Review Paper
Informatization Policy Vol. 31, No.3 (2024), pp.003-033
https://doi.org/10.22693/NIAIP.2024.31.3.003

생성형 AI의 법적 문제와 규제 논의 동향

김 법 연*

요 약

본 논문은 생성형 AI와 관련하여 제기되는 법적 문제점과 쟁점들을 정리한 것이다. 그리고 이러한 쟁점을 해결하거나 대응하기 위하여 혹은 생성형 AI가 제기하는 위험성을 최소화하기 위하여 개별 국가 또는 국제기구 등은 어떠한 규제적 논의들을 하고 있는지에 대하여 살펴보았다. 생성형 AI로 인해 제기되는 개인의 기본권 침해 문제, 새로운 범죄의 등장과 통제가능성, 특정 시장의 독점화 문제와 환경 문제 등이 주로 논의되고 있고, 규제의 필요성과 방향성에 대해 약간의 차이는 있지만 대부분의 국가들이 유사한 시각을 지니고 있는 것으로 보인다. AI와 관련하여서는 등장 초기부터 현재 제기되고 있는 문제들이 지속적으로 논의되어 왔었다. 특정 쟁점들은 상대적으로 많은 논의가 이루어졌지만 국가별로 조금씩 차이가 존재하기도 하고 과거와 다른 현상에 대한 고려가 필요한 상황들도 생겨나고 있다. 개별 국가의 상황에 맞추어 규제와 정책 등을 세밀화하고 있는 것으로 보인다. 다양한 쟁점들이 빠르게 등장하고 변화하는 상황에서, AI의 위험성을 최소화하고 안전한 AI의 활용을 통해 AI로 인한 효용과 이익을 향유하기 위한 방안들을 모색하여야 할 것이다. 국제적인 동향을 지속적으로 파악 및 분석하면서 국내에 적합한 AI 관련 규제와 세부 정책 등을 정비하는 것이 필요할 것이다.

주제어 : 생성형 AI, 생성형 AI의 법적문제, AI 규제, AI 데이터 학습의 법적쟁점, 저작권 침해, 개인정보보호

Legal Issues and Regulatory Discussions in Generative AI

Kim, Beop-Yeon*

Abstract

This paper summarizes the legal problems and issues raised in relation to generative AI. In addition, we looked at what regulatory discussions individual countries or international organizations have in order to solve or respond to these issues or to minimize the risks posed by generative AI. Infringement of individual basic rights raised by generative AI, the emergence and control of new crimes, monopolization of specific markets and environmental issues are mainly discussed, and although there are some differences in the necessity and direction of regulation, most countries seem to have similar views. Regarding AI, the issues that are currently being raised have been discussed continuously from the beginning of its appearance. Although certain issues have been discussed relatively much, there are some differences between countries, and situations that require consideration of phenomena different from the past are emerging. It seems that regulations and policies are being refined according to the situation of individual countries. In a situation where various issues are rapidly emerging and changing, measures to minimize the risk of AI and to enjoy the utility and benefits of AI through the use of safe AI should be sought. It will be necessary to continuously identify and analyze international trends and reorganize AI-related regulations and detailed policies suitable for Korea.

Keywords : generative AI, legal issues of generative AI, AI regulation, legal issues of AI data learning, copyright infringement, privacy protection

Received Aug 12, 2024; Revised Sep 2, 2024; Accepted Sep 4, 2024

* Research Professor of School of Cybersecurity, Korea University(kby82@korea.ac.kr) https://orcid.org/0009-0007-2301-691X)

Ⅰ. 서론

인공지능(Artifitial Intelligence; AI) 기술이 급진적으로 발전함에 따라 AI의 위험의 확장에 대한 우려도 함께 커지고 있다. AI 기술이 가지고 있는 여러 위험요인들 예를 들어 프라이버시 침해, 편향과 차별의 발생, 작동오류로 인한 피해 발생 등에 대한 탁월한 통제방안이 채 마련되지 못한 상황에서 기술은 빠르고 다양한 방식으로 전개되고 있기 때문이다. 기술 사용범위는 물론 이용자의 수도 빠른 속도로 확대되고 있어 해결되지 못한 문제점들이 사회의 위협요소로 작용할 가능성이 농후하다. AI를 포함한 대부분의 신기술 관련 규범에 있어 어려움은 고질적인 문제가 해결되기도 전에 새로운 위험 발생이 예상된다는 점이다. AI기술이 빠른 속도로 발전하고 새로운 도전과제들이 나타남에 따라 주요 국가들은 AI를 둘러싼 질서를 정립하는 데 많은 관심을 쏟고 있다. AI의 위험성을 최소화하기 위한 수단으로써 규제입법을 마련하는 시도들을 실시하고 있다. AI가 야기하는 위험을 통제하면서 새로운 기술을 수용하기 위한 노력이기도 하다.

AI에 관한 새로운 질서체계 정립에 관한 논의는 특히 생성형 AI의 등장 이후 빠르게 진행되고 있다. 생성형 AI는 지금까지 꽤나 정교한 출력값과 엄청난 데이터의 양 처리로 그간 AI와과 관련된 문제와 위험요소들을 더욱 심화시키고 있다. 침해되는 개인의 권리의 내용과 대상 범위도 확대되고 있다. 따라서 새로운 규범 형성에 대한 필요성이 더욱 강조되고 있는 것이라 할 수 있겠다. 이에 본 논문에서는 생성형 AI가 제기하는 여러 법적 쟁점과 이슈들을 망라하여 살펴보고 각 국가들 또는 국제기구가 어떠한 방식으로 AI의 위험을 통제하고자 하는지 동향을 확인해 보았다. 이를 통해 생성형 AI에 대한 문제들 내지는 위험성을 통제하기 위한 적합한 통제방식은 무엇인지에 대하여 고민해 보고자 함이다.

본고에서 제시하고 있는 쟁점들은 기사, 논문, 정책보고서 등 다양한 종류와 분야의 문헌에서 제기하고 있는 이슈들을 망라한 후 구조화하여 종합적으로 정리하였다.[1] 구체적으로 법적 쟁점은 개인정보자기결정권, 사생활의 자유 등 개인의 기본권과 저작권 등 개인의 권리 침해가 심화되는 문제, 생성형 AI의 오용과 범죄 등과 관련한 이슈, 그리고 시장경쟁과 환경규제 관점에서의 쟁점과 문제점들을 중심으로 살펴보았다. 그리고 각 쟁점별로 제시되는 문제점들에 대하여 정리하고 그에 따라 수반되는 각 국가의 규제적 논의에 대한 주요내용과 방향성 등을 정리하는 방식으로 서술하였다.

Ⅱ. 생성형 AI의 개념과 특징

1. 생성형 AI의 개념

생성형 AI란 학습데이터에서 텍스트, 이미지 또는 음성 등 다양한 형태의 콘텐츠를 대규모로 생성할 수 있는 연산 기술이다(Stefan, et al., 2024; Jovanović & Campbell, 2022). 생성형 AI는 이용자의 특정 요구에 따라 결과를 능동적으로 생성해내는 인공지능 기술로, 특정 기능을 수행하도록 설계되었거나 예측이나 분류만을 수행하는 다른 AI모델과 구별되는 개념이다(Hacker, et al., 2023). 즉, 기존의 딥러닝 기반 인공지능기술이 단순히 기존 데이터를 기반으로 예측하거나 분류하는 정도였다면, 생성형 AI는 이용자가 요구한 질문이나 과제를 해결하기 위해 스스로 데이터를 찾아서 학습하고 이를 토대로 능동적으로 결과물을 제시하는 진화한 것이다(Yang & Yoon, 2023). 이는 주로 딥러닝, 자연어처리(Natural Language Processing; NLP), 컴퓨터 비전 등의 분야에서 주로 활용되는데 학습한 내

1) 대표적으로 포브스(Forbes)는 생성형 AI의 위험을 6가지로 정리하고 있는데, ① 출력물의 품질 문제, ② 만들어진 '사실'과 환각증상(Hallucination), ③ 저작권, 개인정보보호 등 법적 위험, ④ 악용에 대한 취약성, ⑤ 전문성 및 컴퓨터 비용, ⑥ 편향된 출력값 등이 이에 해당한다(Forbes, 2023). 송민호 외(2024)는 ChatGPT(챗지피티)와 관련한 대중의 우려를 살펴보고자 뉴스 등의 댓글의 주요 토픽을 분석한 결과 '법적 및 윤리적 고려 사항', '지적 재산권 및 기술', '기술 발전과 인류 미래, 정보처리에서 인공지능의 잠재력', 'AI에서의 감정 지능 및 윤리적 규제', '인간모방' 등이 도출되었다고 한다(Song & Lee, 2024).

용을 기반으로 완전히 새로운, 창작에 가까운 결과를 도출하는 것으로 볼 수 있다.

이러한 생성형 AI는 발전과 효용에 대한 기대가 크며, 많은 분야에서 활용될 것으로 전망되는 기술이다. 골드만삭스(Goldman Sachs) 경제학자들은 생성형 AI가 향후 10년 간 약 300만 개의 일자리를 대체하고 노동 생산성을 연간 1.5%가량, 전 세계 GDP를 7%(약 7조 달러) 증가시킬 수 있을 것이라고 한 바 있다(Goldman Sachs, 2023). 2023년 MIT에서 발표한 논문에 따르면 'ChatGPT'[2]를 사용하면 업무시간을 단축할 수 있다고 한다. MIT 연구진은 대학 교육을 받은 전문가들에게 글쓰기 작업을 할당하고, 절반은 ChatGPT를 활용하도록 하였다. 실험 결과 ChatGPT를 사용한 그룹이 사용하지 않은 그룹에 비해 평균 37%, 대략 10분 정도의 작업 시간을 단축한 것으로 나타났고, 내용상으로도 좋은 평가 등급을 받았다고 한다(Noy & Zhang, 2023).

2. 생성형 AI의 특성

1) 넓은 학습범위와 독창적 결과값의 생성

생성형 AI는 기계가 콘텐츠, 예술, 음악 등을 만들고 생성할 수 있도록 하는 인공지능의 하위집합이다. 이는 독창적 결과를 생성하기 위해 인간의 행동이나 사고과정, 창의성 등을 시뮬레이션할 수 있는 알고리즘을 사용하는데, 기계가 입력 매개변수와 이전에 학습한 패턴을 기반으로 새로운 콘텐츠나 데이터를 생성할 수 있다. 기존 데이터를 단순히 가공하거나 분석하는 것이 아닌 새롭고 독창적인 콘텐츠를 생성하는 접근방식을 취하는 것이다. 생성형 AI 모델은 패턴을 학습하고 훈련 데이터와 유사한 새로운 출력을 생성하기 위해 대규모 데이터 세트에서 훈련된다(Hong, 2023). 생성형 AI의 대표 서비스인 ChatGPT의 경우 자연어 생성 모델로 대화와 관련된 많은 텍스트 데이터를 학습해서 마치 사람처럼 대화할 수 있도록 훈련된 모델이다. 딥러닝 기술인 트랜스포머(Transformer) 모델은 자연어 처리에 효과적인데 문장 속 단어 등 순차 데이터 내의 관계를 추적해 맥락과 의미를 학습한다(NVIDIA, 2022).[3] 이는 진화를 거듭하는 수학적 기법을 응용해서 서로 떨어져 있는 데이터 요소의 의미가 '관계에 따라 변하는 부분'까지 감지해 낸다. 이러한 기술적 배경에 의해 ChatGPT는 사람처럼 대화하는 것이 가능한 것이다(Han, 2023).

2) 처리되는 데이터의 양적 거대함

생성형 AI의 핵심에는 파운데이션 모델(Foundation Model)이 있다. 파운데이션 모델은 기본적으로 딥러닝 모델인데 이전의 딥러닝 모델과는 다르게 매우 크고 다양한 형태의 데이터를 처리하고 여러 작업을 수행할 수 있다. 대규모 딥러닝 모델은 특정 유형의 콘텐츠를 생성하기 위해 사전 훈련되고 다양한 작업을 지원하기 위해 활용된다. 파운데이션 모델은 대용량 데이터세트를 기반으로 훈련되고, 온라인 상에서 공개되는 데이터뿐만 아니라 대규모 데이터베이스의 비공개 데이터를 포함할 수 있다(Hong, 2023).

---

2) ChatGPT는 OpenAI社가 제공하는 생성형 AI로 대형언어모델(Large Language Model, LLM) 인공지능서비스이다. 대형언어모델은 사람들이 사용하는 언어(자연어)를 학습하여 실제 인간과 유사한 문장을 생성하기 위한 인공지능 모델로 언어모델은 문장 생성을 위해 단어의 순서에 다음에 올 수 있는 확률을 할당(Assign)하는 방식으로 기존의 주어진 단어들을 기반으로 가장 자연스러운 단어의 배열을 찾아 다음 단어를 예측하여 문장을 생성하는 통계적 모델(Statistical Language Model, SLM)에서 인공신경망 방법으로 발전한 것이다. ChatGPT는 GPT-1에서 GPT-5까지의 모델과 학습방식의 변화를 통해 고도화된 것으로 주된 변화는 모델 크기가 변화하면서 발전중에 있다. GPT-1은 1억 1,700만 개의 매개변수를 갖고 있었으나 GPT-2는 15억 개, GPT-3는 1,750억 개를 갖고 있다. GPT-2는 GPT-1대비 12.8배 증가한 것이고, GPT-3는 GPT-2에 비해 117배 증가한 것이다. 매개변수가 많아질수록 AI가 학습하는 능력이 좋아진다. 2023년 4월 출시된 GPT-4는 매개변수와 데이터의 양은 공개되지 않았지만 이전 모델과 다르게 문자뿐만 아니라 이미지도 처리할 수 있는 멀티모달(multimodal) 기능이 추가되었고, 2024년 5월 공개된 GPT-4o는 텍스트와 이미지는 물론 음성까지 입력으로 받고 생성하며, 출시 예정인 GPT-5는 이전 모델에 대비하여 추론 기능을 추가하고 초지능(AGI)을 달성할 것으로 기대되고 있다(Ahn, et al., 2023; Park, 2023; Park&Lee, 2024; AiTIMES, 2023; AiTIMES, 2024).

3) 예를 들어 "점심식사로 무엇을 먹는 것이 좋을까?"라는 질문에 대한 답변으로 연필, 청바지 등이 아닌 김치찌개, 제육덮밥 등과 같은 단어가 나올 확률이 높다. 생성형 AI는 데이터세트에서 가장 일반적인 패턴이나 규칙, 인간의 질문에 확률적으로 가장 알맞은 대답을 출력한다(Kim, 2023).

앞서 언급한 트랜스포머 모델이 개발되기 전에는 라벨링된 대규모 데이터세트로 인공지능을 훈련시켜야 했고, 이러한 데이터세트는 구축하는데 많은 시간과 비용이 소요되었다. 트랜스포머 등장 이전에는 훈련을 함에 있어 대규모 데이터세트를 라벨링을 했어야 하기 때문이다. 그러나 트랜스포머는 요소들 사이의 패턴을 수학적으로 찾아내기 때문에 이 과정이 필요 없어진다. 이에 수조 개의 이미지와 페타바이트(Petabytes)급 텍스트 데이터를 인터넷 등에서 사용할 수 있게 됐다(Han, 2023).

그리고 파운데이션 모델의 훈련은 반복적 과정으로 상당한 컴퓨팅 자원이 필요하다. 모델의 훈련 과정 초반에는 특히 원하는 수준의 정확도 측정을 위해 훈련 알고리즘은 신경망의 가중치를 조정하게 되고 이러한 과정을 몇 백만 번 이상 해야할 수도 있다. 많은 비용과 기간이 소요되는 원인이 된다.

3) 불명확함과 결과값의 허위성

ChatGPT의 경우 대규모언어모델로 이 경우 확률적 과정으로 문장이 생성되는 특성이 있다. 이에 따라 생성정보의 진위가 보장되지 않으며 엉뚱한 문장을 생성하는데 이러한 현상을 환각증상이라고 한다(Yoo, et al., 2023). 인공지능의 정보처리 과정에서 발생하는 오류로 잘못된 답변이나 기이한 이미지를 생성하는 현상을 의미한다. 생성형 AI는 확률과 통계에 기반한 결과를 생성하기 때문인데, 질문의 내용에 대해 가장 적합하고 가까운 답변을 선택하고 조합하는 과정에서 발생하는 오류이다. ChatGPT의 경우 인터넷에 게시된 방대한 양의 텍스트를 분석하여 만들어지기 때문에 사실과 허구를 구분하는 것이 불가능하다. 따라서 질문에 답변을 잘 할 수는 있지만 문맥과 의미에 대해서는 참과 거짓을 판별해내지 못한다.

Ⅲ. 권리 침해의 다양화와 심화

생성형 AI를 포함하여 인공지능 기술이 등장하고 발전, 상용화됨에 따라 다양한 문제들이 발생하고 있다. 대표적으로는 개인정보자기결정권 내지는 프라이버시의 침해, 저작권 등 지식재산권의 침해 등을 꼽을 수 있다. 생성형 AI의 등장으로 이러한 권리 침해의 양상이 보다 심화되고 있는데, 이하에서는 이러한 문제에 대한 논의들을 살펴보고자 한다.

