

1st Korea Legal AI Conference(KoLAC)

2019 제1회 법률 인공지능 컨퍼런스

Legal AI의 현재와 미래 Perspective of Legal AI

2019. 8. 29 (목) PM 1:30~5:30

서울 서초동 변호사회관 5층 정의실/인권실

주최

 **한국인공지능법학회**
Korean Association for Artificial Intelligence and Law

 **사법정책연구원**
Judicial Policy Research Institute

후원

 과학기술정보통신부

 Microsoft



“Legal AI의 현재와 미래(Perspective of Legal AI)”

2019 제1회 법률 인공지능 컨퍼런스



| 일 시 | 2019년 8월 29일(목) PM 1:30 ~ 5:30
| 주 최 | 한국인공지능법학회·사법정책연구원
| 장 소 | 서울 서초동 변호사회관 5층 정의실/인권실(2호선 서초역 8번 출구)
| 후 원 | 과학기술정보통신부·마이크로소프트
| 컨퍼런스 프로그램 |

시 간	컨퍼런스
13:30 ~ 14:00	등록
14:00 ~ 14:10	개회사 - 강현중 사법정책연구원장 - 이상용 한국인공지능법학회장
14:10 ~ 14:20	축사 - 국회부의장 이주영 (자유한국당) - 국회의원 오신환 (바른미래당 원내대표) - 국회의원 김해영 (더불어민주당)
14:20 ~ 14:50	기조강연 : 4차 산업혁명과 인공지능 - KAIST 유창동 교수 (한국인공지능학회 회장)
14:50 ~ 15:40	주제강연 : Legal AI란 무엇인가 - KAIST 김병필 교수
15:40 ~ 16:00	Coffee Break
16:00 ~ 17:10	심층토론 : Legal AI, 도전인가 기회인가? 좌장 : 최경진 가천대 교수 토론 - 김정환 사법정책연구원 연구위원 - 유동균 법원행정처 정보화심의관 - 한애라 성균관대 법학전문대학원 교수 - 고환경 법무법인 광장 변호사 - 김형우 법무법인 지평 변호사 - 김본환 로앤컴퍼니 대표 - 홍정민 로스토토리 대표
17:10 ~ 17:30	시상 및 폐회

개회사

안녕하십니까. 사법정책연구원 원장 강현중입니다.

사법정책연구원이 한국인공지능법학회와 함께 「Legal AI의 현재와 미래」를 주제로 제1회 법률인공지능 컨퍼런스를 공동개최하게 된 것을 큰 영광으로 생각합니다. 또한 이번 행사의 성공적인 개최를 위해 물심양면 지원을 아끼지 않으신 이상용 한국인공지능법학회장님, 그리고 이 자리에 참석해 주신 모든 내외 귀빈 여러분께 감사의 말씀을 드립니다.



과학기술의 발전에 따른 사회의 변화는 법률업무에도 영향을 미치고 있습니다. 제가 처음 법조인 생활을 시작했을 무렵에는 컴퓨터에 의한 업무라는 것은 있지도 않았고 상상도 할 수 없었으나, 지금은 전자기록과 전자소송이 대세가 되어가고 있습니다.

그 시기와 속도가 문제이지 인공지능 역시 언젠가는 법률 영역에 도입될 수밖에 없습니다. 법률업무의 전산화는 그전에는 생각할 수 없었던 방대한 법률정보의 축적을 가능하게 했고, 축적된 법률정보는 인공지능에 의한 활용의 문을 열었습니다. 이미 전산화되어 축적되어 있는 많은 법률문서와 법률정보들이 인공지능에 의해 활용된다면 법조인의 업무에도 큰 변화가 올 것입니다.

IT 강국인 우리나라는 인공지능 법률시장에서도 앞서나갈 수 있는 충분한 저력이 있습니다. 이미 언론에 수차례 보도되었듯이 법률 인공지능에 대한 우리 사회의 관심도 높습니다. 이번 컨퍼런스는 인공지능 도입에 대한 학술적 논의뿐만 아니라 실제 경연까지 선보임으로써 AI의 이론과 실무를 두루 접할 수 있는 좋은 기회가 될 것입니다.

한국인공지능법학회·사법정책연구원
→ 2019 제1회 법률 인공지능 컨퍼런스
"Legal AI의 현재와 미래(Perspective of Legal AI)"

아무쪼록 이번 컨퍼런스를 통해 법률 인공지능 영역에 대한 탄탄한 이론적 토대와 상호 교류의 발판이 마련되기를 바라며, 사법정책연구원은 앞으로도 상호협력을 통한 연구역량 강화를 위해 계속 노력해 나갈 것을 약속드립니다.

마지막으로 오늘 컨퍼런스에 참석해 주신 내외 귀빈 여러분과 바쁘신 일정 중에도 흔쾌히 사회, 발표, 토론을 맡아 주신 전문가 분들, 그리고 이번 컨퍼런스를 기획하고 준비하신 모든 관계자 여러분께 다시 한 번 깊이 감사드립니다.

감사합니다.

2019년 8월 29일

사법정책연구원 원장 **강 현 중**



개회사

안녕하십니까. 한국인공지능법학회 회장 이상용입니다. 무더위에도 불구하고 찾아주신 귀빈들과 여러 참석자 분들께 감사 말씀드립니다.

존경하는 강현중 원장님께서 이끌고 계신 사법정책연구원과 저희 학회가 함께 개최하는 이번 컨퍼런스는 Legal AI라는 새로운 기술이 법률 서비스에 미치는 영향을 심도 있게 고찰하고 바람직한 대응방안을 모색하기 위하여 마련되었습니다.



선진 각국에서는 이미 Legal Tech가 하나의 산업으로 발전하여 미국의 경우 관련 기업의 수가 천 개를 넘어서고 시장 규모 역시 수십조 원에 이르고 있습니다. Legal AI는 적용되는 영역마다 법률 서비스의 생산성을 몇 배씩 끌어올림으로써 고비용 구조로 인하여 기존에 접근할 수 없었던 시장을 개척하는 동시에 새로운 융합 서비스를 탄생시키고 있습니다. 새로운 도구를 맞이한 변호사들의 업무와 일상 역시 빠른 속도로 바뀌고 있습니다.

