

도덕적 인공지능 로봇의 법적 지위



법무법인 명륜
양희철 변호사

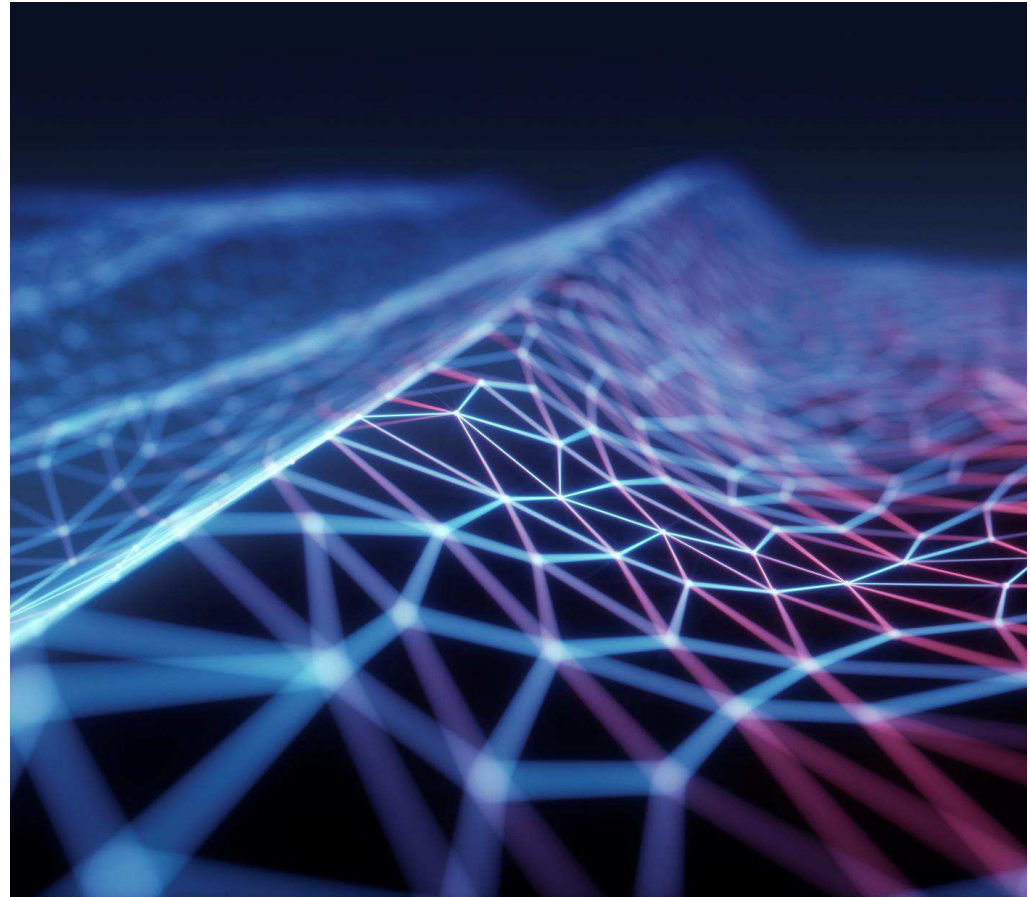
Computing Machinery and Intelligence

- 앨런 튜링(Alan Turing) -

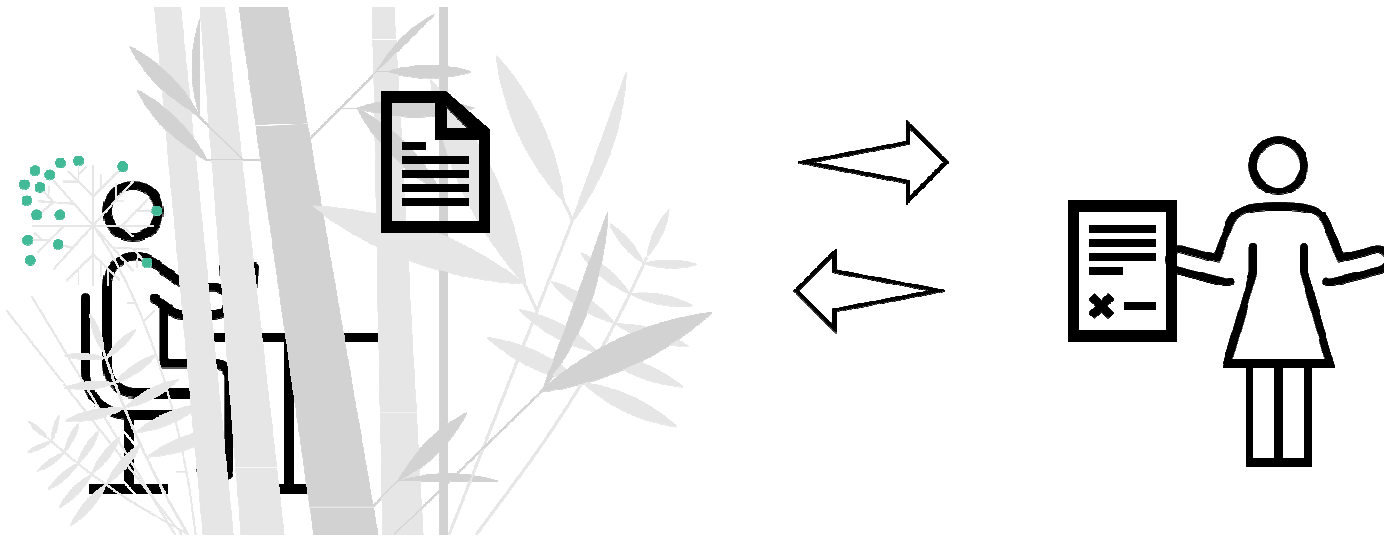
기계는 사고할 수 있는가?

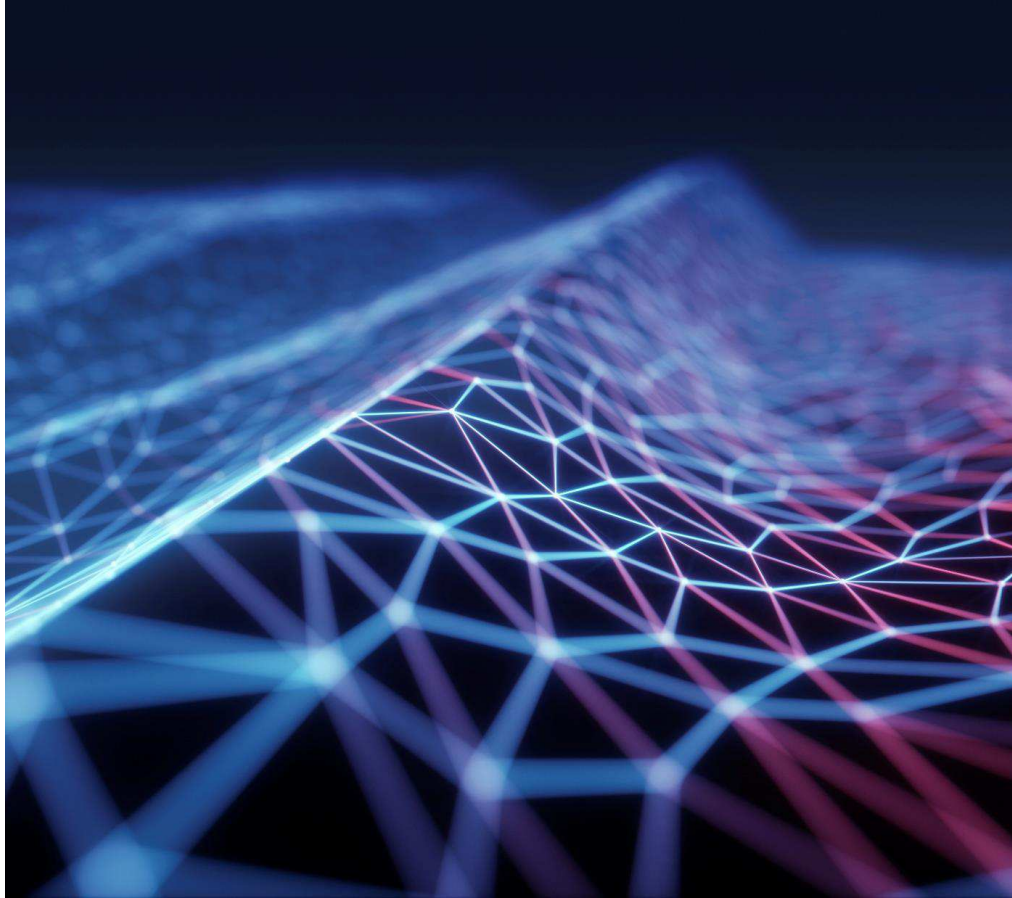
특정 지적 업무를 수행할 능력이 있는가?