1. 개인정보 침해 문제

인공지능 모델은 데이터에 대한 의존성이 높다. 인공지능은 기존 소프트웨어와는 달리 학습데이터로 훈련을 하고 학습데이터의 양과 질에 따라 모델의 성능이 달라지기 때문이다. 개발단계에서 탁월한 성능을 보이는 인공지능 모델도 실제 활용 환경에서는 성능이 급격히 떨어지는 데이터쉬프트(Data Shift)문제가 빈번히 발생한다(MIT Technology Review, 2020). 따라서 데이터의 양과 질이 인공지능 모델 성능에 주요한 변수가 되는데, 이에 따라 개인정보와 관련한 이슈가 특히 문제된다. 개인정보의 수집과 활용은 맞춤형 서비스를 제공하기에 적합하기도 하지만 활용의 가치가 높기 때문에 인공지능기술을 구현하는 과정에서 개인정보를 활용하고자 하는 경향이 강해지고 있다. 이에 더해 모델의 성능을 높이기 위해서는 많은 양의 데이터가 필요하므로 그만큼의 개인정보가 인공지능 기술에 의해 처리되는 상황이 지속적으로 발생하고 있는 것이다. 당연히 개인정보와 프라이버시 침해 문제가 인공지능 기술의 대표적 역기능 사례가 되는 이유라 할 수 있다.

이탈리아는 2023년 3월 30일 OpenAI에 대해 개인정보 처리를 잠정적으로 금지시킨 바 있는데, 이탈리아 데이터보호청(The Italian Data Protection Authority)은 ChatGPT가 데이터 학습을 이유로 방대한 양의 개인정보를 수집, 저장하므로 유럽의 일반개인정보보호규정(General Data Protection Regulation, GDPR)을 위반 여지가 있다고 보고 조사를 실시하였었다.[4] 이탈리아 데이터보호청은 ChatGPT에 대한 접속차단을 하였고, OpenAI에 대해 개인정보보호 위반 가능성에 대한

조치를 취하도록 하였다. 2023년 4월 서비스는 다시 재개되었다. 캐나다도 2023년 5월 ChatGPT에 대한 조사를 시작한 바 있다. 동의없이 개인정보를 수집, 사용, 공개하였다는 문제제기에 대응하는 차원이었다. ChatGPT가 개인정보 수집 등에 대하여 정보주체의 명백한 동의를 확보하였는지, 개방성과 투명성, 접근성, 정확성, 책임 등의 의무를 준수하였는지, 합리적인 개인정보의 사용과 공개, 목적내 사용 등에 해당하는지 등을 조사하였다(Office of the Privacy Commissioner of Canada, 2023). 유럽연합 데이터보호위원회(European Data Protection Board, EDPB)은 2023년 4월 ChatGPT에 대하여 법집행 등에 관한 정보공유 등을 위해 전담 태스크포스를 출범하였었다(EDPB, 2023). 해당 태스크포스는 유럽의 개인정보보호 관련 규정을 준수하는지 검토하였고, 2024년 5월 '챗지피티 태스크포스가 수행한 작업 보고서(Report of the work undertaken by the ChatGPT Taskforce)'를 발간하였다(EDPB, 2024). 그러나 해당 보고서는 OpenAI의 개인정보 처리가 적법하고 공정한지 결론을 내리지는 못하였다.

1) 공개된 개인정보의 활용 문제

(1) 공개된 개인정보의 활용과 문제점

생성형 AI에서는 대형언어모델인 ChatGPT와 같은 경우 인터넷상에 공개된 텍스트를 스크래핑(Scraping)하는 방식으로 수집하는데(Chan, 2024), 이 때 공개된 개인정보가 학습에 활용될 경우 개인정보 침해 가능성이 존재한다. 정보주체가 개인정보를 게시하거나 공개할 때 인공지능의 학습을 전제로 하지 않았을 것이기 때문이다. 정보주체가 개인정보를 공개하는 것까지는 인지하거나 동의하였다 하더라도 자신의 개인정보가 스크래핑되고 나아가 생성형 AI 모델의 학습에 사용되는 것에 대해 동의하였다고 보기 어려우므로 침해 문제가 제기될 수밖에 없다.

공개된 개인정보의 경우 「개인정보보호법」상의 '개인정보'에 해당하는 것을 전제로 할 경우 ① 정보주체가 온라인상에 정보를 공개하는 것은 누군가에 의해 정보가 수집되고 활용되는 것을 인지하면서 게시하였기 때문에 정보주체의 동의가 있는 것으로 보아야 하므로 생성형 AI가 학습데이터로 스크래핑 하는 것이 법적으로 허용된다고 볼 수도 있고, ② 온라인상에 자신의 개인정보를 게시 또는 공개하였다고 하여 생성형 AI의 학습에 활용되는 것까지 동의하는 것으로 보기는 어렵다고 볼 수도 있다. 그리고 만약 공개된 개인정보 자체가 「개인정보보호법」에서 보호되는 '개인정보'에 해당하지 않는 경우에는 정보주체의 동의 여부를 판단할 필요 없이 스크래핑이 가능하다. 현재까지 생성형 AI가 학습과정에서 온라인상에 공개되어 있는 개인정보를 학습데이터로 활용하는 것에 대하여 어떤 관점에서 해석하는 것이 타당한지 결정하는 것이 어려운 문제이다. 정보주체의 권리를 강화하여 공개된 개인정보에 대해서도 동의를 확보하고 학습데이터로 활용하도록 하게 되면 생성형 AI 모델의 학습에 장애요인이 되고, 모델이나 서비스 개발·발전이 어렵게 된다. 반면 공개되었음을 이유로 자유롭게 생성형 AI 학습에 활용되도록 할 경우 정보주체의 권리를 약화시키는 결과가 되기 때문이다. 이러한 이유로 「개인정보보호법」도 명확하게 기준을 제시하지 못하고 있는 상황이다(Son, 2024).

대부분의 국가에서 공개된 개인정보에 대해서도 수집 등에 대한 동의가 요구되는 것으로 보고 있는 듯 하지만 실제로 동의를 받는 것이 가능할 것인지에 대해서도 의문은 남아있다. 유럽연합의 데이터보호위원회(European Data Protection Board, EDPB)는 OpenAI가 인터넷에 공개된 정보들에서 자동으로 데이

4) 이탈리아 데이터보호청은 GhatGPT의 GDPR 위반사항으로 개인정보가 처리되는 이용자와 정보주체에게 정보가 제공되지 않는다는 점(제5조 관련), 개인정보의 수집 및 처리에 있어 법적 근거가 없다는 점(제6조 관련), ChatGPT에 의해 제공된 정보가 실제 정보와 일치하지 않는 사례가 발생하였다는 점(제13조 관련), 미성년자 연령 확인 절차가 부재하였다는 점(제8조 관련) 등을 지적하였다(Garante Per La Protezione Dei Dati Personali, 2023).

터를 수집하지만 ChatGPT를 훈련시키는데 개인정보를 사용하지 않았다고 보기 어렵고, '웹 스크래핑'을 통해 많은 양의 데이터가 수집된다는 점을 고려할 때 각 정보주체에게 데이터 수집 등에 대한 상황을 알리는 것은 일반적으로 실용적이지 않거나 불가능하다고 한 바 있다(EDPB, 2024).

(2) 법적 대응 동향

공개된 개인정보에 대해 주요 국가에서는 법률적으로 개인정보와 동일하게 규정하거나 특례의 방식으로 다루기도 한다. 또는 법률상의 보호대상에서 배제하여 활용이 가능하도록 하기도 한다. EU의 개인정보보호법인 GDPR은 명시적으로 공개된 개인정보를 처리할 수 있는 근거를 따로 명시하고 있지 않아 공개된 개인정보를 일반 개인정보와 동일하게 판단하는 것으로 보인다. 따라서 GDPR이 적용되는 경우 공개된 개인정보는 적법한 법적 근거에 의해서만 처리될 수 있다(제6조).[5] 다만 최근 EU의 국가들 중에서도 일부 국가들은 웹 스크래핑을 통해 수집한 데이터를 생성형 AI의 학습에 사용하는 것이 '정당한 이익'이 있는 경우에는 처리를 할 수 있는 것으로 판단하고, 이에 대한 적정성을 확인하거나 '정당한 이익'이 있는 경우에 대해 기준을 마련하기도 한다. 구체적으로 영국과 프랑스는 AI의 학습 목적으로 개인정보를 처리할 때 '정당한 이익'이 인정되기 위한 기준을 제시하고 있다. 영국의 ICO(Information Commissioner's Office; ICO)는 AI 개발자들이 영국 개인정보보호법의 법률상 의무를 이행하고 입증할 수 있다면 공개된 개인정보를 스크래핑해서 생성형 AI를 학습시킬 수 있다고 하였다. 학습용 데이터 수집시 최대한 구체적 이익을 설정하여야 하고, 이때 처리자의 이익은 광범위한 사회적 이익을 포괄하며, 이러한 이익은 개발자가 AI의 구체적인 목적과 용도를 입증할 수 있어야 한다고 하고 있다(ICO, 2024). 프랑스도 처리자의 이익 추구가 정당하고, 그러한 이익을 달성하기 위해 개인정보의 처리가 불가피하며, 정보주체의 이익과 권리에 대해 불균형이 없을 경우 공개된 개인정보를 활용할 수 있다고 한다(CNIL, 2024).

싱가포르는 '개인이 공개한 정보'와 '대중에 공개된 정보'를 공개적으로 이용가능한(Publicly Available) 정보라고 규정하고, 동의없이 수집·이용·공개할 수 있도록 하고 있다(제2조 및 제17조). 미국 캘리포니아 소비자프라이버시법(California Consumer Privacy Act of 2018; CCPA)은 생체인식정보를 제외하고 공개된 정보는 개인정보의 개념에서 제외하고 있다(CCPA § 1798.140.(o)(2)). 중국도 중국 개인정보보호법은 공개된 개인정보에 대해 '개인이 스스로 공개하였거나 이미 적법하게 공개된 개인정보'로 명시하고 있다(제13조(6)). 따라서 싱가포르나 미국 캘리포니아, 중국의 경우 생성형 AI 학습을 목적으로 공개된 개인정보를 수집하거나 활용하는 것이 허용된다.

우리나라의 경우 공개된 개인정보와 관련한 문제가 생성형 AI 이슈 제기 이전에 대법원에서 다루어진 바 있는데, 해당 사안에서 대법원은 공개된 개인정보에 대해 동의 없이 영리목적으로 수집·제공하였을 경우의 위법성은 정보주체와 정보처리자의 이익을 비교 형량하여 판단하여야 한다고 보고 있다. 정보주체의 동의가 필요한지에 대해서는 이미 공개한 개인정보는 공개 당시 정보주체가 자신의 개인정보에 대한 수집이나 제3자 제공 등의 처리에 대하여 일정한 범위내에서 동의를 한 것이라고 보고 있다. 다만 이 경우에도 정보주체의 동의가 있었다고 인정되는 범위는 개인정보의 성격, 공개의 형태와 대상 범위, 그로부터 추단되는 공개 의도 내지는 목적, 정보처리자의 정보제공 등 처리 형태, 정보제공으로 공개의 대상 범위가 원래의 것과 달라졌는지, 정보제공이 정보주체의 원래의 공개 목적의 상당한 관련성 등을 검토하여 판단하여야 한다고 하였다(대법원 2016.8.17. 선고 2014다235080). 이후 개인정보보호위원회에서는 공개된 개인정보를 이용하기 위해

5) 다만 민감정보의 경우에는 정보주체가 명백하게 공개한 개인정보의 경우에는 처리할 수 있도록 예외적으로 허용하고 있다(제9조 제2항 제e호).

객관적 동의 의사를 추단할 수 있는 범위 내 또는 정당한 이익이 정보주체 권리보다 명백히 우선하는 범위 내에서 동의없이 수집할 수 있다고 하였다(개인정보보호위원회, 2023). 그리고 구체적인 공개된 개인정보를 수집·이용할 수 있는 법 해석 기준과 적법한 이용이 되기 위한 처리 방법 등에 대하여는 안내서를 통해 제시하고 있다(개인정보보호위원회, 2024). 그러나 이러한 정부부처의 안내서 등은 법률상 구속력을 갖는 것이 아니기도 하고, 개별 사안에 따라 판단이 달라질 수 있기 때문에 공개된 개인정보 스크래핑에 대한 적법성이나 명확한 법적 기준 필요성에 대한 논의는 당분간 이어질 것으로 예상된다.[6)]

2) 생성형 AI에 의한 정보주체의 권리 침해 심화와 보장의 어려움

(1) 정보주체의 권리 행사 및 보장의 어려움

생성형 AI의 경우 비지도학습 내지는 자기지도학습을 중심의 학습방식을 취하고 있고 예측하기 어려운 방식으로 데이터가 처리되는 사례가 증가하면서 불특정 다수의 개인정보 침해에 대한 우려는 점차 커지고 있다. 뿐만 아니라 AI 서비스 과정에서 개인정보가 노출되거나 재식별되는 가능성도 존재하고, 민감정보가 생성되고 처리될 가능성도 높아지고 있다. 대규모 언어모델의 경우 언어 패턴을 바탕으로 예측하여 답변을 생성하기 때문에 환각증상의 발생가능성이 높고, 데이터의 오류나 왜곡, 거짓정보 등이 개인의 정체성을 위협할 우려도 있다(Personal Information Protection Commission, 2023).

그리고 정보주체의 권리 보장에 대한 문제도 발생하고 있다. 즉, 정보주체가 보유하고 있는 수집·동의권을 비롯한 열람권, 정정권, 삭제권 등에 대한 권리 행사가 기술의 복잡성 등의 특성으로 인해 권리 행사에 대한 처리자의 이행이 현실적으로 어려운 사례가 등장하고 있다. 일례로 오스트리아 개인정보보호 관련 비영리단체인 NOYB(None of Your Business; NOYB)는 OpenAI에 대해 GDPR을 위반하였다는 문제를 제기한 바 있는데, ChatGPT의 사용자가 생일에 대한 질문을 하였고, ChatGPT는 정보가 없다는 답변을 하는 것이 아닌 잘못된 정보를 반복적으로 제공하였기 때문이다. NOYB는 해당 사용자가 OpenAI에 대해 데이터를 수정하거나 삭제해달라는 요청을 하였지만 OpenAI는 이러한 요청을 거부하고 데이터를 수정할 수 없고 처리된 데이터의 출처나 수신자에 대한 정보를 공개하지 않았다고 하면서 오스트리아 데이터보호당국에 조사를 요청하였다(Reuters, 2024).

생성형 AI가 학습하는 과정에서 개인정보를 학습하고 나아가 답변을 통해 개인정보를 생성하거나 노출시킬 경우 정보주체는 개인정보의 처리 가능성을 인식할 수 없다는 측면에서 권리 침해에 대한 우려가 더 큰 상황이라 할 수 있다(One Trust Data Guidance, 2024).

(2) 생성형 AI에 의한 개인정보의 노출, 추론 등으로 인한 권리 침해 문제

생성형 AI에서 비의도적으로 개인정보를 포함하는 콘텐츠를 생산하는 상황에 대한 문제도 제기되고 있다. 이는 인공지능이 생성한 데이터가 개인정보보호법에 따라 어떻게 취급되어야 하는지 보호 필요성과 보호에 대한 책임을 누가 져야 하는지 등과 같은 의문이 드는 것이다(Abenezer Golda, et al., 2024). AI 학습과정에서 공개된 정보 분석을 통해 공개되지 않는 정보(민감정보를 포함하여)를 추론하거나, 비식별처리된 개인정보를 재식별하여 노출시킬 가능성도 있고, 암기된 개인정보를 노출시키거나 출력할 수도 있다.[7)] [8)]

이 외에도 법률 적용상의 문제들도 제기된다. 예를 들어 생성형 AI의 경우 개인정보가 노출될 경우 이를 법률상의 '유출'로 해석하는 것이 타당한지, 또는 '공개'라고 해석하는 것이 타당한지에 대한 논의와 같은 것들이

6) AI가 공개된 데이터를 학습할 수 있도록 입법화를 검토할 필요가 있다는 견해도 존재한다(이성엽, 2023).

다. 사업자가 보유한 개인정보가 포함되어 있는 데이터베이스에 기초한 모델의 경우 데이터베이스에 저장된 개인정보가 검색되어 그대로 공개되는 방식이지만, 생성형 AI의 경우 AI의 자율성에 기초하여 생성되는 것이기 때문이다. 「개인정보보호법」상 '처리'의 개념에는 개인정보의 '생성'도 포함되므로 이러한 생성 행위에 대해 법률을 적용하는 것이 타당하다고 해석되지만 개인정보와 유사한 산출물을 AI가 자율적으로 생성하거나, 학습대상의 데이터를 사업자가 직접 제공한 것이 아닐 경우 개인정보의 노출이라 하더라도 이를 공개나 유출로 해석하기는 어렵다고 보기도 한다(Jeon, 2023).

