



독일 함부르크 지방법원, 인공지능(AI) 학습 데이터셋(Dataset) 생성을 위한 사진 이미지 복제의 저작권 침해 부정

정태호 | 경기대학교 지식재산학과 교수

- 이 사건의 개요
- 이 사건에서의 법적 분쟁의 주요 사실
- 대상 판결의 주요 내용
- 대상 판결의 검토
- 대상 판결의 시사점

1. 이 사건의 개요¹⁾

원고인 사진작가 Robert Kneschke는 피고인 비영리사단법인 LAION e.V.가 자신의 사진 중 하나를 이미지로 복제하는 것에 대하여 저작권 침해라고 주장하면서 침해행위 금지 등에 관한 소송을 제기하였다.

원고는 이미지 대행 제공 플랫폼인 Bigstock의 마이크로 스톡(Stock) 사진²⁾ 웹사이트를 통해 자신의 사진을 업로드하여 제공해 왔다. 원고가 사진을 이용할 수 있는 비독점적 권리를 라이선스로써 부여한 해당 플랫폼의 사진 웹사이트에는 “제한 사항: (...) 18. 자동화된 프로그램, 애플릿(Applets)³⁾, 봇(Bots)⁴⁾ 또는 이와 유사한 것을 사용하여com 웹사이트 또는 그 안의 콘텐츠에 접근하는 것은, 예를 들어 콘텐츠 다운로드, 인덱싱(Indexing)⁵⁾, 스크래핑(Scraping)⁶⁾ 또는 웹사이트의 콘텐츠 캐싱(Caching)⁷⁾을 포함하여 어떠한 목적이든 금지된다.”와 같이, 사진 이미지를 자동화 프로그램에 이용할 수 없다는 공지 내용이 이용자들이 읽을 수 있도록 이용약관에 기재되어 있었다.

1) Dr. Marcus von Welser, “Hamburg Regional Court Rules on Training Data for Artificial Intelligence”, Vossius & Partner News & Knowledge, 2024.10.16.에서
참조(<https://www.vossius.eu/en/news/detail/hamburg-regional-court-rules-on-training-data-for-artificial-intelligence>)(최종접속일 : 2024.11.4.).

2) 특정 용도로 라이선스가 부여되는 사진 공급임.

3) 작은 응용프로그램.

4) 사람의 의도를 대행해서 자동화할 수 있는 프로그램.

5) 저장되어 있는 데이터에 접근하기 위하여 색인을 사용하는 것.

6) 프로그램이 웹사이트에 접속하여 데이터를 자동으로 추출하는 것.

7) 이미지 등과 같은 콘텐츠를 사용자에게 더 빠르고 효율적으로 제공하기 위해 임시로 저장해두는 것.

피고는 인공지능(AI) 학습을 위해 58억 5천만 개의 이미지와 텍스트가 쌍으로 구성된 데이터셋(Dataset)을 제공하고 있었다. 이러한 데이터셋을 만들기 위해 피고는 미국의 비영리 조직인 Common Crawl⁸⁾의 기존의 데이터셋을 사용하였고, 인터넷상에서 Common Crawl 데이터셋의 이미지를 다운로드했다. 그 후에는 해당 이미지를 확인하여 Common Crawl 데이터셋의 이미지 콘텐츠 텍스트에 대한 설명이 이미지에서 보이는 콘텐츠와 일치하는지 확인했다. 여기서 텍스트 설명과 이미지 콘텐츠가 일치하지 않는 이미지는 필터링하여 삭제했고, 나머지 이미지의 경우에는 이미지 위치의 URL과 이미지 설명인 메타데이터를 추출하여 새로 생성한 데이터셋인 'LAION-5B'로 전송했다. 피고는 해당 데이터셋을 인터넷에서 공개적으로 접근 및 이용 가능한 이미지에 대하여 하이퍼링크가 있는 표의 형태로서 공중에게 무료로 제공했다.

피고는 원고의 분쟁 대상 사진의 이미지 파일⁹⁾을 이상과 같은 Bigstock이라는 이미지 대행 제공 플랫폼의 웹사이트에서 다운로드하여 분석했고, 그 후 추출된 메타데이터가 'LAION-5B'라는 해당 데이터셋으로 전송된 것이었다.

