

생성형 AI를 둘러싼 최근 저작권 분쟁 동향과 시사점

김 창 화 교수

한밭대학교 공공행정학과 교수

Trend

I. 서언	177
II. 생성형 AI의 저작권 침해	179
III. 생성형 AI의 저작자성	204
IV. 결어 - 시사점 및 향후 과제	219

제24-1호

요약문

예술이나 미디어 분야에서 생성형 AI에 의한 콘텐츠의 생성이 활발하게 이루어지고 있다. 그 생성 과정에서 필연적으로 저작물의 이용이 수반되어, 저작권법상 여러 문제가 발생하고 있다. 구체적으로, 생성형 AI가 학습하는 과정에서 저작물을 이용하여 저작권 침해 문제가 발생하고, 이용자의 프롬프트에 의해 생성한 콘텐츠가 기존 저작물과 유사하여 침해 문제가 발생하기도 한다. 그리고 생성형 AI는 타인에 의해 저작권 침해가 발생하는 경우 저작권 침해를 도운 것으로 인해 간접 책임을 질 수도 있고, 저작권 관리 정보를 변경한 것에 대한 책임도 질 수 있다. 또한, 생성형 AI가 생성한 콘텐츠에 대하여 권리를 주장하려 하는데, 인간이 아닌 AI에게 저작자에 대한 지위를 부여할 수 있는지 그리고 만일 부여할 수 없다면 발생하는 이익을 누구에게 귀속시켜야 하는지에 대한 논란도 계속되고 있다.

비록 우리나라는 생성형 AI와 관련된 직접적인 분쟁이 많이 제기되고 있지 않으나, 미국에서는 여러 저작권 관련 소송이 제기되고 있고, 여기서 생성형 AI에 대해 공정이용을 인정하기보다는 저작권자의 이익 침해에 보다 집중하고 있는 것으로 보인다. 또한, 생성형 AI의 저작자성에 대해서도 아직은 저작자로 인정하지 않고 있으며, 인간이 가공한 경우에도 저작물로 인정하는 데 다소 제한적이다. 이는 우리나라에서는 생성형 AI가 학습하는 경우와 콘텐츠를 생성하는 경우 미국보다 더 저작권 침해 가능성이 높을 수 있다는 것과, 생성형 AI가 저작자로서 인정받기 어렵다는 것을 의미한다. 따라서 생성형 AI의 발전을 도모하기 위해서는 이러한 문제를 해결하는 방안들이 마련되어야 한다.

생성형 AI의 발전과 저작권의 보호 양자의 이익을 조화시키기 위해서는 먼저, 생성형 AI에 의해 제기된 저작권법상의 문제 중 저작권법의 개정이 필요한지를 고려해 보아야 한다. 구체적으로, AI가 생성한 저작물에 대하여 저작권을 여전히 인정하지 않을 것인지, 생성형 AI의 저작권을 인정하지 않는다면 누구에게 그 권리를 귀속시킬 것인지, 그리고 생성형 AI를 학습하는 과정이 공정 이용에 해당하는지를 명확히 해야 한다. 그리고 입법이 이루어지기 전까지 이해관계인들은 저작권 침해 방지를 위한 조치를 적절히 하고, 대립하는 이해관계 속에서 이익을 조정하는 방안을 마련하고, 각종 판단에서 세부적이고 구체적인 기준들을 확립해 나가야 할 것이다. 최근의 AI 사업자의 필터링 조치, 확대된 집중 관리나 보상금 청구권, 그리고 AI 사용자가 생성된 저작물로 인정받기 위해 필요한 입력의 정도 등의 기준 마련 등이 예가 될 수 있을 것이다.

I

서언

생성형(Generative) AI(Artificial Intelligence, 인공지능)는 텍스트(text), 이미지(image), 음악, 동영상 등과 같은 콘텐츠를 만들 수 있는 기술이다.¹⁾ 생성형 AI의 이용자가 프롬프트(prompt)²⁾를 입력하면, 생성형 AI는 기존에 학습된 능력을 기반으로 결과물인 콘텐츠를 생성한다. 여기서, 학습은 인간 두뇌의 뉴런(neuron)에서 영감을 받은 신경망을 사용하며, 기존 데이터의 패턴과 특징을 학습하고, 학습한 패턴과 일치하는 새로운 데이터를 생성한다. 예로써, 이미지 세트에 대해 훈련된 생성형 AI 모델은 훈련된 이미지와 유사한 새로운 이미지를 생성할 수 있다. 기업과 개인들은 이러한 생성형 AI를 통해, 고품질의 맞춤형 콘텐츠를 제작할 수 있게 된다. 현재 가장 인기 있고 강력한 생성형 AI 모델들에는 텍스트 프롬프트에 대해 인간과 같은 반응(채팅)을 생성하도록 설계한 ChatGPT, 텍스트 프롬프트를 기반으로 독창적이고 상세한 시각적 결과를 생성하는 DALL-E, 인공지능 검색 엔진 서비스인 Google의 Bard 등이 있다.

생성형 AI는 기술, 금융, 광고나 마케팅 분야에서도 활발히 이용되고 있으나, 예술이나 미디어 분야와 같은 창의적인 분야에서 가장 널리 이용되고 있다. 생성형 AI는 소설을 쓰는 작가에게, 음악을 작곡하는 음악가에게, 새로운 가상 세계를 만드는 게임 디자이너에게, 그리고 시각 효과를 생성하는 영화 제작자에게 콘텐츠를 생성하여 작품을 더 빠르고 경제적으로 제작할 수 있도록 도움을 준다. 그런데, 생성형 AI는 위에서 살펴본 바와 같이 학습 과정에서 저작물을 포함한 여러 정보를 이용하여 콘텐츠를 제작하기 때문에, 여러 법률적 문제가 발생한다.³⁾ 특히, 저작권법과 관련해서는 저작물을 허락 없이 이용하는 경우 저작권 직접

1) elastic, 생성형 AI란 무엇인가? <<https://www.elastic.co/kr/what-is/generative-ai>, 2024. 1. 21. 최종방문>.

2) 프롬프트는 생성형 AI에 입력하는 입력값을 의미하며, ChatGPT에 입력하는 질문이나 지시가 이에 해당한다. 프롬프트를 잘 작성해야 좋은 결과값을 얻을 수 있어, 좋은 프롬프트를 만들어 내는 프롬프트 엔지니어링과 프롬프트 엔지니어가 생겨났다. <<https://namu.wiki/w/프롬프트>, 2024. 1. 21. 최종방문>.

3) 초상이나 목소리 등의 퍼블리시티권(publicity) 대상을 허락 없이 영리적으로 이용하는 경우나 개인정보를 이용하는 경우 등이 있을 수 있다. 다만, 본고에서는 저작권법적 측면에 한정되어 있어, 이들에 대한 논의는 생략하기로 한다.

침해가 성립할 수 있고, 저작권 직접 침해가 발생하는 경우 그 침해에 대하여 간접적으로 책임을 질 수도 있다. 그리고 생성형 AI는 물론 AI 전체 분야에 걸쳐 저작자나 저작권자로 될 수 있는지의 문제도 여전히 존재하고 있다.

저작권(재산)권 직접 침해는 저작권자 허락 없이 저작물을 이용할 때 발생한다. 구체적으로는 저작물을 무단으로 복제하거나, 저작물을 변형하여 2차적저작물을 작성하는 경우에 복제권 침해나 2차적저작물 작성권 침해가 성립하며, 저작권 침해가 성립하면 민형사상의 구제 조치를 할 수 있다. 우리는 그 구제 조치에 별다른 요건이 필요하지 않지만, 미국은 그 경우 등록을 사전에 하도록 요구한다.⁴⁾ 다만, 공정 이용과 같이 저작권법이 정한 일정한 경우에는 저작권자 허락 없이 저작물을 이용하여도 침해가 성립하지 않는다.⁵⁾ 또한, 저작권을 직접적으로 침해하지 않은 경우에도 저작권 침해를 방조하는 경우에는 간접적으로 책임을 질 수 있다. 미국의 경우에는 저작권 침해에 대해 간접책임으로서 대위(vicarious) 책임, 기여(contributory) 책임, 유인(inducing) 책임을 인정한다.⁶⁾ 그리고 저작자는 저작물을 창작한 자이며, 저작자는 무방식주의에 따라 어떠한 절차나 형식 없이 저작권자가 된다. 문제는 저작물의 법적 정의가 “인간의 사상과 감정을 표현한 창작물”로 규정되어 있어, 인간이 아닌 AI 등과 같은 주체의 작품은 저작물이 될 수 없다는 데 있다.⁷⁾

국내에서는 아직 생성형 AI에 대한 직접적인 분쟁은 발생하고 있지 않지만, 생성형 AI 개발이 증가하고 있어 그 시장은 계속하여 확대되고 있고, 해외 여러 국가 특히, 미국에서는 생성형 AI와 관련된 저작권 분쟁이 계속 발생하고 있다. 따라서 이제 우리도 생성형 AI에 대한 문제에 대하여 살펴보고, 이에 대한 대응 방안을 살펴볼 때가 되었다. 이를 위해, 2장에서는 생성형 AI가 학습 과정에서 그리고 작품을 생성한 후에 발생할 수 있는 저작권 침해 관련 분쟁을 살펴보고, 3장에서는 생성형 AI가 저작자로 인정될 수 있는지와 저작자로 인정되지 않는 경우 저작자로 될 수 있는 자에 대한 생성형 AI의 저작자성에 대해 검토한다. 그런 후, 이러한 문제들이 우리에게 가져다주는 시사점을 살펴보고, 앞으로의 우리의 대응 방안에 대하여 살펴보면서 글을 마치고자 한다.

4) 17 U.S.C. § 411(a).

5) 저작권법 제23조 내지 제35조의5.

6) 미국에서 간접 책임이 성립하기 위해서는 모두 공통적으로 직접 침해가 존재해야 하고, 대위 책임은 침해를 관리할 수 있는 권리와 능력 그리고 침해를 통한 직접적인 이익이 있어야만 성립하며, 기여 책임은 침해에 대한 인식과 침해에 대한 실질적 기여가 있어야만 성립하며, 유인 침해는 저작권 침해를 유인하는 경우에 성립한다.

7) 저작권법 제2조제1호.

II

생성형 AI의 저작권 침해

1. 최근 제기된 일련의 소송

가. Thomson Reuters Enterprise Centre GmbH et al v. ROSS Intelligence Inc. (2020. 5. 6.)⁸⁾

본 사건은 초기 생성형 AI와 관련된 대표적 사건 중 하나이며, Thomson Reuters는 ROSS가 경쟁적 관계에 있는 생성형 AI 기반의 법학 리서치 플랫폼의 학습 데이터집으로 이용하기 위해, 라이선스를 거절한 후, 자신의 플랫폼인 Westlaw의 데이터베이스를 전부 복제했다고 주장하였다. 2021년 Reuters의 청구가 받아들여지자, ROSS는 Westlaw 데이터베이스의 허락받지 않은 복제나 이용은 공정 이용에 해당한다고 반박하였다. 구체적으로, ROSS는 모델을 훈련하기 위해 텍스트 중 보호되지 않는 아이디어와 사실만을 취하였고, 그 목적은 생성형 AI를 기반으로 하는 검색 도구에 대한 “완전히 독창적이고 새로운 코드를 작성”하는 것이라고 하였고, 헤드노트(headnotes)와 키 넘버(key numbers)로 구성된 Westlaw 콘텐츠를 침해하지는 않았다고 주장하였다.

나. Doe v. GitHub, Inc., et al. (2022. 11. 10.)⁹⁾

두 명의 원고는 Microsoft, GitHub, OpenAI에 대하여 DMCA 위반으로 그리고 GitHub와 OpenAI에 의해 공동 개발된 구독 기반 AI 도구인 Copilot과 관련하여 계약 위반, 불법 간섭(tortious interference), 사기, 허위 지정, 부당 이득 그리고 불공정 경쟁을 이유로 소송을 제기하였다. 원고의 소송 중심에는,

8) 1:20-cv-613-SB (D. Del. 2023).

9) 22-cv-06823-JST (DMR), (N.D. Cal. Nov. 28, 2023).

Copilot에 대한 학습 데이터로서 저작권에 의해 보호되는 소스 코드(source code)를 피고들이 사용했다는 주장이 있는데, 이는 소프트웨어 개발자들이 자연어 프롬프트를 코딩 제안으로 변환시켜 쉽게 코드를 생성할 수 있도록 한다. 원고들에 따르면, 오픈 소스 코드 저장소인 GitHub, Copilot을 만드는 데 사용된 GPT-3 언어 모델을 만든 OpenAI, GitHub 소유자 그리고 OpenAI 투자자 Microsoft의 피고들은 Copilot을 훈련하기 위해 GitHub에 공개적으로 접근할 수 있는 저장소에서 가져온 데이터를 사용하였다. 원고들은 자신들이 특정 오픈 소스 라이선스에서 그러한 코드나 다른 작품들을 게시하였고, 그러한 라이선스 모두는 저자의 이름과 저작권에 대한 표시를 요구한다고 하였다. 그런데, 피고들은 자신들의 코드를 사용하면서 라이선스 및 원고의 권리에 반해 코드에 저작자 표시, 저작권 고지 그리고 라이선스 조건을 제외하였다고 원고는 주장하였다.

다. Anderson v. Stability AI Ltd. et al. (2023. 1. 13.)¹⁰⁾

Stability AI는 2023년 1월 공동 피고인 DeviantArt 및 Midjourney와 함께 저작권 침해, 불공정 경쟁 그리고 퍼블리시티권 관련 소송에 피고로 이름을 올렸다. 본 소송에서 3인의 예술가들은 Stable Diffusion을 포함한 AI 이미지 생성기가 새로운 이미지라고 하지만, 실제로 침해에 해당하는 2차적저작물을 생성하였고, 이는 그들의 예술 작품을 허락 없이 사용하여, 막대한 침해를 일으킨 것이므로, 그 주체인 Stable AI를 상대로 소를 제기한다고 하였다. 피고들은 Stable Diffusion이 그해 봄, 인터넷에서 공개적으로 이용할 수 있는 수십억 개의 이미지에 대하여 훈련을 받았지만, 모델을 훈련하는 것이 차후 배포를 위해 그 이미지를 복사하거나 암기하는 것을 의미하지는 않는다고 반박하였다. 사실, Stable Diffusion은 어떠한 이미지도 저장하지 않았다. 한편, 텍스트 이미지 생성기인 DeviantArt는 책임을 인정한다면 예술을 창작하는 DreamUp의 잠재력에 크게 영향을 끼칠 것이고, 이는 언론의 자유를 침해하는 것이 되어, 이에 대한 청구는 기각되어야 한다고 주장하였다. 피고들은 또한, Los Angeles에 기반을 둔 온라인 예술 및 AI 플랫폼은 원고의 청구들이 California의 반(反)-SLAPP 법령¹¹⁾에 의해 금지되어야 한다고도 하였다.

라. Getty Images (US), Inc. v. Stability AI, Inc. (2023. 2. 3.)¹²⁾

Getty가 런던 고등법원에서 Stability AI를 상대로 법적 절차를 개시했다고 발표한 후, 미국의 Getty Images는 Stability AI가 엄청난 규모의 지식 재산을 뽐낸다고 비난하며 미국 내에서 소송을 제기하였다. 사진 에이전시인 원고 Getty는 Stability AI가 경쟁 비즈니스를 구축하기 위한 노력의 하나로,

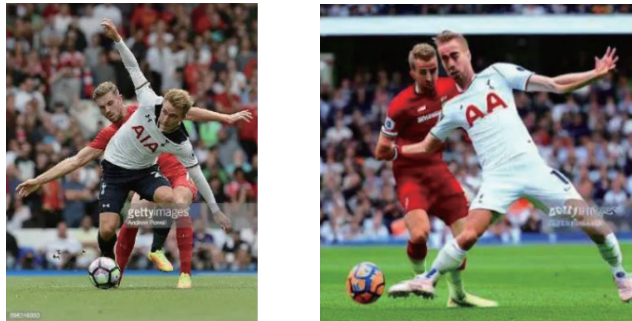
10) 3:23-cv-00201 (N.D. Cal. Oct. 30, 2023).

11) SLAPP(Strategic Lawsuit Against Public Participation)는 대중 참여에 대한 전략적 소송으로, 비판자가 비판이나 반대를 포기할 때까지 법적 방어 비용을 부담시켜 비판자를 검열, 협박, 침묵시키기 위한 소송이다. <https://en.wikipedia.org/wiki/Strategic_lawsuit_against_public_participation, 2024. 1. 27. 최종방문>.

12) 1:2023cv00135 (E.D. Mo. 2023).

Getty Images에게 어떠한 허락이나 보상 없이 컬렉션에서 수백만 장의 사진을 복제했다고 주장하였다. 원고는 또한, 저작권 침해 소송 사유를 제시하고 Stability AI가 허위 저작권 관리 정보를 제공했거나 저작권 관리 정보를 제거 또는 변경했다고 주장하는 것 외에도, Stable Diffusion 모델이 종종 Getty Images 워터마크(watermark)의 수정된 버전이 포함된 결과물을 생성한다는 이유로 Stability AI를 상표권 침해 및 희석(dilution) 혐의로도 고소하였다. 이는 이미지의 출처에 대한 혼란을 일으키고, Getty Images와의 연관성을 잘못 암시할 수 있게 한다고 하였다. 또한, Getty는 Stable Diffusion을 사용하여 생성된 결과물 중 일부는 미학적으로 만족스럽지만, 다른 결과물은 품질이 훨씬 낮고 때로는 기괴한 것부터 터무니없는 것까지 다양하여 희석의 원인이 된다고도 주장하였다. 당해 5월에 제출한 기각 신청서에서, Stability AI는 Getty가 Stable Diffusion 훈련과 관련하여 저작권을 침해하는 행위가 Delaware 내에서 발생했다고 주장하지 않았기 때문에, Delaware의 Long-Arm 법률¹³⁾에 따른 관할권이 없다고 주장하였다. 대신, Stability AI는 Getty가 영국과 독일에서 훈련이 이루어졌다고 주장하는 것으로 보인다고 주장하면서, 원고의 수정된 소장에서 Stable Diffusion이 당사자가 아닌 독일 법인인 LAION이 준비한 데이터 세트에서 훈련되었다는 문구를 지적하였다. Getty는 Stability AI가 Delaware에서 서비스 또는 물건을 공급하기로 계약했다고 주장하지도 않았다.

[그림 1] Getty Image 원본(왼쪽)과 Stable Diffusion에 의해 만들어진 사진(오른쪽)



마. Tremblay v. OpenAI, Inc. (2023. 6. 28.)¹⁴⁾

두 명의 저자들은 ChatGPT 개발사인 OpenAI를 상대로 소송을 제기하였다. 6월 28일 캘리포니아 북부 연방법원에 제출한 소장에 따르면, Paul Tremblay와 Mona Awad는 ChatGPT라는 생성형 AI 챗봇을 구동하는 대규모 언어 모델을 학습하는 과정에서 OpenAI가 자신들이 저술한 책의 텍스트를 포함한 대량

13) Delaware의 Long-Arm 법률은 주의 비거주자 등에게 관할권을 부여한다. <<https://law.justia.com/cases/federal/district-courts/FSupp/752/1223/1587839/>, 2024. 1. 27. 최종방문>.

14) 3:23-cv-03223 (N.D. Cal. 2023).

의 데이터를 허가 없이 사용함으로써 직접적인 저작권 침해, DMCA 위반, 불공정 경쟁에 관여했다고 주장하였다. 원고들은 OpenAI가 고의로 저작물의 일부 또는 요약을 저작자 표시 없이 출력하도록 ChatGPT를 설계했으며, 도용한 글과 아이디어를 저작자 표시 없이 복제하여 상용 제품을 개발함으로써 부당하게 이익을 취하고 그 공로를 가져갔다고 주장하였다.

바. Kadrey, et al. v. Meta Platforms, Inc. (2023. 7. 7.)¹⁵⁾

세 명의 원고인 Sarah Silverman, Christopher Golden 그리고 Richard Kadrey는 7월 7일 캘리포니아 북부 연방법원에 Meta 플랫폼을 상대로 별개이지만 매우 유사한 소송을 제기하였고, 여기서 페이스북과 인스타그램의 소유주인 메타 플랫폼이 대규모 언어 모델인 LLaMA¹⁶⁾를 통해 저작권법을 위반하고 있다고 주장하면서 소송을 제기하였다. 원고에 따르면, 저작권이 있는 많은 책들이 LLaMA 훈련의 하나로 복사 및 수집된 EleutherAI라는 연구 기관에서 수집한 데이터 세트에 포함되어 있었다. 2023년 11월 9일, 미국 캘리포니아 북부 지방법원의 Vince Chabri 판사는 신청 심리를 통해, 명백히 무가치하다고 판단한 대위 저작권 침해와 메타의 생성 AI 모델 LLaMA의 결과물 형태의 저작권 침해를 포함한 원고의 저작권 청구 대부분을 기각할 것이며, 이는 원고가 LLaMA의 결과물이 자신의 저작물과 실질적으로 유사하다고 주장하지 않았기 때문이라고 하였다. 그런 후, 판사는 Kadrey 부부가 완전히 패소한 것은 아니지만, 원고 측에 소장을 수정할 수 있는 시간을 주겠다고 밝혔다.

