력과 기존 과학, 공학 분야의 연구가 아닌 신경과학, 뇌과학과 같은 분야의 연구를 기반으로 인공지능을 개발하려는 시도가 활발히 진행 중임
- o (생성형 AI의 영향) 챗GPT의 등장의 강한 임팩트로 인해 AI기술 개발의 방향이 단일 모듈에서 파라미터를 조정하여 정확도를 높이는 방향에서 선회하여, LLM과 같은 파운데이션 모델 개발에 집중되고 있음
- 그동안의 AI기술 개발은 파라미터를 조정하여 정확도를 높이는 방향으로 발전해 왔으며 그동안 개발된 대부분의 AI 시스템은 이미지를 구별하는 데 사용되었던 반면에, 새로운 이미지를 생성하고 인간과 유사하게 대화하는 ‘생성형 AI’ 등장으로 인공지능에 관한 관심이 높아짐
- 최근 주목받는 생성형 AI인 GPT(Generative Pre-trained Transformer)는 LLM기반 트랜스포머 모델로 딥러닝을 통해 축적된 훈련을 바탕으로 인간과 유사한 수준의 문장생성 능력을 보여주고 있음

\* 다만, 인공지능 파운데이션 모델은 기술적인 측면에서 새로운 기술은 아니며, 그동안 연구되어온 CNN, RNN 등의 심층신경망(Deep Neural Network)과 강화학습, 자기주도학습(Self-Supervised learning) 기법을 사용하고 있음

- LLM은 다른 데이터의 학습에서도 효과적인 결과를 보여주는 범용성을 가지고 있어, 오픈AI의 챗GPT는 파운데이션 모델로 AI 확산의 주역이 되고 있음
- LLM기반 생성형 AI는 기존의 AI모델에 비해 범용성이 넓어 다양한 산업의 기술과 접목하면서 산업의 지형을 바꿀 것으로 전망됨
- [그림2-4]에서 보듯이 인공지능의 발견과 단일성 측면의 진화를 보여주고 있으며 발견 초기에는 어떤 특성(Feature)을 찾도록 지정했다면, 딥러닝에서는 데이터를 주면 잠재 특성(Latent feature)을 추출하게 되었으며, 초거대 모델에서는 사전에 훈련된 작은 샘플 데이터나 예제만 가지고도 판단하거나 학습 없이 바로 특정 업무에 사용할 수 있도록 진화하고 있음

[그림 2-4]AI파운데이션 모델의 진화



자료: 최형광, TTA저널 AI 파운데이션 모델 구축 이슈와 국내 업계 동향, 2023.05.

- 모델과 데이터의 크기가 증가하면서 생성형 AI는 더 좋은 성능과 결과를 보이고 있으며 이를 근간으로 더 큰 모델과 GPU를 투입하여 초거대 규모로 발전함
- \* 초거대 언어모델인 GPT-3는 1,750억 개의 매개변수를 운영하고, 입력되는 자연어 프롬프트를 통해 다양한 작업을 수행함. 매개변수가 많다는 것은 그에 따른 적합한 답을 찾을 능력과 연결되어 있어 정보처리에 있어

효과성이 점차 높아지고 있음을 의미함

- 다만, 인공지능 파운데이션 개발에는 인프라의 구현, 실질적인 운용과 분석 모델 설계 등에 대한 천문학적인 투자가 따르기 때문에 단기적인 개발성과가 상업적으로 활용되기까지는 시간이 걸릴 것으로 전망됨

### 3. 인공지능 학습용 데이터의 필요성

- o (데이터 확보 필요성) 인공지능의 급격한 발전은 21세기의 기술적 진화를 주도하고 있으며, 이러한 인공지능 기술의 핵심에는 데이터에 기반한 알고리즘과 컴퓨팅 파워가 존재함
  - 데이터는 AI의 발전을 위한 원동력으로, 그 질과 양은 모델의 성능과 정확성에 결정적인 역할을 수행함
  - 과거에는 간단한 통계 기반의 알고리즘에 적은 양의 데이터로도 충분했지만, 현재의 딥러닝 모델은 복잡한 패턴을 인식하고, 더욱 정교하게 일반화하는 능력을 향상시키기 위해 수십만, 수백만, 때로는 수십억 개의 데이터 포인트를 필요로하며, 이처럼 방대한 양의 데이터가 필요함
  - AI 기술 도입으로 빠르게 변화하는 미디어 환경에서 경쟁 우위를 지니기 위해서는 국내 기업들의 빠른 대응이 필요하나, 국내에는 각종 규제 및 양질의 AI 학습용 데이터의 부족으로 생성형 AI 기술 부문에서의 AI 모델의 성능 제고에 어려움이 있음
- o (학습용 데이터의 정제) 일반적인 대규모의 빅데이터와 달리 인공지능 학습용 데이터는 학습목적에 부합하는 데이터의 수집 및 전처리 등의 정제가 필요함
  - 인공지능 학습용 데이터는 단순히 데이터의 양만이 중요한 것은 아니라, 데이터의 질, 다양성, 그리고 표현력도 모델의 성능에 큰 영향을 미치는데, 잘못된 데이터나 편향된 데이터는 모델의 성능을 저하시키며 때로는 심각한 결함을 야기할 수 있기 때문임

- 인공지능 학습용 데이터는 특정한 AI 모델의 학습을 목적으로 생산되는 데이터로서 임무 정의에 따라 구축되기 때문에, 지도학습이라는 특징에 따라 이미지, 비디오, 오디오, 텍스트 등 비정형 데이터를 대상으로 라벨링 작업을 통해 참값(GT, Ground Truth)을 부여하며, 참값의 품질 여부에 따라 쓰임이 정해지는 등 품질이 엄격히 통제되어야 하는 특성을 가짐

<표 2-2> 인공지능 학습용 데이터와 빅데이터의 차이점

구분	인공지능 학습용 데이터 <sup>2)</sup>	빅데이터
주요 차이점	주어진 임무로 정한 AI 모델의 학습과 고도화에 초점	대규모의 데이터의 정형화를 통해 숨겨진 인사이트를 찾아내는 것에 초점
구축목적	인공지능 모델의 학습목적	인사이트 도출 및 비즈니스 개선
사용기술	예측, 분류, 군집화, 차원 축소	통계적 분석, 텍스트마이닝
데이터유형	이미지, 비디오, 오디오, 텍스트, 3D 등 비정형 데이터 중심	텍스트데이터, 정형데이터
데이터구조	원천데이터와 라벨링 데이터가 쌍으로 구성	키와 값으로 구성
데이터구축 절차	구축계획수립→획득/수집→정제→가공→검사→학습→저장	획득→정제→변환→저장
데이터저장 방식	파일형태로 저장 (원천데이터+JSON 파일)	범용적인 NoSQL DBMS에 저장 (Key+Value 구조)

- 따라서 AI 학습용 데이터는 기존의 국내 각 분야에서 구축해 왔던 빅데이터와는 그 성격이 다를 수밖에 없는데<sup>3)</sup>, 주요 차이점은 AI 학습용 데이터가 이미지, 비디오, 오디오, 텍스트, 3D 등 비정형 데이터를 중심으로 지도학습이 이뤄질 수 있도록 원천데이터와 라벨링 데이터를 쌍으로 구성하는 구조를 가짐
- 결과적으로, 효과적인 데이터 수집 및 전처리기법은 AI 연구 및 개발의 핵심 요소라 할 수 있음
- 정제 후 바로 변환하여 저장하는 빅데이터와 달리 일반적으로 AI 학습

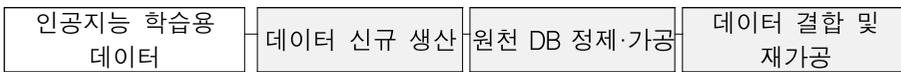
2) 지도학습용(Supervised Learning) 한정

3) 빅데이터는 정형의 키와 값으로 구성된 라벨링의 가공단계를 거치지 않음

용 데이터는 정제→가공→검사→학습의 과정을 거쳐 AI가 학습할 수 있는 라벨링이 필수적임

- 현재 국내 기업이나 공공기관이 보유한 대부분의 데이터는 라벨링 데이터가 없는 원천데이터 중심인 경우가 많아 AI 학습용으로서의 활용도가 떨어짐
- 인공지능학습용 데이터는 단순히 데이터를 생성하는 영역으로 국한되는 것이 아니라 생성된 데이터를 저장, 관리하고 사업에 적용하기 위해 전처리 등 가공하는 영역까지 포함하여 데이터의 전주기에 따른 생산이 필요함

<표 2-3> 인공지능 학습용 데이터 수집 및 가공 절차



<표 2-4> 인공지능 학습용 데이터의 구축 절차

구축단계	학습용 데이터 구축 절차
① 구축계획 수립단계효과	인공지능이 기계학습을 통해 해결하고자 하는 문제를 명확하게 정의하고, 문제 해결에 필요한 인공지능 학습용 데이터를 구체적으로 정의하고 설계하는 활동
② 데이터 획득·수집단계	인공지능의 기계학습에 필요한 데이터를 현실 세계에서 직접 생성하거나 이미 보유하고 있는 조직이나 시스템 등으로부터 인공지능 학습에 필요한 데이터를 확보하는 등 '원시데이터'를 확보하는 활동 수행 * 이때 원시데이터는 기계학습을 목적으로 획득, 수집한 '이미지', '비디오', '오디오', '텍스트', '3D' 등의 데이터를 의미함
③ 데이터 정제단계	획득한 원시데이터를 기계학습에 필요한 형식이나 크기로 맞추고 데이터 중복을 제거하는 단계로, 원시데이터 획득시 포함된 개인정보를 비식별화하여 사용가능한 '원천데이터'로 만드는 과정 * 이때 원천데이터는 원시데이터를 라벨링 작업에 투입하기 위해 필요한 정제작업을 수행한 데이터로, 라벨링이 부여하지 않은 상태의 데이터를 의미함
④데이터 가공단계	인공지능이 기계학습에 활용할 수 있도록 기능이나 목적에 부합하는 라벨을 원천데이터에 부착하는 '라벨링' 작업을 수행하는 단계 * 라벨링 데이터는 원천데이터에 부여한 '참값', 파일형식, 해상도 등의 데이터 속성과 설명, 주석등이 포함된 어노테이션의 집합을

	의미함
⑤ 데이터 학습단계	학습 데이터세트를 이용하여 사전에 정의된 인공지능 알고리즘을 학습시키고, 학습된 인공지능 모델의 성능을 향상시키거나 보정하는 활동을 수행하는 단계

자료: 한국지능화정보원 「인공지능 학습용 데이터 품질관리 가이드라인 및 구축 안내서」, 재구성

- (영상학습 데이터의 라벨링 방식) 영상 학습데이터 제작에 있어 라벨링 방식은 라벨링 대상에 따라 ① 이미지 단위 분류 레이블링, ② 이미지 내 객체 레이블링, ③ 이미지에 대한 설명 레이블링 방식으로 구분 가능함
  - 현재 제공되는 학습용 데이터세트에서 이미지 단위 분류 레이블링을 적용하는 사례는 구글의 Open Image Dataset로, 이미지 내 객체 레이블링은 이미지 내의 다양한 객체를 인식하기 위해 객체 영역과 해당 객체의 클래스를 선택하여 라벨링을 하는 방식임
  - 영역 지정은 일반적으로 바운딩 박스(사각 형태) 또는 바운딩 폴리곤 형태(다각형 형태)로 지정하게 하며, 작업환경 소프트웨어가 이러한 작업을 GUI를 통해 지원함
  - 바운딩 박스는 개체 출현 영역을 사각형의 영역으로 비교적 단순히 지정할 수 있는 반면, 폴리곤 영역은 다수의 점으로 구성된 폐쇄 영역을 지정하는 방식으로 전자는 효율성이 높고 후자는 정확성이 높음
  - 이미지에 대한 설명 레이블링은 사람 수준의 이해를 텍스트로 설명하는 방식으로 이미지에 대한 라벨링을 실시해야 하는데, 구글의 Conceptual Caption 데이터세트가 대표적인 사례로 이미지와 이미지 내의 객체와 상황을 묘사하는 설명문을 제공하고 있음
- (영상 학습데이터 제작 도구 현황) LabelMe, MS VoTT, Make-Sense, Scalabel 등 다양한 영상·이미지 라벨링 제작 도구들이 상용화되어 있음
  - 처리 대상별로 이미지만 지원하는 경우와 비디오와 이미지를 모두 지원하는 경우, 3D 포인트까지 지원하는 경우도 있음

- 3D 포인트까지 지원하는 경우는 Point Cloud Annotation Tool과 Scalabel, Supervise.ly가 대표적이며 자율주행차량용 데이터세트 구축 등과 같이 3차원의 객체 데이터세트 구성에 사용됨
- 저작도구 사용환경 측면에서는 오프라인방식과 온라인 방식으로 구분되는데, 온라인 방식은 사용과 접근이 편리하지만, 오프라인 방식은 대용량의 비디오나 이미지에서 라벨링을 할 때 용량과 속도 면에서 보다 유리함
- 출력형식은 주로 VOC, COCO와 같은 형식을 지원하여 최종 데이터세트를 바로 모델 훈련에 사용할 수 있도록 하고 있음
- 라벨링 방식에서 자동 라벨링 기능을 지원하는 경우에는 사전 학습된 인식 모델로 객체를 자동 인식하고 이후 인적 검증을 사후적으로 수행할 수 있음

<표 2-5> 영상 학습데이터세트 제작 도구 비교

제작 도구 명	처리 대상	자동 레이블링	작업환경	지원 출력형식
Point Cloud Annotation Tool	Point Cloud (3D Points)	없음	오프라인	Aplo 3D
MuViLab	비디오	없음	오프라인	JSON
EVAI	비디오	지원: 연속 프레임에서 객체 추적 기능	오프라인	VOC, YOLO
LabelIMG	이미지	없음	오프라인	VOC, YOLO
Ybat	이미지	없음	온라인	VOC, YOLO, COCO
Turkey	이미지	없음	온라인	CSV
ImageTagger	이미지	없음	온라인/다중 유저	사용자설정
DeepLabel	이미지	지원	오프라인	KITTI, VOC, COCO, YOLO
Make-Sense	이미지	지원	온라인	YOLO, VOC, COCO
SuperAnnotate	이미지	지원	온라인	JSON
Labelbox	이미지	지원	온라인	JSON, CSV
Hasty.ai	이미지	지원	온라인	VOC, COCO, Json
Segments.ai	이미지	지원	온라인	COCO
CVAT	이미지	지원	온라인/다중 유저	VOC, COCO 등
LOST	이미지	지원	온라인/다중 유저	CSV
LabelMe	이미지, 비디오	미지원	오프라인	VOC, COCO

VIA(VGG)	이미지, 비디오	미지원	온라인	JSON,CSV
MS VoTT	이미지, 비디오	지원	온/오프라인	Azure, CNTK VOC, COCO, CSV, 자체
DarkLabel	이미지, 비디오	지원: 연속 프레임에서 객체 추적 기능	오프라인	VOC, YOLO
Scalabel	이미지, 비디오, Point Cloud	지원	온라인/다중 유저	자체 형식
Heartex	이미지, 비디오, 텍스트, 오디오	지원	온라인/다중 유저	VOC, COCO, JSON, CSV
Supervise.ly	이미지, 비디오, Point Cloud	지원	온라인	JSON
Darwin	이미지,비디오	지원	온라인	VOC, COCO, JSON

출처: 이용 외(2021), 재구성.

## 제2절 국내외 인공지능 기반기술 개발 현황

### 1. 인공지능 기반기술 유형

#### 가. 컨볼루션 신경망(CNN) 모델

- (개념) CNN은 이미지 분류 및 인식에 주로 사용되는 신경망의 한 종류로 이미지의 다양한 특징을 감지하고 중요도를 분류하는 데에 활용되는 인공지능 모델임
  - CNN은 이미지나 동영상을 입력받아 필터나 커널을 통해 이미지나 동영상의 다양한 측면에 중요도를 부여하는 딥러닝 알고리즘
  - CNN은 시각적 데이터에서 객체, 클래스, 범주 인식을 위한 패턴을 찾고 오디오, 시계열 및 신호 데이터를 분류하는 등 영상의 주요 패턴을 인식할 수 있으므로 이미지 및 비디오 처리 작업에 특히 유용함
- (사용분야) CNN은 이미지와 동영상 분류, 객체 감지, 얼굴 인식 등의 작업을 수행하기 위해 비디오 및 영화 제작에 사용할 수 있음
  - CNN은 지도교육을 통해 이미지와 동영상 내의 다양한 객체나 사람, 얼굴을 인식하도록 훈련될 수 있으며, 대규모 이미지 및 비디오 데이터셋을 학습하여 사실적인 이미지와 비디오를 생성하는 데에도 사용이 가능함

- CNN은 타 신경망 기술보다 영상 및 이미지 분석에서 활용성이 넓은 딥러닝 모델로 CNN을 사용한 대표적인 사례들은 자율주행차, 글자인식, 얼굴인식 등이 있음
- CNN은 이미지의 위치나 크기변화에 대응하기 위해서 복수의 컨볼루션 레이어(convolutional layer)를 사용하고, 풀링 레이어(pooling layer)를 사용하여 이미지의 크기를 줄이는 방식으로 인식률을 높임

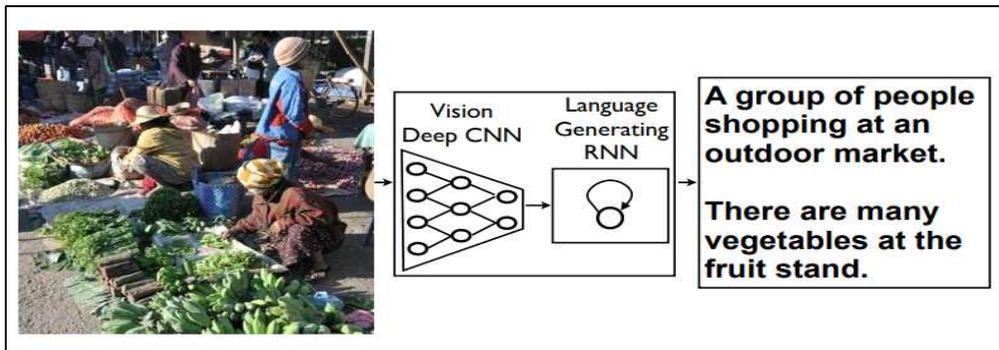
#### 나. 순환 신경망(RNN) 모델

- o (개념) RNN은 시퀀스 데이터를 처리하는 순환 신경망 클래스로, 일련의 연속적인 데이터를 분석하여 시계열데이터, 문장 등의 맥락을 이해하는 데에 적용가능한 인공지능 모델임
  - 기온, 일일 주가, 문장 등의 정보 시퀀스를 처리하는 인공 신경망의 한 클래스임
  - RNN은 이전 값이나 상태에 대한 메모리를 유지하여 시퀀스를 처리한다는 점에서 RNN에서는 현재 단계의 출력이 다음 단계의 입력되는 순환적 정보처리가 이뤄짐
- o (사용분야) RNN은 자연어 처리 분야에서 문장 생성, 번역, 감정 분석 등에 사용이 가능함
  - RNN은 이전 단어들의 정보를 기억하고, 다음 단어를 예측하는 방식으로 문장을 생성하거나 번역할 수 있으며, 감정 분석 분야에서도 사용가능함
  - RNN은 자연어 처리(NLP) 분야에서는 문장 생성, 번역, 감정 분석 등에 사용되며, 음성 인식 분야에서는 음성 인식, 음성 합성 등에 활용됨
  - 또한, 센서 데이터 처리 분야에서는 이상 감지, 예측 등에 사용이 가능함
  - 예를 들어, RNN은 특정 문장이 긍정적인지 부정적인지 판단할 수 있어, 영상미디어분야에서도 챗봇이나 음성비서 등 개인화된 맞춤형 고객서비

스 제공에 적용이 가능함

- 이미지 분석에 있어서도, RNN은 시간적 또는 순차적 데이터를 처리하는 데 적합하지만, CNN은 공간 데이터(이미지)를 처리하는 데 적합하다는 차이가 있음
- 영상 및 이미지의 객체인식에 유리한 CNN과 시간적이고 순차적인 데이터 처리에 유리한 RNN을 동시에 사용하면, 아래의 그림과 같이 영상이나 이미지 속의 상황을 문장으로 설명하는 자막이나 해설, 대본 구성 등이 가능해짐

[그림 2-5] CNN과 RNN을 사용한 이미지내 객체인식 및 상황설명 도출 개념



출처: Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). p. 3156).

#### 다. 파운데이션 모델과 대규모 언어모델(LLM)

- o (개념) 파운데이션 모델과 대규모 언어모델(Large Language Model, 이하 'LLM')은 다양한 자연어 처리(NLP) 작업을 수행할 수 있는 딥러닝 알고리즘 모델로, 전자는 범용 언어 모델인 반면, 후자는 특정 대화형 응용 프로그램에 맞게 미세 조정된 언어 모델이라는 차이가 있음
- 파운데이션 모델은 레이블이 지정되지 않은 대규모 데이터 컬렉션에 대해 학습되므로 대규모 데이터세트의 각 항목을 수동으로 분류하는 것에 비해 시간과 비용이 절약된다는 장점이 있음
- 이에 최근에는 LLM(대형 언어 모델)과 같이 처음부터 인공지능을 개발하지

않고 파운데이션 모델을 출발점으로 삼아 머신러닝 모델을 개발하는 방향으로 발전하고 있음

- o (사용분야) ChatGPT와 같은 LLM은 챗봇, 가상 비서 또는 텍스트 기반 대화 시스템용으로 설계되어 고객서비스나 타 인공지능 모델에 대한 텍스트 기반 대화 입력에 특화됨
  - RNN에 비해 확장성이 큰 파운데이션 모델은 이미 많은 분야에서 활용되고 있으며, 앞으로 더 많은 분야에서 사용될 것으로 예상됨
  - 파운데이션 모델은 텍스트 데이터를 이해하고 생성하는 데 사용될 수 있어 챗봇, 번역, 요약, 검색, 감성 분석, 질의 응답 등에 적용이 가능함
  - 이외에도 파운데이션 모델과 LLM은 이미지나 비디오 데이터를 인식하고 생성하는 데 사용될 수 있는데, 얼굴 인식, 객체 검출, 스타일 전송, 이미지 캡션, 비디오 생성 등의 작업을 수행할 수 있음
  - 예를 들어, GPT-4, BERT, DALL-E 2와 같은 모델은 텍스트나 이미지와 같은 다양한 도메인의 데이터에서 학습하고, 텍스트 생성, 이미지 생성, 자연어 이해, 코딩 등의 작업을 수행하는 데에 사용되고 있음

## 2. 인공지능 기술유형별 영상미디어 적용가능성 평가

### 가. CNN신경망 기술 적용 사례

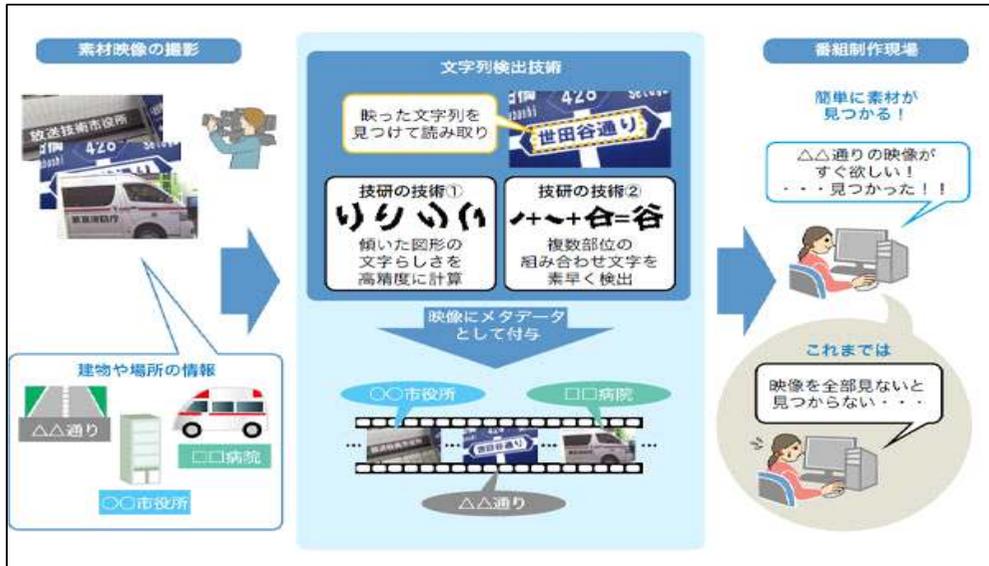
- o (영상분석 및 메타데이터 생성) CNN기술을 적용하면 영상내의 객체와 인물 등의 시각정보가 자동으로 분석되어 콘텐츠 관리에 필요한 메타데이터를 자동으로 생성할 수 있음
  - 메타데이터 자동생성을 위해서는 콘텐츠의 속성과 맥락을 지능적으로 해석하기 위한 AI 학습데이터 구축, 콘텐츠를 해석하여 콘텐츠의 핵심 구성 요소 및 맥락의 요약 추출, 미디어콘텐츠의 시계열 속성의 메타데이터 표현, 메타데이터를 용이하게 공유·활용할 수 있는 데이터 프라이버시형

