

AI와 저작권의 균형을 모색하는 EU·프랑스의 실험

이원

인천가톨릭대학교 문화콘텐츠학과 교수

목차

- 1 AI와 저작권의 갈등
- 2 EU 저작권 지침(2019)의 TDM 규정
- 3 EU AI Act(2024)의 TDM 규정
- 4 프랑스의 TDM 법제와 특징
- 5 EU와 프랑스 TDM 규제의 함의

요약

AI가 전 산업으로 확산하며 핵심 정책 의제가 되었지만, 인간 창작물과의 갈등으로 문화산업에 불안을 초래하고 있다. 생성형 AI는 대량의 저작물을 학습 데이터로 활용하면서 기존 저작권 체계와 충돌하고, 이에 따라 TDM 규율이 저작권 정책의 핵심이 되었다. EU와 프랑스는 연구 목적 TDM은 폭넓게 허용하되 일반 TDM은 권리자의 명시적 이용거부(opt-out)를 전제로 하는 규정을 마련하고, 이를 AI 규제와 연계해 저작권 준수와 데이터 투명성을 강화했다. 전체적으로 EU와 프랑스의 규제는 AI 발전과 저작권 보호 간의 균형을 모색하는 규범적 대응이라 할 수 있다.

1. AI와 저작권의 갈등

2022년 ChatGPT의 등장 이후 인공지능(AI)의 열기는 식을 줄 모르고 있다. 이제 모든 분야에서 AI를 제외하고는 얘기가 되지 않을 정도로 AI는 우리의 삶 곳곳에 스며들고 있다. 전 세계 모든 국가가 AI 산업에 정책 역량을 집중하고 있다. 2025년 우리나라 정부가 AI를 핵심 성장 동력으로 제시한 것도 이 맥락 속에 있다.

생성형 AI는 텍스트·이미지·음악·영상 등 다양한 문화콘텐츠 생산 영역으로 빠르게 침투하며 생산성 향상에 크게 기여하고 있다. 하지만, 동시에 인간 창작 활동의 대체 가능성을 높여 고용 안정성과 문화산업 생태계의 구조적 변동에 대한 우려도 낳고 있다. 실제로 AI 도입이 작가의 생계를 위협한다는 문제 제기는 2023년 미국 할리우드 작가 파업 사례에서도 확인되며, 기술이 제공하는 효율성과 혁신의 이면에 구조적 위험이 공존하고 있음을 시사하고 있다(정재훈, 2023).

AI를 둘러싼 문제는 단순한 고용 감소나 창작 도구의 변화에만 국한되지 않는다. 생성형 AI가 방대한 학습 데이터를 기반으로 작동하기 때문에, AI 모델의 성능 향상을 위해 저작물이 대량으로 수집·분석되는 과정에서 기존 저작권 체계와 충돌이 불가피하게 발생한다. 저작물의 무단 활용, 데이터 수집 과정의 정당성, 기술적 보호조치 무력화 가능성 등이 주요 쟁점으로 떠오르면서, 학습 데이터 단계에서의 규율이 AI 시대 저작권 정책의 핵심으로 부상하고 있다. 이러한 배경 속에서 텍스트·데이터 마이닝(TDM)에 관한 논의가 다시 주목되고 있으며, 특히 AI와 결합된 TDM의 법적 해석과 정책적 방향은 국제적으로도 중요한 논쟁 축이 되고 있다.

TDM은 AI 훈련의 기본 전제가 되는 매우 중요한 기술이다. 우리나라에서도 2020~2021년 TDM 예외 규정을 둘러싼 입법 시도가 있었으나, 상업성 배제 문제, 국제적 정합성 결여, 개념적 모호성 등의 이유로 무산되면서 제도적 공백이 채워지지 않고 있다. 그 사이 AI는 더 고도화되었고 학습 데이터의 확보가 산업 경쟁력의 핵심 요소로 부상함에 따라, TDM 규율은 단순한 기술 문제를 넘어 문화콘텐츠산업의 심장에 비유될 수 있는 저작권 체계의 근본적 재설계를 요구하는 수준으로 확장되었다. 이러한 맥락에서 본고는 특히 EU와 프랑스의 AI 학습 데이터와 저작권의 관계를 규정한 TDM 규제를 살펴보고 시사점을 모색해 보고자 한다.