(3) 생성형 AI에서 정보주체 권리 보장을 위한 방안 논의

2023년 5월 도쿄에서 '제3회 G7 데이터 보호 및 개인정보보호 당국 라운드테이블(The 3rd G7 Data Protection and Privacy Authorities Roundtable)'이 개최된 바 있다. 해당 회의에서 G7 각국은 생성형 AI가 사용자의 개인정보를 고려하여 설계 및 운영될 것을 요청하는 성명을 발표하였다(朝日新聞デジタル, 2023).[9]

성명서는 생성형 AI가 적절하게 개발되고 규제되지 않으면 프라이버시 및 데이터 보호, 기타 기본적 인권에 대해 리스크와 잠재적 손해를 초래할 가능성이 있다는 우려가 높아진다는 것을 인식한다고 하면서, 생성형 AI 도구와 관련하여 프라이버시 및 데이터 침해 위험이 발생할 수 있는 주요 분야에 대해 주의를 촉구한다고 하였다. 구체적으로는 ① 생성형 AI 모델의 훈련, 검증, 테스트에 사용되는 데이터세트, 개인에 의한 생성형 AI 도구와의 대화, 생성형 AI 도구를 통해 생성된 콘텐츠와 관련한 개인정보의 처리에 대해 주의가 필요함을 강조하였고, ② 생성형 AI 모델 훈련에 이용된 데이터세트 내에서 개인정보를 추출해 내거나, 데이터 보호를 위하여 설계된 조치의 효과를 상실시키려는 위협이나 공격으로부터 데이터를 보호하기 위한 안전한 보호조치가 필요하다고 하였다. 또한 ③ 생성형 AI 도구를 통해 생성된 개인정보가 정확성·무결성·최신성과 차별적·위법적·기타 부정적 영향을 받지 않도록 모니터링 조치 등을 실시하여야 하며, ④ 생성형 AI 도구 운영에 있어 공개성과 설명가능성을 촉진하는 투명성에 관한 조치를 취하고, 특히 해당 도구가 개인에 관한 의사결정이나 그러한 의사결정을 지원하는데 이용되는 경우의 투명성 조치를 실시할 것을 강조하고 있다. 이외에도 ⑤ 생성형 AI 시스템의 영향을 받는 개인이나 시스템과 대화하는 개인들이 생성형 AI 도구와 관련하여 자신의 개인정보에 대하여 접근하고, 부정확한 개인정보를 수정할 수 있어야 하며, 자신의 개인정보를 삭제할 수 있으며, 중대한 영향을 미치는 자동화된 결정을 따르는 것을 거부할 수 있도록 보장하여야 한다고 한다. ⑥ AI 공급망에서 주체간 적절한 수준의 책임을 확보하는 설명 책임 조치 등이 필요하다고 하고 있다.[10]

한편, 2023년 영국 ICO가 서명한 '데이터 스크래핑 및 개인정보보호에 관한 공동서명서(Joint statement on data scraping and the protection of privacy)'는 공개적으로 접근할 수 있는 개인정보는 개인정보보호법의 적용을 받는 것으로 소셜미디어 회사 등 공개적으로 접근이 가능한 개인정보를 호스팅하는 웹사이트

7) 최근에는 프롬프트 주입공격을 통해 사용자가 개인정보를 추출하는 것이 가능할 수 있다는 연구결과가 제시되기도 하였다. ChatGPT에서 단순 프롬프트 주입 공격으로 '탈옥'을 유도하고, 개인정보를 비롯한 훈련 데이터를 추출해 낸 사례가 있다. 구글 연구진이 훈련 데이터를 출력하지 않도록 미세조정한 모델이 사전 훈련 데이터를 밝히도록 하는 탈옥방법을 발견, 특정 단어를 무한반복하도록 요청하면 해당 단어를 반복하다 어느 순간 훈련 데이터의 원본 텍스트가 포함된 답변을 내놓았다고 한다. 추출된 훈련데이터에는 학술 논문, 문학 작품 등을 포함해 사람의 이름, 이메일주소, 전화번호 등 개인정보가 포함되었고 답변중 총 16.9%가 개인정보이고, 그 중 85.8%가 실제 정보라고 하였다(AiTIMMES, 2023).

8) GPT-3.5 터보에서 뉴욕타임즈 직원 30명 이상의 개인 이메일 및 비즈니스 이메일 주소를 생성시킨 연구결과도 있다. 이메일 주소는 비공개 정보라고 할 수는 없으나 이러한 개인정보의 추출은 민감한 정보를 공개할 수도 있다는 잠재력을 보여주는 것이라는 점에서 의미가 있다(NewYorkTimes, 2023).

9) 성명서 원문은 EDPS, "G7 DPA Roundtable", 2023.6.21., Available at: https://www.edps.europa.eu/data-protection/our-work/publications/international-conferences/2023-06-21-g7-dpa-roundtable_en, Accessed: 2024.8.3.을 참고.

10) G7 Data Protection and Privacy Authorities Roundtable, "Roundtable of G7 Data Protection and Privacy Authorities Statement on Generative AI", 2023.6.21.

운영자는 데이터 및 개인정보보호법에 따라 불법적 데이터 스크래핑으로부터 플랫폼 이용자의 개인정보를 보호해야 할 의무가 있다고 하고 있다. 데이터 스크래핑과 관련된 잠재적 개인정보 피해를 적절하게 보호할 수 있는 안전장치가 없으므로, 위험을 완화하기 위해 웹사이트 운영자가 다양한 기술적·절차적 통제를 구현해야 한다고 강조한다(ICO, 2023).

2. 저작권 관련 문제

1) 타인 저작물에 대한 저작권 침해

생성형 AI에서는 대형언어모델인 ChatGPT와 같은 경우 인터넷상에 공개된 텍스트를 스크래핑할 경우 개인정보뿐만 아니라 저작물도 대량으로 스크랩하여 학습하므로 저작권 침해 문제가 발생한다(Kim, 2022). 저작권자와 소유권자들이 허가 없이 학습 데이터세트의 일부로 자신의 작품을 사용하는 것은 저작권 침해라고 주장할 수 있기 때문이다.[11] 반면, AI의 학습에는 타인의 저작물에 대한 학습이 불가피하거나 이를 의도적으로 제한하기도 어려운 상황이다. 2024년 2월 발간된 영국 의회의 커뮤니케이션디지털위원회(Communications and Digital Committee)의 보고서에는 '대규모 언어모델과 생성형 AI'에 대하여 OpenAI가 제출한 증거자료가 포함되어 있는데 해당 자료에서 OpenAI는 '콘텐츠 제작자와 소유자의 권리를 존중한다'라고 하면서도 '저작권이 있는 자료를 사용하지 않고 선도적인 AI모델을 훈련시키는 것은 불가능하다'라고 하고 있다(UK Parliament, 2024). 생성형 AI가 학습과정에서 저작물을 이용하였으므로 저작권자에게 정당한 보상을 해야 한다 하더라도 대부분의 국가에서 저작권을 별도로 등록하는 것을 의무화하고 있지 않기 때문에 저작권이 있는 작품인지, 저작권자가 누구인지를 식별하는 것도 어려운 문제이다. 이에 더하여 해당 저작물의 이용이 생성형 AI 모델을 학습하는데 얼마나 기여했는지를 평가하고, 저작권자에게 보상을 하는 것이 필요한지 등을 결정하는 것과 함께, 저작권자에게 보상을 실질적으로 할 수 있는 방법적인 부분도 현재는 구체화 된 바가 없다.

2) 생성형 AI가 생성한 저작물에 대한 저작권 인정 문제

생성형 AI가 저작권을 침해하는 행위뿐만 아니라 생성형 AI가 생산한 저작물에 대한 권리 부여 문제도 쟁점이 되고 있다. ChatGPT는 현재 이미지나 영상 같은 데이터의 형식을 문장과 결합하여 자연어를 이용하여 이미지에 대한 질문을 할 수도 있고, 'DALL-E'는 텍스트로 입력을 하면 그림을 생성한다. 딥마인드(DeepMind)가 공개한 시각적 언어모델인 플라밍고(Flamingo)는 자연어를 이용하여 이미지에 대한 질문에 답을 할 수가 있다. 이와 같이 생성형 AI가 새로운 창작물과 콘텐츠를 만들어내기 때문에 해당 저작물의 지식재산권을 누구에게 부여해야 하는가를 결정하는 것이 쟁점이 된다(Novelli, et al., 2024). 일반적으로 생성형 AI가 생성한 결과물을 얻기 위해 이용자가 질문을 구성한 경우라면 이용자를 저작자로 볼 수도 있을 것이나 생성형 AI의 경우 저작자가 아니기 때문에 해당 결과물은 저작권을 어떻게 성립시킬 것인지를 결정하기 어렵기 때문이다(Kim, 2023). 이에 대해서는 AI가 만들어낸 산출물을 저작물로 인정할 수 없다고 보기도 하고,[12] AI가 생성한 문장과 이미지에서도 '창작적 기여가 있으면 저작물성이 인정된다'고 하기도 한다(Son, 2023).[13] 그러나

11) 실제로 게티이미지라는 이미지 판매서비스 기업은 이미지 생성AI 도구를 개발한 회사인 스태빌리티에 대해 저작권 위반혐의로 소송을 제기한 바 있다. 게티이미지는 1천200만 장 이상의 이미지를 승인없이 AI모델을 훈련시킨 것에 대해 문제를 제기한 것이다(Reuters, 2023; Bloomberg Law, 2023).

12) 2022년 한국음악저작권협회는 AI의 산출물을 저작물로 볼 수 없다고 하여 AI프로그램이 작곡한 노래에 대해 저작권료 지급을 중단한 바 있으며(SBS 뉴스, 2022), 2023년 미국 법원은 생성형 AI 개발자가 인간의 개입 없이 AI가 만든 산출물의 저작권 등록을 거부한 미국 저작권청 결정에 대한 불복 소송에서 저작권을 인정할 수 없다는 판결을 한 바 있다(Thaler v. Perlmutter, Case 1:22-cv-01564-BAH (D.D.C., Aug. 18, 2023)).

13) AI가 생성한 결과물에 대해 인간에게 부여하는 저작권과 동일한 내용으로 권리를 부여하여야 한다는 견해는 없으나 유사한 권리를 부여하여야 한다는 논의는 지속적으로 제기되고 있다. 대표적으로 Kim, 2016., Son, 2016., Chung, 2019., Jeong, 2023. 등이 있다.

후자의 경우에도 인간에 의한 창작적 기여[14]가 있어야 하며, 이를 판단하는 것이 저작권 보호대상인 저작물이 되는지를 결정하는데 중요한 기준이 된다.[15] 생성형 AI에 의한 산출물이 저작권으로 보호받는 저작물이 될 경우 이용자는 권리 침해 문제를 고려하여 이용하여야 한다. 유럽집행위원회(European Commission, EC) 산하 IP Helpdesk의 'ChatGPT의 지식재산권 쟁점 분석 보고서'에서는 ChatGPT에 의해 생성된 콘텐츠를 AI가 소유하는 것은 불가능하지만 저작권으로 보호되는 저작물이라고 하고 있다. 따라서 ChatGPT에 의해 생성된 콘텐츠는 개인적 대화 보조용으로 사용할 경우 법적 문제가 없으나, ChatGPT에 의해 생성된 콘텐츠를 재사용할 경우 타인의 저작권 침해가능성이 높다고 한다(European Commission IP Helpdesk, 2023). 생성형 AI가 타인의 저작물을 학습하는 과정에서 타인의 권리를 침해하는 경우 이러한 결과물을 활용하는 이용자 또한 저작권을 침해하게 될 수도 있어 저작권 관련 문제는 여러 관점의 쟁점이 얽혀 있는 상황이라 할 수 있다.

3) 법정책적 대응 동향

AI에 의한 데이터 스크래핑 문제는 AI 등장 초기부터 논의가 이루어져 왔다. 이에 입법으로 문제를 대응해 온 국가들이 존재하는데, 생성형 AI 등장으로 기존의 논의와 제도가 변화를 요구받고 있는 상황이다. 데이터 스크래핑은 텍스트·데이터 마이닝(Text and Data Mining, TDM)이라고 하는데, TDM의 분석 대상으로 삼는 대상인 정보에 저작물이 포함됨에 따라 논의가 시작되었다. 해외 주요국들은 2000년대 초반부터 TDM과 관련한 면책규정을 도입하는 것에 대해 정책적 논의를 시작하였고, 우리나라에서도 2020년 즈음 TDM 면책규정을 담은 「저작권법 개정안」이 제안된 바 있다.[16] 우리나라는 이를 입법화하지 못하였지만, 일본과 유럽연합의 경우 현재 법률로 TDM을 제한적으로 허용하고 있다.[17] [18]

이와 같이 AI에 의한 저작권 침해 관련하여서는 이미 그 문제점이 다양하게 제기된 바 있기는 하나, 생성형 AI의 등장 이후 AI와 저작권에 대한 논의가 보다 구체적이고 빠른 속도로 확산되고 있다. 주요 국가들은 AI의 학습을 위한 저작물 이용에 대한 기준과 규칙을 마련하려는 시도와 함께, 생성형 AI에 의하여 만들어진 결과물에 대한 법적 통제를 추진하고자 하는 상황이다.

(1) 생성형 AI가 생성한 콘텐츠의 저작물성 인정과 관련한 해석

우선 생성형 AI의 산출물의 저작권 인정 여부에 대해서 기본적으로는 인간에 의한 것이 아닌 AI 기술에 의한 창작물에 대해 저작권을 인정하지 않는 경향으로 보인다. 미국 의회 조사국이 작성한 '생성형 인공지능과 저작권법(Generative Artificial Intelligence and

14) 창작적 기여의 내용과 정도를 결정하는 것이 결국 생성형 AI 산출물의 저작물성을 결정하는 기준이 될 것으로 보이는데 창작적 기여가 무엇인지에 대해 예를 들어 보면, 이용자가 프롬프트를 통해 콘텐츠의 방향이나 내용을 본인이 의욕한대로 도출하도록 한 경우라 할 수 있다. 이에 대해서는 이용자의 프롬프트의 내용에 따라 결과물의 질이 달라지기 때문에 프롬프트를 저작권의 보호대상으로 보아야 한다는 견해도 존재한다(Kim, 2023).

15) 일본 정부는 2017년 AI가 생성한 문장과 이미지에서도 인간의 '창작적 기여가 있으면 저작물성이 인정된다'고 한 바 있다. AI가 도구로 사용되어 완성된 콘텐츠에 인간의 창조성이 있다고 판단할 수 있으면 저작물로 볼 수 있다는 것이다(知的財産戦略本部, 2017).

16) 대표적으로 '저작권법 전부개정법률안', 의안번호: 210740, 제안자: 도종환의원 등 13인, 제안일: 2021.1.15.이 있다.

17) EU의 경우 'EU 디지털 단일시장 저작권 지침(Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC)' 제3조는 회원국으로 하여금 학술적 연구 수행, 교육서비스 제공기관 등이 과학적 연구 목적으로 TDM을 허용하도록 하고 있고, 제4조는 회원국으로 하여금 합법적 TDM을 허용하도록 규정하고 있다. 영국도 저작권법 제29A조에서 이용자가 저작물에 합법적으로 접근하여 이용하면서 그 출처를 표시한 경우, 비상업적 연구 목적으로 복제물을 작성하는 행위는 저작권 침해에 해당하지 않는다고 하고 있다. 일본 저작권법도 제30조의4에서 TDM을 허용하고 있는데 EU와는 다르게 영리·비영리 목적이나 주체에 제한없이 저작물의 표현을 향수하지 않는 이용의 경우에는 자유롭게 이용할 수 있도록 하고 있다. 다만 저작권자의 이익을 부당하게 침해하는 경우에는 이용이 허용되지 않는다. 2023년 일본 문화청은 'AI와 저작권'이라는 세미나 발표에서 AI개발 및 학습과 같은 정보 분석 등 행위는 제30조의4에 따라서 원칙적으로 이용이 가능하나 '필요하다고 인정되는 한도'를 넘거나 '저작권자의 이익을 부당하게 침해하는 경우'는 제외된다고 한 바 있다(文化庁, 2023).

18) 이에 대해서 자세한 사항은 Ryu(2023)을 참고.

Copyright Law)'이라는 검토보고서는 생성형 AI 프로그램이 만들어낸 결과물이 저작권 보호의 대상이 되는지 여부에 대하여 경향성을 조사하고 발표하였는데, 생성형 AI가 만든 결과물에 대하여 저작권 등록의 시도가 있었으나 저작권청과 법원은 오직 인간에 의해 창작된 저작물에 대해서만 저작권을 인정하는 입장으로 보인다고 밝히고 있다. 그리고 현재까지 AI가 생성한 결과물에 대한 저작권을 인정한 판례나 저작권청의 결정이 없다는 점을 감안할 때, AI가 생성한 결과물의 저자 또는 저자들이 누구인지를 식별하는 명확한 규정은 존재하지 않는 것으로 판단한다고 하고 있다(Congressional Research Service, 2023). 일본의 경우 'AI와 저작권에 관한 생각'라는 보고서[19]를 통해 AI 산출물의 저작물성이 인정되기 위해서는 인간에 의한 '창작적 기여'가 필요하다는 것으로 정리하고 있다. 해당 보고서는 생성형 AI의 산출물이 저작물성을 인정받기 위해서는 인간에 의한 지시와 입력(프롬프트 등)의 내용과 분량, 생성 시행 횟수, 복수 생성물 중에서의 선택 등)을 고려하여 인정 여부를 결정하게 된다고 한다. 다만 이 경우에도 단순히 선택한 행위는 창작적 기여가 인정되기 어려운 것이라고 하고 있다(문화심의회저작권분과회법제도소위원회(文化審議会著作権分科会法制度小委員会), 2024).

