JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 10, pp. 3089-3103, Oct. 2026

Check for updates

생성형 AI 저작권 침해의 구제수단으로서 머신 언러닝의 법적 지위에 관한 연구

신 민 지[1] · 노 승 민[2] · 이 미 영[3*]
[1]중앙대학교 융합보안학과 박사과정 [2]중앙대학교 산업보안학과 교수 [3]중앙대학교 융합보안학과 연구교수

Legal Status of Machine Unlearning as a Remedy for Copyright Infringement by Generative AI

Min-Ji Shin[1] · Seungmin Rho[2] · Mi Young Lee[3*]

[1]Ph.D. Student, Department of Security Convergence, Chung-Ang University, Seoul 06974, Korea
[2]Professor, Department of Industrial Security, Chung-Ang University, Seoul 06974, Korea
[3]Professor, Department of Security Convergence, Chung-Ang University, Seoul 06974, Korea

[요 약]

생성형 AI가 특정 저작물을 재현하는 수준에 이르렀으나, 저작권법 제123조는 학습된 모델에 대하여 적용가능한지 여부 또한 분명하게 판단하기 어려운 실정이다. 본 연구는 머신 언러닝을 제123조의 구제수단으로 활용할 경우 그 법적 지위와 적용 요건·절차·증거를 현행법 해석으로 검토하였다. 학습자·배포자·운영자·검색 저장소 관리자·이용자의 역할을 구별하고, 제2항의 "그 밖의 필요한 조치"를 제1항의 실효성을 확보하기 위한 부수조치로 넓게 해석하였다. 모델의 복제물성이 부정되어도 예방청구는 가능하고, 복제물성이 인정되어도 모델 전체의 폐기가 필연적이지 않다. 언러닝 명령은 결과금지를 핵심으로 하여 적극적 조치와 검증 절차로 구성되며, 폐기와 언러닝을 선택적 구제수단으로 제시하였다. 시험 통과는 위반이 관측되지 않았다는 증거일 뿐 이행 자체는 아니다. 끝으로 현행법이 뒷받침하는 부분, 해석상 제안, 합의 또는 입법이 필요한 과제를 구분하였다.

[Abstract]

Although generative AI models can regurgitate specific copyrighted works, Article 123 of the Korean Copyright Act does not contain an effective order against trained models. This study examines, under current Korean law, the legal status of machine unlearning as a remedy for this situation based on Article 123 and the requirements, procedure, and evidence for its application. Distinguishing the roles of trainer, distributor, operator, retrieval store manager, and user, it reads the "other necessary measures" of paragraph (2) broadly as ancillary measures securing paragraph (1). Prevention remains available even when the model is not a copy, and destruction of the entire model is not an outcome even when it is. An unlearning order consists of a result-based prohibition as its core, and it includes positive measures and verification procedures; the choice between destruction and unlearning is decided on a case-by-case basis. Passing a test only indicates that no violation was observed; it does not evaluate the performance itself. Finally, this study separates what current law supports from interpretive proposals and from matters requiring agreement or legislation.

색인어 : 저작권 침해, 집행, 생성형 인공지능, 금지청구, 머신 언러닝
Keyword : Copyright Infringement, Enforcement, Generative AI, Injunctive Relief, Machine Unlearning

http://dx.doi.org/10.9728/dcs.2026.27.10.3089

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 23 September 2026; **Revised** 12 October 2026
Accepted 12 October 2026

***Corresponding Author; Mi Young Lee**
Tel: ※ 개인정보 표시제한
E-mail: miylee@cau.ac.kr

Ⅰ. 서 론

현존하는 생성형 인공지능(AI; artificial intelligence, 이하 '생성형 AI')은 특정 저작물을 어렵지 않게 구현할 수 있다. 독일 뮌헨 지방법원은 2025년 11월 GEMA v. OpenAI에서 단순한 프롬프트만으로 가사가 재현된 사실을 근거로 그 가사가 모델 안에 기억되어 있다고 인정하고, 모델 안에서의 복제와 출력에서의 복제·공중이용제공을 각각 금지하였다[1]. 영국 고등법원은 Getty Images v. Stability AI에서 학습된 모델의 가중치가 저작물을 저장하거나 재현하지 않으므로 침해 복제물이 아니라고 판단하였다[2]. 두 판결의 전제와 평가는 달랐지만 권리자가 구하는 바는 동일하다. 즉, 해당 모델이 자신의 저작물을 다시 출력하지 못하도록 하는 것이다. 국내에서도 지상파 방송 3사가 2026년 2월 OpenAI를 상대로 기사의 복제·전송 금지와 학습 데이터셋의 삭제를 청구하였다고 보도되었다[3]. 모델이 남아 있는 한 같은 저작물이 다시 출력될 수 있고, 모델 전체의 폐기는 한 저작물의 침해를 이유로 다른 모든 이용에도 영향을 끼친다. 모델의 존치와 전면 폐기 사이에 간극이 있는 것이다. 최근에는 이 간극을 메울 중간 조치로 머신 언러닝(machine unlearning), 곧 학습된 모델에서 특정 데이터의 영향이나 특정 출력의 가능성을 사후에 제거하는 기술이 거론되지만, 그 적용은 조건부라는 한계가 있다[4].

우리 저작권법에서 이 청구의 근거가 되는 조항은 제123조다. 제1항은 침해의 정지와 예방을, 제2항은 그에 부수하여 침해행위에 의하여 만들어진 물건의 폐기와 "그 밖의 필요한 조치"를 규정한다[5]. 종래 이 조항은 눈에 보이는 침해 복제물을 전제로 운용되었으나, 권순재가 지적한 대로 학습의 결과로 만들어진 모델과 산출물을 곧바로 침해물이라고 단정하기는 어렵고, 이 괴리는 모델과 산출물이 제123조 청구의 대상이 될 수 있는지의 문제로 이어진다[6]. 나아가 이 조항을 생성형 AI에 실효적으로 적용하기는 어렵다. 그 원인은 다음 세 가지로 정리된다. 첫째, 무엇을 폐기하거나 제거할지 여부가 특정되지 않는다. 학습된 모델 안에서 저작물은 특정 파일로 존재하지 않는다. 둘째, 모델이 저작물의 복제물인지가 판결마다 갈리고, 그 판단이 서기 전에는 제2항의 조치가 어느 대상에 미치는지 정할 수 없다. 셋째, 조치가 이행되었는지를 확인할 수단이 없다. 파라미터를 열어 보아도 저작물의 존부를 읽어낼 수 없고, 시험 결과는 시험한 조건에 대해서만 의미를 가진다. 이로 인해 권리자는 청구권을 가지고도 실효적인 주문을 얻지 못한다. 이것이 본 연구가 다루는 보호의 공백이다.

이에 본 연구는 다음 두 물음에서 출발한다. 특정 저작물의 침해적 재현이 문제 되는 생성형 AI 모델에 대하여 머신 언러닝을 제123조의 구제수단으로 활용한다면, 그것은 제123조의 어느 청구에 어떤 자격으로 포섭되는가. 그리고 어떤 경우에 어떤 요건·절차·증거 아래 적용되는가. 앞의 물음은 언러닝의 법적 지위에 관한 것이고, 뒤의 물음은 그 적용 범위에 관한 것이다. 본 연구는 이 공백을 새로운 청구권으로 메우려 하지 않는다. 그 대신 제123조의 요건과 절차 안에서 언러닝이 어떤 조치로서 어느 범위까지 인정될 수 있는지를 밝히고자 한다. 이를 위해 제123조의 문언과 구조를 해석하고, 기술 문헌이 밝히는 언러닝의 보장 범위를 법적 요건과 증거의 관점에서 검토한다. 해외 판결과 논의는 명령 내용을 해석하는 기준으로 삼아 비교·분석한다.

논문의 구성은 다음과 같다. Ⅱ장에서는 생성형 AI를 통한 침해가 종래의 침해와 어떻게 다른지 살펴 보호의 공백을 확인한다. Ⅲ장에서는 머신 언러닝의 기술적 특징을 정리하고, 이를 토대로 제123조의 적용 가능성과 언러닝의 법적 지위를 도출한다. Ⅳ장에서는 해외 논의와 국내 선행연구를 비교법적으로 검토하고, Ⅴ장에서는 국내법상 적용 가능성과 법적 지위를 검토한 뒤 Ⅵ장에서 결론을 맺는다.

연구의 범위는 다음과 같이 한정한다. 본 연구는 학습된 모델이 저작물의 복제물인지를 먼저 확정하지 않는다. 대신 모델 안에 침해 복제물이 있다고 인정되는 경우와, 그렇지 않더라도 침해적 출력의 위험이 있는 경우를 나누어 청구권의 근거와 명령의 내용이 어떻게 달라지는지를 살핀다. 의거와 실질적 유사성의 증명은 별개의 논점이므로, 침해 또는 구체적 침해 우려가 인정된 이후의 문제만 다룬다. 학습 자체가 공정이용(제35조의5)에 해당하는지, 그리고 텍스트·데이터 마이닝(TDM; text and data mining) 면책을 어떻게 입법할 것인지도 다루지 않는다. 제35조의5는 구제명령이 적법한 이용까지 차단하지 않도록 하는 한계로서만 언급한다. 스타일 모방은 저작권의 보호 대상이 아니므로 논의에서 제외한다[7]. 제125조의 손해배상은 고의·과실과 이미 발생한 손해를 요건으로 하는 별개의 청구이다. 따라서 여기서는 사후의 언러닝이 이미 성립한 책임을 소멸시키지 않는다는 점을 지적하는 데 그친다[5].

Ⅱ. 생성형 AI와 침해: 특수성과 보호의 공백

2-1 생성형 AI의 기억과 재현

생성형 모델은 학습데이터를 데이터베이스처럼 저장하지 않는다. Cooper 등은 파라미터에서 특정 정보 조각을 깨끗하게 식별하여 삭제할 방법은 없다고 정리한다[4]. 그러나 모델이 학습데이터의 일부를 거의 그대로 재생산하는 현상, 곧 기억(memorization)은 실증되어 있다[8]. 뮌헨 지방법원은 학습데이터가 모델 안에 포함되어 출력으로 추출될 수 있는 현상을 기억으로 정의한 뒤, 학습데이터가 알려진 경우에는 단순한 프롬프트에 대한 출력과의 대조로, 알려지지 않은 경우에는 엔트로피·퍼플렉시티 분석으로 기억을 입증할 수 있다고 보았다[1]. 여기서 확인할 것은 재현의 원인이다. Wei 등

은 재현의 원인을 모델이 학습 과정에서 저작물을 기억하여 파라미터로부터 재생하는 경우와, 검색 증강 생성(RAG; retrieval-augmented generation)처럼 추론 시점에 외부 자료가 문맥으로 공급되어 모델이 그것을 옮겨 적는 경우로 구별하였다[9]. 여기에 이용자가 프롬프트에 저작물 원문을 직접 입력하는 경우를 더하면 원인은 셋이 된다. 세 원인은 같은 출력을 낳을 수 있지만 조치의 대상이 다르다. 기억에 의한 재현은 파라미터 또는 출력 계층에 대한 조치를, 검색 증강에 의한 재현은 외부 자료 저장소에 대한 조치를, 이용자 입력에 의한 재현은 입력 계층의 조치나 이용자 자신의 책임 문제로 이어진다. 기억의 존부는 저작물별로 시험을 통해서만 확인되며, 시험은 "재현되지 않았다"는 사실을 보여줄 뿐 "재현될 수 없다"는 보장을 주지 않는다.

