



연속토론회

모두의 AI : 정책과 전략

제1차 AI 사회정책

일 시 2023년 11월 24일 (금) 14시

장 소 국회의원회관 제2간담회의실

소주제 AI를 통한 불평등 격차 축소, 기후위기 대응,
AI 신뢰성 제고 방안 모색



「모두의 AI : 정책과 전략」 연속토론회 소개

■ 목적

- AI 시대의 도래에 따른 산업·사회 등의 변화에 적극 대응하기 위해 AI 기술의 영향과 활용방안을 토론하고 현장의 목소리와 함께 민생과 관련한 ‘정책·입법과제’를 제언하는 연속토론회 추진
- 더불어민주당 국회의원, 시민사회단체, 학계, 관련분야 전문가 중심으로 AI 시대와 기술을 통해 불평등 격차 해소, 기후위기 대응, 정당 역량강화, 산업 경쟁력 강화 등 AI 활용 방안과 해법 모색
- 김대중정부 IT관련 벤처기업 활성화, 노무현정부 IT산업 육성을 통한 정보화기반 지식강국 실현, 문재인정부 IT강국에서 AI강국을 지향한 ICT산업 활성화 정책 등 민주당 정부의 신성장동력으로 AI 산업을 경제 활성화 추진전략으로 계승 및 AI시대 주도 의지 표명

■ 행사 개요

- 시 기 : 2023년 11월 하순 ~ 2024년 1월 중순(3개월)
- 연속토론회(’23.11.24. ~ ’24.1.10., 총 3회)
- 장 소 : 국회의원회관 제2, 8간담회의실
- 행사명 : ‘모두의 AI : 정책과 전략’ 연속토론회
- 주 최 : 민주연구원
- 참 석 : 더불어민주당 국회의원 및 보좌진, 분야 관계자 및 전문가 등

■ 주제

- (제1차) AI 사회정책
- AI를 통한 불평등 격차 해소, 기후위기 대응, AI 신뢰도 제고방안 모색
- (제2차) AI 정당
- 당원을 위한 AI 정책보좌관, 정당업무 효율화, 데이터에 입각한 정책 전략 수립
- (제3차) AI 산업비전
- 초거대AI 기업 육성, AI-X 경쟁력 강화를 위한 특화전략, AI-ESG를 통한 기업의 사회활동



「모두의 AI : 정책과 전략」 연속토론회 소개

■ 개요

시간		순서		주요내용
~14:00		사전등록		
14:00~14:15	15분	개회식	국민의례	진행 사회(이연희 민주연구원 상근부원장)
			축 사	◦홍익표 더불어민주당 원내대표
			인사말	◦정태호 더불어민주당 민주연구원 원장
14:15~14:20	5분	토론회	장내정리 인사말	◦좌장: 이인영 국회의원
14:20~15:05	45분		발제	발제1 ◦주제: AI 디바이드 극복을 위한 모두의 AI 전략 ◦발표: 이원재 LAB2050 시민참여인공지능포럼 운영위원장
				발제2 ◦주제: 기후 테크와 AI ◦발표: 김성하 경기연구원 AI혁신정책센터장
				발제3 ◦주제: 인공지능 신뢰성에 관한 제언 ◦발표: 최재식 KAIST 김재철AI대학원 교수
15:05~15:35	30분		지정토론	◦주제: 기후테크 투자와 스타트업, 그리고 과제 ◦발표: 유서영 소풍벤처스 기후네트워크TF 팀장
				◦주제: 모두를 위한 AI 토론문 ◦발표: 구본권 한겨레 사람과디지털연구소장
				◦주제: 인공지능 신뢰성 확보를 위한 정책 제언 ◦발표: 김윤명 한림국제대학원대학교 교수
15:35~15:50	15분			



Contents

「모두의 AI : 정책과 전략」연속토론회

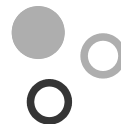
◦ 모두의 AI : 정책과 전략 연속토론회 소개	i
◦ 모두의 AI : 정책과 전략 연속토론회 제1차 프로그램	ii
◦ 목차	iii
◦ 축사	v
◦ 인사말	vii

발 제

1. AI 디바이드 극복을 위한 모두의 AI 전략	1
이원재 LAB2050 시민참여인공지능포럼 운영위원장	
2. 기후 테크와 AI	19
김성하 경기연구원 AI혁신정책센터장	
3. 인공지능 신뢰성에 관한 제언	31
최재식 KAIST 김재철AI대학원 교수	

토 론

1. 기후테크 투자와 스타트업, 그리고 과제	41
유서영 소풍벤처스 기후네트워크TF 팀장	
2. 모두를 위한 AI 토론문	59
구본권 한겨레 사람과디지털연구소장	
3. 인공지능 신뢰성 확보를 위한 정책 제언	65
김윤명 한림국제대학원대학교 교수	



축사

홍 익 표 원내대표
더불어민주당



안녕하십니까. 더불어민주당 원내대표 홍익표입니다. <모두의 AI : 정책과 전략> 연속토론회 개최를 진심으로 축하드립니다. 토론회 개최를 위해 애써주신 민주연구원 정태호 원장님과 관계자 여러분 고맙습니다. 좌장 이인영 의원님과 발제와 토론을 맡아주신 분들께도 감사의 말씀을 드립니다.

디지털 혁명은 코로나 시대를 거치면서 가속화되고 있습니다. AI 기술이 불러온 놀라운 변화는 이제 흔한 일상이 되었습니다. 눈앞의 스마트폰을 바라보면 얼굴을 인식해 화면 잠금을 해제해주고, 장바구니에 넣어둔 채 구매를 망설이던 상품을 자동으로 결제해줍니다. 우리는 AI가 안내하는 길을 따라 차를 몰고, AI가 골라주는 음악을 흥얼거립니다. 우리 일상은 이미 AI로 가득합니다.

산업구조 역시 대변혁을 예고하고 있습니다. 제조업 분야의 생산 공정에서 불량품을 찾아내고, 생산 공정을 관리하는 주체도 AI인 시대에 살고 있습니다. 세계는 이미 로봇, AI가 연계되는 첨단 산업 분야 무한경쟁에 돌입했습니다. 미국, 일본, 독일 등 로봇, AI가 연계되는 공장을 앞다퉈 건설하고 있습니다. AI 대응이 늦어진다면, 그 시간만큼 우리가 가질 세계 경쟁력도 늦춰질 수밖에 없습니다.

전면적인 디지털 혁명에 우리 정치도 능동적으로 대응해야 합니다. 기후 위기, 인구 위기, 지역 소멸 등 AI를 활용할 수 있는 방법은 이미 모색되고 있습니다. 대변화에 맞게 AI와 관련된 민생 정책 개발과 입법 과제 발굴도 뒷받침되어야 합니다. 그러나 여전히 AI의 권리와 책임, 데이터 보호와 활용 등에 대한 명확한 기준이 부족한 상황입니다.

민주당은 국민 누구나 골고루 혜택을 누릴 수 있는 ‘모두를 위한 AI시대’를 선도하겠습니다. AI 기반 산업이 더욱 활성화되고 세계 시장에서 경쟁력을 갖출 수 있도록 법과 제도도 살피겠습니다. 민주연구원의 연속토론회 개최를 진심으로 축하드리며, 참석하신 모든 분의 건강과 행복을 기원합니다.

감사합니다.

2023.11.24.

1더불어민주당 원내대표 **홍 익 표**



인사말

정 태 호 원장
민주연구원



안녕하십니까?

더불어민주당 민주연구원장 정태호입니다.

‘모두의 AI : 정책과 전략’ 연속토론회 제1차 AI 사회정책 분야 토론회에 참석하신 많은 분들께 감사의 말씀을 드립니다. 민주연구원이 주최하는 이번 연속토론회는 ‘AI’와 관련한 사회정책, 정당, 그리고 산업정책 의제들에 대해 토론하고 ‘정책입법과제’를 제언하는 방식으로 진행됩니다.

김대중 정부, 노무현 정부, 문재인 정부를 돌아해보면, 민주당은 늘 미래지향적인 아젠다를 주도해왔다는 자부심이 있습니다. 김대중 대통령께서는 ‘우리나라의 미래는 IT와 문화다’라고 하시며 국민PC, 초등학교 컴퓨터 교육, 초고속인터넷 도입 등 과감한 정책을 시행해 IT산업의 초석을 구축했습니다.

노무현 대통령께서는 정보통신 강국으로의 진입과 벤처육성을, 문재인대통령께서는 IT산업 우위를 바탕으로 전기차, 자율주행자동차, 신재생에너지, 인공지능 등 핵심기술 분야에 적극 투자하며 벤처강국으로 위상을 높였습니다.

이제 우리는 AI시대라는 피할 수 없는 흐름 앞에 서 있습니다. 갈 수 밖에 없는 길이라면 누구든지 보편적으로 활용할 수 있는 환경을 만들어야 합니다. 김대중 정부의 선견지명으로 디지털 격차를 줄였듯이 이제는 AI가 가져올 격차, 즉, AI 활용 유무에 따라 심화되는 불평등, AI기술로 인한 일자리 감소 등에 어떻게 대처할 것인가. 다양한 과제에 직면해 있습니다.

우리 삶의 일부가 되어가고 있는 AI 기술은 우리 사회에 윤리적·사회적 문제를 수반하지만 불평등 해소, 기후문제 해결 등 공공과 민간분야에서 다양한 역할을 수행하고 있습니다. 인공지능 기술을 개발하고 사용하는 과정에서 이해와 책임이 중요합니다. 신뢰할 수 있는 AI와 누구에게나 공평한 기회를 주는 AI 시대를 만드는 것이 정부와 국회의 역할일 것입니다.

오늘 토론회는 AI의 미래 전망을 제시하고, 정부·기업·학계가 함께 지식을 공유하고 AI 기술의 발전과 사회적 가치 이익을 조화시키는 방안에 대해 논의가 이뤄질 것으로 기대됩니다.

오늘 이 자리가 더불어민주당 국회의원과 시민사회관계자, 학계 및 분야 전문가들이 함께 모여 AI 정책들을 제안하고 논의하는 정책 소통의 장이 될 것입니다. 세밀한 정책 설계와 집행이 이뤄질 수 있도록 입법 활동을 추진하고 총선공약으로 제안할 것입니다.

바쁜 국회일정과 당무에도 불구하고 AI 정책에 대한 많은 관심을 가지고 참석해주신 홍익표 원내대표님께 다시 한 번 감사드립니다.

좌장을 맡아주신 이인영 국회의원님과 발제해주신 이원재 LAB2050 시민참여인공지능포럼 운영위원장님, 김성하 경기연구원 AI혁신정책센터장님, 최재식 KAIST 김재철AI대학원 교수님, 토론해주신 유서영 소풍벤처스 기후네트워크TF 팀장님, 구본권 한겨레 사람과 디지털연구소장님, 김윤명 한림국제대학원대학교 교수님께도 다시 한 번 감사드립니다.

고맙습니다.

2023.11.24.

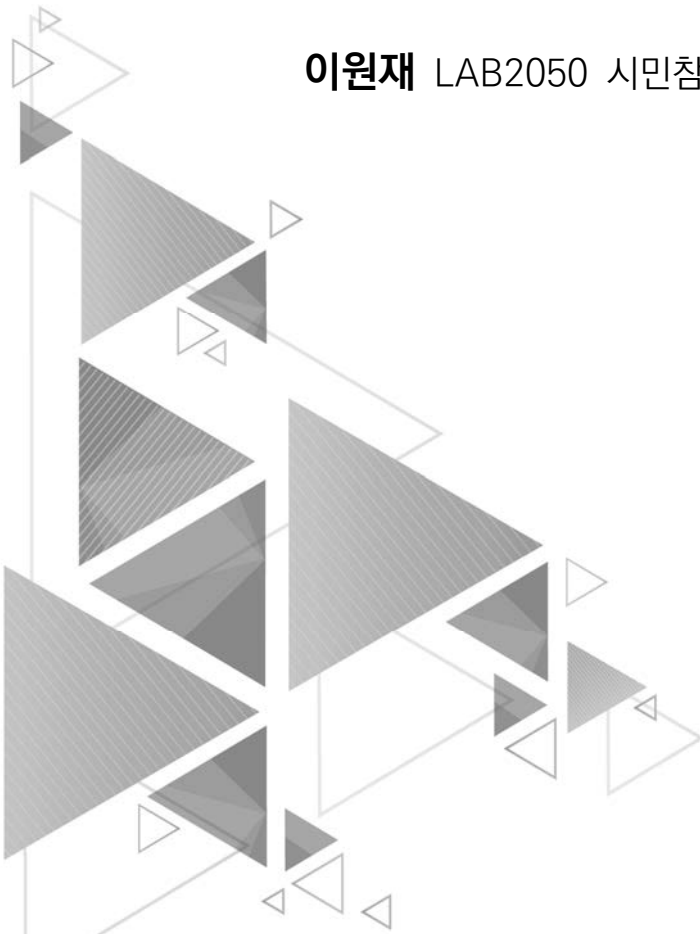
민주연구원 원장 정 태 호



발제 1

AI 디바이드 극복을 위한 모두의 AI 전략

이원재 LAB2050 시민참여인공지능포럼 운영위원장



‘AI 디바이드’ 극복을 위한 ‘모두의 AI’ 전략¹⁾

이원재 LAB2050 시민참여인공지능포럼 운영위원장

요약

이 발표의 목적은 챗GPT 등 생성형AI가 가져올 사회 변화에 대응할 정책담론의 프레임을 제시하는 것이다. 사회적 영향 측면에서 볼 때, AI는 분명한 강약점을 갖고 있다. 인간의 지적 능력을 확장해 생산력을 증대시킬 것이 분명한 반면, 독과점과 격차를 확대한다는 문제점도 있다. 상상하기 어려운 정도의 생산성 향상이 일어나면서, 생산과정의 구조적 변화로 사회경제적 격차가 커질 수 있기 때문이다. ‘생산력 증대’와 ‘격차 확대’는 AI기술의 사회적 명암이다.

이 발표에서는 격차 확대 문제를 ‘AI 디바이드’로 정의하고, 우리 사회가 이 문제를 극복할 전략으로 ‘신뢰할 수 있는 보편적 AI’ 개념의 확립, 기본소득제 등 전환적 사회보장정책 도입, 기후AI에 대한 투자, 기술기업과 협력해 AI주권 정립 등을 제시한다. 국회와 정부는 ‘AI 디바이드’를 기회로 바꿀 비전에 합의하고, 여기 투입할 재원을 AI기술의 수혜자들로부터 마련하는 방안을 우선적으로 논의해야 한다.

목차

1. ‘AI 디바이드’가 온다
2. AI의 정의와 특성
3. SWOT분석: AI의 사회적 영향
4. 대응 전략
5. 누가 지출할 것인가

1. ‘AI 디바이드’가 온다

AI는 놀라운 생산성 향상 도구다. 생성 AI는 더욱 그렇다. MS오피스 코파일럿, 깃허브 코파일럿, 챗GPT의 대화 등 그동안 인간 고유의 것으로 여겨지던 화이트칼라 노동, 지식노동에 엄청난 파장을 몰고 오고 있다.

1) 본 원고는 완결된 원고가 아니며, 2023년 11월 24일 민주연구원 주최 ‘모두의 AI: 정책과 전략’ 토론회의 구두 발표를 위한 초고이므로 전제 및 인용을 금합니다.

생산성 향상의 이면에는 커져가는 격차가 숨어 있다. 물론 디지털 전환 초기부터 경제적 격차는 커지고 있었다. 컴퓨터와 스마트폰 사용이 대중화되며 전세계의 소득과 자산 격차는 커졌다.

이 시기 격차 확대 양상은 이전 시대와는 다르게, 상위 10%의 약진이 두드러졌다. 0.1% 초 거대부자와 그 이하 계층의 격차도 물론 커졌지만, 기술활용능력이 있는 10%의 지식엘리트 계층과 그 이하 계층 사이에도 그에 못지 않게 골이 패였다. 극소수 최상위 계층, 또는 귀족 등 계급적 특권층에게 부가 물리던 이전 시기와는 다른 모습이었다.

이전의 격차가 세습된 계급 또는 형성한 자산의 격차였다면, 디지털 시대의 격차는 상당부분 능력 격차이기 때문이다.

‘정의란 무엇인가’를 썼던 마이클 샌델은 근간 ‘공정하다는 착각’에서 이 문제를 지적한다. 이제 불평등 문제는 자산격차의 문제에서 지식격차의 문제로 넘어갔고, 하버드대 등 명문대 출신의 지식엘리트 계층이 그 지식과 능력을 세습하는 문제가 일어나고 있다고 지적한다. 이를 극복하기 위해 명문대 입학 자격을 추천으로 가려야 한다는 파격적 주장을 하기도 한다.

AI 시대에 이런 지식과 능력 격차는 더욱 커질 것으로 보인다. 해법 찾기는 더욱 복잡하고 어려워질 것이다.

인터넷과 모바일 혁명까지를 포괄하는 ‘디지털 전환’과 비교해 보자. 디지털 전환 초기 우리는 기술로 인한 격차를 ‘디지털 디바이드’라고 불렀다. ‘디지털 디바이드’는 디지털 기기 활용 여부가 격차의 원인이었다. 데스크톱 컴퓨터와 스마트폰이 생산성을 높이는 도구였다.

이 문제의 해법은 상대적으로 단순했다. 빌 게이츠 마이크로소프트 창업자는 저개발국의 어린이들에게 컴퓨터를 보급하는 자선사업을 펼쳤다. 우리나라를 포함해 세계 각국 학교에서는 학교에 컴퓨터를 비치하고 태블릿 피서를 학생들에게 나눠주기도 했다. 도구를 활용할 수 있도록 해주면 격차 문제가 해소된다는 생각이었다.

하지만 AI시대 격차 문제는 좀더 복잡하다. 이전까지의 기술이 단순히 기술을 가진 사람과 갖지 않은 사람 사이의 격차를 벌려놓았다면, AI기술은 기술을 활용할 능력을 이미 보유한 사람과 그렇지 않은 사람 사이에 이미 있는 격차를 더 키울 가능성이 높다.

또한 AI시대 격차의 또 한 가지 차별점은 AI에 내재된 알고리즘을 이해하고 통제할 수 있는 권한이 소수에게만 주어진다라는 것이다. AI는 그 특성상 설계자 또는 학습 데이터의 편향을 그대로 지닐 수밖에 없다. 그런데 이런 점을 이해하고 통제할 수 있는 사람은 거대기업이나 소수의 엔지니어에 한정되고, 대부분의 이용자는 이런 편향을 수동적으로 수용하는 데 그칠 가능성이 높다.

‘디지털 디바이드’와 달리, ‘AI 디바이드’는 기존에 능력을 갖춘 사람과 그렇지 못한 사람, 접근권을 가진 사람과 그렇지 않은 사람 사이의 능력 격차를 더욱 극적으로 벌려놓을 것으로 보인다.

AI는 사람의 능력을 증폭시킨다. 즉 능력이 있는 사람이 더 큰 능력을 갖게 해준다. 그런 점에서 디바이스 중심의 이전 기술과는 다르다. 빗대어 말하자면, AI는 동료보다 이미 일을 두 배 잘 하는 사람이, 이전보다 스무 배 더 잘 할 수 있게 만들어주는 기술이다.

예를 들어 보자. 챗GPT의 플러그인은 데이터만 구해다 주면 입맛에 맞게 분석해주기도 한다. 원래 데이터분석을 하던 사람들이 더 빠르게 많은 작업을 할 수 있게 해주는 도구다. 마이크로소프트 코파일럿은 이미 엑셀과 파워포인트를 잘 사용하는 직원이 엑셀만으로도 엄청나게 복잡한 작업을 해낼 수 있게 도와준다. 아직 엑셀과 파워포인트를 사용할 줄 모르는 직원은 누릴 수 없는 혜택이다.

소프트웨어 개발을 기획할 수 있는 능력이 있는 사람은 일일이 코딩하지 않고도 AI개발자에게 맡겨 좋은 소프트웨어를 개발할 수 있다. 창작능력이 뛰어난 작가는 AI작가나 AI화가와 협업하며 뛰어난 작품을 더 많이 창작할 수 있게 된다.

즉 능력이 있는 사람들이 더 큰 능력을 발휘할 수 있게 만들어주는 게 생성 AI 기술이다. 결과적으로 사람들 사이의 능력 격차는 더욱 커지고, 이에 따른 경제적 격차가 사회 문제로 떠오를 것이다. 이전의 사회경제적 격차와는 질적으로 다른 ‘AI 디바이드’의 특징이다.

AI시대의 격차는 디바이스 시대보다 파괴력은 더 크고, 해법도 복잡하다. 어떻게 해결할 수 있을까?

2. AI의 정의와 특징

본격적으로 ‘AI 디바이드’의 해법을 논하기 전에, 우선 우리가 논의해야 하는 AI를 정의해 보자.

인공지능은 오래 된 기술이다. 수십년 전부터 연구가 진행됐다. 인공지능의 대부로 불리는 마빈 민스키가 존 매카시와 함께 인공지능을 세상에 소개한 ‘다트머스 회의’는 1956년에 열렸다. 인간 최고의 체스 선수를 이겼던 IBM의 딥블루가 나온 게 1990년대 후반이다. 이세돌 9단을 이겼던 알파고 사건도 2016년이였다. 우리는 이미 세탁기에도 냉장고에도 인공지능이 탑재되어 있다는 광고에 익숙하다.

인공지능은 이 모든 것이고, 챗GPT 역시 그저 이런 흐름 위에 있는 하나의 작은 사건으로 여겨질 수도 있다.

하지만 챗GPT는 이전과는 전혀 다른 충격을 가져다 주었다. 이유는 단순하다. 사실상 사람과 구분할 수 없는 수준의 언어를 구사한다는 점 때문이다. 인공지능을 정말 ‘지능’이라고 부를

수 있는지에 대한 평가의 기준으로 통용되던 ‘튜링테스트’를 통과하는 최초의 인공지능으로 평가받기 때문이다. 챗GPT가 처음 장착했던 GPT-3.5부터는 튜링테스트가 의미가 없어졌고, GPT-4에서는 더 분명하게 튜링테스트를 넘어선 것으로 보인다.

즉 챗GPT 시대의 특징은, 말만으로는 인간과 구별할 수 없는, ‘지능을 갖고 말하는 기계’가 대중화됐다는 것이다.

튜링테스트란 앨런 튜링이 제안한 지능기계 식별 방법이다. 앨런 튜링은 1950년 ‘계산 기계와 지성’이라는 논문을 기고했는데, 여기서 ‘어떤 경우에 우리는 기계가 지능을 갖고 있다고 말할 수 있는가’라는, 인공지능과 관련된 본질적 질문을 던진다. 방법은 단순하다. 직접 보지 않고 대화를 나누는데 기계인지 사람인지 구분할 수 없으면 지능을 가진 것으로 간주할 수 있다는 것이다.

알파고 때까지만 해도, 튜링테스트를 논할 정도는 아니었다. 알파고는 앨런 튜링의 기준으로 보면, 지능을 가진 기계라기보다는 바둑기술을 갖고 있는 기계에 가까웠다. 포크레인이나 냉장고에 가까웠다는 이야기다.

그런데 챗GPT는 다르다. 대화 내용만으로는 사람과 거의 구분할 수가 없다. 몇 년 전까지만 해도 불가능하다고 여겨졌던 인공지능 모델이 탄생한 것이다. 최초로 ‘지능 기계’가 나왔다고 볼 수 있는 상황이다.

챗GPT가 가져온 충격의 본질이 여기 있다. 사람과 구분하기가 어렵다. 그런데 사람과의 유사성은 대화 방식에 그치는 게 아니다. 더 본질적으로, 기능 자체도 사람과 구분하기가 어려워진 게 챗GPT이다.

챗GPT 공개 초기에는 ‘틀린 이야기가 많다’는 비판이 많았다. GPT-3.5가 GPT-4로 넘어가고 마이크로소프트 검색엔진 bing이 이 모델을 장착하면서 정확도가 좀더 높아지기도 했다. 다만 이 문제는 어느 정도는 본질적이다.

딥러닝 모델은 본질적으로 사람의 두뇌와 유사한 모델이다. 따라서 일정한 확률로 틀린 이야기를 하는 현상은 이어질 가능성이 높다.

그런데 조금만 생각해 보자. 사람이 꼭 그렇다. 대화 중에 정확한 팩트만을 이야기하는 사람은 없다. 착각하거나 의도적으로 유리하게 만들거나 무성의하거나, 이유가 어느 쪽이든 사람은 틀린 이야기를 하는 경우가 많다.

인간은 어쩌면 ‘틀린 이야기를 할 수 있는’ 특징을 갖고 있기 때문에 오히려 스토리텔링을 할 수 있고 창작을 할 수 있다. 엄격한 팩트만 이야기하는 사람이 상상력을 발휘해 놀라운 창작을 하기는 쉽지 않다.

GPT 모델의 팩트 오류는, 사람의 지능을 닮아서 생기는 일이다. 사람의 지능을 따라가지

못해서 생기는 일이 아니다.

즉 챗GPT는 사람이 못하는 엄청난 일을 해주는 인공지능이 아니다. 꼭 사람처럼 활용할 때 가장 좋은 인공지능이다. 사람처럼 대할 때, 즉 창작과 스토리텔링 등을 주문할 때 가장 큰 가치를 줄 수 있다.

〈인공지능: 현대적 접근〉의 저자 스튜어트 러셀은 인공지능에 대한 모든 논의를 2x2 매트릭스 위에 아래와 같이 담는다.

〈그림 1. AI의 분류〉

인간적 사고	합리적 사고
인간적 행동	합리적 행동

가로축은 ‘인간적’인 면을 지향

하는지, ‘합리적’인 면을 지향하는지를 구분한다. 공학 계산기는 합리적으로 높은 단위의 곱셈과 나눗셈과 지수와 로그함수까지도 담아 계산해낸다. 그러나 인간과 유사한 방식으로 소통하지 않는다. 매우 단순한 형태의 ‘합리적’ 지능을 담고 있다고 할 수 있다. 챗GPT는 수많은 실수를 하지만 사람과 구분할 수 없는 대화체의 글을 쓴다. ‘인간적’ 지능을 담고 있다고 할 수 있다.

세로축은 ‘사고’를 강조하는지, ‘행동’을 강조하는지에 따라 나뉜다. 사고를 강조하는 인공지능은 추론하고 논리를 세우는 지능이다. 행동을 강조하는 지능은 지각(sensing)과 표현(representation)까지를 포괄한다.

이런 네 가지 분류에 따라 인공지능에 대한 논의는 매우 다양하다. 논자에 따라서는 세탁기에 들어 있는 빨랫감 자동 판단 로직부터, 화성을 탐사 중인 로봇 스피릿과 오퍼튜니티의 작동 알고리즘까지도 인공지능에 포함시킬 수 있다.

다만 이 발표에서는 앨런 튜링이 ‘튜링 테스트’에 담았던 인공지능 논의를 따라 ‘인간적 행동’을 하는 에이전트로서의 인공지능을 다루면서 이를 확장한다. 즉 ‘인간과의 소통 가능성’을 인공지능의 필수 요소로 놓고, 이런 인공지능의 사회적 영향 측면을 다룬다. 챗GPT 이후 등장한 생성AI가 여기 해당한다고 보면 된다. 자연어와 이미지를 튜링테스트를 통과할 수 있는 수준으로 생성할 수 있는 거대모델(Large Model)의 등장이 그것이다.

생성AI는 인간과 소통하며 협업할 수 있는 ‘인간적 AI’이면서, 인간 고유의 일이라고 여겨졌던 지식정보관련 노동을 더 효과적으로 할 수 있는 AI이다. 그래서 엄청난 크기의 사회 변화를 가져올 수 있다는 게 발표자의 시각이다.

3. SWOT분석: AI의 사회적 영향

사회적 영향이라는 측면에서 AI는 분명한 강점과 약점을 지니고 있다. 다행히 이를 잘 활용한다면 우리 사회의 문제 해결 능력을 크게 높일 수 있지만, 상당히 높은 확률로 거대한 ‘AI 디바이드’ 시대를 맞아 상상하지 못했던 수준의 격차 확대를 목격하게 될 수도 있다.

강점은 우리 사회의 문제해결능력을 한층 높일 수 있다는 점이다. 특히 개인의 지적 능력을 확장하는 방식으로 이를 수행할 수 있다. 이런 AI는 마치 영화 <어벤져스>에서 토니 스타크와 대화를 하며 다양한 작업을 신속하고 정확하게 수행하는 ‘자비스’와 같다. 인간이 설정한 지적 문제를 인간보다 더 빠르고 정확하게 수행해주는 만능 비서와 유사하게 작동할 수 있다. 물론 개인이 아니라 사회적으로도 그렇게 작동할 수 있을 것이다. 즉 사회적으로 공론화된 문제를 해결하는 과정에서 AI가 성능 좋은 지적 조력자, 또는 씽크탱크로 작동할 수 있을 것이다.

반면 약점은 독과점 가능성이다. 소유 측면을 보면, 빅데이터와 엄청난 크기의 하드웨어를 갖추야 하는 초거대 AI 기술의 특성상, 거대기업 몇 군데서 독점할 가능성이 높다.

활용 측면을 봐도, 이미 기술에 대한 이해를 갖추고 있거나 학습 여력이 있는 엘리트 계층이 AI 활용능력도 독점할 가능성이 높다. 특히 AI의 작동원리에 관한 정보와 그 통제권 역시 소수에게 독점될 가능성이 높다. 여력이 부족한 대다수 서민들과 이들 사이의 사회경제적 격차는 기하급수적으로 커질 수 있다.

사회적으로는 AI를 둘러싼 기회와 위협이 존재한다.

대중민주주의 확대로 지식 욕구와 권리의식이 높아지는 우리 사회 환경은 ‘지능 기계’인 AI 기술에게는 기회로 작동할 것이다. 우리는 높은 수준의 지적 욕구를 지닌 사회에 살고 있다. 또한 사람들의 권리의식은 점점 더 확대되고 있다. 따라서 점점 더 많은 사람들이 학습과 지식 표현에 더 많은 시간을 투입한다.

때로 이런 욕구 증가는 지나친 교육열이나 정치적 극단주의로 나타나기도 한다. 하지만 과거와 견주어 시민의 정치 및 사회참여 의지가 강해지고 있다는 점에서 긍정적으로 해석할 수도 있다. AI는 이렇게 높아지는 시민의 욕구와 권리의식을 충족시키며 빠르게 확산될 수 있는 기회를 맞고 있다.

일자리 불안과 기후위기 등 우리를 위협하는 사회 문제는 AI기술에 대해서는 위협으로 작동할 것이다. AI기술은 잠재적으로 노동시장의 태풍이 될 수 있다. 현대사회의 고급 숙련기술로 여겨지던 ‘지식 노동’을 대체할 수 있기 때문이다. 또한 AI기술은 기후위기를 증폭시키는 이미 ‘기후 악당’으로 지목받고 있다. 고용 불안과 기후위기는 이미 우리 사회가 맞닥뜨린 거대한 사회 문제인데, AI가 이를 극적으로 증폭시키는 주범으로 지목받을 수 있다. 이런 상황은 AI기

술을 억제하자는 여론으로 확대될 가능성이 높은 위협이다.

여기 대응하는 사회적 전략을 네 가지로 구분해 살펴볼 수 있다.

첫째, SO전략, 즉 강점으로 기회를 포착하는 전략이다. 보편적 AI접근권을 전국민에게 보장해 지식의 실질적 민주화를 이루는 전략을 구상할 수 있다. ‘자비스’가 될 수 있다는 강점을 활용해 대중민주주의라는 기회를 포착하자는 전략이다.

둘째, ST전략, 즉 강점으로 위협을 극복하는 전략이다. 기후AI와 돌봄AI 등 사회문제 해결에 적극 투자해 AI를 문제 유발자가 아니라 해결자로 자리매김하는 전략을 구상해 볼 수 있다. 말하자면 사회적 가치를 추구하는 ‘자비스’에 투자해 위협을 극복하자는 전략이다.

셋째, WO전략, 즉 약점을 보완해 기회를 포착하는 전략이다. 우리 사회의 높은 지적 욕구와 기술 활용 능력을 토대로, 국내 AI산업을 키우는 새로운 산업정책으로 글로벌 AI독점이 야기할 문제에 대응해보자는 전략을 구상할 수 있다.

넷째, WT전략, 즉 약점을 보완해 위협에 대응하는 전략이다. 보편적 기본소득제를 포함한 전환적 사회보장정책으로 강력한 사회안전망을 제공해서, 일자리 불안 등의 문제에 개인이 대응할 수 있는 체제를 갖추자는 구상이 가능하다. AI 디바이드를 사회보장을 통해 극복하자는 전략이다.

〈그림 2. AI의 사회적 영향 SWOT분석과 대응전략〉

	강점	약점
	자비스(어벤져스)	독과점과 격차 심화
기회	SO전략	WO전략
대중민주주의와 지식욕구 확대	■ 보편적 AI접근권 보장 ==> 지식의 민주화	■ 소버린AI ==> AI주권 확립
위협	■ ST전략	■ WT전략
일자리 불안과 기후위기 증폭	■ 기후AI, 돌봄AI 등에 투자 ==> 사회적가치 AI의 정립	■ 기본소득제 등 전환적 사회보장정책 도입 ==> 전환 기회 제공

(출처: 발표자)

4. 대응 전략

1) SO전략 - 보편적 AI접근권

‘AI 디바이드’를 극복하는 가장 중요한 전략은 전국민에게 보편적 AI 접근권을 보장하는 일이다. 가장 취약한 계층부터 시작할 수 있다.

실제 사례를 보자. 2023년 4월, 경기도는 ‘AI창작단’을 출범시켜 운영했다. 발달장애인으로 구성된 AI창작단은, 생성AI를 활용한 그림 그리기 훈련부터 시작해 기성 작가들과 공동으로 여는 전시회까지 이어지는 창작지원 프로그램이었다.

