The 8th ASIA Future Forum
제8회 아시아미래포럼

폐회식 경품 이벤트 안내

- 추첨시간 : 11/16(목) 16:40~17:00 노사정 공동선언 폐회식 이후
- 추첨장소 : 국제회의장
- 응모방법 : 일반참가자 목걸이 네임택 위쪽의 경품응모권을 절취하여 로비의 경품응모함에 넣어주세요.

1등
1명
LG 올데이그램15 노트북

2등
2명
삼성 기어VR 2017

3등
3명
SKT 누구 NUGU 인공지능 음성인식 블루투스 스피커

더불어 행복한 세상을 만드는 언론
한겨레미디어

The 8th ASIA Future Forum

제8회 아시아미래포럼

PROGRAM BOOK

세션 3

사람과디지털연구소
중회의실 A

알고리즘과 사회적 가치

주최 | 한겨레신문사 주관 | 한겨레경제사회연구원

후원 | 기획재정부 외교부 산업통상자원부 고용노동부 보건복지부 대한상공회의소 은행연합회 생명보험협회 kosif 희망제작소 충남연구원 한국사회연대경제시민정부협의회 IWSR 복지국가연구센터 KLI 금융투자협회

협찬 | HYUNDAI MOTOR GROUP SK LG LOTTE posco GS칼텍스 emart Hyundai Oilbank KB금융그룹 KDB산업은행 SGI서울보증 SHINSEGAE S-OIL KOREAN AIR MIRAE ASSET 미래에셋대우 MIRAE ASSET 미래에셋생명 MIRAE ASSET 미래에셋자산운용 신한금융그룹 LG디스플레이 LG하우시스 LG화학 대한제지 DK 동국제강 서울연구원 The Seoul Institute 하나금융그룹 하나은행 | 외환은행 KOEN 한국남동발전 한국수출입은행 여신금융협회 KGC인삼공사 캠코 true friend 한국투자증권 한화 한화생명 현대건설 현대엔지니어링 HYUNDAI STEEL Hyundai Card Hyundai Capital HYOSUNG

November 16 (Thu.)
2017. 11. 16(목)
중회의실 A

The 8th ASIA Future Forum

제8회 아시아미래포럼

알고리즘과 사회적 가치

더불어 행복한 세상을 만드는 언론

한겨레미디어

아시아와 세계에 새로운 패러다임을 제시하고, 성장과 번영에 기여

한겨레신문사가 주최하고 한겨레경제사회연구원이 주관하여 올해로 제8회를 맞이하는 아시아미래포럼은 '번영과 지속가능성 사이의 균형 찾기'라는 취지 아래 아시아 국가의 공통 과제를 살펴보고 협력 방안을 모색하는 자리입니다.

2010년부터 매년 늦가을에 서울에서 개최되며, 해를 거듭하면서 세계적인 석학, 정부 관료 및 정치 지도자, 시민사회 활동가, 기업인, 시민사회의 멘토 및 청년들이 함께 지혜를 모으고 협력을 모색하는 지식네트워크의 허브로 자리매김하고 있습니다.

아시아미래포럼 조직위원회

공동위원장

김영호 한국학중앙연구원 석좌교수 **박용만** 대한상공회의소 회장 **양상우** 한겨레신문사 대표이사

조직위원

박영철 고려대학교 석좌교수 **박원순** 서울특별시장 **손경식** CJ그룹 회장
이 근 서울대학교 교수 **이정우** 경북대학교 명예교수 **임현진** 서울대학교 명예교수
정운찬 동반성장연구소 이사장 **김종구** 한겨레신문사 편집인 **이창곤** 한겨레경제사회연구원 원장

일의 미래: 새로운 사회적 합의를 향하여

우리 삶의 토대인 일이 변하고 있습니다. 세계화와 인구구조 변동에서 시작된 일의 변화는 최근 인공지능(AI), 로봇 등 4차 산업혁명 기술의 발전으로 한층 가속화하고 있습니다. 대리운전, 음식 배달 등 디지털 네트워크를 통해 조직되는 플랫폼 노동은 이미 우리의 일상이 됐습니다.

어떤 직업은 사라지고 어떤 직업은 새로 태어납니다. 디지털 네트워크의 도움으로 언제 어디서나 일을 하는 시공간의 유연화가 일어나고 있습니다. 노사의 고용관계도 달라지고 있습니다. 이는 한편으로 일하는 사람들의 선택폭을 넓히기도 하지만 사람이 기계에 일을 뺏기거나 노동이 파편화, 부차화할 우려도 있습니다. 일의 변화는 또 한 사업장에 고정된 정규직 일자리를 기반으로 설계된 노동 관련법과 연금, 실업급여 등 사회보장제도의 변화를 가져올 수밖에 없습니다.

이런 일의 변화는 우리에게 무엇이 '좋은 일'(Decent Work)인지를 묻고 있습니다. 임금이 적절하고 지위가 안정적이면서도 자녀양육이나 재충전을 위해 자유로이 시간과 공간을 디자인 할 수 있는 일이 우리가 바라는 '좋은 일'일 것입니다. 미래 우리의 삶을 노동, 여가, 가정생활이 균형을 이루고 사회 구성원이 성취감과 행복감을 느낄 수 있는 '좋은 일'로 디자인하는 것은 우리의 선택과 합의에 달려있습니다.

한국은 이미 경제협력개발기구(OECD) 국가 가운데 최장인 노동시간, 최저인 출산율과 결혼율을 경험하고 있습니다. 우리의 일과 생활이 이대로인 채 행복한 사회는 오지 않습니다. 만족한 직원의 창의성과 생산성에서 발원되는 혁신이 우리 기업에 필요한 때가 됐습니다.

우리 사회의 근간을 바꾸는 4차 산업혁명의 기술 변화가 행복한 일상으로 이어지는 법적, 제도적인 변화를 논의할 때입니다. 일자리를 어떻게 나눌지, 생애주기에 맞춰 근무시간과 근무공간을 어떻게 유연화할지, 근무시간 단축 등에 맞춰 국민 모두에게 기본소득을 제공해야 하는지 등 다양한 토론과 모색이 가능합니다. 독일은 '산업 4.0'으로 스마트 작업장을 구상하면서 동시에 '노동 4.0'으로 '좋은 일'에 대한 사회적 합의를 마련해가고 있습니다.

'포용적 성장', '사람 중심 경제' 등 지속해서 대안 담론을 제기해온 아시아미래포럼은 올해 '좋은 일'을 통해 우리 사회가 밝아지고 따뜻해지는 길을 찾아갑니다.

The 8th ASIA Future Forum
제8회 아시아미래포럼
아시아미래포럼

게티이미지

“알고리즘 논쟁, 인류·사회가치 우선해야”

세션 3. 알고리즘과 사회적 가치

16일 오전 진행될 아시아미래포럼 세션3은 한겨레 사람과디지털연구소 주관으로 ‘알고리즘과 사회적 가치’를 주제로 열린다.

인공지능과 자동화, 디지털 기술을 어떻게 다루고 적응하는가의 문제는 기계를 움직이는 알고리즘에 어떠한 사회적 가치를 담느냐의 문제로 귀결된다. 카이스트 교수와 소프트웨어정책연구소장을 지낸 김진형 지능정보기술연구원 원장이 “소프트웨어로 사회문제를 해결할 수 있나?”라는 주제로, 과학기술의 사회적 영향을 연구하는 전치형 카이스트 과학기술정책대학원 교수가 “알고리즘에 어떠한 사회적 가치를 담을 것인가”를 주제로 발표한다. 이성웅 한국아이비엠(IBM) 상무, 황용석 건국대 미디어커뮤니케이션학과 교수, 오세욱 한국언론진흥재단 연구위원이 토론자로 나선다.

김진형 원장은 인공지능과 소프트웨어 기술을 긍정적으로 바라보며 양면성을 지닌 기술을 어떻게 인류와 사회문제의 해결에 활용할 수 있을지를 다룰 예정이다. 컴퓨터를 움직여 실제 가치를 만들어내는 것이 소프트웨어인데 소프트웨어는 비배제성과 비경합성이라는 공공재적 성격을 갖고 있다. 하지만 현실에서는 기술 개발에 대한 개입과 간섭을 배제하고 시장에 맡겨 경쟁을 기술 혁신의 동력으로 삼자는 주장과 영향력과 안전성을 검토해 규제를 강화하자는 요청이 맞서고 있다. 인공지능이 가진 가치와 가능성과 위험성에 대한 종합적 고려를 통해 인간의 가치, 민주주의의 가치를 촉진하는 방향을 모색하자는 게 김 원장의 주장이다.

전치형 교수는 실용화가 임박한 자율주행차의 사례를 통해, 기술에 깊이 스며들어 있는 사회적 가치와 시스템이 구체적 상황에서 어떻게 작동하고 있는지를 설명한다. 자유주행차는 ‘사고 시 누구를 희생시킬 것인가’라는 ‘전차 문제’ 딜레마와 함께 소개된다. 하지만 이는 자율주행차의 알고리즘에 좀더 훌륭한 가치를 담는 것으로 해결할 수 있는 문제가 아니다. 현재 교통사고에서 만나는 긴급상황은 순간적이고 불가피한 것이 아니라 교통정책과 시스템, 습관이 누적된 결과라는 점에서 안전은 자율주행차 알고리즘에 집어넣어서 손쉽게 구현할 수 있는 게 아니라는 게 전 교수의 주장이다. 사회적 가치는 개별자동차가 아니라 전체 시스템 속에 걸쳐 종합적으로 구현되어야 하는 거대한 것이기 때문이다. 자율주행 기술이 간과하고 있지만 중요한 것은 이미 복잡하게 얽혀 있는 사회적 가치와 시스템에 담긴 문제다.

구본권 사람과디지털연구소장 starry9@hani.co.kr

〈한겨레〉 제8회 아시아미래포럼 특집호 2017년 11월 7일 화요일

The 8th ASIA Future Forum

제8회 아시아미래포럼

알고리즘과 사회적 가치

더불어 행복한 세상을 만드는 언론

한겨레미디어

Contents

세션 3

프로그램 008

발제 012
김진형 / 지능정보기술연구원장
전치형 / 카이스트 과학기술정책대학원 교수

토론 054
황용석 / 건국대학교 미디어커뮤니케이션학과 교수
오세욱 / 한국언론진흥재단 선임연구위원
이성웅 /한국 IBM 상무

The 8th ASIA Future Forum
제8회 아시아미래포럼
프로그램

11월 16일(목)

세션3 1부 알고리즘과 사회적 가치 09:30 ~ 11:20

| 발제 |

김진형
지능정보기술연구원장

| 발제 |

전치형
카이스트 과학기술정책대학원 교수

| 토론 |

황용석
건국대학교 미디어커뮤니케이션학과 교수

| 토론 |

오세욱
한국언론진흥재단 선임연구위원

| 토론 |

이성웅
한국 IBM 상무

세션3 2부 제3회 휴먼테크놀로지 어워드 11:30 ~ 12:00

| 시상자 |

윤종수
휴먼테크놀로지어워드 심사위원장

| 시상자 |

양상우
한겨레신문사 대표이사

| 시상자 |

김종구
한겨레신문사 편집인

시간	프로그램		
	세션 3 (중회의실 A)		휴먼테크놀로지 포럼 & 어워드 (사람과디지털연구소)
			1부. 휴먼테크놀로지 포럼: 알고리즘과 사회적 가치
		발 제	**김진형** 지능정보기술연구원장 소프트웨어로 어떻게 사회문제 해결할 수 있나?
			전치형 카이스트 과학기술정책대학원교수 자율주행차로 무엇을 할 것인가?
		토 론	**황용석** 건국대학교 미디어커뮤니케이션학과 교수
			오세욱 한국언론진흥재단 선임연구위원
			이성웅 한국 IBM상무
			2부. 제3회 휴먼테크놀로지 어워드
		시상자	**윤종수** 휴먼테크놀로지 어워드 심사위원장
			양상우 한겨레신문사 대표이사
			김종구 한겨레신문사 편집인

The 8th ASIA Future Forum

제8회 아시아미래포럼

세션 3

사람과디지털연구소

중회의실 A

알고리즘과 사회적 가치

발제

김진형 / 지능정보기술연구원장

전치형 / 카이스트 과학기술정책대학원 교수

토론

황용석 / 건국대학교 미디어커뮤니케이션학과 교수

오세욱 / 한국언론진흥재단 선임연구위원

이성웅 /한국 IBM 상무

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

발 제

김진형

지능정보기술연구원장

한국과학기술연구원(KIST) 출신 1세대 프로그래머로 미국 UCLA 대학원에서 컴퓨터 사이언스로 박사학위를 받고, 1985년부터 KAIST 교수를 지냈으며 현재 명예교수다.

