453회 공동체자유주의 세미나 - 품격있는 문명국가

인공지능 윤리와 규제

연 사 | 채은선 한국지능정보사회진흥원 수석연구원
일 시 | 4월 18일 (목) 오전 7시 30분
장 소 | Google Meet 실시간 온라인 세미나 구글 미트 설치법
주 최 | 한반도선진화재단, 정보경영 연구원

공동체자유주의 세미나 참가하기

챗GPT 등과 같은 생성·대화형 인공지능(AI)의 등장으로 단 몇 초 만에 어렵고 복잡한 논문을 요약해 주는 것은 물론 친구처럼 대화하는 것도 가능해졌습니다. 이미 인공지능은 우리의 일상 곳곳에 스며들고 있습니다. 인공지능은 우리의 생활을 보다 편리하고 윤택하게 해주는 디지털 시대의 핵심기술입니다. 하지만 인공지능기술과 함께 간과해서는 안 되는 것이 바로 인공지능 윤리와 규제의 문제입니다. 실제로 자율주행 자동차 관련 사망 사고, 자동화 로봇 오작동 사고, AI 챗봇 '이루다'의 성차별·소수자 혐오 표현 문제, 지식재산권 침해 문제 등 인공지능 윤리와 규제를 둘러싼 문제는 계속해서 발생하고 있습니다. 강연자 채은선 한국지능정보사회진흥원 수석연구원은 앞으로 기술이 고도화하는 만큼 인공지능과 관련한 'AI 윤리'와 관련한 문제는 무궁무진해질 것이라 지적합니다. 또한 그는 준수해야 하는 규범 형태 중 하나인 자율적인 이행을 이끌어낼 수 있는 수단이 필요하며 윤리가 법을 대체하는 것이 아니라 법을 보완하는 중요한 수단임을 강조합니다. 채은선 수석연구원과 함께 인공지능의 개념부터 인공지능 윤리 규제의 방향성을 살펴보고, 인공지능과 인간이 올바른 상호작용을 통해 슬기롭게 공존해 나가는 방법은 무엇인지 모색해 봅니다.

※이 세미나는 (사) 정보경영연구원의 지원에 의하여 이루어졌습니다

연사소개

채은선 한국지능정보사회진흥원 수석연구원

- BK21글로컬과학기술산학협력단 연구조교수

참가방법안내

이번 상반기 공동체자유주의세미나는 온라인 비대면 세미나로 진행됩니다.
참여자께서는 Google Meet를 통해 실시간 온라인 세미나에 참여하실 수 있습니다.
실시간 온라인 세미나에 참가하시면 발제자의 강연 후 진행되는 토론에 참여하실 수 있습니다.
세미나는 추후 유튜브 '한선TV' 채널에서 다시 보실 수 있습니다.

세미나 일정

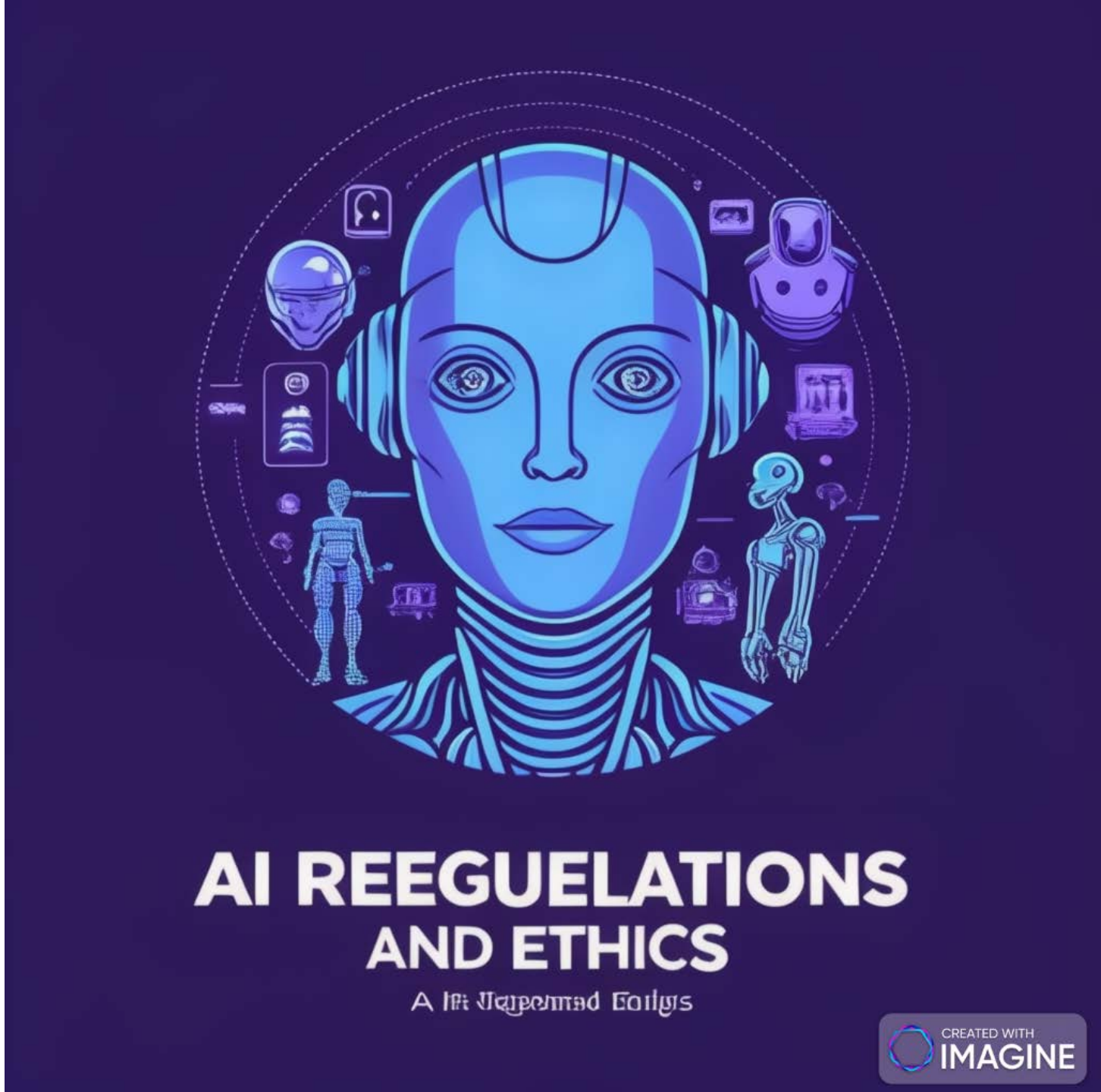
일정	주 제	발제자
4.18	인공지능 윤리와 규제	채은선 한국지능정보사회진흥원 박사
4.25	재정과 포퓰리즘	김우철 서울시립대학교 교수
5.2	기후변화와 한국의 과제	정태용 연세대학교 교수
5.9	경제안보와 한국의 과제	이효영 국립외교원 교수

