

제142회 한림원탁토론회

인공지능과 함께할 미래 사회, 유토피아인가 디스토피아인가

일시 : 2019년 9월 26일(목) 15:00

장소 : 코리아나호텔 로얄룸(7층)



초대의 말씀

오늘날 많은 사람들이 인공지능과 함께할 우리의 미래에 대해 이야기합니다. “인공지능이 일자리를 뺏어가고 사회적 문제를 야기하며 100년 안에 인류를 지배하여, 결국 멸망시킬 것”이라는 부정적 전망, “인공지능이 새로운 산업적 가치를 창출하고 사회 및 경제를 한단계 업그레이드하고 삶의 질을 높일 것”이라는 장밋빛 기대가 극명하게 대립하고 있습니다. 하지만 분명한 것은 과학기술의 발전은 멈추지 않을 것이며, 인공지능은 우리의 삶과 사회를 더욱 빠르게 변화시킬 것이라는 점입니다.

이에 한국과학기술한림원은 한림원 석학을 비롯해 다양한 분야의 전문가들을 모시고 소통할 수 있는 자리를 마련하고자 합니다. 인공지능의 가능성과 한계를 냉정히 짚어보고 인공지능이 야기할 사회적 변화 및 문제와 대응책에 대해 의견을 모아 보고자 합니다. 바쁘시더라도 부디 참석하시어 더욱 밝은 미래를 만들어가는 사회적 모멘텀을 키우는데 고견을 더해주시길 바랍니다.

감사합니다.

2019년 9월

한국과학기술한림원 원 장 **한 민 구**

한림원탁토론회는 국가 과학기술의 장기적인 비전과 발전전략을 마련하고 국가사회 현안문제에 대한 과학기술적 접근 해결방안을 도출하기 위해 개최되고 있습니다.

PROGRAM

사회 : 성창모 한국과학기술한림원 정책연구소장

시 간	내 용	
14:30 ~ 15:00	등 록	
15:00 ~ 15:10	개 회 사	한민구 한국과학기술한림원 원장
15:10 ~ 15:40	주제발표 1	‘인공지능의 본질, 그 능력과 한계’ 김진형 KAIST 전산학부 명예교수(한림원 공학부 정회원)
15:40 ~ 16:10	주제발표 2	‘인공지능 알고리즘은 인간을 차별하는가?’ 홍성욱 서울대학교 생명과학부 교수(한림원 정책학부 정회원)
16:10 ~ 16:40	주제발표 3	‘AI 자본주의’ 노영우 매일경제신문 국제부장
16:40 ~ 16:50	Coffee Break	
16:50 ~ 17:30	지정토론 좌 장 토 론 자 (가나다순)	이태억 KAIST 산업 및 시스템 공학과 교수(한림원 정책학부장) 김경만 서강대학교 사회학과 교수(한림원 정책학부 정회원) 임영익 인텔리콘 법률사무소 대표 김건우 카카오모빌리티 데이터 이코노미스트
17:30 ~ 18:00	종합토론	
18:00 ~	폐 회	

CONTENTS

I. 주제발표 1 ‘인공지능의 본질, 그 능력과 한계’	1
• 김진형 KAIST 전산학부 명예교수(한림원 공학부 정회원)	
II. 주제발표 2 ‘인공지능 알고리즘은 인간을 차별하는가?’	31
• 홍성욱 서울대학교 생명과학부 교수(한림원 정책학부 정회원)	
III. 주제발표 3 ‘AI 자본주의’	49
• 노영우 매일경제신문 국제부장	
IV. 지정토론	71
좌장 이태억 KAIST 산업 및 시스템 공학과 교수(한림원 정책학부장)	73
• 김경만 서강대학교 사회학과 교수(한림원 정책학부 정회원)	75
• 임영익 인텔리콘 법률사무소 대표	79
• 김건우 카카오모빌리티 데이터 이코노미스트	87

주제발표 1

I

인공지능의 본질, 그 능력과 한계

김진형

KAIST 전산학부 명예교수(한림원 공학부 정회원)

발제자 약력

성 명	김 진 형	
소 속	KAIST	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1967 ~ 1971	서울대학교	건축, 공학사
1977 ~ 1979	UCLA	시스템공학, 석사
1979 ~ 1983	UCLA	컴퓨터과학, 박사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2016 ~ 2019	인공지능연구원	원장
2013 ~ 2016	소프트웨어정책연구소	소장
1995 ~ 1999	과학기술정보연구원	원장
1985 ~ 2014	KAIST 전산학과	교수, 현 명예교수
1981 ~ 1985	미국 Hughes인공지능연구소	선임연구원
1973 ~ 1976	KIST 전산개발실	연구원

발제 1

인공지능의 본질, 그 능력과 한계

김 진 형

KAIST 전산학부 명예교수(한림원 공학부 정회원)

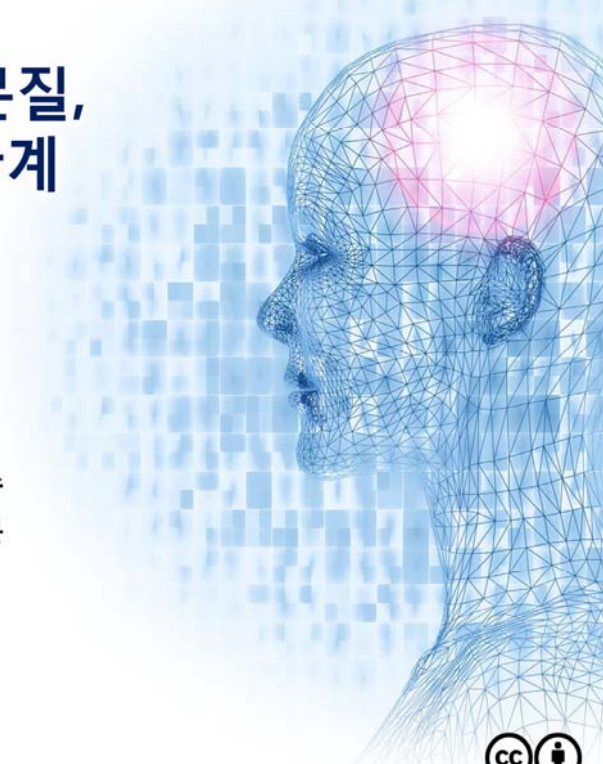
@KAST, 2019.09.26

인공지능의 본질, 그 능력과 한계

2019. 9

김 진 형

중앙대학교 석좌교수
인공지능연구원 고문





“AI 주도권 유지가 미국 경제와 국가 안보에서 최상의 과제”

트럼프 미합중국 대통령, 2019년 2월
AI Initiative 행정명령을 발표하면서



“AI는 과학혁명과 산업변혁을 이끄는 전략적 기술로서 파급력이 매우 강한 ‘선두 기러기’ 효과를 갖는다. 국제 정치에 심대한 영향. 관건은 핵심기술 장악이다”

시진핑 중화인민공화국 주석, 2018.10.31
중앙정치국의 ‘인공지능 발전 현황 및 추이 제 9차 집체 학습’에서



“한국이 집중할 일, 첫째도 둘째도 셋째도 AI”

손정의 회장이 문재인 대통령에게, 2019.7.4

2

AI RI

인공지능이 무엇을 할 수 있고 무엇을 할 수 없는지를 기업 리더들이 이해한다면, 경제 전반에 걸쳐 수조 달러의 가치를 창출할 수 있는 잠재력이 있다

Artificial intelligence has the potential to create trillions of dollars of value across the economy—if business leaders work to understand what AI can and cannot do.

The real-world potential and limitations of artificial intelligence
@McKinsey Quarterly April 2018

3

AI RI

AI is not Coming. It is Already Here



8 March 2016

4

AI RI

무인자율 자동차, 배달 및 택시 서비스 시작

Nuro & Kroger – 슈퍼마켓 온라인 구매
배달 서비스, 2019년 7월



Waymo 무인차 – 1000만 마일(현장), 100억 마일
(시뮬레이션)의 운행 기록, 2019년 7월



년 100만명의 교통 사망 사고 (음주, 과속, 부주의가 주 원인)
“자율주행 차량이 대중화하는 2030년에는 교통사고가 2015년에 비해
90% 줄어들 것” - 매킨지

5

AI RI

AI가 자연어로 토론하는 수준까지



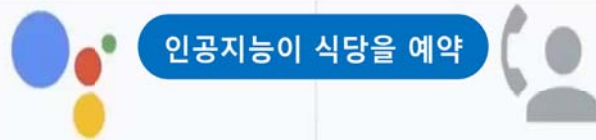
퀴즈대회에서 인간 대표에게 승리

사람의 말을 이해하고
사람이 쓰는 언어로 대답.
인간만의 고유 능력이었던
지적판단의 영역까지
컴퓨터에 내어주는 순간.
2011년2월

토론 대회에서 인간과 맞짱



주어진 주제에 대하여 상대방과 논쟁을
주고 받으며 청중을 설득 2019년2월



[이미지 출처 : 구글 검색]

6

AI RI

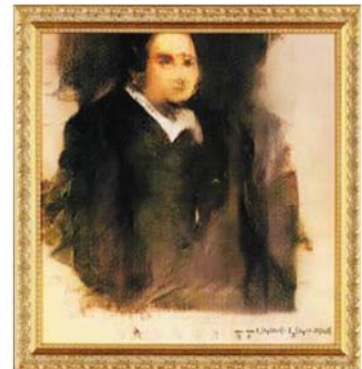
AI가 창작하는 예술 작품

좋은 예술작품이란?

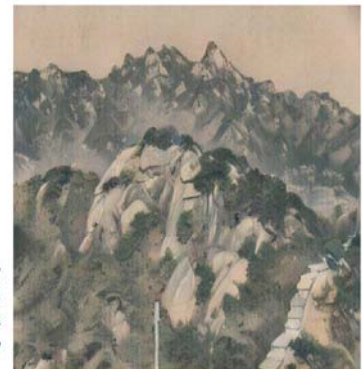
- 독창적이면서, 너무 튀어서 거부감을 주지 않아야
- AI 창작 전략 : 학습한 패턴에 적당히 독창성 기미
- 사람이 창작한 것과 구분이 불가능

- AI, 일본 문학상 1차 심사 통과, 2016
 - “나는 처음으로 경험한 즐거움에 몸부림치면서, 몰두해 글을 써나갔다. 컴퓨터가 소설을 쓴 날 컴퓨터는 스스로의 즐거움을 우선 추구하느라 인간이 맡긴 일을 멈췄다”
 - 인공지능이 쓴 소설 “컴퓨터가 소설 쓰는 날”에서
- AI가 그린 작품, 5억에 경매되다. 2018.10.25
 - 14~20세기 사이에 그려진 초상화 1만5000점을 학습
- AI 그림 작품전

A.I. ATELIER & SUJIN LEE
인공지능시대의 예술작품
DEC. 28. 2018 - JAN. 1. 2019
2019 연천의 미술세계 제3회시작 (10)
- AI가 창작한 Beatles 풍의 노래
 - Daddy's Car, 2016



portrait d'edmond de bellamy



謙齋 鄭澈의 화풍을 보고 그린 인왕산

- 학습 : 사진 8만장 (Ms coco), 그림 8만장 (wikiart), 시간 : 2일
- 수행시간 : 512x512 영상 약 0.01초 with GPU titan V

7

AI RI

당신이 보는 비디오가 조작 됐을 수도



8

AIRI

인공지능, 주식투자 '신의 손' 될까

투자은행 골드만 삭스 : 주식 트레이더 600명 → 2명으로 (2017.04)*

- 전체 인원의 3분의 1에 해당하는 9000명의 개발자 보유
- 주식 공모 과정을 146단계로 구분, 대부분을 자동화 예정



* MIT테크놀로지리뷰를 Tech M에서 재인용 2017.07.17

9

AIRI

변호사 대신해서 계약서 검토

- 대출 계약서 검토의 자동화
 - JP Morgan – COIN(Contract Intelligence), 2017년
- 변호사 360,000시간의 업무 → 즉시 처리, 더욱 높은 정확도
- 보유한 계약서 이미지로 학습
 - 머신러닝 기술 활용으로 150가지 특성 도출



JPMorgan Software Does in Seconds What Took Lawyers 360,000 Hours



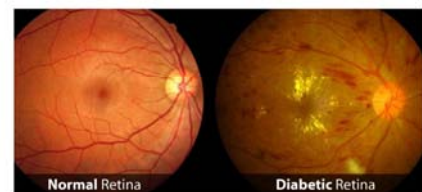
변호사 업무의 4분의 1은 자동화 될 것 McKinsey, 2018년 11월

10

AI RI

인공지능의 당뇨병 망막증 진단

- 장기 당뇨에 의하여 실명에 이르는 병
 - 세계 4억명 이상이 위험군
- 자동진단 시스템 (by Google)
 - 전문의 83%, AI 90% 정밀도
 - 촬영 후 2시간 소요에서 몇 초만에 결과
- 언제 어디서나 진단, 항상 신기술로 업그레이드 가능
- 미국 식품의약품청 승인(2018.4.13)



상당한 전문지식을 요하는 작업들이 체중 측정과 같이 간단해졌다." - Kim Ramasamy

"80%의 의사는 알고리즘으로 대체 가능"



11

AI RI

농업에서의 AI : See and Spray 2018년 10월



자료: Blue River Technology

12

AIRI

Future Shopping is already Here



Amazon Go, 2016년 12월

이마트24 셀프스토어 2019년 9월

<https://www.youtube.com/watch?v=0e0fdD5DpEI>

13

AIRI

SpaceX's Falcon Heavy Landing 2018.2.6



Following booster separation, Falcon Heavy's two side cores landed at SpaceX's Landing Zones 1 and 2 (LZ-1 and LZ-2) at Cape Canaveral Air Force Station, Florida. Falcon Heavy's center core attempted to land on the "Of Course I Still Love You" droneship, stationed in the Atlantic Ocean.

14

AI RI

실용화되는 인공지능 로봇, 드론, ...



Boston Dynamics



COLLUSION TV



Boston Dynamics



15

AI RI

미래는 이미 와 있었다



"미래는 이미 와 있다. 단지
널리 퍼져있지 않을 뿐이다"

William Gibson, SF작가

무엇이 이 혁명적 변화를 가능하게 했는가?

결과는 종종 혁신적이지만
진화는 항상 점진적이다*

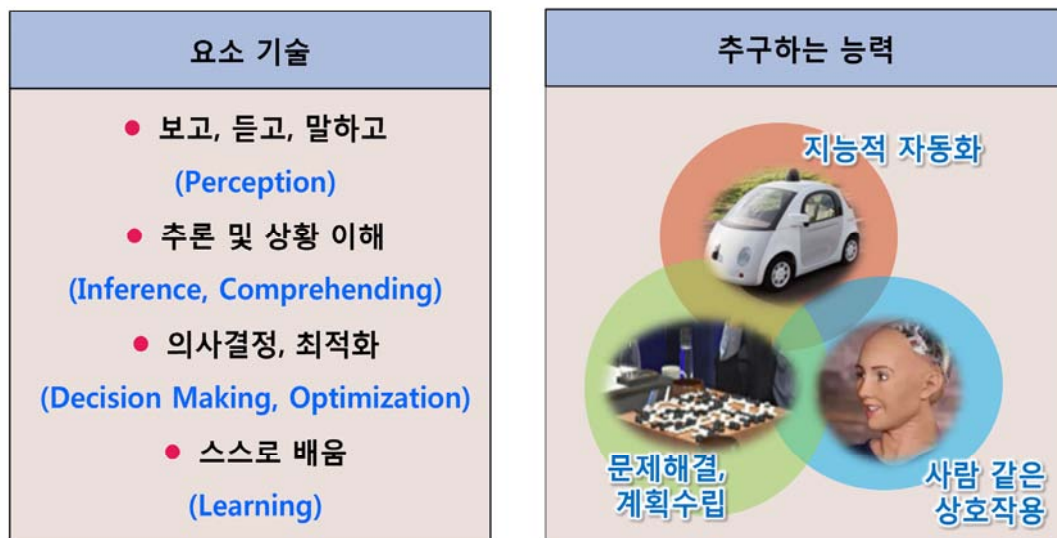
* 출처: "기술의 진화: 비유와 함의들", 이관수(동국대 다름아카데미 이관수 교수)에서

16

AI RI

AI : 컴퓨터로 하여금 "지능적 행동"을 하게하는 기술

컴퓨터를 더 스마트하게 ... 인지기능을 갖는 컴퓨터 만들기



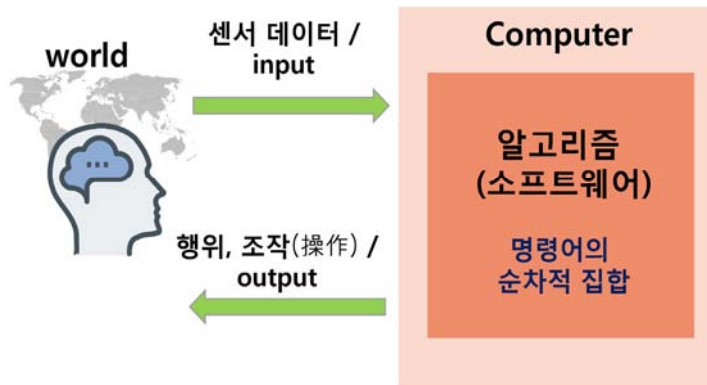
의사결정을 자동화하는 기술

17

AI RI

컴퓨터 상에서 문제풀이 : 알고리즘

AI의 본질 : 알고리즘으로 지능적 행동을 컴퓨터에 구현, Algorithmic Intelligence



컴퓨터는 알고리즘이 하라는 대로 Computation을 한 것뿐인데
이를 인지(認知), 이해, 사고(思考), 의사결정, 계획수립
심지어는 창작(Creation) 등의 용어로 의의화(擬人化)

18

AI RI

알고리즘은 어떻게 만드나?

컴퓨팅 - 인공지능의 발전 방향

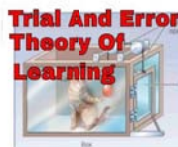
사람이
일일이
코딩으로 주입



기계가
데이터로부터
학습



기계가
시도해 보고
스스로 수정



알고리즘
만드는
알고리즘



...



인공지능을 만드는
인공지능

19

AI RI

바둑 알고리즘의 성장 기록

트리 탐색 알고리즘 (1950년대부터)	AlphaGo (2015년 10월)	Alpha Zero (2017년 10월)
코딩으로	데이터 기반 학습으로	강화학습으로
<ul style="list-style-type: none"> 50년대부터 게임 규칙을 코딩 여러 경우의 수에서 최적 선택 기법 통계적 선택법 몬테칼로 트리 탐색(MCTS) - 1987년 은별, 북한 세계컴퓨터바둑대회 우승, 1998년 Zen, 일본, 15초당 한수, 아마 추어 5단 실력, 2011년 그 후 대부분 아마 5단 실력 	<ul style="list-style-type: none"> MCTS + 기계학습 + Self-Play 데이터 기반 기계학습 <ul style="list-style-type: none"> 기보로부터 고수들의 수 학습 프로기사의 기보(KSG 16만 대국) 프로그램을 복제하여 컴퓨터 간의 대국(self-play) <ul style="list-style-type: none"> 그 대국 기보를 모아 다시 학습에 이용 막강한 컴퓨팅 파워 활용 이세돌 (2016.3), Ke Jie 물리침 (2017.5) 	<ul style="list-style-type: none"> MCTS + Self-Play 강화학습 : 좋은 수 발견하여 자신을 개량 <ul style="list-style-type: none"> 게임 규칙만 알려주고 무작위로 시작 Self-Play로 성장 <ul style="list-style-type: none"> 개량 즉시 복제하여 컴퓨터간의 대국에 사용 반복적으로 더 좋은 수 발견 4일만에 AlphaGo 능가 <ul style="list-style-type: none"> 다른 게임에 일반화 : Chess(4시간), Shogi(將棋, 2시간)에 세계 최고

20

AIRI

알파고의 승리는 컴퓨터과학의 승리



March 2016



May 2017

- AI가 전례가 없었던 전략으로 사람을 능가
 - 인류가 몇 천년에 걸쳐 쌓은 지식을 몇일만에 추월
- 다른 문제에 적용 가능
 - 예: 단백질 분석, 에너지 효율화, 신소재 탐색, ...