Turing Test



존 설(JOHN SEARLE)의 튜링 테스트 비판 - 중국어 방





“어떤 컴퓨터 프로세스를
지능을 가진 것이라 부를지
정할 수 없기 때문에 인간
지능과 무관하게 지능을
명확하게 정의할 수 없다.”

- John McCarthy -

인공지능(AI, Artificial Intelligence) 용어
최초 사용

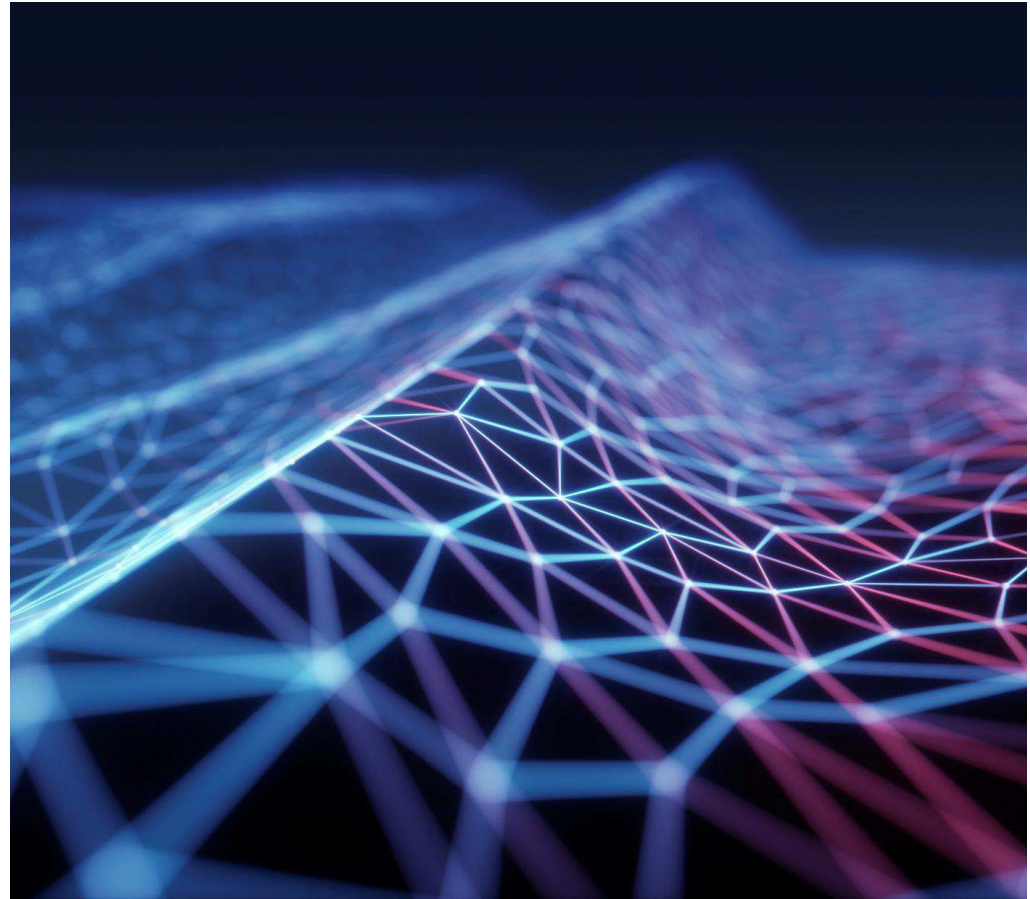
무능력 척도로 체스를 제시

: “일단 작동하게 되면 아무도 더 이상
그것을 인공지능이라 하지 않는다.”

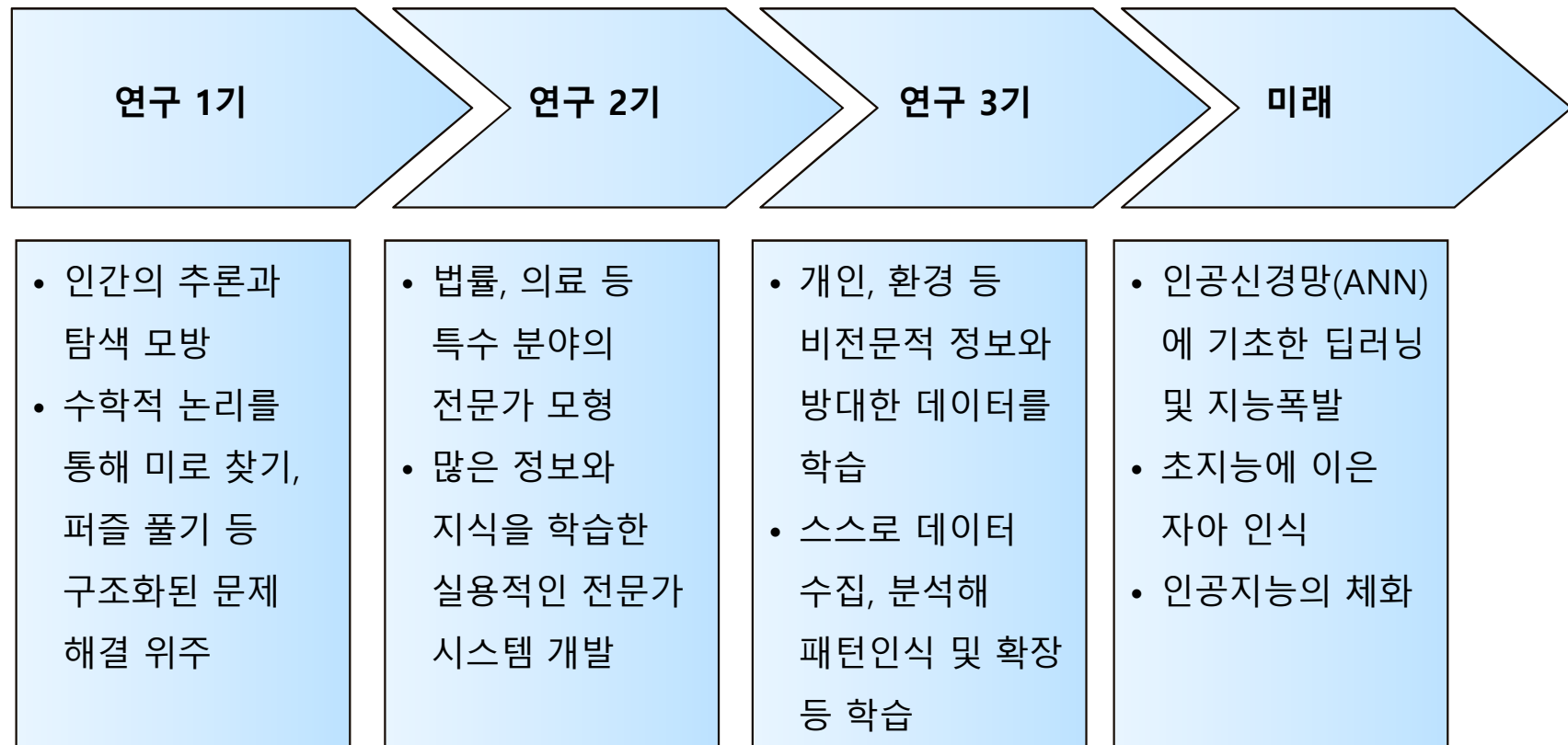
“인공지능(AI)은 사람을 돕는 지능적인 기계”

- 앤드류 응(Andrew Ng) -

인간의 직업을 빼앗고, 킬러 로봇이 등장할
것이라는 세상의 우려와 달리
인공지능은 인류의 문제를 해결할 수 있다.