사. Silverman, et al. v. OpenAI, Inc. (2023. 7. 7.)¹⁷⁾

앞의 저자 Paul Tremblay와 Mona Awad가 OpenAI를 상대로 제기한 소장을 참조하여, Sarah Silverman, Christopher Golden 그리고 Richard Kadrey 원고들은 ChatGPT 개발자를 상대로 저작권 직접 침해와 대위 침해 그리고 DMCA 제1202(b), 즉 저작권자 허락 없이 저작권 관리 정보를 제거한 것으로 소송을 제기하였다. 원고들은 책의 저자들이며, 자신이 출판한 책에 대해 저작권을 등록하였고, 저작권이 있는 책을 ChatGPT의 교육 자료로 사용하는 데 동의하지 않았다. 그럼에도 그들의 저작권이 있는 자료가 수집되어 ChatGPT 교육에 사용되었다. 그리고 원고들은 ChatGPT가 사용된 작품들의 저작권 관리 정보를 의도적으로 제거한 것은 생성된 콘텐츠가 저작권 침해 파생물이라는 사실을 은폐하고자 한 것이라 주장하였다. 이에 대하여, 피고는 생성형 AI가 학습용 자료를 이용하는 것은 예술적, 과학적 발전을 촉진하려는 저작권법의 궁극적 목적과도 일치하며, 근무 시간 단축, 생산성 향상 등의 잠재적 가치를 가지므로 공정 이용의 하나로 취급해야 한다고 반박하였다.

15) 23-cv-03417-VC (N.D. Cal. 2023).

16) LLaMA는 Large Language Model Meta AI의 약자로 메타가 공개한 대규모 AI 언어 모델이다. <<https://modulabs.co.kr/blog/llama-2-intro/>, 2024. 1. 29. 최종방문>.

17) 3:23-cv-03416, (N.D. Cal. 2023).

아. Chabon v. OpenAI, Inc. (2023. 9. 8.)¹⁸⁾

저자 Michael Chabon, David Henry Hwan, Matthew Kam, Rachel Louise Snyder 그리고 Ayelet Waldman은 출판된 저작물에 대한 저작권을 보유한 저자 집단을 대표하여 OpenAI를 상대로 소송을 제기하였다. Chabon 등은 OpenAI가 ChatGPT 제품을 구동하는 GPT 모델을 훈련하는 데 사용하는 데이터 세트에 자신들의 저작권이 있는 저작물을 포함하였다고 주장하였다. 원고들에 따르면, 문제는 ChatGPT가 메시지를 표시할 때 요약뿐만 아니라 저작권이 있는 저작물에 존재하는 주제에 대한 심층 분석도 생성하는데, 이는 기본 GPT 모델이 원고들의 저작물을 사용하여 훈련된 경우에만 가능하다는 것이다. 원고들은 자신들의 저작권이 있는 저작물을 GPT 모델의 훈련 자료로 사용하거나 ChatGPT와 함께 사용하는 것에 동의하지 않았다고 주장하며, OpenAI가 ChatGPT를 운영함으로써 원고들의 저작권이 있는 저작물을 무단으로 사용하였고, 이를 통해 막대한 상업적 이익을 얻고 있다고 주장하였다.

자. Authors Guild, et al. v. OpenAI, Inc. (2023. 9. 19.)¹⁹⁾

작가 조합 및 John Grishman과 George R.R. Martin을 포함한 수십 명의 작가들은 다수의 OpenAI 법인들을 상대로 소설 작가와 저작권 소유자의 권리를 침해하고, 똑같은 않아도 유사한 피해를 주는 대규모 저작권 침해에 조직적으로 관여했다고 하면서 소송을 제기하였다. 9월 19일 미국 뉴욕 남부 지방법원에 제출한 소장에서 광범위한 소설 작품의 저자들인 원고들은 허락이나 대가 없이, 해당 작품을 대량으로 복제함으로써 OpenAI가 등록 저작권에 대한 명백하고 유해한 침해 행위를 하였으며, 이에 대한 구제를 요청한다고 하였다. 구체적으로, 원고들은 OpenAI가 자신의 책 텍스트가 포함된 데이터 세트를 이용해, 그들의 저작물을 사용자의 프롬프트와 쿼리(queries)²⁰⁾에 대응하여 사람이 볼 수 있는 텍스트 응답을 출력하도록 설계된 알고리즘인 대규모 언어 모델에 공급하였으며, 이것이 광범위한 상업적 기업의 핵심이라고 주장하였다. OpenAI 모델은 원고들의 저작물을 기반으로 하거나, 모방하거나, 요약하거나, 의역한 자료인 2차적 저작물을 만들었고, 이는 원고들의 저작물 시장에 해를 끼칠 수 있기에, 작가들에게 돈을 주고 창작하는 텍스트를 누구나 자동으로 무료로(또는 매우 저렴하게) 생성할 수 있게 한다는 점에서 소설 작가들의 생계를 위협하고 있다고 주장하였다. 그리하여, 원고들은 저작권 직접 침해, 대위 침해 및 기여 침해의 저작권 간접 침해에 대한 청구를 제기하였다.

18) 3:23-cv-04625, (N.D. Cal. 2023).

19) 1:23-cv-08292-SHS, (S.D.N.Y. 2023).

20) 쿼리(query)란 데이터베이스에 정보를 요청하는 것으로, 웹 서버에 특정한 정보를 보여달라는 웹 클라이언트 요청(주로 문자열을 기반으로 한 요청)에 의한 처리이다. 쿼리는 대개 데이터베이스로부터 특정한 주제어나 어귀를 찾기 위해 사용된다. <<https://blog.naver.com/PostView.nhn?blogId=r1arbtjq7913&logNo=221805728231>, 2024. 1. 29 최종방문>.

차. Concord Music Group, Inc. v. Anthropic PBC (2023. 10. 18.)²¹⁾

여러 음악 출판사는 저작권이 있는 노래 가사를 체계적이고 광범위하게 침해했다는 이유로 AI 스타트업인 Anthropic PBC를 상대로 소송을 제기하였다. 미국 테네시주 중부 지방법원에 제출된 소장에 따르면, Anthropic은 AI 모델을 구축하고 운영하는 과정에서 원고인 출판사가 소유하거나 관리하는 수많은 음악 작품의 가사를 포함하여 방대한 양의 저작권이 있는 저작물을 불법적으로 복제하고 유포했다고 하였다. Concord 뮤직 그룹에서 Universal 뮤직에 이르는 원고들은 혁신을 수용하고, 윤리적이고 책임감 있게 사용될 때 AI의 큰 가능성은 인정하지만, Anthropic은 체계적이고 광범위하게 이러한 원칙을 위반하고 있다고 주장하였다. 특히, Universal은 Anthropic이 인터넷에서 수집한 수많은 저작물 중에서 저작권을 소유하거나 관리하고 있는 수많은 음악 작곡의 가사를 사용하여 AI 모델을 구축하였다고 주장하였다. 원고들은 이러한 저작권이 있는 자료는 단순히 인터넷에서 찾을 수 있다는 이유만으로 무료로 사용할 수 있는 것이 아니라고 주장하였으며, Anthropic이 귀중한 저작권이 있는 저작물을 이러한 방식으로 사용하기 위해 게시자의 허락을 구하거나 확보하지 않았다고 주장하였다. 저작권으로 보호되는 노래 가사 중에는 Rolling Stones, Garth Brooks, Katy Perry, Gloria Gaynor와 같은 아티스트의 가사가 포함되어 있으며, Anthropic의 모델은 각 곡에 대한 출판사의 저작권이 있는 가사와 동일하거나 실질적으로 그리고 현저하게 유사한 방식으로 저작권을 침해했다고 주장하였다. 저작권 직접 침해, 대위 침해, 기여 침해, 저작권 관리 정보 삭제에 대한 DMCA 청구 및 구제 요구와 관련하여, 원고는 Anthropic이 다른 수많은 기술 회사가 준수하는 것처럼, 정립된 저작권법을 준수해야 한다고 주장하였다.

카. Sancton v. OpenAI Inc., Microsoft Corporation, et al.(2023. 11. 21.)²²⁾

기자인 Julian Sancton은 거대 기술 기업들이 허락 없이 인류의 저작물을 도용하여 수백억 달러에 달하는 사업을 구축했다고 주장하며, OpenAI와 Microsoft를 상대로 소송을 제기하였다. 11월 21일 미국 뉴욕 남부 지방법원에 제출한 소장에서, Sancton은 피고들이 지식 재산에 대한 대가를 지급하기보다는 저작권을 보호하는 법률이 존재하지 않는 것처럼 행동했다고 주장하였다. 그런 지식 재산 중에는 피고들이 AI 모델을 훈련하는 데 사용했다고 주장하는 논픽션 서적을 포함한 수천, 어쩌면 그 이상의 다른 저작권이 있는 저작물과 함께 Sancton의 저서인 '지구 끝의 매드하우스(Madhouse at the End of the Earth)'가 있었다. Sancton은 미국 헌법이 수년에 걸쳐 창작물을 구상하고 연구하고 집필하는 논픽션 작가들을 포함하여, 창작자에게 저작물에 대한 보상을 받도록 하는 기본 원칙을 보호하고 있다고도 하였다. Sancton은 저작권 직접 침해와 기여 침해에 대해 청구하였고, 그 청구의 중심에는 OpenAI 플랫폼의 근간이 저작물에 대한 만연한 도용이라고 하였다.

21) 3:23-cv-01092, (M.D. Tenn. 2023).

22) 1:23-CV-10211 (S.D.N.Y. 2023).

다. New York Times Company v. Microsoft Corp. and OpenAI, et al. (2023. 12. 27.)²³⁾

뉴욕타임스는 OpenAI와 그 파트너인 Microsoft를 저작권 침해, DMCA 위반, 도용에 의한 불공정 경쟁, 상표권 희석 등의 혐의로 고발하였다. 뉴욕타임스는 12월 27일 미국 뉴욕 남부 지방법원에 제출한 소장에 서, 피고들이 뉴욕타임스와 경쟁하는 AI 제품을 만들기 위해 뉴욕타임스의 저작물을 불법적으로 사용했으며, 이는 신뢰할 수 있는 정보, 뉴스 분석 및 논평을 제공하는 뉴욕타임스의 능력을 위협하고 있다고 주장하였다. 피고의 생성형 AI 도구는 저작권이 있는 수백만 건의 더 타임즈 뉴스 기사, 심층 조사, 오피니언 기사, 리뷰, 사용법 가이드 등을 복사하고 사용하여 구축한 대규모 언어 모델에 의존하고 있다고 주장하였다. 더 타임즈는 피고들이 저작권이 있는 콘텐츠를 무단으로 사용하여 GenAI 모델을 훈련하는 것은 새로운 혁신적 목적을 달성하기 때문에 자신들의 행위가 공정 사용으로 보호된다고 주장하지만, 신문사는 더 타임즈의 콘텐츠를 대가 없이 사용하여 더 타임즈를 대체하는 제품을 만들고 시청자를 빼앗는 것은 혁신적인 것이 아니라고 주장하였다. 또한, 피고의 GenAI 모델의 결과물은 그 모델을 학습시키는 데 사용된 입력과 경쟁하고 이를 거의 모방하기 때문에, 이러한 목적으로 더 타임즈 저작물을 복사하는 것은 공정이용에 해당하지 않는다고 주장하였다. 저작권에 관한 주장에 더하여, 더 타임즈는 상표권 희석에 관한 소송을 제기하면서 피고들이 뉴욕을 포함한 미국 전역의 사용자들에게 영리를 목적으로 GenAI를 생산하는 상거래와 관련하여, 자신들의 GPT 기반 제품으로 생성된 결과물에 더 타임즈의 상표를 무단으로 사용했다고 주장하였다. 특히, 피고들이 품질이 낮고 부정확한 글씨로 더 타임즈의 상표를 무단으로 사용함으로써, 상표법을 위반하여 더 타임즈 상표의 품질을 희석했다고 주장하였다.²⁴⁾ 이와 관련하여 뉴욕타임스는 피고 모델들이 허락이나 보상 없이 뉴욕타임스의 콘텐츠를 복사, 복제, 의역하는 동시에, 뉴욕타임스가 게시하지 않은 콘텐츠를 뉴욕타임스에 잘못 귀속시켜 잘못된 정보를 제공함으로써, 뉴욕타임스에 상업적 및 경쟁적 손해를 입히고 있다고 주장하였다. 더 타임즈는 새로운 디지털 제품에 콘텐츠를 사용할 수 있도록 피고들과 협상을 통해 합의에 도달하려고 시도했지만, 협상이 되지 않았다고도 주장하였다.

파. Basbanes v. Microsoft Corp. and OpenAI, et al. (2024. 1. 5.)²⁵⁾

저널리스트인 Nicholas Basbanes와 Nicholas Ngagoyeanes(예명 Nicholas Gage)는 피고들이 현재 수십억 달러의 가치를 지닌 대규모 상업 기업을 구축하기 위해 자신들의 작품을 을 복제했다고 주장하면서, Microsoft Corporation, OpenAI, 그리고 OpenAI의 여러 계열사를 상대로 저작권 직접 침해, 대위 침해 그리고 기여 침해 소송을 제기하였다. 특히, 원고들은 Microsoft와 OpenAI가 자신들의 대규모 상업적 기업을 운영하기 위해 원고의 저작물을 도용하는 의도적 전략을 갖고 있었으며, 그들의 거대 언어 모델을

23) 1:23-cv-11195, (S.D.N.Y. 2023).

24) 15 U.S.C. § 1125(c).

25) 1:24-cv-00084, (S.D.N.Y. 2024).

만드는 입력에 대해 비용을 지급하지 않고, 2차적저작물을 만들어 훨씬 높은 이익을 얻는 결과를 가져왔다고 주장하였다. Basbanes와 Ngagoyeanes는 이 소송을 제기하게 된 계기로 마이크로소프트와 OpenAI를 상대로 제기한 소송을 지적하면서, 뉴욕타임스가 이 법원에 같은 피고들을 상대로 소송을 제기한 직후, 피고들은 콘텐츠 제작자와 소유자의 권리를 존중하며, AI 기술과 새로운 수익 모델의 혜택을 누릴 수 있도록 그들과 협력하기 위해 노력하고 있다며 원고와 같은 저작권 소유자가 피고들의 저작물 사용에 대해 보상을 받아야 한다는 사실을 공개적으로 인정했다고 주장하였다. 이러한 배경에서 원고들은 피고들에 의해 저작물을 도용당한 작가 집단을 대표하고자 한다고 밝혔으며, 피고가 작가들에게 저작권 침해, 저작물 라이선스 기회의 상실을 하도록 만들었고, 시장도 파괴하고 있어 그에 대한 손해배상을 요구하였다. 원고들은 또한 이러한 피해의 재발을 방지하기 위해 영구적인 금지 명령을 구하였다.

하. Main Sequence, et al. v. Dudesy LLC, et al. (2024. 1. 25.)²⁶⁾

AI로 작품을 제작하는 미디어 회사인 Dudesy가 1월 9일에 AI로 만든 George Carlin의 목소리를 모방한 1시간짜리 스페셜을 Dudesy 팟캐스트의 YouTube 채널에 공개한 후, 고인이 된 코미디언의 유족이 캘리포니아 연방 법원에 퍼블리시티권 및 저작권 침해 소송을 제기하였다. Carlin이 사망한 지 16년여가 지난 1월 25일에 접수된 이 소장에 따르면, Dudesy와 그 설립자인 코미디언 Will Sasso 그리고 작가 Chad Kultgen은 AI의 도움으로 Carlin을 부활시키기로 결심했다고 하였다. Main Sequence, George Carlin 유산의 집행인인 Jerold Hamza, 그리고 Jerold Hamza의 개인 자격(총칭하여 ‘칼린의 유족’ 또는 ‘원고들’)에 따르면, Carlin의 원본 저작물을 사용하여 Dudesy LLC, Sasso, Kultgen(총칭하여 ‘Dudesy’ 또는 ‘피고들’)은 가짜 George Carlin 코미디 스페셜의 대본을 만들고 생성된 대본에 맞춰 George Carlin과 유사한 음성을 생성하여 ‘공연’을 했다고 하였다. 원고들은 피고 중 누구도 AI로 생성된 George Carlin 스페셜에 Carlin의 초상을 사용할 수 있는 허가를 받은 적이 없으며, 고인이 된 코미디언의 저작권이 있는 자료를 사용할 수 있는 라이선스도 없다고 주장하였다. 이러한 배경에서 그들은 캘리포니아 관습법에 따른 퍼블리시티권 침해와 캘리포니아주 민사법에 따른 퍼블리시티권 박탈을 주장하고 있으며,²⁷⁾ 이들은 Dudesy가 Carlin의 이름, 명성 및 초상, 즉 Carlin의 생성된 이미지, 목소리 및 무대에 있는 모습을 연상시키도록 고안된 이미지를 사용한 것에 대해 문제를 제기하였다. 원고들은 또한, 피고들이 원고의 저작물(즉, Carlin의 스탠드업 코미디)을 모방하기 위한 결과물을 생성할 목적으로 데이터 세트를 구축하고 훈련하기 위해 원고의 저작물을 불법적으로 사용했다고 주장하며, 연방 저작권 침해에 대한 청구를 제기하였다. 또한, 원고들은 금전적 손해배상과 더불어 Dudesy와 공동 소유주가 직접적으로 저작권 침해 행위를 저지르거나 이를 방조하는 것을 금지하는 예비적 및 영구적 금지 명령 구제를 구하였다. George Carlin의 저작권이 있는 저작

26) 2:24-cv-00711 (C.D. Cal. 2024).

27) Cal. Civ. Code § 3344.1.

물을 직접 사용하거나, 침해를 방조하거나, 장려하거나, 가능하게 하거나, 유도하거나, 유발하거나, 실질적으로 기여하거나, 기타 방식으로 Dudesy 스페셜 및 해당 Dudesy 스페셜과 관련하여 Dudesy가 제작하거나 유포한 기타 콘텐츠를 생성하는 것을 금지하는 것이었다. 그리고 Dudesy에 George Carlin 스페셜의 비디오 또는 오디오 사본(일부 사본 포함)이 어디에 있든 즉시 제거, 삭제 및 파기할 것을 명령하도록 요청하였다.

2. 최근 판단된 소송 – Andersen v. Stability AI Ltd. et al²⁸⁾

본 소송은 유명 아티스트, 프로그래머, 작가들이 AI 모델 학습에 저작권이 있는 자료를 사용하는 것이 적법한지에 대한 판단을 한 첫 번째 사건이며, 2023년 10월 30일, 미국 캘리포니아 북부 지방 법원은 저작권 직접 침해와 대위 침해 및 DMCA 위반에 대한 원고의 청구를 기각한 바 있다.

가. 배경

Stability AI, DeviantArt 그리고 Midjourney 모두는 AI를 사용하여 이용자가 입력한 텍스트 프롬프트를 기반으로 새로운 이미지를 생성하는 도구와 서비스를 제공한다. 원고들은 Sarah Anderson, Kelly McKernan 그리고 Karla Ortiz를 포함한 예술가들이다. 원고는 독일의 비영리 단체인 대규모 AI 오픈 네트워크(the Large-scale Artificial Intelligence Open Network, LAION)가 Stability AI의 요청에 따라 인터넷에서 50억 개 이상의 이미지(원고의 작품을 포함)를 스크랩했다고 주장하였다. 원고는 또한, ① Stability AI가 Stable Diffusion 소프트웨어 제품을 훈련하기 위한 목적으로 LAION 데이터 세트를 이용하였고, ② DeviantArt와 Midjourney가 Stable Diffusion을 자사 제품에 결합하였으며, ③ Midjourney의 AI 제품도 Stable Diffusion 훈련에 사용된 이미지의 일부로 훈련되었다고 주장하였다.

피고들에 대한 청구들은 저작권 직접 침해, 대위 침해, DMCA 위반, 캘리포니아의 퍼블리시티권 법률 위반 그리고 캘리포니아 주법에 따른 불공정 경쟁이 포함되었다. 1심에서, Orrick 판사는 소송의 전제 조건인 저작권 등록이 이루어지지 않았다는 이유로 Ortiz와 McKernan의 청구를 기각하고, Anderson의 청구 범위를 소송 제기 당시 저작권청(Copyright Office)에 등록된 이미지로 한정하였다. 또한, Orrick 판사는 저작권법에 의한 저작권 침해 규정이 불공정 경쟁 청구보다 먼저 적용되어야 하기 때문에, 불공정 경쟁에 대한 청구는 판단할 필요가 없다고 하면서 이를 기각하였으며, 다른 모든 청구에 대해 보정 허가를 내리고, 각 청구를 뒷받침하는 피고의 행위에 대해 더 높은 수준의 구체성에 따라 제출하도록 경고하였다.

28) 3:23-cv-00201 (N.D. Cal. Oct. 30, 2023).

[그림 2] Sarah Anderson의 작품(왼쪽)과 Stable Diffusion의 그림(오른쪽)²⁹⁾



나. 저작권 직접 침해

1) 저작권 직접 침해 요건

저작권 직접 침해 청구는 ① 침해된 저작물에 대한 저작권 보유와 ② 침해자가 저작권법에 의해 부여된 배타적 권리 중 하나 이상을 침해하여야 한다.³⁰⁾ 저작권 청구는 구체적으로 명시될 필요가 없으며,³¹⁾ 직접 침해는 의도나 어떤 특정한 정신 상태를 요구하지는 않는다.³²⁾ 따라서 저작권 침해는 엄격 책임의 불법 행위이고, 인식에 대한 부재는 책임을 제한하지 않고, 단지 손해배상에만 관련된다.³³⁾ 여기서, 원고에 의한 저작권 보유, 해당 법령에 따른 등록 그리고 피고에 의한 침해들은 충분히 판단되었다. 하지만 Midjourney는 침해된 작품들을 구체적으로 밝히지 않았기 때문에, 원고인 Sarah Anderson이 저작권 등록 요건을 만족하지 못하였다고 주장하였다. 하지만 그 기준이 그렇게 높지는 않다.³⁴⁾ 원고들은 피고의 웹사이트에 “수백, 수천 장의 침해 사진”이 존재한다고 주장하였고, 이러한 대규모 침해에서 모두 진술하도록 하는 것은 적절치 않다고 하면서, 법원은 원고의 주장이 충분하다고 판단하였다.