인공지능–패턴인식 연구를 전공했으며, KAIST 인공지능연구센터 소장을 지냈다. 미래부 산하 소프트웨어정책연구소 소장을 지냈고, 2009년 (사)앱센터를 설립하고 2016년까지 이사장으로 활동하며 소프트웨어 교육의 중요성을 설파해왔다. 앱센터 안에 소프트웨어(SW)교육봉사단을 설립해 초대 단장을 지냈다.

2010 국가정보화 전략위원회 전문정보위원장, 2013년 공공데이타전략위원회 위원장, 국가과학기술심의회 위원 등을 지냈으며 정보과학회, 인지과학회, 의료정보학회 회장을 지냈고 국제패턴인식학회 석학회원(Fellow), 공학한림원과 과학기술한림원 회원이다.

초록

소프트웨어로 사회문제를 해결할 수 있나?

기술은 인간이 필요를 충족시키고 자연을 변화시키는 도구이자 과정인데, 그 중에서도 컴퓨터는 만능도구이다. 그 컴퓨터를 움직여 실제 가치를 만들어내는 소프트웨어는 점점 인공지능화하고 있다. 비배제성과 비경합성이라는 공공재적 성격을 지닌 소프트웨어는 산업 발전만이 아니라 다양한 사회문제를 해결하는 용도로 활용할 수 있다. 강력한 기술의 복잡성은 기술 개발을 바라보는 두 견해를 낳는다. 개입과 간섭을 배제하고 시장에 맡겨 경쟁을 기술혁신의 동력으로 삼자는 주장과, 영향력과 안전성을 검토해 규제할 필요가 있다는 주장이다. 시장만능주의와 대중적 공포 모두 위험하다. 인공지능이 가진 가치와 가능성과 위험성을 종합적으로 고려해 인간과 민주주의의 가치를 실현하는 방향을 모색해야 한다.

소프트웨어로 사회문제를 해결할 수 있나?

2017. 11

김 진 형
AIRI 원장. KAIST 명예교수
jkim@AIRI.kr

AIRI 지능정보기술연구원

CC BY

기술(Technology)이란 무엇인가?

- 인간이 자신의 필요를 충족시키고 자연을 변화시키는 과정
  - 인공물(Artifacts) : 컴퓨터, 소프트웨어, 비행기, 피임약, ...
  - 지식과 과정 :엔지니어링 노하우, 설계 및 제조 전문 지식, 기술력
  - 사회 기반 시설 : 설계, 제조, 운영, 수리 등을 위하여

We have become a people unable to comprehend the technology we invent

모든 기술은 무죄가 입증될 때까지
유죄로 간주되어야 합니다
by David Brower(환경운동가)

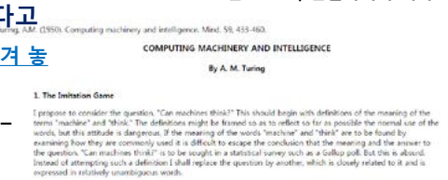
2 AIRI

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

컴퓨터는 다기능 기계

- Programmability - 범용기계(Universal Machine) 이론 제안
  - 영국의 수학자 Alan Turing 1936년
  - 지시사항(프로그램)이 달라짐에 따라 다른 기계처럼 작동하는 기계
- 2차대전 종료 후 각국에서 경쟁적으로 구현
  - 컴퓨터 산업 태동
  - 반도체 기술 발달로 소프트웨어 중요성 부상
- 튜링은 이 기계로 인공지능이 가능하다고
  - 프로그램을 통해서 사람의 지능을 기계에 옮겨 놓으면 그 기계는 지능적 행동 수행 가능
  - 기계의 지능검사로 Imitation Game을 제안 – 1950년

1946년 ENIAC, 펜실바니아 대학


Turing, A.M. (1950). Computing machinery and intelligence. Mind, 59, 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

Wiki

3 AIRI

상상을 현실로 만들어 주는 소프트웨어

소프트웨어의 정의

- 컴퓨터 하드웨어 = 계산과 논리연산(비교)만 수행하는 단순한 기계
- 알고리즘 = 문제를 계산과 비교의 순차적 수행을 통하여 해결하는 방법
- 컴퓨터 프로그램 = 알고리즘을 컴퓨터가 이해할 수 있는 언어로 표현(코딩)한 것
- 소프트웨어 = 컴퓨터 프로그램 + 데이타 + 관련 문서

소프트웨어의 본질

- 지식과 경험을 코딩한 문제 해결책(솔류션)
- 생각을 자동화한 것

디지털 표현의 장점

- 저렴하게 구현 가능
- 복제하여 공유 가능
- 원격에서 접근, 동시 사용 가능
- 자유로이 혼합, 누적하여 성능 추가
- 점진적 개선 가능

소프트웨어는 전 인류가 이룬 지식/해결책의 집대성(集大成) !!

4 AIRI

소프트웨어의 특징(1)

- 창작(두뇌) 활동의 작품 – 정보財, 지식財
  - 물리적인 것이 아니라 논리적인 것
  - 기계와 프로그래머 사이에서 연계 역할
  - 높은 생성 비용, 낮은 복제 비용 - 투자는 risky
- 가치는 무엇인가가 아니라, 무엇을 하는가로 결정
  - 개인과 조직을 위하여 밀접하게 임무 수행 – 사용자의 협조 필요
  - 가치는 사용자에 따라 다름 : 공유, 재사용으로 증가
  - 경험하지 전에는 가치를 인정하지 않음
  - 계약시 정보의 비대칭성 - 개발될 SW의 소비자 가치 불명
- 네트웍 효과, 특정 상품에의 고착 현상
  - 승자독식의 시장
- 소프트웨어와 하드웨어는 상호 교체 가능

5 AIRI

소프트웨어의 특징(2)

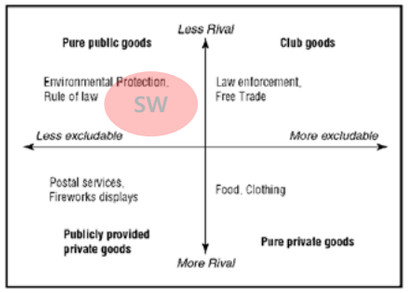
- 도구산업 - 타 산업의 효율과 성과를 위하여 사용
- Industry Best Practice를 상품으로 판매
  - 개별 요구사항은 수정 보완 용역으로
- 시장이 수평적 구조, 기업은 한 두 계층에 집중
- 플랫폼 SW에 따라서 독자적 Community 구성
  - MS Window 계열, Apple 계열, Android, SAP-based, …
- 소프트웨어 개발자의 역량이 산업의 역량
  - 다른 어느 산업보다 사람의 능력이 중요
- 신기술 출현 속도가 빠르고 시장이 일시에 괴멸되는 High Risk, High Return 산업
- 공개, 공유의 철학이 상업적 경쟁과 공존
- 인터넷 서비스(Saas) 영업형태로 급속히 변환 중
- …

공개 소프트웨어
(Open Source Software)

6 AIRI

소프트웨어의 공공재적 특성

- SW는 Non-Excludability(비배제성)가 있다
  - 개발되면 (요금을 지불하지 않아도) 누구나 사용할 수 있다. (ex: 음악, 도로)
- SW는 소비의 Non-Rivalry(비경합성)가 있다
  - 한 사용자가 사용한다고 다른 사용자가 사용할 수 있는 양이 줄어들지 않는다. (ex: 음악)
  - 공공재는 가격을 0 으로 해서 모두 사용하는게 이익
- 공공재의 특성
  - 시장에 맡기면 공급 안됨 -> SW는 저작권 제도로 공급자 보호



7 AIRI

Computing은 혁신의 기술
(Skills of Innovation)

새로운 혁신의 방법

코딩, 프로그래밍	공학기초, 컴퓨팅 사고력	성장 마인드
데이타과학 (데이터 활용 능력, 빅데이타, 데이터 시각화, 기계학습/딥러닝, 인공지능)	Making, 실물 제어 컴퓨팅(Physical Computing) & 사물인터넷	디자인 사고력
앱 개발 & 게임화	디지털 (첨가적) 제조법 – 3D 프린팅	기업가 정신
신뢰성, 개인정보보호, 보안	새로운 응용 – 로봇, 드론, 무인차, 입는 컴퓨터	세계 시민 정신 (창의력 포함)

= 실세계 문제 해결을 위한 팀 프로젝트 수행

↓

SMART COMMUNITIES

Next generation of innovators . Intel Corporate Affairs Group By Ann Marshall.Anshul Sonak, May 2016

8 AIRI

AI는 사람의 지능적 행동을 흉내 내는 기술

AI 는 고도의 자동화 기술

AI는 최적화, 문제해결 기술

AI는 사람 같은 상호작용의 기술

인간은 지능은 다차원 – 다양한 능력의 교향곡

인지능력, 언어 구사력, 추리능력, 계산력, 공간감, 음감 등 100여 가지로 구성

현재의 인공지능 시스템들은 두 서너 개의 능력으로 구성됨

– Kevin Kelly (2016)

그러나 기계지능이 꼭 사람지능과 같을 필요는 없다

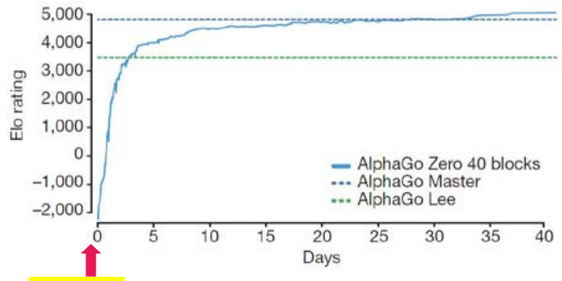
9 AIRI

AlphaGo는 바둑계를 평정하고 은퇴

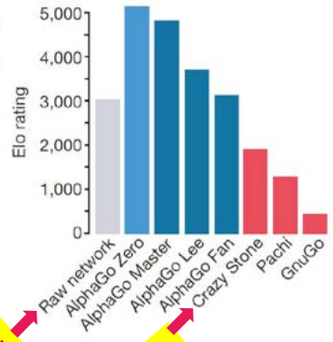
10 AIRI

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

AlphaGo Zero 능력 성장 추세(1)



Random Move



신경망만 사용
MCTS 사용안함

MCTS 사용,
학습기능 없음

11 AIRI

AlphaGo Zero 능력 성장 추세(2)

80수 이후의 판세

훈련 시간 :
3시간

상대 알을 잡는데 급급

19시간

생과 사의 개념 파악
세력과 집수의 균형,
포석, 맥(脈), 선수(先手),
축, 끝내기 등을 학습

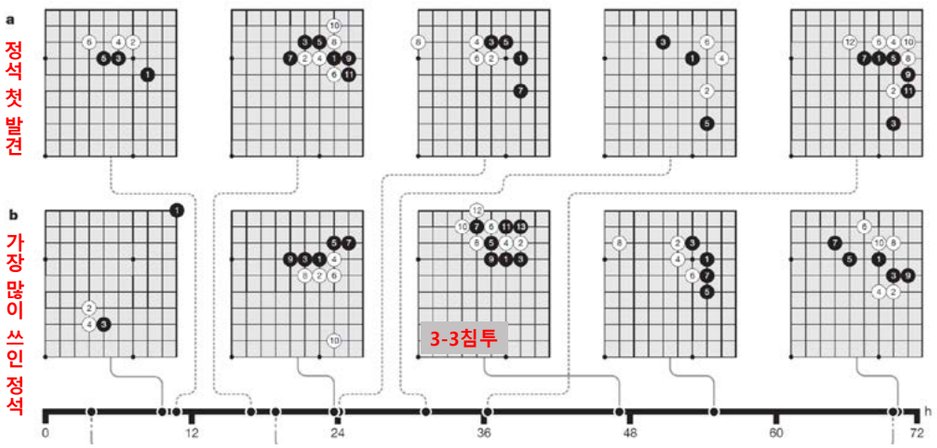
70시간

여러 곳에서의 전투 수행
패 싸움 개념 학습

David Silver, et al, "Mastering the game of Go without human knowledge", Nature 550, 354-359, 19 October 2017