1) 행위와 행위주체의 특수성

제123조는 제1항의 정지·예방청구, 제2항의 "침해행위에 의하여 만들어진 물건의 폐기나 그 밖의 필요한 조치", 제3항의 보전처분(정지, 물건의 압류 그 밖의 필요한 조치)으로 구성된다[5]. 이 가운데 제2항의 조치 청구는 제1항의 청구를 하는 경우에만 구할 수 있는 부수적 청구이다. 따라서 어떤 조치를 구하든 논증은 상대방의 침해행위 또는 침해 우려를 특정하는 데서 시작하여야 한다. 어느 조항이 적용되는지 역시 기술의 종류가 아닌 상대방의 행위를 기준으로 결정된다.

생성형 AI가 운용되는 과정에서는 시간 순서에 따라 세 가지 행위가 문제 된다. 먼저 학습 단계의 복제이다. 데이터셋을 수집·저장하고 학습을 위하여 적재하는 행위는 제2조 제22호의 복제에 해당하며[5],[10], 이 행위는 학습이 끝나면 완료된다. 완료된 행위에 대하여는 정지를 구할 수 없고, 같은 데이터로 추가 학습이 예정된 경우에 한하여 예방을 구할 수 있다. 이미 성립한 복제의 위법성이 사후의 언러닝으로 소급하여 달라지지도 않는다[11].

다음으로 모델 안의 복제물과 그 제공이 문제 된다. 모델에 특정 저작물이 기억되어 있고 그것이 복제물의 고정으로 평가된다면, 그 모델을 내려받게 하거나 서비스로 제공하는 행위는 복제물의 전송 또는 복제에 해당하고, 이것이 현재 진행중인 침해행위가 된다. 그러나 모델을 서버에 보유하고 있다는 상태만으로는 침해가 성립하지 않는다. 저작권법은 복제물의 소지를 배포 목적과 인식이 있는 경우에 한하여 침해로 간주(제124조 제1항 제2호, 제2조 제23호·제10호)하고 있기 때문이다[5]. 뮌헨 지방법원도 "모델 안에서의 복제"와 "파라미터 형태의 저장"을 특정하여 금지하였다[1],[12].

마지막으로 출력의 생성과 제공이다. 모델이 특정 저작물과 실질적으로 유사한 출력을 생성하여 이용자에게 표시하는 행위는 복제 또는 전송에 해당한다. 이 단계의 침해자가 누구인지에 관하여 뮌헨 지방법원은, 운영자가 출력에 대한 행위 지배(Tatherrschaft)를 가지므로 침해자가 되고, 이용자가 출력을 유발한 경우에는 행위 지배가 이용자에게 넘어갈 수 있으나 단순한 프롬프트의 경우에는 그렇지 않다고 판단하였다[1]. 본 연구는 국내법에서도 같은 구별을 적용하고자 한다. 물론 이는 독일 판결의 논증을 참고한 필자의 해석으로서, 이에 관하여 우리 판례가 확립한 귀속 기준은 아직 없다. 또한 단순한 프롬프트였다는 사실은 출력의 기술적 출처를 좁혀 줄 수는 있으나, 그것만으로 행위주체가 곧바로 확정되지는 않는다.

"침해할 우려"를 인정하기 위하여는 추상적 가능성만으로 부족하다. 그 저작물이 실제로 재현된 사실이나 재현될 개연성을 뒷받침하는 시험 결과가 제시되어야 한다. 뮌헨 지방법원이 사용한 출력 대조 방식이[1] 이러한 사실인정의 방법이 될 수 있다. 다만 권순재가 지적하듯 유사한 출력물을 찾아내는 일 자체가 쉽지 않으므로[6], 이 요건은 실무상 권리자에게 부담이 된다.

2) 모델의 복제물성

모델이 저작물의 복제물인지에 관하여 두 판결이 갈린다. 영국 고등법원은 Getty에서 가중치는 침해 복제물이 아니라고 판단하였고[2], 권순재도 학습된 모델의 내부 표현은 통계적 묘사에 불과하여 복제물이 보관되어 있다고 평가할 수 없다고 본다[13]. 반대로 뮌헨 지방법원은 기억된 가사가 파라미터로 분해되어 흩어져 있더라도 유형적 고정이 있다고 보았으며[1], Dornis와 Stober[14], Cooper와 Grimmelmann도[15] 기억된 데이터에 대하여 모델을 복제물로 본다.

본 연구는 이 판단의 차이를 어느 한쪽으로 확정하지 않는다. 다만 제2조 제22호가 "그 밖의 방법으로" 고정하거나 "다시 제작"하는 것을 복제에 포함하여 저장 방식을 한정하지 않는 이상[5], 법원이 특정 저작물의 재현 가능성이 입증된 범위에서 모델 안에 그 저작물의 복제물이 존속한다고 인정할 가능성은 열려 있다. 두 판결은 각각 그 소송의 전제 아래에서 내려진 판단이므로 어느 쪽도 일반 명제로 확장할 수 없으며, 복제물성은 저작물별로 성립한다. 어떤 저작물이 모델에 기억되어 있다는 사실은 복제물성을 인정할 사실적 기초를 제공할 뿐이며, 그 기억이 보호되는 표현의 유형물 고정 또는 재제작에 해당하는지는 별도의 법적 평가이다.

2-2 행위주체의 법적 지위

"제공자"는 하나의 주체가 아니다. 학습자, 모델 배포자, 추론 서비스 운영자, 검색 저장소 관리자, 이용자의 다섯 역할은 한 주체가 겸하기도 하고 서로 다른 주체에 나뉘기도 하며, 역할마다 문제 되는 행위와 이행할 수 있는 조치가 다르다(표 1).

법적 책임과 기술적 통제 가능성을 분리해야 한다. 어떤 주체가 조치를 이행할 수 있다는 사실이 그 주체의 침해책임을 대신하지 못하며, 침해책임이 있는 주체에게도 이행할 수 없는 조치를 명할 수는 없다. 운영자가 가중치를 수정할 수 없다면 그에게 명할 수 있는 것은 출력 계층의 조치와 서비스 중단이며, 파라미터 수준의 조치는 가중치를 보유한 배포자의 책임 근거가 별도로 갖추어진 경우 그에게 구해야 한다. 검색 증

표 1. 역할별 침해행위와 이행 가능한 조치
Table 1. Roles, infringing acts, and measures within each role's control

Role	Act potentially in issue	Basis under Art. 123	Measures within its control
Trainer	Reproduction at training; creation of the model	Completed act: damages; prevention only if further training is planned	Retraining, data-level removal, parameter-level unlearning, retention of training records
Model distributor	Persistence and provision of a copy inside the model (if copy status is found)	Cessation of provision; measures on the model copy it holds; updated release	Parameter-level measures; withdrawal or replacement of the distributed version; no control over copies already held by third parties
Inference service operator	Generation and display of infringing outputs to the public	Cessation or prevention of outputs	Output-layer filters, prompt-layer controls, service suspension; parameter changes only if it holds the weights
Retrieval store manager	Supply of the work to the model at inference; lawful possession of the document is not itself infringement	Cessation or prevention of outputs so caused, only where its own act and a basis of liability are shown	Deletion of the document, or blocking of retrieval to the service concerned; no effect on memorization
User	Inducing reproduction by inputting the work	Direct infringement by the user; provider's liability separate	Not an addressee of measures on the model

강 자료에서 표현이 공급된 사안이라면 모델 단독 시험에서 재현이 없음이 확인되는 한 저장소의 자료에 대한 조치로 족하다. 침해행위를 직접 하지 않는 주체가 "침해하는 자"에 들어오는지는 방조의 문제다. 대법원은 P2P (peer-to-peer) 서비스 제공자의 과실 방조를 인정하였고[16], 하급심은 금지조치를 취하지 아니한 방조자를 제123조 제1항의 "침해하는 자"에 포함시켰다[17]. 또한 대법원은 온라인서비스제공자의 조치의무를 침해의 명백성, 구체적 인식 가능성, 기술적·경제적 관리 가능성이 갖추어진 경우에만 인정한다[18],[19]. 이 법리는 이용자가 원문을 입력하여 재현을 유도한 경우 제공자의 조치의무 요건을 정하는 출발점이 된다. 단순한 프롬프트에 대하여 모델이 기억된 저작물을 재현하는 사안에서 제공자가 복제·전송의 주체로 인정된다면 방조 법리를 거칠 필요가 없으나, 그 인정 자체는 2-1절 1)에서 본 대로 행위주체에 관한 별도의 판단이다. 모든 AI 제공자가 방조자라는 결론도, 모두 직접침해자라는 결론도 이 판결례들에서 도출되지 않는다.

2-3 소결

생성형 AI를 통한 침해는 종래의 침해와 다음 세 가지 점에서 다르다. 침해 복제물이 특정 파일로 존재하지 않고 그 존부는 저작물별 시험으로만 확인된다. 학습단계의 복제는 완료된 행위이고 보유 상태 자체는 침해가 아니어서, 현재의 침해는 제공·전송·출력이라는 행위이며 그 주체는 역할에 따라 갈린다. 같은 출력이라도 원인이 기억인지, 검색 증강인지, 이용자 입력인지에 따라 조치의 대상과 상대방이 달라진다. 이 때문에 침해 복제물의 폐기라는 종래의 주문은 모델에 바로 옮겨지지 않으며, 폐기할 물건이 무엇인지, 폐기가 침해를 끝내는지, 폐기 아닌 조치는 무엇을 근거로 명하는지가 각각 문제 된다. 이 세 물음이 언러닝을 제123조의 구제수단으로 볼 때 답해야 할 물음이며, 그 답은 언러닝 기술이 무엇을 보장하고 무엇을 보장하지 못하는지를 확인한 뒤에야 제123조의 해석으로 얻을 수 있다.