이에 따라 원고는 피고의 해당 데이터셋에서 자신의 사진 이미지를 삭제해달라고 요청했지만, 피고는 공중에게 제공하는 방식이 인터넷에서 공개적으로 사용할 수 있는 이미지 파일에 대한 링크가 포함된 데이터베이스만 유지한 것이라고 하면서 사진의 사본 보유를 부인했다.¹⁰⁾

이에 대하여 원고는 동의 없이 피고의 데이터셋에 있는 자신의 사진 이미지가 AI 학습데이터의 목적으로 사용된 것에 대해서 피고의 저작권 침해를 주장하면서 TDM(Text and Data Mining)¹¹⁾에서의 저작권에 대한 제한이 적용되지 않는다고 그 근거를 언급하였다.

2. 이 사건에서의 법적 분쟁의 주요 사실

여기서 이 사건의 법적 분쟁의 대상이 되는 주요 사실은 이미지 영역에서 생성 AI 모델을 학습시키는 데 사용되며 공개적으로 접근이 가능한 데이터셋인 'LAION 5B'를 만드는 맥락에서 원고의 사진 이미지를 사용한 것이었다. 그러나 해당 데이터셋에는 이미지가 포함되어 있지 않고 인터넷의 다른 곳에서 찾을 수 있는 이미지에 대한 하이퍼링크만 있었다는 특성도 보인다.¹²⁾

이러한 주요 사실을 근거로 하여 함부르크 지방법원은 이하와 같은 내용의 판결(이하,

8) 웹을 크롤링하고 데이터셋을 대중에게 무료로 제공하는 비영리 조직임.

9) 해당 이미지 제공 플랫폼의 워터마크가 찍혀 있는 이미지 파일임.

10) Matt Growcoot, "Court Rules Against Photographer Who Sued AI Dataset for Copyright Theft", PetaPixel, 2024.10.1.

(<https://petapixel.com/2024/10/01/court-rules-against-photographer-who-sued-ai-dataset-for-copyright-theft-germany-laion-robert-kneschke/>)(최종접속일 : 2024.11.4.).

11) 여기서의 TDM은 학습 데이터셋의 생성을 의미함.

12) Heidrich Rechtsanwälte, "Klares Urteil für LAION e.V.: Gericht bestätigt rechtmäßige Nutzung von Bildmaterial für KI-Training", Pressemeldung LG Hamburg Verfahren LAION. 2024.10.(<https://www.recht-im-internet.de/presseanfragen/pressemeldung-laion>)(최종접속일 : 2024.11.3.).

“대상 판결”이라 함)을 선고하였다.

3. 대상 판결의 주요 내용¹³⁾

(1) 원고의 주장

원고는 이상과 같은 분석 프로세스의 맥락에서 복제가 불법이라고 생각하고 피고가 사진을 복제하거나 AI 학습 데이터셋을 만드는 목적으로 복제하도록 하는 것은 원고의 저작권 침해라고 주장하였다.

(2) 피고의 주장

피고는 독일 저작권법(UrhG) 제44b조, 제60d조의 TDM에서의 예외에 근거하여 저작권 침해가 아님을 주장하였다.¹⁴⁾ 즉 비영리 기반으로 운영되고 과학 연구의 진흥에 전념하는 피고에게는 TDM 프로세스의 범위 안에서의 복제가 허용되며, 피고의 행위는 저작권법(특히 제60d조)에서의 저작권 침해의 예외에 해당한다고 피고는 주장하였다.¹⁵⁾

(3) 판결의 내용

① 일시적 복제 여부에 관한 판단

법원은 우선 적법한 이용을 통한 저작권의 제한에 관하여 제6절 제1관¹⁶⁾에 규정되어 있으며 일시적이고 부수적인 복제 행위를 허용하는 독일 저작권법 제44a조¹⁷⁾의 적용 가능성을 검토했다. 법원은 이미지-텍스트 쌍을 분석할 목적으로 복제가 구체적이고 영구적으로 수행되었기 때문에 해당 규정에서의 일시적 복제에 관한 저작권의 제한이 현재 이 사건에는 적용되지 않는다고 판단했다. 즉 저장은 일시적인 것이 아니라 의식적으로 프로그래밍되고 적극적으로 제어되었고, 또한 이미지를 다운로드하는 것은 단순히 분석에 수반되는 2차적 단계가 아니라 주로 분석 목적을 달성하는 독립적인 프로세스였다고 보고 법원

13) LG Hamburg Urteil vom 27.09.2024 – 310 O 227/23.

