



제 428회 국회 (임시회)
제2차 과학기술정보방송통신위원회

인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 일부개정법률안

검 토 보 고

〈인공지능 학습용데이터에 관한 정보 공개 등〉

- 박수현의원 대표발의(의안번호 제2210815호)
- 김기현의원 대표발의(의안번호 제2210903호)

2025. 8.

과학기술정보방송통신위원회

수석전문위원 이복우

목 차

I . 제안경 위	1
II . 제안이유 및 주요내용	1
III . 검토의견	4
1. 인공지능 학습용데이터에 관한 정보 공개 등(안 제31조 제4항 신설 등) ..	4

I. 제안경위

1. 정보통신망 이용촉진 및 정보보호 등에 관한 법률 일부개정법률안(제2210815호)

가. 발 의 자 : 박수현의원 등 12인

나. 발 의 일 : 2025. 6. 13.

다. 회 부 일 : 2025. 6. 16.

2. 정보통신망 이용촉진 및 정보보호 등에 관한 법률 일부개정법률안(제2210903호)

가. 발 의 자 : 김기현의원 등 12인

나. 발 의 일 : 2025. 6. 17.

다. 회 부 일 : 2025. 6. 18.

II. 제안이유 및 주요내용

1. 정보통신망 이용촉진 및 정보보호 등에 관한 법률 일부개정법률안(제2210815호)

현행법은 인공지능 발전과 신뢰 기반 조성을 위한 기본적 사항을 규정하고 있으나, 생성형 인공지능의 학습용데이터 공개 의무 등 창작자 보호를 위한 구체적 규정은 두고 있지 않음.

생성형 인공지능이 무단으로 다양한 창작물을 학습용데이터로 활용하는 사례가 증가하고 있음에도, 창작자는 자신이 제작한 콘텐츠가 어떻게

인공지능 학습에 활용되고 있는지 확인할 수 있는 제도적 장치가 마련되어 있지 않음. 이로 인해 창작자의 권리가 침해될 우려가 있으며, 생성형 인공지능으로부터 창작자를 보호하기 위한 규정이 시급한 실정임.

이에 인공지능사업자가 학습용데이터에 관한 정보를 공개하도록 노력하여야 하고, 저작물등의 권리자가 자신의 저작물등이 학습용데이터로 이용되었는지 확인을 요청하는 경우 학습에 이용되었는지 여부를 확인할 수 있는 절차를 마련하도록 하는 한편, 인공지능사업자들이 투명성확보의무를 자율적으로 준수할 수 있도록 민관협의체를 구성할 수 있도록 하여 인공지능 발전과 창작자 권리 보호 간의 균형을 도모하려는 것임(안 제31조, 제31조의2 신설 등).

2. 정보통신망 이용촉진 및 정보보호 등에 관한 법률 일부개정 법률안(제2210903호)

현행법은 인공지능 발전과 신뢰 기반 조성을 위한 기본적 사항을 규정하고 있으나, 생성형 인공지능의 학습용데이터를 확인할 수 있는 절차 등 창작자 보호를 위한 구체적 규정은 미비한 상황임.

특히 저작물의 권리자는 생성형 인공지능이 자신이 제작한 콘텐츠가 학습용데이터로 활용되고 있음에도 이를 확인할 수 있는 제도적 장치가 없어 권리가 침해될 우려가 있음. 따라서 생성형 인공지능으로부터 저작물의 권리자를 보호하기 위한 규정이 시급한 실정임.

이에 저작물등의 권리자가 자신의 저작물등이 학습용데이터로 이용되었는지 확인을 요청하는 경우 학습에 이용되었는지 여부를 확인할 수 있는 절차를 마련하도록 함으로써 인공지능 발전과 창작자 권리 보호 간의 균형을 도모하려는 것임(안 제31조).

