2026. 10.

EU 권리자의 텍스트·데이터 마이닝 권리 유보 표시 등록의 실현 가능성 평가 연구보고서 발표

한국저작권위원회 통상산업연구팀
강민지 선임

1. 개요

2026년 7월, EU는 「권리자의 텍스트·데이터 마이닝(TDM) 권리 유보 표시 등록의 실현 가능성 평가 연구(Study to Assess the Feasibility of a Registry of Text and Data Mining Opt-outs expressed by Rightholders)」 보고서를 발표했다. 이 연구는 「디지털 단일시장의 저작권 및 저작인접권에 관한 지침」(이하 'DSM 지침') 제4조 제3항[1]에 따른 TDM 권리 유보(opt-out) 표시를 등록하고 조회할 수 있는 시스템(이하 '시스템')을 도입할 경우의 기대 효과와 기술적·운영적 실현 가능성을 평가하기 위해 수행되었다.

연구진은 문헌조사, 이해관계자 인터뷰, 설문, 다자간 워크숍을 통해 근거를 수집했다. 인터뷰에는 시청각, 음악, 출판, 언론, 실연, 시각예술 등 창작 분야의 권리자를 비롯하여 AI 업계, TDM 옵트아웃(opt-out) 솔루션 제공자, 시민사회단체(CSO) 등 다양한 이해관계자가 참여했다.

연구 결과, EU 차원의 시스템은 기존 옵트아웃 표시 수단을 보완하는 수단으로서 도입에 따른 기대 효과가 있을 수 있는 것으로 나타났다. 또한 운영·관리 체계, 신원 확인, 서비스 수준, 감사 기록 등의 조건이 갖춰진다면 하이브리드 방식으로 기술적·운영적 실현이 가능하다고 보았다. 여기서 하이브리드 방식이란 등록은 권리자가 직접 하거나 집중관리단체·출판사 등 분야별·국가별 창구를 통해 하고, 조회는 중앙 한 곳에서 처리하는 방식을 말한다.

2. 현행 옵트아웃 표시의 한계

연구진은 현행 옵트아웃 표시 수단들이 각각 한계를 지니고 있어 권리자의 옵트아웃 의사가 AI 개발자에게 제대로 전달되지 않고 있다고 진단했다.

먼저, 사이트 기반 표시 방식인 robots.txt는 웹사이트 운영자가 사이트 최상위 경로에 두는 텍스트 파일로, 검색엔진이나 AI 크롤러 등 자동 수집 프로그램에 사이트 전체 또는 일부 경로에 대한 접근 허용

1) DSM 지침(Directive (EU) 2019/790) 제4조 제3항: "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.

여부를 알리는 방식이다. 그러나 이는 웹사이트 운영자가 설정하는 방식이기 때문에, 저작물이 다른 사이트로 복제·재배포되는 경우 옵트아웃 표시가 함께 이전되지 않는다는 한계가 존재한다.

다음으로 저작물 기반 표시 방식으로는 TDMRep, IPTC 등이 있다. TDMRep은 웹사이트 단위 또는 개별 콘텐츠 단위로 권리 유보 여부를 기계가독형으로 표시하는 방식이며, IPTC는 이미지 파일 내부에 저작권자와 이용 조건 등의 정보를 기록하는 방식으로 데이터 마이닝 및 AI 학습 등에 대한 이용 허용·유보 여부를 표시하는 항목도 포함하고 있다. 이러한 방식은 개별 파일이나 메타데이터를 통해 opt-out 정보를 표시할 수 있다는 장점이 있으나, 파일 변환·재게시·스크래핑 등의 과정에서 관련 정보가 변경되거나 삭제될 수 있다.

마지막으로 이용조건 등을 문장으로 기재하는 일방적인 선언 방식의 opt-out 표시는 주로 사람이 읽는 형태로 작성되기 때문에, AI 개발자의 자동 수집 프로그램이 이를 일관되게 인식하지 못한다는 문제가 있다.

※ 옵트아웃 표시 수단

robots.txt	웹사이트 운영자가 사이트 최상위 경로에 두는 텍스트 파일로, 검색엔진·AI 크롤러 등 자동 수집 프로그램에 사이트 전체 또는 일부 경로에 대한 접근 허용 여부를 알리는 방식
TDMRep	웹사이트 단위 또는 개별 콘텐츠 단위로 권리 유보 여부를 기계가독형으로 표시하는 방식
IPTC	이미지 파일 내부에 저작권자·이용 조건 등의 정보를 기록하는 방식이며, 데이터 마이닝 및 AI 학습 등에 대한 이용 허용·유보 여부를 표시하는 항목도 포함

3. 등록·조회 시스템의 권고 모델

권고 모델에서 시스템의 역할은 옵트아웃 여부와 조회 기록 등 최소한의 정보를 제공하는 데 한정된다. 즉, 시스템은 저작권 귀속과 관련된 분쟁을 판단하지 않으며, 라이선스 시장을 운영하거나 저작물을 보관하지도 않는다. 또한 AI 개발자가 옵트아웃을 준수하는지를 강제하거나 감시하는 기능도 갖지 않는다. 시스템은 기존 옵트아웃 표시 수단을 대체하는 것이 아니라 이를 보완하는 성격을 지니며, 권리자의 자발적인 등록을 전제로 운영된다.

등록 절차를 보면, 권리자는 직접 또는 집중관리단체·출판사 등 분야별·국가별 등록 창구를 통해 신원과 위임 관계를 확인받은 후, 저작물의 식별자나 URL과 함께 옵트아웃 의사를 등록하게 된다. 이때 활용할 식별자로는 ISBN 등과 함께 국제 표준 콘텐츠 코드(International Standard Content Code, 이하 'ISCC')가 제시되었다. ISCC는 공개된 방식에 따라 권리자, 중개자, 개발자가 저작물의 내용으로부터 직접 생성할 수 있는 식별 코드이다.

AI 개발자는 데이터 수집 단계에 따라 서로 다른 방식으로 시스템을 조회하여 옵트아웃 정보를 확인할 수 있다. 데이터를 수집하기 전에는 수집하려는 사이트의 URL이나 도메인으로 조회하고, 수집한 이후에는 ISCC 등 콘텐츠 기반 식별자로 조회하는 방식이다.

이처럼 등록은 권리자가 직접 또는 분야별·국가별 등록 창구를 통해 하고, 조회는 중앙에서 일괄 처리하는 하이브리드 구조가 권고 모델로 제시되었다. 아울러 같은 저작물에 대해 상충하는 등록이 이루어지거나 권리자, AI 개발자 등으로부터 이의가 접수되는 경우에는 해당 저작물에 '분쟁 중' 표시를 하고 옵트아웃 상태로 임시 표시하도록 하였다.

4. 시스템 도입 방안 및 한계

분야마다 저작물 식별 체계의 성숙도와 권리 관계의 복잡성이 다르기 때문에, 시스템은 식별 체계가 이미 갖춰져 있고 권리 관계가 비교적 단순한 분야부터 도입하여 점차 확대하는 단계적 방식이 권고되었다. 구체적으로는 출판 콘텐츠 중심 분야에서 시작하여 시각 콘텐츠 분야, 음악·시청각 콘텐츠 분야 순으로 확대하는 방안이 제시되었다.

비용 측면에서는 중앙에서 운영하는 조회 기능은 공적 재정지원으로 운영하고, 분야별·국가별 등록 창구는 기존 기관의 예산을 활용하되 필요한 경우 초기 지원을 통해 운영하는 방안이 제시되었다. 또한 권리자의 등록과 AI 개발자의 조회는 원칙적으로 무료로 하되, 대량 이용이나 고도화된 서비스를 이용하는 AI 개발자에게는 이용량 등에 따라 비용을 부담하게 하는 방안도 함께 검토되었다.

등록·조회 시스템이 도입되면 옵트아웃 표시의 실효성과 투명성이 높아지고, 여러 방식으로 흩어져 있는 옵트아웃 표시를 하나의 통일된 형식으로 확인할 수 있게 될 것으로 보았다.

권리자 입장에서는 robots.txt, TDMRep, 이용조건 등 여러 방식으로 옵트아웃을 표시하고 관리해야 하는 부담이 줄어든다. 또한 기존 옵트아웃 표시 수단을 그대로 유지하면서, 시스템에서 통일된 기계가독형 형식으로 옵트아웃 의사를 표시할 수 있게 된다. AI 개발자는 데이터 수집 전에 URL·도메인 단위의 옵트아웃 표시를 확인하고, 수집 후에는 콘텐츠 기반 식별자를 활용하여 확보한 데이터에 옵트아웃 표시된 저작물이 포함되어 있는지 확인할 수 있다.

다만 보고서는 시스템만으로는 해결하기 어려운 한계도 함께 지적했다. 관할권 및 법 해석의 문제, 옵트아웃 표시 이전에 이루어진 TDM 행위에 소급하여 적용할 수 없다는 문제, 제3자로부터 확보한 데이터셋에 포함된 저작물의 옵트아웃 여부를 완전히 확인하기 어렵고, 개별 데이터의 출처와 권리관계 및 적법한 접근 및 이용허락 여부까지 검증할 수 없다는 문제 등은 시스템 도입만으로 해소되기 어렵다는 것이다.

향후 과제로 시스템 시행에 앞서 산업계와의 협력 체계를 갖추고, 운영 방식과 복잡한 사례에 대해 추가로 검토할 필요성이 언급되었다. 특히 이미 운영 중인 식별 체계나 권리 관리 방식을 해치지 않도록 초기 단계부터 분야별 이해관계자가 참여하는 실무협의체(워킹그룹)를 구성할 것이 권고되었다. 아울러 EU 기관, 위탁 운영기관, 공공-민간 파트너십 등 시스템을 누가 운영할지와 비용을 어떻게 회수할지에

대해서도 추가적인 협의가 필요하다. 이 밖에 2차적 저작물의 처리, 허위·무분별한 등록 등 악용 방지, 옵트아웃 이전에 수집된 데이터의 처리 문제 등에 대해서도 법적·기술적 검토가 더 이루어져야 할 것으로 보인다.

참고자료

- European Commission, Study to Assess the Feasibility of a Registry of Text and Data Mining Opt-outs expressed by Rightholders, 2026. 7. 13. (https://data.europa.eu/doi/10.2759/4959861)