

OpenChain-KWG Tooling & Legal SG 21차 회의

Robert Kneschke v. LAION 소송 사례

2025. 4. 8.(화)

ETRI 박정숙

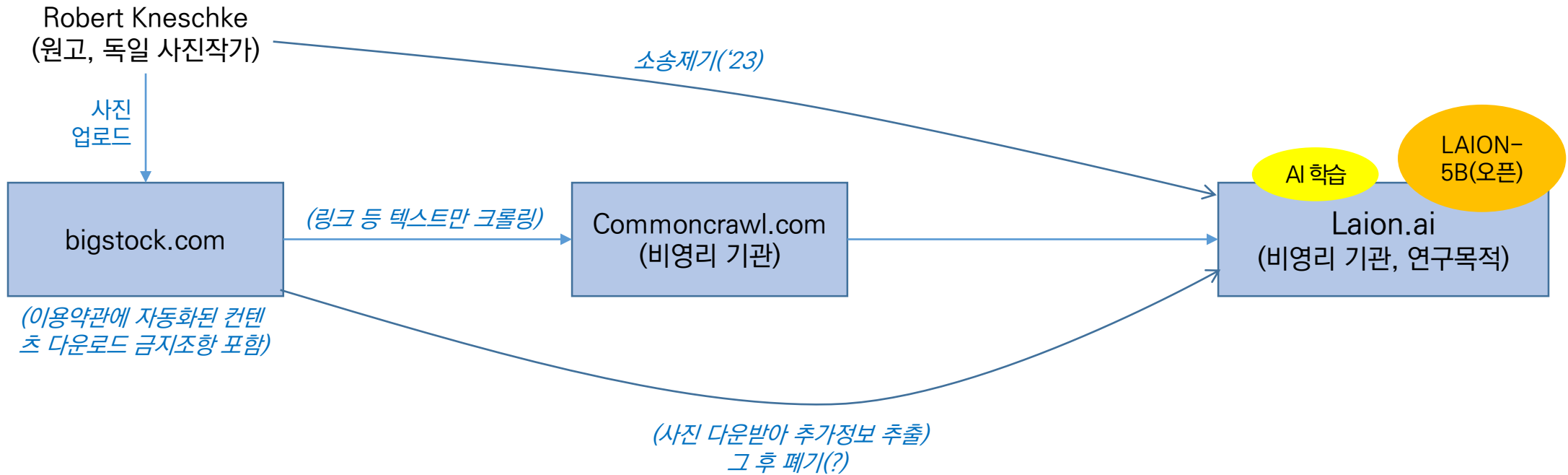
Robert Kneschke v. LAION

- 독일 함부르크 지방 법원에서 진행된 TDM을 다룬 첫번째 판결
 - 연구 기관이 웹사이트에서 이미지를 자동으로 다운로드하고 분석하는 행위가 저작권법 제60d조에 의해 보호받을 수 있음
 - ❖ 저작권법 제60d조는 연구 목적의 TDM 허용, 저작권자의 사전 동의 없이도 복제 가능하도록 규정
- 관련 규정이 국제적 특히 EU 회원국 간에 일관성있게 적용되어야 함을 강조
 - AI 훈련용 데이터 수집/분석은 국경을 초월하는 활동 → 통일된 법적 기준이 중요

Robert Kneschke v. LAION: 경과

- ‘21년, 로버트 크네슈케(독일, 사진작가)는 자신의 사진이 LAION-5B에 무단으로 포함된 사실을 인지
 - ‘LAION-5B’ 데이터셋은 약 58억 개의 이미지 관련 메타 정보로 구성
- 법적 분쟁 시작(’23.4.27., 독일 함부르크 지방법원)
 - 크네슈케는 함부르크 지방법원에 LAION을 상대로 저작권 침해 소송 제기
 - ❖ 자기 사진이 무단 사용된 것에 대해 이의를 제기, 데이터셋에서 자신의 이미지 제거 요구
 - ❖ 자신의 사진이 포함된 웹사이트의 이용약관에 자동화된 콘텐츠 다운로드를 금지하는 조항이 있음에도 불구하고, LAION이 자신의 사진을 무단으로 다운로드하고 분석한 것이 저작권 침해라고 주장
 - LAION은 자신들의 활동이 과학 연구 목적의 TDM이므로 저작권법 제60d조에 따라 허용된다고 반박
- 판결(’24.9.27.): LAION 행위가 저작권 침해 아니라고 판결
 - LAION의 데이터셋 생성 활동은 독일 저작권법 제60d조에 따른 비상업적 과학 연구 목적의 TDM에 해당
- 원고 항소 제기