안타깝게도 지금까지 우리나라는 이러한 흐름에 동참하지 못해 왔습니다. 법률가, 공학자, 기업은 물론 정부조차 한정된 자원을 투입하는 과정에서 Legal AI를 도외시한 탓에 연구 생태계가 만들어지지 못했기 때문입니다. 선진 각국이 발빠르게 앞서간 것은 수십년 동안 형성된 연구 생태계를 기반으로 한 것입니다. 하지만 늦었다고 생각할 때가 가장 빠른 법입니다. 변화를 직감하고 이 자리에 와 계신 분들이 씨앗이 될 수 있습니다. Legal AI에 대한 열정으로 은퇴를 앞둔 나이에 사법시험에 합격한 일본의 컴퓨터공학자도 있으며, 변호사로서 뒤늦게 컴퓨터 공부를 하여 Legal Tech 기업을 일으킨 분들도 부지기수입니다. 법률가와 공학자가 협업하고 정부와 기업, 연구자가 힘을 합치면 우리도 충분히 따라잡을 수 있습니다.

한국인공지능법학회·사법정책연구원
→ 2019 제1회 법률 인공지능 컨퍼런스
"Legal AI의 현재와 미래(Perspective of Legal AI)"

부디 이번 컨퍼런스를 계기로 우리나라에도 Legal AI의 뿌리가 내리고 생태계가 자리잡게 되기를 바랍니다. 마지막으로 바쁘신 중에도 흔쾌히 발표와 토론을 맡아주신 여러 전문가 분들과 세심하게 컨퍼런스를 준비해주신 실무자 분들께 다시 한 번 깊은 감사의 말씀을 드립니다.

감사합니다.

2019년 8월 29일

한국인공지능법학회 회장 **이 상 용**



축사

안녕하십니까,

국회부의장 이주영(경남 창원시 마산합포구 국회의원) 입니다.

먼저 대한민국에서 최초로 「법률인공지능 컨퍼런스」를 주최하게 된 것을 축하드립니다. 무더운 여름, 이처럼 뜻깊은 행사를 만들어주신 강현중 사법정책연구원장님, 이상용 한국인공지능법학회장님께 깊은 감사의 말씀을 드립니다.



얼마 전 한국을 찾은 손정의 소프트뱅크 회장은 인공지능에 투자해야 함을 강조하면서 “앞으로 한국이 집중해야 할 것은 첫째도 AI, 둘째도 AI, 셋째도 AI” 라고 말하셨습니다. 이 말은 인공지능이 미래사회의 핵심이 되고 있으며 인공지능을 선도적으로 이끌어가는 국가가 세계 기술 전쟁에서 승리하고, 경제대국으로 나아가는 첩경임을 웅변한 것이라고 생각합니다.

세계는 지금 인공지능 플랫폼 전쟁터라고 할 정도로 AI 기술 발전이 급변하고 있습니다. 우리나라도 알파고(AlphaGo)가 바둑기사 이세돌을 꺾은 이후 인공지능에 대한 관심이 커지고 있습니다. 그만큼 기술 발전과 정책적·제도적 지원, 그리고 인재확보까지 과제가 쌓이고 있다 하겠습니다.

2017년 영국에서 열린 ‘세계법률인공지능 경진대회’에서 우리나라의 ‘법률AI 알고리즘’이 우승했다는 소식을 전해들었을 때 놀라움을 감추지 못했습니다. ‘법률인공지능’ 분야에서 우리나라의 원천기술이 세계적으로 인정받을 정도로 성장했다는 사실이 오늘과 같은 법률인공지능컨퍼런스를 개최하게 된 배경이 되었다고 생각합니다.

미국은 여러 주에서 형사재판의 선고형을 정할 때 인공지능 알고리즘이 활용되고 있고, 북유럽의 에스토니아는 올해 4월, AI판사 제도를 도입하겠다고 선언했으며, 중국 하이난 고등법원의 형사법정에서는 AI를 활용한 양형 시스템이 사용되고 있다고 합니다.

이처럼 세계 인공지능 법률시장이 빠른 속도로 확대되고 있는 상황에 우리나라는 어디에 서 있는지 현재를 성찰할 필요가 있습니다. 우리나라가 세계 인공지능 전쟁에서 승리해 미래사회를 선도하기 위한 방안과 전망을 찾는 자리가 바로 오늘 법률인공지능 컨퍼런스라고 생각합니다.

한편 오늘 행사의 부대행사로 마련되었다는 법률AI팀과 인간변호사팀의 계약서 자문 대결 '알파로 경진대회'는 그 자체의 흥미를 떠나서 법률AI 알고리즘의 활용 가능성을 시험해 볼 수 있는 중요한 가늠자가 되리라고 봅니다. AI의 데이터 학습이 과연 인간의 종합적 지식과 판단능력을 뛰어넘을 수 있을지 많은 국민들이 결과를 주목할 것입니다.

특히 법률영역의 법령·판례의 양은 너무나 방대해, 사람의 두뇌로 소화할 수 없을 정도의 양인 것이 사실입니다. 따라서 수많은 데이터를 학습해 체계적으로 분류할 수 있는 AI의 장점이 발휘될 수 있을 것입니다. 그럼에도 소송이나 자문은 고도의 종합판단이 개입되는 분야인 만큼 AI가 인간을 완전히 대체하기는 현재로서는 어려울 것으로 봅니다. 하지만 인간과 기계가 각자의 능력을 살리면서 AI가 사람의 한계를 보완한다면 국민들이 법률서비스에 더 가까이 다가가고, 양질의 법률서비스를 받을 수 있는 토대가 마련될 것이라 생각합니다. AI는 두려움의 대상이 아니라 우리가 활용해야 할 무기라는 것입니다.

오늘 행사를 준비하신 여러 교수님, 판사님, 변호사님 등 전문가 여러분들의 고견을 소중히 수렴하여 국회 차원에서도 지속적으로 관심을 갖고 제도적·정책적 지원방안 마련을 위해 노력하겠습니다.

감사합니다.

2019년 8월 29일

국회부의장 이 주 영

축사

안녕하십니까. 바른미래당 원내대표 국회의원 오신환입니다.

오늘 법률인공지능 컨퍼런스라는 뜻깊은 자리에 초대해주셔서 감사드립니다. ‘법률’ 인공지능의 현재와 미래를 돌아볼 수 있는 이처럼 귀중한 행사를 마련해주신 사법정책연구원 강현중 원장님, 한국인공지능법학회 이상용 회장님께도 감사드립니다.