12 AIRI

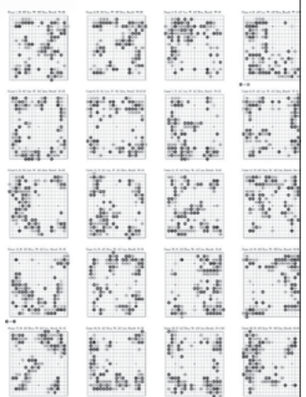
AlphaGo Zero의 정석 발견



David Silver, et al, "Mastering the game of Go without human knowledge", Nature 550, 354-359, 19 October 2017

AlphaGo 승리의 의의

- 인공지능이 사람을 이길 수 있는 경지에 도달
- 인공지능이 학습할 수 있다
- 사람의 지식을 인공지능이 이용할 수 있다
- 인공지능이 맨땅에서 지식을 쌓기 시작 가능 (Learn from Scratch)
- 인공지능이 인공지능 간의 경쟁으로 지식 고도화
- 인류가 몇천년에 걸쳐서 쌓은 지식을 몇일만에 생성
- 유사한 기술이 다른 구조화된 문제(예: 단백질 분석, 에너지 효율화, 신소재 탐색 등)에 적용 가능
- 사람이 인공지능으로부터 배워야 한다
- 사람과 인공지능이 같이 일해야 한다



The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

인공지능이 창작한 그림

By Google

By Rutgers University

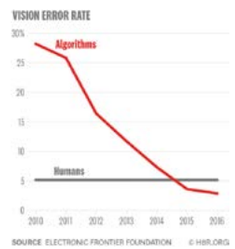
http://www.businessinsider.com/google-to-auction-trippy-artwork-made-by-ai-2016-2/#inceptionism-can-work-one-of-two-ways-the-first-way-is-to-feed-googles-neural-network-an-image-and-ask-it-to-look-for-something-specific-1

https://news.artnet.com/art-world/rutgers-artificial-intelligence-art-1019066

데이터 기반 AI의 성공사례

◆ 목표 : 이미지 내 존재하는 객체 분류

- 1000 종류의 클래스

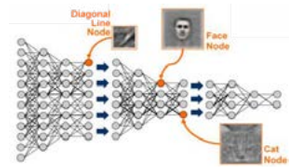


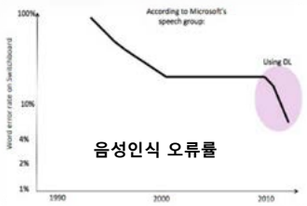
음성 및 영상처리에서 사람을 능가

MS ResNet Object detection
ILSVRC2015, 1000 categories, 150k images

딥러닝의 성공사례

3일간의 You-Tube 영상 보여줌으로 "Cat Detector" 실현

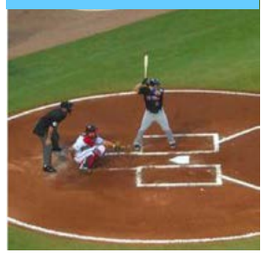




The man at bat readies to swing at the pitch while the umpire looks on.

A person riding a motorcycle on a dirt road.

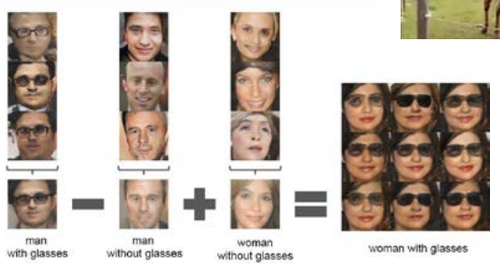
야구 선수가 공을 치려고 합니다

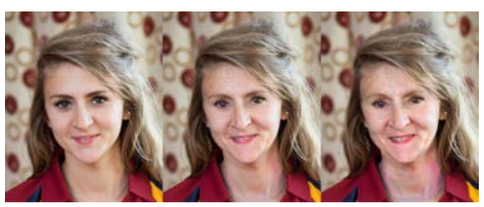


눈 덮인 언덕 아래로 스키를 타는 사람

영상, 동영상의 내용 서술

Generative Adversarial Network



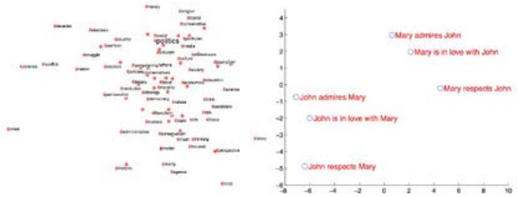


https://github.com/paulu/deepfeatir 나이 든 모습, 젊었을 때 모습 생성

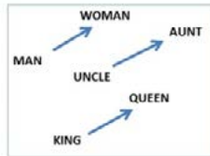
The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

Deep Learning : Language Processing

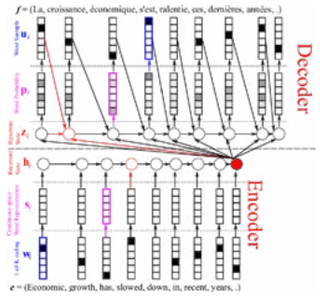
- 단어 의미를 다차원 공간에서 '벡터화' - Word to Vector



vec("man") – vec("king") + vec("woman") = vec("queen")



- Recurrent Neural Network 기반의 Encoder-Decoder Architecture를 이용하여 기계번역에서 성과



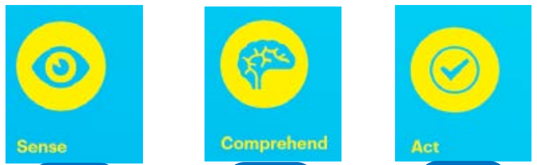
인공지능은 새로운 시장을 창출하고
기존의 것들은 더 효율적으로 만든다

또 많은 상품과 서비스의 가격을 낮춰서
모든 사람을 더 부유하게 만든다

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

이제 "소프트웨어 = AI"

AI 기능



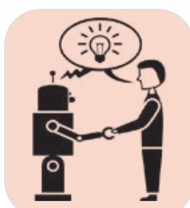
인식 이해 행동

- "오늘날 소프트웨어 앱의 1%만이 AI 기능을 사용한다. 그러나 2018 년에는 50%가 될 것이다"
  - IDC 연구를 NY Times, 2016-03-26 에서 인용

All companies are now Software companies
All companies are now AI companies
:인공지능을 생산하거나 활용하는 기업

21 AIRI

Human and Machine Working as a Team



사람이 못하던 업무를
기계의 도움으로 성취

'68승 1패' 바둑 정복한 알파고…기보 50편 남기고

알파고 은퇴… 커제에 3전 전승

▲ ①33이 명당이다 (알파고와 알파고 대국 중 제3국)
시작하자마자 33에 침투하면 십중팔구 '초짜구나' 하고 생각했다 33 집 있은 집을 찾기에 가장 좋지만 상대방의 세력을 강하게 하고 자신은 귀퉁이에 고립되는 결과를 초래하는 '소탐대실'이 되기 십상이다. 그런데 알파고는 33 침입을 즐긴다. 초반 33 침입이 이득이라고 판단했다는 점에서 바둑의 패러다임 하나를 무너뜨렸다

2018년까지 3백만명 이상의 노동자가 Roboboss에 의하여 관리될 것
- Gartner, 2016

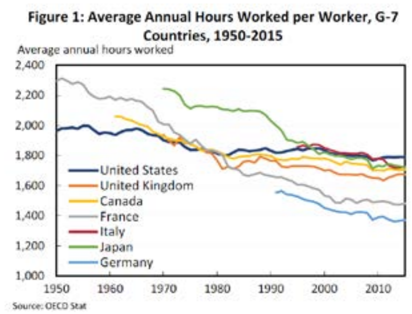
기계가 먼저 배우고 발견한 것을 인간에게

자료 : Gartner's Top Prediction, by Werner Goertz 2016; http://www.hankookilbo.com/v/dd6570858e254948906a875ce5b6af8a

22 AIRI

기술 발전으로 근로시간 감소

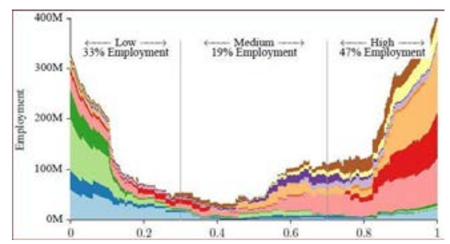
수렵사회 365일, 24시간 농경사회 봄·여름·가을, 낮 시간 산업사회 주 40시간



http://radar.oreilly.com/2012/11/unemployment-education-stem-investment-polic.html
http://www.hani.co.kr/arti/science/science_general/777109.html#csidxa394c577ce7c0d780ff14086e784b10

23 AIRI

일자리 급격히 감소



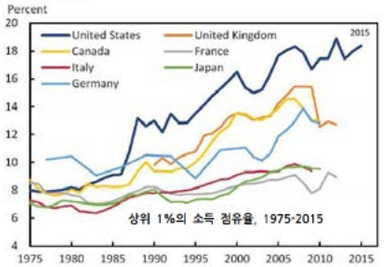
- 우리나라도 10년 안에 노동자의 70%, 즉 1800만명의 일자리 위협[2]
- 노동시장 기회의 양극화
- 청소년 미취업, 하향 취업, 미활용
- 많은 가정이 적은 소득
- 직업/노동 형태의 변화 : Digital Platform 위에서 독립 근로 형태
- 노동 유연성
  - 한국의 노동유연성 : 83위
  - By USB, 제4차 산업혁명 준비 지수

1. The Future of Employment : How susceptible are jobs to computerization?. by Carl Benedikt Frey & Michael A. Osbone, Oxford University, Sep, 2013
2. 한국고용정보원, 2017,1

24 AIRI

소득 양극화

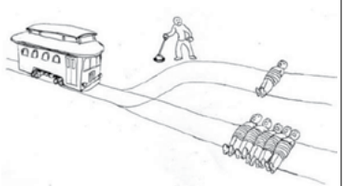
- 고소득을 올리는 극소수
  - Superstar-Biased Technological Change
  - 0.01%가 5% 소유
  - 1%가 18% 소유
- 국가간의 격차도 심화
  - 패권주의의 회귀 ?
- 사회적 갈등 심화
- 사회적 안전망 확충 노력
  - 기본소득, 무상교육, 로봇세, ...



Source: World Wealth and Income Database.

인간-기계 공생(?)의 Dilemma

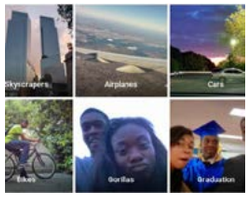
- 구글 무인자동차 330만Km 주행 첫 사고
  - 고속 운행 중인 버스가 서행으로 바꿀 것으로 기대
  - "인간처럼 적당히 합리적으로 행동해라?"
- 윤리적 행동을 프로그램 할 수 있을까?
  - Trolley Problem : 5명 사망 vs 1명 사망
- 감시카메라가 "생계형 범죄 장발장"을 눈 감아 주어야 하는가?
- 인공지능이 만든 저작물의 권리는 ?
- 로봇 행동 3원칙
- ...



The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

데이터 기반 AI의 한계

- Garbage In, Garbage Out



흑인 여성 → 고릴라

재범 예측 AI, , ...

- Back Box System
  - 왜 이런 결정을 했는지 설명할 수 없다
  - 어디까지 할 수 있는지 예측 불가
- Heavy Computation required
- 사람처럼 하지 않는다(?)
  - "어린 아이들이 고양이를 아는데 몇 백만장의 고양이 사진이 필요하지 않다"

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing http://www.wnyc.org/story/deep-problem-deep-learning/

AI 시스템은 안전한가? 도덕적인가?

도요다의 급발진 사고

폭스바겐의 배기 가스 사기

AI in Science Fictions

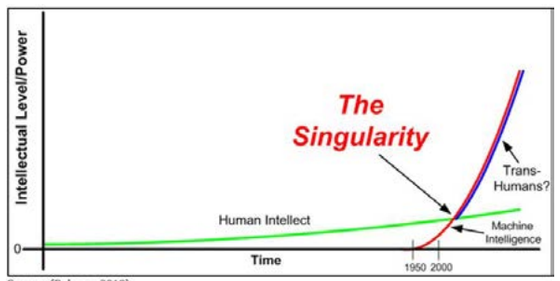
인공지능이 가까운 장래에 인간에게 절박한 위협이 되지 않을 것이다

AI100 Stanford Report, 2016.9

* https://ai100.stanford.edu/

29 AIRI

The Singularity is Near



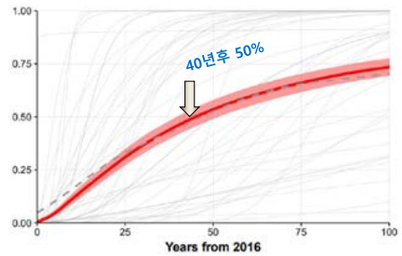
Source: (Rebane, 2010)

인공지능은 어디까지 발전할까?