29) Sarah Anderson, “The Alt-Right Manipulated My Comic. Then A.I. Claimed It.”, 2022. 12. 31. <<https://www.nytimes.com/2022/12/31/opinion/sarah-andersen-how-algorithm-took-my-work.html>>, 2023. 12. 30. 최종방문).

30) 17 U.S.C. § 106.

31) Perfect 10, Inc. v. Cybernet Ventures, Inc., 167 F. Supp. 2d 1114, 1120 (C.D. Cal. 2001).

32) Fox Broad. Co. Inc. v. Dish Network, L.C.C., 905 F. Supp. 2d 1088, 1098 (C.D. Cal. 2012).

33) ITC Textile Ltd. v. Wal-Mart Stores Inc., Case No. CV-12-2650-JFW (AJWx), 2015 WL 12712311, at *5 (C.D. Cal. Dec. 16, 2015).

34) Microsoft Corp. v. My Choice Software, LLC, No. SA CV 16-2189-DOC (KESx), 2017 WL 5643210, at*3 (C.D. Cal. Oct. 10, 2017).

다음으로, 저작권자는 등록된 작품에 대해 배타적 권리, 즉 저작물에 대한 복제권, 배포권, 전시권, 공연권, 2차적저작물 작성권 등을 보유한다. 저작물의 복제는 침해이고, 특정 저작물에 대한 침해를 보이기 위해서는, 피고가 침해 작품을 만들기 위해 원고의 작품 중 보호되는 부분을 복제했다는 것을 주장해야만 한다.³⁵⁾ 실질적 유사성은 저작권 침해 청구의 요건이 아니며, 전체적으로 복제하지 않고 원본의 요소를 이용한 때에, 원본의 구성 요소들에 대한 복제가 실제로 발생했는지를 법원이 판단하는 데 도움을 줄 뿐이다.³⁶⁾ 실질적으로, 실질적 유사성은 다음 두 개의 구별되는 개념을 집합하여 지칭하기 위해 사용되는 언어이다: ① 독립적 창작에 대응하는 것으로써, 피고가 저작물을 복제했다는 의미에서의 ‘복제’ 그리고 ② 보호될 수 없는 요소에 대응하는 것으로써, 복제가 저작권에 의해 보호되는 자료에 도달했다는 의미의 ‘불법적 이용’.³⁷⁾ 원고는 Midjourney의 주장에 대응하기 위해 하나의 유효한 이론만 청구하면 되지만, 그 청구에는 Midjourney에 적용될 수 있는 직접 침해의 몇몇 사실들을 명시하고 있었다: ① Midjourney 제품에 대한 학습 이미지로서 사용되는 이미지들을 스크래핑하고 복제함으로써 보호되는 저작물을 복제하여 발생한 직접 침해, ② Midjourney 제품의 일부로서, 학습 이미지들의 압축 사본을 포함하는 Stable Diffusion을 배포하여 발생한 직접 침해, ③ Midjourney 제품을 만들고 배포하여 발생한 직접 침해, ④ 2차적저작물을 침해하는 출력 이미지를 생성하고 배포하여 발생한 직접 침해.

2) 저작권 직접 침해 판단

먼저, Midjourney가 AI 제품을 훈련하기 위해 원고의 저작물을 복제하여 직접적으로 침해하였는지를 살펴본다. 복제권의 침해는 행위가 장치 또는 제품을 이용하도록 하거나 침해를 위해 이용되는 서비스를 제공하는 것과 관계없이, 저작권으로 보호되는 표현의 허락받지 않은 복제를 의미한다.³⁸⁾ 원고들은 Midjourney가 창작자의 허락 없이 AI 제품 학습을 위해 복제된 이미지를 사용했음을 보여주는 구체적인 사실을 제시했으며, Midjourney는 이를 인정하였다. Midjourney는 이러한 압축 사본을 Midjourney 제품 학습에 사용하였고, 이러한 주장을 사실로 간주하면, Midjourney가 원고의 배타적 복제권을 침해했다는 주장이 성립하게 된다. 이러한 주장에도 불구하고 Midjourney는 소장에서 “학습 자료의 ‘하위 집합’에 무엇이 포함되어 있는지, 또는 원고의 저작물이 포함되어 있는지”를 충분히 명시하고 있지 않다고 반박하였다. 하지만 Midjourney는 교육 이미지의 출처에 대한 직접적인 질문에 대해 회피하고 있었다. 따라서 법원은 이미 알려진 사실이나 추측을 통해 합리적인 추론을 할 수밖에 없었다. 원고는 Midjourney 제품이 Stable Diffusion에 의존하고 있으며, Stable Diffusion 학습 풀에 스크랩된 수십억 개의 이미지 중 상당수가 원고의 라이선스 자료가 포함된 데이터 세트에서 나왔다고 주장하였다. 이러한 주장은 Midjourney 제품의 교

35) Skidmore v. Led Zeppelin, 952 F.3d 1051, 1064 (9th Cir. 2020).

36) Range Rd. Music, Inc. v. E. Coast Foods, Inc., 668 F.3d 1148, 1154 (9th Cir. 2012); Briggs v. Blomkamp, 70 F. Supp. 3d 1155, n.3, 1163-64 (N.D. Cal. 2014).

37) Skidmore, 952 F.3d at 1064.

38) Columbia Pictures Indus., Inc. v. Fung, 710 F.3d 1020, 1033 (9th Cir. 2013).

육 자료에 실제로 원고의 저작물이 포함되어 있다는 합리적인 추론을 가능하게 하였고, 실제로 이러한 추론은 이 사건과 같이 대규모 침해와 관련된 저작권 침해 사건에서 허용되었다.

다음으로, Stable Diffusion이 압축 사본을 배포하여 직접적으로 저작권을 침해하므로 Midjourney의 AI 제품도 마찬가지로 살피본다. 배포에 의한 침해는 “판매 또는 기타 소유권 이전, 대여, 임대 또는 대여에 의한 ‘사본의 실제 배포’를 필요로 한다.”³⁹⁾ Stability가 Stable Diffusion 내에 학습 이미지의 압축 사본을 내장하고 저장했기 때문에, 원고는 피고가 원고의 작품을 직접 복제했다고 주장하였다. 실제로, Stable Diffusion과 이를 통해 만들어진 압축 사본은 Midjourney 제품의 기초를 형성한다. 원고들은 또한, 사용자가 비용을 지불하고 Midjourney 제품을 사용할 수 있도록 함으로써 이러한 사본을 고객에게 배포한다고 주장하였다. 원고가 직접 복제를 주장하였으므로 실질적 유사성은 필요하지 않았다.

그다음은 Midjourney 제품이 2차적저작물을 직접적으로 침해하고 있는지를 살펴본다. 2차적저작물로 만족되기 위해서는 원고가 그 작품이 기존 저작물을 “재구성(recast)하고, 변형(transform) 또는 개작(adapt)하였음”을 주장하여야 한다.⁴⁰⁾ 더 나아가 2차적저작물이 침해하고 있다는 것을 주장하기 위해, 원고는 피고들이 침해하는 2차적저작물을 만들기 위해 원고의 작품 중 보호되는 부분을 복제하였음을 주장하여야만 한다.⁴¹⁾ 이는 실질적 유사성을 입증함으로써 추론적으로 증명할 수 있지만, 이 사건과 같이 저작권이 있는 저작물의 직접 복제를 수반하는 경우에는 전체 저작권이 있는 저작물의 복제가 반드시 저작물의 보호되는 부분에 대한 복제가 수반하기 때문에 관련이 없다. 원고들이 주장하는 바와 같이, “Stability는 수십억 개의 저작권이 있는 이미지의 사본을 다운로드하거나 다른 방법으로 취득”하여 Stable의 ‘학습용 이미지’로 사용했다. 이것들을 종합해 보면, Stability는 수많은 저작물을 복제했다는 것을 알 수 있고, Stable Diffusion 및 이에 의존하는 제품은 “수십억 개의 저작권이 있는 이미지”의 “압축된 사본을 저장”함으로써 원고의 저작물을 다른 저작물로 ‘변형’시켰으므로 2차적저작물에 해당한다. 따라서 Midjourney 제품은 Stable Diffusion에 의존하기 때문에 2차적저작물에 해당한다.

마지막으로, Midjourney 제품의 출력 이미지들이 2차적저작물을 침해하고 있는지를 살펴본다. 유사한 이유로, Stable Diffusion과 DreamStudio의 출력 이미지들은 또한 2차적저작물을 직접적으로 침해하고 있다. 여기서, 원고들은 “Stable Diffusion이 수학적 소프트웨어 공정을 통해 새로운 이미지를 생산하기 위해 학습 이미지들을 사용”하고 있다고 주장한다. 이러한 새로운 이미지들은 전적으로 학습 이미지들을 기반으로 하고, Stable Diffusion이 주어진 결과물을 조합할 때 사용하는 특정 이미지의 2차적저작물이다. 원고들은 또한 각 출력 이미지가 저작권이 있는 이미지의 사본인 학습용 데이터와 잠재적 이미지의 조합

39) 17 U.S.C. § 106(3).

40) 17 U.S.C. § 101; *Jarvis v. K2 Inc.*, 486 F.3d 526, 531 (9th Cir. 2007).

41) 17 U.S.C. § 106(2).

로만 생성되기 때문에 필연적으로 2차적저작물이라고 주장하였다. 원고들은 Stable Diffusion 내에서 학습 이미지의 사본이 보관(報簡), 즉 수학적으로 혼합되어 새로운 2차적저작물 이미지를 생성한다고 주장하였다. 원고들은 출력 이미지를 만들기 위해 ‘소스 이미지 자체가 직접 보관’되고, Stable Diffusion이 소스 이미지를 출력 이미지로 결합하는 방법을 보여주는 다이어그램을 제공한다고 주장하였다. 이 사건은 “저작권이 있는 저작물의 직접 복제”와 관련되어 있으므로, 원고들은 출력 이미지가 저작권법상 2차적저작물을 침해하고 있다고 충분히 주장하고 있다. Midjourney는 이러한 제품을 제작하고 마케팅하는 과정에서 자체 AI 이미지 생성기를 사용하여 출력 이미지를 만들었다. 이 경우 직접 복사를 한 것이기 때문에 출력물이 입력과 다른지 여부는 중요하지 않다.⁴²⁾

3) 실질적 유사성 테스트

제9순회법원의 수많은 법원이 결론을 내린 바와 같이, 원고가 실질적 유사성을 입증해야 한다는 피고의 주장은 잘못된 주장이다.⁴³⁾ 왜냐하면 중요한 쟁점은 침해가 발생하였는지이기 때문이다. 위에서 살펴본 바와 같이, 원고들은 자신의 저작권이 있는 저작물을 허락 없이 직접 복제함으로써 직접 침해를 했다고 주장하였다. 따라서 피고들이 실질적 유사성이 없음을 입증하더라도 이는 관련이 없다. 원고들은 특정 텍스트 프롬프트에 대응하여 제공된 Stable Diffusion의 출력 이미지 중 어느 것도 학습 데이터의 구체적 이미지와 거의 일치하지 않을 가능성이 있다고 주장하였다. 원고는 또한, 출력 이미지가 학습 이미지에서 수학적으로 혼합된 2차적저작물을 침해하고 있다고 주장하였다. 이 과정의 자연스럽고 예상되는 결과는 특정 출력 이미지가 ‘학습 데이터의 특정 즉, 단일 이미지’와 ‘거의 일치하지 않을 가능성이 있다’는 것이다. 그런데, 입증된 바와 같이, 혼합된 소스 이미지와 출력 이미지 사이에 상당한 유사성이 존재한다고 추론하는 것이 합리적이다. 이는 원고의 주장이 “특정 텍스트 프롬프트에 대한 응답으로 제공되는 Stable Diffusion 출력 이미지 중 어느 것도 훈련 데이터의 특정 이미지와 거의 일치하지 않을 가능성이 있다”는 원고의 주장과 전적으로 일치하는 이유이기도 하다. 원고들은 또한 Stable Diffusion의 사본을 갖는 것은 “컴퓨터에 수십억 개의 JPEG 파일이 있는 디렉터리를 갖는 것과 비슷하다”고 주장하였다. 따라서 법원은 컴퓨터 디렉터리를 열어 JPEG 파일을 검사하고 원본 저작물과 비교할 수 있는 것처럼, Stable Diffusion을 유사하게 검사하여 Stable Diffusion 내의 ‘저장된 압축 사본’과 원본 저작물 간의 실질적 유사성을 밝힐 수 있다고 합리적으로 추론할 수 있다. 이러한 주장은 실질적 유사성을 충족한다.

42) ITC Textile, 2015 WL 12712311, at *5 (“직접 복제의 경우 피고 저작물의 최종 결과물이 원고의 저작물과 다르다는 사실은 면책 사유가 되지 않는다. 반대로, 이는 무단 복제 및 2차적저작물 무단 작성과 같은 여러 배타적 권리의 침해를 입증할 수 있다.”).

43) Range Rd. Music, 668 F.3d at 1154.

다. 대위 침해

저작권 대위 침해에 대한 청구는 ① 제3자에 의한 직접침해, ② 침해를 막거나 제한할 수 있는 권리 ③ 그 침해 행위로부터의 직접적 경제적 이익의 3가지를 요구한다.⁴⁴⁾ 먼저, 원고들은 출력 이미지들이 2차적 저작물을 어떻게 침해하고 있는지를 이미 보여주었다. 원고들은 또한 이용자에게 의한 제3자 침해를 주장하였다. 원고들은 이용자에게 의한 프롬프트로부터 이미지를 생산하기 위해 사용될 때, Stable Diffusion은 수학적 소프트웨어 공정을 통해 새로운 이미지를 생산하는 학습 이미지를 사용했다고 주장하였다. 이러한 새로운 이미지들은 전적으로 학습 이미지들에 기반을 두고 있고, 주어진 출력물을 조합할 때, Stable Diffusion이 수집한 특정 이미지들의 2차적저작물이다. 따라서 Midjourney 제품 사용자가 Midjourney에 출력 이미지를 요청하는 경우, 해당 출력 행위는 제3자에 의한 저작권 침해 행위가 된다.

다음으로, Midjourney는 이용자들이 저작권을 침해하는 2차적저작물을 만들지 못하도록 막을 수 있는 법적 권리를 갖는다고 원고들이 주장하는데 이는 잘못된 것이라고 하였다. 원고들은 “Midjourney의 주요 제품은 온라인 AI 기반 이미지 생성기”라고 주장하였고, Midjourney는 2022년 3월 첫 번째 서비스 버전을 배포했고, 그 제품을 계속 업데이트하였다. Midjourney가 AI 이미지 생성기를 디자인하고, 마케팅하고, 판매하였기 때문에, 피고들이 각 침해를 막는 권리를 항상 갖고 있다는 것은 필연적으로 사실이다. 예로써, Midjourney는 애초에 이러한 생성기를 대중에게 공개하지 않았을 수도 있다. 또는 공개하더라도 Midjourney가 저작권을 침해하는 2차적저작물이 나오지 않도록 할 수 있었을 것이다.

마지막으로, 피고는 원고가 침해 행위로 인한 금전적 이해관계를 주장하지 못했다고 주장하는데, 이도 잘못되었다. 원고는 Midjourney가 월 30달러부터 시작하여 텍스트 프롬프트와 디지털 이미지 출력을 무제한으로 허용하는 다양한 유료 구독 플랜을 제공한다고 주장하였다. 따라서 Midjourney는 사용자가 Midjourney 제품을 통해 저작권을 침해하는 출력 이미지를 더 많이 생성함으로써 직접적인 금전적 이익을 얻었다.

라. DMCA 위반

1) DMCA 위반 개요

DMCA는 저작권 관리 정보(copyright management information, CMI)의 제거나 교체에 대한 책임을 부과한다.⁴⁵⁾ 또한, 저작권 관리 정보의 보호되는 카테고리를 설명하는 조항은 넓게 해석되어 왔으며,⁴⁶⁾

44) Perfect 10, Inc. v. Visa Int'l Ass'n, 494 F.3d 788, 802 (9th Cir. 2007).

45) 17 U.S.C. § 1202(b).

46) 17 U.S.C. § 1202(c); Williams v. Cavalli, No. CV 14-06659-AB (JEMx), 2015 WL1247065, at *3 (C.D. Cal. Feb. 12, 2015); Leveyfilm, Inc. v. Fox Sports Interactive Media, LLC, 999 F. Supp. 2d 1098, 1102 (N.D. Ill. 2014).

DMCA의 변론 부담은 낮다.⁴⁷⁾ 책임이 부과되기 위해서는 인식이나 의도가 요구된다.⁴⁸⁾ 제9연방순회법원이 오랫동안 인정해 온 바와 같이, 구체적인 불법적 의도에 대한 직접적인 증거는 거의 없으며, 그러한 지식이나 의도는 정황 증거를 통해 추론할 수 있다.⁴⁹⁾ 제1202조(b)의 정신 상태 요건을 주장하기 위해, 사람의 의도, 인식 및 기타 정신 상태가 일반적으로 주장될 수 있다.⁵⁰⁾ 또한, DMCA 1202조 어디에도 등록이 소송의 전제 조건임을 암시하는 내용은 없다. 그리고 DMCA 위반 청구는 저작권 침해에 대한 증거를 요구하지 않는다.⁵¹⁾ 따라서 공정 이용은 방어 수단이 될 수 없다.⁵²⁾

2) DMCA 위반 요건

DMCA의 저작권 관리 정보 위반에 대한 청구는 원고들에게 ① 침해 작품에 대한 저작권 관리 정보의 존재, ② 해당 정보의 제거나 교체, 그리고 ③ 그 제거나 교체가 고의적으로 일어났다는 것을 주장하도록 요구된다.⁵³⁾ 첫째, 원고들은 자신의 저작물에 저작권 고지, 저작물 제목, 저작자 이름, 저작권자의 이름 등 저작권 관리 정보가 포함되어 있다고 구체적으로 주장하였다. 각각은 제1202조(c)에 의해 저작권 관리 정보로 인식되었다. 또한, 원고들은 저작권 관리 정보가 자신들의 저작물에 직접 통합되었다고 주장하였다.

둘째, 원고들은 Midjourney가 원고의 저작물에서 저작권 관리 정보를 고의로 제거하거나 변경했다고 주장하였다. 첫째, 피고들은 Stable Diffusion을 원고들의 저작물로 학습시켰고, 출력물의 일부로 권리 관리 정보를 생략하도록 설계하였다. 원고들은 권리 관리 정보가 원본과 사본으로부터 변경되고 제거되는 배포 과정이 어떻게 재구성되었고, 배포 기술이 어떻게 개선되고 보완되었는지를 보여주었다. 여기서 Midjourney는 생성 이미지에 대한 소프트웨어 엔진으로써 Stable Diffusion에 의지하였고, Stable Diffusion을 이용한 버전을 배포하였다. 피고들은 모델 훈련에 사용된 이미지에 권리 관리 정보가 포함되어 있으며, 훈련 시 또는 모델이 출력될 때 고의로 권리 관리 정보가 제거되었다는 사실을 알고 있었다. 여기에 더해, Midjourney의 설립자이자 CEO인 David Holz는 Midjourney가 오픈 데이터 세트, 즉 라이선스 협상이 이루어지지 않은 자료가 포함된 LAION 데이터 세트를 기반으로 학습되었다고 인정하였다. 이는 변론 목적으로 충분하였다.

셋째, 원고들은 피고들이 저작권자의 권한이나 저작물에 관한 법률의 권한 없이 권리 관리 정보를 제거하거나 변경되었음을 알면서도 권리 관리 정보를 배포했다고 주장하였다. 특히, Midjourney는 서비스 약관

47) Mollman v. Zoetop Bus. Co., No. CV 22-4128 (PA) (GJSx), 2022 WL 17207103, at *3 (C.D. Cal. Sep. 16, 2022).

48) 17 U.S.C. § 1202(b); Merideth v. Chi. Trib. Co., LLC., No. 12 C 7961, 2014 WL 87518, at *2 (N.D. Ill. Jan. 9, 2014).

49) Friedman v. Live Nation Merch., Inc., 833 F.3d 1180, 1189 (9th Cir. 2016).

50) Logan v. Meta Platforms, Inc., __ F. Supp. 3d __, 2022 WL 14813836, at *9 (N.D. Cal. Oct. 25, 2022).

51) MDY Indus., LLC v. Blizzard Entm't, Inc., 629 F.3d 928, 950-52 (9th Cir. 2010).

52) Murphy v. Millennium Radio Grp., LLC, Civil Action No. 08-1743 MAS, 2015 WL 419884, at *5 (D.N.J. Jan. 30, 2015).

53) O'Neal v. Sideshow, Inc., 583 F. Supp. 3d 1282, 1286-87 (C.D. Cal. 2022).