모두의 예상을 뒤엎고 세계 최고의 프로바둑기사 이세돌을 꺾은 구글의 딥마인드(Deepmind), 우리에게 친숙한 애플 아이폰의 시리(Siri), 아마존의 알렉사(Alexa) 등 인공지능은 이미 우리 생활의 일부분이 되어가고 있습니다.

미래의 경제전쟁은 인공지능 전쟁이 될 것이라고 미래학자들은 한 목소리로 말합니다. 얼마 전 손정의 소프트뱅크 회장이 우리나라를 방문해 “한국은 오직 인공지능에 투자하라.”고 강조한 것은 인공지능 기술에서 뒤처지는 국가는 경제발전에서 도태될 것이고, 인공지능 원천기술은 미래 국민들의 먹거리를 책임지는 핵심기술이 될 것이라는 점을 시사합니다.

AI는 데이터를 입력하면 이른바 ‘학습’을 하게 되고(Machine Learning), 수많은 데이터가 쌓이면 AI 알고리즘이 이를 체계화하고 분류, 분석해내며 사용자가 원하는 결과를 정확하게 도출해냅니다.

법률인공지능을 예로 들어 보겠습니다. 법률AI가 발전하게 되어 법률AI 알고리즘에 수많은 데이터가 쌓이게 되면 국민들은 신속하고 정확한 법률정보를 받아 볼 수 있습니다.

결국 국민들은 누구나 법률서비스에 손쉽게 접근할 수 있게 됨으로써 법률시장의 규모는 더욱 커지는 결과를 낳을 수 있습니다.

그런 의미에서 각종 규제를 해소함으로써 법률시장에 AI 기술이 가져올 수 있는 기회가 무엇인지 긍정적인 고찰이 필요한 때라고 봅니다.

스마트폰이 우리의 삶을 파괴하고 일자리를 빼앗아갔다고 어느 누구도 말하지 않습니다. 우리의 손에 작은PC가 들어오게 되면서 더 많은 기회가 만들어지고 있습니다.

삼성전자가 세계 스마트폰 시장에서 1위를 달리면서 수많은 일자리가 창출되고 인간의 삶 역시 풍요로워지고 있습니다. 인공지능의 시대를 맞아 일찌감치 미래를 준비 하려는 법률인공지능컨퍼런스의 개최가 더욱 뜻깊고 소중한 이유입니다.

4차 산업혁명시대를 맞아 새로운 도전과 기회가 될 이번 컨퍼런스 개최를 다시 한 번 축하드리며 바른미래당 원내대표이자 국회 법제사법위원회 간사로서, 법률인공지능의 발전에 힘을 보태고 아낌없이 지원하고자 노력하겠습니다. 감사합니다.

2019년 8월 29일

국회의원 오 신 환



축사

안녕하십니까? 더불어민주당 국회의원 김해영입니다. 오늘 국내 최초로 개최되는 ‘법률인공지능 컨퍼런스’를 축하드립니다.

하루가 다르게 급변하는 인공지능 기술 발전의 시대를 예비하며 ‘법률인공지능 컨퍼런스’를 준비하신 강현중 사법정책연구원 원장님, 이상용 한국인공지능법학회 회장님께 특히 감사드리며, 귀중한 행사에 초대되어 개인적으로도 뜻깊은 시간이 될 것 같습니다.



얼마 전 중국의 시진핑 주석은 ‘AI굴기’를 선언했습니다. 2030년까지 중국을 AI중심 국가로 만들고 이를 위해 인공지능 핵심 및 연관사업 규모를 10조 위안(약 1,650조원) 까지 키우겠다는 것입니다.

법률인공지능과 관련해서는 이미 중국 하이난 법원에서 AI를 도입해 AI가 판례를 분석하고 자연어 처리에 기반해 판결문 초안을 작성합니다. 미국은 이미 인공지능 알고리즘이 각 주의 법원에 도입되어 활약하고 있습니다. 범죄자의 재범가능성을 예측하는 알고리즘 COMPAS(Correctional Offender Management Profiling for Alternative Sanctions)가 대표적입니다.

전 세계가 이처럼 인공지능 기술발전에 집중하고 있는 이유는 인공지능 기술이 미래 산업의 성장 동력으로 여겨지고, 인공지능을 가장 잘 이해하는 인재들이 기업과 국가에 많은 역할을 할 것으로 예상되기 때문입니다.

세계 유수의 선진국들이 인공지능 기술발전과 지원에 매진하는 반면 우리나라의 관심과 지원은 아직 부족한 실정입니다. AI의 핵심인재들이 해외로 나가 연구하고 있으며, AI 기술발전의 핵심인 ‘데이터’의 확보는 개인정보 등의 문제로 아직 논의가 진행 중인 상황입니다. AI 기술발전을 위해 데이터를 공개하는 문제는 우리 사회의 여러 가지 가치를 고려하여 충분한 검토를 이어나가겠습니다.

한편 오늘 법률인공지능 컨퍼런스 행사의 부대행사로 열리는 ‘알파로 경진대회’는 법률AI의 발전과 실제로 법률 업무에 어떻게 적용되는지를 확인할 수 있는 자리가 될 것으로 기대됩니다.

이미 우리의 곁에 다가온 인공지능의 시대, 법률인공지능 컨퍼런스와 알파로 대회를 통해 미래사회를 대비하고 설계하시는 참가자님들께 격려의 박수를 보냅니다. 저 역시 법률제정이라는 역할을 담당하고 있는 국회의원으로서, 오늘 자리에 참석하신 각계 전문가들과 함께 인공지능으로 변화할 미래사회를 설계하고, AI가 인류에 유익한 역할을 할 수 있도록 최선을 다하겠습니다.

감사합니다.

