

January 26, 2024

ChatGPT, 뉴욕타임스: 법적 혁명

AI를 둘러싼 법적 전쟁(legal war)이 시작되었습니다. 2023년 12월 말, 뉴욕타임스는 OpenAI가 생성형 AI인 ChatGPT를 학습시키기 위해 뉴욕타임스의 보도내용을 “도용”한다고 주장하며 소송을 제기하였습니다. 최신 보도 내용을 학습할 수 없게 되는 경우, ChatGPT의 역량은 심각하게 제한될 것으로 보입니다. 거의 같은 날, 한국신문협회는 네이버의 한글 기반 생성형 AI인 HyperCLOVA X에 대해 이와 유사한 소송을 제기하였습니다. 두 소송을 비롯하여 이와 관련된 사례들은 생성형 AI와 콘텐츠 제작자들에 대한 법적 환경(legal landscape)의 향방을 결정하게 될 것으로 보입니다. 다만 현재로서는 두 소송의 결과를 예측하기는 어렵습니다.

New York Times v. OpenAI / 마이크로소프트

뉴욕타임스는 OpenAI와 마이크로소프트가 ChatGPT의 학습 및 이용과 관련하여 저작권을 침해하였다고 주장하였습니다.¹ ChatGPT는 사용자의 질문을 이해하고 이에 응답할 수 있는 “대규모 언어 모델 생성형 AI(Large Language Model generative AI)” (이하 “LLM AI”)입니다.

LLM AI 개발 단계에서 개발자는 LLM AI를 “학습”시키기 위해 대량의 텍스트를 복제(copy)할 필요가 있습니다. 이 텍스트는 신문, 서적, 블로그, 심지어는 인터넷 카페로부터도 가져올 수 있습니다.

ChatGPT가 뉴스에 대한 질문에 답변을 제공할 수 있는 이유 중 하나는, 뉴욕타임스 기사를 학습하였기 때문입니다. 그 결과, 이용자들은 뉴욕타임스 구독료를 지불하지 않고도 ChatGPT에 질문하여 뉴스를 알 수 있게 되었습니다. 뉴욕타임스는 이것이 원본 보도에 대한 저작권 침해에 해당한다고 주장하는 것입니다.

한국신문협회의 의견 제출

거의 같은 날, 한국신문협회도 네이버에 대해 유사한 우려를 제기하는 의견(이하 “한국신문협회 의견”)을 한국공정거래위원회(이하 “공정위”)에 제출하였습니다.

ChatGPT는 한국어로 된 사용자 질문을 이해하고 이에 대한 답변을 제공할 수 있습니다. 그러나 네이버의 한글 기반 LLM AI인 “HyperCLOVA X”는 ChatGPT보다 6,500배 많은 한글 자료를 학습하였다고 합니다. 따라서 네이버가 한국 검색 엔진 분야에서 Google, Yahoo, Bing을 제쳤듯 HyperCLOVA X는 한국 AI 시장에 상당한 영향을 미칠 수 있을 것으로 보입니다.

¹ Michael M. Grynbaum and Ryan Mac, *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*, The New York Times (Jan. 19, 2024), available at <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>.

저작권 침해: “변형적 이용” v. “시장 대체물”

한국과 미국의 저작권법에서는 저작물의 이용이 “공정한 이용”에 해당하는 경우, 저작권이 있는 자료를 로열티를 지불하지 않고 이용할 수 있도록 허용하고 있습니다.²

“공정한 이용”에는 “목적 또는 성격이 다른 새로운 것을 추가하여 [원저작물]을 새로운 표현, 의미 또는 메시지로 변경하는” 저작권이 있는 자료의 “변형적 이용(transformational uses)”이 포함될 수 있습니다.³ 반면, 원본 저작물에 대한 “시장 대체물(market substitution)”을 만드는 이용은 일반적으로 저작권 침해에 해당합니다.⁴

“변형적 이용”: 수백만권의 책을 스캔한 구글

예를 들어, 2004년 Google은 이전에 하드카피로만 존재했던 수백만 권의 저작권이 있는 책들의 전체 텍스트를 스캔하기 시작했습니다.⁵ Google은 이러한 스캐닝을 통해 사용자가 책 전체를 읽지 않고도 단어와 구절에 대한 전체 텍스트 검색을 수행할 수 있는 온라인 데이터베이스를 구축하였는데, 이 과정에서 저작권 소유자에게 저작권료를 지급하지 않았습니다.