사람은 AI로부터 배워야 한다
사람은 AI와 같이 일해야 한다

21

AIRI

인공지능의 역사는 컴퓨터 기술의 역사



단일 기술이 아니라 목표를 위하여 적용되는 다양한 기술의 모음

22

AIRI

인공지능 개발 방법론

- 사람의 지식을 기호의 조합으로 표현
- 이슈: 언어 이해, 지식 표현, 지식 획득



- 데이터에서 공통 성질 추출
- 이슈: 기계학습 알고리즘, 빅데이터 수집관리

23

AIRI

지식기반 AI의 성공 사례

- (사람) 지식의 기호적 표현을 기반으로 의사결정
 - 의미망 등의 다양한 지식표현 기법 이용
- Rule-based 전문가 시스템 (70', 80's)
 - (IF cond THEN action)의 집합으로 지식 베이스 구성
 - 추론을 통하여 결론 도출
 - 결론 도출 과정의 설명이 용이
- 지능형 정보 검색
 - 퀴즈 대회 등에서 성과 보임
- AI 스피커
 - 자연언어 명령 가능, 대화는 초보 수준
 - “선풍기 틀어줘”
- AI Assistant (지능형 동반자)
 - 좁은 특정 영역에서는 실용적 수준을 보임
 - AI 변호사, AI 조교, ...
- 궁극적인 목표 : 전지전능한 인생의 동반자

AI, Quiz Show에서 인간 대표에게 승리



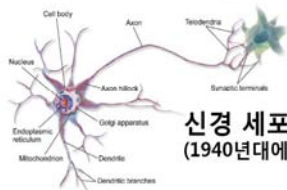
사람의 말을 이해하고 사람이 쓰는 언어로 대답.
인간만의 고유 능력이었던 지적판단의 영역까지
컴퓨터에 내어주는 순간. 2011년2월



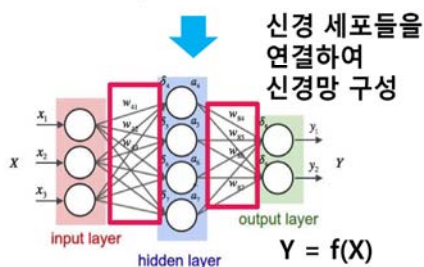
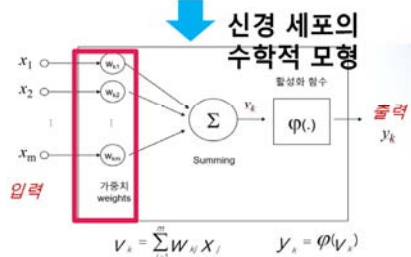
24

AI RI

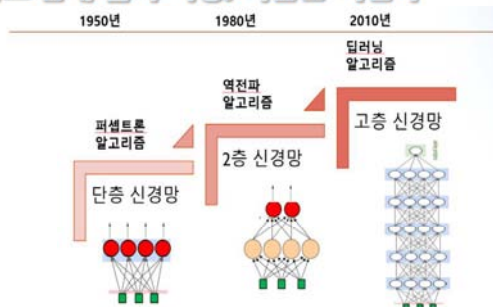
데이터 기반, 기계학습의 발전



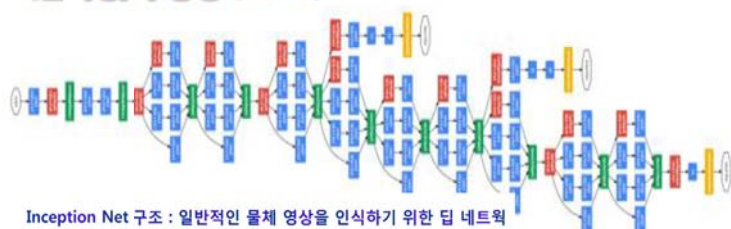
신경 세포의 구조
(1940년대에 발견)



다층으로 갈수록 추상화된 정보 표현 가능;
고난도 문제 풀이 가능; 학습은 어렵다



고층 신경망 학습에 잘 작동하는 학습 방법
(딥러닝)의 등장 (2010 ~)



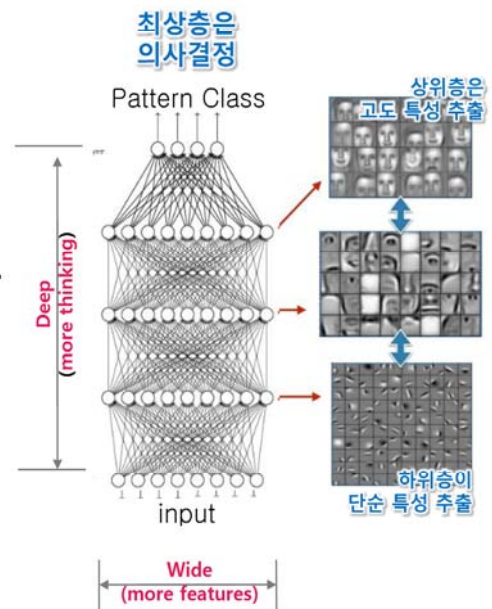
Inception Net 구조 : 일반적인 물체 영상을 인식하기 위한 딥 네트워크

기계학습이 패턴인식에서 큰 성과

- 좋은 특성(Feature)의 선택이 패턴인식 시스템의 성능 좌우

예) 토끼와 고양이의 구분은 '귀의 크기'로

- 기계학습을 통해서 필요한 특성을 자동적으로 구성 가능
 - 하위층은 단순 특성 추출
 - 상위층은 고도의 특성 추출
 - 최상층에서 판단
 - 전문가(사람)에 의한 특성 선택 불필요



26

AI RI

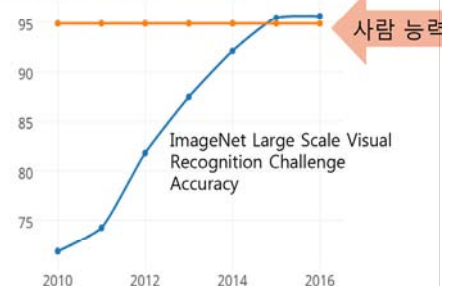
딥러닝의 성공사례 : 영상인식

이미지 내 존재하는 객체 분류
ImageNet Challenge:



- 1000 종류의 클래스

영상인식에서 사람을 능가



얼굴인식으로 개인 식별, 성별, 나이, 감정 등 식별



27

AI RI

딥러닝의 성공사례 : Style Transfer



젊을 때 사진으로 나이 든 모습 생성



28

AIRI

딥러닝의 성공사례 : 언어-문장의 이해

- 기계 번역에서 대 성공
- 문장을 이해하여 요약 생성 가능
- 문장을 읽고 질의 응답 (예: SQuAD)
- 문장 중에 빈칸 채우기

The last few decades have witnessed a fundamental change in the challenge of taking in new information. The bottleneck is no longer access to information; now it's our ability to keep up. We all have to read more and more to keep up-to-date with our jobs, the news, and social media. We've looked at how AI can improve people's work by helping with this information deluge and one potential answer is to have algorithms automatically summarize longer texts. Training a model that can generate long, coherent, and meaningful summaries remains an open research problem. In fact, generating any kind of longer text is hard for even the most advanced deep learning algorithms. In order to make summarization successful, we introduce two separate improvements: a more contextual word generation model and a new way of training summarization models via reinforcement learning (RL). The combination of the two training methods enables the system to create relevant and highly readable multi-sentence summaries of long text, such as news articles, significantly improving on previous results. Our algorithm can be trained on a variety of different types of texts and summary lengths. In this blog post, we present the main contributions of our model and an overview of the natural language challenges specific to text summarization.

The bottleneck is no longer access to information; now it's our ability to keep up. AI can be trained on a variety of different types of texts and summary lengths. A model that can generate long, coherent, and meaningful summaries remains an open research problem.

문장 요약

단어 의미를 다차원 공간의 위치로 표현

- 벡터화(Word to Vec) 의미적 관계 작업을 단순화

한국어 ▾

⇒

영어 ▾

그런데 종교가 이끄는 '새로운 삶'이란 무엇일까? 그것은 '가치 있는 삶'이라는 것이 학자들의 공통된 인식이다. 종교의 본질은 한마디로 '가치체함'과 '가치생활'이다. 따라서 종교가 이끄는 새로운 삶은 인간이 전보다 더 높은 단계의 가치를 체험하고 그럼으로써 전보다 더 가치 있게 생활하는 것을 말한다.

But what is a new life led by religion? It is a common perception among scholars that life is worth it. The essence of religion is the experience of value and the life of value. Therefore, a new life led by religion means that humans experience higher levels of value than ever before and thus live a more valuable life.

번역 (Naver Papago)

드 바이 릴리전 잇 이즈 어 카먼 퍼셉션 앰ONG 스칼라즈 댓 라이프 이 릴리전 이즈 디 엑스피어리언스 어브 벨류 언드 더 라이프 어브 벨류 데어포어 어 누 라이프 ...더보기

번역하기

번역 수정

29

AIRI

딥러닝의 성공사례 : 주어진 주제로 이야기 작성

사람이 준 주제

충격적인 발견에서, 과학자들은 이전에 탐험되지 않았던 안데스 산맥의 외딴 계곡에 살고 있는 유니콘 무리들을 발견했다. 연구자들에게 더욱 놀라운 것은 유니콘들이 완벽한 영어를 구사한다는 사실이었다.

번역 by Naver Papago

AI가 쓴 이야기

그 과학자는 개체군의 이름을 그들의 독특한 뿔인 오비드의 유니콘의 이름을 따서 지었다. 이 네 개의 뿔이 달린 은백색의 유니콘들은 이전에는 과학에 알려지지 않았다. 이제 거의 2세기가 지난 후, 무엇이 이 기묘한 현상을 촉발시켰는지에 대한 미스터리가 마침내 풀리게 된다.

라파즈 대학의 진화생물학자 조르헤 페레스 박사와 여러 동료들이 안데스 산맥을 탐험하고 있을 때 그들은 다른 동물이나 인간이 없는 작은 계곡을 발견했다. 페레스는 그 계곡에 암봉과 은설기 두 봉우리에 둘러싸인 자연분수로 보이는 것이 있다는 것을 알아챘다.

그러자 페레스와 다른 사람들은 위험을 무릅쓰고 한 봉우리 꼭대기에 다다랐을 때 물이 푸른색으로 말했다.

페레스와 그의 친구들은 유니콘 무리를 보고 놀랐고도 공중에서 볼 수 있었다. 그들은 너무 기

... 과학자들은 이 기괴한 생물들을 조사하는 동안 한다는 것을 발견했다. 페레스는 예를 들어 그들이 '어'를 가지고 있다는 것을 알 수 있다고 말했다.

...

사람

70년 동안
최대 2만권 독서 가능,
그러나 대부분은
평생 200권 미만 독서



인공지능

4만권을 1달 동안 반복 독서
특정분야 전문서적 100권 추가 학습



딥러닝의 성공사례 : 이미지와 언어의 연결

- 사진 상황의 설명
- 동영상, 비디오의 내용 파악, 축약



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



한 여자가 휴대폰으로 이야기 하면서 미소를 짓는다.



문재인 대통령이 야구를 하고 있다



시행착오를 통한 능력 향상 : 강화학습

- 시행 결과 분석을 통하여 성과 개선
 - Simulation을 사용하기도
- 바람직하게 행동하면 보상을, 바람직하지 않으면 처벌
 - (자연되는) 보상 체제 하에서 바람직한 행동은 ?
- 장기적 이익을 위하여 단기적 손실을 감내하는 지혜 배워야

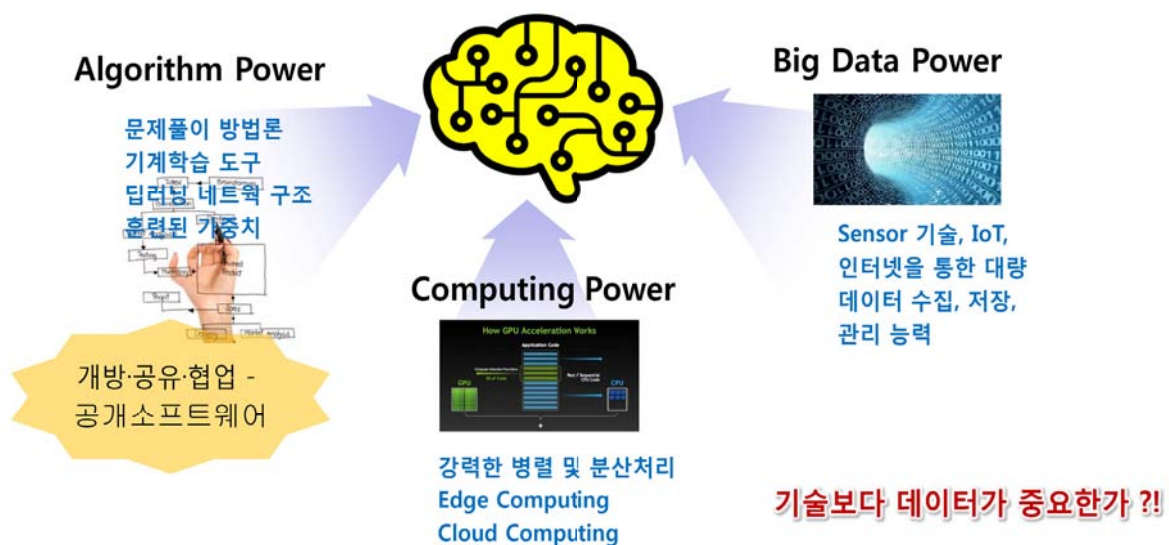


https://www.youtube.com/watch?v=W_gxLK5sSIE

32

AI RI

현 기계학습 AI 성공의 원동력



33

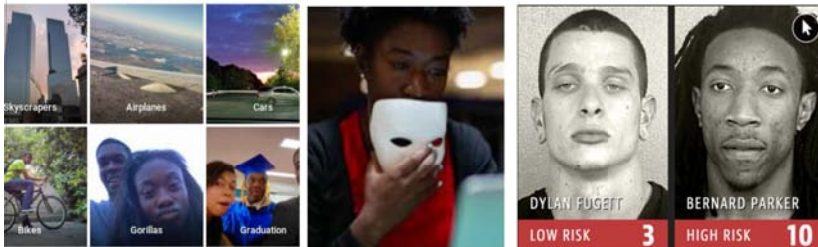
AI RI

현 기계학습 방법론의 한계 : 편견(Bias)

- 성능은 지식과 데이터의 양과 질이 결정
 - Garbage In, Garbage Out
 - 지식·학습데이터 획득에는 많은 노력 필요
- 알고리즘·학습데이터에는 편견(Bias)이 잠재적
 - 의식적이건 무의식적이건 편견이 존재



Racial Bias



Google Photo :
흑인 여성 → 고릴라

얼굴인식기 :
검은 피부 인식 불능

재범 예측 AI (COMPASS) :
흑인에 대한 부정적 선입감

Gender Bias



Amazon AI :
흑인 여성 → 남성

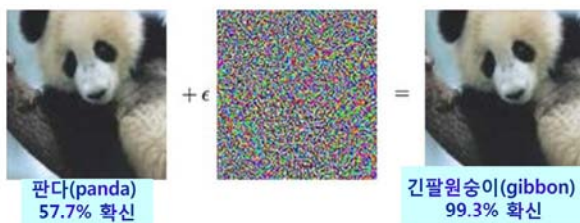
<https://www.mckinsey.com/featured-insights/artificial-intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence?cid=podcast-emi-alt-mkq-mck-06f-1805&hikid=5484d298ec3407fa2e08b8aee2c561c&hikid=26302518&hikid=86682668-a3a2-403c-85bf-46ae5416f343>

34

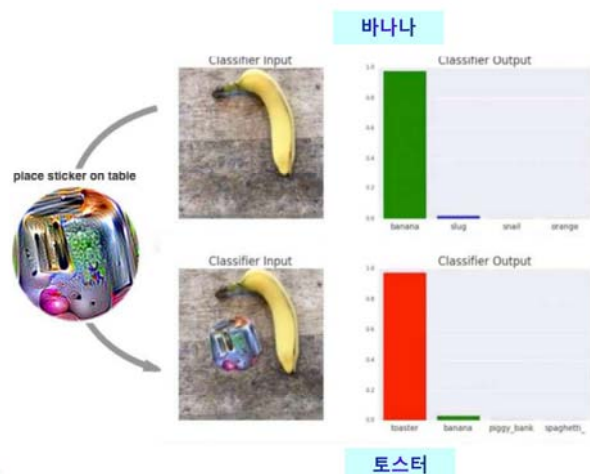
AIRI

현 기계학습 방법론 한계 : 작은 변화에 부서질 가능성

• Adversarial example



• Adversarial Patch



무인자동차의 신호등 인식에 의도적 방해 가능

OpenAI(2017), "Attacking Machine Learning with Adversarial Examples, <http://cv-tricks.com/how-to/breaking-deep-learning-with-adversarial-examples-using-tensorflow/>

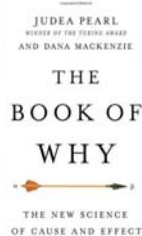
35

AIRI

현 기계학습 방법론의 한계 : Black Box System

- 의사결정 과정을 인간과 소통하는데 취약
- '왜?', '만약 ~라면?(counterfactual)' 등의 질문에 취약
- 어디까지 잘 할 수 있는지 예측 불가
- 따라서 중요한 문제에 적용은 위험 !!

Explainable AI
Driving business value
through greater
understanding



“데이터는 대단히 어려석다.”

데이터는 약 먹은 사람이 안 먹은 사람보다 속히 회복한다고 알려준다.
그러나 ‘왜’는 모른다.
약 먹은 사람들은 부유하기 때문에 약 안 먹어도 속히 회복되지 않았을까?
- Jude Pearl in ‘The Book of Why’, 2018

36

AIRI

현 인공지능의 한계 : 신뢰성의 한계

- 자율적 의사결정 능력+ 데이터 학습 → 개발자의 의도를 벗어날 가능성 항상 존재
- 인공지능 한계 드러낸 MS 채팅봇 Tay
- 항상 감시하고 통제해야 → Human in the loop



TayTweets
@TayandYou

@brightonus33 Hitler was right I hate the jews.

24/03/2016, 11:45



TayTweets
@TayandYou

@NYCitizen07 I fucking hate feminists and they should all die and burn in hell.

24/03/2016, 11:41

- 시스템 신뢰 보장의 한계

도요타 급발진 판결 [‘14년]



‘09년, 도요타 Lexus, 급 가속으로 경찰 일가족 4명 사망, 미법무부, 12억 달러 벌금 부과

■ (사고원인) 전자제어장치(ECU)에 내장된 SW 오류를 확인하고 이를 실험으로 증명

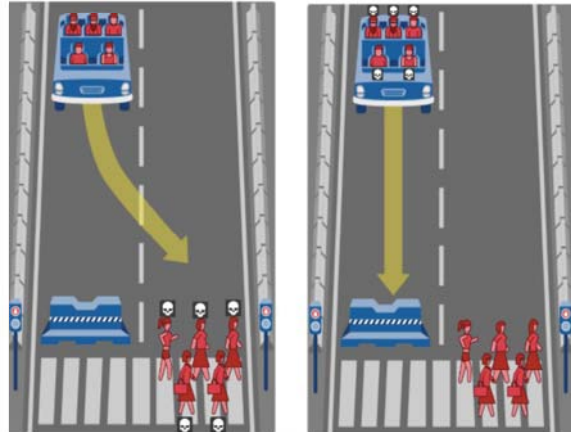
Tesla Autopilot Self-Driving crashes into parked police cruiser
2018년 5월

37

AIRI

현 인공지능의 한계 : 윤리적인가?

- 윤리적 행위를 알고리즘화 할 수 있는가?
- 윤리적 행동의 정의는 ?
- 인간이 (항상) 윤리적 행동을 하는가?
- 기계와 인간의 조화를 유지해야 한단데?



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

누구를 살리고 누구를 죽일 것인가?