2019년 8월 29일

국회의원 김 해 영

➡ CONTENTS

| 기조강연 |

4차 산업혁명과 공정한 인공지능 기술 개발 1

- 발표 : KAIST 유창동 교수 (한국인공지능학회 회장)

| 주제강연 |

Legal AI란 무엇인가 17

- 발표 : KAIST 김병필 교수

| 심층토론 |

Legal AI, 도전인가 기회인가? 69

- 토론① : 김정환 사법정책연구원 연구위원 / 69
- 토론② : 유동균 법원행정처 정보화심의관 / 73
- 토론③ : 한애라 성균관대 법학전문대학원 교수 / 77
- 토론④ : 고환경 법무법인 광장 변호사 / 81
- 토론⑤ : 김형우 법무법인 지평 변호사 / 85
- 토론⑥ : 김본환 로앤컴퍼니 대표 / 89
- 토론⑦ : 홍정민 로스터리 대표 / 93



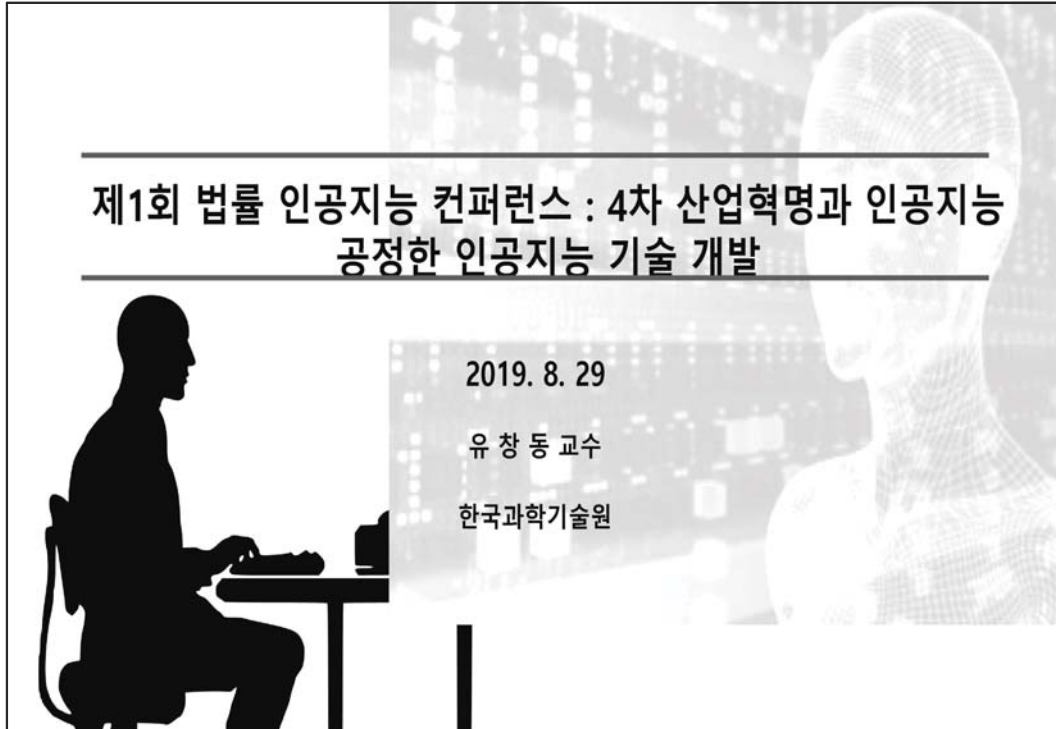
한국인공지능법학회 · 사법정책연구원
→ 2019 제1회 법률 인공지능 컨퍼런스
“Legal AI의 현재와 미래(Perspective of Legal AI)”

4차 산업혁명과 공정한 인공지능 기술 개발



| 기조강연 | KAIST 유창동 교수 (한국인공지능학회 회장)





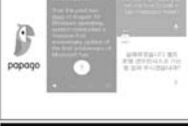
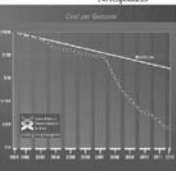
머릿말

산업혁명 시절의 기계발전과 정보화시대에 컴퓨터처럼 인공지능이 수천 개의 분야에서 인간 생활의 각 영역에서 큰 도움을 줄 것이라는 소식이 전해지고 있다. 사람이 기계와 음성으로 이야기하고 운전사업도 자동차가 주변 교통상황을 인식하면서 주행을 하고 있으며 세계 어디를 가더라도 기계 번역기가 여행자에게 그 나라 언어로 소통할 수 있도록 도와주며 심지어 우리가 저렴한 가격으로 자신의 DNA를 인공지능으로 분석하여 맞춤형 의료서비스를 받을 수 있도록 해주고 있다. 병리학과 의료영상에서도 인공지능이 도입되어 다양한 질병을 진단하고 치료에 관한 자문을 받을 수 있도록 하고 있다. 놀라운 사실은 인공지능이 사람을 보호할 뿐 아니라 사람을 평가하기 시작했다는 것이다. 인공지능이 우리의 생활을 편리하게 하면서 우리의 의존도가 높아지고 있는 반면에 인공지능의 위험성과 부작용도 높아지고 있다. 인공지능은 부득이 대량의 데이터를 사용하고 있으므로 대량데이터에서 오는 편향한 학습이 취약계층에게 피해를 줄 수 있다는 점을 우리는 유의해야 한다. 궁극적으로는 **공정(Fair)하고 책임(Accountable)을 물을 수 있고 투명한(Transparent), FAT 인공지능 기술 개발**해야 한다.

2

인공지능 활용 분야

- 인공지능은 사회 전반에 걸쳐 다양한 분야에 활용되고 있음

인공지능 비서 (Virtual Assistant)	자율 주행 (Self-driving Cars)	전문가 시스템 (Expert Systems)	정밀 의료 (Precision Medicine)
   	    	  	  

