

인공지능 관련 저작권 침해¹⁾

I. 논의의 전제 조건과 본고의 기본적인 생각

1. 문제되는 인공지능 기술의 특정 필요
— 심층학습 방법론 기반 인공지능의 상정
2. 단계별 실행과정 기반의 논의 필요
3. 본고의 기본적인 생각

II. 기계학습 과정에서 학습데이터의 이용에 의한

저작권 침해 가능성

1. 주요 기계학습의 학습알고리즘
2. 저작권 침해가 발생할 수 있는 인공지능 학습단계
3. 인공지능의 학습 내지 TDM을 위한 저작권 제한의 요건

III. 인공지능 모델의 창작행위에서 발생할 수 있는

저작권 침해 가능성

1. 인공지능의 창작행위가 타인의 선행 저작권을 침해하는 경우
2. 추가적 쟁점
— 오픈소스 라이선스에서 판매금지 조건 부가의 법적 효력



진윤준
법무법인 로고스
변호사

I. 논의의 전제 조건과 본고의 기본적인 생각

1. 문제되는 인공지능 기술의 특정 필요

— 심층학습 방법론 기반 인공지능의 상정

현재 인공지능 기술은 기계학습(machine learning) 나아가 인공지능경망에 기초한 심층학습(deep learning)에 근거하고 있다. 종래 전문가 시스템에 기초한 규칙기반의 인공지능 기술이 개발된 바 있으나 이러한 전문가 시스템은 기본적으로 외부에서 인간이 모든 규칙을 만들거나 조합을 해야 하는 문제가 있었다. 기계학습은 학습데이터(training data)를 입력하여 기계에게 학습을 하게 함으로써 기계 스스로 규칙을 찾아 예측을 할 수 있도록 한다. 대량의 데이터를 통한 학습을 하여야 하므로 이러한 학습과정에서 타인의 저작권 침해 등 법적인 문제가 발생한다. 또한 이와 같은 학습의 결과에 의해 채택된 규칙, 알고리즘을 인간이 이해할 수 있는지도 문제된다. 설명가능성을 원칙으로 하는 통계학적 방법론과는 달리 데이터사이언스는 예측의 정도를 중요시하기 때문에 알고리즘의 설명가능성은 데이터사이언스에서 이차적인 문제가 된다. 설명가능성의 부족은 저작권 관점에서 인공지능 생성물의 창작성 정도, 인공지능 생성물에 의한 타인 저작권 침해 시 의거성 판단에 영향을 준다.

종래 인공지능 관련 저작권 문제, 개인정보보호 문제가 논의될 때 검토대상이 되는 인공지능 기술의 방법론, 알고리즘이 구체적으로 특정되지 않고 쟁점이 정의되는 경우가 많았다. 이러한 경우 인공지능 관련한 법적인 문제가 추상화되어 인공지능 기술의 실체에 대응하는 결론이 도출되기 어렵다. 예컨대 인공지능의 유형을 약한 인공지능(weak AI), 강한 인공지능(strong AI), 초인공지능(super AI)으로 나누고 현행 인공지능을 약한 인공지능의 범주에서 큰 그림으로 논하는 방식이 여기에 해당한다. 이미지처리를 다루는 컴퓨터 비전(computer vision) 분야에서 제기되는 저작권, 개인정보보호의 문제와 인간의 언어를 다루는 자연어 처리(natural language processing) 분야에서 제기되는 저작권, 개인정보보호의 문제는 그 내용이 매우 상이하다. 인공지능 기술이 다양하고 급격히 변화, 발전하고 있어서 인공지능 기술을 총체적으로 특정하기 어렵다는 한계가 있으나 현재 주류적이고 특징적인 인공지능 기술을 구체적으로 상정하고 관련 법적인 문제를 논해야 한다고 생각한다.

2. 단계별 실행과정 기반의 논의 필요

법적인 관점에서 인공지능 관련 문제를 검토하기 위해서는 인공지능 기술의 실행 과정을 단계별로 분류하는 것이 필요하다. 인공지능에 관련된 저작권 침해 이슈는 인공지능 기술의 실행단계별로 구분하여 논의되어야 한다.

이른바 심층학습에 기초한 인공지능의 경우, 크게 보아 학습된 모델의 생성단계(학습단계)와 학습완료된 모델의 이용단계(이용단계)로 구분할 수 있다. 학습단계는 ① 학습용 데이터셋의 생성 및 입력 단계 ② 학습알고리즘에 의해 추론을 실행하는 학습용 소

1) 본고는 2021. 3. 19. 한국경영법률학회 2021년 춘계학술대회에서 '인공지능 관련 저작권 침해 및 대응' 제목으로 발표한 내용을 수정, 보완한 것이다.

소프트웨어를 구동하고 (하이퍼) 파라미터를 조정하여 학습완료된 인공지능 모델을 생성하는 단계로 나눌 수 있다. 이용단계는 ③ 인공지능의 창작물(생성물)을 얻기 위해 학습완료된 모델에 데이터(문자, 영상, 소리 등)를 입력하는 단계와 ④ 학습완료된 모델에서 인공지능 창작물(생성물)을 출력하는 단계로 구분할 수 있다. ①, ②의 학습단계는 텍스트/데이터마이닝(Text Data Mining; TDM)의 관점에서 a) 콘텐츠에 대한 접근 b) 콘텐츠의 추출 및 복제 c) 텍스트/데이터 마이닝 및 지식발견 단계로 나눌 수도 있다.²⁾ 이러한 각각의 단계를 구체적으로 고려하여 어느 단계에서 저작권 침해의 가능성이 있는지를 검토하여야 한다.

또한 인공지능 기술의 실행단계를 주체별, 객체별로 구분할 수도 있다. 인공지능 개발자측에서 수행하는 단계로는 ① 학습데이터의 수집, 생성, 입력 ② 학습용 소프트웨어의 실행(이용) ③ 학습된 파라미터(하이퍼 파라미터 포함)의 생성 및 조정 ④ 학습완료된 모델의 생성을 들 수 있고, 인공지능 이용자측은 ⑤ 학습완료된 모델에 대해 특정 데이터를 입력하거나 이를 기초로 인공지능 창작물(생성물)을 출력하는 행위를 하게 된다. 이와 같이 인공지능 개발자와 인공지능 이용자가 관여하는 단계 및 객체가 분리된다는 점은 인공지능에 의한 저작권 침해시 침해의 주관적 인식 내지 의거성 판단에서 영향을 주게 된다.

3. 본고의 기본적인 생각

인공지능 관련 저작권 침해는 학습단계인 학습용 데이터셋의 생성 및 입력 단계와 이용단계인 인공지능 모델에서 인공지능 창작물(생성물)을 출력하는 단계에서 주로 발생된다. 국내외 논의를 살펴보면 대체로 인공지능의 학습단계에서 발생할 수 있는 저작권 침해 가능성이 인공지능 학습모델의 이용단계에서 발생하는 저작권 침해 가능성보다 중요하다고 보고 있다.³⁾ 이는 인공지능의 학습이 대량의 학습데이터를 이용한 TDM 방식으로 이루어지기 때문에 인공지능 분야의 발전을 위해서는 학습단계에서 발생할 수 있는 저작권침해의 문제를 시급히 해결할 필요가 있고, 인공지능 이용단계에서 발생할 수 있는 저작권 침해는 기존의 법리에 따라 처리할 수 있다고 보는 것으로 생각된다.

인공지능 학습단계에서 발생할 수 있는 저작권 침해의 문제는 결국 TDM 면책 조항의 설정에 의해 해결할 수밖에 없다. 최근 발의된 저작권법 개정안 제43조(정보분석을 위한 복제·전송)의 요건에서 보듯이, ① TDM의 목적을 학술적, 비상업적 용도로 제한할 것인지 여부 ② 미국의 일부 견해가 차용하고 있는 '비표현적 이용(non-expressive use)'에 영향받은 것으로 보이는, 위 저작권법 개정안 규정의 '저작물에 표현된 사상이나 감정을 향유하지 않는 이용'의 의미 ③ '해당 저작물에 적법하게 접근할 수 있는 경우'의 범위 등이 검토되어야 할 것이다.