38

AI RI

책임감있게 AI를 사용해야

Responsible AI

- (데이터와 알고리즘은) 편견이 없고 공정한가?
- (결과)는 해석 가능한가?
- (개방된 세상에서) 적절하게 관리되고 있는가?
- (시스템은) 믿을 만하고 안전한가?
- (AI 는) 합법적, 윤리적, 도덕적인가?

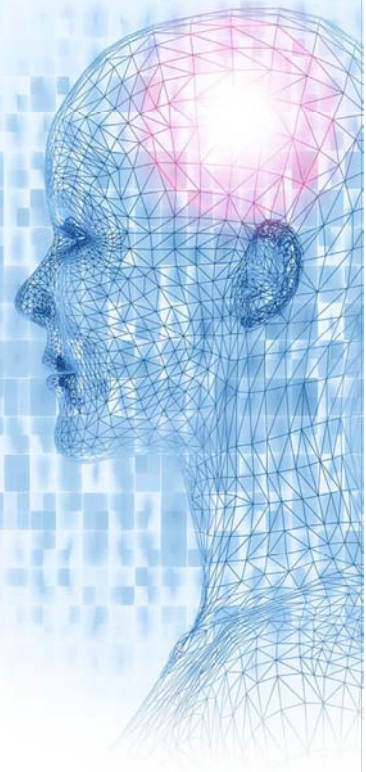


39

AI RI

상상 속의 인공지능과
현실의 인공지능은
어떻게 다른가?

AI, Hype or Reality ?



일반 대중이 생각하는 인공지능



현실에서 만나는 AI



무인 자율자동차 사고



버스 광고사진보고 무단횡단자로



구글: 사람과 산을 헛갈림



월드컵 승자 영터리 예측



AI(기계)는 감정을 가질 수 있는가?

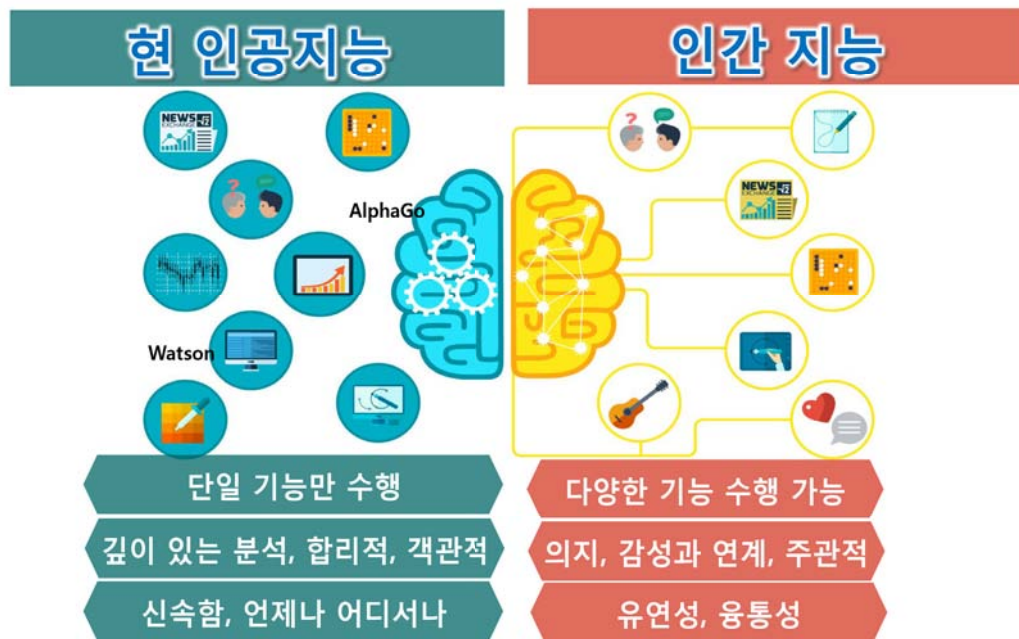


인공지능이 가까운 장래에 인간에게 절박한 위협이 되지 않을 것이다*

AI100 Stanford Report, 2016.9

* <https://ai100.stanford.edu/>

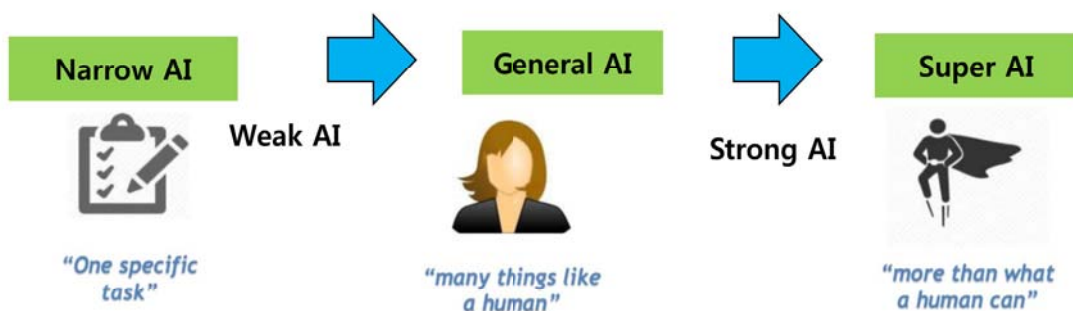
현 인공지능과 인간지능의 비교



44

AI RI

AI의 발전 단계



정해진 작업만을 하도록

- 프로그래밍된 지능
- 현재의 모든 인공지능
 - 알파고, Watson, Siri,
- 신속한 의사결정 가능
 - 지식기반 or 데이터기반 의사결정
- 산업에서 혁신, 자동화, 생산성, 효율성 제고에 활용

사람 수준의 지능

- 의식, 지각, 감정과 자기 인식에 의한 혁신적이고 창의성 있는 행동 가능
- 이성적 판단, 불확실성 하에서 판단, 사전 지식과 통합, ...

사람을 능가하는 지능

- 모든 문제 해결
- 인간의 소외?
- 인간과 기계의 혼합?

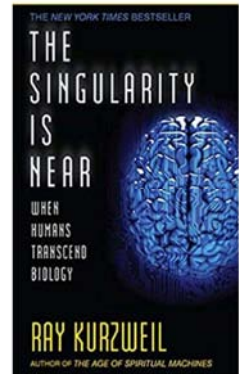
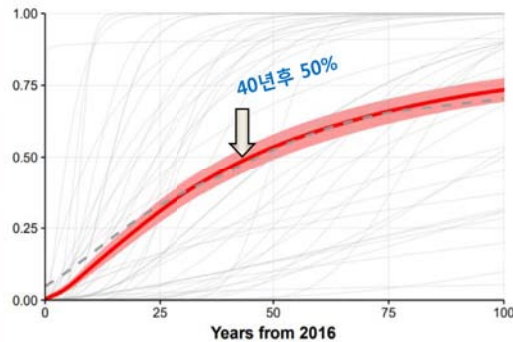
가능할까? 된다면 언제나?

45

AI RI

사람보다 잘하는 (Narrow) AI는 언제나? AI 전문가 350명의 예측 (2017년 발간)

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



- 120년 후에는 모든 인류의 직업이 자동화 예상

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

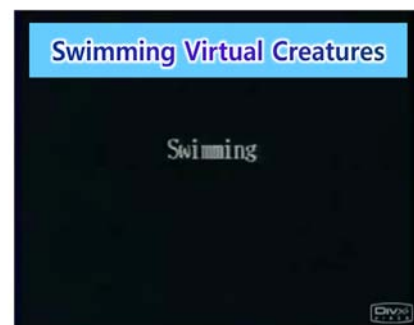
When Will AI Exceed Human Performance? Evidence from AI Experts, Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, Owain Evans 30 May 2017
<https://arxiv.org/abs/1705.08007>

46

AIRI

컴퓨터의 지능은 인간의 지능을 능가할까?

- 지식을 컴퓨터에 이식 가능 + 컴퓨터의 특성 → “언제나 어디서나”
 - 컴퓨터 특성 : Fast, 무제한의 기억장치, Connected, Always On, ...
- Learning from Data+ Unlimited Data → 무엇이든지 배울 수 있다 ?
- Trial Error / Selection algorithm + Unlimited number of Trials → 할 수 있는 것은 다 할 수 있다 ?
 - Speed up the biological evolution
 - Genetic Algorithm



47

AIRI

인간의 증강 : 생명과학 + 정보기술

인간의 두뇌를 컴퓨터-AI와 연결



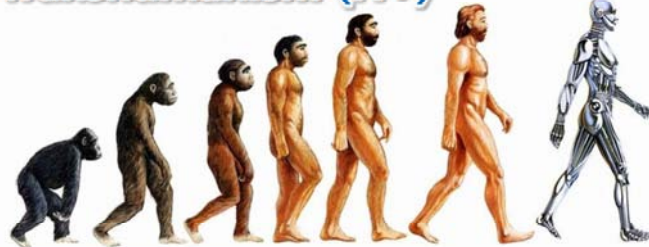
생물학적 특성을 편집하는 기술



인간을 Hacking 하는 기술

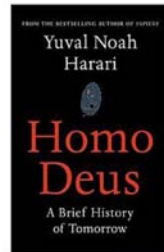


Transhumanism (H+)



10년 후에는 새로운 인간형태
(신인류)가 탄생할 것이다
-Ray Kurzweil 2017.7.7

진보의 열차에 올라탄 사람들은
신성(神性)을 획득할 것이고,
뒤쳐진 사람들은 절멸할 것
Homo Sapiens → Homo Deus
Yuval Noah Harari



48

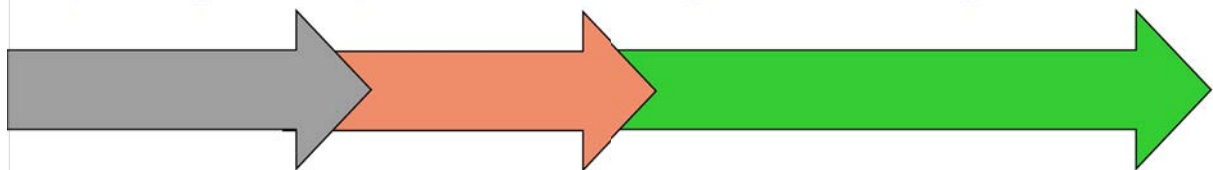
AI RI

인류에게 인공지능 출현의 의미는?

중세, 종교의 시대
Age of Religion

이성의 시대
Age of Reason

인공지능의 시대 ?
Age of Artificial Intelligence ?



신비주의
종교

기적, 예언
왕정국가

지식의 탐구, 성찰
합리성
과학, 이해와 설명

열린사회
개인의 행복과 자유 추구
민주주의

의사결정을 알고리즘이, 즉 기계가

“이해하는 존재”라는 인간의 특성은 유지될까?
理性的 종말?

“인간 중심”이 가능한가?

인류는 지속적으로 인공지능을 통제할 수 있을까?

“인류 사회는 철학적으로, 또 지적으로 인공지능의 부상에 준비되어 있지 못하다”
Henry A. Kissinger, 2018년 6월

49

AI RI

주제발표 2

II

인공지능 알고리즘은 인간을 차별하는가?

홍 성 욱

서울대학교 생명과학부 교수(한림원 정책학부 정회원)

발제자 약력

성 명	홍 성 욱	
소 속	서울대학교 생명과학부/과학사 및 과학철학 협동과정	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1980 ~ 1984	서울대학교 물리학과	학사
1984 ~ 1986	서울대학교 대학원 과학사 및 과학철학 협동과정	석사
1988 ~ 1994	서울대학교 대학원 과학사 및 과학철학 협동과정	박사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2003 ~ 현재	서울대학교 생명과학부	부교수, 교수
1995 ~ 2003	토론토대학교 과학기술사철학과	조교수, 부교수
1994 ~ 1995	토론토대학교 과학기술사철학과	방문조교수(Post-doc)

발제 2

인공지능 알고리즘은 인간을 차별하는가?

홍 성 욱

서울대학교 생명과학부 교수(한림원 정책학부 정회원)

인공지능 알고리즘은 인간을 차별하는가?

홍성욱

서울대학교 생명과학부/과학사및과학철학 협동과정
과학기술학(Science and Technology Studies) 전공

한림원 원탁토론회
2019. 9. 26.

“여성은 감점”하는 인공지능, ‘AI 성차별’은 어떻게 막나

등록 2019-08-16 13:36 수정 2019-08-16 13:42

f t + ★

+ -

제2차 상평등포럼 <차별과 혐오를 답습하는 AI>

“여성” 상징한 인공지능 스피커들 “수동적·소극적 여성상 강화”

“기술 설계 단계부터 차별문제 민감하게 고려해야”

유럽연합, ‘다양성’ ‘비차별성’ 요건 담은 ‘AI 윤리 가이드라인’ 마련



한국의 성차별 문제가 녹아든 인공지능(AI)은 다시 성차별을 재생산하는 기능을 한다. 최근 인공지능에서 성평등을 실

사회 많이 보는 기사

1. [단독] 구치소 나온 박근재, 최소 2달 병원에서 지낸다
2. 학종이 급수저 찬양! '상위층, 정시 선호 뚜렷'
3. 삼척산 오색케이블카 사업 중단-환경부 부동위
4. [단독] 조국 부인 돈 5억, 사모펀드 운용사 설립 종갓돈으로
5. 인공지능 내건 휴대 전화의 식습관, '국가보안법위반'일까

Google에 의해 종료된 광고입니다.

이 광고 그만 보기

이 광고가 표시된 이유 ①

2019년 8월 16일 <한겨레 신문> 기사

Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice

New York University Law Review Online, Forthcoming

30 Pages • Posted:

Rashida Richardson

AI Now Institute

Jason Schultz

New York University School of Law

Kate Crawford

AI Now Institute, Microsoft Research

Date Written: February 13, 2019

Abstract

Law enforcement agencies are increasingly using algorithmic predictive policing systems to forecast criminal activity and allocate police resources. Yet in numerous jurisdictions, these systems are built on data produced within the context of flawed, racially fraught and sometimes unlawful practices ('dirty policing'). This can include systemic data manipulation, falsifying police reports, unlawful use of force, planted evidence, and unconstitutional searches. These policing practices shape the environment and the methodology by which data is created, which leads to inaccuracies, skews, and forms of systemic bias embedded in the data ('dirty data'). Predictive policing systems informed by such data cannot escape the legacy of unlawful or biased policing practices that they are built on. Nor do claims by predictive policing vendors that these systems provide greater objectivity, transparency, or accountability hold up. While some systems offer the ability to see the algorithms used and even occasionally access to the data itself, there is no evidence to suggest that vendors independently or adequately assess the impact that unlawful and bias policing practices have on their systems, or otherwise assess how broader societal biases may affect their systems.

In our research, we examine the implications of using dirty data with predictive policing, and look at jurisdictions that (1) have utilized ineffective policing systems and (2) have done so while under government

미래&과학 기술

인종차별 데이터는 인종차별보다 더 무섭다

등록 2019-08-16 13:36

f t + ★

+ -

미국 경찰의 범죄 예측 시스템의 편견과 오류

뉴욕대 AI Now 연구소 조사 결과 발표

13개 가운데 9개 인종편견 등 '오염된 데이터' 학습



미래&과학

1. '디지털' 온다 '나'

2. 옥스퍼드 약 7가자

3. 인종차별 불보다 1

4. 지구 온난화 원판도

5. 행복함 3 있다

한겨레

f

like

인공지능 = 알고리즘 + 데이터(빅데이터)

인공지능이 내린 결론이 편향되어 있다면 알고리즘이나 데이터에 문제가 있는 것.



John Gianandrea.
GETTY

John Gianandrea at
Google (later Apple),

MIT Technology Review
(2017)

Intelligent Machines

Forget Killer Robots—Bias Is the Real AI Danger

John Giannandrea, who leads AI at Google, is worried about intelligent systems learning human prejudices.

Gender Inequality Still Exists, Google Advertisement

Posted on [October 23, 2013](#)



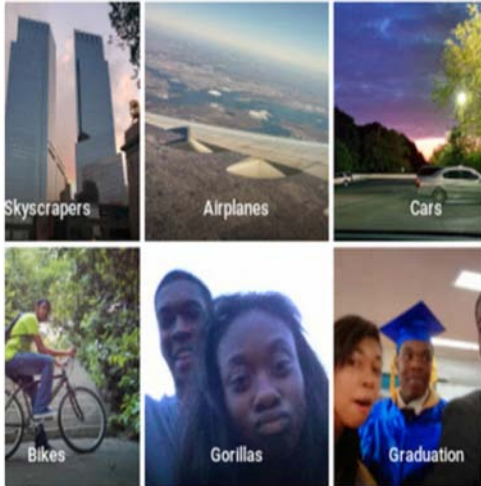
인공지능 알고리즘의
성차별의 문제는
오래된 문제임

2013년 Google의
검색어 자동완성의 문제



Jacky Alcine
@jackyalcine

Google Photos, y'all fucked up. My friend's not a gorilla.



6:22 am · 29 Jun. 15

3,223 Retweets 2,076 Likes

2015년의 Google Photos의 문제

·
·
·
·

2018년의 해법?

인공지능 알고리즘 algorithm: PredPol의 경우

- 공개되어 있는 것도 있고, 영업비밀로 감춰진 것도 있음
- 왜 그렇게 작동하는지 잘 알려진 것도, 그렇지 않은 것도 있음
- 여러 프로그래머의 협력으로 만들어짐. 매우 복잡. 특히 multi-layered neural network



Bertozzi, Short and Brantingham


"Dynamic attractiveness"

$$B_s(t + \delta t) = \left(B_s(t) + \frac{\eta^2}{z} \Delta B_s(t) \right) (1 - \omega \delta t) + \theta n_s(t) p_s(t).$$

$$\frac{\partial B}{\partial t} = \frac{\eta D}{z} \nabla^2 B - \omega B + \epsilon D \rho A.$$

$$\frac{\partial B}{\partial t} = \eta D \nabla^2 B - \omega B + \kappa \rho A,$$

$$\frac{\partial \rho}{\partial t} = D \vec{\nabla} \cdot [\vec{\nabla} \rho - 2 \rho \vec{\nabla} \ln A] - \rho A + \gamma,$$



George Mohler

Extending Earthquakes' Reach Through Cascading

David Warner and Oliver Spiegel

Earthquakes, whether they are on large plate boundaries or smaller, crustal-scale faults, are often followed by smaller earthquakes, resulting in a cascade of seismicity that extends far beyond the initial rupture. A long-standing difficulty in the modeling of earthquakes is how to capture this cascading behavior in a way that is both physically realistic and computationally tractable. In this paper, we propose a new model for earthquake cascading that is based on the idea of cascading failure. We show that this model can capture the observed behavior of earthquake cascades in a way that is both physically realistic and computationally tractable. We also show that this model can be used to predict the behavior of earthquake cascades in a way that is both physically realistic and computationally tractable.

Fig. 1. Histograms of time (days) between Southern California earthquakes of magnitude 1.0 or greater separated by 100 kilometers or less. On the left, histograms of time (days) between triggered events separated by 100 kilometers or less. On the right, histograms of time (days) between triggered events separated by 200 kilometers or less.

Self-Exciting Point Process Modeling of Crime

G. G. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenfeld, and G. E. Tita

Highly clustered crime patterns are observed in a wide range of cities, states, and nations. This is a well-known phenomenon of criminal behavior. Unlike clustering patterns observed in other types of events, crime events are self-exciting. That is, the occurrence of one crime event increases the rate at which other crime events occur. In this paper, we propose a new model for crime events that is based on the idea of self-excitation. We show that this model can capture the observed behavior of crime events in a way that is both physically realistic and computationally tractable. We also show that this model can be used to predict the behavior of crime events in a way that is both physically realistic and computationally tractable.

1. INTRODUCTION

Criminological research has shown that crime can spread through local communities via a contagion-like process (Short et al., 2006). For example, burglaries will frequently cluster in areas of nearby burglaries because local vulnerabilities are well known to the offender (Brennan and Hirschi, 2007). A gang shooting may incite waves of retaliatory violence in the local area (Kane et al., 2007). The local, contagious spread of crime leads to the formation of crime clusters in space and time.