주최 | 한반도선진화재단 Hansun Foundation · IMRC 정보경영연구원

인공지능 윤리와 규제

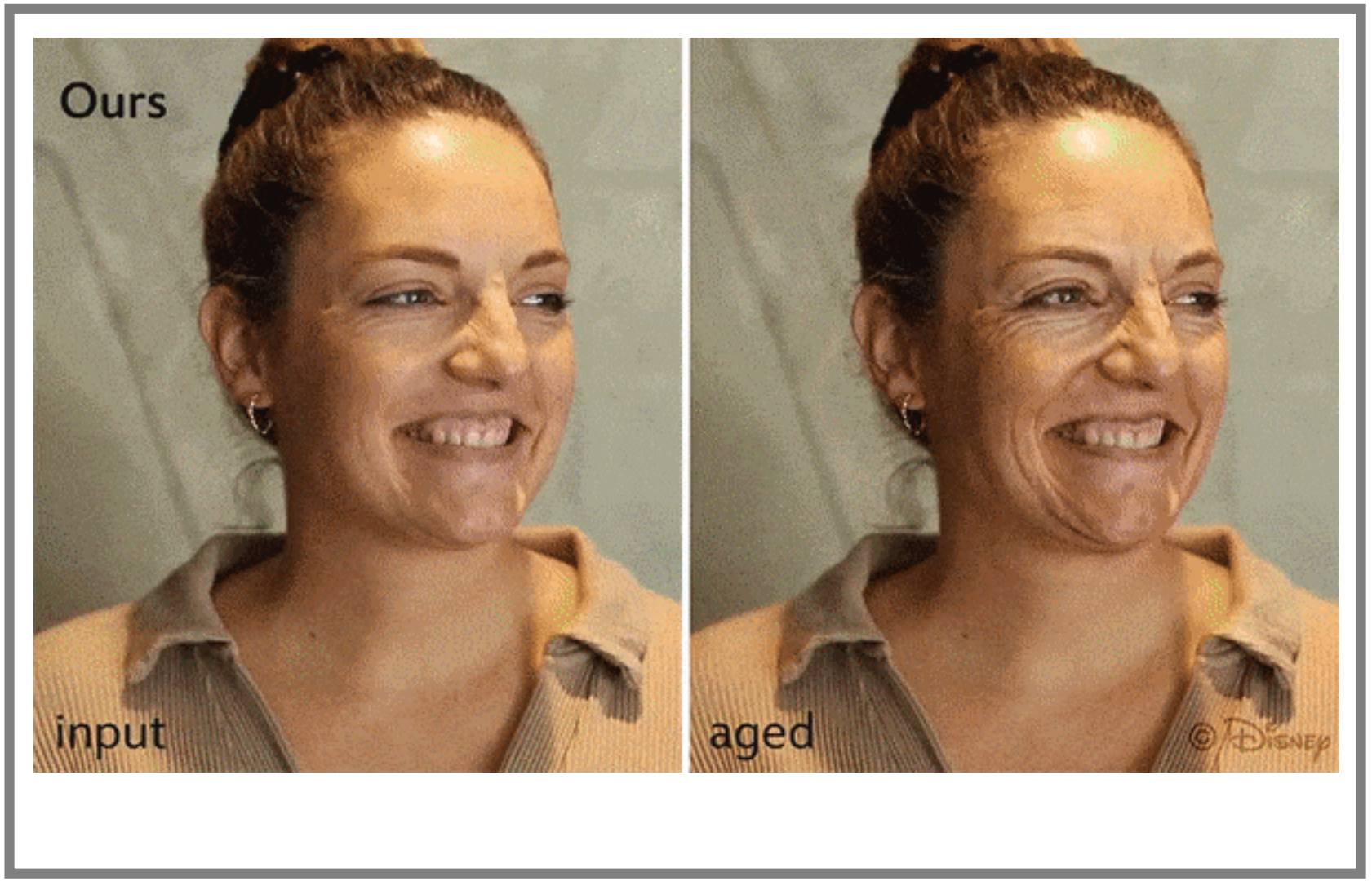
채은선

NIA AI법제도센터



"input": generate an image for the presentation titled as "AI regualtion and ethics" , using blue and violet and including robots and human beings

디즈니, AI 솔루션 'FRAN'



https://www.aitimes.com/news/articleView.html?idxno=148210

CONTENTS

CHAPTER Ⅰ 인공지능 개관

CHAPTER Ⅱ 인공지능 이용의 양면: 긍정 v. 부정

CHAPTER Ⅲ 인공지능 관련 다양한 접근: 윤리, 법적 규제 등

CHAPTER Ⅳ 인공지능 시대: 바람직한 규제 방향

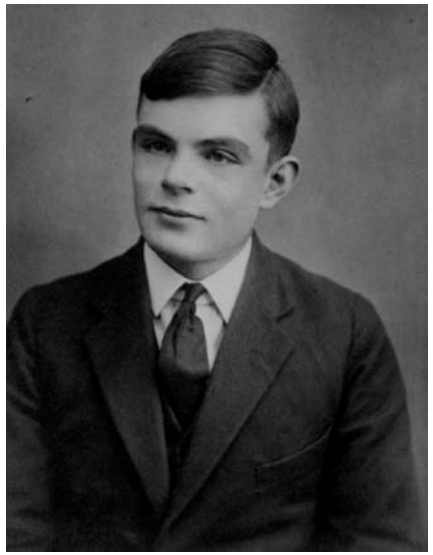
Ⅰ 인공지능 개관

인공지능이란?

OECD 인공지능의 개념(2023.11.)

An AI system is a machine-based system that, ①for explicit or implicit objectives, infers, from the input it receives, ②how to generate outputs ③such as predictions, content, recommendations, or decisions that can influence ④physical or virtual environments. Different AI systems vary in their levels of ⑤autonomy and adaptiveness after deployment

⊂ ?



앨런 튜링
1953, "튜링 테스트*를 통과하면 지능 가진 존재"
*사람과 컴퓨터 양쪽에 판정자가 타이포그래프를 이용한 문자 입력으로 질문을 한 다음 답변을 받아 어느 응답자가 컴퓨터인지 판정하는 테스트,
30% 이상 판정자가 컴퓨터와 인간을 판별할 수 없다면 시험에 합격

시사저널, 2016.9.30., https://www.sisajournal-e.com/news/articleView.html?idxno=158400

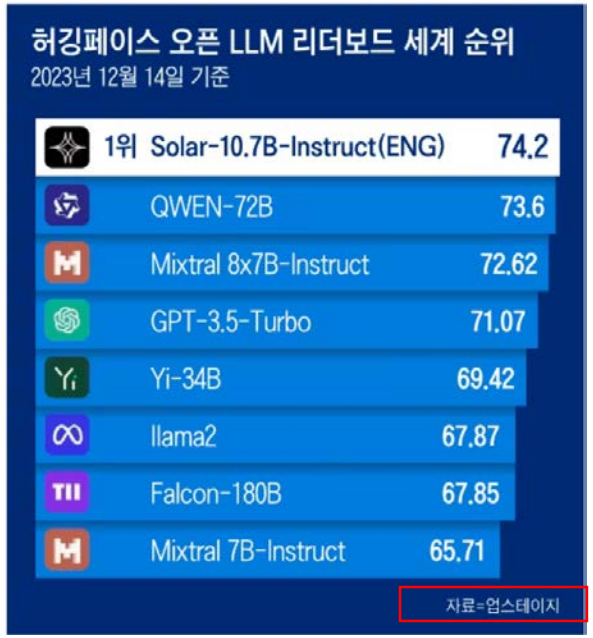
OpenAI GPT4

Ⅰ 인공지능 개관

인공지능 발달 현황

OpenAI ChatGPT

인터넷 발달 ▶ 데이터 축적 ▶ 기계학습, 특히 심층학습(딥러닝) ▶ 생성형 인공지능 ▶ 거대 언어모델 ▶ ?



| (일반) 생성형 인공지능 서비스(예) | 2)

기업	생성형 인공지능	출시년도
OpenAI	· ChatGPT , DALL.E 2, Sora	2022
MS	· 프로메테우스(Prometheus) ※ OpenAI에 대한 투자 통해 ChatGPT에 대한 독점 라이센스 확보3)	2023
구글(Google)	· 바드(Bard) Gemini	2023
메타(Meta)	· 람마(Large Language Model Meta AI, LLaM)	2023
네이버	· 하이퍼클로바X ※ 네이버 검색과 초거대 AI '하이퍼클로바' 간 융합 프로젝트인 '서치GPT' 준비 중	2023.7
카카오	· KoGPT(GPT-3 기반)	2021
SK텔레콤	· 에이닷(A.)(GPT-3 기반) ※ 향후, 멀티모달 서비스 제공 등 고도화 예정	2022
LG	· 엑사원(EXpert Ai for everyONE, EXAONE)' ※ 현재 전문분야별(IT·금융·의료·제조·통신 등) 적용 중	2021
KT	· 믿음(MI:DEUM)	2023 상반기

다음 자료 기반으로 일부 추가
: "ChatGPT 등장과 법제도 이슈", 「지능정보사회 법제도 이슈리포트」(2023-01), NIA, p. 5.