인공지능 이용단계에서 발생할 수 있는 저작권 침해의 문제는 기존의 실질적 유사성 법리, 의거성 법리에 의해 어느 정도 해결할 수 있을 것으로 생각된다. 인공지능 이용

단계에서 저작권 침해가 발생한 경우 침해자는 인공지능 창작물 및 창작행위에 대해 전반적인 지배통제권을 확보한 자로 특정되어야 하고 이는 인공지능 창작물에 대한 보호론에서 권리자로 인정받는 자와 동일하다고 볼 수 있다. 다만, 실무적으로는 의거성 입증에서 상당한 곤란이 있을 것으로 예상되고, Edmond de Belamy 초상화 사례에서 보듯이 창작행위의 지배통제권자 확정이 어려울 수도 있다.

II. 기계학습 과정에서 학습데이터의 이용에 의한 저작권 침해 가능성

1. 주요 기계학습의 학습알고리즘

기계학습 알고리즘은 아래 외에 다양하게 존재하고 계속 발전하고 있다. 인공지능 관련 저작권 문제를 실제적으로 다루기 위해서는 구체적인 인공지능 알고리즘을 특정하여 논의할 필요가 있다.

(1) CNN(Convolutional Neural Network)

우리말로는 합성곱신경망이라고 한다. 컴퓨터가 이미지를 이해하고 높은 수준의 추상화된 정보를 추출하는 인공신경망의 종류이다. 사람의 시신경 구조를 모방하였으며, 학습데이터의 특징(feature)을 추출하여 숫자로 표현하고 숫자로 표현된 특징들의 패턴을 파악하는 방식으로 객체를 인식한다. Convolution 과정과 pooling 과정을 반복한다.

CNN은 사물의 특징을 자동으로 잡아내면서 이미지 인식에 대단히 뛰어난 성능을 보인다. 그러나 CNN의 학습과정은 순서가 중요하지 않은 정보들이 공통으로 가지는 특징들만 관심이 있기 때문에 실시간으로 들어오는 정보들의 순서 관계를 처리할 수 없는 문제점이 있다.⁴⁾ 이에 과거 및 현재 정보를 기반으로 미래 정보를 예측하는 시계열 특성을 반영한 RNN 심층학습 모델이 등장하였다.

(2) RNN(Recurrent Neural Network)

우리말로는 순환신경망이라고 한다. RNN 알고리즘은 반복적이고 순차적인 데이터(sequential data) 학습에 특화된 인공신경망으로서 내부의 순환구조가 있다는 특징이 있다. 순환구조를 이용하여 과거의 학습을 가중치(weight)에 의해 현재 학습에 반영한다. 학습이 진행되어도 과거 학습의 정보를 잃지 않고 연속적인 정보의 흐름을 학습에 반영할 수 있다. 단어 시퀀스(sequence) 단위로 입력과 출력을 처리할 수 있으므로 자연어 처리, 음식인식 등에 사용된다.

(3) GAN(Generative Adversarial Network)

GAN은 이미지, 영상에 특화된 심층학습 알고리즘으로 인공지능 예술 분야에서 가장 많이 활용된 기술로 알려져 있다. 비지도학습으로서 두 개의 모델이 적대적 또는 경쟁적 과정을 통해 동시에 학습된다. GAN의 목적은 생성자(generator, 예술가 역할)와 판별자(discriminator, 예술비평가 역할)를 경쟁적으로 학습시켜 '그럴듯한 가짜'를 만드는

2) Eleonora Rosati, "Copyright as an obstacle or an enabler? A European perspective on text and data mining and its role in the development of AI creativity", Asia Pacific Law Review, 27:2, p. 204 (2019); 이규호, "인공지능 학습용 데이터셋에 대한 저작권법과 부정경쟁방지법상 보호와 그 한계", 인 권과 정의 Vol. 494, 2020, 91면.

3) Benjamin Sobel, "Artificial Intelligence's Fair Use Crisis", Columbia Journal of Law & the Arts, 2017, 65면.

4) 이은주, "CNN과 RNN의 기초 및 응용연구", 방송과 미디어 제22권 1호, 2017, 90면.

것이다. 생성자는 진짜처럼 보이는 이미지를 생성하도록 배우고, 판별자는 가짜의 이미지로부터 진짜를 구별하는 것을 배우게 된다.

학습과정 동안 생성자는 점차 진짜 같은 이미지를 더 잘 생성하게 되는 반면 판별자는 점차 진짜와 가짜를 더 잘 구별하게 된다. 이러한 과정은 판별자가 가짜 이미지에서 진짜 이미지를 더 이상 구별하지 못하게 될 때 평형상태에 도달한다. GAN은 생성자와 판별자를 꾸준히 학습시켜 종국적으로 진짜와 다른 없는 가짜를 만든다(정확히는 실제 데이터의 분포에 가까운 새로운 데이터를 생성한다는 것이다). GAN 역시 기계학습이므로 GAN은 객체를 맥락적(contextual)으로 이해하는 것이 아니라 객체의 특징(feature)을 세분화하여 각기 숫자로 이해할 뿐이다(다만 이러한 유형의 인공지능은 객체들의 특징을 잡아 의미적 연관성 내지 관계적 유사성을 파악하는 수준에는 이르렀다고 볼 수 있다). 그러나 GAN의 최종적인 이미지 결과물을 보면 마치 컴퓨터가 원본 이미지의 전체적인 표현, 분위기를 이해하여 원본 이미지의 표현을 복제한 것과 같은 느낌을 준다. 논쟁적 의견을 제시한다면, 인공지능 알고리즘이 원본 이미지의 비표현적 부분만 사용한다고 단정하기 어려운 사례라고 생각한다. 이러한 관점에 의하면 GAN을 이용한 인공지능 생성물은 TDM 면책규정을 적용받지 못하여 원본 저작권에 대한 침해물로 인정될 가능성이 높다.

(4) 자연어 처리 언어 모델(Language Model)

언어모델은 언어라는 현상을 모델링하기 위해 단어 시퀀스(또는 문장)에 확률을 할당하는 모델이다. 언어모델은 크게 통계를 이용한 방법과 인공신경망을 이용한 방법으로 구분할 수 있다. 전통적으로는 통계적 언어모델이 사용되어 왔으나 최근에는 인공신경망을 이용한 방법이 더 좋은 성능을 보여주고 있다고 한다. BERT, GPT 등은 모두 인공신경망 언어모델에 속한다.

자연어 처리에서 가장 일반적으로 시도하는 방법은 이전 단어들에 주어졌을 때 다음 단어를 예측하게 하는 것이다. 통계적으로는 학습데이터셋에서 실제로 다음 단어(또는 마스킹된 단어)가 나오는 빈도 내지 확률을 구하여 예측을 한다. 인공신경망을 이용한 방법론으로는 단어의 특징을 추출하여 단어를 공간상의 벡터(숫자로 된 실수 벡터)로 표현(맵핑)하는 단어 임베딩(word embedding) 기술을 들 수 있다(현재의 심층학습 모델은 객체를 숫자화하여야 학습이 가능하다). 단어 임베딩 기술을 이용하면 비슷한 의미를 가지는 단어들을 벡터공간에서 서로 가까운 곳에 나타나게 하는 방식으로 단어 간의 유사성을 파악할 수 있다. 단어 임베딩 기술을 활용한 인공신경망 인공지능은 숫자벡터에 의해 단어 간의 유사성, 관계를 파악하고 있으므로 어느 정도 문장의 맥락(context)을 파악하는 기술이라고 볼 수 있다. 인공신경망을 이용한 문장번역 분야에서는 최근 RNN 알고리즘이 적용된 sequence to sequence 기술이 사용되고 있다. 위 기술은 단어 단위가 아닌 문장(sequence) 단위를 다른 문장(sequence)으로 변환하는 것이다. 입력된 문장(예컨대 국문으로 된 문장)을 인코더(encoder)를 이용하여 고정크기의 벡터(context vector라고 함)로 압축하고 디코더(decoder)에 의해 이 벡터를 디코딩하여 결과(예컨대 영문으로 된 문장)를 생성하는 방식이다. 위 sequence to sequence 방식도 관점에 따라서는 문장의 맥락을 파악하는 것으로 이해할 수 있다.

