

미국 AI 산출물 식별 및 이력정보에 관한 미국의 법안

고려대학교 법학전문대학원 교수
이대희

1. 법안의 제출

2024년 7월 11일, Cantwell 의원 등 3명의 미국 연방상원의원은 “COPIED Act”라는 명칭의 법안¹⁾을 제출하였음. 이 법안은 대체로 콘텐츠의 이력(provenance, 출처) 및 인공지능(AI) 산출물의 파악, AI 시스템 이용자가 콘텐츠 이력을 포함시킬 수 있도록 하는 것 등을 내용으로 하고 있음. 이 법안의 제안자인 Cantwell 상원의원은 AI 생성 콘텐츠에 대해서는 투명성이 상당히 필요하고, 이력 및 워터마크 절차에 의하여 기자, 예술가, 음악가 등 창작자들이 자신의 콘텐츠를 다시 통제할 수 있도록 할 것이라고 설명하고 있음.²⁾ 공동 제안자인 Blackburn 상원의원은 AI에 의하여 아무런 동의없이 모든 개인들을 대상으로, 이들의 모습을 유사하게 흉내내는 딥페이크(deepfake)가 만들어 지고 있고, 이런 콘텐츠를 통해 이익을 얻기도 하는 상황임을 언급하며, COPIED Act는 예술가와 실연자들과 같은 평범한 대상들을 딥페이크와 기타 부진정한 콘텐츠로부터 더욱 훌륭하게 방어할 수 있도록 하는 중요한 조치를 취한 것이라고 설명하고 있음.³⁾ 또 다른 공동 제안자인 Heinrich 상원의원은, COPIED Act에 의하여, AI가 생성하는 해롭고 기망적인 콘텐츠를 단속하고, 동의를 받지 않고 AI 콘텐츠를 생성하는 것으로부터 기자와 예술가들을 더욱 훌륭하게 보호하는 것을 지원하는데 매우 필요한 기술적인 도구가 제공되는 것이라고 설명하고 있음.⁴⁾

이와 같이 COPIED Act는 콘텐츠 제작자 내지 저작자를 보호하는 것을 내용으로 하고 있지만, 저작권자와 AI 개발자 간에 추를 저작권자 쪽으로 약간 이동시킬 수 있는 내용을 포함하고 있기 때문에 비판도 제기되고 있음.

2. 법안의 내용

COPIED Act는 첫째, 이 법안은 연방기관으로 하여금 AI 생성 콘텐츠 파악, 워터마크, 콘텐츠 이력정보 등에 대하여 연구·개발하거나 표준을 제정할 의무를 부과하는 것을 내용으로 하고 있음(§ 5).

1) 편집 및 딥페이크 미디어로부터의 콘텐츠 출처 보호 및 무결성(Content Origin Protection and Integrity from Edited and Deepfaked Media Act). S. 118th Congress 2nd Session. Cantwell, Blackburn, Heinrich.

2) US Senate Committee on Commerce, Science, & Transportation, Cantwell, Blackburn, Heinrich Introduce Legislation to Increase Transparency, Combat AI Deepfakes & Put Journalists, Artists & Songwriters Back in Control of Their Content (July 11, 2024).

3) Id.

4) Id.

둘째, 생성형 AI 툴(tool, AI 시스템이나 앱) 제공자는 사용자가 AI 산출물에 대한 (연방기관에 의하여 개발되거나 표준이 제정된) '콘텐츠 이력 정보'를 표시할 수 있도록 하여야 함. 그런데 이러한 의무는 AI 산출물에 대해서뿐만 아니라 AI 툴에 의하여 생성되거나 상당히 수정된 디지털 형태의 콘텐츠에 대해서도 적용됨(§ 6). 콘텐츠 이력정보(content provenance information)는 디지털 콘텐츠의 출처나 이력(history)을 기록한 정보로, '최신의 기술에 의한, 기계가 읽을 수 있는 정보'로 정의됨(§ 2).