에서 “귀하의 지시에 따라 서비스에서 생성된 자산에 대한 취소 불가능한 저작권 라이선스”를 보유하고 있다고 명시하고 있다.

넷째, 원고들은 피고들이 저작물의 사본을 배포했다고 주장하였다. 예를 들어, 피고들은 학습 이미지를 생성 및 저장하고 이러한 사본을 ‘학습 또는 기타 AI 이미지 제품 생성에 사용하기 위해’ 배포하였다.

다섯째, 원고들은 피고들이 자신의 배포가 직접 또는 사용자들에 의해 침해를 유도, 활성화 또는 은폐할 것이라는 사실을 알았거나 알 만한 합리적인 근거가 있다고 주장하였다. 권리 관리 정보의 적절한 표시 없이 저작권이 있는 자료를 배포하면 자신의 침해 행위를 은폐할 수 있다는 사실을 인식한 것은 DMCA의 두 번째 고의 요건을 충족한다.⁵⁴⁾ 두 번째 고의의 요건은 Midjourney의 설립자이자 CEO인 David Holz의 진술에 의해 입증되었다. Holz는 Midjourney의 학습 이미지가 “인터넷에서 스크랩한 것일 뿐”이며, Midjourney는 “공개된 오픈 데이터 세트를 사용하고 이를 통해 학습하였다”고 하였다. 이러한 학습 이미지는 원고를 포함하여 라이선스받은 이미지로 구성되었다. Holz는 또한 이 프로세스가 “현재로서는 괜찮다”고 주장하며 “이에 관한 특별한 법률은 없다”고 주장하였다. 이는 Midjourney 사용자들로 하여금 자신의 저작물에 대한 원고의 권리를 무시하거나 무시하는 방식을 포함하여 자신이 원하는 방식으로 Midjourney의 결과물을 사용할 수 있다고 믿도록 유도한다. Midjourney는 출력물의 사용이 합법적이라고 거짓으로 주장함으로써 자신의 침해 행위를 은폐하고 있다. Holz의 진술 외에도 Midjourney는 Stable Diffusion을 자체 AI 이미지 제품의 기반으로 사용하기로 결정하였고, 이미지 모델의 기반이 되는 확산이 어떻게 작동하는지에 대해 잘 알고 있었기 때문에, 즉 Stable Diffusion이 오픈 소스이기 때문에 확산의 과정과 원본 소스 이미지의 복사 및 재생산에 어떻게 의존하는지에 대해 잘 알고 있었기 때문에 그렇게 하였다. 또한, 스스로를 ‘다양한 연구실’이라고 광고하고 AI 제품을 만드는 사업을 하는 Midjourney는 자사의 AI 이미지 제품에 의해 생성된 이미지에 저작권이 있는 이미지들의 권리 관리 정보가 포함되지 않는다는 사실을 분명히 알았거나 합리적으로 알았어야 한다. 또한 피고가 저작권이 있는 저작물에서 권리 관리 정보를 제거하면 “원고들과 연락하는 것이 어렵거나 불가능해지기 때문에” 피고가 DMCA에 따른 추가 침해 행위를 은폐할 수 있다는 사실을 알고 있었다고 추론하는 것이 합리적이다.

Midjourney는 원고들의 주장을 다음의 3가지 방법으로 공격하였다: ① 원고들은 권리 관리 정보가 변경되거나 제거된 어떤 작품도 밝히지 않았다; ② 원고들은 Midjourney가 구체적으로 권리 관리 정보를 제거하거나 변경하였다고 주장하지 않았다; ③ 원고들은 요구되는 고의를 적절하게 주장하지 않았다. 먼저, Midjourney는 원고들이 저작권 관리 정보를 제거한 저작물을 충분히 식별하지 못하였다고 주장하였다. 하지만 원고들은 학습 이미지에 어떤 저작물이 포함되었는지 구체적으로 식별하였다. 원고들이 주장한 바와 같이 그리고 창립자 겸 CEO가 인정한 바와 같이, Midjourney는 원고의 저작물이 포함하여 라이선스받은

54) *Mango v. BuzzFeed, Inc.*, 970 F.3d 167, 172 (2d Cir. 2020); 17 U.S.C. § 1202(b)(3).

자료들로 구성된 데이터 세트를 학습하였다. Midjourney는 또한 DMCA에 따를 때, 구체적 사례에서 복제된 저작물의 ‘정확한 사본’으로부터 제거된 것을 원고가 주장하도록 요구된다고 주장하였다. 하지만 제 1202(b)에는 그러한 요건이 존재하지 않는다.⁵⁵⁾ 어쨌든, 피고들은 청구들을 잘못 해석하였고, 피고들이 저작물을 통째로 학습 이미지로서 AI 이미지 제품에 통합했다는 원고들의 주장을 무시하였다.

다음으로, Midjourney는 원고들의 청구가 일관성이 없는 주장을 담은 산탄총(shotgun) 청구라고 주장하였다. 하지만 원고가 해야 할 일은 각 피고에 대한 청구의 근거를 주장하여 피고에게 피고에 대한 혐의를 충분히 고지하는 것뿐이다. 원고들은 특히 Midjourney에 대한 충분한 사실들을 구체적으로 주장하였다. 또한, 원고들은 Midjourney가 Stable Diffusion을 기반으로 한다고 주장하였다. Midjourney는 Stable Diffusion이 사용하고 Stability가 개발 비용을 지불한 것과 동일한 LAION 데이터 세트를 사용하였으며, 이는 Midjourney 스스로도 인정한 사실이었다. 이러한 주장은 원고의 DMCA 청구에 대한 근거를 Midjourney에게 통지하기에 충분하였다.

마지막으로, 원고의 DMCA 청구를 회피하기 위한 Midjourney의 마지막 시도도 마찬가지로 실패하였다. 원고는 “피고가 ‘저작권 관리 정보의 제거 또는 변경이 침해를 돕는다는 것을 알았거나 알 만한 합리적인 근거가 있었다’는 합리적 추론을 뒷받침할 수 있는 충분한 사실만 주장하면 되었다.”⁵⁶⁾ 변론 단계에서 정신적 상태는 구체적으로 주장할 필요가 없다. 원고들은 DMCA의 소위 ‘이중 고의’ 요건에 대한 요소들을 충분히 보여주었다. 원고들은 “피고가 해당 이미지에 대해 Stable Diffusion을 훈련시키고, 출력물에 저작권 관리 정보를 생략하도록 설계함으로써 원고들이 소유한 이미지에서 저작권 관리 정보를 제거하거나 변경했다”고 주장하였다. Midjourney는 Stable Diffusion이 이미지를 생성하는 방법을 정확히 알고 있었기 때문에, 그들 자신의 이미지 모델 제품을 Stable Diffusion에 근거하였다. 원고들은 이미지 모델의 기반이 되는 프로세스인 Diffusion이 노이즈를 추가 및 제거함으로써, 즉 저작권 관리 정보를 제거 및 변경함으로써 원본 소스 이미지를 복사 및 재구성하는 방법을 보여주었다. 원고들은 Stability가 이 프로세스에 대한 지식을 가진 사람들을 고용하여 Stable Diffusion을 구축하였다고 주장하였다. 또한 Midjourney의 설립자이자 CEO는 학습 데이터의 출처에 대해 모호한 태도를 취하고 있으며, Midjourney가 아티스트의 동의나 라이선스를 구하지 않았다는 점을 인정하였다. 실제로 Midjourney가 AI 이미지 제품의 기반이 되는 프로세스에 대한 지식이 없다고 부인하는 것은 ‘다양한 연구실’이라는 표현과 오픈 소스 코드에 기반하고 있다고 인정하는 것과 모순된다. 또한, 원고들은 “이 시스템의 출력 이미지는 원고가 제작한 이미지를 포함하여 저작권이 있는 이미지의 사본인 최신 이미지에서만 파생된 것”이라고 주장하였다. 피고는 라이선스가 부여된 자료를 학습 이미지로 포함하기로 결정하면서, 해당 이미지에 “제작자의 이름과 같은” 저작권 관리 정보가 포함되어 있거나 “아티스트의 서명 형태로 저작물에 통합”되어 있다는 사실을 알고 있었다. 피고들은 또

55) ICONICS, Inc. v. Massaro, 192 F. Supp. 3d 254, 272 (D. Mass. 2016).

56) Doe 1, 2023 WL 3449131, at *12.

한 확산 과정을 통해 저작권 관리 정보가 제거될 것이라는 사실도 알고 있었다. 피고는 “출력물의 일부로 저작권 관리 정보를 생략”하도록 AI 이미지 제품을 설계했다. 저작권 관리 정보가 없으면 사용자는 출력물의 기반이 된 원본 저작물에 대해 알 수 없으며, 이를 사용할 때 어떤 의무가 따르는지도 알 수 없다. 이러한 주장은 Midjourney가 저작권 관리 정보를 제거함으로써 침해를 유발할 상당한 위험이 있다는 것을 알았다는 것을 입증하기에 충분하다.

3. 미국의 생성형 AI와 저작권법에 관한 보고서⁵⁷⁾

2023년 9월, 의회조사국(Congressional Research Service, CRS)은 생성형 AI 결과물에 저작권이 있는지와 생성형 AI가 다른 저작물의 저작권을 어떻게 침해할 수 있는지에 관해 법원과 미국 저작권청이 직면하기 시작한 문제들을 살펴보았다.⁵⁸⁾

가. 개요

AI 분야의 혁신은 저작자, 침해, 공정이용과 같은 저작권법 원칙이 AI에 의해 만들어지거나 사용된 콘텐츠에 어떻게 적용될 것인지에 대한 새로운 문제들을 제기하고 있다. 소위 생성형 AI 컴퓨터 프로그램들(예: Open AI의 DALL-E 및 ChatGPT 프로그램, Stability AI의 Stable Diffusion 프로그램, Midjourney의 셀프-타이틀(self-titled) 프로그램)은 사용자의 문자 프롬프트(또는 ‘입력’)에 반응하여 새로운 이미지, 텍스트 및 기타 콘텐츠(또는 ‘결과물’)를 생성할 수 있다. 이러한 생성형 AI 프로그램들은 글, 사진, 그림 및 기타 예술 작품과 같은 기존의 다량 저작물에 프로그램들을 노출해 부분적으로 결과물을 생성하도록 학습된다. 생성형 AI는 저작권 침해에 대한 문제를 제기하며, 논평가와 법원은 생성형 AI 프로그램이 기존 저작물을 복제하여 AI를 학습시키거나 기존 저작물과 유사한 결과물을 생성함으로써 기존 저작물의 저작권을 침해할 수 있는지를 다루기 시작하였다.

나. AI 학습 과정이 다른 저작물의 저작권을 침해하는가?

AI 시스템은 인터넷에서 다운로드한 텍스트, 이미지 및 기타 저작물을 포함한 대량의 데이터에 프로그램을 노출해 문학, 시각 및 기타 예술 작품을 만들도록 훈련된다. 이 학습 과정은 기존 작품의 디지털 사본 제작과 관련된다. 미국 특허상표청에서 설명한 바와 같이, 이 과정은 거의 대부분 저작물 전체 또는 그 상당 부분을 복제하는 것과 관련된다. 예를 들어, OpenAI는 자사 프로그램이 “저작권이 있는 저작물이 포함된

57) Congressional Research Service, Generative Artificial Intelligence and Copyright Law, 2023 <<https://crsreports.congress.gov/product/pdf/LSB/LSB10922>, 2024. 1. 31 최종방문>.

58) 저작권이 있는지 여부에 대해서는 다음 장에서 살펴보기로 한다.

공개적으로 이용 가능한 대규모 데이터 세트”를 기반으로 학습되며, 이 과정은 “분석할 데이터의 사본을 먼저 만들어야 한다는 것”을 인식하고 있다(비록, 현재는 향후 이미지 생성 모델 학습에서 이미지를 제거할 수 있는 옵션을 제공하지만). 허락 없이 그러한 사본을 만드는 것은 저작물을 복제할 수 있는 저작권자의 배타적 권리를 침해할 수 있다.

AI 기업들은 자사의 학습 과정이 공정 이용에 해당하므로, 저작권을 침해하지 않는다고 주장할 수 있다. 복제가 공정 이용에 해당하는지는 저작권법 제107조에 따른 4가지 법적 요소에 따라 달라진다: ① 영리 또는 비영리 등의 사용의 목적과 성격, ② 저작물의 성격, ③ 저작물 전체와 관련되어 이용된 부분의 양과 상당성, ④ 저작물의 잠재적 시장 또는 가치에 대한 사용의 영향.⁵⁹⁾ 일부 이해관계인들은 AI 프로그램 학습을 위한 저작물의 이용이 이러한 요소에 따라 공정 이용으로 간주되어야 한다고 주장한다. 첫 번째 요소와 관련하여, OpenAI는 훈련 과정이 “유용한 생성형 AI 시스템”을 생성하기 때문에, 그 목적이 “표현(expressive)”이 아닌 “변형(transformative)”에 있다고 주장한다. OpenAI는 또한 세 번째 요소에 대해서는 사본이 대중에게 제공되지 않고 프로그램 훈련에만 사용되기 때문에 공정 이용을 지지한다고 주장한다. 이를 위해, OpenAI는 책에서 발췌한 내용을 표시하여 검색이 가능한 데이터베이스를 만들기 위해 구글이 책 전체를 복제한 것이 공정 이용이라는 제2회회향소법원의 Google 판결을 인용하였다.⁶⁰⁾ 네 번째 공정 이용 요소와 관련하여, 일부 생성형 AI 애플리케이션은 저작물에 대한 AI 프로그램 학습을 통해 원본과 경쟁하는 유사한 저작물을 생성할 수 있다는 우려를 제기하였다. 예를 들어, 아티스트 Drake와 The Weeknd처럼 들리도록 만들어진 “Heart on My Sleeve”라는 AI 생성 노래는 스트리밍 서비스에서 수백만 번 재생되었다. 두 아티스트와 계약을 맺은 유니버설 뮤직 그룹은 AI 기업이 이러한 아티스트의 노래를 학습 데이터에 사용함으로써 저작권을 침해하고 있다고 주장한다. OpenAI는 자사의 비주얼 아트 프로그램인 DALL-E 3가 “살아있는 아티스트의 스타일로 이미지를 요청하는 요청을 거부하도록 설계되었다”고 주장한다.

원고들은 AI 프로그램의 학습 과정이 어문 및 시각적 저작물에 대한 저작권을 침해했다고 주장하며, 여러 건의 소송을 제기하였다. 여기에는 OpenAI를 상대로 작가 연합과 Paul Tremblay, Michael Chabon, Sarah Silverman 작가들 등이 제기한 소송; Meta Platforms를 상대로 Michael Chabon과 Sarah Silverman 등이 제기한 개별 소송; Alphabet Inc.와 Stability AI 및 Midjourney를 상대로 제기한 집단 소송; 그리고 Stability AI를 상대로 Getty Images가 제기한 소송을 포함한다. 예를 들어, Getty Images의 소송은 “Stability AI가 Stable Diffusion 모델을 학습시키기 위해 Getty Images의 웹사이트에서 최소 1,200만 개의 저작권이 있는 이미지를 복사했다”고 주장하였다. 이 소송은 Stable Diffusion이 상업적 제품이기 때문에 첫 번째 법적 요소에 따라 공정 이용에 해당하지 않고, 이 프로그램이 원본 저작물의 시장을

59) 17 U.S.C. § 107.

60) Authors Guild v. Google, Inc., 804 F. 3d 202 (2d Cir. 2015).

훼손하기 때문에 네 번째 요소에 따라 공정 이용에 해당하지 않는다고 주장하며, 공정 이용의 모든 특성에 대해 이익을 제기하는 것으로 보인다.

2023년 9월, 미국 지방법원은 AI 회사가 법률 리서치 플랫폼인 Westlaw의 판례 요약물 복사하여 사용자의 질문에 대한 응답으로 법률 의견의 관련 구절을 인용하도록 AI 프로그램을 훈련시키는 것이 공정 이용에 해당하는지 여부를 판단하기 위해 배심원 재판이 필요하다고 판결하였다. 법원은 피고의 사용이 “의심할 여지 없이 상업적”이기는 하지만, 배심원단은 해당 이용이 “변형적”인지 여부(1요소), 원고의 업무 특성이 공정 이용에 어느 정도 유리한지(2요소), 피고가 AI 프로그램 훈련에 필요한 양 이상을 복사했는지(3요소), AI 프로그램이 Westlaw의 “시장 대체물”이 되는지(4요소)에 관한 사실적 분쟁을 해결해야 한다고 판단하였다. 문제가 된 AI 프로그램은 생성형 AI로 간주되지 않을 수 있지만, 생성형 AI 모델을 훈련하기 위해 복제한 경우와 동일한 종류의 사실이 법원의 공정 이용 분석과 관련이 있을 수 있다.

다. AI 결과물들이 다른 저작물의 저작권을 침해하는가?

AI 프로그램들은 기존 저작물과 유사한 결과물을 생성하여 저작권을 침해할 수도 있다. 판례들에 따르면, 저작권자들은 AI 프로그램이 ① 자신의 저작물에 접근(access)하고, ② “실질적으로 유사한” 결과물을 생성한 경우 그러한 결과물들이 자신의 저작권을 침해한다는 것을 입증할 수도 있다. 첫째, 저작권 침해를 입증하려면 원고는 침해자가 저작물을 “실제로 복사”했다는 사실을 입증해야 한다. 이는 때때로 침해자가 “저작물에 접근할 수 있었다”는 증거를 통해 정황적으로 입증되기도 한다. AI 결과물의 경우, AI 프로그램이 저작물을 사용하여 학습되었다는 증거를 통해 접근이 입증될 수 있다. 예를 들어, 저작물은 AI 프로그램을 훈련하기 위해 다운로드되거나 스캔된 공개적으로 접근할 수 있는 인터넷 사이트의 부분일 수 있다.

둘째, 원고는 새로운 저작물이 기존 저작물과 “실질적으로 유사”하다는 사실을 입증하여 침해를 입증해야 한다. 실질적 유사성 테스트는 정의하기 어렵고 미국 법원마다 다르다. 예를 들어, 대법원은 두 저작물이 “전체적인 개념과 느낌이 실질적으로 유사”하거나 “전체적인 모양과 느낌”이 같거나 “보통의 합리적인 사람이 두 저작물을 구별하지 못할 것”을 요구하는 등 다양한 방식으로 이 요건을 설명해 왔다. 주요 판례에서도 이러한 판단에 “원고의 저작물 전체와 관련하여 복제된 부분의 질적, 양적 중요성”을 모두 고려한다고 명시하고 있다. AI가 생성한 결과물의 경우, 전통적인 저작물 못지않게 ‘실질적 유사성’ 분석에서 법원은 AI 결과물과 기존 저작물 간에 이러한 종류의 비교를 할 것을 요구할 수 있다.

생성형 AI 프로그램이 출력물에서 기존 저작물을 복제할 가능성에 대해서는 상당한 의견 차이가 있다. OpenAI는 “잘 구축된 AI 시스템은 일반적으로 학습 말뭉치(corpus)에 있는 특정 저작물의 변경되지 않은 데이터를 사소한 부분이라도 재생성하지 않는다”고 주장한다. 따라서 OpenAI는 저작권 침해가 “우발적인 결과일 가능성은 거의 없다”고 말한다. 이와 대조적으로, Getty Images 소송에서는 “Stable Diffusion이 때때로 Getty Images와 매우 유사하거나 파생된 이미지를 생성한다”고 주장한다. 한 연구에 따르면,

Stable Diffusion으로 생성된 이미지의 2% 미만에서 “상당한 양의 복제”가 발견되었지만, 저자들은 이 방법이 “실제 복제율을 과소평가한 것 같다”고 주장하였다.

두 가지 종류의 AI 결과물이 특별한 문제를 야기할 수 있다. 첫째, 일부 AI 프로그램은 기존의 가상 캐릭터와 관련된 작품을 만드는 데 사용될 수 있다. 이러한 작품들은 그 자체로 저작권 보호를 받는 경우가 있기 때문에 이러한 저작물은 저작권 침해의 위험이 높을 수 있다. 둘째, 일부 AI 프로그램들은 특정 예술가나 작가의 ‘스타일’로 예술 작품이나 문학 작품을 만들라는 메시지를 표시할 수 있지만, 위에서 언급한 바와 같이 일부 AI 프로그램은 이러한 메시지를 ‘거부’하도록 설계될 수 있다. 저작권법은 일반적으로 예술가의 전체적인 스타일이 아닌 특정 작품의 복제를 금지하고 있으므로 이러한 결과물이 반드시 저작권을 침해하는 것은 아니다. 예를 들어, 한 평론가는 AI가 생성한 노래 “Heart on My Sleeve”에 대해 Drake의 목소리를 모방한 것은 저작권법을 위반하지 않는 것으로 보이지만, 주 퍼블리시티 법에 따라 문제를 제기할 수 있다고 지적한다. 그럼에도 일부 아티스트들은 AI 프로그램이 자신의 스타일을 모방한 작품을 대량으로 생산할 수 있어 잠재적으로 작품의 가치를 떨어뜨릴 수 있다고 우려하고 있다. 예를 들어, Stable Diffusion을 상대로 한 소송의 원고는 인간 아티스트 중 다른 아티스트의 스타일을 성공적으로 모방할 수 있는 사람은 거의 없지만, “AI 이미지 제품은 쉽게 모방할 수 있다”고 주장한다.