현장의 모습은 이랬다. 강의실에는 네 명의 발달장애인 참여자들이 앉아 있었다. 강의실 앞 대형화면에서는 강사의 컴퓨터 화면을 보여주고 있었다. 강사는 생성AI를 활용해 만든 ‘경기도 AI창작단’ 창작 도구 페이지를 띄워 두었다. 그림을 그려주는 생성AI를 활용해 개발한 그림 초안 창작 도구였다.

강사는 연신 ‘무엇을 그리고 싶어요?’라는 질문을 던졌다. 참여자들이 그리고 싶은 상상 속의 장면을 이야기하면, 강사는 그 내용을 넓은 화면에 보이도록 타이핑을 해서 보여주었다.

강사의 추가 질문이 이어졌다.

홍학이 사진처럼 보이면 좋을까요? 아니면 어떤 화가의 그림처럼 그려지면 좋을까요? 몇 마리나 있으면 좋을까요? 홍학의 배경으로는 무엇이 있으면 좋을까요? 동물의 숲에 나오는 코끼리는 혼자 있나요 아니면 여럿이 같이 있나요? 멋진 풍경에는 호수가 있나요? 아니면 산길이 있나요? 사람들이 얼마나 있는 풍경이죠?

참여자들은 신이 났다. “사진같이 보이지 않으면 좋겠어요!” “피카소와 달리 그림처럼 그리면 어떨까요?” “호수가 있는 풍경이면 좋겠어요.” “사람은 없이, 조용한 곳이면 좋겠네요!”

강사는 참여자들이 한 말을 하나씩 정리해 준다. 일부는 프롬프트 창에 넣어 그림 초안을 그리고, 큰 화면에 띄워 보여주며 어떻게 바꿔 나갈지에 대해 참여자들과 토론한다. 그리고 참여자들이 각자의 프롬프트를 만들도록 유도한다. 그리고 각자의 컴퓨터에 띄워져 있는 창작 도구에 입력하도록 한다.

컴퓨터를 다룰 줄 아는 몇몇 참여자는 자신의 창에 프롬프트를 써넣기 시작한다. 보조강사는 참여자들 옆에 앉아 한 명씩 대화하며 프롬프트를 넣는 과정을 기술적으로 돕는다. 마음에 드는 그림 초안이 나왔을 때, 그 프롬프트를 강사에게 전달한다. 강사는 다시 그 프롬프트로 그림을 생성해 큰 화면에 띄워 모두에게 보여주고 토론을 유도한다.

참여자들은 원래 그림에 관심이 많고, 실제 그림을 그리는 활동을 하던 이들이었다. 하지만

기획하고 창작하는 능력은 높지 않았다. 누군가 초안을 그려 주면, 그에 맞춰 꾸미거나 유사하게 그리는 활동을 주로 했었다.

이전까지 AI는 사람의 손발이 될 수 있을 것으로 여겨졌다. 예를 들면 자동차를 자율주행차로 만들어 단순작업인 운전을 자동화해줄 것이라고 여겨졌다.

그런데 챗GPT와 같은 생성AI는 우리에게 전혀 다른 기대를 가져다 준다. 손발이 아니라 두뇌가, 단순작업이 아니라 기획과 창작작업을 자동화할 수 있을 것이라는 기대다.

원래 창작이나 기획 능력은 불평등했다. 높은 수준의 교육을 받고 지능이 높은 이들의 전유물이었다. 소수의 창작가나 기획자가 낸 아이디어로 세상이 움직이는 것으로 여겨졌다. 대다수 시민은 이들의 기획에 따라 충실한 손발이 되어 단순한 작업을 ‘노동’이라는 이름으로 이어가며 생계를 꾸리며 살아갔다.

생성AI는 이런 대다수 시민의 능력을 높이는 ‘창작 지능’ 또는 ‘기획 지능’이 될 수 있을까? 바로 AI창작단의 실험이 던지는 질문이다. 발달장애인들이 기존 예술활동에서 한 단계 올라서서, 예술작품을 기획하고 창작하는 과정까지 인공지능의 도움을 받아 수행할 수 있도록 돕는 프로그램이다.

챗GPT 시대는 우리 사회에 어떤 영향을 끼칠까? 우선 눈에 띄는 것은 공포다. 역사적으로 이전의 많은 기술혁신은 많은 이들에게 풍요와 편리를 가져다 주었지만, 한편으로는 불평등을 확대하는 결과를 낳았다. 인간 노동의 일부가 자동화되며 일자리에 구조적인 변화가 나타나기도 한다.

반면 더 많은 사람들이 인공지능을 활용할 수 있게 된다면, 인간의 능력을 획기적으로 향상시켜 오히려 능력의 격차가 줄어들 것이라는 생각도 해볼 수 있다.

챗GPT 시대는 그렇지 않아도 커져가던 불평등을 속절없이 더 커진 사회가 될까? 아니면 지식과 타고난 능력의 격차를 어느 정도까지 좁힐 수 있는 사회혁신을 가능하게 할까? 그래서 기존의 재분배 정책이 건드리지 못했던 불평등의 영역인, 능력의 불평등 또는 지능의 불평등까지도 완화할 수 있을까?

가장 어려운 사람들로 먼저 구성된 경기도 AI창작단은 이런 질문에 대해 답하려는 실험이다. 인공지능 기술이 더 고른 기회를 가져오는 사회로 이어지도록 만들려는 정책적 노력의 출발점이었다.

경기도의 ‘AI창작단’ 사례는 가장 취약한 계층도 적절한 기술습득 과정을 거치고 코칭을 통한다면 이전과 질적으로 다른 능력을 발휘할 수 있다는 점을 보여준다. 발달장애인으로 시작했지만, 노인 어린이 등 대상을 확대하며 결국 전국민에게 보편적인 AI 훈련기회 및 코칭 서비스를 제공하는 방향으로 나아갈 수 있을 것이다. 궁극적으로는 전국민에게 ‘보편적 AI접근권’을

보장해 능력 격차를 완화하는 비전을 세울 수 있을 것이다.

김대중 전 대통령 시절 우리나라는 전국에 초고속인터넷망을 설치하는 대규모 투자를 감행하는 동시에, 농어촌과 어르신들에게까지 컴퓨터와 컴퓨터교육을 제공해 인터넷 사용을 보편적 권리로 격상시키려는 노력을 진행했다. 우리 나라가 인터넷 강국이 되는 바탕이 되는 정책이었다. AI시대에는 더 강력한 보편적 권리 운동이 필요하다.

2) ST전략 - 기후 AI, 돌봄 AI

AI가 효율성을 높이는 것은 사실이나, 사회문제를 오히려 일으키고 있다는 지적은 AI기술에는 중요한 위협이다. 특히 ‘기후악당 기술’이라는 오명은 AI기술이 극복해야 할 위협이다.

실제로 AI를 훈련시키고 이용하는 데 엄청난 탄소배출이 일어난다는 게 전문가들의 의견이다. 국제에너지기구에 따르면, 전세계 데이터센터에서 배출하는 온실가스는 전체 온실가스의 1%를 차지한다. 장비의 발열은 큰데 온도를 적정하게 유지해야 하는 데이터센터의 특성 탓이다.

챗GPT 등 생성AI는 이를 증폭시킨다. 구글의 자체 조사에서는 AI가 이 회사 전력의 15%를 쓴 것으로 나타났는데, 그 규모는 인구 50만명인 미국 애틀랜타의 전체 가구 전력 소비량과 맞먹는다고 한다. 그런데 단순 검색에 비해 챗GPT 등 생성AI는 검색 1회마다 4~5배의 추가 작업이 필요하다고 한다. ‘기후 악당’이라는 오명을 뒤집어쓰기 쉬운 구조다.

그러나 AI는 탄소 배출을 줄이는 데도 높은 성능을 보여줄 수 있으리라는 기대도 있다. 기후테크의 일종인 ‘기후AI’에 대한 투자가 필요한 이유다.

2023년 10월 19일, 임팩트 투자 전문 업체 소풍벤처스는 기후분야 전문가들과 AI전문가들과 기업과 투자자들을 초청해 ‘2023 클라이밋 테크 스타트업 서밋’을 열고 ‘기후위기 문제해결에 AI기술의 활용 가능성’을 논의했다. 에너지 효율화에 기여하는 다양한 AI기술이 논의됐다. 스타트업 ‘식스티헤르츠’는 태양광 및 풍력 발전량을 AI를 활용해 예측해 지도 형태로 제공한다. 이를 활용해 전력의 공급과 수요의 균형을 맞출 수 있게 된다. 네이버와 SK텔레콤은 데이터센터 냉방시스템을 AI로 제어해 에너지 효율을 높인다.

다양한 기업들의 대응도 논의됐다. 아마존웹서비스(AWS)는 데이터센터 전력의 90%를 재생에너지로 전환하는 목표를 세우고 실행을 추진 중이다. 마이크로소프트는 10억 달러(1조 4천 억원) 규모의 기후혁신기금을 설립해 기후AI 등 기술을 활용해 기후문제를 해결하려는 기업에 투자하고 있다. 아마존웹서비스는 기후테크 스타트업 육성 프로그램도 운영 중이다.

이 밖에도 고령화 시대에 대비하는 돌봄AI, 산업재해 가능성이 있는 위험한 작업을 대체하는

로봇기술 등에서 AI가 창출하는 사회적 가치가 드러날 수 있다. 이런 사회문제 해결 AI에 적극 투자하는 게 위협을 극복하는 전략이다.

구체적으로는 기후AI에 투자하는 녹색금융은행의 설립, 돌봄AI 등 사회적 가치를 창출하는 AI기술에 투자하는 사회혁신기금의 설치 등의 대안을 고려할 수 있다.

3) WO전략 - 소버린AI

AI시대의 독점의 정도와 그 폐해는 이전의 기술보다 더 심해질 수 있다. 특히 초거대언어모델이 플랫폼화되고 독점화되면 인간의 사고를 지배할 가능성이 있다는 지적이 이어지고 있다. 플랫폼의 특성상 전세계를 단일 또는 몇 개의 초거대 언어모델이 지배할 수 있게 되고, 인류가 하나의 언어모델을 사용하면서 그 모델이 학습한 사회문화적 배경에 차차 수렴되어갈 수 있다는 이야기다. 결과적으로 각 국가와 문화권이 독자성을 잃어버릴 수도 있다는 무서운 경고다. 일부 과장이 있을 수도 있지만, 완전히 배제할 수는 없는 시나리오다.

현재 초거대 언어모델을 개발한 기업이 있는 국가는 미국, 중국, 이스라엘, 그리고 우리나라 정도다. 우리 사회는 높은 지식 욕구와 독특한 문화에 힘입어 ‘국산 검색사이트’를 사용하고 있는 몇 안 되는 나라 중 하나다. 중국 등 몇몇 국가를 제외하면 상당수 국가들에서 구글이 압도적으로 독점적 지위를 지니고 있다.

이런 독특한 우리 사회의 특성을 토대로, 우리나라 국적AI기업을 지켜가는 전략을 고려해볼 수 있다. 초거대 언어모델 기업도 가능하지만, AI-X 즉 다양한 전문분야에 특화된 AI 기업에도 주목할 필요가 있다. 물론 이런 전략은 재벌기업을 키운 과거 개발연대 산업정책의 형태를 그대로 가져가서는 곤란하다. 거대 AI 기업을 중심으로 다양한 스타트업 생태계가 함께 발전할 수 있도록, AI시대 산업정책의 세부 전략을 구상해야 할 것이다.

4) WT전략 - 전환적 사회보장정책

AI가 증폭시킬 경제적 불평등 확대는 ‘AI 디바이드’의 핵심 문제가 될 것이다. 20세기의 사회보장정책으로는 한계가 분명하다. 새로운 형태의 유연한 사회보장정책이 필요하다.

고용 기간과 급여 수준을 연동시킨 형태의 사회보험(국민연금, 고용보험 등)을 고용과 독립적인 급여 형태인 보편적 사회수당(아동수당, 기초연금, 보편적 기본소득 등)으로 전환하는 전략이 대표적이다. 고용노동정책의 대상을 플랫폼노동자, 1인 사업자, 학생, 전환기 노동자 등으로 확대하는 논의를 진행하고, 보편적 기본소득 정책실험을 확대해야 한다.

또한 기존의 직업 분류가 융합되고 급변하면서, 한 사람이 다양한 직업을 가져야 하는 방향

으로의 환경 변화가 예상된다. 이에 따라 유연한 형태의 교육훈련 시스템이 필요하다. 아동청 소년교육에서 평생학습으로 교육훈련의 중심을 옮기며 대규모 자원을 투입할 필요가 있다. 또한 전통적 직업 분류와 규제 위주로 구성되어 있는 고용보험의 직업훈련 지원 시스템을 혁신할 필요가 있다.

궁극적으로는 사회보장 강화를 통해 누구나 생애 어느 때라도 적절한 학습 시간을 가질 수 있으며 직장을 옮겨다닐 수 있도록 소득을 보장해줘야 한다. 이를 바탕으로, 인간은 종합적 판단력과 통찰력을 중심으로 일하며 전통적인 실무는 상당부분 AI기술을 활용하는 방향으로 노동과정을 변화시켜 나가야 한다. 어찌 보면 인간은 전통적으로 기업가가 해왔던 일을 하게 되는 셈이다.

지난 100년 동안 자본주의는 '모든 사람이 노동자인 사회'를 지향하며 발전해 왔다면, 앞으로 100년 동안 우리는 '모든 사람이 기업가인 사회'로 차차 변화해 나가게 될 것이다. 여기에 대비해 사회보장제도와 교육훈련체계를 손질해 나갈 필요가 있다.

GPT가 대표하는 거대언어모델은 태생적으로 인간보다 사물에 대한 이해도가 낮은 게 사실이다. 언어모델은 대상 자체를 인식하는 게 아니라, 단어와 단어 사이의 이음새만 인식하며 글을 만들어낸다. 언어가 묘사하는 세상 그 자체는 인식하지 못한다.

딸기를 예로 들어 보자. 챗GPT는 '딸기'를 모른다. 다만 딸기에 대한 여러 가지 설명, 예를 들어 '붉은 색의 단 맛이 나는 과일' 등을 언어로만 이해한다.

그런데 딸기란 무엇인가? '붉은 색의 단 맛이 나는 과일'이라는 언어만 알면 딸기를 아는 것인가? 딸기의 촉감, 딸기의 기억, 딸기의 맛 등등을 느껴보지 않고 글로만 딸기를 이해하면 딸기를 아는 것일까? 그렇지 않다.

결국 다섯 살 어린아이가 가장 고도로 발달한 인공지능보다 '딸기'를 훨씬 더 잘 안다고 할 수 있다. 어린아이는 '딸기'나 딸기에 대한 설명을 아는 것이 아니라, 딸기 그 자체를 알기 때문이다.

이런 면에서 챗GPT의 한계는 명확하다. 그래서 챗GPT를 비판하는 사람들은, 세계를 이해하지 못하는 이 모델은 성공할 수 없다고 한다. 어린아이보다도 못한 인식 능력으로는 결국 할 수 있는 일이 없을 것이라는 이야기다.

어쩌면 사람에게 늘 가까이 있고 느낄 수 있는 것일수록, 사람만이 이해할 수 있다. '딸기'처럼 단순해 보이는 사물이 어쩌면 챗GPT에게는 가장 이해하기 어려운 것일 수 있다.

그런데 다시 한번 생각해 보자. 우리는 예컨대 '나스닥시장'을 어떻게 이해하고 있는가? 나스닥시장을 딸기처럼 풍부하게 이해하고 있는가? 언어로만 이해하고 있는 것은 아닌가?

각종 법률지식은 어떨까? '막스 베버의 사회학'은 어떤가? '에피쿠로스학파의 행복론'은 어떤가?

실은 사람도 알고 있는 지식의 많은 부분을 순전히 언어로만 익힌 것 아닌가?

그렇다면 그런 개념을 이해하는 지능은 사람과 기계가 거의 같다고 볼 수 있지 않을까?

GPT도 세상을 이해하고 있다. ‘나스닥시장’이나 ‘한계 효용’처럼 어렵고 추상적인 것일수록 오히려 사람과 비슷한 방법으로 이해한다.

이런 ‘추상적 언어 인식 능력’은 이전에는 인간 고유의 능력으로 여겨졌다. 그래서 ‘딸기’를 다루는 과일장수의 노동보다 ‘나스닥시장’을 다루는 금융인의 노동, ‘에피쿠로스 학파의 행복론’을 다루던 교수의 노동이 더 고급스럽고 인간적인 것으로 여겨졌다. 추상적 개념을 다룰수록 더 많이 존중하고 더 많이 보상했다. ‘화이트칼라 노동’, 즉 ‘지식노동’의 신화는 그렇게 형성됐었다.

챗GPT는 딸기를 모른다. 대신 ‘나스닥시장’은 아주 잘 안다. 그래서 챗GPT는 그래서 화이트칼라 노동에 위협적이다. 여기에 인간 고유의 일이라고 여기던 ‘지식의 생성능력’까지 얻으면서 챗GPT 시대의 인공지능은 노동시장, 특히 지식노동을 뒤흔들 잠재력을 갖게 됐다.

이미 마이크로소프트는 오피스365의 ‘코파일럿 기능’을 발표했다. 워드 엑셀 파워포인트 작업의 상당부분이 자동화된다는 발표다. 또 다른 문서작성서비스 감마앱도 프리젠테이션과 문서와 웹페이지 자동 생성 기능을 내놓았다. 제목만 정하면 좀 어려운 주제에 대한 프리젠테이션 자료도 세련된 디자인으로 1~2분 만에 만들어준다. ‘실무의 종말’ 수준의 직무 대격변이 일어날 것은 확실해 보인다.

이런 격변이 현실화되면, 중간 수준 지식 실무 노동이 빠른 시간 내 소멸하며 격차 확대가 벌어질 것이다. 문서작성이 대부분 자동화될 텐데, 워드 엑셀 파워포인트를 작성하는 일자리나 교육이 의미가 있을까?

전세계 수억명의 노동자들이 영향을 받을 것이다. 기존 일자리는 사라지거나 채워지지 않을 것이다. 새로운 능력을 찾아내지 못한 노동자 계층은 사실상 실업상태로 진입할 것이다. 디자인과 웹페이지 작성 등을 맡고 있는 플랫폼 노동자 상당수가 일거리를 잃어버릴 것이다.

비정규직과 플랫폼 노동 형태로 하고 있는 일이 가장 먼저 위기에 처할 것이다. 고용을 보장 받은 정규직이나 공공부문, 그리고 권력을 가진 면허 전문직들은 여전히 안전할 방법을 찾아낼 것이다.

관리직처럼 결정하고 책임지며 소통하는 일자리가 훨씬 더 많이 필요해질 것이다. 현장에서 사람을 돌보고 소통하는 일이 더욱 중요해질 것으로 보인다. 리더십과 소통 능력이 절실한 직무들이다.

대신 이런 리더들을 보조하는 데 필요한 다양한 업무와 그 인력은 대폭 줄어들며 효율화될 수 있다. 이른바 ‘스태프 조직’이 가장 큰 위기를 맞을 수 있다. 이에 맞춰 기업 조직에도 큰

변화가 생기게 될 것 같다.

2013년, 옥스포드대의 프레이와 오스본은 ‘자동화로 절반 가까운 일이 사라질 것’이라는 예측을 내놓았다. OECD에서는 중간 일자리가 디지털 전환의 가장 큰 영향을 받을 것이라고 발표했다. 많은 사람들은 그들의 예측을 부정했다. 하지만 결국 그 예측들은 맞아들어가고 있다.

불안의 시대를 잘 버티려면, 직업훈련과 고용보험제도와 사회보장제도의 대수술이 필요해질 것이다. 특히 기초연금, 아동수당, 보편적 기본소득제 등 소득보장 정책과 교육시스템 혁신 논의가 필요하다.

5. 누가 지출할 것인가

인공지능은 시민에게 양날의 칼이다. 잘 활용하면 모든 시민이 높은 능력을 갖춘 고도화된 사회로 나아갈 수 있다. 시민 개인이 국회의원이나 공무원의 도움을 받지 않고도, 챗GPT와 대화하면서 좋은 정책을 만들어 제안할 수도 있다. 누구나 특정 사안에 대해 논리를 잘 만들어 전문가들과 치열하게 논쟁할 수 있다. 시민이 정책에 직접 의견을 내고 토론하며 의사결정을 하는, 직접민주주의가 강화될 수 있다. 진정한 ‘지식의 민주화’가 이뤄지는 것이다.

하지만 거꾸로 권력이 시민에 대해 고도의 감시를 하게 될 수도 있다. 초거대 인공지능 모델을 소유한 인공지능 플랫폼 기업이 우리의 가치를 지배하게 될 수도 있다. 독점이익이 더욱 커지면서 불평등이 확대되어 ‘AI 디바이드’라는 재앙을 맞을 수도 있다.

결국 정치와 시민사회와 대응하기 나름이다. 격변의 시대는 늘 진보를 가져오기도 하고 역사의 비극적 후퇴를 가져오기도 했다. 챗GPT시대도 예외는 아니다. 어떤 정치, 어떤 시민사회, 어떤 노동을 선택할 것인가. 우리는 지금 사회적 선택의 기로에 서 있다.

그런 점에서 ‘AI 디바이드’를 막고 AI시대를 기회의 시대로 만들기 위해서는 ‘보편적 AI접근권’을 권리와 정책 차원에서 정립해 나가는 것이 필요하고, 전환적 사회보장정책을 도입해야 하며, 기후AI 등에 대규모 투자를 투입해야 하고, AI주권을 지켜야 한다.

그런데 이 모든 일을 하는 데는 막대한 자원이 필요하다. 무엇보다 정부 재정 투입이 필수적이다. 재원 마련 전략을 마련하는 것이 AI시대 국가가 해야 할 가장 중요한 일이다.

기본적으로 사회 변화에 대응하는 재원은, 그 변화로 인해 혜택을 입는 사회 구성원들에게 거두는 것이 합리적이다. 따라서 AI기술의 혜택을 입는 기업이나 개인들에 대한 과세 등의 재원 마련 방안이 반드시 마련되어야 한다. 유럽연합 등 국제사회에서 논의되는 로봇세나 탄소세 등의 논의를 참고하되, 우리 사회에 맞는 전략을 보다 깊이있게 연구할 필요가 있다. 다만 AI시

대를 위한 전환적 재정 지출은, 사라지는 비용이 아니라 미래에 대비한 투자하는 점을 명확히 하며 전략을 마련할 필요가 있다.

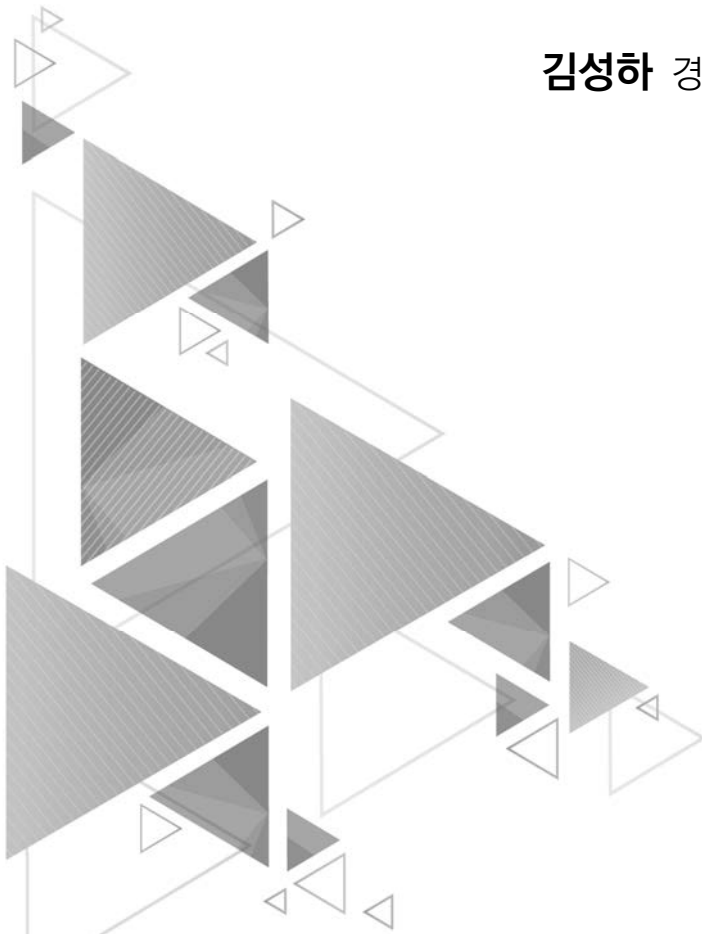
분명한 것은, 미래의 문을 열 열쇠는 시민들에게 있어야 한다는 점이다. 즉 인공지능 접근권과 통제권을 시민들에게 돌려주어야 한다. 이른바 ‘리터러시’와 ‘거버넌스’이다. 이 글에서는 ‘리터러시’를 중심으로 대안을 살펴봤다. 미래로 가는 문을 열 열쇠를 찾아 시민들에게 쥐어줄 책임은, 정치가 짊어지고 있다. 딸기를 아는 시민이 나스닥시장을 이해하고 법제도를 꿰고 있는 챗GPT를 잘 부리며 투표하고 투자하는 세상, 어쩌면 그게 우리가 꿈꾸던 이상적 민주주의 사회일지도 모른다.



발제 2

기후 테크와 AI

김성하 경기연구원 AI혁신정책센터장





목차

I. 2050 탄소중립(net zero by 2050)

II. 「Switch the 경기」 프로젝트

III. 기후 테크(Climae Technology) & AI, 미래사회

I. 2050 탄소중립(net zero by 2050)

GRI 경기연구원

그리스 산불(2023.8)

A Alexandroupoli, dans la région d'Evros, le 22 août 2023.
ALEXANDROS AVRAMIDIS / REUTERS

자료 : https://www.lemonde.fr/planete/article/2023/08/22/incendies-en-grece-les-nouveaux-feux-font-au-moins-deux-morts-et-contrainignent-de-nombreux-habitants-a-evacuer_6186174_3244.html

아프리카 홍수(2023.11)

물에 잠긴 도로 건너는 케냐 뭄바사 주민들



자료 : 연합뉴스(2023.11.18), “아프리카의 불·지역 홍수 사망자 130명으로 늘어”
<https://www.yna.co.kr/view/AKR20231118004900099?input=1195m>

I. 2050 탄소중립(net zero by 2050)

GRI 경기연구원

○ 『탄소중립기본법』

- 제7조제1항 : 정부는 2050년까지 탄소중립을 목표로 하여 탄소중립 사회로 이행하고 환경과 경제의 조화로운 발전을 도모하는 것을 국가비전으로 한다.
- 제10조제1항 : 정부는 (...) 국가 탄소중립 녹색성장 기본계획(이하 “국가기본계획”이라 한다)을 5년마다 수립·시행하여야 한다.

VI. 국가 탄소중립·녹색성장 기본계획

국가비전

2050년까지 탄소중립을 목표로 하여 탄소중립 사회로
이행하고, **환경과 경제의 조화로운 발전**을 도모

국가전략

구체적·효율적 방식으로 온실가스를 감축하는
책임감 있는 탄소중립

민간이 이끌어가는
혁신적인 탄소중립·녹색성장

모든 사회구성원의 공감과 협력을 통해
함께하는 탄소중립

기후위기 적응과 국제사회를 주도하는
능동적인 탄소중립

자료 : 관계부처 합동(2023.4), 「탄소중립·녹색성장 국가전략 및 제1차 국가 기본계획 요약」, p.12.

I. 2050 탄소중립(net zero by 2050)

GRI 경기연구원

○ 국가전략

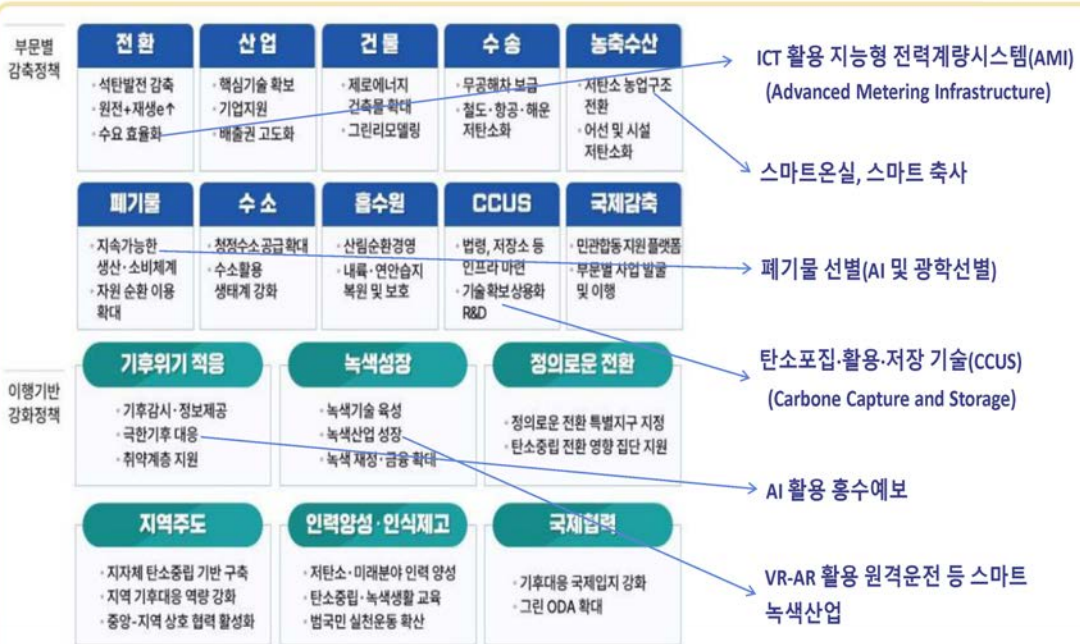


AI 홍수경보,
산불·산사태
조기경보 등
기후위기 감시
인프라 확대

자료 : 관계부처 합동(2023.4), 「탄소중립·녹색성장 국가전략 및 제1차 국가 기본계획 요약」, pp.7~9.

I. 2050 탄소중립(net zero by 2050)

GRI 경기연구원

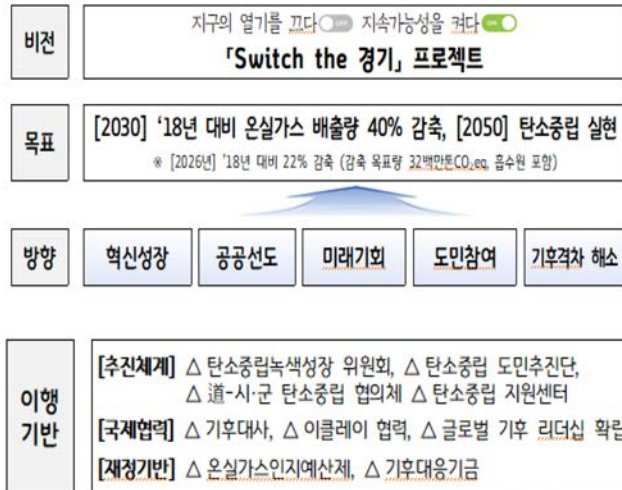


자료 : 관계부처 합동(2023.4), 「탄소중립·녹색성장 국가전략 및 제1차 국가 기본계획 요약」, pp.12~22.

II. 「Switch the 경기」 프로젝트

GRI 경기연구원

◎ 2023.9.20. 경기도, 탄소중립 비전 '스위치 더 경기(Switch the 경기)' 발표



* 경기도 탄소중립 기본계획 수립 (2024년 4월)

자료 : “김동연, ‘스위치 더 경기로 기후위기 대응 앞장서겠다’ ... 2030년까지 온실가스 40% 감축, 보도자료(경기도), 2023.9.20.

II. 「Switch the 경기」 프로젝트

GRI 경기연구원

◎ 8대분야 28개 추진과제

8대 분야 28개 추진 과제	Switch the Energy	Switch the City	Switch the Mobility
	<ul style="list-style-type: none"> • 공공 RE100 • 기업 RE100 • 도민 RE100 • 산업 RE100 	<ul style="list-style-type: none"> • 탄소제로 시티 • 탄소제로 빌딩 • 탄소제로 인프라 • 탄소제로 포레스트 	<ul style="list-style-type: none"> • 그린 모빌리티 • 퍼블릭 모빌리티 • 마이크로 모빌리티 • 스마트 모빌리티
	Switch the Farming	Switch the Waste	Switch the Activity
	<ul style="list-style-type: none"> • 예고 영농 확대 • 농촌 신재생에너지 확대 • 예고 농업기술 개발·보급 	<ul style="list-style-type: none"> • 폐기물 처리기반 확충 • 폐기물 에너지 자원화 • 자원순환 문화 확산 • 플라스틱 다이어트 	<ul style="list-style-type: none"> • 지구동행 실천교육 • 지구동행 실천활동 • 지구동행 전문가 양성
	기후테크 육성	기후위기 적응	
	<ul style="list-style-type: none"> • 기후테크 기업육성 • 기후테크 투자확대 • 기후테크 인력양성 • 기후테크 거버넌스 구축 	<ul style="list-style-type: none"> • 기후재난 피해방지 • 기후격차 해소 	

자료 : “김동연, ‘스위치 더 경기로 기후위기 대응 앞장서겠다’ ... 2030년까지 온실가스 40% 감축, 보도자료(경기도), 2023.9.20.

