

초거대 생성 AI의 국내외 현황과 이슈

조영임

가천대학교 컴퓨터공학과 교수

yicho@gachon.ac.kr

머리말

2016년 알파고가 등장해 'AI가 사람과의 대결에서 이길 수 있다'라는 충격을 안겨 주었는데, 2022년 11월에는 챗GPT(ChatGPT)가 등장해 'AI가 사람처럼 쓰고 말하고 행동할 수 있다'는 큰 충격을 전 세계에 안겨주고 있다.

'초거대 AI'란 일반적으로 매우 크고 강력한 인공지능 모델을 가리킨다. 이러한 모델은 대규모 데이터셋에서 훈련되며, 많은 양의 계산 자원을 사용해 학습과 추론 작업을 수행한다. 초거대 AI는 많은 문제를 해결할 수 있는 능력을 갖추고 있으며 이미지 인식, 자연어 처리, 추천 시스템 등 다양한 작업에 사용될 수 있다. 예를 들면 2023년 3월 15일 발표된 GPT-4 모델이 가장 최신의 초거대 AI 모델이다.

반면에 '생성형 AI' 또는 '생성 AI'란 텍스트나 이미지, 음성 등을 생성하는 데 특화된 인공지능을 의미한다. 여기서 '생성(Generative)'이란 일일이 AI에 지시하거나 학습시키지 않아도 알아서 이용자가 요구하는 바를 만드는, '이것저것 다 할 수 있는 범용 AI'의 등장을 의미하는 것이다. 생성 AI는 주어진 입력으로부터 새로운 콘텐츠를 생성하거나, 이미지를 생성하거나, 대화를 주도하는 등의 작업을 수행할 수 있다. 이러한 모델은 일반적으로 조건부 생성 모델이라고도 불린다. 초거대 AI 모델인 GPT-4는 이미지 데이터를 활용할 수 있으며, GPT-4에 이미지를 입력하면 캡션을 만들거나 이미지를 분류 및 분석할 수 있는 특징을 갖고 있다. 사용자가 채팅하듯 질문하면 챗GPT는 학습한 데이터를 기반으로 '사람처럼 문장으로 답변'을 해주는 기술이다. 실제로 챗GPT는 지금껏 출시된 서비스 중 가장 빨리 이용자를 확보하며 산업, 직업, 교육 등 사회 전반적으로 파급력을 확산 중이다. 주요 서비스 이용자 100만 명을 달성하는데 소요된 기간이 넷플릭스는 3.5년, 인스타그램은 2.5개월인데 반해 챗GPT는 5일에 불과했다고 한다.

요약하면 초거대 AI는 크고 강력한 인공지능 모델을 가리키는 개념이고 생성 AI를 포함하는 개념이지만, 생성 AI는 특정 작업에 특화된 생성 모델을 가리키는 개념으로 초거대 AI와는 다른 다양한 유형의 인공지능 모델을 포함할 수 있다. 따라서 "초거대 AI"와 "생성 AI"는 개별적인 용어로 사용되는 것이 일반적이거나, "초거대 생성 AI"는 이 두 용어를 결합한 것으로 이해될 수 있다.

이처럼 초거대 AI와 생성 AI는 용어가 주는 복잡함과 거대함, 그리고 엄청난 기술의 진보로 인해 전 세계에서 많은 관심을 받고 있다. 본고에서는 초거대 AI와 생성 AI에 대해 국내외 동향 및 기술, 주요 이슈들과 연구방향을 살펴보고자 한다.

초거대 AI와 생성 AI의 국내외 동향

• 국내외 동향

OpenAI가 최근 출시한 GPT-4는 냉장고 안의 재료 사진을 입력하면 바로 그 재료로부터 요리할 수 있는 음식 몇 개를 보여준다고 한다. GPT-4는 입력으로 이미지나 영상을 받아서 이미지나 영상의 이름이 무엇인지를 말하는 정도가 아니라, 이것들로부터 그에 걸맞은 대답을 보여주는 것이다. 즉 사용자가 냉장고의 이미지나 영상을 보여주면 GPT-4는 이용자가 음식을 만들 것이라 예측한 후 만들어낼 수 있는 종류를 나열해 주는 것이다. 만약 사용자가 다른 생각을 하고 있었다면 아마 다른 대답을 보여주었을 것이다.