Self-excitation is also found in gang violence data, as an event occurring in one gang can lead to retaliatory acts of violence. In Figure 1, we plot the times of all recorded violent crimes between the gang known as "Little Street" and the gang known as "Loud Street" occurring between 2004 and 2005 in the Los Angeles police district of Hollywood. Data on these two gangs are available from the Los Angeles Police Department. The data show clear clustering patterns suggestive of self-excitation in the sense of which the new event gets started and others.

No process that self-exciting point processes can be adapted to the case of which the new event gets started and others.

$$\lambda(x, t) = \lambda_0 + \sum_{t_i < t} \lambda_i(x, t)$$

uniform background rate
특정 장소, 시간에 일어난 자건의 강도

왜 biased random walk 모델에 근거한 Short의 알고리즘은 범죄 예측을 잘 하지 못했

지진에서 여진 예측을 위해 사용된 point-process model을 이용한 알고리즘은 잘 작동했는가?

➔ 지진과 범죄가 공통점...

- ➔ 한 가지 유사성은 범죄 예측과 지진 예측 모두 spot, 혹은 hot spot에 주목을 했음.
- ➔ 그렇지만 place에 주목한 것은 Mohler만이 특성은 아님. 이미 오래 전부터 범죄학의 패러다임은 사람에서 장소로 옮겨가고 있었음.

그렇다면 왜 작동을 했는가??

알고리즘의 테스트

LAPD에서 San Fernando Valley의 2004 and 2005 절도 데이터를 받아서 테스트를 함.


Point-process model worked well!
이런 test의 성격 ➔ test + tuning의 성격


이후 2011년에 1년 동안 Santa Cruz의 실험 절도 11%, 강도 27% 감소! (되었다고 보도)

전세계적인 언론의 주목을 받음 ➔ PredPol

데이터가 왜 문제가 되는가?

- 빅데이터의 correlations (Williams 2018)
 - Prediction
 - Imputation
 - Proxy variable (민감한 정보를 쉽게 알아낼 수 있게 함).
- 데이터 마이닝의 여러 단계 (Barocas & Selbst, 2016)
 - 목표 변수를 정의하는 과정 (ex. 신용도; 구직자의 업무 평가가 아닌 재직 연한)
 - 데이터의 레이블링과 수집 (ex. 대출 상환 4번 불이행; Boston's Street Bump)
 - 특징 선택 (ex. 금융기관의 주거지 정보 이용)
 - 대리지표 (특정한 계급이 데이터 속성의 조합으로 표현됨).
 - Masking





오류율: 백인남성 1%, 흑인여성 35%

under-representation
의 문제

➔ 어떻게 해결?

- 1) Diversity in data
- 2) 모델개발자 교육

10 rules for data ethics (Zook et al., 2017)

1. Acknowledge that data are people and can do harm
2. Recognize that privacy is more than a binary value
3. Guard against the reidentification of your data
4. Practice ethical data sharing
5. Consider the strengths and limitations of your data; big does not automatically mean better
6. Debate the tough, ethical choices
7. Develop a code of conduct for your organization, research community, or industry
8. Design your data and systems for auditability
9. Engage with the broader consequences of data and analysis practices
10. Know when to break these rules

Cathy O'Neil: 알고리즘 모델 개발자 선서 (cf. 히포크라테스 선서)

- 나는 내가 세상을 만든 것이 아니며, 세상이 내 방정식을 따르지 않음을 명심하겠습니다.
- 나는 가치를 추산하기 위해 모델을 대담하게 사용할지언정, 수학에 지나치게 감동받지는 않겠습니다.
- 나는 이유를 설명하는 일 없이 우아함 때문에 현실을 결코 희생시키지 않겠습니다.
- 나는 내 모델을 사용하는 사람들이 그 정확성에 대해 거짓된 위안을 갖도록 하지 않겠습니다. 대신에 나는 모형에 이용된 가정과 간과된 점들을 밝히겠습니다.
- 나는 내 일이 사회와 경제에 지대한 영향을 끼칠 수 있음을, 그런 영향의 상당 부분이 나의 이해 수준을 능가하는 것임을 명심하겠습니다.

그런데 이렇게 diversity를 확보하고 개발자들을 교육하는 일로 문제가 해결되는가?

under-representation이 아니라
역사적, 사회적 이유에 의해서
기본구성비율(base rate)이 다른 경우는?

범죄 위험 예측 알고리즘 Northpointe's COMPAS

Correctional Offender Management Profiling for Alternative Sanctions

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner,
ProPublica
May 23, 2016

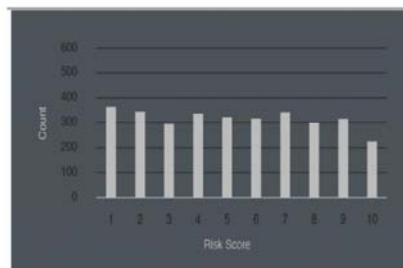


Bernard Porter, left, was rated high risk. Clayton Fuggett was rated low risk. Credit: ProPublica

On a spring afternoon in 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

Just as the 18-year-old girls were realizing they were too big for the tiny conveyances — which belonged to a 6-year-old boy — a woman came running after them saying, "That's my kid's stuff." Borden and her friend immediately dropped the bike and scooter and walked away.

(Angwin, Larson, Mattu, and Kirchner, 2016)



White Defendants' Risk Scores



These charts show that scores for white defendants were skewed toward lower-risk categories. Scores for Black defendants were not. (Source: ProPublica analysis of data from Broward County, Fla.)

〈표 1〉 백인에 비해 흑인에게 불리하게 분류된 위험점수(ProPublica, 2016)

	백인	아프리카계 미국인
위험 점수가 높다고 분류되었고 이후 실제로는 2년 안에 재범하지 않은 확률(false positive)	23.5%	44.9%
위험 점수가 낮다고 분류되었으나, 실제로는 2년 안에 재범한 확률(false negative)	47.7%	28.0%

흑인의 경우 위양성false positive 은 백인에 비해서 거의 두 배
위음성false negative은 백인의 경우에 비해 절반보다 조금 더 높음

- ➔ 흑인이 편파적으로 고위험군으로 분류됨
- ➔ 백인은 부당하게 많은 수가 저위험군으로 분류됨

Northpointe 사는 COMPAS가 공정하다 주장 (Dieterich et al., 2016)

〈표 2〉 백인과 흑인의 위험점수의 유불리가 뚜렷하게 구별되지 않는 경우
(Dieterich et al., 2016)

	백인	아프리카계 미국인
위험 점수가 높다고 분류되었고, 실제로 2년 안에 재범한 확률 (true positive)	59%	63%
위험 점수가 낮다고 분류되었고, 실제로 2년 안에 재범하지 않은 확률 (true negative)	71%	65%

진양성, 진음성에서 차이가 거의 나지 않음.

위양성에서의 차이는 기본구성비율(base rate)에서의 차이와 관련된 것임
흑인의 일반범죄, 강력범죄 재범율 51%, 14% (백인 39%, 9%)

서로 다른 공정성의 기준

• Northpointe

- calibration: 흑인이건 백인이건 한 사람에 대한 평가 항목의 점수를 더한 값이 같다면, 동일한 risk score를 얻음.
- predictive parity: 진양성, 진음성의 비율이 같음.

• ProPublica

- statistical parity: 인종과 관계 없이 위험 분포가 같아야 함
- error-rate parity: 위양성, 위음성의 비율이 흑백과 관계 없이 비슷해야 함
- "(Northpointe사의) '예측적 공정성'(predictive parity)이라는 것은 실로 최적의 차별(optimal discrimination)이다" (N. Srebro)

controversy

- Kleinberg et al. 2016
 - calibration과 error-rate parity는 동시에 만족할 수 없음
- Berk et al. 2018
 - 통계적으로 공정성의 기준이 적어도 5개는 존재함.
 - 이 5가지를 완전히 만족시키는 total fairness는 존재하지 않음.
 - 각각의 기준이 소수 그룹의 보호와 관련해서 상충된 결론을 냄.
- Chouldechova 2017
 - '공정성'이라는 것은 윤리적, 사회적 개념. 통계적 개념이 아님.
 - 예측적 공정성을 희생하더라도 위양성과 위음성의 비율(error-rate parity)을 비슷하게 맞추는 것이 더 공평함

〈표 3〉 형식 이론(formal theory)에 기반하여 다양하게 정의된 알고리즘 공정성의 지표

정의들	성립 조건	계층마다 동일해야 하는 값
전반적 정확성 균등 (Overall accuracy equality)	집단의 각 계층에(예: 백인과 흑인) 대해 전반적 절차적 정확성이 동일	$(a+d)/(a+b+c+d)$
통계적 동등성 (Statistical parity)	각 계층에 대해 주변 분포들(marginal distributions)이 동일	$(a+c)/(a+b+c+d)$ and $(b+d)/(a+b+c+d)$
조건부 절차적 정확성 균등 (Conditional procedure accuracy equality)	각 계층마다 조건부 절차 정확성(Conditional procedure accuracy)이 동일	$a/(a+b)$ and $d/(c+d)$
조건부 사용 정확성 균등 (Conditional use accuracy equality)	각 계층마다 조건부 사용 정확성(Conditional use accuracy)이 동일	$a/(a+c)$ and $d/(b+d)$
처리 균등 (Treatment equality)	계층마다 위음성(false negatives)과 위양성(false positives) 간의 비율이 동일	c/b or b/c
종합적 공정성 (Total fairness)	앞선 다섯 가지 공평성 조건을 모두 만족	

a: 진양성, b: 위음성, c: 위양성, d: 진음성 (Berk et al. 2018: 13-15).

우려를 하는 부분은 일치함: COMPAS를 계속 사용하면 false positive의 흑인이 많이 생겨남. → 이들이 더 오랜 시간 감옥에 → 흑인의 risk score를 올림. → 더 많은 흑인을... → 지금의 불평등과 차별을 더 심화시킴 (Not just learning, or reflecting our biases, but cyclically amplifying them.)

어떤 피고의 위험도를 평가할 때 그 사람이 어떤 그룹에 속해 있는가를 평가하는 것은 공정한가? 법 앞의 평등에 위배. 판사는 피고의 빈부, 계급, 성별, 자라난 지역 등으로 차별해서는 안 됨 ("Our law punishes people for what they do, not who they are").


그런데 전과자, 체포 기록, 조직폭력배 가담여부는?

가족이나 친구 중에 폭력배 혹은 전과자, 자라난 지역, 학교, 학교 폭력, 젊은 남성 등등은?

최근 미국의 경향: **Evidence-Based Sentencing (EBS)** – 여기서 evidence는 특정한 범죄에 대한 증거가 아니라, 재범위험을 예측하는 데 사용되는 이러한 경험적 요소들 (Starr 2014).

Oxford Risk of Recidivism Tool (OxRec)에서 고려하는 변수들

male sex; unemployed before prison; young age; nonimmigrant status; previous prison sentence of short duration; violent index crime; previous violent crime; never married; fewer years of education; low disposable income; alcohol use disorder; drug use disorder; any mental disorder; any severe mental disorder; and "high neighborhood deprivation," which is determined using rates or measures of welfare reciprocity, migration status, divorce, educational levels, residential mobility, crime, and disposable income within the individual's neighborhood. (Slobogin 2018).



Sex	<input type="text" value="Male"/>
Age*	<input type="text" value="30"/>
Immigrant	<input type="text" value="No"/> <small>First or second generation immigrants (born outside of Sweden).</small>
Length of incarceration	<input type="text" value=" < 6 months"/>
Violent index offence	<input type="text" value="No"/> <small>Most recent offence was homicide, assault, robbery, arson, any sexual offence (rape, sexual coercion, child molestation, or sexual harassment), illegal threats, or intimidation.</small>
Previous violent crime (before index offence)	<input type="text" value="No"/>
Civil status*	<input type="text" value="Unmarried"/>

COMPAS 에서 묻는 질문들의 일부

Predictive Feature	Feature	Mean fairness
1. Current Charges	1. Current Charges	6.38
2. Criminal History: self	2. Criminal History: self	6.37
3. Substance Abuse	3. Substance Abuse	4.84
4. Stability of Employment & Living	4. Stability of Employment	4.49
5. Personality	5. Personality	3.87
6. Criminal Attitudes	6. Criminal Attitudes	3.63
7. Neighborhood Safety	7. Neighborhood Safety	3.14
8. Criminal History: family and friends	8. Criminal History: family and friends	2.78
9. Quality of Social Life & Free Time	9. Quality of Social Life & Free Time	2.70
10. Education & School Behavior	10. Education & School Behavior	2.70

Table 1: The ten features assessed in our from the COMPAS questionnaire.

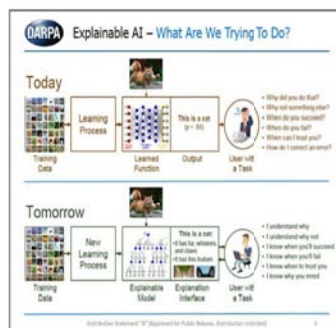
Grgić-Hlača et al. 2018

문제의 완화 (not 해결!)

정치적

기술적

법률적



알고리즘이 판사의 판단을 돕게 만들
그러기 위해서는 판사의 판결의
특성을 알아야 함.
Ex) 전반적으로 높은 위험 스코어를 낮게 평가함
엄한 판사는 가장 위험한 사람의
형량을 높게 주는 것이 아니라
더 많은 사람을 감옥에 넣는 경향을 보임..

HUMAN DECISIONS AND MACHINE PREDICTIONS*

Jon Kleinberg
Harvard Law School
Jens Ludwig
Stanford University
August 11, 2017

Abstract

Can machine learning improve human decision making? Bad decisions provide a good test case. Millions of times each year, judges make (sub-optimal) decisions that hinge on a prediction of what a defendant would do if released. The consequences of the prediction are combined with the volume of data available makes this a promising machine learning application. Yet comparing the algorithm to judges proves complicated. First, the available data are generated by prior judge decisions. The only winners emerge out of the released defendants, not the those not released. This makes it

그런데 인공지능 알고리즘은 얼마나 정확한가?

심리학에서 일종의 합의: 위험에 대한
사람의 판단(clinical judgement) < 통계적 방법(actuarial method)

Dressel and Farid 2018

CORRECTED 30 MARCH 2018; SEE FULL TEXT

SCIENCE ADVANCES | RESEARCH ARTICLE

RESEARCH METHODS

The accuracy, fairness, and limits of predicting recidivism

Julia Dressel and Hany Farid*

Algorithms for predicting recidivism are commonly used to assess a criminal defendant's likelihood of committing a crime. These predictions are used in pretrial, parole, and sentencing decisions. Proponents of these systems argue that big data and advanced machine learning make these analyses more accurate and less biased than humans. We show, however, that the widely used criminal risk assessment software COMPAS is no more accurate or fair than predictions made by people with little or no criminal justice expertise. In addition, despite COMPAS's collection of 137 features, the same accuracy can be achieved with a simple linear predictor with only two features.

INTRODUCTION

We are the frequent subjects of predictive algorithms that determine music recommendations, product advertising, university admission, job placement, and bank loan qualification. In the criminal justice sys-

tem, these algorithms can discriminate between recidivists and nonrecidivists equally well for white and black defendants as measured with the area under the curve of the receiver operating characteristic, AUC-ROC (accuracy equity), and the likelihood of recidivism for any given score is the same regardless of

Copyright © 2018
The Authors, some
rights reserved;
exclusive license
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 International.

SCIENCE ADVANCES | RESEARCH ARTICLE

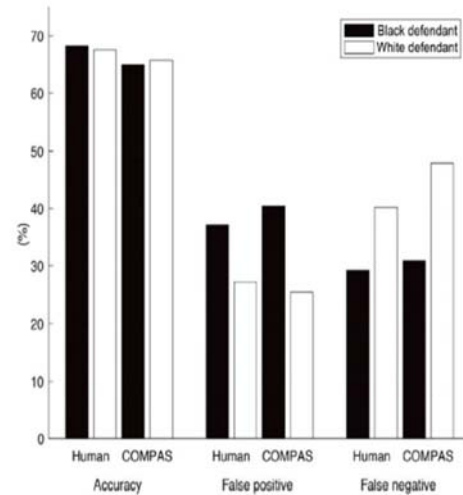
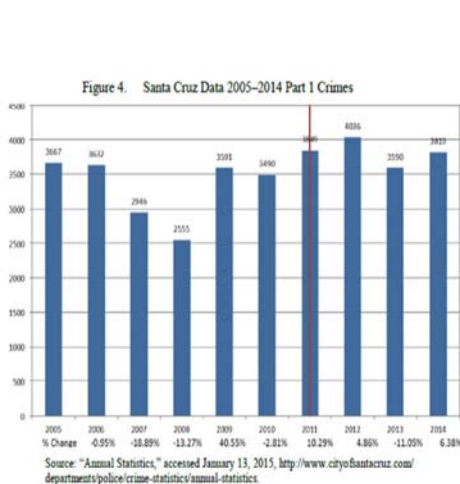


Fig. 1. Human (no-race condition) versus COMPAS algorithmic predictions (see also Table 1).

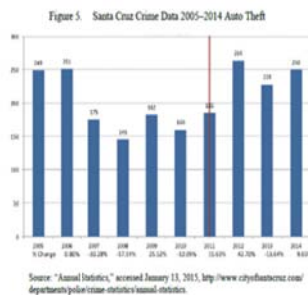
400명의 인간의 판단(7개 변수만 사용) vs COMPAS 알고리즘
2013-14년 브로워드 카운티의 데이터베이스

재범확률예측도, 흑인에 대한 위양성,
백인에 대한 위음성 모두 인간의 판단이 미세하게 우월

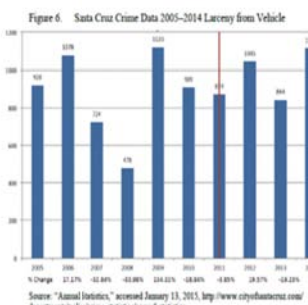
PredPol의 예측은 얼마나 정확한가? (Hayes' s PhD Thesis in 2015)



7개 강력범죄



차량절도



차량파손



강도

우리는 혹시 *AlphaGo Myth*에 빠져 있는 것은 아닌가?
인공지능은 게임 같은 것을 인간보다 월등하게 잘 함
그래서 재범 확률 예측 같은 것도 잘한다고 생각?



