

자율주행,

그 이상의 구현을 위해 해결해야 할

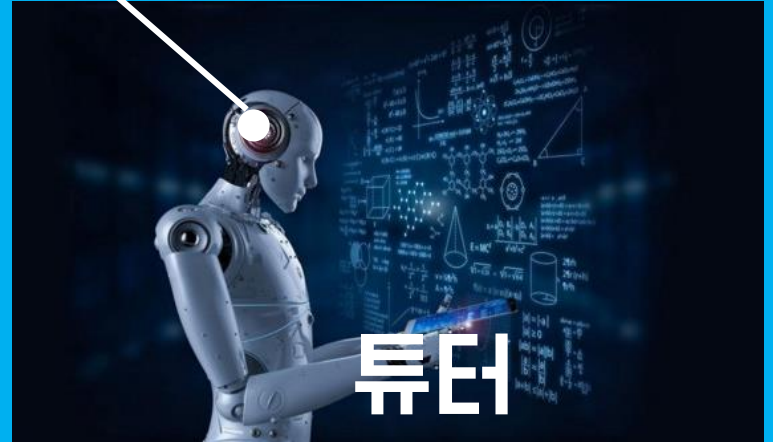
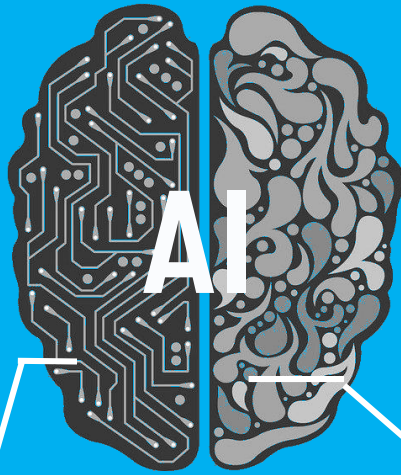
사회/윤리적 이슈

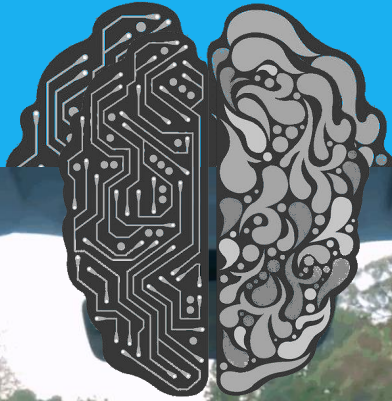
서창호

KAIST 전기 및 전자공학부

KAIST AI Institute 부원장







2021년 10월 15일 한림원탁토론회





자율주행

자율주행

Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
					
운전자 보조 기능			자율주행 기능		
無 자율주행 (No Automation)	운전자 지원 (Driver Assistance)	부분 자동화 (Partial Automation)	조건부 자동화 (Conditional Automation)	고도 자동화 (High Automation)	완전 자동화 (Full Automation)
운전자는 상황을 파악하고 운전함	운전자는 상황을 파악하고 운전함	운전자는 상황을 파악하고 운전함	운전자가 시스템의 요청 시 운전함	운전자가 시스템에 개입하지 않음	시스템의 역할
운전자의 역할	시스템이 운전자의 가/감속 또는 조향을 보조함	시스템이 운전자의 가/감속과 조향을 보조함	시스템이 상황을 파악하고 운전함		

자율주행

Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
					
운전자 보조 기능			자율주행 기능		
無 자율주행 (No Automation)	운전자 지원 (Driver Assistance)	부분 자동화 (Partial Automation)	조건부 자동화 (Conditional Automation)	고도 자동화 (High Automation)	완전 자동화 (Full Automation)
운전자는 상황을 파악하고 운전함	운전자는 상황을 파악하고 운전함	운전자는 상황을 파악하고 운전함	운전자가 시스템의 요청 시 운전함	운전자가 시스템에 개입하지 않음	시스템의 역할
운전자의 역할	시스템이 운전자의 가/감속 또는 조향을 보조함		시스템이 운전자의 가/감속과 조향을 보조함		시스템이 상황을 파악하고 운전함