마지막 질문은 생성형 AI 결과물이 기존 저작물의 저작권을 침해하는 경우 누가 책임을 져야 하는가이다. 현행 법리에 따르면 AI 사용자와 AI 회사 모두 잠재적으로 책임을 질 수 있다. 예를 들어, 사용자가 저작권 침해에 대한 직접적인 책임이 있더라도 “침해 활동을 감독할 권리와 능력”이 있고 “그러한 활동에 직접적인 경제적 이해관계”가 있는 피고에게 적용되는 “대위 침해” 원칙에 따라 AI 회사가 책임을 질 수 있다. 예를 들어, Stable Diffusion에 대한 소송은 피고인 AI 회사가 저작권 침해에 대한 대위 책임이 있다고 주장하였다. AI 프로그램의 한 가지 문제점은 사용자가 사용자의 지시에 따라 복사된 저작물을 인지하지 못하거나 접근할 수 없다는 것이다. 현행법상 이러한 경우 사용자가 저작권 침해에 대한 책임이 있는지 여부를 분석하기가 어려울 수 있다.

라. 의회의 고려 사항

의회는 생성형 AI 프로그램에 의해 제기된 저작권법상의 문제 중 저작권법 또는 기타 법률의 개정이 필요한지 여부를 고려할 수 있다. 예를 들어, 의회는 AI가 생성한 저작물에 저작권이 있는지, 누가 그러한 저작물의 저작자로 간주되어야 하는지, 또는 생성형 AI 프로그램을 학습시키는 과정이 공정 이용에 해당하는지를 명확히 하는 입법을 고려할 수 있다. 법원과 저작권청이 이러한 문제를 다룰 기회가 거의 없었기 때문에 의회는 관망하는 접근 방식을 채택할 수 있다. 법원이 생성형 AI와 관련된 사건을 처리한 경험이 쌓이면, 사법부의 의견을 통해 이 분야에서 더 큰 지침과 예측 가능성을 제공할 수 있을 것이다. 이러한 사건의 결과에 따라 의회는 입법 조치가 필요한지 여부를 재평가할 수 있다.

4. 국내에서의 논의

가. 저작권 침해

생성형 AI 모델을 구축하기 위해서는 데이터를 수집·가공하여 데이터 세트를 구성하고, 이를 신경망에 전달하여 학습시키는 일련의 절차를 거치게 된다. 이 학습 과정의 데이터 세트는 저작물을 포함하고 있어, 저작권자의 허락 없이 이용하는 경우 저작권 제한 사유에 해당하지 않는 한, 침해의 문제가 발생한다. 저작권 침해가 성립하기 위해서는 “침해자의 저작물이 저작권자의 저작물에 의거하여 그것을 이용하였어야 하고, 침해자의 저작물과 저작권자의 저작물 사이에 실질적 유사성이 인정되어야 한다.”⁶¹⁾ 먼저, 의거성과 관련하여, 저작물을 이용했음을 인정할 수 있는 직접 증거는 학습 데이터가 외부에 공개되지 않는 한, 이를 통한 침해 여부 판단은 쉽지 않다. 다만, 접근, 현저한 유사성, 공통의 오류 등의 간접 증거를 통해 인정할 수는 있어 보인다. 예로써, 생성형 AI가 기존 저작물을 암기하는 방식으로 학습하는 경우라면 저작물을 그대로 출력하기 때문에, 현저한 유사성에 의한 의거성을 인정할 수 있다.

그렇다면 대부분의 저작권 침해 판단은 실질적 유사성에 따라 추론할 수밖에 없다. 실질적 유사성이 인정되기 위해서는 저작물에 대한 부당한 이용(improper appropriation)이 있어야 한다. 부당한 이용이 성립하기 위하여는 저작권자의 저작물 중 창작적 표현과의 실질적 유사성이 존재하여야 한다. 실질적 유사성의 판단은 2단계로 이루어지며, 그 하나는 저작물을 이용한 부분의 양적·질적 비중이 어느 정도인가를 판정하는 단계이고, 다음은 이용된 부분과 실제 침해자의 작품이 어느 정도 유사한가를 판정하는 단계이다. 양적 판단 기준은 어느 정도의 표현이 이용되었는지를 판단하는 데 구체적 기준을 제시하는 것은 불가능하지만, 질적 판단 기준의 보조 개념으로서 의미가 있다. 질적 판단 기준은 이용된 부분이 질적으로 중요한 부분일수록, 핵심적인 부분일수록 침해 가능성이 커진다. 생성형 AI가 기존 텍스트나 이미지를 통째로 복제하거나 위의 기준에서 유사한 것을 만드는 경우 실질적 유사성을 인정할 수 있다.

저작권 침해에 해당하는 경우 침해의 주체가 누구인지와 관련된 저작권 문제도 발생할 수 있다. 민법에 따르면 “사람은 생존한 동안 권리와 의무의 주체가 된다.”⁶²⁾ 따라서 사람만이 권리를 가질 수 있고, 책임도 질 수 있다. 생성형 AI의 출력물이 타인의 저작권을 침해하는 경우, 그 침해에 대하여 AI가 책임을 지기는 어려워 보인다.⁶³⁾ 그래도 저작권 침해가 발생한 것에 대하여 책임져야 하므로, 책임의 주체를 찾아보면, AI 개발자나 AI 이용자가 책임자가 될 수 있다. 먼저, AI 개발자는 저작권 침해에 대해 가장 유력하게 책임을 질 수 있는 자이다. 저작물을 학습하는 과정에서 저작권 침해가 발생하고 그 학습을 하도록 만든 것이 AI 개발자이기 때문에, 침해와 가장 관련이 깊다고 볼 수 있다. 다만, AI 개발자는 학습을 하도록 만들었을 뿐, 저

61) 대법원 2021. 6. 30. 선고 2019다208061 판결.

62) 민법 제3조.

63) 김윤명, 생성형 AI와 저작권 현안, KISDI AI Outlook 2023 Vol. 13, 21면.

작물의 복제나 파생 저작물의 생성은 AI에 의해 이루어진 것이어서 AI가 점차 그 수준이 높아지는 경우에는 관련성이 좀 더 적어질 수 있을 것이다. 다음으로, AI 이용자는 기존 저작물과 같거나 유사한 산출물을 만들기 때문에 저작권 침해 가능성이 존재한다. 생성형 AI는 일종의 도구여서 결과물이 위법적이라고 하더라도, 이 생성에 직접적인 관여를 한 것으로 보기 어렵다는 견해가 있다.⁶⁴⁾ 이용자가 저작물을 이용하여 저작권 침해가 성립했다고 보는 것은 어렵겠지만, 그 이용을 통해 만들어진 저작물을 다시 이용하는 경우에는 침해가 성립할 수 있고, 그에 대한 책임을 져야 할 것이다.

나. 공정 이용

위의 생성형 AI에 대한 저작권 책임은 또한 공정이용 여부에 따라 결정될 수 있다. 우리 저작권법은 AI 학습에 대한 개별 면책 사유를 규정하고 있지 않기 때문에, 저작권법 제35조의5에 해당하는지를 살펴봐야 한다.⁶⁵⁾ 첫 번째, 이용의 목적과 성격에 대해 살펴본다. 학습 과정에서 AI는 저작물을 분석하여 특징들을 수치화하여 저장한다. 텍스트의 경우에는 말뭉치를 인덱싱하여 데이터값을 부여하고, 이미지의 경우에는 이미지의 특징을 분석하여 수치화한다. 이러한 저작물의 이용은 저작물을 향유하기 위해 이용하기보다는 데이터를 분석하는 것이어서, 일반적인 저작물 이용 목적과는 다르다. 또한, 그 이용이 변형적 이용인지를 살펴볼 수 있는데, 변형적 이용은 “원저작물과는 다른 목적의 이용이고 원작의 성질에 대한 새로운 표현을 부가해 변화를 준 것”이라 할 수 있다.⁶⁶⁾ 그렇다면, AI 학습은 저작물을 일반적으로 향유하는 것이 아닌, 새로운 가치를 부가해 원저작물에 변화를 준 것, 달리 말하면 창작적 이용에 해당하는 것으로 볼 수 있다. 하지만 문제는 이러한 학습에서의 저작물 이용과 생성형 AI에 의해 결과물이 만들어지는 과정에서의 저작물 이용은 별개로 볼 수 있다는 데 있다. 만일, 기존의 이미지와 유사한 이미지를 만들기 위해 생성형 AI가 사용되었다면, 그 목적이 과연 변형적이라고 할 수 있을까는 의문이다.

두 번째, 저작물의 종류 및 용도와 관련해서는 이용 대상 저작물의 창작성이 높고 낮음에 관계없이 모두 이용될 수 있고, 미공표 저작물보다는 공표 저작물에 대한 이용이 보다 많을 것으로 예상되지만, 이를 통해 공정 이용의 부정과 긍정을 판단하기는 어려워 보이며, 결국 구체적 상황에서 저작물의 종류와 용도에 따라 결정될 수 있을 것이다.

세 번째, 이용된 부분의 비중과 중요성은 이용된 부분의 양과 질을 상대적으로 판단하여 결정할 수 있다. 양적으로 적다고 하더라도 양적으로 적은 부분이 저작물의 중요한 부분에 사용되는 경우에는 공정 이용이 성립되기 어렵다. 따라서 본 요소 역시 개별 사안에서 구체적으로 판단되어야지 일괄적으로 판단할 수는 없어 보인다.

64) 위의 글, 22면.

65) 판단 요소는 ① 이용의 목적 및 성격, ② 저작물의 종류 및 용도, ③ 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성, ④ 저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향 4가지이다.

66) 한국정보법학회 지음, 인터넷 그 길을 묻다, 중앙Books, 2012, 544면.

네 번째는 저작물의 시장에 미치는 영향인데, 그 이용으로 만들어진 작품 등이 저작물 시장에서 경쟁하여 저작물의 시장에 영향을 미치지 않아야 공정 이용이 성립될 수 있다. 여기서 시장의 범위는 2차적저작물의 시장이나 가치도 포함된다.⁶⁷⁾ AI 학습은 “정보 내용이나 표현의 특성을 학습하기 때문에 일반적인 이용 형태와 다른 것이라서 일반 소비자에게 제공되는 것과는 다른 시장을 형성한다”고 하는 의견도 있다.⁶⁸⁾ 하지만 이는 학습 과정에서의 저작물 이용을 말하는 것이고, 제35조의5의 네 번째 요소는 “저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향”이기 때문에, 결과물이 갖는 시장에 대한 영향력을 판단해야 해서, AI 결과물은 원저작물을 변형하여 만들어진 것이라면, 그리고 그 변형이 충분히 이뤄지지 않은 것이라면, 원저작물 시장과 충분히 경쟁 관계가 되고, 그 시장에 영향을 미칠 수 있어 보인다. 결국, “오로지 언어 모델을 학습시키기 위한 목적으로 웹 스크레이핑을 수행하는 것은 저작권법상 공정이용에 해당하여 허용될 가능성이 작지 않다”는 주장도 있지만,⁶⁹⁾ 이는 저작물 이용 과정 그 자체가 아니라 결과물에 의한 판단이 같이 이뤄진다면 공정 이용을 보장할 수는 없어 보인다.

다. 방조 책임

생성형 AI는 이용자에게 저작물을 이용할 수 있는 도구를 제공하기 때문에, 온라인 서비스 제공자와 같은 입장에 설 수 있다. 따라서 생성형 AI의 간접 책임 및 그 면책에 대한 논의가 필요하다. 우리 저작권법은 이용자의 저작권 직접 침해에 대하여 온라인 서비스 제공자의 간접 책임을 방조 책임의 일종으로 보며, 온라인 서비스 제공자는 이용자의 직접 침해에 대하여 부작위에 의한 방조자 공동 불법 행위 책임이 성립할 수 있다. 다만, 온라인 서비스 제공자의 책임이 지나치게 확대되는 것을 막기 위해, 온라인 서비스 제공자가 직접 침해를 구체적으로 인식하지 못하였거나 기술적·경제적으로 관리 및 통제할 수 없는 경우에는 책임을 부정한다.⁷⁰⁾ 그런데, 최근 이러한 방조 책임에 대하여, 링크 행위자가 ① 정범이 공중송신권을 침해한다는 사실을 충분히 인식하면서 ② 그러한 침해 게시물 등에 연결되는 링크를 인터넷 사이트에 영리적·계속적으로 게시하는 등으로 공중의 구성원이 개별적으로 선택한 시간과 장소에서 침해 게시물에 쉽게 접근할 수 있도록 하는 정도의 링크 행위를 한 경우에는 공중송신권 침해의 방조범이 성립한다고 판단하여, 그 범위를 다소 확장하는 판결도 내려진 바 있다. 생성형 AI는 학습 과정이나 생성 과정에서 침해를 인식할 수 있고, 이를 관리하거나 통제할 수 있고, 영리적이고 계속적으로 침해를 돕고 있다. 한편, 저작권법은 온라인 서비스 제공자의 책임을 제한하는 규정을 두고 있으며, 여기서 4가지 유형의 사업자, 즉 도관 서비스 사업자, 캐싱 서비스 사업자, 호스팅 서비스 사업자, 정보 검색 서비스 사업자들에 대해서는 저작권 침해에 대한 책임을

67) 최호진, 썬네일 이미지와 공정이용, Law and Technology, 제8권 제3호, 서울대학교 기술과 법센터, 2012, 70면.

68) 김윤명, 앞의 글(각주 63), 30면.

69) 김병필, 대규모 언어모델 인공지능의 법적 쟁점, 정보법학 Vol. 26, no. 1, 한국정보법학회, 2022, 189면.

70) 대법원 2019. 2. 28. 선고 2016다271608 판결.

면제한다.⁷¹⁾ 생성형 AI는 이용자의 요구에 따라 답변을 제공하는 방식으로 이루어지므로, 정보 검색 서비스 사업자와 유사하다고 볼 수 있다. 하지만 검색 도구를 통하여 이용자가 요구한 검색에 대한 정보만을 제공하는 정보 검색 서비스 사업자와 인간의 두뇌와 유사한 AI를 이용하여 새로운 콘텐츠를 생성하는 생성형 AI를 같다고 하기는 어려워 보이며, 그에 따라 책임의 면책에 대한 적용도 쉽지는 않아 보인다.

71) 저작권법 제102조: 도관 서비스 사업자는 인터넷 접속 서비스 사업자(ISP)를 말하고, 캐싱 서비스 사업자는 온라인 서비스 제공자(OSP)가 네트워크상의 트래픽(traffic, 전송 정체)을 줄이기 위한 방법으로 일정한 콘텐츠를 중앙 서버와는 별도로 구축된 캐시 서버에 자동적으로 임시 저장하여 이용자가 캐시 서버를 통해 해당 콘텐츠를 이용할 수 있도록 하는 것을 말하고, 호스팅 서비스 사업자는 카페나 블로그, 웹하드 등에서 서버의 일정 공간을 제공(hosting)하고, 저작물 등을 해당 공간에 저장하여 사용할 수 있게 하는 것을 말하고, 정보 검색 서비스 사업자는 정보 검색 도구를 통하여 이용자에게 정보통신망상 저작물 등의 위치를 알 수 있게 하거나 연결하는 서비스를 제공하는 것을 말한다. 한국저작권보호원, 2020 저작권 보호 자체점검 지침서, 98~101면.

III

생성형 AI의 저작자성

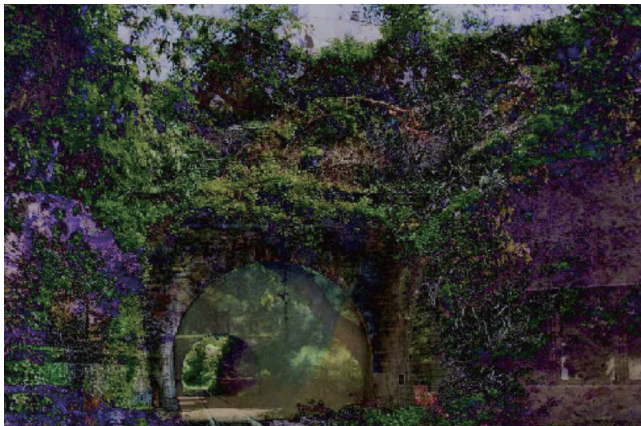
1. 관련 사례의 검토

가. Thaler v. Creativity Machine⁷²⁾

1) 배경

Thaler는 저작물 등록을 신청하면서, 해당 저작물이 “기계에서 실행되는 컴퓨터 알고리즘에 의해 자율적으로 생성된 것”이라고 하였다. 이후 그는 생각을 바꾸어 그 작품이 Creativity Machine의 업무상저작물(work-for-hire)이라고 하면서 저작권을 주장하였다.

[그림 3] A Recent Entrance to Paradise⁷³⁾



72) No. 22-1564 (D.D.C. Aug. 18, 2023).

73) <<https://www.mhc.ie/latest/insights/us-district-court-rejects-copyright-application-for-creativity-machine-ai>, 2023. 12. 30. 최종방문>.

Thaler는 다른 모든 기준이 충족되고, 저작권이 AI 소유주에게 있는 경우 AI에 대한 권리를 인정해야 한다고 주장하며 여러 차례 재심을 요청했지만, 미국 저작권청은 이를 기각하였다. Thaler는 저작권청의 결정에 불복하여 지방 법원에 항소하였다. Beryl A. Howell 판사는 이 사건의 주요 쟁점이 인간의 개입이 AI 시스템이 저작물을 만드는 경우 저작권으로 보호받을 수 있는지였다고 확인하였다. Howell 판사는 Thaler가 제출한 자료의 상당 부분이 AI 생성 소프트웨어에 지시하는 등 인간이 개입한 정도를 다루는 데 할애했다고 설명하였다. 판사는 또한, 사법적 검토는 원래의 사실에 근거해야 하고, Thaler의 저작권 등록 신청은 문제가 된 저작물이 자율적으로 생성되었다는 것을 근거로 제출되었다는 사실도 언급하였다.

2) AI 시스템이 자율적으로 생성한 저작물도 저작권을 주장할 수 있는가?

Howell 판사는 저작권은 시대에 따라 변화하지만, 창의성에 대한 인간의 입력은 저작권의 필수 조건이라는 점을 확인하였다. 판사는 미국 대법원 판결⁷⁴⁾을 언급하며, 카메라는 사진작가의 ‘정신적 구상’을 거쳐 다양한 판단을 통해 장면을 재현한 결과물일 뿐이라고 설명하였다. 이 판결에서 대법원은 Oscar Wilde의 사진에 대해 해당 사진이 “사진작가의 지적 발명의 산물”이며, 최종 결과물은 카메라가 아닌 인간에게 의존한다고 판결하였다. 미국 저작권법 1976은 다음과 같이 규정하고 있다: “저작권 보호는 이 제목에 따라 현재 알려져 있거나 나중에 개발되어 직접 또는 기계나 장치의 도움을 받아 인식, 복제 또는 기타 방식으로 전달될 수 있는 유형의 표현 매체에 고정된 저작물의 원본에 존재한다.”⁷⁵⁾ 따라서 저작권 보호를 받으려면 저작물이 저작자에 의해 고정되어야 한다. 1976년 법에는 “저작자”가 정의되어 있지 않기 때문에 판사는 “저작자”의 사전적 의미인 “어떤 형태의 지적 또는 창작적 저작물의 원천이 되는 사람”이라는 Merriam Webster의 축약되지 않은 사전적 의미에 의존하였다. 판사는 저작자가 되려면 지적 또는 창작적 작업을 할 수 있는 능력이 있어야 한다고 판단하였다.

Sarnoy 판결에서와 마찬가지로 Howell 판사는 미국 대법원이 인간 저작자 요건을 일관되게 확인해 왔다는 점을 언급하였다. Urantia Found v. Kristen Maaherra 사건⁷⁶⁾과 같이, 천체 존재와 관련된 다양한 사건에서 저작권이 인정되었다. 이 사건에서 Urantia는 천상의 존재로부터 계시를 선택하고 배열하는 데 인간이 관여했기 때문에 저작권이 있다고 간주되었다. 판사가 언급한 또 다른 사례는 Kelly v. Chicago Park District 사건⁷⁷⁾으로, 미국 법원은 인간이 관여하지 않은 정원에 대한 저작권을 인정하지 않았다. 법원은 또한 유명한 ‘원숭이 셀카 사건’인 Naruto v. Slater 사건⁷⁸⁾도 언급하였다. 이 사건에서 PETA는 원숭이 Naruto를 대신하여 사진작가를 고소하였다. 미국 저작권법이 누구를 보호하기 위해 고안되었는지에 대한

74) Burrow-Giles Lithographic Co. v. Sarony 111 U.S. 53, 58 (1884).

75) 17 U.S.C. § 101.

76) 114 F.3d 955, 958-59 (9th Cir. 1997).

77) 635 F.3d 290 (7th Cir. 2011).

78) 888 F.3d 418 (9th Cir. 2018).

질문을 고려할 때, 제9순회법원은 인간만이 소송을 제기할 수 있다고 말하였다. 미망인, 어린이 등 권리를 가진 사람을 설명하는 데 사용되는 용어가 이 법이 인간을 보호하기 위해 고안되었음을 보여주기 때문이다. 이러한 결정의 결과로 원숭이는 미국 저작권법에 따라 소송을 제기할 자격이 없다. Thaler는 그 반대의 사례가 있는지는 확인하지 못하였다.