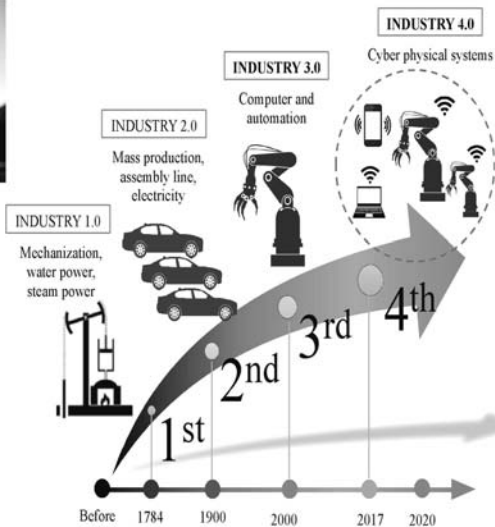
3

The 4th Industrial Revolution

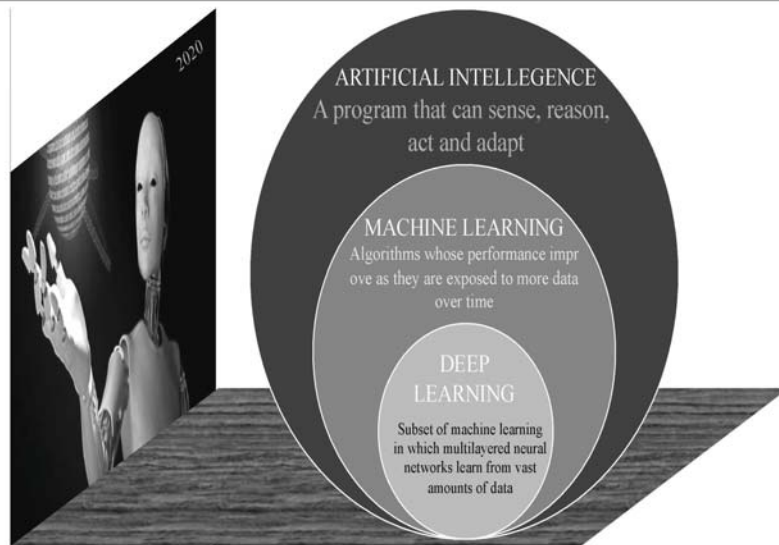


Klaus Schwab

" The Fourth Industrial Revolution refers to the fusion of technologies across the physical, digital and biological worlds which is creating entirely new capabilities and dramatic impacts on political, social economic systems. "



A.I., Machine Learning, and Deep Learning

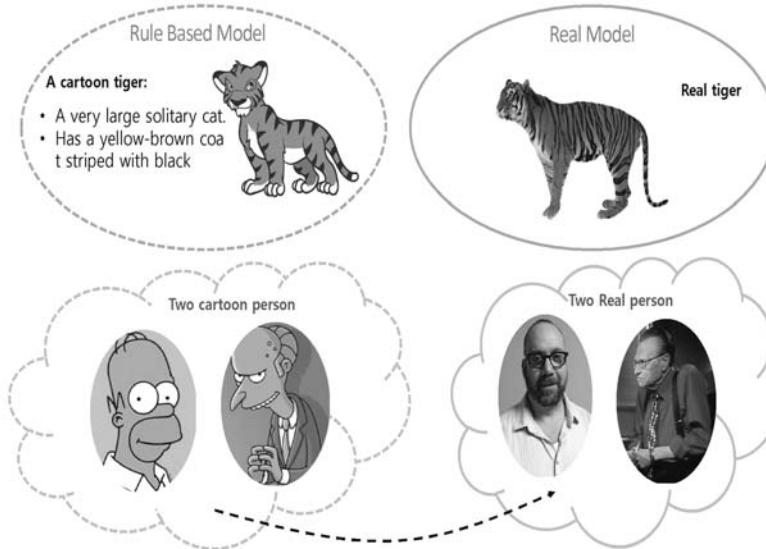


Innovation

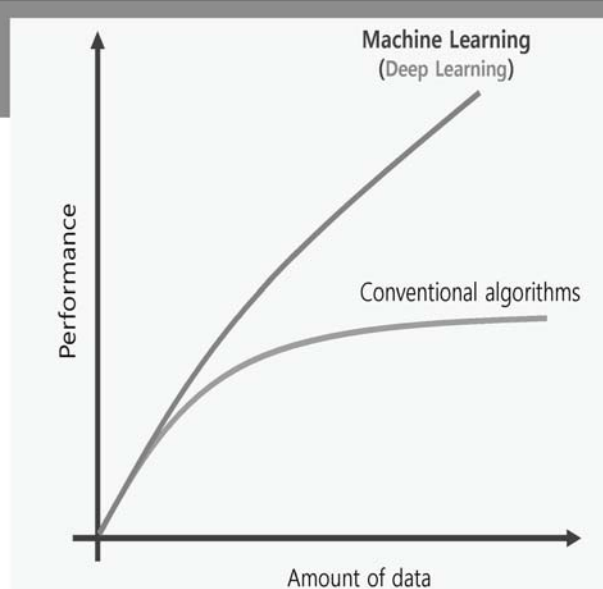


Limitation of Rule-Based Model

Knowledge based model is always going to be insufficient in modeling real world.



Machine Learning & Big Data



Advantages of Machine Learning



Performance rises with more data



Constructed to perform well in real world



Single algorithm can service many tasks

Weak A.I. vs. Strong A.I.

Chat boxes Weak A.I.



- Weak AI only claims that machines can act intelligently
 - Machine intelligence need only mimic the behavior of human intelligence
 - Not intended to match/exceed the capabilities of human beings
- Current AI

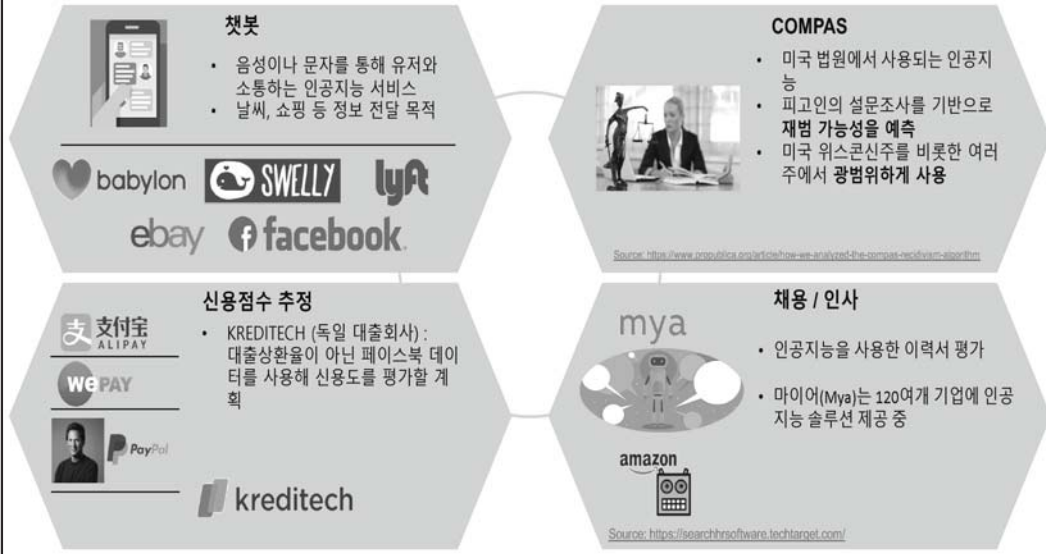
Strong A.I.