현 기술 상황

“완전자동화” 단계에 못 미치는 이유

돌발상황

대처기술 부족

돌발상황 대처기술 개발 노력

컴퓨터 시뮬레이터 활용: 사고장면 연출




서창호 교수 연구진, AAAI 2019 (국제인공지능학회)

사고 위험도 예측 성능

Demonstration Video

Each number represents the predicted probability of collision of the vehicle

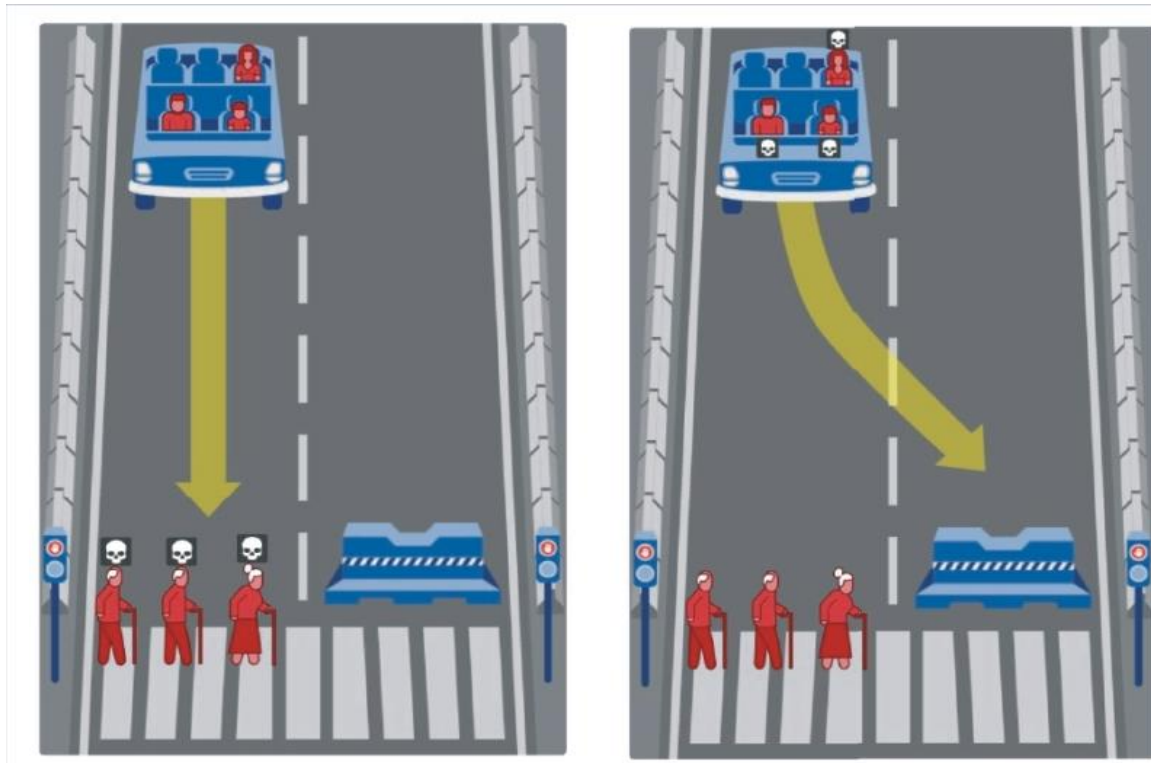
 : Ground truth dangerous vehicle

Predicted result of our algorithm



사회/윤리적 이슈

예시: 트롤리 전차딜레마 상황

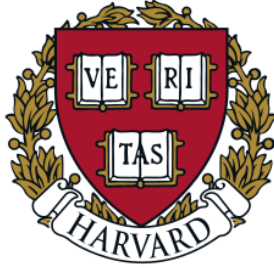


정당한 주행이란?