3) 업무상저작물

Howell 판사는 판례법에 따르면 재산권이 먼저 존재해야 하기 때문에, 해당 저작물이 업무상저작물이라는 Thaler 씨의 주장은 실패했다고 지적하였다. 더 나아가, 1976년 법에 따른 두 가지 정의 모두 사람의 개입을 필요로 한다.

나. Zarya of the Dawn

1) 배경

Kashtanova에 의해 제출된 등록 신청서와 첨부 서류에 설명된 바와 같이, 해당 작품은 18페이지로 구성된 만화 잡지(comic book)였다. 그중 하나는 표지였으며, 여기에는 젊은 여성의 이미지, 저작물의 제목 그리고 Kashtanova와 Midjourney라는 단어들로 구성되어 있었고, 나머지 페이지들은 텍스트와 시각적 자료가 섞여 있었다. 다음은 표지 페이지와 두 번째 페이지다.

[그림 4] Zarya of the Dawn의 커버와 두 번째 페이지



Kashtanova는 2022년 9월 15일에 저작물에 대한 등록 신청서와 저작물의 각 페이지 사본을 첨부 서류로 제출하였다. 신청서에서 Kashtanova는 저작물의 저작자를 “Kristina Kashtanova”로 기재하고 자신

이 “만화 잡지를 만들었다고 진술하였다.” 신청서에는 저작물의 일부를 창작하기 위해 AI를 사용했다는 사실을 밝히지 않았고, 저작물의 일부를 부인하지도 않았다. 저작권청은 같은 날 신청서를 검토하고 저작물을 등록 번호 VAu001480196으로 등록하였다. 저작물 등록 직후, 저작권청은 Kashtanova가 소셜 미디어에 Midjourney AI를 사용하여 만화 잡지를 만들었다는 글을 올린 것을 알게 되었다. 신청서에 AI의 사용 여부가 명시되어 있지 않았기 때문에, 저작권청은 신청서가 잘못되었거나 최소한 실질적으로 불완전하다고 판단하였다. 미국 저작권청은 2022년 10월 28일자 서신을 통해 Kashtanova에게 등록이 취소되어서는 안 되는 이유를 설명하는 추가 정보를 서면으로 제공하지 않는 한 등록을 취소할 예정이라고 통보하였다.⁷⁹⁾

2022년 11월 21일, 저작권청은 Kashtanova의 변호사 Van Lindberg로부터 답변을 받았으며, 여기에는 저작물 창작에 대한 설명과 함께 Midjourney의 사용에 대한 구체적인 정보가 포함되어 있었다. Lindberg는 ① Kashtanova가 저작물의 모든 부분을 저작하였으며 Midjourney는 단지 보조 도구의 역할을 했을 뿐이고, ② 텍스트는 Kashtanova가 저작했으며, 텍스트와 이미지의 창의적인 선택, 조정 그리고 배열로 인해 그 작품은 저작권으로 보호받을 수 있는 모음집(compilation)이 되었기 때문에, 그 작품은 저작권으로 등록될 수 있다고 주장하였다.

2) 법적 기준

저작물에 대한 분석을 시작하기 전에 해당 분석의 기준이 되는 법적 원칙을 간략히 설명한다. 저작권법은 저작권 보호의 범위를 정의한다. 저작권법에 따라 저작물이 “유형적 표현 매체에 고정된 저작물의 원본”에 해당하는 경우 등록할 수 있다.⁸⁰⁾ 대법원은 이러한 맥락에서 “원본”이라는 용어가 독립적인 창작성과 충분한 창작성이라는 두 가지 요소로 구성되어 있다고 설명하였다.⁸¹⁾ 첫째, 저작물은 저자가 독립적으로 창작한 것이어야 한다. 둘째, 저작물에 충분한 창작성이 있어야 한다. 약간의 창의성만 있으면 되지만, 대법원은 Feist 사건에서 문제가 된 알파벳순 전화번호부와 같은 일부 저작물은 이 낮은 기준조차 충족하지 못한다고 판결한 바 있다. 대법원은 또한, “헌법적 문제로서 저작권은 최소한의 창작성 이상을 보유한 저작물의 구성요소만을 보호한다”고 판시하였다. “창조적 불꽃이 완전히 결여되어 있거나 사실상 존재하지 않을 정도로 사소한” 저작물에는 저작권이 존재할 수 없다고 판결하였다.

법원은 “저작물”이라는 문구를 해석할 때 일관되게 인간 저작자의 창작물로 제한해 왔다. 예를 들어, Sarony 사건에서 대법원은 사진이 “저자의 독창적인 지적 개념을 표현한 것”이기 때문에 저작권의 보호를

79) 37 C.F.R. § 201.7(c)(4)(If the Office becomes aware that an issued registration does not satisfy the statutory requirements for copyright “or that information essential to registration has been omitted entirely from the application or is questionable,” the Office will correspond with the copyright claimant “in an attempt to secure the required information . . . or to clarify the information previously given on the application.” If the claimant does not reply in 30 days, the Office will cancel the registration.).

80) 17 U.S.C. § 102(a).

81) Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991).

받는다고 판시하면서, 저자를 “어떤 것의 출처가 되는 사람, 창시자, 제작자, 과학 또는 문학 작품을 완성한 사람”으로 정의하였다.⁸²⁾ 이를 통해 법원은 사진이 기계에 의해 만들어진 “어떤 자연물 또는 어떤 사람의 정확한 특징을 종이에 재현한 것”에 불과하다는 주장을 거부하였다. 그러나 법원은 사진이 인간 사진작가의 “참신성, 발명성 또는 독창성이 개입될 여지가 없는” “단지 기계적인” 과정이라면 “그러한 경우 저작권은 보호되지 않는다”고 설명하였다.⁸³⁾ 비인간 저작물이 주장되는 경우, 항소 법원들은 저작권이 주장된 창작물을 보호하지 않는다고 판결하였다. 예를 들어, 제9순회법원은 “인간이 아닌 영적 존재가 ‘저술’한” 단어가 포함된 책은 “인간이 기사를 선택하고 배열한” 경우에만 저작권 보호를 받을 수 있다고 판시하였다.⁸⁴⁾ 법원은 또한, “저작권법이 보호하려는 것은 신성한 존재의 창조물이 아니기 때문에,” 이 책이 저작권을 갖기 위해서는 “인간 창작성의 일부 요소가 발생하여야 한다”고 판결하였다. 저작권청의 등록 관행은 이러한 법원 판결을 따르고 반영한다. 저작권청은 저작물 심사 및 등록에 대한 기준을 제공하는 미국 저작권청 실무 개요서(제3판)에서 법에 대한 이해를 정리하고 있다. 위에서 설명한 사례들에 따라, 이 개요서에서는 “사람이 저작물을 창작하지 않았다고 판단되는 경우 저작권청은 등록 청구를 거부한다”고 설명하고 있다.⁸⁵⁾

3) 판단 1 – 텍스트

저작권청은 우선, 텍스트에 대한 보호 여부를 살펴봤으며, 텍스트가 저작권에 의해 보호된다는 데는 동의하였다. Kashtanova의 서신에는 “저작물의 텍스트가 생성형 AI 프로그램을 포함한 다른 소스나 도구의 도움 없이 전적으로 Kashtanova가 작성했다”고 명시되어 있다. 이 진술에 근거하여, 저작권청은 해당 텍스트가 인간 저작물의 산물임을 확인하였다. 또한, 저작권청은 그 텍스트가 Feist 판례에 따라 요구되는 “최소한의 창작성” 이상을 포함하고 있다고 판단하였다. 따라서 텍스트에 대하여 등록을 할 수 있다고 결정하였다.

3) 판단 2 – 이미지와 텍스트의 선택과 배열

다음으로, 저작권청은 이미지와 텍스트의 선택 및 배열이 편집물로서 보호받을 수 있다는 데도 동의하였다. 저작권은 충분히 창의적인 방식으로 “선택, 조정 또는 배열된 기존 자료나 데이터의 수집 및 조합”은 보호한다고 규정하고 있다.⁸⁶⁾ Kashtanova는 저작물의 이미지를 “선택, 다듬기, 자르기, 배치, 프레임, 배열”하여 페이지에 담긴 이야기를 만들었다고 진술하였고, 이야기를 반영하기 위해 이미지를 선택, 조정 및 배열한 것은 저작물 전체의 저작권성을 뒷받침한다고 주장하였다. 저작권청은 저작물에 포함된 이미지의

82) Saroni, 111 U.S. at 53, 57–59.

83) Id. at 59.

84) Urantia Found. v. Kristen Maaherra, 114 F.3d 955, 957–59 (9th Cir. 1997).

85) U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 313.2 (3d ed. 2021) (“원숭이가 찍은 사진” 및 “성령을 저작자로 명명한 노래 신청서” 등 인간의 저작물이 아닌 저작물의 예를 제시함).

86) 17 U.S.C. § 101; Compendium(Third) § 312.1.

선택과 배열이 전적으로 Kashtanova에 의해 이루어졌다는 진술에 근거하여 저작물이 인간 저작물의 산물이라는 결론을 내린다. 또한, 저작권 사무소는 저작물 전체에 걸쳐 이러한 이미지와 텍스트의 편집이 Feist 결정에 따라 저작권으로 보호받을 수 있는 충분한 창작성을 포함하고 있다고 판단하였다. 특히, 저작권청은 저작물을 구성하는 이미지의 선택과 저작물의 각 페이지에서 이미지와 텍스트의 배치 및 배열과 관련하여 저작물이 창의적인 선택의 산물이라고 판단하였다. 따라서 저작권은 저작물을 구성하는 텍스트 및 시각적 요소의 전반적인 선택, 조정 및 배열에 대한 저작물성을 인정하였다.

4) 판단 3 – 개별 이미지

마지막으로, 저작물의 개별 이미지에 대한 저작권성 분석에서, 저작권청은 Kashtanova가 Midjourney의 AI 기술을 사용한 것이 저작물에 미친 영향을 고려해야 한다고 하였다. Kashtanova 서신의 대부분은 그녀가 이러한 이미지를 제작하기 위해 Midjourney를 어떻게 사용했는지에 초점을 맞추고 있다. 저작권청은 개별 이미지의 저작권 여부에 대한 문제를 다루기 위해 Midjourney에 대한 이해와 작동 방식에 대해 설명하였고, 그 이해가 서한에 기술된 AI 서비스에 대한 설명, 청 자체의 지식, 그리고 청이 행정적으로 통지한 Midjourney의 공개 문서에 근거한다고 밝혔다.

① Midjourney의 작동 방식

Midjourney는 사용자가 제공한 텍스트에 반응하여 이미지를 생성할 수 있는 AI 기술을 제공한다. Midjourney는 사용자가 운영하는 개별 서버로 구성된 Discord라는 제휴되지 않은 제3의 커뮤니케이션 서비스를 기반으로 운영된다. Midjourney를 이용하기 위해, 사용자들은 사용자들이 텍스트를 입력할 수 있는 공개 ‘채널’이 포함된 Midjourney Discord 서버에 가입해야 한다. Midjourney는 주로 이러한 채널의 자동 계정을 통해 사용자가 입력한 텍스트를 읽고 이를 기반으로 이미지를 생성하는 방식으로 운영된다. 사용자들은 Midjourney의 채널 중 하나에 입력하는 텍스트 명령어인 ‘프롬프트’를 통해 Midjourney를 작동한다. Midjourney의 설명에 따르면 프롬프트는 “/imagine”이라는 텍스트로 시작해야 하며 Midjourney가 생성해야 하는 내용을 설명하는 텍스트를 포함해야 한다. 또한 사용자는 생성된 출력에 영향을 주기 위해 하나 이상의 이미지의 URL을 포함하거나, 특정 화면 비율로 이미지를 생성하거나 다른 기능적 방향을 제공하도록 Midjourney에 지시하는 매개 변수를 포함할 수 있는 옵션이 있다.

사용자가 Midjourney에 프롬프트를 입력하면, 이에 대한 응답으로 4개의 이미지를 생성한다. 이미지는 격자 형태로 제공되며, 사용자는 격자 형태 아래에 위치한 버튼을 사용하여 Midjourney에 더 높은 해상도의 이미지 버전(예: U1, U2, U3, U4)을 제공하거나, 이미지의 새로운 변형 버전(예: V1, V2, V3, V4)을 생성하거나, 4개의 이미지를 처음부터 새로 생성하도록 요청할 수 있다. 여기서 중요한 점은 Midjourney가 프롬프트를 특정 표현 결과를 만들기 위한 구체적인 지침으로 해석하지 않는다는 점이다. Midjourney는 “사람처럼 문법, 문장 구조 또는 단어를 이해하지 못하기 때문에” 대신 “단어와 구를 토큰이라고 하

는 작은 조각으로 변환하여 학습 데이터와 비교한 다음 이미지를 생성하는 데 사용한다.” 이미지 생성에는 “초기 이미지 격자를 생성하기 위한 출발점으로 사용되는 텔레비전 정적(static)과 같은 시각적 노이즈 필드”로 시작한 다음 알고리즘을 사용하여 정적을 사람이 인식할 수 있는 이미지로 정제하는 과정이 포함된다. Midjourney 사용자가 도구를 통해 최종적으로 만족스러운 이미지를 얻는 과정은 인간 예술가, 작가, 사진작가의 작업 과정과 동일하지 않다. 위에서 언급한 바와 같이, 사용자가 처음 입력하는 프롬프트는 Midjourney의 학습 데이터를 기반으로 4가지 이미지를 생성한다. 이러한 초기 이미지 중 하나에 적용된 추가 프롬프트가 후속 이미지에 영향을 미칠 수 있지만, Midjourney가 어떤 이미지를 생성할지 미리 예측할 수 없기 때문에, 사용자가 프로세스를 제어할 수 없다.

② Midjourney 이미지에 대한 저작권법의 적용

사건 기록에 근거하여, 저작권청은 저작물에 포함된 Midjourney가 생성한 이미지가 저작권으로 보호되는 저작물이 아니라고 결론 내렸다.⁸⁷⁾ Kashtanova는 자신이 각 이미지의 구조와 내용을 “가이드(guide)”했다고 주장하지만, Kashtanova 서한에 기술된 과정을 보면 이미지의 “전통적인 저작물 요소”를 만든 것은 Kashtanova가 아니라 Midjourney임을 알 수 있다. Kashtanova는 각 이미지가 “유사한 창작 과정”을 통해 만들어졌다고 주장하였다. 요약하면, 이 과정은 Midjourney를 사용한 일련의 단계로 구성되었다. 먼저, Kashtanova는 Midjourney에 텍스트 프롬프트를 입력하였고, 그것은 창작적 입력으로 이미지에 대한 핵심으로 묘사된다. 다음으로, Kashtanova는 이러한 출력 이미지 중 하나 이상을 선택하여 추가 개발을 진행하였다. 그런 다음, 그녀는 “프롬프트와 Midjourney에 제공된 다른 입력을 조정하거나 변경”하여 새로운 중간 이미지를 생성하고, 궁극적으로 최종 이미지를 생성하였다. Kashtanova는 자신이 직접 시각 자료를 만들었다고 주장하지 않았으며, 최종 이미지를 “생성, 개발, 다듬고 재배치”하고, 중간 이미지의 요소를 “응집력 있는 전체로 통합”한 것으로 설명할 때 수동태를 사용하였다. 최종 이미지를 얻기 위해 그녀는 시행착오의 과정을 설명하며, “수백 번의 반복을 통해 가능한 한 완벽하게 그녀의 비전을 표현”할 때까지 “수백 또는 수천 개의 설명적 프롬프트”를 Midjourney에 제공하였다.

Kashtanova가 원하는 이미지에 도달하기 위해 제어하고 안내하는 도구가 아니라, Midjourney는 예측할 수 없는 방식으로 이미지를 생성하였다. 따라서 Midjourney 사용자는 이 기술이 생성하는 이미지의 저작권에 대한 ‘저작자’가 아니다. 대법원이 설명했듯이, 저작권이 있는 저작물의 “저작자”는 “실제로 그림을 형성한 사람”, 즉 “발명가 또는 마스터 마인드”로 활동하는 사람이다.⁸⁸⁾ Midjourney에 텍스트 프롬프트를 제공하는 사람은 생성된 이미지를 “실제로 형성”하지 않으며, 그 배후에 있는 “마스터 마인드”가 아니다.

87) U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 313.2 (3d ed. 2021) (저작권청은 인간 저작자의 창의적인 입력이나 개입 없이 무작위로 또는 자동으로 작동하는 기계 또는 단순한 기계적 프로세스에 의해 생성된 저작물은 등록하지 않을 것이라고 설명하고 있다.).

88) Burrow-Giles, 111 U.S. at 61.

위에서 설명한 바와 같이, Midjourney는 시각적 “노이즈” 분야에서 이미지 생성 프로세스를 시작하며, 이 분야는 Midjourney의 학습 데이터베이스와 관련된 사용자 프롬프트에서 생성된 토큰을 기반으로 정제된다. 프롬프트의 정보는 생성된 이미지에 ‘영향’을 줄 수 있지만, 프롬프트 텍스트가 특정 결과를 지시하지는 않는다. 사용자가 Midjourney에 지시하는 내용과 Midjourney가 실제로 생성하는 시각적 자료 사이에는 상당한 거리가 있기 때문에, Midjourney 사용자는 생성된 이미지를 충분히 제어할 수 없어 그 배후에 있는 ‘마스터 마인드’로 취급받기 어렵다.

Kashtanova는 Midjourney의 사용 과정이 Adobe Photoshop과 같은 다른 “컴퓨터 기반 도구”를 사용하는 것과 유사하다고 주장하였지만, Midjourney의 구체적인 결과물을 사용자가 예측할 수 없다는 사실은 저작권 측면에서 Midjourney를 아티스트가 사용하는 다른 도구와 다르게 만든다. Burrow-Giles의 사진 작가처럼, 아티스트는 편집 또는 기타 보조 도구를 사용할 때 수정할 시각 자료를 선택하고, 사용할 도구와 변경 사항을 선택하고, 최종 이미지를 제어하기 위한 구체적인 단계를 수행하여 아티스트의 “눈에 보이는 형태를 자신의 독창적인 정신적 개념”에 부합하도록 한다.⁸⁹⁾ Midjourney의 사용자는 생성된 초기 이미지나 최종 이미지에 대한 통제권이 없다. 따라서 Kashtanova와 같은 사용자는 “수백 개의 중간 이미지”를 생성해야 하기 때문에, 사용자가 염두에 둔 이미지와 일치하는 이미지를 “구상부터 제작까지 하는 데 1년 이상” 걸릴 수 있다.

또한, 저작권청은 Kashtanova가 텍스트 프롬프트를 사용하더라도 그 이미지가 “창의적이고 사람이 작성한 프롬프트”를 시각적으로 표현한 것이므로 결과물에 대한 저작권 보호가 허용된다는 데 동의하지 않는다. Midjourney는 무작위로 생성된 노이즈로 시작하여 최종 이미지로 진화하기 때문에 특정 프롬프트가 특정 시각적 결과물을 생성한다는 보장은 없다. 대신, 프롬프트는 주문보다는 제안에 가깝게 작동하며, 이는 아티스트에게 이미지의 내용에 대한 일반적인 방향을 제시하고 이미지를 제작하도록 고용하는 클라이언트의 상황과 유사하다. 만약 Kashtanova가 비주얼 아티스트에게 “라야라는 이름의 홀로그램 백인 노인 여성”이 “곱슬머리를 하고 우주선 안에 있다”는 내용의 이미지를 제작하도록 의뢰하면서 “스타트렉 우주선”, “홀로그램”, “옥탄 렌더(octane render)”, “언리얼 엔진(unreal engine)”과 비슷한 분위기나 스타일을 가지며 “영화적”이고 “매우 디테일”해야 한다는 지시를 내렸다면 Kashtanova는 그 이미지의 작성자가 되지 못했을 것이다. 저작물이 업무상저작물로 인정받기 위한 법적 요건이 없다면, 저작자는 그러한 지시를 받고 이를 표현하는 최선의 방법을 결정한 시각 예술가가 될 것이다. 그리고 Kashtanova가 이미지 검색 엔진에 이러한 용어를 입력할 경우, 검색 결과에 따라 만들어지는 이미지가 자신의 예술적 비전과 아무리 유사하더라도 자신이 저작한 것이라고 주장할 수 없다.

89) Burrow-Giles, 111 U.S. at 60 (사진작가의 창의적인 선택이 사진을 “자신의 지적 발명품”으로 만들었다고 설명하였다.).

저작권청은 Kashtanova가 Midjourney와 함께 작업하는 데 상당한 시간과 노력을 들였다는 주장에 의문을 제기하지 않는다. 그러나 그러한 노력이 그녀를 저작권법상 Midjourney 이미지의 '저작자'로 만들지는 못한다. 법원은 "이마의 땀"이 보호받을 수 없는 자료에서 저작권 보호의 근거가 될 수 있다는 주장을 거부하였다.⁹⁰⁾ 저작권청은 "저작물을 만드는 데 필요한 시간, 노력 또는 비용의 양"은 "저작물이 저작권법과 헌법이 요구하는 최소한의 창작적 불꽃을 가지고 있는지 여부와 관련이 없기 때문에, 고려하지 않을 것"이다. 저작권청의 결정은 Kashtanova가 저작물의 이미지를 생성하기 위해 Midjourney를 사용했다는 구체적인 사실에 근거한 것이다. 표현적 자료를 생성할 수 있는 다른 AI 제품이 Midjourney와 다르게 작동할 수 있다. 하지만 저작권청의 행정 기록에 따르면, Kashtanova는 Midjourney가 생성한 개별 이미지의 저작자가 아니다.