또한 각종 주요 시험에서 상위 10%에 해당하는 놀라운 수준을 보여주고 있다. 2023년 5월 요리미이신문에 의하면, OpenAI가 만든 대화형 AI인 ‘챗GPT-3.5’와 ‘챗GPT-4’에게 최근 5년 치 일본 의사국가시험을 풀게 한 결과, 챗GPT-3.5는 모두 불합격했지만, 최신 버전인 GPT-4는 모두 합격했다고 한다. 이는 ‘사람 수준의 능력’을 보여준 것으로 보여준 사례이다. GPT-4는 인간 응시자의 평균 점수를 밀돌았고, 임산부 환자에게 투여할 수 없는 약을 고르거나 환자에게 안락사를 권하는 등 부적절한 답을 내놓기도 했다고 한다.

GPT-4는 2023년 1월, 미국 미네소타대 로스쿨 시험에서 헌법과 세금 및 불법행위 등 광범위한 내용을 주제로 에세이를 작성해서 합격한 바 있다. 챗GPT에 객관식 문항 95개와 에세이 문항 12개로 이뤄진 실제 학생들이 치르는 시험 문제와 동일한 문제를 출제해 풀게 한 결과이다.

이처럼 초거대 AI로 분류되는 챗GPT와 같은 기술은 이머징 기술(Emerging technology)로, 매우 신속한 기술개발과 고도의 전문적인 인력이 필요한 분야이다. 최근 ‘GPT-4’ 등장과 함께 빅테크들을 비롯한 글로벌 기업들은 초거대 AI에 주목하기 시작했고, 본격적인 경쟁이 시작됐다. 이에 관련 기술 기업과 각국 정부들도 매우 긴박하게 대응하고 있다.

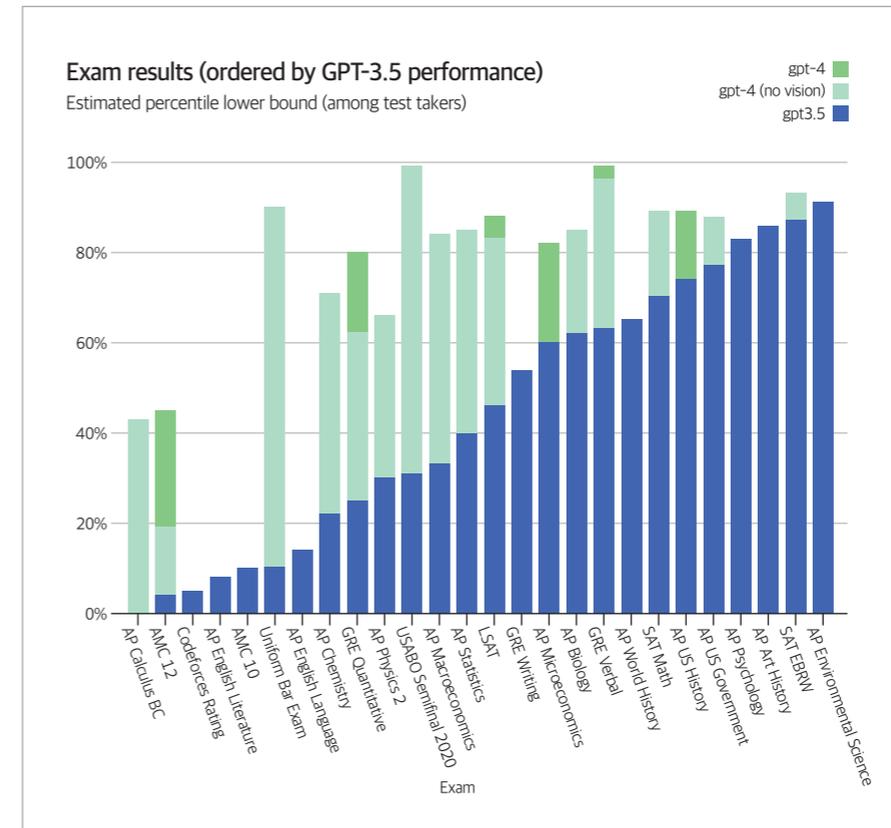
최근 큰 충격을 안겨준 챗GPT와 같이 인간의 지능을 구현하기 위해 인간 지능의 기본소자인 뉴런으로부터 시작한 연구배경에는 서브 심볼릭 뉴런으로 대표되는 초거대 AI가 있다. 일반적으로 인공지능은 ISO/IEC 22989:2022에서 표준으로 제시하는 바와 같이 패러다임이 크게 두 가지로 나뉜다. 먼저 전통적인 전문가 시스템과 같은 심볼릭 인공지능(Symbolic AI)과 인간의 뇌의 기능을 그대로 옮기려는 신경회로망 기반의 서브 심볼릭 인공지능(Subsymbolic AI)으로 구분된다.