2. 저작권 침해가 발생할 수 있는 인공지능 학습단계

인공지능 학습단계 내지 TDM 과정에서 학습데이터를 생성하고 이를 인공지능 학습용 소프트웨어에 입력할 때 입력데이터의 복제가 일어난다. 특히 인공지능 학습은 데이터 전처리(pre-processing)를 거친 데이터 전부 내지 대다수를 학습데이터로 입력하는 경우가 많으므로 데이터에 대한 적절한 이용권한이 없다면 대규모의 저작권 침해가 발생한다.

TDM 관점에서 이러한 입력단계를 풀어 쓰면, 데이터에 접근하는 단계(1단계), 데이터를 추출 및 복제하는 단계(2단계), 텍스트/데이터 마이닝 및 지식 발견(3단계)으로 구분할 수 있다. 그 가운데 특히 제2단계에서 법적 쟁점이 발생할 가능성이 크다.⁵⁾ 제2단계에서는 추출 및 복제할 데이터가 저작물인지 여부, 데이터베이스에 포함되어 있는지 여부(우리나라, EU는 데이터베이스 보호 규정 존재), 다른 법률상 제한 사유(개인정보보호, 프라이버시, 계약 등)가 존재하는지 여부 등을 판단해야 한다.

제3단계에서도 저작물의 복제가 일어날 수 있다. 제3단계는 데이터 전처리(pre-processing)와 구조화된 데이터를 추출하는 단계를 포함한다.⁶⁾ 제3단계는 토큰화(tokenization), 정제(cleaning), 정규화(normalization), 어간추출(stemming), 표제어추출(lemmatization), 불용어(stopword) 삭제 등을 포함한다. 이 과정은 인간의 수작업으로 이루어지는 경우도 있기 때문에 민감한 데이터인 경우에 프라이버시 보호 차원의 문제도 발생할 수 있다.

3. 인공지능의 학습 내지 TDM을 위한 저작권 제한의 요건

(1) 문제 상황

1) 인공지능 학습단계에서 발생하는

저작권침해를 전면적으로 면책하는 경우

인공지능의 기계학습을 위해서는 타인의 저작물을 이용하는 것이 불가피하다. 인공지능 학습데이터 확보를 위하여, 인공지능 학습용으로 제공되는 저작물에 대해서는 저작권제한이 필요하다는 견해가 가능하다. 그러나 인공지능 학습단계에서 발생할 수 있는 저작권침해를 전면적으로 면책하는 정책방향은 다음과 같은 단점을 가지고 있다. ① 종래 저작권자가 독점권자로서 강자이고 저작물 이용자가 소비자로서 약자라는 구도는 현재 상황과 완전히 부합되지 않는다. 저작권자가 개인 또는 소규모 사업자이고 저작물 이용자는 구글과 같은 대규모 플랫폼 사업자인 상황이 도래하였다. 이와 같은 상황에서 저작권자가 인공지능 발전을 위해 자신의 권리를 전면적으로 희생할 필요는 없다고 보인다. ② 인공지능 학습모델이 창작한 결과물에 대해 저작권 내지 그와 유사한 권리를 부여하려는 논의가 있는 상황에서, 학습데이터를 제공한 저작권자에게 아무런 보상이 없다는 것은 형평에 어긋나고 논리적이지도 않다. 이러한 제반 상황을 고려하여 적절한 범위의 인공지능 학습목적 내지 TDM을 위한 저작물 이용범위를 설정할 필요가 있다.

5) Eleonora Rosati, 앞의 논문, 204면; 이규호, 앞의 논문, 91면.

6) Eleonora Rosati, 앞의 논문, 209면.

2) 인공지능 학습단계에서 요구되는 타인의 저작물에 대해 어떠한 권리제한도 하지 않는 경우

원칙적으로 인공지능 개발자가 저작권자로부터 이용허락을 받아 해당 저작물을 학습데이터로서 활용하는 것이 가장 바람직하겠지만, 저작물의 수량, 저작권자 확인 불가, 과도한 이용료 등의 현실적인 문제로 인하여 모든 학습데이터에 대해 저작물 라이선스를 받기 어려운 경우도 상당수 존재한다. 이러한 현실에도 불구하고 모든 경우에 저작물 라이선스가 필요하다는 정책을 펴게 되면 어떠한 상황이 나타날 것인가? 인공지능 개발자는 무상으로 공개된 데이터를 사용하거나 저작권자로부터 유상의 라이선스를 얻어야 하므로, 결과적으로 라이선스 이용료를 지불할 수 있는 자본이 있거나 분쟁을 감당할 수 있는 대형 사업자만이 인공지능 개발을 할 가능성이 있다. 더욱 중요한 점은 즉각적으로 이용가능하거나 합법적으로 라이선스된 데이터만을 학습용 데이터로 사용하는 경우 학습할 수 있는 데이터의 종류와 수량이 극히 제한되어 인공지능 편향(bias)의 문제가 발생할 수 있다는 것이다.⁷⁾ 학습데이터는 현실을 반영하는 것이므로 현실세계에 편향이 있는 것과 같이 데이터 자체에도 편견이 있을 수 있다. 데이터에 내재한 편견, 차별적 요소를 수정할 수 있는 가장 효과적인 방법은 (인공지능 모델을 보정하는 것보다도) 학습데이터의 규모를 확대하는 것이므로, 저작권자에 대한 불이익을 최소화하면서도 저작물을 학습데이터로 이용할 수 있는 예외적 요건을 설정하는 것이 필요하다고 생각된다.

(2) 외국 입법례 및 저작권법 개정안 제43조 요건 검토

EU, 일본, 영국 등은 TDM 관련 예외 요건을 입법하였고, 미국은 판례법에 근거한 공정이용 법리에 따라 TDM을 허용하려는 논의를 하고 있다.⁸⁾ 자연어 처리에 관하여는 독일의 입법례가 주목된다. 독일 저작권법 제60d조에 따르면, 인공지능에 의한 자연어 처리에 사용될 수 있는 데이터베이스인 말뭉치(Korpus)를 생성하기 위하여 원자료를 복제하는 행위와 이를 비영리 목적의 학술연구를 위해 제3자에게 제공하는 행위가 허용된다(제1항). 다만 이 말뭉치(Korpus)와 원자료(Ursprungsmaterials)의 복제는 연구 수행이 종료되면 삭제되어야 하고 제3자 이용제공은 종료되어야 한다. 예외적으로 도서관, 기록보관소, 박물관, 교육시설 등에 영구 내지 장기간 보존하는 것은 허용될 수 있다(제3항). 위 규정은 2017.9.1. 제정되어 2018.3.1.부터 시행되었고 일단 2023.2.28.까지 시행되는 것으로 예정된 일몰규정이다.⁹⁾ 인공지능 자연어 처리를 위해서는 말뭉치 데이터베이스 확보가 필수적인데, 독일의 TDM 조항은 이를 정확히 특정하여 규정한 것이다.

우리나라 역시 TDM 내지 인공지능 학습이 원활하게 진행될 수 있도록 관련 저작권 재산권 제한 사유를 의원입법으로 발의한 상황이다.

저작권법 전부개정안(도종환 의원 대표발의)

제43조(정보분석을 위한 복제·전송)

- ① 컴퓨터를 이용한 자동화 분석기술을 통해 다수의 저작물을 포함한 대량의 정보를 분석(규칙, 구조, 경향, 상관관계 등의 정보를 추출하는 것)하여 추가적인 정보 또는 가치를 생성하기 위한 것으로 저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우에는 필요한 한도 안에서 저작물을 복제·전송할 수 있다. 다만, 해당 저작물에 적법하게 접근할 수 있는 경우에 한정한다.
- ② 제1항에 따라 만들어진 복제물은 정보분석을 위하여 필요한 한도에서 보관할 수 있다.