II. 「Switch the 경기」 프로젝트

GRI 경기연구원

◎ 8대분야 28개 추진과제

Switch the Energy

- 공공 RE100
- 기업 RE100
- 도민 RE100
- 산업 RE100

- '경기 RE100' 비전 + 신재생에너지 정책
- 공공기관 2026년 사용전력 100% 신재생에너지 대체
- 산업단지 입주 기업 민관협력형 태양광 설치 등

Switch the Mobility

- 그린 모빌리티
- 퍼블릭 모빌리티
- 마이크로 모빌리티
- 스마트 모빌리티

- 공공 및 민간 단계적 전기차 보급
- 스마트 모빌리티(예: 독 버스) 적극 도입 확대 등

Switch the City

- 탄소제로 시티
- 탄소제로 빌딩
- 탄소제로 인프라
- 탄소제로 포레스트

- 도시 건축 산림 분야 기후위기 대응
- '개발사업 탄소 총량제' 도입
: 신규개발사업 탄소 흡수량 대체 흡수원 조성
혹은 부담금 납부
- 도로, 버스정류소·차고지, 공영주차장 등 도시 인프라
에너지 자립 추진 등

Switch the Farming

- 예고 영농 확대
- 농촌 신재생에너지 확대
- 예고 농업기술 개발·보급

- 친환경 농업 면적 대폭 확대
- 로컬푸드 직매장 확대
- 저탄소 유통체계 조성 등 저탄소 농업 전환 등

자료 : "김동연, '스위치 더 경기로 기후위기 대응 앞장서겠다' ... 2030년까지 온실가스 40% 감축, 보도자료(경기도), 2023.9.20.

II. 「Switch the 경기」 프로젝트

GRI 경기연구원

◎ 8대분야 28개 추진과제

Switch the Waste

- 폐기물 2탄소1기반 확충
- 폐기물 에너지 자원화
- 자원순환 문화 확산
- 플라스틱 다이어트

- 자원순환분야 기후위기 대응
- 공공기관 1회용품 사용금지
- 바이오가스 활용 신재생에너지 생산
- 폐기물 처리시설 확충

기후테크 육성

- 기후테크 기업육성
- 기후테크 투자확대
- 기후테크 인력양성
- 기후테크 거버넌스 구축

- 2026년 기후관련 새싹기업(스타트업)
100개 발굴 및 지원(성장단계별 맞춤형)

Switch the Activity

- 지구동행 실천교육
- 지구동행 실천활동
- 지구동행 전문가 양성

- 시민들의 자발적인
온실가스 감축 활동 문화조성

기후위기 적응

- 기후재난 피해방지
- 기후적차 해소

- 극단적 재난 피해 최소화
- 기후위기 대응역량에 따른 기후적차 극복

자료 : "김동연, '스위치 더 경기로 기후위기 대응 앞장서겠다' ... 2030년까지 온실가스 40% 감축, 보도자료(경기도), 2023.9.20.

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

탄소 감축 기술과 기후 적응 기술

⇒ 기후 테크(Climate Technology), 기후기술

“기후 AI 테크” (AI solutions for climate change)

완화(mitigation), 적응(adaptation), 회복(resilience)

AI

Fighting Climate Change with the AI for the Planet Alliance

20 May 2022 - Last update: 20 April 2023

(...)

research and advocacy. UNESCO in cooperation with the United Nations Development Programme (UNDP), the UN Office of Information and Communications Technology (OICT), and the AI for Good Foundation along with Startup Inside, Boston Consulting Group (BCG) and BCG GAMMA aims to work with a multidisciplinary and multistakeholder group of AI and climate leaders to drive concrete and measurable action.

(...)

- AI can be used to accelerate efforts to mitigate climate change by

(...)

- AI can be used to accelerate efforts to mitigate climate change by optimizing supply chain networks to simultaneously reduce the energy needed and the costs incurred to ship people and cargo

(...)

- AI can be used to play an important role in climate adaptation and resilience efforts by gathering and analyzing large datasets in real time, leading to the creation of early warning systems for extreme

(...)

○ AI 활용 기후변화 대응 UNESCO, UNDP, OICT, BCG 등 국제 협력

- 실시간 데이터셋(dataset) 수집 분석과 조기 경보 시스템 구축(해수면 상승 등)

- 이동, 운송 등 에너지 소비 및 비용 감소

- 연구·교육 등을 통한 기후 관련 정책 수립 및 결정에 긍정적 영향 등

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

One Concern (<https://oneconcern.com>)

- 공동 설립자

: Ahmad Wani(CEO), Nicole Hu, Timothy Frank

- AI, 데이터 사이언스(data science), 심각한 재난 모델링 활용
“위험을 회복(resilience)으로” 전환

- 정전, 공급망 훼손 등 기업(금융, 보험사 포함) 자산 및 비즈니스 “위험의 정량화” 및 “전략 재구성”을 위한
“회복력계량(지표)” 제공

- 주요 활용 기술

: 데이터 중심 AI(Data-Centric AI)

: 기본 모델(Foundation Models)

: AI 운영(AI Operations)

: 재난 모델링(Catastrophe Modeling)

자료 : 이미지캡처(<https://oneconcern.com/en/products/domino/>)

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

Climavision (<https://climavision.com>) - DALTON AI Platform

- 설립자-CEO : Chris Goode

- Dalton AI + NWP(Numerical Weather Prediction) 수치일기예보 + APIs

- Dalton AI NWP

: 매일 15억 개 글로벌 데이터셋(datasets) 캡처

: AI와 Machine Learning을 통하여 관찰결과를 예측모델로 동화

: 24시간 대응팀

: 유연한 데이터 전달 및 통합 경로

: 완전 자동화

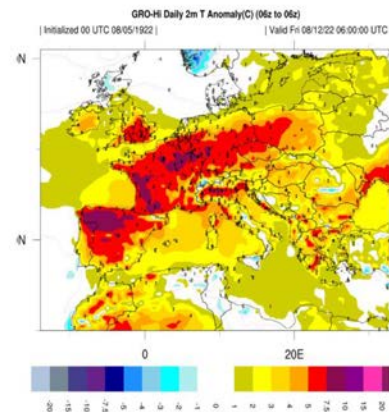
- GRO(Global Radio occultation & Observations) 예측

: 일일 날씨 추적

: 집중호우 확인

: 데이터 수집 증가

: 우주 일기 예보 향상



자료 : 이미지캡처(<https://climavision.com/solutions/weather-forecasting/gro-forecasting/>)

* John Dalton(1766~1844) : 기상학자, 과학자, "기상관측과 소론"(1793).

* API(Application Programming Interface) : 응용체제가 제공하는 함수의 집합체로 한 프로그램이 다른 프로그램과 상호작용하는 방법을 정의하는 일련의 규칙

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

AI 활용 기후테크 사례

자료 : "AI와 데이터로 지구를 지키는 방법! 기후테크(Ctech)", LG CNS 블로그, 2023.2.28.

○ 기후테크 회사 패치(Patch)

- 패치 API + 고객 ERP(Enterprise Resource Planning) 시스템 = 탄소배출량 계산

○ 여행 예약 서비스 사파라(Safara)

- 패치 API 이용 사용자 배출 탄소량 시각화

○ 결제 서비스 업체 애프터페이(Afterpay)

- 패치 API 이용 사용자 온라인 쇼핑시 발생한 탄소 배출량 수치 확인 후 원하는 금액만큼 기후 변화 관련 프로젝트 후원

- 1년 간 기부금 15만 달러 모금

- 클로버리(Cloverly), 카본 인터페이스(Carbon Interface), 쿨러(Cooler) 등 경쟁업체 지속적 증가

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

AI 활용 기후테크 사례(국내)

자료 : "AI 기후테크, 수출로 먹고 살았던 한국에 새로운 기회", 중앙일보, 2023.10.25.

"SK에코플랜트, 엔트와 '탄소회계 & 관리 플랫폼' 고도화 나선다. 벤처스퀘어, 2023.03.31.

○ 엔츠(AENTS)

- 탄소회계 SaaS(Software as a Service) 국내 최초 상용화
- AI 활용 기업의 탄소배출량 측정을 위한 모든 정보 수집 및 탄소 배출량 현황 데이터 생성
- SK에코플랜트와 '탄소배출량 진단 및 감축 서비스 제공' 업무 협약 체결(2023.03.28.)

○ 수퍼빈(SuperBin)

- 폐기물 수거 재생 스타트업
- AI 선별 시스템으로 페트병, 알루미늄, 캔 등의 소재 인식 및 이물질 제거

○ 60Hertz(<https://60hz.io>)

- "햇빛바람지도" : 국내 최대 재생에너지 자원지도
- 약 8만여 개 태양광, 풍력 발전소 확인
- AI 기술로 미래 발전량 예측, 기상 정보 시각화 등

자료 : 이미지캡처(햇빛바람지도) (<https://map.60hz.io>)

* SaaS(Software as a Service) : 서비스형 소프트웨어. 개인이나 기업이 필요에 따라 최적의 소프트웨어를 세분화해 사용할 수 있도록 인터넷으로 제공하는 서비스

III. 기후 테크 & AI, 미래사회

GRI 경기연구원

달라지는 미래 모습

부문	현재	미래
에너지	화석연료 기반 에너지 생산	저탄소 신기술 기반 에너지 생산
산업	탄소 집약적 산업구조	산업의 저탄소 전환
건물	에너지 다소비 건물 다수	성능개선을 통한 에너지 효율 향상
수송	내연기관 중심 수송체계	무공해차 중심 수송체계

·발전 발전 비중
(21) 27.4% → (30) 32.4%

·산업용 발전 비중
(21) 7.5% → (30) 21.6%+*

·배출권거래제 배출요율기준 할당(BM)
(21) 65% → (30) 75%

·그린리모델링(누적)
(22) 약 7.3만건 → (30) 약 160만건

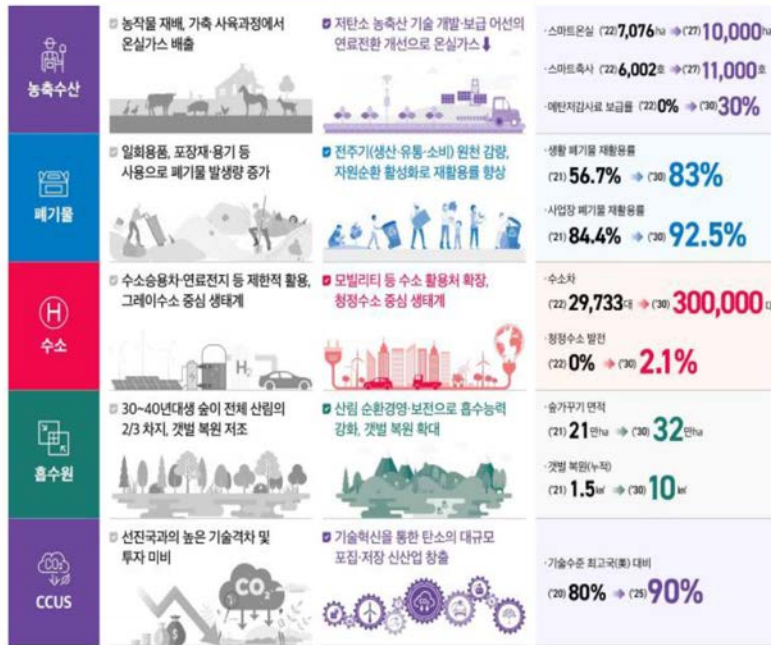
·제로에너지 건축물(누적)
(22) 2,950건 → (30) 약 47,000건

·무공해차 등록 비중(전기차 및 수소차, 누적)
(22) 1.7% (43만대 보급) → (30) 16.7% (450만대 보급)

자료 : 관계부처 합동(2023.4), 「탄소중립·녹색성장 국가전략 및 제1차 국가 기본계획 요약」, p.24.

III. 기후 테크 & AI, 미래사회

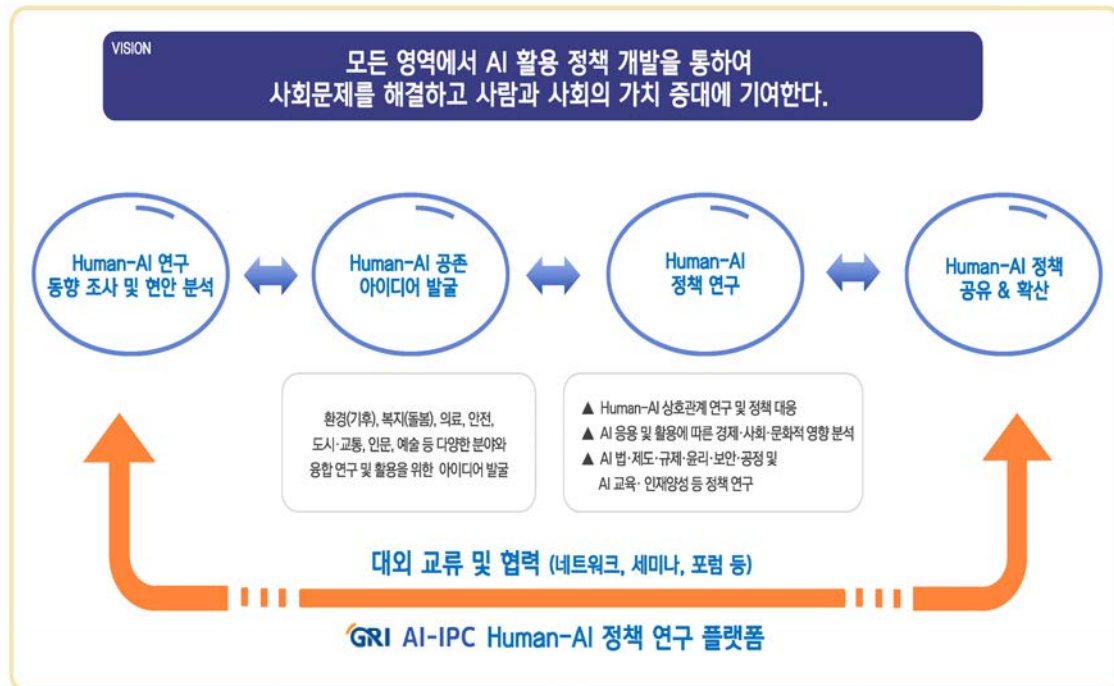
GRI 경기연구원



자료 : 관계부처 합동(2023.4), 「탄소중립·녹색성장 국가전략 및 제1차 국가 기본계획 요약」, p.24.

[첨부] AI 혁신정책센터(Innovation Policy Center) : Human-AI 정책연구 플랫폼

GRI 경기연구원



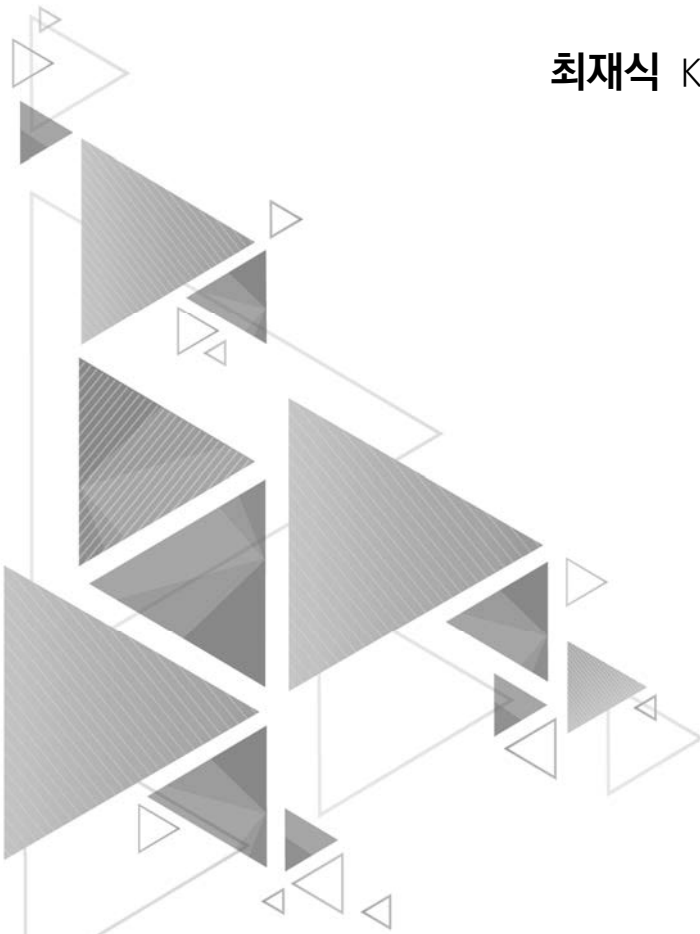




발제 3

인공지능 신뢰성에 관한 제언

최재식 KAIST 김재철AI대학원 교수

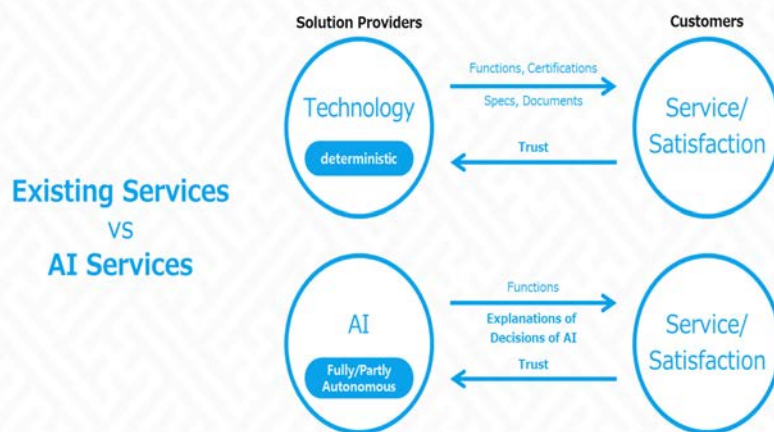


인공지능 신뢰성에 관한 제언

최재식

KAIST, 설명가능인공지능 연구센터장
인이지 대표이사

Why Explainable AI(XAI)?



Right to explanation 13-14 : Customers have the right to receive **proper explanations on the decisions made by AI algorithms**

Fines Up to **4% of total global revenue**
(In high risk AI, the fine will be up to 6% of total global revenue)

EU: General Data Protection Regulations (GDPR)

Items	Contents
Right to be forgotten	17 : When customers do not want, the personal contents should be eliminated
Limit of AI decision	22 : Customers have the right not to be handled by AI algorithm
Right to explanation	13-14 : Customers have the right to receive proper explanations on the decisions made by AI algorithms
Fines	Up to 4% of total global revenue
Enact	2018/05/28

In the area of high risk AI, the fine will be up to 6% of total global revenue

US: NIST AI Risk Management Framework

AI Risk Management Framework

The AI Risk Management Framework (AI RMF) is intended for voluntary use and to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.

As a consensus resource, the AI RMF was developed in an open, transparent, multidisciplinary, and multistakeholder manner over an 18-month time period and in collaboration with more than 240 contributing organizations from private industry, academia, civil society, and government. Feedback received during the development of the AI RMF is publicly available [on the NIST website](#).

[Download the framework](#)

1 Framing risk

[Framing risk](#) includes information on:

- Understanding and Addressing Risks, Impacts, and Harms
- Challenges for AI Risk Management

2 Audience

Identifying and managing AI risks and potential impacts requires a broad set of perspectives and actors across the AI lifecycle. The [Audience](#) section describes AI actors and the AI lifecycle.

3 AI Risks and Trustworthiness

For AI systems to be trustworthy, they often need to be responsive to a multiplicity of criteria that are of value to interested parties. Approaches which enhance AI trustworthiness can reduce negative AI risks. The [AI Risks and Trustworthiness](#) section articulates the characteristics of trustworthy AI and offers guidance for addressing them.

4 Effectiveness of the AI RMF

The [Effectiveness](#) section describes expected benefits for users of the framework.

5 AI RMF Core

The [AI RMF Core](#) provides outcomes and actions that enable dialogue, understanding, and activities to manage AI risks and responsibility develop trustworthy AI systems. This is operationalized through four functions: Govern, Map, Measure, and Manage.

6 AI RMF Profiles

The use-case [Profiles](#) are implementations of the AI RMF functions, categories, and subcategories for a specific setting or application based on the requirements, risk tolerance, and resources of the Framework user.



- US NIST Established the procedure of Trustworthy and Responsible AI Resource Center(AIRC, <https://airc.nist.gov/Home>) to support organizations/institutes to build responsible AI (March 2023).
- **AI Risk Management Framework (AI RMF) assesses the trustworthiness of AI product, service, design, and implementation.**
- 240 organizations (companies, academia, organizations and government) agree the process.

https://airc.nist.gov/AI_RM_F_Knowledge_Base/AI_RM_F

China: forming a national standards for testing large language models

Tech / Policy

China to create and implement national standard for large language models in move to regulate AI, while using its power to transform industries

- The China Electronic Standardisation Institute, under the Ministry of Industry and Information Technology, will enact a local standard for LLMs
- Baidu, Huawei, 360 Security and Alibaba have been enlisted by the institute to lead a special task force that will draw up the new LLM standard



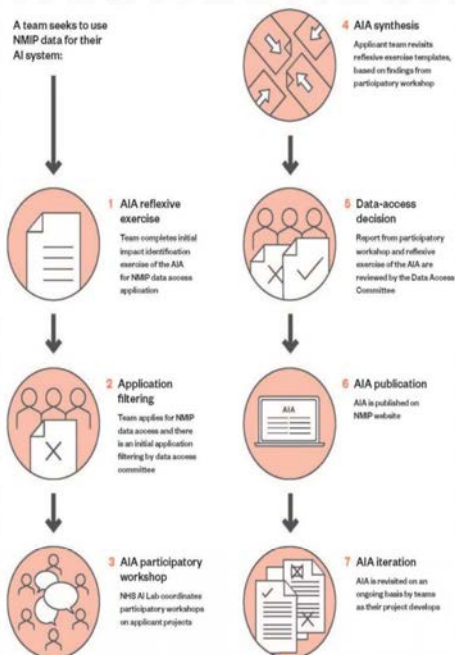
Ben Jiang in Beijing and Ann Cao in Shanghai

Published: 8:15pm, 7 Jul, 2023

Why you can trust SCMP

<https://www.scmp.com/tech/policy/article/3226942/china-create-and-implement-national-standard-large-language-models-move-regulate-ai-while-using-its>

UK: Impact Assessment of AI Health Care



- **Ada Lovelace Institute conducts the Impact assessment of AI algorithm** (Algorithmic Impact Assessment, AIA) on National Health Service (NHS) UK
- All companies need approval from the AIA process when they wish to build AI models with NIMP data of NHS. NHS finally selects the companies and organizations who can access the NIMP data.
- **NHS NIMP AIA process emphasizes the accountability and the transparency of AI models.**

https://www.koraia.org/ctrl/board.php?template=base&com_board_basic=file_download&com_board_id=12&com_board_idx=363&com_board_file_seq=0

Strategy to Realize AI Trustworthy in Korea

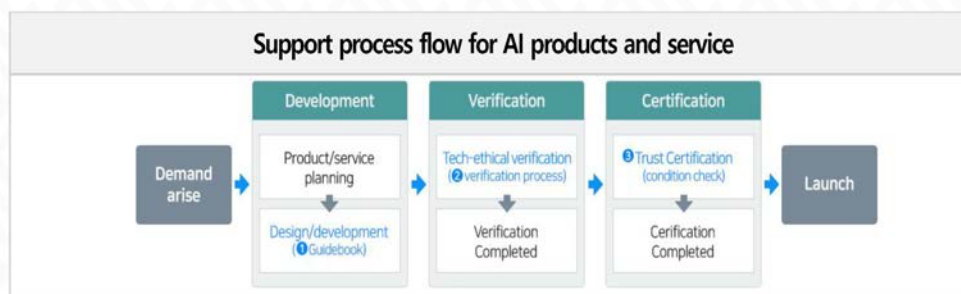
June 2021

Vision, goals, detailed strategies of trustworthy artificial intelligence			
The strategy has the vision of "realize trustworthy artificial intelligence for everyone" and will be implemented step by step until 2025, based on the three pillars of 'technology, system, ethics' and 10 action plans.			
Vision	Trustworthy AI for Everyone		
Goal (~2025)	Responsible use of AI Global no.5	Trustworthy society Global no.10	Safe cyber nation Global no.3
Strategies	Create an environment for trustworthy AI	Lay the foundation for safe use of AI	Spread AI ethics across society
	1) Put in place a systematic process for securing trust for AI products and services 2) Support players in the private sector with securing trust for AI 3) Developing source technology for trustworthy AI	1) Make AI learning data more trustworthy 2) Promote securing trust for high-risk AI 3) Conduct assessment on influence of AI 4) Improve regulations for increased trust for AI	1) Provide strengthened education programs on AI ethics 2) Create and distribute checklists for each stakeholder 3) Operate a platform for policies on ethics

UK ▶ Established 5 codes of ethics (2018 Apr), a guide to using AI in the public sector (2019 June), a guideline for explainable AI (2020 May)

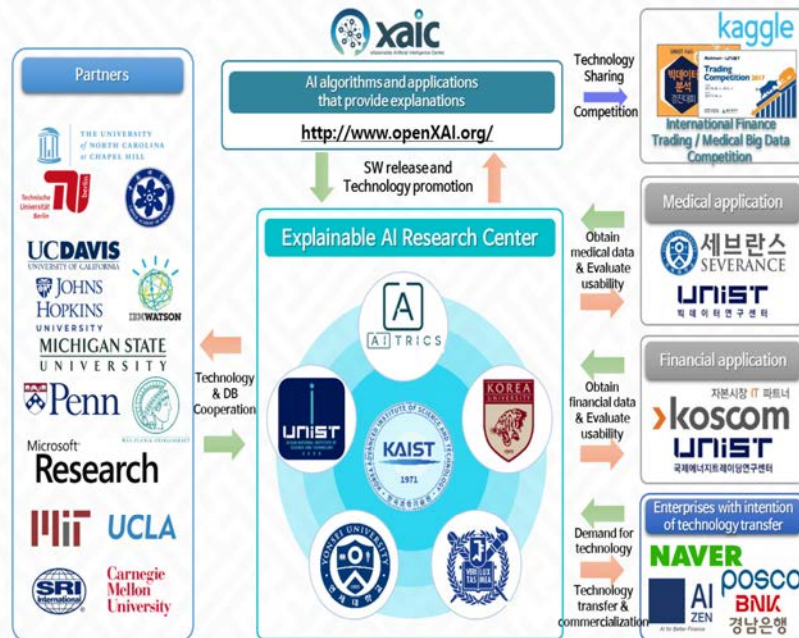
Strategy to Realize AI Trustworthy in Korea

June 2021



Explainable AI Program in Korea

July 2017
- Present



International Standard of XAI

Participant	Title	Organization	Stage	Number	Date	Country
Jaeho Lee	Objectives and methods for explainability of ML models and AI systems	ISO/IEC JTC 1/SC 42	NP	ISO/IEC NP TS 6254	2020-11-16	Switzerland



ISO/IEC JTC 1/SC 42 N 782

ISO/IEC JTC 1/SC 42 "Artificial Intelligence"
Secretariat: ANSI
Committee Manager: Benko Heather Ms.



Official Form 4 - NP - Information technology -- Artificial intelligence -- Objectives and methods for explainability of ML models and AI systems

Document type	Related content	Document date	Expected action
Ballot / Reference document	Project: ISO/IEC NP TS 6254 Ballot: ISO/IEC NP TS 6254 (restricted access)	2020-11-16	VOTE by 2021-02-09

Description

SC 42 N 782 is a NP for ballot to approve the proposal "Information technology -- Artificial intelligence -- Objectives and methods for explainability of ML models and AI systems" and has also been issued via the electronic balloting procedure with the ballot opening on 17 November 2020. SC 42 N 711 is the Draft Document related to the Form 4 contained in SC 42 N 782. Votes should be submitted by 9 February 2021. Any comments submitted with votes should be provided in the standard format.

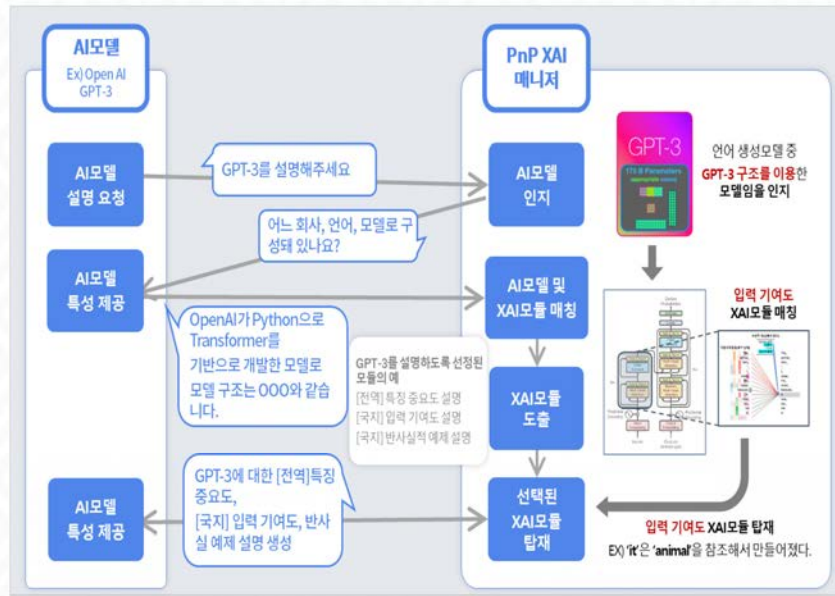


국립전파연구원
National Radio Research Agency

• The First International Standard on XAI Initiated by Korea

Plug and Play Explainable Artificial Intelligence

사용자 맞춤형 플러그앤플레이 방식의 AI모델 설명가능성 제공 기술 개발



개인정보 보호법

[시행 2023. 9. 15.] [법률 제19234호, 2023. 3. 14., 일부개정]

제37조의2(자동화된 결정에 대한 정보주체의 권리 등) ① 정보주체는 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 개인 정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외하며, 이하 이 조에서 “자동화된 결정”이라 한다)이 자신의 권리 또는 의무에 중대한 영향을 미치는 경우에는 해당 개인정보처리자에 대하여 해당 결정을 거부할 수 있는 권리를 가진다. 다만, 자동화된 결정이 제15조제1항제1호·제2호 및 제4호에 따라 이루어지는 경우에는 그러하지 아니하다.

② 정보주체는 개인정보처리자가 자동화된 결정을 한 경우에는 그 결정에 대하여 설명 등을 요구할 수 있다.

③ 개인정보처리자는 제1항 또는 제2항에 따라 정보주체가 자동화된 결정을 거부하거나 이에 대한 설명 등을 요구한 경우에는 정당한 사유가 없는 한 자동화된 결정을 적용하지 아니하거나 인적 개입에 의한 재처리·설명 등 필요한 조치를 하여야 한다.

④ 개인정보처리자는 자동화된 결정의 기준과 절차, 개인정보가 처리되는 방식 등을 정보주체가 쉽게 확인할 수 있도록 공개하여야 한다.

⑤ 제1항부터 제4항까지에서 규정한 사항 외에 장애인회생위원의 직무, 임명, 해임, 임의의 파견, 해임, 해임의 사유, 파견의 사유, 파견의 기간, 파견 및 개인정보가 처리되는 방식의 공개 등에 필요한 사항은 대통령령으로 정한다.

[본조신설 2023. 3. 14.]

[시행일: 2024. 3. 15.] 제37조의2

감사합니다

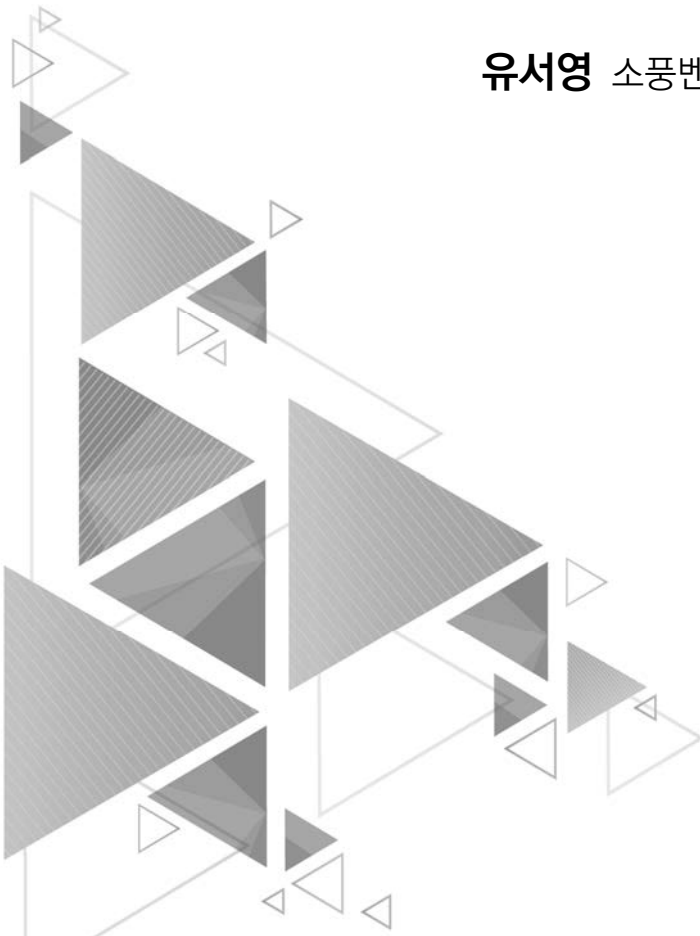
jaesik.choi@kaist.ac.kr
jaesik@ineeji.com



토론 1

기후테크 투자와 스타트업, 그리고 과제

유서영 소풍벤처스 기후네트워크TF 팀장



sopoong the impact capital and accelerator

기후테크 투자와 스타트업, 그리고 과제

유서영 소풍벤처스 기후네트워크TF장



소풍벤처스는 한국, 동남아시아, 북미의 기후테크에 투자합니다.