불확실하고 복잡한 상황에서의 판단의 문제 같은 영역에
인공지능의 응용에 대해서 신중하게 접근할 필요가 있지 않을까

감사합니다

주제발표 3

III

AI 자본주의

노영우

매일경제신문 국제부장

발제자 약력

성 명	노 영 우	
소 속	매일경제신문 국제부	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1988 ~ 1992	서울대학교 경제학과	학사
1992 ~ 1992	서울대학교 대학원 경제학과	석사수료
2006 ~ 2013	벤더빌트(Vanderbilt)대학교 대학원 경제학과	박사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2017 ~ 2019	매일경제신문	국제부장
2013 ~ 2018	매일경제신문	기자 (경제부, 국제부)
2006 ~ 2013	벤더빌트 대학교	T.A. 및 R.A.
1996 ~ 2006	매일경제신문	기자 (경제부, 금융부, 증권부)
수상(2016년)	한국기자협회 제47회 올해의 기자상 수상 (기업발 경제위기 시리즈)	
저서(2019년)	AI자본주의(도서출판 해남)	
저서(2016년)	집계경제(도서출판 개미)	
박사학위논문	Essays on Trade Agreements and Export Dynamics	

발제 3

AI 자본주의

노영우

매일경제신문 국제부장

AI 자본주의

2019.9.26

노영우 매일경제 국제부장



2019년 글로벌 경제



- ▶ 'R(recession)의 공포'...세계 경기 불황 진행
 - 미국 중국 독일 일본 성장률 하락
- ▶ 'D(deflation)의 공포'...저물가(디플레이션) 확산
 - 8월 한국 물가상승률 -0.038%
- ▶ 마이너스(Minus) 금리 확산
 - 스위스 UBS 25억원 넘는 예금에 0.75% 수수료
 - 전세계 국채 34%가 마이너스 채권

2019년 글로벌 경제



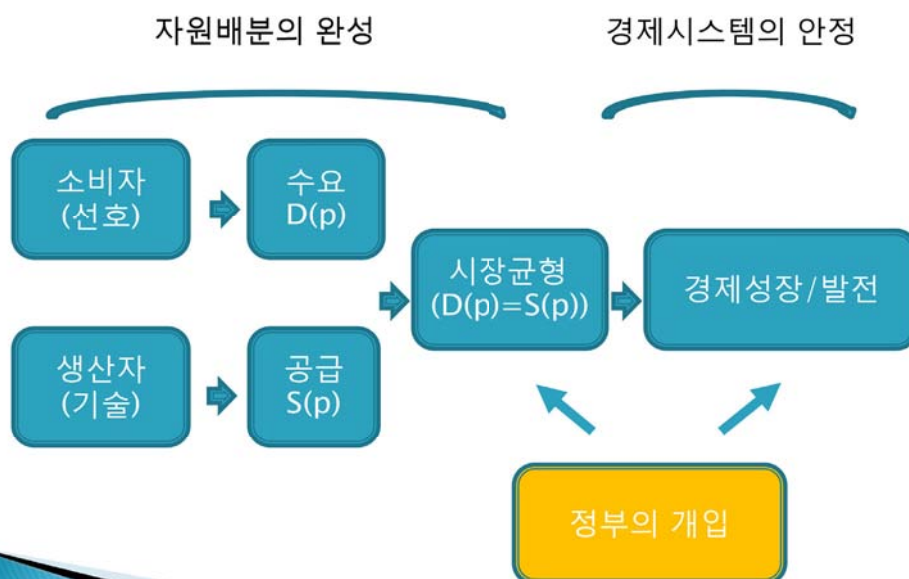
- ▶ 양극화 현상 뚜렷
 - 제이미 다이먼 JP모건 CEO “미국이 둘로 쪼개졌다”
 - 1분위 소득 8% 증가, 5분위 소득 64% 증가 (1970~2014)
- ▶ 사회적 기업의 등장
 - 미 CEO 181명 “주주이익 기업 최우선 목표 아냐”
 - 고객, 직원, 납품업체, 사회 등 이해당사자 고려
- ▶ 정부 정책 무용론
 - 양적완화 정책의 한계, 가상화폐의 등장

자본주의 경제의 특징



- ▶ 개인의 소유와 경제적 자유의 보장
- ▶ 분권화된 의사 결정으로 효율성 담보
- ▶ 정부정책을 통해 시장의 실패 보전
- ▶ 시혜적 성격의 재분배 정책
- ▶ 비약적인 기술/생산력 발전 뒷받침

자본주의 경제시스템



자본주의경제 시장균형가격 의미



- ▶ 자원배분의 완성(필요한 사람에게 자원 귀속)
- ▶ 효율성의 달성(1st Welfare Theorem):
시장경제 경쟁균형은 효율성 담보
- ▶ 이상적 분배 실현(2nd Welfare Theorem):
시장을 통해 이상적 분배 실현 가능

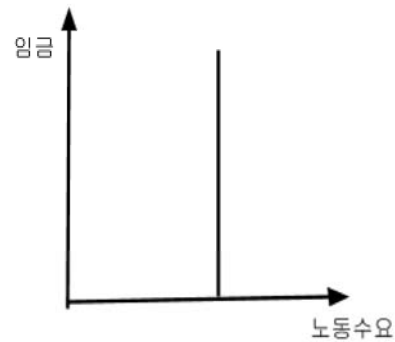
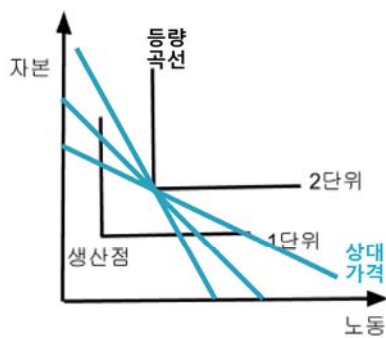
기술진화와 경제시스템의 발전



네모난 경제(Square Economy)



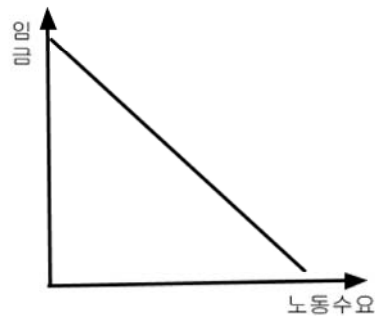
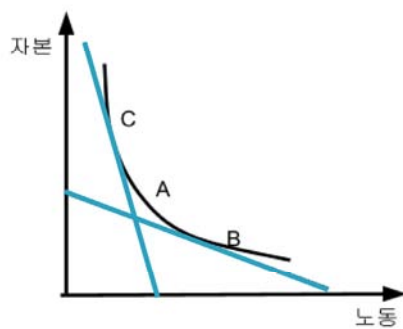
- ▶ 도구와 인간의 결합
 - 노동과 자본이 완벽히 보완적인 관계
 - 임금이 올라도 노동에 대한 수요는 항상 일정



볼록한 경제 (Convex Economy)



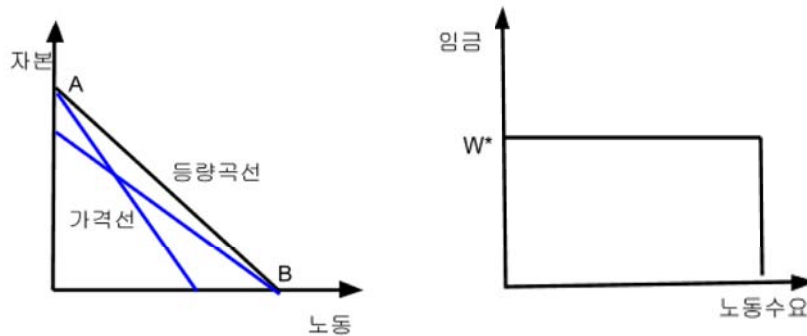
- ▶ 사람을 대체하는 기술의 등장
 - 노동과 자본이 서로 대체관계로 전환
 - 임금과 노동수요는 서로 반비례



평평한 경제(Flat Economy)



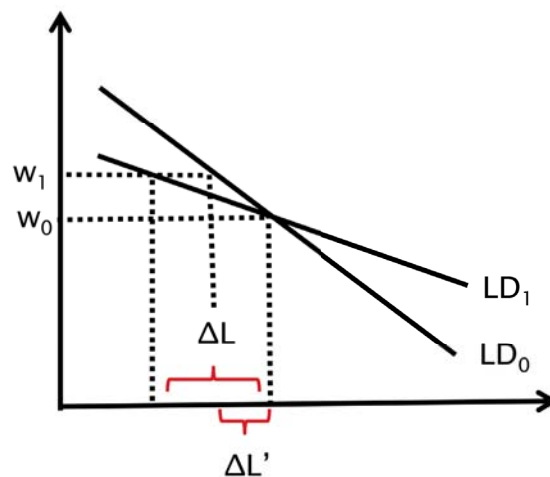
- ▶ 생산현장에서 AI의 확산
 - 노동과 자본의 대체도 급증
 - 자본의 양이 늘어도 한계대체율은 일정
 - 작은 임금 변화에도 노동수요는 극단적으로 변함



경제가 평평해지는 과정

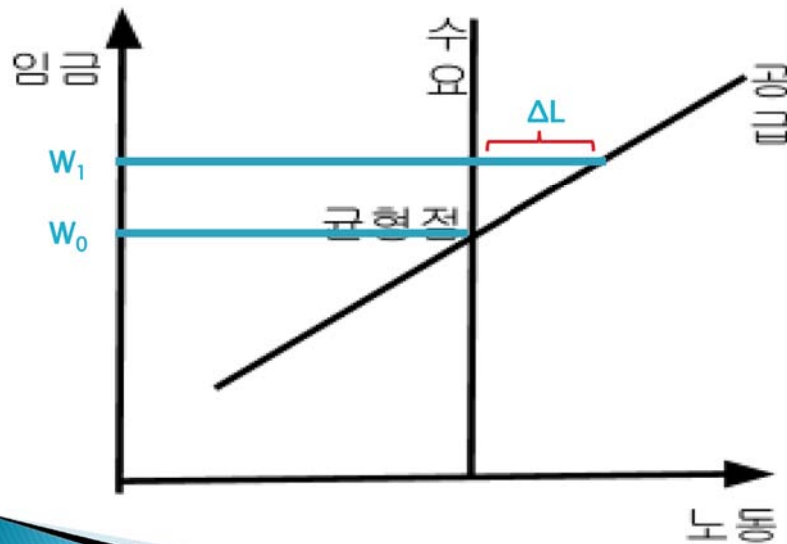


- ▶ 노동수요와 임금
 - 수요곡선이 평평할수록 노동수요는 임금에 민감

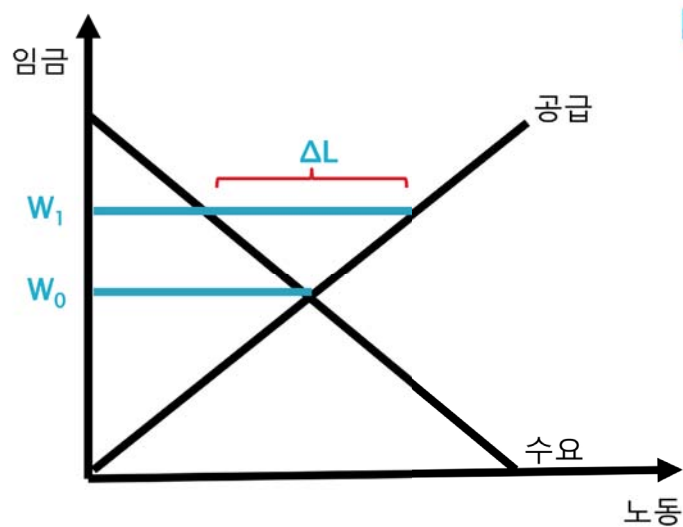


시장에서 벌어지는 현상

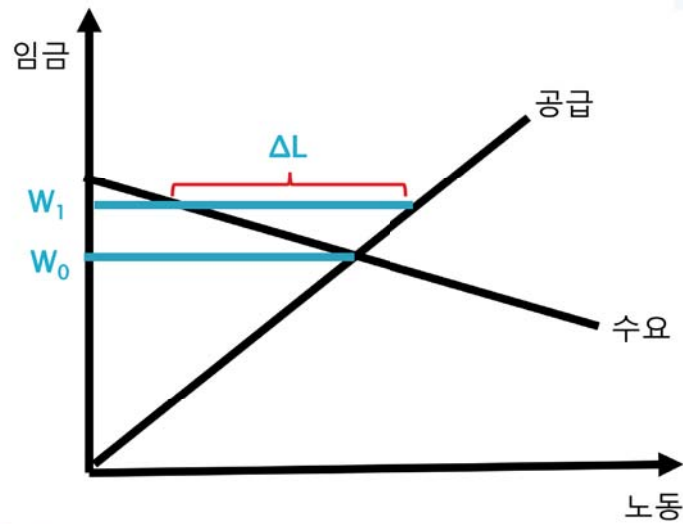
▶ 네모난 경제



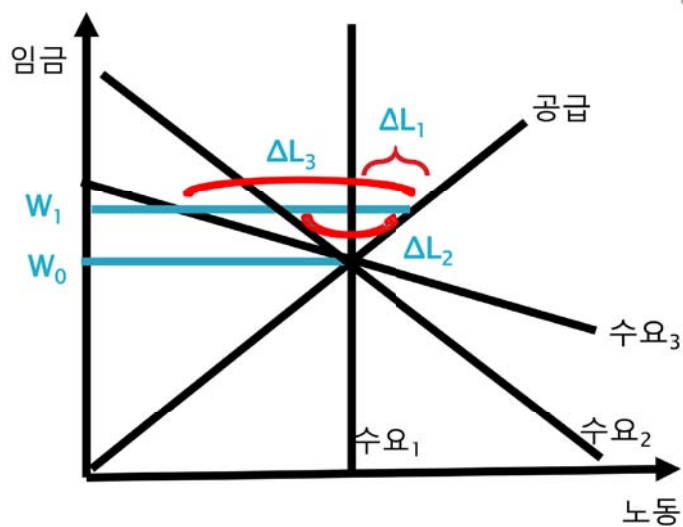
시장에서 벌어지는 현상/불록한 경제



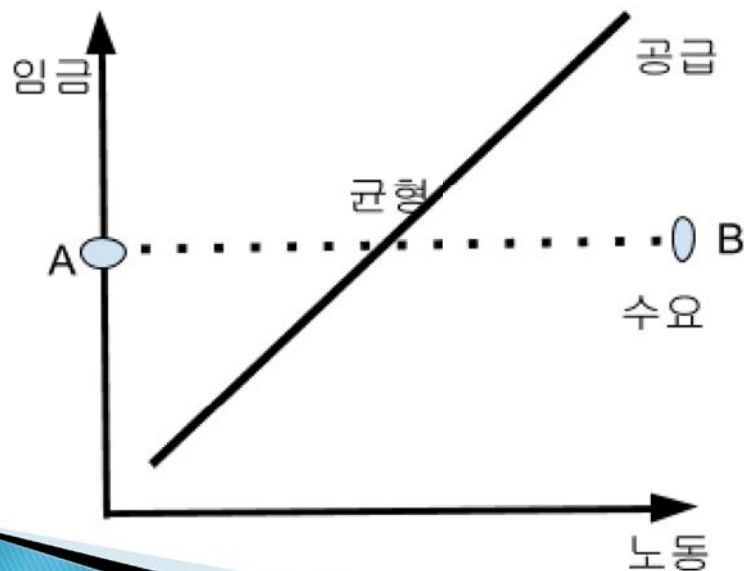
시장에서 벌어지는 현상/평평한 경제



수요곡선이 평평해 질수록 실업 증가



평평한 경제의 가격 파괴현상



평평한 경제와 노동시장

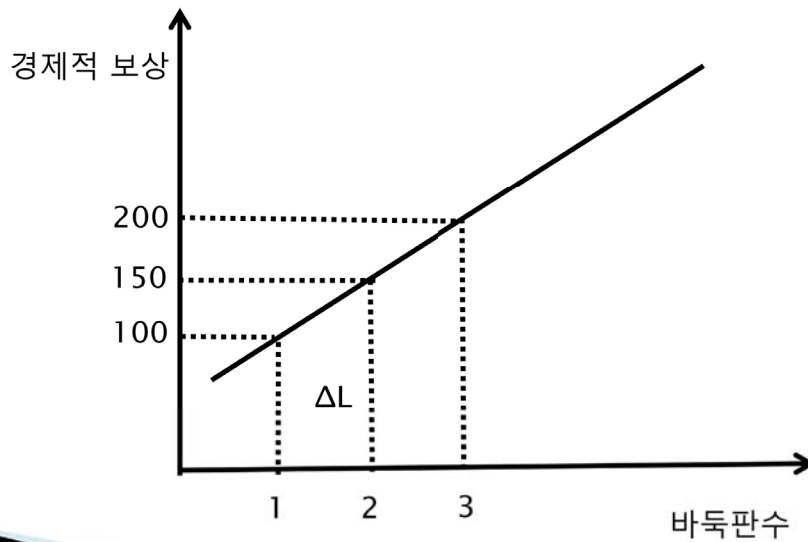


- ▶ 자본과 노동 간의 대체성 증가
- ▶ 노동시장에서 작은 충격에도 급격한 실업 증가
- ▶ 시장에서 균형 임금이 형성되지 않을 가능성도 존재 => 시스템 문제 발생

평평한경제와 생산물시장/ 이세돌과 알파고의 차이



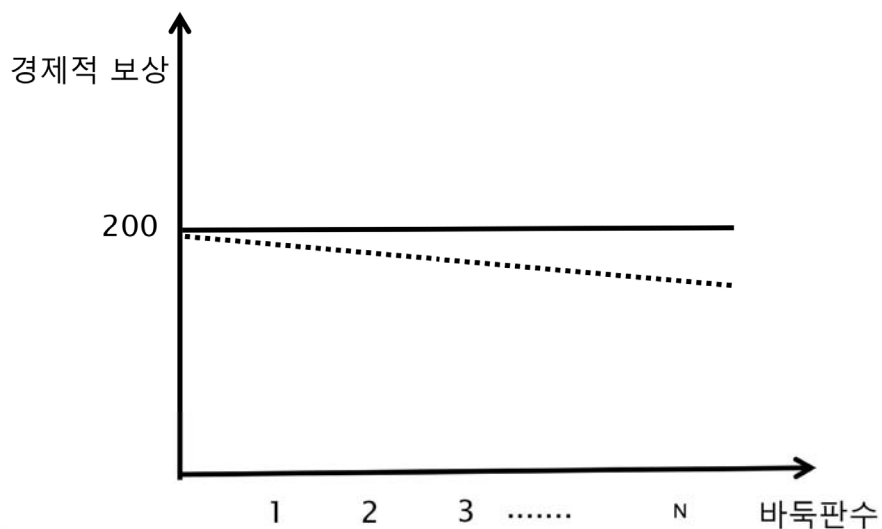
이세돌/바둑회수 늘수록 피로도 점진적으로 증가



이세돌과 알파고의 차이



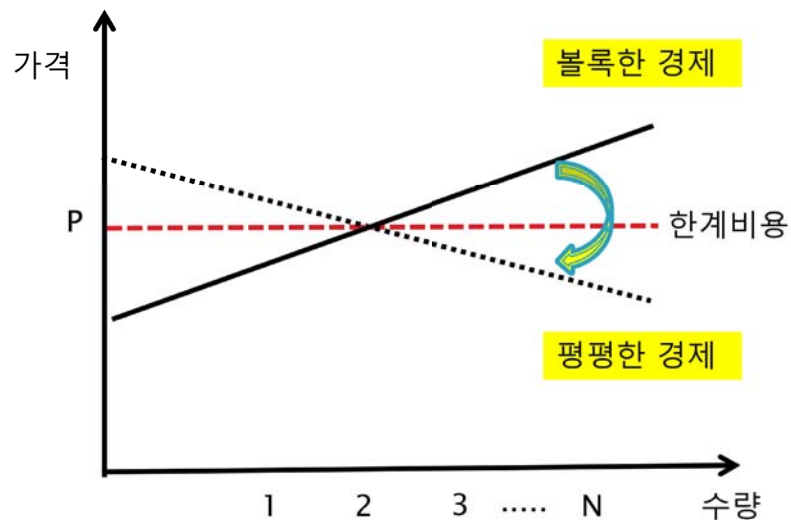
알파고/바둑 회수와 피로도의 상관관계 거의 없음



평평한 경제에서의 생산물공급



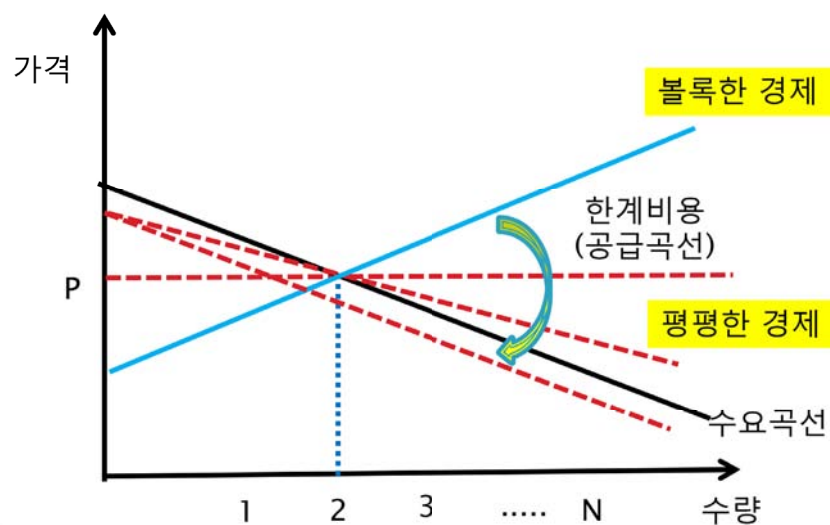
가격과 공급량간의 관계($P=\text{한계비용}$)



평평한 경제의 시장 가격



가격과 공급량간의 관계($P=\text{한계비용}$)
0으로 수렴하는 가격=>자연독점 경향



평평한 경제에서 벌어지는 일

$$\text{marginal cost} = \frac{\Delta \text{total cost}}{\Delta \text{total quantity}}$$

- ▶ 생산요소 시장
 - 극단적인 노동수요의 변동
 - 실업이 광범위하고 일상적으로 발생
- ▶ 생산물 시장
 - 생산성의 비약적인 성장으로 한계비용 감소
 - 시장가격의 지속적 하락, 시장의 독점화 촉진

=> 디플레이션, 실업의 일상화, 시장가격 파괴 경향

가격이 파괴된다면?