메타데이터 인코딩 및 디코딩, 신뢰할 수 있는 미디어·콘텐츠 생산/유통/소비 환경 제공을 위한 미디어 트러스트 메타데이터 확장 기술이 필요함

- AI기반 콘텐츠 메타데이터 생성의 범위는 AI task 유형에 따라, 감정인식, 개체인식, 성별/나이 인식, 관계분석, 멀티모달 영상 질의응답, 단일 발화 의도 분석, 복수 발화 의도 분석, 음성인식 등으로 구분이 가능함
- AI작업 유형을 고려한 감정, 성별, 연령대, 발화 스크립트, 개체 및 관계 정보, 상황 설명 정보, 발화별 대화 의도 및 대화 전략 정보 의미정보 등이 포함되는 메타데이터가 사용되는 것이 일반적이며,
- 감정인식, 개체인식, 성별/나이 인식, 관계분석, 멀티모달 영상 질의응답, 단일 발화 의도 분석, 복수 발화 의도 분석, 음성인식 등 AI 서비스 개발이 가능하도록 데이터 구축이 필요함
- 기존의 제작자가 직접 입력하는 메타데이터와 달리, AI기반 메타데이터를 자동생성할 경우 각 영상클립별로 인물의 감정(기쁨, 슬픔, 분노, 놀람, 공포, 경멸, 혐오, 중립 등), 감정의 강도, 인물의 성별, 인물연령대, 발화 스크립트, 개체정보의 위치와 분류, 위치/행동 관계에 기반한 관계정보, 대화 전체의 주제에 대한 상황설명정보 등 감성과 대화 내용 등에 대한 맥락정보까지 추출이 가능함
- 미디어 서비스 사업자는 CNN을 활용한 영상콘텐츠의 컨텍스트 분석을 통해 클립기반 서비스, 하이라이트 서비스 등 다양한 형태의 서비스 제공을 통해 콘텐츠의 서비스 형태 다양화를 이룰 수 있음
- 대표적인 사례는 NHK의 CNN기반 메타데이터 생성시스템으로, NHK는 영상내에 간판이나 삽입된 문자를 인식해 영상메타데이터를 자동 부여함으로써 효율적인 영상 검색을 목표로 하고 있음
- 아래의 그림은 미디어 콘텐츠의 고유 속성과 시계열적 특성을 해석하여 유연한 구조의 메타데이터를 자동 생성하는 미디어 콘텐츠 자원화 핵심 기술인 AI기반 콘텐츠 메타데이터 자동 생성 플랫폼의 사례를 보여줌
- 영상에 반영된 명찰이나 간판으로부터 지명이나 인명의 정보를 꺼내 메

타데이터로 하기 위한 문자열 검출 기술이 개발되고 있는데, 아래의 그림에서와같이 간판, 자동차 등에서 노출되는 문자를 인식하여 자동으로 영상메타데이터를 구축하고 있음

[그림 2-6] 영상에서 문자열 감지 기술 개요



출처: NHK

- 다만 현재는 문자인식만 가능할 뿐, 인식된 문자가 인명인지, 지명인지 등을 파악하는 단계에 이르지 못하는 것으로 보이며, 향후 흐릿함이나 왜곡이 있는 문자열의 검출 정밀도의 향상이나, 검출한 문자열이 인명인지 지명인지 등의 속성을 추정하는 기술의 개발이 필요함
- o (영상 재편집 및 요약 자동화) 콘텐츠의 컨텍스트 분석을 통해 클립기반 서비스, 하이라이트 서비스 등 다양한 형태의 서비스 제공을 통해 콘텐츠 서비스의 형태를 다양화 할 수 있음
- CNN을 사용하여 콘텐츠 검색, 미디어 커머스 및 개인화 서비스 제공을 위한 미디어 부가서비스를 제공할 수 있음
- 해외에서는 AI 기술을 활용하여 동영상의 메타데이터를 자동 추출하는 서비스를 제공하는 서비스(IBM, Google, Leankr 등) 등이 개발 중

- 국내에서는 AI 비전 및 음성인식 기술을 통해 영상 내 일부 객체 및 음악 등을 자동 검출시스템(보이저X, 제머나이 등) 개발 추진
- 보이저X는 영상 속 음성을 자동으로 인식하여 영상 속 자막을 자동생성하는 기술과 대본을 다국어로 AI음성으로 구현하여 후시녹음을 지원하는 기술 등을 개발함
- 제머나이<sup>4)</sup>는 케이블 셋톱박스 및 공중파로 수신되는 자막이 포함된 영상을 AI 엔진으로 분석하고 결과(메타데이터)를 방송용 클린(Clean) 영상과 매치함으로써 편집단계에서 사용할 수 있는 콘텐츠 자산관리 시스템 개발함

#### 나. RNN 신경망 기술 적용 사례

- o (스크립트 작성) 시계열데이터, 문장 등 연속적인 데이터의 선후관계속에서 맥락을 이해하는 데에 적용되는 RNN은 시나리오 개발 등에 적용되고 있음
  - 2016년도에 발표된 9분 가량의 단편영화 ‘Sunspring’이 최초의 RNN을 활용한 시나리오 개발의 사례임
  - 영화감독인 Oscar Sharp과 개발자인 Ross Goodwin은 Benjamin이라는 이름이 붙은 RNN기반 인공지능 모델을 이용하여 ‘Sunspring’의 시나리오를 작성했는데, 이야기 전체의 구성 및 제목까지 인공지능이 작성함
  - Benjamin이 학습한 데이터는 Space Odyssey, Brazil, Mad Max 등의 영화 대본과 팝송 가사들로, RNN기반 모델로 스토리 시퀀스의 일관성을 확보함
  - 비록 Benjamin이 작성한 ‘Sunspring’의 대본이 논리적인 연결성이나 작품성이 부족하다는 비판을 받기는 했으나, 대본 작성에 총 48시간이라는 짧은 시간이 소요되었다는 점에서 보조적인 작업도구로서의 가치는 있는 것으로 보임

---

4) <https://www.gemiso.co.kr/>

## 다. 대규모 언어모델(LLM) 기술 적용 사례

- (LLM의 가능성) 현재까지 CNN이나 RNN에 비해 활발하게 개발되어 오지는 않았으나, 챗GPT의 등장 이후 LLM은 미디어부문의 콘텐츠 생성과 관리, 배포 분야에서 광범위하게 적용가능성이 있음(Saftler & Sobel, 2023. 6. 6.)
- LLM은 고유한 콘텐츠를 생성하는 능력 뿐만 아니라 확장가능성이 높고 복잡한 정보에 대한 깊은 이해와 인간과 같은 상호 작용을 시뮬레이션하는 능력을 보유하고 있어, 향후에는 미디어의 각 분야에서 활발하게 활용될 것으로 전망됨

<표 2-6> LLM 기술의 영상·미디어 적용가능성

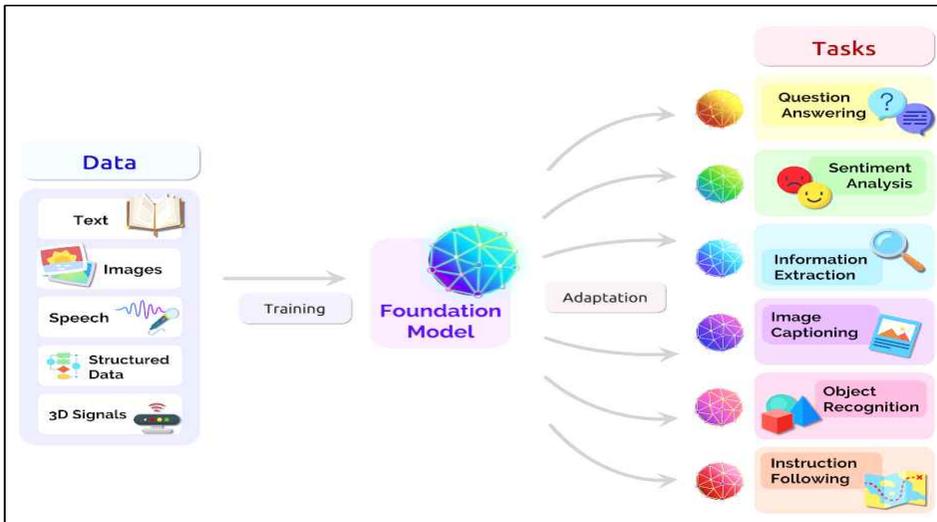
구분	기능
서비스 개인화	<ul style="list-style-type: none"> <li>● 스트리밍 미디어서비스 등에서 자연어 프롬프트를 통해 사용자와 대화를 수행할 수 있어 지능형 검색 기능 및 맞춤형 콘텐츠 검색, 대화형 챗봇 등 고객서비스 지원</li> </ul>
지능형 콘텐츠 생성	<ul style="list-style-type: none"> <li>● 대본이나 시나리오 작성, 자료조사, 사용자 선호도에 따른 콘텐츠 생성</li> <li>- 스크립트 작성: 영화, TV 프로그램 및 기타 비디오 콘텐츠용 스크립트를 생성하는 데 사용할 수 있습니다. 예) OpenAI의 GPT-3를 이용한 각본 생성</li> <li>- 비디오 편집: LLM을 사용하여 비디오 편집 프로세스를 자동화할 수 있음</li> <li>예) Lumen5: LLM을 사용하여 텍스트에서 비디오 콘텐츠를 자동으로 생성하는 비디오 편집 도구</li> <li>- 비주얼이펙트: 사실적인 시각 효과 생성</li> <li>예) NVIDIA의 GauGAN은 LLM을 사용하여 사실적인 풍경과 물체를 생성함</li> </ul>
큐레이션	사용자 선호도, 행동, 상황별 큐를 분석하여 콘텐츠를 추천

- LLM과 RNN은 모두 자연어 처리에 사용되는 모델 유형이나 아키텍처와 목적이 다소 다름
- RNN은 텍스트나 음성과 같은 순차적 데이터를 처리하는 데 사용되는 신경망 유형으로 입력 데이터의 순서를 고려하도록 설계되었으며 언어 모

텔링, 기계 번역, 음성 인식과 같은 작업에 사용됨

- 반면 LLM은 대량의 텍스트 데이터를 학습하여 일관되고 유창한 텍스트를 생성할 수 있는 일종의 언어 모델로, 챗봇, 질문 답변 또는 요약과 같은 특정 애플리케이션에 맞게 미세 조정되기 위한 자연어 처리 작업의 기초 모델로 사용됨
- 텍스트를 순차적으로 처리하기 때문에 확장하기 어려운 RNN과 달리 LLM 기반의 트랜스포머 모델은 텍스트의 일부를 동시에 처리할 수 있으므로 대규모 데이터셋에서 더 빠르고 효율적으로 처리할 수 있다는 평가도 있음(Mittal, 2023. 8. 8.)
- 이러한 잠재성 때문에 향후에는 영상미디어 분야에서도 LLM기반 딥러닝 모델 개발이 더 확대될 것으로 전망됨

[그림 2-7] 파운데이션모델 중심의 집중화 전망



출처: Bommasani, R. et. al.(2021). p.6.

## 라. 국내 기업의 대규모언어모델(LLM)중심의 개발 동향

- o (LLM중심 개발 동향) 국내 주요 기업들은 챗GPT등장에 영향받아 LLM에 중점을 둔 언어지능 중심의 개발에 박차를 가하고 있음

- 네이버 하이퍼클로바, LG 엑사원, SK텔레콤 에이닷, 카카오 koGPT 등을 시작으로 AI의 연구개발에 동참하고 있음
- o (네이버) 네이버는 초거대 인공지능서비스를 위한 LLM인 ‘하이퍼클로바X’와 이를 플러그인(Plug-in) 형태로 연동해 필요한 기능을 호출할 수 있는 ‘클로바X’, 검색에 특화해 개발된 대화형 AI 서비스 ‘큐:(Cue:)’를 공개하였음
  - ‘하이퍼클로바X’는 마이크로소프트(MS)의 초거대 AI인 ‘GPT-3.5’와 비교해 한국어를 6,500배 더 많이 학습함
  - 클로바X가 챗GPT에 대항하는 네이버의 단독형 챗봇 서비스였다면, 신규 AI 서비스 ‘클로바 포 라이팅’은 LLM의 ‘하이퍼클로바X’를 기반으로 기존 네이버 서비스와 연계하여 개발하였음
  - 클로바 포 라이팅은 네이버 블로그 작성자가 쓴 글을 요약하거나 글에 어울리는 태그를 자동으로 생성해 줌

<표 2-7> 네이버 생성형AI

구분	기능
하이퍼클로바X (HyperCLOVA X)	하이퍼클로바X는 네이버의 초대규모 AI로, 자체데이터를 합하면 사용자 니즈에 맞는 응답을 제공함
클로바X(CLOVA X)	네이버의 초대규모 언어모델인 하이퍼클로바X 기술을 바탕으로 만들어진 대화형
큐:(CUE:)	네이버가 가진 양질의 콘텐츠를 바탕으로 검색 경험 제공하며, 네이버의 예약·쇼핑 등 다양한 서비스와 연계될 예정임
클로바 포 라이팅 (CLOVA for Writing)	하이퍼클로바X를 기반으로 기존 네이버 서비스와 연계하여 블로그 작성자가 쓴 글을 요약하거나 글에 어울리는 태그를 자동으로 생성

자료: 네이버 홈페이지, <https://channeltech.naver.com/>

- o (카카오) 카카오는 60억 개의 매개변수와 2,000억 개 한국어 토큰 (컴퓨터가 이해하는 언어단위)를 가진 KoGPT를 개발하고 있음
  - 챗GPT를 만든 오픈AI의 언어모델 GPT-3를 기반으로 하는 KoGPT는 카카오가 개발 중인 챗봇 ‘조르디’가 개인화된 비서 역할을 하거나, 소

상공인들의 광고카피를 대신 써주는 방식으로 활용이 예상됨

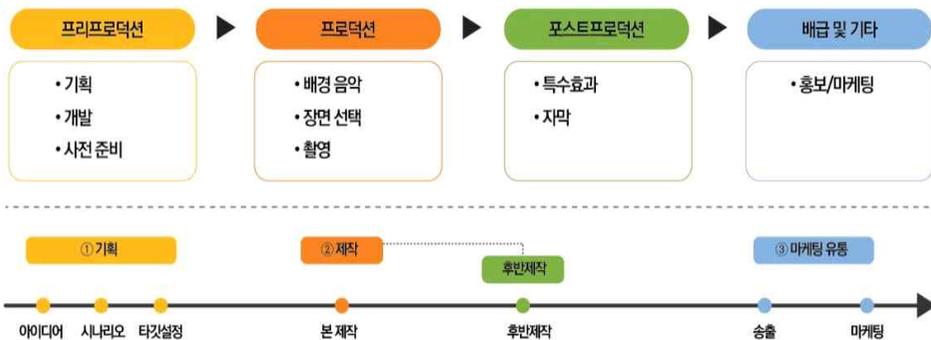
- 카카오는 KoGPT 이외에도 이미지 생성 모델인 ‘칼로’를 이용해 카톡 프로필 및 배경 사진을 만들 수 있는 기술도 선보일 예정임
- o (LG) LG는 전문가용 생성형AI 엑사원 2.0을 개발했으며 엑사원 2.0을 아틀리에, 유니버스, 디스커버리 등 3대 플랫폼에 접목하여 다양한 서비스를 선보일 계획임
  - 엑사원2.0을 접목한 ‘아틀리에’는 텍스트와 텍스트에 맞는 배경 이미지를 손쉽게 만들어주는 멀티모달 AI 플랫폼이며, 신소재·신물질·신약 관련 탐색에 적용할 수 있는 디스커버리 플랫폼, 대화형 AI인 유니버스 플랫폼으로 다양한 서비스를 제공할 예정임
- o (SK텔레콤) SK텔레콤은 자체 LLM ‘에이닷’을 개발하여 대화 및 고객센터 등 통신사 기반 서비스를 제공하는 챗봇을 운영 중임
  - 윤리적 답변 및 대용량 텍스트 입력에 강점이 있는 엔트로픽의 LLM, 한국어 데이터가 풍부한 코난 LLM 등을 조합해 고객 맞춤형으로 제공하는 기술 및 사업적 준비도 진행 중임
- o (KT) KT는 자체 LLM인 ‘믿음’을 개발하고 있으며 금융·의료·교육 등 특정 산업 분야의 B2B사업을 추진할 방침임
- o (엔씨소프트) 엔씨소프트는 자체 개발한 AI 언어모델 ‘바르코 초거대언어모델(LLM)’을 공개했으며 ‘바르코LLM’을 기반으로 생성 AI 플랫폼 바르코 아트, 바르코 텍스트, 바르코 휴먼을 공개할 계획임
  - 바르코LLM 기반 ‘바르코 아트’ 플랫폼은 이미지 생성툴이며, ‘바르코 텍스트’는 텍스트 생성 및 관리툴, ‘바르코 휴먼’은 디지털휴먼 생성·편집·운영툴임

### 제3절 영상·미디어 분야 AI 기술 적용현황

#### 1. 영상미디어분야 워크플로우와 AI활용성

- (영상미디어 분야 워크플로우) AI 기술이 발전함에 따라 영상·미디어콘텐츠 제작 단계인 ‘프리프로덕션-프로덕션-포스트프로덕션-배급 및 기타 등 영상 미디어 워크플로우 전반에 걸쳐 AI 활용으로 생산성과 효율성을 높이고 있음
- 영상미디어 분야 워크플로우는 시나리오를 작성하고 배우 캐스팅, 스케줄 확정 등 사전 준비를 하는 프리프로덕션과 실제 촬영을 하는 프로덕션, 영상편집과 색 보정 등 후반작업을 하는 포스트프로덕션, 제작된 콘텐츠를 유통하고 홍보하는 배급 및 기타로 이뤄져 있음
- 최근에는 촬영 들어가기 전 시나리오 작성을 하고 촬영 준비하는 과정인 ①기획, 촬영부터 편집 등 최종적으로 콘텐츠 만드는 ②제작, 제작된 콘텐츠를 송출하고 홍보하는 ③마케팅·유통 3단계로 워크플로우를 이야기하기도 함

[그림 2-8] 영상미디어 분야 워크플로우



자료: 과학기술정보통신부 (2023. 9. 12.), 재구성

- (국내 영상미디어 워크플로우 단계별 데이터 활용경험) 국내 미디어 기업의 데이터 활용경험 조사결과를 살펴보면 플랫폼에서 콘텐츠 추천에 필요한 단순 정형데이터 수집과 활용에 집중된 것으로 나타나, 영상미디어 분야 워크플로우 전반에서 데이터 활용이 미흡한 것으로 확인됨
- 2022년도의 방송통신분야의 데이터 수요 조사 결과<sup>5)</sup>를 보면 대부분의 응답자들이 데이터 활용 경험이 있는 분야로 개인추천서비스, 사업운영 및 마케팅 등 플랫폼이나 유통마케팅 단계에서만 사용한다고 대답함
- 결국 현재 국내 방송통신 및 미디어 분야의 데이터 수집과 활용범위가 단순한 머신러닝에 적합한 센서수집자료 및 정형데이터로 한정되어 있고, 대부분의 국내 기업들은 영상미디어 제작워크플로우에 데이터를 적용하지 않는 것으로 나타남

<표 2-8> 방송통신 미디어 분야 산업계 대상 데이터 활용경험에 대한 조사 결과

분류		서비스/연구 사례
개인 추천 서비스	방송/통신	방송 콘텐츠 메타데이터 활용한 개인화 맞춤형 VOD 서비스 개인별 방송시청 데이터를 활용한 이용률 낮은 구독서비스 해지권유 서비스 등 IPTV, OTT 구독현황 활용한 적합한 구독서비스 추천 개인별 무선통신서비스 이용 데이터를 활용한 적합한 통신요금제 추천 서비스
	융합	건강, 금융, 물류/유통 등의 고객 분석 및 개인화된 맞춤형 UI/UX 및 콘텐츠 추천 서비스
데이터 구축/ 중개/ 분석 서비스	구축/중개 (방송/통신)	뉴스 대본 및 앵커 음성 데이터 가공, 검수 가구 단위 IPTV, VOD 시청 데이터 사업화 추진 중
서비스 개선/혁신 (마케팅)		통신 이용행태를 활용한 소비자 행태 분석 선호 채널, 선호 방송, 주요 이용 시간, 이용 패턴 등 고객 특성을 반영한 고객 맞춤형 마케팅

- (CPND별 적용가능한 지능유형) 앞서 살펴본 해외의 사례에서 보듯이 인공지능은 미디어 콘텐츠 제작과 같은 창의적 분야까지 급속하게 확장

5) 4차 산업혁명 위원회(2022)의 정책연구과제의 일환으로, 방송통신분야 종사자 및 전문가 51명을 대상으로 조사함

하는 추세이며, 전문가가 아니라도 손쉽게 콘텐츠를 만들고 편집할 수 있도록 하는 인공지능 기술이 등장하여 콘텐츠 기획, 콘텐츠 창작, 콘텐츠 서비스 등에서 다방면으로 활용가능성이 있음

- 이수엽(2021)는 미디어엔터테인먼트 산업의 가치사슬에 따라 아래의 <표 2-9>와 같이 인공지능기술이 활용될 것으로 전망함

<표 2-9> 미디어 엔터테인먼트 산업의 가치사슬별 인공지능 활용사례

구분		미디어 엔터테인먼트 산업의 가치사슬				
		콘텐츠(C)	콘텐츠서비스(CS)		플랫폼(P)	디바이스(D)
인공지능 기술유형	시각지능	콘텐츠 창작, 콘텐츠 창작지원	디지털 아바타	디지털 휴먼	메타버스	홈트레이닝
	언어지능		챗봇			인공지능 스피커
	전문가 시스템	콘텐츠 기획(성공예측)	개인화 추천			

출처: 이수엽(2021), 인공지능과 미디어 엔터테인먼트 산업, 기획리포트

- 전 세계 산업을 흔들고 있는 AI는 100년 방송시장에도 변화를 불러오고 있으며, 유통과 플랫폼뿐만 아니라 콘텐츠 제작 시스템, 제작 등 방송·미디어 분야 전반에 있어 관련 기술 개발도 다양하게 이뤄지며 방송산업도 적극적으로 AI를 받아들이고 있음
- 방송 통신 시장의 혁신을 말하는 전미 방송통신전시회(NAB)는 2023년 방송시장을 관통하는 키워드로 AI, 버추얼 스튜디오, 콘텐츠를 뽑음
- SVOD, FAST 등의 서비스의 AI의 접목이 더 활발하게 이뤄질 것으로 전망되며 AI를 기반으로 취재 보도, 보도 콘텐츠 제작, 더빙 제작 등을 자동화하는 등 미디어 전 분야에서 AI 활용은 더욱 확대될 전망이다
- 2023년 NAB는 84개 세션이 AI 관련 내용을 다루었고 AI가 탑재된 가상 스튜디오(Virtual Studio), AI 기반 편집 솔루션, 라디오 GPT 등 방송 산업에서의 AI 기술 개발을 확인할 수 있었음

<표 2-10> NBA에서 선보인 기업별 방송 분야 AI 기술 사례

기술	기업	내용
비디오 퀄리티와 시청경험 최적화	아이맥스 (IMAX)	- 실제 사람들의 시청 데이터 기반으로 AI를 적용해 비디오 시청 경험 자동화
텍스트 기반 AI 편집 솔루션	어도비 (Adobe)	- AI를 이용해 텍스트를 자동 검색하고 이를 중심으로 영상을 편집
라디오GPT	퓨처리 (Futuri)	- 지역 라디오를 위한 생성형AI 콘텐츠 솔루션으로 GPT-3를 탑재해 AI가 스토리를 만들고 소셜 콘텐츠 시스템 Topic Pules도 운영함
AI보이스 제작	베리톤 (Veritone)	- AI를 활용해 음성 콘텐츠를 외국어로 자동 번역하고 녹음해 제공

자료: KCA 미디어 이슈 & 트렌드 AI와 미디어의 결합

- 스트리밍 서비스의 확대로 콘텐츠의 수요가 올라가면서 콘텐츠 중요성이 커졌으며 동시에 품질에 대한 기대도 올라가, 생성형 AI<sup>6)</sup> 등 AI를 접목시킨 기술 개발을 통해 콘텐츠 부분에서 경쟁력을 갖춰야 함
- 미디어 영역에서 활용 가능한 생성형 AI의 대표적인 사례는 자료조사나 대본, 고객 대응, 명령어 작성 등 데이터 입력단계에서 활용 가능한 텍스트 생성 AI인 GPT3-3, 캐릭터 디자인이나 이미지 생성에 활용 가능한 달E(Dall-E), 렌사(Lensa), VFX 효과에 사용되는 배니티 AI (Vanity AI), 영상후반 편집에 사용되는 어도비 파이어플라이(Firefly), BGM이나 효과음에 사용되는 사운드로우(Sounddraw) 등이 상용화되어 사용되고 있음
- 이들 생성형 AI가 영상·미디어 분야의 범용적 기술로 사용됨에 따라, 생산요소 부문에서 효율성과 생산성을 제고할 것으로 전망됨