국내 최초 임팩트 투자, ESG 전문 투자사(2008년 설립)

국내 및 글로벌 131개 포트폴리오에 투자

기후테크 전문 투자사 (총 47개사, 2022년 21개사 투자)

엑셀러레이터 중 선호도 상위 1.5% (약 400곳 중 6위)

기술창업 TIPS, Seed-TIPS 운용사 (TIPS 선정률 94%)

창업가, 글로벌 기업 등 전문성 보유한 운용인력

대기업, 정부, 학계, NGO 등 광범위한 네트워크

포트폴리오 소재국 현황



[기후테크에 집중 투자]

기후테크
37.5%



[초기기업에 집중 투자]



Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

에너지, 순환경제, 농식품 분야 전반에 걸쳐 초기단계 기후테크 스타트업에 투자합니다.

기후테크 투자 하이라이트

기후테크 포트폴리오 수	48개사
기후 포트폴리오 중 (Pre)Seed 투자 비율	77.1%
기후 포트폴리오 중 리드투자비율	79.2%
기후 포트폴리오 중 후속투자유치율 ¹	63.9%

기후테크 포트폴리오 지도



1. 후속투자유치율: 투자 후 1년이 경과한 기후테크 포트폴리오 중 후속 투자를 유치한 비율 M&A의 경우도 포함
(투자 후 1년 경과 포트폴리오 36개사 중, 후속투자유치 및 M&A된 포트폴리오 23개사)

(자료 데이터 기준시점: 2023년 9월 30일 / 단위: 원, %)

Copyright ©2023. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

분야별 기후테크 주요 포트폴리오 사례

분야별 기후테크 주요 포트폴리오 사례 - 에너지

- 2050년 Net-zero 목표 달성을 위해서는 지금부터 2030년까지 재생에너지 영역에 전 세계적으로 35조 달러, 연간 약 5조 달러 이상 투자 필요, 한편 2023년에는 1.7조 달러가 투자될 것으로 전망, 전년 대비 82% 증가(2022년 1.4조 달러)하며 빠르게 성장 중이나, 여전히 필요 투자금액의 1/3 수준에 그침. (국제재생에너지기구·World Energy Transitions Outlook 2023: 1.5°C Pathway, 국제에너지기구·World Energy Investment 2023)
- 동남아는 연평균 2,100억 달러 투자가 필요하다고 전망되나, 2016~2020년 연평균 투자금액 700억 달러로 1/3수준 (국제재생에너지기구)

Main Portfolios

60Hertz 재생 에너지 발전량 예측 기반 가상발전소(VPP) 개발 운영 누적투자유치액 28억 / 멀티플 3.9배 * 주요성과 국내 최초/최대 규모 재생에너지 발전소 데이터 플랫폼, 한전, SK, GS 등 공공/대기업 VPP 솔루션 제공, CES 혁신상 수상	Hy 국내 최초 액화수소 원천기술 기업 누적투자유치액 213억 / 멀티플 1.4배 주요성과 '22년 120억 수소충전소 사업 수주, '23년 약 200억원 수주 예상 Plug Power, Volvo, 현대차 등 국내외 대기업과 협업
aeonius 2차 전지 배터리 기반 이동형 ESS(Energy storage system) 누적투자유치액 30억 / 멀티플 2.14배 * 주요성과 '23년 인도고 판매 10억 이상 기대, SK EnS 충전서비스 런칭, SKT 매출확약	EA 수소충전소 EPC, P2G(Power to Gas) 설계 시공 전문회사 누적투자유치액 21억 / 멀티플 1.3배 주요성과 한화솔루션, 현대로템, 한전 등 P2G 사업 참여, '22년 매출 38억
SEOLDYNAMICS 40톤급 무인 전기 대형 트럭 개발 스타트업 누적투자유치액 8억 / 멀티플 2.3배 주요성과 삼성, 현대한국조선해양 등 공장 내 이송로봇 수주	SELEX 전기 오토바이 및 배터리 제조, 교체 기술 스타트업 (베트남) 누적투자유치액 \$5.1M (한화 약 66억 원) 주요성과 동남아 배달업 1위 그룹, 라자다 등 배터리 공급
KOMS 부유식 해상 풍력 기술 개발 스타트업 누적투자유치액 4.5억 / 멀티플 - 주요성과 한국에너지기술연구원 50억 R&D 자금 확보	

* 멀티플 기준 시점: 2023.05.31 / 기업가치: 투자 후 기업가치 기준 / 누적 투자금: 소문 투자금 포함 피투자기업의 현재 누적 투자유치금액 / 멀티플: (이 기업별 현재지분 평가액÷투자액)×투자금액
* 표시는 현재 투자유치액에 진행 중인 투자금 포함 금액이 아닌 투자 확정된 금액을 포함하여 계산

Copyright ©2023. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

분야별 기후테크 주요 포트폴리오 사례

sopoong
the impact capital and accelerator

분야별 기후테크 주요 포트폴리오 사례 - 순환경제

- 순환 경제로의 전환은 2030년까지 4조 5천억 달러의 글로벌 시장 성장 기회를 창출 (엑셀큐어, Waste to Wealth)
- 매년 약 1,000억 톤의 천연 자원이 추출, 2050년 필요 자원의 규모가 2배 이상 증가할 것으로 전망 (세계경제포럼, Circularity Gap Report 2023)
- 전 세계 순환성이 2018년 9.1%에서 2020년 8.6%, 2023년 7.2%로 감소세 (세계경제포럼, Circularity Gap Report 2023)

Main Portfolios

Testvalley 선경험-후구매 전자제품 유통 플랫폼 누적투자유치액 114억 / 멀티플 21배 주요성과 누적고객 10만 명, 누적매출 300억 이상, '23년 매출 900억 목표 Exit(부분매각)	RE.harvest 국내 최초 푸드 업사이클링 전문 기업 누적투자유치액 213억 / 멀티플 1.4배 주요성과 리너지 가루 1Kg 당 탄소 11kg 절감 효과, 두레주르, 마스터피자, OB맥주 등 협업 Exit(부분매각)
AETECH AI 기반 폐기물 자동선별 로봇 솔루션 누적투자유치액 30억 / 멀티플 2.14배* 주요성과 95종 재활용 폐기물 최대 99.3% 정확도로 분류, '22년 첫 납품 후 발주 문의 지속	REPLA 재활용 플라스틱 순도 향상을 위한 전처리 솔루션 누적투자유치액 43.5억 / 멀티플 1배 주요성과 바이오링크 양산화 검토, 캐시카우 제품 '플라스틱 순도측정기기' 개발
300 릴라루프 플라스틱과 폐기물 문제를 해결하는 스마트 자원회수 인프라 누적투자유치액 5.5억 / 멀티플 6.4배 주요성과 누적 설치 인프라 755대, 누적 매출 53억, 누적 가입자 8만 명	Microalgae Ark Co 미세조류 스마트팜 기반 천연물 바이오 소재 전문기업 누적투자유치액 20억 / 멀티플 1.7배 주요성과 국내 최대 배양장(500톤) 설비 구축, 종근당 바이오 등 소재 개발인정 진행
RELAY 패션브랜드 B2B 중고마켓 솔루션(Resale as a Service) 누적투자유치액 10억 / 멀티플 1.7배 주요성과 코오롱물, 자전거의류, 유아동복 등 기업 중고마켓 위탁운영	CAPTURE+ 이산화탄소 직접포집과 청정수자원 공급이 가능한 DAC(직접공기포집) 솔루션 (미국) 누적투자유치액 \$6M (한화 약 78억) / 멀티플 1배 주요성과 미국, 이스라엘, 뉴질랜드 등에서 대규모 DAC 프로젝트 진행

*데이터 기준 시점: 2023.05.31 / 기업가치: 투자 후 기업가치 기준 / 누적 투자금: 순출 투자금 포함 자회사기업의 전체 누적 투자금 / 멀티플: (각 기업별 현재가치/투자액) × 투자금

* 멀티플은 현재 투자비용으로 인해 동일 투자금 납입 횟수나 투자 확립된 것을 포함하여 계산

Copyright ©2023. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

분야별 기후테크 주요 포트폴리오 사례

sopoong
the impact capital and accelerator

분야별 기후테크 주요 포트폴리오 사례 - 농식품

- '22년 한 해 글로벌 농식품 벤처투자 296억 달러 (Agfunder, Temasek-Global AgriFoodTech Investment Report 2023)
- 2030년까지 베트남 농업부문 외국인투자 목표치 340억 달러 (베트남 정부 발표, 2022년 12월)
- 러-우 전쟁, 공급망 중단, COVID-19의 여파 등으로 2023년 한 해 3억 4,500만 명 이상이 높은 수준의 식량 불안에 직면할 것으로 추정, 이는 2020년의 두 배 이상 수준 (유엔세계식량계획, 2023)

Main Portfolios

리스토인더 국내 No.1 식품 마감할인 플랫폼 누적투자유치액 114억 / 멀티플 21배 주요성과 '22년 기준 세븐일레븐·씨유 등 가맹점 수 4만 곳, 누적 가입자 수 60만 명, MAU 20만 명, 재구매율 64.5%	AIS 데이터 기반 노지 스마트팜 솔루션 누적투자유치액 27.3억 / 멀티플 3.39배 주요성과 서비스 면적 '22년 65만평 달성, '23년 200만평 (연매출 환산 약 25억) 목표
VANDAL SOFT 국내 No.1 식용곤충 스마트팜 기업 누적투자유치액 4.8억 / 멀티플 9.5배 주요성과 베트남 농장 설립, 베트남 타겟 반려동물 사료 개발	KOF 한국수산기술연구원 데이터 기반 스마트 아쿠아팜 솔루션 누적투자유치액 15억 / 멀티플 3.39배 주요성과 기존 양식장 대비 생산성 10배, 도입의향 양식장 5곳 확보
SWEETEGG 국내 유일 식물성 계란 전문 기업 누적투자유치액 18억 / 멀티플 2배* 주요성과 CU 비엔 에그마요 샌드위치 전 지점 납품, BGF리테일, CJ제일제당 등 대기업 납품	WEMEET 버섯으로 만든 정육 고기 솔루션 누적투자유치액 9.8억 / 멀티플 2.2배* 주요성과 전국 레스토랑 B2B 납품 80곳 돌파, 치킨/식품 브랜드 발주, 납품 논의
NEUROPACK 친환경 기능성 포장재 개발 기업 누적투자유치액 10억 / 멀티플 - 주요성과 롯데중앙연구소와 친환경 기능성 포장재 PoC	채식한끼 국내 1위 채식 중합 플랫폼 누적투자유치액 3.2억 / 멀티플 7.3배 주요성과 등록식당 6,000여 곳, 등록메뉴 약 2만개, 유저리뷰 1.5만개 달성

*데이터 기준 시점: 2023.05.31 / 기업가치: 투자 후 기업가치 기준 / 누적 투자금: 순출 투자금 포함 자회사기업의 전체 누적 투자금 / 멀티플: (각 기업별 현재가치/투자액) × 투자금

* 멀티플은 현재 투자비용으로 인해 동일 투자금 납입 횟수나 투자 확립된 것을 포함하여 계산

Copyright ©2023. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

THEME OF THE YEAR | 2023.10.19 (목) ~ 10.21 (토)

기후기술과 인공지능

AI FOR FIGHTING AGAINST THE CLIMATE CRISIS

DATA & TECHNOLOGY

SHIFT

CLIMATE SOLUTION

DAY 1 | CLIMATE TECH HERE AND NOW

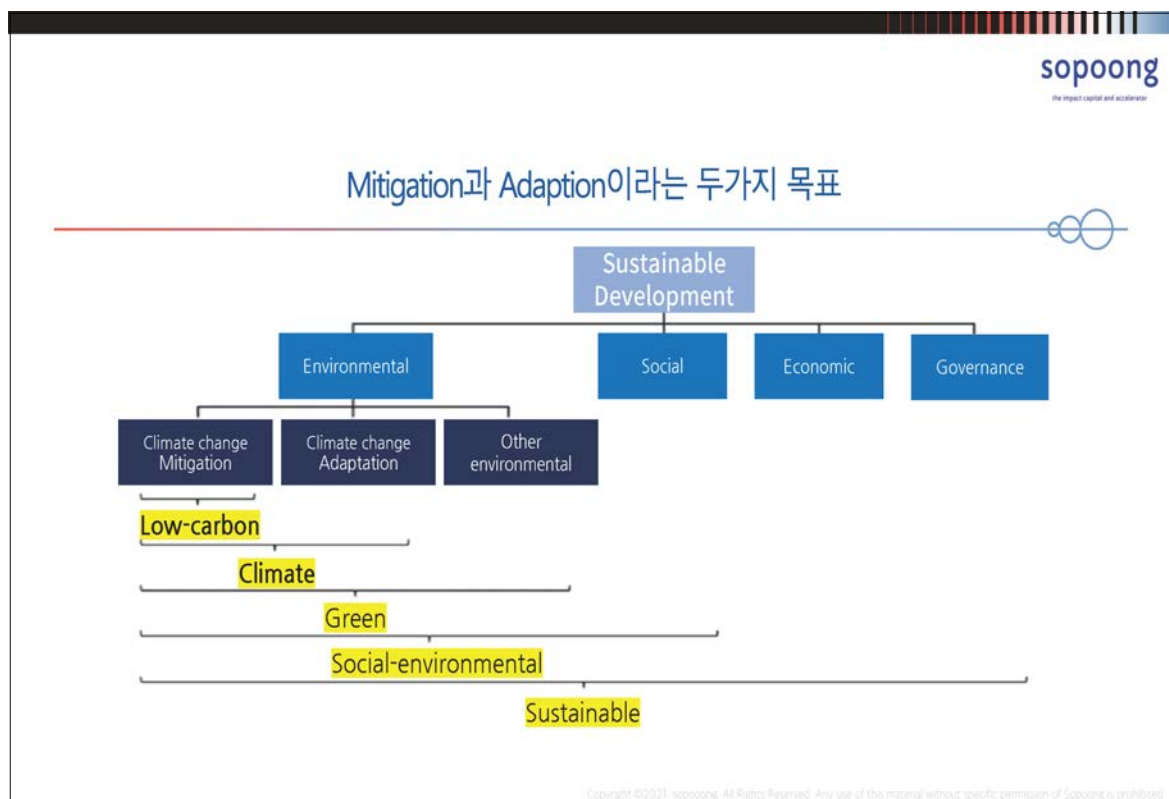
지난 1년간 기후기술 생태계에는 어떤 일들이 있었나?

DAY 2 | HOW AI CAN BE A POWERFUL TOOL IN THE FIGHT AGAINST CLIMATE CRISIS

인공지능은 어떻게 기후문제를 해결하는 강력한 도구가 될 수 있을까?

DAY 3 | ACT ON CLIMATE CRISIS : AI FOR CLIMATE

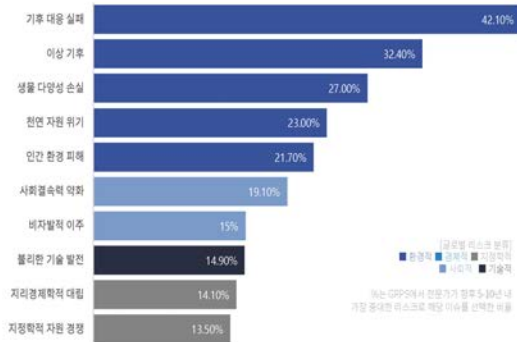
2023 CLIMATE TECH STARTUP SUMMIT



인류가 직면한 가장 큰 문제, 기후위기

- ① IPCC¹는 지구 평균 기온이 산업화 이전 대비 1.5-2°C 초과 시 돌이킬 수 없는 기후 재앙에 직면하게 되며, **2050년까지 Net Zero를 달성**해야만 이를 막을 수 있을 것으로 전망.
- ② 세계경제포럼(WEF)에 따르면 향후 5-10년 내 **가장 중대한 글로벌 리스크 상위 5가지 모두** 기후위기 관련 이슈에 해당함.
- ③ 인류의 생존과 전세계 경제·사회 지속가능성을 위해 **탈탄소사회로의 대전환은 필연적임**.

[향후 5-10년 내 가장 중대한 위협이 될 글로벌 리스크]



IPCC: 기후변화에 관한 정부간 협의체, 인간 활동에 의한 기후변화의 위험을 평가
출처: The Global Risk Report 2022 (WEF은 매년 각 분야 전문가 1,200명 대상 글로벌 리스크 인식조사(SRP)) 실시, 세계 경제에 영향을 미칠 위험(위험)을

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

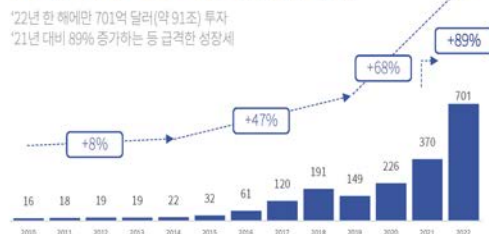
기후테크로 몰리는 글로벌자본

- ① 미·중·유럽 등 경제대국들은 기후위기를 신성장동력 기회로 삼고 **기후테크 산업 주도권 장악**을 하고자 대규모의 투자·정책지원 추진
- ② 벤처투자 시장 전반의 위축 경향과 달리 **기후테크 투자는 급격한 성장**을 보이고 있음.

[미·중·유럽의 기후 관련 예산 및 투자액]

	3,690억 달러(약 480조) → IRA 기후 관련 예산('22~'31년)
	5,460억 달러(약 710조) → '22년 한해 청정에너지 투자액
	2,100억 유로(약 281조) → REPower EU 예산('22~'27년)

[글로벌 기후테크 VC 투자금액 (단위:억 달러)]



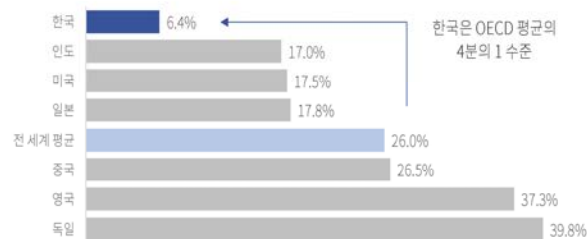
출처: Delving gravity, 2022 Climate Tech VC funding totals \$70.1B, up 89% on 2021. (Piston IQ)

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

기후테크로 몰리는 글로벌자본

- 한국은 세계 8위 온실가스 배출국이자 탄소 다배출 제조업 국가. 제조업·수출 주도 경제로 반세기 만에 눈부신 성장을 이뤄냈으나, 이와 같은 산업 구조로 인해 탄소 중립 이행에 있어 보다 큰 도전을 받고 있음.
- 국제재생에너지기구에 따르면 한국의 신재생에너지 발전량 비율은 6.4%로 전 세계 평균의 4분의 1에 불과, 선진국 뿐 아니라 중국과 인도 등 신흥국에 비해서도 크게 낮은 수준.
- 탄소 규제 확대는 발전·산업 등 여러 산업재와 소비재의 가격을 급격히 증가시켜 수출 중심 경제인 한국에게 새로운 무역장벽으로 작용하며 국가 경제의 위협이 될 것으로 전망.

[2020년 국가별 전력 발전량 대비 재생에너지 발전량 비율]



1. 출처: 한국의 지속가능발전목표(SDG) 이행현황 2023 (통계청)
2. 출처: 넷제로 전환 보고서 (에너지금융연구원, 2022)

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

실물경제와 금융시장 전반의 기후리스크

- 넷제로 사회로의 전환은 농림수산업, 국제 협약을 통한 에너지 이용과 산업생산방식, 수송방식 등 산업, 무역, 경제구조 전반, 일반 시민의 문화와 소비 패턴에까지 영향을 주며 사회경제적으로 전 분야에 광범위한 파급력을 가져올 것으로 예상
- 기상이변으로 인한 피해는 물론 대응을 위한 추가적 비용 역시 천문학적으로 증가할 것으로 예상되는 등 기업 운영과 투자에 있어서 새로운 기회와 리스크가 동시에 등장



출처: 기후리스크와 자산가격의 관계에 대한 조사 및 분석 (자본시장연구원, 조사보고서 23-01)

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

국가 탄소중립-녹색성장 기본계획 중 부문별 주요 목표와 5대 기후테크 영역

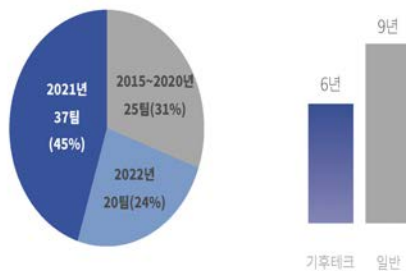


Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

빠른 혁신이 필요한 기후 시장, 스타트업 기회로 대두

- ① '21~'22년 새롭게 등장한 기후테크 유니콘 기업은 총 57팀으로 '15~'20년 누적치가 21팀 인데 반해 최근 성장세가 두드러짐. 또한 유니콘이 되기까지 평균 9년이 소요되는 한편, 기후테크 유니콘은 6년 소요⁴ 이는 기후테크 스타트업들이 시장에서 성장성 및 혁신성을 인정받고 있음을 의미함.
- ② 적극적 정책지원, 인프라 형성과 함께 기술 혁신이 보다 빠르게 이루어지면서 스타트업이 도전할 수 있는 시장 기회들이 보다 증가

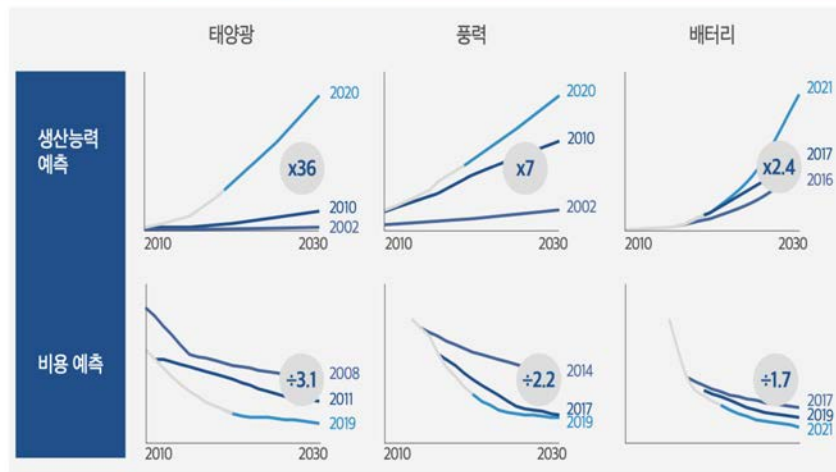
[연도별 신규 기후테크 유니콘 수] [유니콘이 되기까지 소요기간]



4 출처: GLOBAL UNICORN INDEX 2023, Hurun Report (Heron IQ, 2000년~2022년 전 유니콘 기업 1,361개 대상)

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

예측을 뛰어넘는 기술 개발 속도



Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

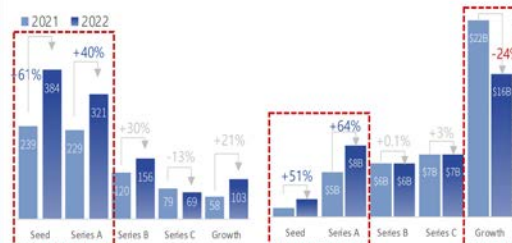
초기 단계로 몰리는 기후 투자

- 기후테크 산업의 태동기에서 전 분야에 걸쳐 기술혁신이 빠르게 이루어짐.
- 글로벌 경제·VC 긴축이 지속됨에 따라 리스크를 축소하고자 분산투자 지향. 대규모 딜은 정체·감소된 한편, **Seed ~ Series A에 대한 투자는 증가.**
- 한편 후기 단계 투자가 신중해진만큼 Dry powder의 여유가 있기 때문에 **초기 기회 포착 시 후속 투자 유치 및 수익률 확보 가능성**이 높아짐.

[기후테크 투자 단계별 투자집행건수(좌측) & 투자금액 합계(우측)]

'21년 대비 '22년은 초기 투자가 강세

- 초기단계에서 투자건수, 투자금액 모두 크게 증가
- Growth 단계 투자는 증가한 한편 전체 투자금액은 축소, 이는 건당 투자 규모가 축소되는 추세이며, 투자결정이 보다 신중해지고 있음을 의미



Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

sopoong
the impact capital and accelerator

국내 및 글로벌 대기업들의 적극적 행보

마이크로소프트 - 브레이크스루에너지 사례

• 2016년 빌게이츠 주도 하에 2050년 온실가스 배출 제로를 목표로 기후테크 투자펀드 'Breakthrough Energy Ventures(BEV)'를 출범, 아마존·소프트뱅크 창업자 등 글로벌 자본가들이 출자하여 10억 달러(한화 약 1.3조) 규모의 1호 펀드, 12.5억 달러(한화 약 1.7조) 규모의 2호 펀드 결성하였으며, 현재까지 **3조원 이상의 펀드** 결성

[BEV 펀드 결성 현황]

BEV 1호 (2016년) 10억 달러 규모	BEV 2호 (2021년) 12.5억 달러 규모	BEV 유럽 (2019년) 1억 유로 규모
-----------------------------	-------------------------------	----------------------------

[BEV 투자 현황]

총 포트폴리오 수 96개	유니콘 기업 수 5개	Exit 기업 수 4개
------------------	----------------	-----------------

[BEV에서 발표한 청정에너지 핵심 기술 후보 목록]

에너지 저장	수소	첨단제조기술	건물에너지 기술
에너지 전력망	CCUS	청안 원자로	생물학적 격리

아마존 - 기후 서약 캠페인 · 기후 서약 펀드 사례

• 2019년 아마존은 글로벌 옵티미즘과 함께 '기후서약' 캠페인 추진. 이는 2050 탄소중립 목표를 10년 앞당겨 2040년까지 달성하자는 선언으로 현재까지 MS, IBM, 유니레버 등 413개 글로벌 기업이 가입하였으며, 기업의 탄소저감 활동과 투자 촉진에 큰 기여를 함.

• 2020년 아마존은 탄소중립을 앞당기는 기후테크 기업에 투자하기 위해 약 20억 달러(약 2.4조원) 규모 '기후서약펀드'를 조성.

[아마존 기후서약펀드 투자현황]

운송 및 물류	투자사태 예시: • 전기 수직이착륙 항공기 제조 스타트업 • 저탄소 콘크리트 제조 스타트업 • 암모니아 전력 기술회사 • 운송용 고밀도 수소솔루션 • 친환경 폐기물 스타트업
에너지 생산·저장·활용	
건설	
제조 및 재료	
순환경제	
식량과 농업	

총 6개국, 22곳 투자

국내 대기업 기후테크 스타트업 지원 사례

[기후테크 펀드 출자]

[기후테크 직접투자 / CVC]

[오픈이노베이션 프로그램(기후 관련 분야)]

1. 출처: Advancing the Landscape of Clean Energy Innovation (Breakthrough Energy 2019)

Copyright ©2021. sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

sopoong
the impact capital and accelerator

Why 기후테크 AI

Complex Data Processing

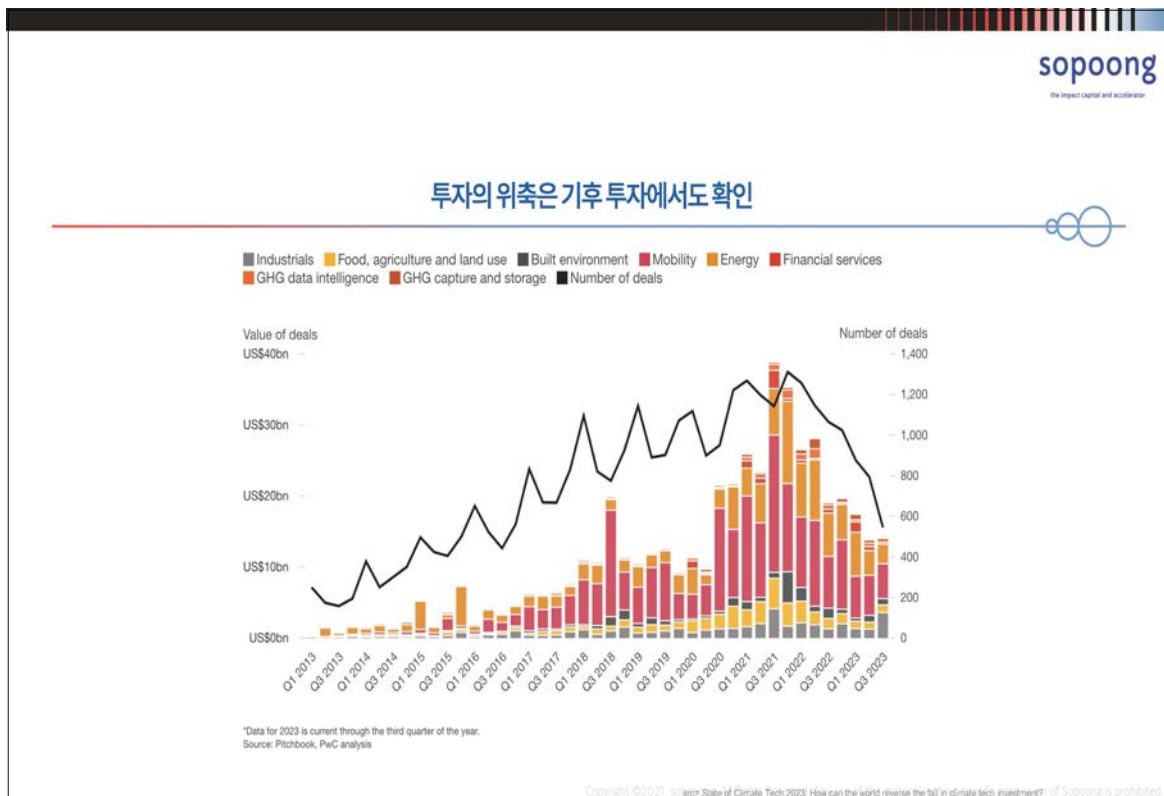
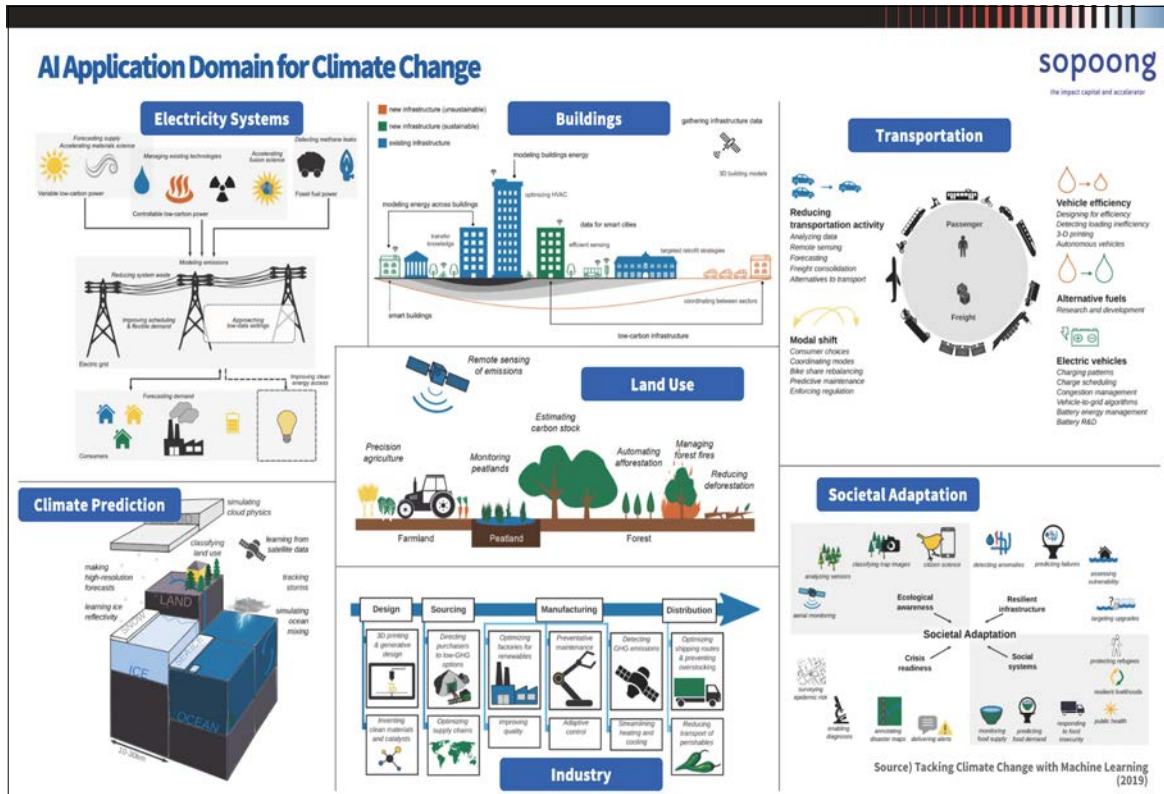
- Integrate & source insights from diverse data sources in real-time
- AI models can handle a wide variety of massive inputs

Adaptive to Environment

- Unforeseen changes result in the need to adapt software & processes
- AI models can self-adjust limiting the need for constant & regular intervention (retraining & transfer learning)

Knowledge Capture

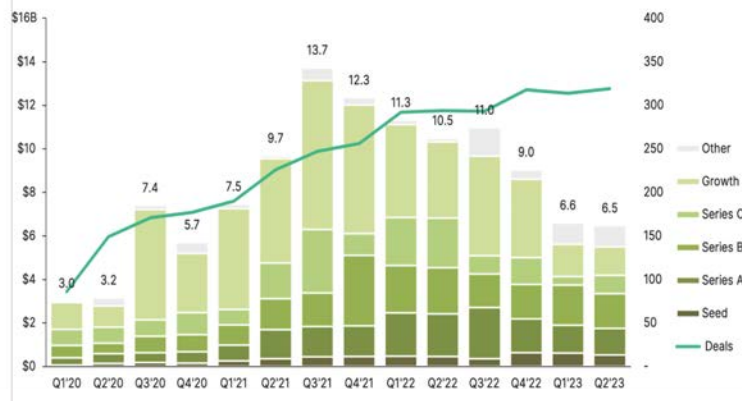
- Capture the expertise from your human workforce
- AI models can capture the deep expertise of professionals in complex domains, empower them through active learning, explainability