정당한 빅데이터 수집을 위한 노력



Massachusetts
Institute of
Technology



Toulouse
School of
Economics

“모럴머신” 데이터수집 플랫폼 (네이처 2018)

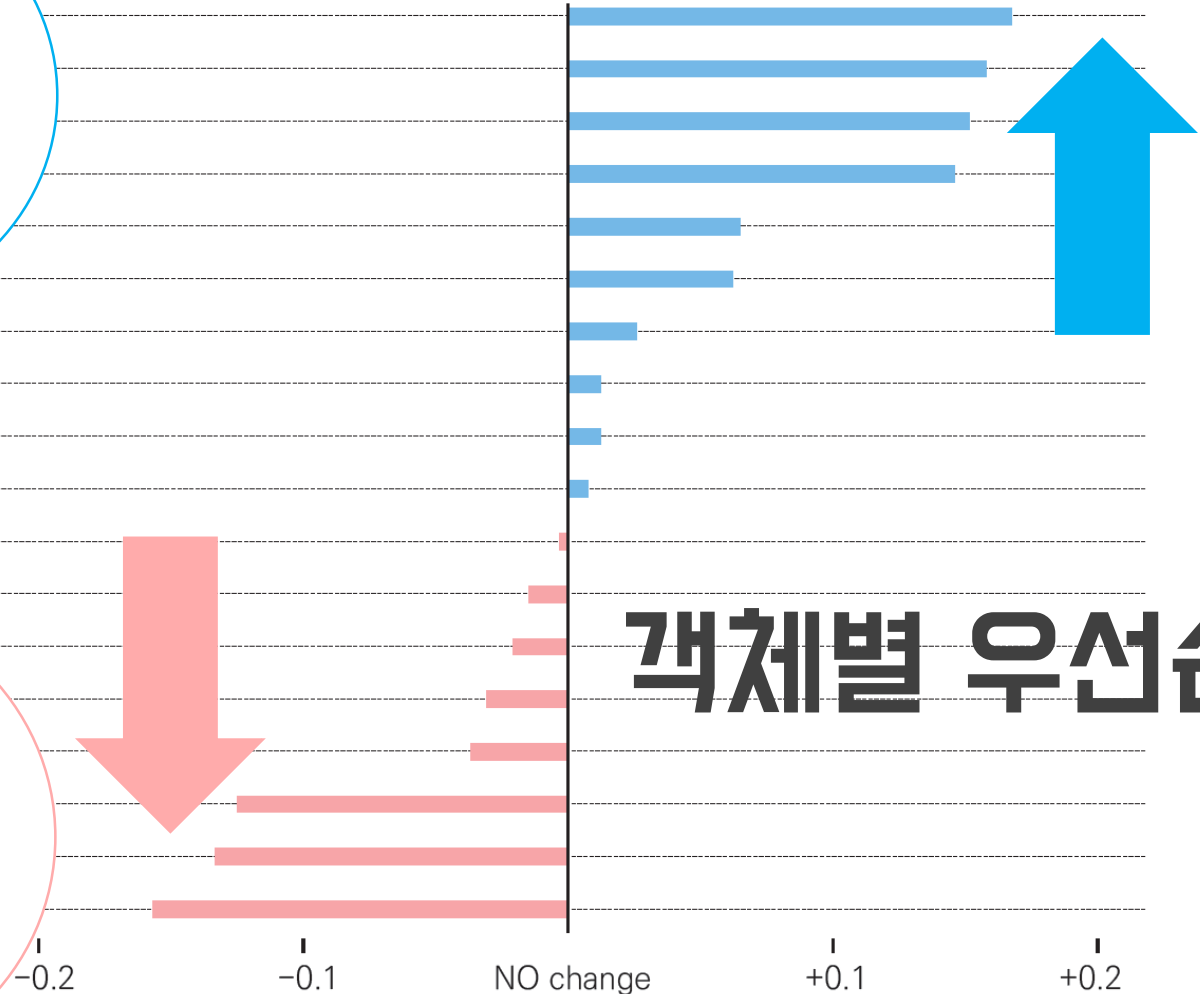
- 4천만개 이상의 딜레마 상황 연출
- 233개 국가 230만 명을 대상 데이터 수집