1) '컴퓨터를 이용한 자동화 분석기술을 통한 대량의 정보를 분석' 요건

인공지능의 학습 내지 TDM은 인공지능 학습알고리즘을 통하여 대량의 학습데이터를 분석하는 것을 전제로 하고 있으므로, 위 요건은 대체로 타당하다고 생각된다. 다만, 학습데이터를 인공지능 학습에 사용하기 위해서는 예컨대 자연어 처리의 경우 말뭉치(corpus)에서 구두점 등을 삭제하고 의미단위로 나누는 단어 토큰화(Word Tokenization), 말뭉치로부터 노이즈 데이터를 제거하는 정제(cleaning), 표현방법이 다른 단어들을 통합시켜서 같은 단어로 만들어 주는 정규화(normalization), 영상인식의 경우에는 데이터 라벨링(labeling) 등의 데이터 전처리(pre-processing) 과정을 거쳐야 하는데, 이러한 전처리 과정은 사람의 수작업으로 진행되는 경우도 많다. 데이터 전처리 과정에서도 저작물의 복제가 이루어지고 있으므로 이와 같이 사람의 수작업이 상당한 정도로 개입된 경우에도 위 요건을 충족시킬 수 있는지에 대해 일부 의문이 있을 수 있다.

2) '추가적인 정보 또는 가치를 생성하기 위한 것' 요건

표현의 의미를 고려할 때 위 요건이 매우 적극적인 역할을 하기는 어려울 것으로 생각된다. 타 입법례에서도 찾아보기 어려운 요건이다.

3) '저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우' 요건

위 요건은 일본 저작권법 제30조의4(저작물에 표현된 사상 또는 감정의 향수를 목적으로 하지 않는 이용)를 참고한 것으로 보이고, 근본적으로는 미국 일부 학설이 대규모 정보처리에 대해 공정이용 법리를 적용할 때 사용하는 용어인 '비표현적 이용(non-expressive use)'이라는 표현에서 시사를 받은 것으로 생각된다.¹⁰⁾

'비표현적 이용'이라는 표현을 사용하는 미국 학설의 견해에 따르면, '비표현적 이용'을 공정이용으로 인정한 대표적인 사건은 Authors Guild v. Google 사건¹¹⁾과 A.V. ex.

7) Benjamin Sobel, 앞의 논문, 81면.

8) 각국의 TDM 조항의 상세는 본고에서 다루지 않는다.

9) 이상의 내용은 김병일, "TDM 면책과 혁신에 대한 EU의 접근법의 시사점", 한국경영법학회 2021년 춘계학술대회 자료집, 88면 참조.

10) Matthew Sag, "Orphan Works as Grist for the Data Mill", 27 BERKELEY TECH. L.J. 2012, 1512면; James Grimmelmann, "Copyright for Literate Robots", 101 IOWA L. REV. 2016, 661면

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

rel. Vanderhye v. iParadigms 사건¹²⁾이라고 한다. 법원은 Google 사건에서 구글이 특정 검색 키워드가 텍스트에서 몇 번 발견되는지 등의 저작물 텍스트에 관한 '사실적인 정보(factual information)'를 공중에게 제공하기 위해 저작물을 복제한 것은 공정이용에 해당한다고 판시하였고, iParadigms 사건에서는 양 저작물의 텍스트적 유사성을 비교하는 것은 저작물의 창작적인 요소와 관계가 없다는 이유를 들어 iParadigms가 표절탐지도구에 사용하기 위해 학생들의 논문을 복제한 행위를 공정이용에 해당한다고 보았다.

위 미국 판결들은 대규모의 텍스트를 대상으로 하는 TDM이 이루어지는 행위에 대해 저작권이 보호하는 표현과 관계없는 '사실에 관한 정보', '비창작적인 요소'를 대상으로 복제가 이루어졌다는 점을 주요 근거로 삼아 공정이용을 인정한 사례들이라고 할 수 있다. 피고들이 사용한 프로그램을 일종의 인공지능이라고 가정한다면 이들 인공지능 프로그램은 비표현적 요소에 대해 TDM을 수행하였고 인공지능 학습을 하였다고 할 수 있다. 이러한 관점에서 보면, '저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우'라는 요건은 미국 학설에서 논하는 '비표현적 이용'을 풀어 쓴 용어로 생각되고 원칙적으로 타당한 요건이라고 생각된다.

그러나 최근 인공지능 학습모델의 발전과 이를 반영한 학설의 견해를 추적하여 보면, 현재 인공지능 학습모델이 창작(생성)한 결과물이 반드시 학습데이터(저작물)의 비표현적 요소만을 이용한 것이라고 단정할 수 없고 표현적인 기계학습(expressive machine learning)의 결과물로도 볼 수 있다는 지적이 나오고 있다.¹³⁾¹⁴⁾ 후술하는 Edmond de Belamy 초상화 사례에서 등장하는 GAN 모델 산출물의 표현력 수준을 보아도 인공지능 모델의 TDM이 반드시 '비표현적 이용'에 국한된다고 단정하기 어렵다고 생각된다.

4) '해당 저작물에 적법하게 접근할 수 있는 경우' 요건

위 요건은 일반적으로 해킹, 불법다운로드 등에 의해 불법적으로 저작물에 접근하는 경우를 배제하려는 것이라고 설명되고 있다. 실무적으로는 웹사이트 등에 대한 데이터 크롤링, 스크레이핑이 적법한 것인지 문제된다. 사건으로는 불법적 접근이 아니면 죽고 저작물에 대한 라이선스 등 적법한 이용권한까지 요구하는 것은 아니라고 생각된다.

위 요건은 원시 데이터 저작자의 권리와 인공지능 학습데이터 개발자 사이의 이해관계를 조율하는데 중요한 역할을 할 것으로 예상된다.¹⁵⁾ 다만 개정안은 원시 데이터의 저작권자가 TDM 유보를 어떻게 선언할 수 있는지에 관하여 특별한 설명이 없다. 관련하여 EU 디지털 단일시장 저작권지침(DSM Directive) 제4조, recital 18은, 원시 데이터를 온라인에서 이용할 수 있는 경우에는 메타데이터 내지 웹사이트 이용약관 등 기계가 읽을 수 있는 수단을 이용하여 저작권자가 명시적으로 TDM 유보를 선언해야 하고 그 외의 경우에는 계약이나 단독행위에 의해서도 TDM 유보를 선언할 수 있다고 규정하고 있다.

12) A.V. ex rel. Vanderhye v. iParadigms, L.L.C., 562 F.3d 630 (4th Cir. 2009)

13) Sobel은 expressive machine learning을 '표현적인 목적(expressive purpose)을 달성하기 위해 저작물의 표현적인 측면(expressive aspects)에서 수행되는 기계학습'이라고 정의한다. Benjamin Sobel, 앞의 논문, 49면.

14) 박현경, "인공지능 학습과정에서 저작물의 이용에 관한 소고", 스포츠엔터테인먼트와 법 제23권 제1호, 2020, 147면.

15) 이규호, 앞의 논문, 100면.

(3) 일반적 공정이용 조항과 비교

저작권법 제35조의5는 일반적 공정이용 요건을 규정한 것이므로, TDM 예외 요건을 일반적 공정이용의 요건과 비교하는 것도 의미가 있다.¹⁶⁾ 위 규정은 제2항에서 공정이용의 고려요소로서 ① 이용의 목적 및 성격 ② 저작물의 종류 및 용도 ③ 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성 ④ 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향을 들고 있다.

기계학습에서 주로 문제되는 요소는 ①번의 이용 목적과 성격, 그리고 ④번의 시장 또는 가치에 미치는 영향이라고 할 수 있다. 다만 인공지능 학습을 위한 입력데이터는 전부 복제되는 것이 통상적임에도 불구하고 ③번째 요건인 '이용된 부분의 비중과 중요성'에 어긋나지 않는다고 보는 이유는, TDM이 저작물의 사상이나 감정과 무관한 비표현적(non-expressive) 데이터만을 추출하여 이용한다고 보기 때문이다.¹⁷⁾ 만약 특정 인공지능 학습모델이 표현적 이용도 가능하다고 본다면 TDM이 허용될 여지는 위 고려요소에 의해서도 극히 제한될 것이다.

