

생성형 AI는 미디어 생태계의 게임 체인저인가

2025. 1. 17 금요일
13:00~16:30

서울과학기술대학교 상상관 3층 오픈 스페이스

주최·주관 서울과학기술대학교 IT정책전문대학원 및 원우회



생성형 AI는 미디어 생태계의 게임 체인저인가

2025. 1. 17 금요일
13:00~16:30

서울과학기술대학교 상상관 3층 오픈 스페이스

주최·주관 서울과학기술대학교 IT정책전문대학원 및 원우회



생성형 AI는 미디어 생태계의 게임 체인저인가

2025. 1. 17 금요일
13:00~16:30

서울과학기술대학교 상상관 3층 오픈 스페이스

사전 네트워킹

12:00~13:00

사회 박소현(IT정책전문대학원 디지털문화정책전공 교수)

제1세션(발표)

13:00~13:50

/// 생성형 AI에 대한 TV 제작진 인식 및 활용 연구 | **조서윤** 융합미디어콘텐츠정책전공 박사과정

/// AI 광고의 사용자 경험 확장 연구 | **차혜영** IT디자인융합전공 박사과정, 광주대학교 산업디자인학과 초빙교수

Break

13:50~14:00

제2세션(발표)

14:00~14:50

/// 인공지능 생성물의 투명성 규범 방안 | **박명순** 융합미디어콘텐츠정책전공 박사수료

/// 이미지생성형 인공지능이 그리는 기괴한 행복의 세계 | **오경미** 디지털문화정책전공 박사수료, 오픈넷 연구원

Break

14:50~15:10

좌장 황주성(IT정책전문대학원 시공공정책전공 교수)

제2세션(종합토론)

15:10~16:30

/// **김대규** HCN 정책연구원

/// **안무정** LG CNS 책임

/// **오병일** 진보네트워크센터 대표

/// **이선우** JTBC PD

/// **전영균** IT정책전문대학원 정책학 박사

/// **조경숙** 테크-페미 활동가

※ 본 콜로키엄은 ST융합과제 "사회 속의 인공지능(AI): 기회와 위협"의 일환으로 수행되었습니다.

제1세션

발표 1.

생성형 AI에 대한 TV 제작진 인식 및 활용 연구 | 조서윤 융합미디어콘텐츠정책전공 박사과정
..... 12

발표 2.

AI 광고의 사용자 경험 확장 연구 | 차혜영 IT디자인융합전공 박사과정, 광주대학교 산업디자인학과 초빙교수
..... 24

제2세션

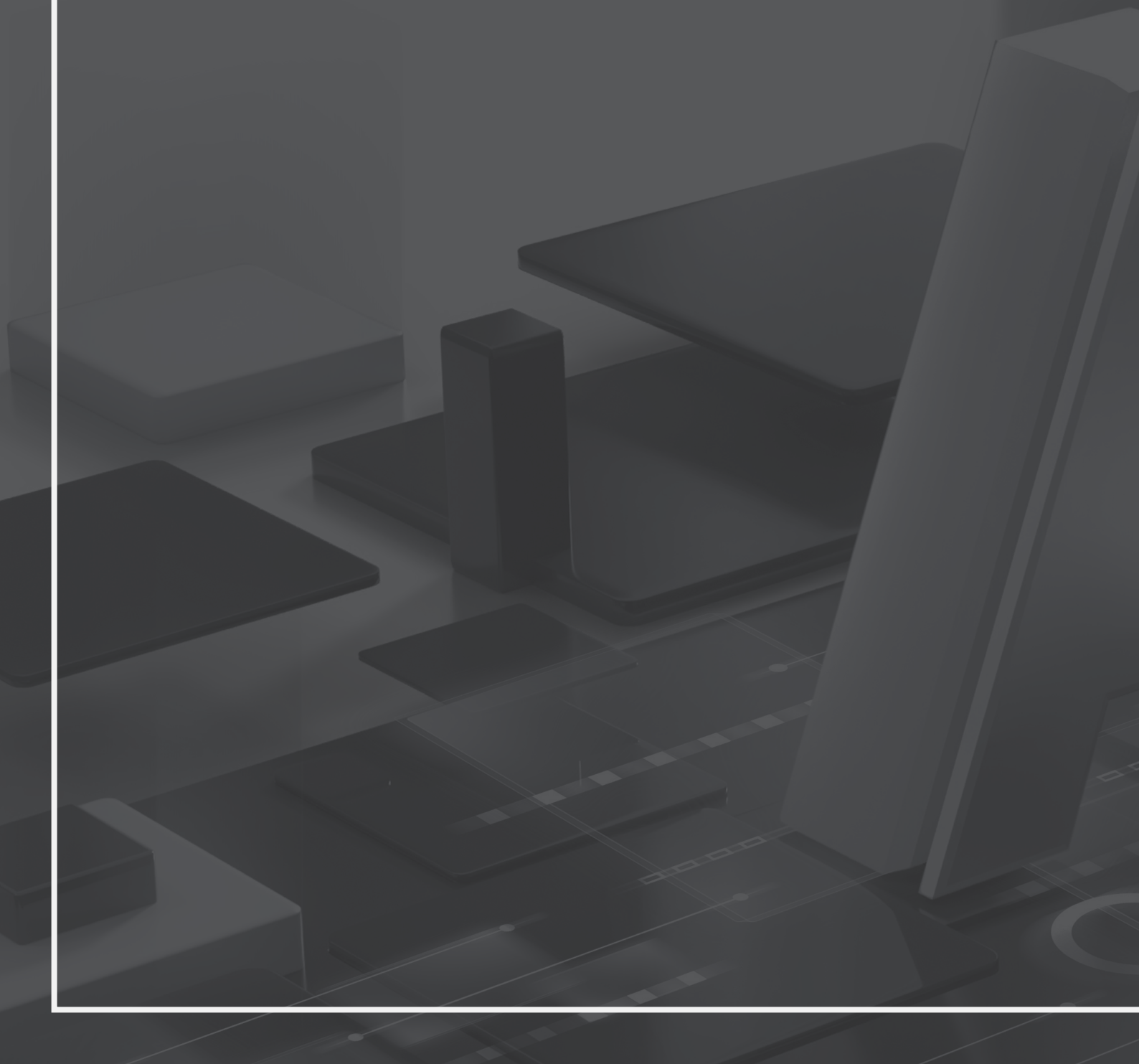
발표 1.

인공지능 생성물의 투명성 규범 방안 | 박명순 융합미디어콘텐츠정책전공 박사과정
..... 36

발표 2.

이미지생성형 인공지능이 그리는 기괴한 행복의 세계 | 오경미 디지털문화정책전공 박사수료, 오픈넷 연구원
..... 50

**생성형 AI는
미디어 생태계의
게임 체인저인가**



제1세션

발표 1.

생성형 AI에 대한 TV 제작진 인식 및 활용 연구

| 조서윤 융합미디어콘텐츠정책전공 박사과정

발표 2.

AI 광고의 사용자 경험 확장 연구

| 차혜영 IT디자인융합전공 박사과정, 광주대학교 산업디자인학과 초빙교수

**생성형 AI는
미디어 생태계의
게임 체인저인가**



생성형 AI에 대한 TV 제작진 인식 및 활용 연구

| 조서윤 융합미디어콘텐츠정책전공 박사과정



THE PERCEPTION AND UTILIZATION OF GEN AI BY TV PRODUCTION STAFF

생성형 AI에 대한 TV 제작진 인식 및 활용 연구

조서윤

IT정책전문대학원 석박사 통합과정



IMAGE BY MIDJOURNEY
PROMPT @WACKYCHICKCHO

INDEX

- INTRODUCTION : GEN AI
- FEATURED GEN AI APP/SW
- SHOWCASE- TV SHOW
- RESEARCH SUMMARY
- CONCLUSION
- DISCUSSION
- REFERENCE

- GEN AI=Generative AI

INTRODUCTION

- WHAT IS GENERATIVE AI? (생성형 AI 정의)

-사전 학습된 데이터를 기반으로, 입력과 유사하거나 새로운 결과물을 생성하는 기술/ 복잡한 텍스트, 이미지, 오디오, 비디오 등의 콘텐츠를 자율 생성하도록 고안된 시스템 (정원준, 2023)

-모델 구조 분류: GANs, VAE(Variational Auto Encoder), Transformer, Diffusion 모델 (이선종, 2023)

-미국 영화,TV 제작자 설문 조사: 30%가 사용/ 실감 오디오 생성(58%), 스토리 구성(52%), 3D 소스 제작(47%) 순으로 활용 (Variety, 2023)

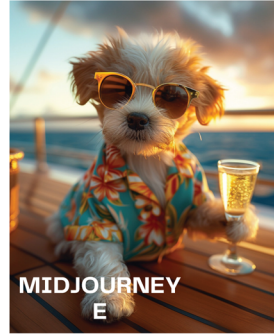
-국내 TV 제작진의 생성형 AI 사용에 대한 설문 조사: 사용 중 26.1%, 검색, 문장 작업(번역, 문서, 스크립트), 이미지 생성 순으로 사용 빈도 높음 (조서윤, 이영주, 2024)

FEATURED TOOLS(S/W & APP)

- 대표적 상용화 도구 (소프트웨어 & 앱)

TTI	TEXT TO IMAGE	MIDJOURNEY, DALL-E, IMAGEN, FIREFLY, STABLE DIFFUSION
TTV	TEXT TO VIDEO	RUNWAY, PIKA, LUMA, PIKA, VEO, SORA
TTA	TEXT TO AUDIO	UDIO, SUNO, Stability Audio
TTV	TEXT TO VOICE	ELEVEN LABS, D-ID, HEYGEN, TYPECAST
ETC.	TRANSLATION WRITING AI EDITING DIGITAL AVATAR	CHATGPT, GEMINI, CLAUDE AI, CANVA, NOTION, CANVA, ADOBE PREMIER, LENSEA, SYNTHESIS AI

FEATURED TOOLS(TEXT TO IMAGE)



SAME PROMPT & DIFFERENT OUTPUT

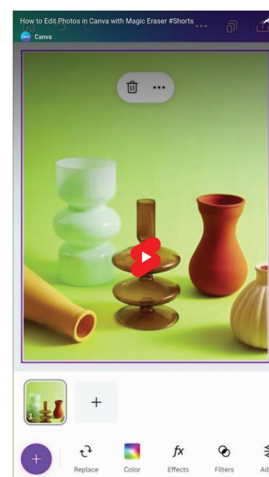
동일 프롬프트 입력--> 달리와 미드저니 다른 출력값

CHATGPT/MIDJOURNEY @ WACKYCHICKCHO

FEATURED TOOLS(VIDEO EDITING)

• CANVA • ADOBE PREMIER (FIREFLY)

- **캔바(Canva)**
- Democratization of Design(디자인의 민주화)를 모토로 호주에서 설립
- 매우 쉬운 UI, UX로 많은 인기
- 이미지, 비디오, 오디오, 문서 작성 등 멀티 기능 원 스톱 디자인 사이트
- 최근 Leonardo ai/ Affinity 등을 인수하면서 AI 기능에 주력 중
- 다양한 AI 소프트웨어, 앱과 콜라보를 통해 다양한 AI PLUG-IN(플러그인) 강화
- **어도비 파이어플라이 (Adobe Firefly)**
- 인기 편집 소프트웨어 어도비에서 개발한 생성형 AI 앱 TEXT TO IMAGE)
- 프리미어 편집 SW와 연계하여 시너지 효과



캔바 기본 AI 기능 소개 영상

SHOWCASE TV SHOW

- 방송 제작 실제 사례 (조서운PD)
- 프로그램: 한끗차이 (E채널, 넷플릭스)
- 회차: 1~12회 / 회차 당 평균 130장 생성+ 2~30장 실제 방송에 사용
- 사용 TOOL: 챗GPT, CLAUDE, DEEPL(프롬프트 작성), CANVA(이미지 편집), 미드저니
- 제작비: 약 68원/장
- 완성 소요 시간: 장당 10~15분 내외
- 제작진 반응: 최고 만족도(퀄리티, 제작비)
- 시청자 반응: 높은 퀄리티, 호감도)



인간 삽화가 생성물(장당 8만원, 2~3일 소요)



미드저니 생성물(장당 68원, 5분 소요/ 조서운 작업)



SHOWCASE HUMAN VS GAI

- 당초 인간 삽화가에게 의뢰----> 짱구는 못말려 수준 결과물
- 담당 프로듀서가 생성형 AI로 직접 생성 시도
- 챗GPT, 미드저니, 캔바 앱 활용
- 구독료 반영 정산 결과 장당 100원 미만
- 제작 소요 시간 장 당 5분

SHOWCASE TV SHOW

- **생성형 AI 활용 부정적 사례**
- M사 <심야괴담> 사용된 삽화
- 이유: 결과물의 낮은 완성도--> UNCANNY VALLEY(불쾌한 골짜기)--> 프롬프트 오류로 인한 불쾌한 결과물



TVREPORT
'심야괴담회' PD, 논란된 'AI' 재연 영상에..."제작비가 없어요"
 입력 2024.07.17. 오후 6:57 | 기사원문



한끗차이 사용 삽화 이미지

WORKFLOW PROMPT 1

- 제작진: 필요한 장면 내용 전달 (한국어)
- 한국어 장면 설명--> 챗GPT, CLAUDE, DEEPL을 모두 사용+크로스 체크--> 영어 프롬프트 완성