Ⅲ. 구제수단으로서 머신 언러닝의 적용

3-1 머신 언러닝의 기술적 특징

1) 제거 목적과 개입 지점

언러닝이라는 이름 아래에는 목적이 다른 두 기술군이 있다. 하나는 특정 학습데이터의 영향을 파라미터에서 제거하는 것이고, 다른 하나는 모델이 특정 출력을 내지 못하게 억제하는 것이다[4]. 파라미터를 바꾸었다는 사실이 영향의 제거를 보증하지 않으므로 목적 축에 개입 지점 축(데이터와 학습 과정, 모델 파라미터, 추론과 출력 계층)을 더해야 한다. 영향 제거의 기준은 삭제 대상을 뺀 데이터로 다시 학습한 모델이다[11]. 그러나 재학습은 대형 모델에서 사실상 불가능하고[4], 분할 재학습은 학습 설계 단계부터 예정되어야 한다[4],[20]. 근사 언러닝을 생성형 모델에 적용한 결과는 실험적 성과로 보고되는 데 그친다[4],[21]. 출력 억제 가운데 개념 소거는 개념의 이름만으로 가중치를 미세조정하여 그 개념의 생성을 억제할 뿐[22] 학습데이터의 영향이 제거되었음을 보이지는 않으며, 시스템 프롬프트·디코딩 시 필터·출력 대조 필터는 추론·출력 계층에 개입할 뿐 모델 자체를 바꾸지 않는다. Cooper 등은 이 방법들이 특정 콘텐츠의 생성 가능성을 낮출 뿐 그것이 결코 생성되지 않는다고 보장하지는 않는다고 정리한다[4]. 표 2는 두 축으로 방법을 배열하였다.

두 기술군은 서로를 대신할 수 없다. 출력을 억제하더라도 정보는 여전히 파라미터에 남는다. 반대로 재학습을 하더라도

표 2. 목적과 개입 지점에 따른 언러닝 방법의 분류
Table 2. Unlearning methods by purpose and point of intervention

Method	Purpose	Point of intervention	Provable result	Residual risk
Retraining without the target data	Influence removal	Data and training process	Contribution of the target data is zero by construction	Cost at scale; generalization may still yield similar outputs
Sharded retraining(SISA)	Influence removal	Data and training process(designed in advance)	Zero contribution for the retrained shard	Not applicable post hoc to a model not so designed
Approximate unlearning	Influence removal	Model parameters	Proximity to the retrained model under stated assumptions; otherwise experimental	Definition contested; guarantees mostly for classifiers; concept resurgence after later fine-tuning
Concept erasure by weight fine-tuning	Output suppression	Model parameters	Reduced generation of the concept in tests	Information remains in parameters; resurgence; scope may exceed the work
Prompt- or decoding-level filters	Output suppression	Inference and output layer	Blocked outputs under tested conditions	Bypass; no effect on the model; may block lawful outputs
Retrieval store deletion	Removal of external source	Data(retrieval store)	The document is no longer retrievable	Says nothing about memorization inside the model

모델은 학습데이터를 넘어 일반화하기 때문에, 유사한 출력을 결코 내지 않는다고 보장할 수 없다[4]. 실험 결과도 이를 뒷받침한다. 개념을 소거한 뒤 무관한 이미지로 미세조정하기만 해도 소거된 개념이 되살아나는 개념 재출현이 보고되었고[23], 언러닝된 지식이 추론 단계에서 문맥 내 학습을 통하여 다시 들어올 수 있다는 점도 확인되었다[24]. 그러므로 두 기술군의 구별은 조치의 후보를 분류하는 데 의미가 있다. 어느 기술이 제123조의 어느 항에 대응하는지는 청구권의 근거와 명령의 내용을 따져 정하여야 한다.

2) 저작물과 제거 집합, 검증의 한계

법이 보호하는 것은 저작물의 표현이고 기술이 다루는 것은 데이터 샘플과 파라미터다. 따라서 언러닝 명령이 무엇을 제거하는지는 다음 네 요소(이하 '특정 4요소')로 특정된다. ① 권리 범위(어느 저작물의 어느 권리), ② 보호되는 표현(사실·아이디어·스타일은 제외)[7], ③ 제거 데이터 집합(그 표현을 담은 샘플과 중복 수록본·변형본의 범위), ④ 검증 출력 조건(어떤 프롬프트 조건에서 어떤 출력이 나오면 침해적 재현으로 볼 것인가)이 그것이다. 기술적 제거 집합은 법적 보호 범위보다 넓어질 수도 좁아질 수도 있다. 특정 샘플을 제거하더라도 같은 정보가 다른 샘플에 남아 있으면 모델은 여전히 그 정보를 생성할 수 있고[4], 작가명이나 장르명 단위로 개념을 소거하면 보호되지 않는 스타일과 사실까지 함께 억제된다[4]. 제거 집합의 넓이는 보호되는 표현을 기준으로 정당화되어야 하며, 특정 문자열이 출력되었다는 사실만으로 모든 맥락의 출력을 위법으로 처리하는 검증 체계는 인용·비평·보도처럼 제35조의5로 허용될 이용까지 차단하므로, 그러한 검증 체계 아래에서는 명령의 범위가 침해의 범위를 넘게 된다[5],[9]. "제거"가 의미하는 바는 "저작물 전부를 잊게 하는 것"이 아니라 "특정 저작물의 보호되는 표현이 특정 조건에서 침해적으로 이용되는 것을 막는 것"이며, 위의 특정 4요소는 명령의 대상을 특정하는 항목이자, 대상 버전과 함께 이행 여부를 판단하는 시험 조건의 항목이 된다.

언러닝이 법적 조치로 인정되려면 그 이행 여부를 제3자가 확인할 수 있어야 한다. 그런데 Thudi 등에 따르면 언러닝은 알고리즘 수준에서만 정의되므로, 감사 가능한 형태로 주장할 수 있는 것은 외부 검증을 허용하도록 설계된 알고리즘을 사용하였다는 사실에 그친다[25]. 벤치마크 연구에서도 모든 지표에서 우수한 방법은 없었고, 언러닝된 내용이 다시 생성되지 않는다고 보장할 수도 없었다[9]. 더욱이 기술 문헌의 평가 축이 그대로 저작권 침해의 요건이 되는 것도 아니다[26]. 이러한 점을 종합하면 법적으로 활용할 수 있는 검증의 내용은 상당히 제한된다. 우선 어떤 알고리즘을 어떤 절차로 어떤 데이터에 적용하였는지를 확인할 수 있다. 다음으로 삭제 대상이 여전히 추출되거나 추론되는지를 시험할 수 있다. 다만 그 시험 결과는 시험한 조건에서 재현되지 않았다는 관측에 머물며, 앞으로도 재현될 수 없다는 보장으로 볼 수는 없다.

3-2 제123조의 적용 가능성

1) "그 밖의 필요한 조치"의 법적 범위

제2항의 "침해행위에 의하여 만들어진 물건의 폐기나 그 밖의 필요한 조치"는 두 가지로 읽을 수 있다. 협의의 해석은 "그 밖의 필요한 조치"를 폐기와 같은 성격의 조치, 곧 침해행위에 의하여 만들어진 물건에 대한 다른 처분으로 한정하고, 광의의 해석은 이를 제1항 청구의 실효성을 확보하기 위한 부수조치 일반으로 넓혀, 정지·예방의 목적에 필요한 작위·부작위까지 포함시킨다. 이 둘은 확립된 학설의 대립이 아니라 필자가 검토하는 해석의 대안이다. 체계적으로 보면 광의

의 해석이 타당하다. 제2항의 조치 청구는 제1항의 정지·예방을 실효적으로 만드는 부수 청구이므로 그 목적에 필요한 조치라면 대상이 물건이 아니어도 배제할 이유가 없고, 제3항이 압류와 성격이 다른 조치를 병렬로 둔 것도 같은 취지로 이해된다. 하급심 판결도 제2항의 조치로 서비스 제공의 중지를 명하거나[17] 서비스·프로그램 전체의 금지를 필요성·비례성을 이유로 배척하는 등[27],[28], 제2항의 조치를 물건의 처분에 한정하지 않고 행위의 금지까지 포함하는 것으로 해석하면서도 그 범위를 필요성과 비례성으로 제한하고 있다. 제1항은 행위에 대한 금지이고 제2항은 물건 또는 상태에 대한 조치라는 서술은 두 항의 대표적 기능을 말한 것이지 엄격한 구분이 아니다.

본 연구는 광의의 해석을 취하나, 협의로 해석하더라도 다음 두 명제는 독립적으로 성립한다. 첫째, 복제물성이 부정되어도 예방청구는 가능하다. 출력에 의한 침해는 복제물성과 무관하게 성립하고, 예방에 필요한 조치는 광의로 해석하면 제2항으로, 협의로 해석하더라도 예방청구의 내용 자체로 구할 수 있기 때문이다. 둘째, 복제물성이 인정되어도 폐기가 필연적이지 않다. 제2항은 폐기와 그 밖의 조치를 병렬로 두므로 어느 조치가 "필요한" 것인지는 별도로 판단된다.

2) 물건성·복제물성·관련성

제2항의 "침해행위에 의하여 만들어진 물건"은 물건일 것(물건성), 그 안에 저작물의 복제물이 있을 것(복제물성), 침해행위에 의하여 만들어졌을 것(관련성)의 세 요건으로 나뉜다. 학습단계의 복제가 침해라는 사실은 관련성을 채우지만 복제물성까지 채우지는 않는다. 침해 학습을 거쳐 만들어진 모델이라도 그 안에 특정 저작물의 복제물이 있는지는 저작물별 기억의 입증으로만 확인되는 별개의 사실이며, 두 사실을 같은 것으로 취급하면 "침해 학습으로 만들어졌으므로 모델 전체가 침해물"이라는 논리의 비약이 생긴다. 이 세 요건은 필자가 제2항의 문언을 분설한 것으로서 판례가 이를 요건으로 명시한 바는 없다. 세 요건이 요구되는 범위도 한정된다. "침해행위에 의하여 만들어진 물건"은 폐기의 대상을 정하는 말이므로 폐기와 물건에 대한 제거 조치에는 세 요건이 모두 요구되지만, 광의의 해석에서 제1항에 부수하는 "그 밖의 필요한 조치"는 물건을 전제하지 않으므로 물건성과 복제물성이 요구되지 않으며, 그 대신 정지·예방에 필요한 조치인지에 대한 심사가 따른다.

서버에 기록된 수치 데이터인 모델이 민법 제98조의 물건에 해당하는지에 관하여[29] 저작권법은 별도의 규정을 두고 있지 않다. 하급심은 디지털 파일의 폐기·삭제를 명한 바 있으나(다만 데이터베이스제작자권·영업비밀 사건은 저작권 침해를 근거로 한 것이 아니다)[30]-[32], 저장되었다는 이유만으로 데이터 자체가 물건이 된다고 단정할 수도 없다. 폐기가 침해 복제물의 존속을 끝내는 조치라는 목적에서 출발하면 조치의 대상은 데이터 자체도 저장매체 자체도 아닌 특정 기록 상태, 곧 상대방이 관리하는 특정 서버에 기록된 특정 버전의 모델 사본이다. 관건은 유형·무형이 아니라 그 기록 상태 안에 복제물이 존재하는지 여부이며[2], 서울중앙지방법원이 "게시판 기타 유·무형물"에서의 삭제 청구를 대상이 특정되지 않았다는 이유로 각하한 것은[32] 이 요건이 실무에서 어떻게 작동하는지를 보여준다.