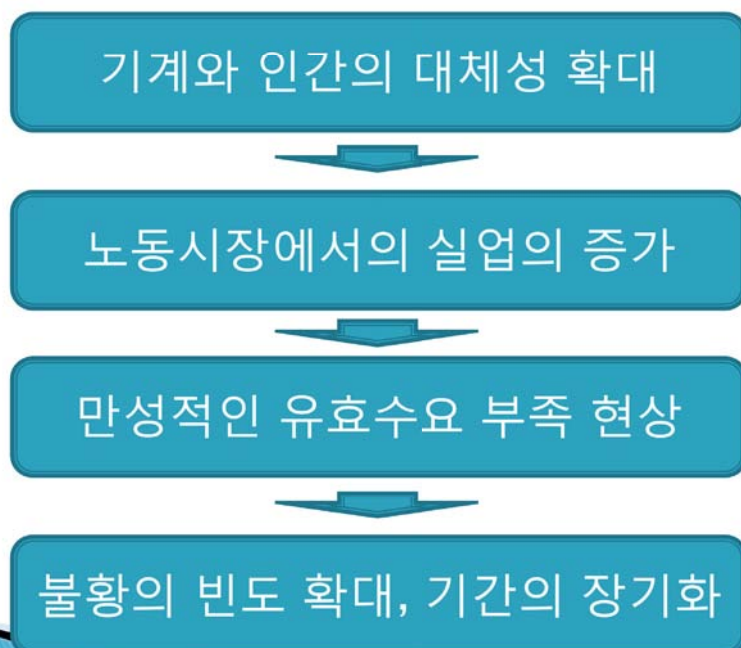
- ▶ 자원 배분의 실패
 - 만성적인 수요 공급 불일치
- ▶ 시장을 통한 자원배분 실현 불가능
- ▶ 경제의 효율성 달성에 실패
- ▶ 분권화 된 의사 결정 방식의 문제점 노출

거시경제환경 변화/자본주의



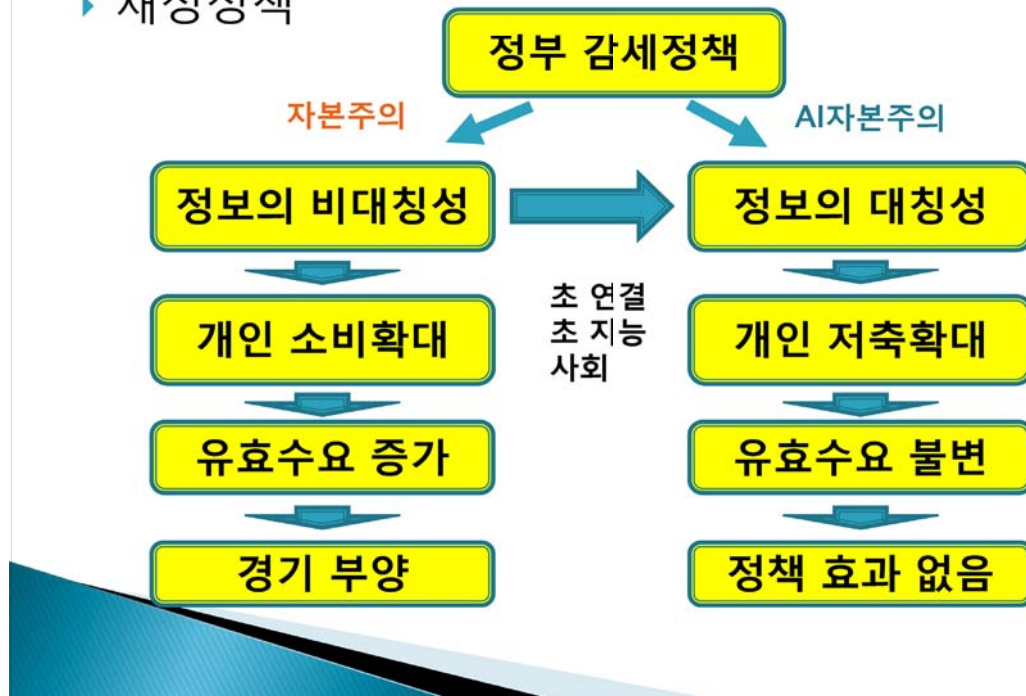
- ▶ 경제주체의 이중성
 - 기업주+소비자, 근로자+소비자
 - 기업주+근로자+소비자
- ▶ 분권화 된 의사 결정으로 효율성 달성
- ▶ 유효수요의 부족은 재정·통화 정책으로 해결
 - 케인즈주의적 경제정책으로 경기 진폭 완화

거시경제구조 변화/AI자본주의



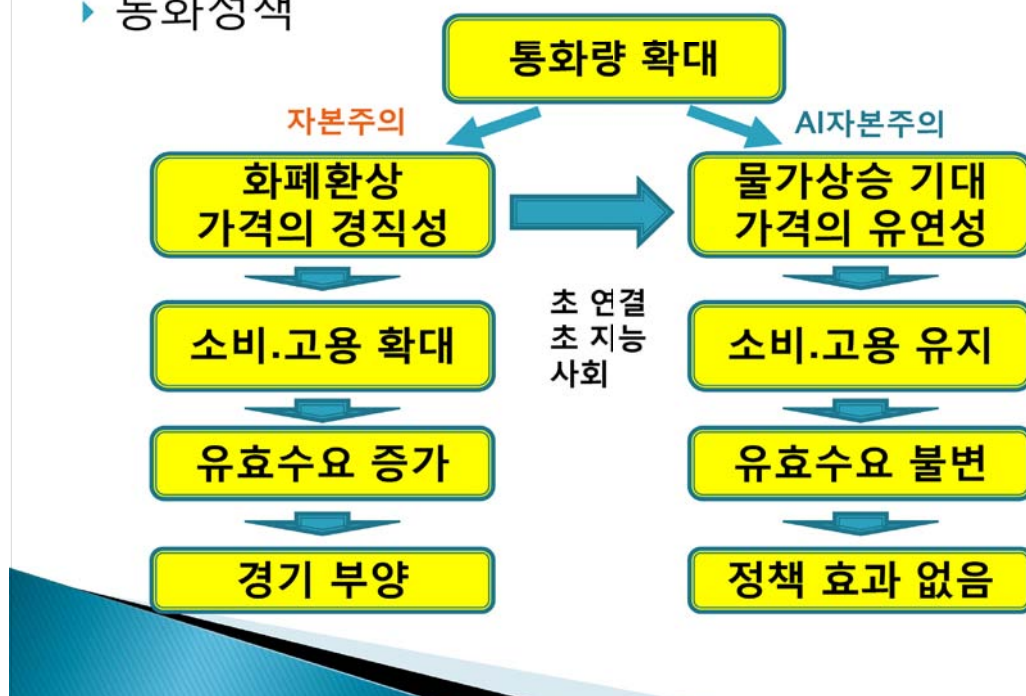
경제정책 무력화/AI자본주의

▶ 재정정책

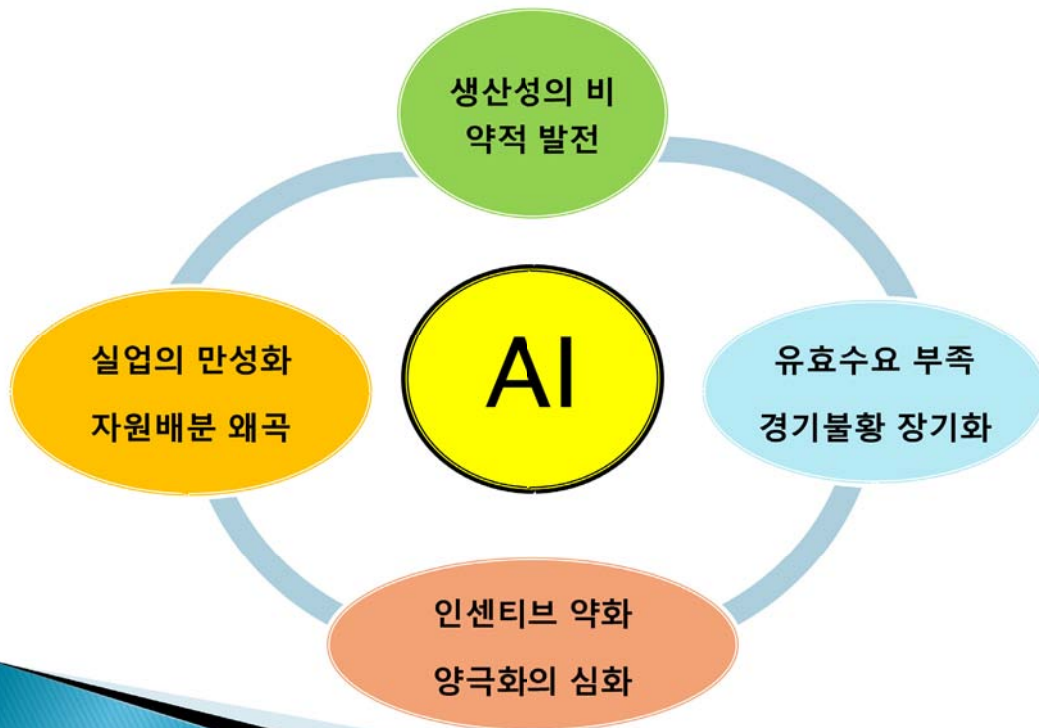


거시경제정책 무력화/AI자본주의

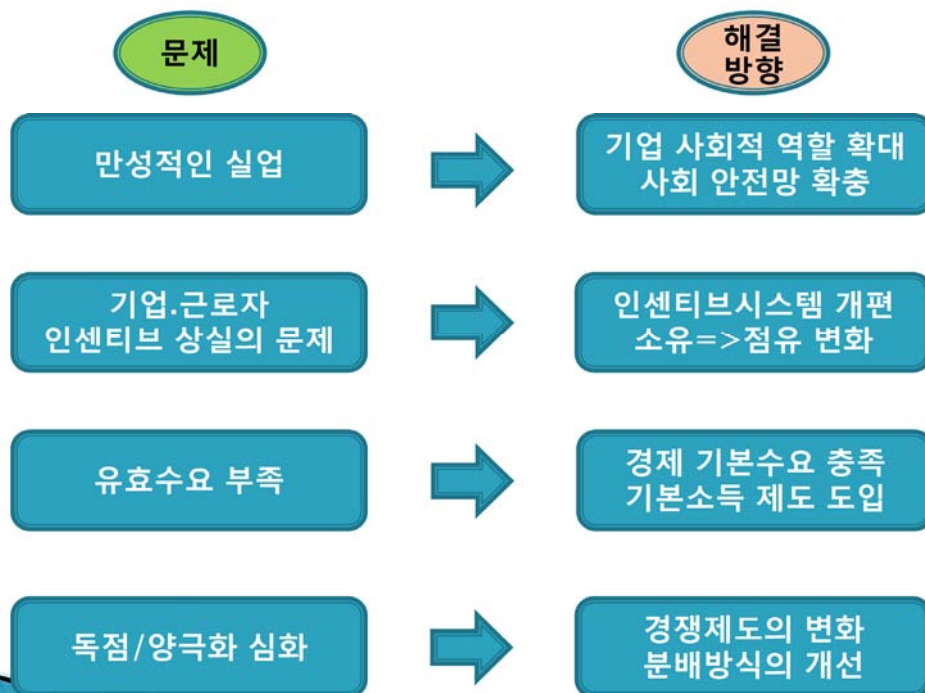
▶ 통화정책



AI자본주의서 벌어질 현상

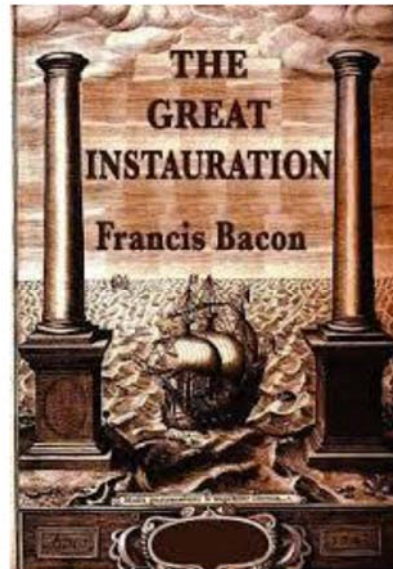


AI자본주의 문제 해결 노력



맺음말

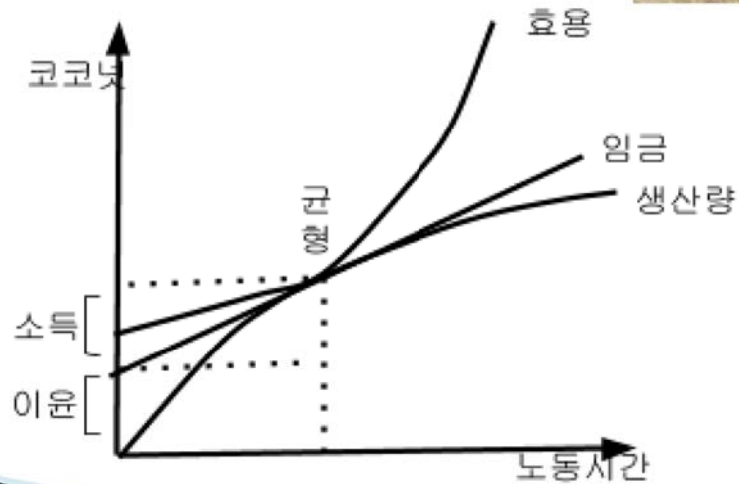
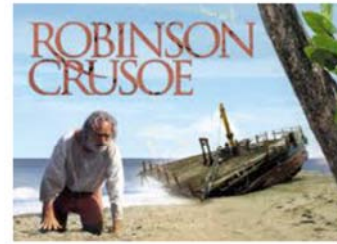
- ▶ 대변혁의 시대
- ▶ AI자본주의는 자본주의와 질적으로 다른 경제 시스템
- ▶ 경제시스템 변화에 대한 고민
- ▶ 경쟁=>공감으로의 인식변화
- ▶ 사회적 대타협의 필요성



감사합니다!

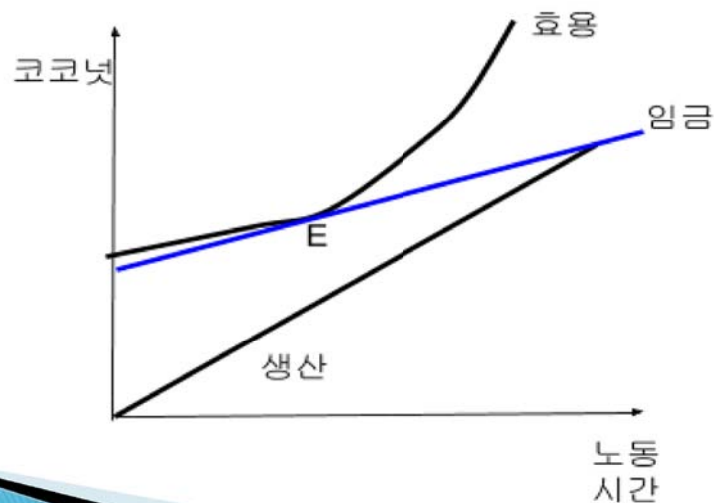
#부록/거시경제 균형

- ▶ 로빈슨 크루소 모델/ 불룩한 경제
- ▶ 분권화된 의사결정의 효율성



#부록/거시경제 균형/AI자본주의

- ▶ 로빈슨 크루소 모델/ 평평한 경제
- ▶ 분권화된 의사결정의 비효율성



IV

지정토론

좌장 약력

성 명	이 태 역	
소 속	KAIST (한국과학기술원)	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1976 ~ 1980	서울대학교	산업공학, 학사
1980 ~ 1982	KAIST	산업공학, 석사
1986 ~ 1991	Ohio State University (Columbus)	산업 및 시스템 공학, 박사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2017 ~	KAIST 교육원 및 교수학습센터	교육원장 겸 교수학습센터장
2017 ~	대한산업공학회	회장
2016 ~	교육부	정책자문위원
2015 ~	한국과학기술한림원 정책학부	정회원
2011 ~ 2015	KAIST Education 3.0 추진단/교수학습혁신센터	추진단장/센터장
2010 ~ 2012	대통령직속 규제개혁위원회	위원
2006 ~ 2007	지식경제부 신성장동력기획단	위원
1999 ~ 2001	KAIST 과학기술도서관/ 정보시스템연구소/ 국가과학기술전자도서관(NDSL)	관장/소장/설립관장
1991 ~ (2006~2013, 2016~2017)	KAIST 산업 및 시스템 공학과	교수, 학과장

토론자 약력

성 명	김 경 만	
소 속	서강대학교 사회학과	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1977 ~ 1981	서강대학교 경제학과	학사
1983 ~ 1985	University of Chicago	사회학 석사
1985 ~ 1989	University of Chicago	사회학 박사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2016.01. ~ 2016.08.	Yale University, Center for Cultural Sociology	Fulbright Fellow
2016 ~ 2018	한국과학기술 한림원	정책학부 부학부장
2011~	Today Social Science (Canada)	편집위원
2010 ~ 2018	『사회과학연구』	편집위원장
2009	경암학술문화재단	경암학술상 (인문/사회부문)
2008	한국사회학회	저술상
1995 ~ 1997	삼성 SDS	자문위원
1989	University of Chicago	Stouffer / Star Research Award

토론문

인공지능과 사회적 이슈

김 경 만

서강대학교 사회학과 교수(한림원 정책학부 정회원)

오늘 김진형 교수, 홍성욱 교수, 그리고 노영우 기자의 AI 에 관한 발표는 AI 에 대한 대략적으로 낙관적인 견해(김진형), AI 의 positive 한 측면을 인정하면서도 그것이 가지고 있는 bias 를 지적하는 견해(홍성욱), 마지막으로 AI 에 의해서 재편되고 있는 새로운 경제의 특성(노영우)을 부각시키고 있음.

토론에서는 AI 에 관한 이 세 story 가 간과했지만, 우리가 반드시 고려해야할 AI 가 초래한, 그리고 초래할 사회적 결과를 논의 할 것임.

첫째는 인간의 동기유발과 결속력(Human motivation and Solidarity)의 문제이고 둘째는 정보의 비대칭과 독점으로부터 야기되는 지배와 권력(Information Asymmetry, Domination and Power)의 문제임.

I. 동기유발과 결속력(Human motivation and Solidarity)의 문제

이미 100년 전에 공학자였던 Frederick Taylor 는 *The Principles of Scientific Management* (1911) 에서 인간의 모든 작업을 최소단위로 분해(decompose) 해서, 측정, 이를 바탕으로 한 보상; 과학적 원리를 통해서 구명함으로써 생산성을 극대화할 수 있다고 주장 (유명한 time and motion studies: manless management 를 추구)

마치 인간이 없어도 생산성을 늘리고 benefit이 증가할 수 있다는 주장

그러나 사람들을 “사람으로” (treat people as human beings) 대할 때 생산성이 증진될 수 있다는 반론이 등장. 즉 Human Relations Movement.

이미 1936년에 Chester Barnard. *The functions of the Executive* ; 조직의 사회심리학=이해와 결속력, 연대감이 조직의 성공의 Key.

1920년대 Hawthorne study 에서부터 최근의 Bank of America 연구; 가장 작업능

를과 생산성이 높은 팀은 informal interaction 을 통한 solidarity 증진을 한 팀.

