

SW·저작권 동향리포트

〈제2023-8호〉 2023년 6월 25일

정책/제도

생성형 AI와 법적·윤리적 쟁점

■ 개요

- 세계적으로 올해의 가장 큰 이슈는 생성형(Generative) AI로 오픈 AI가 개발한 챗GPT를 시작으로 구글의 바드(Bard)와 MS의 Bing, 구글의 검색서비스에도 해당 서비스가 도입됨으로써 현재까지도 전 세계 AI 기술 산업에 큰 영향을 미치고 있음
- 과거의 인공지능 서비스는 다양한 데이터를 기반으로 추론 가능한 선택지를 제시하여 왔으나, 현재의 AI는 보다 현실적인 결과물을 제안하면서 인간의 의도를 파악해 결과를 도출시킨다는 점에서 기존 AI 모델과 차이를 가짐
 - * 기존 AI 모델인 달리(Dall-E)와 미드저니(Midjourney)도 입력창인 프롬프트에서 자연어로 입력한 내용을 파악해 그에 따른 결과물 도출이 가능했지만, 결과물로서 이미지를 만들어낸다는 점에서 텍스트 형태의 콘텐츠를 생성하는 챗GPT와 차이점 존재
- 생성형 AI의 영향으로 최근 출간되는 책과 그림, 디자인 등에서는 생성형 AI의 기술이 폭넓게 활용되고 있으며, 이렇게 만들어진 콘텐츠들은 AI의 학습과정에서 사용되는 데이터의 저작권 문제 및 생성된 콘텐츠의 법적·윤리적 문제가 부각되고 있음. 각 국가에서도 생성형 AI와 관련된 다양한 법적 분쟁 및 소송이 등장하고 있어 규제에 대한 목소리가 높아지고 있는 가운데 생성형 AI가 가져오는 산업적 영향과는 별개로 저작권과 관련된 법제도적 문제점을 고찰할 필요가 있음
 - * 최근에는 학교 과제, 논문, 연설문, 보고서 등의 작성에도 생성형 AI인 챗GPT가 사용되면서 대필에 의한 악용의 문제가 심화되고 있어, 국제머신러닝학회는 AI도구를 활용한 논문 작성을 제한하고, 네이처 및 사이언스 등의 과학학술지들도 논문에 사용 시에는 반드시 AI를 사용하였다는 출처를 표기해야 할 것을 제시하고 있음

■ 주요내용

1. 생성형 대규모 언어 모델의 등장

- 대규모 언어 모델(Large-scale Language Model, LLM)은 인터넷에 공개된 수많은 대규모 텍스트 데이터 세트를 바탕으로 학습하여 언어에 대한 다양한 지식을 갖고 다양한 언어와 문맥을 이해하고 생성하는 작업 수행이 가능. 전통적인 AI 모델은 학습 데이터를 입력해 데이터의 속성을 파악한 뒤 입력된 데이터가 어떠한 데이터인지 추론해 결과 데이터를 출력하는 방식이었으나, 언어 모델은 학습한 텍스트를 바탕으로 다른 텍스트를 통계적인 확률로서 예상하는 방식을 사용
- 오픈AI에서 개발한 생성형 AI인 챗GPT는 인터넷에 공개된 대규모의 데이터를 바탕으로 학습 과정을 거쳐 만들어진 Generative Pre-trained Transformer 모델로서 구글의 트랜스포머 모델에서 분화된 모델. GPT는 트랜스포머와 비지도 학습(Unsupervised Learning basics)을 결합한 것으로 별도의 레이블 된 데이터 없이도 학습이 가능
- AI 모델은 크롤링(Crawling) 등의 방식으로 대량의 데이터를 수집하거나, 공개된 정보를 데이터셋으로 구축해 지속적으로 학습하고 업그레이드하므로, 챗GPT에 프롬프트 형태의 명령어를 문장 또는 단어 등으로 입력하면 GPT가 명령어의 맥락을 분석해 그에 맞는 결과를 생성
- 구글의 CEO인 선다 피차이는 챗GPT를 심각 단계인 코드레드(Code Red) 수준으로 보고, 구글 검색엔진에 미칠 수 있는 상황을 모니터링 중이라고 밝힌 바 있으나 지난 5월 구글의 채팅형 AI인 바드를 공식적으로 출시. 바드 역시 다른 AI 모델에 공통되는 한계이자 문제인 데이터의 저작권 처리가 불분명하다는 점을 살펴볼 때 법률적인 분쟁이 예상됨
 - * 소프트웨어 분야에서 Code Red는 즉각적인 주의가 필요한 치명적인 소프트웨어 버그 또는 시스템 오류를 설명하는 데 사용되며, 추가 손상을 방지하거나 정상 작동을 복원하기 위해 신속하게 해결해야 하는 심각한 문제가 있음을 나타냄