3) 언러닝의 법적 지위

이상의 검토에 따르면 언러닝은 원칙적으로 제123조 제1항의 정지·예방청구를 이행하는 수단이다. 다만 모델 안에 복제물이 인정되는 범위에서는 제1항 청구에 부수하는 제2항의 제거 조치가 될 수도 있으며, 그 이행 여부는 결국 증거로 판단할 문제이다. 이러한 정리는 특정 기술을 특정 조항에 대응시키려는 것이 아니다. 언러닝이 구제수단으로 발령될 경우 명령에 담길 내용을 나누고, 각각의 법적 근거를 확인하려는 것이다. 우선 특정한 침해적 출력이나 제공을 금지하는 의무는 제1항의 정지·예방청구 그 자체이며, 어떤 수단을 쓸지는 상대방이 선택한다(가). 다음으로 특정 기록 상태에 대하여 적극적 조치를 취할 의무는 제2항의 조치이거나, 제1항을 넓게 해석할 경우의 예방조치에 해당한다(나). 끝으로 자료 제출과 검증에 관한 의무는 소송 절차상의 것과 판결 후의 보고로 나누어 볼 수 있다(다). 이들이 각각 독립한 청구권이라는 뜻은 아니며, 문제는 각 의무를 어떤 근거에서 어느 범위까지 명령에 포함할 수 있는지에 있다.

언러닝을 구제수단으로 보는 경우 그 핵심은 (가)의 결과금지이다. 그 이유는 다음과 같다. 첫째, 제1항의 문언은 정지와 예방이라는 결과를 말할 뿐 수단을 지정하지 않으므로 특정 알고리즘의 지정은 제1항이 예정한 내용이 아니다. 둘째, 어느 기술이 그 결과를 달성하는지는 방법·모델·시점마다 다르므로, 특정 알고리즘을 지정하면 그 알고리즘의 한계가 곧 명령의 한계가 된다. 셋째, 수단의 선택을 상대방에게 맡기면 상대방은 자신의 자원과 능력에 가장 적합한 조치를 고를 수 있다. 유럽사법재판소는 UPC Telekabel에서 결과만을 정한 금지명령의 경우 상대방이 가장 적합한 조치를 선택할 수 있다고 보았고[33], 그 조치는 적법한 정보 접근을 침해하지 않도록 엄격히 한정(strictly targeted)되어야 하며[33], 그 적정성을 제재 전에 법원에서 다툴 수 있어야 하고[33], 완전한 차단이 불가능하더라도 무단 접근을 심각하게 단념시키는 효과가 있으면 족하다고 하였다[33]. 접속중개자에 관한 면책 부분은 직접침해자인 모델 제공자에게 옮길 수 없으나[33], 명령 내용에 관한 이 네 기준(이하 'Telekabel 기준')은 상대방의 책임 구조와 무관하게 제1항 해석의 기준이 된다.

결과금지만으로 부족한 경우에 (나)의 적극적 조치명령이 더해지며, 그러한 경우는 두 가지이다. 하나는 침해행위로 만들어진 모델 안에 특정 저작물의 복제물이 있다고 인정되고 그 제공·전송에 대한 정지청구 또는 추가 학습·출력에 대한 예방청구가 성립하는 경우다. 이때 제거는 제1항 청구에 부수하는 제2항의 조치이며, 결과금지만으로는 복제물이 남아 위험이 계속되므로 그 필요성이 인정될 수 있다. 다만 복제물

성이 인정되었다는 사실만으로 제거의 필요성과 범위가 정해지는 것은 아니며, 그 필요성과 범위는 조치의 효과와 부담을 사안별로 비교하여 정해진다. 다른 하나는 결과금지가 반복하여 위반되거나 선택된 수단이 명백히 부적합한 경우다. 어느 경우든 적극적 조치명령은 결과금지에 더해지는 것이지 대체하는 것이 아니며, 특정 알고리즘의 실시를 명하는 것은 추가 사유가 있을 때로 한정된다. 이는 본 연구의 제안이지 판례가 확정한 우선순위가 아니다. 정리하면, 같은 저작물에 대하여 제1항의 결과금지와 제2항의 부수조치는 병행하며, 차이는 복제물성의 인정 여부에 따라 제2항 조치의 대상이 존재하는지에서 생긴다. 복제물이 인정되면 제공·전송의 정지와 사본에 대한 부수적 제거가, 복제물성이 부정되거나 입증되지 않으면 예방청구와 그에 필요한 조치가, 검색 저장소가 원인이면 그 행위·책임이 제시된 당사자에 대한 저장소 조치가, 이용자 입력이 원인이면 이용자의 직접침해와 제공자의 조치의무 요건이 각각 대응한다. 어느 기술로 이를 이행할지는 이 병행 구조 안에서 필요한 범위를 넘지 않는 것을 고르는 문제다.

Ⅳ. 비교법적 검토

4-1 EU 집행지침과 유럽사법재판소

유럽연합(EU; European Union) 지식재산권 집행지침은 제10조의 시정조치(회수·배제·폐기, 제3항의 비례성과 제3자 이익 고려)와 제11조의 금지명령을 별개 조문에 둔다[34]. Rubí-Puig와 Wolsing은 언러닝이 모델 안의 침해 요소를 수정·제거하는 한 제10조 시정조치의 디지털 대응물로 볼 수 있고, 특정 결과를 기술적으로 방지하도록 명하는 한 제11조 금지명령이 적합한 것으로 보인다[35]. 시정 대상과 금지행위를 구별하면서 두 조치를 결합할 수 있다는 점은 제123조 제1항·제2항의 관계를 읽는 비교 기준이며, 제1항의 결과금지와 제2항의 부수조치를 병행시킨 구조는 집행지침의 이 구별과 궤를 같이한다.

유럽사법재판소는 결과만을 정하고 수단을 맡기는 명령의 내용 요소를 UPC Telekabel에서 제시하였고 그 명령이 넘어서는 안 되는 경계를 Scarlet Extended에서 정하였다. 그 내용은, 모든 통신을 무기한으로 감시하고 아직 창작되지 않은 저작물까지 대상으로 하는 필터링 시스템의 설치 명령은 영업의 자유와 정보의 자유를 침해하여 기본권 사이의 공정한 균형을 존중하지 못한다는 것이다[36]. 여기서 배제된 것은 모든 필터링이 아니라 무차별·무기한·예방적 필터링 시스템이다. 이를 언러닝 명령에 적용하면, 특정 저작물의 특정 표현에 대한 한정된 검증과 조치는 이 문제에 걸리지 않지만, 권리자의 모든 저작물이나 장래의 저작물까지 상시 감시하도록 명하거나 특정 문자열의 출현을 맥락과 무관하게 차단하도록 명하면 제35조의5로 허용될 이용까지 막게 되어 같은 문제가 생긴다. 국내 실무도 같은 구별을 한다. 서울고등법원은 적법한 파일을 걸러 낼 절차가 있으면 적극적 필터링이 논리상으로 적법한 유통까지 금지하는 것은 아니라고 하였다[17]. 조치의 방식이 아니라 적법한 이용을 걸러 낼 절차가 갖추어졌는지가 기준이며, 언러닝 명령에서 그 절차는 검증 출력 조건에 해당한다. 이 축은 명령의 범위를 제한하는 원리이지 사업자가 특정 이용을 가능하게 하는 기능을 유지하여야 할 일반적 의무의 근거는 아니며, 본 연구는 이용자에게 새로운 청구권을 인정하려는 것이 아니다.

4-2 미국의 알고리즘 파기와 그 한정

모델의 폐기는 미국에서 “알고리즘 파기”라는 이름으로 먼저 논의되었다. Li는 연방거래위원회가 모델 삭제를 명한 사례를 근거로, 알고리즘 파기를 부당이득 반환과 독수독과 법리에 기초한 실효적 구제수단으로 평가하였다[37]. 이에 대하여 Samuelson은 생성형 AI 저작권 소송 16건 가운데 4건이 모델의 파기를 청구하였다고 정리하면서, 배포된 사본을 모두 파기하기는 어렵고 모델은 일반적으로 저작물의 인식 가능한 표현을 담고 있지 않다는 한계를 지적하였다[38]. Rubí-Puig와 Wolsing도 비슷한 문제의식을 보인다. 이들은 알고리즘 파기를 개별 청구인에게 허용하면 소량의 침해를 지렛대로 과도한 합의를 강요하는 홀드업 문제가 생길 수 있다고 본다. 그리하여 알고리즘 파기는 체계적·대규모 침해에 한정하여야 하고, 출력 단계의 침해에는 적합하지 않다고 한다[35]. 이러한 논의는 모델 전체의 폐기를 체계적·대규모 침해로 한정하고 개별 저작물의 침해에는 보다 좁은 조치를 우선하여야 한다는 점에서, 제123조 제2항의 “필요한” 조치를 판단할 때 참고가 된다.

4-3 국내 선행연구와 시사점

국내 문헌은 주로 침해의 성립(의거 요건, 요건사실의 증명책임, 산출물에 의한 침해 유형)[39]-[41]과 사전 면책(TDM 면책 규정의 설계)을 다루어 왔다. 기술적 보호조치의 관점에서는 김종성 등이 학습 불가능한 예제(unlearnable examples)를 중심으로 학습 단계에서 저작물을 보호하는 사전적 조치의 실효성을 검증하고 이를 옵트아웃과 연계한 보호 틀을 제안하였다[42]. 본 연구는 이와 달리 학습이 이미 끝난 모델에 대한 사후적 구제를 다루며, 사전적 보호조치가 작동하지 않았거나 우회된 뒤에 비로소 문제 되는 단계를 대상으로 한다는 점에서 그 연구와 보완 관계에 있다. 본 연구가 검색한 범위(한국학술지인용색인(KCI; Korea Citation Index) 등재지, 2023년 이후)에서 구제 단계를 정면으로 다룬 연구는 권순재의 연구 외에는 확인되지 않으며, 그 역시 이 문제를 “향후 사법적 판단이 필요한 미해결 문제”로 남겨두었다[6]. 인접 분야로는 대법원 2025마6304 결정[43]을

계기로 한 특허 분야의 연구가 있다. 권보원은 특허침해 금지 명령의 특정성과 집행 회피 사이의 긴장, 그리고 사후 관리형 집행을 논하였고[44], 류시원은 변경된 물건에 대한 집행에서 동일성 판단을 분석하였다[45]. 본 연구는 이들 연구를 후속 버전의 동일성 문제에서 참고하되, 저작권법상 현행 감독권한의 근거로 원용하지는 않는다.

해외에서는 Cooper 등과 Marino 등이 언러닝이 정책과 규제가 기대하는 역할을 하지 못한다는 점을 기술 문헌의 관점에서 경고하였다[4],[11]. 구제수단론의 관점에서 본 연구와 가장 가까운 것은 Rubí-Puig와 Wolsing의 연구로, 이들은 사건별 분석과 조치의 결합을 결론으로 제시하였다[35]. 본 연구는 권순재가 제기한 물음을 이어받는다. 다만 모델의 복제물성에 관한 그의 부정적 평가를 확정적 전제로 삼지 않고 하나의 조건으로 다루며, Rubí-Puig와 Wolsing의 사건별 분석을 한국법의 조문·절차·집행 단계에 맞추어 구체화한다는 점에서 선행연구와 구별된다.