투자 총액은 40% 감소, 투자 건수는 8% 증가

\$13B in climate tech funding in H1'23, down 40% compared to prior year

Venture capital investment (\$B) and deal count (#) in climate tech by stage



Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

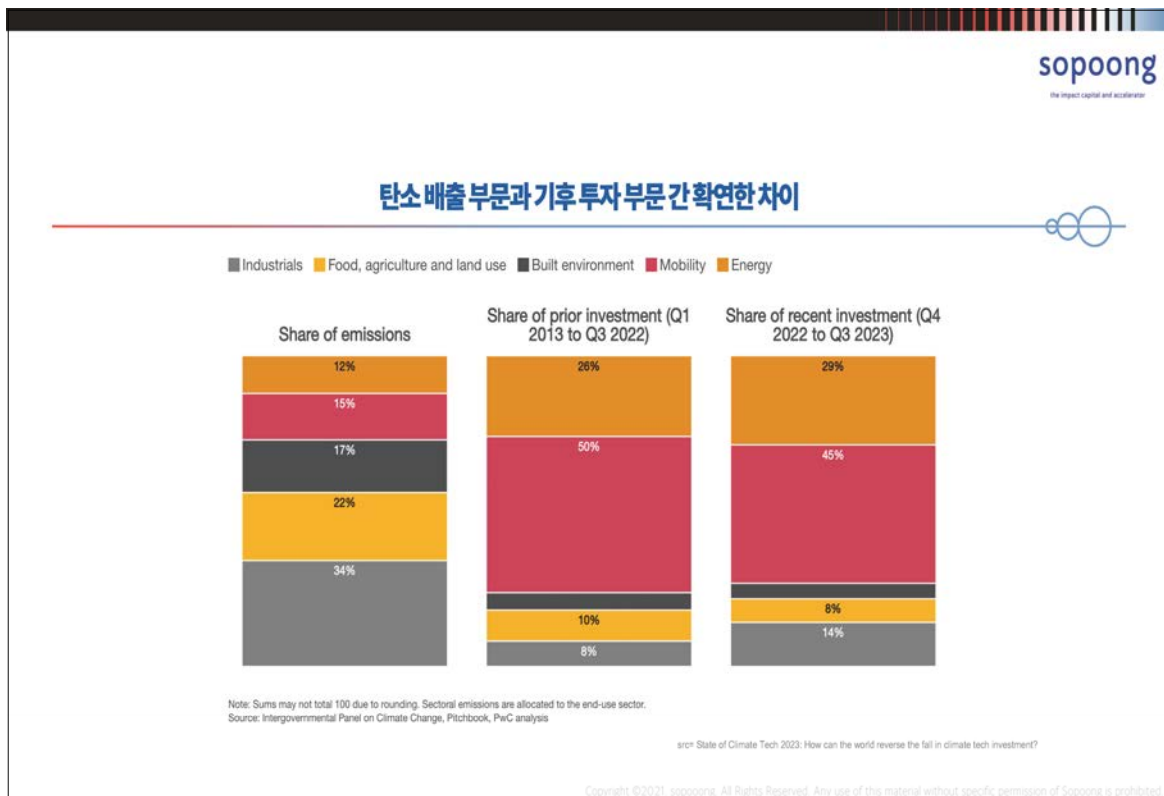
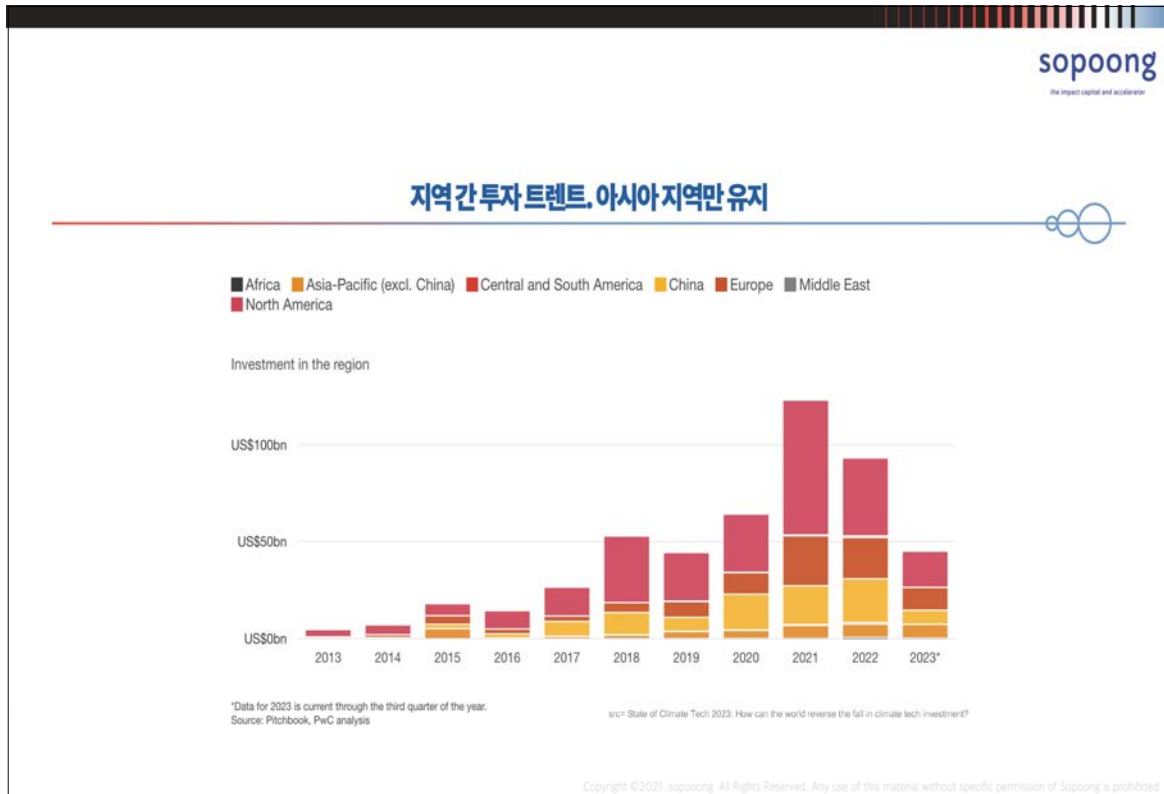
시장은 침체지만 기후테크의 점유율은 증가



*Data for 2023 is current through the third quarter of the year.
 Note: See the second footnote, below, for more on how the percentages in this chart are calculated.
 Source: Pitchbook, PwC analysis

s17c State of Climate Tech 2023: How can the world reverse the fall in climate tech investment?

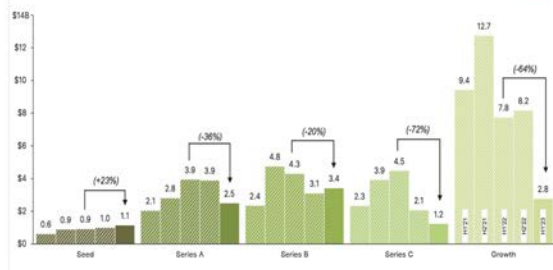
Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.



초기 투자는 여전히 호황, 후기 투자는 경색

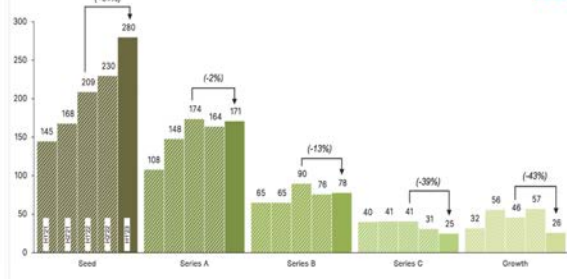
Series B through Growth funding fell >50% YoY

Change in venture capital investment by stage (\$B)



Early stage deal activity up 20%, Mid & Late deal activity declined

Change in venture capital deal count by stage (#)



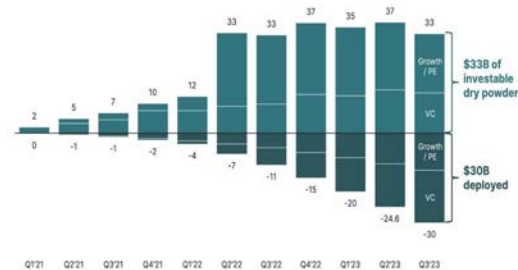
src= Global Landscape of Climate Finance by Climate policy institute

Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

투자 가능 자금(Dry powder)는 계속 늘어

\$33B of investable dry powder for climate

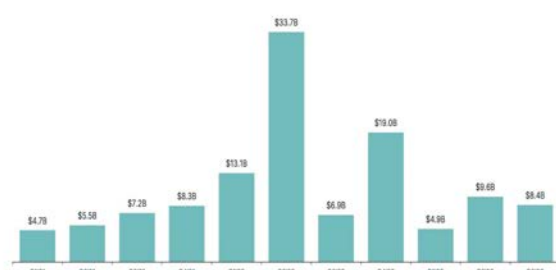
Cumulative investable vs cumulative deployed capital from announced climate-specific venture and Growth / PE funds (\$B)



Note: VC represents VC and CVC. Q3'23 based on data available until September 7 2023.

Fundraising back to pre-2022 levels

Sum of new fund's AUM by final close announcement quarter (\$B)

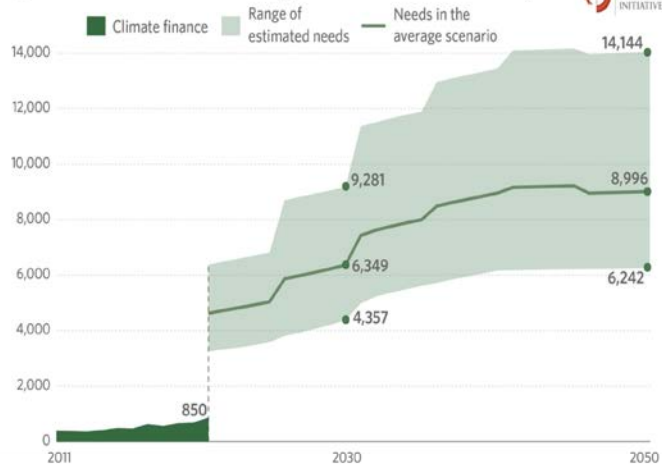


src= Global Landscape of Climate Finance by Climate policy institute

Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

Net Zero까지는 여전히 부족한 자금

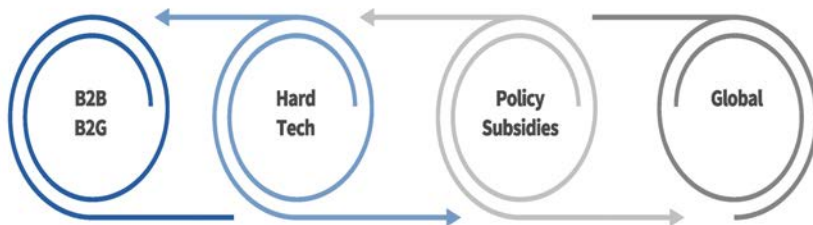
Figure 2: Global tracked climate finance flows and the average estimated annual climate investment need* through 2050



src= Global Landscape of Climate Finance by Climate policy institute

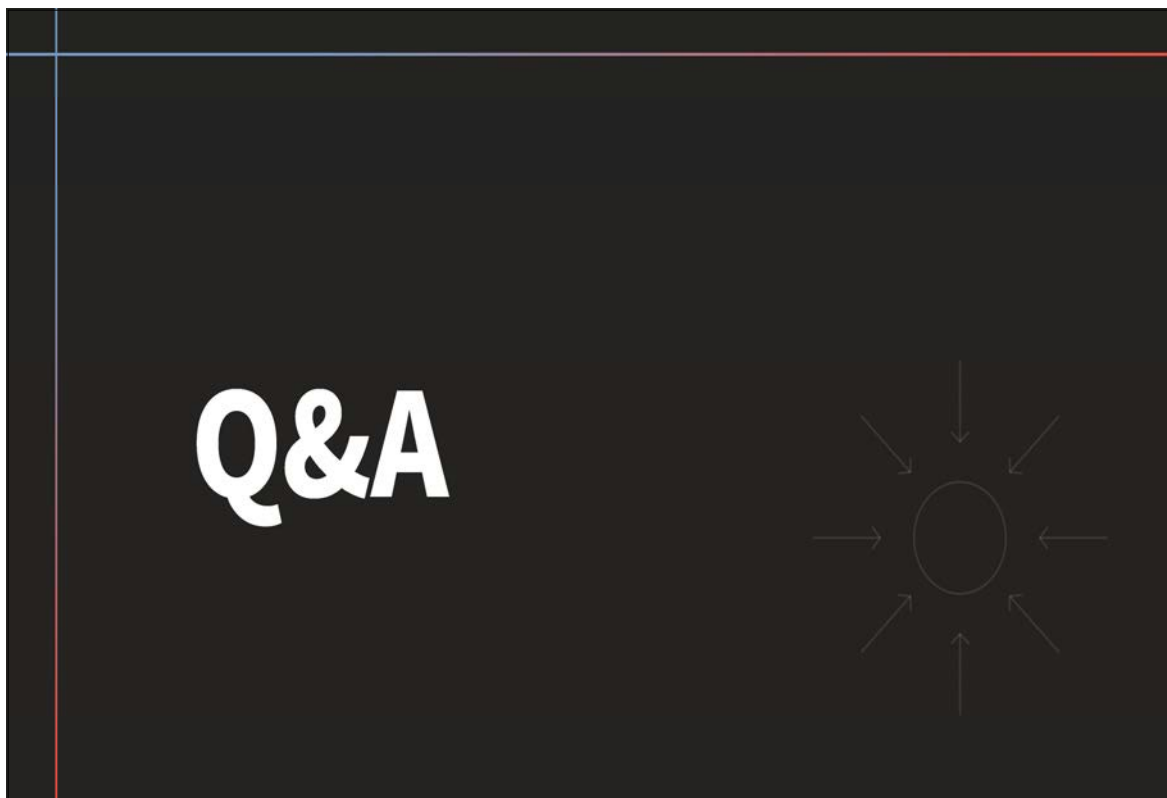
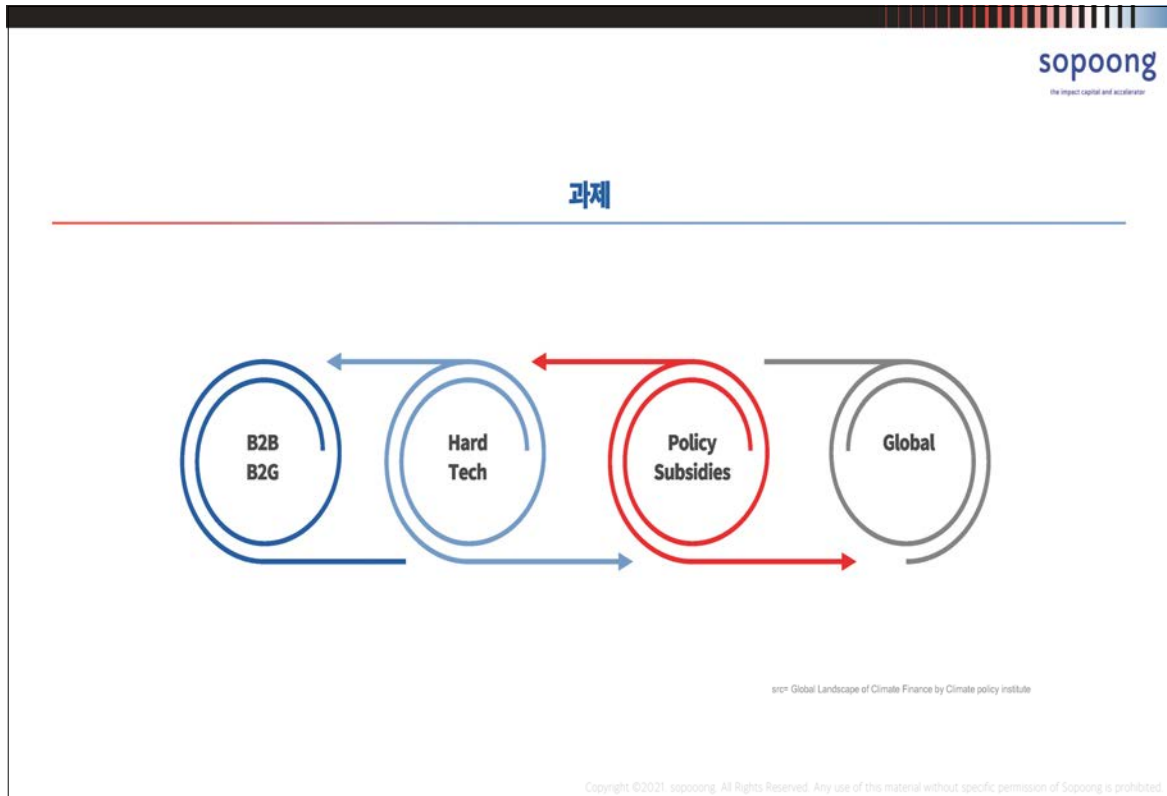
Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

기후테크의 특징



src= Global Landscape of Climate Finance by Climate policy institute

Copyright ©2021, sopoong. All Rights Reserved. Any use of this material without specific permission of Sopoong is prohibited.

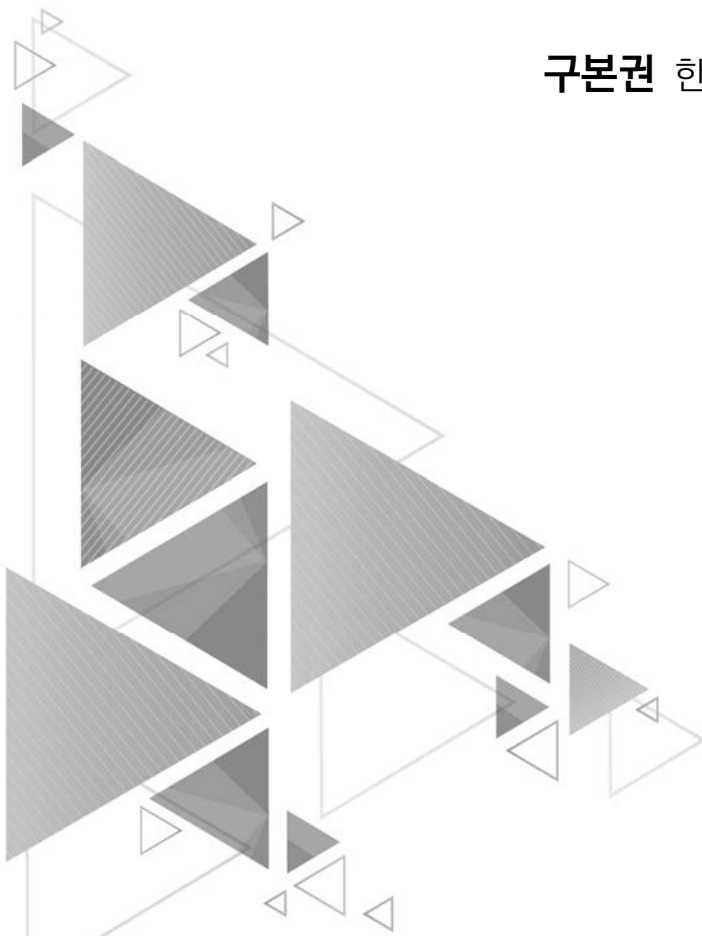




토론 2

모두를 위한 AI 토론문

구본권 한겨레 사람과디지털연구소장



‘모두를 위한 AI’ 토론문

구본권 한겨레 사람과디지털연구소장

‘모두의 AI : 정책과 전략’이라는 토론회 주제는 인공지능 시대에 우리 사회가 직면한 과제를 핵심적으로 요약한다. ‘모두의 AI’라는 표현은 현재의 인공지능이 모두를 위한 것도 아니고 모두의 소유도 아닌 상황이라는 인식을 담고 있다. 그렇다면 현재는 누구의 소유인가라는 문제의식에서 토론문을 시작해본다.

마이크로소프트(MS) 공동창업자 빌 게이츠는 2023년 2월 “인공지능 시대가 열렸다”는 글을 올려 챗지피티가 그래픽사용자환경(GUI)만큼 혁명적인 기술이라고 주장했다. 그래픽사용자환경은 그 이전까지 소수 전문가들이 프로그램 명령어를 통해 조작하던 컴퓨터를 만인의 도구로 만든 혁신기술이다. 직관적 사용법의 그래픽사용자환경이 실질적으로 정보화시대를 연 것처럼, 챗지피티는 누구나 인공지능을 자유롭게 다양한 용도로 활용하는 시대를 불러올 것이라는 얘기다.

누구나 스마트폰과 앱을 조작하듯 자유롭게 인공지능을 사용하는 환경은 산업과 일자리, 교육, 예술창작 등 사회 전 분야에 충격적 변화를 예고한다. 생성 인공지능을 활용해 생산성을 높이고 업무를 자동화하려는 다양한 시도는 이미 시작됐고 일자리가 없어지고 교육에서 영어 인용과 표절이 일반화할 현상에 대한 불안과 우려도 높다. 무엇보다 사람보다 뛰어난 인공지능의 등장은 공동체의 구성원리인 민주주의체제를 근본적으로 위협하는 요인으로 작용할 가능성이 크다. 왜냐하면 공동체에서 가장 강력하고 합법적인 권력을 만들어내는 구조이자, 강한 권력집단을 견제하도록 설계되어 있는 민주주의 시스템은 현재처럼 인공지능이 사회의 기본적인 환경이 되는 상황을 전혀 고려하지 않고 만들어졌고 운영되어 왔기 때문이다.

대의 민주주의는 언론과 표현의 자유를 통해 형성되는 여론 기반의 민의 수렴장치를 요청하는데, 인공지능으로 인해 구조적 변화가 발생하고 있다. 소셜미디어와 인공지능이 등장하기 전까지 정보와 여론은 사람에 의해 만들어졌지만 인공지능으로 인해 근본적 변화가 생겨났다. 앞으로는 사람이 만들어내는 정보와 여론보다 인공지능과 같은 기술에 의해 자동 생성되는 분량이 훨씬 많아지고 개인과 사회에 큰 영향력을 끼칠 가능성이 있다. 근대 계몽주의와 시민혁명을 겪으며 형성된 근대 민주주의는 인식과 판단의 주체로 독립된 인간을 상정한다. 그런데 글쓰기와 논리구성, 판단 등 인간만이 수행하던 지적 활동 영역에서 평균적 인간보다 뛰어난

능력의 인공지능이 등장한 것이다.

민주주의는 공동체가 최악을 피하고 차악을 선택할 수 있게 해주는, 불완전하지만 현실적인 정치권력 창출 방법으로 합의된 사회계약이다. 민주주의는 오랜 기간에 걸쳐 사상과 현실의 영역에서 사회 각 집단이 갈등 조정과 투쟁, 혁명을 거치고 토론과 합의를 통해 만들어낸 결과다. 인공지능은 인간만을 인식과 사회적 행동의 주체로 여겨온 민주주의라는 사회적 시스템의 전제를 흔들며 새로운 관점을 요청하고 있다. 인공지능처럼 의식은 없지만 인간과 유사한 지적 능력을 수행하는 존재가 등장해 개인과 사회에 영향을 끼치고 있는 상황이기 때문이다. 인공지능으로 인한 변화중에서도 모든 사람에게 장기적이고 포괄적인 영향을 끼칠 영역은 사회가 가장 거대한 권력을 다루는 방식, 즉 정치체제에 끼칠 영향이다. 민주주의에 생성형 인공지능이 끼칠 영향은 세 가지 차원에서 살펴볼 수 있다.

첫째, 탈진실 현상을 심화시킨다.

챗지피티로 대표되는 생성인공지능의 주요 특징은 거대언어모델 학습을 통해 문답이 이뤄지는 상황과 담화의 맥락을 이해하고 특정 단어 뒤에 배치될 확률이 높은 단어를 예측해 제기하는 기능이다. 다음에 이어질 단어를 예측하는 기술이라는 점에서 검색어나 문자메시지 자동완성과 유사하지만, 놀라운 정확도로 이용자들을 놀라게 하고 있다. 하지만 이는 언어철학자 존 설이 ‘중국어 방’ 논증을 통해 주장한 것처럼, 인공지능이 정확하고 뛰어난 답변을 하는 것처럼 보여도 실제로는 아무것도 모르는 상태에서 단어를 높은 확률로 구성하는 것일 뿐이다. 챗지피티는 그럴듯한 문장을 제시하지만 그 의미를 전혀 이해하지 못하고 있을 뿐만 아니라 사실이 아닌 황당한 이야기를 자연스러운 문장과 줄거리로 만들어낸다. 챗지피티의 주요 기술적 특징 중 하나는 허위정보를 진짜처럼 묘사하는 ‘환각(Hallucination)’ 현상이다.

인공지능을 활용한 이미지 조작에 활용되는 딥페이크 기술과 이용자들을 현혹하는 정보를 무차별적으로 유통하는 소셜미디어가 일상화하면서 정보 생태계에서 사실과 가짜의 구분이 점점 어려워지고 있다. 탈진실 현상을 부추기는 세력들에게는 생성형 인공지능이 인화력 높은 새로운 도구로 쓰일 수 있다.

2016년 영국의 브렉시트 국민투표와 미국의 대통령선거에서 허위정보가 범람하고 여론과 정치적 판단에 큰 영향을 끼치는 현상이 나타났고, 그해 말 영국 <옥스퍼드 사전>은 ‘탈진실’(Post Truth)을 ‘올해의 단어’로 선정했다. 탈진실 현상은 2016년만 두드러진 현상이 아니고 정보사회에서 갈수록 일상화하고 있다. 민주주의를 위협하는 기술적 환경이다. 현실 정치에서 유력 정치인들은 자신의 생각과 다른 사실이나 뉴스에 대해 공공연히 ‘가짜뉴스’라고 주장하며 지지자들을 선동하고 있는 현실이다.

미국의 컨설팅업체 가트너는 2017년 10월 ‘미래 전망 보고서’를 통해 “2022년이 되면 선진국 대부분의 시민들은 진짜 정보보다 거짓 정보를 더 많이 이용하게 될 것이다”라고 예측했는데, 현실이 된 셈이다. 사실에 기반한 인식과 여론 형성을 어렵게 만드는 기술이 등장하고 사람들이 진짜 정보보다 가짜 정보를 더 많이 만나고 이용하게 될 것이라는 점은 민주주의의 기반을 흔드는 문제다.

둘째, 기계에 의한 편견과 차별이 고착화할 수 있다.

업무 자동화도구인 인공지능은 사회 각 영역에 침투해 비효율을 개선하고 편의성을 높이지만, 기존의 기술과 달리 판단의 근거와 이유를 설명하지 않는다. 지금까지 대부분의 도구는 인간이 설계한 방식대로 작동했고 조작자는 그 원리를 이해하고 조작했다. 그런데 이세돌-알파고 대국에서 알파고는 왜 그 자리에 돌을 놓았는지를 설명할 수 없다. 알파고 설계자나 조작자도 설명할 수 없다. 어떤 연산과정을 통해 결과가 도출됐는지 알지 못하지만 결과가 효과적이기 때문에 사람들은 적극적으로 채택한다.

인공지능 기계학습은 뇌 구조를 모방한 신경망 형태의 은닉층(hidden layer)에서 딥러닝을 통해 진행된다. 인공지능의 심화신경망은 서로 복잡하게 연결된 수백개의 계층에서 수백만개의 매개변수들이 상호작용하는 구조로, 사람이 인지하는 게 불가능하다. 인공지능의 결과와 효율성은 탁월해졌지만 알고리즘은 더 불투명해지고 인간 인지 영역을 넘어섰다. 인공지능은 그 처리과정을 알 수 없는 블랙박스 기술이다. 인공지능은 미국에서 채용, 직원 평가, 법원판결 등에 널리 활용되고 있지만 유색인종과 여성 등 기존에서 사회에서 불이익을 받아온 집단을 구조적으로 차별하는 문제를 일으켰다. 사람에 의한 차별이나 편견과 달리 기계의 의해 자동적으로 이뤄지고 설명되지 않는다는 점은 사회적 문제를 개선하기보다 악화시키고 고착화할 우려가 높다. 판단의 근거와 논리를 알지 못하면 개선할 방법이 요원하다. 인공지능의 알고리즘 의사결정에 대한 투명성과 접근성, 설명가능성을 요구하기 위한 기술적·법적 시도가 있지만, 이는 영역이 제한적이고 심화신경망 방식의 딥러닝 현재 기술 구조에서는 구조적으로 불가능한 요구일 따름이다.

셋째, 위임받지 않은 집단이 거대 권력을 행사한다.

민주주의는 법률을 통해 주권자들이 자신들을 다스릴 최고의 권력을 스스로 만들어내는 구조이고, 이는 임기제 대통령과 국회의원처럼 동의절차를 거친 일시적 위임의 형태로 구성돼 있다. 그런데 오늘날 사람들에게 가장 큰 영향력을 끼치는 거대 권력은 민주적 제도로 선출된 세력이 아니라 정보기술과 거대 플랫폼을 설계하고 소유한 빅테크와 그 설계자들이 행사하고

있다. 2023년 3월 세계적인 인공지능 전문가와 저명인사 등은 “앞으로 6개월간 인공지능 시스템 개발을 중단하자”는 내용의 미국 비영리단체 ‘삶의미래연구소(FLI)’ 명의의 공개 서한에 서명하며, 인공지능이 사회에 끼칠 잠재적 위험성을 경고했다. 1000여명 넘는 서명자들은 ‘지적으로 우월한 비인간 지성(nonhuman minds)’이 인간을 대체하도록 개발하느냐의 문제를 “위임받지 않은 기술 지도자들에게 위임해서는 안 된다”고 주장했다.

블랙박스과 같은 인공지능은 인간이 통제할 수 없는 기술에 과의존하는 현상에 대한 우려를 낳고 있다. 전기나 화학 등 필수적이고 보편적인 기술이라 해도 대개 기술은 인간의 통제 아래 있었다. 그런데 디지털 인공지능 기술은 다르다. 인간이 도구에 대한 창조자로서 우위를 지닌 상태에서 통제가능한 관계를 맺는 게 아니라, 그 관계가 역전될 가능성이 커지고 있다. 인간이 도구에 더 많은 일상을 의존하게 되었지만 사람은 도구에 의해 알지 못하고 조종당할 우려다. 그 인공지능은 소수의 기술 전문가와 빅테크에 의해 통제되고 있는 현실이다.

2023년 11월 중순 1주일 동안 오픈AI에서 공동창업자이자 최고경영자인 샘 올트먼에 대한 해임과 내부반발, 복귀 등 일련의 사태 진행에서 드러난 것은 거의 모든 사람들의 생활과 사회구조 전반에 지대한 영향을 끼칠 기술 개발의 방향과 영향력에 대한 사안이 소수의 내부자들에 의해서 결정되고 있다는 걸 보여줬다. 특히 700여명에 불과한 일개 기업 안에서도 자신들의 명운을 좌우하는 사안에 대해 접근성과 투명성, 논의 참여가 제대로 이뤄지지 않았음이 드러났다.

민주주의는 링컨의 게티스버그 연설문처럼 사회 구성원 모두를 위한 것이어야 하지만, 디지털과 인공지능이 거대한 권력이 되고 있는 상황에서 구조적 위기를 맞고 있다. 구성원 모두의 것이 아니라(of the people) 인공지능과 플랫폼은 이를 운영하는 빅테크 기업과 국가권력의 소유가 되고 있으며, 시민에 의해서 운영되는(by the people) 것이 아니라 보이지 않는 알고리즘과 설명할 수 없는 딥러닝 인공지능에 의해서 작동하고 있다. 사회 구성원 모두를 위한(for the people) 것이 아니라 거대 권력 집단과 그 소유집단의 이익을 위해서 개발, 운영되고 있다. 그 결과가 사회의 여론형성 구조 왜곡과 디지털 격차(digital divide)로 나타나고 있다.

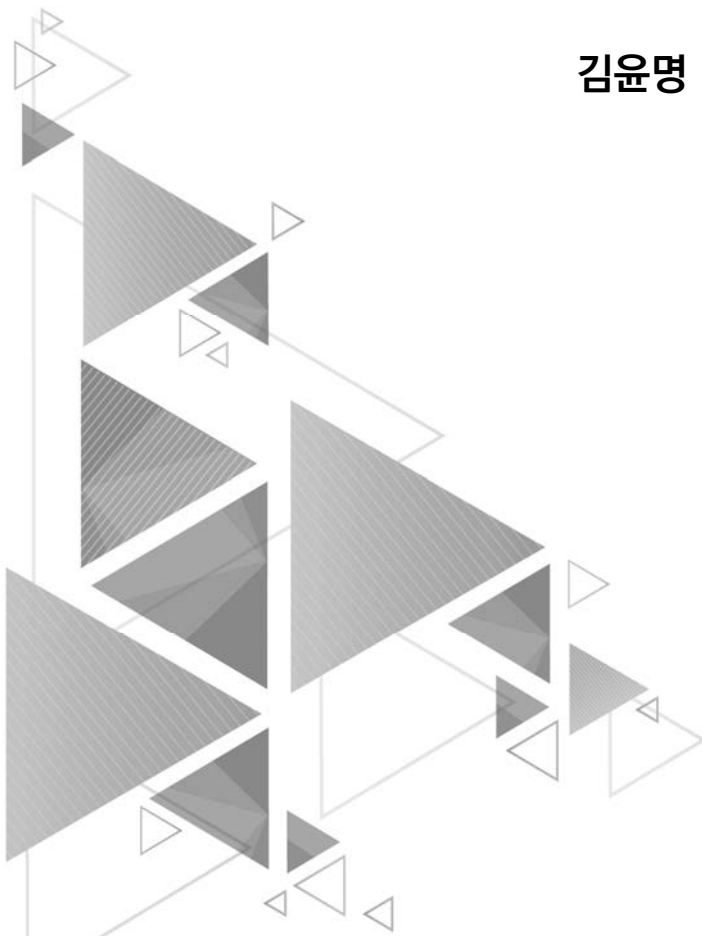
‘모두를 위한 인공지능’을 위해서는 위임받지 않은 거대 기술권력이 민주주의 사회구조를 왜곡하는 현실을 개선하기 위한 구체적 정책과 전략이 필요함을 의미한다. 이원재 발제자의 제안대로, 인공지능에 대한 시민적 리터러시와 거버넌스 논의가 필요한 시기이다. 시민과 사회, 산업 전체에 거대한 영향을 끼칠 인공지능 기술을 모두를 위한 도구와 환경으로 만들기 위해서는 무엇보다 기술에 대한 시민적 접근성과 투명성 확보를 통해 사회 전체적 논의로 가져오려는 시도가 필요하다.