한국신문협회 의견에서 인용된 판례를 보면, 미국 법원은 다음과 같이 판시하였습니다:

“검색 가능한 전체 텍스트 데이터베이스를 생성하는 것은 전형적인 변형적 이용이다... 단어 검색의 결과는 그 단어가 사용된 페이지(및 책)와는 그 목적과 성격, 표현, 의미를 달리한다.”⁶

책 전체를 복제하는 것은 일반적으로 저작권 침해에 해당하는데, 그 이유는 복사본을 읽는 것이 해당 책의 잠재적 판매를 “대체”하기 때문입니다. 그러나 Google이 제공한 “전체 텍스트 검색 기능”의 경우 “책[그 자체]을 대체하는 역할을 하지 않는다”는 것입니다.⁷ 즉 Google의 대량 복제는 “변형적 이용”이므로 “공정한 이용”에 해당하며, 이에 Google은 저작권료를 지급할 필요가 없었던 것으로 보입니다.

“시장 대체물”: 앤디 워홀의 예술적인 사진 복제

한편, 미국 대법원은 지난 해 변형물로 보이는 예술품이 원저작물의 “시장 대체물”으로 기능할 경우 저작권을 침해할 수 있다고 판시하였습니다.⁸

앤디 워홀은 예술적, 법률적 의미에서 “변형적 복제의 전형”이라고 불려왔습니다.⁹ 앤디 워홀은 잡지에 실린 음악가 프린스의 사진을 실크스크린으로 복사하여 “오렌지 프린스(Orange Prince)”를 만들었습니다. 워홀이 “오렌지 프린스”

2 미국 저작권법(U.S. 1976 Copyright Act) 제107조, 한국 저작권법 제35조의 5 및 한국신문협회 의견(미국 사례를 관행적 권위로 인용함) 참조.

3 *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579, 114 S. Ct. 1164, 1171, 127 L.Ed. 2D 500(1994).

4 *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528, 555, 143 S. Ct. 1258, 1290, 215 L.Ed. 2D 473 (2023).

5 *Authors Guild, Inc. v. HathiTrust*, 755 F.3D 87, 96 (2d Cir. 2014).

6 *Authors Guild, Inc. v. HathiTrust*, 755 F.3D 87, 97 (2d Cir. 2014).

7 상동, 99-100 (강조표시 추가).

8 *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 561, 143 S. Ct. 1258, 1293, 215 L.Ed. 2d 473 (2023); 한국신문협회 의견 III.2 참조(이 사건 인용).

9 상동, 592 (Kagan 대법관의 반대이견) (Warhol의 작품을 “변형적 이용”의 예로서 이용한 대법원 판례 인용).

가 다른 잡지에 실릴 수 있도록 허락하자, 프린스 사진을 촬영한 원 사진작가는 소송을 제기하였습니다.



한국신문협회가 인용한 사건에서 법원은 다음과 같이 설명하였습니다.

“원저작물의 목적과 동일하거나 매우 유사한 목적을 달성하기 위해 원저작물을 이용하는 경우, [시장에서] 해당 저작물을 대체하거나 갈음할 가능성이 더 높다.”¹¹

원저작물인 사진은 프린스에 관한 한 잡지에 실렸고, 그 후 다른 잡지 회사가 프린스에 대한 다른 기사에 “오렌지 프린스”를 이용하기 위해 워홀에게 돈을 지불한 상황에서, 법원은 “두 [이미지]가 완전한 대체물이 아니더라도” “동일한 목적을 공유”했으므로 워홀측에 저작권 침해에 대한 책임이 있다고 판단한 것입니다.¹²

LLM AI의 보도물 이용: “시장 대체물”인가, “변형적 이용”인가

미국 법원과 공정위가 미국과 한국 신문사측 주장에 대하여 어떻게 판단할 것인지 현재로서는 알기 어렵습니다.