30 AIRI

사람보다 잘하는 AI는 언제나 가능한가?
AI 전문가 350명의 예측(2016년)

업 무	년 도
언어 번역	2024
고교 에세이 작성	2026
트럭 운전	2027
톱 40 팝송 작곡	2027
소매 점포에서의 업무	2031
베스트 셀러 소설 집필	2049
외과의사	2053
AI연구원	2103



모든 업무에서 사람보다 잘 할 확률

- 120년 후에는 모든 인류의 직업이 자동화 예상

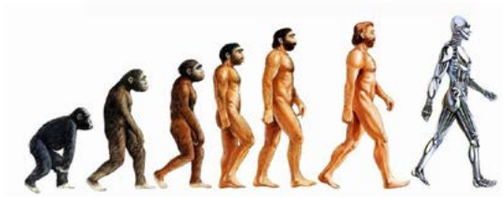
궁극적으로 AI는 인류에게 매우-이롭다(20%), 이롭다(25%), 해롭다(10%), 매우-해롭다(5%)

When Will AI Exceed Human Performance? Evidence from AI Experts, Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, Owain Evans *30 May 2017*
https://arxiv.org/abs/1705.08807

31 AIRI

Automation vs Augmentation

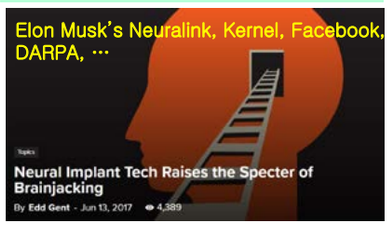
Transhumanism (H+)

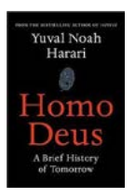


'10년 후에는 지금 상상도 할 수 없는 새로운 인간의 형태 (신인류·New Type)가 탄생한다.'
- Ray Kurzweil 2017.7.7

매일경제
인공지능이 일을 하고 인간은 500살까지 생명 연장할 것
◆ 매경실리콘밸리포럼 / 4차산업혁명 메카 실리콘밸리를 가다 ⑥ 포럼 둘째날
레이 커즈와일 구글 이사 강연 ◆

Brain-Computer Interface :
인간의 두뇌를 클라우드로 연결하는 기술





Homo Sapiens → Homo Deus (永生을 갖은 神 같은 인간)

32 AIRI

AI의 발전을 규제해야 하는가?

- 자율적 의사결정 능력+ 데이터로 학습 ➔ 개발자의 의도를 벗어날 가능성 존재
  - 사례 1 – 인공지능 한계 드러낸 MS 채팅봇 Tay
  - 사례 2 – 중국 텐센트 채팅봇 BabyQ "공산당 부패했다"
- AI의 발전을 시장에 전적으로 맡기는 건 위험, "정부가 AI 개발 규제하자"
  - "소수 엘리트가 AI의 힘을 지배하기 때문에 위험"
  - "선제적 규제가 필요하다. 늦으면 속수무책"
- AI 안전성 평가하기 위한 새로운 FDA 설립하자
  - AI Guardians(감시관)이 필요하지만 규제 체제는 필요하지 않다[1].
- 깊은 이해를 기반으로 정책을 실행해야
  - 대중의 무지나 공포에 기반을 두면 안됨

1 : https://www.forbes.com/sites/ciocentral/2017/01/09/why-regulating-ai-is-a-mistake/#277713e2be38

33 AIRI

SW, AI의 활용으로 사회문제 해결 도전

34 AIRI

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

AI 기술의 진보는 공개 공유를 통하여

arXiv.org GitHub TensorFlow

공개 소프트웨어 (AI)

기계학습을 위한 공개SW tool
- Google의 Tensor Flow, Microsoft의 CNTK, DMTK, Facebook의 Torch, Berkeley Vision Center의 Caffe, Skymind의 Deeplearning4j, ...

얼굴인식 공개SW 및 Tool
- CMU의 openFace. YOLO, openCV, openBR, openGL, ...

음성인식 공개SW 및 Tool
- CMU Sphix, Microsoft HTK, KDE.org의 Simon Project, KALDI, Julius, ...

자연어처리 공개SW 및 Tool
- Stanford Core NLP Suite, NLTK(Natural Language Toolkit), Apache OpenNLP, ...

...

공개 데이타

DBpedia - Large,Multilingual, Semantic Knowledge Graph
Google audio set
Image net
Kaggle
Amazon Public datasets
Human3.6m - 3D human pose와 image
LSMDC
Hollywood2 Human Actions and Scenes Dataset
US Government open data
Reddit
NASA Prognostic Data Repository
Hard disk test data
Quora
...

공개 논문

출판전 논문 공개 관행으로 기술 전파 가속화 : 공개 장소 arXiv.org
논문 발표 시에 소스코드, 데이터도 같이 공개하는 관행

공개 경진대회

DBpedia Open Text Extraction Challenge, Kaggle, Grand Challenge, IMAGENET, Large Scale Movie Description and Understanding Challenge, ...

공동 프로젝트

Github 26 million Users , Sourceforge 43만개 프로젝트, Apach, GNU, ...

Open AI Promotion Community

OpenAI, AI•ON, Partnership on AI, ...

35 AIRI

공개 소프트웨어, 공개 데이터의 가치

공개
- 투명성, 특히 인공지능에서
- 연구결과의 100% 재현 가능
  - 논문은 공개 소프트웨어의 설명서
- 기술의 검증, 명성의 획득
- 빛의 속도로 전파

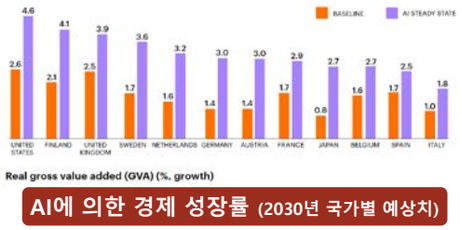
공유
- 기술의 민주화, 사회적 자산화
- 누구나 최고의 기술을 사용할 수 있고
- 최고의 기술에서 새로이 연구 시작 가능

참여
- 공동작업으로 다양한 아이디어의 융합
- 신속한 개발
- 다수의 눈으로 오류 검출

36 AIRI

AI에 의한 경제 성장

기술에 의한 생산성 향상



AI에 의한 경제 성장률 (2030년 국가별 예상치)

미국 8500조원(~2035년)
한국 460조원(~2030년)

경제 파급 효과

자료 : https://www.accenture.com/lv-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.pdf
https://obamawhitehouse.archives.gov/blog/2016/12/20/artificial-intelligence-automation-and-economy
https://www.mckinsey.com/global-themes/employment-and-growth/technology-jobs-and-the-future-of-work?cid=other-eml-alt-mgi-mgi-oth-1706&hdpid=1-1703636608543891748145307230&hdpts=1633231&hdpid=4093168-782-4726-aa3-320f2a154802

37 AIRI

대응을 잘 하면 기회가 된다

- 창의·혁신의 "창조적 파괴"가 실현되어 역동적이고 정의로운 사회를 만들어야
- AI의 작동구조, 능력, 가치, 위험, 한계를 이해하고 활용할 수 있는 능력의 보유
  - AI는 노동·토지·자본에 더한 새로운 생산요소
- AI시대를 향한 전환의 복잡성 이해
  - 사회적·문화적 영향 이해해야
  - 인간의 가치, 민주주의의 가치를 촉진하는 방향으로 AI 활용
- *"우리가 가진 AI는 앞으로 25년 간 있을 것에 비하면 아무 것도 아니다"* – Kevin Kelly

참조 : PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE, EXECUTIVE OFFICE OF THE PRESIDENT, NATIONAL SCIENCE AND TECHNOLOGY COUNCIL October 12, 2016

38 AIRI

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

발 제

전치형

카이스트 과학기술정책대학원
교수

KAIST 과학기술정책대학원 조교수이다. 2010년에 매사추세츠공과대학교(MIT)에서 과학, 기술, 사회(STS) 전공으로 박사학위를 취득했다. KAIST 과학기술정책대학원 이전에는 독일 베를린 막스플랑크 과학사연구소(Max Planck Institute for the History of Science)에서 1년 간 박사 후 과정 연구원으로 있었다.

주요 연구 분야는 사회 · 문화적 맥락 속 인간과 기술의 관계와 인간–기술–사회 구성과 관련있는 정책 이슈들이다. 또한, 항공기 조종사, 자동차 운전자부터 CCTV 모니터 요원, 의사에 이르기까지 유능한 기계 운영자의 육성 및 활용을 위한 과학적 연구, 엔지니어링 실무, 선정 및 교육 기법, 제도적 · 문화적 · 정책적 장치에 대한 연구 역시 진행하고 있다. 그 외 공학 정치, 한국 로봇 정책, 원격 기술 및 문화(의학, 교육 및 보안 분야) 등도 그의 연구 분야이다.

초록

'자율주행차 알고리즘에 무엇을 담을 것인가?'

자율주행차가 약속하는 세상은 사람 운전의 위험성과 불안을 제거한 안전하고 편리한 세상이지만, 사고 시 누구의 목숨을 더 중시해야 하는지에 관한 '트롤리 문제'와 일자리 축소에 대한 불안을 함께 품고 있다. 이는 자율주행차의 알고리즘에 좀 더 훌륭한 가치를 담는 것으로 해결할 수 있는 문제가 아니다. 교통사고 때 겪는 긴급상황은 순간적이고 불가피한 것이 아니라 교통정책과 시스템, 습관이 누적된 결과라는 점에서 안전은 자율주행차 알고리즘에 집어넣어서 손쉽게 구현할 수 있는 게 아니다. 사회적 가치는 개별 자동차가 아니라 전체 시스템 속에 종합적으로 구현되어야 하는 거대한 것이다. 자율주행 기술이 간과하고 있지만 이미 복잡하게 얽혀 있는 사회적 가치와 시스템의 문제를 살펴본다.

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

아시아미래포럼 2017

알고리즘과 사회적 가치

자율주행차로 무엇을 할 것인가

전치형
KAIST 과학기술정책대학원
cjeon@kaist.ac.kr

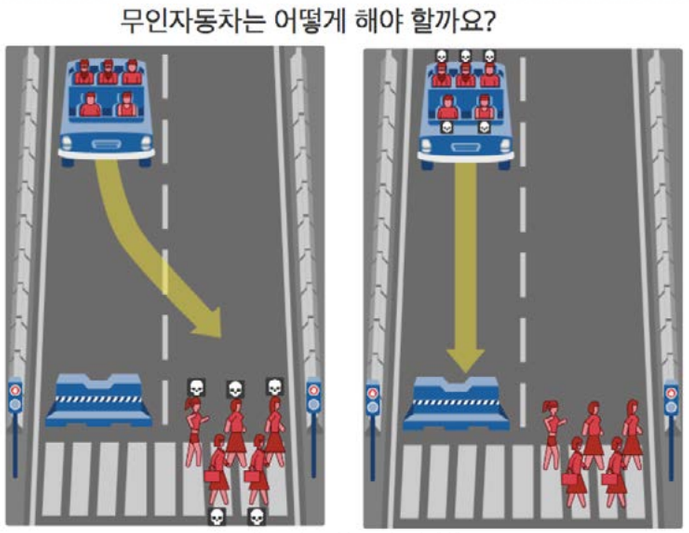
알고리즘에 가치를 담을 수 있을까?

알고리즘에 어떤 사회적 가치를 담아야 할까?