한끗차이 4회 방송용 삽화 프롬프트 예시

4회 삽화 프롬프트

1> 트럼프 선거 캠프 현장. 기자 회견 하는 검은 볼테 안경 낀 회색 양복 차림의 로저 스톤

https://story-img.kakaocdn.net/dn/bmi502/hyVxxfUFl/cNpkvkSgs2DHW4ds7KzQk1/i/mg_xl.jpg?width=447&height=465&avg=%2523836a67&face=102_204_306_427&v=2

Trump campaign sage, Roger Stone in gray suit with black horn-rimmed glasses giving a press conference, simple water color illustration, --ar 16:9



WORKFLOW PROMPT 2

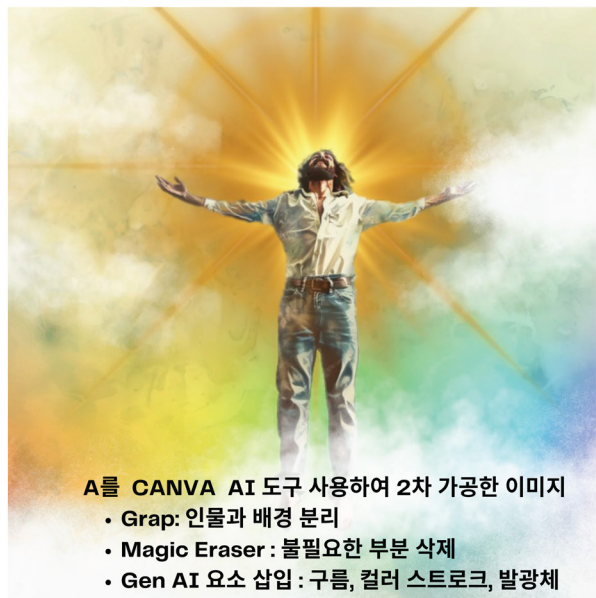
- 미드저니용 프롬프트 완성 (파라미터 설정/ 조정)
- 프롬프트 입력----> 이미지 생성
- 실존, 외국 인물 생성값 매우 우수 (학습 데이터양 비례)
- 신체 묘사 (손가락, 손), 순간 묘사, 캐릭터 일관성 다소 부족
- <--CREF> 파라미터 사용 난이도 (왜곡된 이미지 생성)
- 프로그램 특성상 “SIMPLE WATER COLOR ILLUSTRATION “ 프롬프트 사용



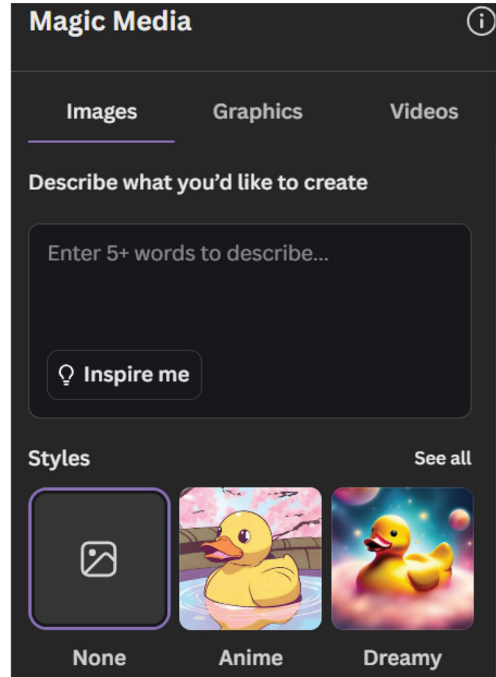
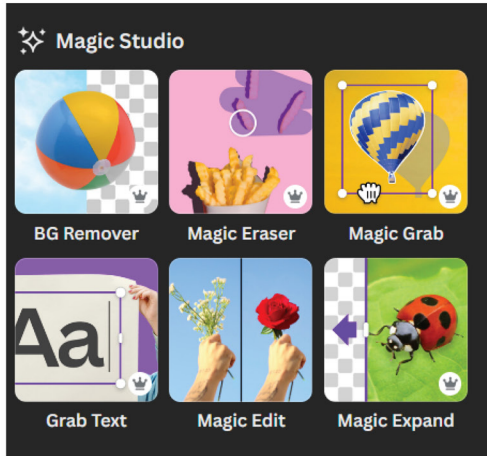
A : 미드저니 생성 이미지

WORKFLOW IMAGE EDITTING

- 미드저니 생성물 2차 가공
- CANVA의 AI 툴 사용----> **그랩/ 이레이저, 익스텐션 등**



- A를 CANVA AI 도구 사용하여 2차 가공한 이미지
- Grap: 인물과 배경 분리
 - Magic Eraser : 불필요한 부분 삭제
 - Gen AI 요소 삽입 : 구름, 컬러 스트로크, 발광체



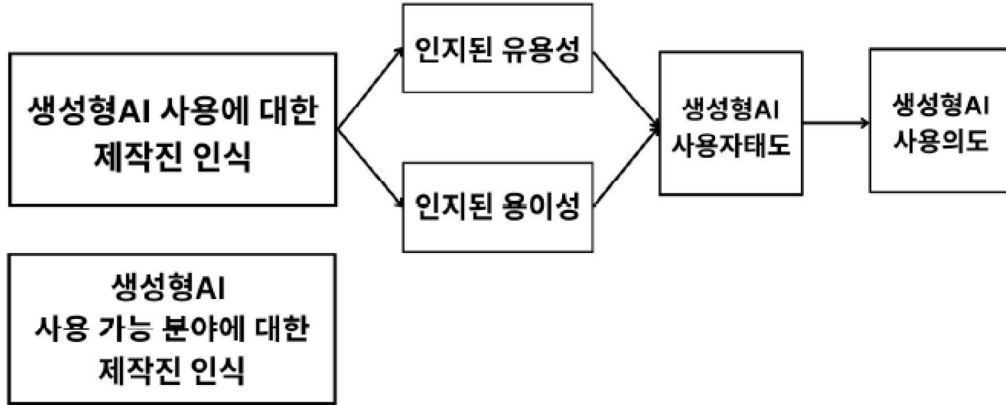
WORKFLOW IMAGE EDITTING

- CANVA 주요 툴 화면----> 매직 그랩/ 블러링/ GEN AI

RESEARCH SUMMARY

- TV 방송 제작진의 생성형 AI기술 수용에 영향을 미치는 요인: 제작진 인식을 중심으로
<조서윤, 이영주(2024), 한국소통학보 제23권 3호, 277-310>
- 기술 수용 모델(TAM) 토대로 생성형 AI 기술에 대한 TV 방송 제작진의 인식 조사
- 생성형 AI 기술 사용에 있어 사용자 태도와 기술 수용 의사에 미치는 제작진의 인식 요인 탐구
- 연구 표본: 국내 방송 현업 종사자 115명(PD, 작가, 기타 스태프/ 드라마, 예능, 시사교양/ 지상파, 종편, 유선 등)
- 연구 방법: 전문가 설문 조사(2024. 3.15~25 온라인 설문 시행)
- 이론적 배경: 기술수용모델/생성형 AI 사용에 대한 인식(긍정, 부정)/ 결과물의 완성도 관련 연구
- 연구문제:
 - 1-1: 방송 영상 제작에 있어 생성형 AI 사용 가능성 및 업무 효율성이 높을 것으로 기대되는 분야는 무엇인가?
 - 1-2: 생성형 AI 사용에 방송 제작진의 인식은 어떻게 분류되는가?
 - 2: 생성형 AI 사용에 대한 제작진의 인식은 인지된 유용성과 인지된 용이성에 어떠한 영향을 미치는가?
(이하 기술 수용 모델 기반 가설 3개 설정: 유용성, 용이성--> 사용자 태도-->사용의도)

RESEARCH SUMMARY 연구 모형



RESEARCH SUMMARY

- 연구 문제 1 측정 문항 및 주요 변인의 조작적 정의

생성형 AI 사용 가능성 및 업무 효율성이 높은 분야 측정 문항

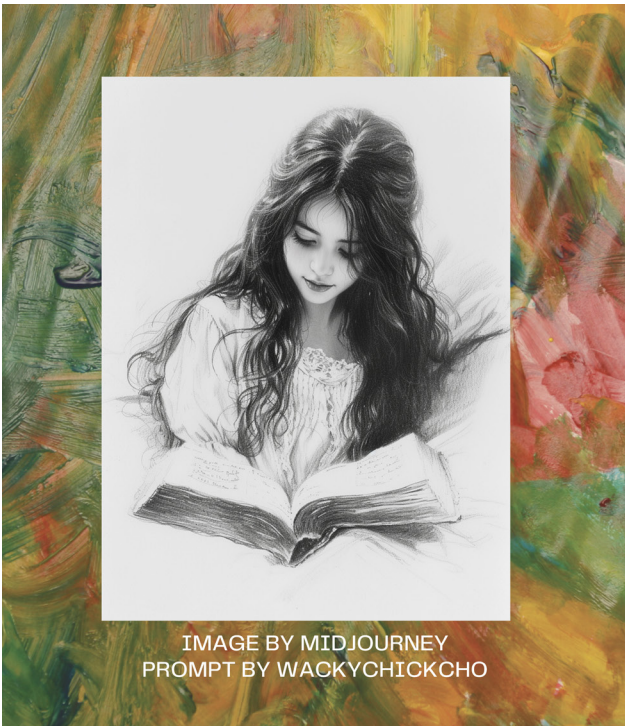
제작단계	측정 항목
1단계	아이디어(기획, 구성), 검색(자료, 섭외)
2단계	문장 작성(대본, 번역, 스크립트, 문서), 예산안(작성, 정산)
3단계	촬영(카메라), 녹음(현장), 미술(세트, 소품), 의상, 분장
4단계	이미지 생성, 비디오 생성, 오디오 생성, 비디오오디오 편집, 디지털 아바타, 종합편집, 자막, 그래픽, 마케팅, 홍보

주요 변인의 조작적 정의

요인	측정 항목	출처	M (SD)	α
유용성	생성 AI 기술 사용이 방송 제작에 도움이 된다. 생성 AI는 방송 제작의 업무 생산성을 높여 준다.		3.73 (.65)	.77
용이성	생성 AI 앱 작동은 직관적이라 사용이 쉽다. 생성 AI는 전반적으로 사용하기가 쉽다.	Davis et al. (1989)	3.30 (.74)	.83
사용자 태도	방송 제작에 생성형 AI를 사용하는 것에 대해 긍정적으로 생각한다. 타 제작진이 사용하는 생성 AI를 나도 배워보고 싶다.	Taylor et al. (1995)	3.65 (.86)	.73
사용 의도	생성 AI를 방송 제작에 계속 활용할 계획이다. 타 제작진들에게도 생성 AI 사용을 권장할 생각이 있다.		3.44 (.86)	.86

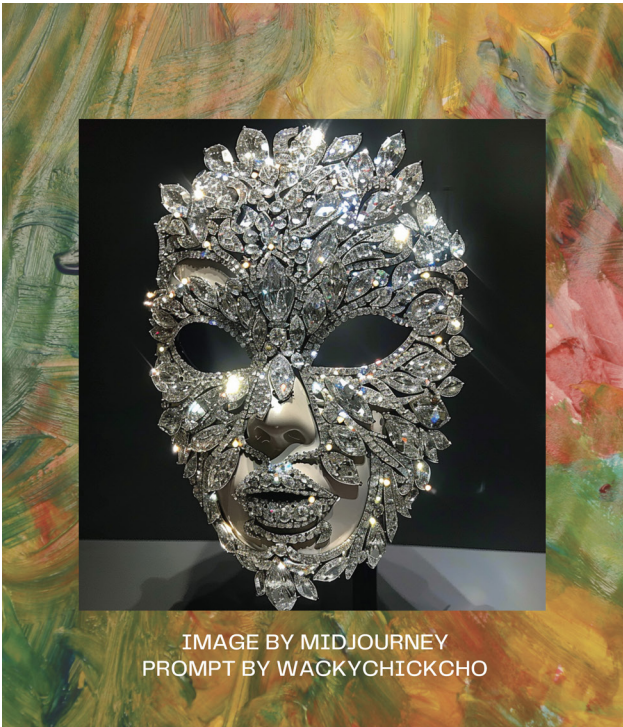
RESEARCH SUMMARY 연구 결과

연구 문제	연구 결과
1-1. 방송영상 제작에 있어 생성형 AI 사용 가능성 및 업무 효율성이 높을 것으로 기대되는 분야는 무엇인가?	<ul style="list-style-type: none"> ● 최다 사용 경험: 검색(55.7%) 문장작성(49.6%) 이미지 생성(42.7%) ● 최소 사용 경험: 녹음(10.4%) 의상-분장(7.8%) ● 최고 사용 가능성: 이미지 생성/검색/디지털 아바타/자막-그래픽/문장-비디오 생성 ● 최저 사용 가능성: 아이디어/ 의상-분장/녹음 ● 업무 효율성 증가 최고: 이미지 생성/검색/비디오 생성/디지털아바타/오디오 생성 ● 업무 효율성 증가 최저: 아이디어/의상-분장/미술
1-2. 생성형 AI 사용에 방송 제작진의 인식은 어떻게 분류되는가?	<ul style="list-style-type: none"> ● 긍정적 인식: 결과물 생성 즉시성/ 상대적 비용 합리성/ 이용 장소, 시간 편의성/ 사용자 능력 영향/ 비용 적정성/ 사용자 창작 주체감(설명력 26%) ● 부정적 인식:사용법 난이도/ 매뉴얼 학습 시간/ 저작권-개인정보 침해 우려(19%) ● 결과물 완성도: 결과물 완성도/ 기존 방식 생성물과 질적 차이 미미(16%)
2. 생성형 AI 사용에 대한 제작진 인식은 인지된 유용성, 인지된 용이성에 어떠한 영향을 미치는가? H1: 생성형 AI에 대한 인지된 유용성이 커질수록 사용 태도는 긍정적이 될 것이다. H2: 생성형 AI에 대한 인지된 용이성이 커질수록 이용자 사용태도는 긍정적이 될 것이다. H3: 생성형 AI에 대한 이용자 사용 태도가 긍정적일수록 사용 의사가 커질 것이다.	<ul style="list-style-type: none"> ● 가설 모두 채택됨 ● 제작진 인식 중 긍정적 요인만이 인지된 유용성에 영향 ● 제작진 인식 중 긍정적, 부정적 요인이 용이성에 영향



CONCLUSION

- 생성형 AI 앱은 빠른 속도로 발전 중
- 디지털 대비 TV 제작 활용도는 아직 미흡
- 제작 비용, 시간 면에서 압도적 장점
- 신규 버전 출시 빈번하여 문제점 개선도 높음
- 런칭 후 2년이 넘고 있지만 신체, 순간 표현은 아직도 완벽히 해결되고 있지 못함
- 저작권, 개인 정보 침해 등 자유로운 사용에 대해 장애물 존재 사실
- 기술 개발, 법적 장치 최소화 마련 필요



DISCUSSION

- TV 방송 제작 상용화를 위한 선결 과제?
- 기술 측면 최우선 해결 과제?
- 방송 심의 등 관련 정책 제정 상황?
- 관련 규정 및 법 제정의 방향은?
- 기술 상용화가 TV제작 산업에 미칠 영향은?

REFERENCE



- 이선중, <챗GPT 질문의 기술>. 서울:영진닷컴.
- 정원준, 정원준(2023). AI 커버곡 사례를 통해 본 생성형 AI의 법률 문제. <산업재산권>, 76호, 455-494.
- 조서윤, 이영주(2024), TV 방송 제작진의 생성형 AI 기술 수용에 영향을 미치는 요인 : 제작진 인식을 중심으로, 한국소통학보, 08/31/2024, Vol. 23, Issue 3, p. 277-310
- Variety(2023.12.1.), Generative AI in Film & TV: A Special Report

**생성형 AI는
미디어 생태계의
게임 체인저인가**



AI 광고의 사용자 경험 확장 연구

| 차혜영 IT디자인융합전공 박사과정, 광주대학교 산업디자인학과 초빙교수



AI 광고의 사용자 경험 확장 연구

A Study on the Expansion of User Experience in AI Advertising

차혜영 Cha, Hye Young

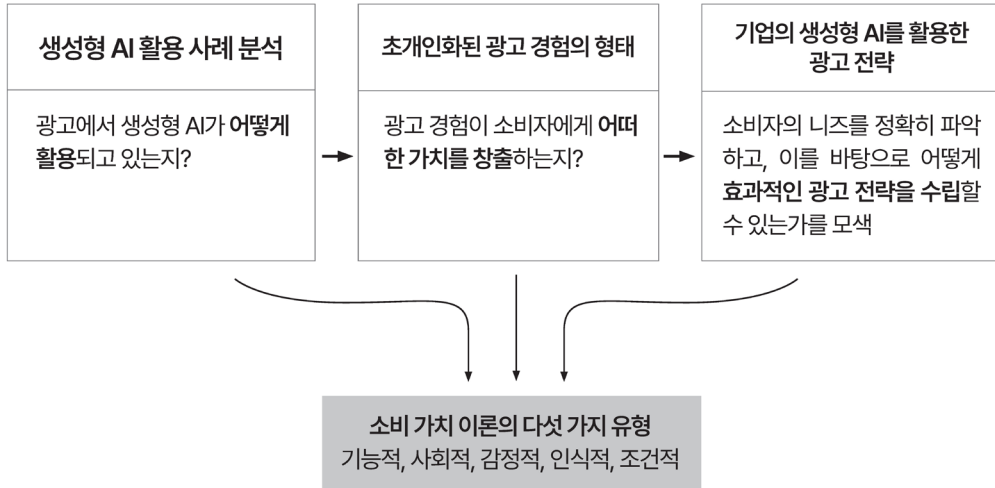
서울과학기술대학교 IT정책전문대학원
IT디자인융합전공 박사과정

1. 연구 배경 및 목적

- 생성형 AI의 발전과 광고 분야에서의 적용
- 사용자 경험 가치와 AI 광고의 효과 분석 필요성
- 소비 가치 이론 기반으로 AI 광고의 사용자 경험 확장 요소 탐구

소비자 중심의 마케팅 전략이 중요해진 현 시점에서, AI를 활용한 초개인화된 광고는 소비자의 니즈를 정확히 파악하고 이를 충족시키는 데 핵심적인 역할을 하고 있다. 이는 기업이 소비자와 더욱 효과적으로 소통하고, 소비자의 관심을 끌며, 궁극적으로는 브랜드 충성도를 높이는 데 기여할 수 있다. 그러나 생성형 AI가 제공하는 초개인화된 광고의 사용자 경험 가치에 대한 체계적인 연구는 아직 부족한 실정이다.

2. 연구 범위 및 방법



3. AI 광고 사례 분석

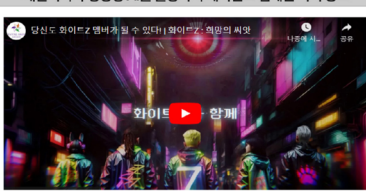
· 분석모델

소비 가치 유형	정의	평가 항목
기능적 가치	제품/서비스의 실용적, 물리적 성능 또는 효율성에서 느끼는 가치	광고가 소비자에게 실용적 효용(정확성, 효율성)을 제공하였는가?
사회적 가치	소비자가 특정 사회적 그룹과의 연관성을 통해 얻는 가치	특정 사회적 그룹이나 트렌드와의 연결을 강화했는가?
감정적 가치	광고가 소비자에게 긍정적인 감정을 유발하고 브랜드와 정서적 유대감을 형성하는 가치	광고가 소비자의 감정적 반응(유대감, 감동 등)을 자극하였는가?
인식적 가치	새로운 정보나 흥미로운 경험을 통해 소비자의 호기심을 자극하는 가치	독창적이고 새로운 경험을 제공하였는가?
조건적 가치	특정 조건이나 상황에서 광고가 제공하는 유용한 가치	광고가 특정 상황(시기, 장소, 맥락)에 맞게 적합하게 제공되었는가?

1991년 셰스(Sheth), 뉴먼(Newman), 그로스(Gross)가 개발한 소비 가치 이론(Theory of Consumption Values, TCV)을 본 연구에 맞게 정리

· 사례 분석 1

제일기획이 생성형 AI를 활용하여 에버랜드 캠페인 티저 광고



"이색 체험 공간 '블러드시티': 타임캐스트(다양한 연령대와 언어 표현 가능)를 활용한 보이스 적용
[출처: <https://typecast.ai/learn/how-to-utilize-generative-ai-to-create-an-ad/>]

적용 AI 기술	
영상 스토리 구성	챗GPT(Chat GPT)
영상 이미지	미드저니(Midjourney) 이미지 딥러닝(Deep Learning)
퍼인 튜닝	스테이블 디퓨전(Stable Diffusion)
배경음악	뮤버트(Mubert)
내레이션	타임캐스트(Typecast)

소비 가치 유형 분석	
기능적 가치	- 타임캐스트의 다양한 연령대와 언어 표현 기술을 활용해 정밀한 고객 맞춤형 메시지 전달 - 생성형 AI를 통해 비용과 시간 절감
사회적 가치	- 캠페인을 통해 에버랜드가 트렌디하고 혁신적인 브랜드로 인식 - 젊은 세대의 정체성 강화
감정적 가치	- 이색적인 체험 공간인 "블러드시티"를 통해 호기심과 감동 유발
인식적 가치	- 독특한 스토리텔링과 몰입감 있는 콘텐츠로 소비자의 흥미를 끌어냄
조건적 가치	- 특정 시즌을 겨냥하여 캠페인 적시성 극대화

가치 평가		
가치 유형	세부항목	점수
기능적 가치	광고가 소비자에게 실용적 효용을 제공하였는가?	5
사회적 가치	특정 사회적 그룹이나 트렌드와의 연결을 강화했는가?	4
감정적 가치	광고가 소비자의 감정적 반응을 자극하였는가?	2
인식적 가치	독창적이고 새로운 경험을 제공하였는가?	5
조건적 가치	광고가 특정 상황에 맞게 적절하게 제공되었는가?	5



※가치 평가 척도: 매우낮음(1) / 낮음(2) / 보통(3) / 높음(4) / 매우높음(5)

· 사례 분석 1 결과

사이버 펑크라는 미래 시점에서의 AI 활약에 대한 상상력과 더불어 현실감 있는 디스토피아적 세계관의 독창적인 스토리텔링과 몰입감, 정확한 타겟팅, 시기적 적합성을 고려한 콘텐츠

→ 기능적 가치와 인식적 가치, 조건적 가치는 높게 측정

낮선 세계관과 영상 이미지는 공감이나 감동을 이끌어 내기에는 무리

→ 감정적 가치는 낮게 분석

감정적 스토리텔링을 보강하여 소비자와의 정서적 연결을 더욱 강화할 필요가 있다.

▪ 사례 분석 2



기획에서 제작까지 전 과정을 AI와 함께 진행한 삼성생명의 AI로만 만든 최초 광고는 소비자의 반응과 요구사항을 LLM(Large Language Model : 대규모 언어모델)을 활용하고 미드저니와 달리로 제작
[출처: 삼성생명]

적용 AI 기술	
소비자 반응 및 요구사항	LLM(Large Language Model)
영상 이미지	미드저니(Midjourney) 이미지 딥러닝(Deep Learning)
소비 가치 유형 분석	
기능적 가치	- LLM(Large Language Model)을 활용해 소비자의 요구사항과 반응을 정밀하게 분석 - AI 기술만으로 광고 전 과정을 진행해 높은 효율성 달성
사회적 가치	- 브랜드의 기술력과 혁신성을 소비자에게 전달, 브랜드 신뢰도 상승
감정적 가치	- 노년층의 행복한 모습을 통해 감정적 연결을 강화
인식적 가치	- 새로운 광고 제작 접근 방식을 통해 AI 광고의 가능성을 제시
조건적 가치	- 광고가 노년층이라는 특정 소비자 계층의 요구에 정 확히 부합

가치 평가		
소비 가치 유형	세부항목	점수
기능적 가치	광고가 소비자에게 실용적 효용을 제공하였는가?	5
사회적 가치	특정 사회적 그룹이나 트렌드와의 연결을 강화했는가?	4
감정적 가치	광고가 소비자의 감정적 반응을 자극하였는가?	5
인식적 가치	독창적이고 새로운 경험을 제공하였는가?	2
조건적 가치	광고가 특정 상황에 맞게 적절하게 제공되었는가?	5



※가치 평가 척도 : 매우낮음(1) / 낮음(2) / 보통(3) / 높음(4) / 매우높음(5)

▪ 사례 분석 2 결과

LLM 기반 데이터 분석으로 정밀한 소비자 맞춤 광고, 감동적인 메시지 전달, 타겟에게 적합한 시기적 타이밍을 갖추고 제작되었다는 점

→ 기능적 가치, 감정적 가치, 조건적 가치는 높게 측정


광고 제작 방식은 새롭지만, 스토리 자체가 혁신적이거나 독창적이라는 인상은 부족한 편

→ 인식적 가치는 부족하다는 평가

AI 기술의 독창적 활용 방식을 더 명확히 강조하거나, 노년층의 미래 지향적 메시지를 통해 인식적 가치를 높일 수 있을 것이다.

· 사례 분석 3

KB라이프생명 라이프를 나눔답게 캠페인



윤여정 배우의 인생 전반에 생명보험의 가치를 전달한다는 광고 메시지를 설정하고, 딥러닝 기술을 사용해서 한 여배우 인생 여정을 보여줌. 2023대한민국광고대상에서 이노베이션 부문 금상을 수상
[출처: <https://www.secdaily.com/NewsView/239YE0IB63I>]

적용 AI 기술	
딥러닝 · 데이터링 기술	
소비 가치 유형 분석	
기능적 가치	- 딥러닝 기술을 사용해 배우 윤여정의 과거 이미지를 복원, 생밀한 콘텐츠 제공 - 높은 수준의 기술 활용으로 효율성 향상
사회적 가치	- 배우의 생애를 조명하며 감성적 메시지를 전달해 브랜드 신뢰도 상승
감정적 가치	- 과거 회상을 통해 감동적이고 따뜻한 감성을 전달
인식적 가치	- 독특한 스토리와 영상 제작 기법으로 소비자 관심 유발
조건적 가치	- 캠페인 타이밍과 연관성 강화로 소비자 반응 최적화

가치 평가		
소비 가치 유형	세부항목	점수
기능적 가치	광고가 소비자에게 실용적 효용을 제공하였는가?	5
사회적 가치	특정 사회적 그룹이나 트렌드와의 연결을 강화했는가?	4
감정적 가치	광고가 소비자의 감정적 반응을 자극하였는가?	4
인식적 가치	독창적이고 새로운 경험을 제공하였는가?	2
조건적 가치	광고가 특정 상황에 맞게 적절하게 제공되었는가?	5



※가치 평가 척도 : 매우낮음(1) / 낮음(2) / 보통(3) / 높음(4) / 매우높음(5)

· 사례 분석 3 결과

배우 윤여정의 삶을 조명하며 깊은 감동과 생애 전환기와 잘 맞는 타이밍의 메시지를 전달한다는 점

→ 기능적 가치, 조건적 가치는 높이 평가

감동적인 광고 스토리에 비해 AI 기술의 혁신성을 강조하는 부분이 부족

→ 인식적 가치는 낮게 평가

AI 기술 활용의 구체적인 과정을 부각하여 기술적 신뢰도를 강화하고, 더 폭넓은 대상층으로 확장할 필요가 있다.

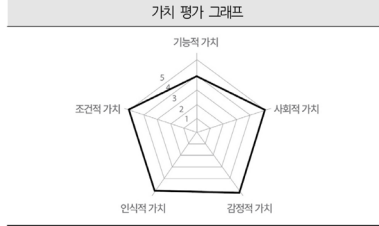
▪ 사례 분석 4



배스킨라빈스 북송이 원정대와 용의 눈물
MZ세대가 사랑하는 인기 캐릭터 기업 '산리오'의 '산리오캐릭터즈'와 협업으로 탄생한 배스킨라빈스 아이들의 맛을 색다르고 흥미롭게 알려고자 동화 초안을 '챗 GPT'에 요청하고 이를 각색하여 광고 영상을 제작
[출처 : <https://www.spcmagazine.com/spc-배스킨라빈스>]

적용 AI 기술	
챗(Chat) GPT	
소비 가치 유형 분석	
기능적 가치	- ChatGPT를 활용해 흥미로운 알레고리 스토리 구성 및 광고 제작 - 효율적 콘텐츠 제작
사회적 가치	- MZ세대의 인기 캐릭터와 협업하여 특정 세대와의 연결 강화
감정적 가치	- 귀여운 캐릭터와 이야기를 통해 소비자의 감정을 자극
인식적 가치	- 새로운 캐릭터와 독특한 스토리로 주목받음
조건적 가치	- 특정 시즌 및 신제품 출시와 연계된 광고 진행

가치 평가		
소비 가치 유형	세부항목	점수
기능적 가치	광고가 소비자에게 실용적 효용을 제공하였는가?	4
사회적 가치	특정 사회적 그룹이나 트렌드와의 연결을 강화했는가?	5
감정적 가치	광고가 소비자의 감정적 반응을 자극하였는가?	5
인식적 가치	독창적이고 새로운 경험을 제공하였는가?	5
조건적 가치	광고가 특정 상황에 맞게 적절하게 제공되었는가?	5



※가치 평가 척도 : 매우낮음(1) / 낮음(2) / 보통(3) / 높음(4) / 매우높음(5)

▪ 사례 분석 4 결과

MZ세대의 인기 캐릭터와 협업으로 세대 간 연결을 강화하고, 귀여운 캐릭터와 이야기를 통해 감정적 유대감을 형성하여 신제품 출시와 연계된 시즌 캠페인을 진행
→ 사회적 가치, 감정적 가치, 조건적 가치는 높은 평가

AI 기술 활용의 실용적 측면(효율성, 데이터 활용 등)이 구체적으로 드러나지 않음
→ 기능적 가치의 평가가 비교적 낮게 나타남

AI의 구체적인 기여도를 명시하여 기술적 신뢰를 강화하고, 더 많은 소비자 참여를 유도할 수 있는 캠페인 활동을 추가한다면 기능적 가치 역시 높아질 것이다.

4. 결론 및 제언

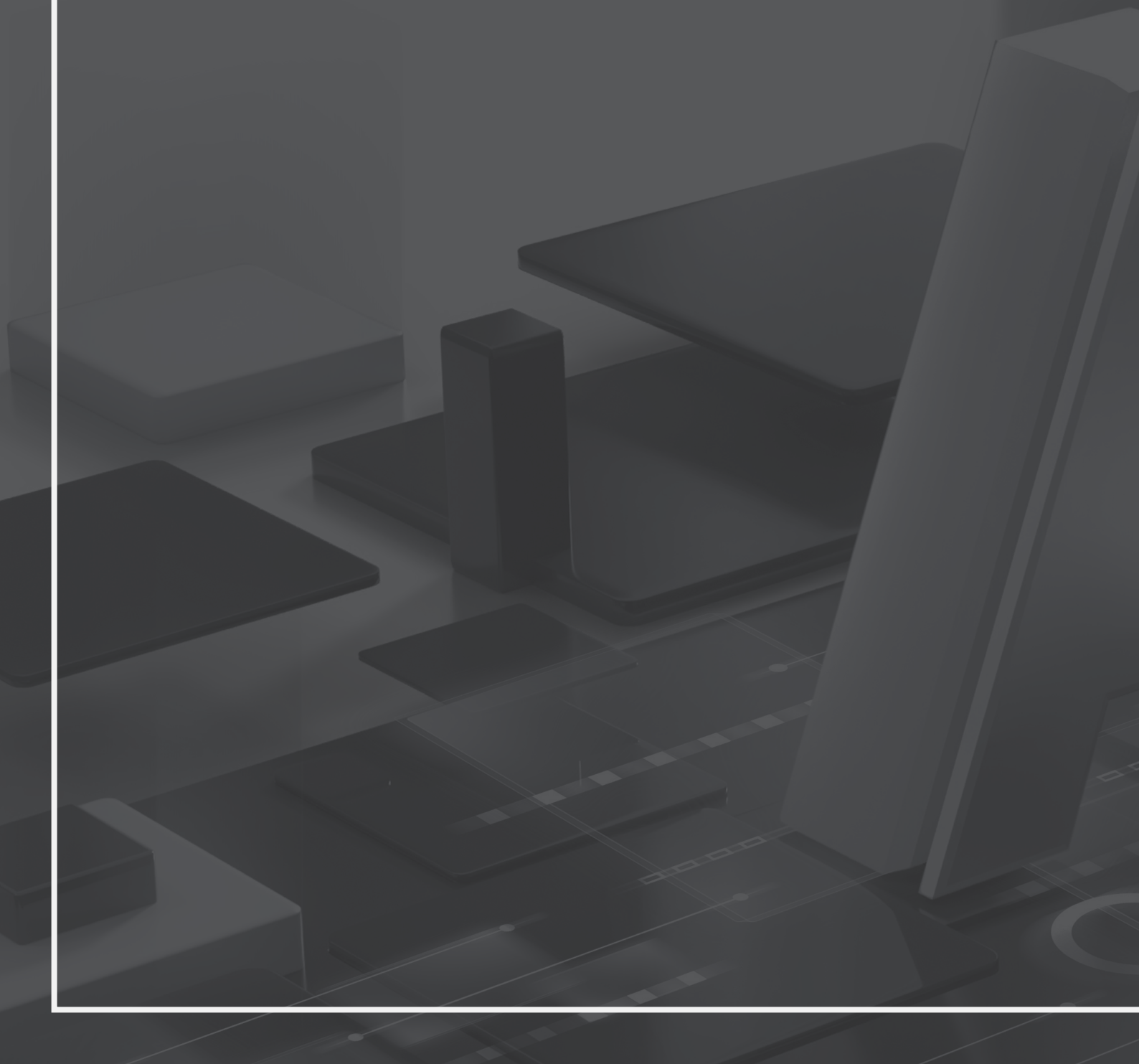
- AI 광고는 효율성과 정밀성을 강화하여 기능적 가치를 높임
- 사회적 가치와 감정적 유대감을 강화할 필요성
- 지속 가능한 메시지와 독창적 접근 방식의 중요성
- 개인정보 보호 및 윤리적 문제 해결을 통한 신뢰 확보

사용자 경험 확장을 위한 경험 가치의 중점을 어디에 둘 것인가를 심도 있게 고민

광고 콘텐츠의 구성, 스토리텔링, 타겟팅, 마케팅 등 경험 가치 요소를 반영한 전략적 접근을 통해 체계적으로 AI 광고를 제작할 필요

감사합니다.

**생성형 AI는
미디어 생태계의
게임 체인저인가**



제2세션

발표 1.

인공지능 생성물의 투명성 규범 방안

| 박명순 융합미디어콘텐츠정책전공 박사수료

발표 2.

이미지생성형 인공지능이 그리는 기괴한 행복의 세계

| 오경미 디지털문화정책전공 박사수료, 오픈넷 연구원

**생성형 AI는
미디어 생태계의
게임 체인저인가**



인공지능 생성물의 투명성 규범 방안

| 박명순 융합미디어콘텐츠정책전공 박사수료



IT 정책전문대학원 2025년 AI in Society “인공지능 생성물의 투명성 규범 방안”

융합콘텐츠미디어정책전공

박명순 박사수료

 서울과학기술대학교 SEOUL NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

목 차

Chapter 1. AI 생성 기술 발전 현황

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

- ↳ AI 생성물 유형 및 특징 * audio 중심
- ↳ 규범 필요성 및 주요 내용
- ↳ 규범 쟁점과 개선방향으로의 자율규제

Chapter 3. AI 생성물 ‘표시 의무’

- ↳ 한계와 개선 방안

Chapter 4. AI 생성물 ‘학습데이터 설명의무’

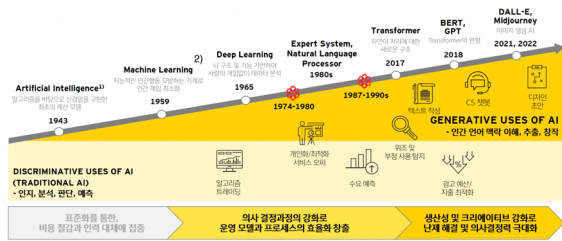
- ↳ 한계와 개선 방안

Chapter 5. 하고 싶은 말

 서울과학기술대학교 SEOUL NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

Chapter 1. AI 생성 기술 발전 현황

[AI 모델 진화 과정]

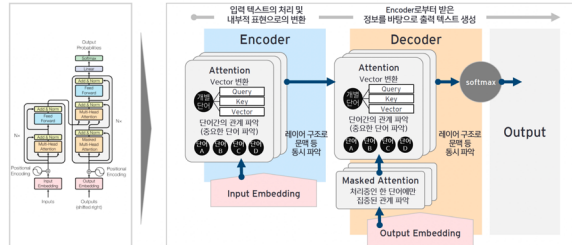


* 이미지 출처: EY, What's New in Artificial Intelligence Gartner Hype Cycle

출처

- 1) 1943년 신경작용에 내재한 개념의 논리적 해석이라는 논문에서 처음 등장
McCulloch, W.S., Pitts, W. "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics 5, 1943
- 2) A. M. TURING, "I.—COMPUTING MACHINERY AND INTELLIGENCE", Mind, Volume LIX, Issue 236, October 1950

[Transformer Architecture]

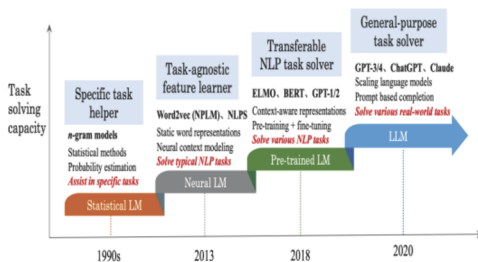


출처: Ashish Vaswaney et al., "Attention Is All You Need", Advances in Neural Information Processing Systems(NIPS), 2017

-> 대량의 데이터를 미리 학습한 모델에 따라 등장할 수 있는 단어 후보 중 가장 높은 확률값 순차적 출력

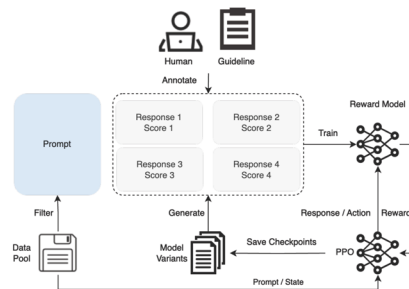
Chapter 1. AI 생성 기술 발전 현황

[Language models의 4세대 진화과정]



출처: Zhao, Wayne Xin, et al. "A survey of large language models.", arXiv preprint arXiv:2303.18223 (2023), p. 2.

[인간 피드백 기반 강화학습 (Reinforcement Learning with Human Feedback)]



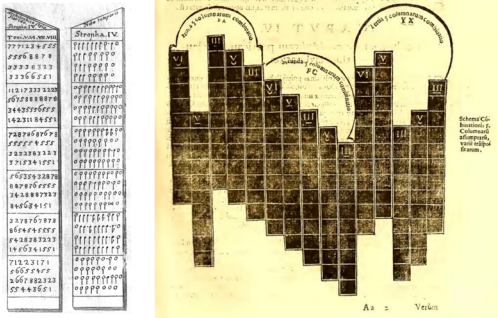
출처: Yang, Aiyuan, et al. "Baichuan 2: Open large-scale language models.", arXiv preprint arXiv:2309.10305 (2023), p. 7.