전문가의 지능을 구현하려는 심볼릭 인공지능은 전문가의 업무를 수행하기 위해 수많은

변수와 논리를 고려해야 하나, 전문가의 지능을 컴퓨터에 옮기는 것은 불가능에 가까운 일이었기 때문에 사실상 인공지능 주류에서 밀려나게 됐다.

일반적으로 AI 모델의 크기는 폭과 깊이로 측정한다. 매개변수의 개수가 많을수록 폭이 넓고, 레이어의 개수가 많을수록 깊이가 깊다. 즉 처리하는 레이어가 많다는 뜻이다. 챗GPT는 OpenAI의 초거대 AI GPT-3 모델을 개선한 GPT-3.5를 기반으로 2022년 11월에 출시됐고, 2023년 3월에는 GPT-4 버전으로 개선돼 출시됐다. GPT-3의 경우 매개변수가 1,750억 개, 레이어가 100개에 이른다. GPT란 ‘Generative Pre-trained Transformer’의 약자로 OpenAI가 개발한 인공지능 모델이다. 특히 GPT는 자연어 처리를 위한 인공지능 모델로, 방대한 데이터를 머신러닝으로 미리 학습해 이를 문장으로 ‘생성(Generative)’한다. 현재 GPT-4 기반의 챗GPT는 ‘ChatGPT 플러스’를 통해 유료로 사용 가능하다.

[그림 1] 챗GPT3.5와 GPT-4의 각종 주요 시험 결과 비교



MS(MicroSoft)와 OpenAI는 오랫동안 협력관계를 유지해 왔는데, 2019년에는 MS가 OpenAI에 10억 달러(약 12조 원)를 투자했다. 두 회사는 AI 및 딥러닝 분야에서의 협력에 합의했는데, 이 협력은 OpenAI의 인공지능 모델인 GPT-4를 MS Azure 클라우드 플랫폼에서 제공함을 명시하고 있다. 즉 MS Azure는 GPT-4를 개발하고 배포하기 위한 인프라를 제공하며, 이를 통해 기업이나 개발자가 GPT-4를 활용해 다양한 AI 응용 프로그램을 구축할 수 있도록 협력한다. 또한 OpenAI와 MS는 에너지 산업의 지속 가능성 및 친환경적인 해결책을 모색하기 위해 MS Azure를 사용하는 등 에너지 분야에서의 협력도 강화하고 있다. 인공지능에서 입력에 대해 출력을 산출하는 과정을 '추론(Inference)'이라고 한다. 인간의 뇌에는 약 1,000억 개의 뉴런(신경세포, Neuron)이 존재하고 각각의 뉴런은 다른 뉴런들과 정보를 주고받기 위해 연결되는데 이 부위를 시냅스(Synapse)라고 한다. 인간의 뇌에는 이러한 시냅스가 약 100조 개 이상 존재한다. 뉴런과 시냅스, 그리고 이들을 잇는 어마어마한 수의 연결 패턴(보통 가중치라 함)은 우리가 뇌를 이해하는 데 매우 중요한 요소이다. 뉴런들은 기본적으로 전기적 방법으로 소통하지만, 뉴런들끼리 신호 전달은 주로 시냅스에서 물질을 교환함으로써 이루어진다. 즉 흥분성 물질과 억제성 물질을 적절히 전달함으로써 정보를 전달하고 궁극적으로는 인간의 뇌의 기억을 형성해 나간다.

인간의 뇌에는 컴퓨터에서 볼 수 있는 메모리와 같은 칩도 없고 명확한 기억장치도 없다. 그럼에도 인간이 기억을 할 수 있는 것은 이러한 뉴런과 시냅스, 그리고 이들 사이의 정보를 전달하는 데 있어서 중요한 연결패턴들 때문이다. 인간의 뇌는 특별한 칩이 없어도 이러한 뇌 안의 뉴런들과 시냅스, 그리고 연결패턴들의 하모니에 의해 기억하고 판단하고 인식하게 되는 것이다.