2. EU 저작권 지침(2019)의 TDM 규정

1) TDM 정의와 지침의 의의

2019년 EU 저작권 지침은 디지털 환경에서 EU의 기존 저작권 정책을 획기적으로 혁신한 것으로 평가되고 있는 중요한 지침이다(이원, 2019). 이 지침은 대량 데이터 분석이 필수적인 AI 시대를 맞아 TDM을 처음 체계적으로 규율한 법령이라는 점에서 중요한 전환점이 된다. 지침 제2조 제2항은 TDM을 “디지털 형식의 텍스트와 데이터를 자동 분석하여 패턴·경향·상관관계를 도출하는 기술”로 정의하고 있다. 이는 곧 AI 학습의 핵심 과정으로 작동한다. 이러한 정의는 TDM이 단순한 열람이나 복제가 아니라 ‘정보 도출’을 목표로 하는 분석 행위라는 점을 명확히 하여, 저작물 이용의 새로운 유형을 법적으로 포착하려는 시도라 할 수 있다.

2) 연구 목적 TDM 예외 규정

TDM이 저작물 복제·추출을 전제로 하기 때문에 EU는 연구 목적과 일반 목적을 구분해 별도 규정을 마련했다. 먼저 지침 제3조는 연구기관 및 문화유산기관이 과학적 연구를 위해 합법적으로 접근 가능한 저작물에 대해 TDM을 수행할 수 있도록 허용하고 있다. 여기서 합법적 접근은 구입이나 구독 등 정당한 권한에 기초해 저작물에 접근하는 경우를 포함한다. 연구를 위해 생성된 복제물은 적절한 보안 수준에서 저장될 수 있으며(제2항), 권리자는 자신의 저작물이 저장된 네트워크의 보안과 무결성을 유지하기 위한 조치를 취할 수 있지만 이 목적을 넘어서 연구기관의 TDM 활동을 방해해서는 안 된다(제3항). 또한 연구기관과 권리자가 협력하여 모범사례를 마련하도록 권고해 분쟁 예방과 기술적 보호조치 충돌을 완화하려는 내용을 담고 있다(제4항). 제3조는 결국 공공적 성격의 연구행위에 한정하여 저작권을 예외적으로 제한한 규정으로, AI 연구 생태계의 기반을 마련한 조항이라 평가할 수 있다.

3) Opt-out 기반 일반 TDM 규정

한편 제4조는 특정 기관에 한정되지 않은 ‘일반 TDM’ 규정을 두고 있으며, 바로 이 조항이 AI 산업에서 가장 큰 영향을 미친다. 제4조는 누구든 합법적으로 접근 가능한 저작물에 대해 TDM을 위해 필요한 복제·추출을 할 수 있도록 허용하지만, 중요한 조건이 있다.

바로 권리자가 “명시적으로 이용을 거부하지 않은 경우”에만 적용된다는 점이다(제3항). 즉, 권리자가 기계가 읽을 수 있는 형태로 ‘이용 거부(opt-out)’를 표시하면 TDM 예외가 적용되지 않으며, 이때는 사전 이용 허락을 받아야 한다. 반대로 권리자가 별도의 거부 표시를 하지 않으면 개발자, 기업, 연구자 등 누구나 허락 없이 저작물을 TDM에 활용할 수 있다. 이 때문에 제4조는 사전이용허락을 의미하는 옵트인(opt-in) 방식보다 훨씬 개발자에게 유리한 규범으로 평가되며, EU가 AI 산업의 데이터 접근성을 일정 부분 보장하려는 정책적 선택을 한 것으로 해석된다. 다만 이는 권리자의 통제권을 약화시키는 효과가 있어 데이터 무단 활용 논란을 가져올 수 있지만, 예외의 범위와 한계를 명확히 하여 법적 예측 가능성을 높였다는 점에서 의의가 있다.