Robert Kneschke v. LAION: 사실 관계도



Robert Kneschke v. LAION: 정보 수집 및 가공 과정

- LAION-5B 학습데이터 수집 방법
 - Common Crawl(비영리 조직)의 제공 정보 활용
 - ❖ 정기적으로 인터넷의 ‘백업’ 또는 ‘이미지’를 생성. 링크를 통해 접근 가능한 모든 웹페이지를 텍스트 형태로 복제
 - Common Crawl의 데이터 수집 방식
 - ❖ 웹페이지의 텍스트 내용 복제
 - ❖ 이미지, 비디오 등 비텍스트 데이터는 직접 저장하지 않음. 대신 콘텐츠에 대한 링크를 포함한 웹페이지의 소스 코드 저장
 - Common Crawl은 수집한 데이터셋을 자체 웹사이트에서 제공
 - ❖ 데이터셋은 웹페이지의 소스코드를 포함하고 있어, 연구자들이 인터넷의 구조와 내용을 분석하는 데 사용 가능

- LAION의 데이터 처리 과정
 - LAION은 Common Crawl이 제공하는 이 데이터셋을 활용하여 자체적인 이미지 데이터셋 생성
 - ❖ Common Crawl 데이터셋에서 이미지 링크 추출: LAION은 Common Crawl 데이터에서 이미지 파일에 대한 링크만을 필터링
 - ❖ 이미지 다운로드 및 분석: 추가 정보 얻기 위해, 수집한 링크 통해 실제 이미지를 다운로드하고, 자체 개발한 AI 모델을 사용하여 이미지 분석
 - ✓ 이미지 설명
 - ✓ 워터마크 유무
 - ✓ 청소년 유해 콘텐츠 포함 여부
 - ❖ 데이터셋 구성: 최종적으로 LAION이 만든 데이터셋은 일종의 표 형태로, 각 행에는 이미지 링크와 해당 이미지에 대한 추가 정보 포함

Robert Kneschke v. LAION: 관련 법률(1/2)

- DSM 지침 (Digital Single Market Directive)

- 2019년 EU는 디지털 단일 시장 저작권 지침(DSM Directive) 채택, 2021.6.7.부터 EU 회원국들에서 시행
- 예외: TDM(텍스트 및 데이터 마이닝)에 대한 두 가지 예외 규정 포함(제3조, 제4조)