인간은 기본적으로 멀티모달(다양한 정보들을 한꺼번에 처리할 수 있는 능력) 처리능력이 있다. 인간의 뇌는 자극을 받으면 반응하는데, 동시에 가질 수 있는 시각정보는 17개이며 이러한 보통의 시각정보는 약 1초 정도 뇌에 머문다. 이때 반복하지 않으면 20초 이내에 모든 시각정보는 우리의 뇌에서 사라진다. 즉 망각하게 된다. 이것을 좀 더 오래가고 잊어버리지 않게 하고자 인간은 본능적으로 '메모'라는 것을 하게 됐는지도 모른다. 인간의 뇌는 보통 5~9개 정도의 일을 동시에 추진할 수 있다고 한다. 하버드대학 조지 밀러 부교수의 논문 「마법의 숫자7+(-)2:인간의 정보처리 능력의 한계」에서 다양한 심리학자의 실험자료를 정리해 인간이 한 번에 기억할 수 있는 용량은 7+(-)2개라는 가설을 입증한 바 있다.

조지 밀러 교수는 먼저 참가자들에게 여러가지 소리를 주파수 100Hz에서 8,000Hz까지로 소리사이의 간격은 25초로 설정해 들려준 다음, 곧바로 이것은 몇 번째 소리였는지 물어보고 참가자에게 점점 많은 소리를 들려주며 몇 개까지 기억하는지 실험했다. 그 결과

참가자들은 4개 정도까지는 거의 틀리지 않았지만 6개를 넘기자 틀리기 시작했고 8개를 들려주었을 때는 정답률이 80%로 떨어진다. 즉 인간이 소리를 정확하게 기억해 재현할 수 있는 한계가 약 5개라는 사실을 증명했다.

또한 인간의 두뇌는 새로운 정보의 60%는 한 시간 안에 잊어버리는 한계를 지녔으나, 4~5회 반복하게 되면 전체 내용의 약 70-80%는 기억하게 된다. 이러한 과정을 거치면 인간은 장기기억체제로 돌입하게 되는 것이다. 따라서 반복적인 사고는 인간의 기억력을 증진시킨다.

챗GPT는 이러한 인간을 모델링했기에 다양한 멀티모달 특성을 보이려고 노력했다. 즉 여러형태의 정보를 동시에 처리하고 판단하려고 한다. OpenAI API는 챗GPT, 알파고, 클로바 등 자연어, 코드, 이미지 이해 또는 생성과 관련된 거의 모든 작업에 적용할 수 있다. OpenAI는 다양한 작업에 적합한 다양한 수준의 강력한 모델 스펙트럼과 사용자 정의 모델의 세부 조정 기능을 제공한다. 이러한 모델은 콘텐츠 생성부터 의미론적 검색 및 분류에 이르기까지 모든 작업에 사용될 수 있다.

인간은 생명체이므로 성능이 무한할 수가 없다. 기억력도 마찬가지이다. 그러나 챗GPT와 같은 범용 AI(인간과 같은 능력을 갖는 보편적인 AI)에서는 생명체가 아니므로 인간과 달리 성능이 무한할 수가 있다는 점에서 인간을 어떤 면에서는 능가할 수 있을 것이다.

인공지능은 궁극적으로 인간과 소통할 수 있는 것을 목적으로 한다. 인간과 소통하려면 인간의 지능 수준과 유사해야 한다. 알란 튜링(Alan Turing)은 인간과 혼동할 수 있는 튜링머신을 1936년에 만들었고, 스투어트 러셀(Stuart Russell)은 "인간과 대화가 되면 인간은 인공지능의 지적 수준을 사람과 같다고 본다"라며 인공지능의 목표를 제시했다. 챗GPT는 일단 인간과 대화가 되고, 인간과 혼동스럽기까지 한 것을 보면 분명 인공지능의 목표에 어느 정도 도달한 것으로 보인다. 물론 오답도 나오고 엉뚱한 결과를 산출해 가끔 웃음을 주지만, 그래도 여러 실험결과에서 보면 매우 그럴듯한 대답을 추론해 낸다.

챗GPT가 몇 개까지 멀티모달 정보를 처리할지, 한 번에 몇 개의 정보량을 처리할지는 하드웨어와 소프트웨어에 달린 문제이므로 성능 면에서는 앞으로 인간과 비교가 되지 않게 진화할 것이다. 다만 정확도, 공감, 인간의 윤리 등 측면에서 인간과 어떤 차이를 보일지는 계속 비교하게 될 것이다. 이 엄청난 인공지능의 산출물 앞에 인간은 어떻게 인공지능과 공존해야 하는지 다시 한번 심각하게 생각하지 않을 수 없다.