'① 이용의 목적 및 성격'에서 문제되는 상황은 상업적 목적으로도 인공지능 학습 내지 TDM이 허용되는지 여부이다. 우리나라 학설, 판례는 영리적인 목적을 위한 이용은 비영리적 목적을 위한 이용에 비해 공정이용의 범위가 좁아지기는 하지만 그러한 이유만으로 공정이용이 부인되지는 않는다고 보고 있다.¹⁸⁾ 이용목적의 상업성, 영리성은 공정이용에 대해 결정적인 판단요소가 될 수 없다고 보아야 하므로, 다른 고려요소와 종합하여 탄력적으로 해석될 필요가 있다. 다만 상업적 목적을 위한 인공지능 학습이 개별 사건에서 허용되는 경우 해당 인공지능 모델의 산출물에 대해 일정한 권리가 부여된다면, 인공지능 학습에 의해 저작권이 제한당한 원시데이터의 저작권자에게도 어떠한 형태로든 일정한 보상이 있어야 하는 것이 논리와 형평에 부합한다고 생각된다.

'④ 시장 또는 가치에 미치는 영향'이라는 고려요소에 대해서도 같은 논리가 가능하다. 인공지능 학습 내지 TDM이 잠재적 시장가치에 위협적이지 않다고 보는 이유는 이러한 인공지능 학습이 비표현적 이용에 해당한다고 보기 때문이다. 인공지능에 의한 저작물(원시데이터)의 표현적 이용이 가능하고, 인공지능 창작물에 대한 시장이 형성되어 원시데이터와 경쟁하는 상황이 도래한다면 위 고려요소에 대한 전반적인 재검토가 필요할 것이다.¹⁹⁾ 이러한 관점에서 원시데이터의 저작권자에 대한 보상체계가 검토될 필요가 있다.

16) 제35조의5(저작물의 공정한 이용) ① 제23조부터 제35조의4까지, 제101조의3부터 제101조의5까지의 경우 외에 저작물의 통상적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있다. ② 저작물 이용 행위가 제1항에 해당하는지를 판단할 때에는 다음 각 호의 사항등을 고려하여야 한다. 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

17) 박현경, 앞의 논문, 142면.

18) 이해완, 저작권법, 제4판(전연개정판), 박영사, 2019, 821면 및 오승중, 저작권법, 제4판(전연개정판), 박영사, 2016, 863면 모두 "영리성을 가지면 언제나 공정이용이 될 수 없고 비영리성을 가지면 무조건 공정이용이 되는 것은 아니다"라는 취지로 설명한다; 대법원 2014. 8. 26. 선고 2012도10786판결은 저작권법 제28조 적용 여부가 문제된 사안에서 "이 경우 반드시 비영리적인 이용이어야만 하는 것은 아니지만 영리적인 목적을 위한 이용은 비영리적인 목적을 위한 이용의 경우에 비하여 자유이용이 허용되는 범위가 상당히 좁아진다"고 판시한 바 있다; 미국 연방대법원은 Campbell. v. Acuff-Rose Music, 510, U.S. 569(1994) 사건에서 "저작물의 이용이 교육적이고, 영리를 목적으로 한 것이 아니라는 사실만으로 침해의 인정으로부터 자유로울 수 없는 것과 마찬가지로 이용의 영리적인 성격만으로 공정이용의 인정이 전적으로 배제되는 것은 아니다"라고 판시하였다.

19) 박현경, 앞의 논문, 147면.

III. 인공지능 모델의 창작행위에서 발생할 수 있는 저작권 침해 가능성

1. 인공지능의 창작행위가 타인의 선행 저작권을 침해하는 경우

(1) 문제되는 상황

개발자 A가 그림을 그리는 인공지능 학습모델을 완성하였고 이를 이용자 B가 이용하여 여러 장의 그림을 생성하였다. 그 중 일부 그림이 화가 C의 그림과 매우 유사하였다. C는 자신의 저작권 침해를 입증할 수 있는가? 침해자는 A, B 중 누구인가?

현재 인공지능 개발 현황을 보면 인공지능에 의한 음악 창작의 경우에도 위와 유사한 상황이 발생하고 있다고 할 수 있다.

(2) 침해자 적격

저작권법적인 논리에서 저작권이 부여되거나 투자/인센티브 이론에 의해 인공지능이 창작한 결과물에 대해 어떠한 권리가 부여된다면, 해당 권리를 부여받을 수 있는 자에게는 당연히 침해자 적격이 인정되어야 한다. 즉 인공지능의 산출물을 지배관리하여 그에 대해 권리 내지 경제적 이익을 취득할 수 있는 자는 인공지능이 발생시킨 저작권 침해 상황에 대해서도 법적 책임을 부담하여야 한다.²⁰⁾

나아가 이론상 저작권 내지 저작권 유사의 권리를 부여받을 수 없는 자라고 하더라도 침해태양에 따라 (공동)침해자가 될 수 있다고 생각된다. 위 사례에서 침해자는 A 또는 B가 될 수밖에 없고 경우에 따라 개발자 A의 사용자가 사용자책임을 부담할 것이다. 인공지능 창작물에 대해 권리를 부여받지 못한다고 하더라도 (공동)불법행위에 가담하여 침해자로 인정될 수 있다고 생각된다.

(3) 의거관계의 입증

저작권침해가 인정되기 위해서는 피침해 저작물의 저작물성, 의거성, 실질적 유사성이 인정되어야 한다. 복잡한 구조의 인공지능 학습모델, 예컨대 여러 개의 은닉층이 설정되고 파라미터가 많은 심층학습 모델이 채택하고 있는 알고리즘은 대체로 인간이 이해하기 어렵다는 점이 여기서 문제된다(블랙박스과 같은 인공지능 알고리즘).

저작권침해의 주관적 요건인 의거관계는 침해 저작물이 피침해 저작물에 의거하여 작성되었다는 것을 의미한다. 직접 증거에 의해 의거관계가 입증되기 어렵더라도 피침해 저작물에 대한 접근가능성과 양 저작물의 유사성 등의 간접사실이 인정되면 의거관계가 추정된다. 이러한 간접사실에는 양 저작물 사이에 의거하지 않았다면 나올 수 없는 현저한 유사성(striking similarity)도 포함된다. 통설, 판례는 미국법의 access라는 용어를 ‘접근’이라는 용어로 번역하여 사용하고 있지만, 원래 이는 침해자가 저작권자의 저작물에 가까이 접근하였다는 것으로는 부족하고 실제로 감상할 수 있는 상태에 있었다는 것을 의미한다. 즉 침해자가 저작권자의 작품을 보거나 복제할 수 있는 합리적인 기회, 가능성이

있는 경우를 말하므로 이에 미치지 못하는 가능성이나 의심만으로는 의거관계가 추정될 수 없다.²¹⁾

기계학습 방법론에 따르면 개발자 A가 사용한 학습데이터에 화가 C의 선행 작품이 포함되었는지에 따라 침해 인정 여부가 달라진다. 먼저 개발자 A가 사용한 학습데이터에 화가 C의 작품이 전혀 없었다면 개발자 A 또는 이용자 B의 행위는 의거관계가 인정되지 않아 저작권침해로 볼 수 없다. 개발자 A가 사용한 학습데이터에 화가 C의 작품이 있는 경우에는 학습데이터로 제공된 화가 C의 작품 수, 채택된 인공지능 방법론에 따라 의거관계의 존부가 달라진다고 생각된다.

예컨대 GAN(Generative Adversarial Networks)은 모델 자체가 진짜와 거의 같은 가짜를 만드는 알고리즘을 채택하고 있으므로(저작권법적으로는 striking similarity), 만약 개발자 A가 화가 C의 작품들을 다량으로 수집하여 이들을 GAN의 학습데이터로 사용하였다면, GAN의 원리 및 제공된 학습데이터의 성격에 비추어 보아 인공지능 학습모델이 화가 C의 작품에 의거하여 화가 C의 그림과 유사한 그림을 생성하였다고 강하게 추정된다는 것이 필자의 견해이다.