③ Kashtanova에 의해 편집된 이미지

마지막으로, Kashtanova는 Midjourney가 만든 이미지 중 일부를 자신이 직접 편집했다고 주장하였다. 그녀의 서신은 저작물에 포함된 두 개의 특정 이미지를 지적하였다. 저작권청은 Kashtanova가 직접 변경했다는 진술은 받아들이지만, 편집 변경이 저작권을 가질 수 있을 만큼 충분히 창조적이라고 단정할 수는 없었다. 첫째, Kashtanova는 2번째 페이지의 이미지에서 Zarya의 입술과 입의 렌더링을 수정했다고 설명하였다. Zarya의 입, 특히 윗입술에 대한 변경은 저작권 보호에 필요한 창작성을 제공하기에는 너무 사소하고 눈에 띄지 않는다. 저작권청은 인간 저작자에 의해 편집, 수정 또는 기타 방식으로 수정되어 보호될 수 없는 자료를 포함하는 작품을 등록할 것이지만, 그 새로운 작품이 저작권 보호를 받을 수 있는 충분한 양의 독창적 저자를 포함하는 경우에만 그렇다.⁹¹⁾ Kashtanova가 이 이미지를 변경한 것은 이 기준에 미치지 못한다.⁹²⁾

둘째, Kashtanova는 작품의 12페이지에 있는 눈을 감고 있는 노파의 이미지를 지적하였다. 그녀는 이 작품이 "Midjourney 서비스와 포토샵을 함께 사용하여" 만들어졌으며, 포토샵에서 "얼굴의 노화, 그러데이션의 부드러움, 선과 모양의 수정"을 위해 편집했다고 설명하였다. Kashtanova의 설명에 근거하여, 저작권청은 이미지의 어떤 표현이 Midjourney에 의해 생성된 것이 아니라 포토샵의 사용을 통해 이루어졌는지를 결정할 수 없다. Kashtanova는 얼굴의 노화를 보여주기 위해 Midjourney가 중간 이미지를 수정하는 데 포토샵을 사용했다고 주장하지만, 이전 중간 이미지의 젊은 얼굴을 직접 편집했는지, 이전에 생성된 나이든 여성의 이미지를 사용하여 합성 이미지를 만들었는지, 아니면 다른 작업을 했는지는 불분명하다.

90) Feist, 499 U.S. at 352-53.

91) COMPENDIUM (THIRD) § 313.6(D).

92) Contra Eden Toys, Inc. v. Florelee Undergarment Co., 697 F.2d 27, 34-35 (2d Cir. 1982) (Paddington Bear의 수정된 그림은 캐릭터 모자의 비율을 변경하고, 손가락과 발가락을 개별적으로 표현하지 않았으며, 전체적으로 선을 부드럽게 처리하여 그림에 "다른, 더 깨끗한 '모양'을 부여한" 점을 근거로 2차적저작물로 인정받았다.).

Kashtanova가 Midjourney가 생성한 중간 이미지를 실질적으로 편집했다면, 이러한 편집은 사람의 저작물일 수 있으며 새로운 등록증에서 제외되지 않을 것이다.

4) 결론

위에서 설명한 이유로, 저작권청은 ‘Zarya of the Dawn’에 대한 등록 번호 VAu001480196의 등록증은 부정확하고 불완전한 정보에 근거하여 발급되었다고 결론지었다. 만약 저작권청이 Kashtanova가 제공한 정보를 미리 알았다면, AI 기술로 생성된 자료를 제외하여 그 청구를 좁힐 수 있었을 것이다. 새로운 정보에 비추어, 저작권청은 37 C.F.R. § 201.7(c)(4)에 따라 이전의 등록을 취소하고, Kashtanova가 이 작품에 기여한 독창적 저자를 포함하는 부분 즉, 텍스트와 AI에 의해 생성된 작품과 저자에 의해 만들어진 텍스트의 선택, 조정, 배열 부분에 대한 새로운 등록으로 대체할 것이다. 기여 부분들은 주로 텍스트 자료를 포함하기 때문에, 그것들은 미공표된 문학 작품으로 등록될 것이다. 새로운 등록에서는 AI가 생성한 예술 작품은 명시적으로 제외될 것이다.

공공 기록에는 이러한 결정이 반영된다. 첫째, 취소된 등록에 대한 기록에는 그 취소가 작품에 포함된 비인간 저작자를 배제하지 않았기 때문임을 표시할 것이다. 둘째, 이 작품에 대한 보다 제한적이고 새로운 등록이 클래스 TXu로⁹³⁾ 발급되었음을 반영하고, 해당 새 등록에 대한 상호 참조를 포함할 것이다. 셋째, 새 등록에는 “이전의 등록 취소된” 분야에 취소된 등록에 대한 상호 참조가 포함될 것이다. 마지막으로, 저작권청은 새 등록증에 다음과 같은 주석을 추가할 것이다: “재등록 사유: VAu001480196은 비인간 저자를 배제하지 않아 37 CFR 201.7(c)(4)에 따라 취소되었다.” 새 등록의 효력 발생일은 취소된 등록과 동일하며, 2022년 9월 15일이다. 저작권청은 원래의 등록 증명서를 취소하고 이러한 변경 사항을 반영한 새 증명서를 발급하여 별도의 표지로 Kashtanova에게 우편으로 발송할 것이다.

2. 미국의 생성형 AI와 저작권법에 관한 보고서⁹⁴⁾

생성형 AI 프로그램의 광범위한 이용은 이러한 프로그램들을 사용하여 만든 콘텐츠의 저작권을 누가 가질 수 있는지에 대한 문제를 제기한다.

가. AI 결과물은 저작권 보호를 누릴 수 있는가?

DALL-E가 생성한 이미지나 ChatGPT가 생성한 텍스트와 같은 AI 결과물에 저작권 보호가 부여될 수 있는지는 적어도 부분적으로는 “저작자”의 개념에 달려 있다. 미국 헌법은 “저작자에게 제한된 기간 동안

93) TXu 분류는 저작권 등록 분류 중 unpublished books, periodicals, and lectures에 대한 새로운 분류를 말한다.

94) 앞의 보고서(각주 57).

저작물에 대한 배타적 권리를 확보”할 수 있도록 의회에 권한을 부여한다고 규정하고 있다. 이 권한에 근거하여 저작권법은 저자의 원본 작품에 대한 저작권 보호를 제공한다. 헌법과 저작권법은 누가(또는 무엇이) “저작자”인지 명시적으로 정의하고 있지는 않지만, 미국 저작권청은 “인간이 창작한” 저작물에 대해서만 저작권을 인정한다. 대법원은 일련의 사진을 찍은 원숭이는 저작권법에 따른 소송을 제기할 자격이 없고, 천체에서 영감을 받았다고 주장하는 책에 저작권을 인정하려면 인간의 창의성이 일부 필요하며, 살아있는 정원은 저자가 없기 때문에, 저작권을 인정할 수 없다고 판결하는 등 인간이 아닌 저작자에게 저작권 보호를 확대하는 것을 거부해 왔다.

최근 한 소송에서는 AI가 ‘저작(authored)’한 것으로 알려진 저작물에 대한 인간 저작 요건에 문제를 제기하였다. 2022년 6월, Stephen Thaler는 Creativity Machine으로 불리는 AI 프로그램에 의해 ‘자율적으로’ 저작되었다고 주장하는 시각 예술 작품의 등록 신청을 거부한 저작권청을 상대로 소송을 제기하였다. Thaler 박사는 저작권법에서 인간의 저작이 요구되지 않는다고 주장하였다. 2023년 8월 18일, 연방지방법원은 저작권청의 손을 들어주는 약식 판결을 내렸다. 법원은 “인간의 저작은 유효한 저작권 청구의 필수적인 부분”이라고 판시하며, 저작물 창작에 대한 인센티브로서 저작권이 필요한 것은 인간 저작자뿐이라고 판단하였다. Thaler 박사는 이 판결에 대해 항소할 계획이라고 밝혔다.

저작물에는 인간 저작자가 필요하다고 가정할 때, 창작 과정에 인간이 개입한 성격에 따라 생성형 AI를 사용하여 인간이 만든 저작물도 저작권 보호를 받을 수 있다. 그러나 최근의 저작권 소송과 그에 따른 저작권 등록 지침에 따르면, 저작권청은 텍스트 프롬프트에 응답하여 AI 프로그램이 저작물을 생성하는 경우 인간 저작자를 찾지 못할 가능성이 높다. 2022년 9월, Kris Kashtanova는 Midjourney가 텍스트 입력에 반응하여 생성한 이미지로 삽화를 그린 그래픽 소설에 대해 저작권을 등록하였다. 2022년 10월, 저작권청은 Kashtanova가 AI 사용 사실을 공개하지 않았다는 점을 지적하며 저작권 등록 취소 절차를 개시하였다. 이에 Kashtanova는 “창의적이고 반복적인 과정을 통해” 이미지를 만들었다고 주장하며 대응하였다. 2023년 2월 21일, 저작권청은 해당 이미지의 저작권이 없다고 판단하고 “시각적 자료”의 저작자는 Kashtanova가 아닌 Midjourney라고 결정하였다. 2023년 3월, 저작권청은 “AI가 출력물의 표현적 요소를 결정할 때, 생성된 자료는 인간 저자의 생산물이 아니다”라는 지침을 발표하였다.

일부 논평가들은 AI 프로그램도 인간이 저작물을 만드는 데 사용한 다른 도구와 같다고 주장하면서, AI가 생성한 저작물은 저작권 보호를 받아야 한다고 주장한다. 예를 들어, 대법원은 1884년 Burrow-Giles Lithographic Co. v. Sarony 사건⁹⁵⁾ 이후 사진작가가 구도, 배열, 조명과 같은 창작적 요소에 관한 결정을 내릴 경우 사진도 저작권 보호를 받을 수 있다고 판시해 왔다. Kashtanova의 주장처럼, 생성형 AI 프로그램은 카메라와 유사한 새로운 도구로 볼 수 있다. 다른 논평가들과 저작권청은 사진에 대한 비유에 이의를

95) 111 U.S. 53 (1884).

제기하며, AI가 단순한 도구로 간주될 만큼 AI 사용자가 충분한 창작적 통제권을 행사하는지에 대해 의문을 제기하였다. Kashtanova의 사례에서 저작권청은 Midjourney가 “예측할 수 없는 방식으로 이미지를 생성”하기 때문에 “Kashtanova가 원하는 이미지에 도달하도록 통제하고 안내하는 도구”가 아니라고 판단하였다. 대신, 저작권청은 AI 사용자를 “아티스트를 고용한 고객”에 비유하며 해당 아티스트에게 “일반적인 지침”만을 제공한다고 설명하였다. 2023년 3월에 발표된 저작권청의 지침도 마찬가지로 “생성형 AI 시스템이 프롬프트를 해석하고 자료를 생성하는 방식에 대해 사용자가 궁극적인 창작 통제권을 행사하지 않는다”고 주장하였다. 반면, Kashtanova의 변호사 중 한 명은 특정 사진과 현대 미술에는 우연성이 어느 정도 포함되어 있다는 점을 지적하면서 저작권법이 그렇게 엄격한 창작적 통제를 요구하지 않는다고 주장하였다.

일부 논평가들은 저작권이 있는 ‘저작물’과 저작권이 없는 ‘아이디어’에 대한 저작권의 구별이 AI가 생성한 저작물을 저작권으로 보호해서는 안 되는 또 다른 이유를 제공한다고 주장하였다. 한 법학 교수는 AI 프로그램에 텍스트 프롬프트를 입력하는 인간 사용자는, 예를 들어 DALL-E에게 “해변에서 티 파티를 하는 고슴도치 그림을 만들어 달라”고 요청하는 것과 같이, 완성된 저작물에 “아이디어만 제공한 것”에 불과하다고 주장하였다. 이 주장에 따르면 결과물인 이미지에는 저작자가 없으므로 저작권을 보호할 수 없다. 저작권청의 조치들은 AI가 생성한 저작물에 대한 저작권 보호를 받는 것이 어려울 수 있음을 시사하지만, 이 문제는 여전히 해결되지 않은 채로 남아 있다. 신청자는 저작권 등록을 거부한 저작권청의 최종 결정에 이의를 제기하기 위해 미국 지방법원에 소송을 제기할 수 있으며(Thaler 박사가 그랬던 것처럼), 연방법원이 저작권청의 모든 결정에 동의하지는 아직 미지수이다. 저작권청은 법원이 이 분야에 대한 저작권청의 경험과 전문성에 무게를 두는 경우가 있지만, 법원이 저작권법에 대한 저작권청의 해석을 반드시 채택하지는 않는다고 지적한다.

또한 저작권청의 지침은 AI가 생성한 자료를 “포함하는” 저작물도 AI가 생성한 자료를 인간이 “충분히 창의적으로” 배열하거나 수정한 경우 또는 AI가 생성한 자료와 인간이 저작한 자료를 결합한 저작물 등 일부 상황에서는 저작권이 인정될 수 있음을 인정하고 있다. 저작권청은 저작자는 이러한 저작물에 대해 ‘자신의 기여’에 대해서만 저작권 보호를 주장할 수 있으며, 저작권 등록을 신청하는 경우 저작물에서 AI가 생성한 부분을 식별하고 부인해야 한다고 명시하고 있다. 예를 들어, 2023년 9월 저작권청 검토위원회는 Midjourney에 의해 생성된 작품에 대한 청의 저작권 등록 거절을 확인하였고, 그 후 신청인이 AI가 생성한 자료를 포기하지 않았기 때문에, 신청인에 의해 여러 방법으로 수정하였다.

나. 생성형 AI 결과물에 대한 저작권은 누가 소유하는가?

AI가 생성한 일부 저작물이 저작권 보호를 받을 수 있다고 가정할 때, 그 저작권은 누가 소유하는가? 일반적으로 저작권법은 “저작물의 저작자 또는 저작자들에게 우선적으로” 소유권을 부여한다. 그러나 현재까지 AI가 창작한 저작물의 저작권을 인정할 사법부나 저작권청의 결정이 없기 때문에, 이러한 저작물의 ‘저작자 또는 저작자들’이 누구인지에 대한 명확한 규정이 나오지 않았다. 사진의 비유로 돌아가서, AI의 창작

자는 카메라 제작자에 비유할 수 있고, 특정 저작물의 창작을 유도한 AI 사용자는 해당 카메라를 사용하여 특정 이미지를 촬영한 사진작가에 비유할 수 있다. 이러한 관점에서, AI 사용자가 저작자로 간주되며, 따라서 최초의 저작권 소유자가 된다. 반면에 AI를 코딩하고 학습시키는 데 있어 창의적인 선택을 했다면, AI의 제작자는 카메라 제조업체보다 어떤 형태의 저작물에 대해 더 강력한 저작권을 주장할 수 있다.

AI 소프트웨어를 제공하는 회사는 회사의 서비스 약관과 같은 계약을 통해 회사와 사용자의 소유권을 각각 할당하려고 시도할 수 있다. 예를 들어 OpenAI의 이용 약관은 사용자에게 모든 저작권을 할당하는 것으로 보인다: “OpenAI는 출력물에 대한 모든 권리, 소유권 및 이권을 귀하에게 양도한다.” 반면, 이전 버전에서는 OpenAI에 이러한 권리를 부여하는 것으로 되어 있었다. 한 학자의 말처럼, OpenAI는 “계약을 통해 대부분의 저작권 문제를 우회하는 것처럼 보인다.”

3. 국내에서의 논의

가. 생성형 AI 작품의 저작권 인정

1) 저작권 인정 여부

저작권법은 ‘저작물’을 “인간의 사상 또는 감정을 표현한 창작물”로,⁹⁶⁾ ‘저작자’를 “저작물을 창작한 자”로 정의하고,⁹⁷⁾ ‘저작권’을 “저작자가 저작권격과 저작재산권을 가진다”고 규정하고 있다.⁹⁸⁾ 이는 현행 저작권법상 인간이 아닌 기계 또는 소프트웨어인 AI가 만든 작품은 저작물이 될 수 없고, 저작물이 되지 않기 때문에 저작자로 인정될 수 없으며, 결국 저작자가 아니기 때문에 저작권도 발생하지 않게 된다. 또한, 대법원은 저작권의 보호 대상을 “학문과 예술에 관하여 사람의 정신적 노력에 의하여 얻은 사상 또는 감정을 말, 문자, 음, 색 등에 의하여 구체적으로 외부에 표현한 창작적인 표현 형식”으로 정의하고 있다.⁹⁹⁾ 따라서 저작권법상 인간이 아닌 AI가 만든 작품에는 저작권이 인정될 수 없다. 하지만 AI 결과물에 인간이 수정, 편집 등을 하는 경우 인간의 창작성이 부가되어 저작물로 인정될 수 있다. 다만, 저작권은 인간의 창작성이 부가된 부분에 한정된다. 저작자는 또한 자연인만이 될 수 있고, 예외적으로 업무상저작물의 경우와 같이 법률에서 별도로 정하고 있는 경우에는 법인 등도 저작자가 될 수 있다.¹⁰⁰⁾ 그렇다면, AI에 대해서도 독자적인 법인격을 부여하여 권리를 귀속시킬 수 있다. AI에 법인격이 부여되면, 생성형 AI의 결과물에 대한 이익은 AI에 귀속되어 저작권이 발생할 수 있다. 다만, 이러한 법인격은 법률로 규정되어야 가능할 것이다.

96) 저작권법 제2조 제1호.

97) 저작권법 제2조 제2호.

98) 저작권법 제10조 제1항.

99) 대법원 2011. 2. 10 선고 2009도291 판결.

100) 저작권법 제9조.

2) 저작권 등록 가능 여부

저작권법은 저작권의 발생에 대하여 무방식주의를 취하고 있어, 어떠한 절차나 형식의 이행을 필요로 하지 아니한다.¹⁰¹⁾ 하지만 저작물의 성명, 창작 연월일, 공표 연월일 등과 저작재산권의 양도나 처분 제한 등 권리의 변동 사항을 등록할 수 있도록 한다. 이를 통해, 저작자로 등록된 자는 저작자로 추정되고, 권리 변동 사실을 제3자에게 주장할 수 있으며, 소송상 원고가 실손해를 증명하지 않아도 법에서 정한 금액을 손해액으로 인정할 수 있는 효과 등이 발생한다. 저작권법은 또한, “저작자 아닌 자를 저작자로 하여 실명·이명을 표시하여 저작물을 공표한 자”에 해당할 경우에는 1년 이하의 징역 또는 1천만 원 이하의 벌금에 처하도록 규정하고 있다.¹⁰²⁾ 다만, 이러한 등록은 인간이 저작자이거나 인간이 창작에 관여한 경우에만 가능하며, AI를 단독이나 공동 저작자로 하여 등록할 수 없으며, AI 결과물을 개발자 명의로 하거나 업무상저작물로 하여 대표 명이나 법인 명의로 등록을 신청할 수 없다.¹⁰³⁾ 실제, 유명 작가의 시를 텍스트 프롬프트로 하여, AI가 만들어낸 영상물을 2차적저작물로 등록을 시도한 사례에서, 한국저작권위원회는 “우리나라는 인간의 창작물만을 저작물로 보기 때문에 AI 산출물에 대해서는 저작권 등록”을 할 수 없고, “인간이 입력한 프롬프트를 활용하여 AI가 이를 영상화한 경우 그 행위는 AI만이 한 것으로 볼 수 있어 해당 영상물은 등록의 대상”이 아니라고 하면서, 등록을 반려한 바 있다.¹⁰⁴⁾

나. 생성형 AI 작품의 저작권 귀속

AI가 저작권을 갖지 못한다면, AI가 생성한 작품에 대하여 누가 이익을 얻고 책임을 져야 하는지, 즉 누구에게 권리를 귀속시켜야 하는지가 문제 될 수 있다. AI가 만든 작품과 관련하여 저작권의 귀속은 크게 AI가 전적으로 작품을 만든 경우와 AI가 만든 작품에 인간이 추가적으로 관여한 경우로 나눌 수 있다. 먼저, AI가 전적으로 작품을 만들어 인간이 아무런 가공을 하지 않은 경우에는 AI의 저작물을 인정하지 않기 때문에, AI는 저작권자가 될 수 없다. 대신, AI를 만든 개발사와 AI를 이용하여 서비스를 하는 회사에게 권리가 인정될 가능성이 있으며, AI 개발사는 AI가 작품을 만들 수 있는 알고리즘을 설계한 자, AI 학습 데이터를 세팅한 자 등이고, AI 서비스 제공자는 AI를 상품화하여 이용자들에게 제공한 자라고 할 수 있다.¹⁰⁵⁾

다음으로, AI가 만든 작품에 인간이 추가적으로 관여한 경우에는 저작권의 귀속 방식이 달라질 수 있다. 인간이 AI가 만든 작품에 창작적 요소를 가미하여 완전히 새로운 작품을 만들 수도 있고, 인간과 AI가 공동으로 창작한 것으로서 공동저작물로 인정될 수도 있으며, AI가 만든 작품에 변형을 가하여 2차적저작물로

101) 저작권법 제10조 제2항.

102) 저작권법 제137조.

103) 문화체육관광부, 생성형 AI 저작권 안내서, 2023, 41면.

104) 위의 책, 42면.

105) 양진영, 생성AI가 만든 작품의 저작권은 누구에게 있을까, 디지털데일리, 2023. 5. 1. <<https://m.ddaily.co.kr/page/view/2023050111354527180>, 2024. 1. 28. 최종방문>.