2. 대규모 언어 모델의 윤리적 쟁점

- AI 모델은 공개된 데이터를 기반으로 학습하기 때문에 인간의 다양한 오류나 문제점을 판단하지 않고 답습함으로써 편향적이고 차별적인 결과를 가져올 수 있으므로 AI 모델을 교육하는데 사용되는 초기 데이터를 신중하게 선택하는 것이 중요함. 챗GPT는 초창기에는 정확하지 않거나

부적절한 답변으로 문제점이 제기되면서 현재는 내용 자체를 필터링하고 있지만 모든 내용을 필터링 하지는 못하여 여전히 부적절한 내용이 포함될 가능성이 있음

○ 챗봇 기능은 인터넷에 게시된 방대한 양의 텍스트를 분석해 만들어지기 때문에 사실과 허구를 구분할 수 없으며 여성과 유색 인종에 편향된 텍스트를 생성할 가능성이 있음. 챗GPT도 인터넷의 다양한 정보를 바탕으로 학습하고 스스로가 정보에 대한 진정성을 확인할 능력이 없으므로 언어적인 체계에서의 답변은 가능해도 문맥과 의미의 참과 거짓을 혼동하며 환각효과(Hallucination)에 빠질 우려가 있음

○ 미드저니(midjourney)와 같은 이미지 생성 모델은 다양한 이미지를 만들어 주지만 학습 데이터에 음란물이 포함되었는지 여부를 확인하기 어려워 음란물 생성에 오용될 가능성이 존재. 특히 이미지 모델이 작동하는 End-of-End 방식에서는 이용자에 의해 AI 모델이 오염되지 않도록 시스템을 운용할 필요가 있으며, 인간의 의도적인 오용에 대해서는 사용자에게 책임을 물을 수 있음

* End-of-End방식: 입력에서 출력까지 파이프라인 네트워크 없이 신경망으로 한 번에 처리하는 것을 의미하며, 중간단계에서 별도의 작업이 이루어지지 않기 때문에 효과적인 처리가 가능

○ 정부는 대화형 정보통신서비스에서 아동을 보호하기 위해 정보통신망법을 개정하면서 정보통신망법상 정보통신서비스 제공자가 제공하는 대화형 서비스가 문제되지 않도록 해야 한다는 일종의 주의의무를 부과. 개정된 정보통신망법은 인공지능 서비스에 대한 규제 수준은 높지 않지만 관련 사건이 발생할 때마다 규제 당국이 다양한 규제 방안을 강구할 것을 의미

* 정보통신망법 제44조의8(대화형 정보통신서비스에서의 아동 보호) 정보통신서비스 제공자는 만 14세 미만의 아동에게 문자·음성을 이용해 사람과 대화하는 방식으로 정보를 처리하는 시스템을 기반으로 하는 정보통신서비스를 제공하는 경우에는 그 아동에게 부적절한 내용의 정보가 제공되지 아니하도록 노력해야 한다.

○ 챗GPT가 생성한 결과물을 둘러싸고 다양한 법적·윤리적 이슈가 제기된다는 점에서 인공지능 윤리의 범위에 대해서도 고민할 필요. 글쓰기에 대해 어떠한 저작권법적 이슈와 윤리적인 이슈를 다룰 것인지에 대한 현실적인 논의가 필요하며, 전문가의 논의를 벗어나 직접적 이해 당사자인 이용자의 입장에서의 논의도 필요