비교법적 검토에서 얻을 수 있는 것은 개별 결론보다 논증의 방식이다. 시정 대상과 금지행위를 구별하면서도 두 조치를 결합하는 방식, 결과를 정하고 수단은 상대방에게 맡기되 적법한 이용을 보호하는 방식, 폐기를 좁게 한정하는 방식, 그리고 저작물별·조건별로 사실을 확인하는 방식이 그러하다. Ⅴ장에서는 이러한 방식을 제123조의 조문·절차·집행 단계에 적용하여 검토한다.

Ⅴ. 국내법상 적용 가능성과 법적 지위 검토

5-1 국내법상 적용 가능성 검토

1) 조치의 선택과 비례성

명령의 내용이 정해지면 그것을 이행할 수 있는 기술은 여러 가지가 있으며, 그 가운데 필요한 범위를 넘지 않는 것이 비례성에 따라 선택된다. 결과금지가 명령의 전부라면 출력 계층의 조치, 파라미터 수준의 억제, 영향 제거가 모두 후보이고 선택은 상대방에게 맡겨진다. 모델 안 복제물의 제거를 명한 경우 출력 계층의 조치는 파라미터에 정보가 남으므로 그 자체로는 적합하지 않다[4]. 검색 저장소가 원인인 사안에서도 기억과 검색 증강 양쪽에 적용되는 출력 단계의 방법이 있으므로 출력 계층의 조치가 적합성을 잃지는 않는다[9]. 어느 후보든 재현이 다시 나타날 수 있으므로[4],[23] 어떤 조치가 적합한지는 기술의 종류만으로 정해지지 않고 이행 증거(시험 결과와 부작용)와 함께 판단된다.

제2항이 먼저 규정한 조치는 폐기이다. 그러나 언러닝이 검증되었다고 하여 폐기청구가 곧바로 배제되는 것은 아니며, 반대로 언러닝이 실패하였다고 하여 폐기가 당연히 정당화되는 것도 아니다. 어느 조치가 "필요한" 것인지는 사안마다 각 조치의 효과와 부담을 비교하여 정하여야 한다. 하급심은 서비스 자체의 금지나 프로그램 배포의 금지가 필요한 범위를 넘는다고 보아 침해 파일 단위로 조치를 좁힌 바 있다[27],[28]. 이러한 판단은 모델 전체의 폐기와 특정 저작물에 대한 조치 사이의 선택에도 그대로 적용될 수 있으며, 알고리즘 파기를 대규모·체계적 침해로 한정한 논의(4-2절)도 같은 취지이다.

비교할 때 고려할 요소로는 침해의 범위, 반복성과 확산 위험, 서비스의 구조, 대체 수단의 효과, 권리자에게 남는 손해, 조치의 비용, 적법한 기능·표현의 손실[9], 제3자 의존(집행지침 제10조 제3항, 전문 24)[34]을 들 수 있다. 침해의 범위와 관련하여서는 한 저작물이라도 그 재현이 서비스의 핵심 기능이라면 비중이 커진다. 서비스의 구조에서는 블랙박스 모델인지 공개 가중치 모델인지, 검색 증강과 이용자 입력이 개입하는지를 살펴야 한다. 제3자 의존에는 배포된 사본, 파생 모델, 도입 사업자에 대한 영향이 포함된다. 이들 요소는 침해 방지의 효과와 권리 제한의 정도를 견주어 보는 기준이 된다. 덜 침해적인 조치로 같은 효과를 거둘 수 있다면 그 조치가 필요한 조치이다. 기능 손실이 폐기와 비슷한 정도라면 언러닝을 폐기보다 덜 침해적인 조치라고 보기 어렵다. 반대로 조치 후에도 권리자에게 남는 손해가 크다면 더 강한 조치의 필요성이 인정될 것이다.

권리자 측에서는 출력 억제가 파라미터에 정보를 남긴다고 반론할 수 있으며, 이 지적은 타당하다. 그러나 필터만으로 명령이 충족된다는 논증은 애초에 제거를 명하지 않고 결과금지만 발령된 사안에 한정되는 것이며, 복제물성이 인정되더라도 제거 명령이 없으면 마찬가지다. 언러닝이 재현 불가를 보장하지 않는다는 반론에 대하여는 폐기가 아니라 5-1절 4)의 재검증 의무와 재발 시의 재청구로 대응할 수 있다. 국내 실무도 상대방이 스스로 삭제·차단하였다는 사정만으로 조치의 필요를 부정하지 않는다[27],[30]. 조치의 비용은 여러 고려요소 가운데 하나일 뿐이며, 침해 상태를 해소하는 비용은 원칙적으로 침해자인 상대방이 부담한다(제10조 제2항)[34].

2) 명령의 특정, 잠정 조치, 후속 버전

특정되지 않은 명령은 판결 주문의 특정성을 갖추지 못하여 집행할 수 없으므로, 언러닝이 구제수단으로 적용되려면 저작물, 모델과 버전, 서비스 구성, 관리 범위가 특정되어 있어야 한다[32]. 저작물은 3-1절 2)의 네 요소로 특정하며, 뮌헨 지방법원이 금지의 대상을 별지 목록의 가사로 한정한 것이 그 실무 형태다[1]. 모델과 버전의 특정은 명령의 층위에 따라 다르다. 결과금지의 대상은 금지행위의 동일성으로 정해지므로 다른 버전을 썼다는 이유만으로 명령 밖으로 벗어나지 않고, 특정 사본에 대한 작위명령의 대상은 기록 상태의 동일성으로 정해지며, 가중치의 복사·미세조정·어댑터 추가·증류 등의 계보는 그 증거일 뿐 유일한 기준이 아니다. 대법원은 특허침해금지가처분에서 채권자가 금지 대상 물건을 집행이 가능할 정도로 구체적으로 특정하여야 하며 변경 물건이 집행 대상에서 벗어나는지는 실체상의 주장이어서 집행

에 관한 이의로 다툴 수 없다고 하였다[43]. 본 연구는 명령의 층위별로 대상의 기준이 주문에 특정되어 있어야 하고, 변경된 버전이 그 기준에 드는지는 집행권원의 해석 문제로서 집행 단계에서도 판단되며, 그 증명책임은 채권자에게 있다고 본다. 변경된 버전이 새로운 침해인지가 다투어지는 경우, 이는 실체 판단의 문제로서 새로운 청구에 의하여야 한다. 관리 범위는 제공자가 관리하는 사본·서비스·갱신본까지이며, 오픈소스로 배포된 모든 사본을 파기할 수는 없고[38] 제3자 보유 사본에 대한 조치는 그 제3자의 침해행위를 별도로 특정하여 구해야 한다.

제3항의 압류는 “침해행위로 말미암아 만들어진 물건”에 대한 보전처분이고, 서버 자체는 모델 사본의 저장 매체에 불과하다. 따라서 압류 대상성은 3-2절 2)의 세 요건이 그 사본에 관하여 갖추어졌는지에 달려 있다. 요건이 갖추어지더라도 서버 전체의 압류는 필요한 범위를 넘으므로, “그 밖의 필요한 조치”로서 비례성에 맞는 것은 모델 사본의 격리나 해당 출력의 차단이다. 잠정 조치는 결과금지에 한정되고, 침해 시점의 모델 버전과 기록을 보존하는 조치가 이에 더해질 필요가 있다. 모든 후속 버전을 자동으로 집행 대상에 포함하는 해석은 취하지 않으며, 제3자 보유 모델에 대한 전면적 회수·원격 삭제 의무는 본 연구에서는 논외로 한다.

3) 검증과 증거, 증명책임

이행을 확인할 자료는 상대방의 지배 영역에 있으며, 소송 중의 자료 확보는 민사소송법 제344조의 문서제출명령이 기본이다[46]. 저작권법 제129조의2의 정보제공명령과[5] 제129조의3의 비밀유지명령은[5] 자료제출청구권이나 계속적 감사권을 창설하지 않는다. 제출 대상은 공개 자료, 비밀유지명령 아래 열람되는 내부 기록, 감정인이 상대방의 통제 아래 시험하고 결과만 보고하는 가중치·비공개 데이터의 세 층으로 나뉜다. 판결 확정 후 계속하여 시험 결과를 보고하고 후속 버전을 재검증할 의무는 제2항의 “그 밖의 필요한 조치”에 포함될 수 있다는 해석이 가능하나 이를 인정한 판례는 확인되지 않았고, 3-2절 1)의 결정들은 제2항의 조치가 물건의 처분에 한정되지 않는다는 논거일 뿐이므로 이 부분은 필자의 해석상 제안이다. 기간·대상·상대방·종료 조건이 특정되지 않은 보고의무는 주문의 특정성을 갖추지 못하며, 특정은 집행 가능성의 요건이지 권한의 근거가 아니다. 사적 합의는 그 자체로 집행권원이 되지 않지만, 재판상 화해로 변론조서에 기재되면 확정판결과 같은 효력(민사소송법 제220조)을 가지므로 특정된 보고·재검증 의무는 그 경로로 집행권원이 될 수 있다[46].

이행 평가의 증거는 네 유형으로 나뉜다. ① 대상과 버전의 동일성(해시, 배포 기록, 서비스 구성 기록). ② 조치의 실행(적용한 절차, 설정, 실행 기록). 문서화가 없는 이행 주장은 검증할 수 없지만[25], 기록만 있으면 이행이 확정되는 것도 아니다. ③ 결과(재현 시험). 재현 시험은 미리 공개되지 않은 검증 프롬프트, 모델 단독 조건과 검색 결합 조건의 구별, 접근 권한과 실행 조건의 기록을 요하며, 뮌헨 지방법원의 출력 대조와[1] 벤치마크의 기억 축이[26] 그 설계 기준이다. ④ 부작용(적법한 출력, 사실 지식, 일반 기능의 손실)[9]. 네 유형이 뒷받침하는 결론은 하나가 아니다. 비재현 관측, 데이터 기여 제거, 침해적 이용 방지는 서로 다른 명제이고 각 증거는 서로 다른 범위까지만 뒷받침한다. 그 대응 관계를 아래와 같이 표로 정리하였다.

본 연구는 “재현 불가가 검증되었다”는 표현을 사용하지 않는다. 시험 결과는 시험한 조건에서의 비재현을 보여줄 뿐이고, 어느 시험 범위와 잔존 불확실성이 충분한지는 법원이 정하며 본 연구는 수치 임계값을 제시하지 않았다[17]. 결과금지에서 시험 통과는 그 범위에서 위반이 관측되지 않았다는 증거이지 곧 이행이 아니며, 이행의 인정은 명령의 내용·대상·기간과 시험의 범위에 비추어 법원이 판단한다. 적극적 조치명령의 이행은 ②와 ③의 결합으로 판단하며 출력이 나오지 않는다는 사실만으로 충분하지 않다.