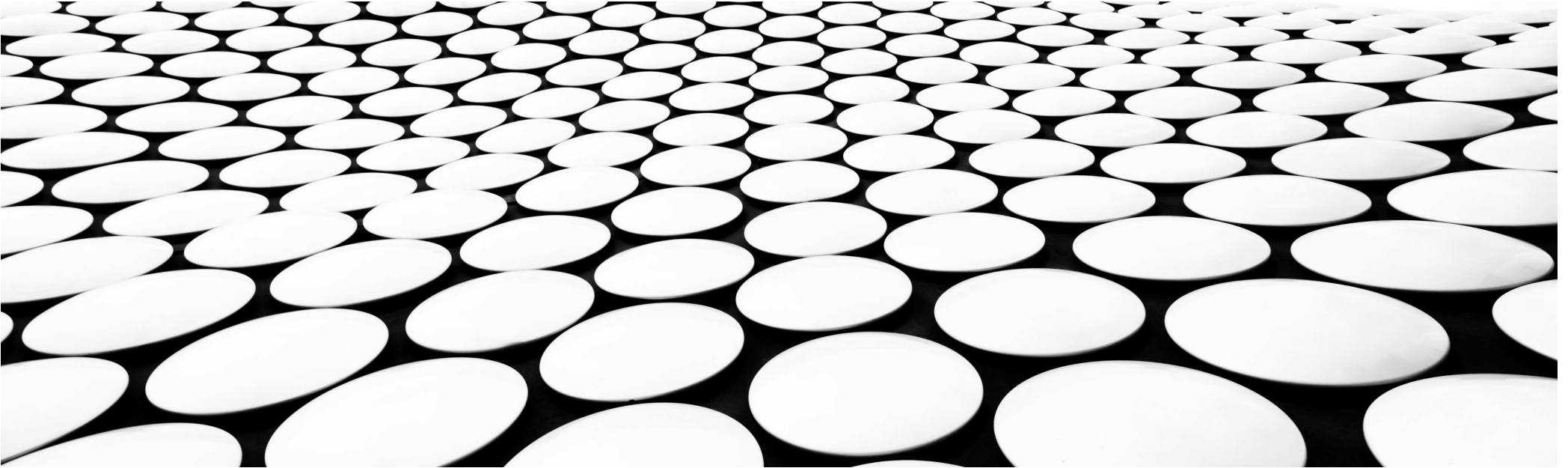


인공지능 연구 과정



1. 우리 곁의 인공지능

챗봇, OTT, 자율주행자동차, 의료 인공지능, 인공지능 판사...



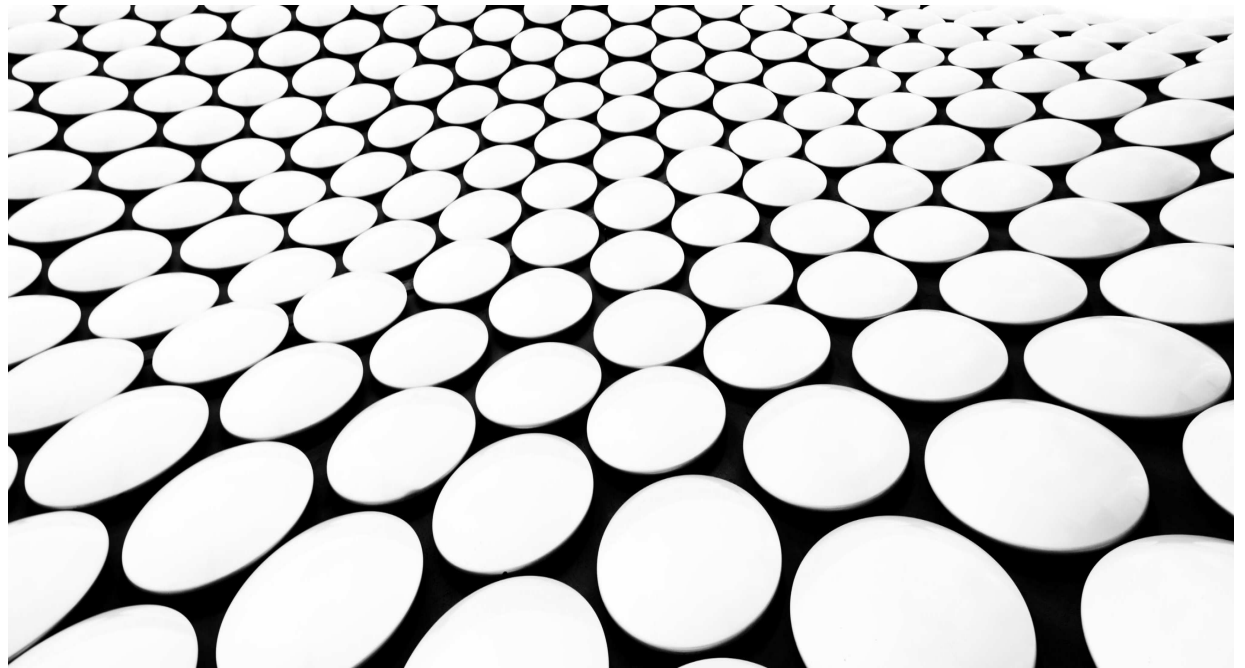
챗봇(CHAT BOT) 서비스

미국 마이크로소프트(Microsoft)
챗봇 '테이'

: 인증차별, 혐오 발언

스캐터랩 챗봇 '이루다'

: 학습데이터의 개인정보 침해
문제 / 사회적 약자 차별, 혐오
발언



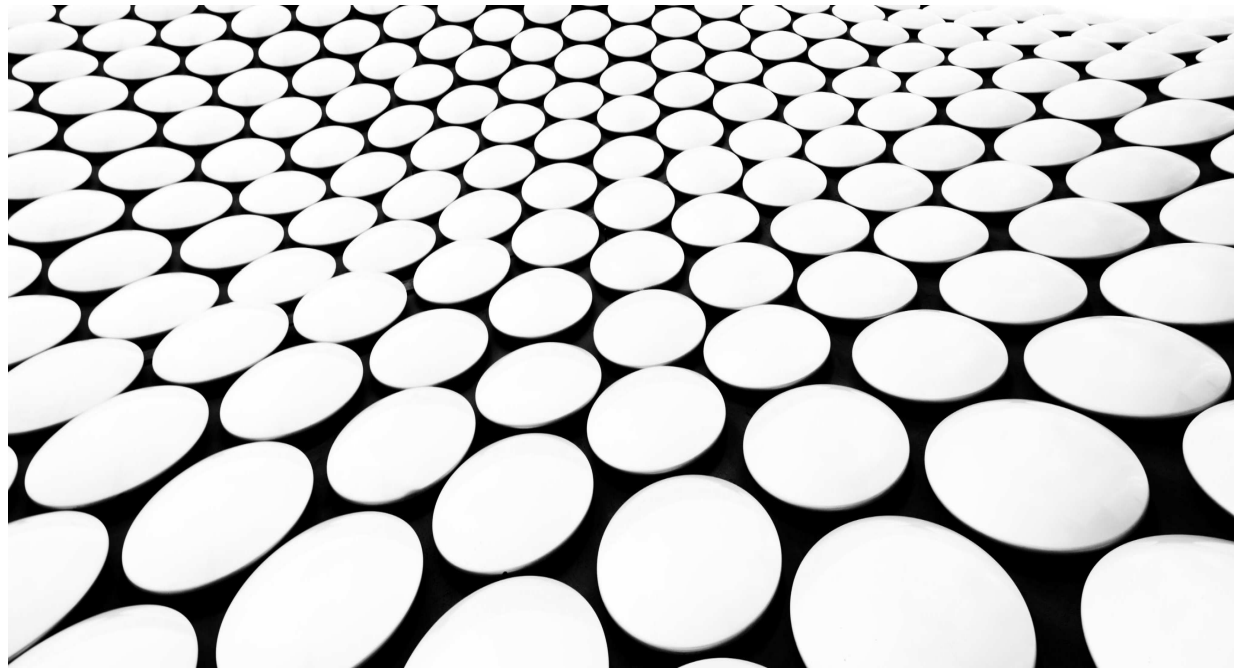
추천 알고리즘

미국 할인마켓 타겟(Target)

: 여고생에게 유아 용품 할인 쿠폰 발송

OTT 서비스(넷플릭스, 유튜브)