14) 이러한 조항은 후술하지만 DSM 지침(디지털단일시장의 저작권 및 관련 권리들에 관한 EU 지침 2019/790)의 제3조 및 제4조를 독일법에 구현한 것임.

15) Heidrich Rechtsanwälte, 앞의 글.

16) Abschnitt 6 Schranken des Urheberrechts durch gesetzlich erlaubte Nutzungen(제6절 적법한 이용을 통한 저작권의 제한) Unterabschnitt 1 Gesetzlich erlaubte Nutzungen(제1관 적법한 이용).

17) § 44a Vorübergehende Vervielfältigungshandlungen(일시적 복제행위)

Zulässig sind vorübergehende Vervielfältigungshandlungen, die flüchtig oder begleitend sind und einen integralen und wesentlichen Teil eines technischen Verfahrens darstellen und deren alleiniger Zweck es ist (일시적 복제행위는 순간적으로 또는 수반되어 행하여지고, 기술적 과정에서 필수적이고 본질적인 부분을 나타내며, 그 유일한 목적이,)

1. eine Übertragung in einem Netz zwischen Dritten durch einen Vermittler oder(네트워크상에서 중개자를 통한 제3자들 간의 전송 또는)

2. eine rechtmäßige Nutzung(합법적 이용일 때)

eines Werkes oder sonstigen Schutzgegenstands zu ermöglichen, und die keine eigenständige wirtschaftliche Bedeutung haben(저작물 또는 그 밖의 보호대상에 대한 위와 같은 행위들을 가능하게 하고, 독자적으로 경제적 의미를 가지지 아니하는 경우에 허용된다.).

은 독일 저작권법 제44a조의 적용을 부정하였다.¹⁸⁾

② 통상적(상업적) 목적의 TDM 및 기계 판독 가능한 이용 제한 문구에 관한 해석

독일 저작권법 제44b조¹⁹⁾의 적용 가능성과 관련하여 법원은 이 사건에서의 데이터셋을 만들기 위한 분석은 이미지 콘텐츠와 텍스트 설명 간의 일치성 등에 관한 상관관계를 얻는 것을 목표로 하므로, 여기서의 복제가 기본적으로 독일 저작권법 제44b조 제1항²⁰⁾에 따른 TDM에 해당한다고 판단했다. 그리고 독일 저작권법 제44b조 제2항²¹⁾의 통상적(상업적) 목적의 TDM에 관한 저작권 제한 규정이 충족된 것으로 보아 피고의 복제 행위에는 문제가 없다고 보았다. 그러나 법원은 저작권법 제44b조 제3항²²⁾에 따라 사진 이미지 대행 제공 플랫폼의 이용 제한에 관한 이용유보의 가능성이 있기 때문에 이러한 제한 규정의 적용에 의문을 표명했다. 즉 해당 사진 이미지 대행 제공 플랫폼의 웹사이트에 작성된 약관에서는 웹스크래핑을 비롯한 자동화된 접근을 금지했다. 그런데 이러한 이용 제한에 관한 이용유보가 저작권법 제44b조 제3항에 따른 유효한 이용유보로 간주 될 만큼 충분히 구체적이고 기계 판독이 가능한 방식으로 작성되었는지에 대해서는 의문점이 제기되기도 하였으나,²³⁾ 법원은 일단 이용유보에 관한 법적 요구사항의 준수가 명확하게 구체적으로 선언되었다고 보았으며, ‘자연어’로 이러한 이용유보를 정확히 이해할 수 있는 기술적 능력이 현재 충분히 있다고 보이므로, ‘기계어’가 아니라 ‘자연어’로 된 이러한 이용유보의 문구도 기계가 판독이 가능한 것으로 해석할 수 있다고 판단하였다.