3. EU AI Act(2024)의 TDM 규정

1) 위험 기반 규제 체계와 시행 구조

EU 저작권 지침(2019)이 처음으로 체계적인 TDM 규정을 마련했다면, 2024년 제정된 EU AI Act는 범용 AI(이하 ‘GPAI’) 모델을 대상으로 저작권 준수 의무와 학습 데이터 요약 공개 의무까지 부과함으로써 투명성과 감독 체계를 한층 강화했다는 데 의의가 있다. EU AI Act는 2024년 7월 12일 EU 관보에 게재되어 2024년 8월 1일 발효되었으며, 단일 법령이지만 실제 시행은 2025년부터 2027년까지 네 단계에 걸쳐 순차적으로 진행된다. 2025년 2월에는 ‘수용 불가 위험 AI’ 금지 규정이 먼저 시행되었다. 2025년 8월에는 GPAI 모델 규정이 적용되었다. 이후 2026년에는 ‘고위험 AI’ 의무 규정 대부분이 적용되며, 2027년에 고위험 AI 제품의 안전 구성 요소에 관한 의무가 최종적으로 시행된다. 이러한 다층적 시행 구조는 AI 시스템을 위험 등급에 따라 구분하고, 기술적, 사회적 영향력이 큰 범용 모델에 대해 보다 엄격한 규제를 적용하는 방식으로 마련되었다.

2) GPAI 개발자 의무와 저작권 규정

EU AI Act에서 가장 주목되는 부분은 GPAI에 관한 제5장의 규정이다. GPAI는 고위험 AI 시스템의 구성요소로 사용되거나 다양한 분야에 통합될 가능성이 크기 때문에, EU

는 GPAI가 시장에 배포되기 이전 단계부터 개발자에게 구체적인 법적 의무를 부과한다. 제 53조 제1항에 따르면 GPAI 개발자는 네 가지 의무를 부담한다. 첫째, 모델의 훈련 및 테스트 과정, 평가 결과를 포함한 기술 문서를 작성하고 갱신하여 유지해야 한다. 둘째, GPAI를 자사 AI 시스템에 통합하려는 다운스트림 제공자(downstream provider)에게 모델을 이해하고 규정을 준수할 수 있도록 필요한 정보를 제공해야 한다. 셋째, EU 저작권법 준수를 위한 내부 정책을 마련해야 하며, 특히 EU 저작권지침 제4조(3)에 따라 ‘이용 거부(opt-out)’ 의사를 최신 기술로 확인, 준수해야 한다. 넷째, AI 사무국(AI Office)이 제시한 양식에 따라 GPAI 훈련에 사용된 콘텐츠의 ‘상세한 요약’을 공개해야 한다는 점이다.

이 중에서 저작권과 가장 직접적으로 연결되는 규정은 제53조 제1항 (c)와 (d)다. (c)항은 학습 데이터가 저작권 보호대상인 경우, TDM 예외 또는 제한에 해당하지 않는다면 권리자의 사전 이용 허락이 필요하다는 원칙을 명확히 하였다. EU AI Act의 전문 105는 생성형 모델에 방대한 텍스트, 이미지, 데이터에 접근해야 하며 이들 상당수가 저작권 보호 대상이므로, 이를 임의로 사용할 수 없다는 점을 강조하고 있다. 또 (d)항은 EU AI Act가 추가한 TDM 관련 새로운 의무로서, GPAI 훈련에 사용된 자료의 ‘상세한 요약’에는 “저작권법으로 보호되는 텍스트 및 데이터를 포함해야 한다.”라는 점을 규정하고 있다. 이는 투명성 확보를 위한 조치로, AI 기업의 영업비밀 보호 필요성은 인정하면서도 학습 데이터의 출처와 구성에 대한 일정 수준의 공개를 의무화한 것이다.