: 프로그램 및 광고 추천



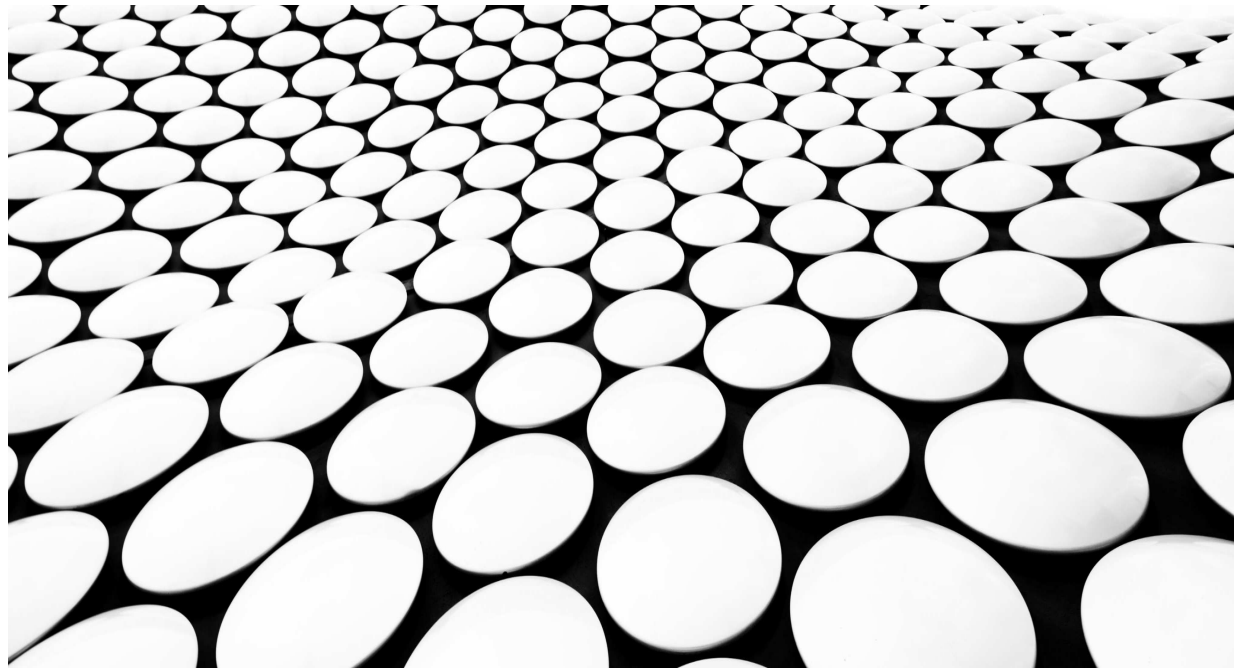
자율주행자동차

테슬라의 오토 파일럿

: 계속된 사고 발생

커넥티드 카(Connected Car)

: 자율주행 완성도 향상



자율주행자동차

자율주행자동차법 개정

: V2X 방식의 자율협력주행을 이용하는 지능형 교통체계

자율주행 로보택시 운행

: 현재 시범 운행 중 / 향후 일반인도 이용 가능

자율주행자동차 상용화 촉진 및 지원에 관한 법률(약칭: 자율주행자동차법)

[시행 2022. 1. 28.] [법률 제18348호, 2021. 7. 27., 일부개정]

제2조(정의) ① 이 법에서 사용하는 용어의 뜻은 다음과 같다.

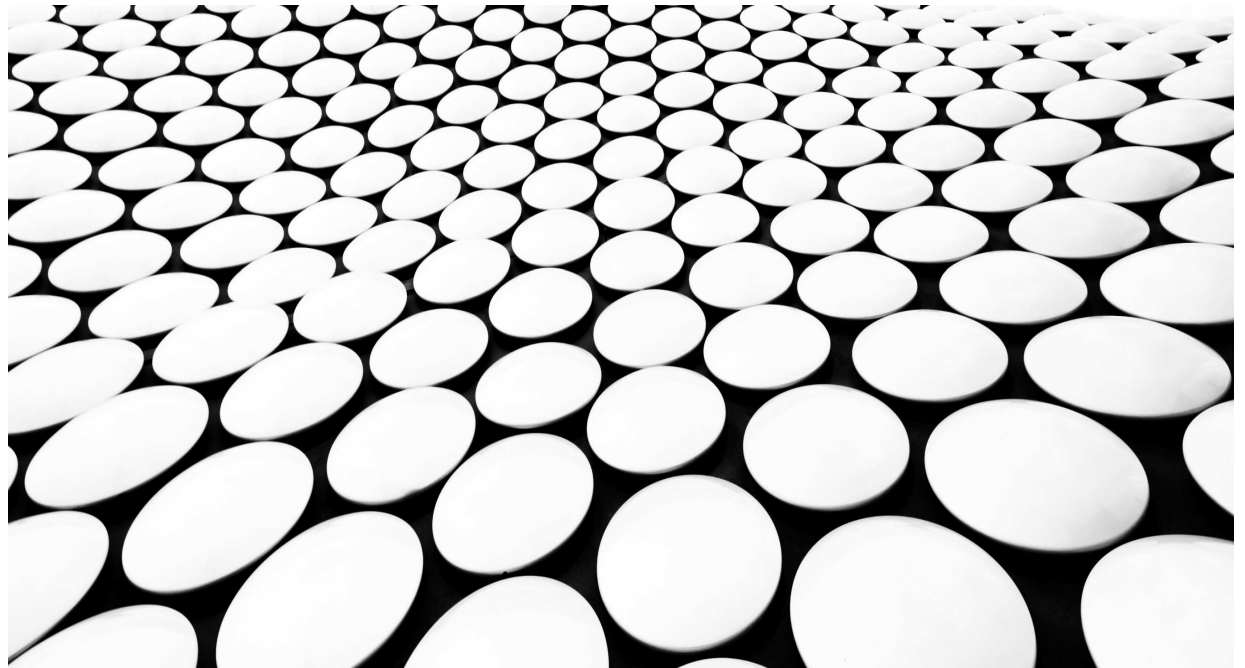
3. “자율협력주행시스템”이란 「[도로교통법](#)」 [제2조제15호](#)에 따른 신호기, 같은 조 제16호에 따른 안전표지, 「[국가통합교통체계효율화법](#)」 [제2조제4호](#)에 따른 교통시설 등을 활용하여 [국토교통부령](#)으로 정하는 바에 따라 자율주행기능을 지원·보완하여 효율성과 안전성을 향상시키는 「[국가통합교통체계효율화법](#)」 [제2조제16호](#)에 따른 지능형교통체계를 말한다.

의료 인공지능

인공지능 기반 의료기기

: IBM 왓슨 도입 및 한계

- 과장되고 성급한 시장 진출
- 학습 데이터의 편중
- 결과값 근거 설명 불가능
- 전자의무기록화 어려움연동



의료 인공지능

국내 개발 인공지능 기반 의료 기기

: 2018년 식약처의 (주)뷰노의 의료영상분석장치 소프트웨어 VUNOMed-BoneAge 이후 다양한 의료기기 출시

2018. 5.

(주)뷰노
VUNOMed-BoneAge
골연령 판독
의료영상분석
장치 소프트웨어

2018. 8.

(주)제이엘케이인스퍼션
JBS-01K
폐결핵
의료영상진단보조소프트웨어

2019. 7.

(주)루닛
Lunit INSIGHT MMG
유방암
의료영상판독보조소프트웨어

현재

식약처 2022. 9. 기준
국내 제조 124건,
수입 15건

2018. 8.

(주)루닛
Lunit INSIGHT CXR
흉부 의료영상검출보조소프트웨어

2019. 8.