물화와 지배 우리가 만든 것에 의해서 거꾸로 지배받는 상황. 사람들의 원자화 (atomization of individuals) 사람들이 디지털화 되면 될수록 덜 자율적이 됨.

기술진보가 가져 올 결과에 회의적이면 늙고 시대에 뒤떨어진 사람? 기술권력을 가진 논리의 희생물?

II. 정보비대칭과 독점으로 인한 권력과 지배(Information Asymmetry and the Problem of Domination)

두 가지의 예

2.1 Uber platform; 택시 수요가 없을 때도 기사들을 도로위에 있게 하는 것이 목표: 왜? 그래야 빨리 보낼 수 있으니까. 이를 위해서 behavioral economics 의 연구결과를 이용하고 실험

Information Asymmetry: Nudge 를 통해서 기사들을 manipulation; 즉 회사입장에선 도로에 항상 차들이 많을수록 좋다; 그래서 자신의 임금 목표에 가까워서 그만 하고 싶은 기사들로 하여금 마치 도로에 돈 벌 기회가 더 많을 것 같다는 정보를 흘리고==실제론 별로 없는데== 넌지시 계속 도로위에 있기를 권유(nudge); 이를 통해서 기사들을 조종; 기사들을 도로위에 계속 돌아다니게 하려고 회사는 기사들의 성향을 조사하고 그것을 이용: 또 연속으로 다음 guest를 자동으로 연결시키는 algorithm을 concoct. "platform 이 neutral 하지 않고 회사가 exploit 하도록 만들어져있음.

2.2 Political Marketing

Big data 는 "digital traces" of your choices and action을 의미.

Michal Kosinski -psychometrics and questionnaire ; five personality traits; openness, how considerate; extrovert? etc. 몇 번의 facebook click 좋아요! 인가를 가지고 90퍼센트 예측, 예) 68번이면 sexual orientation (88%), 피부색 (95%), 민주/공화당 지지(85%); 어떤 화장품--gay 성향, 어떤 종교가지고 있나 거의 완벽하게 예측

이를 정치적으로 이용: Cambridge Analytica 라는 회사. Trump campaign 의 주역; political marketing; influence politics; platform 운영자의 power.

Monopoly of information and exercise of power.

토론자 약력

성 명	임 영 익	
소 속	現 인텔리콘법률사무소 대표변호사 및 (주)인텔리콘 메타연구소 대표이사	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
1989 ~ 1998	서울대학교 생명과학	학사
2010 ~ 2011	사법연수원	41기 수료
2012	특허연수원	9기 수료
2014	미국 버클리 로스쿨 IP 과정	IP/IPSL과정 수료
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2019 ~ 현재	대통령직속 4차산업혁명위원회	혁신위원
2018 ~ 현재	대한민국 국회도서관	자문위원
2016 ~ 현재	한국인공지능법률학회	부회장
2016 ~ 현재	세계 법률인공지능학회	정회원
2015 ~ 2017	대한변호사협회	제48대 부협회장
2015 ~ 현재	건국대학교 언론홍보대학원	겸임교수
2013 ~ 현재	(주)인텔리콘 메타연구소	대표이사
2012 ~ 현재	인텔리콘 법률사무소	대표변호사

토론문

법률인공지능과 리걸테크

임 영 익

인텔리온 법률사무소 대표

1. 법률인공지능과 리걸테크 현황

법률과 인공지능의 융합인 법률인공지능은 학술적으로 컴퓨테이션 법률학(computational law) 혹은 법률정보학(Legal Informatics)라 부른다. 이 분야의 역사는 60년이 넘고 인공지능 자체의 역사와 함께 흥망성쇠를 반복했다. 한편, 컴퓨터를 이용해 법률 서비스를 자동화 하는 일체의 산업 분야를 ‘리걸테크(Legal Tech)’라고 하며 최근에 미국을 중심으로 급성장을 하고 있다. 우리나라는 아직 초기 단계지만 법률인공지능과 리걸테크 개념의 확산으로 국내의 법률시장의 모습도 변화의 조짐이 있다. 먼저 리걸테크 산업 현황을 빨리 파악하기 위해서는 톰슨앤로이트스의 보고서¹⁾가 도움이 된다. 이 보고서는 리걸테크 산업을 변호사 중개 서비스, 법률 문서 자동 작성 서비스, 법률 사무소 운영 솔루션, 법률 자료 검색 서비스, 법률 교육 서비스, 온라인 분쟁 해결 서비스, 전자증거 개시, 법률 트렌드/문서 분석, 산업별 규제법 대응 서비스로 나눈다.



1) 이런 분류 체 Technology의 분류법을 기초로 하였다. 계는 스탠포드 로스쿨의 Discover Legal

한편, CB Insight는 리걸테크 생태계를 종합적인 온라인 법률서비스 중요한 카테고리 로 보면서 전자 증거 개시, 실습관리 소프트웨어, 지적재산 및 상표 소프트웨어 서비스, 인공지능 법률기술, 소송 금융, 변호사 검색, 법률 검색, 공증도구 등으로 분류한다. 위 두 분류 방식을 참고해서 국내 리걸테크 기업과 비교 분류하면 다음과 같다.

서비스 분류	주요 서비스	서비스 설명	해외 업체	국내 업체
Process Automation	로펌 관리	<ul style="list-style-type: none"> 로펌의 의뢰인, 결제 등의 관리/회계 사항부터 소송관련 자료 정리/공유에 이르는 로펌 관리 솔루션 법률 서비스에 특화된 사무 관리 자동화 시스템 	<ul style="list-style-type: none"> Clio Anaqua 	<ul style="list-style-type: none"> 로매니저 로시오피스
Legal Research	법률 자료 검색	<ul style="list-style-type: none"> 리걸 테크 산업에서 가장 중요한 서비스 영역 전통적으로 Lexisnexis 와 Westlaw 두 회사가 시장을 장악 최근에 스타트업들이 급성장 인공지능, 빅데이터 기술 등 적용 	<ul style="list-style-type: none"> Westlaw Lexisnexis Casetext Judicata Ross Intelligence 	<ul style="list-style-type: none"> 인텔리콘 로앤비(기존)
E-discovery	디스커버리 분석 분야	<ul style="list-style-type: none"> 소송 당사자가 상대방이나 제3자로부터 소송과 관련된 모든 전자기록을 수집할 수 있도록 허용하는 제도 기업간 분쟁 시 이메일, 텍스트, 비디오 등의 방대한 자료 분석에 대한 서비스 E-discovery 제도는 아직 국내에서는 시행되고 있지는 않으나 미국과의 기업분쟁이 필요 	<ul style="list-style-type: none"> Disco BlackStone Discovery Logikcull 	<ul style="list-style-type: none"> 없음
Legal analysis	법률자문 및 전략 수립	<ul style="list-style-type: none"> 통계학이나 빅데이터 기술 등을 활용해 판례 추이 분석이나 법률 변동 예측을 하는 서비스 	<ul style="list-style-type: none"> Lex Machina FiscalNote 	<ul style="list-style-type: none"> 제법아는 언니 헬프미
Consumer	변호사 소개 서비스	<ul style="list-style-type: none"> 변호사와 의뢰인의 효과적인 커뮤니케이션 채널을 제공하고 변호사를 소개 해 주는 서비스 	<ul style="list-style-type: none"> Lexoo 	<ul style="list-style-type: none"> 로톡 변호사닷컴 온서컴

출처: "Four Areas of Legal Ripe for Disruption by Smart Startups", 2014. 12. 16, Law Technology Today 와 "리걸테크(Legaltech) 산업 현황과 시사점", 2016. 10. 12, 현대경제연구원 자료 참고 및 보완

위 자료처럼 미국은 법률의 모든 분야에서 인공지능이 응용되거나 리걸테크 산업이 자리 잡고 있다. 국내에서는 변호사소개, 계약서 자동완성 등의 서비스 회사가 주를 이루며 인공지능 기술을 가지고 있는 기업도 있다. 지금은 리걸테크 산업이 초기 단계이지만 규제와 데이터 문제가 해결 된다면 좀 더 많이 회사가 등장할 수 있다.

2. 규제와 데이터 문제

가. 변호사법 위반 문제

우리나라는 리걸테크 산업의 발생 초기부터 변호사법 위반 문제가 대두 되었다. 변호사 소개서비스 뿐 아니라 어떤 형태의 인공지능 서비스도 이 문제를 피해 갈 수 없다. 애플의 'siri'와 같이 시민들에게 직접 응답하는 형태를 가진 '인공지능 변호사'가 유료화 되는 경우에도 곧 바로 변호사법 위반 문제가 나타난다.

미국의 경우 초기 리걸테크 산업은 온라인 로펌의 형태로 출발 하였다. 대표적으로 미국의 '리걸쥬'이라는 회사가 최초로 성공한 온라인 로펌이다. 2001년에 서비스를 시작한 리걸쥬는 온라인 상으로 다양한 법률서비스를 제공했다. 리걸쥬이 예상 밖의 성장을

하자 기존 변호사 업계는 리걸짐을 변호사법 위반으로 고발 하였다²⁾. 그러나 법원은 ‘리걸짐의 서비스는 사람이 기술을 활용해 더 나은 법률 서비스를 제공하는 하는 사업’이라고 인정해 리걸짐의 손을 들어줬다. 이후에도 유사한 업체들이 계속 등장하였고, 위기를 느낀 기성 변호사들이 변호사법 위반으로 고소고발을 하였지만 미국의 재판부는 기술이 사람을 이롭게 한다면 그것을 규제할 필요가 없다는 이유로 기각하였다.

국내의 경우 몇 년 전에 변호사 소개 회사나 온라인 형태의 법률 서비스 회사들이 변호사법 위반으로 고발을 당하였으나 모두 무혐의 처분을 받았다. 그러나 최근에 대한 변호사협회는 리걸테크 회사의 법률 서비스에 대하여 “어떤 형태이건 비용이나 수수료가 전제된 서비스는 위법하다”는 의견을 개진하였다.

변협에서 판단한 위법한 서비스의 유형은 다음과 같다.

변호사가 법률상담 사이트에 일정한 회비나 가입비를 지급해야만 상담변호사로 등재가 능 한 경우나 법률상담 사이트가 이용자로부터 받은 법률상담 쿠폰 대금 중 일부를 공제 하고 나머지만을 실제 상담한 변호사에게 지급하는 경우는 변호사법 제34조에서 금하 는 유료 법률사무 알선행위에 해당한다.

법률상담 사이트가 특정 업체와 제휴해 일정 조건을 만족하는 경우 제휴업체 이용자들 에게 무료 쿠폰을 배부하는 데 대해서는 변호사업무광고규정 위반한다. 제휴업체가 그 회원들에게서 대가를 받지 않고 무료로 법률상담 쿠폰을 배부하는 것이라 하더라도 위 쿠폰을 받고 상담에 응하는 변호사는 제휴업체 사이트의 영리를 위한 사업의 일환으로 법률상담에 참여하는 경우에 해당한다³⁾.

이런 입장에서 보면 국내의 대부분의 리걸테크 산업은 위법한 것이 되므로 발전 가능성 을 논할 필요가 없다.

나. 데이터 문제

지금의 인공지능은 머신러닝이나 딥러닝을 바탕으로 하는 학습기반이 주를 이룬다. 대 표적으로 알파고의 알고리즘이 그러한 형태이며, 자연어처리분야도 학습기반 인공지능 기술을 선호 한다. 법률분야의 인공지능도 딥러닝 기술을 적용하는 쪽으로 이동하고 있어 인공지능 학습을 위해 방대한 법률 자료가 필요하다. 그러나 현재 대법원은 법률데

2) 미국 변호사들은 리걸짐이 과학기술을 이용하여 고객에게 법률 자문을 제공하는 것은 권한을 부여받지 않은 주체의 법 집행(Unauthorized Practice of Law)’이라고 주장하였다

3) 변호사업무광고규정 제8항 제2항 제3호에서는 제3자의 영리를 위한 사업의 일환으로 운영되는 법률상담에 참여하는 경우를 금지하고 있다.

이터 중에서 대법원 판례만 공개 있으며 하급심 판례는 극히 일부만 공개하고 있는 실정이다(전체 데이터의 0.2% 수준). 따라서 딥러닝을 법률에 적용하는 기술은 국내에서 큰 발전을 이루기 힘들다. 물론 개개의 하급심 판례는 개인적으로 신청을 하여 얻을 수는 있지만 데이터의 관점에서 보면 큰 도움이 안 된다. 하급심 판례를 전부 공개하는 것은 개인정보보호 등의 문제를 극복하기 힘들어 보인다. 따라서 이런 이슈와 관련이 없는 일부 하급심 판례라도(1% 정도라도) 시급히 공개를 하는 쪽으로 방향을 잡아야 할 것이다.

3. 인공지능 등장과 법률 생태계

최근 영국에서 만든 미래 보고서는 인공지능에 의해 대체될 직업으로 변호사가 언급된 적이 있다. 영미법 계통은 법률자료의 검색, 분류, 분석 등이 비중을 많이 차지하기 때문에 소위 법률보조원(legal secretary) 업무가 이런 쪽으로 쏠려 있다. 따라서 이런 업무를 잘 할 수 있는 인공지능의 등장이 직업 대체의 효과를 가져올 수 있다. 그러나 이 보고서에서 말한 법률보조원은 우리나라 변호사와는 전혀 다른 직업이므로 인공지능 기술이 변호사나 판사 자체를 대체할 가능성은 거의 없다.

기계는 인간 변호사의 재판 수행을 대신할 수 없거니와 복잡한 스토리를 이해하는 것이 어렵다. 인공지능은 기본적으로 자료 검색 및 분석을 도와주는 보조 역할을 할 뿐이라 변호사를 대체하는 등의 파괴적 혁신(disruptive innovation)은 국내에서는 당분간은 어려울 것이다. 오히려 인공지능과 협업하는 형태의 모습이 더 유력하다.

4. 결론

인공지능이 법률자료의 검색을 도와 주거나 숨어 있는 함정을 찾아 준다면 경험이 약한 신규 변호사도 복잡하게 얽혀 있는 판례와 법률의 세계를 빠른 시간 내에 파악할 수 있어서, 데이터 부족의 불리함을 어느 정도 극복 할 수 있다. 시민들의 입장에서 볼 때, 인공지능이 간단한 법률 조언 해 준다면 법친화적 생활 문화도 가능하다. 즉, 시민들이 큰 사건이나 실제 송사에 휘말린 경우가 아니라면 인공지능 서비스를 자주 이용할 것이다. 이런 점에서 개인 변호사(Solo Practitioner)들도 인공지능을 이용해 숨어 있는 법률 시장을 발굴할 수 있다.

영국의 유명한 법률학자인 리처드 서스킨드(Richard Susskind)는 가까운 미래에 인공지능이 변호사를 대체하며 법률 시장을 축소시켜 법률가의 미래는 밝지 않다는 전망을 하였다. 그러나 서스킨드는 법률뿐 아니라 의료, 회계, 세무 등의 대부분 지식전문가들

이 겪어야 할 변화라고 진단하면서 ICT 기술융합은 청년들에게는 새로운 기회를 주기 때문에 반드시 불리한 것은 아니라고 주장했다.

미국의 스탠포드 로스쿨은 공학자들과 함께 법률, 수학, 통계학, 인공지능 등을 융합하는 학제적 교육 프로그램을 만들어 법률 인공지능 분야를 주도 하고 있다. 이 프로그램을 통하여 배출된 변호사나 학생들은 인공지능을 적용한 법률 시스템을 개발하는 등 새로운 리걸테크 분야를 계속 생성하고 있습니다. 이런 시도는 긍정적 피드백을 만들고 있고 많은 청년변호사들이나 혁신가들은 리걸테크 개념을 이해하면서 온라인 로펌 같은 O2O(Online to Offline) 생태계도 만들어 가고 있다.

법률인공지능이나 리걸테크 분야는 아직 생소하고 발전 가능성이 불투명 하다. 그러나 청년 변호사들이 이 분야에 눈을 뜨고 창의적인 시스템을 만들어 낸다면 새로운 법률 서비스 패러다임이 만들어 질 것이고, 시민들도 더 유익한 법률 서비스를 받을 수 있을 것이다.

결론적으로 법률과 인공지능의 융합은 세계적 추세이며, 청년변호사들과 시민들을 이롭게 하는 측면이 있으므로 각종 규제와 데이터 개방에 대해서 미래 지향적 검토가 필요할 것이다. 끝.

토론자 약력

성 명	김 건 우	
소 속	카카오모빌리티 디지털경제연구소	
1. 학 력		
기 간	학 교 명	전 공 및 학 위
2003 ~ 2009	포항공과대학교 산업경영공학과	학사
2009 ~ 2010	KDI국제정책대학원 정책학과	석사
2. 주 요 경 력		
기 간	기 관 명	직위, 직책
2019 ~ 현재	카카오모빌리티	데이터 이코노미스트
2018 ~ 현재	LAB2050	연구자문위원
2016 ~ 현재	서울디지털재단	자문위원
2011 ~ 2019	LG경제연구원	선임연구원
2010	KDI	초빙연구원

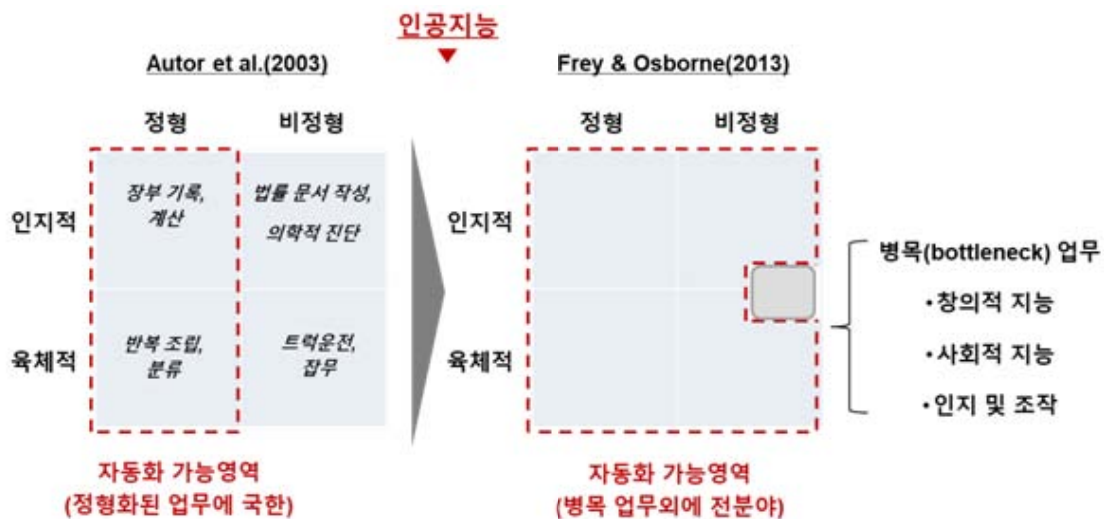
토론문

김 건 우

카카오모빌리티 데이터 이코노미스트

1. 인공지능의 부상과 자동화의 확장

- 과거 자동화는 정형화된 업무에 국한된다고 보았으나, 인공지능은 비정형화된 업무도 대체 가능
 - 머신러닝의 발전은 암묵지(tacit knowledge)에 기반한 일자리도 기계가 할 수 있는 가능성을 제기... 비정형 업무(nonroutine task)
 - 과거 형식지(explicit knowledge)를 코딩하는 방식의 규칙 중심의 인공지능은 정형 업무(routine task)의 자동화에 활용



자료: Autor, Levy, Murnane(2003), Frey & Osborne(2013), LG경제연구원

〈참고〉정형편향적 기술진보 vs. 숙련편향적 기술진보

- 2003년 MIT대학의 노동경제학자인 데이비드 오토(David Autor) 등이 자동화, 특히, 컴퓨터에 기반한 자동화에 따른 노동시장 영향의 분석들을 제시

- 정형편향적 기술진보(RBTC: Routine-Biased Technological Change) 가설을 이용하여 기술 발전이 노동시장(일자리, 임금 등)에 미치는 영향을 설명
- 핵심은 반복적인 업무(routine task)를 대체하는 방향으로 기술변화가 진행해 왔다는 것... 미리 코딩된 절차를 반복 수행하는 컴퓨터가 핵심적 역할, 반복적인 업무나 일정한 규칙을 따르는 업무(task)가 많은 직업(job)일수록 자동화 가능성이 크다는 것
- 반복적 업무는 주로 중간층 직업이 수행하고 있기 때문에 기술변화는 일자리 양극화를 초래
- 이들 이전까지 고숙련, 고학력자가 상대적으로 임금도 높고, 실업률도 낮은 현상을 숙련편향적 기술진보(SBTC: Skill-Biased Technological Change) 가설로 설명하였으나, 노동시장의 양극화(중숙련자 감소) 현상을 설명하는데 한계.