제3조. 과학 연구 목적을 위한 TDM	제4조. TDM에 대한 예외 또는 제한
<ol style="list-style-type: none"> 1. 회원국은 지침 96/9/EC 제5조(a) 및 제7조(1)항, 지침 2001/29/EC 제2조, 이 지침 제15조(1)항에 규정된 권리에 대한 예외를 연구 기관 및 문화 유산 기관이 과학 연구 목적으로 합법적으로 접근할 수 있는 작품 또는 기타 주제에 대한 텍스트 및 데이터 마이닝을 수행하기 위해 복제 및 추출하는 경우를 규정해야 한다. 2. 제1항에 따라 작성된 작품 또는 기타 주제의 사본은 적절한 수준의 보안으로 저장해야 하며 연구 결과의 검증을 포함한 과학 연구 목적으로 보관할 수 있다. 3. 권리자는 작품 또는 기타 주제가 호스팅되는 네트워크 및 데이터베이스의 보안 및 무결성을 보장하기 위한 조치를 적용할 수 있다. 이러한 조치는 해당 목표를 달성하는 데 필요한 범위를 넘어서는 안 된다. 4. 회원국은 권리자, 연구 기관 및 문화 유산 기관이 2항 및 3항에 각각 언급된 의무 및 조치의 적용과 관련하여 일반적으로 합의된 모범 사례를 정의하도록 장려해야 한다. 	<ol style="list-style-type: none"> 1. 회원국은 지침 96/9/EC 제5조(a) 및 제7조(1)항, 지침 2001/29/EC 제2조, 지침 2009/24/EC 제4조(1)(a) 및 (b)항, 이 지침 제15조(1)항에 규정된 권리에 대해 텍스트 및 데이터 마이닝 목적으로 합법적으로 접근 가능한 작품 및 기타 주제를 복제 및 추출하는 것에 대한 예외 또는 제한을 규정해야 한다. 2. 단락 1에 따라 이루어진 복제 및 추출은 텍스트 및 데이터 마이닝 목적에 필요한 기간 동안 보관할 수 있다. 3. 단락 1에 규정된 예외 또는 제한은 해당 단락에 언급된 작품 및 기타 주제의 사용이 온라인에서 공개적으로 제공되는 콘텐츠의 경우 기계 판독 가능한 수단과 같이 적절한 방식으로 권리자가 명시적으로 예약하지 않은 경우에 적용된다. 4. 이 조항은 이 지침 제3조의 적용에 영향을 미치지 않는다.
<ul style="list-style-type: none"> • 대상: 연구 기관 및 문화유산 기관에만 적용 • 목적: 오직 과학 연구를 위한 목적으로만 허용 • 권한: 저작권자의 사전 허가가 필요 없으며, 어떤 형태의 보상도 요구되지 않음 • 접근 조건: 합법적으로 접근할 수 있는 데이터에만 적용(예: 구독, 라이선스 등) • 제한: 민간 기업의 결정적인 영향력 하에 있는 기관은 제외 	<ul style="list-style-type: none"> • 대상: 모든 개인이나 단체에 적용 • 목적: 모든 목적(상업적 목적 포함)의 TDM에 적용 • 권한: 저작권자가 명시적으로 권리를 유보하지 않은 경우에만 적용 • 접근 조건: 합법적으로 접근할 수 있는 데이터에만 적용 <ul style="list-style-type: none"> ❖ Opt-out 메커니즘: 저작권자는 '적절한 방식'으로 권리를 유보 가능(예: 온라인 콘텐츠의 경우 기계가 읽을 수 있는 형식) • 데이터 보관: TDM 목적으로 복제물을 보관 가능

Robert Kneschke v. LAION: 관련 법률(2/2)

- 독일은 DSM에 따라 국내 저작권법 개정

- 제44b조

- ❖ 일반적 목적의 TDM에 대한 예외 규정 신설

- ❖ 모든 목적의 TDM을 허용하지만, 저작권자가 명시적으로 거부(opt-out)할 수 있는 권리 인정

- 제60d조

- ❖ 과학 연구 목적의 TDM에 대한 기존 예외 규정 확대

- ❖ 비상업적 과학 연구 목적의 TDM에 대해 더 넓은 자유 부여, 저작권자의 opt-out 권리 인정 **않음**

Robert Kneschke v. LAION: 핵심 쟁점

• TDM 예외 규정의 적용 범위

- TDM 예외 규정: 저작권법에서 특정 조건 하에 저작권자 허락없이 저작물을 사용가능하도록 하는 규정
 - ❖ 연구나 기술 발전을 위해 대량의 데이터를 분석할 필요가 있는 경우에 적용
- Kneschke v. LAION 소송에서는 AI 학습을 위한 데이터셋 생성이 이 예외 규정에 해당하는지가 쟁점
 - ❖ 연구 목적으로 웹사이트의 텍스트를 자동 수집하고 분석하는 것이 저작권 침해인지, 예외에 해당하여 허용되는지 판단

• 비상업적 과학 연구 목적의 정의

- LAION이 주장하는 비상업적 과학 연구가 무엇인지, LAION 활동이 이에 해당하는지가 논점

• 저작권자의 ‘opt-out’ 권리의 유효성

- Opt-out: 저작권자가 자신의 작품이 TDM에 사용되는 것을 거부할 수 있는 권리
- 이 권리를 어떻게 행사할 수 있고, 어떤 형태의 거부 의사 표시가 유효한지