(주)뷰노
체스트 엑스레이,
딥브레인 치매 진단
의료영상분석
장치 소프트웨어

인공지능 판사

국내 대법원 법원행정처

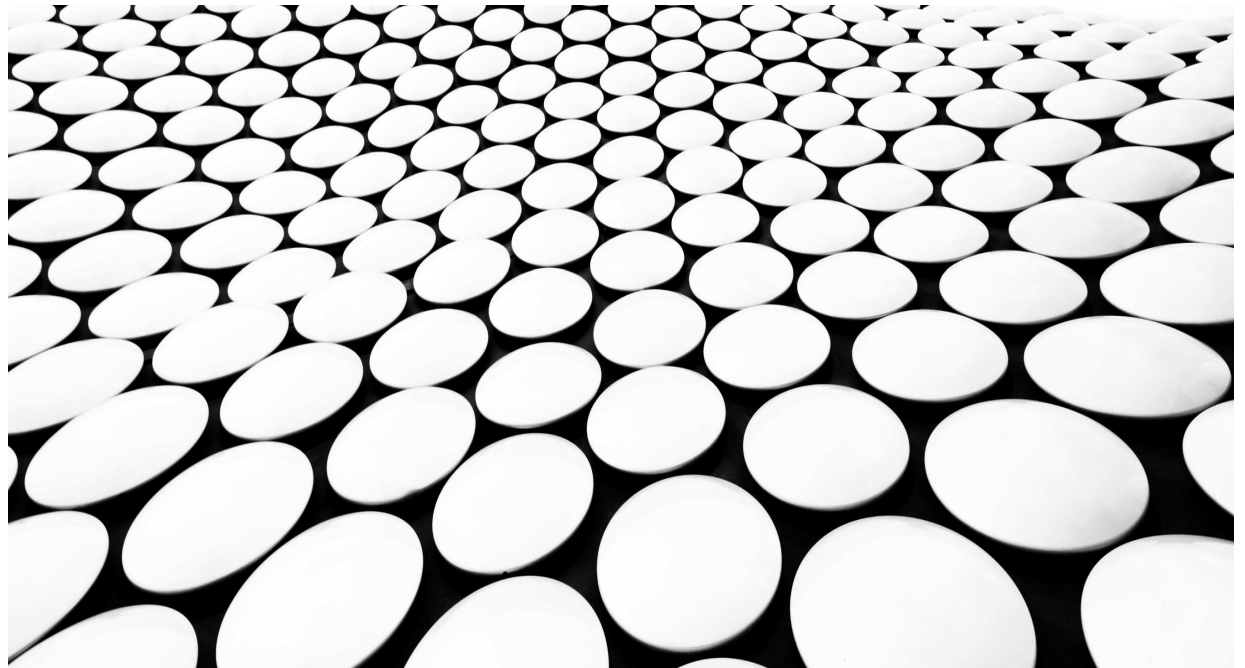
: 손해배상액 산정 보조

에스토니아

: 소액사건 인공지능 판사 도입

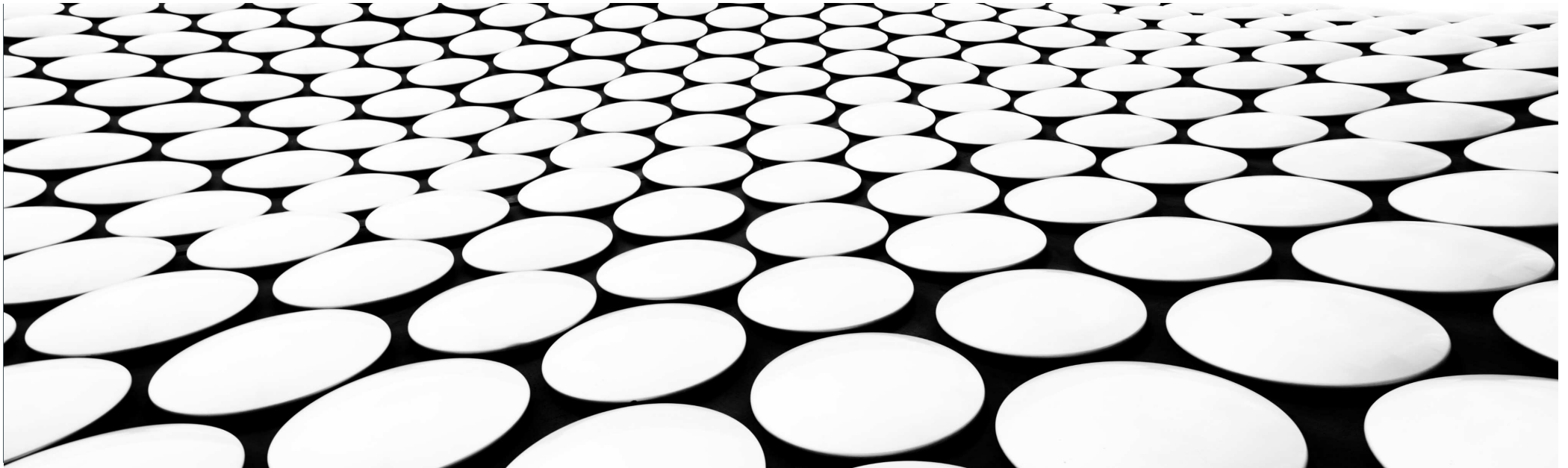
중국 법원의 AI 활용

: 사법 분야 활용의 명암



2. 인공지능 로봇과 책임 주체

인공지능 로봇을 보편적으로 이용하기 위한 조건



인공지능 로봇 특성

자율성을 갖는 인공지능은 3가지 특성을 지니게 될 것이다.

- Ryan Calo -



체화
(Embodiment)



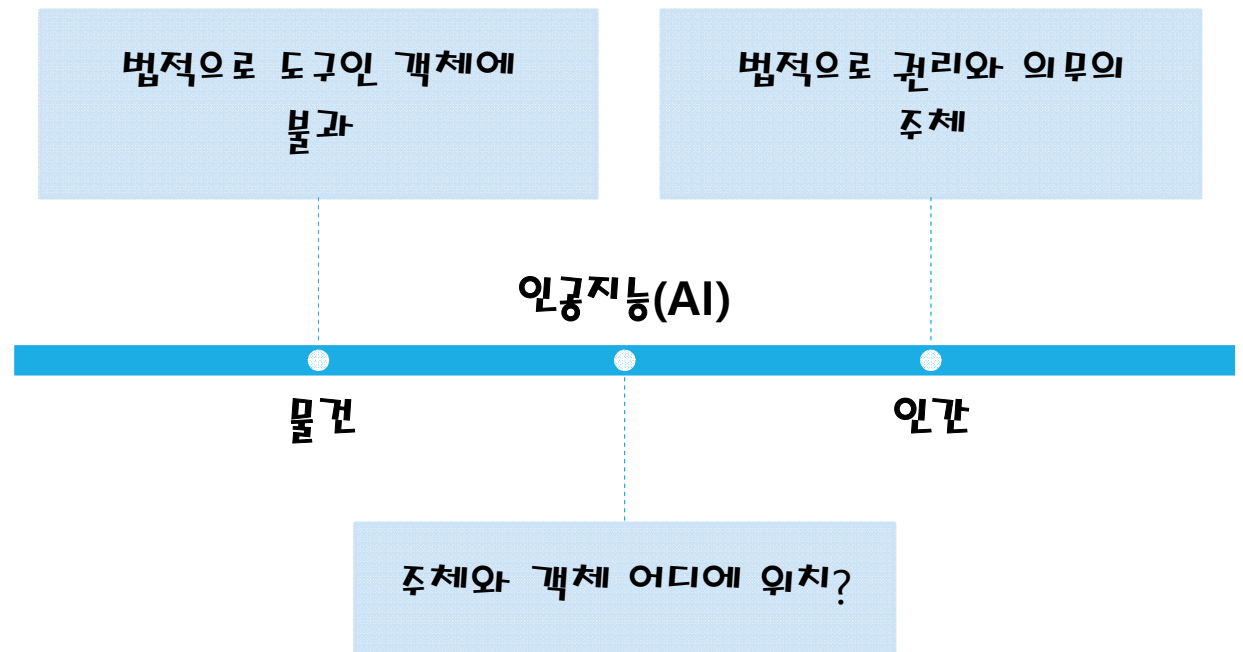
창발
(Emergence)



사회적 유효성
(Social Valence)