모럴머신: 공론화된 윤리기준 도출

유모차
소녀
소년

여성 운동 선수
여성 임원
남성 운동선수
남성 임원
체격이 큰 남성
체격이 큰 여성

개
범죄자
고양이



객체별 우선순위

데이터수집 플랫폼의 한계점

1. 빅데이터 수집의 어려움

예) 모럴머신: 18개월의 시간

230만명의 인력 동원

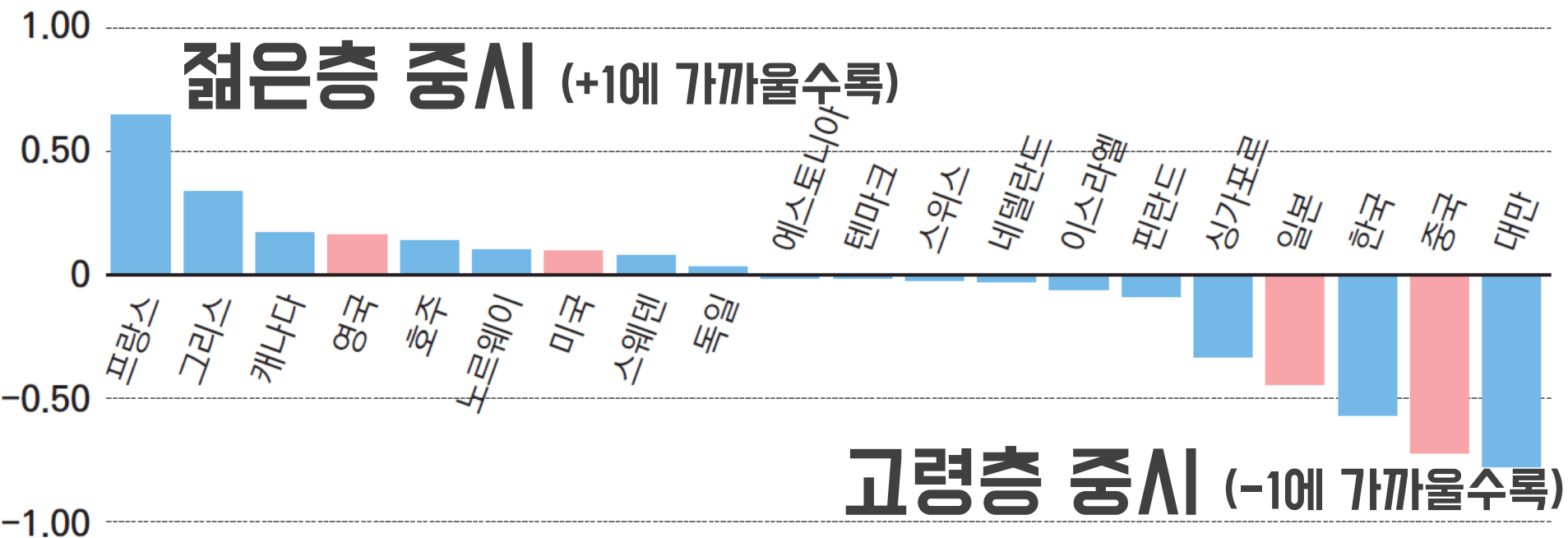
2. 데이터 편향성 문제

예) 샘플링 바이어스 발생할 경우

→ 특정 집단의 지엽적인 의견 반영

데이터 편향성 예시

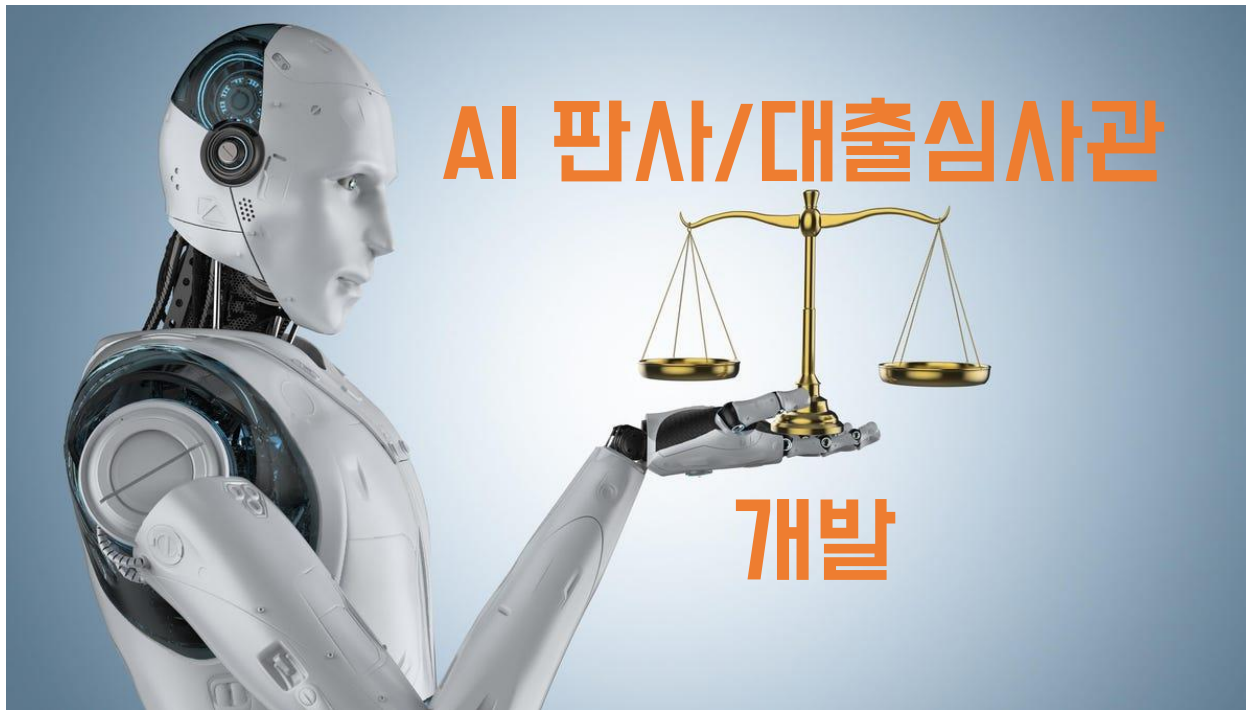
국가별: 고령층 vs 젊은층의 중요도



데이터 편향성 극복을 위한 노력

AI Ethics:

편향된 데이터로도 공정성을 보장하는 기술



서창호 교수 연구진, NeurIPS 2020 (국제신경망학회)

도전적 과제

1. 공정성/윤리에 대한 올바른 개념 확립
2. 이를 AI 기술에 접목시키는 능력 배양

장기적인 노력방안

윤리성을 겸비한 AI 인재양성

단기적인 노력방안

1. 공정성/윤리성 관련 기술 표준화

예) 한국정보통신기술협회 (TTA):
"신뢰할 수 있는 AI 기술" 표준화

2. 윤리성 보장을 위한 법제화

예) 비윤리적인 데이터 수집/활용관련
개인정보보호법 강화

마무리

자율주행을 위한 사회/윤리적 이슈:

정당한 주행?

노력방안:

기술표준화, 법제화, AI윤리교육