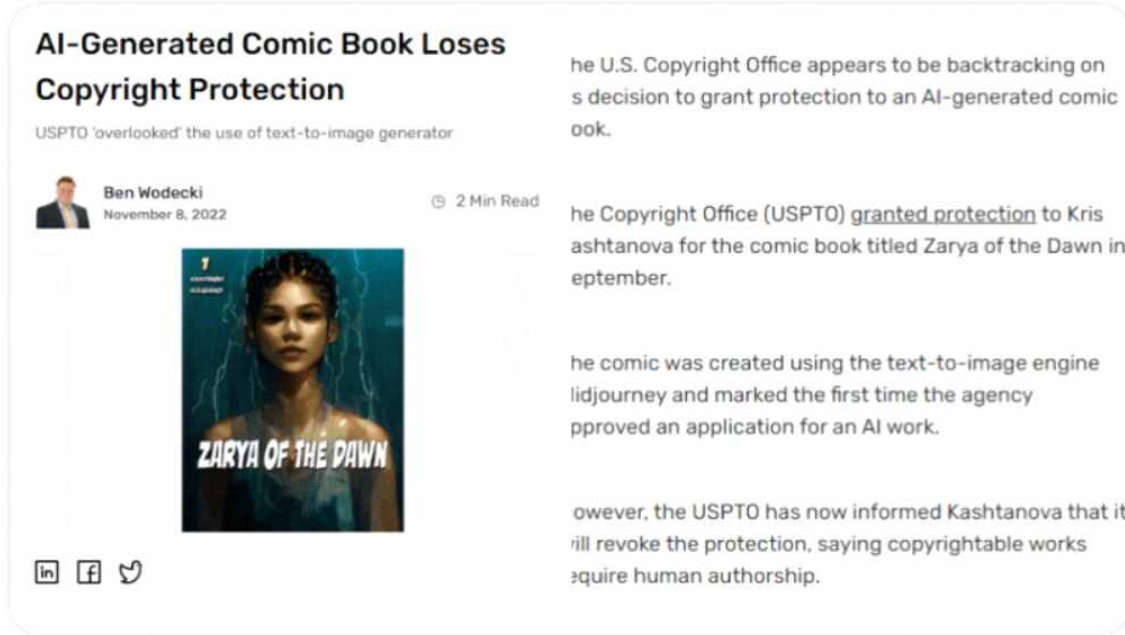
○ 챗GPT가 생성한 결과물을 자신의 것으로 표시하는 것은 표절 또는 연구윤리에 위배될 수 있으나, 저작권이 인정되기 어려운 상황에서는

저작권 침해로 보기 어려움. 챗GPT를 이용한 작성여부를 판별해주는 GPT를 개발하는 의견이 있지만 기능의 완벽을 보장할 수 없으므로 챗GPT를 이용하지 않은 경우에도 오인하여 판별하는 문제가 발생할 우려가 있음

3. 저작권 관련 쟁점

- 최근에는 생성형 AI와 같은 기계가 작성한 결과물의 저작권 여부에 대한 관심이 높아지고 있으며, 챗GPT가 생성한 결과물은 인간의 창작 수준을 넘어선다는 평가도 있는 반면, 결과물에 인간의 상상과 감정이 담겨있다고 보기 어렵고 창작성이 없어 저작권으로 보호하기 어렵다는 입장이 대립
 - * 창작성이란 완전한 의미의 독창성을 요구하는 것은 아니라고 하더라도, 창작성이 인정되려면 적어도 어떠한 작품이 단순히 남의 것을 모방한 것이어서는 안 되고 사상이나 감정에 대한 창작자 자신의 독자적인 표현을 담고 있어야 한다. 누가 하더라도 같거나 비슷할 수밖에 없는 표현, 즉 작성자의 창조적 개성이 드러나지 않는 표현을 담고 있는 것은 창작물이라고 할 수 없다.(대법원 2021. 6. 30. 선고 2019다268061판결)
- 저작권 등록은 실질 심사가 아닌 형식적인 심사로 진행되지만, 중요한 요건은 사람이 창작했는지의 여부에 대한 확인이며 이러한 사실을 기재하지 않는 경우에는 등록이 취소될 수 있음. 미국 저작권청(Copyright Office)이 미드저니로 만든 그림 소설인 ‘새벽의 자리야(Zarya of the Dawn)’를 저작물로 등록한 사례를 보면 등록 당시에 AI를 활용했다는 내용을 기재하지 않아 형식적인 심사에서는 저작권청에 등록이 가능했으나, 추후 작가가 AI로 생성했다는 사실을 SNS에 밝히며 등록이 취소됨
 - * 해당 사실을 인지한 미국 저작권청은 2022년 10월 작가에게 AI에 의해 제작된 내용의 소명을 요청했으나 창작과정에서 AI가 관여하지 않았다는 점을 명확하게 소명하지 못하여 등록이 취소된 사례가 있음. 다만 미드저니로 제작한 이미지에 대한 등록이 취소된 것이고 텍스트 및 편집물에 대한 등록은 유효한 것으로 인정
- 저작권은 창작과 동시에 발생되므로 등록을 취소하더라도 저작자가 창작한 것이라면 저작권 자체를 부인할 수는 없으며, 만약 저작물이 성립되지 않을 경우에는 해당 저작물은 누구라도 자유롭게 이용 가능. 미국 저작권청은 ‘새벽의 자리야’ 사건 이후로 AI 기술이 인간의 개입 없이 자율적으로 저작물을 생성하는 경우에는 해당 저작물은 저작권 보호의 대상이 아니라는 인공지능 창작에 의한 가이드라인을 제정