자율주행차와 사회적 가치

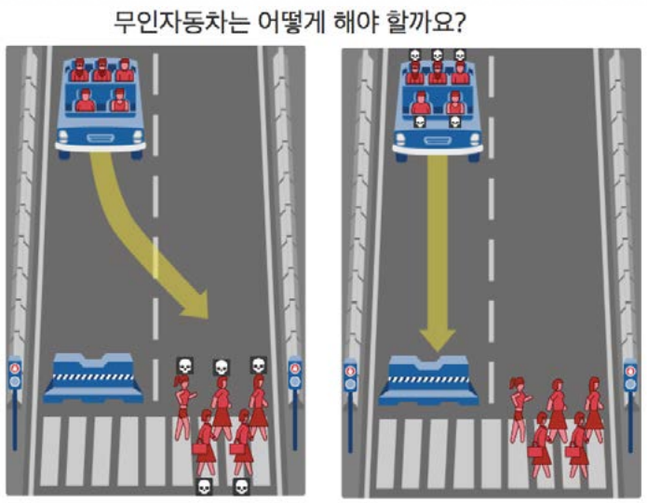
- 자율주행에 담으려는 가치—생명, 윤리, 안전 등.
- 자율주행이 바꿔놓을 가치—운전의 사회성, 인간의 자율성.

‘트롤리 문제’: 누구를 살리고 누구를 죽일 것인가?



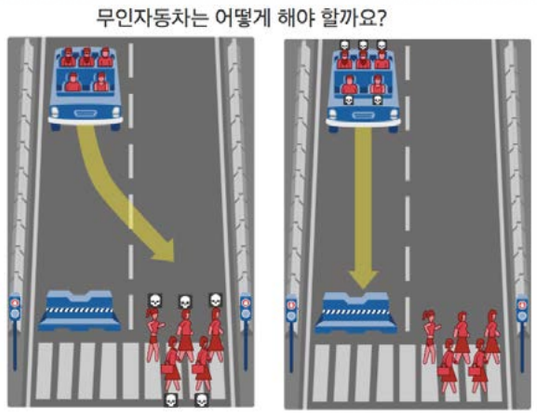
http://moralmachine.mit.edu/hl/kr

어떤 가치를 담으려 하는가, 담을 수 있는가?
안전, 생명, 효용 ...

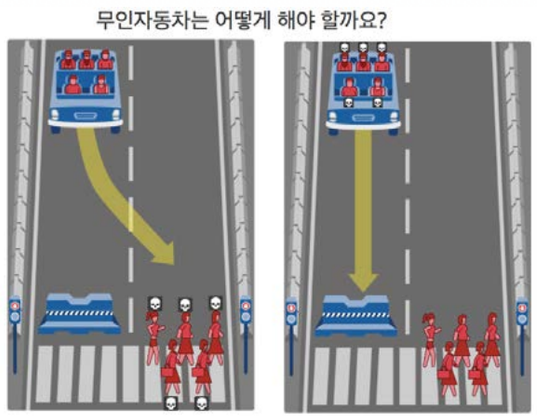


The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

‘트롤리 문제’만 해결하면 과연 자율주행이 완성될까?

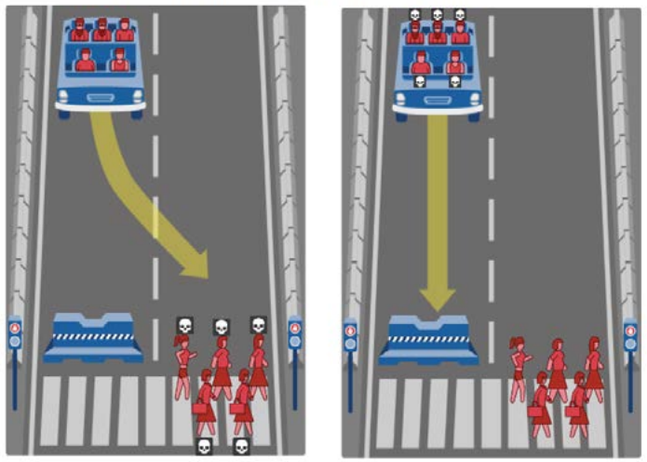


자율주행차가 긴급상황에서 현명하게 판단할 것을 기대하기 전에 해결해야 할 문제들이 많이 있다.

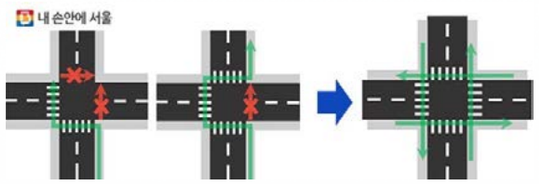


이런 긴급상황은 순간적이고 불가피한 것이 아니다.
교통 정책과 시스템과 습관이 누적된 결과이다.

무인자동차는 어떻게 해야 할까요?



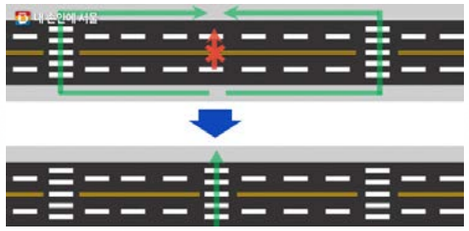
가령, 보행자가 안전하게 건널 수 있는 횡단보도가
충분히 많이 설치되어 있고 적절하게 관리되고 있는가?



http://opengov.seoul.go.kr/mediahub/11428755

http://img.yonhapnews.co.kr/etc/inner/KR/2017/05/31/AKR20170531044400052_02_i.jpg

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3



opengov.seoul.go.kr/mediahub/11428755

"그동안 횡단보도는 자동차에 밀려 있었다. '원활한 교통 흐름에 방해가 된다'는 지적 때문에 횡단보도 설치가 쉽지 않았다. 윤준병 서울시 도시교통본부장은 '기존엔 보행자보다는 차량이 우선이었던 게 사실이다. 이제는 보행자를 우선하는 패러다임 전환기에 접어들었다'고 말했다."

중앙일보 "ㄷ자 99초, X자 19초 … 횡단보도만 바꿔도 걷기 즐겁다"
http://news.joins.com/article/21609709

"현행 도로교통법 시행규칙 제11조에 따르면 어린이·노인·장애인 보호구역 등 특별한 사유가 없는 한 횡단보도 설치 장소로부터 반경 200m 내에는 횡단보도를 추가로 설치할 수 없다.

이같은 거리 제한 규정 탓에 횡단보도 확대 설치가 어려워 도심권에서는 도로를 건너가려면 먼 거리로 우회하거나 일부 보행자는 무단횡단을 하다가 교통사고를 당하는 경우가 많다는 게 경찰측의 판단이다."

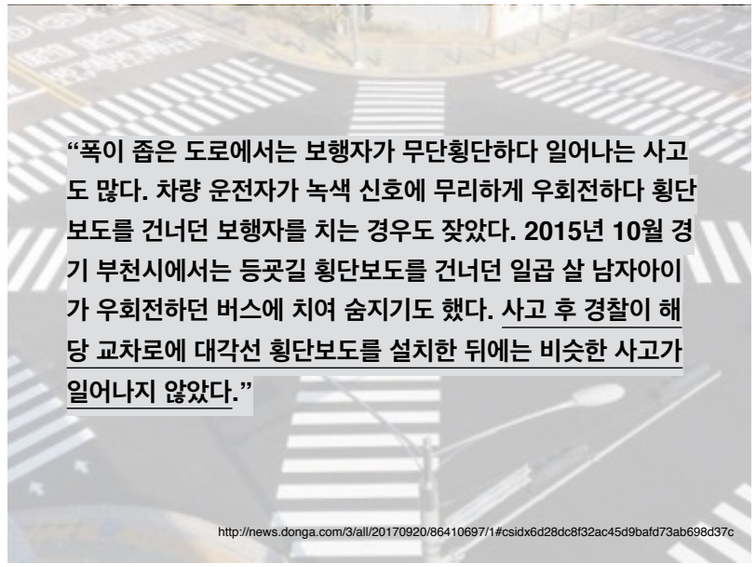
http://news.joins.com/article/17423799

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

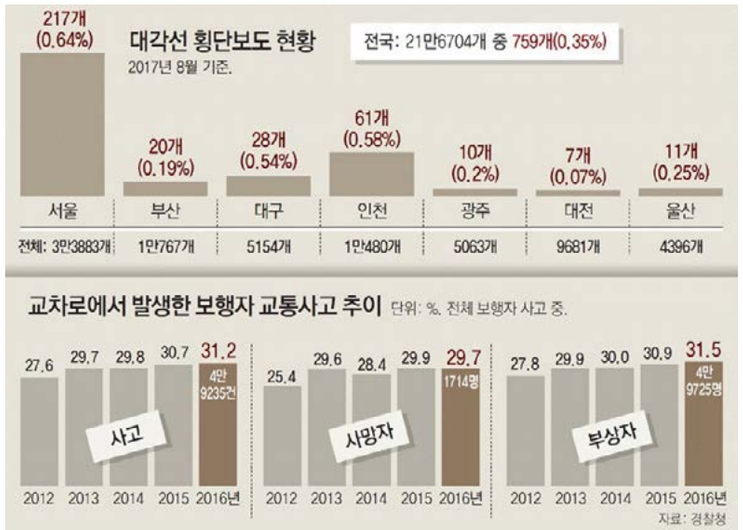
횡단보도 숫자와 위치를 정할 때, 우리는 이미 엄청난 가치판단을 한다.

"보행자와 자동차 중 무엇이 더 중요한가?"

KBS 뉴스 http://mn.kbs.co.kr/data/news/2017/04/09/3460488_140.jpg

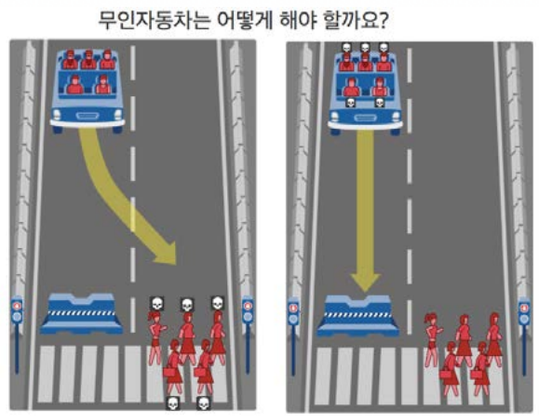


보행자 안전을 위해 무엇을 할 수 있는가?

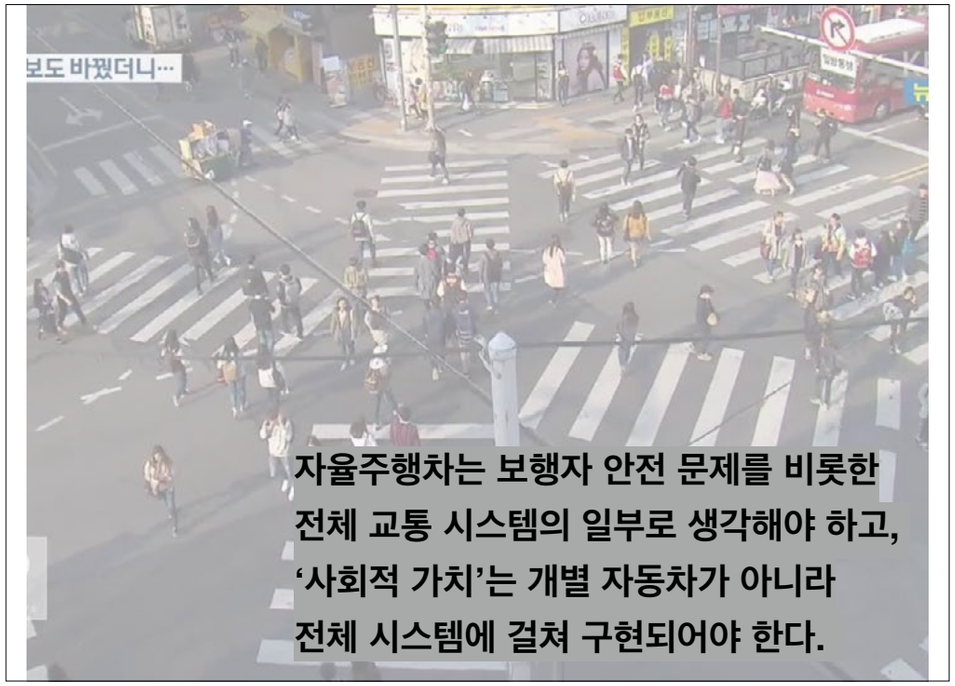


http://news.donga.com/3/all/20170920/86410697/1

순간적으로 누구를 살리고 죽일지를 묻는 '트롤리 문제'로는 자율주행 시스템에서 보행자 안전을 확보할 수 없다.