불가능하거나 부담이 과도한 조치는 최초 명령 단계에서 배제되고, 성과가 부족하거나 실행이 입증되지 않은 조치는 확정된 명령의 이행·위반 단계에서 명령의 내용에 따라 판단되며, 명령 자체를 다시 정하지는 않는다. “불가능”은 하급심

표 3. 증거와 그것이 뒷받침하는 결론
Table 3. Evidence and the conclusions it supports

Evidence	Non-reproduction observed	Data contribution removed	Infringing use prevented
Identity of work, version, and configuration (hashes, release records)	Precondition only	Precondition only	Precondition only
Execution of the measure (procedure, settings, logs)	No	Supports, if the method's effect is documented for that procedure	No
Reproduction tests (undisclosed prompts; model-only vs. retrieval-combined; access conditions recorded)	Yes, within the tested conditions	Weakly: absence of output is consistent with, but does not prove, removal	Yes, within the tested conditions and layer
Side-effect measurements (lawful outputs, facts, general utility)	No	No	Measures the scope of loss (fact); whether the loss reached uses outside the order is a separate legal assessment

이 온라인서비스제공자의 면책사유를 좁게 읽은 것처럼 [17],[28] 좁게 읽어야 하며, 파라미터 수준의 제거가 불가능하다는 사정은 그 조치를 명하지 않을 이유일 뿐 결과금지 자체를 면하게 하지 않는다.

증명책임은 하나의 기준으로 일률적으로 배분하기 어렵다. 저작권 침해소송의 요건사실 구조[40]을 언러닝 명령에 적용해 보면 다음과 같이 나누어 볼 수 있다. 이 가운데 판례로 확립되지 않은 부분은 해석상 제안이다.

권리자는 침해와 구체적 우려를 증명하여야 한다. 예방청구에서는 재현의 사실이나 재현될 개연성을 뒷받침하는 시험 결과가 이에 해당한다. 시험이 실패한 사실, 그리고 집행 단계에서 새 버전이 명령의 객관적 범위에 포함된다는 사실도 권리자가 증명한다.

상대방은 조치를 실행하였다는 사실을 증명하여야 한다. 이행 자료를 제출하는 단계에서 시험 대상이 명령의 대상 버전과 같다는 사실, 기술적으로 불가능하다는 사실, 더 좁은 조치로 충분하다는 불비례 항변의 요건사실도 상대방이 증명한다. 이처럼 명령 대상의 동일성에 관한 증명책임을 단계에 따라 달리 정하는 것은, 간접강제결정의 집행에서 대상 해당성을 채권자가 증명하도록 한 대법원 판례와 같은 구조이다[47]. 한편 증거가 상대방에게 편재되어 있다는 문제는 문서제출명령, 감정, 사실상 추정으로 해결하여야 하며, 별도의 근거 없이 객관적 증명책임을 전환할 수는 없다.

4) 집행, 기간, 증거보존, 재발

민사집행법 제260조는 작위채무의 대체집행(민법 제389조 제2항)을[29],[48], 제261조는 간접강제를 규정한다[48]. 어느 집행방법에 의할지는 조치의 내용에 따라 다르며, 필터 설치나 저장소 문서 삭제는 대체집행이 가능한 경우가 있고, 파라미터 수준 조치가 부대체적인지는 제3자의 대행 가능성, 상대방의 협력 필요, 접근 절차, 명령이 결과를 정했는지를 고려하여 판단하여야 한다. 결과금지는 부작위채무이며 판결절차에서 동시에 간접강제를 명할 수 있다[49],[50]. 대법원 2016다268695 판결은 집행 조건의 성취는 채권자가, 대상의 부존재는 채무자가 증명하고, 의무이행 기간을 정한 가처분결정은 그 기간이 경과하면 집행권원으로서의 효력이 없다고 하였다[47]. 이를 언러닝 명령에 적용하면 결과금지의 위반은 채권자가 재현의 사실과 그 출력이 금지의 범위에 든다는 사실로 증명하고, 적극적 조치의 불이행은 채권자가 이행 요구와 이행 부재를 보이되 조치가 이미 실행되었다거나 대상이 존재하지 않는다는 사실은 채무자가 증명한다. 기간은 가처분의 효력기간, 적극적 조치의 이행기한, 보고의무의 존속기간, 간접강제금의 발생기간을 나누어야 한다. 위 판결이 다룬 것은 의무 존속기간을 정한 가처분이고, 일정 기한 안에 조치를 마치라는 이행기한은 기한이 지나면 이행지체가 될 뿐이어서, 간접강제결정이 있는 경우 배상금이 발생하는 것이지 의무가 소멸하는 것은 아니다.

모델과 기록을 먼저 파기하면 검증에 필요한 증거가 사라진다. 반대로 검증을 이유로 문제의 모델을 계속 서비스하게 하면 권리 구제가 늦어진다. 이 충돌은 서비스에서의 사용 중단(제3항의 잠정 조치 또는 제1항의 정지·예방청구)과 소송·검증을 위한 제한적 보존을 구별함으로써 해결할 수 있다. 민사소송법 제375조와 제376조의 증거보전은 미리 증거조사를 하지 아니하면 그 증거를 사용하기 곤란한 사정이 있을 때 소 제기 전에도 증거조사를 할 수 있도록 한다[46]. 침해 시점의 모델 버전과 조치 전후의 기록은 갱신으로 쉽게 소멸하므로 이 요건을 충족할 수 있다. 그러나 증거보전은 상대방에게 보관 의무를 부과하는 제도가 아니다. 따라서 특정 사본을 일정 기간 보관하도록 명하는 계속적 보존명령에는 별도의 근거가 필요하다. 제3항의 "그 밖의 필요한 조치"로 모델 사본의 격리·보존을 명할 수 있다고 해석할 여지는 있으나, 이를 인정한 판례는 확인되지 않는다. 본 연구는 이를 해석상 제안과 당사자 합의의 영역으로 남겨 두었다. 보존이 삭제보다 우선하는 것은 그 대상·통제자·목적·기간이 특정된 경우로 한정된다.

명령 이행 후 같은 저작물이 다시 출력되는 상황은 셋으로 나뉜다. ① 같은 모델이 추가 학습이나 미세조정 뒤 같은 표현을 재현한다(개념 재출현)[23]. ② 이용자가 저작물 원문을 새로 입력하여 재현을 유도한다(문맥 내 재도입)[24]. ③ 외부 검색 자료에서 표현이 공급된다[9]. 셋을 모두 "언러닝 실패"로 분류해서는 안 되며, 기술적 원인과 책임 귀속은 별개다. ①은 명령의 대상 버전에서 명령이 정한 출력 조건 안에서 발생하면 결과금지의 위반이고, 대상 밖의 버전에서 발생하면 5-1절 2)의 승계 문제다. ②는 2-2절에서 본 조치의 무의 발생 요건이 문제 된다. 다만 프롬프트의 유형은 사실이지 위반의 기준이 아니며, 위반 여부는 그 출력이 대상 버전에서 명령이 정한 시험 조건 안에서 발생하였는지로 판단한다. 원문 입력이 위반에서 제외되는 것은 명령이 이를 시험 조건에서 제외하였거나 상대방이 그 입력이 다른 원인임을 증명한 경우이므로, 시험 프롬프트의 유형이 명령에 특정되어 있을 때에만 원문 입력에 의한 재현이 위반에서 구별된다. ③은 저장소 관리자의 별개 행위로서 별도의 책임 근거를 요하되 같은 절차에서 공동으로 다룰 수 있다. 재검증 의무는 대상 버전, 재검증을 촉발하는 변경, 확인할 표현과 시험 범위, 보고 상대방, 종료·재검토 조건의 다섯 가지가 특정된 범위에서만 집행 가능한 의무가 되며, 특정되지 않은 재검증 의무는 집행할 수 없다. 가처분에는 사정이 바뀐 때의 취소 절차(민사집행법 제301조, 제288조 제1항 제1호)가 있으나[48], 확정된 본안판결의 금지명령을 사후에 관리·변경하는 절차는 현행법에 없고 그러한 제안이 특허 분야에서 나와 있을 뿐이다[43],[44].

5-2 머신 언러닝의 법적 지위와 한계점

이상의 검토를 종합하여 현행법상 머신 언러닝의 법적 지위를 정리하면 다음과 같다.

표 4. 현행법이 뒷받침하는 부분, 필자의 해석상 제안, 합의·입법 과제
Table 4. What current law supports, what the author proposes, and what requires agreement or legislation

Item	Supported by current law and case law	Author's interpretive proposal	Requires agreement or legislation
Result-based prohibition (Art. 123(1))	Cessation where reproduction occurred; prevention where concrete risk is shown	Tests as the means of showing concrete risk	–
Positive measures on the model copy held (Art. 123(2))	Measures beyond physical destruction accepted in lower courts; ancillary to a para (1) claim	Removal of a copy inside the model where copy status is found; mere possession is not infringement; necessity and scope examined separately	–
Court-ordered specific algorithm	No case law	Only where an additional reason is shown; result-based prohibition as the default	–
Document production and expert examination (Civil Procedure Act Art. 344 ff.; Copyright Act Arts. 129-2, 129-3)	Yes, for specified records, tests, and version information; confidentiality orders	Three tiers of materials; verifier within expert examination	Standing audit rights; access beyond expert examination
Post-judgment reporting and re-verification	No case law; specification is a condition of enforceability, not a source of power	Arguable as an "other necessary measure" if version, trigger, scope, recipient, and termination are specified	Unspecified continuing duties; effect on third parties; agreed reporting unless recorded as a judicial settlement (Civil Procedure Act Art. 220)
Effect on successor versions	Enforcement against modified objects turns on the specification of the title (2025Ma6304, patent)	Result-based prohibition: identity of the prohibited act; measures on a copy: identity of the record state, lineage as evidence	Automatic inclusion of all future versions
Copies held by third parties	No; separate claims against those third parties	Not proposed	Not created by this article
Evidence preservation (Civil Procedure Act Arts. 375-376; Art. 123(3))	Early examination of evidence, expert examination, copies	Isolation and preservation of a model copy as a provisional measure under Art. 123(3)	Continuing preservation duties; retention exceptions to deletion orders
Enforcement and periods (Civil Execution Act Arts. 260-261)	Substitute execution or indirect compulsion by content of the measure; creditor proves the object falls within the title; validity period, deadline, and accrual distinguished	Factors for non-substitutability of parameter-level measures; separate duration of a reporting duty	–

특정 저작물의 침해적 출력·제공을 금지하는 결과금지는 제123조 제1항의 정지·예방청구 그 자체이다. 재현의 사실 또는 재현될 개연성을 뒷받침하는 시험 결과로 구체적 우려가 인정된다면 현행법으로도 발령할 수 있다. 이때 수단의 선택은 상대방에게 맡겨지며, 그 위반은 부작위채무의 위반으로서 판결절차에서 간접강제와 함께 다루어진다.