토론 3

인공지능 신뢰성 확보를 위한 정책 제언

김윤명 한림국제대학원대학교 교수



인공지능의 신뢰성 확보를 위한 정책 제언

김윤명 한림국제대학원대학교 교수

며칠 동안 ChatGPT를 서비스하는 OpenAI CEO 샘알트만의 퇴출과 복귀 이슈가 AI업계의 화두가 되었다. 안전한 AI를 위한 노력과 수익성을 위한 가치충돌이 이루어진 것이라는 견해도 있다. 안전한 AI를 만들고, 신뢰할 수 있는 AI를 서비스하는 것, 그리고 AI를 악용하지 않는 것이 AI개발자, 서비스제공자 및 이를 이용하는 이용자의 책무이다. AI와 같이, 미지의 기술에 대한 입법자 내지 규제론자의 입장은 법적 안정성을 유지할 수 있도록 하자는 것이다. 그렇기 때문에 규제가 선행될 가능성이 크다. 인공지능으로 인하여 발생할 수 있는, 그렇지만 구체적인 그림자에 대한 우려는 크다.

정부에서도 안전한 AI환경과 더불어 AI패권을 위한 정책방안에 대해 고민하고 있다. 그렇지만, 국민의 안전을 위한 정책은 무엇보다 중요하다. EU AI법(안)과 미국 바이든 대통령의 행정 명령, 중국의 생성형AI규제법 등 많은 내용이 위험기반의 규제지향적인 규범이다. 그러한 규제와 기술개발 사이에서 기업은 또다른 AI겨울을 맞이할 수 있다. 정책이나 입법에서도 기술에 대한 규제가 아닌 그 기술이 활용된 서비스나 비즈니스모델에 대한 문제를 찾아내는 것이 필요하다.

사업자들이 문제라고 지적하는 것은 규제의 불투명성에 있다. 규제가 높은 것도 문제일 수 있지만, 명확하지 않는 규제정책이나 그림자규제에 있다. 이러한 규제는 규제권자가 책임을 지지 않겠다는 것과 다름이 없다. 적어도 정책이나 입법은 구체성이 있어야 하는 이유이다.

다만, 사업자들도 정확한 데이터를 제시하여야 한다. 적기조례(Red Flag Act)로 알려진 The Locomotives on Highways Act 1861는 영국의 자동차산업을 망가트렸을까? 적어도, 가솔린 자동차와 같은 내연기관을 규제하지는 않았다. 적기조례는 2톤이 넘는 28인승 증기기관 버스에만 적용되었기 때문이다. 이 적기조례가 영국의 자동차 산업의 발전을 저해했다는 연관성은 어디에서 나왔을까? 후세 삽화가의 역사에 대한 오해나 무지에서 출발했다면, 그리고 AI시대에 또다시 적기조례를 들고나오면 무지에 대한 조롱거리가 되지 않을까? 한 가지 더, 적기조례의 제정이유는 2톤의 증기기관 버스가 도로를 망가트리고, 증기기관의 불안전성으로 폭발하면서 인명사고자 잦았기 때문에 국민의 안전을 위해 규제가 이루어졌다는 점이다.

국민의 안전이 무엇보다 중요하다. 우리 헌법도 국민의 안전을 위한 책무를 규정하고 있다. 그에 따른 구체적인 안전권을 담보할 수 있는 입법이 요구된다. SW안전에 대한 규정이 있지만,

디지탈 사회의 안전이나, AI시대의 국민의 안전을 위한 규정은 어디에도 없다. AI정책이나 입법은 국민의 안전과 신뢰받을 수 있는 기업이 같이 할 수 있는 방향을 제시하여야 할 것이다.

1. 생성형 AI의 규제 이슈와 정책방안

알파고(AlphaGO)와 이세돌의 대국에 대중은 관찰자 입장이었다며, OpenAI의 ChatGPT에 대중은 사용자로서 직접 관여하고 있다. 그 만큼 인공지능이 생활의 일부가 되고 있음을 보여 준다. 그렇지만, 생성형 AI는 부정확한 정보나 알고리즘의 편향 등에 대한 문제도 여전하다. 특히, 환각효과를 가져오거나 인간의 규범과 충돌할 경우에는 규제가 작동될 가능성이 크다. 다만, 알고리즘을 포함한 다양한 인공지능 기술을 규제적 측면에서 접근할 경우, 신산업으로서 인공지능산업의 경쟁력을 약화될 수밖에 없다. 그렇기 때문에 기술에 대한 정책방안은 기본적으로 기술 발전을 저해해서는 안된다. 자칫 기술규제는 기술경쟁력을 저하시킬 수 있기 때문이다.

알고리즘의 오류, 편향이나 AI생성물의 내용상의 오류를 희석화할 수 있는 방안은 보다 많은 데이터를 제공함으로써 편향적인 결과가 나올 수 있는 확률을 줄이는 것이다. 이러한 문제를 극복하기 위해서는 다양한 데이터를 확보하여, 학습데이터로 활용할 수 있어야 한다. 물론, 이 과정에서 권리자에 대한 보상체계도 고민해야 한다. 공정이용에 따라 플랫폼 사업자가 사용한 데이터에 따라 얻은 수익을 독점하는 것은 합리적이라고 보기 어렵기 때문이다. 그동안 인공지능과의 경쟁에서 인간이 우위에 설 수 있는 영역이 창작성이나 예술적인 분야로 알려져왔으나, 생성형 AI가 가져오는 법률문제는 인간의 창작과 AI의 결과물과 차이를 구분하기 쉽지 않다.

〈새벽의 자리아〉와 같이, 미국 저작권청에 등록되었던 생성형 AI를 이용한 콘텐츠가 등록이 취소된 이유와 같이 등록자가 그 사실을 고지하지 않을 경우에는 확인하기가 쉽지 않다는 것을 보여준다. 이러한 한계로 미국 저작권청은 AI생성물에 대한 등록 가이드라인을 공개하였지만, 보다 구체적인 논의를 통해 AI 생성물의 저작권을 어떻게 처리할 것인지 결론을 내릴 필요가 있다. 생성형 AI모델은 인간이 만들어낸 것이라는 점에서 현행 규제체계에서 보면, 문제점도 상존하는 것은 사실이다. 따라서, 규제와 산업의 균형적인 접근을 통해 AI가 보다 인간의 의사 결정에 합리적으로 작동할 수 있는 투명하고 신뢰가능하도록 하는 것이 필요하다.

2. AI안전성 확보와 규제

생성형AI의 문제로 환각(hallucination)을 들고있으나, 어쩌면 AI가 인간을 모방한다는 점에서 당연한 것이기도 하다. 즉, 인간의 대화에도 허풍과 거짓이 담겨있고, 사고하는 방식이나 학습도 기계와 다르지 않다. 그렇기 때문에 AI가 오답을 내거나 다른 결정을 내리는 것도 인간의 연장이라는 측면에서 그 자체가 문제는 아니다. 다만, 그러한 환각이나 오류가 지속될 경우 인간사회에 미치는 영향은 결코 작지 않을 것이기 때문에 규제 필요성이 제기된다. AI는 미지의 기술이기도 하지만, 어떤 방향으로 진화할지는 알 수 없다. 다만, 미래는 현재의 연속이라는 점에서 AI를 둘러싼 다양한 이해관계에 따라 달라질 수 있다. 규제의 필요성을 제기하거나 탈규제를 주장하는 입장이 상이하다. EU의 AI법안은 상당히 구체적인 규제를 담고있으나 미국은 AI법안 자체보다는 공정거래법, 통신법 등에서 기업의 투명성을 강조해왔다는 점에서 규제가 없다고 보기 어렵다. 우리도 AI기본법(안)을 준비 중에 있으나, 선언적이고 ‘선허용 후규제’ 원칙에 따라 사실상 규제는 전무하거나 불투명하다고 볼 수 있다. 기업의 입장에서 규제의 문제는 높거나 낮은 것에 있는 것이 아니라, 불투명함에 있다. 향후, AI의 특성상 규제는 강화될 것이 명확하다. 그런 면에서 규제라는 것은 명확해야, 기업은 그에 따른 경영전략을 수립할 수 있기 때문이다. 정책당국도 여러 진흥정책을 산발적으로 발표하기 보다는 지속가능하고 관리 가능한 정책을 펴야할 것이다.

3. 알고리즘 공개와 영업비밀을 침해하는가?

ChatGPT와 같은 생성형(generative) AI가 AI의 새로운 흐름을 바꾸고 있다. 판단형 알고리즘은 채용이나 범죄예측 등의 활용에 사용되며 학습데이터가 갖는 한계로 인하여 편견 가능성이 크다. 또한, 추천형 알고리즘은 편견을 확증할 수 있다는 우려도 크다. 특정 정보를 추천하면서 다른 정보가 배제되는 필터버블(filter bubble) 현상이 발생할 수 있기 때문이다. 이러한 편견 등의 발생은 데이터 기반의 기계학습이 갖는 한계이기도 하다. 학습데이터로 가공되어지는 공개된 정보에는 세상의 모든 편견이나 오류가 그대로 담기기 때문이다. 인간의 문제가 투영된 데이터를 기반으로 하는 AI모델도 인간의 문제를 그대로 학습하게 된다. 그렇기 때문에 알고리즘에 의한 사회적 문제는 지속적으로 제기될 것이다. 예를 들면, 채용이나 신용도 등을 판단하는 알고리즘과 이를 학습하는 데이터에 편향이 담겨있는 경우, 또는 기본권의 침해가 이루어져 삶의 가치를 훼손하는 경우 등이다. 알고리즘으로 인하여 발생하는 다양한 문제를

해결하기 위한 방안 중 하나로서 알고리즘 공개를 들 수 있다. 이에 대해 플랫폼사업자는 알고리즘이 영업비밀에 해당하기 때문에 공개할 수 없다는 입장이다. 공공부문에서도 그 입장은 다르지 않다. 문제는 알고리즘 규제론자들은 알고리즘 공개에 대한 합리적인 기준을 제시하지 못한다는 점이다. 알고리즘 공개에 대한 설득력을 얻기 위해서는 정당성을 확보하기 위한 방안을 제시해야 한다. 즉, 알고리즘 공개에 대한 구체적인 이유나 방법을 규제기관이 플랫폼사업자에게 제공해주어야 한다. 규제기관은 명확한 정책목표를 달성하고, 플랫폼사업자들은 예측가능성을 가져야하기 때문이다. 이에 알고리즘 공개에 대한 구체적인 방안과 사업자들이 우려하는 영업비밀과의 비교교량을 통해 신뢰가능한 알고리즘을 위한 합리적인 방안을 제시하여야 한다.

4. 데이터 공정이용을 위한 정책 방안

데이터 수집과 이용은 인공지능 또는 기계학습의 처음과 끝이다. ChatGPT를 보면, 데이터는 기계를 인간의 능력에 비견될 수준으로 만들고 있다. 기계학습을 위한 데이터 생산이나 확보 과정에 대한 공정이용의 판단에서 상업적인 목적이 당연히 배척되는 것은 아니다. 영국, 독일 및 EU에서도 연구학술 등 비영리 목적을 위한 데이터 마이닝을 위해 저작권재산권 제한규정을 도입하고 있으며, 일본은 보다 적극적이다. 일본의 적극적인 입법은 우리나라나 미국과 같이 포괄적 공정이용 규정이 없는 이유이기도 하지만, 인공지능 산업을 선도하겠다는 의지를 보여주는 것이다. 2020년, 우리나라도 정보분석을 위한 제한규정 도입을 위한 저작권법 개정안이 발의되었다. 이로써 사업자에게 예측가능성을 높여줄 수 있을 것이다. 다만, 개정안의 입법화는 권리자의 반발이 예상되고, 시간이 걸릴 수 있다. 따라서, 현행 저작권법에 따른 공정이용을 통해 데이터 크롤링과 TDM 등 기계학습이 공정이용에 해당하는지 살펴보았다. 결론적으로, 인간의 이용행태와는 다르다는 점 등을 들어, 공정이용에 해당할 수 있다고 보았다. 다만, 공정이용에 따른 타인의 저작물을 이용하여 독점적인 과실을 사업자에게 전부 귀속하는 것이 타당한 것인지는 의문이다. TDM이나 기계학습으로 인하여 발생하는 기계의 저작물 이용을 통해 사업자가 얻는 수익에 대한 보상 체계를 고려하지 않는 것은 공정한 경쟁환경에 대한 심대한 결과를 가져올 가능성도 배제할 수 없기 때문이다. 참고로, 박정 의원은 2017년 TDM을 허용하는 저작권법 일부 개정안을 발의한 바 있다. 이때, 구체적인 논의가 이루어졌더라면 인공지능 학습데이터 확보가 보다 원활했을 것이고, AI발전에 기여했을 것이다. 물론, 권리자에 대한 보상으로써 데이터보상청구권을 부여하는 것도 필요하다.

〈정책 자료〉

1. 알고리즘 공개와 영업비밀 보호 간의 긴장관계

1. 알고리즘 공개 논의

ChatGPT와 같은 생성형(generative) AI가 AI의 새로운 흐름을 바꾸고 있다. 판단형 알고리즘은 채용이나 범죄예측 등의 활용에 사용되며 학습데이터가 갖는 한계로 인하여 편견 가능성이 크다. 또한, 추천형 알고리즘은 편견을 확증할 수 있다는 우려도 크다. 특정 정보를 추천하면서 다른 정보가 배제되는 필터버블(filter bubble) 현상이 발생할 수 있기 때문이다. 이러한 편견 등의 발생은 데이터 기반의 기계학습이 갖는 한계이기도 하다. 학습데이터로 가공되어지는 공개된 정보에는 세상의 모든 편견이나 오류가 그대로 담기기 때문이다. 인간의 문제가 투영된 데이터를 기반으로 하는 AI모델도 인간의 문제를 그대로 학습하게 된다. 그렇기 때문에 알고리즘에 의한 사회적 문제는 지속적으로 제기될 것이다. 예를 들면, 채용이나 신용도 등을 판단하는 알고리즘과 이를 학습하는 데이터에 편향이 담겨있는 경우, 또는 기본권의 침해가 이루어져 삶의 가치를 훼손하는 경우 등이다. 알고리즘으로 인하여 발생하는 다양한 문제를 해결하기 위한 방안 중 하나로서 알고리즘 공개를 들 수 있다. 이에 대해 플랫폼사업자는 알고리즘이 영업비밀에 해당하기 때문에 공개할 수 없다는 입장이다. 공공부문에서도 그 입장은 다르지 않다. 문제는 알고리즘 규제론자들은 알고리즘 공개에 대한 합리적인 기준을 제시하지 못한다는 점이다. 알고리즘 공개에 대한 설득력을 얻기 위해서는 정당성을 확보하기 위한 방안을 제시해야 한다. 즉, 알고리즘 공개에 대한 구체적인 이유나 방법을 규제기관이 플랫폼사업자에게 제공해주어야 한다. 규제기관은 명확한 정책목표를 달성하고, 플랫폼사업자들은 예측가능성을 가져야하기 때문이다.

2. 알고리즘 공개에 따른 논란

가. 알고리즘 공개에 대한 반대론

알고리즘 공개에 반대하는 논거는 알고리즘을 대중(公衆)에게 공개하는 것으로 이해하고

있다는 점이다. 물론, 공중에게 공개하지 않더라도 공개가 가져오는 의미에 대해 부정적인 입장인 것도 사실이다. 알고리즘이 공중에게 공개될 경우에 이용자들은 해당 알고리즘을 회피하여 부정행위를 할 수 있다는 점, 알고리즘을 경쟁사업자가 활용함으로써 경쟁력이 약화될 수 있다는 점도 지적된다. 한편으로는 이해할 수 있는 지적이지만, 이러한 우려는 과도한 면이 있는 것도 사실이다. 무엇보다, 모든 알고리즘을 공개의 대상으로 하는 것은 아니다. 예를 들면, 생명에 위협적이거나 인격에 대한 침입이나 기본권을 제한하는 등의 물리적, 신체적 위험성이 높은 알고리즘을 대상으로 한다. 그 대상으로는 EU AI법(안)이나 우리나라 국회에 제안된 ‘인공지능기본법(안)’¹⁾에 규정된 금지된 AI나 고위험 인공지능 유형이 제시될 수 있을 것이다. 물론, 이 것만이 대상이 되는 것은 아니다. 개별 법률에서 위험성이 있거나 안전을 위해 필요로 하는 서비스 등에 적용되는 알고리즘도 대상이 되어야 할 것이다.

기술중심론자들의 주장은 기술이 경제발전을 이끌어가기 때문에 함부로 기술 규제를 해서는 안된다는 것으로, 기업의 투자를 이끌어내기 위해서는 규제보다는 투자할 수 있는 환경을 마련하는 데 있다는 것이다. 그렇지만, 기술개발이나 투자는 규제와 상관없이 또는 규제를 넘어서기 위해 진행되는 것으로서 규제 때문에 기술개발이 어렵다는 주장은 설득력을 얻기 어렵다. 물론, 기업의 입장에서는 알고리즘이 공개될 경우에 이를 악용하는 경우가 문제될 가능성은 있다. 예를 들면, 알고리즘이 공개된다면 알고리즘을 개선하려는 동기가 사라질 수 있다는 것이다. 특허의 경우, 기술은 공개하되 발명자에게 독점사용권을 주면서 동기부여를 제시한다. 그렇기 때문에 무조건적인 공개는 알고리즘 고도화 기회 상실로 연결될 수 있다.²⁾ 소스코드로서 알고리즘인 경우라면 경쟁력이 상실될 가능성도 있으며, 유사코드 형태의 기준이나 방법 등을 공개할 경우에는 이를 회피하는 경우도 예상된다. 따라서, 공개의 방법이나 수준에 대해서는 정책적 설계가 요구된다. 전체를 공개할 것인지, 아니면 핵심적인 사항을 공개할 것인지, 그러한 공개를 공중에게 할 것인지 등에 관한 사항이다. 실제, 알고리즘을 공개한다는 것은 공중에게 공개하는 것으로 보기 어렵다. 사업자들이 우려하는 사항이고, 이렇게 공개하는 것 자체가 경쟁질서를 훼손할 가능성도 배제하지 못하기 때문이다. 따라서, 특정 서비스나 법률에서 목적으로 하는 수준을 달성할 수 있는 정도까지 공개하는 방안도 고려할 수 있을 것으로 생각된다. 즉, 알고리즘 공개를 누구나 접근할 수 있는 수준으로 공개하자는 것은 아니다. 전문적이고 객관적인 정부기관을 두어 예외적인 경우에만 접근할 수 있도록 한다면 영업비밀이 해제되거나 유출될 가능성은 없다. 따라서, 플랫폼사업자들의 우려는 해소될 수 있을 것이다.

1) 2023.2.14일, 과학기술정보방송통신소위 제1차 회의에서 기존 발의된 법안을 폐기하고 위원회 단일안으로 제정하기로 의결한 바 있다.

2) 모정훈, “알고리즘 공개에 관하여”, 『디지털사회』 제36호, 2021.10.20. ISSN 2586-3525.

알고리즘이 데이터와 결합됨에 따라 블랙박스화 되면서 발생하는 기술적 어려움에 관하여 본다. 고전적인 알고리즘과 달리 기계학습을 통해 고도화되는 알고리즘은 데이터가 달라질 경우에 해당 알고리즘도 달라질 수 있다. 알고리즘은 데이터에 따른 가변성이 큰 특성이 있다. 그렇기 때문에 알고리즘 공개에 대한 논의가 쉽지 않은 것은 데이터와 알고리즘을 공개한다는 개념과 기준이 명확하게 수립되지 않을 경우에는 그 효과를 담보하기가 쉽지 않다는 점이다. 콤파스(COMPAS) 사건에서 퍼블리카가 알고리즘의 오류를 찾아낼 수 있었던 것은 공공 부문에서 제공되는 것이기 때문에 해당 알고리즘이 정보공개 청구 대상이었다는 점이다. 즉, 정보공개 청구를 통해 관련 자료를 확보하고 분석할 수 있었기 때문에 가능했던 것이다. 물론, 알고리즘 공개가 정보공개법의 대상이 되도록 하거나 민간부문의 알고리즘도 공개대상이 되도록 ‘알고리즘 공개법’이 제정하는 것도 하나의 방안이라고 보나, 데이터와 알고리즘을 구분해서 분석할 수 있을지 의문을 제기한다.³⁾ 실상, 추천알고리즘, 검색알고리즘 및 기사배열 알고리즘 등 수많은 알고리즘이 데이터와 결합하여 사용되기 때문이다. 이러한 기술적 어려움외에 알고리즘이 공개되어도, 그에 따른 문제를 해결하기 어렵고 오히려 파생되는 이슈로 시장을 파괴한다는 주장도 있다.⁴⁾ 역설적으로, 네이버는 뉴스알고리즘을 공개한 바 있다.⁵⁾ 이러한 공개는 그동안 반대논거로 제시했던 것에 대한 자기부정이라는 점에서 알고리즘 공개 확산을 가져올 수 있는 단초가 될 수 있을 것이다.

나. 공개의 반론에 대한 반론

알고리즘 공개 주장은 알고리즘으로 인하여 발생할 수 있는 문제를 규제하기 위한 방안 중 하나이다. 무엇보다, 처리과정을 확인하기 쉽지 않다는 알고리즘 자체가 갖는 블랙박스 속성에 따른 것이다. 즉, 딥러닝에 의한 AI모델은 개발자도 어떻게 구현되는 것인지, 결과가 나오는지 알 수 없다. 블랙박스를 열기위한 방안으로써 알고리즘을 공개할 경우에 전문가를 통해 해당 알고리즘을 분석하여 문제점을 파악하고 수정이나 개선을 권고할 수 있기 때문이다. 물론, 그 과정에서 위법성이 확인될 경우에는 법적인 제재도 검토하게 될 것이다. 알고리즘이 미치는 영향력이 크기 때문에 긍정적인 측면은 권장을 하고, 부정적인 결과를 가져오는 경우에는 개선토록 하자는 것이다.

EU에서는 수용하기 어려운 인공지능은 개발이나 활용을 금지하고 있으며, 고위험 인공지능에 대해서는 다양한 규제장치를 두고 있다. 금지되는 AI는 사람에게 해를 끼치는 행위를 하기

3) 모정훈, “알고리즘 공개에 관하여”, 『디지털사회』 제36호, 2021.10.20. ISSN 2586-3525.

4) 박재형, “알고리즘 설명하라” 온플법 통과 앞두고 갑론을박 여전, 투데이이코노믹 2021.11.10.일자.

5) 금준경, 네이버 뉴스 알고리즘 공개 내용 살펴보니, 미디어오늘 2023.7.8.일자.

위해 AI가 동원되는 경우다. AI를 이용해 사람의 행동을 유해한 방식(잠재의식을 활용)으로 조작하거나, AI를 악용, 연령이나 신체적·정신적 장애 등 특정집단의 취약성을 악용하는 경우다. 또한 행정기관 등이 개인의 행동 특성을 기반으로 신뢰도를 · 평가 분류하기 위한 경우나 법 집행 목적으로 공공장소에서 실시간 원격 생체인식정보 시스템을 이용하는 경우 등도 금지된다. EU의회에서 추가된 금지유형으로는 민감하거나 보호되는 속성, 특성 또는 그에 따른 추론 등을 통해 자연인을 분류하는 행위다. 또 개인의 과거 범죄 특성을 평가하고, 프로파일링 등에 기반한 범죄 발생을 예측하거나, 범죄 재범 · 위험 평가를 AI를 이용해 하는 행위도 금지된다. CCTV영상을 이미지 스크래핑하여 얼굴인식 데이터베이스를 구축 확장하는 행위도 금지된다. 또한 법집행이나, 국경관리, 직장 및 교육기관에서 감정 추론 등을 위해 AI를 활용해서도 안 된다. 이 밖에 공공장소 녹화 영상을 분석할 목적으로 이용해서도 안 된다. EU AI법(안)에서 ‘고위험 AI’로 분류되는 경우에는 것은 다양하다. 우리도 인공지능기본법안을 입법 중에 있지만, ‘선허용 후규제’라는 원칙하에 실효적인 규제를 두고있지는 않다.

이러한 위험성에 따른 논의에 있어서, 알고리즘의 사용을 금하고 있지만 실상 공개에 대해서는 영업비밀성을 이유로 반대하고 있다. 그렇지만, 영업비밀이 기업의 중요한 자산이지만, 재산권은 제한될 수 있다는 점이 헌법의 기본원칙이다. 즉, 공공복리를 위하여 재산권이 제한될 수 있는 것이다. 만약, 침해되는 법익이 크다고 판단되면 알고리즘을 공개할 수 있을 것이고, 그러한 경우라면 영업비밀성을 유지할 수 있는 범위내에서 전문적인 기관을 통해 공개토록 하는 것이다. 해당 기관은 비밀유지의무를 부담해야하며, 그러한 조건하에 알고리즘을 임치(任置)할 수 있다. 만약, 그렇지 않고 오해하는 것처럼 공중에 공개할 경우에는 알고리즘을 오남용할 가능성이 크기 때문이다. 그렇지만, 현재로서는 알고리즘 공개를 주장하면서 안전하게 관리할 수 있는 방안을 포함하여 구체적인 방식을 제안하지 못하고 있다. 이러한 주장은 사업자들의 우려를 가져오게 되며, 공개제도가 추구하는 가치를 몰각시키고 있다. 따라서, 보다 구체적인 방안을 제시하는 것이 알고리즘 규제에 대한 정당성을 확보할 수 있는 방안이다.

다. 정리

정리하자면, 알고리즘의 공개에 대해서는 사업자는 영업비밀을 포함한 기업비밀이라는 점을 들어 공개를 반대하는 입장이다. 이러한 주장에 따라, 유관부처를 포함한 많은 전문가들도 동의하는 경향성을 갖는다.⁶⁾ 다만, 영업비밀은 헌법에 부여된 재산권의 하나에 불과하며, 재산권

6) ‘정보통신서비스 제공자의 알고리즘 자료 제출 의무화’를 규정한 정보통신망법 개정안에 대해 공정거래위원회는 “영업비밀에 해당하는 알고리즘을 매년 정부에 제출하도록 하는 것은 사업자에 대한 과도한 규제로서 오히려 산업의 혁신을 저해할 우려가 있다”는 의견을 제출하였다고 한다. 조기열, 정보통신망법 일부개정안 검토보고, 국회과학기술정보방송통신위원회, 2021, 6면.

은 공익적 목적 등 다양한 경우에 제한된다. 즉, 영업비밀이 국민의 기본권을 포함하여 생명이나 신체를 해하는 경우에는 중대한 기본권의 침해라는 점에서 다른 재산권과 마찬가지로 공개라는 형태로 제한되어야 한다는 것이다. 또한, 영업비밀은 사회적 상당성이 있는 경우에 한하여 보호받는다.⁷⁾ 따라서, 알고리즘을 활용하는 것이 헌법적 가치를 훼손하거나 법률에 위반되는 것이라면 국가의 개입이 필요하다. 개입의 유형은 규제나 영업비밀의 공개 등을 들 수 있다. 이처럼 위법한 알고리즘을 영업비밀로 보호한다면 국가는 평등권, 사회적 약자의 보호 등에 대한 헌법적 책무를 방기하는 것이기 때문이다.⁸⁾ 다만, 우려하는 지점이 경쟁업자에게 영업비밀 등이 공개됨으로써 기업경쟁력이 약화된다는 우려에 따라서, 제한적이고 엄격한 절차에 따른 공개요건을 갖출 경우라면 상황은 달라질 수 있다. 알고리즘의 영업비밀성을 훼손하지 않는 범위내에서 알고리즘을 공개하거나 이에 준하는 방법으로 문제를 해결할 수 있는 방안을 찾는 것도 필요하다.

3. 알고리즘 공개의 입법 방안

가. 알고리즘 공개의 입법시 고려사항

현행 법제하에서 알고리즘을 공개를 강제할 수 있는 방안을 찾기 어렵다. 최근 개정된 개인정보보호법이나 신용정보법에서는 정보주체의 알고리즘 설명요구권이나 적용거부권을 규정하고 있으나, 알고리즘 공개에 대한 구체적인 규정을 두고 있지 않다. 다만, 제한적으로 공공영역에서는 정보공개법을 통해 정보공개를 청구할 수 있을 뿐이다. 공공기관의 AI 채용시스템에 대한 공개청구가 인용된 바 있기 때문에 가능성은 있으나 여전히 정보공개법은 제한적이라는 점에서 한계가 있다. 더욱이, 민간 영역에 대해서는 알고리즘 공개를 요구할 수 있는 근거를 찾기 어렵다. 앞서 살펴본 바와 같이 기업은 자사의 영업비밀이라는 이유로 공개를 꺼리고 있는 실정이다. 물론, 알고리즘으로 인한 개인의 피해 정도가 크거나 광범위하게 기본권을 제한하는 경우에는 영업비밀이라고 하더라도, 일정한 조건하에서 알고리즘 공개를 강제할 수 있음을 살펴보았다.

물론, 알고리즘을 구성하는 소스코드나 데이터의 공개가 가져오는 위험성이 크다는 점에서 사업자의 우려를 불식시키기 위해서는 제도적인 뒷받침이 필요하다. 전문적인 국가기관에 알고리즘이나 데이터를 임치하도록 해야 한다. 알고리즘을 공개해야할 구체적인 사유와 기준을

7) 사회적 상당성이 없는 영업비밀에 대해서는 경제적 유용성이 인정되지 않는다는 것은 송영식 외 6인, 『송영식지적소유권법(하)』, 육법사, 2008, 458면.

8) 同挀로는 이제희, “알고리즘 취급에 대한 법적 논의”, 『공법학연구』 제19권 제3호, 2018, 326면.

사업자에게 제시하여야 하며, 이를 위반할 경우에는 가중처벌할 수 있는 근거를 두어야 한다. 무엇보다, 알고리즘 위법적인 공개와 사용에 대해 처벌할 수 있어야 한다. 허가받지 않는 경우에는 누구라도 임치되거나 기탁되는 알고리즘이나 데이터에 대해 임의로 접근할 수 없도록 해야 한다. 즉, 일정한 요건하에 제한적으로 접근이 허용될 뿐이다. 기본적으로 내부적인 관리체계는 영업비밀보호법상 요건에 따른 절차를 수립해야 한다. 특히, 알고리즘을 다루는 사람은 비밀유지의무와 함께 공무원 의제를 통하여 엄격하게 통제되어야 하며, 그에 따른 책임을 지도록 해야 한다. 이처럼, 알고리즘 공개에 대해 엄격하게 관리하고 통제할 수 있는 법률이 제정되어야 알고리즘을 공개할 수 있을 것이다

나. 알고리즘공개법의 제정

공공부문에서 채용된 알고리즘의 공개는 현행 정보공개법을 통해 가능하다. 그러나, 비공개 사유에 해당한다는 이유로 공개를 거부할 가능성이 크다. 따라서, 알고리즘 활용이 기본권을 해치는 등 필요할 경우에는 알고리즘 공개에 대해 구체적으로 규정할 필요가 있다. 반면, 민간 부문에 대해서는 차별에 관한 경우, 편견에 대한 부분, 채용에 관한 경우 등 다양하기 때문에 개별 법률의 개정을 추진하는 것은 어려운 일이다. 따라서, 가칭 ‘알고리즘공개법’과 같은 특별법 제정을 통해 구체적인 내용이나 수준, 전문기관 등에 대해 규정하는 것이 바람직하다. 알고리즘공개법에서는 영업비밀에 대한 특례조항을 포함하는 것이 필요하다. 영업비밀은 공익목적으로 제한될 수 있는 것이라는 점에서 공개하는 것이 낫다고 판단되는 것은 공개하도록 하는 것이다. 절차나 방법 등에 대해 명확하게 규정해야 한다는 점에서 특별법을 통해서 구체화하는 것이 필요하다. 따라서, 알고리즘공개법에서 알고리즘을 공개에 관한 구체적인 내용을 규정할 필요가 있다. 공개된 알고리즘은 기본권 침해할 시정하거나 시장경쟁, 공정성 확보 등의 공개 목적이 달성될 경우에는 즉시 폐기해야 한다. 또한, 공개된 영업비밀의 목적외 활용을 금지해야 한다. 만약, 소스코드나 데이터가 부정한 방법 등으로 유출되거나 경쟁질서를 훼손하는 경우에는 관련자에 대해 엄격하게 처벌할 수 있어야 한다. 이러한 우려에 대해서는 앞서 살펴본 EU AI법(안)과 같이, 전문기관이 해당 서버에 접속하여 데이터나 알고리즘을 확인할 수 있는 권한을 부여함으로써 공개에 대한 우려를 불식시키는 것도 고려할 수 있다.⁹⁾

9) 다만, 알고리즘공개법의 구체적인 구성이나 전문기관의 역할 등 세부적인 사항은 정책적인 사항으로 구체적인 연구를 통해 도출해내는 것이 바람직하다고 본다.