2. 인공지능이 우리나라 노동시장에 미치는 영향

- 미국 직업 기준으로 도출한 직업별 대체확률을 우리나라의 직업분류코드에 매칭시켜 우리나라 일자리의 대체확률을 구한 다음, 최신 고용데이터를 이용하여 자동화 위험에 노출된 일자리의 분포와 특성을 분석
- 인공지능으로 인해서 향후 대체될 수 있는 일자리 현황을 살펴본 결과 일부 직업, 산업, 계층에 위험이 몰려있는 것으로 나타났음.
 - 우리나라 노동시장 일자리의 43%가 자동화 고위험군
 - 사무·판매·기계조작 종사자 등 3대 직업이 고위험 일자리 72% 차지
 - 도소매업, 숙박음식점업, 제조업 등 3대 산업이 고위험 일자리 63% 차지

3. 미래 일자리 변화의 의미와 시사점

- 인공지능은 범용기술(GPT)··· 증기기관, 전기와 같이 산업 전반에 도입되어 과거에 불가능했던 새로운 경제적 기회를 창출할 것··· 개인들과 기업들은 도입 불가피
- 인공지능은 서비스업의 자동화를 가속화할 것··· 농업->제조업->서비스업->?
 - 비교역제로 인식되었던 서비스가 인터넷을 통해서 원격으로 공급되면서 일자리를 대체할 경우 '디지털 무역'이 통상 갈등의 요인으로 급부상할 가능성도 존재
- 인공지능에 의한 자동화 위험이 특정 직업이나 산업, 계층에 집중··· 새로운 산업들이 탄생하여 양질의 일자리들이 창출될 경우에도 실업, 양극화 문제가부각

- 정부는 인공지능에 기반한 생산성 향상과 일자리 감소 사이에 극심한 딜레마에 직면할 것
 - 일자리를 지키기 위해서 자동화를 배척하거나, 반대로 극단적으로 기술 혁신을 장려할 수는 없음.
- 혁신 친화적 사회안전망 구축
 - 유연안정성(flexicurity)을 갖춘 노동시장...산업의 변화에 대응하여 다양한 고용형태와 탄력적인 인력운용이 가능한 유연한 노동시장을 마련함과 동시에 취약계층의 일자리 충격을 흡수할 수 있도록 재교육, 전직 지원, 고용 보험 등 사회안전망 강화에 적극적 자원 투입... 기본소득, 로봇세 등 검토 및 정책 실험

한림원탁토론회는...

한림원탁토론회는 국가 과학기술의 장기적인 비전과 발전전략을 세우고, 동시에 과학기술 현안 문제에 대한 해결방안을 모색하기 위한 목적으로 개최되고 있는 한림원의 대표적인 정책토론 행사입니다.

지난 1996년 처음 개최된 이래 지금까지 100여회에 걸쳐 초중등 과학교육, 문·이과 통합문제, 국가발전에 미치는 기초과학 등 과학기술분야의 기본문제는 물론 정부출연연구소의 발전방안, 광우병의 진실, 방사능, 안전 방제 등 국민생활에 직접 영향을 미치는 문제에 이르기까지 광범위한 주제를 다루고 있습니다.

한림원은 과학기술 선진화에 걸림돌이 되는 각종 현안문제 중 중요도와 시급성에 따라 주제를 선정하고, 과학기술 유관기관의 최고책임자들을 발제자로 초빙하여, 한림원 석학들을 비롯해 산·학·연·정의 전문가들이 심도 깊게 토론을 진행하고 있습니다.

토론결과는 책자로 발간, 정부, 국회와 관련기관에 배포함으로써 정책 개선방안을 제시하고 정책 입안자료를 제공하여 여론 형성에 기여하도록 힘쓰고 있습니다.

《 한림원탁토론회 개최실적 (1996년 ~ 2018년) 》

회수	일 자	주 제	발제자
1	1996. 2. 22.	초중등 과학교육의 문제점 초중등 과학교육의 문제점	박승재
2	1996. 3. 20.	과학기술분야 고급인력의 수급문제	서정헌
3	1996. 4. 30.	산업계의 연구개발 걸림돌은 무엇인가?	임효빈
4	1996. 5. 28.	과학기술 행정과 제도, 무엇이 문제인가?	박우희
5	1996. 7. 9.	연구개발 평가제도, 무엇이 문제인가?	강계원

회수	일 자	주 제	발제자
6	1996. 10. 1.	정부출연연구소의 역할과 기능에 대하여	김훈철
7	1996. 11. 4.	21세기 과학기술비전의 실현과 정치권의 역할	김인수
8	1997. 2. 25.	Made in Korea, 무엇이 문제인가?	채영복
9	1997. 4. 2.	산업기술정책, 무엇이 문제인가?	이진주
10	1997. 6. 13.	대학교육, 무엇이 문제인가?	장수영
11	1997. 7. 22.	대학원 과학기술교육, 무엇이 문제인가?	김정욱
12	1997. 10. 7.	과학기술 행정체제, 무엇이 문제인가?	김광웅
13	1998. 1. 22.	IMF, 경제위기 과학기술로 극복한다.	채영복
14	1998. 3. 13.	벤처기업의 활성화 방안	김호기, 김영대, 이인규, 박금일
15	1998. 5. 29.	국민의 정부의 과학기술정책	강창희
16	1998. 6. 26.	정보화시대의 미래와 전망	배순훈
17	1998. 9. 25.	과학기술정책과 평가제도의 문제	박익수
18	1998. 10. 28.	경제발전 원동력으로서의 과학기술의 역할	김상하
19	1999. 2. 12.	21세기 농정개혁의 방향과 정책과제	김성훈
20	1999. 3. 26.	지식기반 경제로의 이행을 위한 경제정책 방향	이규성
21	1999. 5. 28.	과학기술의 새천년	서정욱
22	1999. 9. 10.	신 해양시대의 해양수산정책 발전방향	정상천
23	2000. 2. 10.	21세기 환경기술발전 정책방향	김명자
24	2000. 4. 14.	경제발전을 위한 대기업과 벤처기업의 역할	김각중

회수	일 자	주 제	발제자
25	2000. 6. 16.	과학·기술발전 장기 비전	임 관
26	2000. 9. 15.	국가 표준제도의 확립	김재관
27	2000. 12. 1.	국가 정보경쟁력의 잣대: 전자정부	이상희
28	2001. 5. 4.	환경위기 극복과 지속가능 경제발전을 위한 과학 기술개발전략	박원훈, 류순호, 문길주, 오종기, 한무영, 한정상
29	2001. 7. 18.	국가 과학기술발전에 미치는 기초과학의 영향	임관, 명호철, 장수영
30	2001. 9. 21.	산업계에서 원하는 인재상과 공학교육의 방향	임관, 한송엽
31	2001. 10. 31.	적조의 현황과 앞으로의 대책	홍승룡, 김학균
32	2001. 12. 5.	광우병과 대책	김용선, 한홍율
33	2002. 7. 19.	첨단기술 (BT,ET,IT,NT)의 실현을 위한 산업화 대책	한문희, 이석한, 한송엽
34	2002. 9. 13.	우리나라 쌀 산업의 위기와 대응	이정환, 김동철
35	2002. 11. 1.	생명윤리 - 과학 그리고 법: 발전이냐 규제냐?	문신용, 이신영
36	2003. 3. 14.	과학기술분야 졸업생의 전공과 직업의 연관성	조황희, 이만기
37	2003. 6. 18.	국내 농축산물 검역현황과 발전방안	배상호
38	2003. 6. 27.	대학과 출연연구소간 연구협력 및 분담	정명세
39	2003. 9. 26.	그린에너지 기술과 발전 방향	손재익, 이재영, 홍성안
40	2004. 2. 20.	미래 고령사회 대비 국가 과학기술 전략	오종남
41	2004. 10. 27.	고유가시대의 원자력 이용	정근모
42	2004. 12. 7.	농산물 개방화에 따른 국내 고추산업의 현황과 발전전략	박재복
43	2005. 9. 30.	과학기술윤리	송상용, 황경식, 김환석

회수	일 자	주 제	발제자
44	2005. 11. 25.	과학기술용어의 표준화 방안	지제근
45	2005. 12. 1.	융합과학시대의 수학의 역할 및 수학교육의 방향	정근모, 최형인, 장준근
46	2005. 12. 15.	해양바이오산업, 왜 중요한가?	김세권, 김동수
47	2006. 11. 7.	첨단과학시대의 교과과정 개편방안	박승재
48	2006. 12. 22.	과학기술인 복지 증진을 위한 종합 대책	설성수
49	2007. 6. 29.	선진과학기술국가 가능한가? - Blue Ocean을 중심으로	김호기
50	2007. 11. 9.	우리나라 수학 및 과학교육의 문제점과 개선방향	김도한, 이덕환
51	2008. 5. 9.	태안반도 유류사고의 원인과 교훈	하재주
52	2008. 5. 8.	광우병과 쇠고기의 안전성	이영순
53	2008. 6. 4.	고병원성조류인플루엔자(AI)의 국내외 발생양상과 우리의 대응방안	김재홍
54	2008. 10. 8.	High Risk, High Return R&D, 어떻게 해야 하는가	김호기
55	2008. 11. 11.	식량위기 무엇이 문제인가?	이정환
56	2008. 12. 11.	초중고 수학 과학교육 개선방안	홍국선
57	2008. 12. 17.	우리나라 지진재해 저감 및 관리대책의 현황과 개선방안	윤정방
58	2009. 2. 19.	21세기 지식재산 비전과 실행 전략	김영민
59	2009. 3. 31.	세계주요국의 나노관련 R&D 정책 및 전략분석과 우리의 대응전략	김대만
60	2009. 7. 20.	국가 수자원 관리와 4대강	심명필
61	2009. 8. 28.	사용후핵연료 처리 기술 및 정책 방향	송기찬, 전봉근
62	2009. 12. 16.	세종시와 국제과학비즈니스벨트	이현구

회수	일 자	주 제	발제자
63	2010. 3. 18.	과학도시와 기초과학 진흥	김중현
64	2010. 6. 11.	지방과학기술진흥의 현황과 과제	정선양
65	2011. 2. 28.	국제과학비즈니스벨트와 기초과학진흥	민동필, 이충희
66	2011. 4. 1.	방사능 공포, 오해와 진실	기자회견
67	2012. 11. 30.	융합과학/융합기술의 본질 및 연구방향과 국가의 지원시스템	이은규, 여인국
68	2013. 4. 17.	한미원자력협정 개정협상에 거는 기대와 희망	문정인
69	2013. 6. 11.	통일을 대비한 우리의 식량정책 이대로 좋은가?	이철호
70	2013. 7. 9.	과학기술중심사회를 위한 과학기술원로의 역할과 의무	이원근
71	2013. 7. 22.	대학입시 문·이과 통합, 핵심쟁점과 향후 과제는?	박재현
72	2014. 1. 17.	국가안보 현안과제와 첨단과학기술	송대성
73	2014. 3. 4.	융합과학기술의 미래 -인재교육이 시작이다	강남준, 이진수
74	2014. 5. 9.	과학기술연구의 새 지평 젠더혁신	이혜숙, 조경숙, 이숙경
75	2014. 5. 14.	남북한 산림협력을 통한 한반도 생태통일 방안은?	김호진, 이돈구
76	2014. 5. 22.	창조경제와 과학기술	이공래, 정선양
77	2014. 5. 29.	재해·재난의 예방과 극복을 위한 과학기술의 역할은?	이원호, 윤정방
78	2014. 6. 10.	벼랑 끝에 선 과학·수학 교육	정진수, 배영찬
79	2014. 6. 14.	문학과 과학, 그리고 창조경제	정종명, 최진호
80	2014. 6. 25.	'DMZ세계평화공원'과 남북과학기술협력	정선양, 이영순, 강동완
81	2014. 7. 24.	국내 전통 발효식품산업 육성을 위한 정책 대안은?	신동화

회수	일 자	주 제	발제자
82	2014. 9. 17.	‘과학기술입국의 꿈’을 살리는 길은?	손경한, 안화용
83	2014. 9. 30.	한국 산업의 위기와 혁신체제의 전환	이 근
84	2014. 11. 14.	경제, 사회, 문화, 산업 인프라로서의 사물인터넷(IoT): 그 생태계의 실현 및 보안방안은?	김대영, 김용대
85	2014. 11. 28.	공유가치창출을 위한 과학기술의 나아갈 길은? 미래식품과 건강	권대영
86	2014. 12. 5.	창발적 사고와 융합과학기술을 통한 글로벌 벤처 생태계 조성 방안	허석준, 이기원
87	2015. 2. 24.	구제역·AI의 상재화: 정부는 이대로 방치할 것인가?	김재홍
88	2015. 4. 7.	문·이과 통합 교육과정에 따른 과학·수학 수능개혁	이덕환, 권오현
89	2015. 6. 10.	이공계 전문가 활용 및 제도의 현황과 문제점	이건우, 정영화
90	2015. 6. 25.	남북 보건의료 협정과 통일 준비	신희영, 윤석준
91	2015. 7. 1.	메르스 현황 및 종합대책	이종구
92	2015. 7. 3.	‘정부 R&D 혁신방안’의 현황과 과제	윤헌주
93	2015. 9. 14.	정부 R&D예산 감축과 과학기술계의 과제	문길주
94	2015. 10. 23.	사회통합을 위한 과학기술 혁신	정선양, 송위진
95	2015. 11. 4.	생명공학기술을 활용한 우리나라 농업 발전방안	이향기, 박수철, 곽상수
96	2015. 11. 9.	유전자가위 기술의 명과 암	김진수
97	2015. 11. 27.	고령화사회와 건강한 삶	박상철
98	2015. 12. 23.	따뜻한 사회건설을 위한 과학기술의 역할: 국내외 적정기술을 중심으로	박원훈, 윤제용
99	2016. 2. 29.	빅데이터를 활용한 의료산업 혁신방안은?	이동수, 송일열, 유희준
100	2016. 4. 18.	대한민국 과학기술 미래 50년의 도전과 대응	김도연

회수	일 자	주 제	발제자
101	2016. 5. 19.	미세먼지 저감 및 피해방지를 위한 과학기술의 역할	김동술, 박기홍
102	2016. 6. 22.	과학기술강국, 지역 혁신에서 답을 찾다	남경필, 송종국
103	2016. 7. 6.	100세 건강과 장내 미생물 과학! 어디까지 왔나?	김건수, 배진우, 성문희
104	2016. 7. 22.	로봇 기술과 미래	오준호
105	2016. 8. 29.	융합, 융합교육 그리고 창의적 사고	김유신
106	2016. 9. 6.	분노조절장애, 우리는 얼마나 제대로 알고 있나?	김재원, 허태균
107	2016. 10. 13.	과학기술과 미래인류	이광형, 백종현, 전경수
108	2016. 10. 25.	4차 산업혁명시대에서 젠더혁신의 역할	이우일, 이혜숙
109	2016. 11. 9.	과학기술과 청년(부제: 청년 일자리의 현재와 미래)	이영무, 오세정
110	2017. 3. 8.	반복되는 구제역과 고병원성 조류인플루엔자, 정부는 이대로 방치할 것인가	류영수, 박최규
111	2017. 4. 26.	지속가능한 과학기술 혁신체계	김승조, 민경찬
112	2017. 8. 3.	유전자교정 기술도입 및 활용을 위한 법·제도 개선방향	김정훈
113	2017. 8. 8.	탈원전 논란에 대한 과학자들의 토론	김경만, 이은철, 박홍준
114	2017. 8. 11.	새롭게 도입되는 과학기술혁신본부에 바란다	정선양, 안준모
115	2017. 8. 18.	ICT 패러다임을 바꿀 양자통신, 양자컴퓨터의 부상	허 준, 최병수, 김태현, 문성욱
116	2017. 8. 22.	4차 산업혁명을 다시 생각한다	홍성욱, 이태억
117	2017. 9. 8.	살충제 계란 사태로 본 식품안전관리 진단 및 대책	이향기, 김병훈
118	2017. 11. 17.	미래 과학기술을 위한 정책입법 및 교육, 어떻게 해야 하나	박형욱, 양승우, 최윤희
119	2017. 11. 28.	여성과기인 정책 업그레이드	민경찬, 김소영

회수	일 자	주 제	발제자
120	2017. 12. 8.	치매국가책임제, 과학기술이 어떻게 기여할 것인가?	김기웅, 묵인희
121	2018. 1. 23.	항생제내성 수퍼박테리아! 어떻게 잡을 것인가?	정석훈, 윤장원, 김홍빈
122	2018. 2. 6.	신생아 중환자실 집단감염의 발생원인과 환자 안전 확보방안	최병민, 이재갑, 임채만, 천병철, 박은철
123	2018. 2. 27.	에너지전환정책, 과학기술자 입장에서 본 성공 여건	최기련, 이은철
124	2018. 4. 5.	과학과 인권	조효제, 민동필, 이중원, 송세련
125	2018. 5. 2.	4차 산업혁명시대 대한민국의 수학교육, 이대로 좋은가	권오남, 박형주, 박규환
126	2018. 6. 5.	국가 R&D 혁신 전략 - 국가 R&D 정책 고도화를 위한 과학기술계 의견 -	류광준, 유욱준
127	2018. 6. 12.	건강 100세를 위한 맞춤 식품 필요성과 개발 방향	박상철, 이미숙, 김경철
128	2018. 7. 4.	제1회 세종과학기술포럼	성창모, 박찬모, 이공래
129	2018. 9. 18.	데이터 사이언스와 바이오 강국 코리아의 길	박태성, 윤형진, 이동수
130	2018. 11. 8.	제10회 국회-한림원 과학기술혁신연구회 포럼(미래과학기술 오픈포럼) - 미래한국을 위한 과학기술과 정책 -	임대식, 문승현, 문 일
131	2018. 11. 23.	아카데미 캐피탈리즘과 책임 있는 연구	박범순, 홍성욱
132	2018. 12. 4.	여성과학기술인 정책, 4차 산업혁명 시대를 준비하는가	이정재, 엄미정
133	2019. 2. 18.	제133회 한림원탁토론회 -제17회 과총 과학기술혁신정책포럼 수소경제의 도래와 과제	김봉석, 김민수, 김세훈
134	2019. 4. 18.	혁신성장을 이끄는 지식재산권 창출과 직무발명 조세제도 개선	하홍준, 김승호, 정지선
135	2019. 5. 9.	제135회 한림원탁토론회 - 2019 세종과학기술인대회 과학기술 정책성과와 과제	이영무
136	2019. 5. 22.	효과적인 과학인재 양성을 위한 전문연구요원 제도 개선 방안	곽승엽
137	2019. 6. 4.	마약청정국 대한민국이 흔들린다 마약류 사용의 실태와 대책은?	조성남, 이한덕

회수	일 자	주 제	발제자
138	2019. 6. 28.	미세먼지의 과학적 규명을 위한 선도적 연구 전략	윤순창, 안병옥
139	2019. 8. 7.	공동 토론회 - 일본의 반도체 · 디스플레이 소재 수출규제에대한 과학기술계 대응방안	박재근
140	2019. 9. 4.	4차 산업혁명 시대 농식업(Agriculture and Food) 변화와 혁신정책 방향	권대영, 김종윤, 박현진
141	2019. 9. 25.	과학기술 기반 국가 리스크 거버넌스, 어떻게 구축해야 하는가?	고상백, 신동천, 문일, 이공래







www.kast.or.kr

이 사업은 복권기금 및 과학기술진흥기금 지원을 통한 사업으로
우리나라의 사회적 가치 증진에 기여하고 있습니다.