③ 과학적인 연구 목적의 인정

궁극적으로 법원은 과학적인 연구 목적으로 TDM을 허용하는 독일 저작권법 제60d조²⁴⁾

18) Dr. Martin Meggle-Freund, “LG Hamburg, 310 O 227/23”, ipweblog, 2024.10.29.
(<https://ipweblog.de/2024/10/lg-hamburg-310-o-227-23/>)(최종접속일 : 2024.11.3.).

19) § 44b Text und Data Mining(TDM)

(1) Text und Data Mining ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen(텍스트 및 데이터 마이닝은 특히 패턴, 추세 및 상관관계에 대한 정보를 획득하기 위해 하나 이상의 디지털 또는 디지털화된 저작물을 자동으로 분석하는 것이다.).

(2) Zulässig sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Die Vervielfältigungen sind zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind(텍스트 및 데이터 마이닝을 위해 합법적으로 접근할 수 있는 저작물의 복제가 허용된다. 복제품은 텍스트 및 데이터 마이닝에 더 이상 필요하지 않으면 삭제된다.).

(3) Nutzungen nach Absatz 2 Satz 1 sind nur zulässig, wenn der Rechtsinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online zugänglichen Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt(제2항 제1문에 따른 이용은 권리보유자가 유보하지 않은 경우에만 허용된다. 온라인으로 접근할 수 있는 저작물에 대한 이용유보는 기계 판독이 가능한 형태인 경우에만 유효하다.).

20) TDM은 패턴, 추세 및 상관관계에 대한 정보를 획득하기 위해 하나 이상의 디지털 또는 디지털화된 저작물을 자동으로 분석하는 것이라는 내용임.

21) TDM을 위해 합법적으로 접근할 수 있는 저작물의 복제가 허용되며, 복제품은 TDM에 더 이상 필요하지 않으면 삭제된다는 내용임.

22) TDM에 의한 저작물의 복제 이용은 저작권자가 그 이용허락을 유보하지 않은 경우에만 허용되는데, 온라인으로 접근할 수 있는 저작물에 대한 이용유보는 기계 판독이 가능한 형태인 경우에만 유효하다는 내용임.

23) Dr. Martin Meggle-Freund, 앞의 글.

24) § 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung(과학적인 연구 목적의 TDM)

(1) Vervielfältigungen für Text und Data Mining (§ 44b Absatz 1 und 2 Satz 1) sind für Zwecke der wissenschaftlichen Forschung nach Maßgabe der nachfolgenden Bestimmungen zulässig(텍스트 및 데이터 마이닝을 위한 복제(제44b조 제1항과 제2항 제1문)는 다음 조항에 따라 과학적인 연구 목적으로 허용된다.).

의 제한이 이 사건에 적용되므로, 앞서 다룬 동법 제44b조에 대한 최종적인 검토가 불필요하다고 결정했다. 피고는 과학적인 연구 목적을 가진 비영리법인으로 활동했으며 어떠한 상업적 목적도 추구하지 않았다. 이 사건 관련 데이터셋이 연구용으로서 무료로 제공되었기 때문에 법원은 앞서 검토한 동법 제44b조의 규정에 관계없이 복제 행위를 다루는 독일 저작권법 제60d조에서 요구하는 사항을 최종적인 판단의 근거로서 확인했다.

결국 피고인 LAION e.V.의 이 사건에서의 분쟁 대상이 된 데이터셋을 과학적인 연구 목적으로 인정하여, 저작권법상 TDM에 관한 예외 규정의 적용을 확인했고, 이에 따라 최종적으로 법원은 AI 학습 데이터셋 생성을 위한 이미지 복제가 독일 저작권법 제60조(d)에 따른 TDM에 해당한다고 보아 저작권 침해를 부정하는 판단을 하였다.

4. 대상 판결의 검토

이상과 같이 함부르크 지방법원은 피고인 LAION e.V.가 원고인 Robert Kneschke의

(2) Zu Vervielfältigungen berechtigt sind Forschungsorganisationen. Forschungsorganisationen sind Hochschulen, Forschungsinstitute oder sonstige Einrichtungen, die wissenschaftliche Forschung betreiben, sofern sie(연구기관은 복제할 권리가 있다. 다음 중 어느 하나에 해당하는 경우 과학적인 연구를 수행하는 대학, 연구소 또는 그 밖의 기관은 연구기관에 해당한다.)