Ⅲ. 검토 의견

1. 인공지능 학습용데이터에 관한 정보 공개 등(안 제31조 제4항 신설 등)

가. 개정안 요지

- 인공지능사업자는 저작물등¹⁾의 관리자가 자신의 저작물등이 학습용데이터로 이용되었는지 확인을 요청하는 경우 이를 확인할 수 있는 절차를 마련하여야 함(박수현의원안 제31조 제5항 신설, 김기현의원안 제31조 제4항 신설).
- (박수현의원안) 인공지능사업자가 생성형 인공지능을 이용한 제품 또는 서비스를 제공하는 경우 대통령령으로 정하는 바에 따라 학습용데이터에 관한 정보를 공개하도록 노력하여야 함(안 제31조 제4항).
- (박수현의원안) 정부는 생성형 인공지능의 학습용데이터 공개 및 저작물등 이용에 관한 사항 등을 논의하기 위하여 필요한 경우 인공지능사업자, 저작권자 등과 협의체를 구성·운영할 수 있음(안 제31조의2 신설).
- (박수현의원안) 과학기술정보통신부장관은 인공지능사업자가 다음의 의무를 위반하였음을 발견하였거나 그 혐의가 있음을 알게 된 경우와, 관련 신고를 받거나 민원이 접수된 경우 인공지능사업자에 대하여 관련 자료를 제출하게 하거나, 소속 공무원으로 하여금 필요한 조사를 하게 할 수 있음(안 제40조 제1항).

1) 「저작권법」 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

1. “저작물”은 인간의 사상 또는 감정을 표현한 창작물을 말한다.

7. “공중송신”은 저작물, 실연·음반·방송 또는 데이터베이스(이하 “**저작물등**”이라 한다)를 공중이 수신하거나 접근하게 할 목적으로 무선 또는 유선통신의 방법에 의하여 송신하거나 이용에 제공하는 것을 말한다.

- 안 제31조 제4항에 따른 학습용데이터 정보 공개 노력 의무
- 안 제31조 제5항에 따른 저작물등의 학습용데이터 이용 여부 확인 절차 마련 의무

나. 개정안 조문대비표

현행	박수현의원안	김기현의원안
법률 제20676호 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 제31조(인공지능 투명성 확보 의무) ① ~ ③ (생략) <u><신설></u>	법률 제20676호 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 제31조(인공지능 투명성 확보 의무) ① ~ ③ (현행과 같음) ④ <u>인공지능사업자는 생성형 인공지능을 이용한 제품 또는 서비스를 제공하는 경우 대통령령으로 정하는 바에 따라 학습용데이터에 관한 정보를 공개하도록 노력하여야 한다.</u> ⑤ <u>인공지능사업자는 「저작권법」 제2조제7호에 따른 저작물등(이하 “저작</u>	법률 제20676호 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 제31조(인공지능 투명성 확보 의무) ① ~ ③ (현행과 같음) ④ <u>인공지능사업자는 「저작권법」 제2조제7호에 따른 저작물등의 권리자가</u>

현행	박수현의원안	김기현의원안
	물등”이라 한다)의 권리자가 자신의 저작물등이 학습용데이터로 이용되었는지 확인을 요청하는 경우 이를 확인할 수 있는 절차를 마련하여야 한다.	자신의 저작물등이 학습용데이터로 이용되었는지 확인을 요청하는 경우 이를 확인할 수 있는 절차를 마련하여야 한다.
④ 그 밖에 제1항에 따른 사전고지, 제2항에 따른 표시, 제3항에 따른 고지 또는 표시의 방법 및 그 예외 등에 관하여 필요한 사항은 대통령령으로 정한다.	⑥ ----- ----- ----- ----- -----예외, 제4항에 따른 공개 방법 및 절차, 제5항에 따른 확인 절차 등에----- ----- -----.	⑤ ----- ----- ----- ----- -----예외, 제4항에 따른 확인 절차 등에----- ----- -----.
<신설>	제31조의2(민관협의회 구성) ① 정부는 생성형 인공지능의 학습용데이터 공개 및 저작물등 이용에 관	

현행	박수현의원안	김기현의원안
	<u>한 사항 등을 논의하기 위하여 필요한 경우 인공지능사업자, 저작권자 등과 협의체를 구성·운영할 수 있다.</u> <u>② 제1항에 따른 협의체의 구성·운영 등에 필요한 사항은 대통령령으로 정한다.</u>	
제40조(사실조사 등)	제40조(사실조사 등)	
① 과학기술정보통신부장관은 다음 각 호의 어느 하나에 해당하는 경우에는 인공지능사업자에 대하여 관련 자료를 제출하게 하거나, 소속 공무원으로 하여금 필요한 조사를 하게 할 수 있다. 1. <u>제31조제2항·제3항, 제32조제1항·</u>	① ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- 1. <u>제31조제2항부터 제5항까지</u>-----	