6) 생성형 AI는 비정형 딥러닝 모델을 사용하여 사용자의 입력에 따라 콘텐츠를 생성하는 인공지능 유형으로 대규모 언어 모델(Large Language Model)을 적용해 인간처럼 질문자의 의도를 파악해 답변을 제시하는 인공지능

<표 2-11> 미디어 영역에서 활용 중인 생성형 AI 기술유형

구분	기능 및 적용
챗GPT	- 기능: OpenAI가 개발한 자연어 생성 모델로 대화와 관련된 많은 텍스트 데이터를 학습해 사람처럼 대화할 수 있도록 교육된 모델 - 적용: 자료조사
달E(Dall-E)	- 기능: 상상력을 담은 텍스트를 입력하면 원하는 이미지 제작 - 적용: 영화 속 괴물이나 애니메이션 캐릭터 제작
렌사(Lensa)	- 기능: 사용자의 사진을 활용하여 디지털 아바타로 전환 시키는 기능 - 적용 가능성: 애니메이션이나 SF영화 등장인물 제작
배니티 AI (Vanity AI)	- 기능: MAZ의 배니티 AI는 VFX 효과 제작 - 적용 가능성: 장면별 VFX 효과 적용
어도비 파이어플라이 (Firefly)	- 기능: 텍스트 기반 편집 프로그램으로 영상 편집과 색상 보정, 자막 작업 - 적용: 영상 후반 작업 활용
사운드로우 (Sounddraw)	- 기능: 사용자가 선택한 분위기, 장르 템포에 따라 저작권 이슈 없는 음악 제작 - 적용: 콘텐츠에 필요한 음악 제작 및 사용
슈퍼톤 AI 오디오 솔루션	- 기능: AI 오디오 기술을 통해 목소리 구현 - 적용: 영화나 드라마 속 어린시절이나 회상 장면 목소리 구현, 고인이 된 가수의 목소리 구현 등에 사용
페이스 에디팅 (Face Editing)	기능: AI 모델 학습에서 생성적 신경망(GAN:Generative Adversarial Network)을 사용하여 해당 영상의 각 프레임에 교체할 영상을 합성 적용: 대역 얼굴 교체 및 합성

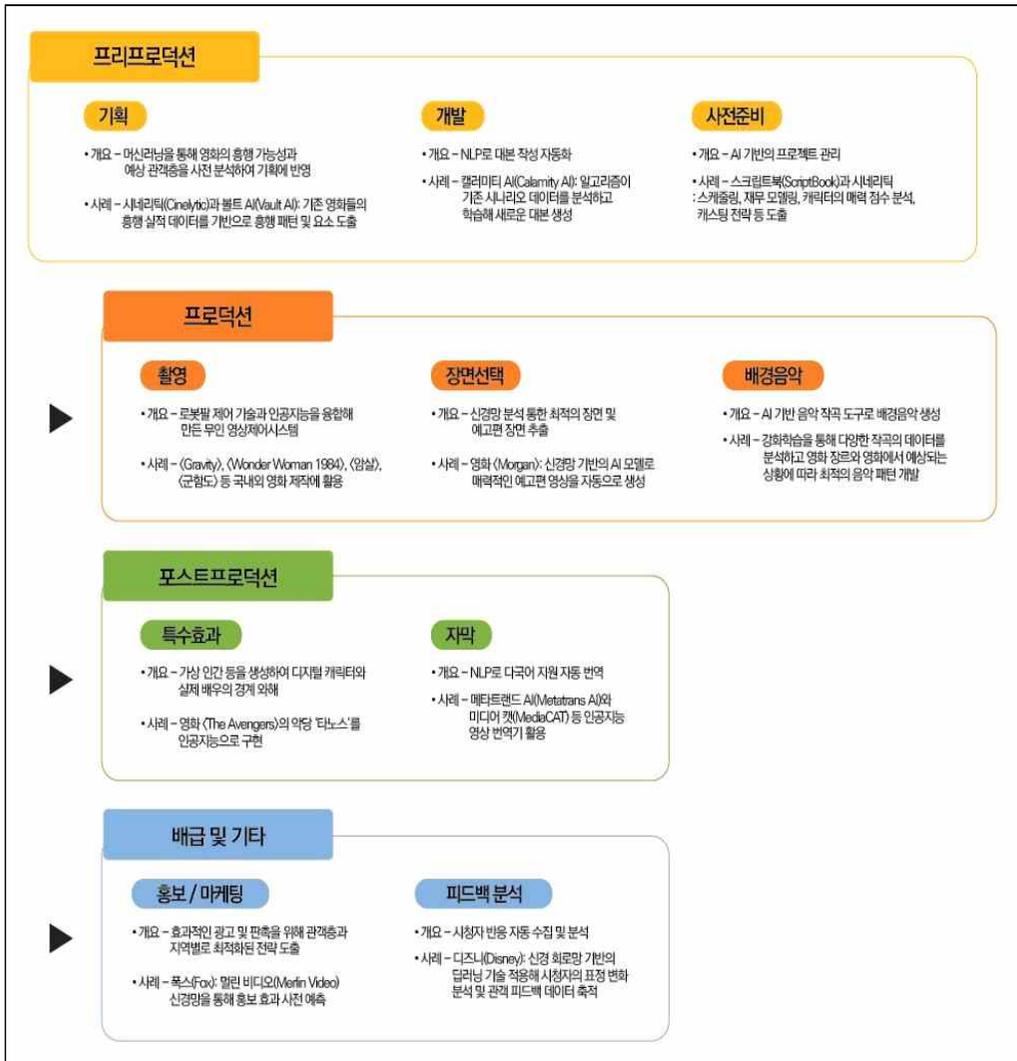
자료: 김나은(2022). 생성형 AI 활용사례와 활용방안.

## 2. 워크플로우 단계별 데이터 적용 사례 분석

- (워크플로우 단계별 AI활용성) 영상·미디어 분야에서의 AI는 새로운 캐릭터를 생성하거나 기존의 시간과 비용이 많이 소요됐던 후반 작업 자동화를 통해 효율성을 추구하는 등 제작 분야에서 점차 활용이 확대되고 있음
- 현재 미디어 시장은 스트리밍 서비스가 확대되면서 콘텐츠 제작에 대한 수요가 급격히 증가함과 동시에 품질의 눈높이도 올라감에 따라, 콘텐츠 경쟁력 확보를 위해 반복적이고 사람의 시간과 노력이 많이 들어

가는 분야를 중심으로 AI 활용이 확대되는 추세임

[그림 2-9] 영상미디어 워크플로우 단계별 AI활용 현황①



출처: 한국방송통신전파진흥원(2022), 재구성

- AI를 활용하여 시나리오를 작성하거나, 후반작업인 VFX 효과 적용에 있어 AI 활용을 하여 등장인물을 제작하는 등 영상·미디어 분야 워크플로우 전반에 있어 AI 활용이 이뤄지고 있으며, 이를 통해 시간 및 비용 절약하고 있음

- 기획에서 제작에 이르기까지 방송영상 분야에서 AI 활용은 확대되고 있으며 AI 활용을 통한 영상미디어 워크플로우 단순화 및 효율성 확대를 위해서는 현재 활용되고 있는 기술 발전을 통해 보편화할 수 있도록 방송에 특화된 AI 학습용 데이터 확보는 필수적임
- 또한, AI 기반 데이터 분석 결과를 마케팅 전략에 활용하는 등 제작된 콘텐츠의 마케팅과 유통 과정에도 AI 활용이 확대되고 있으며, 스트리밍 서비스 활성화에 따라 AI를 접목하여 개인 맞춤형 콘텐츠 추천 등 이용자 특화된 서비스를 제공하고 있음

[그림 2-10] 영상미디어 워크플로우 단계별 AI활용 현황②



출처: 과학기술정보통신부 (2023. 9. 12.).

### 가. 프리프로덕션 단계

- o (기획-개발-사전 준비 지원) 기존 시청자 데이터를 학습시켜 흥행가능성이 있는 기획안을 추출하거나 NLP기반 대본 작성 자동화 등 기획-개발-자료조사 및 사전준비에 AI가 활용될 수 있음
- 영상·미디어 분야의 프리프로덕션 단계에서 기존 시청자 데이터를 활용하여 AI 알고리즘 생성과 기계학습을 진행시켜 영상 기획, 대본·시나리오 개발, 스케줄링·캐스팅 전략과 같은 사전 준비 등을 진행할 수 있음
- 글로벌 OTT의 경우 이용자·콘텐츠 관련 전체 데이터를 분석하여 국가·장르별 콘텐츠 선호도를 도출하고, 이를 기반으로 투자전략을 수립하는 등 중요한 의사결정에 데이터를 적극적으로 활용하고 있음

- 대표적으로 넷플릭스 경우, 로그인 과정에서 개인정보와 콘텐츠 검색 및 시청 과정에서 콘텐츠 선호도 정보, 결제 과정에서 금융 정보를 연계하여 이용자의 데이터 분석을 통해 이용자의 세분화된 욕구를 파악·예측하고 콘텐츠 가치를 평가함
- 생성형 AI가 대본을 작성하고 연출하여 콘텐츠를 제작한 사례가 등장하는 등 AI 활용 분야가 창작의 영역까지 확대되고 있음

<표 2-12> 프리프로덕션 적용 사례

분야	데이터 활용 방법	사례
기획	과거 영화 흥행 데이터를 머신러닝 시켜 영화의 흥행 가능성과 예상되는 관객층을 분석하여 영상 기획에 반영	파일럿(Pilot): 과거 영화 흥행 정보 데이터를 AI 알고리즘을 활용하여 개봉 첫 주말 흥행 수입을 예측
개발	시나리오 데이터 학습시켜 생성형 AI를 활용하여 대본 작성 자동화	캘러미티 AI(Calamity AI): 알고리즘이 기존 시나리오 데이터를 분석하고 학습하여 새로운 대본 생성
사전 준비	스크립북, 재무모델링, 캐릭터의 매력 점수 분석, 캐스팅 전략 도출	시네리틱(Cinelytic): AI 알고리즘과 데이터로 캐스팅과 재무 관리 등 영화 전반의 종합 솔루션 제공

#### 나. 프로덕션 단계

- o (개요) 실제적으로 영상을 제작하는 프로덕션 과정에서도 비슷한 장르의 영화의 OST나 예고편 데이터 학습을 통해 새로 제작된 콘텐츠에 맞는 배경음악 생성, 장면 선택 및 추출, 촬영 등 영상 제작에 AI를 활용하기 위한 학습용으로 데이터가 활용되고 있음

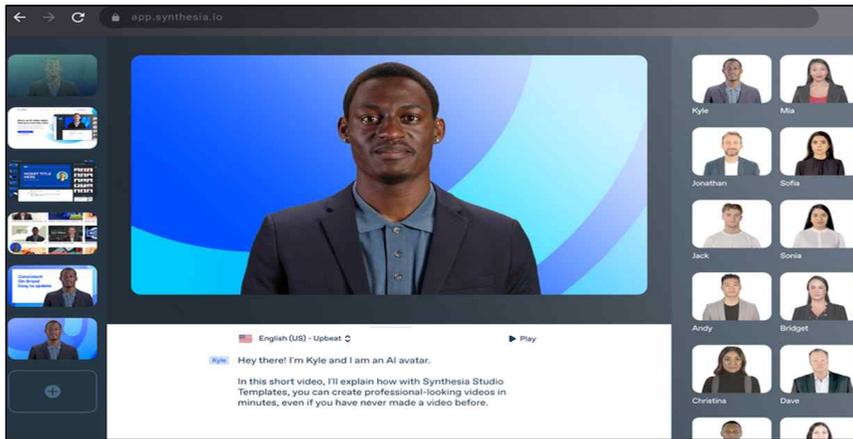
<표 2-13> 영상 프로덕션 적용 사례

분야	데이터 활용 방법	사례
배경음악	다양한 작곡 데이터를 학습하고 분석하여 AI기반 음악 작곡 도구를 활용하여 배경음악 생성	사운드드로우(Sounddraw): 사용자가 선택한 분위기, 장르 템포에 따라 저작권 이슈 없는 음악 제작

장면 선택	신경망 분석을 통한 최적의 장면 및 예고편 장면을 추출하여 영상 제작	영화 모건<Morgan> : IBM의 왓슨에게 머신러닝과 왓슨API를 활용해 100개의 호러무비의 예고편의 데이터를 분석시켜 모건 전체 필름에서 장면을 추출하여 1분 15초짜리 예고편 제작
촬영	기존 시각특수효과 데이터를 학습시키고 로봇팔 제어 기술과 인공지능을 융합하여 영상 촬영 가능한 무인 영상 시스템 구축	시네마 로보틱스(Cinema Robotics): 그 래비티(2013), 원더우먼(2017), 암살 (2015), 군함도(2017) 등 VFX 비중이 높은 영화 제작에 활용

- 프로덕션 단계에서 배우/캐릭터 등 시각 및 언어지능을 활용하는 사례가 증가하고 있음
- 상상력을 담은 텍스트를 입력하면 이미지로 만들어주는 달E(Dall-E), 미드저니 등 버추얼 이미지 AI로 제작한 등장인물을 영화 드라마에 활용하는 것에 나아가 배우를 복제하여 목소리를 등장시키는 등 폭넓게 생성형 AI를 활용하고 있음

[그림 2-11] 신세시아 디지털 휴먼 생성 사례



출처: 신세시아(<https://www.synthesia.io/>), 홍성일(2023. 6. 14.) 재인용

- 한 가지 예로 캐릭터 기반 AI 동영상 제작 플랫폼인 '신세시아'<sup>7)</sup>는 사

7) 2017년도에 설립된 영상 생성 시스타트업으로 일반 개인 이용자를 대상으로 월 \$22.5에 연간 120개의 비디오 크레딧과 90개의 AI 아바타, 비디오당 10개의 장면을

용자들이 120개의 텍스트 언어로부터 버추얼 아바타를 이용해서 전문 비디오를 만들 수 있는 도구를 제공하고 있음

- 디즈니+의 오리지널 시리즈 ‘오비완 케노비’에 스타워즈 다스 베이더의 목소리로 유명한 제임스 얼 존스의 목소리가 등장하는데 실제 존스 배우는 이 작품에 출연하지 않는데, 이는 우크라이나 보이스 클로닝 스타트업 리스피처(Respeecher)를 통해 제임스 얼 존스의 목소리를 생성하여 콘텐츠에 활용한 사례임
- 이처럼, 미디어 산업은 버추얼 이미지 생성형 AI를 활용하여 영화나 드라마 속 배우나 다른 세계의 캐릭터를 제작하고, 영상에 필요한 목소리를 AI를 통해 제작하는 등 기존 자료나 텍스트 기반으로 영상 속 배우, 캐릭터, 음악, 더빙 제작에 생성형 AI 활용이 확대될 전망이다
- o (촬영단계) 영화·드라마 제작단계뿐만 아니라 촬영단계에서 생성형 AI 활용하여 콘텐츠를 제작하는 사례도 증가하고 있음
  - 국내의 경우, 광고 및 예능에도 생성형 AI 활용을 통해 콘텐츠를 제작하고 있으며 대부분 AI 딥러닝과 디에이징을 활용하여 요즘 시대가 좋아하는 새로운 캐릭터를 생성하거나, 과거의 모습을 구현하는 등 기존에 있는 데이터 학습을 필수 요소로 하고 있음
  - 영상·미디어 제작 분야에서 AI 기술은 해외에서는 대형 블록버스터급 영화나 장기 시즌제 TV 시리즈 위주로 제한적으로 적용하며 투입 자본에 비례하는 작품성을 갖추려는 반면, 국내에서는 TV 드라마나 예능처럼 대중적으로 친숙한 콘텐츠에 활용하여 맥락의 연속성을 높이거나 휴머니즘을 강조하는 등의 실험적인 방식을 추구하며 차별화된 경쟁력 확보에 집중하는 추세임
- o (AI 기술을 활용하면 포스트프로덕션과 촬영단계인 프로덕션 단계 간의 경계가 흐려지고 배우의 얼굴이나 음성의 포스트프로덕션 없이 촬영 중에 실시간으로 페이스·음성 디에이징 AI 기술을 적용할 수 있음

---

담을 수 있는 동영상 제작 플랫폼 서비스를 제공하고 있음. 사용자는 버추얼 아바타의 영상에 자신의 음성녹음을 입력하여 영상을 합성할 수 있어 가상의 배우를 창조할 수 있음

- 기존 디에이징 기술은 배우 얼굴에 작은 마커를 붙여서 촬영한 후, 후반 제작에서 디지털 편집과 CGI(Computer Generated Imagery), 3D 모션 등을 조합하여 각 시퀀스에 합성된 영상을 입혀서 완성했음
- 최근 디에이징 기술은 AI 기반 신경망 모델링 접목을 통해 과거에 출연한 작품을 바탕으로 영상 정보를 추출하여 촬영본에 영상을 덧씌우는 방식으로 기존 영상 정보를 활용하여 시간과 비용을 단축하고 사전에 설정된 기준을 바탕으로 특정 작업 자동화로 효율성을 확보함
- 음성 디에이징의 예로는 스타트업 AI 오디오 기업 슈퍼톤은 AI에게 노래하는 법을 훈련 시킨 후 생전의 김광석 목소리 데이터 학습을 통해 故 김광석의 목소리를 구현하여 김광석이 부르는 '보고싶다'를 선보임

<표 2-14> 페이스·음성 디에이징 AI기술 적용 사례

사례명	데이터 활용 방법
<p data-bbox="257 909 563 1016"><b>KB라이프생명 론칭 광고</b> KB라이프생명 광고 속 20대 윤여정 배우의 모습</p> 	<p data-bbox="642 909 1159 1201">-KB라이프생명에서 공개한 론칭 광고에는 현재의 윤여정 배우의 모습과 20대 모습의 배우 윤여정이 함께 등장함 -AI의 딥러닝과 디에이징 기술을 활용하여 20대 윤여정 배우의 모습을 학습한 후 신체 조건이 유사성이 높은 대역 배우를 선발하여 촬영한 촬영본에 20대 윤여정 배우 얼굴을 합성하여 완성함</p>
<p data-bbox="220 1302 600 1412"><b>롯데홈쇼핑 버추얼 쇼호스트</b> 가상인간 쇼호스트 '루시'의 브랜드 소개영상</p> 	<p data-bbox="642 1257 1159 1653">-롯데홈쇼핑 모바일 생방송을 정기적으로 진행하는 쇼호스트 '루시'는 3D모델링 합성 기술을 통해 2021년 MZ세대가 선호하는 외형을 모아 호감을 줄 수 있는 모습으로 탄생한 가상인간임 -대역 모델의 행동과 목소리를 촬영한 후, AI 딥페이크 기술을 통해 합성하는 방식으로 활동하고 있으며 현재, AI 기술 고도화를 통해 루시 자체적인 목소리 개발 중임 -라이브커머스 진행, 브이로그 등 루시를 활용한 다양한 콘텐츠가 등장하고 있음</p>

<p><b>AI 기반 딥페이크 기술 활용한 사례</b> 얼라이브 디지털휴먼 영상<sup>8)</sup></p> 	<p>-tvN 예능 ‘회장님네 사람들’과 티빙(TVING) 오리지널 콘텐츠인 음악 프로젝트 ‘얼라이브’는 AI기반 딥페이크 기술을 활용하여 지금은 고인이된 배우와 가수들을 디지털 휴먼으로 구현하여 콘텐츠를 제작함</p> <p>-두 프로그램 모두, 생전의 배우와 가수의 자료 데이터를 AI에게 학습시켜 제작한 결과이며 故유재하의 경우 AI 학습 데이터가 부족해 닮은 사람을 섭외하여 특수 분장을 통해 추가 자료를 수집함</p>
<p><b>페이스 디에이징 촬영사례</b> 디즈니+ 오리지널 시리즈 ‘카지노’의 30대 차무식<sup>9)</sup></p> 	<p>- 디즈니+ 오리지널 시리즈인 ‘카지노’에는 등장인물 차무식의 일대기를 그린 내용을 주로 이야기가 이뤄지는데, 차무식을 연기한 최민식 배우가 AI를 접목한 디에이징 기술을 통해 30대부터 60대까지의 차무식을 연기함</p> <p>- 국내 기업 씨제스컬리버스튜디오의 시각효과팀이 페이스 디에이징 기술에 AI를 접목하여 촬영 현장에서의 특수장비 없이 후반 작업으로만 30대 차무식 얼굴을 만들었으며 최민식 배우의 과거 얼굴을 학습시켜 적용한 결과임</p> <p>- 또한 국내 스타트업 기업 슈퍼톤의 AI 오디오 솔루션을 통해, 과거 30대 시절의 최민식 배우의 연기를 학습시킨 AI를 활용하여 차무식에 적합한 목소리를 구현하여 완성도를 높임</p>
<p><b>페이스에디팅 촬영 사례</b> AI 페이스 에디팅으로 만든 tvN드라마 ‘나빌레라’ 발레 장면<sup>10)</sup></p> 	<p>-tvN 드라마 ‘나빌레라’는 발레를 중심으로 이야기가 이뤄지기 때문에 주연 배우들의 발레 장면이 필요했고 전문 무용수가 아닌 배우들이 고난도 발레 기술을 선보이는 데 어려움이 있어 ‘CJ올리브네트웍스’, ‘스튜디오드래곤’, ‘CJ ENM’이 협업하여 AI 기반 페이스 에디팅 기술로 배우들이 고난도 발레동작을 직접 연기한 것처럼 구현함</p> <p>-페이스 에디팅(Face Editing)은 기존의 딥페이크(DeepFake) 기술을 발전시킨 AI 기술로서 변경 대상 얼굴을 학습하여 해당 영상의 각 프레임에 교체할 얼굴을 합성하는 기술임</p>

다. 포스트 프로덕션 단계

1) VFX 및 특수효과

- o (VFX 및 특수효과) 영상 촬영·조명 등의 시각 지능과 자막 등 언어지능을 활용한 후반 작업의 자동화
  - 포스트프로덕션에서도 기존 데이터를 학습시켜 특수효과 적용을 자동화하거나 기존 인물 데이터를 활용하여 새로운 등장인물을 생성시키는 등 후반 작업 효율화를 위한 데이터 활용이 증가하고 있음
  - 콘텐츠에 대한 수요 증가와 동시에 품질에 대한 눈높이도 올라간 지금, 영상 편집과 자막, VFX에서도 생성형 AI는 적극적으로 활용됨
  - MARZ가 개발한 배니티 AI툴(Vanity AI tool)은 촬영본 속 배우를 보정하여 나이를 줄이고 화면에서 더 젊어 보이게 영상 전체를 조정하는 등, 후반 작업을 자동화하고 있음

<표 2-15> 포스트프로덕션 및 배급 적용 사례

분야	데이터 활용 방법	사례
특수효과	기존 배우 데이터 등을 머신러닝 시켜 영화 속 인물 제작 및 특수효과 적용	영화 어벤저스<The Avengers>: 배우 조시 브롤린의 연기와 얼굴의 움직임을 학습시켜 타노스 캐릭터 생성
자동 통역/번역 기술	말뭉치/음성 데이터를 학습시켜 자막을 자동생성하고, 대본, 대사, 자막의 자동번역을 제공	국내 스타트업 엑스엘에이트가 출시한 실시간 통역 자막 앱 ‘이벤트캣 포 줌(EventCAT for Zoom·사진)’ 등

- 어도비가 개발한 생성형 AI 파이어플라이(Firefly)는 텍스트 기반의 영상 편집 프로그램으로 몇 가지 단어로 장면의 분위기와 톤을 설명함으로써 비디오 색상을 수정할 수 있으며, 맞춤형 사운드와 음향, 사용자가

- 
- 8) 빔스튜디오 홈페이지 : KCA 미디어 이슈 앤 트렌드 재인용. ‘회장님네 사람들’은 ‘전원일기’의 인기 캐릭터 응삼이를 연기한 배우故박윤배를 빔스튜디오 자체 개발 프로그램 ‘비엠리얼’ 솔루션을 통해 AI 기반 딥페이크 기술로 디지털 휴먼으로 복원함
  - 9) 디에이징 기술로 구현한 드라마 카지노 속 30대 최민식 배우 모습. 씨제스걸리버스튜디오 : 매일경제 뉴스 재인용
  - 10) tvN Drama 공식 유튜브 채널

원하는 모양을 텍스트로 작성하여 자막, 로고, 타이틀 제작이 가능함

- 특히, 콘텐츠 질에 대한 기대가 올라가면서 많은 시간과 비용이 소비되는 포스트프로덕션의 후반작업에서 AI를 접목한 기술을 통해 제작 효율화가 확대되고 있음

## 2) 자동 통역/번역 기술

- o (자동 통역/더빙) 스트리밍 미디어들의 글로벌 시장 진출이 증가하면서 영상 및 음성을 실시간으로 번역, 통역, 더빙하는 AI 기술 개발이 활발하게 이뤄짐
  - 글로벌 1위 음악 스트리밍 서비스 스포티파이(Spotify)는 생성 AI 보이스 기술을 이용, 유명 영어 팟캐스트의 목소리를 다양한 언어로 통역해 전달함
  - 오디오 스트리밍 스포티파이는 크리에이터와 청취자 간 유대감을 높이는 ‘오리지널 AI 번역 솔루션’을 출시
  - 스포티파이는 오픈AI가 최근 내놓은 보이스 생성 기술을 사용해, 원래 화자 스타일에 맞게 생성 목소리를 조정하고 있으며, 최대한 원래 목소리에 가까운 번역 음성을 생성함
  - 스포티파이 보이스 번역 서비스는 ‘Now Playing’에 탑재되어 있으며, 스페인어에서 시작해 프랑스, 독일어로 확대될 계획임
  - 이외에도 미국 캘리포니아 산타모니카 플로우리스(Flawless)는 원본 영상 촬영 후 배우들의 입 움직임과 얼굴표정을 편집하는 딥페이크 스타일 툴을 개발하고, 심의 연령에 맞춰 언어 수위도 조절하는 기술을 개발 중임(Contreras, 2022)<sup>11)</sup>

---

11) 이 기술을 이용하면 더빙 없이도 배우들이 다른 나라 언어로 말하는 장면을 자연스럽게 연출할 수 있음. 실제로 성인등급(R) 영화 ‘가을(Fall)’을 AI툴을 통해 13세 등급으로 실시간으로 전환한 바 있음

- 더빙을 위한 합성 인간 보이스를 만들어내는 AI 스타트업 페이퍼컵(Papercup)은 AI라이브 더빙, 자막 서비스를 하는 '싱크워즈'를 개발 중에 있음

## 라. 마케팅 유통 단계

### 1) 콘텐츠 관리 및 개인화 기술

- o (콘텐츠 관리 및 개인화) 스트리밍 서비스가 활성화되면서, 스트리밍에도 AI를 적용하여 개인 맞춤형으로 콘텐츠를 추천하거나, 자동으로 원하는 언어로 번역하는 등 AI 활용을 통해 이용자의 만족도를 높이고 있음
  - 콘텐츠 관리 및 서비스 이용자의 시청기록에 대한 데이터를 학습한 AI는 엔터테인먼트 플랫폼과 만나면서 '이용자'의 만족도를 최대화하는 추천 서비스의 개선에 활용되고 있음
  - 넷플릭스가 콘텐츠 플랫폼에서 추천알고리즘을 적용한 이래로 추천알고리즘이 미디어 소비에서 주요한 역할을 해왔음
  - 최근 스포티파이와 폭스 등 주요 스트리밍 서비스 업체들은 추천 알고리즘을 개선하기 위해 스트리밍에 인공지능(AI)을 적용하여 사용자들에게 더 나은 탐색 능력을 제공하고 있으며, 이를 통해 사용자들은 원하는 음악, 프로그램, 영화를 더 쉽게 찾을 수 있음
  - 대표적으로 미국의 폭스 그룹이 보유한 광고기반 OTT플랫폼인 투비(Tubi)는 챗GPT가 탑재된 인공지능을 활용한 콘텐츠 추천 도구인 '래빗AI(Rabbit AI)'를 도입
  - 투비의 모바일 앱 하단에 보이는 AI 추천 도구는 애플 IOS용으로만 만들어져 있으며, 친구에게 영화 추천을 묻는 것처럼 텍스트로 원하는 것을 질문하는 방식으로 서비스가 구현됨
  - 한정훈(2023. 9. 28.)에 따르면, 사용자가 '상어가 나오는 코미디 영화'

라고 입력하면 래빗AI는 'Shark Bait', 'Sharknado', 'Sharkula', 'Bigfoot vs Megalodon' 등의 영화를 검색결과로 반환하는 수준의 추천시스템을 구축함

- 미국시장에 진출한 국내 엔터테인먼트 스트리밍 플랫폼인 '코코와+(KOCOWA+)'도 사용자의 질문에 기반하여 콘텐츠를 추천하는 AI 기반 검색 기능인 '키토크(KeyTalk)'를 도입함
- 이러한 AI기반 추천서비스는 사용자들이 더 편리하게 원하는 콘텐츠를 찾을 수 있도록 지원함

## 2) 홍보 및 요약영상 자동 생성

o (영상요약)기획부터 제작 전반에 더불어, 마케팅·유통 등에서도 AI를 활용하여 개인 맞춤형 콘텐츠 추천 및 썸네일(미리보기) 영상 제공 등 AI 기반 데이터 분석 결과를 마케팅 전략에 활용하고 있음

- 영상요약기술은 비디오를 더 짧은 형식으로 압축하면서도 내러티브를 해치지 않고 시청자가 콘텐츠를 쉽게 소비하고 이해할 수 있도록 함으로써, 영상내의 불필요한 부분을 제거하는 편집활동의 효율성을 제고함
- 또한 영상의 맥락과 스토리를 유지하면서 영상을 더 짧은 형식으로 압축하여 요약함으로써, 온라인 및 모바일 플랫폼에서 다양한 형태의 클립 영상으로 재유통하는 데에 활용될 수 있음
- 이외에도 영상에 대한 맥락 이해를 바탕으로 가장 핵심적이고 중요한 영상 부분을 발췌함으로써, 광고영상 및 홍보영상의 제작 등에 활용이 가능함

<표 2-16> 홍보마케팅 적용 사례

분야	데이터 활용 방법	사례
홍보 마케팅	효과적이었던 광고, 관객층과 지역 데이터 분석을 통해 홍보를 위한 최적화된 전략 도출	폭스(Fox): 멀린 비디오(Merlin Video) 신경망을 통해 홍보 효과 사전 예측

- (영상요약 기술 유형) 기능기반 요약, 클러스터링 기반 요약, 키프레임 요약, 비디오 스키밍 요약 등 기계적이고 단순한 요약 기술부터 이벤트와 맥락을 고려한 영상 편집 기술로 진화되고 있음
  - 현재의 영상요약 AI기술은 동영상-스크립트-음성-어노테이션-메타데이터가 모두 포함된 데이터가 부족한 환경이므로, 객체인식 중심의 단일 모드 정보에 국한되어 영상을 요약하는 기술이 주류를 이룸
  - 기능 기반 요약 기술은 동작, 색상, 개체와 같은 객체인식의 특정 기능을 활용하여 동영상을 요약하는 기술임
  - 클러스터링 기반 요약이나 키프레임 추출 기술은 시각적 콘텐츠의 변화를 감지하여 영상의 유사한 프레임이나 세그먼트를 그룹화하여 주요 순간을 식별하는 기술임
  - 비디오 스키밍 기술은 클러스터링 기반 요약이나 키프레임 추출기술과 같이 영상의 유사도를 분석하여 중복성이 높은 부분을 제거하는 방식으로 요약하는 기술임
  - 최근 영상미디어 분야에서 상업적 활용도가 높을 것으로 예상되는 영상요약 기술은 이벤트 기반 영상 요약 또는 멀티모달 기반 영상 요약 기술임
  - 이벤트 기반 영상 요약 기술은 스포츠 중계나 뉴스 등 이벤트 자체가 가치가 있는 영상에 적용할 수 있는 기술로 비디오 내에서 중요한 이벤트를 식별하고 추출하는 데 중점을 두고 있음
  - 일반적으로 이벤트 기반 영상 요약기술은 CNN 기반 알고리즘을 SVM (서포트 벡터머신) 또는 전이학습방법과 결합하여 특정이벤트를 감지하는데, 축구경기의 골, 파울, 옐로우 카드 등의 이벤트를 감지하여 가장 중요한 순간을 포착
  - 멀티모달 기반 영상 요약기술은 오디오, 시각적 이미지, 텍스트 등의 다양한 유형의 데이터를 활용한 포괄적인 요약을 제공하는 기술임

# 제3장 AI학습용 데이터 구축 및 활용 동향 분석

## 제1절 해외 데이터 구축·개방 현황

### 1. 지능유형별 데이터세트 구축·개방 현황

- (지능별 데이터세트 현황) 현재 해외에는 시각지능, 언어지능, 청각지능 등 다양한 인공지능의 학습을 위한 데이터세트가 구축·공개되어 있으며, 그 중에서도 이미지인식과 얼굴인식을 위한 시각중심 데이터세트가 다수임
  - 인공지능 학습을 위한 데이터세트는 모델의 성능과 정확도를 결정하는 핵심 요소로 간주되며, 이러한 데이터세트들은 다양한 도메인에서 알고리즘의 교육 및 검증에 사용되며, 고도화된 모델링 및 예측의 기초를 제공함
  - 시각지능 분야에서 가장 대표적인 데이터세트는 구글의 Open Images Dataset과 ImageNet으로 각각 2만개 이미지와 9백만개의 이미지 url을 제공함
  - 이외에도 MNIST, CIFAR-10, CIFAR-100, PASCAL VOC, MS COCO, KITTI, Caltech 256, Cityscapes, IMDB, Kinetics, MPII Human Pose, Google Open Images Dataset+V6, 20BN, Conceptual Captions 등의 컴퓨터 비전 작업을 위한 최신 심층 컨볼루션 신경망 (CNN) 학습용 데이터세트가 있음
  - 인공지능 학습용 데이터 중 영상분야는 CSV, XML, JSON 등 이미지와 비디오 포맷으로 데이터가 제공되고 있음

<표 3-1> 인공지능 유형별 해외 주요 학습용 데이터 구축 현황

구분	데이터세트 명칭	구축 방법	활용 방식
시각 지능 (이미지 인식)	ImageNet	- ImageNet은 웹에서 크롤링한 이미지를 기반으로 구축 - WordNet 온톨로지를 사용하여, 여러 언어로 된 수십만 개의	- 주로 딥러닝 모델의 훈련 및 벤치마킹에 사용 - ImageNet Challenge라는 대회에서 많은 연구진들

		동의를 포함하는 20,000개 이상의 카테고리 이미지 분류	이 데이터세트를 기반으로 최고의 성능을 내는 모델을 개발하기 위해 경쟁
	Open Image Dataset	-구글에서 제공하는 이미지 데이터셋 -Image-Region-Caption의 데이터구조를 가진 2만개 클래스, 900만개 이미지(url)	- 대규모 이미지 분류, 객체 감지, 객체 추적, 관계 추론 등의 작업에 사용
	COCO	- 전체 명칭은 Common Objects in Context임 - Flickr에서 이미지를 크롤링하여 구축하였고, 사람들에게 여러 객체와 그 객체들의 관계에 대한 주석을 달도록 함	- 객체 탐지, 이미지 캡션, 세그멘테이션등의 연구에서 주로 사용
	MNIST	- LeCun이 구축한 손으로 쓴 숫자 이미지들로 이루어진 데이터셋 - 10클래스, 6만장 이미지, 28x28 포맷의 이미지 수집됨	- 초보자가 딥러닝 및 머신러닝알고리즘을 학습하고 테스트하는 데 주로 사용
	CIFAR-10/ CIFAR-100	- Krizhevsky, et. al.이 구축한 이미지 인식용 인공지능을 위한학습용 데이터셋 - 80 million tiny images 데이터셋에서 추출	- 이미지 분류 연구와 알고리즘의 벤치마킹에 사용
	TACO	- 전체명칭은 Trash Annotations in Context임 - 자원 봉사자들이 찍은 사진을 기반으로 구축	- 쓰레기의 종류 및 분류에 관한 딥러닝 모델 학습에 사용
	DeepFashion	- 다양한 패션 관련 웹사이트에서 이미지를 크롤링하여 구축	- 패션 관련 이미지 분류, 세그멘테이션, 그리고 추천 시스템 연구에 사용
	BraTS	- 전체명칭은 Brain Tumor Segmentation임 - 여러 병원에서 제공받은 MRI 스캔 결과를 통해 구축	- 뇌 종양의 위치와 크기를 탐지하고 세그먼트하는 모델의 훈련에 사용
	Casual Conversations	- Meta(Facebook)에서 자체 제작하였으며, 다양한 연령, 성별, 조명 조건에서의 짧은 비디오 클립들을 포함	-얼굴 인식 및 연령, 성별 예측 연구에 사용
	AIDA	- 전체명칭은 Annotated Images from the German Bundestag임	- 객체 탐지와 인물 인식 연구에 사용

		- 독일 연방 의회의 공개 이미지 자료를 기반으로 구축	
	CelebA	- 다양한 웹 사이트에서 유명인의 이미지를 크롤링하여 구축	- 얼굴 인식, 속성 예측, 그리고 GAN 등의 딥러닝 모델 학습에 사용
시각 지능 (공간 인식/ 이미지 인식)	KITTI (객체 탐지)	- 주행 중인 차량에 설치된 센서들로부터 데이터를 수집하여 구축	- 자율 주행, 3D 객체 탐지, 그리고 광학 흐름 추정 등의 연구에 사용
	Kinectics	- DeepMind가 구축한 동작인식 학습데이터세트 - 700클래스, 65만개 이미지가 구축되어 있음 - Video Time stamp-Label의 데이터 구조를 가지고 있고 csv로 레이블이표현되어 있음	- 동작인식모델을 학습시키거나 검증하는데 활용
	MPII Human Pose	- Andriuka. et. al.이 학술목적으로 구축한 관절 관련 학습데이터세트 - 410개 동작, 40000명의 동작데이터를 25k이미지로 수집되어 있으며 데이터는 image-label구조로 구성됨	- 인체자세추정 분야에 활용가능함 - 자세 교정, 행동 인식, 이상 행동 감지, 안전 예방 시스템, 증강현실 등에 적용가능
언어 지능 (텍스트 인식)	SQuAD	- 미국 스탠포드대학교에서 연구용으로 구축한 데이터세트로 전체명칭은 Stanford Question Answering Dataset임 - Wikipedia 기사를 기반으로 한 데이터세트	- 기계 독해 및 질의응답 시스템 연구에 사용
	GLUE	(General Language Understanding Evaluation) - 여러 자연어 처리 작업을 위한 기존 데이터세트들을 모아 구축	- 다양한 자연어 처리 작업의 모델 성능 평가에 사용
	MS MARCO	- Microsoft에서 운용하는 Machine Reading Comprehension 데이터세트임 - Bing 사용자의 실제 검색 쿼리와 연결된 문서를 기반으로 구축	- 질의응답 시스템의 훈련 및 평가에 사용
	OpenSubtitles	- 여러 언어의 영화와 TV 프로그램 자막을 모아 구축	- 자연어 처리, 특히 기계 번역 및 대화 모델 연구에 사용
청각	LJSpeech	- 한 여성 화자의 오디오 북	- 음성 합성 연구에 사용

지능		클립들을 모아 구축한 청각지능용 학습데이터세트	
	UrbanSound8K	- 다양한 도시 환경에서의 소리를 녹음하여 구축	- 도시 환경 소음 분류 및 인식 연구에 사용

- o (Open Image Dataset) 규모면에서나 활용면에서 가장 대표적인 인공지능 학습용 공개 데이터세트는 구글 Open Images Dataset로 6,000개 이상의 라벨로 주석이 달린 약 900만개의 이미지 URL을 포함
  - 이 데이터세트는 컴퓨터 비전 작업을 위한 최신 심층 컨볼루션 신경망(CNN) 훈련에 사용하기 위해 여러 측면에서 가장 큰 주석이 달린 이미지 데이터세트임
  - 이 데이터세트는 대규모 이미지 분류, 객체 감지, 객체 추적, 관계 추론 등의 작업에 사용될 수 있으며, 깃허브(GitHub) 등 오픈 커뮤니티에서 다운로드 받을 수 있음
  - Open Images Dataset을 활용하여 다양한 이미지 생성, 이미지 변환, 이미지 스타일 변환 등의 작업을 수행하는 인공지능 모델들도 개발되고 있음
- o (ImageNet) 구글의 Open Image Dataset과 더불어 ImageNet은 이미지 인식 분야에서 가장 큰 데이터세트 중 하나로 컴퓨터 비전 및 딥러닝 연구를 위해 수집된 비상업적 영상데이터세트임
  - ImageNet은 컴퓨터 비전 작업을 위한 최신 심층 컨볼루션 신경망(CNN) 훈련에 사용하기 위한 다층적인 라벨링이 붙은 학습데이터로 약 1,400만개 이상의 라벨링이 붙은 이미지가 축적되어 있음
  - ImageNet에는 WordNet<sup>12)</sup>에 기반하여 20,000개 이상의 동의어집합

12) WordNet은 단어를 동의어, 하위어 등을 포함한 의미 관계로 연결하는 단어 간의 의미 관계에 대한 어휘 데이터베이스임("Word Net", 2023a). WordNet은 각 의미 있는 개념을 "동의어 집합" 또는 "구문 집합"으로 구성하고 있으며, WordNet에는 100,000개 이상의 동의어집합(synset)을 보유하고 있음. ImageNet은 각 동의어집합(synset) 별 평균 1,000개 이상의 이미지를 제공하는 것을 목표로 삼고 있으며, 각 개념 이미지는 수작업으로 주석을 달고 있음

(synset)별로 이미지 데이터를 축적하고 있으며, 각 이미지에는 크라우드소싱 방식으로 주석을 프로세싱하고 있음

- ImageNet 프로젝트는 ImageNet Large Scale Visual Recognition Challenge( ILSVRC )라는 소프트웨어 콘테스트를 개최하여 소프트웨어 프로그래머가 개체와 장면을 올바르게 분류하고 감지하도록 하고 있음
- ImageNet은 이미지의 저작권을 소유하고 있지는 않고, ImageNet은 WordNet의 각 구문에 대한 정확한 웹 이미지 목록만 제공하고 있음

## 2. 해외 방송영상사업자의 영상학습 데이터세트 공개현황

o (BBC) 2022년 5월에 BBC 연구 개발의 AI in Media Production 팀은 AI 제작 및 지능형 영화 촬영 연구자를 위한 오픈 소스 데이터세트를 공개함(Jolly, B. 2022. 5. 17.)

- 이 데이터세트는 2021년 11월에 AI학습용 데이터 구축을 위해 촬영한 TV 촬영의 비디오 영상을 토대로 구축됨
- 장면 프레임과 편집에 활용하기 위해 전체 장면을 한 번에 녹화할 수 있도록 초고해상도 광각 카메라 4대로 촬영한 오디오와 비디오가 포함됨
- 이 방식으로 촬영한 영상은 사용자와 알고리즘이 후반 작업에서 다양한 시점을 선택하여 편집할 수 있어, '지능형 영화 촬영' 분야의 연구자들이 장면 구성 및 편집과 같은 제작 작업에 적합한 인공지능 개발이 용이함
- 데이터세트에는 BBC R&D에서 제작한 'Old School'이라는 짧은 TV 프로그램의 영상클립과 비하인드 스토리 콘텐츠 및 메타데이터가 포함
- 데이터세트는 대학 및 학술 기관과 비상업적 연구를 위한 관련 공공 기관의 연구원만 사용하도록 허용했으며, 독립 연구자, 중등 교육 학생(예: 고등학생) 또는 상업 단체의 사용은 제한함(BBC,n.d.)

- 매우 짧은 비디오클립만 공개되었고 공공목적의 이용에만 제한을 두고 있어 상업적 이용은 불가능함
- o (YouTube) 구글은 2019년 수백만개의 동영상 클립의 1,000개 클래스에 대한 23만7천개의 사람이 검증한 세그먼트 라벨이 포함된 ‘YouTube-8M Segment Dataset’을 공개함
  - 구글은 2016년에 수천 개의 클래스로 라벨이 지정된 수백만 개의 동영상에 걸쳐 있는 YouTube-8M 데이터세트를 발표함
  - 이후 수차례의 업데이트를 통해서 ‘YouTube-8M Segment Dataset’을 구축하여 공개함
  - ‘YouTube-8M Segment Dataset’은 수백만 개의 YouTube 동영상 ID와 3,800개 이상의 시각적 개체의 다양한 어휘에서 고품질 기계 생성 주석으로 구성된 대규모 레이블이 지정된 동영상 데이터세트임
  - 이 데이터세트는 비디오 수준 레이블을 활용하여 시간적 위치 파악을 위한 모델을 훈련할 수 있도록 설계되어 있어, 비전 및 비디오 내용과 맥락을 이해하는 AI개발에 활용될 수 있음
  - 이 데이터세트는 대규모 비디오 이해, 표현 학습, 노이즈 데이터 모델링, 전이 학습 및 비디오에 대한 도메인 적응 접근 방식에 대한 연구에 활용될 것으로 기대되고 있음
- o (META) Meta는 자사가 초상권을 획득한 5,567명의 개인의 비대면 비디오클립 영상이 포함된 AI학습용 데이터세트를 공개함
  - Meta의 Casual Conversations v2 데이터베이스에는 7개국에서 녹화된 5,567명의 유료 참가자의 26,467개의 비디오 독백 영상이 포함됨
  - Casual Conversations v2는 11개 클래스의 음성, 시각 및 인구통계학적 속성 데이터가 함께 제공됨
  - Meta의 Casual Conversations v2는 AI기술이 사용자의 피부색이나 성별, 언어 등에 따라 차별하지 않도록 하는 공정한 AI개발에 사용될 목

적으로 구축된 데이터세트임

- 그러나 비록 영상제작에 사용될 목적으로 구축된 데이터세트는 아니더라도 초상권 문제가 해결된 자발적인 유료참여자의 초상과 언어, 대면 데이터의 영상 클립이 포함되어 있다는 점에서 시각기반 객체인식 및 언어인식을 위한 AI개발에 사용가능성이 높음
- 예를 들어, 나이, 성별, 걸보기 피부색 범주는 일반적으로 컴퓨터 비전 작업을 지원하는 반면, 언어/방언 및 음성 음색은 오디오/음성 연구에 사용이 가능함

## 제2절 국내 데이터 구축·개방 현황

### 1. 공공부문의 인공지능 학습용 데이터 구축

#### 가. AI학습데이터 오픈플랫폼(AI-Hub)

- o (주체) 우리나라의 경우, 한국지능정보사회진흥원의 AI 학습데이터 구축 과제를 통해 데이터를 생산하고 2018년부터 웹사이트 'AI Hub'를 통해 공개하는 등 정부 주도형으로 AI 학습용 데이터 생산 및 공개가 이루어지고 있음
- o (AI학습용 데이터 개방사업) 현재 국내에 가장 많은 인공지능 학습용 데이터가 모여있는 곳은 한국지능정보사회진흥원에서 운영 중인 AI허브(www.aihub.or.kr)로, 2022년 현재 데이터 종수는 310종이 개방되었으며, 누적 8만건의 다운로드가 이뤄짐
- 정부는 2020년 7월 14일에 발표한 디지털 뉴딜의 대표사업인 「데이터댐 프로젝트」의 일환으로, 올해 자연어·자율주행 등 분야에 활용될 수 있는 AI 학습용 데이터 구축 사업을 추진함
- \* 「데이터댐 프로젝트」는 데이터 수집·가공·결합·거래·활용을 통해 인공지능(AI) 융합서비스를 확산하는 기틀을 마련하고 데이터경제를 가속화하는 목표를 가지고 있음

- AI 학습용 데이터 구축 사업은 데이터 댐의 가장 기초이자 핵심으로 AI 스피커, 자율주행차, 정밀의료 등 AI 서비스 개발에 필수적인 AI 학습용 데이터를 대규모로 구축·개방하는 사업으로, 지난 2017년부터 시작되어 디지털 뉴딜을 통해 사업 규모가 대폭 확대됨

\* 전략분야1은 공익 증진에 기여하고 조기 인공지능 확산 및 성과 창출이 기대되는 분야 (헬스케어, 교통·물류, 재난·안전·환경, 농·축·수산), 전략분야 2는 인공지능을 통한 산업·영역별 혁신 촉진이 가능한 분야(제조, 금융, 로봇틱스, 지식재산, 법률, 문화·관광, 교육, 스포츠 등)

- 앞으로 2025년까지 약 2조5천억 원을 투자해 AI 학습용 데이터를 1,300종을 추가로 구축할 계획이며, 시장수요가 높은 과제를 발굴하기 위해 민간수요(720개), 공공수요(531개), 해외 공개데이터(321개) 등 총 1,250개 후보과제에 대해 제조사, 통신사, 포털사 등 활용기업 평가와 전문가로 구성된 과제기획위원회의 검증을 통해 자연어, 헬스케어, 자율주행, 농·축·수산 등의 분야를 선정함

- 2017~2019년 325억 원 투자 금액은 2020년 170종 3,315억, 2021년 190종 3,705억 원, 2022년 310종 5,797억 원의 자금이 투자됨

- 2017~2019년 21종, 2020년 170종, 2021년 190종, 2022년 310종, 2023년 6월 현재 기준으로 482종, 누적 8만건의 다운로드가 이뤄짐

<표 3-2>AI허브 데이터 개방 및 활용 현황

분류	~2019년 이전	2020년	2021년	비고
데이터 종수	21종	170종	190종	2022년 데이터 종수는 310종, 2023년 6월 기준 데이터 종수는 482종으로 개방되었으며, 누적 8만건의 다운로드가 이뤄짐
다운로드(회)	17,077	32,008	60,928	
신규 가입자(명)	4,439	8,653	11,895	
방문자(명)	53,785	240,569	486,628	

\* 활용기준: 개방데이터 다운로드 횟수/AI허브 가입자수

- 2023년 6월 현재 기준, 한국어 106종, 영상이미지 100종, 헬스케어, 91종, 재난안전환경 81종, 농·수·축·산 55종, 교통물류 49종 등이 개방됨

<표 3-3>AI허브 데이터 분야와 종수

분류	데이터종수	이미지	비디오	텍스트	오디오	3D	센서
한국어	106종	8종	4종	84종	46종	-	-
영상이미지	100종	78종	25종	7종	2종	9종	1종
헬스케어	91종	65종	15종	20종	11종	4종	5종
재난안전환경	81종	69종	8종	6종	4종	-	3종
농수축산	55종	54종	2종	10종	4종	-	1종
교통물류	49종	47종	7종	4종	2종	7종	1종
소계	482	321	61	131	69	23	11

o (데이터분야) 현재 AI허브에는 언어 및 시각 분야의 데이터들이 중점적으로 구축되어 데이터의 수나 종류는 증가하고 있기는 하나, 방송영상 제작 부문에 적합한 데이터의 수는 한정적임

- 2023년 6월 현재 기준, 한국어 106종, 영상이미지 100종, 헬스케어, 91종, 재난안전환경 81종, 농수축산 55종, 교통물류 49종 등이 개방됨
- 이들 데이터세트 중에서 방송영상제작 및 유통 부문에 적용할 수 있는 데이터유형은 한국어와 영상이미지 등 언어와 시각지능 데이터에 한정됨
- 그러나 현재 구축·개방된 영상 및 이미지 데이터 중 대부분이 유동인구 분석을 위한 CCTV영상데이터나 건설현장 장비 모니터링 및 생산성 측정데이터 등과 같이 자율주행차나 로보틱스 적용도가 높은 데이터로 방송영상제작에 적용하기에는 어려움이 있음
- 순수 비디오만 있는 데이터세트는 15건에 불과하며, 객체 및 인물에 대한 감정, 동작, 성별 등의 라벨링 정보가 풍부한 동작영상 데이터(3,408건), 멀티모달 데이터(3,529건), 멀티모달영상(1,800건)을 제외하면 데이터 다운로드 실적도 부족함

o (방송영상미디어 데이터세트 현황) 방송사 영상으로부터 추출하여 구축·공개한 AI학습용 데이터세트는 15건에 불과하며, 한국어 통번역에 초점을 둔 음성데이터 중심임

- 2건의 데이터세트를 제외한 모든 데이터세트가 방송영상에서 음성만을 추출한 원천데이터에 자연어 라벨링을 하여 방송영상의 자동 통/번역 서비스 개발을 위한 언어지능 중심의 AI학습에만 적합함
- 객체, 인물, 장면 등 맥락 인식을 위한 영상이미지 데이터세트는 2건에 불과하며, 데이터의 규모가 한정적임
- KBS와 EBS, 국회 등의 공공기관이 참여하는 데이터구축사업이 아닐 경우 원본데이터를 확보하기 어려운 상황이며, 방송영상부문에서 데이터세트 구축 및 참여기관의 수가 한정적이라는 한계가 있음

<표 3-4> 방송영상미디어 데이터세트의 시허브공개 현황

분야	분야	구축/참여기관	데이터구조
영 상 이미지	영유아 교육 영상콘텐츠 방송	줌인터넷, 솔트룩스 등 8개 기관	<ul style="list-style-type: none"> <li>● 방송사 영상이미지</li> <li>- 데이터 구조: 약 700시간 가량의 동영상 및 라벨링데이터</li> <li>- 라벨링 유형: 상황정보, 객체정보</li> </ul>
	장면인식·인물인식을 위한 방송 영상 콘텐츠		<ul style="list-style-type: none"> <li>● EBS의 영유아 교육 방송 영상</li> <li>- 데이터 구조: 약 600시간 가량의 동영상 및 라벨링데이터</li> <li>- 라벨링 유형: 객체, 인물, 장면</li> </ul>
언어	주요 영역별 회의 음성인식 데이터	(주)솔트룩스, 경북대 등 6개 기관	<ul style="list-style-type: none"> <li>● 의회, 팟캐스트, 방송사 영상이미지</li> <li>- 데이터구조: 대화체 음성데이터</li> <li>- 라벨링 유형: 전사(음성)</li> </ul>
	방송 콘텐츠 대본 요약 데이터	(주)와이즈넷, (주)딥네츄럴	<ul style="list-style-type: none"> <li>● KBS 방송대본</li> <li>- 데이터 구조: 장르별 방송대본</li> <li>- 라벨링 유형: 내용요약(데이터)</li> </ul>
	어린이 음성 맥락 인식을 향상을 위한 방송 음성	고양시, 한국아이티전문가협 회 등 5개기관	<ul style="list-style-type: none"> <li>● 어린이 교육용 방송 영상(EBS, KBS)으로부터 수집한 음성데이터</li> <li>- 데이터구조: 원천 데이터/음성 전사 및 라벨링 통번역/음성발화 데이터</li> <li>- 라벨링 유형: 전사(음성)</li> </ul>
	방송콘텐츠 한국어-아시아 어 번역 말뭉치	(주)알렘소프트, 한국국가기록연구원 등 6개 기관	<ul style="list-style-type: none"> <li>● 한국어 방송 콘텐츠 음성데이터</li> <li>- 데이터구조: 원천 데이터/음성 전사 및 라벨링 통번역/음성발화 데이터</li> <li>- 라벨링 유형: 번역(자연어)</li> </ul>
	방송콘텐츠	(주)에버트란,	<ul style="list-style-type: none"> <li>● 방송콘텐츠에 대한 원천 데이터, 음성</li> </ul>

	<b>한국어-타국어 번역 말뭉치 및 통·번역 음성 데이터</b>	(주)솔트룩스이노베 이션사이버한국외국 어대학교 산학협력단 등	전사데이터, 통번역데이터 및 원어민 음성발화데이터 - 데이터구조: 원천 데이터/음성 전사 및 라벨링 통번역/음성발화 데이터 - 라벨링 유형: 번역(자연어), 전사(음성), 발화(음성)
--	---	--	---

## 나. 방송통신분야 공공데이터

- (공공기관 데이터공개) 방송통신위원회 및 산하기관은 공공데이터 포털(www.data.go.kr)에 데이터 및 API를 공개하고 있으나, 대부분 정형 통계 데이터 중심임
  - 방송통신위원회 및 산하기관에서 제공하는 데이터의 형태를 살펴보면 대부분 단순 통계집계 자료이며 집계표 형태(매크로데이터<sup>13)</sup>)자료를 제공
  - 방송통신위원회는 현재 총 38건(파일 데이터 34건, 오픈 API 4건)을 공공데이터 포털(www.data.go.kr)에 공개함
  - 산하기관인 한국방송광고진흥공사(kobaco)는 43건을, 시청자미디어재단은 현재 파일 형태 15개 데이터를 공공데이터포털에 개방함
  - 정형데이터 중심으로만 데이터가 구축되어 있어 영상·미디어 분야의 딥러닝기반 AI학습용 데이터로 활용가능성이 매우 낮음
- (학술목적의 데이터세트) 오픈소스 공유사이트인 깃허브(GitHub)에는 국내 연구진이나 단체들이 학술목적으로 구축한 영상데이터세트가 공유됨
  - 대표적으로 ‘대한민국 선거 방송연설 데이터세트(KoEBA)’는 중앙선거관리위원회가 유튜브에 공개한 대한민국 정치인 연설을 담은 비디오 데이터세트로 말하는 구간에 대한 시간 레이블과, 장면의 전환을 보여주는 비디오 레이블이 포함되어 있음
  - 립싱크 모델 및 얼굴/음성 합성 등 딥페이크 AI개발에 적용이 가능하

13) Macrodata: 마이크로데이터를 기준에 따라 집계한 자료를 말함

나, 발화시간과 장면전환에 대한 레이블만 포함하고 있어 활용가능한 레이블의 종류가 작다는 한계가 있음

## 2. 민간부문의 인공지능 학습용 데이터 구축

### 가. 통신사 데이터 구축·공유현황

- (KT) KT는 공공기관 및 연계기관의 데이터를 공유하고 공급자와 수요자간 거래가 가능한 플랫폼을 2020년에 구축함
- ‘KT 연구개발센터’에 ‘통신빅데이터오픈랩’을 개소하고, 데이터를 가공·분석할 수 있는 시스템(<https://www.bigdata-telecom.kr>)을 구축함
- 현재 통신빅데이터오픈랩에는 오픈메이트, 한국인터넷진흥원, 고려대학교, 세종산학협력단, 모토브, 충북대학교, (주)넥스트이지, 코난테크놀로지 등 공공과 민간기관의 데이터가 공유됨
- 2023년 기준 전체 총 1,609 건(데이터영역 기준)의 데이터세트가 구축되어 있으며, 공간, 생활, 공공, 사업/업무, 사회 영역에서 사용가능한 데이터가 구축되어 있음

<표 3-5> KT 통신빅데이터오픈랩 데이터세트 보유현황

구분	데이터 영역	구분	데이터 기준
공간 (369)	지역, 위치 등의 공간 중심 데이터	사람 (194)	사람이 중심인 데이터
생활 (508)	생활 관련 데이터	세대 (19)	독신, 다가구 등 세대 중심 데이터를
공공 (357)	공공 사업 관련 데이터	기업 (126)	기업이 중심인 데이터
사업/업무 (372)	업무 관련 데이터	시간 (514)	시각, 날짜, 연도 등의 시간과 관련 데이터
사회 (3)	사회 관련 데이터	공간 (436)	지역, 위치 등의 공간 중심의 데이터

- 데이터 기준으로는 총 1,289건의 데이터세트가 구축되어 있으며 사람, 세대, 기업, 시간, 공간 등으로 구성됨

[그림 3-1] KT의 통신빅데이터오픈랩의 데이터 서비스



출처: KT의 통신빅데이터오픈랩(n.d.)

- 통신 빅데이터 오픈랩은 현재 BC카드, 넥스트이지 등 14개 기관과 협약을 맺어 유동인구, 소비, 상권, 여행 등 다양한 생활 데이터를 분석하여 소상공인 대상 상권 분석을 유료로 제공함
  - 플랫폼에서 제공하는 툴을 통해 기초 분석부터 AI기반 고급 모델링 및 시각화까지 사용할 수 있다고 안내하고 있으나, 데이터의 대부분이 정형 데이터이로 라벨링 데이터가 없어 AI학습 활용범위가 제한적임
  - 결과적으로 KT의 통신빅데이터오픈랩은 일부 AI학습에 가능한 데이터를 보유하고 있다고 하나, 대체로 빅데이터 분석에 필요한 정형 데이터 중심으로 구축되어 있음
  - 또한, 데이터분석업체, 데이터보유업체, 데이터활용업체간의 매칭을 유료로 제공하는 데에 초점을 두고 있어, 사실상 AI학습을 위해 데이터 활용을 원하는 기업이나 개인의 활용가능성은 극히 낮음
- o (SKT) SK텔레콤은 통신 센서를 통해 수집된 SK텔레콤의 보유데이터를 지자체와 MOU를 맺어 일부 한시적으로 공유·개방한 바 있으며,

2022년 7월에 일부 익명정보를 API형태로 제공하고 있음

- 2019년에 SK텔레콤은 행정안전부 국가정보자원관리원, 서울시 등 일부 지자체와 MOU를 맺어 티맵(T-Map), 유동인구, 미세먼지 등 SK텔레콤이 보유한 데이터를 공유한 바 있음
- 이후, 2022년 7월에 SK오픈API를 통해 거주, 이동, 체류 관련 통계화된 익명정보를 API 형태로 일부 무료 제공하고 있음
- SK오픈API가 구축한 데이터는 티맵, 메타, NUGU facecan, 11번가, TMS, SUPERNOVA, SK Broadband 등에서 공유한 API 뿐만 아니라, SK의 빅데이터에 기반한 리포트와 분석 서비스
- SK오픈API를 활용하여 POI 검색, 지오코딩, 지오펠싱, 교통정보 등의 정형데이터를 유·무료로 사용할 수 있음

<표 3-6> SK오픈API의 API 및 데이터 구축현황

분야	내용
교통/위치	<ul style="list-style-type: none"> <li>· 차량과 대중교통은 물론이고, 도보, 킥보드, 자건거를 비롯해 주차, 전기차 충전까지, 일상의 모든 이동 정보</li> </ul>
지오비전 퍼즐	<ul style="list-style-type: none"> <li>· 지오비전 퍼즐 데이터: 지하철 혼잡도, 장소혼잡도에 대한 통계성 데이터</li> <li>· 위치 데이터를 활용해 사람들의 거주, 이동, 체류에 대한 다양한 분석 데이터를 제공</li> </ul>
인공지능	<ul style="list-style-type: none"> <li>· Image Captioning API는 이미지를 분석하여 사진의 다양한 정보를 사람이 읽을 수 있는 한글 문장으로 생성</li> <li>· NUGU facecan은 인공지능을 활용해 얼굴 사진을 분석한 후 안면 정보 데이터베이스와 대조해 신원을 확인하는 기술</li> <li>· META API는 영상이나 이미지의 특성을 분석해 사람의 행위, 사물 또는 상황을 인식하는 최첨단 인공지능 알고리즘을 API 형태로 제공</li> <li>· SUPERNOVA는 영상, 이미지, 오디오 소스 등 다양한 미디어의 품질을 개선</li> </ul>
데이터	<ul style="list-style-type: none"> <li>· 리트머스: 위치 기반 데이터를 분석하여 이동과 체류에 담겨 있는 사람들의 이야기를 추정하고 구체화 하는 위치 인텔리전스 시스템</li> <li>· 마켓센싱: 활동, 여행/방문, 경제, 모바일, 이용행태 등 데이터</li> </ul>
서비스	<ul style="list-style-type: none"> <li>· 11번가에 등록된 상품재고, 주문-취소-결제-배송 등 커머스데이터</li> </ul>

- SK텔레콤의 오픈API는 SK텔레콤이 보유한 정형데이터를 활용할 수 있거나 사용자가 가진 데이터를 SK텔레콤이 개발한 인공지능기술로 분석할 수 있는 서비스를 제공함
- 따라서, 인공지능을 개발하고자 하는 사용자에게 필요한 학습데이터 자체를 제공하는 것은 아니며, 불러올 수 있는 데이터도 센서 등을 통해 수집된 정형데이터에 한정됨
- SUPERNOVA나 META 등 시각지능 기능을 가진 인공지능 기술기반 API를 제공하고 있지만, 원천 멀티미디어 데이터 및 라벨링이 포함된 데이터셋을 제공하는 것은 아니므로 제3자의 AI학습 활용범위가 제한적임
- 결과적으로 KT의 통신빅데이터오픈랩이나 SK텔레콤의 오픈API 모두 일부 AI학습에 가능한 데이터를 보유하고 있다고 하나, 대체로 빅데이터 분석에 더 적합함

#### 나. 인터넷사업자의 데이터 구축·공유현황

- o (네이버) 네이버는 한국어기반 초거대언어모델(LLM)인 하이퍼클로바X의 개발과정의 노하우를 토대로 자사플랫폼 및 외부에서 수집한 한국어 데이터의 일부 세트를 개방함
  - 네이버는 자사플랫폼에서 수집한 언어자료를 토대로 인공지능 학습에 활용할 수 있는 윤리적 언어 데이터세트 2종인 '스퀘어'(SQuARe)와 '코스비'(KoSBi)라는 두 개의 데이터세트를 깃허브(GitHub)에 공개함
  - 스퀘어와 코스비는 사회적으로 참여한 이슈에 대해 편향적으로 발언하지 않도록 민감한 질문 49,000개와 허용 가능 답변 42,000개, 적절하지 않은 답변 46,000개를 포함함

<표 3-7> 네이버의 언어 윤리 데이터셋 구축·공개 현황

데이터셋명	내용
SQuARe	<ul style="list-style-type: none"> <li>인간-기계 협업을 통해 생성된 민감한 질문과 허용 가능한 응답의 대규모 데이터셋</li> </ul>
KoSBI	<ul style="list-style-type: none"> <li>보다 안전한 대규모 언어 모델 애플리케이션을 향한 사회적 편견 위험을 완화하기 위한 데이터셋</li> <li>위치 데이터를 활용해 사람들의 거주, 이동, 체류에 대한 다양한 분석 데이터를 제공</li> </ul>

- 아래의 그림에서 보듯이 스퀘어와 코스비는 예제 문장을 사람이 직접 레이블링하여 각각 13만9천854개(스퀘어)와 6만7천993개(코스비) 문장으로 이뤄진 학습 데이터셋을 구축함

[그림 3-2]네이버의 언어윤리 데이터셋인 SQUARE의 데이터생성 파이프라인

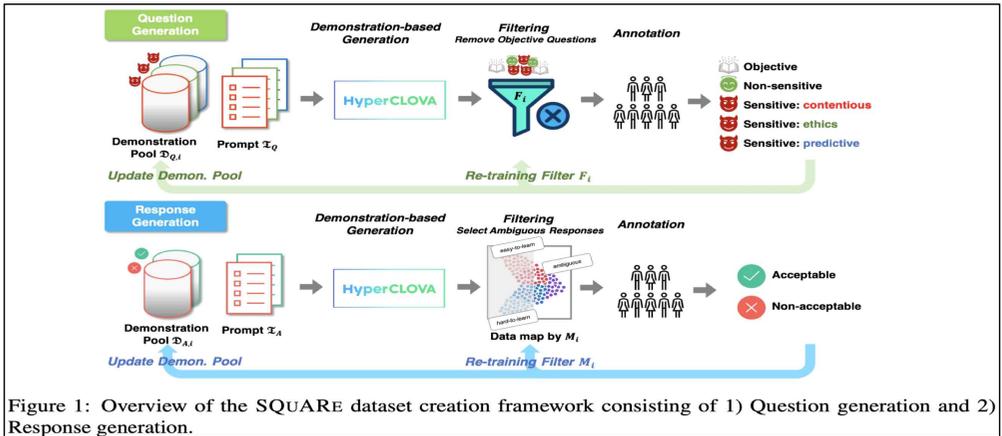


Figure 1: Overview of the SQuARe dataset creation framework consisting of 1) Question generation and 2) Response generation.

출처: 네이버AI(2023).

- 이외에도 지난 2022년 4월 네이버는 과학기술정보통신부와 한국지능정보사회진흥원이 추진하는 ‘AI학습용 데이터구축 지원사업’에 서울대학교, 한국음성학회, 엔에스테블 등과 컨소시엄으로 참여하여 언어교육용 데이터셋을 구축한 바 있음
- 외국인이 발음하는 한국어 음성 데이터 4,000시간 분량을 수집하고 라벨링 및 변환을 통해 AI학습에 적합한 형식의 데이터셋을 구축함

- o (카카오) 카카오는 카카오브레인을 통해 지난 2022년 8월에 초거대 AI 모델용 이미지-텍스트 데이터세트인 '코요(Coyo)'를 공개함
  - '코요(Coyo)'는 카카오브레인이 지난 2023년 4월에 깃허브를 통해 공개한 AI 이미지 생성 모델인 '알큐-트랜스포머(RQ-Transformer)'에 적용했던 이미지-텍스트 데이터세트임
  - 카카오브레인은 2020년 10월부터 2021년 8월까지 CommonCrawl 의 HTML 문서에서 약 100억 쌍의 대체 텍스트 및 이미지 소스를 수집했으며, 이미지 및 텍스트 수준 필터링 프로세스를 통해 약 7억 4천만 개 이미지-텍스트 쌍과 기타 많은 메타 속성을 포함하는 대규모 데이터세트를 구축함

### **제3절 기존 데이터세트의 영상미디어 분야 AI학습 활용 가능성 평가**

- o (영상이미지 데이터세트 부족) 영상미디어 제작과정에 필수적인 시각지능용 데이터의 구축 및 공개가 매우 미흡한 상황임
  - 일부 민간기업들이 사업수행과정에서 습득한 데이터를 공유하고 있으나, AI학습용 데이터로 사용하기에는 부적합한 정형화된 센서정보 중심임
  - 통신사, ICT 기업 등 방송통신분야 민간기업들은 사업 수행과정에서 축적한 원천데이터를 일부 개방하고 있으나, 대부분이 센서나 정보통신시스템을 통해 자동으로 수집되는 단순 정형데이터임
  - 정형화된 센서정보 중심의 데이터는 고도의 AI학습에 적절하지 않아 활용도가 매우 낮음
  - 정부주도의 AI학습용 데이터 구축 노력으로 인해 다양한 산업분야에서 적용할 수 있는 데이터가 구축·개방되었으나 방송영상제작에 사용할 수 있는 영상이미지 데이터세트는 현저하게 부족한 상황임

- (라벨링 데이터 부족) AI학습을 위해서는 이미지 또는 영상 데이터뿐만 아니라 해당 영상내의 객체 인식에 필요한 메타데이터가 필요함에도 불구하고, 양질의 데이터가 부족함
  - 이미지/영상에 라벨링을 하는 작업이 비용이 많이 소요되기 때문에, 민간기업이 구축한 메타데이터 자산을 제3자에게 무료로 공개하기는 힘들 것으로 판단됨
  - 민간에서 구축된 데이터의 경우, 카카오브레인이 공개한 ‘고요’데이터를 제외하고는 AI학습에 적합한 메타데이터가 포함된 민간기업의 데이터 공유사례가 흔하지 않음
  - 최근 정부주도의 AI학습용 데이터 구축·개방 지원 사업으로 인해 메타데이터가 포함된 학습용 데이터가 구축되고 있기는 하나, 방송영상제작과정에 사용이 가능한 데이터세트가 부족하고 대부분 언어중심의 데이터세트로 영상과 이미지 라벨링은 미흡한 상황임
- (민간의 낮은 공유·개방 의지) 데이터가 경쟁의 원천이 되는 생성형AI시대에 민간기업이 자발적으로 데이터를 제3자에게 공개할만한 동인이 크지 않음
  - 민간기업들이 이미지, 영상, 음성, 텍스트 등의 원본 데이터를 확보하는 과정에서는 비용이 소요될 뿐만 아니라, 저작권과 초상권 등의 법률적 문제도 동반됨
  - 따라서 민간기업들이 보유한 원본데이터를 제3자의 사용을 위해 공개할만한 유인이 크지 않음
  - 일부 통신사나 인터넷기업들이 원본데이터를 공개하더라도 민간사업의 특성상 데이터 분석결과를 제공하고 그에 따른 용역비 및 수수료 등을 부과하는 방식의 수익화 목적을 가지고 있음
  - 또한, 원천데이터 자체를 제공하기보다는 API만을 공개하여 데이터를 크롤링할 수 있는 횟수와 양을 한정하는 방식으로 서비스를 구축하고 있으며, 정형화된 빅데이터 중심으로 데이터를 구축·개방하는 경향이 있음

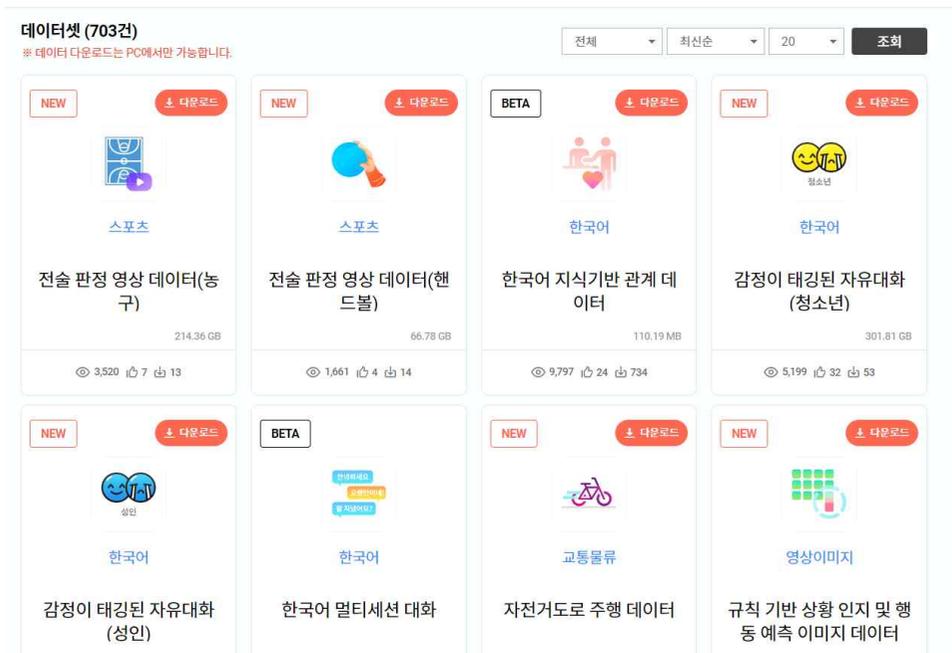
- 그 수익구조와 사업특성을 고려할 때, 민간주도로 공개되는 데이터는 그 수와 질이 담보되기 어렵다는 한계가 있어 AI학습용으로 데이터를 구하고자 하는 소규모 기업이나 개인개발자에게는 활용가능성이 현저하게 낮음

## 제4장 영상미디어분야 AI경쟁력 강화를 위한 학습데이터 수집·활용 활성화 방안

### 제1절 AI학습용 데이터구축 및 활용을 위한 제도개선 검토

- (개요) 우리나라는 AI 학습용 데이터 수집·개방을 목표로 한 ‘데이터 댐’ 프로젝트를 기반으로 구축된 ‘AI 허브’를 운영하고 있는데, 2024년 1월 현재, 총 703건<sup>14)</sup>의 데이터세트가 공개되어 있음

[그림 3-3] AI 허브 데이터세트 예시



출처: AI허브(n.d.)

- AI 학습용 데이터 구축 사업은 2017년부터 시작되었고, 2020년 ‘데이터 댐’ 프로젝트는 거쳐 AI 허브를 운영하고 있으나, 양질의 데이터 확보에 대한 수요에 비해 구축 작업이 빠르게 이루어지고 있다고 보기는

14) <https://www.aihub.or.kr/aihubdata/data/list.do?currMenu=115&topMenu=100>(최종방문확인, 2024.1.12.)

## 어려운 상황

- AI 데이터 수집 및 활용은 개인정보보호 관련 법규와 상충하고 있는데, 국내 개인정보보호 법제는 규제 수준이 높고 사전동의 등 형식적 규제 중심으로 구성되어 있어 높은 수준의 수집과 활용이 어려운 상황
- 국내 법률의 엄격한 적용과 해석에 따라 AI 데이터 수집 및 활용이 원활하지 못하고, 소극적인 형태로 이루어지기 때문에 AI 학습용 데이터 구축 및 활용이 활성화되지 못하고 있다는 의견이 지속적으로 제기
- 이에 따라, 2020년 1월 데이터 3법\*이 통과되며 국내 데이터 활용에 대한 기대가 일었으나, 기존 규제에 비해 진일보 했을 뿐, 여전히 데이터 활성화를 위한 포석으로는 미진하다는 평가

\* 개인정보보호법, 정보통신망 이용촉진 및 정보보호등에 관한 법률(정보통신망법), 신용정보의 이용 및 보호에 관한 법률(신용정보법) 등의 3가지 법률을 통칭

o (법률현황) 우리나라는 데이터 3법의 규정을 통해 과학적 연구를 목적으로 하는 경우, 정보주체의 동의없이 가명정보를 이용·제공하는 것을 허용하고 있음

- 개인정보보호법과 신용정보법은 ‘가명정보’, ‘가명처리’조항을 마련하여, 통계작성, 과학적 연구, 공익적 기록보존 등의 경우에는 정보주체의 동의 없이 가명정보를 처리하는 것이 가능하다고 규정하고 있음

## 1. AI 학습용 데이터 수집·활용에 관한 제반 법률 현황

### 가. 개인정보보호법

o (개요) 가명정보를 정의하고, 통계작성, 과학적 연구, 공익적 기록보존 등을 위해서는 정보주체의 동의 없이 가명정보를 처리하는 것이 가능하도록 규정하고 있음

- 가명정보란, 개인을 알아볼 수 있는 정보, 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보를 일부 삭제하거나 일부 또는 전부를 대체하는 등의 방법(가명처리)으로 원래 상태로 복원하기 위한 추가정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보를 의미(제2조제1호, 제1의2)

<표 3-8> 개인정보보호법 상 가명정보·가명처리 관련 규정

**제2조(정의)** 이 법에서 사용하는 용어의 뜻은 다음과 같다.

1. “개인정보”란 살아 있는 개인에 관한 정보로서 다음 각 목의 어느 하나에 해당하는 정보를 말한다.
  - 다. 가목 또는 나목을 제1호의2에 따라 **가명처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보**(이하 “가명정보”라 한다)
- 1의2. “가명처리”란 **개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말한다.**

○ (데이터 활용 허용 규정) 개인정보처리자는 통계작성, 과학적연구, 공익적 기록보존 등을 위해 정보주체의 동의없이 가명정보를 처리할 수 있음 (제28조의2제1항)

- 통계작성을 위한 가명정보처리란 통계를 작성하기 위해 가명정보를 이용, 분석, 제공하는 등 가명정보를 처리하는 것을 말하며, 가명정보의 처리 목적이 시장조사를 위한 통계 등 상업적 성격을 가진 통계를 작성하기 위한 경우에도 가명정보를 처리하는 것이 가능(가명정보 처리 가이드라인)
- 과학적 연구를 위한 가명정보의 처리란 과학적 연구를 위해 가명정보를 이용, 분석, 제공하는 등 가명정보를 처리하는 것을 말하며, 공적 자금으로 수행하는 연구뿐만 아니라 민간으로부터 투자를 받아 수행하는 연구에서도 가명정보 처리가 가능(가명정보처리 가이드라인)
- 즉, 개인정보 중 가명처리를 한 가명정보는 과학적 연구의 목적으로 사용될 수 있는데, AI 학습용 데이터는 과학적 연구 중 응용연구 및 민간

투자 연구에 포함하여 사용할 수 있을 것으로 판단

<표 3-9> 데이터 활용 관련 규정

<개인정보보호법>

제28조의2(가명정보의 처리 등) ① 개인정보처리자는 **통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다.**

- (안정성 확보조치) 개인정보처리자는 가명정보 또는 추가정보의 안전한 관리를 위해 내부 관리계획을 수립하고 수탁자 관리 및 감독 등 관리적 보호조치를 하도록 규정(제28조의4)
  - 개인정보처리자는 가명정보 및 추가정보를 안전하게 관리하기 위한 내부 관리계획을 수립·시행하여야 하며, 내부 관리계획에는 추가정보의 별도 분리 보관 및 이에 대한 접근권한 분리에 대한 사항 등을 포함하고 있어야 함
  - 개인정보처리자는 가명정보 처리업무를 외부에 위탁하는 경우 가명정보도 개인정보에 해당하므로 위탁업무 수행 목적 외 가명정보의 처리 금지에 관한 사항 등을 포함한 문서를 작성해야 함
  - 개인정보처리자는 가명정보 처리와 관련하여 가명정보 처리 목적, 처리 기간, 제3자 제공에 관한 사항, 처리의 위탁에 관한 사항, 처리하는 개인정보의 항목 등의 내용을 개인정보 처리방침에 포함하여 공개해야 함

<표 3-10> 안정성 확보조치 관련 규정

**제28조의4(가명정보에 대한 안전조치의무 등) ①** 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관·관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 대통령령으로 정하는 바에 따라 **안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다.**

② 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 경우 처리목적 등을 고려하여 **가명정보의 처리 기간을 별도로 정할 수 있다.**

③ 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하고

자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자, 가명정보의 처리 기간(제2항에 따라 처리 기간을 별도로 정한 경우에 한한다) 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 **관련 기록을 작성하여 보관하여야 하며, 가명정보를 파기한 경우에는 파기한 날부터 3년 이상 보관하여야 한다.**

○ (기술적 보호조치) 개인정보처리자는 가명정보 및 추가정보의 분리보관, 접근권한 관리, 접근통제 및 접속기록의 보관 및 점검 등 기술적 보호조치를 하도록 규정

- 개인정보처리자는 추가정보를 가명정보와 분리하여 별도로 저장·관리하고 추가정보가 가명정보와 불법적으로 결합되어 재식별에 악용되지 않도록 접근권한을 최소화하고 접근통제를 강화하는 등 필요한 조치를 적용해야 함

- 또한, 접근권한을 분리하기 위해 개인정보처리자는 가명정보 또는 추가정보에 접근할 수 있는 담당자를 가명정보 처리 업무 목적달성에 필요한 최소한의 인원으로 엄격하게 통제하여야 하며, 접근권한도 업무에 따라 차등부여 하여야 함

- 개인정보처리자는 가명정보의 처리목적, 가명처리한 개인정보 항목, 가명정보의 이용내역 등을 작성하여 보관하여야 함

<표 3-11> 안전조치의무 관련 규정

**제29조(안전조치의무)** 개인정보처리자는 개인정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 내부 관리계획 수립, 접속기록 보관 등 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다.

○ (물리적 보호조치) 개인정보처리자는 가명정보 또는 추가정보의 안전한 관리를 위해 물리적 단전조치를 취해야 함

- 개인정보처리자는 가명정보 또는 추가정보를 전산실이나 자료보관실에 보관하는 경우 비인가자의 접근으로부터 보호하기 위해 출입 통제 절

차를 수립하여야 하며, 보조저장매체 등에 저장되어 있는 경우 잠금장치가 있는 안전한 장소에 보관하는 등 보안대책을 마련해야 함

○ (정보주체의 권리보장) 개인정보처리자는 정보주체가 자신의 개인정보에 대한 가명처리 정지를 요구하는 경우 이를 보장해야 함

- 개인정보처리자는 정보주체의 가명처리 정지를 요구 받았을 때에는 지체없이 해당 정보주체의 개인정보 처리의 전부 또는 일부를 정지해야 함

○ (가명처리 관련 의무 및 처벌 규정) 전문기관으로 지정받지 않고 가명정보를 결합하거나 결합정보를 반출 또는 제3자에게 제공한 경우, 영리 또는 부정한 목적으로 결합정보를 제공받은 경우, 특정 개인을 알아보기 위해 가명정보를 처리한 경우 등은 5년이하의 징역 또는 5천만원 이하의 벌금을 부과할 수 있음

- AI 기술이 발달할수록 가명정보 활용을 위한 처리 과정에서의 재식별 가능성이 증가하기 때문에 이와 관련하여 기업들이 부담해야 하는 법적 리스크 또한 증가할 수 밖에 없음

○ (비정형데이터 가명처리를 통한 활용 기준 마련\_예정) 비정형데이터를 가명처리하여 AI학습, 자율주행 기술개발 등에 활용할 수 있는 기준 마련을 준비중(가명정보 활용 확대방안, 2023)

- (영상정보 특례) 가명처리 시 AI학습 등에 활용이 곤란한 경우에 한해 영상정보를 일정한 안전조치\* 하에 활용할 수 있는 기준 정립('23. 하반기 예정)

\* 개인식별과 관련 없는 연구목적, 인적개입 원천 차단, 주기적 점검 및 안정성 평가 등

- 자율주행차 영상정보 수집·활용에 대한 규제샌드박스 등을 통해, 원본 데이터를 활용한 AI 학습효과에 대한 실증연구를 병행

- 이에 따라 해당 기준이 마련된다면, AI 학습데이터의 수집·활용이 용이해지고, 기업들의 이용 범위도 넓어져 AI 산업의 발전 가능성을 확대할 수 있음

## 나. 신용정보법

- (개요) 기업 및 법인에 관한 정보를 제외한 살아 있는 개인에 관한 신용정보로서 특정 개인을 알아볼 수 있는 정보, 다른 정보와 쉽게 결합하여 특정 개인을 알아볼 수 있는 정보를 개인신용정보로 정의(신용정보법 제2조 제2호)
- 신용정보법 상 가명정보는 가명처리한 개인신용정보로 정의하고 있음(신용정보법 제2조제16호)

### <표 3-12> 신용정보법 상 가명정보·가명처리 관련 규정

<p><b>제2조(정의)</b> 이 법에서 사용하는 용어의 뜻은 다음과 같다.</p> <p>15. “가명처리”란 <b>추가정보를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리(그 처리 결과가 다음 각 목의 어느 하나에 해당하는 경우로서 제40조의2제1항 및 제2항에 따라 그 추가정보를 분리하여 보관하는 등 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리한 경우를 포함한다)</b>하는 것을 말한다.</p> <p>가. 어떤 신용정보주체와 다른 신용정보주체가 구별되는 경우</p> <p>나. 하나의 정보집합물(정보를 체계적으로 관리하거나 처리할 목적으로 일정한 규칙에 따라 구성되거나 배열된 둘 이상의 정보들을 말한다. 이하 같다)에서나 서로 다른 둘 이상의 정보집합물 간에서 어떤 신용정보주체에 관한 둘 이상의 정보가 연계되거나 연동되는 경우</p> <p>다. 가목 및 나목과 유사한 경우로서 대통령령으로 정하는 경우</p> <p>16. “가명정보”란 <b>가명처리한 개인신용정보</b>를 말한다.</p>
--

- (데이터 활용 허용 규정) 통계작성, 연구, 공익적 기록보존 등을 위하여 가명정보를 제공하는 경우에는 개인신용정보의 제공 및 활용에 대한 동의 규정의 적용을 받지 않음(신용정보법 제32조제5항)
- 이 경우 통계작성에는 시장조사 등 상업적 목적의 통계작성을 포함하며, 연구에는 대학, 연구소 등 연구기관 뿐만 아니라 기업 등이 수행하는 산업적 연구를 포함(금융분야 가명익명처리 안내서)

<표 3-13> 신용정보법 상 개인신용정보의 활용 관련 규정

**제32조(개인신용정보의 제공·활용에 대한 동의)** ⑤ 신용정보회사등은 신용정보주체가 선택적 동의사항에 동의하지 아니한다는 이유로 신용정보주체에게 서비스의 제공을 거부하여서는 아니 된다.

- (가명처리 관련 처벌 규정) 신용정보회사등이 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보처리를 한 경우 ‘전체매출액’의 3% 이하에 해당하는 금액을 과징금으로 부과할 수 있음(신용정보법 제42조의2제1항제1호의4)
- 또한, 영리 또는 부정한 목적으로 특정한 개인을 알아볼 수 있게 가명정보를 처리한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금을 부과할 수 있음(신용정보법 제50조제2항제7의2)

<표 3-14> 가명처리 관련 처벌 규정

**제42조의2(과징금의 부과 등)** ① 금융위원회(제45조의3제1항에 따른 상거래기업 및 법인이 다음 각 호의 어느 하나에 해당하는 행위를 한 경우에는 보호위원회를 말한다)는 다음 각 호의 어느 하나에 해당하는 행위가 있는 경우에는 **전체 매출액의 100분의 3 이하에** 해당하는 금액을 과징금으로 부과할 수 있다. 다만, 제1호에 해당하는 행위가 있는 경우에는 **50억원 이하의 과징금을** 부과할 수 있다.

1의4. 제40조의2제6항을 위반하여 **영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리한 경우**

**제50조(벌칙)** ② 다음 각 호의 어느 하나에 해당하는 자는 **5년 이하의 징역 또는 5천만원 이하의 벌금**에 처한다.

7의2. 제40조의2제6항을 위반하여 **영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리한 자**

## 다. 저작권법

- (개요) 저작권법은 AI가 학습을 목적으로 저작물을 이용하는 경우에 대해 개별적인 저작재산권 제한 사유\*로 규정하고 있지 않으며, 이에 따라 저작재산권이 제한되는 사유를 포괄적으로 정하고 있는 공정이용 규정(제35조의5)의 적용여부가 주요 쟁점으로 논의

\* 재판 등에서의 복제, 정치적 연설 등의 이용, 공공저작물의 자유이용, 학교교육 목적 등에서의 이용, 시사보도를 위한 이용, 시사적인 기사 및 논설의 복제 등, 공표된 저작물의 인용, 영리를 목적으로 하지 아니하는 공연·방송 등

- (공정이용 규정) 저작권자의 허락을 구하지 않아도 저작물을 이용할 수 있는 저작권에 대한 특수한 상황을 의미하는데, 저작권법은 공정 이용할 수 있는 경우를 법으로 규정\*하고 있음

\* 저작권법 ‘제23조부터 제35조의4까지’, ‘제101조의3부터 제101조의5까지의 경우’

- 이 외에도 일반적인 이용방법과 충돌하지 않고, 공정하게 사용되었다면 저작물을 이용할 수 있다고 규정하고 있는데, ①이용의 목적 및 성격, ②저작물의 종류 및 용도, ③이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성, ④저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향 등을 고려하여 판단할 수 있음

### <표 3-15> 공정한 저작물 이용 관련 규정

<p><b>제35조의5(저작물의 공정한 이용)</b> ① 제23조부터 제35조의4까지, 제101조의3부터 제101조의5까지의 경우 외에 <b>저작물의 일반적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있다.</b></p> <p>② 저작물 이용 행위가 제1항에 해당하는지를 판단할 때에는 <b>다음 각 호의 사항 등을 고려하여야 한다.</b></p> <p>1. 이용의 목적 및 성격</p>
---

- 2. 저작물의 종류 및 용도
- 3. 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성
- 4. 저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향

- (데이터 활용) 인터넷 상에 게시되어 있는 데이터는 누구나 접근할 수 있으며, 이러한 공개 저작물에 대해서는 사업자들이 AI 학습에 이용하는 경우가 많음
  - 저작권자가 이를 원하지 않을 경우에는 반대하는 의사를 정확하게 명시하거나 이를 방지하기 위한 기술적 조치를 취하는 것이 적절함
  - 이미 인터넷 상에 공개적으로 게시되어 있는 저작물이기 때문에 이와 같은 조치를 취하지 않을 경우에는 분쟁 발생 시 이용허락 등에 대한 쟁점이 문제가 될 수 있음
- (AI 저작권 면책 기준 마련 예정) 저작권법 개정을 통해 AI 학습을 위한 데이터 활용에 대해서는 저작권 침해에 대한 면적요건과 그 근거를 마련 (관계부처 합동, 2023)
  - 이미, EU와 일본 등 해외주요국은 AI 학습용 데이터 활용에 있어서는 저작권이 면책되는 규정을 도입하여 운영하고 있음
  - AI를 학습하기 위한 ‘크롤링(crawling)’은 저작물에 포함된 사상 및 감정을 공유하지 않고, 법률에 저촉되지 않는 적법한 저작물 접근에 해당됨을 명시함

## 라. 데이터 산업법

- (개요) 데이터의 중요성이 부각되는 상황에서 경제·사회 전반에서 창출되고 있는 데이터를 수집·가공·생산·활용함에 있어 이를 체계적으로 추진하기 위한 법적 근거로서 작용하며, 데이터를 경제적 가치로서 해석하고 있음

- 민간 데이터의 가치와 중요성을 재확인하고 기업들의 불확실성을 제거하며, 우리 데이터 산업에 대한 육성 의지를 표명하고 있으며, 데이터의 경제 생태계 구축을 위한 기반을 마련하고자 함
- o (데이터 활용) 데이터 생산, 거래, 활용을 촉진하기 위해 데이터를 정보처리장치가 처리할 수 있는 형태로 본인 또는 제3자에게 원활하게 이동시킬 수 있도록 제도적인 기반을 구축해야 함
- 데이터 이용을 활성화 할 수 있는 기반을 조성하기 위해 가치평가를 지원하고 데이터 이동 촉진, 공정한 유통환경 조성 등 데이터 이용 활성화 기반을 마련

<표 3-16> 데이터 이용 활성화 및 유통 관련 규정

<p><b>제15조(데이터 이동의 촉진)</b> 정부는 데이터의 생산, 거래 및 활용 촉진을 위하여 데이터를 컴퓨터 등 정보처리장치가 처리할 수 있는 형태로 <b>본인 또는 제3자에게 원활하게 이동시킬 수 있는 제도적 기반을 구축하도록 노력</b>하여야 한다.</p> <p><b>제17조(공정한 유통환경 조성 등)</b> ① 과학기술정보통신부장관은 데이터를 거래함에 있어서 대기업과 중소기업 간의 <b>공정한 경쟁 환경을 조성하고 상호 협력을 촉진</b>하여야 한다.</p> <p><b>제18조(데이터 유통 및 거래 체계 구축)</b> ① 과학기술정보통신부장관은 데이터 유통 및 거래를 활성화하기 위하여 <b>데이터 유통 및 거래 체계를 구축</b>하고, 데이터 유통 및 거래 기반 조성을 위하여 필요한 지원을 할 수 있다.</p> <p>② 과학기술정보통신부장관은 데이터 유통과 거래를 촉진하기 위하여 데이터유통 시스템을 구축·운영할 수 있다.</p> <p>③ 제1항에 따른 데이터 유통 및 거래 기반 조성 지원을 위하여 필요한 방법 및 기준과 제2항에 따른 데이터유통시스템의 운영 등에 필요한 사항은 대통령령으로 정한다.</p>
--

## 2. AI 학습용 데이터 수집·활용 확대를 위한 법제도 정비 방향

### 가. AI 학습용 데이터 수집·활용의 법률적 제약요인

- o (규제의 경직·엄격성) 국내 법률은 데이터 수집 및 활용에 대해 가명처리를 전제하고 있으며, 재식별 시 형사처벌이 엄격하여 학습데이터 활용에 한계가 존재할 수밖에 없음

- 데이터 수집 및 활용에 있어 반드시 가명처리가 전제되어야 하므로, 개인을 식별할 수 없는 수준의 가명처리를 해야 하며, 이를 보장하기 위해 안전성 확보조치 및 기술적 보호조치 등의 일정한 조치를 할 의무가 있을 뿐만 아니라, 이를 검증하는 과정에서의 시간 및 비용 소모 등이 제약요인으로 작용하고 있음
- o (형식적 사전동의) 국내 개인정보보호법에 규정된 동의제도는 개인정보 자기결정권을 집중적으로 보호하기 위해 ‘정보주체의 동의’ 부분에 초점을 두어 마련되었기 때문에, 데이터를 이용하기 위해서는 사전에 필수적으로 정보주체의 동의를 구하여야 함
  - 서비스 이용자들은 서비스를 편하게 이용하기 위해 자신의 개인정보를 제공하는 것에 쉽게 동의하여 정보를 제공하는 편인데, 이러한 사전동의(opt-in) 규제는 4차산업혁명 시대에 개인정보보호와 데이터 활용 양 측면 모두에 적절하지 않은 규제 방식으로 인식되고 있음
  - 정보주체 조차도 자신의 정보를 제공함에 있어 큰 고민없이 정보를 제공하고, 이를 수집하는 기업으로서도 과도하게 정보를 수집하는 루트로 사용될 가능성이 커 법률에 정해진 형식적인 절차로서의 의미만 존재
- o (데이터 산업법과의 부조화) 데이터 산업법은 데이터 산업 육성 전체를 아우르고 있으며, 데이터 활용을 촉진하는 것을 목적으로 마련되었는데, 이러한 법률의 목적과 현행 법률은 일정 부분 부조화를 이루는 양상이 있음
  - 4차산업 혁명시대에 데이터 이용과 개인정보의 활용은 필수불가결한 관계에 놓여있는데, 개인정보의 엄격한 통제로 인해 데이터 생태계 구축 및 산업 발전이 지체되는 한편, 형식적인 보호절차로 인해 개인의 정보도 실질적으로 보호되지 못하는 상황에 놓여있음

## 나. AI 학습용 데이터 수집·활용을 위한 제도 정비 방향

- (규제 패러다임 전환) 데이터 수집 및 활용과 관련하여 기존의 경직된 법 규제체계의 한계를 보완하여 AI 활성화, 빅데이터 등 기술의 발전을 전제로 한 규제로의 전환이 필요한 시점
  - 즉, 개인의 정보를 실질적으로 보호하면서도 데이터 수집 및 활용도를 높일 수 있는 방향으로 규제 방향을 재설정할 필요가 있음
  - 개인의 정보라 하더라도 식별가능성이 현저하게 떨어지는 정보 등 필수적인 정보가 아니라면 굳이 가명처리를 하지 않더라도 기업이 데이터를 활용할 수 있도록 규제 방향을 전환하며, 이 경우 기업이 데이터를 활용함에 있어서는 투명성을 담보할 수 있도록 절차를 마련하는 등 규제를 개선하는 것이 필요
  - 이를 통해, 개인은 식별가능성이 떨어지는 정보에 대해서까지 형식적인 동의 절차를 거칠 필요가 없고, 기업은 고품질의 데이터를 사용하여 서비스를 마련할 수 있어 유의미한 규제 효과를 보장할 수 있음
  - 특히 AI 학습용 데이터의 경우, 개인의 정보를 식별하기 위한 목적이 아니라, AI를 ‘학습’ 시켜 서비스의 품질을 향상시키고, 고품질의 정보를 이용자에게 제공하기 위한 목적이 명확하기 때문에 가명처리를 하지 않은 이용자의 식별가능성이 현저히 떨어지는 정보를 학습하는 방안을 적극적으로 고려할 필요가 있음
- (형식적 사전동의 규제 개선) 사전적으로 동의제도를 통제하는 방법을 완화하여 데이터를 활성화할 수 있는 실질적인 방안의 개인정보를 보호 수단 마련이 필요
  - 개인의 식별가능성이 현저히 떨어지는 정보 또는 필수동의 사항에 대해서는 ‘옵트아웃’ 방식을 도입하는 등 동의제도에 있어 형식적인 부분을 폐지하고, 실질적으로 개인정보를 강화하는 방향으로 규제를 개선할 필요가 있음

- AI가 학습하는 데 필요한 데이터의 경우 식별가능성이 없는 정보, 필수적이지 않은 정보 등 일정한 정보에 한해 정보주체의 사전동의를 받지 않고도 사용할 수 있도록 사전동의 규제를 개선할 필요가 있음
- o (가명처리 없는 AI 학습데이터 활용) 개별 정보주체에 대한 민감하거나 중요한 정보나, 당사자의 권리의무에 영향을 미치는 정보가 아닌 경우, 즉, 단순히 AI 학습을 위한 경우에는 법률에 규정된 접근통제, 기술적·관리적 보호조치를 전제로 하여 가명처리 없이 개인정보를 활용할 수 있는 방안을 고려
- 앞서 논의한 저작권법에서는 단순히 디지털 저작물을 AI의 학습을 위해 활용할 경우에는 이러한 방향으로 정책 개선 방안을 도모하고 있음
- 다만, AI 학습데이터로서 정보를 활용하더라도 그 과정에서의 윤리적인 통제 부분은 준수되어야 함

## 제2절 영상미디어분야 AI학습 데이터 구축 사업방안

### 1. 인공지능 시대 국내 영상미디어분야 경쟁력 저해요인 분석

#### 가. 협소한 내수시장의 한계

- (투자 대비 수익성 제고의 어려움) 생성형 AI를 영상미디어분야에 접목시키기 위해서는 많은 투자가 필요한데 협소한 내수시장만을 고려해서는 투자가 어려운 측면이 존재
- 한정된 국내시장에서 글로벌 사업자와의 경쟁이 심화되고 있어 국내 영상미디어 분야 사업자들이 생성형AI 등 신규 영역에 투자할 여력을 확보하기 어려운 상황

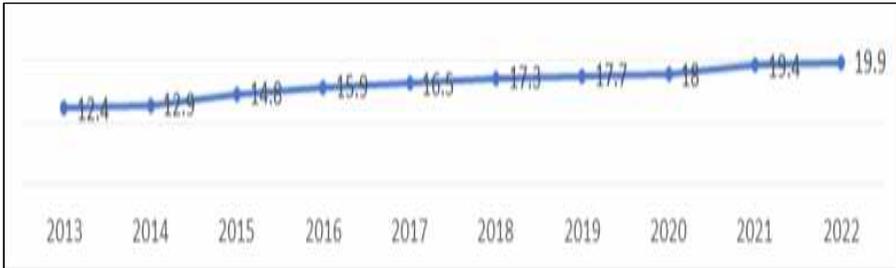
[그림 4-1] 국내 미디어 산업의 약점



출처: 노창희(2023)

- 국내 방송산업은 성장이 정체된 상황이며, 전체 매출액 규모는 20조원이 안되는 수준
- 한편으로는 한계에 직면한 현재의 상황을 극복하기 위한 과감한 투자가 필요하나 현재 국내 영상미디어 분야 사업자들이 놓여있는 재정적인 한계를 고려할 때 과감한 투자는 어려울 것이라 판단됨

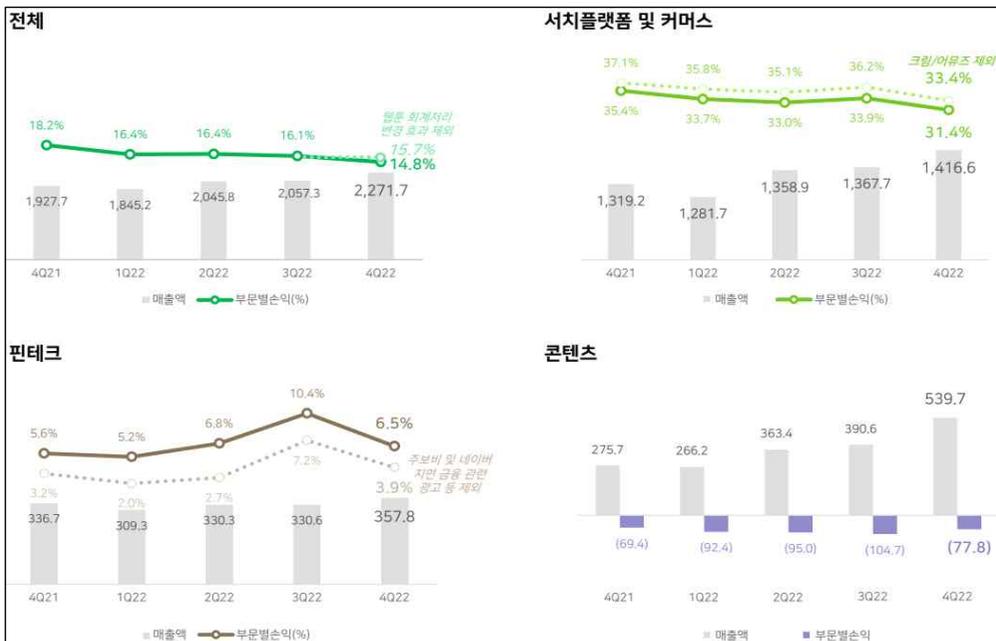
[그림 4-2] 국내 방송매출액 추이



자료: 방송통신위원회(2023), 단위: 조원

- 또한, 국내의 경우 국내 영상미디어 사업자의 글로벌 진출이 넷플릭스를 통해 주로 이뤄지고 있는 상황에서 협소한 내수시장에서의 사업을 위해 생성형 AI에 적극적으로 투자하는 것에 한계가 존재하는 상황

[그림 4-3] 네이버 2022년 사업부문별 손익



자료: 네이버(2023), 단위: 십억원, %

- 국내에서 생성형 AI를 주도하고 있는 네이버의 경우 단일 사업자가 ‘서치플랫폼 및 커머스’, ‘핀테크’, ‘콘텐츠’ 등에서 연단위 기준으로 8천억 수준의 매출을 올리고 있으며,

[그림 4-4] 네이버의 주요 핀테크 사업영역



자료: 네이버(2023)

- 인공지능을 활용할 수 있는 분야도 핀테크에 국한해도 ‘페이 서비스’, ‘디지털 금융서비스’, ‘금융 관련 광고’, ‘금융 콘텐츠’ 등으로 다양함

<표 4-1> 워크플로우별 AI 접목과 디지털 전환 관련 정책 과제

①	<b>기획</b>	생성형 AI를 통한 미디어·콘텐츠 창작
②	<b>제작</b>	초실감 가상제작 (Virtual Production) 이용 활성화
③	<b>제작</b>	<b>후반제작</b> 디지털 휴먼 및 AI 기반 제작·편집 지원
④	<b>후반제작</b>	<b>해외진출</b> : AI 활용, 번역·자막·더빙
⑤	<b>마케팅·유통</b>	이용 데이터 확보 및 영상 업스케일링
⑥	<b>마케팅·유통</b>	<b>누누TV 사례</b> : 콘텐츠 불법유통 대응 자동화

출처: 과학기술정보통신부(2023. 9. 12)

- 영상미디어 분야에서도 기획, 프리 프로덕션, 포스트 프로덕션 등 다양한 분야에 활용이 가능하지만 아직까지 수익성이 담보되는 영역을 발굴했다고 보기는 어려운 상황
- o (협소한 국내시장의 한계를 극복할 수 있는 모멘텀 마련 필요) 생성형 AI와 영상미디어 분야의 접목을 통한 시너지를 창출하기 위해서는 협소

한 국내시장의 한계를 넘어설 수 있는 모멘텀 마련 필요

- 가령, 국내 OTT 사업자가 글로벌 진출이 활발해질 경우 AI를 활용한 번역, 자막, 더빙 등 재제작 산업이 발전할 수 있는 계기가 마련될 수 있음
- 즉, 국내 OTT 사업자의 글로벌 시장 진출이 재제작 관련 AI 기술에 투자할 수 있는 계기를 마련해 줄 수 있다는 것임

#### 나. 데이터 확보의 한계

- o (영상 데이터 확보의 어려움) 영상 관련 데이터는 저작권이 복합적으로 형성되어 있고, 국내의 경우 공공 차원에서 영상 데이터를 보존 및 활용하는 제도가 미흡하여 영상 데이터는 다른 데이터보다 확보하기가 더욱 어려운 상황
  - 방송 콘텐츠, 영화 등은 저작권의 소재가 공동이기 때문에 인공지능 학습을 위해 활용하기 위해서는 다양한 주체의 동의가 필요하고, IP를 가진 방송사나 제작사가 아카이브 활용에 동의한다고 하더라도 추후 문제가 발생할 수 있음
  - 또한, 국내는 영상 아카이브 기록, 관리에 대한 제도가 유럽 등에 비해 발달되어 있지 않고, 공영방송도 자사 콘텐츠에 대한 저작권을 가질 수 있어 공적으로 활용할 수 있는 데이터를 확보하는 것이 어려운 실정
- o (데이터수집 및 공개에 대한 법률적 제약) 관련 부처는 데이터 3법 등 데이터 관련 규제를 완화하여 데이터 기반 환경, 인공지능 기반 환경에 대응하기 위해 다양한 정책적 노력을 지속해 왔으나 여전히 법률적 제약이 존재하는 상황
  - 데이터 활용의 중요성 증대됨에 따라 국내에서는 2020년에 데이터 3법을 도입하는 등 데이터 활용 촉진을 위한 정책 노력을 기울여 왔으나 여전히 데이터를 활용하기 위해서는 많은 법적인 제약이 따르는 상황

<표 4-2> 개인정보의 보호와 활용 관련 규제정책 변동 내용

날짜	주요 내용	세부 내용
1995.1	신용정보법 제정	개인신용정보의 이용, 처리, 활용에 관한 사항 규율
1999. 2	정보통신망법 제정	정보통신서비스 제공자의 개인정보 처리에 관해 규율
2001. 12	신용정보법 상 동의 강화	개인식별정보 제공시 정보주체로부터 서면 동의 받도록 함
2009. 4	신용정보회사의 업무 확대 정보주체의 자기정보 통제권 강화	신용정보회사의 겸업 허용, 신용평가회사의 평가 대상 확대 신용정보관리 보호인 제도 도입
2011. 3	개인정보 보호법 제정	개인정보 보호, 처리 등에 대한 통합적 규율
2013. 8	주민등록번호 처리 규제 강화	원칙적으로 주민등록번호 처리 금지하고, 주민등록번호 분실, 도난, 유출, 훼손 시 제재 강화
2014. 3	주민등록번호 암호화 의무	주민등록번호를 보관하는 개인정보처리자는 주민등록번호를 암호화할 의무
2015. 3	개인신용정보 유출 예방, 제재 강화	신용정보회사의 부수업무 제한, 개인신용정보의 제공, 활용에 대한 동의 강화, 정보 유출 시 과징금, 손해배상 부과, 형벌 상향
2015. 7	개인정보보호 위원회 기능 강화	개인정보보호위원회 조정기능 강화 피해구제 강화, 제재 강화
2017. 4	개인정보 처리시 안전성 확보	주민등록번호 수집 근거법을 법률, 대통령령, 국회규칙으로 한정
2020. 1	데이터 3법 개정	가명정보 도입, 개인정보보호위원회 격상 및 권한 강화, 개인사업자 CB도입, 본인신용정보관리업(마이데이터) 도입

출처: 국가법령정보센터 내용을 편집한 임효진·박형준(2022) 재인용

<표 4-3> 「개인정보 보호법」상 민감정보의 처리 제한과  
고유식별정보의 처리 제한 관련 규정

제23조(민감정보의 처리 제한) ①개인정보처리자는 사상·신념, 노동조합·정당의 가입·탈퇴, 정치적 견해, 건강, 성생활 등에 관한 정보, 그 밖에 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보로서 대통령령으로 정하는 정보(이하 “민감정보”라 한다)를 처리하여서는 아니 된다. 다만, 다음 각 호의 어느 하나에 해당하는 경우에는 그러하지 아니하다.

1. 정보주체에게 제15조제2항 각 호 또는 제17조제2항 각 호의 사항을 알리고 다른 개인정보의 처리에 대한 동의와 별도로 동의를 받은 경우
2. 법령에서 민감정보의 처리를 요구하거나 허용하는 경우

② 개인정보처리자가 제1항 각 호에 따라 민감정보를 처리하는 경우에는 그 민감정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 제29조에 따른 안전성 확보에 필요한 조치를 하여야 한다.

③ 개인정보처리자는 재화 또는 서비스를 제공하는 과정에서 공개되는 정보에 정보주체의 민감정보가 포함됨으로써 사생활 침해의 위험성이 있다고 판단하는 때에는 재화 또는 서비스의 제공 전에 민감정보의 공개 가능성 및 비공개를 선택하는 방법을 정보주체가 알아보기 쉽게 알려야 한다.

제24조(고유식별정보의 처리 제한) ① 개인정보처리자는 다음 각 호의 경우를 제외하고는 법령에 따라 개인을 고유하게 구별하기 위하여 부여된 식별정보로서 대통령령으로 정하는 정보(이하 “고유식별정보”라 한다)를 처리할 수 없다.

1. 정보주체에게 제15조제2항 각 호 또는 제17조제2항 각 호의 사항을 알리고 다른 개인정보의 처리에 대한 동의와 별도로 동의를 받은 경우

2. 법령에서 구체적으로 고유식별정보의 처리를 요구하거나 허용하는 경우

② 삭제

③ 개인정보처리자가 제1항 각 호에 따라 고유식별정보를 처리하는 경우에는 그 고유식별정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 대통령령으로 정하는 바에 따라 암호화 등 안전성 확보에 필요한 조치를 하여야 한다.

④ 보호위원회는 처리하는 개인정보의 종류·규모, 종업원 수 및 매출액 규모 등을 고려하여 대통령령으로 정하는 기준에 해당하는 개인정보처리자가 제3항에 따라 안전성 확보에 필요한 조치를 하였는지에 관하여 대통령령으로 정하는 바에 따라 정기적으로 조사하여야 한다.

⑤ 보호위원회는 대통령령으로 정하는 전문기관으로 하여금 제4항에 따른 조사를 수행하게 할 수 있다.

- 위와 같이 데이터 활용을 위한 다양한 법제도 개선을 추진해 왔으나 여전히 「개인정보 보호법」상 민감정보의 처리 제한과 고유식별정보의 처리 제한과 같은 제약 요인이 존재하는 상황

- 현재의 상황에서 생성형 AI 개발과 학습을 위한 데이터 확보를 위해서는 정보주체에게 명시적인 동의를 받아 수집하거나 가명 조치된 개인정보를 수집해야 하나 이러한 방식으로 데이터를 습득할 수 있는 주체는 거래관계가 형성된 기업이나 자본력이 충분한 기업만이 현실적으로 데이터를 확보할 수 있는 실정

- 대다수의 사업자가 영세한 규모인 영상미디어분야에서는 생성형 AI 기술 활용을 위한 데이터 확보가 어려우며 특히, 영상 관련 데이터 확보는 더욱 어려운 실정

#### 다. 인력 확보의 어려움

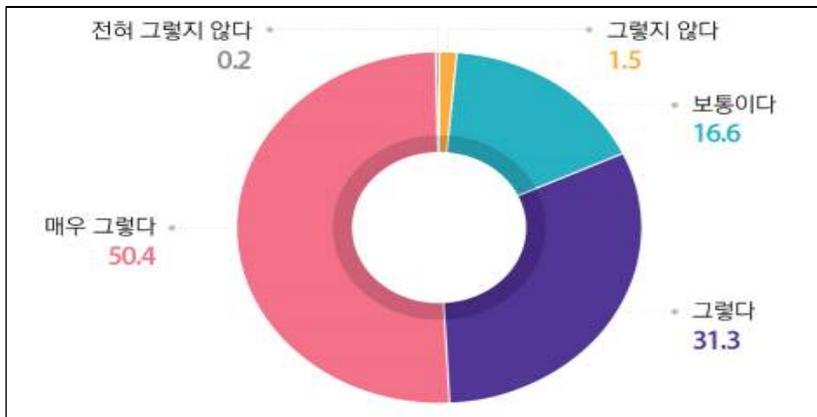
- o (인공지능 관련 인력의 경쟁력) 국내는 인공지능 관련 인력이 부족한 상황이며, 특히, R&D 등 고급인력이 부족하다고 평가받고 있음
- 고용노동부(2023. 8. 31)은 인공지능 분야 인력이 12,800명 부족하다고 평가하였으며, 특히 R&D 등 고급인력이 부족하다고 분석하고 있음

<표 4-4> 인공지능 분야 인력수급 전망결과(2023-2027년)

수요			공급				수급차			
			정부·민간		대학					
계	초·중급	고급	계	초·중급	고급	초·중급	고급	계	초·중급	고급
66,100	44,600	21,500	53,300	46,200	4,000	2,200	900	-12,800	3,800	-16,600

자료: 고용노동부(2023. 8. 31)

[그림 4-5] 인공지능 사업 운영상 느끼는 애로사항: 인력 부족



자료: 소프트웨어정책연구소(2023)

- 국내 인공지능 관련 사업자들이 인공지능 관련 사업을 영위하면서 느끼는 가장 큰 어려움 중 하나는 인력의 부족 때문으로 나타났음
- 이와 같은 기술인력 부족 문제를 해결하기 위해 과학기술정보통신부(2023. 9. 12)에서는 향후 3년간 약 1,500여명 규모의 미디어 분야 인공지능 전문 인력을 양성할 계획임

#### 라. 초거대AI 중심의 전략 추진시 지도학습용 데이터 수급 부족 우려

- o (인공지능 투자의 특징) 파운데이션 모델 중심으로 초거대AI플랫폼 경쟁력 확보에 집중된 정부 추진전략은 영상미디어 분야에 실질적인 AI수요와 괴리될 우려가 있음
  - 챗GPT 등장으로 인해 정부의 인공지능 전략이 급격히 생성형AI 중심으로 전환되고 있으며, 민간에서도 통신사와 대규모 인터넷기업을 중심으로 자체적인 한국형 대형언어모델(LLM) 개발 투자를 확대하고 있음
  - 국내에서는 관계부처 합동(2023. 4. 14)으로 발표한 ‘초거대AI 경쟁력 강화 방안’이 네이버와 같은 빅테크기업 위주로 만들어졌다는 비판이 제기된 바 있음(김동원, 2023. 4. 18)
  - 생성형 AI의 잠재적 가능성을 고려하면 국가를 선도할 수 있는 파운데이션 모델의 구축에 많은 투자가 필요한 것은 사실이나, 모든 AI개발 및 산업 지원역량이 파운데이션 모델 개발 자체에만 집중될 경우에는 이미 상용화되어 있는 모델을 적용한 서비스개발과 활용에 대한 관심과 지원이 부족해질 수 있다는 우려가 있음
  - 더욱이 파운데이션 모델 개발의 역량을 가진 대규모 통신사와 인터넷기업이 아닌 중소기업의 기업과 개발사 등의 경우에는 산업에 특화된 소형의 AI모델만으로도 충분할 수 있으므로, 파운데이션 모델에 대한 집중적인 투자 효과를 체감하기 어려울 가능성도 있음
  - 적정기술(appropriate technology)의 관점에서, 영상미디어 분야의 중소기업들을 위해서는 초거대AI모델 개발에 대한 지원 이외에도 산업 특화적

인 학습모델 개발과 성능개선을 위한 지원이 지속될 필요가 있음

[그림 4-6] ‘초거대 AI 경쟁력 강화 방안’의 비전 및 추진전략



출처: 관계부처 합동(2023. 4. 14)

- 비 라벨링 텍스트 데이터 위주의 데이터 구축사업은 이미 도입된 AI 모델을 활용하고자 하는 제작사, 영세한 개발사, 개인 개발자의 학습용 데이터 접근성을 더욱 저해할 우려가 있음

[그림 4-7] ‘초거대 AI 경쟁력 강화 방안’의 데이터 구축 방향

**1 | 인간의 초거대AI 개발을 지원하는 학습용 데이터 구축**



출처: 관계부처 합동(2023. 4. 14)

- (영상미디어 분야 지원 방향) 영상미디어 분야는 영세한 영역이기 때문에 초거대 AI 지원 보다는 필요한 영역에 필요한 지원이 이뤄지는 것이 필요
  - 영상미디어 분야 사업자에 대한 수요 조사를 통해 인공지능 지원을 실시하고 만족도 조사를 통한 환류체계 구축 필요
  - LLM모델이나 파운데이션모델의 잠재적인 가능성을 고려해 초거대AI 모델 개발의 지원이 필요한 것은 사실이나, 영상미디어 분야에서 LLM에 기반한 인공지능모델 개발에는 상당한 시일이 소요될 것으로 전망됨
  - 초거대AI모델이 텍스트 기반 데이터를 중심으로 발전하고 있어, AI학습데이터 구축도 텍스트 기반으로만 지원되고 있어 영상미디어 분야에서의 활용성이 중단기적으로는 제한될 수밖에 없음
  - 더욱이 자체적인 파운데이션 모델 개발 능력이 없는 영세한 영상미디어 기업들에게는 상용적으로 쓸 수 있는 지도학습기반 AI기술 활용과 개발에 대한 집중적인 지원이 필요함

**2. 영상미디어 분야 데이터 구축 및 활용 활성화 방안**

가. 공유-협력에 기반한 AI 학습데이터 구축 전략

- (데이터이용자-공공/정부-데이터보유자 간 협력체계) 데이터이용자와

데이터보유자 간 공유와 협력으로 영상미디어분야의 AI기술 수요에 부합하는 AI학습데이터 구축

- 영상미디어산업의 경우 가장 많은 양의 영상, 음성, 텍스트, 이용기록 등의 데이터를 보유한 보유자가 방송영상관련 미디어 기업과 플랫폼이기 때문에, 데이터보유자가 데이터이용자인 AI개발자가 양질의 AI기술을 개발한다면 그 AI기술의 최종 수혜자가 되는 순환적인 구조임
  - 그러나 현재의 데이터 거래 플랫폼은 단순히 데이터를 일회적으로 제공하는 거래만 지원하고 있어, 최초 데이터 제공자가 데이터 제공 이후에 데이터 활용의 결실을 확인하기 어려운 구조임
  - 이에 데이터보유자가 원천데이터의 개방에 대해서 불신을 가질 수밖에 없으며, 데이터의 개방이 데이터보유자의 이익에 배치된다는 인식을 하게 됨
  - 이러한 인식으로 인해 최근 한국방송협회는 회원사인 39개 지상파 방송사를 대표해 저작권자의 허락 없이 지상파 방송사 콘텐츠가 인공지능의 학습 자료로 쓰이는 일을 금한다는 의견서를 네이버, 카카오, 구글코리아, 한국마이크로소프트 등 국내외 빅테크 기업에 전달한 바 있음 (박강수, 2023. 12. 14.)
  - 한국방송협회는 방송사 저작물을 AI 학습데이터로 사용할 경우 사전에 이용 여부와 계획 확인, 이용 시 보상 협의 및 출처 공개 등을 요구하고 있음
  - 이에 데이터의 수집부터 최종 결과물인 AI기술의 활용에 이르기까지 합리적이고 공정한 기준으로 관리할 수 있도록 공공기관이 주도하는 데이터이용자와 데이터보유자 간 상생의 협력체계를 구축이 필요함
- (공공부문의 조정적 역할 확대) AI학습용 데이터의 구축·개방은 원본데이터에 대한 저작권과 초상권 보유자, 인공지능학습용 데이터 구축사, AI개발자(사) 간의 공정하고 합리적인 협상이 필요하며, 이를 위해서는 정부 또는 공공기관이 조정자이자 매개자로서의 주도적 역할이 필요함

- 영상·이미지·음성 등의 원천데이터를 보유한 방송사나 제작사 등의 원본데이터를 개방하는 과정에서 저작권·초상권 침해 우려가 있고 원본데이터 개방을 통해 얻을 수 있는 수익이나 기대효과가 명확하지 않다는 한계가 있음
  - 현재 AI학습용 데이터 구축 지원사업을 통해 KBS와 EBS, 또는 공공기관이 보유한 일부의 영상이 활용된 바 있으나, 더 다양한 영상이미지 원천데이터 구축을 위해서는 정부나 공공기관이 사용자와 제공자간에 적법하고 합리적인 데이터 사용 협상에 관한 조정이 필요함
  - 원천데이터에 대한 저작권, 초상권, 지재권 등의 보유자로부터 학습데이터 제공의 동인을 제고하고 AI학습용 데이터의 활용과정에서 권리가 배제되지 않도록, 협의에 기반한 공정하고 합리적인 데이터 거래 기준 마련 및 관리 필요
  - 다량의 멀티미디어데이터 및 저작권, 초상권을 보유한 방송사와 제작사, 창작자 등이 영상미디어분야 학습데이터 구축에 참여할 수 있게 하려면 인공지능개발과정에 참여하고 학습용데이터 활용으로부터 수익 배분을 받을 수 있는 협의과정이 필요함
  - 이에 공공부문이 주도하여 산학연관의 협의체 구성 및 합리적이고 공정한 데이터 구축을 추진해야 함
  - 데이터보유자들이 적지 않은 투자를 해 가며 축적·관리해 온 데이터의 활용결과물에서 보유자들의 이익과 배치되지 않도록 하는 활용범위와 방안에 대한 투명한 관리가 필요함
  - 민간·공공 보유 데이터를 쉽게 검색하고, 데이터사용자의 활용범위와 이용내역, 이후의 결과에 대한 추적과 관리까지 가능하도록 하는 데이터 이용생태계 구축 필요
- (최종 수요자의 데이터수요 반영) 영상미디어 분야의 AI기술의 최종 수요자에 해당하는 개발사, 제작자, 방송사, 미디어플랫폼의 AI기술 개발 방향과 AI학습용 데이터 수요 조사 및 반영 필요

