

# AI산업 경쟁력 강화를 위한 생성형 AI 규제 방향



이성엽  
고려대학교  
기술경영전문대학원 교수



## 챗GPT의 등장

2016년 인공지능 알파고가 이세돌 9단과 겨룬 바둑에서 완승하면서 엄청난 반향을 일으킨 지 6년여 만인 지난해 11월, 대화형 인공지능인 챗GPT가 등장하면서 또 한 번 세계에 큰 충격을 주고 있다.

미국 스타트업 '오픈AI'가 출시한 챗GPT(Chat Generative Pre-trained Transformer)는 사람의 언어를 이해하고, 사람의 질문에 꼭 맞는 유용한 답을 하는 대화형·생성형 AI로, 일반 대중에게 AI의 일상화라는 경험을 제공하고 있다. 2007년 손 안의 컴퓨터 시대를 연 아이폰, 2016년 인간 바둑기사를 이긴 알파고의 충격 이후 6년 만에 챗GPT가 새로운 인공지능 시대를 열고 있다.

대화형·생성형 AI 모델인 챗GPT의 등장과 함께 인류는 그동안 영화에서나 보던 로봇, 인공지능의 위험성이 현실화할 것으로 보고, AI에 대한 규제 논의를 본격화하고 있다. 유럽 연합(EU), 미국, 한국 등 주요 국가가 생성형 AI에 대한 법적 규제 논의를 시작하고 있다. EU가 가장 먼저 입법화를 진행 중인 가운데 미국, 한국도 입법안을 마련하고 국회에서 논의를 진행 중이다.

그러나 법적 규제의 본격 도입 이전에 생성형 AI의 특성을 이해하고, AI산업의 경쟁력 강화 차원을 고려한 규제 방향을 설정하는 과정이 중요하다.

## AI 특성과 규제 현황

생성형 AI인 대표 격인 챗GPT는 문장을 생성(generative)하는 AI 모델로 사전학습(pre-trained)을 통해 훈련받았으며, 트랜스포머(transformer)라는 구조에 기반한 인공지능 언어모델이다. GPT는 ‘사전 훈련된 생성 변환기(generative pre-trained transformer)’를 뜻한다. 결론적으로 생성형 AI란 대규모 데이터의 패턴을 자기지도 학습하여 명령어(prompt)에 따라 새로운 텍스트, 코드, 이미지, 음악, 영상 등의 콘텐츠를 생성하는 인공지능으로 정의할 수 있다. 패턴을 인식하고 예측하도록 설계한 기존 AI 시스템과 달리 생성형 AI는 새로운 콘텐츠를 생성한다.

생성형 AI 모델의 특성은 다음 몇 가지로 볼 수 있다. 첫째, 대규모 언어모델(large language model)이라는 점이다. 우선 언어모델(LM)은 문장 생성을 위해 다음에 올 단어의 확률을 할당하는 모델이다. 대규모 언어모델은 사람들이 사용하는 언어를 학습해 실제 인간과 유사한 문장을 생성하기 위한 언어모델로, 점차 규모가 커지며 초거대 AI로 진화하고 있다. 둘째, 생성형 AI 모델은 콘텐츠를 창조한다. 기존 AI는 이용자의 특정 요구에 따라 결과를 생성해 낸다. 즉, AI의 역할은 데이터 분석, 활용 등 인간의 행위를 대체하거나 보완하는 데 그친다. 생성형 AI는 자가학습 알고리즘으로 새로운 디지털 이미지, 영상, 음성, 텍스트 등을 ‘창조’한다. 셋째, 범용 인공지능(Artificial General Intelligence: AGI)이라는 점이다. 이는 인간처럼 종합적으로 사고·학습·추론하는 인공지능으로, 텍스트 이해·생성, 자연어 처리, 이미지 분류, 예측, 추론 등 다양한 과제(태스크)를 동시에 처리하는 인간의 능력을 모방한다. 기존의 AI는 대화, 질의응답, 기계번역 등 태스크별 개별 모델로 개발됐으나, 생성형 AI 모델은 인간이 하는 다중 작업(멀티태스킹)처럼 하나의 모델로 여러 태스크를 동시 처리한다. 넷째, 기존 검색 모델과의 차별성이다. 검색 모델이 키워드 검색을 통해 정보를 제공하지만, 생성형 AI 모델은 사용자 친화적인 상호작용을 통해 질문을 이해하고 대답하는 방식이다. 생성형 AI의 부상으로 검색 시장 핵심 가치가 검색 효율성에서 생성한 답변의 신뢰성으로 변하고 있다.