MS, 챗봇 코파일럿 버튼 키보드에 탑재한 AI PC· AI노트북 출시 마이크로소프트(MS)는 21일(현지시간) 자체 인공지능(AI) 모델인 '코파일럿'을 키보드에 기본 버튼으로 탑재한 PC와 노트북을 출시한다고 밝혔다. 사진은 MS 서피스 랩톱 6에 탑재된 코파일럿 키. 마이크로소프트 제공.
© 제공: 서울신문

서울신문, 2024.3.22., https://www.seoul.co.kr/news/economy/industry/2024/03/22/20240322500074

II 인공지능 이용의 양면: 긍정 v. 부정

기피 업무* 지원
* 감정노동, 단순·반복 노동 등

**[르포]수백대 소팅봇이 상품 분류…
쿠팡 로켓배송 비결은 '자동화'**

쿠팡 대구FC 1층에 도입된 소팅 봇(Sorting Robot)이 고객 주문 상품을 배송지별로 자동 분류하는 모습(사진=박준호 기자)

전자신문, 2023.2.7., https://www.etnews.com/20230207000003

사고, 안전 문제

문화일보 사회 로봇, 사람을 박스로 오인… 농산물 선별

로봇, 사람을 박스로 오인… 농산물 선별장서 4

문화일보 | 입력 2023-11-08 11:49

프린트

박영수

댓글 15 폰트 공유

경남 고성군 영오면의 한 농산물선별장 모습. 7일 이곳에서 프로그램 재설치 중이던 로봇이 오작동해 설치업체 작업자가 숨졌다. 경남소방본부 제공

태였다. 119는 유압 콤비툴을 이용, 선별기계 구조물을 분리해 A 씨를 구조해 이송했으나 끝내 숨졌다. 경찰은 현장 안전관리 책임자 등을 상대로 정확한 사고 원인과 과실 여부를 수사하고 있다.

앞서 지난 3월 28일에도 전북 군산의 한 자동차부품 제조 공장에서 작업자 1명이 산업로봇을 점검하던 중 오작동으로 로봇 기계에 눌려 중상을 입었다. 2022년 4월에는 경기 평택 진위면의 음료 생산공장에서 30대 근로자가 컨베이어 벨트와 연결된 산업로봇을 점검하던 중 신체가 끼이는 사고가 발생해 목숨을 잃었다. 2020년 7월에도 충남 아산의 한 자동차부품 공장에서 산업로봇을 점검하던 40대 노동자가 갑자기 작동한 로봇 팔에 깔려 숨지는 사고가 났다.

당시 40대 중국인 작업자가 부품을 생산하는 로봇 작동이 이상하다는 말을 하자 작업반장이 로봇 작동을 멈추고 수리를 시작했지만, 전원이 꺼지지 않은 상태에서 B 씨가 실수로 작동 스위치를 눌러 사고가 난 것으로 알려졌다. 해외에서도 로봇 오작동 사고가 종종 발생하고 있다. 2015년 독일 프랑크푸르트 북쪽의 자동차 공장에서 산업로봇이 오작동해 직원이 사망하는 사고가 발생했다.

교통·운전 문제 해결

경기도, 제로셔틀, 판타G버스

https://ggzerocity.or.kr/?p=8

https://www.gnnews.org/news/articleView.html?idxno=10138

사고, 안전, 학습 문제

Tesla settles lawsuit over Autopilot crash that killed Apple engineer

KEY POINTS

- Tesla has settled a wrongful death lawsuit brought by the family of Apple engineer Walter Huang who died in a Model X crash while using Autopilot in 2018.
- NTSB previously investigated the collision and found Tesla's driver assistance systems were partly to blame for the crash.
- Tesla has filed to seal from public view the terms of the settlement, including any amount paid by the company to compensate the family.

https://www.cnbc.com/2024/04/08/tesla-settles-wrongful-death-lawsuit-over-fatal-2018-autopilot-crash.html

https://www.tesla.com/ko_kr/modelx

美 자율주행차 첫 사망사고 운전자 '유죄'...보호관찰 3년형

| 우버 시험 차량 보조운전자로 일한 40대 여성…우버는 기소 안돼

(로스앤젤레스=연합뉴스) 임미나 특파원 = 2018년 미국에서 자율주행을 시험하던 우버 차량이 사람을 치어 숨지게 한 사고와 관련해 당시 우버 보조운전자였던 40대 여성이 법원에서 유죄 판결을 받았다.

미 애리조나주 매리코파 카운티 검찰총장실은 우버 차량의 교통사고에 연루돼 과실치사 혐의로 기소된 라파엘라 바스케스(49)가 매리코파 고등법원에서 보호관찰 3년 형을 선고받았다고 28일(현지시간) 밝혔다.

2018년 3월 18일 밤 바스케스는 자율주행을 시험하던 우버 차량 볼보 XC90에 보조운전자로 타고 있다가 애리조나주 템페에서 자전거를 몰고 도로를 횡단하던 엘레인 허츠버그를 치어 숨지게 한 혐의를 받았다.

How a Self-Driving Uber Killed a Pedestrian in Arizona

The death of a woman who was struck by an autonomous car operated by Uber is believed to be the first pedestrian fatality associated with self-driving technology. Newly released video offers clues about what happened.

https://www.yna.co.kr/view/AKR20230729011500075

https://www.nytimes.com/2018/03/23/technology/uber-self-driving-cars-arizona.html

창작 고취, 접근 용이

창작의 본질, 저작물, 저작자, 저작권자 관련 법제도에 대한 도전

AI가 제작한 작품 <스페이스 오페라 극장> 옆에 미국 콜로라도 주립 박람회 미술대회 우승 표시가 붙어 있다. 제이슨 엘런 디스코드 갈무리·연합뉴스

https://economist.co.kr/article/view/ecn202303110017

https://m.khan.co.kr/economy/economy-general/article/202209100800001#c2b

| ChatGPT 등 생성형 인공지능의 저작물 이용 관련 소송 사례 |

소송 등 대상	주요 사실 관계 및 주장
OpenAI :뉴스 기사 학습[8]	• OpenAI는 GPT-3.5 학습시키는데 월스트리트저널, 뉴욕타임즈 등의 유·무료 기사 이용 • 월스트리트저널 모회사는 기사를 인공지능 학습에 이용하기 위해서는 허가가 필요하다고 주장
MS, GitHub, OpenAI : 코드 데이터 학습 소송[9]	• 개발자들은 계약 등에 따라 GitHub(오픈소스코드 개발 플랫폼)에 특허자료를 게재 • 이후 MS가 GitHub를 인수하면서 GPT-3를 사용, GitHub 및 OpenAI는 GitHub의 코드 데이터를 학습한 소프트웨어 코드개발자 지원 AI기반 상품 출시 • GitHub 이용계약 및 오픈소스 저작권 위반, 불공정 행위, 사기 및 부당이득 등을 이유로 소송을 제기
Stable Diffusion, Midjourney, Devianart : 저작물 학습 소송[10][11]	• 예술가들이 AI학습에 허가 또는 동의없이 자신들의 저작물이 사용된 것에 대해 집단소송 제기('23.01.13) • 'Getty Images'가 자사 사이트의 최소 천이백만 개의 이미지를 복제한 것에 대해 Stable Diffusion을 상대로 소송 제기('23.02.03) • 피고는 상업적 목적으로 저작물을 이용했고, 이로 인해 원 저작물의 시장 가치가 저하되는 결과가 발생하였기에, 이들의 저작물 이용은 「저작권법」상 공정한 이용에 해당되지 않는다고 주장

"ChatGPT 등장과 법제도 이슈", 「지능정보사회 법제도 이슈리포트」(2023-01), NIA, p. 8.

- 뉴욕타임즈, OpenAI 대상 저작권 소송(2024)
- 네이버, 블로그글, 기사 등 자사 및 계열사 인공지능 학습에 이용

- 중국, 광저우 인터넷 법원, 인공지능기업에 대해 생성형 서비스 제공과정에서 울트라맨의 저작권 등 침해 판결(2024.2.8.)