우선, 뉴욕타임스와 한국신문협회는 ChatGPT와 HyperCLOVA X가 기존 보도를 대체하는 “시장 대체물”을 만들어냈다고 주장합니다.

- LLM AI는 기자들의 저작권이 있는 기사를 분석하여 뉴스에 대한 질문에 답변을 제공하며,
- 이용자들은 LLM AI를 통해 동일한 소식을 알 수 있기 때문에, LLM AI는 “기존 보도와 동일한 정보를 제공”할 목적이 있다고 볼 수 있다는 것입니다.¹³

LLM AI가 원본 기사를 “완전히 대체”할 수는 없습니다. 그러나, 신문사들은 LLM AI가 자사 웹사이트에서 수익 창출이 가능한 트래픽을 “도용”하고 저작권을 침해한다고 합리적으로 주장할 수 있어 보입니다.¹⁴

그러나 다른 한편으로는, LLM AI는 사용자가 뉴스와 상호작용하는 방식 자체에 “변화”를 가져온다는 해석도 가능해 보

10 상동, fig. 6.

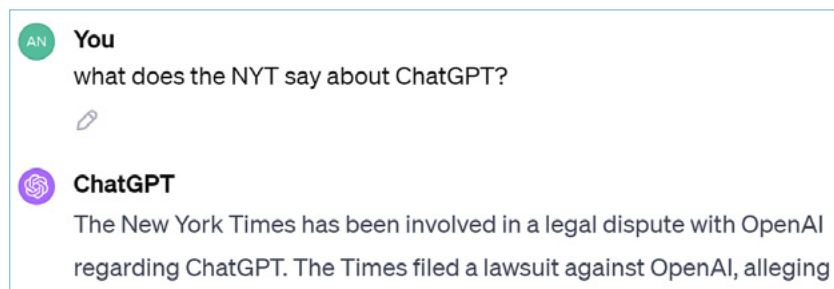
11 상동, 528(인용문 생략)

12 상동, 536.

13 뉴욕타임스 소장, ¶108.

14 Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 536 (2023).

입니다. Google 사안만 보더라도, 원래 독자들은 항상 책 전체를 꼼꼼하게 읽어야 특정 텍스트를 찾을 수 있었습니다. 그러나 Google의 비라이선스 스캐닝을 통해 독자들은 전체 텍스트 검색으로 몇 초 내에 텍스트를 찾을 수 있게 되었습니다. 마찬가지로, LLM AI는 저작권이 있는 기사의 정보에 접근하는 다른 방법을 제공한다고 주장할 수 있습니다. 예를 들어, 뉴욕타임스 v. OpenAI 사건에 대해 알고 싶은 사람들은 뉴욕타임스 웹사이트를 탐색하거나, 기존 검색 엔진을 사용하여 해당 주제에 대한 기사를 찾아 읽을 수 있습니다. 그러나 ChatGPT를 사용하는 경우 사용자는 간단히 질문하기만 하면 보다 직접적으로 답변을 얻을 수 있게 됩니다.¹⁵



따라서 OpenAI와 네이버는 LLM AI가 이용자들이 뉴스 기사로 전달되는 정보와 상호작용하는 방식을 “변화”시킨다는 점을 들어 “공정한 이용”에 해당한다고 주장할 여지가 있어 보입니다.¹⁶