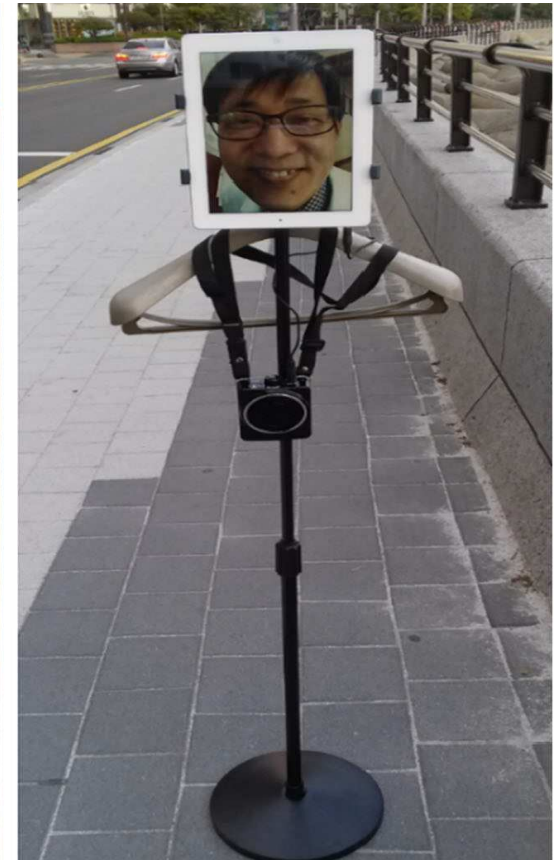
도구와 자유의지 사이

인공지능 로봇은 물건인 도구일까 아니면 인간처럼 자유 의지를 갖는 존재일까?



로봇에게도 권리를
인정할 수 있을까?

2014년 3월부터 4월까지 배
일한 교수에 의해 진행된 대
한민국 20개 도시에서 실험



실험 참가자들의 권리 인정 답변 평균 점수

[5점 만점]

	평등 대우	평등 보호	평등 서비스	신체의 자유	표현의 자유	선거권	공무담임권
고정 영상통화 장치 구경한 참가자	1.79	1.91	2.76	2.14	2.36	1.30	1.17
원격인식로봇 구경한 참가자	2.53	2.51	2.60	2.33	2.79	2.06	1.60
원격인식로봇 작동시킨 참가자	2.76	2.50	2.81	2.38	2.72	2.16	1.88

2. 인공지능 로봇과 책임 주체

민사 책임

- 계약상 책임
 - 도구 vs 주체
- 불법행위 책임
 - 도구 vs 주체

형사 책임

- 터미네이터 형사처벌?
 - 가능 vs 불가능
- 대안적 형사책임

법적 책임 관련 보다 상세한 논의는 양희철, "인공지능(AI) 로봇의 법적 지위 고찰 - 인공지능 로봇의 인격과 법적 책임 귀속 -", 서울대학교 대학원, 2020년 8월 법학석사 학위논문 참조

설명 가능한 인공지능(XAI, EXPLAINABLE AI)

블랙박스 문제



설명 가능한 인공지능 도입 필요성

해석가능성



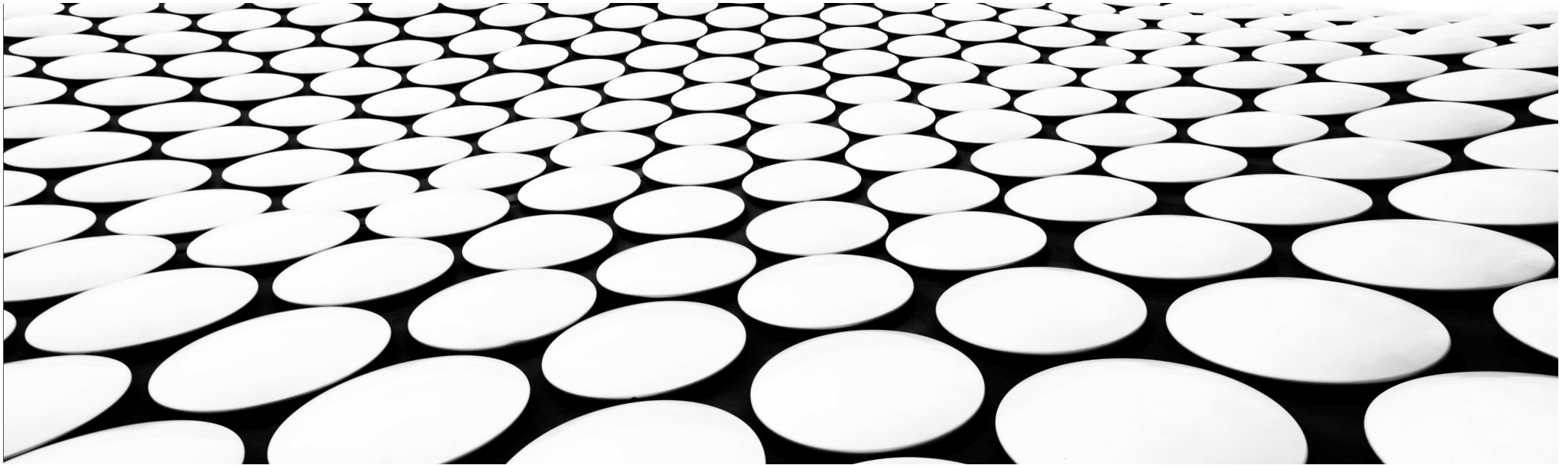
완전무결성

- 모델 비종속적 설명 (Model -agonistic)
: 블랙박스에 대한
귀납적 추론 기반

- 모델 종속적 설명 (Model -specific)
: 매개변수, 아키텍처
관련 지식 기반

3. 책임을 지는 인공지능

법적 주체와 객체



법인격이란?

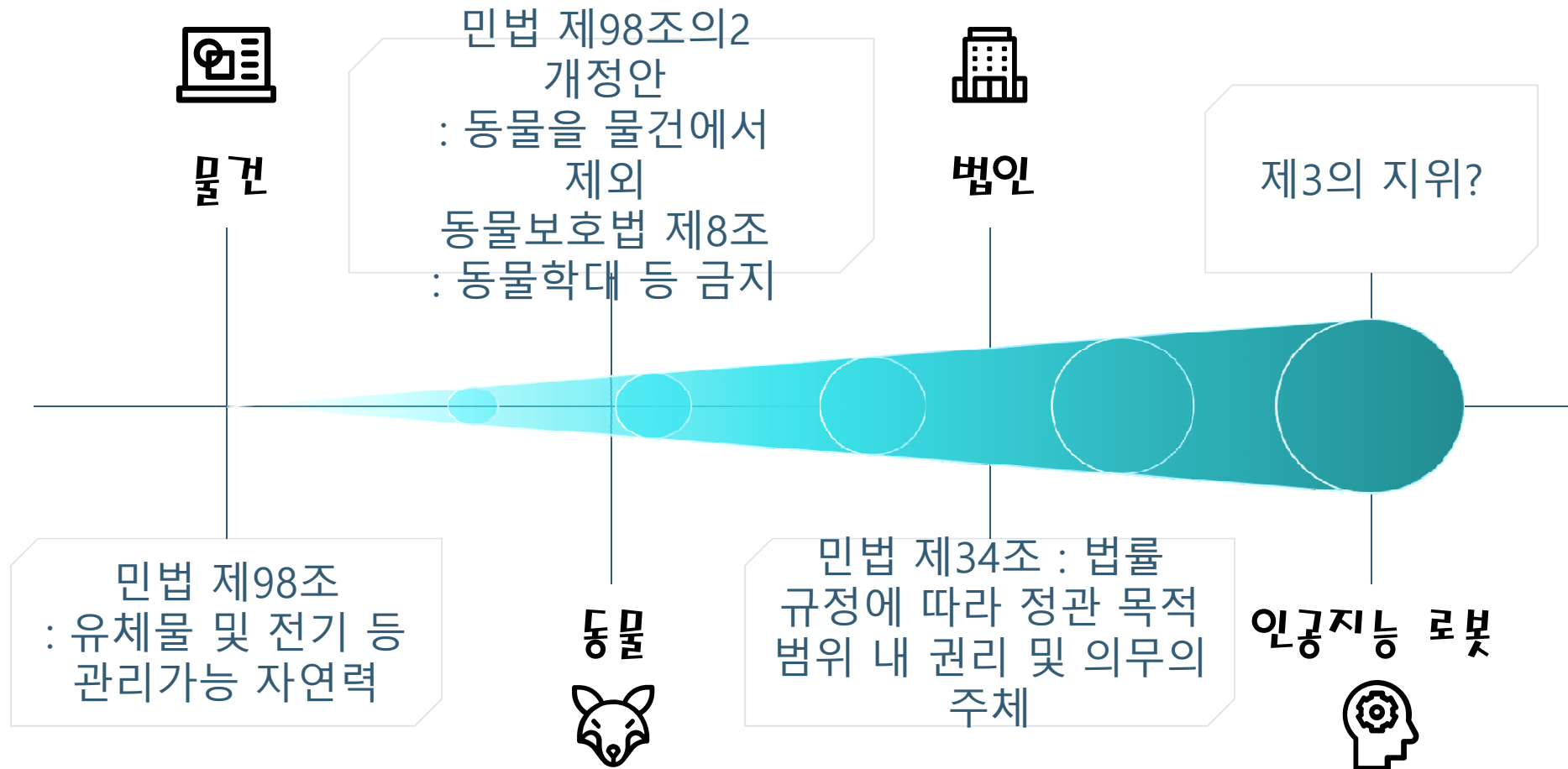
생래적으로 법인격을 가진 자연인

인공지능 로봇?