Robert Kneschke v. LAION: 판결 주요 내용

- 판결: 함부르크 지방법원은 원고의 청구는 적법하지만 이유가 없다고 판결
- 1) 원고 사진의 속성
 - 사진저작물(저작권법 제2조 제5항)이 아니라 일반 사진(제72조제1항, 저작인접권)으로 판단. 저작인접권자도 복제권을 가지므로 복제권 침해에 대하여 침해금지청구권(저작권법 제97조 제1항) 행사 가능하지만 피고의 복제행위는 법률상 제한 규정을 통하여 정당화될 수 있음
- 2) 일시적 복제 해당 여부
 - 피고의 복제행위는 일시적 복제행위의 제한 규정(제44a조)에 의해서는 허용 안됨. 이미지 파일은 특정한 SW를 사용하여 분석하기 위한 목적으로 다운로드되었으므로 분석 수행에 수반되는 단순한 과정이 아니라, 분석의 전 단계에서 의도적이고 적극적으로 조정된 작성 과정으로 봄
- 3) 일반적(상업적) 목적의 TDM 해당 여부
 - 다운로드가 TDM 목적으로 수행되었으므로 저작권법 제44b조 제2항의 상업적 목적의 제한 규정에 해당할 수 있다고 봄
 - 제44b조 제2항에 의해서 피고의 복제 행위는 정당화되지만, 제44b조 제3항의 이용유보(제한의 예외)에 의해서 정당화되지 않는다고 봄
 - ❖ 권리자가 유보를 선언하지 않은 경우에만 TDM이 정당화됨(온라인 접근인 경우 기계가 읽을 수 있는 방식으로 선언되어야 유효)
- 4) 과학적 연구 목적의 TDM 해당 여부
 - 피고의 복제는 저작권법 제60d조의 제한 규정에 해당한다고 판단(연구 기관의 과학적 연구 목적의 TDM을 위한 복제는 허용)
 - AI 시스템을 학습하기 위한 토대가 될 수 있는 데이터셋의 구축은 과학적 연구로 볼 수 있다고 판단
 - ❖ 데이터셋이 무료로 공개되어 AI 분야의 연구자들에게 제공되었다는 것만으로도 충분
 - 데이터셋이 영리 기업에서 학습이나 AI 시스템의 추가 개발을 위해 사용되었는지 여부는 중요하지 않다고 판단

Robert Kneschke v. LAION: 향후 전망

- 판결의 의의
 - DSM 지침의 TDM 제한 규정이 AI 학습 데이터 수집에 적용될 수 있는지 여부에 대한 EU 최초의 사례
 - 이미지뿐만 아니라 텍스트, 코드, 동영상, 음악 등 모든 종류의 데이터에 적용될 수 있음
- 판결 대상은 학습을 위한 데이터셋의 구축 과정에서 발생하는 복제행위에 초점
 - TDM 제한 조항이 적용될 수 있는 단계를 학습용 데이터셋 구축 단계로 제한
- 법원은 AI 학습을 위한 데이터셋 구축을 과학적 연구의 범주에 포함
 - AI 개발 과정을 연구 활동으로 인정한 것
 - EU 인공지능법 제53조 제1항 c가 DSM 지침 제4조 제3항의 권리유보를 명시적으로 언급 → 인공지능법과 저작권법 간 연계성 인정
 - 자연어로 선언된 사용 유보도 제44b조 제3항 제2문에서 의미하는 기계 판독이 가능하다고 간주
- TDM 제한 규정의 해석 및 적용에 대한 논쟁 촉발
 - 학계의 다수의견과 상당한 거리
- 향후 상급심 법원이나 유럽 사법재판소의 판결이 이 문제에 대한 최종적인 해석을 제공할 것으로 예상

참고자료

- Lisa Käde, “To Mine or Not To Mine”, <https://jbb.de/to-mine-or-not-to-mine/>, JBB, 2024.10.27.
- 장학성, “To Mine or Not To Mine: 독일 법원이 AI 시대의 저작권 딜레마에 내린 판결”, Openchain-kwg, 2024.11.6.
- 판결문, <https://openjur.de/u/2495651.html>, 2024.9.27.
- Common Crawl, <https://commoncrawl.org/latest-crawl>
- LAION-5B, <https://laion.ai/blog/laion-5b/>
- 박희영, “[이슈리포트] 2024-28-[독일] 생성형 인공지능 모델 학습과 TDM 제한(박희영)”, 한국저작권위원회, 2024.10.31.

감사합니다



ETRI 오픈소스센터
Open Source Center