기타 AI 관련 분쟁

생성형 AI는 이미 “핫뉴스”를 넘어 권리에 대한 법률적 이해에 지각 변동을 일으키고 있으며, 다음과 같은 사례들이 이를 방증하고 있습니다.¹⁷

- **생체정보**: 이용자의 실제 얼굴을 기반으로 아바타를 만드는 AI 기업이 생체정보에 관한 개인정보보호법을 위반하였다는 혐의로 소송을 당했습니다.¹⁸
- **초상권과 퍼블리시티권**: 한 “딥페이크(deep fake)” AI 크리에이터는 유명인의 초상을 무단으로 이용하여 당황스러운 가짜 이미지를 만든 혐의로 고소를 당하기도 했습니다.¹⁹
- **사진 및 이미지**: Getty Images는 한 AI 이미지 생성 회사가 자신의 사진을 무단으로 사용했다며 소송을 제기하였습니다.²⁰
- **서적**: George R.R. Martin (『왕좌의 게임』 저자)과 같은 저자들은 OpenAI가 ChatGPT를 학습시키기 위해 자신의 책을 도용했다고 소송을 제기하였습니다.²¹

15 ChatGPT와 HyperCLOVA X는 모두 응답과정에서 실수를 범한다. 더 심각한 문제는 LLM AI가 종종 존재하지 않는 사실과 사례를 참조하여 ‘환각’을 일으키는 것으로 알려져 있다는 점이다. 그럼에도 불구하고, 두 제품 모두 신뢰성을 지속적으로 높이고 있다.

16 Cf. *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 99 (2d Cir. 2014) (종종 원문을 인용하고 요약하는 서평(書評)은 일반적으로 저작권 침해에 해당하지 않는다고 설명함).

17 *International News Serv. v. Associated Press*, 248 U.S. 215 (1918).

18 *Flora v. PRISMA Labs, Inc.*, 2023 WL 5061955 (N.R., 2023).

19 *Young v. NeoCortex, Inc.*, 2023 WL 6166975 (C.D.Cal., 2023).

20 *Getty Images (US), Inc. v. Stability AI, Inc.*, U.S. District Court for the District of Delaware, No. 1:23-cv-00135, Date Filed: Feb. 3, 2023.

21 *Authors Guild v. OpenAI Inc.*, 1:23-cv-08292, (S.D.N.Y.). Date Filed: Sept. 19, 2023.

한국 기업들은 무엇을 해야할까?

2024년 1월에 발표된 국제통화기금(IMF) 보고서에 따르면 AI는 곧 전 세계 일자리의 40%에 영향을 줄 것이라고 합니다.²² 이 수치는 한국과 같은 선진국에서는 전체 일자리의 60%까지 증가할 것으로 보입니다. 결국 AI가 법조계를 포함한 모든 산업군에 영향을 미칠 것임에는 이론의 여지가 없습니다.

AI 관련 법(the law of AI)은 매일 변화하고 있습니다. AI 개발자들의 경우, 소송이 제기되기를 기다렸다가 대응하는 것만으로는 변화하는 환경에 충분한 대비를 하기 어렵습니다. 법률자문을 통해 법적 변화를 예측하고, 그에 따른 사업 계획을 수립할 수 있어야 합니다.

한편 콘텐츠 제작자의 경우, AI는 권리의 법률적 정의뿐만 아니라 시장 전반의 성격도 변화시킬 수 있음을 주의하여야 합니다. 단기적으로 한국 콘텐츠 제작자들은 생성형 AI와 관련하여 어떤 권리를 가질 수 있는지, 그리고 이러한 권리를 어떻게 보호할 수 있을지 고려해보아야 합니다. 미국, 한국 또는 유럽 법원에 소송을 제기하거나 미국 내 집단 소송에 참여하는 것도 방법이 될 수 있습니다.

[영문] “ChatGPT, the New York Times, and the KFTC: A Legal Revolution”?

22 Mauro Cazzaniga et al., “Gen-AI: Artificial Intelligence and the Future of Work”, INTERNATIONAL MONETARY FUND Staff Discussion Notes No. 2024/001, <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>.

관련 구성원

김중세

외국변호사(미국 Texas주)

T 02.3404.0473

E andrew.kim@bkl.co.kr

강동욱

변호사

T 02.3404.0158

E tony.dw.kang@bkl.co.kr