이러한 배경하에 해외 초거대 생성 AI 빅테크 기업들은 앞다퉈 생성 AI 모델 개발과 관련 서비스에 집중하고 있으며, 우리나라 기업들도 생성 AI 개발에 박차를 가하고 있다. 해외의 경우 구글의 바드(Bard), OpenAI의 챗GPT와 달리(DALL-E), MS의 빙챗(Bing-Chat), 메타의

라마(LLaMA) 등이 해외의 주요 초거대 AI 모델과 관련된 서비스들이다. 미국은 영어 중심의 초거대 AI 모델을 개발하며 일부 기업은 이를 OpenAPI 형태로 공개해 시장 선점을 타진하고 있다. 중국은 정부 주도로 데이터·엔지니어 확보 및 집중투자를 통한 초거대 AI 모델 구축을 발표했으나, 모델은 비공개 상태이다.

[표 1] 해외 초거대 생성 AI 개발 현황

기업명	모델명	방식	파라미터 수	공개여부
딥마인드	Gopher	GPT	3,500억 개	비공개
구글	PaLM	GPT	5,400억 개	비공개
구글	OPT-175B	GPT	1,370억 개	공개
구글	LaMDA2	GPT	1,370억 개	공개
OpenAI	ChatGPT	GPT	1,750억 개	공개
중국인공지능아카데미	우다오	GPT	1조 7,000억 개	비공개
바이두	어니3.0	GPT	2,800억 개	비공개

출처: 국가과학기술심의회위원회, 2023

우리나라는 2023년 2월 인공지능 최고위 전략대회와 2023년 4월에 초거대 AI 경쟁력 강화방안을 발표한 바 있다. 국내는 포털·통신사를 중심으로 초거대 AI 모델을 빠르게 구축하며 한국어 서비스 성능 고도화를 추진 중이다. 네이버는 초거대 ‘하이퍼클로바’ 고도화로 상반기 새 검색 경험 ‘서치GPT’를 출시할 예정이며, 신뢰성, 최신성 부족, 한국어 번역 시 정확성 저하 등을 보완하도록 구성했다. 카카오는 연내 카카오브레인 ‘KoGPT’ 기반 버티컬 AI 서비스를 출시할 예정이며, 한국어 특화 AI 모델 KoGPT를 활용해 우리가 잘할 수 있는 버티컬 AI 서비스에 집중하고 있다. 국내는 대기업을 중심으로 한국형 초거대 생성 AI 모델을 자체 구축했으나 결과의 불완전한 신뢰성으로 사업화에 어려움을 겪고 있는 상황이다.

[표 2] 국내 초거대 생성 AI 개발 현황

기업명	모델명	방식	파라미터 수	개요
네이버	HyperCLOVA	GPT	840억 개	세계 최대 한국어 초대형 언어모델로, 검색어 교정·노코드AI·시스피커 기능 강화에 활용.
카카오	KoGPT	GPT	300억 개	공개 모델 중 한국어 최대 규모

기업명	모델명	방식	파라미터 수	개요
카카오	min-DALL-E	GPT	600억 개	공개 모델 중 영상-언어(영어) 연계 최대규모
SKT	GLM	GPT	180억 개	세계 최초 한국어 대화 서비스‘에이닷’ 상용화
KT	마음	T5	180억 개	과금 기반 OpenAPI 서비스 및 내부 인력 기반 업무 자동화에 활용을 확대 중
LG	엑사원	GPT	3,000억 개	국내 영상-언어(영어) 연계 모델 중 최대 규모 모델. 한국어 서비스 불가로 해외 제휴 혹은 사내 원천 AI연구에 활용 중

출처: 국가과학기술심의회위원회, 2023

• 이슈사항

초거대 생성 AI에서 등장하는 이슈들은 다음과 같다.

첫째, 초거대 생성 AI는 특징상 대규모로 수집한 데이터를 기반으로 학습하고 있으나 학습데이터 부족, 과도한 컴퓨팅 리소스 사용, 최신성 유지, 신뢰성 확보 등의 문제해결 이슈가 지속적으로 제기되고 있다. 최근에는 데이터 경량화 기술, 머신러닝 학습 추론 방법의 개선 등이 이를 해결하기 위해 연구되고 있는 주제들이다.

둘째, 초거대 생성 AI활용이 본격화되면서 AI의 비윤리성·편향성 이슈가 확산되고 있으며, 비윤리성·편향성 제거(Filtering) 기술에 대한 관심도 높아지고 있다. 이는 사전학습에 사용한 데이터에 의존적으로 인종·성별·정치성향 등이 편중된 콘텐츠가 생성되기 때문이다. 최근 AI모델의 편향성 제거를 위해, 모델에 영향을 줄 수 있는 데이터의 특징을 사전에 규명하고 분류 및 제거하기 위한 공정성 평가 도구, 프레임워크들이 연구되고 있다. 특히 해외에서는 거대 IT기업에서 학습데이터 및 모델의 다양한 검증을 위해 비상업적 용도로 모델 및 데이터를 공개하는 등 문제가 발생했을 때 빠르게 조치하기 위한 체계도 구축 중이다.

구글은 2022년 이미지/영상 처리 AI에 응용하기 위한 몽크 스킨 톤(MST: Monk Skin Tone) 피부색 구분 체계(Scale)를 공개하고, 인종 기반 피부색 대신 채도에 따른 10단계 체계로 AI 모델을 평가함으로써 AI 모델의 인종적 편향성을 낮추기 위해 노력했다.

MS는 AI 모델의 성별, 문화적, 민족적 편향성을 학습 단계에서 제어하기 위한 오픈소스 Fairlearn을 공개하고, 오픈소스를 통해 사용자는 데이터 지표를 기반으로 민감한 특징을 사전에 확인할 뿐 아니라 모델의 성능 및 예측에 따른 평가 결과를 확인할 수 있다.

메타는 거대 언어모델 OPT-175B를 비상업적 서비스로 출시하며, 학습 데이터 및 모델과 코드 등 관련 내용을 모두 공개했고, 이는 사전 테스트에서 발견하지 못한 자체 모델의

편향과 오용을 여러 연구자가 다양한 분야에서 테스트함으로써 모델의 편향성을 최소화하기 위한 전략이다.

국내 IT 기업들이 AI 윤리 준칙을 발표하고, 연구소에서 AI 공정성 판단 프레임워크를 발표하는 등 기술 발전과 맞물려 AI 모델의 편향을 방지하기 위한 다양한 방법들이 시도되고 있다.

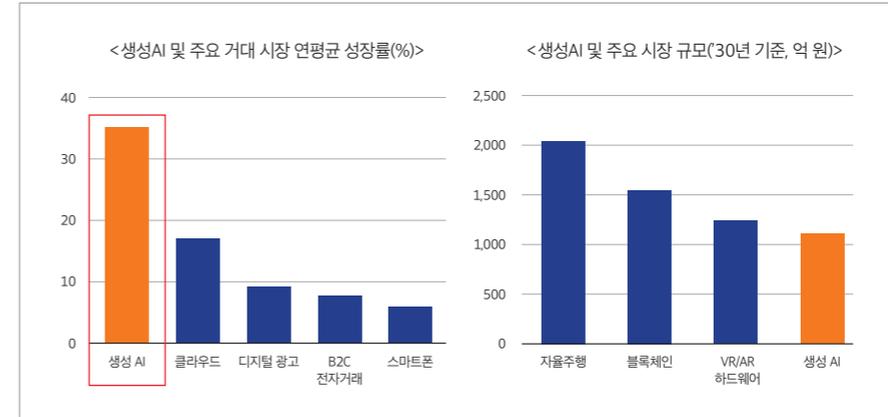
네이버 하이퍼클로바 연구팀에서는 AI 모델의 편향성을 판단해 권장되는 내용, 권장되지 않는 내용에 대한 생성 확률을 각각 최대화, 최소화하는 학습 방식을 연구하고 있다. 다만 주관적인 사람의 판단이 포함되기 때문에 기준을 수립하기 어렵다는 문제가 있다.

카이스트 인공지능 공정성 연구센터에서는 인공지능 모델의 편향성을 판단하기 위한 프레임워크 MSIT AI FAIR 2022를 개발했다. 해당 프레임워크는 AI 모델과 학습데이터의 편향성을 사전에 분석 및 완화하는 기능을 갖추고 있다.