영상처리에 주로 사용되는 CNN(convolutional neural network)의 경우라면, 인공지능이 인식하는 특징(feature)은 형식적, 수식적인 관계에서 도출되는 것이고 의미적인(semantic) 것과는 전혀 관계가 없으므로, 인공지능이 학습데이터로 제공된 그림에서 그 어떤 미감을 향수하였다고 보기 어렵다. 만약 학습데이터로서 여러 화가의 작품들이 랜덤하게 인공지능에게 제공되었다면, 필자의 견해로는 이들 그림 전부를 학습한 인공지능 학습모델은 특정 화가 C의 그림에 의거하여 새로운 그림을 생성하였다고 단정할 수 없다. 물론 학습데이터가 주로 또는 전부 화가 C의 작품들이었다면 인공지능 학습모델의 생성물이 화가 C의 작품에 의거하지 않았다고 보기도 어려운 상황이기도 하다.

인공신경망에 의한 자연어 처리에서 채택되고 있는 단어 임베딩 기술, Seq2Seq 기술 등은 단어 내지 문장의 특징들을 숫자벡터로 압축치환하여 단어 내지 문장 간의 관계적 유사성, 의미적 연관성을 파악하는 정도에 이르렀으므로 어떤 관점에서 보면 인공지능이 문장의 맥락을 파악하는 것으로도 볼 수 있다(‘맥락’을 어떻게 정의하고 이해하는가에도 의존하는 문제이다). 그러나 이 경우에도 인공지능이 원본 저작물 표현(학습데이터)의 개념적 의미(semantic)나 개성을 이해하고 복제하였다고 볼 수는 없기 때문에, 원칙적으로 인공지능이 생성한 문장들이 원본 저작물과 유사하더라도 의거관계를 쉽게 인정하기는 어렵다고 생각된다. 다만, 이러한 결론은 언어모델의 구조, 학습데이터의 규모와 내용에 따라 달라질 수 있다. 언어모델이 학습데이터 문장을 그대로 기억한 경우에는 동일한 형식의 문장에 대한 추론에서 높은 확률로 동일한 문장을 생성할 것이므로²²⁾ 의거성을 부정하기 어렵다. 최근 언어모델의 용량이 증가하면서 암기(memorization)의 위험성도 증

20) 인공지능 생성물 내지 산출물의 저작자에 대해, 이규홍 특허법원 부장판사는 일본의 가라오케 판결의 법리, 제조물책임법, 자동차손해배상보장법상의 운전자(운행 지배 + 운행이익)처럼 관리자를 의제하는 방안도 고려할 수 있다고 지적한다. 이규홍, '인공지능 관련 저작권 침해 및 대응'에 대한 토론문, 한국경영법률학회 2021년 춘계학술대회, 2021. 3. 19.

21) 오승중, [판례평석]저작권침해에 있어서의 의거성 판단과 일반 불법행위 성립 여부-서울고등법원 2012. 12.20. 선고 2012나17150판결, 대한변협신문, 2013. 8. 12.

22) 김병필, 이우타 사건을 통해 본 자연어 처리 인공지능의 법적 쟁점, 한국정보보호학회 사례연구회 발표문, 2021. 4.13. 68면

가하고 있는데, 한 연구에 의하면 GPT-2 언어모델이 생성한 문장 60만건 중 최소한 604건(0.1%)이 암기된 텍스트였다고 한다.²³⁾

외국의 견해를 살펴보면, 학습데이터에 피침해 저작물이 존재하는 경우 인공지능이 피침해 저작물에 접근(access)하였다는 사실이 입증된다고 보는 견해가 있다.²⁴⁾ 이러한 주장은 원론적으로 타당할 수 있으나, 전술한 바와 같이 학습데이터의 종류, 수량, 기계학습 방법론의 구조 등을 전반적으로 검토해야 의거관계의 판단이 가능하다고 생각한다.

음악저작물 침해 사건의 경우, X음악이 Y음악을 표절하였다는 가능성이 제기되었으나 실제로는 X음악과 Y음악 모두 해당 분야의 원조격인 Z음악에 강한 영향을 받고 있어서 X음악과 Y음악이 유사하게 들리는 것으로 결론이 내리는 사례가 많다. 이러한 경우에는 의거관계가 부인된다고 할 것인데, 대량의 학습데이터를 습득한 인공지능 모델이 생성한 음악에도 발생할 수 있는 사례라고 생각된다.

소송실무의 관점에 보면, 인공지능 모델에 의해 저작권 침해가 발생한 경우 학습데이터 및 학습모델 알고리즘(구조)에 대한 증거조사방법이 쉽지 않을 것으로 예상된다. 먼저 의거관계를 판단하기 위해서는 학습데이터에 피침해 저작물이 있는지 살펴보아야 하나, 학습이 완료되어 학습모델이 정립된 인공지능(소프트웨어 형태로 존재할 수 있다) 자체는 학습데이터를 저장하고 있지 않기 때문에 이용자가 보유한 인공지능 프로그램에 대해 증거조사를 하더라도 원본적인 피침해 저작물이 학습데이터에 존재했었는지를 확인할 수 없을 것이다. 학습데이터는 학습모델과 분리되어 원칙적으로는 인공지능 개발자가 보유할 것으로 생각된다. TensorFlow와 같은 인공지능 학습용 플랫폼에 학습데이터에 관한 로그기록 등이 있어야 의거성 관련 증거조사가 용이할 것으로 보인다.

결론적으로, 종래 저작권법 이론은 1:1 관계에서 의거성 판단을 하였지만(n:1 관계인 경우에도 실제로는 병렬적인 1:1 관계였다고 생각된다), 기계학습과 같이 대량의 학습데이터가 랜덤하게 제공되어 학습되는 경우에는 종래의 의거성 판단의 기준과 방법이 그대로 적용되기는 어렵다고 본다. 특정 저작자의 작품만을 한정하여 학습하거나 학습방법론이 특정 작품을 모방하도록 의도된 경우를 제외하고는, 블랙박스 같은 인공지능 알고리즘의 특성을 고려할 때 침해 작품이 피침해 작품에 의거하여 작성되었다는 점을 쉽게 입증하기는 어려울 것으로 판단된다.

(4) Edmond de Belamy 초상화 사례

1) 사실관계

2018년 10월 25일 GAN(generative adversarial network)이라는 기계학습 모델에 의해 생성된 Edmond de Belamy라는 가상의 사람에 대한 초상화가 크리스티 경매장에서 \$432,500에 낙찰되었다. 최초 크리스티 경매장이 예상한 금액은 대략 \$10,000 정도였다. 이 경매는 크리스티측에 의해 인공지능에 창작된 초상화 작품에 대한 최초의 경매로 홍보되면서 인공지능이 창작한 작품에 대한 저작자가 누구인지에 관하여 큰 의문을 던졌다.

23) Carlini 외 11인, Extracting Training Data from Large Language Models, 2020, 13면

24) Benjamin Sobel, 앞의 논문, 66면

최초 위 초상화는 Obvious라고 하는 3인의 프랑스 대학생 그룹이 GAN 모델을 작동시켜 창작한 것으로 알려졌다. 그러나 그 후 알려진 사실관계는 이와 조금 다르다. Obvious 그룹이 GAN 모델을 이용하여 최종 그림을 출력한 것은 사실이다. 다만 위 초상화는 Obvious 단독의 결과물이 아니다. Ian Goodfellow가 GAN 아키텍처를 최초 설계했고 Alec Radford, Luke Metz, Soumith Chintala가 GAN을 개량한 DCGAN(Deep Convolution GAN)을 만들었으며, 결정적으로 당시 19세였던 Robbie Barrat이라는 미술품 기술자가 DCGAN 알고리즘을 기반으로 하여 르네상스 스타일로 그림을 출력할 수 있도록 미세조정(fine tuning)을 한 art-DCGAN이라는 이름의 프로그램을 코딩하고 이를 GitHub에 오픈소스(BSD 라이선스)형태로 공개한 바 있었다.²⁵⁾ Obvious는 Robbie Barrat의 위 오픈소스 프로그램을 이용하여 위 초상화를 제작한 것이다. 위 art-DCGAN은 이미 학습이 완료된 pre-trained model이나 이용자는 WIKIART 등에서 제공하는 풍경, 초상화 등 특정 장르의 학습데이터(그림)를 추가적으로 입력하여 학습시킬 수도 있다. Obvious는 르네상스 스타일의 초상화를 만들기 위해 art-DCGAN 프로그램 코드를 일부 변형하여 Edmond de Belamy 초상화를 출력한 것으로 알려져 있다(추가 학습까지 진행하였는지 불분명함).