셋째, 누구든지 불공정하거나 기망행위를 위하여 고의로 콘텐츠 이력정보를 제거, 수정, 변경, 무력화하는 것은 금지됨 [§ 6(b)(1)]. 플랫폼도 이력 정보를 제거 등을 함으로써 이용자가 이러한 정보에 접근할 수 없도록 하는 것도 금지되며, 다만 보안연구를 위하여서만 허용됨[§ 6(b)(2)]. 저작권 관리정보의 제거 등이 금지되는 것과 같이 이력정보의 제거 등도 금지하는 셈임.

넷째, 콘텐츠 이력 정보를 가진 저작물이나 이력 정보가 제거·분리되었다는 것을 알거나 알 수 있었던 저작물을, AI 또는 알고리즘을 사용하는 시스템을 훈련시키거나 AI 산출물을 생성·수정하기 위하여, 사용하는 것은 금지됨. 이에 해당하지 않기 위하여서는 해당 저작물을 소유하는 자의 동의를 얻거나 보수 지급 등 해당 저작물 이용에 관한 조건을 준수해야 함(6(c)).

다섯째, COPIED Act의 집행은 불공정하거나 기망적 행위를 규율하는 연방거래위원회(Federal Trade Commission, FTC)나 주 검찰총장에 의하여 이루어짐(§ 7).

3. 시사점

COPIED Act는 연방기관으로 하여금 콘텐츠 이력정보, AI 산출물 식별, 워터마크 등에 대한 연구개발이나 표준제정을 의무화시키면서, 이력정보가 부착된 저작물 이용을 금지시킴으로써, AI가 어떻게 작동하는지에 대하여 가시성을 제공하고, 학습에 대한 투명성을 제공하고, AI 산출물 파악을 용이하게 하고(§ 2), 궁극적으로 AI의 창작물 절취를 제어하는 것을 목적으로 함.

COPIED Act는 법안 제안자들이 설명하는 것처럼 저작권자를 보호하는 것을 내용으로 하고 있고, 이에 따라 음악산업 분야에서는 상당한 지지를 받고 있음.⁵⁾ AI 시스템 개발자는 시스템 사용자가 AI 시스템에 의하여 생성되거나 상당히 수정된 디지털 형태의 저작물 유형에 대해서도 콘텐츠 이력정보를 부착할 수 있도록 하여야 하고, 콘텐츠 이력정보를 가진 저작물 이용을 학습데이터로 사용하는 것이 금지됨.

5) Public Knowledge, The COPIED Act Is an End Run around Copyright Law.

COPIED Act가 적용되는 ‘콘텐츠(covered contents)’는 저작권법 규정(§ 102)상 저작물(work of authorship)의 디지털 형태로 정의되고(§ 2), AI 개발자가 사용자로 하여금 이력정보를 부착할 수 있도록 하는 이러한 저작물은 AI 틀에 의하여 생성되거나 상당히 수정된 것이라고 하면서[§ 6(a)(2)(A)], 학습데이터로 사용하는 것을 금지하는 것은 이력정보를 가지고 있는 저작물의 디지털 형태임[§ 6(c)]. 따라서 COPIED Act가 적용되는 콘텐츠는 저작물인 것처럼 규정하고, 이력정보를 부착할 수 있는 콘텐츠는 AI에 의하여 생성되거나 상당히 수정된 콘텐츠로 규정함으로써, 저작권법상 보호되는 저작물인 것으로 보이거나 AI에 의하여 생성 또는 상당히 수정된 콘텐츠인 것처럼 보이기도 함. COPIED Act의 제안자들이 설명한 것에 의하면, 이러한 콘텐츠는 저작물에 해당하는 것으로 보임. 따라서 COPIED Act에 대해서는 여러 비판이 행하여지고 있음.

미국에서 AI 저작권 소송이 많이 제기되어 있고 소송에서 공정이용 여부가 다투어지고 있지만, 저작물을 학습데이터로 이용하는 것이 공정이용인지 여부는 불확실함. 그런데 COPIED Act에 의하면, 저작권자는 자신의 콘텐츠에 이력정보를 부착할 수 있고, 이러한 이력정보가 부착된 콘텐츠는 학습데이터로 이용할 수 없게 됨. 따라서 학습데이터 이용이 공정이용에 해당하는지 불확실하거나, AI 개발자들이 공정이용이라고 주장하는 상황에서, 저작권자는 이력정보를 부착시킴으로써 자신의 콘텐츠 사용을 금지할 수 있게 되어 학습데이터 이용에 대한 공정이용을 부인하는 것이 됨. 이력정보가 부착된 콘텐츠를 이용하기 위하여서는 콘텐츠 소유자의 동의를 얻거나 사용료 지급 등 콘텐츠 이용조건을 준수하여야 함. 바로 여기에서 COPIED Act가 EU DSM 저작권 지침⁶⁾에서 텍스트-데이터마이닝(TDM) 예외에서 이용자가 권리유보를 한 것(opt-out)과 동일하다고 비판받고 있음.