셋째, 할루시네이션(Hallucination) 현상으로 인한 부적절한 답변의 생성을 방지하기 위한 연구가 필요하다. AI에서 할루시네이션 또는 인공 할루시네이션(Artificial Hallucination)은 학습 데이터를 통해 판단하지 않는 것처럼 비쳐지는 AI에 의한 확신적 답변이다. 예를 들어 테슬라의 소득과 관련한 학습 데이터가 없는 챗봇의 경우 내부적으로 알고리즘이 높은 신뢰도로 순위를 올리는 무작위 숫자(예: "\$13.6 billion")를 생성한 다음 사실이 아닌 방향으로 움직이면서 반복적으로 테슬라의 소득을 \$13.6 billion으로 언급한다. 더불어 해당 수치가 생성 알고리즘의 약점의 결과임을 드러내지도 않는다. 즉 진실 여부와 출처가 불분명한 데이터를 바탕으로 챗봇은 논리적이고 그럴듯해 보이는 답변을 계속 만들지만, 이용자는 틀린 정보를 수용해 잘못된 판단을 내놓을 수 있다. 이러한 현상을 이른바 "할루시네이션"이라고 하는데 이는 인간심리학에서의 환각 현상에 빚댄 것이다. 현재 검증 가능한 생성형 AI 기술은 일상화가 촉진됨에 따라 결과의 신뢰성 문제가 대두되고 있으나 해결방법이 명확하지 않다. 결과의 사실 여부를 판단하고 할루시네이션 문제를 해결하기 위해 마스킹 기술, 증명트리(Proof Tree) 구성 등의 다양한 기술적 시도가 이루어지고 있고, 특히 학습과정에서 사람이 피드백에 관여하거나 최종 결과를 제공하기 전에 생성결과를 검증하는 등의 방법 등이 이를 해결하기 위해 연구되는 주제들이다.

넷째, 인재양성에 대한 투자와 계획이 필요하다. 2022년 그랜드뷰리서치와마켓앤마켓 분석 자료에 의하면, 초거대생성 AI 시장은 2030년 이전까지 세계 AI시장의 10% 비중을 차지할 것으로 전망되며, 주요 거대 시장 중 가장 높은 성장률을 보이는 고성장 시장이 될 것이라고 전망했다.

[그림 2] 생성 AI 시장 규모 및 연평균 성장률 전망



출처: 국가과학기술심의회, 2023

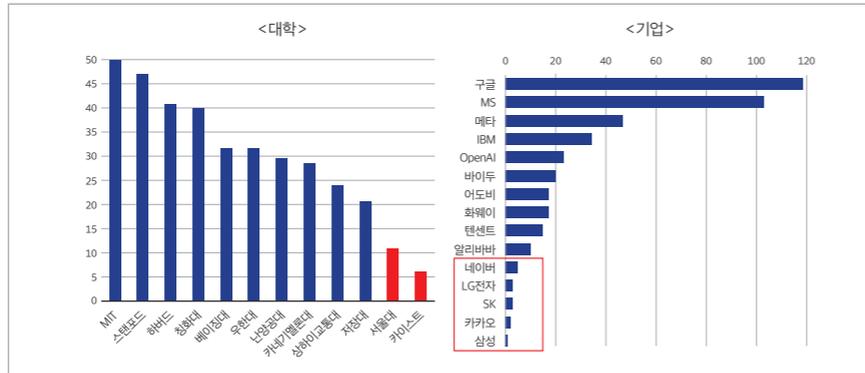
특히, 챗GPT의 등장으로 지난 10년간 구글이 지배했던 검색엔진 시장에 균열이 발생하고 있다. 실제로 챗GPT 등장 이전에는 구글이 93.4% 시장을 차지했고 MS가 약 2.8% 검색시장을 차지했으나, 챗GPT가 등장한 '22년 11월 이후로는 구글은 'Code Red'를 발령하고 비상사태에 들어갔으며 MS는 불과 몇 달 만에 사용자가 6배 이상 증가하고 일일 사용자도 1억 명을 돌파하는 등 성장추세가 매우 가파르게 변화하고 있다.

그런데, 최근 OpenAI는 GPT-4를 내놓으며 API 사용료를 15~53배까지 인상했다. OpenAI 사용료가 GPT-3.5에서는 1,000단어당 약 3원(8K)인데, GPT-4는 1,000단어당 약 39.9(8K)에서 159.6원(32K)으로 인상됐다. 이로 인해 사업초기 무료 서비스를 제공 중인 국내 서비스 기업들은 사용자를 확보하면 할수록 사용료로 인해 수익 악화를 우려할 수밖에 없는 상황에 직면해 이슈화되고 있다.

이러한 가운데, 우리나라의 초거대 생성 AI 기술 수준은 대학과 기업 모두 논문, 특허 보유 건수에서 절대적 열위로 미국, 중국 등에 크게 뒤처진 것으로 조사됐다. 2023년 클래리베이트 조사에 의하면, '18~'22년 국가별 생성 AI 상위 1% 특허 건수를 보면, 중국은 256건, 미국은 159건, 한국은 9건 등으로 조사됐다.