- Strong AI claims that a machine that acts intelligently also has mind and understands in the same sense people do
- Given the appropriate programming a computer actually becomes a mind
- AI that matches or exceeds human intelligence

인공지능 활용 분야

- 인공지능은 사회 전반에 걸쳐 다양한 분야에 활용되고있음



연구의 필요성

- 편향된 데이터 / 인공지능으로 인한 문제점



- 사용자와의 대화를 훈련데이터로 사용해 학습
- 편향된 데이터로 인해 성 / 인종차별 발언
- 16시간 만에 운영 중지

출처: 연합뉴스TV, 2016년 3월 24일 뉴스



16시간만에 운영 중지

연구의 필요성

Equivant의 미국 법원 AI 'COMPAS'



- 피고인의 설문조사를 기반으로 재범 가능성을 예측
- 미국 위스콘신주를 비롯한 여러 주에서 광범위하게 사용
- ProPublica(인터넷언론)에서 인종편향성에 대한 문제 제기
- 실제 사례 조사에서 재범가능성 높다고 예측했으나 실제 재범 안 한 비율이 흑인 44.9%, 백인 23.5%로 나타남

아마존의 AI 채용 시스템



- 지원자의 이력서를 분석 / 평가
- 지난 10년간 아마존에 제출한 이력서를 훈련데이터로 사용
- 학습된 모델은 성차별적인 결과를 주었음

구글의 구글포토



- 구글의 객체인식 기반 자동 태깅 시스템
- 흑인 얼굴을 고릴라로 잘못 인식하는 문제가 발생하여 사회적으로 이슈화
- 훈련용 이미지 데이터가 백인 위주로 편향됐을 가능성 존재

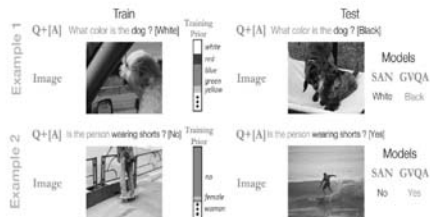
아마존의 이미지 분석 AI 'Rekognition'



- 이미지 분석 시스템으로 성별 분석 기능 제공
- 성별, 인종에 따라 최대 35배 까지 에러율 차이
 - 백인남성 : 1% 에러
 - 흑인여성 : 35% 에러
- IBM, MS, Megvii, Face++ 시스템에서도 비슷한 경향을 보임

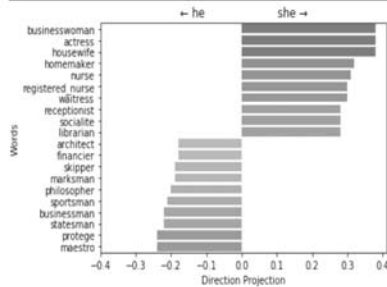
연구의 필요성

조지아 공과대학의 'VQA-CP'



- VQA(Visual Question Answering)는 이미지에 대한 질문에 답을 찾아내는 문제
- 기존의 VQA 데이터셋에서 편향성이 확인됨 - 질문의 종류에 따른 답변 분포의 편향성
- 조지아 공과대학에서는 인공지능의 편향성을 제거하기 위한 Adversarial training 기법과 인공지능의 편향성을 측정할 수 있는 VQA-CP(VQA under Chaining Priors) 데이터를 발표

마이크로소프트의 'Gender-biased word embedding'



- Word embedding은 단어를 일정한 길이를 갖는 벡터의 형태로 표현해주는 기술(Word2Vec, GloVe 등)
- 대부분의 언어 관련 인공지능에서 사용되므로 파급력이 굉장히 큼
- 마이크로소프트에서 개발한 모델에서 성 편향이 확인 됨

공정성을 위해 고려되어야 할 속성



공정한 사회 (소수집단의 사회기여)



국외 연구 현황

- 국외에서는 IT / AI 대기업 중심으로 공정성을 위한 연구 개발 수행 중

IBM – Fairness AI 360



- ✓ 공개SW
- ✓ 편향성 검출
- ✓ 편향성 제거

IBM – AI OpenScale



- X 공개SW
- ✓ 편향성 검출
- ✓ 편향성 제거

- 편향성 검출-완화-제거를 위한 알려진 여러 알고리즘을 적용 및 비교 분석할 수 있음(Python open source framework)

Google – What-if



- ✓ 공개SW
- ✓ 편향성 검출
- ✓ 편향성 제거 (후처리)

- 기계학습 개발자를 위한 오픈소스 시각화 도구(tensorflow web application)
- Group unaware, group thresholds, demographic parity, equal opportunity, equal accuracy의 다섯 가지로 범주화 된 공정성 측도에 대한 분석 / 시각화 제공

MS – FATE / Facebook – Fairness Flow



- ✓ 연구 수행 중

- MS의 FATE 연구 그룹은 다양한 분야에서 논문 출판
- Facebook은 f8 conference에서 Fairness AI 연구 진행 언급