3) 무료·오픈 GPAI 모델의 차등 규제

한편, EU AI Act는 무료·오픈 형태의 GPAI 모델에 대해 차등적인 규제를 적용한다. 즉, 무료로 제공되거나 모델 가중치가 공개된 오픈 GPAI는 기술 문서(a)와 사용설명서(b) 작성 의무에서 면제된다. 그러나 저작권 관련 의무(c)와 훈련데이터 상세 요약 공개 의무(d)는 그대로 적용된다. 이는 오픈소스 모델이라도 학습 데이터 사용 단계에서 권리자의 opt-out 의사를 확인해야 하며, 투명성 확보를 위해 데이터 요약 공개는 필수라는 점을 분명히 한 것이다. 반면 시스템적 위험이 높은 GPAI 모델은 이러한 면책이 허용되지 않고, 더 강한 감독과 보고 의무가 적용된다. EU는 이를 통해 혁신적인 오픈 생태계는 보호하되, 대규모 상업적 모형이 발생시키는 사회적 위험에 대해서는 강화된 규제를 적용하는 균형적 체계를 모색하고 있다.

4. 프랑스의 TDM 법제와 특징

1) EU 지침의 국내법 수용

프랑스는 EU 저작권 지침(2019)을 여러 법령을 통해 순차적으로 국내법에 수용하였다. 그중 2021년에 제정된 법률명령(2021-1518)은 프랑스 지식재산권법전(Code de la propriété intellectuelle, 이하 CPI)¹⁾에 제L122-5-3조를 새로 삽입함으로써 프랑스 법제에 TDM에 관한 규정을 도입하였다. 이 조항은 EU 저작권 지침의 제3조와 제4조를 기반으로 이를 보다 명확하고 구체적인 방식으로 확장하고 있다. 즉 프랑스는 EU 지침의 주요 내용을 충실히 반영하되, 자국의 환경을 고려하여 특정 조항에 대해 세부적인 내용을 추가한 것이다.

2) 프랑스 TDM 규정의 특징

프랑스의 TDM 관련 규정의 주요 내용은 다음과 같다.

○ TDM 예외 기관 범위의 구체화

EU 지침이 “연구기관과 문화유산 기관”이라고 규정한 데 비해, 프랑스는 이를 공공 도서관, 박물관, 기록보관소, 영화, 시청각·음향 유산 보관기관 등으로 구체화하였다. 또한 이 기관들을 대신하거나 그 요청에 따라 민간 주체가 TDM을 수행하는 경우에도 TDM 예외 적용이 가능하도록 하여, 비영리적 민·관 파트너십을 통한 TDM 활용도 폭넓게 허용하였다. 이는 EU 규정보다 폭넓게 연구협력 방식을 인정한 것으로, 실제 연구 개발 환경에서 발생하는 다양한 공동 프로젝트를 제도적으로 뒷받침하는 장점이 있다. 다만 연구결과를 특정 기관의 주주나 관계기업이 독점적으로 이용하는 것은 허용되지 않아 공익성을 명확히 유지하도록 하였다. 또한 관리자와 연구기관·보관기관은 모범사례 협약을 체결할 수 있도록 규정하여 기술적 보호조치로 인한 갈등을 예방하는 장치를 마련했다(김경숙, 2021).

1) 프랑스 지식재산권법전은 저작권뿐만 아니라 특허, 상표, 디자인 등 산업재산권을 포함하는, 소위 '지식재산(IP)' 전반을 아우르는 법전이다.

○ 기계가 읽을 수 있는 형태의 opt-out 적용 방식

CPI는 EU 지침 제4조의 원칙을 그대로 수용하여, 저작자가 온라인에서 기계가 읽을 수 있는 방식으로 명시적 거부 의사를 표시한 경우, 누구든지 해당 저작물을 TDM에 이용할 수 없다고 규정하고 있다. 이는 opt-out 원칙을 충실히 반영한 것으로, 저작자가 TDM 활용을 통제할 수 있는 최소한의 권리를 보장한다. 반대로 ‘명시적’ 거부 표시가 없다면 이 용허락 없이 TDM이 가능하므로 연구자, 개발자에게는 법적 예측 가능성을 제공한다.

○ 복제물 보관·폐기 규정의 차별화

프랑스 규정은 복제물 보관 규정에서 EU 지침보다 더 정확한 구분을 제시한다. 연구기관 등의 공익기관은 TDM 과정에서 생성된 복제물을 적절한 보안 하에 보관할 수 있다. 그러나 일반 이용자가 제3항에 따라 TDM을 수행하는 경우, 생성된 복제물은 TDM이 완료되면 반드시 폐기해야 한다. 이는 공익적 연구기관과 일반 사용자를 명확히 구분하여, 연구 목적의 지속적 검증을 가능하게 하면서도 상업적, 일반적 데이터 축적의 남용 가능성을 차단하는 조치이다.