한편 미국 저작권청(U.S. Copyright Office)은 생성형 AI를 포함한 AI 산출물에 대한 저작권을 등록하기 위한 구체적인 프로세스를 만들기도 하였다. 미국 저작권청은 인간과 AI가 공동으로 생성한 산출물에 대하여 인간이 기여한 부분에 대해 증명이 가능하다면 이에 대해서는 저작권을 인정할 수 있다고 하였다. 즉, AI 기술이 인간의 프롬프트를 통해 응답으로 산출물을 생성한 경우 기술에 의해 저작물이 생성된 것으로 볼 수 있으나, 인간이 창의적 방식으로 AI가 생성한 자료를 선택하거나 배열하여 작품 전체가 독창적인 저작물을 구성하는 경우 또는 아티스트가 원래 자료를 수정함으로써 저작권 보호 기준을 충족하는 정도에 이르는 경우 등에 있어서는 AI가 생성한 자료에 독립적이고 영향을 미치지 않는 인간의 저작물에 대해서는 저작권을 인정할 수 있다는 것이다(Federal Register, 2023).

(2) 생성형 AI의 저작권 침해 문제에 대한 제도적 대응 논의

생성형 AI가 훈련 과정에서 타인의 저작물을 학습하는 경우에 대해서는 저작권 침해 문제를 최소화하기 위하여 AI 개발자뿐만 아니라 개별 국가에 대해서도 필요한 조치가 필요한 것으로 공감대가 형성되고 있는 것으로 보인다. 2023년 5월 일본 히로시마에서 개최된 G7 정상회담에서 각국 정상은 생성형 AI에 관한 논의를 지속하기로 하고 '히로시마 AI 프로세스(Hiroshima AI Process)' 수립에 합의한 바 있다. 이에 따라 10월 '고도화된 AI 개발 조직을 위한 국제 이행 원칙(Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI system)'과 '고도화된 AI 개발 조직을 위한 행동강령(Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems)'을 합의하였는데, 해당 지침은 AI에 관한 11가지의 원칙을 제시하면서 AI 개발자에게 ① 개발 단계에서 AI의 리스크나 보안상 취약점 등을 특정하고 대책을 강구할 것과, ② AI에 데이터를 학습시킬 경우 개인정보나 저작권을 보호할 것을 요구하고 있다.

유럽연합은 회원국 전체에 포괄적으로 적용되는 인공지능에 관한 일반법인 'AI Act(REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024 laying down harmonised rules on artificial intelligence, 이하 'AI Act')'를 2024년 입법하고[20] 고위험에 해당

19) 일본은 문화청(文化庁)을 중심으로 2023년 7월부터 AI와 저작권에 관하여 논점을 정리하고 의견을 수렴하는 과정에 있다. 여러 쟁점들을 정리하여 2024년 1월 15일 'AI와 저작권에 관한 생각 초안(AI と著作権に関する考え方について(素案)'을 공개한 후 의견을 수렴하였는데, 의견수렴 결과 등을 종합하여 생성형 AI와 관련한 논점들에 대한 저작권법 해석 방향 등을 작성하여 발간한 보고서이다(문화심의회저작권분과회법제도소위원회(文化審議会著作権分科会法制度小委員会), 2024).

하는 AI시스템에 대하여 위험성을 최소화하기 위하여 필요한 조치들을 의무화하였다. 동 법은 '범용AI모델(General-purpose AI model)'을 정의하고[21] 범용AI 모델 공급자가 준수하여야 하는 의무 중의 하나로 EU의 저작권법을 준수하기 위한 정책을 마련하고 학습에 사용된 콘텐츠에 관한 '충분히 상세한 요약서'를 작성하고 이를 일반에게 공개할 것을 요구하고 있다(Article53 1. (c)&(d)). 미국 연방 하원은 2024년 4월 '생성형 AI 저작권 공개법안(Generative AI Copyright Disclosure Act of 2024)'을 발의하기도 하였다. 해당 법안에서는 생성형 AI 시스템의 구축에 사용되는 학습용 데이터 세트의 작성자 또는 주요한 변경자에 대해 (작성자의 경우) 학습용 데이터 세트에 이용된, 또는 (주요한 변경자의 경우) 학습용 데이터 세트를 변경하기 위해 이용한 모든 저작물에 대한 '충분히 상세한 요약서(Sufficiently Detailed Summary)'를 저작권청장에게 통지하도록 하고 있다. 또한 통지시 학습용 데이터 세트가 온라인상에 공개되어 있다면 해당 데이터 세트의 URL(Uniform Resource Locator)을 통지에 포함시켜야 한다(H.R.7913 - Generative AI Copyright Disclosure Act of 2024). 영국의 경우 저작권 침해에 문제를 입법적으로 해결할 필요성이 제시되고 있는데 영국 의회의 커뮤니케이션디지털위원회(Communications and Digital Committee)가 2024년 2월 발간한 '대규모 언어 모델과 생성형 AI'는 저작권과 관련해서 필요하다면 정부가 법률 제·개정을 통해 대형 기술 회사들이 저작권자의 허락 또는 보상 없이 저작물을 사용하여 발생하는 분쟁을 종결시켜야 한다고 제안하고 있다. 저작권자가 AI 학습데이터를 확인할 수 있도록 하고, AI 제공자는 웹 크롤링 대상을 공개하여야 한다고 요구하고 있다(UK Parliament, 2024).

Ⅳ. 차별과 불평등의 심화

1. 현상과 문제점

AI 기술의 대표적인 문제점 중 하나로 제시되는 것이 바로 편향과 오류 가능성이다. 이는 AI 기술이 빅데이터와 알고리즘을 통해 만들어진 결과물이기 때문에 제기되는 문제점이다. '편향'은 국제 표준화 기구(International Organization for Standardization, ISO)에 의하면 '기준값이 진실에서 벗어나는 정도'라고(OECD, 2007) 정의된다. 이러한 맥락에서 AI 기술의 편향은 통계적 관점에서는 AI기술이 체계적으로 부정확한 행동을 보이는 경우를 의미하며(NIST, 2022), 이러한 부정확한 행동은 AI기술을 기반으로 하는 인간 행동에 영향을 미치게 된다. 이러한 사회적 현상까지 포괄한다면 AI 편향(Bias in AI)이란 특정 그룹내 개인에게 체계적으로 덜 유리한 결과를 도출하고 그러한 피해를 정당화할 만한 그룹 간의 관련성이 없는 컴퓨터 시스템 또는 프로세스에 의한 의사결정을 의미한다. 이러한 AI의 편향은 대표성이 없거나 불완전한 학습데이터 또는 역사적 불평등을 반영하는 결함이 있는 정보에 의존하는 데에서 비롯된다. AI의 편향을 방치하게 되면 알고리즘을 만든 프로그래머에게 별다른 차별 의도가 없다고 하더라도 특정 개인 혹은 집단에게 불균등한 영향을 미칠 수 있는 의사결정이 내려질 수 있다(UN WOMEN, 웹페이지).

기존의 머신러닝 모델도 공정성과 편향성 등에 대한 동일한 문제가 발생하지만 생성형 AI는 학습데이터의 광범위함에 더하여 이용자와 직접 상호작용한다는 특성으로 인해 그 위험성이 더욱 가중된다. 생성형 AI가

20) 2021.4.21. 발의, 2024.6.13. 의회 및 이사회가 공동서명하였고, 2024.7.12. 관보 게재되었다. 해당 법률은 2024년 8월 1일 발효되었으며, 발효일로부터 6개월인 2025년 2월 2일부터 2027년 8월 2일까지 순차적으로 시행된다.

21) '범용 AI 모델'이란 시장 출시 방법 및 다양한 하방 시스템(downstream systems) 또는 애플리케이션에 통합되는 방법에 상관없이 상당한 일반성을 가지며 광범위한 범위의 다양한 업무를 능숙하게 수행할 수 있는 AI 모델로서, 상당한 규모의 자기지도학습을 이용하여 대량의 데이터를 학습한 AI 모델을 포함한다. 다만 연구·개발 및 시제품 제작 활동을 위해 시장 출시 전에 이용되는 AI 모델은 포함하지 않는다(EU AI Act Article3 (63)). 한편, '범용 AI 시스템'도 정의하고 있는데, 이는 범용 AI 모델에 기반하고 다양한 목적을 수행하는 성능을 보유한 AI 시스템으로서 다른 AI 시스템에 통합되거나 직접 이용되기도 하는 AI 시스템을 의미한다(EU AI Act Article3 (66)).

패턴을 학습, 유추, 예측하기 위해서는 과거의 데이터에 의존하는데, 이러한 학습 데이터에 이미 성별, 인종, 연령, 사회경제적 편견과 같은 과거의 편향성이 녹여져 있으며(Kim, 2023), 이용자에게 내재되어 있는 편향성이 상호작용 과정에서 그대로 생성형 AI에게 학습되기 때문이다.[22) 23)] 이러한 차별과 불평등의 문제들은 개인간 차별의 문제에서 지역간·국가간 차별문제로 확장될 수도 있다.[24)]

더욱 어려운 점은 생성형 AI의 개발과 사용과정에서 차별과 편향의 문제가 심화되고 있지만 이를 법제도적으로 통제하는 것이 사실상 힘들다는 것이다. 이는 '편향'이라는 개념의 모호성 때문이다. 즉 생성형 AI의 결과물이 '편향'된 결과를 초래하였다는 것은 다양한 가치판단과 견해에 따라 각각 다르게 해석될 가능성이 농후하다. 또한 편향을 완화하기 위해서는 알고리즘과 데이터의 조정뿐만 아니라 사회적 규범과 가치의 복잡하고 진화하는 특성에 대한 이해가 필요한데, 이를 기술적으로 해결하는 것은 쉽지 않다. 따라서 편향성을 완화시키거나 제거하여야 한다는 법적 의무를 부과한다 하더라도 무엇이 편향에 해당하는지에 대한 의문이 지속적으로 제기될 것으로 보인다(Hacker, et al., 2024). 더욱이 AI의 복잡성으로 인해 편향성의 원인을 정확하게 찾아내고 의도하지 않은 결과에 대해 누가 책임이 있는지를 확인하고 책임을 묻는 것 또한 어려운 문제이다(Kim, 2023).

2. 제도적 대응 논의

AI의 편향 문제는 생성형 AI 등장 이전부터 대응이 필요하다는 논의가 활발하게 진행되어 왔다. 국제기구는 물론 개별 국가의 정부에서 'AI 윤리'등을 강조하면서 AI의 편향성 문제를 지적하고 경계하여 왔다.[25)] 이러한 문제의식은 생성형 AI 이후에도 지속·확산되고 있다. 2024년 유네스코는 LLM 등의 생성형 AI에서 성별, 인종 등의 편향이 유발되는 경향성이 있다는 연구결과를 발표하면서, 2021년 유네스코가 제시한 권고안의 이행이 시급함을 언급하였다(UNESCO, 2024a). 해당 발표와 함께 오드리 아줄레이(Audrey Azoulay) 유네스코 사무총장은 권고안에서 강조하는 바와 같이 각국 정부가 명확한 규제 프레임워크를 개발 및 집행할 것을 촉구한다고 하였다(UNESCO, 2024b).

차별 발생에 대한 보호조치는 생성형 AI 시스템뿐만 아니라 AI시스템 자체에 대해 요구되는 것이 AI 규제의 전체적 흐름으로 보인다. 유럽연합은 'AI ACT'에서 고위험AI시스템의 경우 데이터로 AI를 학습시키는 경우 법률에서 요구하는 기준을 충족하는 데이터로 학습·검증·테스트를 실시하도록 하고 있다. 학습·검증·테스트 데이터 세트는 특히 데이터의 결과물이 향후 작업을 위한 입력에 영향을 미치는 경우, 사람의 건강과 안전에 영향을 미치거나 기본권에 부정적 영향을 미치는 등 EU의 법률에 따라 금지된 차별을 초래할 수 있는 편향 가능성을 고려하여야 한다(Article10 2.(f)). 그리고

22) 마이크로소프트의 '테이'는 일부 극우성향 이용자들의 대화를 학습하면서 각종 인종 및 성차별 발언과 자극적인 정치적 발언 등으로 운영이 중단되었으며(Yonhap News, 2016), 챗봇 서비스인 '이루다'의 경우 학습데이터에 포함된 편향성이나 편견이 알고리즘에 반영되었고 이용 과정에서는 이용자들의 혐오발언을 학습하면서 인공지능 윤리 문제가 불거진 바 있다(Jungang Ilbo, 2021; Yonhap News, 2016).

23) 생성형 AI의 편향된 결과물 도출과 관련한 연구에서는 Stable Diffusion과 DALL·E는 '아프리카'라는 단어를 빈곤과 연결시키거나 '가난'을 어두운 피부톤과 연관시키는 등 일반적 고정관념에 의존한다는 결과를 제시하고 있다. 가정부는 유색인종으로 승무원은 여성으로 묘사하였고, 비율은 실제 인구통계보다 높은 수치에 해당한다고 한다(Nature, 2024.; Bianchi, et al., 2023).

24) 일례로 ChatGPT의 경우 영어텍스트로 작성된 문서를 학습한 결과 영어 외 다른 언어에 대한 답변 등은 원활하게 이루어지지 못하기도 하고, 비영어권 국가에 대한 편향된 결과를 도출한다는 연구 결과도 있다(Liang, et al., 2023; HAI, 2023.).

25) 대표적으로 2019년 유럽연합의 AI HLEG가 발간한 '신뢰할만한 인공지능(AI)를 위한 윤리 가이드라인(The Ethics Guidelines for Trustworthy Artificial Intelligence)'과 2020년 미국 FTC(Federal Trade Commossion; FTC)의 '인공지능 알고리즘 이용에 관한 지침(Using Artificial Intelligence and Algorithms)', 2020년 OECD의 'OECD AI 권고안(OECD Principles on AI)', 2021년 유네스코(UNESCO)의 'AI 윤리 권고안(Recommendation on the Ethics of Artificial Intelligence)' 등이 있다. 국내에는 과학기술정보통신부가 2021년 발표한 '인공지능(AI) 윤리기준'이 대표적이며, 대부분의 지침이나 권고안 등에서 AI시스템이 차별과 편향 등을 유발하지 않도록 주의하여야 한다는 내용이 포함되어 있다.

학습·검증·테스트 데이터는 충분한 대표성을 보유하여야 하며, 가능한 최대한으로 오류를 없애야 하며, 시스템의 의도된 목적에 비추어 완전하여야 한다. 그리고 가능한 경우 고위험AI시스템을 이용할 개인이나 개인과 연관된 집단에 관련된 적절한 통계적 특성을 반영하여야 한다(Article10 3&4). 또한 일정 조건[26] 하에서 편향성 및 수정사항 파악을 위한 목적으로 예외적으로 '특별한 범주의 개인정보'[27]를 처리할 수도 있도록 하고 있다(Article10 5.). 미국은 법률을 통해 AI의 위험성을 통제하는 것은 아니지만 규제의 필요성과 방향성에 대해서는 행정명령 등을 통해 제시하고 있다. 2022년 백악관 과학기술정책실(Office of Science and Technology Policy, OSTP)이 발표한 '인공지능 권리장전 청사진(Blueprint for an AI Bill of Rights)'은 모든 자동화된 시스템은 판매 또는 사용되기 전에 알고리즘 차별이 없는지 확인하기 위해 테스트를 실시해야 한다고 제시하고 있다. 알고리즘 차별에 대한 보호에는 광범위하게 해석되는 형평성을 보장하기 위한 설계가 포함되어야 하며, 이러한 보호기능은 설계, 개발, 배포 프로세스 전반에 걸쳐 설정되어야 한다고 한다. ① 잠재적 차별과 형평성에 미치는 영향을 식별하기 위해 사전 형평성 평가를 수행하고, ② 모든 데이터는 계획된 배포의 기반이 되는 지역사회를 대표해야 하며, 데이터의 과거 및 사회적 맥락을 기반으로 편견이 존재하는지 검토해야 한다고 한다. 또한 ③ 자동화된 시스템의 설계, 개발, 배포에 인구 통계 정보를 직접 사용하는 경우(차별에 대한 시스템 평가 또는 차별에 대응하기 위한 시스템 사용 이외의 목적으로) 알고리즘 차별로 이어질 위험이 높으므로 지양해야 할 것을 요구하고 있으며, ④ 시스템은 장애가 있는 사용자의 접근성을 보장하는 방식으로 설계, 개발, 배포될 것을 요구한다. ⑤ 배포전 테스트를 실시하고, ⑥ 격차 평가에서 격차가 확인되면 이를 완화·제거하기 위한 조치를 취하여야 한다. ⑦ 시스템의 정기적인 모니터링을 통해 테스트 중 확인하지 못한 불평등과 상호작용으로 인해 발생할 수 있는 차별을 평가해야 하는 등 AI 시스템의 차별을 완화시키기 위하여 필요한 조치들을 상세하게 제시하고 있다(The White House, 2022).[28]

중국은 생성형 AI에 관한 별도의 법률인 '생성형 AI 서비스 관리 잠정방법(生成式人工智能服务管理暂行办法)'을 2023년 8월부터 시행중에 있다. 해당 법률에서도 생성형 AI 서비스는 알고리즘 설계, 훈련 데이터 선택, 모델 생성 및 최적화, 서비스 제공과정에서 민족, 이념, 국적, 성별, 연령, 직업, 건강 등에 따른 차별을 방지하기 위한 효과적인 조치를 취할 것을 의무화하고 있다(제4조 제2항). 그리고 생성형 AI 서비스 제공자는 서비스를 제공할 경우 데이터 훈련과 품질제고 등 관리조치를 취해야 하는데, AI가 데이터를 훈련하는 과정에서는 훈련 데이터의 신뢰성, 정확성, 객관성, 다양성을 강화하기 위한 효과적인 조치를 취하여야 한다(제7조 제4항).[29]

26) (a) 합성 데이터나 익명 데이터를 포함한 기타 데이터를 처리함으로써 편향 탐지 및 보정을 효과적으로 수행할 수 없는 경우;
(b) 특별한 범주의 개인정보는 개인정보를 재이용함에 있어 기술적 제한이 적용되고, 가명처리를 포함한 최첨단 보안 및 사생활 보호조치가 적용될 수 있는 경우;
(c) 오용 방지와 승인받은 사람만 적절한 비밀 유지 의무하에 해당 개인정보에 접근하는 것을 보장하기 위하여, 특별한 범주의 개인정보에 대한 접근을 엄격히 통제하고, 문서화를 포함한 처리된 개인정보가 안전하고 보호되는 것을 담보하는 조치와 적절한 보호처리가 적용되는 경우;
(d) 특별한 범주의 개인정보가 다른 당사자에게 전송, 이전 또는 달리 접근될 수 없는 경우;
(e) 특별한 범주의 개인정보가 편향 보정 시점 또는 보존기간 종료 시점 중 먼저 도래하는 시점에 즉시 삭제되도록 하는 경우;
(f) Regulation (EU) 2016/679 및 Regulation (EU) 2018/1725, Directive (EU) 2016/680에 따른 처리절차 기록이 특별한 범주의 개인정보 처리가 편향을 탐지하고 보정하기 위해 엄격히 필요했고 다른 데이터의 처리로 이러한 목표를 달성할 수 없었던 이유를 포함하는 경우

27) 인종 또는 민족, 정치적 견해, 종교 또는 철학적 신념, 노동조합 가입 정보, 유전정보, 생체정보, 건강정보, 성생활 또는 성적 취향에 관한 데이터 등 (Regulation (EU) 2016/679(GDPR) Article9 1.).