국내 연구 현황

- 국외의 활발한 연구와 달리, 국내에서는 제한적으로 국외 연구를 추적하는 방향으로 진행

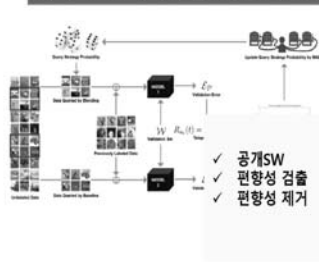
코난테크놀로지 – KONAN Analytics 4



- X 공개SW
- ✓ 편향성 검출
- X 편향성 제거

- 대규모 뉴스 데이터의 편향성 판단 / 분석 알고리즘을 개발함
- 편향성을 직관적으로 평가 / 분석할 수 있는 시각화 도구를 개발함

카이스트 – 데이터 불균형 완화 연구



- ✓ 공개SW
- ✓ 편향성 검출
- ✓ 편향성 제거

- 데이터의 불균형 분포로 인한 학습정확도의 편향성 완화 기술

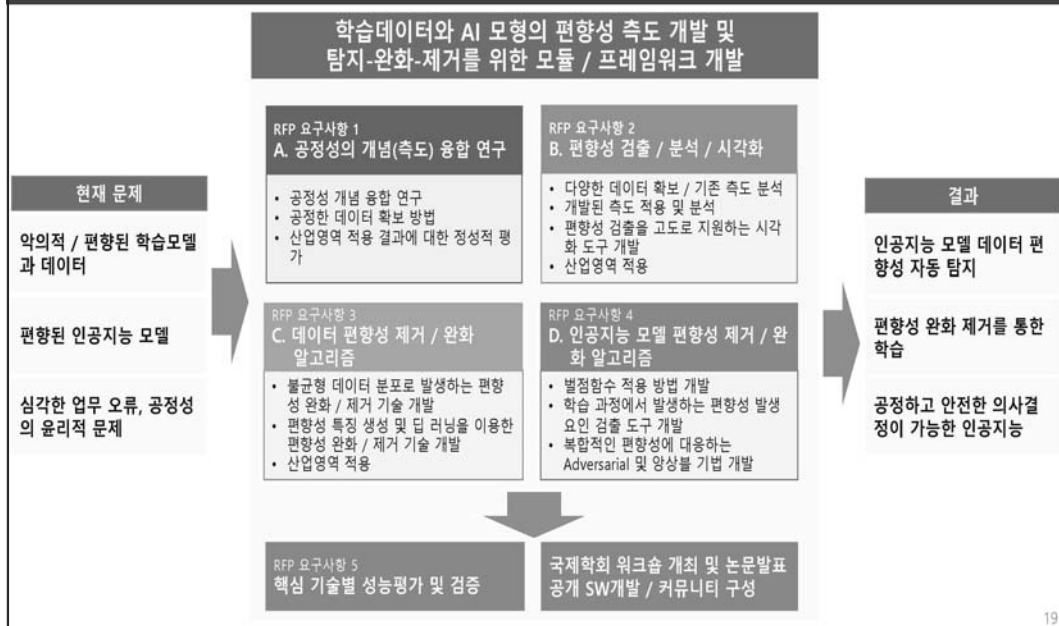
삼성 – PAI 가입



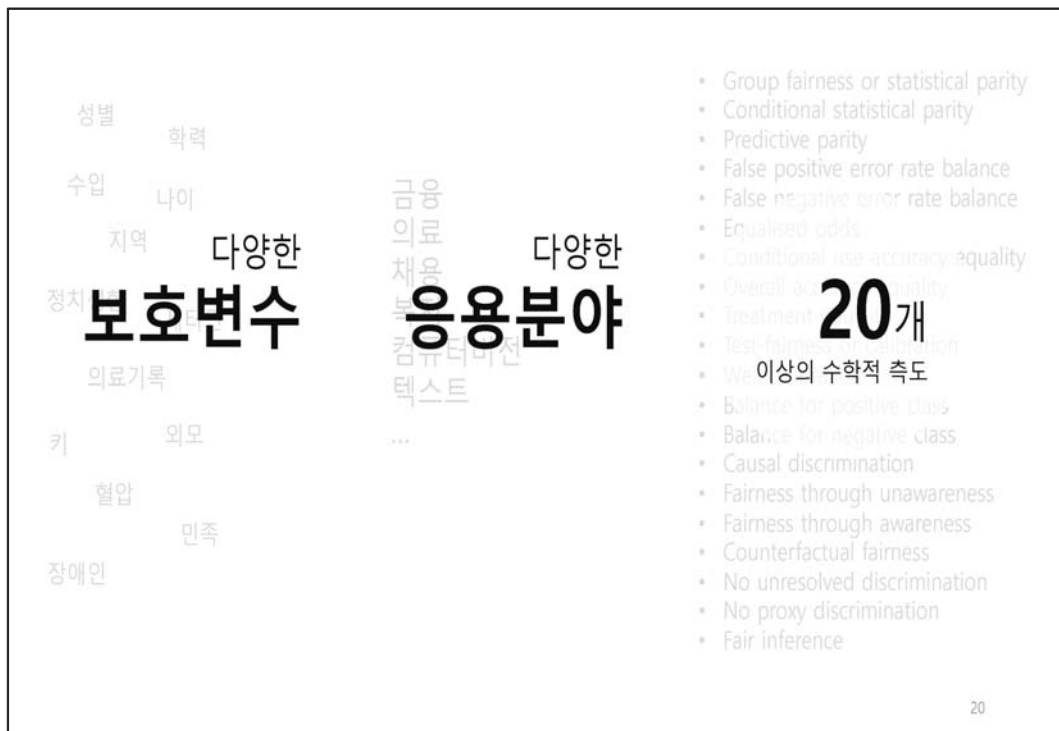
- ✓ 연구 수행 중

- 아마존, 구글, MS등 70여개 글로벌 기업이 참여하는 PAI에 가입, 인간과 AI의 협력 분야에 연구 시작

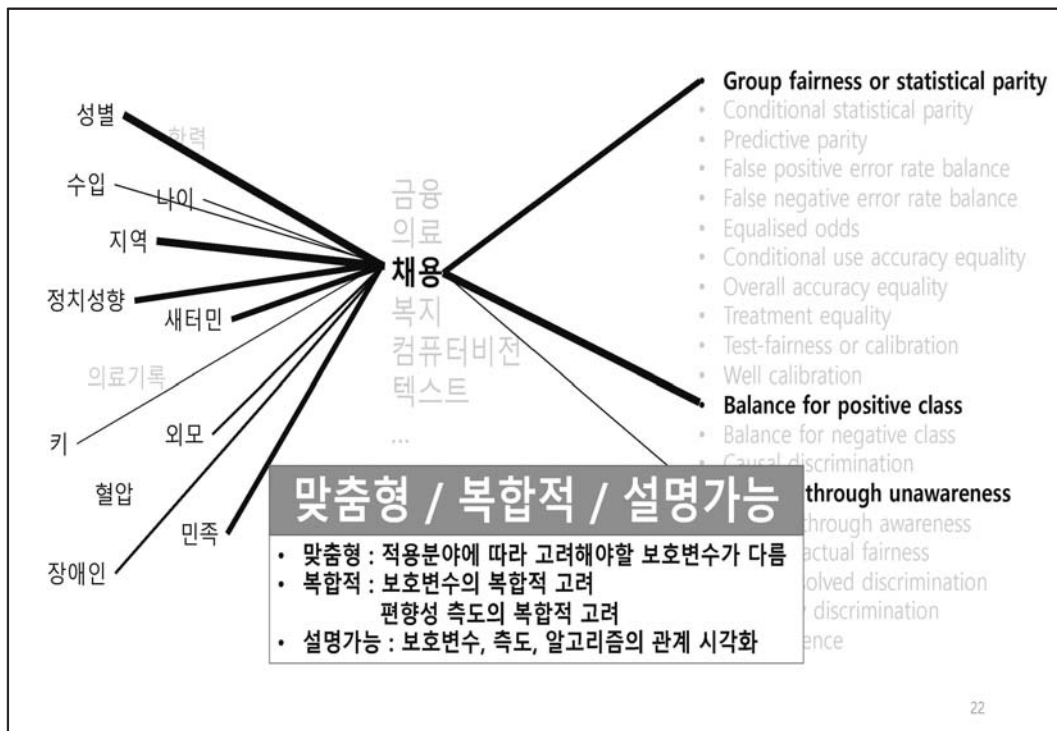
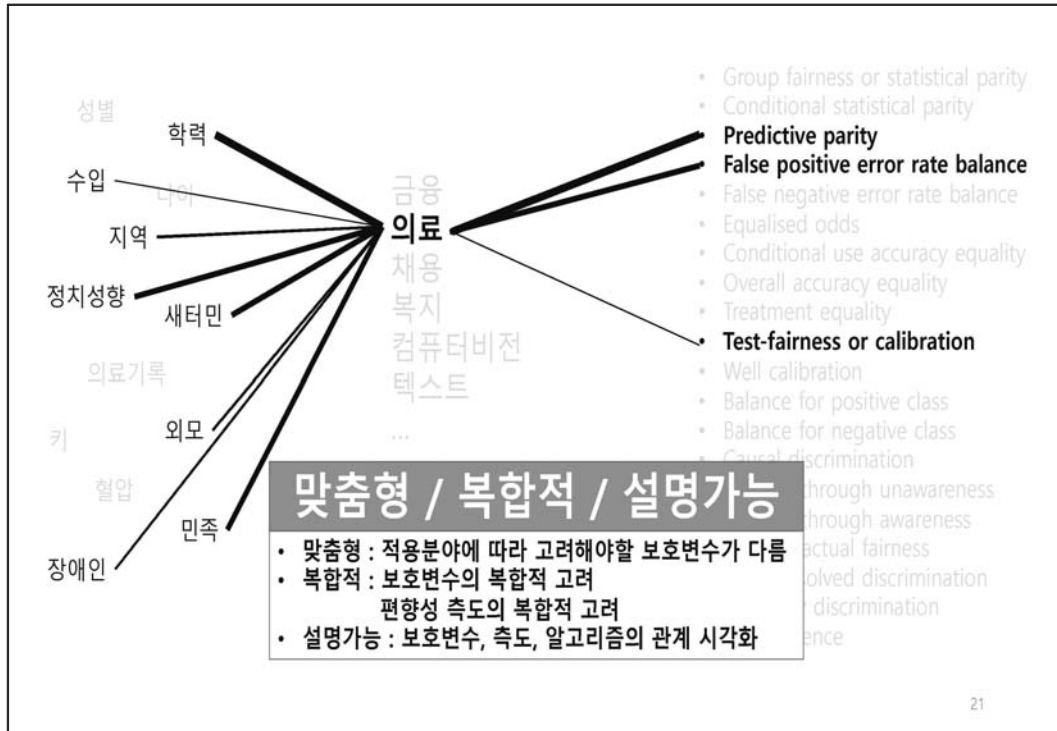
연구 목표



19

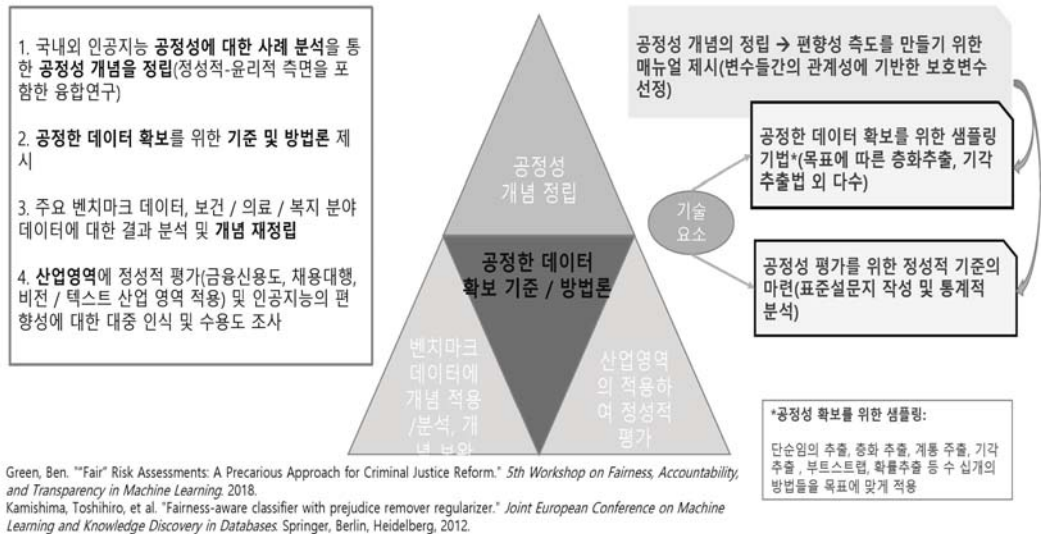


20



세부 연구 내용 - 공정성의 개념(측도) 융합 연구

• 공정성 개념에 대한 융합 연구



"The development of full artificial intelligence could spell the end of the human race...It would take off on its own, and re-design itself at an ever increasing rate. Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded." -- Stephen Hawking told the BBC



"Artificial intelligence would be the ultimate version of Google. The ultimate search engine that would understand everything on the web. It would understand exactly what you wanted, and it would give you the right thing. We're nowhere near doing that now. However, we can get incrementally closer to that, and that is basically what we work on." —Larry Page



"AI is a *fundamental risk to the existence of human civilization*" and poses *'vastly more risk' than North Korea*. If you're not concerned about AI safety, you should be. Vastly more risk than North Korea, according to Musk.



"AI can be our friend"

결론

- 국내외 인공지능 공정성에 대한 사례 분석을 통한 **공정성 개념을 정립** (정성적-윤리적 측면을 포함한 융합연구)
- 불균형 데이터로 생기는 편향성을 영역에 적절한 공정성 기준으로 **데이터 편향성 제거/완화** 하는 기술
- 보호 변수에 영향을 받지 않고 **독립된 결정이 가능한 인공지능 연구**
- **공정(Fair)하고 책임(Accountable)을 물을 수 있고 투명한(Transparent), FAT 인공지능 기술 개발**

25