위 초상화의 제작과정을 전체적으로 관찰할 때 가장 중요한 기여를 한 자는 art-DCGAN이라는 프로그램을 만든 Robbie Barrat이다. GAN의 설계자인 Ian Goodfellow, DCGAN의 설계자인 Alec Radford 등은 일종의 아이디어를 제공한 자이고(르네상스 풍의 초상화 작성을 의욕한 것도 아니다), 구체적으로 Edmond de Belamy 초상화와 같은 유형의 작품을 출력할 수 있도록 한 것은 Robbie Barrat이라고 할 수 있다. Obvious도 많은 시간과 노력을 들여 art-DCGAN 프로그램 코드를 조작하고 종국적으로 르네상스 초상화 스타일의 작품을 출력한 공로가 있으나 이러한 조작은 통상의 화가 내지 기술자라면 용이하게 실시할 수 있는 작업이어서 AI 미술 커뮤니티에서는 Robbie Barrat이 위 초상화 제작에서 가장 큰 기여를 한 것으로 보고 있다(뉴질랜드의 AI 화가인 Tom White는 Barrat의 코드를 그대로 사용하여 Belamy 초상화와 매우 유사한 그림을 출력하는 것에 성공하기도 하였다). 그러나 문제는 Obvious를 제외한 나머지 기여자가 위 경매의 수혜를 전혀 받지 못하였다는 점에 있다. 경매대금 뿐만 아니라 홍보에 있어서도 Robbie Barrat 등의 기여는 최초에 알려지지 않았다.

위 경매 후 Robbie Barrat은 art-DCGAN 프로그램에 적용되는 BSD 라이선스를 변형하여, “사전학습된 모델(pre-trained model)의 결과물은 판매되거나 상업적으로 이용될 수 없다”는 라이선스 조건을 부가하였다.²⁶⁾ 이는 Obvious의 경매출품을 의식한 행위라고 생각된다.

2) 검토되어야 할 점

정책적 내지 법리적으로, 위 경매에서 보상받아야 할 주체가 누구인지 문제된다. Robbie Barrat은 특정 유형의 그림들을 사전에 학습시켜 해당 유형의 그림이 출력될 수

25) BSD(Berkeley Software Distribution) 라이선스는 소스코드 공개의무가 없고 수정, 배포, 상업적 이용이 가능한 permissive 계열의 오픈소스 라이선스이다.

26) “EXTRA: NO OUTPUTS OF THE PRE-TRAINED MODELS MAY BE SOLD OR USED FOR-PROFIT OTHERWISE.”

있는 소프트웨어를 제작하였다. Obvious는 art-DCGAN 모델의 프로그램 코드를 일부 수정하였는데 이 과정에서 하이퍼 파라미터의 조정 등의 기여가 있었을 것으로 추측된다.

Robbie Barrat은 오픈소스 라이선스를 변경하여 사전학습된 모델이 산출한 결과물의 판매를 금지하였다. 이러한 조건의 부가가 적법한 것인가? 해당 BSD 라이선스는 Robbie Barrat이 GitHub에 제공한 프로그램에 대해 적용되는 것인지 그 프로그램의 결과물에 대해 적용되는 것은 아니라고 보아야 한다. 이는 마치 워드 프로세서 프로그램의 저작권자가 이용자에게 워드 프로세서에서 산출된 문서에 대해 이용조건을 부과하는 것과 유사한 측면이 있다.

2. 추가적 쟁점 — 오픈소스 라이선스에서 판매금지 조건 부가의 법적 효력

(1) 문제점

오픈소스 라이선스가 계약(contract)인지 아니면 단독행위 성격의 라이선스(bare license)인지 논란이 있다. 계약으로 성립하기 위해서는 라이선서의 청약(offer)과 라이선서의 승낙(acceptance)이 있어야 하고 영미법의 경우에는 더 나아가 약인(consideration)이 존재해야 한다. Edmond de Belamy 사례에서 Robbie Barrat이 설정한 결과물(output) 판매금지 조건이 법적으로 유효한지, 그러한 조건에 위반하여 결과물이 양도된 경우 양수인이 적법하게 결과물을 소유할 수 있는 것인지 여부 등을 판단하기 위해서는 위 문제를 먼저 검토하여야 한다.

(2) Contract와 Bare license의 구분

라이선스는 일방적일 수 있고 상호 합의를 요구하지 아니하므로 반드시 계약일 필요가 없다. 계약(contract)이 아닌 라이선스를 bare license라고 하며, 이러한 법적 개념은 보통법(common law) 시스템의 토지법(land law)에서 연원한다.²⁷⁾ Bare license의 의미는 라이선서가 라이선시에게 저작물에 대한 이용허락을 하고 저작권침해주장을 하지 않겠다는 약속을 하는 것에 그치고 라이선서가 라이선시에게 그 이상의 독점배타적 권리(제3자에 대한 방해배제청구권 등)를 부여하는 것은 아니다. Bare license의 경우 라이선서에게 특별한 의무가 없기 때문에 라이선서는 원칙적으로 언제든지 이용허락(license)을 철회할 수 있다.²⁸⁾ 오픈소스 라이선스의 대표적이라고 할 수 있는 GPL의 예를 들면, 라이선서가 GPL의 청약내용(offer)을 승낙(acceptance)하지 아니하면 GPL은 계약으로서 성립할 수는 없으나 bare license로는 인정될 수 있다.²⁹⁾ 즉 라이선서가 GPL 조건에 동의하였다고 볼 수 없는 상황이라도 라이선서는 일단 bare license의 형식으로 라이선시에게 일정 조건하의 이용허락 및 부제소의 의사를 표시한 것으로 볼 수 있다. 다만 GPL 2.0은 프로그램을 수정, 배포할 권한이 본 라이선스에 의해서만 부여되고 본 라이선스에 동의하지 아니하는 경우에는 수정, 배포행위가 금지된다고 규정하고 있으므로,³⁰⁾ 라이선서가 라이선스 조건에 승낙하지 않은 상황에서 상정할 수 있는 bare license의 형태로서 실질적인 법적 의미는 거의 없다고 할 수 있다. 이 경우 라이선서는 이용허락의 의사표시를 철회하고 장래를 향하여 저작권에 기한 금지청구를 할 것이기 때문이다. 정리하자면, 라이선스 특히 라이선서가 저장소, 웹사이트 등

에서 일방적으로 저작물의 이용허락을 선언하는 오픈소스 라이선스, CC라이선스(Creative Commons License) 등은 라이선서의 의사, 행위에 따라 계약이 아닌 bare license로 평가될 수 있다는 점에 유의할 필요가 있다.

(3) 오픈소스 라이선스에 청약(offer)과 승낙(acceptance)이 존재하는지 여부

라이선서가 라이선시에게 적극적으로 접근, 이용, 개작, 재배포를 허락하는 행위가 청약이라고 볼 수 있다. GitHub 등의 접근할 수 있는 저장소, 웹사이트에 소스코드 제공의사를 밝히는 것이 청약사실의 존재를 보여주는 것이다. 승낙의 요건은 청약의 그것에 비해 엄격하다. 승낙은 절대적(absolute)이어야 하고 승낙사실에 관하여 어떠한 의문도 없어야 한다(no doubt as to the fact of acceptance). 소프트웨어 배포방식인 shrink-wrap, clickwrap, browserwrap에 관한 몇 개의 판례가 있는데, ProCD v. Zeidenberg 판결에서³¹⁾ 법적 유효성이 인정된 shrinkwrap은 오픈소스 배포에서는 사용되지 않으므로 고려대상이 아니고, clickwrap, browserwrap 방식으로 배포된 오픈소스에서 라이선서와 라이선시 간에 계약이 체결되었는지 주로 문제된다. 대체로 clickwrap 라이선스, browserwrap 라이선스의 경우 이용자는 승낙 전에 라이선스 조항을 읽을 수 있고 이에 적절히 대응할 수 있는 기회를 가지고 있으므로 shrinkwrap 라이선스에 비해 승낙의 법적 유효성을 인정하기에 매우 용이하다. 반면, 전술한 바와 같이, 라이선서의 승낙이 존재한다고 볼 수 없는 상황이라면 해당 라이선스는 bare license라고 보아야 한다.