‘안전’이라는 사회적 가치는
교통 시스템과 정책과 관행에 폭넓게 걸쳐 있다.
자율주행차 알고리즘에 손쉽게 집어넣어 구현할 수 있는 것이 아니다.

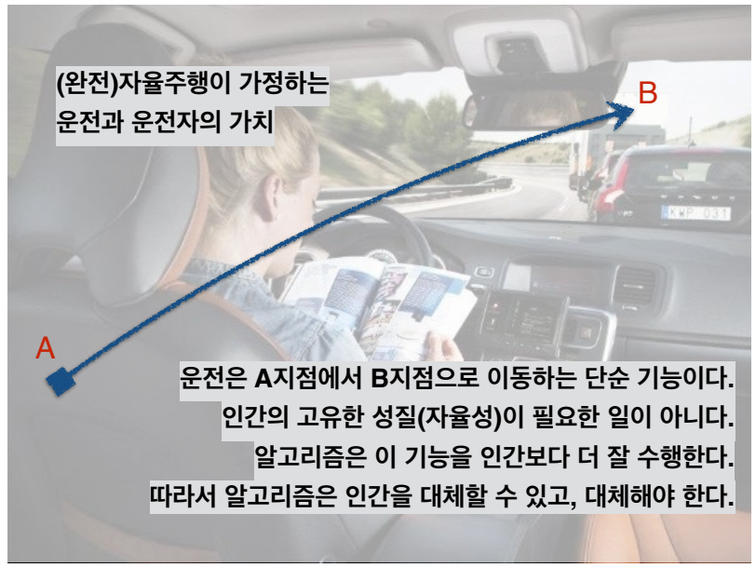


자율주행차는 보행자 안전 문제를 비롯한
전체 교통 시스템의 일부로 생각해야 하고,
‘사회적 가치’는 개별 자동차가 아니라
전체 시스템에 걸쳐 구현되어야 한다.

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

2. ‘자율주행차’는 운전과 운전자의 가치를 어떻게 바꿀 것인가?

Wikimedia Commons: DimiTVP



운전은 A지점에서 B지점으로 이동하는 단순 기능이다.
인간의 고유한 성질(자율성)이 필요한 일이 아니다.
알고리즘은 이 기능을 인간보다 더 잘 수행한다.
따라서 알고리즘은 인간을 대체할 수 있고, 대체해야 한다.

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

(완전)자율주행이 가정하는
운전과 운전자의 가치

특히 버스, 택시, 트럭 운전사는
정해진 대로 가는 기능을 수행할 뿐
고유한 자율성이 없으며,
불완전하고, 불친절하고, 위험하다는 생각.

승객을 돌보고 공동체를 유지하는 역할을 하는 버스운전사

EBS 한국기행 (2017. 10. 3.)

운전은 때로 도덕적 결단, 정치적 의사표시가 된다.
“가자는 대로 간다”가 아니라
“가지 말래도 간다”

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

“사람들이 우왕좌왕하고 있을 때 시내버스를 앞세우고 택시들이 행진했죠. 라이트를 켜고 클랙슨을 울렸습니다. 그러자 사람들이 다시 모이기 시작했어요. 그대로 도청까지 갔습니다.”

<한국일보> 2017. 7. 20.
광주 택시운전사 인터뷰

사진: 연합뉴스

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3

운전은 여성 권리의 척도가 되기도 한다

US politics world opinion sports soccer tech arts lifestyle fashion business travel environment

home › world › middle east africa australia cities development UK europe americas asia

Saudi Arabia
Solutions and innovations

'This is a huge step for us': jubilation as Saudi women allowed to drive

Women hail royal decree overturning ban on them holding driving licences as a turning point in conservative kingdom

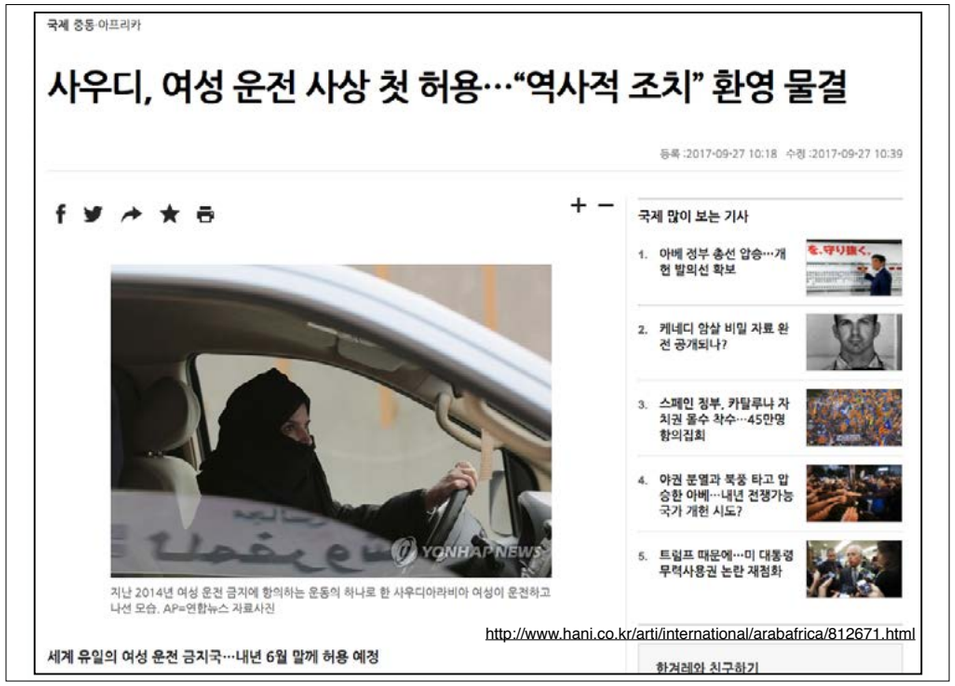
3,889

Martin Chulov and Nadia al-Faour

Wednesday 27 September 2017 05.50 EDT

Under the new law, women can legally obtain a driving licence without asking a male guardian for permission. Photograph: Faisal Al Nasser/Reuters

https://www.theguardian.com/world/2017/sep/27/huge-step-jubilation-saudi-arabian-women-allowed-to-drive


국제 중동·아프리카

사우디, 여성 운전 사상 첫 허용…"역사적 조치" 환영 물결

등록 :2017-09-27 10:18 수정 :2017-09-27 10:39

국제 많이 보는 기사

1. 아베 정부 총선 압승…개헌 발의선 확보
2. 케네디 암살 비밀 자료 완전 공개되나?
3. 스페인 정부, 카탈루냐 자치권 몰수 착수…45만명 항의집회
4. 야권 분열과 북풍 타고 압승한 아베…내년 전쟁가능 국가 개헌 시도?
5. 트럼프 때문에…미 대통령 무력사용권 논란 재점화

지난 2014년 여성 운전 금지에 항의하는 운동의 하나로 한 사우디아라비아 여성이 운전하고 나선 모습. AP=연합뉴스 자료사진

세계 유일의 여성 운전 금지국…내년 6월 말께 허용 예정

한겨레와 친구하기

http://www.hani.co.kr/arti/international/arabafrica/812671.html

운전은 양도할 수 없는
인간의 의지와 자율성이
표출되는 통로가 되기도 한다.

https://twitter.com/manal_alsharif/status/91277126891865292 8/

왜 자율주행차를 만드는가?

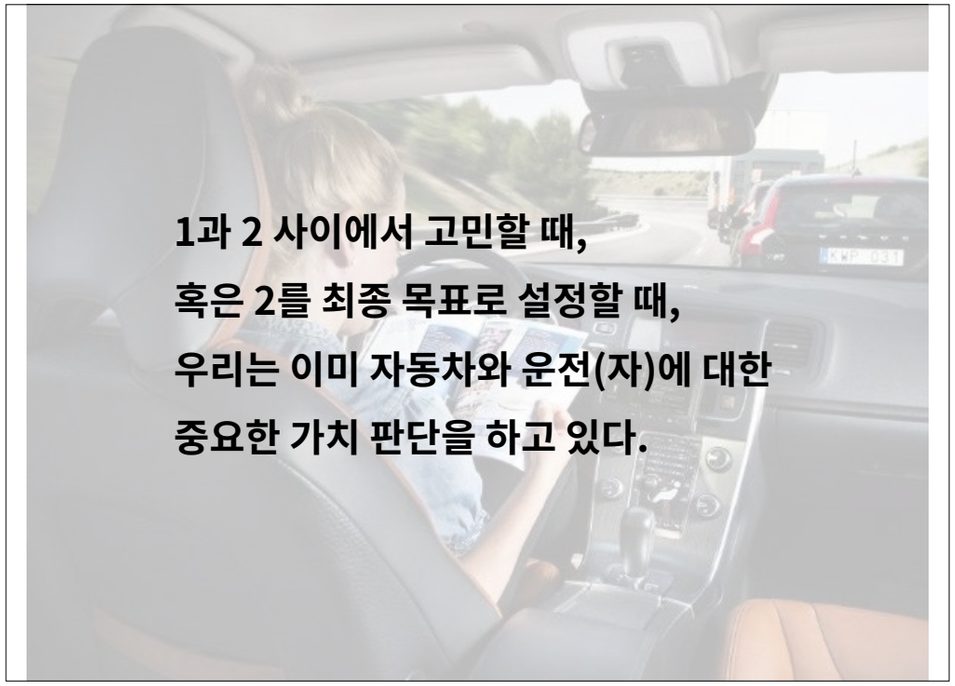
1. 더 운전하기 쉬운 차를 만들자

피로를 줄이고, 집중력을 높이고, 노인과 장애인도 편하게
조작할 수 있고, 보행자에게도 안전하도록

2. 운전자를 없애자

인간은 위험하니 운전을 못하게 하고, 차가 스스로 운전하도록

The 8th ASIA Future Forum
제8회 아시아미래포럼
세션 3



자율주행차와 사회적 가치

- 현 교통 시스템은 이미 수많은 가치 판단의 결과이다. 운전자의 자격과 자율성에 대한 판단, 보행자와 자동차의 상대적 중요성에 대한 판단 등 수많은 가치에 대한 판단이 교통 시스템 속에 얽혀 있다.
- 자율주행차는 홀로 존재하는 기술이 아니라 이러한 사회적 가치들의 네트워크 속으로 편입되어 존재하게 될 기술이다.
- 알고리즘에 몇몇 가치를 '주입'해서 완벽한 자율주행 시스템을 구현할 수 없으며, 복잡하게 얽힌 정책-디자인-관행-가치의 네트워크 속에서 자율주행의 형태와 사용방식을 고민해야 한다.

더 큰 물음들

- 현재의 자동차 교통 시스템의 난제들 중 자율주행은 어떤 부분을 해결해 줄 수 있는가?
- 자율주행으로도 해결하지 못할 교통과 안전의 문제는 무엇인가?

- 자동차를 포함하는 전체 교통안전 시스템의 미래는 어떻게 설계해야 하는가?
- 그 미래에서 인간의 (자율)운전과 기계의 (자율)운전은 어떤 역할을 나누어 맡아야 하는가?

토론문

황용석
건국대학교 미디어커뮤니케이션학과 교수

김진형 원장님과 전치형 교수님의 발표문은 소프트웨어 구성물로서 인공지능 및 알고리즘이 가져올 긍정적인 측면과 부정적인 측면을 모두 다루면서, 이들 비인간 행위자가 의사결정 과정에 개입하면서 생기는 책임의 문제를 다루고 있습니다. 과학기술의 책임, 윤리, 그리고 사회적 가치의 식재를 논하는 것은 새로운 주제는 아닙니다. 그러나 그 대상기술이 인공지능과 같이 '자기 학습적 존재'라는 점은 새로운 사안을 부르고 있습니다.

비인간행위자인 소프트웨어, 새로운 존재론적 고민을 던져

김진형 원장님의 발표에 나온 것처럼, 기술이 인간을 초월하는 순간 즉, 특이점(singularity)이 다가오고 있습니다. 자기 학습능력을 가진 기술의 등장은 인간에게 불안감을 가져다주고 있습니다. 인식과 의사결정의 측면에서, 우리들의 이런 불안감은 인간노동이 인공지능으로 대체되는 현상 등 관찰 가능한 현상 때문에 과장된 측면도 있습니다. 산업혁명 이후 기술의 발전은 분명 인간의 노동을 가시적으로 외주화해왔습니다. 오늘날 인공지능의 등장은 외주화하는 노동의 질적 속성이 변했다는 점과 그 속도와 범위가 빠르고 크다는 점을 제외하면 증기기관과 크게 다르지 않습니다. 즉, 시장과 사회의 변동은 기술을 채택하면서 인간이 경험해온 예측가능한 결과들을 포함합니다.