현행	박수현의원안	김기현의원안
제2항 또는 제34조제1항에 위반되는 사항을 발견하거나 혐의가 있음을 알게 된 경우	----- ----- ----- ----- -----	
2. 제31조제2항·제3항, 제32조제1항·제2항 또는 제34조제1항의 위반에 대한 신고를 받거나 민원이 접수된 경우	2. 제31조제2항부터 제5항까지----- ----- ----- ----- -----	
②·③ (생략)	②·③ (현행과 같음)	

다. 검토사항

- 최근 사진을 특정 화풍으로 변환하는 생성형 인공지능이 주목을 받으면서, 저작물을 인공지능 학습용데이터로 이용하는 것이 저작권 침해에 해당하는지에 대한 논란이 제기된 바 있음.
- 현행 「저작권법」은 저작물의 복제, 배포, 2차적저작물 작성 등의 방법으로 저작재산권을 침해한 자를 처벌하되²⁾,

2) 제136조(벌칙) ① 다음 각 호의 어느 하나에 해당하는 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처하거나 이를 병과(併科)할 수 있다.

1. 저작재산권, 그 밖에 이 법에 따라 보호되는 재산적 권리(제93조에 따른 권리는 제외한다)를 복제, 공연, 공중송신, 전시, 배포, 대여, 2차적저작물 작성의 방법으로 침해한 자

저작물의 일반적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있도록 하고 있음(저작물의 공정한 이용)³⁾.

- 그런데 저작물을 인공지능 학습용데이터로 이용하는 것이 공정한 이용에 해당하는지에 대해 견해가 대립하고 있고, 저작권자가 자신의 저작물등이 인공지능 학습용데이터로 활용되었는지를 직접 확인할 수 없어 저작권 침해를 주장하기 어려운 현실임.
- 이에 개정안은 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하 “인공지능기본법”)에 인공지능 학습용데이터의 정보 공개 및 저작물등의 이용 여부 확인 절차 등을 규정함으로써 저작권자의 권리를 보호하는 동시에 학습용데이터의 투명성과 신뢰성을 제고하려는 것임.
- 다만 개정안에 대해서는 다음의 사항을 고려할 필요가 있음.
- 첫째, 인공지능기본법은 인공지능 등에 관한 기본법으로서 다른 법률에 특별한 규정이 있는 경우 그에 따르도록 하고 있고⁴⁾,

3) 제35조의5(저작물의 공정한 이용) ① 제23조부터 제35조의4까지, 제101조의3부터 제101조의5까지의 경우 외에 저작물의 일반적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에는 저작물을 이용할 수 있다.

② 저작물 이용 행위가 제1항에 해당하는지를 판단할 때에는 다음 각 호의 사항등을 고려하여야 한다.

1. 이용의 목적 및 성격
2. 저작물의 종류 및 용도
3. 이용된 부분이 저작물 전체에서 차지하는 비중과 그 중요성
4. 저작물의 이용이 그 저작물의 현재 시장 또는 가치나 잠재적인 시장 또는 가치에 미치는 영향

4) 제5조(다른 법률과의 관계) ① 인공지능, 인공지능기술, 인공지능산업 및 인공지능사회(이하 “인공지능등”이라 한다)에 관하여 다른 법률에 특별한 규정이 있는 경우를 제외하고는 이 법에서 정하는 바에 따른다.

현재 저작자 및 저작권 등의 보호와 저작권 관련 분쟁의 조정 등 저작권에 관한 일반적인 사항은 「저작권법」에 규정되어 있으므로, 저작물의 인공지능 학습용데이터 활용에 관한 사항을 「저작권법」과 인공지능기본법 중 어느 법률에 포함하는 것이 법체계상 적절한지에 대해서는 추가적인 논의가 필요할 것으로 보임⁵⁾.