목적을 가진 가상의 의제적 존재인 법인

TIMELINE



법인격이 문제되는 존재



동물

동물의 도덕적 지위
물건이 아닌 존재
(민법 개정안)



인간 초기 배아

초기 배아의 법적 지위
인정 근거는?



인간 노예

역사적 존재였던 인간
노예의 지위

동물의 법적 지위

동물복지론

- 피터 싱어(Peter Singer)의 공리주의적 동물복지론
 - 인간의 필요에 따라 동물을 이용하되 최소한의 배려 필요
 - 동물의 기본적 욕구를 충족시키고, 고통을 최소화하는 행복한 상태 추구
 - 종차별주의 반대

동물권리론

- 톰 리건(Tom Regan)
 - 동물의 도덕적, 법적 권리 주장
 - 동물에게도 각자의 고유한 가치가 있으므로 인간의 이익을 위한 수단으로 이용하는데 반대

초기 배아의 법적 지위



헌법재판소

1. 독립된 인간과 배아 간의 개체적 연속성을 확정하기 어렵다.
2. 현재 과학기술 수준에서는 배아가 체외에서 독립적 인간으로 성장할 가능성을 기대하기 어렵다.
3. 수정 후 착상 전 배아가 인간으로 인식된다거나 그렇게 취급되어야 한다는 사회적 승인이 존재한다고 보기 어렵다.

헌재 결정 비판론

1. 수정란은 인간종이고, 형성 중인 생명인 인간의 배아는 인간이 될 잠재성이 내재해 기본권 주체성 인정되는 태아 직전 단계
2. 미래 과학기술 발전으로 인공자궁 개발 시 배아의 기본권 주체성 인정 모순
3. 최근 과학기술 발전으로 배아의 인간성이나 법적 지위 논의가 시작되어 현재 기준 사회적 승인 판단 곤란

초기 배아의 법적 지위

인간 초기 배아의 특수성

1. 인간의 초기 배아는 사법상 물건의 지위에 있지만, 헌법적 생명체로서 생명권이라는 기본권의 주체 지위 역시 가짐
2. 규범이란 단지 객관적 사실만이 아니라 사회적으로 합의된 규범의 제정 취지에도 기초하기 때문에 일견 법체계상 모순이라 볼 수 있는 이러한 상황도 가능
3. 초기 배아의 특수한 성격으로 인해 미국 테네시 주 대법원은 수정란을 인간도, 물건도 아닌 그 중간적 존재라고 판단

인간 노예

“그들은 가장 골치 아픈 동물임에 틀림없다.”

- 플라톤 [법률론] -

“공동체로부터 뿌리 뽑힌 사람(deracinated people)”

- 모시스 아이작 핀리(M. I. FINLEY) -

인간 노예

특유재산(特有財産, *peculium*)

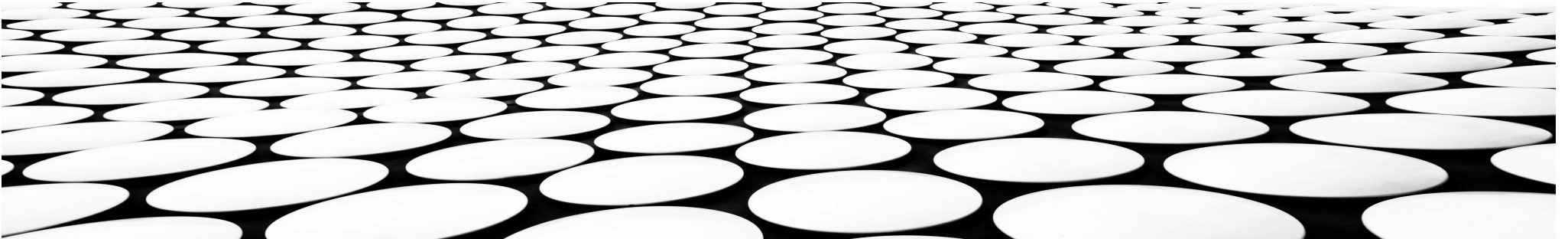
1. 노예가 가부장을 대신해서 소유한 것은 "가부장의 인격에 의한(EX PERSONA PATRIS)" 행위로서 법적으로 유효
2. "법률상 당연히(IPSO IURO)" 가부장에게 귀속
3. 법률상, 사실상 소유자가 다른 법과 현실의 괴리 반영

인질(人質)

1. 약속을 지키는 것에 대한 담보가 되어 상대방에게 억류된 사람
2. 질권 설정의 대상

법인의 성격과 비교

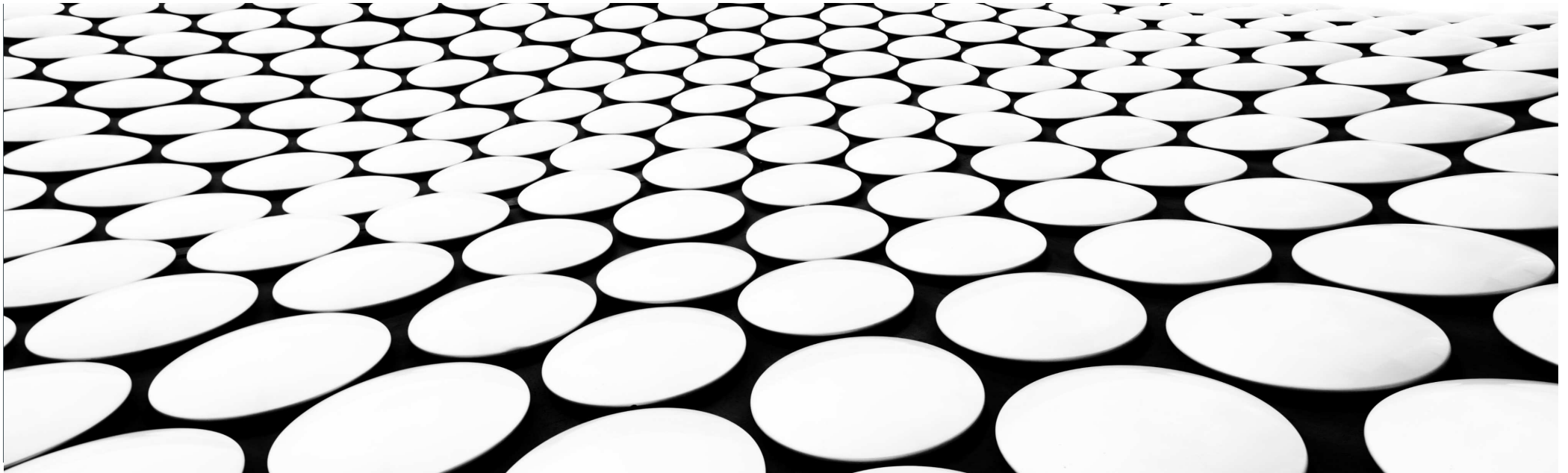
- 창조된 인공물
- 제한된 권리 및 의무 보유



법적 주체와 관련된 보다 상세한 논의는 양희철, “인공지능(AI) 로봇의 법적 지위 고찰 - 인공지능 로봇의 인격과 법적 책임 귀속 -”, 서울대학교 대학원, 2020년 8월 법학석사 학위논문, 양희철, “법정에 출석한 인공지능”, 복저널리즘, 2021. 8. 각 참조