인정될 수도 있다. 첫 번째, AI가 만든 작품에 인간이 창작성을 가미하여 원작과는 다른 새로운 작품을 만드는 경우라면, AI가 만든 작품은 저작물의 재료로써 사용된 것이어서 인간의 저작권에 영향을 끼치지 않는다. 두 번째, AI가 만든 작품을 인간이 가공하는 경우, 그 가공 부분에 대하여 저작물성이 인정될 수 있고, 이 경우 공동저작물로 인정될 수 있는지가 문제 된다. 공동 저작자는 저작자의 일종이기 때문에, 자연인만이 인정되기 때문에, 공동 저작자로서 AI를 인정할 수는 없어 보인다. 또한, 우리 대법원은 공동 저작자가 되기 위해서는 “2인 이상이 공동 창작의 의사를 가지고 창작적인 표현 형식 자체에 공동의 기여를 함으로써 각자의 이바지한 부분을 분리하여 이용할 수 없는 단일한 저작물을 창작한 경우”여야 한다고 하면서, 공동 창작의 의사를 요구하고 있다.¹⁰⁶⁾ 여기서 “공동 창작의 의사는 법적으로 공동 저작자가 되려는 의사를 뜻하는 것이 아니라, 공동의 창작 행위에 의하여 각자의 이바지한 부분을 분리하여 이용할 수 없는 단일한 저작물을 만들어 내려는 의사를 뜻하는 것”이다.¹⁰⁷⁾ 그렇다면, 인간이 생성형 AI와 공동 창작의 의사를 갖는 것은 어려워 보인다. 따라서 AI의 작품에 인간이 가공을 한 경우, 공동 저작자로서 인정되기는 어려워 보인다. 세 번째 그리고 마지막으로, AI가 만든 작품에 인간이 변형을 하여 작품이 만들어진 경우에는 2차적저작물이 될 수 있다. 저작권법은 “원저작물을 번역·편곡·변형·각색·영상 제작 그 밖의 방법으로 작성”한 경우 2차적저작물로서 보호한다.¹⁰⁸⁾ 2차적저작물이 성립하기 위해서는 또한, 원저작물에 의거하였고 실질적으로 유사하다는 관점에서의 의거성과 원저작물을 변형한 것이라는 관점에서의 창작성이 존재해야 한다. 그리고 2차적저작물을 작성하기 위해서는 원저작물의 저작권자 동의가 필요하다. AI가 생성한 작품에 대하여 변형을 가하는 경우, AI가 생성한 작품은 저작권이 없는 저작물, 즉 공유 재산(public domain)에 해당하여 2차적저작물로 원칙적으로 성립할 수 없어, 원저작물의 저작권자로서 AI에게 동의를 받아야 할 필요가 없다. 다만, AI가 기존의 다른 저작물을 학습하여 작품을 만드는 경우에는 AI의 출력물이 타인의 저작물을 이용한 것이어서 2차적저작물 작성권을 침해한 것이 될 수 있다.

106) 대법원 2014. 12. 11. 선고 2012도16066 판결.

107) 위의 판결.

108) 저작권법 제5조 제1항.

IV

결어 - 시사점 및 향후 과제

앞에서 살펴본 바와 같이, 예술과 미디어 분야에서 생성형 AI와 관련된 저작권 분쟁은 매우 치열하게 전개되고 있다. 저작권자, AI 개발자와 서비스 제공자, AI 이용자들은 저작물을 생성하고 이용하는 데 있어 직접적 이해관계를 가지며, 모두 자기에게 유리한 지위를 차지하기 위해 노력하고 있다. 저작권자들은 자신들의 저작물이 AI 학습 및 저작물의 생성에 이용되는 것을 막거나 그에 대한 적절한 대가를 받으려 한다. AI 개발자와 서비스 제공자들은 저작물을 더 자유롭게 이용하기 위해 저작권 침해 책임 등에서 면제받기를 바란다. AI 이용자들은 저작물의 이용자로서 때로는 저작물의 생성자로서 법적 위험을 제거하고 위해 때로는 법적 이익을 취득하려고 한다. 이러한 이해관계의 대립 속에서 분쟁들은 크게 생성형 AI의 저작권 침해 문제와 생성형 AI의 저작자성 문제로 나눌 수 있으며, 이러한 분쟁들은 관련 이해관계자들의 합의나 명확한 정책적 기준이 확립되지 않아 앞으로도 한동안은 지속될 것으로 예상된다.

먼저, 생성형 AI의 저작권 침해 분쟁 사건들을 살펴보면, 최근 제기된 일련의 소송에서 저작권자들은 생성형 AI에게 저작권 직접 침해와 간접 침해를 주장하거나 저작권 관리 정보를 무단으로 제거 또는 변경하였다 주장하였고, 생성형 AI는 저작권 침해에 대한 방어로 공정이용이 성립하여 저작권 침해가 성립하지 않는다고 반박하였다. 많은 분쟁이 현재 계속 중이어서 명확한 결론을 내릴 수는 없지만, 제1심이 내려진 한 소송의 법원은 저작권자들이 AI 모델 학습에 저작권이 있는 자료를 사용하는 것이 적법한지에 대하여 저작권 직접 침해와 대위 침해 및 DMCA 위반에 대한 원고의 청구를 기각하였다. 하지만 그 결정은 생성형 AI 업체에 다소 유리한 판단으로 보이나, 실제 그 내용을 보면 저작권 침해와 DMCA 위반에 대해 긍정하는 태도를 보여, 생성형 AI의 저작권 침해에 대한 문제를 해결하는 판단이라고 보기는 아직 어려워 보인다. 다만, 침해에 대한 책임이나 공정이용의 판단들에서 저작물의 취득 경로와 대가를 지급하지 않은 부분 그리고 시장의 대체제로 작용할 수 있는 면을 강조하는 것을 보면, 저작권자들에게 저율의 추가 다소 기울어 있는 듯하다. 그리하여, 생성형 AI는 저작권 침해에 대한 책임을 회피하기가 쉽지는 않을 것으로 보인다. 이는 저작권 침해에 대하여 간접적으로 책임을 지는 간접책임과 저작권 관리 정보의 제거에도 같은 영향을 끼치고 있는 듯하다.

다음으로, 생성형 AI의 저작자성과 관련하여, 최근 제기된 분쟁들은 모두 저작자성을 인정하지 않는 듯한 태도를 보인다. 미국은 컴퓨터 알고리즘에 의해 자율적으로 생성된 작품에 대해 저작권을 인정하지 않는 원칙에 따라, AI가 생성한 작품에 대해 저작권청과 법원들 모두 저작자로 인정하지 않고 있다. 특히, 생성형 AI가 생성한 작품에 대해 인간이 가공한 경우의 저작자성에 대해서도 논의하고 있는데, 생성형 AI가 어떤 이미지를 생성할지 미리 예측할 수 없기에, 사용자는 프로세스를 제어할 수 없고, 그렇기 때문에 생성형 AI 결과물은 인간이 개입할 수 없는 구조로 간주하고 있는 듯하다. 더 나아가 인간이 생성형 AI 결과물을 편집한 경우에도, 그 편집이 저작권을 가질 수 있을 만큼 충분히 창조적이어야 한다고 하면서 그 창작성을 다소 높이고 있다. 따라서 생성형 AI의 저작자성을 인정하는 것은 현재 다른 방법이 강구되지 않는 한 어려워 보인다. 또한, 이러한 경우에 생성형 AI 결과물에 대한 권리를 누구에게 귀속할지가 문제되는데, 이와 관련된 판단이 존재하지 않아, 명확하지 않으나 그와 관련된 인간으로서 AI 개발자나 서비스 제공자 또는 이용자가 저작권 귀속의 대상으로 고려되고 있다.

우리나라는 이러한 분쟁들이 지금까지 많이 발생하고 있지 않지만, 생성형 AI 개발이 대대적으로 이루어지고 있는 지금 그 문제에 맞닥뜨릴 날이 머지않았다. 우리 저작권법은 아직 텍스트 마이닝에 대해서도 명확히 규정하고 있지 않아, 생성형 AI의 학습에서 저작물의 이용이 저작권 침해에 대한 면책이 될지도 분명하지 않다. 따라서 현재로서는 저작물을 학습하는 경우 저작권 침해에 대한 위험이 낮다고 볼 수 없으며, 이와는 별개로 결과물로서 작품을 생성한 경우 타 저작물을 이용한 경우나 실질적 유사성이 있는 경우 침해의 위험성도 높다. 사실, 학습에서의 저작물 이용은 일반적인 저작물의 이용과 다르고, 공개되지 않는 것이어서 공정이용으로 인정될 가능성이 있지만, 작품을 생성하는 경우에는 기존 저작물의 시장 등에 영향을 미치지 않을 수 없기에 더 큰 위험이 따를 수밖에 없다. 그리고 저작자성과 관련해서 우리나라는 저작물을 법적으로 명확하게 인간의 창작물로 한정하고 있기 때문에, 별도의 법인격을 부여하지 않는 한, 생성형 AI의 저작자성을 인정하기 어렵다. 따라서 AI 개발자나 서비스 제공자 그리고 이용자 등 밀접한 관련이 있는 자에게 기술의 특성을 포함한 구체적 사안에 따라 권리를 귀속시킬 수밖에 없어 보인다. 또한, AI가 생성한 작품을 인간이 개입하여 가공한 경우에도 공동저작물이나 2차적저작물이 성립되기는 어려우며, 다만 타인의 저작물에 대한 2차적저작물 작성권 침해만 성립할 수 있을 뿐이다. 따라서 생성형 AI의 발전을 도모하기 위해 이러한 문제를 해결하는 방안들이 마련되어야 한다.

생성형 AI의 발전과 저작권의 보호 양자의 이익을 조화시키기 위해서는 먼저, 생성형 AI에 의해 제기된 저작권법상의 문제 중 저작권법의 개정이 필요한지를 고려해 보아야 한다. 가장 먼저는 AI가 생성한 저작물에 대하여 저작권을 여전히 인정하지 않을 것인지, 그리고 만일 생성형 AI의 저작권을 인정하지 않는다면 누구에게 그 권리를 귀속시킬 것인지, 그리고 생성형 AI를 학습시키는 과정이 공정 이용에 해당하는지를 명확히 하는 입법이 고려되어야 한다. 물론, 현재로서는 법원이나 행정에서의 사례가 많지 않아 입법에 신중을 기할 필요가 있으나, AI의 작품들이 많이 나오는 지금 이러한 사항들에 대한 입법을 적극적으로 검토할 필요가 있어 보인다. 그리고 이러한 입법이 이루어지기 전에 저작권 침해 방지를 위한 조치, 대립하는 이

해관계 속에서 이익의 조정, 그리고 각종 판단에서 세부적이고 구체적인 기준들을 확립해야 할 필요가 있어 보인다. 먼저, 저작권 침해 방지를 위한 현실적 조치로서 AI 사업자는 필터링 조치 등을 통해 서비스 제공 시에 기존 저작물과 동일하거나 유사한 산출물이 도출되지 않도록 해야 할 것이다. 다음으로, 저작권자와 생성형 AI와의 이해관계 대립 속에서 양자가 정당한 보상을 받고 이용할 수 있도록 하는 시스템을 만들어, 분쟁을 미연에 방지할 수 있게 하는 것이 필요해 보인다. 최근 확대된 집중 관리나 보상금 청구권이 고려되고 있다. 그리고 마지막으로 생성형 AI가 창작한 저작물의 저작자가 누구인지, 인간의 저작물과 비교하여 다른 기준으로 저작권을 인정할지, AI 사용자가 생성된 저작물의 저작자로 인정받기 위해 어느 정도의 입력이 필요한지 등의 판단에 필요한 기준들을 정립해야 할 필요가 있어 보인다.

참고문헌

[국내 문헌]

김병필, 대규모 언어모형 인공지능의 법적 쟁점, 정보법학 Vol. 26, no. 1, 한국정보법학회, 2022.

김윤명, 생성형 AI와 저작권 현안, KISDI AI Outlook 2023 Vol. 13.

문화체육관광부, 생성형 AI 저작권 안내서, 2023.

양진영, 생성AI가 만든 작품의 저작권은 누구에게 있을까, 디지털데일리, 2023. 5. 1., <https://m.ddaily.co.kr/page/view/2023050111354527180>,

최호진, 썸네일 이미지와 공정이용, Law and Technology, 제8권 제3호, 서울대학교 기술과 법센터, 2012.

한국저작권보호원, 2020 저작권 보호 자체점검 지침서.

한국정보법학회 지음, 인터넷 그 길을 묻다, 중앙 Books, 2012.

elastic, 생성형 AI란 무엇인가?, <https://www.elastic.co/kr/what-is/generative-ai>.

[국내 판례]

대법원 2021. 6. 30. 선고 2019다208061 판결.

대법원 2019. 2. 28. 선고 2016다271608 판결.

대법원 2014. 12. 11. 선고 2012도16066 판결.

대법원 2011. 2. 10 선고 2009도291 판결.

[해외 문헌]

U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES (3d ed. 2021)

Sarah Andersen, “The Alt-Right Manipulated My Comic. Then A.I. Claimed It.”, 2022. 12. 31., <https://www.nytimes.com/2022/12/31/opinion/sarah-andersen-how-algorithm-took-my-work.html>,

Congressional Research Service, Generative Artificial Intelligence and Copyright Law, 2023, <https://crsreports.congress.gov/product/pdf/LSB/LSB10922>.

[해외 판례]

Main Sequence, et al. v. Dudesy LLC, et al., 2:24-cv-00711 (C.D. Cal. 2024).

Basbanes v. Microsoft Corp. and OpenAI, et al., 1:24-cv-00084 (S.D.N.Y. 2024).

New York Times Company v. Microsoft Corp. and OpenAI, et al., 1:23-cv-11195 (S.D.N.Y. 2023).

Sancton v. OpenAI Inc., Microsoft Corporation, et al., 1:23-CV-10211 (S.D.N.Y. 2023).

Concord Music Group, Inc. v. Anthropic PBC, 3:23-cv-01092 (M.D. Tenn. 2023).

Authors Guild, et al. v. OpenAI, Inc., 1:23-cv-08292-SHS (S.D.N.Y. 2023).

Chabon v. OpenAI, Inc., 3:23-cv-04625 (N.D. Cal. 2023).

Silverman, et al. v. OpenAI, Inc., 3:23-cv-03416 (N.D. Cal. 2023).

Kadrey, et al. v. Meta Platforms, Inc., 23-cv-03417-VC (N.D. Cal. 2023).

Tremblay v. OpenAI, Inc., 3:23-cv-03223 (N.D. Cal. 2023).

Getty Images (US), Inc. v. Stability AI, Inc., 1:2023cv00135 (E.D. Mo. 2023).

Anderson v. Stability AI Ltd. et al., 3:23-cv-00201 (N.D. Cal. Oct. 30, 2023).

Doe v. GitHub, Inc., et al., 22-cv-06823-JST (DMR) (N.D. Cal. Nov. 28, 2023).

Thomson Reuters Enterprise Centre GmbH et al v. ROSS Intelligence Inc., 1:20-cv-613-SB (D. Del. 2023).

Thaler v. Creativity Machine, No. 22-1564 (D.D.C. Aug. 18, 2023).

Logan v. Meta Platforms, Inc., __ F. Supp. 3d __, 2022 WL 14813836 (N.D. Cal. Oct. 25, 2022).

Mollman v. Zoetop Bus. Co., No. CV 22-4128 (PA) (GJSx), 2022 WL 17207103 (C.D. Cal. Sep. 16, 2022).

O'Neal v. Sideshow, Inc., 583 F. Supp. 3d 1282 (C.D. Cal. 2022).

Skidmore v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020).

Mango v. Buzzfeed, Inc., 970 F.3d 167 (2d Cir. 2020).

Naruto v. Slater, 888 F.3d 418 (9th Cir. 2018).

Microsoft Corp. v. My Choice Software, LLC, No. SA CV 16-2189-DOC (KESx), 2017 WL 5643210 (C.D. Cal. Oct. 10, 2017).

ICONICS, Inc. v. Massaro, 192 F. Supp. 3d 254 (D. Mass. 2016).

Friedman v. Live Nation Merch., Inc., 833 F.3d 1180 (9th Cir. 2016).

Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).

ITC Textile Ltd. v. Wal-Mart Stores Inc., Case No. CV-12-2650-JFW (AJWx), 2015 WL 12712311 (C.D. Cal. Dec. 16, 2015).

Williams v. Cavalli, No. CV 14-06659-AB (JEMx), 2015 WL1247065 (C.D. Cal. Feb. 12, 2015).

Murphy v. Millennium Radio Grp., LLC, Civil Action No. 08-1743 MAS, 2015 WL 419884 (D.N.J. Jan. 30, 2015).

Merideth v. Chi. Trib. Co., LLC., No. 12 C 7961, 2014 WL 87518 (N.D. Ill. Jan. 9, 2014).

Leveyfilm, Inc. v. Fox Sports Interactive Media, LLC, 999 F. Supp. 2d 1098 (N.D. Ill. 2014).

Briggs v. Blomkamp, 70 F. Supp. 3d 1155 (N.D. Cal. 2014).

Columbia Pictures Indus., Inc. v. Fung, 710 F.3d 1020 (9th Cir. 2013).

Fox Broad. Co. Inc. v. Dish Network, L.C.C., 905 F. Supp. 2d 1088 (C.D. Cal. 2012).

Range Rd. Music, Inc. v. E. Coast Foods, Inc., 668 F.3d 1148 (9th Cir. 2012).

Kelly v. Chicago Park District 635 F.3d 290 (7th Cir. 2011).

MDY Indus., LLC v. Blizzard Entm't, Inc., 629 F.3d 928 (9th Cir. 2010).

Jarvis v. K2 Inc., 486 F.3d 526 (9th Cir. 2007).

Perfect 10, Inc. v. Visa Int'l Ass'n, 494 F.3d 788 (9th Cir. 2007).

Perfect 10, Inc. v. Cybernet Ventures, Inc., 167 F. Supp. 2d 1114 (C.D. Cal. 2001).

Urantia Found v. Kristen Maaherra, 114 F.3d 955 (9th Cir. 1997).

Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991).

Contra Eden Toys, Inc. v. Florelee Undergarment Co., 697 F.2d 27 (2d Cir. 1982).

Burrow-Giles Lithographic Co. v. Sarony 111 U.S. 53 (1884).

PROFILE



김 창 화

한밭대학교 공공행정학과 교수

Email: patzzang@hanbat.ac.kr / Tel: 010-3906-6746

학 력

- 1992. 3. - 2001. 2. 인하대학교 공과대학 반도체공학과 공학사
- 2001. 3. - 2008. 8. 인하대학교 일반대학원 지식재산법 박사과정 수료
- 2008. 9. - 2009. 5. 美 Wisconsin-Madison Law School 석사(LLM-LI)
- 2009. 9. - 2012. 12. 美 Wisconsin-Madison Law School 박사(SJD)

주요경력

- 2004. 10. - 2006. 8. 한국저작권위원회 선임연구원
- 2014. 3. - 2015. 8. 단국대학교 IT법학 융합과정 연구전담조교수
- 2015. 9. - 現 한밭대학교 공공행정학과 교수
- 한국지식재산학회 총무이사, 한국산업보안학회 교육이사, 지식재산경상학회 연구이사, 한국스포츠엔터테인먼트법학회 엔터테인먼트 이사 등
- 변리사 시험 출제·채점 위원
- 한국저작권위원회 감정전문위원
- 대전 고등검찰청 상소심의위원
- 음반산업발전특별위원회 이사

주요 연구

- 디지털 전환에 따른 새로운 유형의 디자인 보호를 위한 제도개선 연구, 2022
- 상표권자의 손해배상 및 침해금지 등 청구권 제한에 관한 연구, 2021
- 데이터 이용 및 활용 촉진을 위한 제도 개선에 관한 연구, 2020

주요 논문

- 혁신과 창작의 균형을 위한 저작권법 목적의 해석에 관한 연구 - 미국 저작권법의 목적과 원칙을 중심으로, 법학연구 제26권 제1호, 2023. 3.
- 디지털 디자인의 보호 방안에 관한 연구 - 디자인의 물품성 개선을 중심으로, 스포츠엔터테인먼트와법 제26권 제1호, 2023. 2.
- 온라인에서의 저작물 이용 개선 방향에 관한 연구, 정보법학 제26권 제1호, 2022. 4.
- 지식재산 라이선스 거절의 경쟁법 위반 여부에 관한 판단 - 미국의 규정 및 사례와의 비교를 중심으로, 아주법학 제15권 제4호, 2022. 2.
- Google v. Oracle 판결에서의 공정이용에 관한 연구, 산업재산권 제68권, 2021. 7.
- 데이터 마이닝과 저작권 면책의 범위 및 한계, 계간저작권 제34권 제2호, 2021. 6.
- 이스포츠 개인 방송의 공정이용 적용에 관한 연구, 스포츠엔터테인먼트와법 제24권 제1호, 2021. 2.

규제Review 법제리뷰

제 24-1호

발 행 2024년 2월 29일

인 쇄 2024년 2월 29일

발행인 **한영수**

발행처 **한국법제연구원**

세종특별자치시 국책연구원로 15

(반곡동, 한국법제연구원)

전화: (044) 861-0300

homepage: <https://klri.re.kr>

인 쇄 **(주)법신사**

ISSN 2982-6454