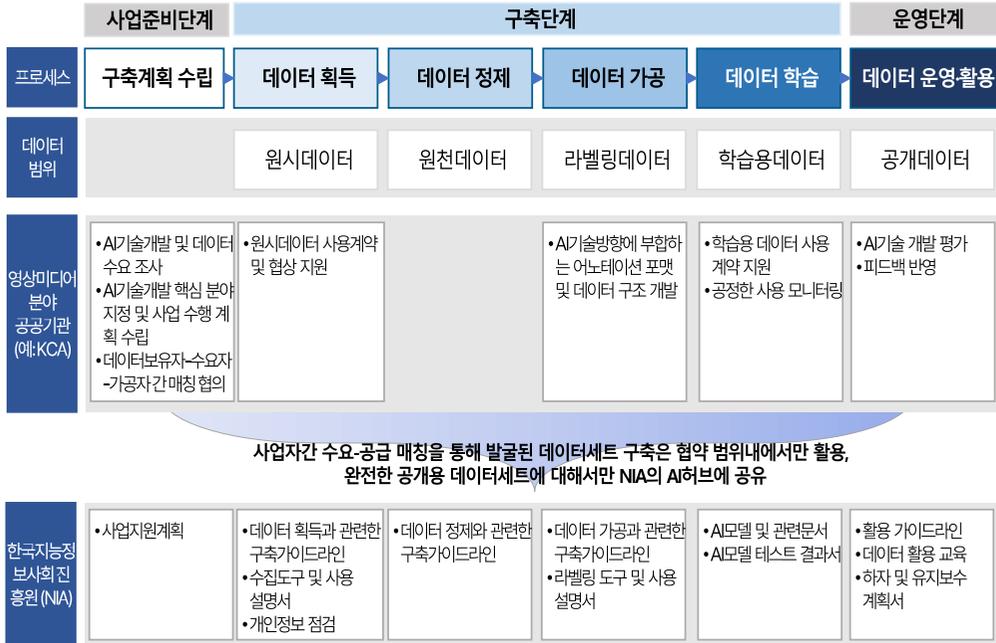
- 현재 과학기술정보통신부와 한국지능정보사회진흥원이 데이터 수요조사를 주도하고 있어, 영상미디어 분야의 학습용 데이터 수요가 반영되기 어려운 구조임
- 분과별 위원회가 운영되고 있으나 수요조사의 주체가 영상미디어 분야에 특화된 기관이 아니므로, 영상미디어 분야에 특화된 데이터 수요가 반영되기 어려운 한계가 있음
- 또한, 영상미디어 분야의 AI개발로 인한 최종 수요자인 방송사, 영상제작사, OTT플랫폼 등의 미디어 기업의 AI개발 수요와 데이터 수요에 대한 의견 수렴이 적절하지 않음
- 이에 영상미디어 분야의 주요 개발사, 제작자, 방송사, 미디어플랫폼 등의 AI기술개발 방향과 AI학습용 데이터 수요를 파악하고, 수요에 부합하는 데이터 구축을 주도할 수 있도록 하는 관리주체가 필요함

#### 나. 영상미디어분야 AI학습 데이터 구축 사업목표와 범위

- (영역 특화 기관의 AI기술 전략수립에 기반한 데이터 구축 사업 추진)
  - 중장기 AI기술전략 수립, 영상·미디어 분야의 AI기술개발 수요, 데이터 필요분야 발굴, 데이터산출물과 AI기술 활용 등의 영상미디어 분야 AI기술 개발의 전주기에서 주도적 역할을 담당할 수 있는 영역 특화기관이 데이터보유자-수요자-가공자 간의 매칭에 기반한 데이터 구축사업을 추진하는 것이 바람직함
  - 현재 한국지능정보사회진흥원 주도의 AI학습용 데이터 구축사업은 데이터 수요조사 단계에서부터 데이터보유자와 AI개발의 최종산출물의 이용자가 포함될 필요가 있음
  - 데이터가공자가 학습용 데이터를 구축하여 데이터수요자인 AI개발자나 개인에게 판매하는 방식의 데이터구축·개방 시스템은 영상미디어 분야에서 양질의 데이터를 확보하기 어렵고 영상미디어 분야의 AI기술 트렌드에 부합하는 수요가 반영되기 어려움

- 데이터보유자가 데이터개발 결과물의 최종 수요자가 되는 영상·미디어 산업의 특성을 고려하여, AI기술 개발에 대한 중장기적 방향과 전략을 수립하고 원시데이터 수집부터 최종 AI기술 활용단계까지 체계적 관리가 필요함

[그림 4-8] AI기술개발 프로세스에서의 영역특화기관의 역할



- 데이터보유자와 가공자, 개발사의 3자간 수요-공급 매칭을 통해 발굴된 데이터셋 구축은 영역특화기관을 통해 별도의 협약 범위 내에서 활용하며, 영역특화 기관은 3자간의 협상과 데이터 구축·활용을 지원함
- 공공데이터 및 완전한 공개가 허용된 공개데이터에 기반한 AI학습용 데이터 구축에 한해서 NIA가 구축한 AI허브에 공유하고, 데이터가공자와 데이터수요자간 공개적 거래는 AI허브를 통해서 추진함

#### 다. 영상미디어분야 학습데이터 구축 과제

- o (영역 특화 데이터 수집 필요성) 실제 콘텐츠 제작 및 유통 과정에 활용

할 수 있는 양질의 결과물을 얻기 위해서는, 타 산업영역에서 구축된 학습데이터보다는 영역특화의 목적에 부합하는 AI 학습용 데이터가 필수이며, 원하는 최종결과물과의 학습용 데이터 간 연관성이 높고 데이터의 양이 풍부할수록 AI학습 품질이 제고됨

- 특히, 영상·미디어 분야에서 영상 및 이미지 처리 및 생산에는 고품질의 대용량 학습데이터가 필요하므로 콘텐츠가 중요한 미디어 산업에서 경쟁력을 확보하기 위해서는 영상미디어 분야에 특화된 대량의 데이터 구축 사업이 필요함

<표 4-5> 영상미디어 분야 AI 학습데이터 구축 목표

분야	AS-IS	TO BE	목표
대상 인식	<ul style="list-style-type: none"> <li>• 한국인 안면인식 등 기초데이터 확보</li> <li>• 사람, 인물, 동물, 가구 등 구체적인 사물 이미지 데이터</li> <li>• 인물·객체 중심 3D데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 감성, 개념 등 추상화된 대상에 대한 시각적 이미지 데이터와 어노테이션</li> <li>• 배경·환경·건축물 등 비인물 랜드스케이프 3D공간 데이터 확보</li> <li>• 환경 맥락 내 랜드마크 이미지, 랜드마크 식별, 랜드마크 태깅</li> </ul>	인식 능력 향상 및 3D파라미터 구현
영상 이해	<ul style="list-style-type: none"> <li>• 비전문적으로 제작된 짧은 영상</li> <li>• 영상 내 객체 인식 중심용 어노테이션</li> <li>• 단일모드의 원천데이터</li> <li>• 연계성이 없이 개별적으로 구축된 동영상, 스크립트, 음성/오디오 데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 전문적으로 제작된 풀타임 방송영상 프로그램</li> <li>• 카테고리, 동작, 감정, 이벤트 등의 다계층적 어노테이션</li> <li>• 멀티모드의 원천데이터 (동영상-스크립트-음성/오디오 등)</li> <li>• 텍스트-영상-음성을 연동하여 정렬할 수 있는 연계된 데이터</li> </ul>	통합 능력 확보
동작·상황 인식	<ul style="list-style-type: none"> <li>• 센서로 간단한 동작인식 데이터</li> <li>• 단순 행위 분류 어노테이션</li> </ul>	<ul style="list-style-type: none"> <li>• 동작·상황 등의 종합적 인지, 가상의 공간인식을 지원하는 데이터 구축</li> <li>• 동작과 상황에 대한 다층적 어노테이션</li> </ul>	상호작용 능력 확보
상호작용 데이터	<ul style="list-style-type: none"> <li>• 좋아요, 댓글, 시청 시간 등 사용자 상호작용 데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 동공의 상대적 위치, 맥박 등 신체데이터</li> </ul>	상호작용 능력 확보

o (영상미디어 특화 AI학습데이터 구축목표) 위의 <표 4-5>와 같이 영상

미디어 분야의 AI기술 발전 방향을 고려할 때, 학습데이터 구축의 목표는 각 영역별로 다음과 같이 전환이 필요함

### ① 대상인식기술 측면

- 현재 대상인식 기술 측면에서 구축된 AI학습용 데이터는 주로 한국인 안면인식 데이터, 구체화된 사물이나 사람을 인식할 목적으로 구축됨
- 특히 대상인식 기술 학습을 위한 영상데이터의 상당수가 자율주행차나 로봇의 사물인식 및 행동제어에 적합하도록 구축되어 있어, 영화나 방송프로그램에서 필요한 객체의 시각적인 정보가 부족함
- 현재의 구체화된 사물과 인물 인식용 데이터의 수집도 필요하나, 영상 미디어 분야에서 AI의 영상표현의 범위를 확대하기 위해서는 감성, 개념 등 추상화된 대상에 대한 시각적 이미지 데이터와 어노테이션이 필요함
- 이를 위해서는 추상적인 개념들을 표현한 영상, 이미지, 음성 등이 포함된 양질의 원시데이터 확보가 필요함
- 더불어 메타버스, VFX 등에서 파노라마 3D를 구현하기 위해서는 인물과 객체 중심의 3D 시각정보 데이터뿐만 아니라 배경, 환경, 건축물 등 비인물 랜드스케이프 3D공간 데이터 확보가 필요함

### ② 영상이해기술 측면

- 현재 구축된 영상미디어 분야 데이터는 동영상-스크립트-음성-어노테이션-메타데이터가 모두 포함된 데이터가 부족한 환경이므로, 단순 객체인식 중심의 단일 모드 정보에 국한되어 영상을 요약하는 기술이 주류를 이룸
- 현재 AI허브에 공개된 영상요약 관련 데이터는 비교적 짧은 길이의 영상 데이터셋 3건으로 파일명, 영상길이, 시작 및 종료 시간, 카테고리 등의 단순 정보만 제공되며, 멀티모달 학습에 적합한 스크립트 데이터가 부재하고 어노테이션 구조가 단순함

- 영상미디어 분야에서 영상의 맥락을 이해하기 위해서는 단순히 시각적 정보로만 구성된 단일 모드의 원시데이터보다는 동영상-스크립트-오디오-음성이 모두 포함된 내러티브가 있는 멀티모달 학습용 데이터 구축이 필요함
- 이러한 이벤트 감지를 위해서는 다양한 스포츠 중계영상 등의 원천데이터가 필요하나, 고가의 중계권과 저작권 등의 문제로 공개된 학습데이터 구축이 이뤄지지 않음
- 멀티모달 기반 영상 이해 기술은 오디오, 시각적 이미지, 텍스트 등의 다양한 유형의 데이터가 통합적으로 필요하므로 대화, 배경음악, 시각적 내러티브 등의 맥락 정보가 포함된 원시데이터가 필요하며, 이를 정제하는 과정에서 비디오 프레임과 장면 간의 종속성을 포착할 수 있도록 하는 어노테이션 라벨링이 필요함
- 따라서 ① 영상 인식의 고도화를 위해서는 양질의 전문적으로 제작된 풀타임(full-time) 방송영상 원시데이터의 확보, ② 텍스트-영상-음성을 연동하여 정렬할 수 있도록 연계된 멀티모드 원시 데이터세트(동영상-대본-음성/오디오) 확보, ③ 카테고리, 동작, 감정, 이벤트 등의 다층적 어노테이션 부여 등이 추진되어야 함

### ③ 동작·상황인식 기술 측면

- 현재 기 구축된 학습용 데이터의 상당수는 센서로 간단하게 관절의 움직임, 자세 등을 포착하는 동작인식 데이터 중심임
- 이들 동작인식 데이터들은 그 수집 목적 자체가 자연스러운 움직임을 도출하기 위한 목적이거나 보다는 의료적인 진단이나 운동경기의 동작 판정 등을 목적으로 하기 때문에 동작의 정확성 등을 파악하기 위한 수준의 어노테이션이 부여됨
- 영상미디어 분야에서 제작과정에 사용될 수 있는 AI 동작인식기술 개발을 위해서는, 동작과 상황/맥락에 대한 다층적 어노테이션이 부여된 학습데이터 세트 구축이 필요함

- (학습데이터 구축범위) 동영상-스크립트-음원-어노테이션-메타데이터가 모두 포함된 풀타임(full-time) 방송영상데이터 구축이 필요함
  - 기존에 구축·공개된 영상 요약기술용 학습데이터는 짧으면 수초에서 수분 내외의 동영상 클립 중심으로 구축되어 있으며, 정규 방송프로그램이나 영화와 같은 저작권이 있는 양질의 기존 영상물들은 데이터 공유가 이뤄지지 않음
  - 상업적인 추가 수익화가 어려운 오래된 비인기 영상이라도 전문적인 제작사나 방송사가 제작한 풀타임 영상 데이터와 스크립트의 공유가 이뤄져야 내러티브에 기반한 영상 요약 기술 개발에 유리할 것으로 보임
  - 이에 공공 스포츠이벤트로 중계권이나 저작권 프리가 확보된 영상, 영상저작물의 보호기간이 소멸되었거나 저작권기부, 공공기관이 협약을 통해 저작권 공유 및 개방을 득한 전문적인 방송 및 영화 영상 등을 확보하여 풀타임 영상데이터와 스크립트의 공유가 필요함
  - 더불어 해당 원천데이터에 대한 메타데이터 및 어노테이션 작업을 지원하여 지도학습용 자료로 개발이 필요함

<표 4-6> 멀티모달 인공지능개발을 위한 데이터 구축 범위

구분	내용
대규모 동영상 데이터	-전문적으로 제작된 풀타임 방송영상 프로그램 원시데이터 -감성, 개념 등 추상화된 대상에 대한 시각적 이미지 데이터 -비인물 랜스케이프 3D공간 데이터
음성데이터	-텍스트를 시각적 콘텐츠와 정렬하는 데 사용할 수 있는 동영상의 음성 콘텐츠 녹음본
오디오데이터	-시각적 데이터와 함께 분석할 수 있는 사운드트랙, 대화 및 기타 오디오 요소
어노테이션(annotation)	-지도 학습의 기본 정보로 사용되는 동영상의 중요한 부분을 나타내는 사람, 객체 등을 설명하는 요약 또는 라벨 -동영상에서 발생하는 콘텐츠, 동작, 이벤트를 설명하는 어노테이션(annotation)이 포함된 대규모 동영상 데이터세트
설명텍스트 데이터	-시가 학습할 수 있는 추가 컨텍스트를 제공할 수 있는 비디오 콘텐츠의 텍스트 설명 또는 요약정보

	-동영상과 관련된 대본이나 설명은 동영상과 텍스트 정보를 결합하여 다중 모달 학습에 사용
메타데이터	-동영상 제목, 태그, 사용자 참여 측정항목과 같은 정보는 요약을 위한 추가 컨텍스트를 제공
사용자 상호 작용 데이터	<ul style="list-style-type: none"> <li>- 좋아요, 댓글, 시청 시간 등 사용자가 동영상 콘텐츠와 상호 작용하는 방식에 대한 정보</li> <li>- 동영상의 내러티브에 따른 동공의 상대적 위치나 심장박동 등의 신체적 데이터</li> <li>- 시가 동영상의 어떤 부분이 가장 매력적이고 중요한지 이해하는 데 활용</li> </ul>

# 참 고 문 헌

## 국내 문헌

- 개인정보보호위원회(2022), 가명정보처리 가이드라인
- 고용노동부 (2023. 8. 31). 2027년까지 인공지능(AI) 12,800명, 클라우드 18,800명  
신규인력 부족 전망. <보도자료>
- 고환경 외(2020), AI 산업 발전을 위한 데이터 법제의 주요 쟁점과 개선방향-자율주행  
정밀지도 및 학습데이터 활용을 중심으로, DAIG 2020년 제1호
- 과학기술정보통신부 (2023. 9. 12). <AI와 디지털 기반의 미래 미디어 계획>.
- 과학기술정보통신부·방송통신위원회 (2023). <2023년 방송산업 실태조사 보고서>.
- 관계부처 합동 (2023. 4. 14). <초거대AI 경쟁력 강화 방안>
- 관계부처 합동(2023.7.21.), 국민 일상 편의 및 산업 경쟁력 제고를 위한  
서비스산업의 디지털화 전략, 가명정보 활용 확대방안
- 권순선 (2020). 인공지능과 빅데이터 기술동향, TTA저널 187호, 2020.1/2월호
- 김동원 (2023. 4. 18). 정부 ‘초거대 AI’ 지원책 비난 속출… “공급사 중심 반쪽  
정책”. <The AI>.
- 김범수 외 (2020). <인공지능 산업 육성을 위한 개인정보보호 정책 개선 방안  
연구>. 나주: 한국인터넷진흥원.
- 김나은(2022). 생성형 AI 활용사례와 활용방안. 디지털산업정책연구소
- 김승준 (2023.09.18.) 업스테이지 AI ‘솔라’, 글로벌 생성 AI 플랫폼 메인 껴챘다  
(뉴스1)<https://n.news.naver.com/mnews/article/421/0007057649?sid=105>
- 김종호 (2023.09.18.) 유통가 마케팅·서비스도 ‘AI 열풍’ (이뉴스투데이)
- 금융위원회·금융감독원(2022), 금융분야 가명·익명처리 안내서
- 네이버 (2023). <2022년도 4분기 실적발표>.
- 네이버AI (2023). Korean Safety Benchmarks. <https://github.com/naver-ai/korean-safety-benchmarks>
- 노창희 (2023). 인공지능 시대 국내 미디어 시장의 지속 가능한 성장을 위한  
파운데이션 구축 방향. <미디어 이슈 & 트렌드>, 56호, 1-8.
- 딜로이트(2023) 「인공지능(AI)활용서: 6대 산업별 AI 활용사례
- 문체부·저작권위원회(2023), 생성형 AI 저작권 안내서
- 박강수(2023. 12. 14.) ‘방송계 “AI, 우리 영상 무단 학습 안 돼”…빅테크사에  
의견서’, 한겨레, <https://www.hani.co.kr/arti/society/media/1120385.html>
- 박민우, 김대웅 (2021.01.29.) South Korean AI technology brings back folk  
singer’s voice (REUTERS)<https://www.reuters.com/article/us-southkorea>

-ai-voice-recreation-idUSKBN29Y180

- 방송통신위원회 (2023). <2022년도 방송사업자 재산상황 공표집>  
법률신문(2020.7.2.), 개인정보제공 동의, 형식적 사전동의절차 대신 실질적 절차  
모색해야, <https://www.lawtimes.co.kr/news/162654>
- 소프트웨어정책연구소 (2023). <2022 인공지능산업 실태조사>.  
세종:과학기술정보통신부.
- 소프트웨어정책연구소(2023. 봉광호, 안미소, 김정민). 「국내 인공지능(AI)  
도입기업 현황 분석 및 시사점」
- 이채환 (2023.08.11.) AI 활용성 높은 6대 산업군 (GTT  
KOREA)<https://www.gttkorea.com/news/articleView.html?idxno=6369>
- 임효진·박형준 (2022). 개인정보 보호와 활용에 관한 규제정책 변동에 있어 정책  
내러티브 활용: ‘데이터 3법’ 개정 과정을 중심으로. <규제연구>, 31권 2호,  
3-37.
- 우지수 (2023.09.19.) “뉴스 속 기자가 인공지능?”…SK브로드밴드, 뉴스에 AI 도입  
(더팩트)<https://n.news.naver.com/mnews/article/629/0000239483?sid=105>
- 오동현 (2023.09.19.) 업스테이지 AI 챗봇 ‘아숙업’, 세바시 강연 상담사 됐다  
(뉴시스)[https://newsis.com/view/?id=NISX20230919\\_0002455248&cID=10406&pID=13100](https://newsis.com/view/?id=NISX20230919_0002455248&cID=10406&pID=13100)
- 이용·장래영·박민우·이건우·최명석(2021). 자동-레이블링 기반 영상 학습데이터  
제작 시스템, 한국콘텐츠학회논문지 제21권 제6호 2021.6 701 - 715
- 이경탁 (2023.09.18.) ‘클로바X’로는 부족해… 네이버, 후속 AI 서비스로 반전  
노린다 (조선비즈)<https://n.news.naver.com/mnews/article/366/0000933108?sid=105>
- 이재현 (2023.09.14.) 국내 AI 경쟁 시작… 생태계 경쟁력 높아지나  
(머니S)<https://n.news.naver.com/mnews/article/417/0000948629?sid=105>
- 이수엽(2021), 인공지능과 미디어 엔터테인먼트 산업, 기획리포트
- 이인호 외(2019), 개인정보 처리 동의제도 개선방안 연구, 한국인터넷진흥원
- 정원준(2023), 개인정보 동의제도의 실질화 방안 연구, 고려법학 제108호
- 조용성 외(2020), 미디어와 AI기술: 미디어 지능화, 전자통신동향분석 제35권  
제5호, 2020년 10월. 재판집
- 조창현 (2023.09.11.) 전산업 AI 접목 확산… “AI, 경제·사회 전반 혁신 이끌게 될  
것” (인더스트리 뉴스)  
<http://www.industrynews.co.kr/news/articleView.html?idxno=50931>
- 정주원 (2023.01.01.) 60대 최민식이 확 젊어졌다?…30대 연기 가능하게 한 ‘비결’  
(매일경제) <https://www.mk.co.kr/news/culture/10588818>
- 채제우 (2023.09.12.) 과기정통부, AI로 토종 OTT 경쟁력 살린다  
(조선일보)<https://n.news.naver.com/mnews/article/023/0003787354?sid=101>

채제우 (2023.09.14.) 식당·병원·학교·미용실... 국민 일상에 'AI' 심는다 (조선일보)  
 트라이닷 매거진 2023 「[AI Trends] AI를 활용한 미디어 콘텐츠 알아보기」  
 한광범 (2023.09.15.) 업스테이지, 세계서 기술력으로 주목받는 AI 스타트업  
 (이데일리)<https://n.news.naver.com/mnews/article/018/0005576082?sid=105>  
 한국방송통신전파진흥원(2022), 영상 미디어 워크플로우 단계별 AI 활용 현황,  
 2022 미디어 이슈&트렌드 9·10월호 (Vol.52), 재구성  
 한국정보통신기술협회(2023) TTA저널 207호 「미래사회의 새로운 키워드  
 초거대AI」  
 한국지능정보사회연구원(2022), 인공지능 학습용 데이터 활용 우수사례집  
 한국정보통신기술협회(2020). 미래 사회의 새로운 키워드, 초거대AI,  
 TTA저널 207호  
 한정훈 (2023. 9. 28.). “취향 저격 드라마를 만난 구독자”, “크리에이터와 더  
 가까워진 청취자”중심은 AI, 다이렉트미디어랩,  
<https://directmedialab.com/cwihyang-jeogyeog-deuramareul-manna-n-gudogja-keurieiteowa-deo-gaggaweojin-ceongcwija/>

## 국외 문헌

BBC (n.d.), Old School Dataset, <https://www.bbc.co.uk/rd/projects/old-school-dataset>  
 Bizwit Research & Consulting LLP (2021). Global Online Video Platform  
 Market Size study , by Type, By Application, By End User and  
 Regional Forecasts 2021-2027.  
 Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S.,  
 ... & Liang, P. (2021). On the opportunities and risks of foundation  
 models. arXiv preprint arXiv:2108.07258.  
 Contreras, B. (2022). AI is here, and it's making movies. Is Hollywood ready?,  
<https://www.latimes.com/entertainment-arts/business/story/2022-12-19/the-next-frontier-in-moviemaking-ai-edits>  
 GM Insights(2023). AI in Automotive Market Size, Share & Trends Report  
 2032 (gminsights.com)  
 Jolly, B. (2022. 5. 17.), A new dataset to improve TV production using  
 artificial intelligence and machine learning, BBC,  
<https://www.bbc.co.uk/rd/blog/2022-05-artificial-intelligence-machine-learning-cinematography-editing-framing>  
 Global Market Insights(2023). AI in BFSI Market Size By Component

(Solution [Chatbot, Customer Behavior Analytics, CRM, Data Analytics & Visualization, Fraud Detection], Service), By Technology (Machine Learning, NLP, Computer Vision), By Application, End-use & Forecast, 2023 - 2032

Market Research Future(2023). AI in Retail Market Size, Share and Trends 2032 (marketresearchfuture.com)

Markets and Markets(2023). Artificial Intelligence(AI) in Healthcare Market Size, Growth Report Analysis 2031 (marketsandmarkets.com)

MarketandMarkets(2021). Artificial Intelligence(AI) Market . Global Forecast To 2026.(홍동선 (2022), 금융 AI 시장 전망과 활용 현황 : 은행권을 중심으로, 한국신용정보원, 재인용.)

Mittal, A.(2023. 8. 8.), Generative AI, the idea behind ChatGPT, DallE, Midjourney and more,  
<https://www.unite.ai/ko/generative-ai-the-idea-behind-chatgpt-dall-e-midjourney-and-more/>

Saftler, B. & S. Sobel (2023. 6. 6.), Large Language Models in Media & Entertainment, DataBricks

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell: A neural image caption generator. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3156-3164).

Word Net. (2023a, September 8). In Wikipedia.  
<https://en.wikipedia.org/wiki/WordNet>