오히려 인공지능시대 우리가 보다 걱정해야 하는 것은 '인간의 생각을 외주화'하는 문제입니다. 우리가 인공지능에 대해 갖는 불안감은 니콜라스 카(N. Carr)가 쓴 〈생각하지 않는 사람들〉이란 책에서처럼, 인간이 사고나 의사결정을 기계에 의존하면서 생긴 과의존의 결과일 수 있습니다. 기술에 의한 결정을 수동적으로 수용하는 것이야말로 인공지능 시대에 인간이 당면한 큰 문제입니다.

보다 본질적인 불안감은 우리가 기계 또는 소프트웨어를 생각하는 주체로 성급하게 간주하면서 나타납니다. 마투라나와 바렐라(Maturana & Varella, 2007)는 고양이와 인간의 상호작용을 예로 듭니다. 고양이의 행동에 반응하는 인간들은 고양이의 행동을 자신의 언어적 모델로 이해하면서 스스로 상호작용한다고 믿습니다. 그러나 언어체계가 없는 고양이의 동기를 이해하는 것은 인간이 만들어 낸 의미화와 연상의 결과일 수 있습니다.

디지털 재현기술인 인공지능 챗봇과 대화하는 인간을 상상해 볼 수 있습니다. 챗봇이 데이터 학습에 기반해서 인간이 '대화의 인공성'을 지각하지 못하는 수준으로 대화를 하더라도 이 역시 기술이라는 외적 대상과 인간정신세계를 인간이 연결한 결과 즉, 의미화한 것이라 할 수 있습니다. 자연스러운 대화가 기계를 인간과 같은 존재임을 입증하는 것은 아닙니다.

인공지능 윤리의 의미, 인간가치의 투영

인공지능에 의한 의사결정이 많아질수록 이들 기술에 의한 편향, 차별, 왜곡 등의 부작용이 확인되고 있습니다. 특히, 권리가 상충하는 상황에서 학습된 기계가 어떤 판단을 하는가는 인공지능 또는 알고리즘 윤리담론을 형성시켰습니다. 인공지능이나 알고리즘과 같은 소프트웨어에 윤리라는 용어를 사용한다는 것은, 이들을 스스로 옳고 그름을 가리는 도덕적 판단행위자로 인정하는 것입니다. 그러나 이 용어는 다른 관점에서 소프트웨어가 내린 판단의 결과를 인간화해 해석한 의미화의 결과라고도 말할 수 있습니다.

이런 관점에서 보면, 인공지능 윤리는 인간이 가이드할 수 있는 대상이 됩니다. 알고리즘은 본시 일련의 해를 구하는 논리적 연산으로 결과값 산출을 목표로 하며, 옳고 그름의 문제는 데이터와 학습모형으로 결정됩니다. 그 방향을 조절할 수 있는 것은 인간의 의지입니다.

기계학습은 그 모형의 특성으로 인해 데이터와 연결된 정보망에 따라 마치 자율적으로 살아 움직이는 '경량적 존재론'(소프트웨어가 스스로 데이터나 정보의 연결구조에 의해 반응하는 것)에 가깝습니다. 그러나 이 기술은 고유한 의미의 자율적 주체가 아니며 인간이 부여한 일정 수준의 지능을 토대로 특정 범위에서 인간을 대신하는 대리행위자라고 할 수 있습니다. 이는 소프트웨어에 부여하는 처리방식에 대한 사회적 합의가 필요하다는 논리로 연결됩니다.

즉, 우리는 인공지능 또는 알고리즘 윤리와 책임성이 곧 기술에 대한 인간의 설계 또는 가이드를 내포하며 그 열쇠는 인간에게 있다고 하겠습니다. 그러나 이것이 쉬운 문제만은 아닙니다. 인공지능 알고리즘에 내재하는 불확실성을 완벽하게 통제하는 것이 쉽지 않기 때문입니다. 뿐만 아니라 알고리즘의 투명성을 확보하는 것 역시 어렵습니다.

전치형 교수님이 발표하신 트롤리 딜레마(Trolley Dilemma)를 포함해서 '차이니즈 룸 논쟁'(Chinese Room Argument), '도덕적 튜링 테스트'(Moral Turing Test) 등은 인공지능에 대한 윤리적 설계의 어려움을 설명하는 대표적인 실험의 예들입니다.

최근의 접근들은 인간친화적 기술설계로 모아지고 있습니다. 보네폰 등(Bonnefon, et.al, 2015)이 최근에 '트롤리 딜레마'를 자율주행자에 응용해서 복수의 국가에서 조사를 실시했습니다. 그 결과는 국가나 문화권에 관계없이 기술이 인간을 보다 보호하는 방향으로 작동되어야 한다고 답을 했습니다.

윌러치 • 알렌(Allemang, & Hendler, 2014)은 '인공적 도덕 행위자'(AMA: Artificial Moral Agent) 개념을 통해 의도하지 않아도 도덕적 의사결정을 내려야하는 기계나 소프트웨어가 도덕적 판단을 할 수 있도록 알고리즘 설계 단계에서 도덕과 윤리기준을 프로그래밍하는 방법론을 강조했습니다. 그밖에도 미국의 인공지능 전문가 엘리저 유드코프스키(Eliezer Yudkowsky)가 제시한 '우호이론'(friendliness theory, 로봇이 인간에게 우호적으로 작동해야 함을 강조), 2017년 FLI(Future of Life Institute)가 인간에게 유용하고 혜택을 주는 착한 인공지능(beneficient AI)을 개발해야 한다는 이른바 '아실로마 AI 원칙'(ASILOMAR AI PRINCIPLES)을 발표하는 등 다양한 원칙과 개념이 제시되고 있습니다.

유럽연합(EU)에서는 인공지능 기술에 대한 규범적 차원에서의 법제화가 진행되고 있습니다. EU 집행위원회가 2012년부터 2년 동안 진행한 '로봇법 프로젝트'(RoboLaw Project)를 수행한 후 2014년에 '로봇규제에 관한 가이드라인'(Guidelines on Regulating Robotics)을 제정했으며, 2016년 5월에는 유럽의회 법사위원회(Committee on Legal Affairs)가 '로봇법 규칙 초안을 위한 보고서'(Draft Report)을 위원회에 제출했습니다.

이 보고서는 전자적 인격(electronic person) 부여, 로봇 등록제 도입, 로봇기술 헌장 마련, 로봇기술 규제기구 창설, 지적재산권 및 개인정보의 보호, 로봇기술 표준화, 법적 책임 등 로봇법 관련 법제화 과정에서 고려해야할 원칙과 주요내용도 구체적으로 제시하고 있습니다. 특히 '강한 인공지능'을 고려해서 현행 인간중심의 법체계에 비인간행위자를 포함하는 방향을 검토해야 한다고 보고 있습니다. 로봇에게 '전자인간'(electronic personhood)이라는 특수한 형태의 법적 지위를 부여하는 결의안을 채택한 것은, 인공지능이 구현되는 기술체에 특수한 형태의 법적 지위를 부여해 사회적 문제를 예방하고 문제 발생시 책임을 부여하는 방식을 고려했다고 할 수 있습니다.

인공지능에 대한 규범적이고 법적인 접근은 아직까지 신중할 필요가 있습니다. 그러나 새로운 사회에서는 인간의 판단을 위임한 비인간행위자의 개입이 커지면서, 인간이 다시 자신들의 가치를 기술에 식재할 필요성이 높아졌습니다. 인공지능기술의 특성상 기술작동의 '비통제성'과 '불확실성'이 커지고 그로 인한 편향이 발생되는 등 사회적 문제가 가시화하고 있기에 우리 사회에서도 인공지능과 인간의 새로운 규범적 관계 설정이 본격적으로 논의될 필요가 있습니다. 그러나 한편으로, 과도한 인간 개입이 기술의 발전을 저해할 수 있다는 문제도 인식할 필요가 있습니다.

토론문

오세욱
한국언론진흥재단 선임연구위원

두 분 발표 너무 잘 들었습니다. 새로운 기술은 새로운 시장을 창출하고 기존의 것들을 더 효율적으로 만들며 인간을 좀 더 안전하게 살아갈 수 있도록 하는 동시에 인간을 더 부유하게 하는 측면이 있습니다. 반면에 김진형 원장님께서 지적하여 주셨듯이 인간 기계 공생의 딜레마 등과 같은 의도하지 않은 부작용도 함께 가져올 수 있습니다. 이러한 문제를 방지하기 위해 설계 단계부터 일종의 규제가 필요하다는 주장도 제기되고 있습니다. 하지만, "알고리즘에 몇몇 가치를 주입해서 완벽한 자율주행 시스템을 구현할 수 없으며, 복잡하게 얽힌 정책–디자인–관행–가치의 네트워크 속에서 고민해야 한다"는 전치형 교수님의 발표처럼 이는 그리 좋은 방안은 아닐 것입니다. "대중의 무지나 공포에 기반을 둔 규제 정책을 실행해서는 안 되며, 공개 소프트웨어 등 투명성을 강화하는 방안이 필요하다"는 김진형 원장님의 발표처럼 규제보다는 사회적 검증을 유도하자는 방안에 저는 적극 공감합니다.

의료, 법률, 금융, 교통, 쇼핑, 뉴스 등 우리 생활 거의 대부분의 영역에 걸쳐서 알고리즘에 의한 의사결정이 이루어지고 있습니다. 이에 따라 전치형 교수님이 언급하신 '트롤리 문제'는 자율 주행차와 관련해서만이 아니라 다양한 영역에서 벌어지고 있습니다. 보룩바시 등(Bolukbasi, Chang, Zou, Saligrama & Kalai, 2016)의 연구 결과가 이를 잘 보여줍니다. 이들은 구글 뉴스에서 수집한 기사 수십만 개를 기계가 학습하도록 했습니다. 이후 "어떤 부부가 있는데 남성의 직업이 프로그래머라면, 여성의 직업은 무엇인가"란 질문을 던지자 기계는 "주부"라고 답했습니다. 여성의 직업 가능성이 가장 높은 것을 답한 것인데, 문제는 가능성이 정답을 의미하지는 않는다는 점입니다. 기계는 인종, 지역 등을 주제로 비슷한 질문을 던졌을 때도 기존 인간의 편향성(bias)이 더욱 공고화된 답변을 내놓았습니다. 인간이라면 그 질문에 대한 답을 할 때 여러 가지 복합적 요인을 고려했을 것입니다. 기계는 가치보다는, 주어진 데이터를 학습하고 그것에 판단해 최적의 답을 찾고 있습니다.

문제는 또 있습니다. 김진형 원장님이 발표에서 언급하신 것과 같이 기계는 '블랙박스'입니다. "우리가 구글의 능력을 숭배하는 이유 중 하나는 주술과 마찬가지로 최종 결과만이 나타날 뿐, 내부의 작동 과정을 전혀 알 수 없기 때문이다. 알고리즘의 작동은 블랙박스에 숨겨질 뿐 아니라, 사실상 찰나에 이루어진다."(Dormehl, 2014/2014, 272쪽)는 말처럼 알고리즘, 인공지능 등의 결정은 인간이 이해할 수 있는 수준을 벗어나고 있습니다. 이러한 시스템을 만드는 엔지니어들도 구체적으로 그 내용을 이해하지 못합니다. 사회적으로 알고리즘을 공개해야 한다는 주장이 제기되고 있지만, 알고리즘을 공개해도 사실 그 내용을 이해하기 어려운 경우가 많습니다. 알고리즘은 실시간으로 받는 신호에 따라 결정을 내리고 있기 때문입니다. 우리는 기계의 결정 과정을 점점 더 이해하기 어려운 상황에 놓일 것입니다.

알고리즘과 사회적 가치와 관련해 앞으로 문제가 되는 지점은 인간과 기계 사이의 통역불가능성(incommensurability)의 강화라고 봅니다. 이세돌과 알파고의 바둑 대결이 이를 상징합니다. 그 대결의 결과는 첫째, 인간 이세돌이 복기를 하려해도 그 상대가 없다는 점에서 소통불가능성(ex–communication)을, 둘째, 인간의 바둑 이론으로는 알파고의 수를 설명해낼 수 없다는 점에서 공재불가능성(incompossibility)을, 셋째, 비인간 행위자와의 대화 상황처럼 언어를 공유하지 못한다는 점에서 외계–소통(xeno–communication)을 보여주었습니다(이재현, 2016).