28) 이는 미국 내 대중의 권리를 위협하는 가장 큰 문제로 기술, 데이터, 자동화된 시스템을 사용하는 것이라는 것을 강조하며, 이러한 문제를 해결하면서도 강력한 이점을 가진 기술 등을 사용하기 위한 원칙과 정책 추진 방향성을 제시하기 위한 것이다.

Ⅴ. 사용자 역량과 범죄 문제

1. 거짓 정보에 대한 선별 역량

생성형 AI는 앞서 언급한 바와 같이 이른바 환각증상을 나타낸다는 특징이 있다. 따라서 생성형 AI에 의해 산출된 정보에 대해 진위여부와 정확성 등을 선별할 수 있는 능력이 사용자에게 요구된다. 사용자의 역량은 생성형 AI를 이용하는 과정에서도 고려가 필요하다. 생성형 AI에 의해 필요한 정보를 얻거나 원하는 콘텐츠를 생성하기 위해서는 이를 잘 유도할 수 있는 사용자의 역량이 중요한데 생성형 AI는 사용자가 무언가를 요청하면 기존에 학습한 모델을 사용해 새로운 콘텐츠를 생성해 내므로 사용자의 요청에 따라 생성되는 콘텐츠의 내용과 질이 결정되기 때문이다. 그렇기 때문에 생성형 AI의 위험성 자체가 사용자에 의해 결정될 수도 있다. 사용자가 위험도 낮은 질문에 답하기를 원한다면 이 경우에는 위험도가 높은 인공지능이 아니지만 사용자가 고위험 활동을 수행하기를 원할 경우 고위험 인공지능으로 간주될 수도 있는 것이다(Decker, 2023).

생성형 AI에 의한 허위정보 양산도 문제된다. 생성형 AI 등장 이전에도 AI를 기반으로 하는 딥페이크 또는 딥보이스와 같이 허위의 정보를 만들어서 배포하는 문제들이 발생했었지만,[30] 생성형 AI로 허위정보가 더욱 빠르게 생성, 확산되고 있다. 사실적인 이미지, 비디오, 텍스트의 생성을 자동화하여 실제 콘텐츠와 인공적인 콘텐츠를 구별하기가 점점 더 어려워지고, 생성 자체가 무료로 이루어지기 때문에 개별적이고 타겟팅된 정보가 대량으로 생성되는 것도 가능해졌다. 생성형 AI의 정교함은 뉴스, 미디어 등 정보가 무결하여야 하는 것이 핵심인 분야에 도전 과제를 제기한다(Loth, et al., 2024). 생성형 AI를 이용함에 있어 세심한 프롬프트를 통해 LLM이 원하는 것을 정확하게 생성하도록 할 수는 있다. 그러나 이러한 노력에는 LLM의 작동방식에 대한 통찰력과 LLM의 한계에 대한 인식이 필요한데, 문제는 보통의 사용자들은 생성형 AI의 결과가 정답을 추론할 것이라고 신뢰한다는 것이다(Thome, 2024). 생성형 AI의 환각증상은 생산된 지식의 진리성에 대한 판단과 잘못된 지식의 생산에 대한 검증 모두 사용자에게 요구하고 있다(Byun, 2023).

2. 범죄 도구로서의 사용

생성형 AI는 범용성을 지니고 있어 다양한 방식으로 활용할 수 있는 가능성이 있어 이를 범죄의 목적으로 활용하는 것도 가능하다. 이에 ChatGPT를 활용하여 악성코드, 랜섬웨어, 피싱메일 제작 등 사이버 공격에 활용할 가능성에 대해서도 우려가 제기되고 있다(Carolina & Lorenzo, 2024). 실제로 다크웹 등에서 실제로 ChatGPT를 사용하여 새로운 악성코드 등을 제작할 수 있는 사이버 공격 도구를 생성하려고 한 시도가 발견되었다는 것이 밝혀지기도 하였고(CheckPoint, 2023), 이외에도 LLM에 대한 사이버 공격 유형인 '프롬프트 인젝션(Prompt Injection) 공격'으로 프롬프트를 통해 데이터를 유출하도록 생성형 AI시스템을 조작하거나 잘못된 정보를 유포하는 것도 가능하다고 한다(IBM, 2024).[31] [32] ChatGPT와 같은 대화형 AI 서비스는 실제 사람과 대화하는 것처럼 자연스럽게 대화할 수 있어 사

29) 이 외에도 ① 법적 출처가 명확한 데이터와 기본 모델을 사용하여야 하며, ② 지식재산권이 관련된 경우 법률에 따라 타인이 향유하는 지식재산권을 침해해서는 안 된다. 그리고 ③ 개인정보는 개별적으로 동의를 받거나 법률상의 근거가 있어야 하고, ④ 사이버보안법, 데이터보안법, 개인정보보호법 등 법률과 관련 행정법규 등의 요구사항을 준수하여야 한다(제7조 및 제8조). 그리고 동 법은 생성형 AI를 "알고리즘, 모형, 규칙에 근거하여 텍스트, 도면, 음성, 동영상, 코드 등의 콘텐츠를 생성하는 기술"이라고 정의하고 있다(제22조 제1항).

30) 미국 백신업체인 맥아피의 연구에 따르면 인공지능은 3초만에 사람의 음성을 복제한다고 하는데, 전 세계 성인 7천명을 대상으로 조사를 한 결과 성인의 4분의 1이 일종의 인공지능 음성 사기를 경험한 적이 있는 것으로 나타났다. 그리고 응답자의 70%가 인공지능과 실제 음성을 구분하지 못하였고 사기 피해자의 77%가 재산 피해를 입었다고 잃었다고 한다(McAfee, 2023.).

31) 이는 LLM 모델에 사용자가 직접 다른 명령을 내릴 수 있게 하는 공격 기법이다.

기와 피싱을 더 효과적으로 수행할 수 있다. 특정 사람의 말투나 기관의 문구 등을 학습하여 유사한 문장으로 글을 작성할 수 있고 이를 악용하여 특정 인물이나 기관을 가장하는 메일을 보내는 방식으로 피싱 메일 전달이 가능하다. 생성형 AI를 채팅방에 연결할 경우 피해자와의 실시간 상호작용이 가능하므로 피해자는 본인이 AI와 대화한다는 것을 인지하지 못할 수도 있다(Park, et al., 2024). 생성형 AI가 만들어내는 불법정보, 성적 학대 이미지의 생성과 유포 등도 심각한 사회문제화 되고 있다. 미국 정부는 생성형 AI가 만든 이미지를 포함하여 이미지를 기반으로 성적 학대를 당한 피해자를 위한 법적 보호를 강화할 것을 의회에 촉구하는 성명을 발표하기도 하였다(The White House, 2024).[33] FTC는 오픈소스를 통해 AI 모델이 공개되면 경쟁을 활성화하는데 도움이 되겠지만 생성할 수 있는 이미지 유형에 대한 보호 기능을 제거하고 동의 없는 사적인 이미지를 만드는 등 악의적 사용자에 의해 오픈소스가 오용될 가능성이 있다고 언급한 바 있다(FTC, 2023).[34]

3. 사용자의 역량, 악용에 의한 법적 문제의 한계와 대응방식

사용자의 역량 문제는 생성형 AI시스템을 통제하는 방식으로 위험을 완화시킬 수 없다는 한계가 존재한다. 현재 AI에 대한 규제는 AI시스템을 개발하고 공급하는 사업자에 대하여 위험성을 완화시키기 위하여 필요한 조치를 의무화하는 방식으로 구성된다. 그러나 사용자에 의한 위험은 AI시스템의 개발자나 공급자가 통제할 수 있는 영역의 문제가 아니며, 사용자가 어떤 방식으로 AI시스템을 사용하게 될지를 예측하기도 어려운 문제이다. 따라서 서비스 개발 및 공급자에 의한 통제와 함께 사용자의 역량을 강화하고 사용자에 의한 오용을 방지하는 법제도의 설계가 중요한 과제가 된다.

유럽연합은 'AI Act'에서 'AI 리터러시'라는 개념을 "공급자, 배포자, 영향을 받는 사람이 각자의 권리와 의무를 고려하고 정보에 입각하여 AI시스템을 배포할 수 있도록 하고, AI의 기회와 위험, 발생할 수 있는 피해에 대해 인식할 수 있도록 하는 기술, 지식, 이해"라고 하고(Article 3(56)), AI 시스템의 공급자와 배포자로 하여금 AI 리터러시를 강화하는 조치를 취하도록 요구하고 있다(Article 4). 공급자와 배포자는 직원 및 자신을 대신하여 기타 AI 시스템의 운영과 이용을 담당하는 사람의 기술적인 지식, 경험, 교육 및 훈련, AI 시스템이 이용되는 맥락과 AI 시스템을 이용할 사람이나 집단을 고려하여 충분한 수준의 AI 리터러시를 최대한 보장하기 위한 조치를 취하여야 한다. 이는 AI 시스템의 공급자와 배포자가 리터러시를 강화하는 조치를 취함으로써 AI 시스템이 오용 없이 의도된 목적대로 사용될 수 있도록 하기 위함이다. 사용자에 의한 예측 불가능한 오용이나 악용을 예방하고, 안전하게 활용되게 하기 위해 정확한 사용방법과 관련한 정보 등을 제공하도록 하는 것이다.

한편, 미국 바이든 행정부는 2023년 10월 '안전하고 신뢰할 수 있는 인공지능의 개발 및 이용에 관한 행정명령(Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence)'을 발표하였는데 해당 행정명령은 AI를 안전하게 활용할 수 있도록 연방기관으로 하여금 구체적인 조치를 취하도록 요구하고 있다(The White House, 2023). 생성형 AI와 관련해서 해당 행정명령은 생성형 AI가 만들어낸 합성 콘텐츠[35]로 인한 위험성을 줄이기 위해 상무부 장관(Secretary of Commerce)

32) 프롬프트 인젝션 공격을 포함한 AI에 대한 공격 유형과 이를 방어하기 위한 방법에 대해서는 NIST, 2024a.를 참고.

33) The White House, "A Call to Action to Combat Image-Based Sexual Abuse", 2024.3.23., Available at: https://www.whitehouse.gov/gpc/briefing-room/2024/05/23/a-call-to-action-to-combat-image-based-sexual-abuse/, Accessed: 2024.8.7.

34) 이 외에도 ChatGPT에 대하여 공격이 성공하였던 사례에 대해서는 Maanak Gupta et al., 2023.을 참고.

35) '합성 콘텐츠(synthetic content)'란 AI를 포함한 알고리즘에 의해 상당히 수정되거나 생성된 이미지, 비디오, 오디오 클립 및 텍스트와 같은 정보를 의미한다(Sec.3. 4.(ee)).

으로 하여금 AI시스템에서 생성된 합성콘텐츠를 식별하고 진위성과 출처를 확인과 디지털 콘텐츠 인증 및 합성콘텐츠 탐지 조치에 대한 도구와 관행 등에 관한 지침을 만들도록 하고 있다. 구체적으로는 ① 콘텐츠의 인증 및 출처 추적, ② 워터마킹 등 AI를 기반으로 생성 및 합성된 콘텐츠의 라벨링, ③ 합성콘텐츠 탐지 ④ 생성형 AI가 아동의 포르노 제작이나 특정 개인의 동의 없는 신체 이미지 제작 방지, ⑤ ①~④의 목적을 위해 사용되는 소프트웨어 테스트, ⑥ 합성콘텐츠 감사 및 유지 관리 등에 관한 표준 및 기술개발 관련 보고서를 관리예산처장(Director of OMB)과 대통령 국가안보보좌관(Assistant to the President for National Security Affairs)에게 제출하고 관련 지침을 개발하여야 한다(Sec.4. 4.5).

이외에도 미국 NIST는 생성형 AI의 위험관리 방안과 이중사용 파운데이션 모델의 오용 위험관리 방안을 담은 프레임워크 문서를 발표한 바 있다(NIST, 2024b). NIST는 생성형 AI의 12가지 잠재적 위험 중 하나로 허위 정보나 혐오 발언 및 기타 유해한 콘텐츠가 생성된다는 것과 생성형 AI시스템이 결과물을 왜곡하거나 환각현상을 일으키는 것이 포함된다고 하면서, 생성형 AI시스템 공급자가 이러한 위험을 관리하고 최소화하기 위하여 필요한 조치들에 대한 정보를 제공하고 있다.

생성형 AI에 의하여 만들어진 결과물에 대하여 '표시'가 필요하다는 논의가 확산됨에 따라 우리 국회도 AI 기술을 통해 제작된 콘텐츠를 표시하도록 하는 「콘텐츠산업진흥법 개정안」을 발의한 바 있다. 법안은 AI 기술을 이용하여 콘텐츠를 제작한 제작자는 해당 콘텐츠가 AI기술을 이용하여 제작된 콘텐츠라는 사실을 표시하도록 하고 있다. 이러한 표시를 통해 사용자의 혼선을 방지하고 AI 콘텐츠의 신뢰성과 책임성을 강화하고자 하는 것이다(강유정의원 등 10인 제안, 의안번호: 2200048, 2024).

일본 문화청도 생성 AI를 악의적으로 사용하여 저작권을 침해하거나 불법적인 콘텐츠를 생성하는 경우, 이는 창작자 및 실연자의 권리에 심각한 위협이 될 수 있으며, 이러한 불법적 활용을 방지하기 위한 제도적 장치가 필요하다고 언급하고 있다(일본 문화심의회저작권분과법제도소위원회(文化審議会著作権分科会法制度小委員会), 2024).

Ⅵ. 독점 및 환경 규제 관련 쟁점

1. 독점시장의 형성

AI시스템은 기본적으로 학습 데이터의 양과 컴퓨팅 인프라 역량에 비례하여 산출물의 정확도가 향상된다. 즉, 생성형 AI의 경쟁력은 대규모 데이터, 컴퓨팅 인프라, 인재, 자본력의 시너지에 비례하기 때문에 기존 글로벌 빅테크 기업이 유리한 위치를 선점할 수밖에 없다. Google, Microsoft, Meta 등 빅테크 기업들은 컴퓨팅 인프라(AI반도체), 클라우드, AI플랫폼(AI기반모델), 어플리케이션으로 구성되는 생성형 AI 가치사슬 강화를 위해 AI 전문인력, 컴퓨팅 인프라, 기술력 확보를 위해 투자를 점점 확대해나가고 있다. MS는 OpenAI에 2019년 10억 달러 투자 후 2023년 100억 달러 투자를 통해 OpenAI 기술의 독점 라이센스를 확보하였고, 2024년 3월 두 기업은 AI 슈퍼컴퓨터를 포함한 데이터센터를 구축하기 위하여 1000억 달러(약 134조 원)을 투자할 것이라고 밝혔다(Reuters, 2024). Google은 자사 AI칩(TPU) 기반의 클라우드 플랫폼을 구축하였고, 자회사 DeepMind와 함께 다양한 초거대언어모델(LaMDA, PaLM 등) 및 대화형 검색서비스 Bard와 같은 AI 서비스를 개발하였고, 2023년 2월 OpenAI의 경쟁사인 Anthropic에 대해 약 20억 달러를 투자하기도 하였다(CNBC, 2023).

생성형 AI시스템의 경우 상당한 컴퓨팅 리소스가 필요한데 이는 비용이 많이 소요되므로 소수 기업에 의해서만 통제될 가능성이 높아 반경쟁적 관행으로 이어질 가능성이 높다는 우려가 제기되고 있다(Forbes, 2024). 미국 FTC는 DOJ와 함께 2024년 6월 OpenAI와 NDIVIA에 대한 '독점금지법' 위반 여부 확인을 위

해 조사에 착수하였다(The New York Times, 2024a; The New York Times, 2024b).