EU DSM 지침의 TDM 예외규정에 의하면, 저작권자가 TDM 예외, 곧 학습데이터 이용을 유보하면 AI 개발자는 저작물을 이용할 수 없는데(§ 4), EU AI Act는 범용 AI(GPAI) 모델의 공급자로 하여금, EU의 저작권법, 특히 ‘디지털 단일 시장 저작권 지침(DSM 지침)’상의 TDM을 위한 예외규정의 요건인 저작권자의 권리 유보⁷⁾를 식별 준수하는 정책방향을 마련하도록 하는 의무를 부과하고 있음[§ 53.1(c)]. 결국 AI 개발자가 학습데이터를 이용함에 있어서 저작권자의 유보를 확인하여야 하고, 유보가 존재하지 않는다면 저작물을 학습데이터로 이용할 수 있음. 자신의 저작물이 학습데이터로 이용되는 것을 희망하지 않는 저작권자는 유보를 함으로써 학습데이터로 이용하지 못하도록 의사표시를 하는 것이 되고, 그럼에도 불구하고 이용하고자 한다면, 저작권자에게 이용허락을 받아야 함. 궁극적으로 저작물 이용에 대한 저작권자의 ‘유보’가 학습데이터 이용의 허용 여부를 좌우하는 것이 됨. COPIED Act는 저작권자가 자신의 저작물에 이력정보를 부착하는 것이 바로 유럽에서 권리가 자신의 저작물이 TDM 목적을 위하여 이용되는 것에 대하여 유보를 한 것에 해당한다고 비판되는 이유가 바로 여기에 있음. 유럽에서는 저작권의 일반적 제한사유인 공정이용이 존재하지 않지만, 공정이용이 존재하고 학습데이터 이용이 공정이용에 해당하는지 여부가 판가름나지

6) 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, PE/51/2019/REV/1.

7) DSM 지침은 TDM을 위한 예외로서, 저작물에 대하여 적법하게 접근하는 경우, TDM 목적을 위하여 필요한 한도에서 저작물의 복제와 추출을 허용하고 있는데, 다만 저작권자가 적절한 방식으로 이러한 저작물 이용을 명시적으로 유보한 경우에는 예외가 적용되지 않도록 규정하고 있다(§ 4).

않은 미국에서, 공정이용을 부인하는 듯한 COPIED Act에 대하여 저작권법을 회피하는 법안이라고 비판받는 이유가 여기에 있음.

AI와 관련된 저작권 쟁점은 기술적인 사항을 고려하여야만 해결이 가능함. 그런데 COPIED Act가 의도하는 바와 같이 이력정보나 워터마크, AI 산출물이나 AI를 이용한 수정 여부의 파악, 학습데이터의 이용 등에 가이드라인, 표준, 모범행위기준의 개발이 용이하지 않음. 전제요건이 충족되기 어려운 상황에서, 전제요건에 해당하는 이력정보를 콘텐츠에 부착하여 이를 보호하려는 것이 현실적으로 얼마나 이행가능한가 여부가 관건이 됨. 만약 한국에서도 COPIED Act와 같은 법안을 고려한다면 이러한 사항을 반드시 고려하여야 할 것임.

참고자료

US Senate Committee on Commerce, Science, & Transportation, Cantwell, Blackburn, Heinrich Introduce Legislation to Increase Transparency, Combat AI Deepfakes & Put Journalists, Artists & Songwriters Back in Control of Their Content (July 11, 2024)

Public Knowledge, The COPIED Act Is an End Run around Copyright Law.

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, PE/51/2019/REV/1.