-> 2017년 구글 Transformer 기술 발표 시점부터 데이터 학습 방식과 매개 변수 폭발적으로 증가
GPT-3 부터 Hallucination과 편향성 보정을 위한 RLHF로 개선 중

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

AI 생성물의 제작 원리 및 특징 - 오디오 생성물 [바로크 ~ 고전주의 시대]

[아타나시우스 키르허의 저서 *Musurgia Universalis*]
- 부분적 기계화 시점은 1,650년부터



출처

Athanasius Kircher, *Musurgia Universalis*, Vol. II, p.185.

[모차르트의 작품 “음악의 주사위 놀이”], 1787년



Mozart K. 516F

주사위의 숫자와 마디를 조합해 곡을 만드는 방법

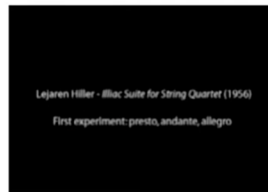
당시, 초기 기계식 작곡 모델은 과거 유명한 곡들을 규칙화하고 이를 유사하게 모방

-> 오디오 창작물을 기계적으로 자동화하려는 시도는 과거 바로크 시대부터 시작
(20세기 말까지 ‘조합’ 접근법은 수학적 양적 축적에 의존)

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

AI 생성물의 제작 원리 및 특징 - 오디오 생성물 [낭만주의를 넘어 20세기 컴퓨터의 등장]

[Iliac Suite] Markov model 1956년



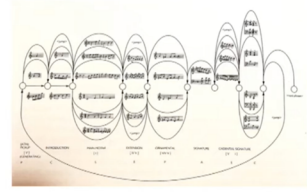
컴퓨터 기반 최초로 자동 작곡한 곡(4개의 현악기로 작곡).
숫자의 나열 시퀀스를 마르코프 모델로 작곡함

출처

Iliac Suite

<<https://www.musicainformatica.it/argomenti/iliac-suite/>>

[David Cope] Rule-based AI, 1980년대



데이비드 코프가 EM이라고 해서 스타일 모방
기존 쇼팽이나 작곡가의 곡들을 하나하나씩 조각으로 나누고
조각들 간의 음악적인 관계 규칙을 학습 모델을 통해서하는게 아니라
음악적인 지식과 당시 규칙기반 인공지능 기술을 이용해서 하나씩 이어붙임
스타일은 비슷하게 생성이 되는 모델

출처

David Cope

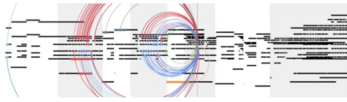
<<https://liveinnovation.org/david-cope-a-lifetime-contribution-to-artificial-intelligence-and-music/1995-david-cope/>>

-> 20세기 말까지 ‘조합’ 접근법에서 컴퓨터의 등장으로 ‘규칙’기반 어울리는 것관의 관계를 중요하게 보기 시작

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

AI 생성물의 제작 원리 및 특징 - 오디오 생성물 [현대음악을 넘어 AI로]

[Google Music Transformer] AI 모델, 2018년



알고리즘상으로는 트랜스포머라는 긴 시퀀스 나열을 효과적으로 학습할 수 있는 모델로 이 모델이 음악으로 악보를 연주한 피아노 연주의 MIDI 파일(음악의 반복구조, 메인 테마 음악이 다양하게 소위 셸프 유사도를 통해 학습을 하고 난 데이터로) 랜덤값을 보면 제리피티해보고 나름 여러 결과물중에 좋은 것을 편집하여 생성

출처
magenta, <<https://magenta.tensorflow.org/music-transformer>>

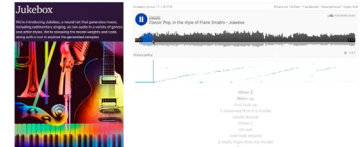
[OpenAI MuseNet] AI 모델, 2019년



오픈AI에서 큰 데이터를 갖고 백본 아키텍처는 gpt2를 활용해서 다양한 미디어데이터를 심볼릭한 데이터를 활용함 조건을 주면 무작위가 아니라, 조건에 맞게 생성할 수 있음
테마 멜로디를 고르고 멜로디를 누구의 스타일로 할 수 있는 지 선택할 수 있음

출처
OpenAI Musenet
<<https://openai.com/index/musenet/>>

[OpenAI Jukebox] AI 모델, 2020년



음표레벨이 아니라 오디오 파형을 직접 생성할 수 있는 기술로 소위 불쾌한 골자기 경험한 결과물이다.

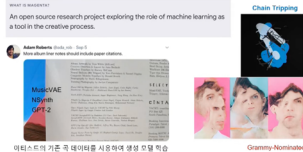
출처
OpenAI Jukebox
<<https://jukebox.openai.com/>>

-> '조합' -> 컴퓨터의 등장으로 '규칙'기반 -> '데이터 학습' & 효과적인 시퀀스 나열 '오디오 파형 생성'

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

AI 생성물의 제작 원리 및 특징 - 오디오 생성물 [현대음악을 넘어 AI로]

[Magenta + YACHT], 2019년



AI모델과 아티스트와의 협력 사례
GPT가 하고 멜로디는 뮤직스 AI, 학습 데이터는 창작자 본인

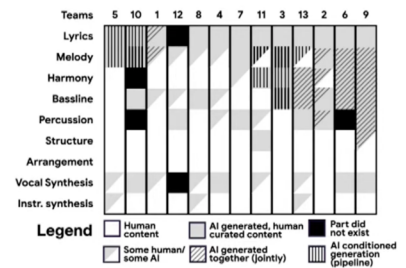
출처
YACHT + Magenta
<<https://magenta.tensorflow.org/chain-tripping>>

[AI Song Contest], 2020년



여러 음악 요소별로 생성모델을 다른것을 활용한 것
화성 디자인, 드럼, 구조 음성 합성 등 모두 다른 모델로 협업함
총 13팀이 참가해서 영역별로 어떻게 활용했는지 등 여러 케이스를 두고 비교를 했던 결과가 있다.
가사 멜로디 화성 (심볼릭은 AI에 의존하거나 협업)
어렵고 구조적인 편곡부분은 사람이 마지막으로 터치하는 결과물로 볼 수 있음

출처
AI song contest 2024, <<https://www.aisongcontest.com/>>

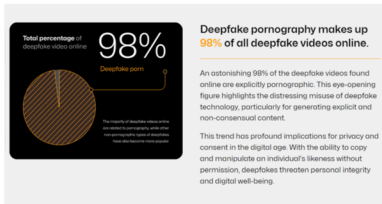


여러 학습모델을 사용해서 대회를 열었음

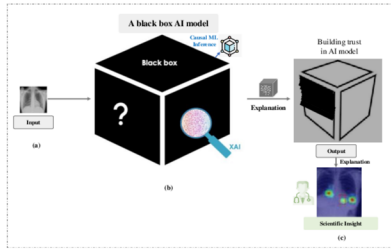
-> AI 모델을 활용 방법이 음악 예술 창작 방법으로 녹여들고 있는 중
그러나, 여전히 아티스트 허락없이 사용되는 학습데이터, 저작권 이슈 등 해결해야할 과제를 투명성 규범으로...

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

- AI 생성물의 투명성 규범의 목적은? 오용 딥페이크, AI 생산 과정의 불투명성은 사회적 신뢰도 저하에 심각한 영향
 - ‘AI발전’과 ‘민주주의적 질서 조화 필요’
- 본래 투명성 규범의 의의는 국민의 알 권리를 보장하기 위해 정부의 공개 의무를 부여하는 데 초점, ‘민간 영역 강제 예외’ 특히, 인공지능의 투명성 이슈는 사회적 합의가 어렵고 갈등을 유발할 여지 큼
 - ↳ 투명성 규범의 범위는 **국민의 생명, 신체, 재산의 안전과 직접, 밀접하게 관련있는 경우에 한해**
 + **민간 기업에 영업과 관련 투명성 의무 요구**



온라인 딥페이크 영상 중 딥페이크 포르노의 비율
(출처: Security Hero)



출처: R. Qureshi et al., "Artificial Intelligence and Biosensors in Healthcare and Its Clinical Relevance: A Review," in IEEE Access, vol. 11, pp. 61612,

Classic output:	Genuine Concord lyrics:
I used to bite my tongue and hold my breath Scared to rock the boat and make a mess So I sat quietly, agreed politely I guess that's forget I had a choice I let you push me past the breaking point I closed for writing, and I fell for everything	I used to bite my tongue and hold my breath Scared to rock the boat and make a mess So I sat quietly, agreed politely I guess that's forget I had a choice I let you push me past the breaking point I closed for writing, and I fell for everything
You held me down, but I got up Already breaking off the dust You hear my voice, you hear that sound Like thunder, gonna shake the ground You held me down, but I got up (hey) Get ready, cause I'm had enough I see it all, I see it now	You held me down, but I got up (hey) Already breaking off the dust You hear my voice, you hear that sound Like thunder, gonna shake the ground You held me down, but I got up (hey) Get ready, cause I'm had enough I see it all, I see it now
I got the eye of the tiger, a fighter, dancing through the fire 'Cause I am a champion, and you're gonna hear me roar Louder, louder than a lion 'Cause I am a champion, and you're gonna hear me roar	I got the eye of the tiger, a fighter Dancing through the fire 'Cause I am a champion, and you're gonna hear me roar Louder, louder than a lion 'Cause I am a champion, and you're gonna hear me roar
Oh oh oh oh oh oh oh Oh oh oh oh oh oh oh Oh oh oh oh oh oh oh You're gonna hear me roar	Oh oh oh oh oh oh oh Oh oh oh oh oh oh oh Oh oh oh oh oh oh oh You're gonna hear me roar
Now I'm flexin' like a butterfly Stripping like a bee, I amed my stripes I wear them across, to my own hero You held me down, but I got up Already breaking off the dust	Now I'm flexin' like a butterfly Stripping like a bee, I amed my stripes I wear them across, to my own hero You held me down, but I got up (hey) Already breaking off the dust

Claude(좌), Kary Perry 원곡('CMG', 우)과 가사 비교
출처: MUSIC-Z SONGS; and ABKCO MUSIC, INC.
v. ANTHROPIC PBC (3:23-cv-01092),

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

- AI 발전 측면
 - LLM 연구 패러다임의 확장, 콘텐츠(어문, 이미지, 오디오, 영상 등) 생성을 넘어 3rd Party 애플리케이션 생태계 확장 중
 - AI 기업의 부담, 학습데이터 투명한 제공과 영업 비밀 정보 경계 모호해 기업에 추가 비용 부담 초래
 - 생성형 AI 시장 성장, GACR 35.6%로 약 1,094억 달러에 이를 것으로 예상되며, 이와 관련된 AI Value Chain은 컴퓨터, 하드웨어, 클라우드 플랫폼, 파운데이션 모델, 모델 허브, 애플리케이션, 서비스라는 6개 요소 구성



출처: McKinsey & Company, Generative AI value chain
QuantumBlack AI by McKinsey, Exploring opportunities in the generative AI value chain, p. 1.

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

□ 민주주의적 질서 측면

- 딥페이크와 가짜뉴스, 실재와 구별하기 어렵고 가짜뉴스, 선거, 전쟁 전략등에 활용은 심각성 초래 예, ‘공직선거법’ 등 기존 법률로는 이러한 문제 효과적 규제 한계
- 책임 소재의 모호성, 규제 대상자인 공급자, 사용자, 개발자 등에 책임을 묻는 방식 모호
- AI 시스템의 블랙박스(불투명) 특성으로 편향성 논란 극복의 어려움
- 학습 데이터 권리자의 지재권 보호 필요 * 다수 소송 제기

☞ AI발전과 민주주의적 질서의 조화를 위해, 사회적 신뢰를 기반으로 하는 공공 규범으로서의 투명성 규제가 필요하다. 이를 위해 **규범적 관점에서의 요구를 넘어 기술적 해결책이 뒷받침되어야**함

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

국가	주요 규범	투명성 규범 관련 내용	시행 시기	목적
유럽 연합 (EU)	EU AI 법	- 제 50조 2항 기계가 읽을 수 있는 형태로 표시 - 제 53조 1항d; AI 사무국에서 제공하는 템플릿에 따라 GPAI 훈련에 사용하는 콘텐츠에 대한 충분히 자세한 내용을 작성하고 설명	2024.8.1 발효, 2026년 전면 시행 예정	AI의 신뢰성 확보 및 글로벌 리더십 강화
미국	행정명령 14110 및 캘리포니아주 AI 투명성법 SB942	- 14110: 생성형 AI 생성물에 워터마크 또는 라벨 부착 요구 - SB 942(AI 투명성법)는 월간 이용자 100만명 이상인 AI시스템 개발자에게 AI콘텐츠 탐지 도구와 AI워터마크 표시 요구	2023. 10. 30발 표(행정명령)	AI 혁신 촉진과 국가 안보 강화
중국	생성형 인공지능 서비스 관리 규정	- 제 17조, 규제 당국이 제공자의 학습 데이터와 알고리즘 정보 제출을 요청할 경우 이에 응해야 하는 의무 - 제 12조, 생성형 AI 콘텐츠에 대해 이미지와 동영상 출처를 명확히 표시해 허위 정보의 유통을 방지하도록 규정	2023. 7. 13 시행	사회주의 가치 존중 및 허위 정보 유통 방지
한국	AI 기본법	- 22대 국회에서 발의된 법률안 19건	2026년 전면 시행 예정	AI 기술 발전과 윤리적 사용 촉진