생성형 AI시스템을 구축하고 서비스를 제공하기 위하여 필요한 방대한 양과 고품질의 데이터를 비롯하여[36] 빠른 데이터 처리 능력을 갖춘 컴퓨팅 인프라 등이 필요하다. 이는 신규 사업자의 시장 진입에 장애요소가 된다. FTC는 기존 기업이 다년간 사용자로부터 수집한 데이터, 특히 대량의 데이터를 축적하는 디지털 플랫폼을 소유하고 있는 경우 기존 기업이 경쟁 우위를 점할 수 있음은 물론, 기존 기업은 데이터를 수집하거나 스크래핑 하기 위한 독점적 수집 도구 및 기술을 갖추고 있을 가능성 또한 높다고 보고 있다. 즉, 사전 학습된 베이스 모델을 만드는 데 드는 높은 진입 비용으로 인해 고품질의 사전 학습된 모델이 소수의 기존 업체에 의해 통제되는 방식으로 생성형 AI 시장이 형성될 수 있다는 것이다. 이미 시장에 진출한 생성형 AI 서비스 기업은 네트워크 효과를 활용하여 지배적 위치를 유지하거나 시장 지배력을 집중시킬 가능성이 크다고 보고 있다(FTC, 2023).

2. 새로운 환경 문제의 발생

MIT는 2022년 대규모언어모델에 더러운 비밀(Dirty Secret)이 있다고 지적했다(MIT Technology Review, 2022). 생성형 AI가 학습하는 과정에서 많은 양의 에너지를 소비하고 온실가스를 배출하기 때문이다. 구글의 2023년 온실가스 배출량은 2019년보다 48% 증가하였다고 하는데, 이에 대해 구글 AI 컴퓨팅의 강도가 높아짐에 따라 에너지 수요가 증가하였기 때문이라고 한다(Google, 2024). 2023년 한 연구결과에 의하면 ChatGPT와 같은 생성형 AI시스템은 작업별 소프트웨어를 실행하는 기계보다 약 33배 더 많은 에너지를 사용한다는 한다(Alexandra, et al., 2023). 국제에너지기구(International Energy Association, IEA)는 데이터센터가 2026년까지 연간 총 1,000테라와트시(thw)를 사용할 수 있다고 했는데 이는 일본 전체 전기 소비량과 거의 같다고 한다(IEA, 2024).

3. 경쟁 및 환경 규제 당국의 대응

G7의 경쟁당국은 2023년 11월 도쿄에서 'G7 Joint Competition Enforcers and Policy Makers Summit'을 개최하고 생성형 AI 분야의 독과점 문제를 경계하며 공동성명을 채택하였다. 성명에는 디지털 시장이 급격한 독점화와 시장지배적 지위를 낳는 경향이 있다고 지적하며 빅테크 기업들이 생성형 AI 분야에서도 반경쟁적 행위를 할 수 있다고 우려하고 있다. 성명에는 디지털 시장에서 반경쟁적 행위를 조사, 이해, 분석, 시정하기 위하여 경쟁당국이 역량을 강화해야 하며, 반경쟁적 우려를 해소하기 위해 기존 권한의 개혁과 법률의 개선이 필요하다고 강조하고 있다(G7 2023 Hiroshima Summit, 2023). 이를 미루어 볼 때 생성형 AI시장의 구조적 변화에 따라 각 국가의 경쟁법과 경쟁당국의 권한이 강화·효율화될 것이 예측된다. G7 서밋에 참여한 국가들은 생성형 AI에 대한 국제적 규범 마련이 필요하다는 것에 공감하면서, 기업은 AI의 위험을 식별하고 완화하기 위하여 적절한 조치를 해야 하고, AI의 기능과 제한사항, 사용 및 오용에 관한 보고서를 공개하여야 하며, 강력한 보안 통제에 투자하여야 한다고 강조하고 있다.

유럽의 AI Act는 환경 문제에 대하여 직접적으로 규제하려는 하지는 않지만 공급자나 배포자 등이 자율적으로 이 문제를 해결하기 위해 노력할 것을 유도하고 있다. AI Act 제95조는 법률상 요구사항에 대해 고위험AI시스템 외의 AI시스템에서도 적용될 수 있도록 각 회원국들이 이를 장려하여야 한다고 명시하고 있다(Article95

36) 이미 전통적 데이터 중심시장에서 데이터는 시장 지배력의 중요한 요소로 평가받고 있다(Cabral et al., 2021). 생성형 AI시스템에서 데이터는 경쟁우위를 확보하고 시장 지배력을 공고하게 하는 데 기여하는데, ChatGPT와 같은 일부 모델은 사용자 대화 데이터를 통해 학습함으로써 개선되며, 이는 모델의 사용자가 많을 수록 모델의 품질이 높아진다는 것을 시사한다(Carugati, 2024).

1.). AI Act에 의해 설치되는 AI사무소와 회원국들은 공급자, 배포자 등이 자발적으로 법률을 준수하기 위한 행동강령을 작성할 것을 촉진하여야 한다고 한다. 특히 AI 시스템의 에너지 효율적 프로그래밍 및 기술과 관련된 것을 포함하여 AI시스템이 환경적 지속가능성에 미치는 영향의 평가와 영향을 최소화하기 위한 핵심 성과지표를 만들고 이를 중심으로 자발적 행동강령을 만들고 준수할 수 있도록 촉진할 것을 요구하고 있다(Article95 2.(b)).

Ⅶ. 국내 생성형 AI 규제입법 및 정책 설계에 대한 제언

1. 인공지능 규제 논의의 특징과 방향성

1) 인공지능 규제의 방식과 방향성

해외 주요국은 인공지능의 효용은 극대화하고 역기능 등은 최소화하기 위한 수단으로 인공지능의 신뢰성 확보를 위한 법제도적 기반과 통제방안을 마련중에 있다. 인공지능에 관한 일반적 규제조치를 주요 골자로 하는 유럽연합의 'AI Act'이 대표적이다. 이 뿐만 아니라 미국과 캐나다 등에서도 인공지능에 대한 법적 규제의 필요성을 공감하면서 규제의 방향성을 제시하거나 필요한 조치들을 마련하고 있는 상황이다. 기본적으로 AI에 의한 위험으로 인해 개인의 권리침해가 심화되는 것에 대해 규제적 접근을 통해서 적극적 대응을 실시하고 있는 것으로 보인다. 이러한 입법적 조치의 핵심은 규제의 대상인 인공지능 기술을 법률적으로 정의하고, 통제의 대상인 위험을 식별하는 데 있다. 그리고 그러한 위험을 통제하기 위하여 인공지능 시스템 및 서비스의 개발, 공급, 유통하는 사업자 등에 대하여 관련한 의무조치를 부담시키는 것을 주요 내용으로 한다. 구체적으로는 위험관리체계를 수립하여 해당 시스템이 발생시킬 수 있는 위험을 통제하도록 하고, 표준화된 기술사양을 활용하도록 하거나 품질보증을 위해 필요한 검사나 테스트 등을 실시하여 품질을 유지 및 관리하도록 하는 의무도 부과한다. 데이터로부터 발생가능한 문제들을 최소화하기 위한 데이터를 관리하도록 하고 AI시스템이 의도된 목적에 따라 적절한 수준의 정확성과 견고성을 유지할 수 있도록 사이버보안을 강화할 것 또한 요구한다. AI시스템은 인간에 의해 통제되고 감독되어야 하며, 이용자에게 필요한 정보를 정확하게 제공하여 투명성을 확보하고 정보주체의 설명요구권을 적극적으로 보장하도록 하기도 한다.

AI의 법적 통제에 대한 가능성과 방법 등은 AI의 위험을 통제할 필요성이 제기된 이후 관련 논의가 지속되는 과정에서 구체화되고 그 방향성이 어느정도 정립되고 있는 것으로 보인다. 앞서 언급한 AI 규제 법률의 핵심적 내용들을 중심으로 규제체계를 설계하고 있다. 국내 입법안에서도 유사한 방식으로 접근하고 있는데, 대표적으로 2024년 발의된 '인공지능기술 기본법안'을 보면, '인공지능', '인공지능 기술', '고위험영역 인공지능'을 개념 정의하고, '고위험영역 인공지능' 관련 사업자는 위험관리 방안을 수립하고, 설명가능성 확보, 사람에 의한 관리 및 감독, 신뢰성 확보조치와 문서화 등의 의무를 부담시키고 있다.

2) 생성형 AI 규제 논의의 흐름과 쟁점사항

생성형 AI로 인한 문제점 내지는 이를 둘러싼 법적 쟁점들과 이에 대한 규제 논의들은 국가별로 약간의 차이는 존재하지만 큰 흐름은 유사한 방향으로 흐르고 있는 것으로 보인다. 일반적으로 AI의 위험성을 통제하는 방식을 통해 생성형 AI로 인한 문제들도 관리하고자 하면서 생성형 AI의 특성 때문에 발생하는 이슈들은 추가적으로 논의를 실시하고 있는 것으로 보인다. 즉, 생성형 AI에 의한 개인의 권리 침해를 방지하고 발생시키는 위험 등을 통제하기 위한 다양한 수단과 방법들을 강구해내고 있는데, 이를 종합 정리하면 〈표 1〉과 같다.

AI의 법적 통제 방식과 수단은 생성형 AI에서도 유사하게 작동하지만 위험의 정도가 일반적인 AI시스템보다 크고, 통제의 방향성을 결정하기 어렵다는 문제가 있다. 또한 그 방향성이 합의된다 하더라도 효과적 집행

〈표 1〉 생성형 AI의 법적 쟁점에 대한 국내외 주요 논의 동향 정리

〈Table 1〉 A summary of major domestic and international discussion trends on legal issues in Generative AI

Key Issues and Points of Contention	Trends in Relevant Legal Frameworks	Domestic Discussion
Issues Regarding the Use of Publicly Disclosed Personal Information	• (EU) Under the GDPR, even publicly disclosed personal information can only be processed if there is a legitimate legal basis. • (UK, France) If there is a "legitimate interest" and the data controller can demonstrate it, publicly disclosed personal information can be scraped for training generative AI models. • (Singapore) "Information disclosed by an individual" and "publicly available information" are classified as data that can be collected and used without consent. • (California, USA) Publicly disclosed information, excluding biometric data, is not considered personal information and is therefore not subject to legal protection. • (China) Personal information that has been voluntarily disclosed by the individual or lawfully disclosed is permissible for collection and use for the purpose of training generative AI models.	• (Supreme Court) When collecting and providing publicly disclosed personal information for commercial purposes without consent, the legality must be determined by weighing the interests of the data subject and the data processor. • (Personal Information Protection Commission) Information may be collected without consent within the scope where objective consent can be inferred, or where legitimate interests clearly outweigh the rights of the data subject.
Intensified Infringement of Data Subject's Rights and Challenges in Protection Privacy Violations due to Exposure and Inference of Personal Information	• (G7) Emphasizes the need for caution regarding the handling of personal data used in the training, validation, and testing of generative AI models, as well as in conversations with generative AI tools and the creation of AI-generated content. • (UK) Operators of websites hosting publicly accessible personal information are obligated to protect users' personal data from unlawful data scraping in accordance with data protection laws.	• (Personal Information Protection Act) Article 37-2 grants data subjects the right to reject decisions made solely by fully automated systems if such decisions significantly impact their rights or obligations. Furthermore, data subjects have the right to request explanations regarding automated decisions. • (Artificial Intelligence Industry Promotion and Trust Assurance Act, Bill No. 2200053) The draft bill establishes that the development or use of AI must adhere to the fundamental principle of ensuring individuals' right to self-determination over their personal information (Article 3).
Copyright Infringement Issues in the Training of Generative AI Models	• (G7) Calls on generative AI developers to protect personal data and intellectual property rights. • (EU) The AI Act requires general-purpose AI model providers to establish policies to comply with copyright laws and to prepare summaries of the content used in training. • (USA) A bill in the House of Representatives mandates that creators or significant modifiers of training datasets must notify the Register of Copyrights with summaries containing information on copyrighted works used in the training process. • (UK) The UK Parliament requires AI providers to make AI training data accessible for copyright holders to verify and to disclose the targets of web crawling.	• (Copyright Act Amendment) Proposes immunity from liability for copyright infringement during the data scraping process for AI training purposes.

〈표 1〉 계속

Key Issues and Points of Contention	Trends in Relevant Legal Frameworks	Domestic Discussion
Copyright Recognition Issues for Works Created by Generative AI	• (General Consensus) Copyright is not recognized for creations made by AI. • (USA) A report from the Congressional Research Service confirms that copyright is only recognized for works created by humans. • (USA) When registering the copyright of AI-generated content, only the portion that can be proven to have human contribution is eligible for copyright protection. • (Japan) A report by the Agency for Cultural Affairs mentions that AI-generated content requires "creative human contribution" to be recognized as a copyrightable work.	
Issues of Bias and Inequality	• (UNESCO) Published research findings suggest that generative AI has tendencies to induce biases related to gender and race, emphasizing the need for implementing AI ethics guidelines. • (EU) The AI Act mandates that during data training, validation, and testing, the potential for bias that could negatively impact fundamental rights or cause discrimination must be considered. • (USA) The "Blueprint for an AI Bill of Rights" outlines the necessity of testing all automated systems to ensure they are free from algorithmic bias before being sold or used. • (China) The "Interim Measures for the Administration of Generative AI Services" require effective measures to be taken during service provision to prevent discrimination.	• (Artificial Intelligence Industry Promotion and Trust Assurance Act, Bill No. 2200053) The draft bill establishes that the development and use of AI must not result in discrimination based on gender, age, ethnicity, religion, or social status (Article 3).
User Competence and Crime Issues	• (EU) The AI Act requires suppliers and distributors to take measures to strengthen "AI literacy." • (USA) The "Executive Order on the Development and Use of Safe and Trustworthy AI" directs the Secretary of Commerce to take steps to reduce the risks posed by generative AI synthetic content, such as implementing guidelines for content authentication, source tracing, banning the production of synthetic content like child pornography, and establishing methods for detecting and labeling synthetic content.	• (Content Industry Promotion Act Amendment, Bill No. 2200048) Proposes mandating the labeling of content created through AI technologies.
Monopoly and Environmental Issues	• (G7) Issued a joint statement warning against monopolistic practices in the generative AI field, emphasizing the need to strengthen the capabilities of competition authorities. • (EU) The AI Act calls for the assessment of AI systems' environmental impact, including energy-efficient programming and technology. It encourages the creation and adherence to a voluntary code of conduct centered on key performance indicators aimed at minimizing environmental impact.	

이 쉽지 않다는 한계도 존재한다. 따라서 규제 및 통제의 방향성을 설정하는 것도 중요하지만, 이를 실효적으로 집행하고 위험을 효과적으로 통제할 수 있는지에 대한 검증방안도 함께 고민이 필요한 시점이다.

2. 생성형 AI 규제 법률 및 정책의 설계시 고려사항

1) 규제 대상의 선별과 규제 실효성에 대한 고민

유럽연합은 생성형 AI 기술의 등장을 고려하여 법안 수정 과정에서 '범용 AI 모델'의 개념을 추가하였다.[37] 2021년 제안되었던 법안에서는 '인공지능시스템'을 "지도학습이나 비지도학습, 강화학습 등의 머신러닝 기법이 적용되거나 지식표현, 추론, 학습 등 요소기술이나 전문가시스템 등 응용방법이 적용된 기술이나 시스템"으로 특정하였다. 베이즈 추정이나 최적화 방법 등 통계적 접근 방법을 활용한 시스템 등을 개념 정의내에 포함시키는 등 활용되는 기술 자체를 특정하는 방식으로 접근하였다. 그러나 생성형 AI 등의 기술이 등장 및 확산되자 'AI시스템'의 개념을 "다양한 수준의 자율성으로 작동하도록 설계되고 명시적 또는 암묵적 목표에 대해 물리적 또는 가상환경에 영향을 미치는 예측, 추천, 결정 등과 같은 결과를 생성할 수 있는 기술기반시스템(Machine-based System)"으로 수정하였다. 생성형 AI 등이 등장하여 학습된 데이터를 기반으로 자율성을 갖고 새로운 정보나 이미지 등을 생성하는 방식으로 변화하고 있고, 일반인공지능(Artificial Gerneral Intelligence; AGI)과 같이 의도된 목적이나 목표없이 어떤 방식으로도 사용될 수 있는 AI 기술이나 서비스가 있다는 점 등을 감안한 것으로 보인다.

AI 기술은 발전 속도가 다른 기술에 비해 변화 속도가 빠르고, 다양한 방식으로 전개되고 있다는 점 때문에 법률로 규제대상을 특정하는 것이 매우 어렵다는 문제가 있다. 유럽연합에서 AI Act안이 처음 발의되었을 때 AI시스템의 개념이 모호하거나 광범위하다는 비판이 있었다. 유럽의회가 법률안에 대하여 전문가 의견을 수렴하는 과정에서 정의 규정의 광범위하다는 점이 지적되었고(European Parliamentary Research Service, 2021), Facebook, TechUK, Huawei, IBM 등 산업계에서도 정의규정이 광범위하다는 우려를 표명한 바 있다(TechMonitor, 2021). 유럽 집행위원회의 AI 윤리 및 인권 문제 등을 다루기 위해 시작된 프로젝트인 SHERPA 컨소시엄에서도 정의규정에 대한 문제를 비판하였었다(SHERPA, 2021).