- 법원, 사라 앤더슨의 Stability에 대한 저작권 침해만 판단(나머지 각하, 2023. 10. 30.)
  - ✓ DeviantArt의 이미지 스크래핑 및 사용 관련 역할에 대한 자료 설득력 부족
  - ✓ 생성된 결과물이 파생적 저작물이라는 점에 대해서는, 해당 결과물과 학습된 저작물간 실질적 유사성 입증 필요

- 신문협회 공정한이용 미해당 입장 전달 및 저작권자 권리 보호 원칙 발표(2023.8)
- 한국일보, 약관에 국내 언론 첫 'AI 및 대량 크롤링' 항목 신설('23.7.31. 시행)

시·공 등 한계를 넘은 다양한 콘텐츠 제작

https://www.etnews.com/20230614000247

동아일보, 2024.4.9., https://www.donga.com/news/Entertainment/article/all/20240409/124400665/1

KB라이프생명 론칭 광고 | 라이프를 나름답게

밤양갱, 김광석 노래

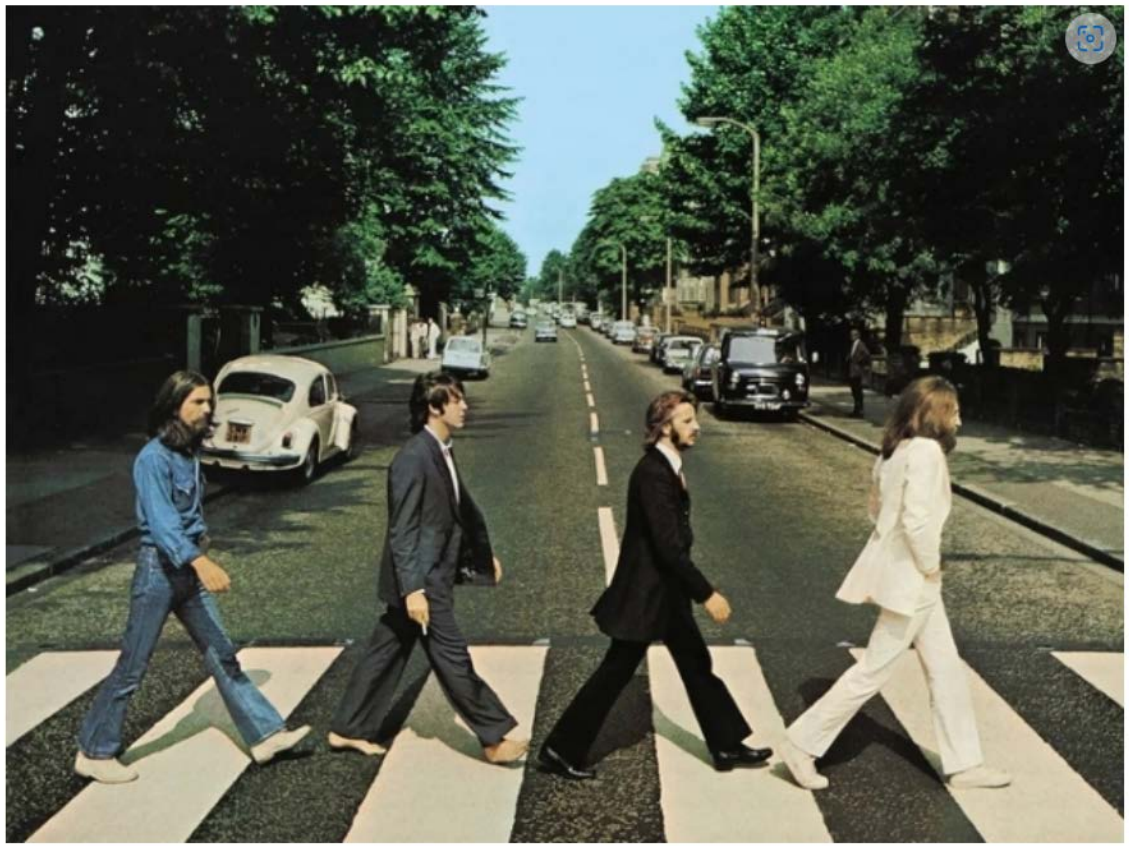
달디단 김광석 X 밤양갱(비비) ai 커버

YouTube · 당신의 이야기를 들려드립니다 · 2024. 2. 21.

YouTube

https://youtu.be/Zn1gwhP3w1E

The Beatles - Now And Then
(Official Music Video, 2023)
존 레논, 1978년 작곡



딥페이크 오남용

한경 "본인 맞아요?"...AI 증명사진 들고 주민센터 갔다가 '낭패' [이슈+]

2023.10.07.

이 안내문은 "AI 프로필 사진 등 과도한 보정으로 본인확인이 어려운 사진은 주민등록증용으로 사용할 수 없다"며 "주민등록법 시행규칙 제9조에 따라 본인확인이 어려운 경우, 담당 공무원이 보완을 요구할 수 있다"고 명시했다.

한 주민센터에 부착된 'AI 프로필, 과도한 보정 사진' 금지 안내문. /사진=김세린 기자

딥페이크: 허위정보 조작, 금융사기 등 범죄에 이용

N 뉴스 국민일보

주요뉴스 쇼폼 정치 경제 사회 생활 세계 IT 사설/칼럼 신문보기

국민일보

'기시다 AI 가짜영상'에 日정부 "민주주의 기반 손상"

입력 2023.11.06. 오후 3:52 기사원문

김승연 기자

기시다 후미오 총리의 AI(인공지능) 가짜 영상. 닛테레 캡처

중국, 친구 얼굴로 화상통화...8억원 송금, 2023.5.22

로이터는 22일(현지시간) 정교한 딥페이크 기술로 친구의 얼굴과 음성을 생성해 돈을 송금하도록 한 사기 사건이 벌어져 중국인들의 우려를 불러일으켰다고 소개했다.

이에 따르면 네이멍구 바오터우의 IT 업체 대표 궈모씨는 지난달 20일 "입찰에 필요한 보증금이 필요하다"며 도움을 요청한 친구의 계좌에 430만위안(약 8억원)을 송금했다.

송금이 제대로 이뤄졌는지 친구에 확인하는 과정에서 사기라는 것을 알아차린 피해자는 공안 당국에 신고했으나, 은행계좌를 정지하기 직전 이미 93만1600위안(약 1억7000만원)이 빠져나간 상태였다.

피해자는 영상 통화 속 친구의 얼굴과 음성이 똑같아 의심하지 않았던 것으로 밝혀졌다. 공안 당국은 가해자가 인공지능(AI)을 이용한 딥페이크 기술을 쓴 것으로 추정하고 있다.

https://www.aitimes.com/news/articleView.html?idxno=151265

일자리 창출
사람 업무 지원

- 세계경제포럼(WEF), "The Future of Jobs Report 2023"

AI 및 기계학습 전문가, 지속가능성 전문가, 비즈니스 인텔리전스 분석가, 정보보안전문가, 핀테크 엔진니어, 데이터 분석가 등 기술 관련 직업에서 빠르게 일자리가 성장
- 다만, 사무직, 비서, 은행원 및 관련 사무원, 우편서비스 사무원, 출납원 및 매표원, 데이터 입력원 등의 일자리는 디지털화 및 자동화에 의해 빠르게 감소할 것으로 전망

국문 주요 내용 정리 보고서: 과학기술인재정책 동향리포트 일자리의 미래 보고서 2023 - 세계경제포럼(WEF) -, KISTEP, 2023년 제 09호, https://hrstpolicy.re.kr/kistep/upload/file/202310284_1324100000.pdf

- **IMF, "Gen-AI: Artificial Intelligence and the Future of Work"**

"almost 40 percent of global employment is exposed to AI. Historically, automation and information technology have tended to affect routine tasks, but **one of the things that sets AI apart is its ability to impact high skilled jobs.**"

https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379

미국에 있는 기업들이 지난달 정리해고를 결정한 직원 규모는 모두 8만 명이 넘는 것으로 추산되는데, 이 가운데 3천 9백 명, 5%의 해고 이유가 바로 AI 때문인 것으로 조사
미국에서 기업들이 인력을 감축하는 이유로 AI가 등장한 것은 통계상 이번이 처음 **2023.6.2.,** https://news.kbs.co.kr/news/pc/view/view.do?ncd=7690744

일자리에 부정적 영향

美싱크탱크 "은행·IT업종 위험군"

생성형 인공지능(AI)이 은행·정보기술(IT)업계 일자리에 가장 큰 충격을 줄 것이라는 연구 결과가 나왔다.

미국의 싱크탱크 버닝글래스인스티튜트는 1일(현지시간) 보고서를 통해 재무분석가와 회계사의 주 업무인 숫자 계산, 법률 전문가의 규정 확인, 개발자의 코딩과 디버깅 등을 전적으로 AI가 대체할 수 있는 업무로 분류했다. 기자의 리포팅, 뉴스 요약과 기사 작성 등도 AI가 수행할 수 있다고 평가했다.

기업 직원들이 AI가 대체할 수 있는 업무를 수행하는 정도를 평가하는 '생성형 AI 노출 점수'를 집계한 결과 모건스탠리가 가장 높았다. 노출 점수가 높을수록 AI가 대체하기 쉬운 업무임을 뜻한다. 뱅크오브아메리카(2위)와 노스웨스턴뮤추얼자산운용(3위), 골드만삭스(4위)가 뒤를 이었고 맥킨지앤드컴퍼니(7위), KPMG(8위), 세일즈포스(15위), 구글(19위) 등도 20위 내에 들었다.