모델 사본의 제거·격리와 같은 적극적 조치는 제1항 청구에 부수하는 제2항의 "그 밖의 필요한 조치"에 해당한다. 따라서 모델 안에 복제물이 인정되는 범위에서 필요성과 비례성 심사를 거쳐 명할 수 있다. 제2항의 조치를 물건의 처분에 한정하지 않은 하급심 실무가 이러한 해석을 뒷받침한다. 반면 특정 알고리즘의 실시를 명하는 것에 관하여는 판례가 없으므로, 이를 추가 사유가 있는 경우로 한정하자는 것은 해석상 제안에 머문다.

검증 절차 중 소송 과정에서의 문서제출·감정과 비밀유지명령은 현행법이 뒷받침한다. 그러나 판결 확정 후의 계속적 보고·재검증 의무는 사정이 다르다. 이를 제2항의 조치로 볼 여지는 있으나 인정한 판례가 없고, 대상 버전·촉발 조건·시험 범위·보고 상대방·종료 조건이 특정되어야 비로소 집행 가능한 의무가 된다. 당사자 사이의 합의도 재판상 화해로 조서에 기재되어야 집행권원이 된다.

현행법 해석만으로는 닿지 않는 부분이 한계점이다. 상대방의 시스템에 대한 계속적 감사권, 확정판결의 금지명령을 사후에 관리·변경하는 절차, 제3자가 보유한 사본에 대한 전면적 회수·원격 삭제, 모든 후속 버전을 자동으로 집행 대상에 포함하는 효력, 특정 사본을 일정 기간 보관하도록 하는 계속적 보존명령은 제123조의 해석으로 창설할 수 없고 당사자의 합의 또는 입법을 요한다. 표 4는 이를 현행법이 뒷받침하는 부분, 저자의 해석 제안, 합의·입법 과제로 나누어 정리한 것이며, 판례로 확인되지 않은 부분은 해석상 제안으로 표시하였다.

Ⅵ. 결 론

본 연구는 특정 저작물의 침해적 재현이 문제 되는 생성형 AI 모델에 대하여 머신 언러닝을 제123조의 구제수단으로 활용할 경우, 그 법적 지위와 적용 요건 및 범위를 검토하였다. 검토는 기술적으로 완전한 삭제가 가능하다는 전제를 두지 않고 현행법 해석의 틀 안에서 이루어졌다. 논의의 출발점은 침해의 특수성이었다. 생성형 AI를 통한 침해에서는 침해 복제물이 특정 파일로 존재하지 않으므로, 그 존부를 저작물별 시험으로만 확인할 수 있다. 학습 단계의 복제는 이미 완료된 행위이고 모델을 서버에 보유하는 것만으로는 침해가 되지 않으므로, 정지청구의 대상은 현재의 제공·전송·출력이다. 또한 같은 출력이라도 그 원인이 기억인지, 검색 증강인지, 이용자 입력인지에 따라 조치의 대상과 상대방이 달라진다. 상대방은 학습자·배포자·운영자·검색 저장소 관리자·이용자 가운데 누구에게 책임과 통제 가능성이 있는지를 확인하여 특정하여야 한다. 이때 프롬프트의 유형은 하나의 사실로 고려될 뿐, 행위주체나 위반 여부를 정하는 기준이 되지는 못한다. 언러닝 기술은 영향 제거와 출력 억제라는 서로 다른 목적의 두 기술군으로 나뉘는데, 어느 쪽도 재현 불가를 보장하지 못한다. 그러므로 법이 활용할 수 있는 것은 어떤 알고리즘을 어떤 데이터에 적용하였는지에 대한 확인과, 시험 조건에서 재현되지 않았다는 관측에 한정된다.

이러한 전제에서 본 연구가 밝힌 내용은 다음과 같다.

먼저 제123조의 해석이다. 제2항의 “그 밖의 필요한 조치”는 제1항의 실효성을 확보하기 위한 부수조치로 넓게 해석하는 것이 법의 체계와 하급심 실무에 부합한다. 어느 해석을 취하더라도 복제물성이 부정된다고 하여 예방청구가 불가능해지는 것은 아니며, 복제물성이 인정된다고 하여 모델 전체의 폐기가 필연적인 것도 아니다. 물건성·복제물성·관련성은 각각 별개의 요건이고, 제거 조치는 제1항 청구에 부수하여 상대방이 관리하는 특정 기록 상태를 대상으로 명할 수 있다.

다음으로 언러닝의 법적 지위이다. 언러닝을 구제수단으로 명하는 경우 그 내용은 결과금지, 적극적 조치, 검증 절차로 나뉜다. 언러닝은 원칙적으로 제1항의 정지·예방청구를 이행하는 수단이며, 복제물이 인정되는 범위에서만 제1항 청구에 부수하는 제2항의 제거 조치로 확장된다. 그 이행 여부는 증거에 의하여 판단한다. 명령의 중심은 결과금지에 있고, 적극적 조치나 특정 알고리즘의 지정은 추가 사유가 있을 때 덧붙여진다.

끝으로 국내법상의 적용이다. 명령의 내용에 따라 후보 기술이 정해지면 그 가운데 비례성에 맞는 조치가 선택되며, 폐기와 언러닝 사이의 선택은 여러 고려 요소에 비추어 사건별로 판단한다. 적법한 이용의 보호는 명령 범위를 제한하는 원리로 작용할 뿐, 이용자에게 별도의 청구권을 부여하는 근거는 되지 않는다. 이행 여부는 동일성·실행·결과·부작용에 관한 증거로 판단한다. 시험을 통과하였다는 사실은 위반이 관측되지 않았다는 증거로서 의미를 가질 뿐, 그 자체로 이행을 뜻하지는 않는다. 증명책임은 침해와 위반에 대하여는 권리자가, 이행·불능·불비례에 대하여는 상대방이 부담한다. 다만 명령 대상의 동일성에 관하여는 이행 자료를 제출하는 단계와 집행 단계를 나누어 배분한다. 가처분의 효력기간, 조치의 이행기한, 보고의무의 존속기간, 간접강제금의 발생기간도 서로 구별하여야 한다. 재발 사례는 명령의 층위에 따라 기존 명령의 위반, 명령 대상에 해당하는지가 다투어지는 변경 버전, 별개의 새로운 침해로 나누어 다룬다. 비교법적 검토에서 참고한 것은 외국의 결론 자체보다 그 논증의 방식이었다.

본 연구의 기여는 새로운 조문이나 일반적 언러닝 청구권을 창설하지 않고도 제123조에 근거하여 조치의 선택, 명령의 특정, 자료의 접근, 이행의 판단을 연결하는 틀을 현행법 해석으로 구성한 데 있으며, 판례로 확인되지 않은 부분은 표 4에 해석상 제안으로 표시하였다. Rubí-Puig와 Wolsing이 사건별 분석과 조치의 결합을 결론으로 제시한 것과 방향을 같이하되[35], 본 연구는 그 분석을 한국법의 조문·절차·집행 단계로 내려 구체화하였다. 한계는 셋이다. 국내에는 AI 학습이나 언러닝에 관한 판례가 확인되지 않아 해석의 참조 사례는 침해물 폐기·삭제 명령과 간접강제에 관한 국내 판례, 그리고 독일·영국·EU의 판결에 그쳤고, 검증 방법은 기술 문헌의 평가 축을 규범으로 번역한 것이어서 어느 시험 범위와 잔존 불확실성이 이행 판정에 충분한지는 사건의 축적을 기다려야 하며, 학습 자체의 적법성과 사전 면책의 설계는 범위 밖에 두었다. 모델의 복제물성에 관한 두 갈래와 침해 입증의 곤란은 본 연구가 조건으로 처리한 선결문제로서 그 해결에 따라 제2항 조치의 비중이 달라진다. 후속 연구의 접점은 둘이다. 본 연구가 해석론으로 세운 명령의 특정 요건과 재검증 의무의 다섯 요소는 언러닝 의무를 조문화할 때의 요건으로 옮겨질 수 있으나 그 조문이 TDM 면책과 어떻게 결합할지는 별도의 검토를 요하고, 학습 완료 모델에서 개인정보의 삭제를 요구하는 것이 모델 언러닝을 뜻한다는 논의가 시작된 만큼[51] 저작물과 개인정보가 같은 모델 안에서 같은 기술로 제거되어야 한다면 두 법의 이행 판단 기준은 서로를 참조하여 설계되어야 한다. 아울러 국내외 판결과 언러닝 평가 방법이 축적됨에 따라 명령의 특정 요건과 재검증 기준이 어떻게 달라지는지를 추적하는 추이 연구가 필요할 것이다.

감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-정보통신방송혁신인재양성사업의 지원을 받아 수행된 연구임(IITP-2026-RS-2023-00266605) & 이와 더불어 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2025-00518960) & 이와 더불어 본 연구는 문화체육관광부 및 한국콘텐

츠진흥원의 2026년도 문화체육관광 연구개발사업으로 수행되었음(과제명: 블록체인 기술 기반 SW 저작권 보호를 위한 유통·관리 플랫폼 기술개발 및 인재양성, 과제번호: RS- 2023-00228867).

참고문헌

[1] Landgericht München I (Munich Regional Court I), 42 O 14139/24 (GEMA v. OpenAI), Rendered on November 11, 2025, BeckRS 2025, 30204.

[2] High Court of Justice of England and Wales, Getty Images (US) Inc. v. Stability AI Ltd [2025] EWHC 2863 (Ch), Rendered on November 4, 2025.

[3] Law Times. Three Terrestrial Broadcasters Sue OpenAI: ‘Unauthorized Training on News Infringes Copyright [Internet]. Available: https://www.lawtimes.co.kr/news/216712.

[4] A. F. Cooper, C. A. Choquette-Choo, M. Bogen, K. Klyman, M. Jagielski, K. Filippova, ... and K. Lee, “Machine Unlearning Doesn’t Do What You Think: Lessons for Generative AI Policy and Research,” arXiv:2412.06966, 2025. https://doi.org/10.48550/arXiv.2412.06966

[5] Copyright Act (Republic of Korea), Act No. 432, Jan. 28, 1957, as amended by Act No. 21336, Feb. 10, 2026.

[6] S. Kwon, “A Study on the Practical Problems of Copyright Infringement Litigation Related to AI Training – Focusing on Pending Cases in the U.S. District Court for the Northern District of California,” *Hanyang Law Review*, Vol. 41, No. 4, pp. 121-162, December 2024. https://doi.org/10.18018/hylr.2024.41.4.121

[7] S. Choi, “Artistic Works in the Era of Artificial Intelligence in the Context of the Copyright Law and Unfair Competition Prevention Act – Focusing on the Issue of Generative AI’s Style Copying,” *Journal of Korea Information Law*, Vol. 29, No. 2, pp. 87-116, 2025.