1. nicht kommerzielle Zwecke verfolgen(비상업적인 목적을 추구하는 경우),

2. sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder(모든 수익을 과학적인 연구에 재투자하는 경우)

3. im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind(국가 공인 계약의 범위 내에서 공익을 위해 활동하는 경우).

Nicht nach Satz 1 berechtigt sind Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat(연구기관에 특정한 영향을 미치고 과학적인 연구 결과에 우선적으로 접근할 수 있는 민간 기업과 협력하는 연구기관은 제1문에 따른 권리가 없다.).

(3) Zu Vervielfältigungen berechtigt sind ferner(또한 다음의 경우 복제할 권리가 있다.)

1. Bibliotheken und Museen, sofern sie öffentlich zugänglich sind, sowie Archive und Einrichtungen im Bereich des Film- oder Tonerbes (Kulturerbe-Einrichtungen)(일반 대중의 접근이 허용된 도서관과 박물관, 기록보존소 및 시청각유산 부문의 시설(문화유산기관)),

2. einzelne Forscher, sofern sie nicht kommerzielle Zwecke verfolgen(상업적 목적을 추구하지 않는 개별 연구원).

(4) Berechtigte nach den Absätzen 2 und 3, die nicht kommerzielle Zwecke verfolgen, dürfen Vervielfältigungen nach Absatz 1 folgenden Personen öffentlich zugänglich machen(비상업적 목적을 추구하는 제2항과 제3항에 따라 권리를 부여받은 사람은 다음의 사람에게 제1항에 따른 복제에 공개적으로 접근할 수 있도록 할 수 있다.):

1. einem bestimmt abgegrenzten Kreis von Personen für deren gemeinsame wissenschaftliche Forschung sowie(공동의 과학적인 연구를 위해 특정적으로 분류된 사람들의 집단)

2. einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung(과학적인 연구의 질을 검증하기 위한 개별 제3자).

Sobald die gemeinsame wissenschaftliche Forschung oder die Überprüfung der Qualität wissenschaftlicher Forschung abgeschlossen ist, ist die öffentliche Zugänglichmachung zu beenden(공동의 과학적인 연구 또는 과학적인 연구의 품질 검증이 완료되는 즉시 일반 대중에게 접근이 가능하도록 제공하는 것은 종료되어야 한다).

(5) Berechtigte nach den Absätzen 2 und 3 Nummer 1 dürfen Vervielfältigungen nach Absatz 1 mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung aufbewahren, solange sie für Zwecke der wissenschaftlichen Forschung oder zur Überprüfung wissenschaftlicher Erkenntnisse erforderlich sind(제2항과 제3항 제1호에 따라 권한을 부여받은 사람은 과학적인 연구 또는 과학적인 발견의 검증을 위해 필요한 경우 무단 사용을 방지하는 적절한 보안 예방 조치와 함께 제1항에 따라 복제물을 보관할 수 있다.).

(6) Rechtsinhaber sind befugt, erforderliche Maßnahmen zu ergreifen, um zu verhindern, dass die Sicherheit und Integrität ihrer Netze und Datenbanken durch Vervielfältigungen nach Absatz 1 gefährdet werden(권리보유자는 제1항에 따라 복제물에 의해 네트워크와 데이터베이스의 보안 및 무결성이 저해되는 것을 방지하도록 필요한 조치를 취할 권한이 있다.).

사진 이미지를 데이터셋 생성에 이용한 것이 독일 저작권법 제60d조에 따른 저작권 침해 예외의 적용이 인정된다고 판결했다.