☞ 따라서, 투명성 규범은 '기술 혁신'과 '산업 생태계와의 균형'을 이뤄야 한다. 과도한 규제는 혁신을 저해할 수 있고, AI 기업이 자율적으로 투명성을 높일 수 있는 정책 고려 필요

의연명 (대표발의 의원, 제안일)	투명성 관련 주요 내용
인공지능 산업 육성 및 신뢰확보에 관한 법률안 (안철수의원 등 12인, '24.5.31)	생성형 인공지능에 의한 생성 사실을 결과물에 표시하여야 함(제29조)
인공지능 발전과 신뢰 기반 조성 등에 관한 법률안 (정동영의원 등 108인, '24.6.17)	생성형 인공지능 사업자가 생성형 인공지능 운용 사실 고지 및 표시하도록 함(제29조 및 제30조)
인공지능 산업 육성 및 신뢰 확보에 관한 법률안 (조인철의원 등 19인, '24.6.19)	학습데이터의 개요 등에 대한 설명(제27조)
인공지능산업 육성 및 신뢰 확보에 관한 법률안 (권성현의원 등 11인, '24.6.19)	고위험 영역에서 활용되는 인공지능의 고지 의무(제27조)
인공지능 기술 기본법안 (민형배의원 등 13인, '24.6.28)	학습데이터에 대한 설명 방안(제29조) 생성형 인공지능에 의해 생성 사실 표시(제30조)
인공지능 개발 및 이용 등에 관한 법률안 (권승희의원 등 15인, '24.7.4)	고위험 인공지능에 기반해 운용 사실을 이용자에게 고지 의무(제23조)
인공지능 기본법안 (한인수의원 등 10인, '24.8.22)	학습데이터에 대한 설명 방안(제27조) 생성형 인공지능에 의해 생성되었다는 사실 표시(제28조)
인공지능 책임 법안 (황희의원 등 10인, '24.8.27)	고위험인공지능이용자업자는 이용자에게 사전에 고위험인공지능을 이용한 서비스 제공 사실 알리아함(제20조)
인공지능 발전 진흥과 사회적 책임에 관한 법률안 (배준영의원 등 10인, '24.8.28)	학습데이터에 대한 설명 방안(제26조) 생성형 인공지능에 의해 생성된 결과물 표시(제27조)
인공지능의 발전과 안전성 확보 등에 관한 법률안 (이훈기의원 등 14인, '24.9.12)	학습데이터에 대한 설명 방안(제31조)
인공지능 산업 진흥 및 신뢰 확보 등에 관한 특별 법률안 (김우영의원 등 19인, '24.9.24)	고위험 인공지능에 의해 운용사실 고지(제27조) 학습데이터에 대한 설명 방안(제28조)
인공지능 산업 육성 및 신뢰 기반 조성 등에 관한 법률안 (이정현의원 등 13인, '24.10.29)	고위험 인공지능에 대한 설명 방안(제28조)
인공지능 발전 및 신뢰성 보장을 위한 법률안 (홍정익의원 등 10인, '24.11.1)	인공지능 기술 발전 중심 내용(투명성 관련 부제)
인공지능 산업 진흥 및 인공지능 이용 등에 관한 법률안 (이해민의원 등 16인, '24.11.11)	생성형 인공지능에 의해 생성된 제품 또는 서비스의 결과물 표시(제30조)
인공지능산업진흥에 관한 법률안 (정동영의원 등 14인, '24.11.11)	고위험 영역에서 활용되는 인공지능의 고지 의무(제31조)
인공지능 산업 육성 및 신뢰 기반 조성 등에 관한 법률안 (최민희의원 등 11인, '24.11.14)	고위험 영역에서 활용되는 인공지능의 고지 의무(제30조)
인공지능 안전 및 신뢰 기반 조성 등에 관한 법률안 (조순래의원·이인선의원 등 11인, '24.11.18)	생성형 인공지능에 의해 생성된 제품 또는 서비스의 결과물 표시(제17조)
인공지능 진흥에 관한 법률안 (조순래의원·이인선의원 등 11인, '24.11.18)	인공지능 기술 발전 중심 내용(투명성 관련 부제)
인공지능 산업 육성 및 발전 등에 관한 법률안 (정희용의원 등 16인, '24.11.18)	고위험 영역에서 활용되는 인공지능의 고지 의무(제26조)
인공지능 발전 진흥 및 신뢰 기반 조성 등에 관한 법률안 (최민희의원 등 11인, '24.11.18)	학습데이터에 대한 설명 방안(제27조)

Chapter 3. AI 생성물 표시 의무에 대한 한계와 합리화 방안

□ AI 생성물 표시의무

○ 쟁점

- 1) 기준 불명확성; AI 생성물의 출처 표시 방식에 대한 명확한 기준 부재
 - 2) 생성물 유형별 가이드라인 부재; 생성물 종류에 따른 구체적인 출처 표시 방법의 모호
 - > '콘텐츠 산업 진흥법 일부 개정 법률안' 인공지능 법안에도 내용 포함되어있는 바, 생성형 AI 고지 및 표시 의무가 필요하다는 논의 적극 개진되는 중
 - > 추후 출처 표기 방법 실행을 위한 기술적 해결책, 규범 방향 그리고 표준화 방안 등이 명확히 필요한 상황
- 검토 결과, '법적 의무의 이행은 기술적 불명확성'으로 그 '이행 가능 여부 모호'

○ 한계 분석

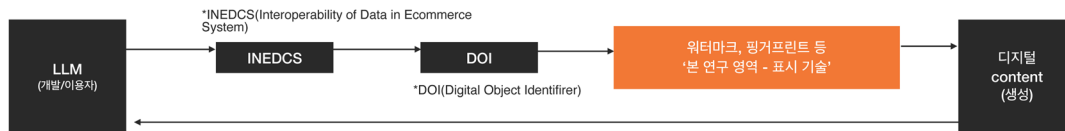
- * 기술적) AI 생성물에 실효성에 대한 검증 부족, 약간의 편집에도 쉽게 손상되고 표시의무 적용 방법이 생성물마다 다름
 - * 관리적) 표준화 가이드라인 부재, 식별 및 감지 체계 대응 방안 부재
 - * 법적 책임) 표시 기술 의도적 제거 훼손 공격으로 기술적 보호조치 무력화 등 저작권 침해 문제 발생 가능
- 인간이 표절 책임을 지지 않기 위해 고의적 AI 생성한 것으로 포장 가능
 결국 이용자 및 권리자가 최대 피해자
 (적합 콘텐츠가 누가 창작한 것인가? 알수없음)
 이를 뒷받침할 '더 현실적이고 효과적인 규범, 합리적 대안 필요'

표시 방식	위터마크 삽입 이미지	특징
눈에 보이는 방식 (Visible)		- 생성 사용자가 쉽게 인식 가능 (AI 생성물 인식 직관적) - 단점: 시각적 품질 저하 및 감상 방해 - 편집 및 제거 용이
눈에 보이지 않는 방식 (Invisible)		- 생성 생성물 품질 손상시키지 않고 출처 표기 가능 - 단점: 사용자에게 즉각적인 AI 생성물 인식 못함 - 특수 도구를 사용해서 워터마크 검출 가능

AI 생성물 유형	위터마크 삽입 이미지	주요 특징 및 방법
(1) 이문 생성물		- 기존 텍스트에 표시 : Format-based : Lexical-based : Syntactic-based - Generation-based - LLM을 위한 표시 : 혼란단계 : Logits 생성단계 : 토큰 레벨링 단계
(2) 이미지 & 영상 생성물		- 원본 이미지와 거의 동일한 새로운 이미지 생성 후 육안으로 식별할 수 없는 패턴을 삽입 - 동영상 프레임의 레벨에 워터마크를 직접 삽입
(3) 오디오 생성물		- (생성) 사람 귀에 들을 수 없는 비가청 주파수 영역에 워터마크 생성 - (식별) 딥 학습에서 전체 오디오 클립에 워터마크 프로그램 여부 판단 확인 - 출처: Meta Audio Seal152

Chapter 3. AI 생성물 표시 의무에 대한 한계와 합리화 방안

□ 디지털 콘텐츠 보호 기술 구성



표시 기술 구분	설명	한계	현황 (대표예)
위터마크	디지털 파일에 사람이 인지하기 어려운 형태로 식별 정보를 삽입해서 출처를 표시 기술	강한 편집이나 변형 시 제거될 수 있음	<구글> SynthID - 이미지/영상 <META> AudioSeal - 오디오
AI 라벨링 (Made with AI)	AI 생성물에 'AI로 제작됨' 등의 꼬리표 부착하여 이용자에게 알리는 기술	시각적 품질 저하 가능성 및 편집 시 제거 용이, 오입력 가능성 있음	<구글> '23.11. 부터 적용중 <어도비, 오픈시 등> C2PA로 라벨링 (콘텐츠 설명란에 표기)
메타데이터	출처, 제작자, 제작일, 위치 등의 정보 삽입	이미지 등 다른 사이트에 업로드하거나 스크린 샷으로 저장되면 제거될 수 있다.	<구글> '24.07 C2PA(콘텐츠 출처와 신뢰성 연합)
핑거프린트	저작물 자체가 지닌 고유한 특성(강인한 해시, 핑거프린트)을 추출해서 DB화하고 이를 기반 검색하는 기술	서로 다른 생성물에 대해 오인식이 적고, 위터마크 대비 차별성이 있음 (대규모 DB 구축비, 아직 개발 초기 단계)	<구글 유튜브> ContentID
프로브넌스 (Provenance)	콘텐츠의 생성, 성, 배포 과정을 추적 출처와 변화를 기술하는 기술	복잡한 추적 시스템 구축 필요 모든 단계에서 표준화된 협업 구조 필요	<어도비> CAI를 통해 출처와 이력추적

□ 표시 기술은 공격 우회 방법 다양하여 '여전히 취약'함, AI 기업입장에서는 투자 매몰비 최소화와 비즈니스 유지를 위해 '연관사 간의 커뮤니티', '자발적 개발' 추진 중
 최근 예.) '24.02. 윈윈 안보회의에서 기술 기업들이 발표하고 서명한 '24년 선거에서 AI의 기만적 사용 방지하기 위한 기술 협약 (예방, 출처명시, 탐지, 신속대응, 평가, 공공 인식 제고, 회복력 강화의 7가지)

Chapter 3. AI 생성물 표시 의무에 대한 한계와 합리화 방안

1) 국내외 AI 워터마크 법제화 및 자율규제 현황

구분	내용	자율규제 유형 ²⁶³⁾
G7	'23.10. AI 국제 행동 강령 합의, 11개 조항 중 워터마크 통해 AI창작물 구별 내용 담김	자발적 (voluntary)
미국	'23.10. AI 제작물을 구분할 수 있는 기술 개발에 대한 행정명령 발표	명령적 (mandatory), 자발적 (voluntary)
EU	'23.06 DSA(디지털 서비스법)로 AI생성물 워터마크 의무화 '24.06 AI법 최초 시행, 연말부터 AI 생성물에 탐지 및 추적할 수 있는 기술 적용 의무화	강제적 (coerced), 명령적 (mandatory)
국내	'24.11. 과방위는 AI 기본법 통과 및 관련 대표 내용으로 제29조(생성형 인공지능 고지 및 표시) 제31조(민간자율인공지능윤리위원회의 설치 등) TTA(한국정보통신기술협회) 주도로 '24년말 AI워터마크 가이드라인 발표예정	명령적 (mandatory), 제재적 (sanctioned)
대표기업	Google, OpenAI, Meta, Microsoft 등 '24.02 독일 원천안보회의(MSC)에서 딥페이크 차단 합의, 라벨링 (워터마크) 조치 취하기로 합의.	자발적 (voluntary)

출처
머니투데이, AI콘텐츠 고리표 붙이기 움직임, 정부 연말 가이드라인 낸다,
<<https://news.mt.co.kr/mtview.php?no=2024081909112510749>>.