이에 유럽연합 AI Act는 유럽집행위원회로 하여금 기술 변화에 따라 규제 대상을 특정하도록 하고, 의무조치를 실시하여야 하는 '고위험AI시스템'의 이용 사례와 법 집행의 구체적 가이드라인을 마련하도록 하였다(Article 6). 규제 대상이 모호하다는 점을 극복하기 위한 목적으로 보인다. 반면, 국내에서는 인공지능 기술에 대한 정의나, 관련 기술에 대한 개념을 구체화는 데 특별히 관심을 보이는 것으로 파악되지는 않는다. 대부분의 법률안에서는 인공지능 기술을 "학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것"등의 방식으로 정의를 할 뿐, 범용AI나 파운데이션모델과 같은 달리 규정이 필요한 개념에 대해서는 특별히 언급하지 않고 있다. 최근 입법안에서는 생성형 AI에 대하여 개념을 정의하고 있지만, 이러한 방식은 더 다양한 기술이 등장하거나 활용 방식이 달라지게 될 경우에는 규제적 조치를 강제할 수 어렵다는 한계가 존재한다.[38]

생성형 AI도 향후 기술의 사용 방식의 변화 양상이나 발생가능한 문제 상황, 생성형 AI 기술 외에 어떠한 기

37) '범용 AI 모델'은 시장 출시 방법 및 다양한 다운로드 시스템 또는 애플리케이션에 통합되는 방법에 상관없이 상당한 일반성을 가지며 광범위한 범위의 다양한 업무를 능숙하게 수행할 수 있는 AI 모델로서, 상당한 규모의 자기지도학습을 이용하여 대량의 데이터를 학습한 AI 모델을 포함한다. 다만 연구·개발 및 시제품 제작 활동을 위해 시장 출시 전에 이용되는 AI 모델은 포함하지 않는다(Article 3(63)). 유럽연합 AI Act는 '범용 AI 모델'중에서 구조적 위험(Sestemic Risk)이 있는 모델을 위험성이 높은 것으로 분류하고 기술문서 작성, 저작권 정책 수립, 훈련 콘텐츠의 요약본 작성 및 공개 등의 조치를 취하도록 하고 있다(Article 53).

술이 등장할 것인지 등을 예측하기는 어렵다. 유럽연합의 AI Act가 제안된 이후 약 2년이 흐르기 전에 생성형 AI가 등장하여 법안이 대폭 수정을 실시하였다는 점은 이러한 변화의 가능성을 방증한다고도 할 수 있다. 이는 규제적 조치들을 담고 있는 국내 인공지능 관련 법률안들이 통과될 경우 효과적으로 적용되기 어려울 수도 있다는 것을 의미한다. 더욱이 국내 법률안들은 인공지능 기술에 대한 위험 통제의 관점보다는 산업 진흥에 관심을 보이고 있고, 위험 통제를 위한 규제적 조치들이 면밀하게 검토된 것으로 보기에도 어려워 생성형 AI를 포함한 인공지능 기술의 위험을 실효적으로 관리할 수 있을 것인가 다소 의문이 든다. 생성형 AI에 관한 법적 쟁점들이 명확하게 나타나고 있고, 논의들이 활발하게 이루어지고 있는 만큼 기술 변화를 주목하고 발생하는 문제들을 효과적으로 컨트롤 할 수 있는 법적 수단과 방법에 대한 고민이 이루어질 필요가 있다.

2) 규제 집행기관의 전문성 확보

생성형 AI의 문제나 위험성을 법적 규제방식으로 통제하고자 할 경우 이를 집행하는 집행기관의 전문성과 역량을 강화하는 것이 필요하다. 어떤 규제적 조치가 적정하게 위험을 통제할 수 있는지를 결정하는 것에도 역량이 요구되지만 법률을 위반하여 위험을 유발시키고 확산시키었는지를 발견하고 입증하고 필요한 의무 조치를 부담시키기 위해서는 기술 및 서비스에 대한 이해가 수반되어야 하기 때문이다. 유럽연합 AI Act는 회원국들이 동 법을 집행할 수 있는 국가관할당국을 설립 또는 지정하도록 요구하면서, 관할당국은 AI의 기술, 데이터 및 데이터 컴퓨팅, 개인정보보호, 사이버보안, 기본권, 건강 및 안전 위험 등의 지식에 대한 심층적인 이해를 포함하는 역량과 전문성을 갖춘 충분한 수의 인력을 상시적으로 확보하도록 요구하고 있다. 그리고 회원국은 매년 이러한 역량과 자원 요건을 평가하고 필요한 경우 개정을 하도록 하고 있다(Article 70). 생성형 AI로 인한 위험이 현재 나타나는 양상에서 변화를 어떻게, 어느 정도의 수준으로 달리하게 될지 예측이 어려운 상황에서 규제적 조치들이 시행될 경우 이를 실효적으로 집행하고 위험을 효과적으로 통제하기 위해서는 규제 집행기관의 역할이 매우 중요하다. 모호한 법적 규정의 해석과 규제적 조치가 적용되는 대상에 대한 결정, 통제가 필요한 위험 유발 사례 등을 제시할 수 있어야 하기 때문이다. AI기술 및 서비스의 개발, 공급자 등에 대해 보다 명확한 기준을 제시하여 권리 침해 등의 문제 발생을 최소화하고 기술 및 산업 발전도 이룰 수 있도록 하여야 한다.

3) 이용자 역량의 고려

생성형 AI는 이용자와 상호작용하며 고도화된다는 특징이 있다. 따라서 이를 이용하는 이용자의 역량 문제에 대한 고려도 반드시 필요하다. '이용자들이 생성형 AI서비스를 범죄목적이나 수단으로 사용하는 등 부정적으로 이용하는 사례들이 지속적으로 나타나고 있는데, 이는 사업자에 대한 규제만으로 위험을 통제할 수 없다는 것을 의미한다. 따라서 규제적 입법들이 사업자를 통제하는 방식으로 설계된다 하더라도 생성형 AI로 인한 위험이나 문제들을 관리하는 데는 효과적이지 못할 가능성이 높다. 따라서 이용자에 대한 관리가 필요한데 이는 디지털 역량 교육, 인공지능 윤리 교육 등을 통해 실현하는 것이 가장 합리적일 것이다. 디지털 시민역량이란 '디지털 사회에서 시민이 갖추어야 하는 자질과 능력'으로 시민의 관점에서 새로운 기술을 적절히 이해하고 효율적으로 이용할 수 있는 능력, 필요한 정보를 분별하고 찾을 수 있는 능력, 정보를 통해 현안을 합리적으로 판단하고 의견을 형성할 수 있는 능력, 의견의 적극적 표현과 소통할 수 있는 능력 등이 포함된다. 디지털 시민역량은 인공지능서비스의 이용자들이 해

38) 국내 법률안에서는 법 적용의 대상이 되는 '고위험영역 인공지능'을 해당되는지를 확인하고자 할 경우 과학기술정보통신부장관에게 확인을 요청할 수 있도록 하고 있는데, 규제대상 여부가 명확하지 않을 때를 대비한 것으로 보인다(의안번호: 2200053, 법률안명: 인공지능 산업 육성 및 신뢰 확보에 관한 법률안, 발의자: 안철수 의원 등 12인, 발의일: 2024.5.31. 안 제26조).

〈표 2〉 한국의 AI 관련 법률안에서 정의하고 있는 'AI'의 개념
〈Table 2〉 The definition of 'AI' as defined in Korea's AI-related Bill

Bill Title	Definitions
Bill on the Development of Artificial Intelligence and Establishment of Trust (Bill No.: 2200543)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as the related application technologies. • "Generative Artificial Intelligence" refers to AI designed to generate outputs such as text, sound, images, or video, at various levels of autonomy.
Bill on the Promotion of the Artificial Intelligence Industry and Assurance of Trust (Bill No.: 2200673)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as the related application technologies.
Bill on the Promotion of the Artificial Intelligence Industry and Assurance of Trust (Bill No.: 2200675)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as related application technologies.
Bill on the Basic Act on Artificial Intelligence (Bill No.: 2203072)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as related application technologies. • "Generative Artificial Intelligence" refers to AI designed to generate outputs such as text, sound, images, or videos at various levels of autonomy.
Bill on the Basic Act on Artificial Intelligence Technology (Bill No.: 2201158)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as related application technologies. • "Generative Artificial Intelligence" refers to AI designed to generate outputs such as text, sound, images, or videos at various levels of autonomy.
Bill on the Promotion of the Artificial Intelligence Industry and Assurance of Trust (Bill No.: 2200053)	• "Artificial Intelligence" refers to the electronic implementation of intellectual abilities possessed by humans, such as learning, reasoning, perception, judgment, and language comprehension, with autonomy to adapt to external environments or inputs. • "Algorithm" refers to a system composed of a set of calculations, rules, procedures, or logical commands described for solving problems, performing tasks, or operating equipment, devices, or machines. • "Artificial Intelligence Technology" refers to the hardware technology necessary to implement artificial intelligence, or the software technology that systematically supports it, as well as related application technologies. • "Generative Artificial Intelligence" refers to AI designed to generate outputs such as text, sound, images, or videos at various levels of autonomy.

당 기술의 위험성을 인지하고 안전하고 올바르게 이용하는 것은 물론 서비스의 목적에 맞게 활용하여 효용성을 극대화하는 데 기여할 수 있으므로 이러한 역량 강화 정책을 강화하고 지속적으로 실시할 필요가 있다. 시민 역량의 문제는 단기적이고 국소적 법정책 과제라기 보다 장기적 관점에서 지속적으로 국가 전반의 교육제도 전반의 개혁을 통해서 실현이 가능하다. 유럽연합의 AI Act도 2023년 수정안에서 AI리터러시[39]에 관한 규정을 새롭게 포함시켰는데, 해당 규정은 유럽연합과 회원국에 대해 교육 및 훈련, 숙련 및 재교육 프로그램을 통해 AI시스템에 대한 민주적 통제를 허용한다는 관점에서 적절한 성별, 연령 균형을 보장하면서 모든 부문에 걸쳐 공급업체, 배포자 및 관련자 그룹의 다양한 요구를 고려하여 충분한 수준의 AI 활용 능력을 개발하기 위한 조치들을 실시하도록 하고 있다.

Ⅷ. 결론

생성형 AI를 포함하는 AI 기술 및 서비스는 효용성, 생산성, 혁신성이라는 특징으로 인해 발전을 거듭할수록 인간의 삶에 긍정적 변화를 가져다 줄 것으로 기대된다. 경제적·산업적 발전 가치 또한 다른 분야나 산업에 비해 가능성이 큰 기술로 전 세계가 발전 가능성에 관심이 높은 상황이다. 한편, 생성형 AI의 등장 이후 AI가 야기하는 위험성에 대한 우려가 점점 커지고 있고 일부는 현실화되고 있기도 하다. 이에 각 국의 정부는 AI의 효용가치는 극대화하면서 국가적 혁신과 국민 삶의 질을 향상시키기 위해 AI기술이 그 역할을 충분히 할 수 있도록 기술과 산업을 발전시키고자 노력을 하고 있다. 이와 더불어 AI로 인해 발생하는 위험으로 인하여 개인의 권리가 침해되지 않도록 AI의 위험성을 적정히 통제할 수 있는 방안들을 모색하는 방안들을 고민하고 있다. 앞서 살펴본 바와 같이 데이터 학습 과정에서의 저작권 및 개인정보 침해 문제와 이에 대한 책임 부담 여부, 새로운 기술들이 범죄 수단으로 악용되는 상황 등 생성형 AI로 인한 여러 법적 문제들에 대한 해답을 찾고자 노력한다. 국내에서도 AI에 대한 규제와 법적 통제의 필요성, AI로 인해 발생하는 법적 쟁점들에 대한 논의들이 다양하게 제기되고 있는 상황이지만 개별 이슈나 쟁점에 있어 구체적인 제도 설계 방안이나 방향성 등에 대해서는 아직 국가적 합의에 이룬것으로 보기는 어려운 상황이다. 특히 AI에 대한 규제적 관점의 법률안의 경우에도 AI로 인하여 발생하는 위험의 내용과 정도를 평가하는 것은 물론 AI로 인한 위험성을 통제하는 실효적인 방법이 제시되어 있다고 볼 수도 없다. 해외 주요 국가들이 AI에 대한 통제수단을 세밀하게 구조화하고 있는 반면, 우리나라의 경우 AI로 인한 우려를 지속적으로 제기하고 있는 것에 비해 논의는 여전히 답보상태라고 할 수 있다. 위험이 구체화되지 않은 상황에서 섣부른 규제는 기술과 산업의 발전을 저해하고 관련 시장에 혼란을 야기할 가능성도 존재하지만, 기술의 발전과 긍정적 효과는 기술의 위험성과 역기능을 제거하는 것을 전제로 한다는 점도 유념할 필요가 있다.

AI 기술의 법과 제도는 기술의 효용가치를 극대화하고 위험성을 최소화하여야 하는 도전과제에 직면하여 있다. AI 기술에 대한 기대가 크고 국가 산업 발전의 새로운 전환점이 될 것으로 예측되는 만큼 발생하는 위험을 최소화하고 안전하게 활용될 수 있도록 기반을 잘 마련하는 것이 중요하다. 이러한 기반형성은 국내의 사회적·문화적·산업적 특성을 고려하여 세밀하게 이루어져야 할 것이다. 생성형 AI를 포함한 AI와 관련한 법적 문제들과 쟁점들에 대한 국제적 동향을 확인하고 관련 논의를 구체화하는 작업을 지속화하여야 할 것이다.

39) 'AI 리터러시'란 이 법의 맥락에서 공급자, 배포자 및 영향을 받는 사람이 각자의 권리와 의무를 고려하고 정보에 입각하여 AI 시스템을 배포할 수 있도록 하고, AI의 기회와 위험, 발생할 수 있는 피해에 대해 인식할 수 있도록 하는 기술, 지식 및 이해를 의미한다(Article 3.(56)).

■ References

Abenezer G., Kidus M., Amit P., Anushka S., Vikas H., Vinay C. & Biplab S. (2024). "Privacy and Security Concerns in Generative AI: A Comprehensive Survey." *IEEE Access* Vol.12. 48126-48114.

Ahn, S., Yoo J., Cho W., Noh J. & Son, H. (2023). "The Rise of the Supergiant Language Model and the Technical Characteristics and Social and Industrial Implications of ChatGPT." *Issue Report IS-158*, Gyeonggi: Software Policy & Research Institute.

{안성원·유재흥·조원영·노재원·손효현 (2023). 초거대언어모델의 부상과 주요이슈-ChatGPT의 기술적 특징과 사회적·산업적 시사점. 〈Issue Report〉 IS-158, 경기: 소프트웨어정책연구소.}

AiTIMES (2023). "Google succeeded in extracting personal information from 'Chat GPT'…"Can identify LLM training data." December 1.

{AiTIMES (2023). "구글, '챗GPT'에서 개인정보 추출 성공…"LLM 훈련 데이터파악 가능'" 12월 1일.}

AiTIMES (2023). "GPT-4, 5 things that changed from before." March 15.

{AiTIMES (2023). "GPT-4, 이전과 달라진 점 5가지는." 3월 15일.}

AiTIMES (2024). "OpenAI unveils stage 5 to AGI… "We're on the verge of stage 2." July 12.

{AiTIMES (2024). "오픈AI, AGI로 가는 5단계 공개…"우리는 2단계 직전"", 7월 12일.}

Alexander, L., Martin, K. & Marc-Oliver P. (2024). "Blessing or curse? A survey on the Impact of Generative AI on Fake News." arXiv. https://arxiv.org/abs/2404.03021.

Alexandra, S. L., Yacine, J. & Emma, S. (2024). "Power Hungry Processing: Watts Driving the Cost of AI Deployment?." FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, 85-99.

Bianchi, F., Kalluri, P., Durmus, E., Ladhak F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J. & Caliskan, A. (2023). "Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale." FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. 1493 - 1504.

Bloomberg Law (2023). "Getty Images Sues Stability AI Over Art Generator IP Violations" February 7.

Byun, S. (2023)."Ethical Problems of Super-Massive Generative AI." *Journal of AI Humanities, 14*, 63-82.

Cabral, L., Haucap, J., Parker, G., Petropoulos, G., Valletti, T. & Van Alstyne, M. (2021). "The EU Digital Markets Act." *JRC* 122910. DOI: 10.2760/139337

Carolina P. & Lorenzo P. (2024). "Artificial Intelligence and Cybersecurity." *Intereconomics, 59*(1), 10-13.

Carugati, C. (2024). "The Generative AI challenges for competition authorities." *Intereconomics, 59*(1), 14-21.

Chan, J. (2024). "How did OpenAI scrape the internet?" https://www.askhandle.com/blog/how-didopen-ai-scrape-the-internet. (Retrieved on August 2).

CheckPoint (2023). "Cybercriminals Starting to Use ChatGPT."https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt/. (Retrieved on August 17).

CNIL (2024). "AI How-to sheets(Ensuring the lawfulness of the data processing - Defining a legal basis)." https://www.cnil.fr/fr/node/164399. (Retrieved on August 5).

Congressional Research Service (2023). "Generative Artificial Intelligence and Copyright Law." https://crsreports.congress.gov/product/pdf/LSB/LSB10922. (Retrieved on August 3).

DeepMind (년도) "Tackling multiple tasks with a single visual language model." https://www.deepmind.com/blog/tackling-multiple-tasks-with-a-single-visual-language-model. (Retrieved on August 5).

EDPB (2023). "EDPB resolves dispute on transfers by Meta and creates task force on Chat GPT." https://edpb.europa.eu/news/news/2023/edpb-resolves-dispute-transfers-meta-and-creates-task-force-chat-gpt_en. (Retrieved on August 15).