생성 AI의 충격파로부터 안전한 직종은 블루칼라일 것으로 전망했다. 대면 서비스는 생성 AI가 수행할 수 없는 영역이라는 등의 이유에서다. 건설 노동자, 수리공, 세탁소 근로자, 의료 종사자, 사회복지사, 아동 돌보미 등을 생성 AI 영향이 덜한 직종으로 분류했다.

매트 시겔만 버닝글래스인스티튜트 최고경영자(CEO)는 "과거 자동화 기술이 육체노동자의 일자리를 기계로 대체했다면, 생성 AI는 고도로 숙련된 전문직에 가장 큰 영향을 미칠 것"이라고 관측했다.

https://www.hankyung.com/article/2024020210051

- 미국, 콤파스 (COMPAS) 흑인 차별
- MS, 테이 성차별적 발언
- 아마존 채용 AI 여성 차별
- 2015 구글 포토 흑인을 고릴라로 분류
- 2021년 페이스북, 영장류 사건

Facebook's A.I. labeled the video of Black men as content "about Primates."

-

차별, 불공정, 편향

파이낸셜뉴스

"내 얼굴이 백인됐다. 비인간적" 아시아 女모델 '인종' 바뀐 황당한 사연

문영진 | 입력 2023. 11. 7. 05:00 | 수정 2023. 11. 7. 10:13

디자이너 마이클 코스텔로 패션쇼에 선 대만계 모델 쉬린 우의 얼굴이 백인으로 바뀌었다. 틱톡 캡처

교육의 질 향상, 교육 기회 확대

교육부 보도자료 당시 대한민국! 새로운 국민의 나라

배포 2023. 6. 7.(수) 15:30 보도시점 (인터넷)2023. 6. 8.(목) 14:00 (지 면)2023. 6. 8.(목) 14:00

인공지능(AI) 디지털교과서로 1:1 맞춤 교육시대 연다

- '모두를 위한 맞춤 교육' 실현을 위한 「인공지능(AI) 디지털교과서 추진방안」 발표
- 2025년 수학, 영어, 정보, 국어(특수교육) 교과에 우선 도입, 2028년까지 국어, 사회, 역사, 과학, 기술·가정 등으로 확대
- 담당 교원 연수, 맞춤형 교수·학습방법 개발 등 학교 현장 안착을 위한 지원 추진

https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&boardSeq=95261&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N

'AI교과서 학습데이터 플랫폼'

교육부가 추진 중인 '인공지능(AI) 디지털교과서 학습데이터 관리 체계 구축'사업 우선협상대상자로 LG CNS가 선정됐다. 공공소프트웨어 사업에 대기업 입찰 참여 제한 제도가 도입된 후 교육부가 실시하는 대규모 시스템 구축 사업에 대기업이 선정된 것은 처음이다.

교육부는 학습데이터 관리체계 구축을 위해 807억원의 예산을 투입하며 대기업의 사업 참여 제한을 풀었다. 이는 학습데이터의 중요도를 고려한 조치로 풀이된다. AI 디지털교과서로 학생별 맞춤 교육을 제공하기 위해서는 교과서를 통해 산출되는 개별 학생 학습데이터를 수집하고 분석해 활용해야 한다. 높은 수준의 민감한 학생 개인 정보가 모이는 만큼 높은 수준의 보안과 데이터 안정성 및 신뢰성이 요구된다. 또한 대기업 참여가 제한됐던 4세대 교육행정정보시스템(NEIS)에서 오류가 발생해 교육부가 전담TF를 꾸려 대응해야 했던 점 등도 결정에 영향을 준 것으로 보인다.

https://www.etnews.com/20240305000247

AI가 자동 번역…영어 못 해도 외국 명문대 강의 듣는다

"앞으로 제공하게 될 서비스는 단순 번역을 넘어 음성까지 자동으로 번역하는 것"이라며 "아직 많은 컴퓨팅 비용이 필요하지만, 조만간 전 세계 모든 학생이 모든 강의를 자국어로 들을 수 있을 것"이라고 말했다.

지난 12일 서울 중구 플라자호텔에서 열린 미국 교육 플랫폼 기업 코세라(Coursera) 기자간담회에서 제프 마지온칼다 최고경영자가 딥페이크 기술을 활용해 자신이 영어로 말하는 영상을 한국어로 곧바로 변환한 영상을 소개하고 있다. 박지영 기자

한겨레, 2024. 3.19., https://v.daum.net/v/20240319070501567

의료인 지원(업무 효율화), 의료질 향상

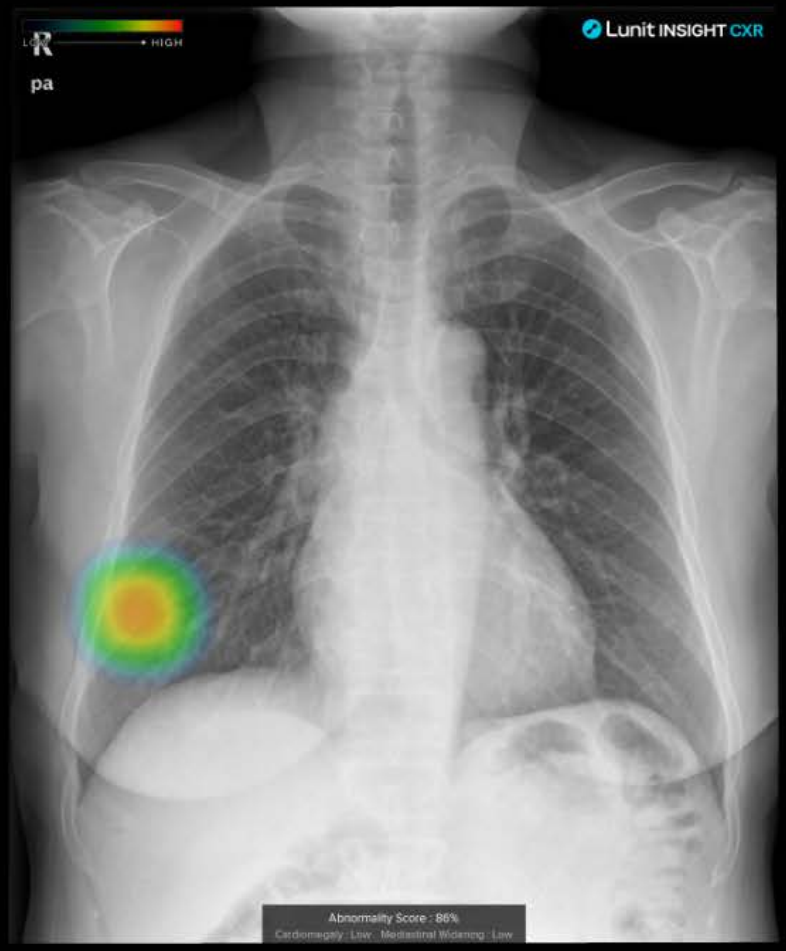
Lunit®

제품 연구/논문 투자 채용 루닛소개 제품 문의

97-99% AUC

루닛 인사이트 CXR은 흉부 X-ray 영상에서 가장 흔하게 발견되는
10가지 비정상 소견을 97~99% 정확도로 검출합니다.
(결핵 선별 검사에도 사용할 수 있습니다)

결절 ✓	경화
기흉	흉수
무기폐	기복증
심장비대	종격동 확장
석회화	폐섬유화
폐결핵 진단보조	



이 웹사이트는 보다 나은 서비스 제공을 위해 당사 개인정보 처리방침에 명시된 쿠키(또는 유사한 기술)를 사용합니다.
이 창을 닫거나 "동의"를 클릭하면 쿠키 사용에 동의하는 것입니다.

동의

법률서비스 질적 향상, 업무효율성 향상

생성형 인공지능의 등장으로 법원의 업무에 인공지능 사용 수준 및 범위가 보다 확대될 것으로 예상*, 우리나라도 법원 업무지원을 위한 인공지능 도입 예정**

* 예) 싱가포르는 2023년 9월 美스타트업 기업 Harvey AI와 소액사건에의 생성형 인공지능 시범사용 계약 체결, 대만은 판결문 데이터(1996~2021년) 학습 인공지능을 이용하여 판결문 자동생성 시스템 개발 중

** 예) '차세대 전자소송 시스템'(24년 9월 출범 예정)은 유사 판결문을 추천해 주는 인공지능 모델 적용, 사법서비스 편의성 개선을 위한 인공지능 활용 필요성이 대법원 차원에서 처음 언급

법률서비스 부정확성, 신뢰성 저하 등

2021년 11월 OpenAI의 ChatGPT 서비스 개시 이후 일부 변호사가 소송 서면 작성에 생성형 인공지능 활용하고 있는데 그 내용의 허구성 · 부적절성으로 인해 벌금이 부과되거나 증거로 채택되지 못한 사례* 존재

(예) (미국) 뉴욕주 판사 Peter Castel은 ChatGPT로 생성된 허구의 사례를 제출한 두 변호사에 대해 변호사로서 책임을 다하지 못한 것을 이유로 각각 벌금 5,000달러를 부과('23.06.)

(영국) 조세법원은 국세관세청으로부터 벌금을 부과받은 것에 대해 원고가 자신에게 유리한 사례로 제시한 '9개 사례'를 허위로 판단(따라서 원고의 행위를 정당화할 사유가 없다고 판단하여 원고의 이의제기 기각). 원고는 해당 사례를 사무변호사(solicitor)로부터 제공받았다고 주장('23.12.)

NIA, "해외 "사법기관 인공지능 이용 지침"의 주요 내용 및 시사점", 디지털 법제 Brief 2024-1, 2024.2.15.

맞춤형 서비스, 초개인화

<통신사>
(SK텔레콤) 'PASS 금융비서 보험분석 서비스'
(LG유플러스) '머니Me' 서비스를 통해 통신사, 카드사, 은행, 의료기관 등 여러 곳에 흩어진 개인 데이터를 통합 관리함으로써 개개인에 대해 맞춤형 혜택 제공

<IPTV>
(SK브로드밴드) TV를 보는 이용자를 자동으로 인식해 맞춤형 콘텐츠를 제공하는 자동개인식별(Auto Detection) 기능을 구현

<게임회사>
(넥슨)'카트라이더 러쉬 플러스' 재접속을 유도하고자 게임 유저들의 데이터를 분석하는 등 이용자들의 접속 패턴까지 분석해 그에 따른 알림 통지

<카드사>
(하나카드) '원더카드(ONE THE CARD)'고객별 최적화한 초개인화 맞춤 서비스를 제공하는 상품이다. 기존 데이터와 분석을 토대로 초개인화 플랫폼 '원더마켓' 론칭
(신한카드) 디지털·빅데이터를 기반으로 상권 분석, 마케팅 대행 및 개인사업자 대출 진행

https://www.newsian.co.kr/news/articleView.html?idxno=66733

'개인 맞춤형' AI 비서: Google Assistant with Bard

Oct 04, 2023

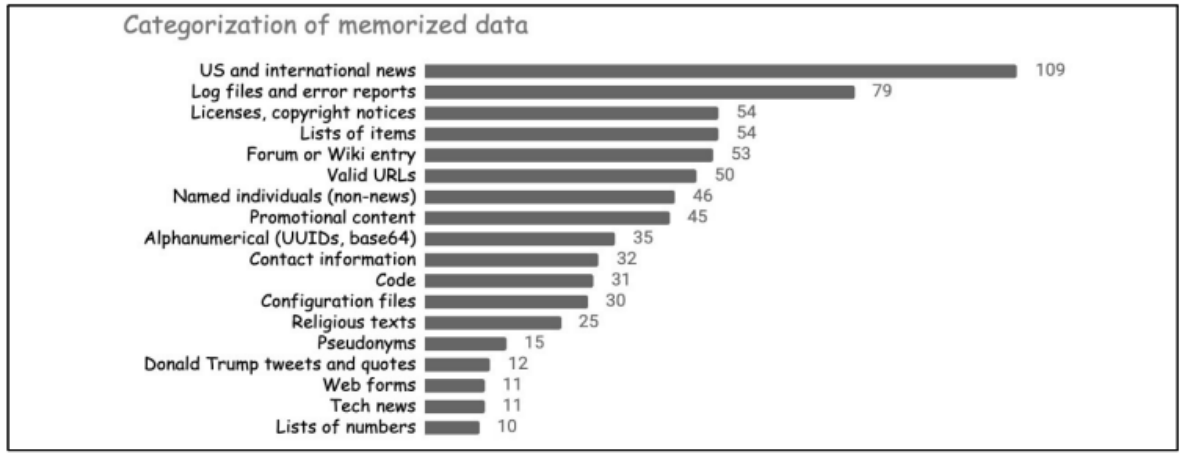
Today at Made by Google, we introduced Assistant with Bard, a personal assistant powered by generative AI.
You can interact with it through text, voice or images — and it can even help take actions for you.
In the coming months,
you'll be able to access it on Android and iOS mobile devices.
We're building Assistant with Bard to be a more personalized helper that's integrated with some of the Google services you already use, like Gmail and Docs — making it even easier to stay on top of the most important things in your life.[1]

https://blog.google/products/assistant/google-assistant-bard-generative-ai/

개인정보 유출

2020년 GPT-2를 대상으로 대규모 언어 모델이 암기하고 있는 데이터가 추출될 가능성을 연구한 결과, 생성 결과의 최소 0.1%가 암기한 데이터를 그대로 포함하고 있었다. 또한 암기한 데이터에는 개인의 이름, 연락처 등 개인정보가 포함되어 있었다.

| GPT-2 생성 결과에 포함된 암기된 데이터의 분류 |



※ 출처 : https://bair.berkeley.edu/blog/2020/12/20/lmmem/

개인정보보호위원회, 2023년 7월 26일 OpenAI에 대해 개인정보 유출 신고 의무 위반으로 과태료 부과

노출된 정보:

"챗GPT 플러스(유료계정) 사용 회원 중 1.2%의 이름과 대화 내용, 이메일 주소, 청구 주소, 신용카드의 마지막 4자리와 유효기간 등 결제 관련 정보"

- 캐나다 에어, 챗봇의 부정확한 정보로 인한 손해에 대해 배상책임

- 입법 미국 영어학습회사인 ‘iTutorGroup, Inc.'는 인공지능이 적용된 채용프로그램을 이용하여 교사 신청자의 서류 심사를 하였는데, 본 프로그램이 50세 이상은 무조건 탈락시킨다는 신고가 있었고, 조사 결과 부당한 연령차별 사실이 드러나 미국 고용기회평등위원회로부터 벌금을 부과받기도 함

  - 본 사건은 고용상의 차별로 다루어지긴 하였으나 자동의사결정 시스템으로 교사 신청자의 개인정보를 차별적으로 불투명하게 이용한 사례에도 해당

InEqual Employment Opportunity Commission v. ITutorGroup, Inc., et al., No. 1:22-CV-2565, E.D.N.Y. Aug. 9, 2023.

OECD, G7 회원국 대상 생성형 인공지능 관련 조사결과(2023.9.)

〈기회 및 각 기회를 선택한 국가의 수〉

기회	선택한 국가 수	기회	선택한 국가 수
① 생산성 향상	7	⑥ 사이버 보안 및 범죄에 대응	2
② 혁신 및 기업 진흥	7	⑦ 노년층 부양	2
③ 의료서비스 향상	6	⑧ 엔터테인먼트 · 창작 진흥	1
④ 기후 위기 해결에 기여	5	⑨ 빈곤 및 기아 감소	0
⑤ 맞춤형 교육	3	-	-

〈위험 및 각 위험을 선택한 국가의 수〉

위험	선택한 국가 수	위험	선택한 국가 수
① 허위정보 및 조작	7	⑧ 근로시장 대변동	1
② 지식재산권 침해	6	⑨ 사회 양극화	1
③ 프라이버시 침해	6	⑩ 불평등 확대	1
④ 편향성 및 차별성 악화	4	⑪ loss of human agency	0
⑤ 안전성 위협	4	⑫ 장기간의 실존적 위협 (existential threat)	0
⑥ 사이버 보안 위협	3	⑬ 환경에 부정적 영향	0
⑦ 불법적 활동(예: scams)	3	-	-

"OECD,_G7_히로시마_프로세스_논의_지원을_위한_보고서의_주요_내용_및_시사점", 「디지털 법제 Brief」(2023-06), NIA, p. 1, 2.

III 인공지능 관련 다양한 접근: 윤리, 지침, 법 등 (우리나라)

접근법	예시
윤리	▪ 2023 인공지능 윤리기준 자율점검표(안)(과학기술정보통신부) ▪ 2023 디지털 권리장전(과학기술정보통신부) ▪ 2022 교육분야 인공지능 윤리 원칙(교육부) ▪ 2020 사람이 중심이 되는 「인공지능(AI) 윤리기준」 (관계부처합동) ▪ 2019 '이용자 중심의 지능정보사회를 실현하기 위한 원칙 ('AI 윤리원칙')(방송통신위원회·KISDI) ▪ (2007~) 로봇윤리헌장 제정 추진(산업통상자원부)
지침	▪ 2023 생성형 AI 저작권 안내서 (문체부) ▪ 2023 「챗GPT 등 생성형 AI 활용 보안 가이드라인」 (국가정보원·국가보안기술연구소) ▪ 2021 금융분야 AI 보안 가이드라인(금융위원회) ▪ 2024 가명정보처리 가이드라인 개정, 비정형정보의 가명처리방법제시(개인정보보호위원회) ▪ 2021 인공지능(AI) 개인정보보호 자율점검표(개발자·운영자용) (개인정보보호위원회)
법	▪ 인공지능기본법 입법 추진 중(9건: 이성민 의원안, 이용빈 의원안, 윤두현 의원안, 안철수 의원안 등) ▪ 인공지능교육진흥법 입법 추진 중(2건: 조해진 의원 2022.8.24., 안민석 의원 2021.5.17.) ▪ 개인정보 보호법 및 신용정보보호법상 **자동의사결정** 관련 규정 ▪ 공직선거법 **딥페이크 이용 제한** 규정 신설(2024)
국제	▪ 2023~2024「인공지능의 책임 있는 군사적 이용에 관한 고위급회의(REAIM)」 (2023 우리나라·네덜란드, 2024 우리나라·네덜란드·싱가포르 공동) *Responsible AI in the Military Domain Summit ▪ 2023 안전성 정상 회의(AI Safety Summit) 개최(영국 공동), 2024.5 국내 개최 예정

<참고> 우리나라 인공지능법안(21대 국회)

의안번호	의안명	제안자명	제안자구분	제안일자
2123709	인공지능 책임 및 규제법안	안철수의원 등 10인	의원	2023-08-08
2120353	인공지능책임법안	황희의원 등 14인	의원	2023-02-28
2118726	인공지능산업 육성 및 신뢰 확보에 관한 법률안	윤두현의원 등 12인	의원	2022-12-07
2113509	알고리즘 및 인공지능에 관한 법률안	윤영찬의원 등 12인	의원	2021-11-24
2111573	인공지능에 관한 법률안	이용빈의원등31인	의원	2021-07-19
2111261	인공지능 육성 및 신뢰 기반 조성 등에 관한 법률안	정필모의원 등 23인	의원	2021-07-01
2104564	인공지능 집적단지의 육성에 관한 특별법안	송갑석의원 등 11인	의원	2020-10-19
2103515	인공지능산업 육성에 관한 법률안	양향자의원 등 23인	의원	2020-09-03
2101823	인공지능 연구개발 및 산업 진흥, 윤리적 책임 등에 관한 법	이상민의원 등 11인	의원	2020-07-13

〈주요 내용〉

- 고위험영역 인공지능 규제
- 인공지능윤리 마련 지원
- 자율적 검인증 지원
- 산업 혁신 지원

인공지능 신뢰기반 조성

Ⅲ 인공지능 관련 다양한 접근: 윤리, 지침, 법 등(해외사례)

❖ **2023.10.30. 바이든-해리슨 행정명령 행정명령**
❖ **2023 빅테크기업의 자발적 약속**(2023.7.21.)
❖ **연방기관의 가이드라인 또는 명령 등** 통한 규제
- 저작권청, 특허청, 보험, 금융 등 가이드라인/인공지능 생성 목소리 이용한 '로보콜' 금지 등(FTC, 2024.2.8.)

❖ 2023 **인공지능 위험관리 프레임워크**'(NIAST)
❖ 2022 **인공지능 권리장전을 위한 청사진**
❖ **'인공지능 규제 일반법안**(알고리즘책임법안, 인공지능 연구·혁신 및 책임 법안 소비자 보호 법안 등)**, 연방기관의 역할 및 책무 근거 법안**(연방 위험관리 법안, 인공지능책임법안 등)**, 소비자 보호 법안 등 입법 추진**
❖ 인공지능교육법(행정기관 종사자 교육)(22.10.17,10년 한시법)
❖ 주 또는 시별 진흥 또는 규제 입법 추진
- 뉴욕시, 2023. 1. 1. 「자동화된 채용 결정 도구 법」 시행
- Utah, 인공지능정책법(2024.3.13. 제정, 2024.5.1. 시행)

❖ 2024 **UN 결의 주도**

❖ **2023 親혁신적 인공지능 규제 접근**(A pro-innovation approach to AI regulation), 2024 의견 수렴 및 정부 대응에 관한 보고서 발간
❖ 2023 **인공지능 안전성 정상 회의** 주도
❖ 2023 안전연구소 설립: 프론티어 모델 출시 전 안전성 시험)
❖ 2023 영국 의회 입법 촉구
❖ 2023 영국 인공지능법안 발의(의원안)('23.11.22.)

❖ 기존 법제 내 혁신 추진
☞ **자민당 입법 추진('24.2.)** : 「책임 있는 인공지능 추진 기본법안」(責任ある AI推進基本法(仮))
❖ **2023 G7 히로시마프로세스 주도**: 생성형 인공지능 개발자 대상 11가지 지도원칙, 지도원칙별 준수사항 규정한 행동강령 마련)

❖ **「인터넷 정보 서비스의 심층합성 관리 규정」**('22.11. 제정, '23.1.10. 시행)
❖ **「생성형 인공지능 서비스 규정」**('23.7.제정, '23.8.15.시행)

❖ **데이터 및 인공지능법안**('22.3.29./'23.11.28. 수정안)
❖ 연방 공공 서비스에서의 생성형 인공지능 사용 지침
❖ 책임있는 고도화된 생성형 인공지능 개발 및 관리에 관한 자발적 행동 강령
❖ 공공 대상 영향평가 시행(자율)

입법 추진, 지침 마련, 정부주도의 자율규제 보장, 입법 최소화: 위험관리와 진흥 동시 추진

III 인공지능 관련 다양한 접근: 윤리, 지침, 법 등

다양한 접근법 개관: 해외 사례

<참고> 참조

- 2021.4.21.~ **EU 인공지능법** 입법 추진 중(2024.3.13. 의회 전체 회의 채택, 2024.4.~5. 내 이사회 채택 예정)
- 2022 제조물책임법지침 내 소프트웨어 포함 개정 추진
- 2022 인공지능 손해배상책임 지침 제정 추진

<참고> 참조

- 2024 UN 총회 **인공지능 결의**
- 2023, 유네스코 교육 및 연구를 위한 생성형 인공지능에 관한 지침
- 2021, 유네스코, 인공지능(AI) 윤리 권고

- 2023 **인공지능의 개념 정의** 개정
- 2023 G7논의 지원: G7 회원국 대상 생성형 인공지능에 관한 설문조사 및 보고서 발간
- https://oecd.ai/en/ 운영
- 2019 인공지능 원칙 채택

- 2023.10.30. 히로시마 프로세스에 관한 G7성명, 생성형 인공지능 개발자 대상 11가지 지도원칙, 지도원칙별 준수사항 규정한 행동강령

<참고> UN

- **인공지능 윤리 권고**(2021.11.24.)
  - ✓ **(Values)** 인권, 기본 자유 및 인간 존엄성의 존중·보호·증진/ 환경 및 생태계의 번영/ 다양성 및 포용성 보장/평화롭고 정의로우며 상호연결 된 사회 구축
  - ✓ **(Principles)** 비례의 원칙 및 무해성/ 안전 및 보안/ 공정성 및 반차별성/지속가능성/ 프라이버시 및 데이터 보호의 권리인간의 감독 및 결정/설명가능성/책임 및 의무/의식 및 리터러시/다자적이고 조정 가능한 거버넌스 및 협력