[8] N. Carlini, J. Hayes, M. Nasr, M. Jagielski, V. Sehwag, F. Tramèr, ... and E. Wallace, “Extracting Training Data from Diffusion Models,” in *Proceedings of the 32nd USENIX Conference on Security Symposium*, Anaheim: CA, pp. 5253-5270, August 2023.

[9] B. Wei, W. Shi, Y. Huang, N. A. Smith, C. Zhang, L. Zettlemoyer, ... and P. Henderson, “Evaluating Copyright Takedown Methods for Language Models,” arXiv:2406.18664, 2024. https://doi.org/10.48550/arXiv.2406.18664

[10] Supreme Court of Korea Decision, 2016Da20916, Rendered on November 15, 2018.

[11] B. Marino, M. Kurmanji, and N. D. Lane, “Position: Bridge the Gaps Between Machine Unlearning and AI Regulation,” arXiv:2502.12430, 2025. https://doi.org/10.48550/arXiv.2502.12430

[12] Landgericht München I (Munich Regional Court I), 42 O 763/25 (GEMA v. Suno), Rendered on July 31, 2026, BeckRS 2026, 17961; Press Release No. 16/2026, July 31, 2026.

[13] S. Kwon, “A TDM Exceptions and Fair Use in the Context of AI Training – Focusing on the Changed Aspects of Copyright Problems Following the Advent of the Generative AI,” *Journal of Korea Information Law*, Vol. 28, No. 2, pp. 25-56, 2024.

[14] T. W. Dornis and S. Stober, “On the Application of Law in the AI Engine Room – A Critical Look at ‘GEMA/OpenAI’,” *GRUR*, p. 620, 2026.

[15] A. F. Cooper and J. Grimmelmann, “The Files Are in the Computer: On Copyright, Memorization, and Generative AI,” arXiv:2404.12590, 2025. https://doi.org/10.48550/arXiv.2404.12590

[16] Supreme Court of Korea, Decision 2005Da11626, Rendered on January 25, 2007.

[17] Seoul High Court, Decision 2006Ra1245, Rendered on October 10, 2007.

[18] Supreme Court of Korea, Decision 2009Da4343, Rendered on March 11, 2010.

[19] Supreme Court of Korea, Decision 2016Da271608, Rendered on February 28, 2019.

[20] L. Bourtoule, V. Chandrasekaran, C. A. Choquette-Choo, H. Jia, A. Travers, B. Zhang, ... N. Papernot, “Machine Unlearning,” in *Proceeding of the 2021 IEEE Symposium on Security and Privacy (SP)*, San Francisco: CA, pp. 141-159, May 2021. https://doi.org/10.1109/SP40001.2021.00019

[21] M. K. Sung, K. Lee, J. Lee, and J. Y. Han, “An Investigative Study on the Evaluation of Efficiency and the Model Verification of Machine Unlearning,” in *Proceeding of the 5th Korea Artificial Intelligence Conference*, pp. 171-172, 2025.

[22] R. Gandikota, J. Materzyńska, J. Fiotto-Kaufman, and D. Bau, “Erasing Concepts from Diffusion Models,” in *Proceeding of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, pp. 2426-2436, October 2023. https://doi.org/10.1109/ICCV51070.2023.00230

[23] V. M. Suriyakumar, R. Alur, A. Sekhari, M. Raghavan, and A. C. Wilson, “Unstable Unlearning: The Hidden Risk of Concept Resurgence in Diffusion Models,”

arXiv:2410.08074, 2024. https://doi.org/10.48550/arXiv.2410.08074

[24] I. Shumailov, J. Hayes, E. Triantafillou, G. Ortiz-Jimenez, N. Papernot, M. Jagielski, ... and E. Bagdasaryan, "UnUnlearning: Unlearning Is Not Sufficient for Content Regulation in Advanced Generative AI," arXiv:2407.00106, 2024. https://doi.org/10.48550/arXiv.2407.00106

[25] A. Thudi, H. Jia, I. Shumailov, and N. Papernot, "On the Necessity of Auditable Algorithmic Definitions for Machine Unlearning," in *Proceeding of the 31st USENIX Security Symposium*, Boston: MA, pp. 4007-4022, August 2022.

[26] W. Shi, J. Lee, Y. Huang, S. Malladi, J. Zhao, A. Holtzman, ... and C. Zhang, "MUSE: Machine Unlearning Six-Way Evaluation for Language Models," arXiv:2407.06460, 2024. https://doi.org/10.48550/arXiv.2407.06460

[27] Seoul Central District Court, Decision 2007GaHap67109, Rendered on April 23, 2008.

[28] Seoul Central District Court, Decision 2008KaHap968, Decided on August 5, 2008.

[29] Civil Act (Republic of Korea), Act No. 471, Feb. 22, 1958, as Amended by Act No. 21454, Mar. 17, 2026.

[30] Seoul High Court, Decision 2016Na2019365, Rendered on April 6, 2017.

[31] Seoul High Court, Decision 2014Na2011824, Rendered on January 11, 2018.

[32] Seoul Central District Court, Decision 2021GaHap573167, Rendered on October 13, 2023.

[33] Court of Justice of the European Union, Case C-314/12, UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH and Wega Filmproduktionsgesellschaft mbH, ECLI:EU:C:2014:192, Rendered on March 27, 2014.

[34] European Parliament and of the Council of the European Union, "Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the Enforcement of Intellectual Property Rights," *Official Journal of the European Union*, L 157, pp. 45-86, April 2004.

[35] A. Rubí-Puig and D. Wolsing, Remedying Copyright Infringement in Generative AI. What Role for Algorithmic Destruction, Machine Unlearning and Damages?, in *The Cambridge Handbook of Generative AI and IP in Europe*, Cambridge, UK: Cambridge University Press, 2026 (forthcoming); SSRN Working Paper No. 5465935, 2025. https://doi.org/10.2139/ssrn.5465935

[36] Court of Justice of the European Union, Case C-70/10, Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM), ECLI:EU:C:2011:771, Rendered on November 24, 2011.

[37] T. C. Li, "Algorithmic Destruction," *SMU Law Review*, Vol. 75, No. 3, p. 479, 2022. https://doi.org/10.25172/smulr.75.3.2

[38] P. Samuelson. How to Think About Remedies in the Generative AI Copyright Cases [Internet]. Available: https://www.lawfaremedia.org/article/how-to-think-about-remedies-in-the-generative-ai-copyright-cases.

[39] S. Ryu, "Rethinking Dependence (Copying) Requirement, Part II: Do Generative AI Models Copy Training Data?," *The Justice*, Vol. 209, pp. 472-526, August 2025. https://doi.org/10.29305/tj.2025.8.209.472

[40] S.-H. Park, "A Study on Burden of Proof and Averment in Copyright Infringement Suit in the Case of Generative AI," *Seoul National University Law Review*, Vol. 66, No. 1, pp. 165-220, 2025.

[41] E. J. Jeon, "A Study on Copyright Infringement by AI-Generated Works," *Journal of Korea Information Law*, Vol. 29, No. 1, pp. 50-86, 2025.

[42] J. Kim, S. Park, Y. Hwang, B. Oh, J. Lee, and J. Hong, "A Study on the Improvement of Unlearnable Examples and Technical Protection Measures for Copyright Protection in the Generative AI Era," *Journal of Digital Contents Society*, Vol. 26, No. 12, pp. 3507-3520, December 2025. https://doi.org/10.9728/dcs.2025.26.12.3507

[43] Supreme Court of Korea, Order 2025Ma6304, Rendered on September 29, 2025.

[44] B. W. Kwon, "From Order to Management: An Equitable Redesign of Patent Injunctions − Lessons from the Supreme Court Decision 2025Ma6304," *The Justice*, Vol. 213, pp. 322-371, May 2026. https://doi.org/10.29305/tj.2026.05.213.322

[45] S. Ryu, "Civil Execution Against a Target Product Altered After a Patent Infringement Preliminary Injunction − Review of Supreme Court's 2025Ma6304 Decision," *Journal of Industrial Property*, No. 82, pp. 1-46, December 2025. https://doi.org/10.36669/ip.2025.82.1

[46] Civil Procedure Act (Republic of Korea), Act No. 547, Apr. 4, 1960, as Amended by Act No. 19516, July 11, 2023.

[47] Supreme Court of Korea, Decision 2016Da268695, Rendered on June 24, 2021.

[48] Civil Execution Act (Republic of Korea), Act No. 6627, Jan. 26, 2002, as Amended by Act No. 20733, Jan. 31, 2025.

[49] Supreme Court of Korea, Decision (en banc) 2020Da248124, Rendered on July 22, 2021.

[50] Supreme Court of Korea, Decision 2013Da50367, Rendered on November 28, 2013.

[51] D. H. Lee, “A Study on Measures to Resolve the Tension Between Artificial Intelligence (AI) and Personal Data Protection Norms,” *The Justice*, Vol. 212, pp. 254-281, February 2026. https://doi.org/10.29305/tj.2026.02.212.254



신민지(Min-Ji Shin)

2022년 : 단국대학교 법과대학 (법학사)
2024년 : 중앙대학교 대학원
(석사-보안경영학 산업보안관리 전공)
2021년～2022년: LNB 특허사무소
2023년～2026년: 중앙대학교 일반대학원 공공갈등정보학과 교육조교
2022년～현 재: 중앙대학교 산업보안연구소 연구원
2024년～현 재: 중앙대학교 대학원 융합보안학과 박사과정
※관심분야 : 인공지능법, 데이터법, 지식재산권법, 산업보안 법·정책



노승민(Seungmin Rho)

2001년 : 아주대학교
(공학사-컴퓨터공학)
2003년 : 아주대학교 대학원
(공학석사-정보통신공학)
2008년 : 아주대학교 대학원
(공학박사-정보통신공학)
2008년～2009년: Carnegie Mellon University 박사후연구원
2010년～2011년: 고려대학교 전기전자전파공학부 연구교수
2012년: 백석대학교 정보통신공학부 조교수
2013년～2018년: 성결대학교 미디어소프트웨어학부 조교수
2019년～2020년: 세종대학교 소프트웨어학부 조교수
2021년～현 재: 중앙대학교 산업보안학과 교수
※관심분야 : 신뢰할 수 있는 인공지능, 인공지능 보안, 머신 언러닝, 데이터 프라이버시, 산업보안



이미영(Mi Young Lee)

1995년 : 울산대학교
(디자인학사-산업디자인학)
2004년 : 부산대학교 대학원
(공학석사-영상정보공학)
2011년 : 부산대학교 대학원
(공학박사-영상정보공학)
2012년～2019년: 세종대학교 디지털콘텐츠학과 연구교수
2019년～2023년: 세종대학교 소프트웨어학과 대우교수
2023년～현 재: 중앙대학교 융합보안학과 연구교수
※관심분야 : 인공지능, 컴퓨터 비전, 생성형 AI, 데이터 프라이버시, 머신 언러닝, 인공지능 보안