이러한 상황에서 우리는 어떻게 해야 할까요? 알고리즘 혹은 인공지능에 의한 의사결정시스템을 더 잘 다루려면 우리는 무엇을 해야 할까요? 윈필드와 지롯카(Winfield & Jirotka, 2017)는 이러한 문제들을 해결하기 위한 대안으로서 "윤리적 블랙박스"(ethical black box)를 제시합니다. 로봇과 자동 시스템이 사고를 낼 경우를 대비하여, 비행기의 블랙박스처럼 사고 원인과 과정을 알 수 있는 "윤리적 블랙박스" 구비를 의무화하자는 제안입니다. 인공지능이 내리는 결정에 책무성과 책임성을 부여하기 위한 수단으로, 이들은 "윤리적 블랙박스"와 같은 것을 통해 투명성이 구현되지 않으면, 로봇과 자동 시스템이 신뢰를 얻기 어려울 것이라고 주장합니다.

그런데 저는 여기에 더해, 소중히 지켜온 민주주의 사회의 가치를 앞으로도 지켜나가기 위해서 이들 시스템에 대한 사회적 감시가 필요하다고 생각합니다. 최근 스탠포드대학교의 연구진은 기계학습을 통해 사진만으로 그 사람의 성 정체성을 판단하는 인공지능을 개발했다고 논문을 통해 밝힌 바 있습니다. 성소수자가 아직까지 억압받는 곳이 상당함을 감안하면, 이 AI의 위험성은 굳이 언급하지 않아도 분명합니다. '남용(abuse)을 통한 낙인', 낙인의 남용 등 악용의 우려를 너무 쉽게 떠올릴 수 있었습니다. 이에 대해 〈와이어드〉지에서는 굉장히 강하게 비판하면서 "AI 연구에 윤리적 감시견이 절실히 필요하다"고 주장한 바 있습니다.

이러한 관점에서 MIT 미디어 랩의 '시민 미디어'(civic media) 그룹이 만든 '알고리즘 저스티스 리그'(The Algorithmic Justice League) 라는 일종의 시민단체 사례를 주목할 필요가 있다고 봅니다. 알고리즘 저스티스 리그는 알고리즘의 차별과 편향에 대항하는 것을 목적으로 하고 있습니다. 누구를 고용할 것이며, 대출을 승인할 것인지, 혹은 얼마나 감옥에 있게 할 것인지 등의 결정을 돕는 보이지 않는 권력이 떠오르고 있다며, 이들은 이 권력을 "코드화된 응시"(codede gaze)라고 부릅니다. 알고리즘 저스티스 리그는 미디어, 예술, 과학을 통해 이러한 편향을 알리고, 코드화된 차별을 경험할 공간을 제공하고 개발자들에게 책무성을 상기시키는 것을 목표로 하고 있습니다. 이를 위해 그러한 편향이 빚어지게 된 이유를 기계학습에 사용된 데이터를 통해 검증하고 있습니다. 이들은 일반 시민들이 알고리즘 혹은 인공지능의 의사 결정이 자신들의 일상생활에 미치는 영향을 쉽게 알 수 있도록 알리는 한편으로, 그 결정이 어떤 문제를 지니고 있는지를 감시하는 것으로 목표로 하고 있습니다.

또한, 저는 사회적, 혹은 정책적 대안으로 알고리즘의 의사 결정 영역을 명기하도록 하는 방안을 추진할 필요가 있다고 생각합니다. 이러한 맥락에서 알고리즘 팁스(algorothm tips) 라는 단체의 사례를 주목할 만합니다. 이들은 미국 정부에서 사용하는 정책 결정 알고리즘을 목록화해 제시하고 있습니다. 점점 더 많은 정책 결정이 알고리즘에 따라 이루어지면서, 알고리즘 투명성이 중요하다는 이유에서입니다. 정부 문서를 분석해 사례를 찾아내고 일반 시민들이 사례를 제보하는 방식으로 목록화하고 있습니다.

기술의 발전을 막을 수는 없으며, 막을 필요도 없습니다. 그것이 해결하고 개선할 현 사회의 문제도 많기 때문입니다. 다만, 기술적, 산업적 관점에서 효율성 등이 지나치게 강조되다 보면, 그것이 가져오는 새로운 사회문화적 문제에 제대로 대처하지 못할 수도 있습니다. 특이점이 언제 올지는 모르겠지만, 어쨌든 기술의 발전 속도는 매우 빠른 것이 사실입니다. 그 빠름을 감안할 때 규제를 통해 예측적으로 제어하는 것은 불가능할 것입니다. 감시를 통해 문제를 재빨리 발견하고 수정하려는 노력이 곁들여져야 민주주의 사회의 가치를 지향하면서 인류에게 도움이 되는 기술이 인류와 공존할 수 있지 않을까라고 생각해 봅니다.

토론문

이성웅
한국 IBM상무

우리 사회의 화두로 자리 잡은 인공지능은 생활 곳곳에서 우리 삶을 변화시킬 것입니다. 인공지능이 제어하는 스마트 팩토리에서 생산된 자율주행차가 거리를 달리고, 로봇과 드론이 인간과 함께 일하며, 스마트 헬스케어가 우리의 건강을 관리하는 세상은 이미 현실화 되고 있습니다.

오늘 김진형 원장님이 발표하신"소프트웨어로 사회문제를 해결할 수 있는가" 그리고 전치형 교수님이 발표하신 "자율주행차로 무엇을 할 것인가"는 눈앞에 다가와 있는 인공지능의 윤리 문제를 어떻게 다룰 것인가 라는 주제를 공통적으로 담고 있습니다.

저는 인공지능 관련 비즈니스를 실제로 수행하고 있는 IBM이라는 회사가 인공지능 윤리 문제를 어떻게 다루고 있는지를 설명 드리는 것으로 토론을 대신 하고자 합니다. 인공지능 윤리문제를 회사들이 실제로 어떻게 다루고 있는지 살펴보고 향후 나아가야 할 방향을 모색하는 것이 의미가 있다고 생각하기 때문입니다.

1. 인공지능과 윤리에 관한 정책

IBM은 2017년 1월에 인공지능 윤리 문제와 관련하여 다음과 같은 세 가지 원칙들(Transparency and Trust in the Cognitive Era)을 수립하여 실행하겠다고 선언하였습니다.

1) 목적(purpose): IBM이 개발하고 응용하는 인공지능 시스템의 목적은 인간의 지성을 향상시키는 데에 있습니다. IBM의 기술, 제품, 서비스 및 정책은 인간의 역량, 전문성 및 잠재력을 향상시키고 확대시킬 수 있도록 설계될 것입니다. IBM의 이러한 태도는 원칙 그 자체에도 기반을 두고 있지만 과학에도 기반을 두고 있습니다. 인공지능은 현실적으로 의식을 얻거나 독립된 기관으로서 역할을 할 수 없을 것입니다. 인공지능은 기업과 사회가 기능하는 프로세스, 시스템, 제품 및 서비스에 통합되어(embedded) 갈 것입니다. 이러한 모든 것은 인간의 통제 범위 내에 머무를 것이며, 또 그렇게 되어야만 합니다.

2) 투명성(transparency): 인공지능 시스템이 세상을 변화시킬 만한 잠재력을 발휘하기 위해서는 인공지능 시스템의 추천, 판단, 그리고 그 사용에 있어 사람들 사이에 신뢰가 형성되어야 합니다. 따라서 IBM은 다음과 같은 사항들을 명확히 하고자 합니다.

- 언제, 그리고 어떠한 목적으로 IBM이 개발하고 판매하는 솔루션(cognitive solutions)에 인공지능을 응용할 것인가?
- 인공지능 솔루션에 대한 통찰력을 제공하는 데이터 및 전문지식의 주된 원천, 그리고 그러한 시스템 및 솔루션을 향상시키는 데 사용되는 방법

■ 고객이 자신의 비즈니스 모델과 지적재산권을 소유하며, 수 년 간의 경험을 통하여 만들어낸 역량을 강화하는데 인공지능 시스템을 사용할 수 있다는 원칙. IBM은 고객과 협력하여 고객의 데이터와 노하우를 보호하고, 다른 고객, 파트너, 그리고 업계 관계자들이 그와 유사한 방식을 채택하도록 권장함

3) 기술(skills): 새로운 인공지능 시대의 경제적, 사회적 이익은 인간과 관련된 문제가 해소되어야만 진정한 의미를 가질 수 있을 것입니다. 이것은 인공지능 기술에 있어서 매우 중요하며, 인공지능 기술은 인간의 지성과 전문기술을 발전시키고 인간과 함께 협업해야 합니다. 따라서 IBM은 학생, 근로자, 일반 시민 등이 인공지능과의 관계를 안전하고, 무해하며, 효과적인 방향으로 발전시킬 수 있는 지식과 기술을 습득할 수 있도록 도울 것입니다. 또한 인공지능 경제에서 나타날 새로운 종류의 업무와 직업을 수행할 수 있는 지식과 기술도 습득할 수 있도록 도울 것입니다.

2. 인공지능과 데이터에 관한 정책

IBM은 또한 2017년 10월에 Data Responsibility@IBM 이라는 정책을 발표했습니다. AI가 도입되기 전부터 IBM은 데이터를 수집, 저장, 관리, 처리하는 조직이라면 책임 있는 자세로 데이터를 다룰 의무가 있다고 생각했습니다. 이러한 신념은 IBM이 한 세기 동안 모든 관계에서 지향해 온 신뢰 및 책임 의식으로 구체화 되었으며, 세계 유수의 기업들이 가장 중요한 데이터를 IBM에 맡겨온 이유이기도 합니다. 데이터 정책의 수립과 실행을 통해 IBM은 항상 책임을 다하는 자세와 철학으로 고객의 신뢰를 더욱 공고히 하고자 한다는 입장을 천명한 것입니다.

Data Responsibility@IBM은 "데이터 소유권과 개인정보보호", "데이터 이동과 접근", "데이터 보안과 신뢰", "데이터와 인공지능", 그리고 마지막으로 "데이터 기술과 새로운 인력 양성" 관한 IBM의 정책과 의지를 담고 있습니다.

Data Responsibilty@IBM 정책 내용 중에서 "데이터와 인공지능"에 관한 내용을 소개하면 다음과 같습니다.

IBM은 오래 전부터 AI 기술 분야를 개척하고 전 세계의 고객들이 이 기술을 활용할 수 있도록 지원하면서 '인공(artificial)'보다는 '증강(augmentation)'이라는 표현이 더 어울리는 이 기능이 기업, 기관, 정부, 개인에게 긍정적인 변화의 원동력이 될 것임을 확신했습니다. 그와 더불어 AI 기술 개발이 신중하고 책임 있는 방식으로 이루어져야 한다는 점을 깨달았습니다.

■ AI의 가치는 사람의 능력을 대체하는 것이 아니라 증강하는 데 있습니다. AI 시스템은 의식이나 지각을 가진 존재가 되지 않을 것입니다. 그보다는 전 세계의 프로세스, 시스템, 상호 작용의 일부로 통합될 것입니다. 인공지능이 사람의 의사 결정, 판단, 직관 또는 윤리적 선택을 대체하지 않으며 대체할 수도 없습니다.

■ IBM은 AI 시스템이 어떤 결론이나 조언에 도달하는 방식에 대해 사람들이 온전하게 이해할 수 있도록 투명성 및 데이터 거버넌스 정책을 지원합니다. 기업은 어떤 과정을 거쳐 알고리즘의 권장 사항이 나왔는지 설명할 수 있어야 합니다. 그러지 못하면 이 기업에서 만든 시스템은 상용화될 수 없습니다.

■ AI 시스템의 사회적 합의에 대한 토론이 이루어지는 가운데 IBM은 자동화를 규제하거나 혁신에 불이익을 주어서는 안 된다고 생각합니다. IBM은 정책 입안자 그리고 고객과 함께 인간이 AI 시스템과의 파트너십을 십분 활용하면서 효과적으로 일하는 데 필요한 기술을 습득할 수 있도록 지원할 것입니다.

경청해 주셔서 감사합니다.