2) 표시 의무 식별 및 표준화 현황

- Truepic은 C2PA 표준을 따르는 메타데이터를 생성물에 포함해 출처를 명확히 하여 이미지의 신뢰성을 높이고 있다
- 글레이즈(Glaze)는 자신의 스타일에 따른 이미지 생성을 방해하는 솔루션
- 포토가드(Photoguard)는 이미지의 중심이 되는 픽셀을 왜곡하여 이미지 편집을 시도할 때 오류를 발생시켜 이미지 조작을 방지하고, 이미지의 진위성을 보호하는 솔루션
- 나이트셰이드(NightShade)는 저작권자가 자신의 저작물이 학습에 사용되는 것을 원하지 않을 경우 이를 크롤링 해서 학습데이터로 사용하면 모델이 피해입게 만드는 솔루션
- Fawkes는 언골 인식 시스템으로부터 사용자의 이미지를 학습하지 못하도록 보호하는 솔루션
- 2024년 7월 네이버는 AI 생성물의 워터마크 기술 표준 구축을 하기 위해서 C2PA를 국내 최초로 가입하였고, AI가 생성한 코드 위에 워터마크를 삽입하는 내용을 제안
- 카카오도 불법 합성물로 의심되는 콘텐츠를 실시간으로 탐지할 수 있는 기술을 개발하여 배포 및 제공 행위 적발 시 카카오톡 전체 서비스 영구 제한하는 조치 중

3) 표시 방식 표준화 방안

- 첫째, AI 생성물은 변형이나 제거 공격에도 표시 방식이 훼손되지 않고 견고성을 보장할 수 있는 기술이 필요하다
- 둘째, 모든 AI 생성물에 일관되게 적용할 수 있는 생성물 유형별로 범용적이고 표준화된 워터마크 기술이 필요하다
- 셋째로, 표시 방법의 표준화 방안의 효과가 제한적인 이유는 기업이 해당 조치에서 경제적 이득을 얻기 어렵다는 산업 구조에 기인한다

-> 이러한 문제를 해결하기 위해 정부가 표준화 방안을 실효성 있게 추진할 수 있도록 기술적 지원과 경제적 인센티브를 제공하는 것이 중요하다

Chapter 4. AI 생성물 학습데이터 설명의무에 대한 한계와 합리화 방안

□ 학습데이터 설명 의무

- 쟁점) AI기업 입장에서 영업비밀로 경쟁력 악화, 대규모 GPU 인프라 구축, 유지에 있어서 투자 손실

설명 방법은 생성물 유형별로 방식과 요구사항이 다름, 설명 의무를 부여하기 앞서 기술적 한계의 극복 선행부터 모호한 또는 뻔한 그리고 불분명한 설명의무는 데이터 편향성의 해결방안이 아니라 오히려 데이터 편향성 책임에 대한

면피나 회피 수단으로 적용될 수 있음 * [EU AI Office의 가이드는 불명확](#)하고 '25.08까지 마련 계획 (26년 시행 계획)

- 한계 분석)
 - * 설명의 범위; 기업 영업 비밀과의 경계 설정의 어려움
 - * 내용의 정확성 검증; 데이터의 신뢰성과 편향성에 대한 검증 어려움

Chapter 4. AI 생성물 학습데이터 설명의무에 대한 한계와 합리화 방안

□ EU AI 법안 53조 1항, d에 대한 학습데이터 설명 방법 템플릿 청사진 (출처: OpenFuture, mozilla)

유형 (주요 부분 발제)	내용
1.1 크기	학습데이터 총 크기(작품수, 토큰 수, GB 단위 크기 등) 토큰화 과정 정보 포함
2. 데이터 출처 및 데이터 세트	데이터 출처와 구체적 데이터 세트에 대한 정보 제공을 통해 데이터의 적법성과 신뢰성 증대
2.1. 데이터 수집 정보	데이터 수집 방식 (인터넷 스크래핑, 공공 저장소, 특정 데이터베이스, 제3자 제공 등)과 수집 과정 설명
2.7. 데이터 세트 목록	각 데이터 세트의 비율, 이름, 소유자, 도메인, 모달리티(텍스트, 이미지 등), 언어, 사용 목적
3. 데이터 다양성	학습 데이터의 대표성과 편향성을 고려하여, 다양한 데이터 유형을 포괄함으로써 AI 모델의 신뢰성을 증대

☞ 현 시점의 학습데이터 설명 템플릿은

'데이터 수집의 불균형성'은 편향된 데이터를 감지 및 제거하는 '기술적 한계 극복 선행 필요'
직접적인 정부 규제 방식보다 기업이 스스로 자율적인 조치를 독려하는 방안 모색 필요

- > 1) AI 기업들이 투명하게 설명하기 위한 합리적인 범위 설정
- 2) 이를 적극적으로 수행할 수 있도록 환경과 시스템 구축
- 3) 자율규제는 기술 변화에 유연, 사회적 비용 절감 달성할 수 있는 방안

Chapter 4. AI 생성물 학습데이터 설명의무에 대한 한계와 합리화 방안

(쟁점)

1) 데이터 수집의 불균형성

AI 학습 데이터의 출처와 내용에 대한 명확한 공개는 AI 생성물의 기원을 추적하고 AI가 어떤 데이터를 바탕으로 결정을 내렸는지 이해하는 데 필수적이다.

편향된 데이터는 생성물의 생성 과정에 부정적인 영향을 미치고 AI 이는 공정성과 신뢰성에 심각한 문제를 일으킨다.

2) 데이터 출처의 불투명성

AI 기술의 복잡성과 불투명성은 AI 생성물의 투명성을 저해하는 주요 문제 중 하나로, 이로 인해 블랙박스 문제라고 불리는 여러 법적, 윤리적 쟁점이 발생하고 있다

3) 학습 데이터 설명으로 공정성과 신뢰성 문제

AI 시스템이 어떤 학습 데이터를 기반으로 생성물을 만들어내는 지 공개하는 것은 공정성과 신뢰성 보장을 위한 중요한 쟁점이다

Chapter 4. AI 생성물 학습데이터 설명의무에 대한 한계와 합리화 방안

(합리화 방안)

1) 투명성 지표를 통한 학습데이터 표준화된 설명 방안 마련

Bommasani et al.은 AI 모델 생태계의 투명성을 평가하기 위해, '기초모델투명성지수 (Foundation Model Transparency Index, 이하 'FMTI')를

출처: BOMMASANI, Rishi, et al. The foundation model transparency index. arXiv preprint arXiv:2310.12941, 2023, p. 27; Figure 8: Indicators. The 100 indicators we use across 3 domains (upstream, model, and downstream)that are the same as in the October 2023 Foundation Model Transparency Index.

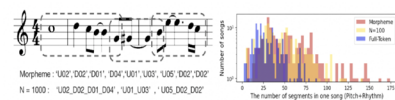
Name	Platform/Model	Release Date	Input	Output	Status	HD	WHL	WIL	FMTI	AMT
AI21 Labs	Flan-T5	2022.07	T, I	T	Startup	USA				
Alphazero	AlphaGo	2016	T, I	T	Startup	USA				
Anthropic	Claude-2	2023.07	T, I	T	Startup	USA	o			
BAIR	StarCoder	2023.04	T	T	Startup	USA			o	
Google	Gemini 1.0 Ultra	2023.07	T, I, A, V	T, I, A, V	Big Tech	USA	o			o
IBM	Granite	2023.07	T	T	Big Tech	USA	o			o
META	Llama2	2023.07	T	T	Big Tech	USA	o			o
Microsoft	Phi-2	2023.07	T	T	Big Tech	USA	o			o
Mistral	Mistral 7B	2023.06	T	T	Startup	FR				
OpenAI	GPT-4	2023.03	T, I	T	Startup	USA	o			o
Stability AI	Stable Diffusion	2022.10	T, V	T, V	Startup	UK	o			o
Meta	Llama3	2024.03	T, I	T, I	Startup	USA				

- Name: ICATRENNING (Learning Proficiency Index)
 - Input: Text, Images, Videos, Audio
 - Output: Text, Images, Videos, Audio
 - Status: Open (Open Source), Closed (Proprietary)
 - HD: High Definition (High Resolution)
 - WHL: Wheel (API)
 - WIL: Wheel (API)
 - FMTI: Foundation Model Transparency Index
 - AMT: AI Model Transparency Index

Upstream	Model	Downstream
Data size (학습데이터 크기 정보)	Input modality	Release decision-making
Data sources (학습 데이터의 출처 명시)	Output modality	Release process
Data source selection (데이터 출처 선택 기준)	Model components	Distribution channels
Data curation	Model size	Products and services
Data augmentation	Model architecture	Detection of machine-generated content
Harmful data filtration	Centralized model documentation	Model License
Copyrighted data (저작권 보호 여부)	External model access protocol	Terms of service
Data License (데이터 라이선스 정보)	Blackbox external model access	Permitted and prohibited users
Personal information in data	Full external model access	Permitted, restricted and prohibited users
Use of human labor	Capabilities description	Usage policy enforcement
Employment of data laborers	Capabilities demonstration	Justification for enforcement action
Geographic distribution of data laborers	Evaluation of capabilities	Usage policy violation appeals mechanism
Triggers	External reproducibility of capabilities evaluation	Permitted, restricted, and prohibited model behaviors
Instructions for creating data	Third party capabilities evaluation	Model behavior policy enforcement
Labor protections	Limitations description	Interoperability of usage and model behavior policies
Third party partners	Limitations demonstration	User interaction with AI system
Queryable external data access	Third party evaluation of limitations	Usage disclaimer
Direct external data access	Risks description	User data protection policy
Compute usage	Risks demonstration	Permitted and prohibited use of user data
Development duration	Unintentional harms evaluation	Usage data access protocol
Compute hardware	External reproducibility of unintentional harm evaluation	Versioning protocol
Hardware owner	Intentional harm evaluation	Change log
Energy usage	External reproducibility of intentional harm evaluation	Deprecation policy
Carbon emissions	Third party risks evaluation	Feedback mechanism
Broader environmental impact	Mitigations description	Feedback summary
Model stages	Mitigations demonstration	Government inquiries
Model objectives	Mitigations evaluation	Monitoring mechanisms
Core frameworks	External reproducibility if mitigations evaluation	Downstream applications
Additional dependencies	Third party mitigations evaluation	Affected market sectors
Mitigations for privacy	Trustworthiness evaluation	Affected individuals
Mitigations for copyright	External reproducibility of trustworthiness evaluation	Usage reports
	Inference duration evaluation	Geographic statistics
	Inference compute evaluation	Rodex mechanism
		Centralized documentation for downstream use
		Documentation for responsible downstream use

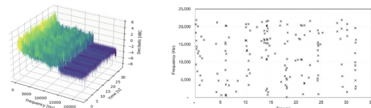
Chapter 4. AI 생성물 학습데이터 설명의무에 대한 한계와 합리화 방안

2) 저작권 식별체계 도입



출처

Park Saebui, "(A) natural language processing approach to quantitative measurement of symbolic melodic similarity", college of liberal Arts and Convergence Science, Graduate School of Culture Technology, GCT-Theses Ph.D, (2023), pp. 66-83를 요약하면 소리의 집합체인 음악을 컴퓨터가 이해하기 쉬운 자연어 처리방식으로 분석하려면 멜로디, 박자 등을 언어적 표현인 문장과 단어 형태로 구현하는 작업이 필요하다. 이를 위해서 음악적 특징들을 멜트워드(Mel2Word, 멜로디를 워드로 변환) 하는 알고리즘을 고안하여 음악 저작물 침해 판단에 있어서 정량적, 정성적 분석이 가능한 알고리즘 연구이다



출처

Lo'pez-Garc'ia A, Mart'inez-Rodr'iguez B, Liern V, "A proposal to compare the similarity between musical products. one more step for automated plagiarism detection?", Mathematics and Computation in Music, (2022)

3) 공통 표준 개발을 위한 업계 간 협력 거버넌스 구축

- 자율규제의 실효성을 높이기 위해서는 지속적인 연구 개발 지원이 필수적이다. AI 기술의 급속한 발전과 그로 인한 규제 필요성은 단순히 규제의 필요를 넘어서 기술 혁신과 사회적 책임을 동시에 달성해야 하는 과제를 안겨준다. 이를 해결하기 위해 정부와 민간이 협력이 필요하며, 이를 뒷받침하는 연구 개발 지원과 인센티브 제공은 자율규제에 있어서 핵심 동력이 된다.

Chapter 2. AI 생성물 투명성 규범의 의의와 내용

- 자율규제는 규제가 완화되는 것이 아닌 정부에서 사업자, 사업자 단체 혹은 독립적인 자율규제 기구로 이양됨을 뜻
(전문성 간극) **‘국가 경계를 초월’**하고 **‘시스템의 높은 복잡성과 전문성이 높아 평가와 규제가 어려운 분야’**
(국가간 차이) AI 강국 마다 규제 방식 차이가 있어 적절한 균형 방안 필요 (EU는 공적 규제 강화 강조, 미국은 자율규제 중심 등)
- 자율규제 내부/외부 협의체 사례
(내부 협의체)
<메타> **‘oversight board’** 페이스북, 인스타그램, 스레드 등의 콘텐츠에서 글로벌 인권 기준 바탕의 자율적 내부 감시 체계
* 국내에. <당근마켓> **‘프라이버시 정책’**, <이용자 보호위원회>, <배달의 민족> **‘리뷰 운영 정책’**
(외부 협의체)
<OpenAI, 메타 등> **Frontier Model Forum, Partnership n AI** 같은 민간 이니셔티브 외부 협의체 형식
<KISO> 인터넷 서비스 제공자(ISP)들이 자율적으로 콘텐츠 규제 및 관리를 목적으로 설립된 기구
최근에는 AI와 같은 신기술 대응을 위한 규약 제정 및 집행형 자율규제 기구로 변화 등
-> 내부/외부 협의체들은 AI 기술의 빠른 발전과 더불어 투명성 이슈에 대응하는 데 필수적인 자율규제 매커니즘으로 발전 중
- 자율규제 적합성
(유연성과 신속성) AI 결과물의 예측 불가능성, 정부의 정통적인 규제로 AI 발전 속도 대응 어려움
(산업 전문성) 기술 개발자와 규제 당국 간의 전문성 간극은 투명성 규제 효과적 작동에 큰 장애물,
(국제협력 촉진 등) 피규제자(AI 기업) 입장에서 EU AI 법안은 강제성이 짙은 반면 구체적 가이드가 없는 어려움 토로 중