[그림1] <새벽의 자리아> 저작권 무효



- 최근 챗GPT의 사용을 표기한 논문이나 글이 증가하고 있으나 이는 챗GPT라는 알고리즘이 만들어 낸 것으로 저작권을 인정하기 어려움에도 저자와 챗GPT를 공동저작자로 표기하고 있음. 저작권법에서는 저작물은 인간의 사상 또는 감정을 표현한 창작물로 정해 창작성을 요구하고 있으므로 이러한 표기 형태는 저작권법에서는 허용되지 않음. 또한 저작권법은 저작자 아닌 자를 저작자로 표시하는 것을 금지하고 있으므로 사람이 아닌 자를 저작자로 표시하는 것은 법적으로 성립이 어려움
 - * 공동저작자는 '해당 저작물을 창작하기 위한 공동의 의사를 바탕으로 분리해 이용할 수 없는 저작물인 공동저작물을 작성한 다수의 저작자'를 의미하므로, 챗GPT는 저작자의 요건인 사람이 아니라는 점과 공동의 의사를 가지고 창작을 한 것이 아니라는 점 등을 이유로 공동저작자가 되는 것이 불가능
- 이처럼 대규모 언어모델이 갖는 한계로는 첫째, 학습 데이터에 포함된 정보만을 이용해 학습하고 생성하기 때문에 학습 데이터에 의존한다는 점, 둘째, 생성된 문장에 일관성이 없거나 잘못된 정보를 제공하게 되는 일관성과 신뢰성 문제, 셋째, 생성된 문장에 혐오 발언이나 차별적 발언 등을 포함시켜 사회적으로 문제를 일으키는 등 환각효과를 포함한 윤리적 문제, 넷째, 학습과 생성에 많은 컴퓨팅 자원이 필요하지만 이를 충족하지 못하는 부족한 컴퓨팅 자원의 한계, 다섯째, 인간이 갖고 있는 직관이나 감성을 포함하지 못하며 인간의 창의성과 상호작용을 대체할 수 없다는 점 등 여러 가지 해결이 필요한 문제들이 지적되고 있음
- 저작권법적 입장에서 생성형 AI는 새로운 창작도구 역할을 할 수 있다는

기대와 함께 기계학습 과정에서 타인의 저작물을 이용해 학습될 경우 저작권 문제가 확산된다는 우려가 동시에 존재. 이처럼 새로운 기술에 대해서 법은 양면적인 해석이 가능하고 사회적 합의를 통해 기술에 대한 사회적 수용성을 확대 가능. 이러한 과정에서 저작권법의 개정은 이뤄져왔고 새로운 비즈니스 모델이 확산되면서 다양한 영역에서 혁신을 이끌어 냈으므로 기술에 대한 규제는 기술 자체의 차단이 아닌 예견되는 문제를 다룰 수 있는 수준으로 관리가 필요

■ 시사점

- 생성형 AI가 가져오는 법률문제로 등록자가 AI 활용 사실을 고지하지 않을 경우 인간의 창작과 AI 결과물의 차이를 구분하기 쉽지 않기 때문에 미국 저작권청은 AI 생성물에 대한 등록 가이드라인을 공개하였지만, 이는 실무적인 지침일 뿐 입법적인 판단으로 보기 어려움. 보다 구체적인 논의를 통해 AI 생성물의 저작권을 어떻게 처리할 것인지 결론을 내릴 필요가 있으며, 가이드라인 등의 도출을 통해 향후 법적 안정성을 가져올 것으로 기대
- 챗GPT의 환각효과를 가져오는 결과나 일반적인 알고리즘이 제시하는 결과가 인간의 규범과 충돌할 경우에는 규제가 작동될 가능성이 있으므로 부정확한 정보나 알고리즘의 편향 등에 대한 문제도 해결해야 할 이슈. 알고리즘을 포함한 다양한 인공지능 기술에 대해 규제적 측면에서 접근할 경우 신산업으로서 인공지능 산업의 경쟁력은 약화될 수밖에 없으므로 기술에 대한 정책 방안은 기술 발전을 저해하지 않는 범위 내에서 운영
- 알고리즘의 편향과 내용상의 오류를 희석화 하는 방안으로는 보다 많은 데이터를 제공함으로써 편향적인 결과가 나오지 않도록 확률을 축소 시켜야 하며, 다양한 데이터를 확보해 학습데이터로 활용. 이를 위해 저작권법에 따른 공정이용이나 저작권법 개정안에 따른 정보 분석 등에 대한 제한 규정이 통과되도록 하고, 플랫폼 사업자가 사용한 데이터를 통해 얻은 수익을 독점하지 않도록 권리자에 대한 보상체계도 확보
- 생성형 AI의 문제는 생성물 자체만의 문제가 아닌 국가의 시스템에 미치는 영향이 크다는 점을 고려하여 심도 깊은 정책 설계가 필요. 이러한 문제를 선제적으로 논의하고자 문체부는 'AI-저작권법 제도개선 워킹그룹'을 발족 '하고 AI 학습데이터에 사용되는 저작물의 원활한 이용 방안, AI 산출물의