따라서 대상 판결을 통해 피고는 독일 저작권법 제60d조에 따른 저작권 침해의 예외 혜택을 받았다고 볼 수 있다. 이에 관하여 피고가 학습 데이터셋을 무료로 공개하기 때문에 해당 데이터셋의 생성을 비상업적이고 과학적인 연구를 위한 것으로서 대상 판결이 판단한 것이라고 볼 수 있다.²⁵⁾

결국 이상과 같은 대상 판결을 통해 다음과 같은 내용들을 확인할 수 있다.²⁶⁾ 첫째, 일시적이거나 부수적이지 않은 복제 행위는 독일 저작권법 제44a조의 제한 사항에 포함되지 않는다. 따라서 이 사건처럼 분석 목적으로 저작권이 있는 콘텐츠를 대상으로 프로그래밍하여 저장하고 처리하는 것은 동법 제44a조의 의미 내에서의 일시적이거나 부수적인 이용이 아니라고 보아야 할 것이다. 둘째, 독일 저작권법 제44b조 제2항의 제한 규정은 TDM을 위한 디지털 저작물의 복제를 허용한다. 다만, 이것은 동법 제44b조 제1항에 따라 패턴, 추세 및 상관관계에 대한 정보를 얻기 위한 자동 분석 목적에 부합하는 경우에만 한한다. 즉 이미지-텍스트 관계와 같은 작업 구성요소 간의 상관관계를 인식하는 것은 저작권의 제한을 받지 않고 복제 등을 할 수 있다는 일종의 특혜를 부여하는 적격 요건이 된다고 볼 수 있다. 셋째, 동법 제44b조의 복제가 허용되는 '상관관계'에는 저작물의 지적인 콘텐츠를 반드시 이용할 의도 없이 저작물의 구성요소 간의 상관관계를 획득하는 것도 포함된다고 볼 수 있다. 넷째, 이상과 같은 TDM 관련 저작권의 제한 규정의 적용을 배제하려면, 동법 제44b조 제3항의 의미 내에서 이용제한(이용유보)에 관한 약관을 기계가 판독이 가능한 방식으로 명시적인 선언을 해야 한다. 이러한 약관의 문구는 TDM에 대한 제한 사항으로 자동으로 명확하게 인식될 수 있도록 구체적으로 표현되어야 하는데, '기계어'로 표현되는 것 뿐만 아니라, '자연어'로 표현되더라도 기계가 판독하여 이해 가능한 기술적 상황이라면 이러한 약관 문구의 효력을 인정할 수 있다는 것을 확인할 수 있다. 즉 지금까지 저작권법 제44b조에서 요구하는 기계가 판독이 가능한 이용약관에 '자연어' 선언이 충분한지 여부에 대해 논란이 있었는데, 함부르크 지방법원이 대상 판결을 통해 이에 대하여 정리하는 판단을 하였다고 볼 수 있다. 결국 대상 판결은 AI 애플리케이션이 이미 자연어를 이해하고 처리할 수 있으므로, 여기에는 자연어로 작성된 약관을 해석할 수 있는 AI 시스템도 포함된다는 점을 고려하여 이와 같은 판단을 하였다고 볼 수 있다.²⁷⁾ 마지막으로 다섯째, 독일 저작권법 제60d조의 제한 규정은 TDM을 위해 저작권으로 보호되는 저작물의 복제를 허용하는데, 다만, 이러한 작업은 비영리 연구기관에 의해 과학적인 목적으로 수행되고 상업적 목적은 추구되지 않아야 한다는 것을 확인할 수 있

25) Norton Rose Fullbright, "Infringement risk relating to training a generative AI system", Global Publication, (2024.7.)(<https://www.nortonrosefulbright.com/en/knowledge/publications/ef8d8cce/infringement-risk-relating-to-training-a-generative-ai-system>)(최종접속일 : 2024.11.4.).

26) Dr. Martin Meggle-Freund, 앞의 글 참조.

27) Dr. Saskia Ostendorff, "Foto-graf muss Nutzung seines Bildes in KI-Daten-bank dulden", Wolters Kluwer Legal Tribune Online, 2024.10.15.(<https://www.lto.de/recht/hintergruende/h/kuenstliche-intelligenz-ki-urheberrecht-text-data-mining-lg-hamburg-310o22723>)(최종접속일 : 2024.11.3.).

다. 그리고 이에 따라 만들어진 데이터셋을 연구 목적으로 대중에게 공개한다고 해서 해당 규정의 이러한 제한의 적용이 배제되는 것은 아니라는 점도 아울러 확인할 수 있다.