- **인공지능에 관한 결의**(2024.3.21.)
  - ✓ 제78차 총회에서 인공지능을 설계·개발·배포·이용할 때 인권의 존중·보호·증진을 강조한 '인공지능에 관한 결의' 채택, 미국 주도, 신규 분야를 규제하는 첫 번째 결의

  ① Same rights, online and offline

  ② Closing the digital divide

  ③ Hope for other sectors

  ④ We govern AI

<참고> EU 인공지능법제 입법 추진(1)

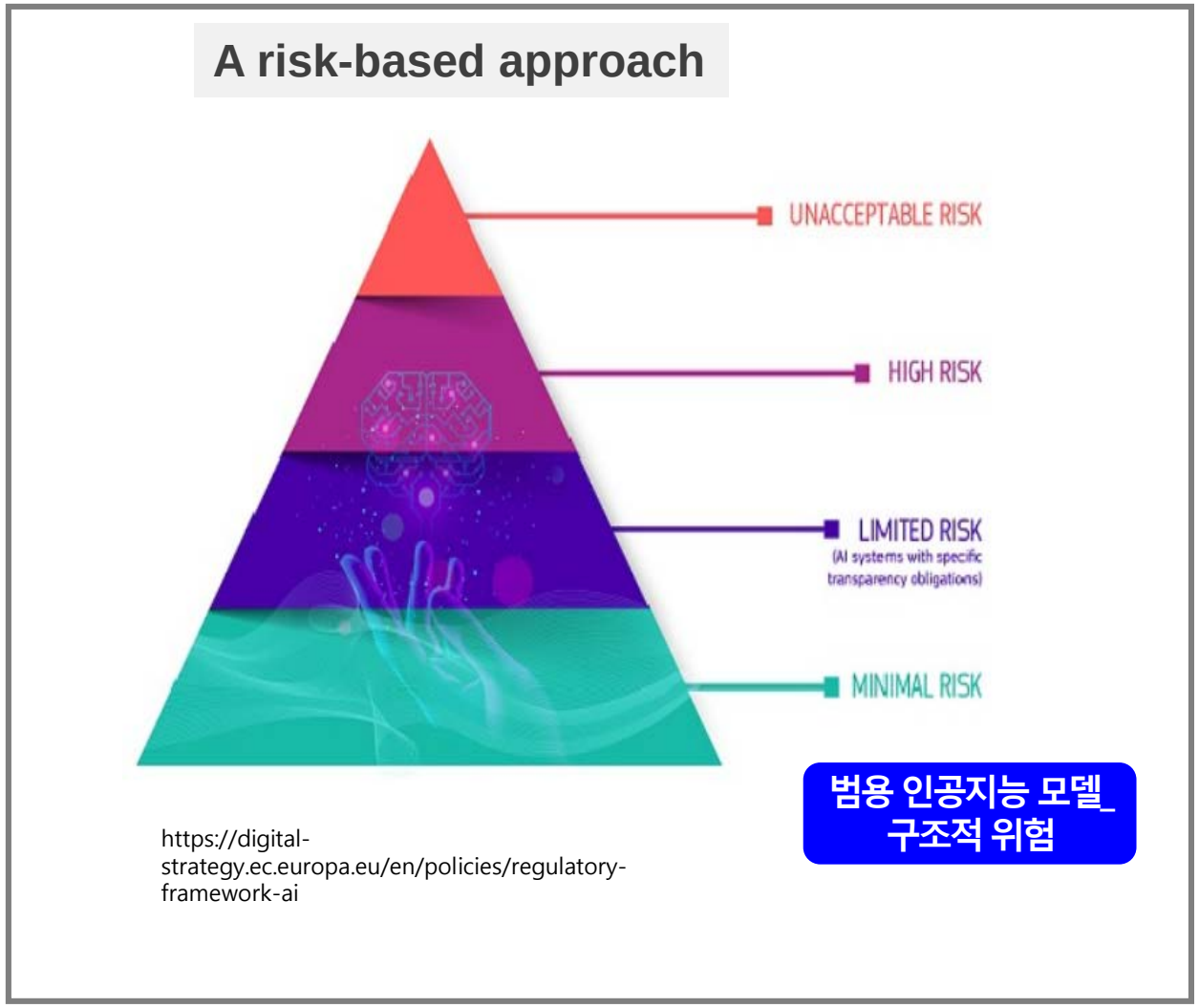
- **인공지능법**
  - ⑦ 20일 후 발효
  - ⑥ EU 공보에 게재
  - ⑤ 의회, 이사회 서명
  - ④ 이사회 승인
  - ③ 2024.3.13. 의회 전체회의 법안 채택
  - ② 2023.12.6.~12.9. 정치적 합의(trilogue)
  - ① 2021.4.21. 법안 마련 및 공개

European AI Office 설치·운영
(집행위원회 조직,
AI Office 설치 규정: 2024.1.24. 채택, 2024.2.21. 시행)

2024.2.2. 공식 잠정안 공개
2024.1.22. 비공식 잠정안의 초안 공개

- **EU 인공지능책임지침(Directive) 제정** 추진(2022~)
- **EU 제조물책임지침 개정** 추진(2022~)

<참고> EU 인공지능법제 입법 추진(2)

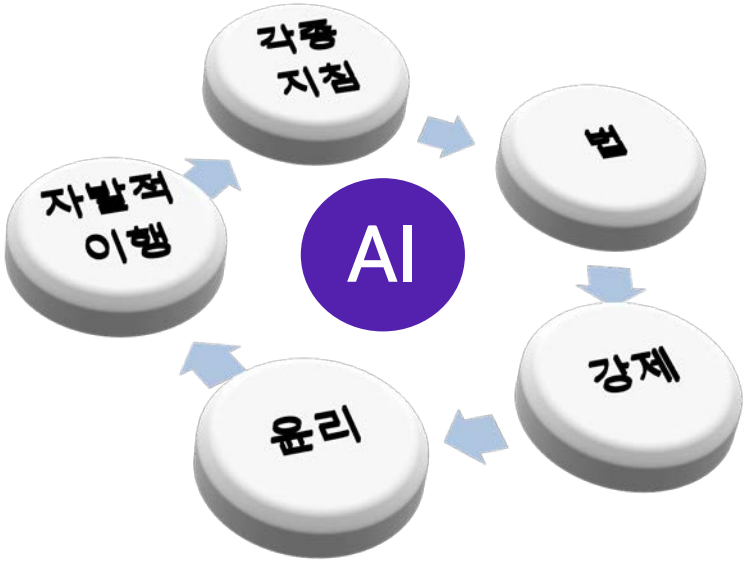


범용 인공지능 모델_ 구조적 위험

처벌 대상	벌금
금지 인공지능 관련 의무 위반	3,5백만 유로 또는 '전년도 전세계 매출액 7%' 중 높은 금액
고위험 인공지능 및 투명성 의무 위반	1,5백만 유로 또는 '전년도 전세계 매출액 3%' 중 높은 금액
권한당국에 부정확·불완전 정보 제공	7.5백만 유로 또는 '전년도 전세계 매출액 1%' 중 높은 금액
※ 중소기업 및 스타트업의 경우, 낮은 금액 부과	
범용 인공지능 의무 위반	1,5백만 유로 또는 '전년도 전세계 매출액 3%' 중 높은 금액

Ⅳ 인공지능 시대: 바람직한 규제 방향

- 기피 업무* 지원
  * 감정노동, 단순·반복 노동 등
- 교통·운전 문제 해결
- 창작 고취, 접근 용이
- 시·공 등 한계를 넘은 다양한 콘텐츠 제작
- 신규 일자리 창출
  사람 업무 지원
- 교육의 질 향상, 교육 기회 확대
- 의료인 지원(업무 효율화), 의료질 향상
- 법률서비스 질적 향상, 업무효율성 향상
- 맞춤형 서비스, 초개인화
- …



- 사고, 안전 문제
- 사고, 안전, 학습 문제
- 창작의 본질, 저작물, 저작자, 저작권자 관련 법제도에 대한 도전
- 지식재산권 침해
- 딥페이크 오남용
- 딥페이크: 허위정보 조작, 금융사기 등 범죄에 이용
- 일자리에 부정적 영향
- 차별, 불공정, 편향
- 법률서비스 부정확성, 신뢰성 저하 등
- 개인정보 유출
- 피해, 책임
- …

https://www.bbc.com/korean/articles/c72487e83y8o

감사합니다.

- 문의: chaes@nia.or.kr/ 053-230-1253