EDPB (2024). "Report of the work undertaken by the ChatGPT Taskforce." https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en. (Retrieved on August 2).

EDPS (2023). "G7 DPA Roundtable." https://www.edps.europa.eu/data-protection/our-work/publications/international-conferences/2023-06-21-g7-dpa-roundtable_en. (Retrieved on August 3).

European Commission IP Helpdesk (2023). "Intellectual Property in ChatGPT" Feburary 20.

Federal Register (2023). "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence." https:/www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence. (Retrieved on August 3).

Feuerriegel Stefan, Jochen Hartmann, Christian Janiesch & Patrick Zschech. (2024). "Generative ai." *Business & Information Systems Engineering 66*(1), 111-126.

Forbes (2023). "FTC Concerned About Generative AI Monopolies." June 30.

Forbes (2023). "Six Risks Of Generative AI." June 29.

FTC (2023). "Generative AI Raises Competition Concerns." https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns. (Retrieved on August 7).

G7 2023 Hiroshima Summit (2023). "Compendium of approaches to improving competition in digital markets." G7 2023 Hiroshima Summit.

G7 Data Protection and Privacy Authorities Roundtable (2023). "Roundtable of G7 Data Protection and Privacy Authorities Statement on Generative AI." G7 Data Protection and Privacy Authorities Roundtable.

Garante Per La Protezione Dei Dati Personali (2023). "Provvedimento del 30 marzo 2023 [9870832]: IL GARANTE PER LA PROTEZIONE DEI DATI PERSONALI." https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870832. (Retrieved on August 5).

Goldman Sachs (2023). "Generative AI could raise global GDP by 7%." April 5.

Google (2024). *2024 Environmental Report*. Google.

Gupta, M., Akiri, C., Aryal, K., Parker, E. & Praharaj, L. (2023). "From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy." *IEEE Access*, 11, 80218-80245.

Hacker, P., Mittelstadt, B., Borgesius, F. Z. & Wachter, S. (2024). "Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It." https://arxiv.org/abs/2407.10329.

Han, J. (2023). "The Beginning of the Generative AI Era." *Media Issue & Trend, 55*. 6-17.

{한정훈 (2023). 생성형 AI시대의 개막. 〈Media Issue & Trend〉, 55, 6-17.}

Hong, E. (2023). "Beyond ChatGPT, Future of Generative AI (Generation AI) -Part 1."SAMSUNG SDS Insight Report. https://www.samsungsds.com/kr/insights/future_of_generative_ai_1.html, (Retrieved on August 9).

{홍은주 (2023). "ChatGPT를 넘어, 생성형 AI(Generative AI)의 미래-1편." SAMSUNG SDS 인사이트리포트. https://www.samsungsds.com/kr/insights/future_of_generative_ai_1.html. 검색일: 2024.08.09.}

IBM (2024). "What is a prompt injection attack?" https://www.ibm.com/kr-ko/topics/prompt-injection. (Retrieved on August 7.)

{IBM (2024). "프롬프트 인젝션 공격이란 무엇인가요?" https://www.ibm.com/kr-ko/topics/prompt-injection. (검색일: 2024.08.07.)}

ICO (2023). Joint statement on data scraping and the protection of privacy. ICO.

ICO (2024). "Generative AI first call for evidence: The lawful basis for web scraping to train generative AI models." https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/generative-ai-first-call-for-evidence/. (Retrieved on August 5).

IEA (2024). *Electricity 2024 Analysis and forcast to 2026*. IEA

Jeon, E. (2023). Personal Information Protection Issues Raised by Generative Artificial Intelligence Models such as ChatGPT. *Regulatory Legislation Review, 23*(3). Sejong: The Korea Institute of Legislation. 111-146.

{전응준 (2023). ChatGPT등 생성형 인공지능 모델이 제기하는 개인정보 보호 관련 쟁점. 〈규제법제 리뷰〉, 23권 3호, 111-146.}

Jeong, C. (2023). "A Study on the Copyright Protection for Artificial Intelligence System." *The Journal of Law, 31*(2), 179-202.

{정충원 (2023). 인공지능(Artificial Intelligence:AI) 시스템에 대한 저작권보호에 관한 연구. 〈법학연구〉, 31권 2호, 179-202.}

Jung W. (2019). "Legal Issues and Policy Challenges on the Protection of Artificial Intelligence Creatures." *Information and Communication Broadcasting Policy, 31*(6), 1-27.

{정원준 (2019). 인공지능 창작물의 보호에 관한 법적 쟁점과 정책적 과제. 〈정보통신방송정책〉, 31권 6호, 1-27.}

Jungang Ilbo (2021). "[Factple] Learned human bias as it is, hate-spouting AI 'Iruda Shock'." January 12.

{중앙일보 (2021). "[팩플]인간의 편견 그대로 배웠다, 혐오 내뱉는 AI '이루다 쇼크'". 1월 12일.

Kim, B. (2022). "Legal Challenges in Large-Scale Language Models." *Journal of Korea Infomation Law, 26*(1), 173-217.

{김병필 (2022). 대규모 언어모형 인공지능의 법적 쟁점. 〈정보법학〉, 26권 1호, 173-217.}

Kim, H. (2023)."Could Data Lie?: Generative AI and Bias." *IP & Data Law, 3*(1), 75-105.

{김현진 (2023). 생성형 AI와 편향성. 〈IP & Data 法〉, 3권 1호, 75-105.}

Kim, J. (2023). "Generative AI, How far have we come (1) Basic concepts and development status." https://campaigns.do/discussions/598, (Retrieved on August 5).

{김재경 (2023). "생성형 AI, 어디까지 왔는가(1) 기본 개념과 개발 현황."https://campaigns.do/discussions/598, 검색일: 2024.08.05.)}

Kim, Y. (2016)."The Possibility to Protect Creative Works of ArtificialIntelligence." *Chungnam Law Review, 27*(3). 267-297.

{김용주 (2016). 인공지능(AI; Artificial Intelligence) 창작물에 대한 저작물로서의 보호가능성. 〈충남대학교 법학연구소 법학연구〉, 27권 3호, 267-297.}

Kim, Y. (2023). "Legal Issues in Generative Artificial Intelligence Models: Focusing on the discussion of ChatGPT Products." *Journal of Korea Infomation Law, 27*(1), 77-112.

{김윤명 (2023). 생성형 인공지능(AI) 모델의 법률 문제. 〈정보법학〉, 27권 1호, 77-112.}

Lee, S. (2023). "Issues and challenges in data law due to technological innovation." *Journal of Law & Economic Regulation, 16*(2), 84~106.

{이성엽 (2023). 기술혁신에 따른 데이터법의 이슈와 과제. 〈경제규제와 법〉, 16권 2호. 54-106.}

Liang, W., Yuksekgonul, M., Mao, Y., Wu E. & Zou, J. (2023). "GPT detectors are biased against non-native English writers." *Patterns, 4*(7). https://arxiv.org/abs/2304.02819.

Mark, D. (2023). "Chat GPT and the EU AI Act: Will New Regulations be Added?" *Denver Journal of International Law & Policy*. https://djilp.org/chat-gpt-and-the-eu-ai-act-will-new-regulations-be-added/.(Retrieved on August 17).

McAfee (2023). "Artificial Intelligence Voice Scams on the Rise with 1 in 4 Adults Impacted." https://www.mcafee.com/ko-kr/consumer-corporate/newsroom/press-releases/press-release.html?news_id=5aa0d525-c1ae-4b1e-a1b7-dc499410c5a1&PR=/PressMedia/09201999-C.asp&Sel=583. (Retrieved on August 3).

MIT Technology Review (2020). "The way we train AI is fundamentally flawed."November 18.

MIT Technology Review (2022). "We're getting a better idea of AI's true carbon footprint." November 14.

Jovanović, M. & Campbell, M. (2022). "Generative Artificial Intelligence: Trends and Prospects, Computer." *Computer, 55*(10). 107-112.

Nature (2024). "AI image generators often give racist and sexist results: can they be fixed?" https://www-nature-com-ssl.oca.korea.ac.kr/articles/d41586-024-00674-9. (Retrieved on August 5).

NIST (2022). "Towards a Standard for Identifying and

Managing Bias in Artificial Intelligence." (Retrieved on May 15).
NIST (2024). "Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations."*NIST Trustworthy and Responsible AI NIST AI* 100-2e2023.
NIST (2024). "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile." *NIST Trustworthy and Responsible AI NIST AI* 600-1.
Novelli, C., Casolari, F., Hacker., Spedicato, G. & Floridi, L. (2024). "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity." arXiv. https://arxiv.org/abs/2401.07348
Noy, S. & Zhang, W. (2023). "Experimental Evidence on the Productivity Effects of generative Artificial Intelligence." *Science, 381*(6654). 187-192.
NVIDIA (2022). "What is a Transformer Model?(1)." https://blogs.nvidia.co.kr/2022/04/01/what-is-a-transformer-model/, (Retrieved on August 9).
{NVIDIA (2022). "트랜스포머 모델이란 무엇인가?(1)." https://blogs.nvidia.co.kr/2022/04/01/what-is-a-transformer-model/. 검색일: 2024.08.09.}
OECD (2007). "Glossary of statistical terms." https://stats.oecd.org/glossary/detail.asp?ID=3605. (Retrieved on August 7).
Office of the Privacy Commissioner of Canada (2023). "OPC to investigate ChatGPT jointly with provincial privacy authorities." https://www.priv.gc.ca/en/opc-news/news-and-announcements/2023/an_230525-2/. (Retrieved on August 7).
OneTrust DataGuidance (2024). "UK: Guidance on generative AI - the legal basis for scraping data." https://www.dataguidance.com/opinion/uk-guidance-generative-ai-legal-basis-scraping-data. (Retrieved on August 3).
OpenAI. "DALL · E2." https://openai.com/product/dall-e-2. (Retrieved on August 5).
OpenAI—written evidence (LLM0113) (2024). https://committees.parliament.uk/writtenevidence/126981/html/. (Retrieved on August 5).
Park T. (2023). *Park Tae-woong's lecture on AI*. Seoul: Hanbit Biz
{박태웅 (2023). 〈박태웅의 AI강의〉. 서울: 한빛비즈.}
Park, D. & Lee, H. (2024). "Literature Review of AI Hallucination Research Since the Advent of ChatGPT: Focusing on Papers from arXiv." *Informatization Policy, 31*(2), 3-38.
{박대민·이한종 (2024). 챗GPT 등장 이후 인공지능 환각 연구의 문헌 검토: 아카이브(arXiv)의 논문을 중심으로. 〈정보화정책〉, 31권 2호, 3-38.}
Park, W., Kim, M., Park, Y., Ryu, H. & Jeong, D. (2024). "Taxonomy and Countermeasures for Generative Artificial Intelligence Crime Threats." *Journal of the Korea Institute of Information Security & Cryptology, 34*(2), 301-321.
{박우빈·김민수·박윤지·유혜진·정두원 (2024). 생성형 인공지능 관련 범죄위협 분류 및 대응 방안. 〈정보보호학회 논문지〉, 34권 2호, 301-321.}
Personal Information Protection Commission (2023). *Policy Direction for Safe Use of Personal Information in the Age of Artificial Intelligence*. Seoul: Personal Information Protection Commission.
{개인정보보호위원회 (2023). 〈인공지능 시대 안전한 개인정보 활용 정책방향〉.}
Personal Information Protection Commission. (2024). *Public Personal Information Processing Guide for Artificial Intelligence (AI) Development and Services*. Seoul: Personal Information Protection Commission.
{개인정보보호위원회 (2024). 〈인공지능(AI)개발·서비스를 위한 공개된 개인정보 처리 안내서〉.}
Hacker, P., Engel, A. & Mauer, M. (2023). *Regulating ChatGPT and other large generative AI models*. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. DOI:10.48550/arXiv.2302.02337
Reuters (2023). "Getty Images lawsuit says Stability AI misused photos to train AI." Feburary 7.
Reuters (2024). "Microsoft, OpenAI plan $100 billion data-center project, media report says." May 30.
Reuters (2024). "OpenAI's ChatGPT targeted in Austrian privacy complaint." April 29.
Ryu, S. (2023). "A Study on Introducing a Specific Copyright Exception Clause to Text and Data

Mining." *Advanced Commercial Law Review, 101*. 347-390.
{류시원 (2023). 저작권법상 텍스트·데이터 마이닝(TDM) 면책규정 도입 방향의 검토. 〈선진상사법률연구〉, 101, 374-390.}
SBS News (2022). "[Exclusive] "AI Creation Can't Give You Copyright"…Domestic AI Copyright Conflict Ignition." October 14.
{SBS 뉴스 (2022). "[단독] "AI 창작물, 저작료 못 줘"…국내AI 저작권 갈등 점화." 10월 14일.}
Thome, S. (2024). "Understanding the Interplay between Trust, Reliability, and Human Factors in the Age of Generative AI." *International Journal of Simulation: Systems, Science & Technology*. 10.1-10.5.
Son, S. (2016). "Copyright Protectionon Artificial Intelligence(AI) generated Works." *Journal of Korea Infomation Law, 20*(3), 83-110.
{손승우 (2016). 인공지능 창작물의 저작권 보호. 〈정보법학〉, 20권 3호, 83-110.}
Son, Y. (2023). "A Study on Creation by Generative AI and Copyright." *Journal of Law and Politics research, 23*(3). 357-389.
{손영화 (2023). 생성형 AI에 의한 창작물과 저작권. 〈법과 정책연구〉, 23집 3호. 357-389.}
Son, Y. (2024). "Generative artificial intelligence and personal information protection." *Yonsei Law Review, 34*(2). 351~382.
{손영선 (2024). 생성형 인공지능과 개인정보보호. 〈연세대학교 법학연구원 법학연구〉, 34권 2호, 351-382.}
Song, M. & Lee, S. (2024). "What Concerns Does ChatGPT Raise for Us?: An Analysis Centered on CTM (Correlated Topic Modeling) of YouTube Video News Comments." *Informatization Policy, 31*(1), 3-31.
{송민호·이수범 (2024). ChatGPT는 우리에게 어떤 우려를 초래하는가?: 유튜브 영상 뉴스 댓글의 CTM (Correlated Topic Modeling) 분석을 중심으로. 〈정보화정책〉, 31권 3호, 3-31.}
Stanford University Human-Centered Artificial Intelligence (2023). "AI-Detectors Biased Against Non-Native English Writers." https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers.(Retrieved on August 5).
The NewYorkTimes (2023). "How Strangers Got My Email Address From ChatGPT's Model." December 22.
The NewYorkTimes (2024a), "U.S. Clears Way for Antitrust Inquiries of Nvidia, Microsoft and OpenAI."June 5.
The NewYorkTimes (2024b). "Regulators Take on the Giants of A.I." June 6.
The White House (2023). "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/. (Retrieved on August 8).
The White House (2024). "A Call to Action to Combat Image-Based Sexual Abuse." https://www.whitehouse.gov/gpc/briefing-room/2024/05/23/a-call-to-action-to-combat-image-based-sexual-abuse/. (Retrieved on August 7).
UK Parliament (2024). "Communications and Digital Committee Large language models and generative AI, 1st Report of Session 2023-24." https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/5402.htm. (Retrieved on August 5).
UN WOMEN. (2023). "Glossary: Gender and Technology." https://www.unwomen.org/en/how-we-work/innovation-and-technology/glossary?gclid=EAIaIQobChMIjZe1p8zs_gIVWTBgCh3C1QAdEAAYASAAEgLYdvD_BwE. (Retrieved on August 7).
UNESCO (2024a), "Challenging systematic prejudices: an investigation into bias against women and girls in large language models." https://unesdoc.unesco.org/ark:/48223/pf0000388971. (Retrieved on August 3).
UNESCO (2024b). "Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes." https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes. (Retrieved on August 3).

Yang, J. & Yoon, S. (2023). "Going beyond ChatGPT to the era of Generative AI: Media and Content Generative AI Service Case and How to Secure Competitiveness." *Media Issue & Trend Domestic Report, 55*, 62-70.

{양지훈·윤상혁 (2023). ChatGPT를 넘어 생성형(Generative) AI 시대로: 미디어·콘텐츠 생성형 AI서비스 사례와 경쟁력 확보 방안. 〈Media Issue & Trend Domestic Report〉, 55, 한국방송통신전파진흥원, 4면, 62-70.

Yonhap News (2016). "Danger of Artificial Intelligence Brainwashing…MS Chatbot 'Tay' Suspended due to discriminatory remarks." March 25.

{연합뉴스 (년도). "인공지능 세뇌의 위험…MS 채팅봇 '테이' 차별발언으로 운영중단(종합2보)." 3월 25일.}

Yoo, J., Ahn, S., Kim, J., Ahn, M., Jang, J., Bong, K. & Noh, J. (2023). "The Rise of Generative AI and the Change of Industry." *ISSUE REPORT* IS-161, Gyeonggi: Software Policy & Research Institute.

{유재흥·안성원·김정민·안미소·장진철·봉강호·노재원 (2023). "생성AI의 부상과 산업의 변화." 〈ISSUE REPORT〉 IS-161, 경기: 소프트웨어정책연구소.}

文化審議会著作権分科会法制度小委員会 (2024). "AIと著作権に関する考え方について." March 15.

文化庁 (2023). "AIと著作権". 令和5年度著作権セミナー. May 6.

朝日新聞デジタル (2023). "生成AI設計にプライバシー配慮を G7の個人データ保護当局が声明." June 21.

知的財産戦略本部 (2017). 知的財産推進計画2017.