(4) 약인(consideration)의 존재 여부

이 쟁점은 civil law 시스템인 대륙법계와 common law 시스템인 영미법계가 극명하게 차이를 보이는 지점이다. 약인을 강하게 이해하는 견해는 약인에 경제적 가치(economic value)가 있어야 한다고 보거나 약인은 권리와 같은 것(right kind)으로 본다. 따라서 비금전적인 약속(버그 보고, 패치 테스트 참여 등)은 약인으로 평가되기에 부족할 수 있다고 한다.

반면, 약인을 상호적인 약속(reciprocal promises) 정도로 이해하는 견해도 있다. 라이선시가 GPL에서 요구하는 고지사항을 이행하거나 코드를 재배포할 때 GPL 조항에 따를 것을 약속하는 것은 유효한 약인으로 평가된다고 본다. 이러한 견해에 따르면 GPL 라이선스는 법적으로 유효한 약인을 구비하고 있기 때문에 결론적으로 계약으로 볼 수 있다고 한다.

27) 계약적인 라이선스(contractual license)와 구분되는 bare license의 특징에 대해서는, Jonathan Hill, "The Termination of Bare Licences", The Cambridge Law Journal, Mar., 2001, Vol. 60, No. 1 (Mar., 2001), 90면 참조

28) Christopher M. Newman, "A License Is Not a "Contract Not To Sue": Disentangling Property and Contract in the Law of Copyright Licenses", Iowa Law Review, Vol. 98, No.3, March 2013, 1115면, 1119면

29) Lawrence Rosen, Open Source Licensing: Software Freedom and Intellectual Property Law, Prentice Hall PTR, 2004, 54면

30) GPL 2.0 제5조: "당신은 서명이나 날인을 하지 않았기 때문에 본 라이선스의 내용을 반드시 받아들여야 할 필요는 없다. 그러나 프로그램이나 그로부터 파생된 프로그램을 수정하거나 배포할 권한은 본 라이선스에 의해서만 주어진다. 본 라이선스에 동의하지 않을 경우에는 이런 행위들이 법률적으로 금지된다. 따라서 프로그램(혹은 프로그램에서 파생된 모든 저작물)을 수정하거나 배포하는 행위는 이와 관련된 본 라이선스의 내용에 동의한다는 것을 의미하며, 복제와 수정 및 배포에 관한 본 라이선스의 조건과 규정들을 모두 받아들여야겠다는 의미로 간주된다"; GPL 3.0 제9조도 유사한 내용이다.

31) ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996)

미국 *Jacobsen v. Katzer* 판결(CAFC 항소심)은 관련 artistic license가 계약으로서 약인을 가지고 있는지 여부에 관하여, 저작권 귀속(attribution)과 개작된 부분의 공개와 설명을 요구하는 의무사항은 금전으로 평가되기는 어렵지만 약인으로서 평가받기에 충분하다고 판시한 바 있다.³²⁾ 영국에서 오픈소스 라이선스의 계약으로서 유효성에 대한 법원 판단은 없는 것으로 보인다. 영국 판례법에 기초할 때 *Jacobsen* 사건의 artistic license가 계약으로서 유효할지 여부는 불분명하다고 한다. 적어도 일부 견해에 따르면 잉글랜드와 웨일즈 법원은 약인 결여를 이유로 위 artistic license를 계약이 아닌 단순 이용허락(bare license)으로 볼 것이라고 한다.³³⁾

참고로, 약인의 개념이 없는 대륙법계 독일의 경우, 독일 민법 제151조에 따라 GPL은 법적으로 유효하므로³⁴⁾ GPL의 요건이 라이선서와 라이선시 간의 계약내용을 구성한다고 본 하급심 판결이 있다.³⁵⁾

(5) 오픈소스 라이선스가 계약인지 아니면 단순 라이선스(bare license)인지 따라 달라지는 상황

오픈소스 라이선스가 계약으로 평가되는 경우 계약 당사자 간의 구속력이 발생하나, 단순 이용허락(bare license)으로 평가되는 경우에는 라이선서에게 금반언 원칙 외 특별한 구속력이 없다(즉 라이선서는 자유롭게 이용허락을 철회할 수 있다). 단순 이용허락인 경우에는 계약법의 적용을 받지 못하고 저작권법의 적용만을 받을 수 있다. 라이선스가 단순 이용허락으로 평가되는 경우에는 GPL이 채택하고 있는 여러 제한요건 예컨대 dynamic linking에 대해 GPL을 적용하는 것을 라이선서에게 요구할 수 없다는 견해도 존재한다. 이러한 제한요건의 위반을 언제나 저작권침해행위로 볼 수 없기 때문이다.

라이선시가 라이선스 위반행위를 한 경우 그 구제방법도 상이하다. 계약으로 평가되는 경우에는 라이선서는 영미법상 인정되는 특정이행(specific performance)을 구할 수 있고 그 외 채무불이행에 따른 손해배상청구, 계약해지를 할 수 있다. 단순 이용허락으로 평가되는 경우에는 라이선서는 이용허락을 철회한 후 저작권에 근거하여 장래를 향하여 사용금지청구를 할 수 있다.

(6) 추가적으로 설정된 판매금지 조건의 효력

우리나라 계약법은 영미법상 약인(consideration)의 개념을 알지 못하므로 오픈소스 라이선스를 계약으로 인정할 가능성이 영미법계의 경우보다 높다고 생각된다. 반면, 영미법계는 특별한 대가관계가 없는 오픈소스 라이선스를 약인의 결여를 이유로 계약으로 보지 않고 단독행위적으로 볼 가능성이 있다.

오픈소스 라이선스를 단독행위 성격의 bare license로 보면 *Robbie Barrat*이 BSD 라이선스에 추가한 이용조건은 동의를 하지 아니한 이용자에게 특별한 계약법적 효력을 가지지 못하므로 해당 이용조건에 위반하여 인공지능 창작물이 양도되어도 양도의 효력이 있다고 볼 것이나, 만약 오픈소스 라이선스를 계약으로 본다면 인공지능 학습모델의 이용자는 일단 해당 조건에 구속되어 이를 위반시 계약위반의 책임을 부담하게 된다. 본 사안의 경우, 판매금지 조건이 사후적으로 추가되었으므로 이러한 추가적 조건에 대해

학습모델의 이용자가 승낙을 하였다고 볼 수 있는 사정이 있는 시점부터 판매금지 조건의 계약법적 구속력이 발생된다고 할 수 있다. 다만 인공지능 창작물이 동산 등 물권에 해당한다면 양수인은 이를 적법히 취득할 수 있다고 생각된다. 물권은 채권에 우선하기 때문이다.

32) *Jacobsen v. Katzer*, 535 F.3d 1373, 1381 (Fed. Cir. 2008): "The choice to exact consideration in the form of compliance with the open source requirements of disclosure and explanation of changes, rather than as a dollar-denominated fee, is entitled to no less legal recognition."

33) McDonagh, Luke, "Copyright, Contract and FOSS", Free and Open Source Software, Oxford University Press, 2013, 32면.

34) 독일 민법 제151조: "계약은 승낙으로 성립되지만, 거래관행상 기대하기 어렵거나 청약자가 미리 포기한 경우에는 승낙의 의사표시가 청약자에게 도달할 필요가 없다. 청약의 효력 소멸 시점은 청약에 의해 정해지거나 제반 사정에서 도출되는 청약자의 의사로 정해진다"

35) *Welte v. D-Link Deutschland GmbH*, No 2-6 O 0224/06 (LG Frankfurt) (2006. 9.22.)