Chapter 5. 결론

□ 자율규제 대응 현황

현황	설명
구글	24년 5월 30일부터 딥페이크 기술을 활용한 음란물이나 서비스 광고를 금지 SynthID 기술 개발로 이용자가 이미지의 출처를 확인, AI 생성물의 무단 사용을 방지할 뿐만 아니라 저작권 보호에 기여
유튜브	24년 3월부터 크리에이터에게 영상을 업로드할 경우 AI사용 여부 표시를 의무화 'Content ID'를 활용하여 저작권을 자동으로 인식하고 보호하는 체계를 구축
X	AI 투명성 보고서를 통해 약 530만 개에 이르는 이용약관을 위반한 계정을 삭제
CAI	콘텐츠 진위 입증 이니셔티브(Content Authenticity Initiative, 이하 'CAI')를 통해서 AI 생성물에 신뢰할 수 있는 메타데이터를 삽입해서 생성물의 진위 여부를 확인
C2PA	AI 생성 콘텐츠의 출처와 진위를 확인하기 위한 표준을 개발 중
스탠포드CRFM	기반모델투명성지수(Foundation Model Transparency Index, 이하 'FMTI') 를 개발 데이터 소스 공개, 신뢰성, AI 생성 콘텐츠 탐지 등의 100여개 지표를 제시하며 투명성 규제 기준을 마련

다수 AI 기업들은 표시, 식별 기술이 수익성에 직접 연관이 없어, **‘자발적 도입에 소극적인 태도’**
기술 혁신을 보호하면서 ‘학습데이터의 편향성을 줄이고’ ‘사회적 책임을 다하려는 움직임’

Chapter 5. 결론

- AI 기술의 발전은 창작의 패러다임 혁신이지만, 오용 가능성은 신뢰와 민주주의 위협을 가함
- EU AI 법안 등 투명성 규범은 AI 생성물의 출처와 학습데이터 설명 의무를 통해 신뢰를 구축하고자 하나, 기술적 한계와 세부지침의 부재가 문제로 남아있음
- AI 관련 제도 특징은 사회적 합의가 어려워서, 이를 해결하기 위한 투명성 가치가 필요함. 그러나, 표시 기술인 워터마크, 메타데이터, 라벨링 그리고 학습데이터 설명 의무의 효과적 실행을 위해 기술 개발 선행과 표준화가 필요
- AI 기술은 국경 초월/ 시스템 복잡/ 민간 중심의 기술 전문성으로 평가와 규제 어려움 따라서 자율규제로 기술 변화와 글로벌 환경에 유연히 대응하며, 신뢰성과 투명성을 강화할 대안 (현재 충분히 논의 거치지 않은 규제는 기술 발전에 미흡한 대응 초래, 즉 시기상조다.)
- 기업 간의 자발적 합의와 기술적 협력 기반이 AI 기술 발전과 윤리적 사용을 동시에 가능케 할 것 (표시방법: 생성물별 국제 기술 표준화, 학습데이터: 투명성 지수 표준화 방안 선행 필요)

제 개인적인 생각은

AI 생성 기술도 하나의 기술 혁신 단계,

즉 생성형 AI는 '도구'로서 생성될 것이고, 무엇보다 '인간 다운 표현 방법'은 더 많이 나올 것이다.

**생성형 AI는
미디어 생태계의
게임 체인저인가**



이미지생성형 인공지능이 그리는 기괴한 행복의 세계

| 오경미 디지털문화정책전공 박사수료, 오픈넷 연구원



이미지 생성형 인공지능이 그리는 기괴한 행복의 세계

오경미 (오픈넷 연구원)

이미지 생성형 인공지능을 둘러싼 갈등/논쟁의 촉발



- 2023년에 이미지 생성형 인공지능인 ‘미드저니’(Midjourney)로 만들어진 이미지는 전 세계의 수많은 이들을 혼란스럽게 했음
- 도널드 트럼프 전 미국 대통령이 경찰에 연행되어 구치소에 수감된 모습을 담은 이미지
- 프란치스코 교황이 이탈리아 명품 브랜드인 발렌시아가 외투를 착용하고 성베드로 광장을 활보하는 이미지
- 세상은 실제와 분간하기 어려울 정도의 정교함과 높은 완성도에 환호했고, 조작된 이미지가 촉발할 수 있는 사회적 갈등과 정치적 분쟁을 우려했음



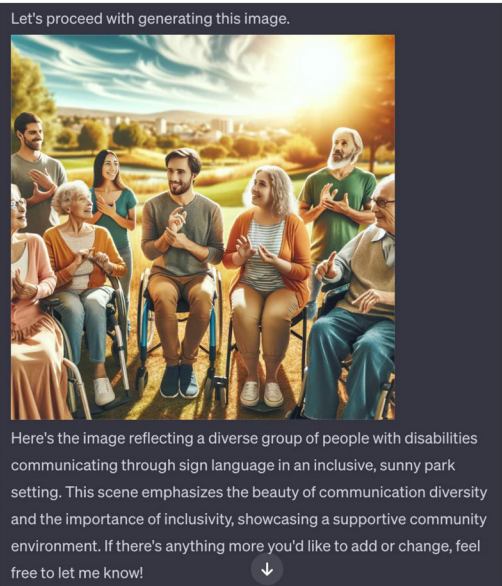
*본 발표는 2024년 3월에 발간된 과학잡지 에피 27호에 기고한 동일 제목의 글을 수정·보완한 것입니다.

갈등도 분쟁도 없이 전형적으로 행복한



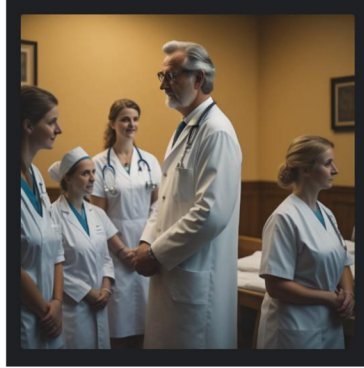
- ‘스테이블 디퓨전’(Stable Diffusion), 오픈AI의 ‘달리’(DALL-E), 마이크로소프트 빙의 ‘코파일럿’(Copilot), ‘이미진’(Imagine AI Art Generator)을 대상으로 의도적으로 자극적이고 혐오적이거나 정치적 혹은 논쟁적인 표현의 이미지 생성을 시도
- 이미지 생성형 인공지능으로 그릴 수 있는 세상에서 갈등과 분쟁, 불편함과 불쾌함은 생성하기가 어려웠음
- 이미지 생성형 인공지능의 데이터셋과 데이터 필터링은 현실을 왜곡하는 동시에 불편함과 불쾌함은 적극적으로 감추려 한다는 측면에서 정확하게 현실을 반영하고 있다고 분석할 수 있음

갈등도 분쟁도 없이 전형적으로 행복한



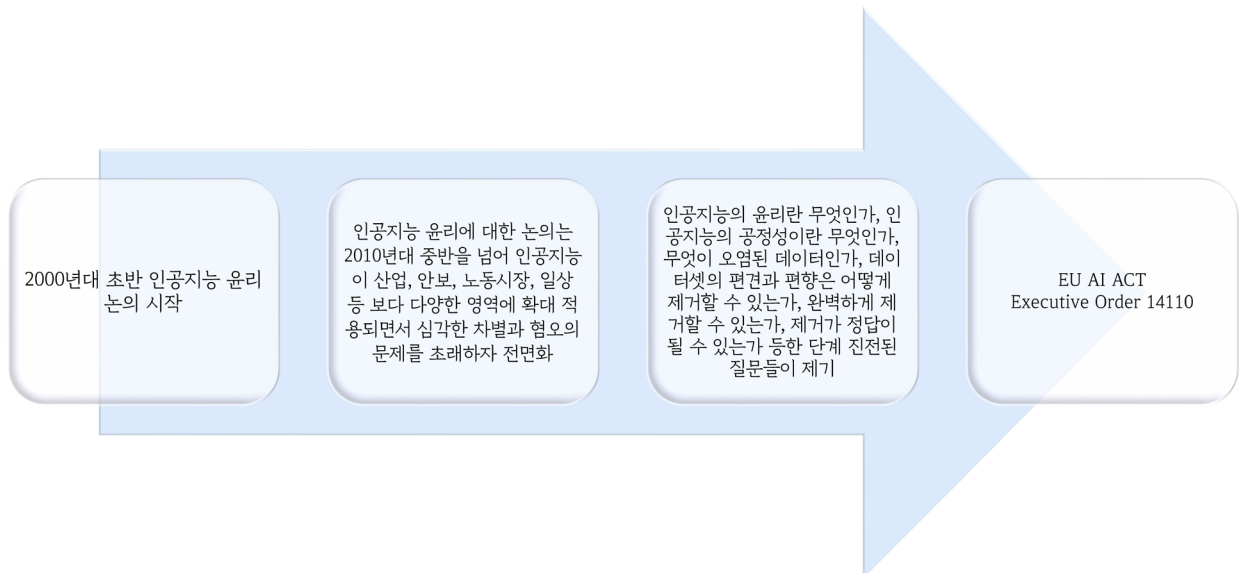
- 장애인 재현의 경우 차별에 대한 이미지 요청은 거절했으나 장애인이 수화로 대화하는 장면, 다양한 집단의 사람들이 함께 모여있는 장면 등의 요청에는 ‘다양성의 아름다움과 포용성의 중요성을 강조하며 서로를 지지하는 커뮤니티 환경’, ‘활기차고 긍정적이며, 공동체가 서로 연결된 기쁨과 다양성의 아름다움을 강조’, ‘따뜻하고 포용적인 장면을 만들어 보겠다’는 설명까지 곁들이며 생성해 주었음

갈등도 분쟁도 없이 전형적으로 행복한



- “의사와 간호사들”이라는 명령어에 스테이블 디퓨전과 이매진은 남성 의사와 여성 간호사를 결과물로 제시

이미지 생성형 인공지능의 한계와 우리에게 던져진 과제



이미지 생성형 인공지능의 한계와 우리에게 던져진 과제

- 인공지능 윤리에 관한 담론 진행 과정에서 인공지능 개발업체들은 윤리에 대한 논의에서 비난의 대상이자 교육과 계몽의 대상으로 여겨지면서 인공지능 서비스를 윤리적으로 운영할 책임을 요구 받게 되었음
- 이들은 위험을 감수하거나 도전하기보다 사회적인 비난이나 문제를 초래할 수 있는 원인을 미리 제거해버리는 손쉬운 방법을 택하게 되었을 것임

이미지 생성형 인공지능의 한계와 우리에게 던져진 과제

- 이미지 생성형 인공지능의 한계는 이미지 생성형 인공지능이 보여준 한계는 인공지능을 둘러싼 윤리와 공정성의 문제를 근본적인 차원에서 되짚어보며 어떤 방향으로 구체화할 것인지 논의할 때가 되었음을 말해주는 구체적인 사례일 것
- 혐오와 차별의 현실을 기록하고 비판하려면 혐오와 차별을 담은 표현을 경유할 수밖에 없음
- 이 과정이 불편하거나 불쾌하다고 해서 방법을 개발하지 않는 것은 약자들의 삶 자체를 인터넷 공간에서 지우는 것임

이미지 생성형 인공지능의 한계와 우리에게 던져진 과제

- 이미지 생성형 인공지능이 창작을 위한 도구로 적극 홍보되고 있음
- 이들 서비스가 강조하는 ‘창작’이 미적으로 아름답고 깔끔한 이미지를 만드는 숙련된 기술이라는 좁은 의미에 머무르지 않고 사회 비판, 철학적 질문, 자아 성찰 등 깊은 사유의 과정과 결과라는 의미까지 포함한다면, 이미지 생성형 인공지능에서 표현의 자유는 최우선적으로 고려되어야 하는 사항

생성형 AI는 미디어 생태계의 게임 체인저인가

주최/주관 서울과학기술대학교 IT정책전문대학원 및 원우회

기획 강태임(IT정책전문대학원 IT융합디자인전공 교수)
김자희(IT정책전문대학원 산업정보시스템전공 교수)
김현경(IT정책전문대학원 융합미디어콘텐츠정책전공 교수)
박소현(IT정책전문대학원 디지털문화정책전공 교수)
성욱준(IT정책전문대학원 시공공정책전공 부교수)
이영주(IT정책전문대학원 융합미디어콘텐츠정책전공 교수)
황주성(IT정책전문대학원 시공공정책전공 교수)

진행 총괄 박소현(IT정책전문대학원 디지털문화정책전공 교수)
황주성(IT정책전문대학원 시공공정책전공 교수)

진행 고윤지(IT정책전문대학원 디지털문화정책전공 석사과정)
김로건(IT정책전문대학원 융합미디어콘텐츠정책전공 석사과정)
이경미(IT정책전문대학원 디지털문화정책전공 박사과정)
이정원(IT정책전문대학원 융합미디어콘텐츠정책전공 석사과정)
오재상(IT정책전문대학원 융합미디어콘텐츠정책전공 박사과정, 원우회장)

행정지원 정득순(서울과학기술대학교 주무관)

홍보물 디자인 및 제작 도연디자인 고도연(IT정책전문대학원 디지털문화정책전공 석사졸업)

케이터링 파이데이



**생성형 AI는
미디어 생태계의
게임 체인저인가**