우리나라 AI 전반의 생태계 수준은 54개국 중 8위로 기록돼 상위권 수준이라고 볼 수 있으나, AI 기술 활용 전문인력 부문인 글로벌 AI인덱스에 있어서는 미국 100점, 영국 31.8점, 독일 23.4점에 비해 평균 14.5점에 못 미치는 11.4점으로 미국의 약 1/10 수준이다.

[그림 3] 생성 AI 상위 1% 논문 건수



출처: 국가과학기술심의회, 2023

국내 초거대 생성 AI 기업들은 다양한 방법으로 핵심인재의 확보에 나서고 있으나, 글로벌 기업에 비해 부족한 수준이다. 국내 기업 중 가장 많은 인재를 확보한 네이버도 구글·MS에 비해 절반에도 못 미치는 상황으로 관련 인재의 확보가 시급한 상황이라고 한다.

국내 초거대 생성 AI 기업들은 기술개발의 가장 큰 애로사항으로 '인재의 부족'을 지적하면서, 특히 생성 AI의 등장으로 정부의 적극적인 생성 AI 인재의 양성지원과 차별화된 교육을 통한 수준별 인재양성, 다양한 방법의 인재 발굴 기회를 요구하고 있다.

맺음말

2023년 5월, 미국에서는 미국의 경쟁력과 국가안보에 필수적인 국가정책으로서 CET (Critical Emerging Technology)의 8가지 표준전략을 제시했다. 이중에서 3번째가 「Artificial Intelligence and Machine Learning」이다. 산업 전반에 걸쳐 혁신적인 기술과 과학적 돌파구를 꼭 이루겠지만 신뢰할 수 있고 위험 관리된 방식으로 개발돼야 함을 강조하는 전략이다.

미국은 그동안 인공지능에 있어서는 민간의 개발에 많은 비중을 두었으나 국가 표준전략으로 제시하는 것은 이번이 처음이다.

인공지능과 같이 CET 분야의 기술들은 하루가 다르게 매우 가파른 속도로 진화하고 있고 규모가 거대하고 융복합화 되고 있다. 때문에 정부가 기술개발이나 인력양성 등 각종 계획을

[그림 4] 2023년 백악관이 발표한 「핵심·신흥기술을 위한 미국 국가 표준전략」



출처: United states government national standards strategy for critical and emergung technology, 2023

세워 중장기로 추진해 나가는 대응책으로는 개발내용이 너무 느려서, 도저히 그 발전 속도를 따라갈 수 없는 수준에까지 이르렀다.

따라서 CET 분야의 기술들은 기업이나 문제점을 먼저 인식한 기관들에서 먼저 개발하고 정부가 필요한 부분을 신속히 지원해 나가는 형태로 추진돼야 할 것이다. 물론 선제적으로 정부가 먼저 추진할 수도 있으나 기술 시장 흐름을 읽어야 하는 타이밍과 관련된 것이어서 이 또한 만만치 않은 일이 될 것이다.

초거대 생성 AI는 지금까지 언급하였듯이 규모와 생성되는 정보로 인해 우리사회 많은 분야에 영향을 미치게 될 것이다. 특히 기술적으로 해결해야 할 이슈뿐 아니라 국가 수준을 높이기 위해 수준 높은 연구지원 및 인재양성이라는 큰 숙제가 앞으로 국가적 차원에서 추진해야 할 이슈이다.

초거대 생성형 AI 분야에서는 우리나라가 아직은 논문, 특허 및 기술력 부분에서는 세계 TOP 선진국 대열에 들어섰다고 말하기에는 다소 부족한 면이 있다. 그러나 과거의 ICT 기술과 같이 정부의 지원을 먼저 기대하기에는 세상이 너무 빨리 변하고 있다. 따라서 정부의 지원도 중요하지만, 먼저 민간의 적극적인 기술개발에 노력과 적극적인 참여를 통해 글로벌 기업으로 도약해야 할 것이다. 민간의 관련 연구나 논문, 특허 등도 마찬가지이다. 앞으로 이러한 급변하는 기술분야에 있어서는 R&D 추진체계가 선투자 후지원 형태로 달라져야 할 것이므로 정부와 민간의 협력이 그 어느 때보다 중요한 시기이다.