II. 생성형 AI의 규제 이슈와 정책방안

1. 변곡점에 서있는 생성형AI

AI의 대중화를 가져온 알파고(AlphaGO)와 이세돌의 대국에 대중은 관찰자 입장이었던 것과 달리, OpenAI가 제작한 ChatGPT에 대중은 사용자로서 직접 관여하고 있다. 마이크로소프트(MS)의 창업자인 빌게이츠(Bill Gates)는 “사용자 질문에 놀라울 정도로 인간과 유사한 응답을 제공하는 챗봇 ChatGPT가 인터넷의 발명만큼 중요하다”¹⁰⁾고 했으며, 이미지넷(imageNet)을 구축한 페이페이리(Fei-Fei Li) 교수는 “위대한 변곡점(Great Inflection Point)”¹¹⁾이라고 했다. 초기 인터넷은 메인 서버간의 통신망 연결에 불과하였으나, 90년대 웹 프로토콜과 웹 브라우저가 개발되면서 일반인들도 쉽게 인터넷을 이용할 수 있는 웹의 시대가 도래하면서 인터넷은 혁신서비스로서 역할을 해왔다. 이처럼, 인터넷 혁신 이상으로 ChatGPT를 평가하고 있으며, 검색서비스에도 영향을 미칠 것으로 예상된다. 인공지능 서비스는 다양한 데이터를 기반으로 하여, 추론할 수 있는 선택지를 인간에게 제시해왔다. ChatGPT도 추론을 위한 것이라는 점은 변함없다. 다만, 보다 현실적인 결과물(contents)을 제안하면서 인간의 의도를 파악하여 결과를 도출시킨다는 점에서 기존의 AI모델과는 차이가 있다. 물론, 달리(Dall-E) 2나 미드저니(midjourney)라는 AI모델은 입력창인 프롬프트(prompt)에서 자연어로 입력한 내용을 파악하여 그에 따른 결과물을 만들어낸다. ChatGPT가 텍스트 형태의 콘텐츠를 만들었다면, 앞의 두 모델은 이미지를 만들어냈다는 점에서 차이가 있다.

앞으로 작가와 출판사는 새로운 경쟁자로서 AI모델을 대면할 가능성이 크다. AI모델을 활용하여 다른 형식의 콘텐츠를 생성해낼 가능성도 크다. 다양한 분야에서 ChatGPT와 공동으로 작업했다는 표시의 글들이 출간되고 있다. 글쓰기라는 작가만의 벽을 ChatGPT 등 인공지능이 낮춰주고 있다. 다만, 해결해야 할 문제는 학습과정에 사용되는 데이터의 저작권 문제와 생성된 콘텐츠와 관련된 법적, 윤리적 문제이다. 특히, 데이터를 수집하는 과정과 기계학습 과정에서 발생할 수 있는 저작권법과 AI과의 충돌은 문명사적 이기(利器)에 대한 규제로서 작용할 것인지, 아니면 수용할 것인지에 대한 경계에 있다고 해도 과언이 아니다. 규제에 대한 가능성도 높아지고 있는데, 특히 EU에서도 ChatGPT에 대한 규제가능성을 내비치고 있으며,¹²⁾ 이탈리아 등에서는 접속을 차단하고 있다.¹³⁾

10) Rachel More, “Microsoft co-founder Bill Gates: ChatGPT ‘will change our world’”, REUTERS, February 10, 2023.

11) Fei-Fei Li, “AI’s Great Inflection Point”, Generative AI: Perspectives from Stanford HAI, Stanford HAI, 2023, p.4.

12) Foo Yun Chee and Supantha Mukherjee, “Exclusive: ChatGPT in spotlight as EU’s Breton bats for tougher AI rules”, REUTERS, February 4, 2023.

13) Elvira Pollina and Supantha Mukherjee, “Italy curbs ChatGPT, starts probe over privacy concerns”, REUTERS, April 1, 2023.

2. ChatGPT가 산업에 미치는 영향

가. 현황

ChatGPT는 출판, 이미지 생성, 계약서, 연설문 등 다양한 영역에서 활용되고 있다. 현재 공개된 ChatGPT는 텍스트에 특화된 AI모델이지만, 다양한 보고서, 계약서, 에세이, 시나 소설 등을 만들어내고 있다. 특히, 코딩은 상당 수준의 소스코드(source code)를 생성하고 있다. ChatGPT 모델을 구축한 OpenAI는 달리라는 이미지 생성 모델을 통해 다양한 유형의 이미지를 만들어낼 수 있는 비즈니스 모델을 운용 중이다. 미드저니(midjourney)¹⁴⁾라는 이미지 생성모델로 만든 그림이 미국 Colorado State Fair art competition에서 수상한 사건이 있었으며¹⁵⁾, 이를 계기로 AI모델이 만든 콘텐츠에 대해서는 표절이나 저작권 침해를 이유로 수상을 금지해야한다는 여론이 일기도 했다.¹⁶⁾

[그림] Theater D'Opera Spatial(스페이스 오페라 극장)



* source : Jason M. Allen with midjourney(2022)

14) Leap Motion 사에서 모델링한 미드저니는 OpenAI의 DAL-E, stable Diffusion과 유사하게 텍스트 설명으로 이미지를 생성하는 AI모델이다.

15) "AI-generated artwork wins prize in art competition in US! Find out how", The Economic Times, Sep 05, 2022.

16) Kevin Roose, An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy, New York Times, Sept. 2, 2022.

제작자는 출품할 때에 미드저니와 본인의 이름(Jason M. Allen via Midjourney)을 표기함으로써 자신의 작업이 AI를 사용하여 만들어졌다는 것을 밝혔으며, 누구도 속이지 않았다고 한다.¹⁷⁾ 실제, 위 그림이 미드저니를 이용하여 제작한 것이 밝혀진 이후, 경쟁자들의 문제제기에 대해 심사위원들은 위 이미지에 대해 수상을 철회할 만한 사항이 발견되지 않았고, 작성자도 AI를 활용하였다는 것을 표시했다는 점이 소명되어 수상을 철회하지 않았다.¹⁸⁾

무엇보다, 영화나 광고에서 가장 중요한 것은 스크립트이며, ChatGPT가 스크립트를 만들어 낸다는 점에서 영상제작의 핵심적인 역할을 하는 것으로 볼 수 있다. 물론, 이미 동영상을 만들어내는 다양한 AI모델이 개발되어있기 때문에 이를 활용할 수도 있다. 더욱이, OpenAI도 동영상을 생성하는 AI모델을 공개할 것이라고 하므로,¹⁹⁾ 영상 콘텐츠 생성 부문에서도 논란이 일 것으로 생각된다. 음악분야에서는 구글에서 공개하려던 계획을 철회한 뮤직 LM(Music LM)이라는 음악생성 AI모델처럼, 다양한 음원을 기계학습한 모델도 있다. 물론, 다양한 음악생성 모델이 있지만 상용화는 저작권 이슈가 있기 때문에 쉽지 않은 분야이기도 하다. 구글의 경우도 음악저작물은 다른 유형의 콘텐츠 보다는 저작권 침해 내지 표절에 보다 민감한 측면이 있기 때문에 공개를 유보한 것이다.

나. 향후 전개

ChatGPT를 활용한 다양한 비즈니스 모델이 개발될 것으로 보인다. 이미, 글쓰기나 코딩은 인간의 능력을 앞서는 것으로 판명되고 있다는 점에서 ChatGPT의 한계를 논하기도 쉽지 않다. 구글이나 네이버 등이 공개하는 모델도 ChatGPT의 능력에 견줄 수 있다는 점에서 인공지능의 본질이 아닌 하나의 모델에 관심이 집중되고 있다. ChatGPT 모델 자체가 텍스트 위주의 생성모델이지만, 영상을 위한 텍스트 제작을 쉽게 해준다는 점에서 영상과 음성 모델과 결합된 다양한 비즈니스 모델도 가능하다.

AI모델이 직접 콘텐츠를 생성하는 것도 가능하지만, 무엇보다 중요한 것은 콘텐츠 제작을 위한 기획안과 구체적인 스토리(story)라고 할 것이다. ChatGPT는 다면적인 스토리텔링이 가능하므로 다양한 스토리 구성을 할 수 있다는 점에서 향후 콘텐츠산업은 스토리 중심으로 재편될 것이다. 이 과정에서 그래픽이나 프로그래밍도 AI모델이 구성한다는 점에서 인간의 창

17) Kevin Roose, An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy, The New York Times, Sept. 2, 2022

18) Paul DelSignore, AI Art Wins Competition And Sparks Controversy, <https://medium.com/mllearning-ai/ai-art-wins-fine-arts-competition-and-sparks-controversy-882f9b4df98c> (2023.2.1. access)

19) Connie Loizos, That Microsoft deal isn't exclusive, video is coming and more from OpenAI CEO Sam Altman, TechCrunch, January 18, 2023.

작적 기회가 충돌할 수 있다는 우려도 있다. 콘텐츠 시장에서 AI모델은 저작권 이슈에서 자유로울 수 없을 것으로 보이며, 현재 음악은 저작권 침해소송에 가장 많은 분야 중 하나이다. 무엇보다, 게임콘텐츠나 드라마, 음악 등은 인간에 의한 창작적 우위가 있는 분야라는 점에서 당분간은 우리나라가 경쟁력을 유지할 수 있는 분야라고 할 것이다. 이러한 환경에서 AI모델을 활용함으로써 창작 기회를 높이는 분야도 생겨날 것으로 보인다.

3. 생성형AI 관련 이슈

가. LLM에 대한 문제제기

ChatGPT가 이용되는 도메인(domain) 영역에 따라 다양한 법적, 윤리적 이슈가 제기될 수 있을 것이다.²⁰⁾ LLM은 법적 측면에서 몇 가지 문제점이 있다. 첫째, LLM은 학습 데이터에 대한 개인정보 보호문제가 있다. LLM이 학습 데이터로부터 개인정보를 추출하여 이를 저장하고 활용할 가능성이 있기 때문이다. 이에 대한 대응방안으로는 학습 데이터에 포함된 개인정보를 삭제하는 등의 방법이 있다. 둘째, LLM이 생성한 문장이 혐오, 차별, 저작권 침해 등 법적 문제를 일으킬 수 있다. 이는 LLM이 자연어를 생성할 때, 학습 데이터 내에 포함된 문장들을 참고하기 때문이다. 이를 방지하기 위해서는 모델 학습 시 학습 데이터의 선정과 전처리를 신중히 해야하며, LLM을 이용하여 생성된 문장들을 검증하는 시스템을 구축하는 등의 방안이 필요하다. 셋째, LLM이 생성한 문장이 인간의 작성물과 구분하기 어려울 수 있다. 이는 저작권 문제를 일으킬 가능성을 내포하며, 누군가가 LLM을 이용하여 다른 사람의 작성물과 유사한 문장을 생성하면 이는 저작권 침해가 될 수 있다. 이를 방지하기 위해서는 LLM이 생성한 문장이 어디까지가 학습 데이터에서 유래한 것인지를 알아내는 기술적인 방법이 필요하다. 넷째, LLM을 이용한 자동화된 텍스트 생성이 악용될 수 있다. 이는 허위 정보 생성, 광고 정보 생성, 사기 등 다양한 범죄 행위에 악용될 수 있다. LLM을 이용한 범죄 행위를 방지하기 위해서는 LLM 사용의 규제나 적극적인 모니터링, 법적 제재 등의 대응이 필요하다. 이를 위해서는 정부나 관련 단체에서 LLM 사용의 규제나 감시 시스템을 구축하고, 법적 제재를 강화하는 등의 정책적 대응이 필요하다. LLM으로 인한 문제와 한계를 극복하기 위해서는 기술적인 발전 뿐만 아니라 법적인 측면에서도 적극적인 대응이 필요함을 알 수 있다.

20) Zaremba, Adam and Demir, Ender, ChatGPT: Unlocking the Future of NLP in Finance (January 13, 2023). Available at SSRN: <https://ssrn.com/abstract=4323643> or <http://dx.doi.org/10.2139/ssrn.4323643>

나. 환각효과에서 자유로울 수 있을까?

AI모델은 공개된 데이터를 기반으로 학습하기 때문에 인간의 다양한 오류나, 문제점을 판단하지 않고 답습함으로써 편향적이고 차별적인 결과를 가져올 수 있다. 따라서, 독성이 있거나 편향된 콘텐츠를 포함하지 않도록 이러한 모델을 교육하는 데 사용되는 초기 데이터를 신중하게 선택하는 것이 중요하다.²¹⁾ 이외에도 부적절한 내용이 담겨있는 경우이다. ChatGPT는 내용 자체를 필터링하고 있지만, 모든 내용을 필터링할 수 있는 것은 아닌 것으로 보인다. 따라서, 부적절한 내용이 담길 가능성이 높으며 이는 교육현장에서 사용할 경우에는 교육 방법이나 내용에서 문제될 가능성도 있다.

또한, OpenAI에서도 밝힌 바와 같이 결과물에 대해서 정확하지 않는 정보가 생성될 수 있다고 고지하고 있다. ChatGPT의 결과물에서도 사실과 다른 내용을 사실처럼 표현하고 있다는 점에서 환각효과가 문제될 수 있음을 알 수 있다. 환각 효과(hallucination effect)란 사람이 실제로 존재하지 않는 것을 보고, 듣고, 느끼고, 냄새 맡고, 맛보는 것과 같이 현실에 대한 거짓 또는 왜곡된 인식을 경험할 때 발생하는 현상이다. 환각은 정신 장애(예: 정신 분열증, 우울증 또는 불안), 물질 사용(예: 알코올, 약물 또는 투약) 또는 의학적 상태(예: 파킨슨병, 간질 또는 편두통)와 같은 다양한 요인으로 인해 발생할 수 있다. 챗봇의 기능은 인터넷에 게시된 방대한 양의 텍스트를 분석하여 만들어지기 때문에 사실과 허구를 구분할 수 없으며 여성과 유색 인종에 편향된 텍스트를 생성할 수 있다.²²⁾

ChatGPT도 인터넷의 다양한 정보를 바탕으로 학습했고, 스스로 정보에 대한 진정성을 확인할 수 있는 능력이 없기 때문에 언어적인 체계라는 면에서 답변은 잘 할 수 있더라도 문맥과 의미에 대해서는 참과 거짓을 혼동할 수 있다. 이러한 면에서 볼 때, ChatGPT도 환각효과(hallucination effect)에서 자유롭기는 어려울 것으로 보인다.²³⁾

21) Michael Chui et. al, "Generative AI is here: How tools like ChatGPT could change your business", Mckisey, 2022.12.20.

22) Cade Metz, "OpenAI to Offer New Version of ChatGPT for a \$20 Monthly Fee", The Newyork Times, Feb. 1, 2023

23) Michael Chui et. al, "Generative AI is here: How tools like ChatGPT could change your business", Mckisey, 2022.12.20.

4. 생성형 AI모델에 대한 규제는 필요한가?

가. EU AI법안의 생성형AI 규제

(1) 추가 규제 내용

수정된 EU AI법안에서는 AI 시스템을 “다양한 수준의 자율성으로 작동하도록 설계되었으며, 명시적 또는 암묵적 목표를 위해 물리적 또는 가상 환경에 영향을 미치는 예측, 추천 또는 결정과 같은 결과를 생성하는 기계 기반(machine-based) 시스템”으로 정의하고 있다. 또한, 파운데이션 모델을 대규모 데이터로 사전 훈련된 AI 모델로서 출력이 범용적이어서 다른 다양한 시스템 개발의 기초로 쓰일 수 있는 모델로 정의하고 있다. 다만, 파운데이션 모델(foundation model)을 고위험으로 분류하지는 않았으나, 모델 위험 관리 및 기술문서 제공 등의 엄격한 요구사항 규정하고 있다. 특히, 생성 AI에 사용되는 파운데이션 모델의 경우, 위법한 콘텐츠 생성 방지, 학습데이터의 저작물 정보 공개의무 등을 부과하고 있다.

(2) 생성 AI에 사용되는 파운데이션 모델의 추가 의무²⁴⁾

GPT 모델과 디퓨전(Diffusion) 모델과 같이 문자, 이미지, 영상 등을 생성하는 AI 시스템에 사용되는 파운데이션모델의 공급자는 위 일반적인 파운데이션모델 제공자의 의무와 함께 추가 의무를 부담한다. 즉, 생성 AI 파운데이션 모델의 공급자는 이용자에게 AI 시스템 활용 사실을 알려야 하고, 위법한 콘텐츠 생성을 방지할 수 있도록 모델을 설계·개발·학습해야 한다.

또한, EU 의회 AI법안은 생성 AI 서비스의 저작권 침해를 방지하기 위한 수단으로서, 생성 AI 파운데이션 모델이 저작권에 의해 보호되는 학습데이터를 이용할 경우, 이에 대한 정보를 문서화하여 공개하도록 하고 있다.

나. ChatGPT에 대한 사회적 반향

ChatGPT에 대한 사회적인 반향은 긍정적이거나 부정적이거나 다양하다. 먼저, 긍정적인 면을 본다. ChatGPT가 생성하는 결과물은 인간의 창작이나 사고의 능력을 넘어서기 때문에 문명사적으로 획기적인 전환점이 될 수 있을 것으로 보인다. 다양한 창작활동의 도구로 활용할 수 있다는 점에서 AI의 민주화나 창작의 민주화를 가져올 수 있을 것이다. 또한, 인간의 상상을 기술을 활용하여 구현할 수 있도록 한다는 점에서 기술적 사상(idea)을 높일 수 있을 것이다.

반면, GPT모델의 독점은 특정 사업자에게 인터넷의 공개된 정보를 통해 확보된 AI모델이라

24) 법무법인 태평양, EU AI 법안의 EU의회 수정안 주요내용, 2023.5.26.

는 점에서 공공재적 성격의 학습데이터를 특정 기업이 독점하는 것이 공정한 것인지는 의문이다. 독점적으로 사용될 경우에는 다양한 법적 이슈에 휩싸일 수 있다고 보며, 현재 미국에서 진행 중인 집단소송도 영향을 미칠 수 있을 것으로 생각된다. ChatGPT 활용에 대한 전문가 집단의 가이드라인을 제시함으로써, 인간이 ChatGPT에 의지하는 경향성을 차단할 필요가 있다. 인간의 의사결정 내지 다양한 일을 AI가 대신하는 경우, 기술에 의지하는 경향성이 높아진다는 점에서 인간의 지적 활동은 AI에 종속될 가능성도 있기 때문이다.

다. 인간의 규제에서 자유로운가?

기술에 대한 규제는 지양될 필요가 있다. 연혁적으로 기술이 가져오는 장점보다는 문제점을 크게 키우는 경향이 있기 때문이다. AI에 대한 규제도 마찬가지이다. 다양한 이슈가 발생하는 것은 사회현상의 일반적인 경향이기도 하다. 수많은 법과 제도가 운용중에 있음에도 범죄를 포함한 사회적 문제가 발생하는 것은 이를 증명한다. 따라서, 법제도의 미비한 점이 있을지라도 해석과 적용을 통하여 최대한 규제할 수 있다. 그렇기 때문에 별도의 새로운 규제보다는 기존 규제나 법제를 통해 해결할 수 있는 방안을 찾는 것이 기술 발전과 그에 따른 사회적 효용 측면에서 효과적이다.

ChatGPT가 출시되면서 다양한 이슈가 생겨나고 있음은 앞서 살펴본 바와 같다. 따라서, 이에 대한 규제의 필요성도 제기되고 있다. EU는 집행위원회 차원에서 규제필요성을 제기하고 있다. EU는 2021년 제안된 AI법안에서 위험의 정도에 따라 4가지로 AI모델을 유형화하고 있다. 무엇보다, EU AI법안은 생체 인식 및 대규모 모니터링을 활용하는 시스템뿐만 아니라 건강, 안전 및 기타 중요한 측면에 중대한 영향을 미칠 수 있는 시스템과 같이 개인에게 특정 위험을 초래하는 AI 시스템을 규제하는 것을 목표로 한다. 이에 따라, EU AI법안에 따라 AI 시스템은 네 가지 범주로 나뉜다.²⁵⁾ 먼저, 수용할 수 없는 인공지능은 사용이 금지된다. 둘째, 고위험 AI 시스템(High-risk AI systems)으로 생체 인식을 사용하는 시스템, 대규모 모니터링 및 중요한 의사 결정 시스템과 같이 개인에게 높은 위험을 초래하는 AI 시스템을 말한다. 셋째, 제한된 위험 AI 시스템(Limited-risk AI systems)으로 고객 서비스 또는 저위험 의사 결정에 사용되는 시스템과 같이 개인에게 제한된 위험을 초래하는 AI 시스템을 말한다. 넷째, 최소 위험 AI 시스템(Minimal-risk AI systems)으로 기본 패턴 인식 또는 간단한 데이터 처리에 사용되는 시스템과 같이 개인에게 위험을 최소화하거나 전혀 주지 않는 AI 시스템이다. 이처럼 EU AI법안은 수용할 수 없는 AI는 사용할 수 없도록 하고 있으며, 고위험 AI 시스템을 배포하기

25) EU AI법안에 대해서는 김윤명, 「블랙박스를 열기위한 인공지능법」, 박영사, 2022, 128~131면 참조.

전에 위험 평가를 거쳐야 하며 조직은 이러한 시스템이 위험을 최소화하고 개인의 권리와 자유를 보호하는 방식으로 설계, 개발 및 운영되도록 보장해야 한다. 제한된 위험 및 최소 위험 AI 시스템에는 덜 엄격한 요구사항이 적용된다.

OpenAI 내에서도 규제 필요성을 언급하고 있다.²⁶⁾ OpenAI의 CTO는 타임지(TIME)와의 인터뷰에서, 정부 개입이 혁신을 늦출 수 있다는 두려움 때문에 정책 입안자와 규제 당국이 개입하기에는 너무 이르지 않았느냐는 질문에 “너무 이르지 않다. 이러한 기술이 가져올 영향을 고려할 때 모든 사람이 참여하기 시작하는 것이 매우 중요하다.”고 답했다. 또한, “OpenAI와 같은 회사는 이를 통제되고 책임 있는 방식으로 대중의 의식에 가져오는 것이 중요하다. 그러나 우리는 소수의 사람들이고 이 시스템에 훨씬 더 많은 정보가 필요하고 기술을 넘어서는 훨씬 더 많은 정보가 필요하다. 확실히 규제 기관과 정부 및 기타 모든 사람이 필요하다.”고 답했다. 구체적인 규제라기 보다는 규제의 필요성이 있는 정도라는 점을 언급하고 있음을 알 수 있다. OpenAI공동 창업자였던 일론머스크도 규제의 필요성을 강조한다.²⁷⁾ 최근에 6개월간의 개발을 멈추자는 주장을 하였으나, 중국 등이 반사적 이익을 누릴 것이라는 이유 등으로 무산된 바 있다.

저작권법, 특허법 등 지식재산법제나 민법이나 형법에서는 행위 주체를 사람으로 한정하고 있다는 점에서 인간이 아닌 동물, 자연현상, 인공지능이 만들어낸 것에 대해서는 권리를 인정하지 않고 있다. 참고로, 특허의 인정여부가 다투어진 사례로는 AI 발명가인 다부스(DABUS)가 발명한 것에 대해 특허출원을 하였으나 각국은 AI의 발명자성을 인정하지 않고 있다.²⁸⁾ 이는 특허법에서 발명자를 인간으로 한정하고 있기 때문에 인간이 아닌 다부스에 대해서 발명자로서 지위를 인정할 수 없는 것으로 출원자체를 거절하고 있기 때문이다. 인공지능의 발명 주체성, 저작 주체성에 대해 논의하였으나 인공지능이 기획부터 창작을 한 것이 아닌 사람의 관여에 의하여 이루어지고 있기 때문에 주체성을 인정하지 못한 것이다. 행위 주체성이 인정되지 않기 때문에 권리를 침해하더라도 그 책임이 인정되기 어렵다. 향후, 헌법, 민법 및 형법 등 기본법의 개정을 통해 인간이외에 대한 법인격을 인정할 것인지에 대한 방안을 찾아야 할 것으로 판단된다. 참고로 EU에서는 로봇에 대한 법인격을 부여할 것인지에 대해 2017년부터 논의를 하고있으나 구체적인 결과가 도출되지 못하고 있다.

26) STEVE MOLLMAN, “ChatGPT must be regulated and A.I. ‘can be used by bad actors,’ warns OpenAI’s CTO”, Fortune, February 6, 2023.

27) Ryan Browne, “Elon Musk, who co-founded firm behind ChatGPT, warns A.I. is ‘one of the biggest risks’ to civilization”, CNBC, FEB 15 2023.

28) 이에 대한 논의는 권보원, “인공지능을 발명자로 볼 수 있을까? -DABUS 판결에 부쳐”, 「사법」 통권 59호 사법발전재단 (2022) 참조.

5. ChatGPT 등 신기술에 대한 규제정책

가. 신기술에 대한 정책 방향

기술에 대한 정책방안은 기본적으로 기술 발전을 저해해서는 안된다는 것이다. 자칫 기술규제는 기술경쟁력을 저하시킬 수 있기 때문이다. 또한, 규제당국은 규제보다는 현장의 의견을 반영하여 제대로 이용할 수 있는 환경을 조성할 수 있도록 최소한의 가이드라인을 제시해야한다. 기계학습에 사용된 데이터셋의 저작권 이슈에 대해서는 해소된 것이 아니라는 점에서 향후, 저작권 침해소송이 제기될 가능성도 다분하다. 데이터의 크롤링이나 TDM, 기계학습 과정은 공공성을 인정받을 가능성도 있다는 점에서 공정이용(fair use) 항변을 주장할 수 있을 것이며, 실제 인정받을 가능성도 배제할 수 없다.²⁹⁾ 다만, 저작권자의 권리를 침해하면서도 수익까지 독점하는 것은 형평성의 원칙에 위배될 수 있다는 점에서 확대된 집중관리(Extended Collective License)나 보상금청구권을 고려해야할 것으로 생각된다.³⁰⁾ 표절은 저작권 여부와 상관없이, 윤리적인 의무이기 때문에 출처표시를 해야할 것이며, 이는 학회 등의 패컬티(faculty)에서 자율규제 형식으로 대응해야할 것이다. 데이터에 담긴 편향 등 기본권 침해 소지가 있는 내용에 대해서는 필터링이나 모니터링을 통해 수시로 점검해야 한다. 최종적인 이용자 단에서 서비스를 이용할 수 있도록 하는 경우에는 더욱 조심할 필요가 있다.

나. 생태계 편중의 심화

벤처 생태계가 OpenAI라는 글로벌 기업에게 편중되는 현상이 발생하고 있다. 글로벌 서비스를 위한 목적이 아니더라도, 한글을 사용하는 데 있어서 한계가 없기 때문이다. 국내에서 LLM을 구축한 네이버, LG, 카카오 등이 있지만, 글로벌 기업들과 경쟁할 수 있는 경쟁력으로 인식하고 있는 한글데이터는 누구나 이용할 수 있는 환경이라는 점에서 국내 기업은 물론 글로벌 기업에서도 사용할 수 있다는 점에서 한글에 대한 경쟁력을 크지 않을 것으로 보인다.

국내 벤처생태계를 확산하기 위해서는 국내 LLM에 연동되는 서비스를 이용할 수 있는 환경이 마련되어야 한다. 정부차원에서 LLM 구축을 위한 정책을 발표하였지만, 한글 데이터만의 문제가 아니라는 점을 확인할 필요가 있다. 벤처생태계의 속성이 글로벌 환경에서 유리하다는

29) OpenAI나 구글 등이 LLM 모델을 공개하는 것은 다양한 데이터가 활용되었음에도 불구하고, 미국 저작권법 제107조에 따른 공정이용에 따른 면책이 가능하다는 판단에 따른 것으로 생각된다. 실제, 우리 저작권법 제35조의5에 따른 공정이용을 적용하더라도 동일한 결과가 나올 것이라고 생각한다. 이러한 견해에 대해서는 김윤명, “데이터 공정이용”, 「계간 저작권」, 제141호, 한국저작권위원회(2023), 5~53면 참조.

30) 확대된 집중관리에 대해서는 오승중, “확대된 저작권 집중관리제도에 관한 연구”, 「홍익법학」 제8권 제1호, 홍익대학교 법학연구소(2007); 박기태, “저작물 이용환경 변화에 따른 ECL 제도 도입에 관한 연구”, 「계간 저작권」 vol.34, no.2, (통권 134호) 한국저작권위원회(2021) 참조.

점이나 시장의 확장성을 위해서도 글로벌 플랫폼에 연동하는 것을 선호할 수 있기 때문이다. 따라서, 정부는 국내 LLM을 이용할 수 있는 바우처제도 등의 지원체계를 확대하는 것도 고려할 필요가 있다.

다. 생성형 AI에 대한 규제정책

저작권법적으로 볼 때, 생성형 AI는 새로운 창작도구로서 역할을 할 수 있을 것으로 기대되면서도 기계학습 과정에서 타인의 저작물을 이용하여 학습된 경우에는 저작권 문제로 확산될 수 있다는 지적도 제기될 수 있다. 이처럼, 새로운 기술에 대해서 법은 양면적인 해석이 가능하다. 결국, 사회적 합의를 통해 기술에 대한 사회적 수용성을 높이게 된다. 이러한 과정에서 저작권법의 개정은 이루어져왔고, 새로운 비즈니스 모델이 확산되면서 다양한 영역에서 혁신을 이끌어냈다. 그렇기 때문에 기술에 대한 규제는 기술자체의 차단이 아닌 예견되는 문제를 다룰 수 있는 수준으로 관리되어야 한다.

또한, 환각효과는 국내 LLM 보다는 해외 LLM에서 나타날 가능성이 높다. 우리나라에 대한 잘못된 정보가 나오는 것은 우리나라와 관련된 한글 데이터가 부족하기 때문으로 보이며, 이를 해소하기 위해서는 해외 사업자가 한글데이터셋을 활용할 수 있도록 정책적 체계를 마련할 필요가 있다. 참고로, 2023.5월에 일본 자민당에서 정부에 제안한 AI백서인 'AI국가 전략'에는 학습데이터의 편향성 문제에 대응하기 위해 국내외 AI기반 모델에 일본 관련 데이터를 적극적으로 제공하여, 일본 관련 학습데이터 비율을 높이겠다는 내용이 포함되어있다.³¹⁾ 국내 기업의 경쟁력 보다는 국가 경쟁력을 제고하는 차원에서 접근할 필요가 있음을 시사한다.

라. 정책 제언

AI는 미지의 기술이기도 하지만, 어떤 방향으로 진화할지는 알 수 없다. 다만, 미래는 현재의 연속이라는 점에서 AI를 둘러싼 다양한 이해관계에 따라 달라질 수 있다. 규제의 필요성을 제기하거나 탈 규제를 주장하는 입장이 상이하다. EU의 AI법안은 상당히 구체적인 규제를 담고 있으나 미국은 AI법안 자체보다는 공정거래법, 통신법 등에서 기업의 투명성을 강조해왔다는 점에서 규제가 없다고 보기 어렵다. 우리도 AI기본법안을 준비 중에 있으나, 선언적이고 '선허용 후규제' 원칙에 따라 사실상 규제는 전무하거나 불투명하다고 볼 수 있다. 그렇지만, 국내

31) 데이터 바이어스 문제에 대응하기 위해 국내외 기반 모델에 대해 적절한 일본 관련 데이터를 적극적으로 제공하는 등 일본과 관련된 학습 데이터의 비율을 높여 나갈 것. 또 정부 주도로 일본어 코퍼스(대역용으로 일본어 문장이 구조화된 데이터베이스)의 작성·활용을 추진하는 것. AIホワイトペーパー ～AI新時代における日本の国家戦略～, 日本自民党, 2023.4, 7頁.

기업은 EU의 GDPR처럼 EU의 AI법안을 따르지 않을 수 없다. 이처럼, 국내 기업은 2중적인 경영전략을 수립할 수밖에 없다는 점에서 부담이다. 따라서, 국내의 규제가 명확하게 제시되어야 한다. 규제의 문제는 높거나 낮은 것에 있는 것이 아니라, 불투명함에 있다. 기업의 입장에서 규제 또는 정책이 명확해야 그에 따른 경영전략을 수립할 수 있기 때문이다. 관련 부처에서는 이러한 점을 고려하여, 정책 및 입법을 추진하는 것이 필요하다.

Ⅲ. 생성형AI의 저작권 침해와 공정이용

1. 학습데이터셋(dataset)의 저작권 이슈

AI모델을 위해 만들어지는 데이터셋은 인터넷에 공개된 정보를 크롤링하여 제작되고 있으며, ChatGPT도 인터넷 정보를 이용하여 제작된 것임을 밝히고 있다. 인터넷에 공개된 정보는 이용허락 조건에 맞지 않게 이용할 경우라면 저작권 침해로부터 자유로울 수 없으며, 개인정보가 포함된 경우에는 개인정보보호법 위반에 해당할 수 있다. 물론, 저작권이 없거나 허락된 경우라면 침해로부터 자유로울 수 있기 때문에 가급적 저작권 이슈에서 자유로울 수 있는 데이터를 사용해야하나, 사실상 이를 확인하는 것은 어려운 작업이다. 일례로 GPT-2에서는 학습된 데이터를 그대로 복제해서 출력함으로써 개인정보에 대한 가명 처리 등이 되지 않은 것을 확인하였으며,³²⁾ 저작권 등도 유사하게 이용할 것으로 추정된다. 구글이 뮤직LM을 공개하려다 취소하였으며, 이는 해당 AI모델의 표절율이 1%에 이르는 저작권 침해 우려에 따라 취소된 것이라고 한다.³³⁾

표절 판단에 있어서, 음악저작물은 일반적으로 가락(melody), 리듬(rhythm), 화성(harmony)의 3가지 요소로 구성되고, 이 3가지 요소들이 일정한 질서에 따라 선택·배열됨으로써 음악적 구조를 이룬다. 음악저작물의 창작성 여부 판단은 음악저작물의 표현에서 가장 구체적이고 독창적인 형태로 표현되는 가락을 중심으로 하여 리듬, 화성 등의 요소를 종합적으로 고려하여 판단하여야 한다.³⁴⁾ 데이터셋의 저작권 이슈 등에 대해서는 공정이용(fair use)에 따라 침해에 대한 면책을 주장할 수 있으며, 이러한 주장은 구글의 도서 디지털화와 관련한 Google v. The Author's Guild 사건³⁵⁾에서 비소비적 이용(non-consumptive)이라는 점을 들어 공정이용으로 인정된 바 있다. 국내에서도 다음(현, 카카오)이 썸네일 검색을 제공함으로써 저작권을 침해하기는 하였지만, 시장대체성이 없었다는 점, 더 많은 정보를 제공한다는 공익적 목적으로 활용된다는 점에서 공정한 이용(저작권 제28조(인용))³⁶⁾임을 확인받아 면책된 바 있다.³⁷⁾

32) Nicholas Carlini et. al, "Extracting Training Data from Large Language Models", <https://arxiv.org/abs/2012.07805> (2023.2.1. access)

33) Kyle Wiggers, "Google created an AI that can generate music from text descriptions, but won't release it", TechCrunch, 2023.1.28.