법적 지위 문제 및 저작권 제도에서의 인정 여부, AI 기술 활용 시 발생하는 저작권 침해와 책임 규정 방안 등을 논의하고 있으며, 현행 저작권법 내에서 활용될 수 있는 ‘저작권 관점에서의 AI 산출물 활용 가이드(안)’도 마련할 예정

참고자료

- 생성형 AI 모델의 법적 쟁점
<http://file.npplus.co.kr/ebook/sw/sw108/#p=54>
- 생성형 AI 관련 법적 문제
<https://blog.naver.com/jinnysuh00/223134901469>
- AI 창작물 어디까지 인정? 고민 커진 플랫폼, 운영지침 고심
<https://www.ddaily.co.kr/page/view/2023051914114964186>
- Chat GPT와 저작권
<https://news.mt.co.kr/mtview.php?no=2023041316201844918>
- 오픈소스를 학습한 생성형 AI 관련 법적 분쟁
https://blog.naver.com/kcc_press/223045277963
- “모든 것이 붓으로 귀결되는 미래” 생성형 AI에 대한 10가지 우려
<https://www.itworld.co.kr/news/277867>
- “새벽의 자리야” 결국 저작권 무효: AI-Generated Comic Book Could Lose Copyright Protection
<https://aibusiness.com/ml/ai-generated-comic-book-loses-copyright-protection>

SPC 'ANGEL' 통계

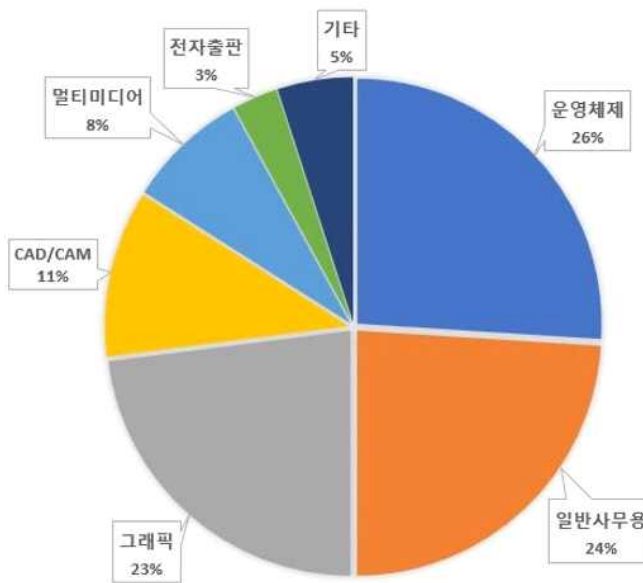
월 1회(매 25일) 제공

한국소프트웨어저작권협회, 불법복제 SW 제보 'ANGEL' 서비스 5월 통계 현황

- 한국소프트웨어저작권협회(SPC)가 지난 5월 한 달간('23. 5. 1. ~ 5. 31.) 'ANGEL (불법제보)' 서비스를 분석한 결과, 기업 또는 개인의 불법복제 SW 사용 제보는 총 62건으로 나타났으며,
- SW 용도별로는 운영체제 16건(26%), 일반사무용 15건(24%), 그래픽 14건(23%), 설계(CAD/CAM) 7건(22%), 멀티미디어 5건(8%), 전자출판 2건(3%), 기타 3건(5%) 순으로 제보가 접수된 것으로 나타남

[그림] SPC 'ANGEL(불법제보)' 서비스 2023년 5월 통계 현황

2023. 5. 불법복제 소프트웨어 제보 통계
- SW 용도별 제보 건수 -



* 한국소프트웨어저작권협회는 2018년 11월부터 제보시스템과 제보 방식의 편의성을 개선한 불법복제 SW 제보 시스템 'ANGEL(불법제보)' 서비스를 운영하고 있음

SW·저작권 동향리포트는 매월 10일, 25일에 발간됩니다.
다음 SW·저작권 동향리포트 <제2023-9호> 발간일은 7월 10일입니다.