다음은 AI 규제를 살펴보자. 가장 앞서가는 곳은 EU이다. 2021년 4월 EU 집행위원회가 AI 법안의 초안을 발표했고, 2022년 12월에는 각료이사회가 수정안을, 지난 5월에는 의회가 수정안을 발표했다. 수정안에 GPT로 대표되는 파운데이션 모델(foundation model) 규제를 추가했다. 수정안에 따르면, 파운데이션 모델은 대규모 데이터에 의해 학습되고, 출력의 범용성을 고려하여 설계됐으며, 다양한 작업에 적용할 수 있는 AI 모델로 정의한다. 파운데이션 모델 공급자는 모델의 위험관리, 신뢰성관리를 위한 검증, 모델을 공급받아 활용하는 하위 사업자의 서비스 내용까지 고려한 모델 기술문서와 지침의 작성, 제공 등의 의무가 부과된다. 특히, 생성형 AI 파운데이션 모델 공급자는 이용자에게 AI 시스템 활용 사실을 알려야 하고, 위법한 콘텐츠 생성을 방지할 수 있도록 모델을 설계·개

발해야 하며, 또한 저작권 보호를 받는 학습데이터를 이용할 경우, 이에 관한 정보를 문서화해 공개하도록 하고 있다. 미국 의회는 2022년 2월 ‘알고리즘 책임법안(Algorithmic Accountability Act of 2022)’을 발의했다. 자동화된 의사결정 시스템의 투명성과 책임성을 확보하기 위해 기업이 이용자와 규제기관에 ‘자동화된 의사결정 프로세스’가 어떻게 활용되는지를 인식할 수 있도록 관련 정보를 제공하도록 하고, 기업이 의사결정의 영향 평가를 지속적으로 수행하도록 하는 내용이다.

한국은 지난 2월 14일 국회 과학기술정보방송통신위원회 법안심사소위에서 ‘인공지능산업 육성 및 신뢰 기반 조성에 관한 법률안(대안)’을 통과시켰다. 법안은 인간의 생명과 안전 및 기본권 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용하는 부분을 ‘고위험 영역 인공지능’으로 정하고, 인공지능 사업자에게 이용자에 대한 고위험 영역 인공지능 사용 사실의 고지의무, 신뢰성 확보 조치, 인공지능 도출 최종 결과 등에 대한 설명의무를 부여하고 있다. 다만, 법안은 누구든지 인공지능 기술과 알고리즘의 연구·개발 및 인공지능 제품 또는 인공지능 서비스 출시 등과 관련한 행위를 할 수 있도록 보장함을 원칙으로 하여, 우선 허용, 사후 규제라는 네거티브 규제 원칙을 선언하고 있다.

## AI 규제의 방향

규제는 위험의 통제와 관리를 목표로 하고 있다는 점에서 생성형 AI가 가져올 위험을 살펴보면, 다음과 같다. 첫째, 챗GPT에 과도하게 의존하는 경우, 표절이나 대필 등 저작권 침해, 기술 발달에 따른 교육 격차, 학습능력 저하 문제가 발생할 가능성이 있다. 둘째, 전문직의 일자리를 대체할 가능성이 있다. 챗GPT로 인해 인간은 논문 작성, 프로그래밍, 번역, 교정 등 지식 노동을 위한 시간과 비용을 획기적으로 절감할 것이다. 셋째, 대규모 학습 데이터를 기반으로 답변하기 때문에 잘못된 정보나 편향된 콘텐츠를 생성할 수 있다. 이제 우리의 과제는 이런 우려를 해소하고 인간에게 유용한 AI가 되도록 하는 방안을 강구하는 일이다. 가장 중요한 부분은 기업 자율적으로 AI가 작동하도록 하는 규칙인 알고리즘을 투명하고, 공정하게 설계하도록 하는 일이다. 또한 AI 윤리 교육의 강화도 필요하다. AI 개발자는 물론, 제조자, 운전자, 이용자 모두 AI를 인간에게 유용하게 활용하려는 윤리의식을 가져야 한다.

다음, 최후 수단으로 경성규제(hard law), 즉 법적 구속력이 있는 규제를 고려할 수 있다. 다만, 경성규제를 도입하는 경우에도 생성형 AI에 대해서는 원칙중심 규제, 네거티브 규제, 국내 규제와 국제 규제의 조화, 위험 기반 규제 등을 고려해야 하며, 결론적으로 위험에 상응하는 규제 도입으로 생성형 AI 기술의 혁신을 저해하지 않는 범위에서 위험관리가 이뤄지는 비례적 규제 원칙을 중요하게 고려해야 한다.○