34) 대법원 2015. 8. 13. 선고 2013다14828 판결.

35) Authors Guild v. Google, Inc., 804 F.3d 202 (2nd Cir. 2015)

36) 저작권법 제28조(공표된 저작물의 인용) 공표된 저작물은 보도·비평·교육·연구 등을 위하여는 정당한 범위 안에서 공정한 관행에 합치되게 이를 인용할 수 있다.

37) 대법원 2006. 2. 9. 선고 2005도7793 판결

2. 생성형 AI모델의 저작권 침해

가. 침해 주체

(1) 이용자인가, ChatGPT인가?

민법은 사람의 권리능력의 존속기간에 대해 사람은 생존한 동안 권리와 의무의 주체가 된다고 규정하고 있다(제3조). 이처럼, 살아있는 사람만이 권리와 의무의 주체가 된다는 점을 명확히 함으로써 저작권의 발생이나 침해에 있어서도 주체가 된다. 특히, 대규모 데이터의 경우에는 데이터의 속성상 클 수밖에 없다. 무엇보다, 대규모 언어모델은 그 인공지능망 내부에 학습 데이터를 암기(memorization)하고 있다가 그대로 재현해 낼 위험이 있다. 그 결과 학습 데이터에 포함된 개인정보를 재생성하는 등의 문제가 발생할 가능성도 존재한다.³⁸⁾ 이처럼, ChatGPT를 이용하는 과정에서 저작권 문제를 포함한 다양한 법적, 윤리적 이슈가 있음을 살펴해보았다. 다만, 현행 법률에서는 문제에 대한 책임을 사람에게 지우는 법률 시스템에 따라 ChatGPT를 매개하여 발생하는 문제의 주체도 사람이어야 한다. 따라서, 기본적으로 ChatGPT가 저작권을 갖지 못하는 것처럼 저작권 침해의 주체도 되지 못한다. 그렇다면, 저작권 침해는 발생했지만 책임지는 주체가 없게되는 상황이 발생할 수 있다. 따라서, ChatGPT를 생성하여 운용하는 OpenAI 또는 이를 생성한 이용자에게 물을 수 있을 것이다.

(2) 이용자인가, OpenAI인가?

다만, 이용자에게 물을 수 있는 것인지는 의문이다. 이용자는 OpenAI가 제공하는 일종의 도구인 ChatGPT를 이용한 것이기 때문에 조작을 통한 불법적인 결과물을 만들어낸 것이 아닌 행위의 결과에 대한 책임을 지기 어렵다. 설령, 생성해낸 결과물이 위법적인 것이라고 하더라도, 이 생성에 직접적인 관여를 한 것으로 보기 어렵다는 점에서 책임을 지우기는 어렵다. 물론, OpenAI도 중립적인 형태의 기술이라는 점을 주장할 수 있을 것이나, 실상 기계학습 과정에서 문제되는 데이터를 정제해야할 작위의무를 저야하는 것이 인터넷에서 생성되거나 유통되는 정보에 대한 최소한의 필터링 의무를 부담지울 수 있을 것이다.

그렇지만, 이러한 주장은 법리적으로 명확하지 않다. 따라서, 책임의 주체는 서비스를 제공한 사업자에게 지우는 것이 합리적이다. OpenAI가 될 것이라는 점에서 OpenAI를 침해 주체로 볼 수 밖에 없다. 권리발생과는 다르게, 실질적으로 서비스 과정에서 일어나는 것이므로 OpenAI에게 부담지우고자 한다.

38) 김병필, “대규모 언어모델 인공지능의 법적 쟁점”, 『정보법학』 vol.26, no.1 한국정보법학회(2022), 176면.

나. 침해 요건

학습데이터는 공개하지 않는다. 영업비밀을 이유로 하는 것이 아니라면, 하지 못하는 경우가 대부분일 것이다. 앞의 그림??에서 살펴본 바와 같이, 개인정보 침해물, 저작권 침해물, 경쟁사업자가 제공하는 정보에서 크롤링한 데이터 등을 기반으로 제작되었을 가능성이 무척이나 높기 때문이다.³⁹⁾

(1) 의거성

인공지능이 내부적으로 데이터를 어떻게 처리하는 것은 알 수가 없다. 이러한 이유 때문에 블랙박스라고 표현하기도 한다. 실상, 데이터가 공개되지 아니한 이상, 의거성을 확인하기 어렵다. 결과물을 통하여 역으로 유사성을 판단할 가능성도 있지만, 해당 데이터의 침해여부를 판단하기 쉽지 않다. 다만, 경우에 따라서 기계학습 과정에서 데이터를 암기하는 방식으로 학습이 이루어지는 경우라면 암기문 그대로 출력함으로써 저작물이 사용된 것을 확인할 가능성도 있다.⁴⁰⁾ 이러한 경우는 예외적인 경우에 해당하기 때문에 침해 여부 판단을 위한 기술적인 방법론을 별도로 강구하지 않는 이상, 어려울 것이다.

학습데이터가 외부에 공개되는 경우가 아니라면 침해여부 판단의 방법은 없다. 다만, 기계번역 결과에 나타나는 표현이나 특별한 단어의 번역 일치성이나 오류의 번역이 유사한 경우라면 추론할 수는 있다. 다만, 의거성을 확인할 수 있는 방법은 실질적 유사성을 토대로 분석할 수는 있다. 예를 들면, 특정 화가의 화풍이 그대로 나타나는 경우에는 화풍 자체는 저작권 침해로 보기어렵지만, 접근가능성에 따른 의거성을 주장할 수 있기 때문이다.

(2) 실질적 유사성

ChatGPT를 포함한 생성형 AI가 만들어낸 결과물이 타인의 저작권을 침해하였는지는 하나하나 분석해보지 않는 이상은 확인하기 어렵다. 접근가능성도 직접적으로 확인이 어렵기 때문에 실질적인 유사성을 바탕으로 추론할 수밖에 없다. 데이터의 파편들이 담겨있는 경우도 있겠지만, 앞서 살펴본 바와 같이 데이터를 암기하여 그대로 출력하는 경우도 발견된다는 점에서 의거성을 확인할 수 있다. 만약, 파편들이라도 확인하기 어려울 경우에는 실질적으로 유사한지를 판단할 수 있을 것이다. 참고로, 번역저작물에 있어서 대법원은 “번역저작물의 개개 번역 표현들을 구성하고 있는 어휘나 구문과 부분적으로 유사해 보이는 어휘나 구문이 대상 저작물

39) 블랙박스라는 의미는 내부적인 처리과정을 알 수 없다는 점에서 이름지워진 것이지만 정말 블랙박스를 열면 책임 하나를 따질 수 있기 때문에 비행기의 탑재된 블랙박스와 같은 의미도 담겨있다.

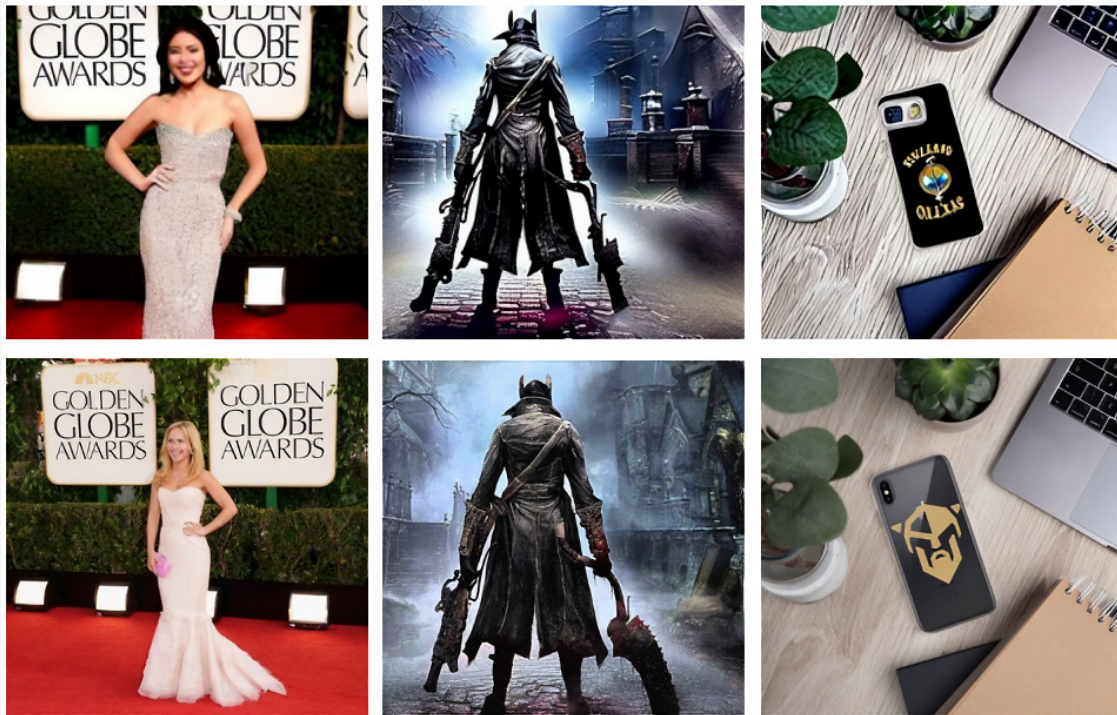
40) Nicholas Carlini et. al, “Extracting Training Data from Large Language Models”, <https://arxiv.org/abs/2012.07805> (2023.2.1. access)

에서 드문드문 발견된다는 사정만으로 바로 번역저작물과 대상 저작물 사이에 실질적 유사성이 있다거나 번역저작물에 대한 번역저작권이 침해되었다고 단정할 수는 없고, 그 실질적 유사성을 인정하기 위해서는 대상 저작물에서 유사 어휘나 구문이 사용된 결과 번역저작물이 갖는 창작적 특성이 대상 저작물에서 감지될 정도에 이르러야 한다”⁴¹⁾고 판시하고 있다. 고호풍의 이미지 학습을 할 경우에는 고호의 작품을 기계학습에 이용했다는 것을 추론할 수 있다. 다만, 고호는 저작권이 만료되어있기 때문에 저작권 보호에 대한 논란은 없을 것이나 저작권이 유효한 저작물을 기계학습에 이용하여 그 화풍을 학습하게 했다면 다른 결론에 이르게 된다. 즉, 화풍은 저작권의 보호범위에 포함되지 않기 때문에 결과물이 저작권 침해라는 것을 주장할 수 없으나, 학습과정에서 저작물이 사용된 사실은 부인하기 어려울 것이라는 점이다. 이러한 경우도 극히 예외적인 경우일 것이나 이러한 방식으로 침해 여부의 추론은 가능하다.

다. 침해가능성에 대한 추론

ChatGPT를 포함한 생성형 AI모델이 만들어낸 결과물이 학습데이터를 통째로 복제하는 수준이 반복되는 경우라면 실질적 유사성을 확인할 수 있을 것이다. 특히, 미드저니 등 이미지 생성 AI가 만들어낸 이미지는 유사한 것을 내보이는 경우가 많다. 실제, 아래 <그림>과 같이 생성형 AI모델이 원본 이미지와 유사하게 복제된 이미지를 생성해내놓는 경우도 발생하고 있다. 이러한 현상은 학습데이터가 방대하기 때문에 발생한다. 즉, 데이터의 특징을 학습한 것이 아닌 일부 학습데이터 자체를 암기해버리기 때문에 발생하는 현상이다.

41) 대법원 2007. 3. 29. 선고 2005다44138 판결.

[그림] 스테이블 디퓨전(stable diffusion)에서 복제된 이미지⁴²⁾

* source : TechCrunch(2022)

그렇지만, 기계학습은 특징점을 학습하기 때문에 전통적인 저작물의 이용방식과는 다른다는 주장은 설득력을 얻기 어렵게 되고 있다. 이와 같은 사례에서 볼 때, 실질적 유사성을 인정할 수 있는 가능성이 높다. 다만, 수백억건의 데이터 중 극히 일부라는 것만이 확인된다면 미세 이용요건에 따라 면책될 가능성도 배제할 수는 없다.

3. OSP 면책 여부

가. OSP 면책규정의 도입 취지

저작권법은 “온라인서비스제공자(OPS)”를 2가지 유형으로 정의하고 있다. 첫째, 이용자가 선택한 저작물등을 그 내용의 수정 없이 이용자가 지정한 지점 사이에서 정보통신망을 통하여 전달하기 위하여 송신하거나 경로를 지정하거나 연결을 제공하는 자, 둘째, 이용자들이 정보통신망에 접속하거나 정보통신망을 통하여 저작물 등을 복제·전송할 수 있도록 서비스를 제공하거나 그를 위한 설비를 제공 또는 운영하는 자 등이다.

42) 위가 AI가 복제하여 현시(display)한 이미지이며, 아래가 학습데이터로 사용한 원본 이미지이다.

ChatGPT가 이용자의 요구에 따라 생성한 콘텐츠가 타인의 저작권을 침해하는 경우에는 누가 책임을 지는지는 의문이다. 왜냐하면, ChatGPT는 하나의 도구로서 이용자의 요구에 따라 중립적으로 결과를 보여주는 것에 불과하기 때문이다. 다만, ChatGPT가 자체적으로 구축하는 과정에서 저작권을 침해한 것이 확인된다면 OpenAI가 직접적인 책임을 지는 구조가 될 것이다. 다만, 침해책임에서 ChatGPT가 OSP로서 위치를 갖는다면, OSP의 면책규정에 따라 책임에서 제외될 가능성도 고려할 필요가 있다. OSP가 인터넷 서비스의 발전과 이용자가 다양한 정보를 이용할 수 있는 혜택을 주기 때문에 정책적으로 면책규정을 도입한 것이다.

나. ChatGPT서비스 제공과 OSP면책

저작권법에 따른 OPS로서 OpenAI는 ChatGPT 제공에서 저작권법에 따른 면책가능성도 있다. 저작권법은 OSP에 대해서는 면책을 두고 있으나, 이 경우에도 이용자가 올리는 게시물 등에 대한 방조책임을 지우는 것이기 때문에 직접적인 침해에 대해서는 면책주장은 해당되지 않는다. ChatGPT나 기계번역과 같이 OSP의 서비스 유형이 저작권법에 규정된 OSP면책 유형에 해당하지 않을 것으로 보인다. 무엇보다, ChatGPT는 이용자의 요구사항에 대해 답변을 제공하는 방식일 경우에는 정보검색과 유사한 서비스로 볼 수 있다.

서비스의 성격도 이용자가 요구하는 프롬프트(prompt)의 내용을 해석하여 그에 맞는 답변을 가공하여 제공하는 것이라면 정보검색과 유사하기 때문에 저작권법상 OSP 면책유형에 포함된다. 따라서, ChatGPT가 생성하는 결과에 대해서는 면책가능성이 높다. 다만, 검색도구를 통하여 이용자에게 정보통신망상의 저작물 등의 위치를 알 수 있게 하거나 연결하는 행위로 규정하고 있어서, 엄격하게 해석할 경우에는 적용가능성이 배제될 수도 있다.

[표] 생성형 AI에 해당하는 유형

구분	내용	생성형 AI
제1유형 도관	내용의 수정 없이 저작물등을 송신하거나 경로를 지정하거나 연결을 제공하는 행위 또는 그 과정에서 저작물등을 그 송신을 위하여 합리적으로 필요한 기간 내에서 자동적·중개적·일시적으로 저장하는 행위	X
제2유형 공간제공	서비스이용자의 요청에 따라 송신된 저작물등을 후속 이용자들이 효율적으로 접근하거나 수신할 수 있게 할 목적으로 그 저작물등을 자동적·중개적·일시적으로 저장하는 행위	X
제3유형 정보검색	복제·전송자의 요청에 따라 저작물등을 온라인서비스제공자의 컴퓨터에 저장하는 행위 또는 정보검색도구를 통하여 이용자에게 정보통신망상 저작물등의 위치를 알 수 있게 하거나 연결하는 행위	O (대화형 검색)

* source : 연구자 작성

만약, 제3유형에도 적용이 배제될 경우에는 OSP면책규정을 적용받지 못한다는 점에서 생성형 AI가 일반인을 대상으로 하는 서비스에는 어려움이 예상된다. 따라서, 이러한 경우에 가능성이 있는 주장은 ChatGPT 서비스가 공정이용 요건에 해당하는지 여부이다.

4. 공정이용 여부

가. 공정이용 논의의 필요

ChatGPT를 제공하는 OpenAI는 저작권법이 규정하고 있는 OSP에 해당한다. 다만, 서비스 자체는 저작권법 제102조에 해당하는 유형으로 보기 어렵다. 따라서, 면책될 수 있는지 여부는 공정이용 규정에 따른 법원의 판단뿐이다. 기본적으로 OSP가 제공하는 서비스 그 자체로 수익을 발생하는 것은 아니다. 이용자의 정보활용을 위한 것을 목적으로 한다. 물론, OSP의 서비스 품질을 높이는 목적으로 활용되는 점도 부인하기는 어렵다. 다만, 검색과 마찬가지로 서비스는 저작자의 저작권을 침해할 목적으로 한다기 보다는 저작물에 담겨있는 정보를 파악함으로써 얻는 사회적인 이익이 크다는 점을 무시하기 어렵다. 썸네일 검색을 제공하는 포털사업자에 대하여 다음과 같이 공정이용에 해당한다는 이유로 면책한 바 있다. 이러한 이유와 더불어, “오로지 언어모델을 학습시키기 위한 목적으로 웹 스크레이핑을 수행하는 것은 저작권법상 공정이용에 해당하여 허용될 가능성이 작지 않다”⁴³⁾는 주장도 설득력을 얻는다.

나. 공정이용인지에 대한 판단

저작물의 통상적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있다. 다만, 저작물 이용 행위가 앞의 경우에 해당하는지를 판단할 때에는 다음 각 호의 사항 등을 고려하여야 한다.

(1) 데이터 이용으로서 기계학습

인간의 뇌구조를 모델링하여 구현한 인공지능 신경망은 인간의 것과 유사하다는 점에서 기계학습은 인간의 학습과정과 크게 다르지 않다. 기계학습 과정에서 인공지능은 저작물을 분석하여 특징들을 수치화하여 저장한다. 수치화한다는 것은 인간의 뇌에 저작물을 복제하는 것이 아닌 인간이 이해할 수 있는 특징만을 기억하는 것을 의미한다. 이미지의 경우, 해당 이미지의 특징을 분석하여 수치화하며, 텍스트의 경우는 말뭉치(corpus)를 인덱싱하여 데이터 값을 부여한다.

43) 김병필, “대규모 언어모델 인공지능의 법적 쟁점”, 『정보법학』 vol.26, no.1 한국정보법학회(2022), 189면.

이 과정은 저작물의 의미를 이해하거나 활용하는 것이 아닌 단어나 문장의 구성을 분석하는 것이다. 이는 저작권법이 의도하는 인간의 저작물 이용방식과는 차이가 있다. 분석된 결과물은 저작물 그 자체가 아닌 저작물에 담겨있는 특성, 패턴, 스토리, 구조 등의 것이다. 따라서, 인공지능이 학습하는 것은 메모리에 복제하는 것이 아닌, 데이터를 이해하는 상태로 분석하고 추상화하는 상태이기 때문에 복제가 일어나는 것으로 보기 어렵다.⁴⁴⁾ 인공지능의 학습과 관련하여 적용할 수 있는 공정이용에 관한 규정이다. 공정이용에 대한 고려에서 필요한 것은 헌법 제22조에 따른 창작자의 권리를 보호하는 것의 해석이다. 공정이용은 헌법상의 창작자 보호저작재산권의 제한규정이외에도, 저작물의 통상적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있기 때문이다.⁴⁵⁾

실제 사례를 보면, Fox사의 콘텐츠를 DB화하여 검색이 가능하도록 제공하는 TV아이스사에 대해 Fox사는 저작권 침해를 소송을 제기했고, 피고는 공정이용 항변을 주장했다. 방송저널리즘의 평가 및 비평 등 폭스사와는 전혀 다른 목적을 위해 비디오클립이나 검색(snippet)에 접근했다는 점에서 동 서비스는 사회적, 공공적 편익을 제공하는 중요한 공공성을 가진다고 판단했다.⁴⁶⁾ 이미 Google의 북서비스에 대해 원작의 전체에 대해 검색을 제공하는 것은 변형적 이용이며, 시장의 대체를 가져오지 않는다면 공정이용으로 판단했다.⁴⁷⁾ 동 판결에서 법원은 Google이 스캔한 책을 새로운 유형의 연구 도구를 만드는 데 사용했기 때문에 변형적이라고 보았다. 또한, Google이 스캔한 책을 사용하는 것이 저작권 보유자에게 시장에 해를 끼치지 않으며 프로젝트의 이점이 잠재적인 해악보다 더 크다고 판단했다. TDM(text and data mining)과 관련하여 중요한 점은 법원이 텍스트 및 데이터 마이닝에 대해 언급하면서 비소비적(non-consumptive) 연구 목적을 위한 도서의 TDM은 변형적 사용이며 공정사용에 따라 보호된다고 본 것이다.

(2) 통상적인 이용인지 여부

기계학습은 정보를 분석하여 그 패턴이나 특징값을 찾아내어 이용하는 것이기 때문에 인간의 이용 방식과는 다르다는 것은 앞서 살펴보았다. 즉, 인간의 저작물 이용과 달리, “저작물 그 자체를 향유하는 것이 아니라 단지 정보를 습득하고자 그 저작물을 구성하는 언어나 기호 등을 통계적으로 분석하는 경우에는 그 저작물 등을 복제하거나 번역 등 필요한 형태로 변환할

44) 만약, 기계학습 과정이 저작권 침해행위로 본다면 인공지능의 학습은 불가능할 수 있다. 아니면 빅데이터를 확보할 수 있는 인터넷기업들만이 경쟁에서 살아남을 수 있을 것이다.

45) 홍승기, “데이터마이닝 면책 입법 방향에 대한 의문”, 『경영법률』, 제32권 제4호, 한국경영법률학회(2022), 33면.

46) Fox News Networ, LLC v. TVEyes, Inc., 43 F. Supp. 3d 379 (S.D.N.Y 2014); 박성호, “텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작재산권의 제한”, 『인권과 정의』 vol. 494, 한국법학원(2020), 56면.

47) Authors Guild v. Google, Inc., 804 F.3d 202 (2nd Cir. 2015)

수 있다”⁴⁸⁾. 기계학습은 저작물 등의 복제나 단순한 2차적 저작물의 작성이 아닌 창작적 이용(creative use)⁴⁹⁾이라는 점에서 보면 공정이용에 해당할 가능성이 높다. 저작권법의 목적은 문화의 창달이며, 기존의 저작물을 향유하는 과정에서 새로운 창작적 표현을 만들어낼 수 있는 동인으로 작용하게 된다. 창작적 이용을 인정하는 수단으로써 공정이용은 저작권법의 목적규정을 통해 확인할 수 있다. 미연방대법원은 변형적 이용을 “새로운 표현, 의미 또는 메시지를 가지고 원저작물을 변형해, 다른 목적 또는 다른 성질을 가지고 원저작물의 표현에 무언가 새로운 것을 추가한 경우”⁵⁰⁾라고 판시했다. 변형적 이용이라면 “원저작물과는 다른 목적의 이용이고 원작의 성질에 대한 새로운 표현을 부가해 변화를 준 것”⁵¹⁾에 해당한다. 이처럼 기계학습은 새로운 가치를 부여함으로써 저작자가 의도했던 가치 이상을 더해주는 경우라면 이는 공정이용으로 판단될 가능성을 높이는 것이다. 이러한 맥락에서 기계학습은 데이터의 특징값을 분석해내어 새로운 가치를 만들어낼 수 있는 모델을 구축한다.

(3) 종류 및 용도의 적합성

기계학습은 인공지능의 성능을 향상시키기는 것이다. 알고리즘이 프로그래밍 된 바대로 데이터의 패턴이나 특징을 인식하고 분석하여, 의도한 결과를 만들어내거나 또는 의도성을 가지고 학습하는 것이다. 기계학습 과정은 용도라는 것이 저작물을 향유하는 과정과는 다른 용도라는 점에서 인간의 이용과는 본질적인 차이가 있다. 따라서, 공정이용 규정에서의 용도와 기계학습에서의 용도는 다른 기준점에서 봐야하며 저작물의 유형에 따라 달리 봐야하는 것은 맞다. 인간이 학습하는 것은 다양성 확보를 위한 것인 것처럼, 기계학습도 인공지능의 다양한 기능의 향상을 위한 것으로 궁극적으로는 인간의 사고와 유사한 범용 인공지능을 개발하기 위한 것으로 볼 수 있다. 정리하자면, 기계학습과 문화의 향유는 기본적인 용도나 목적이 상이하다. 따라서, 용도의 차이라는 점에서 본다면 기계학습은 저작물을 향유하는 것이 아닌 데이터에 담겨진 패턴과 특징값을 찾는 것이기 때문에 공정이용에 해당한다.

(4) 비중 및 중요성

기계학습에 의한 학습데이터의 이용의 경우, 사실 저작물이나 기능 저작물뿐만 아니라 비록 예술 저작물이라고 하더라도 저작물의 문예적, 심미적 가치를 이용하는 것이 아니라 그 속의

48) 임원선, 『실무자를 위한 저작권법』, 저작권위원회(2014), 231~232면.

49) 공정이용인지 여부에서 저작물의 변형적 이용(transformative use)에 대해 판단한다. 여기서는 이를 창작적 이용이라고 표현하나, 그 의미는 변형적 사용과 다르지 않다. 정상조, “딥러닝에서의 학습데이터와 공정이용”, 『LAW & TECHNOLOGY』, 제16권 제1호 (통권 제85호), 서울대학교 기술과법센터(2020), 12면.

50) Campbell v. Acuff-Rose Music, Inc., 510U.S.569(1994).

51) 한국정보법학회 지음, 『인터넷, 그 길을 묻다』, 중앙Books(2012), 544면.

데이터로서의 비표현적 가치를 이용하는 변형적 이용이 많기 때문에 저작물의 종류 및 용도가 크게 영향을 주지는 않을 것이다.⁵²⁾ 학습데이터의 이용은 투입단계와 산출단계의 원저작물 이용 분량이 서로 다를 수 있기 때문에 산출단계의 경미하거나 부수적인 이용으로 비중면에서 공정이용에 해당한다.⁵³⁾ 참고로, 일본 저작권은 컴퓨터를 이용한 정보처리에 의한 새로운 지식 또는 정보를 창출함으로써 저작물이용의 촉진에 이바지하는 정보검색서비스라거나 정보해석서비스는 타인의 공개된 저작물의 이용용도 및 분량에 비추어 저작권자의 이익을 부당하게 해치지 않는 소위 경미한 이용으로서 적절한 이용에 해당한다.⁵⁴⁾

(5) 시장대체성

시장대체성 여부에 대해 살펴본다.⁵⁵⁾ 시장대체성의 범위는 “원저작물 자체뿐만 아니라 2차적 저작물의 시장이나 가치도 포함된다”⁵⁶⁾고 한다. 기계학습은 인간의 이용이 아닌 정보 내용이나 표현의 특성을 학습하기 때문에 일반적인 이용 형태와 다를뿐더러, 일반 소비자에게 제공되는 것과는 다른 시장을 형성하게 된다. 기계학습은 인공지능의 알고리즘을 고도화하기 위한 것에 불과하기 때문이다.⁵⁷⁾ 다만, 기계학습을 위한 별도의 정보(빅데이터)를 구축하여 제공한다면 이는 시장대체성을 인정받을 수 있다. 일본은 이러한 상황을 입법론으로 정리했다. 즉, 정보분석을 위한 빅데이터 등의 이용을 저작권 제한규정으로 입법화했다.⁵⁸⁾ 이에 따라, 인공지능을 학습하는 과정에서 이루어지는 저작물의 습득 자체는 학습 매커니즘이지 저작물을 복제하여 배포하는 것으로 보기 어렵다. 또한, 인공지능의 학습 형태에서 빅데이터 등의 정보를 이용하는 것은 “저작물 등을 구성하는 언어나 기호 등의 요소들 또는 그들이 관계 등을 분석하려는 것일 뿐 그 저작물 등 자체를 이용하고자 하는 것이 아니고, 그 분석의 결과물을 그 저작물

52) TDM과 관련해서는 TDM 분석 결과에서 원저작물을 감득하기 어렵기 때문에, 저작물의 성질에 관한 요소에는 큰 비중을 두지않고 있다고 한다. 김경숙, “TDM 관련 저작권법 개정안의 비판적 고찰”, 『경영법률』 제31권 제3호, 한국경영법률학회(2022), 122면.

53) 저작권법 제35조의3(부수적 복제 등) 사진촬영, 녹음 또는 녹화(이하 이 조에서 “촬영등”이라 한다)를 하는 과정에서 보이거나 들리는 저작물이 촬영등의 주된 대상에 부수적으로 포함되는 경우에는 이를 복제·배포·공연·전시 또는 공중송신할 수 있다. 다만, 그 이용된 저작물의 종류 및 용도, 이용의 목적 및 성격 등에 비추어 저작권자의 이익을 부당하게 해치는 경우에는 그러하지 아니하다.

54) 정상조, “딥러닝에서의 학습데이터와 공정이용”, 『LAW & TECHNOLOGY』, 제16권 제1호 (통권 제85호), 서울대학교 기술과법센터(2020), 18~19면.

55) Robert Merges, Peter Menell, Mark Lemley, 『Intellectual Property in the New Technological Age』, Wolters Kluwer, 2012, p.646.

56) 최호진, “썸네일 이미지와 공정이용”, 『Law and Technology』, 제8권 제3호, 서울대학교 기술과법센터(2012), 70면.

57) 물론, 수많은 인공지능에 탑재할 목적으로 이용했다면 시장대체성을 상실할 가능성도 부인하기 어렵다.

58) 일본 저작권법 제47조의7(정보해석을 위한 복제 등) 저작물은, 전자계산기에 의한 정보해석(다수의 저작물 기타의 대량의 정보로부터, 당해 정보를 구성하는 언어, 음, 영상 기타의 요소와 관련된 정보를 추출, 비교, 분류 기타의 통계적인 해석을 행하는 것을 말한다. 이하 이 조에서 같다)을 하는 것을 목적으로 하는 경우에는, 필요하다고 인정되는 한도에서 기록매체에의 기록 또는 번안(이에 의하여 창작한 2차적저작물의 기록을 포함한다)을 할 수 있다. 다만, 정보해석을 하는 자의 이용에 제공하기 위해 작성된 데이터베이스저작물에 대하여는 그러하지 아니하다.

등과는 전혀 별개로서 그에 원 저작물이 드러나지 않으므로 그 저작물 등의 통상적인 이용과 충돌하거나 저작자의 정당한 이익을 부당하게 침해할 우려가 적다”⁵⁹⁾고 볼 수 있다. 인공지능의 학습과 유사하게 적용할 수 있는 기존 사례는 썸네일 검색이다. 인터넷에 공개된 정보를 크롤링하여, 이를 데이터베이스화하고 검색어가 입력되면 해당 정보를 제공하는 것은 기계학습 매커니즘과 유사기 때문이다. 썸네일 검색은 그 결과를 보여주는 것이지만, 크롤링은 기계학습을 위한 데이터 수집 내지 수집된 데이터를 인덱싱하여 관리 값을 부여하기 때문이다. 기계학습은 특징값을 분류해내는 과정이라는 점에서 차이가 있지만, 정보를 분석하여 분류하는 과정과는 크게 다르지 않다. 물론, 정보를 분석하여 이용가능한 상태에 놓인 것은 공개되거나 출시된 것이 아니기 때문에 시장대체성을 논하는 것이 타당하지 않다는 지적도 가능하다. 그렇지만, 시장대체성을 해당 저작물의 이용과정에서 고려하는 예측에 대한 판단이기 때문에 이를 부인할 필요는 없다. 썸네일 형태로 검색결과에 노출되는 것도 정보의 위치를 알려주는 것으로 공익적 성격으로써 공정이용이 인정되고 있고,⁶⁰⁾ 대법원도 같은 취지로 저작권 침해를 부인한 바 있다.

59) 임원선, 『실무자를 위한 저작권법』, 저작권위원회(2014), 232면.

60) 구글검색엔진의 높은 수준의 변형적 이용과 사회적 편익을 제공한다는 점에서 공정이용에 해당한다고 판시한 바 있다 (Perfect10, Inc. v. Amazon, Inc., 508 F.3d 1146(9th Cir, 2007)).

memo