한편으로 대상 판결에서의 판단 과정에서 기존에 독일 저작권법 제44b조 및 동법 제60d조의 제한 사항을 만들 때 AI를 고려하지 않았기 때문에 TDM이 AI 학습 데이터셋 생성에는 적용되지 않는다는 반론이 있기도 하였다. 그런데 이러한 반론에 대해서는 유럽 인공지능법(AI ACT) 제53조 제1항 (c)²⁸⁾가 DSM(Digital Single Market) 지침²⁹⁾ 제4조³⁰⁾를 참조 및 언급하면서 인공지능 학습을 위한 데이터셋의 생성도 TDM에 대한 저작권의 제한(예외)에 속한다는 점을 강조한 것을 근거로 하여 법원에서 대상 판결의 내용과 같이 판단을 한 것이라고 볼 수 있다. 따라서 이 사건에서의 AI 학습 데이터셋의 생성에도 TDM에 대한 저작권의 제한 규정이 적용될 수 있다는 설득력이 있는 논리를 제공해 준 점도 대상 판결의 의의로 볼 수 있다.³¹⁾

5. 대상 판결의 시사점

대상 판결은 함부르크 지방법원은 하급심이기는 하지만, AI 학습 데이터셋에 대해서 독일 저작권법에 구현된 DSM 지침 제3조³²⁾ 및 제4조에서 규정된 저작권 침해에 대한 TDM 예외의 적용에 관하여 유럽 최초의 판결을 선고했다는 데 그 의의가 있다.³³⁾ 즉 대상 판결은 AI 학습 데이터셋의 생성과 관련하여 피고의 접근 방식의 합법성을 확인하고 저작권 침해를 주장하였던 원고의 주장을 기각했다.³⁴⁾

28) Article 53 Obligations for providers of general-purpose AI models(범용인공지능모델 제공자의 의무) 1. Providers of general-purpose AI models shall(범용인공지능모델 제공자는 다음 사항에 따른다.): (c) put in place a policy to comply with Union copyright law, and in particular to identify and comply with, including through state of the art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790 (유럽연합 저작권법을 준수하고 특히 최신 기술을 통한 것을 포함하여 지침(EU) 제2019/790호 제4조(3)에 따라 표시된 권리유보를 확인하고 준수하기 위한 정책을 시행한다.);

29) 디지털단일시장의 저작권 및 관련 권리들에 관한 지침(Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC).

30) Article 4 Exception or limitation for text and data mining(TDM에 관한 예외 또는 제한) 1. Member States shall provide for an exception or limitation to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, Article 4(1)(a) and (b) of Directive 2009/24/EC and Article 15(1) of this Directive for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining(회원국은 텍스트 및 데이터 마이닝을 목적으로 합법적으로 접근할 수 있는 저작물 및 기타 주제의 복제 및 추출을 위해 지침 96/9/EC, 지침 2001/29/EC 제2조, 지침 2009/24/EC 제4조 (1) (a) 및 (b), 지침 15조 (1)에 규정된 권리에 대한 예외 또는 제한을 규정해야 한다.). 2. Reproductions and extractions made pursuant to paragraph 1 may be retained for as long as is necessary for the purposes of text and data mining(제1항에 따라 이루어진 복제 및 추출은 텍스트 및 데이터 마이닝의 목적상 필요한 기간 동안 보관할 수 있다). 3. The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online(제1항에 규정된 예외 또는 제한은 온라인에 공개되는 콘텐츠의 경우 저작물 및 기타 주제의 이용이 권리자가 적절한 방식으로 명시적으로 유보하지 않은 경우에 적용된다). 4. This Article shall not affect the application of Article 3 of this Directive(이 조항은 이 지침 제3조의 적용에 영향을 미치지 않는다.).

31) Dr. Saskia Ostendorff, 앞의 글.

32) Article 3 Text and data mining for the purposes of scientific research(과학적 연구 목적의 TDM)로서 이 사건에서의 독일 저작권법 제60d조로 구현됨.

33) Dr. Saskia Ostendorff, 앞의 글.