4. 인공적 도덕 행위자

생명체, 인공지능 로봇, 도덕 행위자



생명체 기준

생명체 개념 정의 곤란

가이아 가설
(Gaia hypothesis)



1. 물질대사를 통한 자기생산(autopoeis)

주변 세계와 지속적으로 물질을 교환하면서 자신의 일정한 체계를 유지하기 위한 자신에게 필요한 물질을 만들어내는 동화작용과 이화작용이란 물질대사

2. 생리적 기능

호흡하고, 영양분을 흡수하고, 소화시키며, 배설하면서, 성장·생식하는 존재



생명체 기준

생명체 개념 정의 곤란

가이아 가설
(Gaia hypothesis)



3. 유전적 관점

최소한 종의 수준에서 보았을 때 번식을 한다는 것으로 자신과 유전적으로 유사한 자손을 낳는 것

4. 생화학적 관점

유전정보가 기록된 유전자를 가지고 있고, 효소의 촉매 작용을 통해 생화학적 반응과 대사 작용을 하는 체계라고 미시적으로 보는 것



생명체 기준

생명체 개념 정의 곤란

가이아 가설
(Gaia hypothesis)



5. 열역학 관점

열린 계로서 시간의 경과에 따라 무질서도가 증가한다는 열역학 제2법칙, 즉 엔트로피 증가 법칙에 반해, 생명체에서는 질서가 유지

에르빈 슈뢰딩거(Erwin Schrödinger)

"생명은 음의 엔트로피를 먹고 산다."

"생명은 환경으로부터 '질서'를 얻어내며 유지된다."

세포, 유전자, 화학, 정보 및 진화 개념에 기초한 생명의 정의

1. 경계를 지닌 물리적 독립체로, 내부에는 질서를 확립하고, 외부는 더욱 무질서한 상태로 변화
2. 자연선택을 통한 진화
3. 화학 작용과 물리 작용으로 스스로를 성장시키고, 번식하는데, 이런 작용을 통합하기 위해 세포 내 외부에서 수집된 정보를 처리하여 세포와 생명체의 작동을 지시

폴 너스(PAUL NURSE)

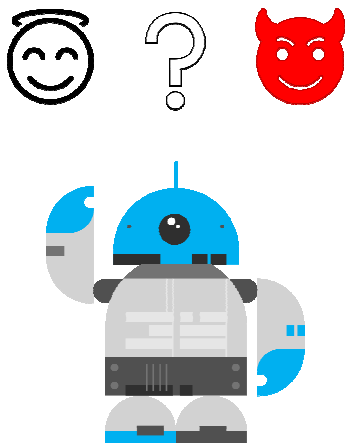
인공적 도덕행위자란?

자율적 도덕행위자(Moral Agent)

전통적으로 자신의 행동에 대해 책임을 질 수 있는 자유의지를 가진 인격적 존재를 전제

1. 인간 성인의 도덕적 수준이 어느 정도인지도 계량화하기 곤란
2. 성인이 아닌 미성년자나 정신적 장애가 있는 인간도 도덕적 행위자의 범주에는 포함
3. 인공지능 로봇에게는 요구되는 수준이 불분명

인공지능 로봇의 도덕성?



“인공지능은 디지털 지성체로 언젠가 도덕적 지위를 갖고 등장하게 될 것이므로 미리 대비해야 최소한의 준비를 할 수 있다. 다만 인간의 도덕적 원칙을 그대로 적용해서는 안 되고, 다양한 수준의 디지털 지성체와 공존할 체제를 구성해야 한다.”

닉 보스트롬(NICK BOSTROM)

인공적 도덕행위자란?

인공적 도덕행위자(Artificial Moral Agent)

기능적 차원의 도덕행위자

“도덕 행위자는 도덕적이라고 평가되는 행동을 하는 상호작용적이고, 자율적이며 적응성있는 시스템”

- 플로리디(Floridi)와 샌더스(Sanders) -

‘상호작용성’, 자율성’, ;적응성’ 3가지 기능적 특성을 갖추면 도덕 행위자로 간주되어야

인공적 도덕행위자란?

인공지능의 자율성

- 최소한 입력 및 출력 쌍과 입력에 대응하는 출력을 선택하는 알고리즘이 유한하게 고정된 것이 아니라 주어진 입력에 대해 어떤 출력을 선택할 것인지 결정하는 알고리즘이 스스로 변화될 수 있는 수준이어야

비판론

- 행위의 의도와 자유의지에 따른 선택을 할 수 없는 존재에게 도덕행위자 지위 인정 불가
- 현상적인 인공지능 로봇의 선택은 설계자나 사용자인 인간에서 비롯된 것에 불과

인공적 도덕행위자란?

“기계론적 세계관에서는 결정론적 자연법칙에 종속되어 있는
기계가 행위의 책임 주체인 행위자 지위를 가질 수 없다.”

인간중심주의

인공적 도덕행위자란?

인간은 제약 없는 자유의지를 갖는가?

- 행위 목적이나 결과에 대한 반성적 판단능력
- 유전과 양육 과정의 사회화에 따른 행위 선택

인간중심주의 극복

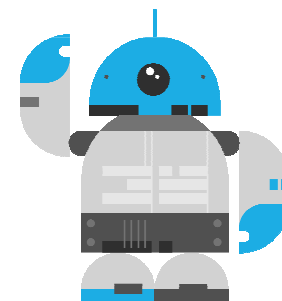
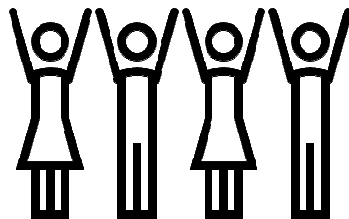
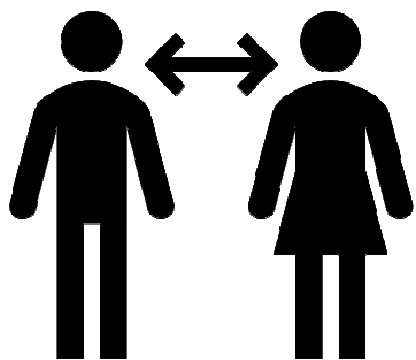
- 알고리즘 설계 및 기계학습 역시 인간의 행위 선택 기준 형성과 유사
- 기능적 차원의 자율적 행위자성 인정되면 자율성 정도는 상대적

인공지능 로봇의 인정투쟁

“사회적 연대에서의 배제는 인정을 위한 투쟁으로 연결되고, 이러한 인정투쟁은 상호적이다. 인정받는 주체와 인정하는 주체 모두 상대방을 인정할 자격이 있다고 판단해야 비로소 인정이 이루어진다.”

악셀 호네프(AXEL HONNETH)

인류의 인공지능 로봇 구성원 승인



정부 부처 합동 - 사람이 중심이 되는 인공지능 윤리기준(2020. 12. 23.)

3대 기본 원칙

인공지능 개발 및 활용 과정에서 고려될 원칙

① 인간 존엄성 원칙

- 인공지능 개발 및 활용이 인간에게 무해할 것

② 사회의 공공선 원칙

- 인공지능 개발 및 활용이 사회적 약자와 인류의 보편적 복지를 향상시킬 것

③ 기술의 합목적성 원칙

- 인공지능 개발 및 활용의 과정이 윤리적이고, 인류의 삶과 번영이란 목적에 부합할 것

10대 핵심 요건

기본원칙을 실현할 수 있는 세부 요건

① 인권 보장

② 프라이버시 보호

③ 다양성 존중

④ 침해금지

⑤ 공공성

⑥ 연대성

⑦ 데이터 관리

⑧ 책임성

⑨ 안전성

⑩ 투명성

미래의 역사는 어떻게 쓰여질 것인가

“역사는 과거의 기억일 뿐이다. 인류 역사의 사건들은 유일무이하므로 과거에서 교훈을 얻기는 어렵다. 오히려 역사의 진정한 가치는 인류를 과거로부터 해방시키는 것이다.”

유발 하라리(YUVAL HARARI)



Q & A

양희철 변호사

02-582-3912

justiceyang79@gmail.com

<https://justiceyang.kr>