- 둘째, 인공지능시스템 개발 과정에서 어떠한 학습데이터가 이용되었는지는 해당 시스템의 품질 및 특성과 밀접하게 관련되어 있으므로, 학습데이터에 관한 정보 일체를 공개하도록 하는 경우 기업의 영업비밀을 침해할 우려가 있다는 점을 고려할 필요가 있음.
- 셋째, 인공지능, 특히 생성형 인공지능 학습에 사용되는 데이터의 규모가 방대하다는 점을 고려할 때, 학습데이터에 특정 저작물등이 포함되어 있는지를 추적·확인하는 것이 현실적으로 불가능하거나 어려운 측면이 있음⁶⁾.
- 넷째, 박수현의원안은 안 제31조 제4항의 학습용데이터 정보 공개 노력의무를 위반한 자에 대해서도 자료제출요구 및 사실조사를 할 수 있도록 규정하고 있는데,

어떠한 경우에 노력의무를 다하지 않았다고 볼 수 있는지 명확한 기준을 마련하기 어렵다는 점을 고려할 때 사업자에게 과도한 부담을

5) 참고로, 개정안과 유사한 취지에서 인공지능 활용 과정에서 저작물을 허락 없이 이용하는 사례를 방지하기 위하여 교육·조사·연구 등 비상업적·공익적 목적의 정보분석에 한정하여 저작물을 복제·전송할 수 있도록 명시하려는 「저작권법 일부개정법률안」(의안번호 제2208071호, 김태선의원 대표발의)이 현재 문화체육관광위원회에 계류되어 있음.

6) 백지연, ‘생성형 AI의 학습데이터 공개 관련 논의와 입법 과제 - 특정 화풍의 이미지 생성 사안을 중심으로-’, “국회입법조사처 이슈와 논점” 제2355호, 2025.5.7., p. 3

초래할 우려가 있다고 보임.

- 과학기술정보통신부는 창작자의 권리를 보호하려는 취지에는 공감하나 학습용데이터 공개는 기업의 영업비밀 및 기술자산과 관련된 사항으로 국내 인공지능산업계에 과도한 규제로 작용할 수 있고, 인공지능 학습 과정에서 개별 저작물의 이용 여부를 확인하는 것은 기술적으로 매우 복잡하고 많은 비용이 소요될 뿐만 아니라 기업들이 계속적으로 저작권 분쟁에 연일 불확실한 상황에 놓이게 됨에 따라 국내 인공지능산업의 위축이 우려되며, 미국, 일본, 유럽연합 등 주요국에서도 학습용데이터 공개 관련 의무는 규정하지 않거나 제한적으로 규정하고 있어⁷⁾, 주요국의 규제 동향과 법체계적 정합성을 고려하여 신중한 검토가 필요하다는 입장임.
- 문화체육관광부는 창작자 보호라는 개정안의 취지를 분명히 하기 위하여 학습용데이터 정보 ‘공개 노력’ 의무를 ‘공개’ 의무로 강화할 필요가 있고, 생성형 인공지능 외의 인공지능 학습 과정에서도 저작물등이 이용될 수 있으므로 인공지능 전체에 대하여 학습용데이터 정보를 공개하도록 하여야 한다는 의견을 제출함.
- 한국인터넷기업협회는 생성형 인공지능에 사용되는 학습데이터의 양이 방대하고 그 출처가 다양하여 기술적으로 데이터의 분리·추적이 어렵고, 해외사업자에 대하여 관련 규제를 집행하기 어려운 현실을 고려할 때

7) 유럽연합 인공지능법은 생성형 인공지능 학습데이터의 요약서를 공개하도록 하고 있으며, 영국, 중국, 일본 등의 경우 학습용데이터의 공개나 개별 저작물 이용 여부를 사후적으로 확인하도록 하는 규정은 없다는 의견임.

국내 기업의 경쟁력 저하가 우려된다는 의견을 제출함.

- 한국방송협회는 학습데이터 정보 공개가 권리자의 알 권리 및 대가 청구와 관련되어 있음을 고려할 때, 학습데이터 정보 공개를 권고사항이 아닌 의무사항으로 규정할 필요가 있고, 인공지능 학습을 위하여 저작물을 사용하는 경우 사전에 권리자에게 허락을 받아야 함(옵트인)에도, 개정안에 따른 저작물등의 이용 여부 확인 절차는 인공지능사업자가 사전 허락 없이 저작물등을 학습데이터로 이용하고, 사후적으로 권리자가 요구하는 경우에만 해당 저작물등을 학습데이터에서 제외(옵트아웃)할 수 있는 것으로 해석될 여지가 있어 우려된다는 의견을 제출함.

문 의 처
02)6788-5230