함부르크 지방법원은 이 사건의 저작권이 있는 사진 이미지의 복제가 불법적인 복제에 해당한다는 점을 발견하지 못했다. 즉 해당 법원은 피고인의 행위가 과학적 연구 목적의 TDM에 관한 독일 저작권법 제60d조의 저작권의 제한 사유에 따른 것이라고 판단했다. 해당 법원에 따르면, 학습데이터를 추출하기 위한 목적으로 이미지를 복제하는 것은 법적인 책임에 대한 면제 조항의 의미 내에서 TDM이라고 보았는데, 이에 따라 이러한 대상 판결은 그 이후의 생성형 AI의 후속 학습도 이러한 예외에 포함되는지 여부에 관한 검토의 여지를 열어 두었다는 점에서도 그 의의를 찾을 수 있겠다.³⁵⁾

한편으로 대상 판결은 과학적 연구의 맥락에서 공개적으로 접근가능한 데이터를 합법적으로 사용할 수 있는 중요한 기반을 마련했으며, 앞으로도 오픈소스 이니셔티브(Open Source Initiative)를 촉진하는 데 상당한 기여를 할 수 있음을 확인시켜 주면서, 특히 독일에서 AI 개발을 촉진하는 법적 기반이 될 수 있을 것이라고 긍정적으로 대상 판결의 의의를 보는 견해도 언급되고 있다.³⁶⁾

분명히 대상 판결은 관련 분야에서의 역사적 관점에서 중요한 사례이고, 이것은 DSM 지침에 포함된 TDM 예외에 대한 첫 번째 법적인 테스트였으며, 이에 대해서는 아직 최종적인 결론이 내려진 것이 아니라고 보아야 할 것이다. 그런데 대상 판결의 그 법적인 영향력 자체는 매우 제한적일 수 있으며 다른 국가의 법원, 심지어 독일의 다른 법원들도 하급심 판결인 대상 판결과는 다른 결론을 내릴 수도 있을 것이다.

한편으로 대상 판결은 데이터셋의 생성을 위한 TDM의 이용만 다루고 AI의 실제적인 학습 과정에 대해서는 구체적으로 다루고 있지 않아 상대적으로 좁은 범위에서 이루어진 판단이라는 점이 강조되기도 한다. 이에 따라 현재로서는 AI 학습 데이터셋의 제작자가 계속될 AI 관련 저작권 분쟁에서 하나의 작은 승리만을 거두었다고 볼 수도 있다. 따라서 이에 더 나아가서 AI 모델 학습 자체에 대한 유럽연합 내의 소송의 제기는 이 사건의 다음 단계가 될 것이라고 기대되기도 한다.³⁷⁾

그런데 대상 판결이 사진 업계와 더 넓은 창작 커뮤니티의 우려를 불러일으킬 것이라는 견해도 제기되는바, 대상 판결의 판단에 의해 제60d조가 자신의 작품에 대해 가질 수 있는 사진작가 등 창작자의 통제권을 제한하게 되는 결과가 초래된다는 점을 그 이유로 하고 있다.³⁸⁾

대상 판결은 AI 학습 데이터셋 생성에 대한 법적인 명확성을 제공하고, 오픈소스 커뮤니티와 AI 연구 발전에 긍정적인 영향을 미칠 것으로 예상하여 볼 수 있다. 그리고 대상 판결은 AI 개발과 저작권법 사이의 균형을 모색하는 중요한 선례가 될 것으로 보인다. 특히 대상 판결은 오픈소스 AI 개발과 과학적 연구 목적의 데이터 이용에 대한 법적 근거를 제공했다는 점에서 가장 큰 의미가 부여될 수 있겠다.

34) Heidrich Rechtsanwälte, 앞의 글.

35) Heidrich Rechtsanwälte, 앞의 글.

36) Heidrich Rechtsanwälte, 앞의 글.

37) Andres Guadamuz, "LAION wins copyright infringement lawsuit in German court", TechnoLlama, 2024.9.28. (<https://www.techollama.co.uk/laion-wins-copyright-infringement-lawsuit-in-german-court>)(최종접속일 : 2024.11.4.).

38) Andres Guadamuz, 앞의 글.

원고는 대상 판결에 대하여 불복하는 항소를 제기할 수 있는데, 현재까지는 아직 항소 여부를 확인할 수 없지만, 항소가 제기될 경우에는 이에 대하여 어떠한 판단이 새롭게 이루어질지 계속 주의 깊게 그 추이를 살펴볼 필요가 있을 것이다.