5. EU와 프랑스 TDM 규제의 함의

EU와 프랑스의 TDM 규제는 AI 학습 데이터를 둘러싼 저작권 갈등이 단순한 기술 문제가 아니라, “문화콘텐츠산업의 심장에 비유될 수 있는 저작권 체계의 근본적 재설계”를 요구하는 수준으로 확장되었음을 잘 보여준다. EU 저작권 지침(2019)은 연구 목적 TDM과 일반 TDM을 분리해 규율함으로써, AI 연구 생태계의 기반을 마련하는 동시에 저작물 이용의 새로운 유형을 법적으로 포착했다. 특히 공공적 성격의 과학 연구에 한하여 저작권을 예외적으로 제한하는 반면, 권리자의 명시적 opt-out을 전제로 일반 TDM을 허용함으로써, AI 연구·교육 분야의 공공성과 상업적 AI 개발 영역에서의 권리자 선택권을 조화시키는 균형적 모델을 제시하였다.

EU AI Act(2024)는 이러한 TDM 규율을 GPAI 모델에 직접 연결시켜, 저작권 준수 의무와 학습 데이터 ‘상세한 요약’ 공개 의무를 부과함으로써 투명성과 감독 체계를 한

층 강화했다. GPAI 개발자에게 기술 문서 작성, AI 모델 통합자에 대한 정보 제공, opt-out 확인 및 준수, 훈련데이터 요약 공개를 요구하는 구조는 AI 기업의 영업 비밀 보호 필요성을 인정하면서도 학습 데이터의 출처와 구성을 일정 수준 공개하도록 하는 장치로 기능한다. 무료·오픈 GPAI에 대해서도 저작권 관련 의무와 데이터 요약 공개를 그대로 유지하면서, 혁신적인 오픈 생태계를 보호하되 대규모 상업적 모형의 사회적 위험에는 강화된 규제를 적용하는 방향을 모색하고 있다.

프랑스는 EU 지침을 국내법으로 수용하면서 TDM 예외가 적용되는 기관 범위를 세부적으로 규정하고, 비영리적 민·관 파트너십까지 포함시켜 실제 연구협력 방식을 폭넓게 인정하였다. 동시에 기계가 읽을 수 있는 형태의 opt-out을 통해 저작자가 TDM 활용을 통제할 수 있는 최소한의 권리를 보장하고, 공익기관과 일반 이용자를 구분해 복제물 보관과 폐기를 다르게 규율함으로써 연구 목적의 지속적 검증과 데이터 남용 방지 사이의 균형을 도모하였다. 이러한 EU와 프랑스의 TDM 규제는 생성형 AI의 고도화 속에서 학습 데이터 확보 필요성과 저작권 보호, 데이터 접근성과 권리자 통제권, 그리고 AI 산업의 성장과 문화산업 생태계의 안정성 사이의 긴장을 조정하려는 규범적 실험이라는 점에서 중요한 함의를 가진다.

참고문헌

- 김경숙(2021), "TDM 관련 저작권법 개정안의 비판적 고찰 : 기술적보호조치 관점에서의 비교법적 검토", 경영법률, 31권 3호, 한국경영법률학회, p. 119.
- 이원(2019), "유럽연합 디지털 단일시장 저작권 지침의 개혁적 함의", 문화정책논총, 33권 2호, 한국문화관광연구원, p. 29.
- 이원(2025), "AI 학습 데이터와 저작권 : EU · 프랑스의 TDM 규제 분석과 시사점", 계간 저작권, 38권 3호, 한국저작권위원회, 75-105.
- 정재훈(2023), "AI 글쓰기와 소설의 미래 : 박금산 장편소설 AI가 쓴 소설을 중심으로", 비평문학, 90, 한국비평문학회, p. 257.
- European Union(2019), Directive 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE.
- European Union(2024), Regulation(EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance).