



제 439회 국회 (정기회)
제1차 과학기술정보방송통신위원회

인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 일부개정법률안

검 토 보 고

〈인공지능시스템 결과물 표시 훼손·위작·변작 제재〉

■ 김대식의원 대표발의(의안번호 제2218919호)

2026. 9.

과학기술정보방송통신위원회

수석전문위원 이지민

목 차

I . 제안경위	1
II . 제안이유 및 주요내용	1
III . 검토의견	2
1. 인공지능시스템 결과물 표시 훼손·위작·변작 제재 등	2

I. 제안경위

가. 발 의 자 : 김대식의원 등 10인

나. 발 의 일 : 2026. 5. 11.

다. 회 부 일 : 2026. 5. 12.

II. 제안이유 및 주요내용

현행법은 인공지능시스템을 이용한 가상의 음향, 이미지 또는 영상 등의 결과물이 인공지능시스템을 이용하여 생성되었다는 사실을 이용자가 명확하게 인식할 수 있는 방식으로 고지 또는 표시할 의무를 인공지능사업자에게 부과하고 있음.

그러나, 현행법은 고지 또는 표시의 대상을 한정하고 있지 않아 유통 과정의 가공 등으로 인하여 해당 결과물이 인공지능시스템을 이용하여 생성되었다는 사실이 이용자에게 전달되지 아니하고 사람이 직접 만들거나 실제로 존재하는 정보로 유통될 위험이 있음.

따라서, 해당 결과물 자체에 코드 또는 문자·부호 등을 삽입한 비가시적 또는 가시적 워터마크를 생성하여 인공지능시스템을 이용한 생성 사실을 고지·표시하도록 하고, 이와 같은 고지·표시를 훼손 또는 위작·변작하는 행위를 금지하면서 위반 시 2년 이하의 징역 또는 2천만원 이하의 벌금에 처함으로써 인공지능 투명성 규제의 사각지대를 해소하려는 것임(안 제31조제3항·제4항 및 제42조제2항 신설).

Ⅲ. 검토의견

1. 인공지능시스템 결과물 표시 훼손·위작·변작 제재 등

가. 개정안의 내용

- 개정안은 인공지능사업자가 인공지능시스템을 이용하여 실제와 구분하기 어려운 가상의 음향, 이미지 또는 영상 등의 결과물(이하 “딥페이크 생성물”)을 제공하는 경우 ‘이용자가 명확하게 식별할 수 있도록 해당 결과물 자체에 코드 또는 문자·부호 등을 삽입하는 방식’으로 고지 또는 표시하도록 구체적으로 규정하고(안 제31조제3항), 해당 고지 또는 표시의 훼손·위작·변작을 금지하며, 이를 위반한 경우 2년 이하의 징역 또는 2천만원 이하의 벌금형 대상으로 하려는 것임(안 제31조제4항 및 제42조제2항 신설).

현행	개정안
제31조(인공지능 투명성 확보 의무) ①·② (생략) ③ 인공지능사업자는 인공지능시스템을 이용하여 실제와 구분하기 어려운 가상의 음향, 이미지 또는 영상 등의 결과물을 제공하는 경우 해당 결과물이 인공지능시스템에 의하여 생성되었다는 사실을 이용자가 명확하게 <u>인식할 수 있는 방식으로</u> 고지 또는	제31조(인공지능 투명성 확보 의무) ①·② (현행과 같음) ③ ----- ----- ----- ----- ----- ----- -----식별할 수 있도록 해당 결과물 자체에

현행	개정안
표시하여야 한다. 이 경우 해당 결과물이 예술적·창의적 표현물에 해당하거나 그 일부를 구성하는 경우에는 전시 또는 향유 등을 저해하지 아니하는 방식으로 고지 또는 표시할 수 있다. <신설> ④ (생략) 제42조(벌칙) (생략) <신설>	<u>코드 또는 문자·부호 등을 삽입하는 방식으로</u>----- -----. -----. ④ 누구든지 제3항에 따른 고지 또는 표시를 훼손하거나 위작 또는 변작하여서는 아니 된다. ⑤ (현행 제4항과 같음) 제42조(벌칙) ① (현행 제6조 제1항 제2호의2와 같음) ② 제31조제4항을 위반하여 고지 또는 표시를 훼손하거나 위작 또는 변작한 자는 2년 이하의 징역 또는 2천만원 이하의 벌금에 처한다.
부칙	
이 법은 공포 후 6개월이 경과한 날부터 시행한다.	

다. 검토의견

- 현행법 제31조제3항은 인공지능사업자로 하여금 인공지능시스템을 이용하여 딥페이크 생성물에 해당하는 결과물을 제공하는 경우 인공지

능시스템에 의하여 생성되었다는 사실을 이용자가 명확하게 인식할 수 있는 방식으로 고지 또는 표시하도록 하고 있음.

- 그리고, 인공지능사업자가 제31조제3항의 표시 의무를 위반한 경우 제40조제1항과 제3항에 따라 사실조사 및 중지·시정명령의 대상이 되며 해당 명령을 이행하지 않는 경우 3천만원 이하의 과태료¹⁾를 부과함.
- 이는 생성형 인공지능의 발전 및 확산에 따라 이를 활용하여 생성된 사진·영상 등에 의한 명예훼손 및 허위조작정보 유통 등의 문제가 발생하고 있어 이에 따른 피해와 부정적인 사회적 영향을 최소화하기 위한 것임.
- 그러나, 인공지능사업자가 인공지능으로 생성한 결과물임을 표시하더라도 이후 이용자가 그 결과물을 유통하거나 가공하는 과정에서 표시를 제거하거나 인식이 어려울 만큼 훼손하는 경우에 대해서는 법적으로 금지하는 명문의 규정이 없고 처벌 대상도 아닌 상황임.

1) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

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. 제31조제2항·제3항, 제32조제1항·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 과학기술정보통신부장관은 제1항에 따른 조사를 위하여 필요한 경우 소속 공무원으로 하여금 인공지능사업자의 사무소·사업장에 출입하여 장부·서류, 그 밖의 자료나 물건을 조사하게 할 수 있다. 이 경우 조사의 내용·방법 및 절차 등에 관하여 이 법에서 정하는 사항을 제외하고는 「행정조사기본법」에서 정하는 바에 따른다.

③ 과학기술정보통신부장관은 제1항 및 제2항에 따른 조사 결과 인공지능사업자가 이 법을 위반한 사실이 있다고 인정되면 인공지능사업자에게 해당 위반행위의 중지나 시정을 위하여 필요한 조치를 명할 수 있다.

제43조(과태료) ① 다음 각 호의 어느 하나에 해당하는 자에게는 3천만원 이하의 과태료를 부과한다.

3. 제40조제3항에 따른 중지명령이나 시정명령을 이행하지 아니한 자

- 이에 개정안은 인공지능사업자에게 딥페이크 생성물에 대한 표시 의무를 부과하는 것에서 나아가 이용자를 포함한 모두에게 그 표시를 훼손하거나 위작·변작하는 것을 금지하고, 이를 위반한 경우 2년 이하의 징역 또는 2천만원 이하의 형사처벌 대상으로 규정함으로써 인공지능의 투명성을 확보하려는 표시 의무 규정의 실효성을 제고하려는 것임.
- 또한, 개정안은 딥페이크 생성물에 대한 표시 방식을 ‘해당 결과물 자체에 코드 또는 문자·부호 등을 삽입하는 방식’으로 법률에 구체적으로 명시하여 수범자의 예측가능성을 높이려는 것임.
- 인공지능으로 만들어진 결과물이 정보통신망에서 수시로 유통되고 그 과정에서 인공지능 생성물이라는 표시가 변질·오염될 우려가 있다는 점을 고려할 때, 인공지능사업자에 대한 제재만으로는 딥페이크 생성물 표시 의무를 규정한 당초의 입법 취지를 달성하기 어려우므로, 개정안과 같이 이미 인공지능사업자가 표시한 내용을 의도적으로 숨기거나 변형하는 것을 제재할 수 있는 법적 근거를 마련할 필요성이 인정될 수 있을 것으로 보임.
- 다만, 개정안은 다음과 같은 점을 고려할 필요가 있음.
- 첫째, 현재는 인공지능사업자가 딥페이크 생성물에 표시를 하지 않더라도 과태료 부과 대상일 뿐 형사처벌 대상으로 규정하지 않고 있는데, 안 제31조제4항 및 제42조제2항과 같이 불특정 다수인 이용자의 표시 훼손·위작·변작 행위에 대해서 형사처벌 하도록 하여 사업자보다 무거

운 제재를 하는 것이 법익 균형에 맞지 않는 과도한 제재에 해당하지 않는지 검토가 필요한 것으로 보임.

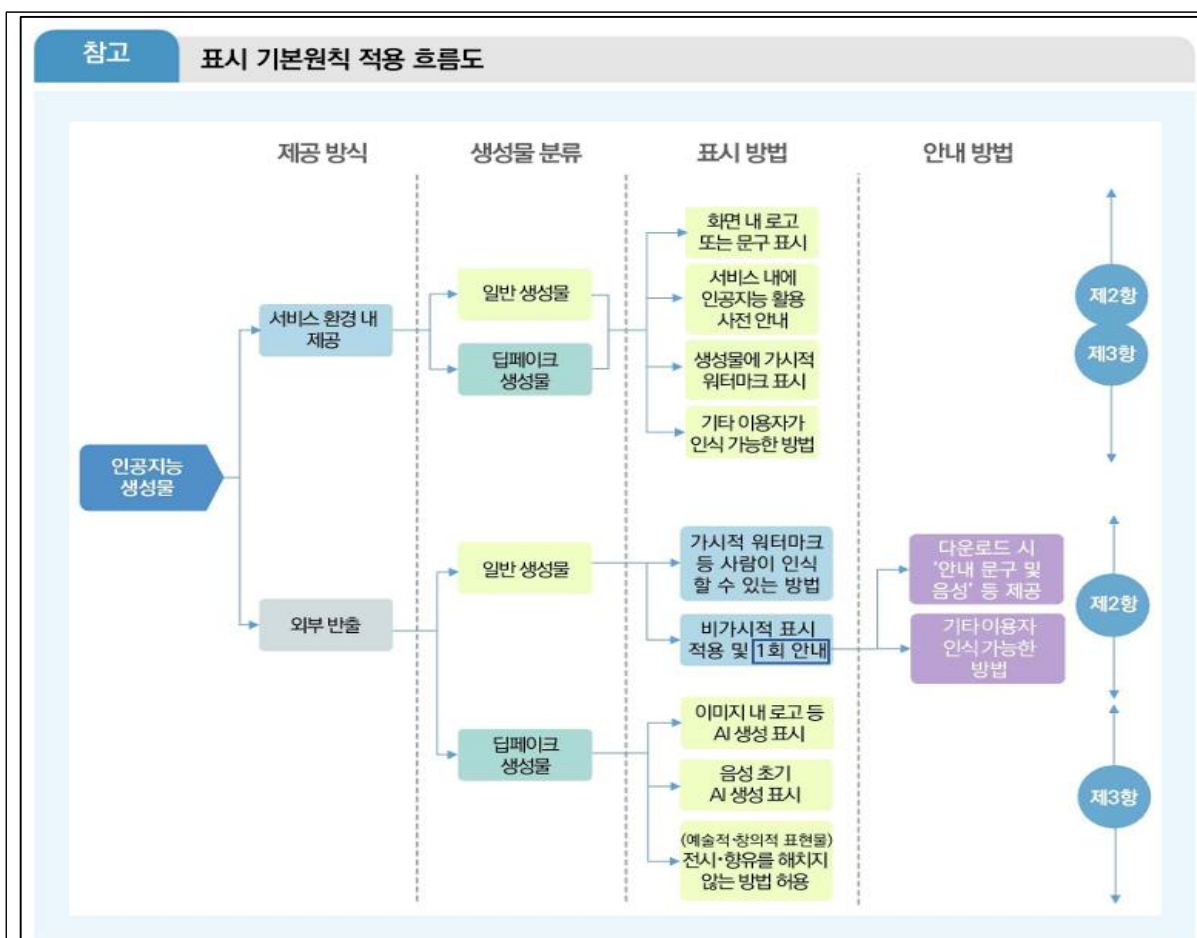
- 현행법에 따르면, 인공지능사업자가 딥페이크 생성물 표시 의무를 이행하지 않은 경우 1차적으로 사실조사 및 시정명령의 대상이고, 시정명령을 이행하지 않는 경우 과태료가 부과됨.
- 둘째, 안 제31조제4항 및 제42조제2항에 따르면 의도적으로 인공지능 생성물 표시를 훼손하거나 위작·변작하는 경우뿐만 아니라 고의가 없이 단순 과실로 훼손한 경우에도 형사처벌 대상이 될 수 있는데, 의도하지 않는 훼손에 대해서도 처벌하는 것이 적절한지 검토가 필요함.
- 개정안에 따라 금지되는 표시의 훼손, 위작, 변작 중 위작 및 변작은 행위자의 의도를 전제로 하나 훼손은 의도 없이도 발생할 수 있고 특히, 정보통신망에서 반복적으로 유통되는 과정에서 훼손될 가능성이 있다는 점을 고려할 때,

의도하지 않는 훼손을 고의적인 훼손·위작·변작과 동일한 형사처벌 대상으로 규정하는 것은 과도한 측면이 있음.
- 셋째, 안 제31조제3항은 딥페이크 생성물의 투명성 고지·표시 방식을 결과물 자체에 부호 등을 삽입하는 방식으로 구체화하려는 것인데, 현재 정부가 가이드라인을 통해 허용하는 방식에 비하여 고지·표시 방식이 축소될 것으로 보여, 이에 대한 검토가 필요함.
- 현행법에 따른 사업자 의무의 구체적인 이행기준을 제시하고 있는

「인공지능 투명성 확보 가이드라인」에서는 결과물 자체에 표시를 삽입하는 방식뿐만 아니라 결과물과 연계된 별도의 표시 등 다양한 방식도 허용하고 있음.

- 구체적으로 ① 결과물을 다운로드·공유 등 서비스 환경 외부로 반출하는 기능을 제공하는 경우에는 반드시 ‘결과물 내에 삽입’하는 방식을 적용하되, ② 결과물을 서비스 이용 환경 내에서만 제공하는 경우는 화면 UI에 표시, 사전 안내 등 간접적인 고지·표시 방법을 폭넓게 인정하고 있음.

<인공지능 투명성 확보 표시 기본원칙 적용 흐름도>



자료 : 과학기술정보통신부

- 즉, 딥페이크 생성물이더라도 서비스 환경 내에서 제공되는 경우에는 결과물 자체에 표시를 삽입하지 않고 서비스 이용 과정에서 이용자가 인식할 수 있게 고지·표시하는 방식도 허용됨.
- 따라서, 안 제31조제3항과 같이 딥페이크 생성물 표시방법을 결과물 자체에 삽입하는 방식으로 한정할 경우, 현재 허용되고 있는 표시수단의 범위보다 표시·고지 방식이 축소될 수 있고 이에 따라 사업자의 의무 이행에 대한 부담이 증가할 우려가 있음.
- **넷째**, 현행법 제31조제2항은 일반적인 인공지능 생성물에 대해서는 가시적·비가시적 표시를 모두 허용²⁾하는 반면, 제3항은 딥페이크 생성물에 대해서는 이용자가 쉽게 인식할 수 있도록 가시적 표시를 중심으로 규율하고 있음
- 그런데, 안 제31조제3항에서 예시로 규정한 방식 중 ‘코드를 삽입하는 방식’은 이용자가 직접 인식하기 어려운 비가시적 표시방식으로 해석될 가능성도 있으므로, 딥페이크 생성물에 대해서는 가시적 표시를 의무화한 현행 규율체계 및 입법취지와 부합하는지에 대한 검토가 필요할 것으로 보임.
- **넷째**, 현재 개정안과 유사한 취지로 인공지능 생성물을 활용한 정보가 정보통신망에 게시되는 경우 게시자가 생성물 표시를 훼손하거나

2) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

제31조(인공지능 투명성 확보 의무) ② 인공지능사업자는 생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 경우 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 한다.

위조·변조하는 행위를 금지하는 「정보통신망 이용촉진 및 정보보호 등에 관한 법률 일부개정법률안」³⁾이 우리 위원회에 계류 중으로, 인공지능 생성물 표시 훼손 등에 대한 제재의 근거를 현행법과 정보통신망법 중 어디에 규정할지 입법정책적인 검토가 필요함.

다. 정부 및 관련 단체 의견

1) 과학기술정보통신부: 신중검토

- 개정안은 규제 대상의 포괄성으로 인해 콘텐츠 산업 등 인공지능 생성물을 활용하는 산업의 위축을 초래할 수 있을 뿐만 아니라 일반 국민에게도 광범위하게 적용되는 사항으로, 인공지능기본법 시행 초기라는 현재 상황을 고려할 때 신규 규제 도입은 신중한 검토가 필요함.

2) 방송미디어통신위원회: 신중검토

- 인공지능기본법은 인공지능사업자를 중심으로 AI 서비스 자체의 안전·신뢰성 확보를 위해 최소한의 규제를 두고 있어 AI 이용자 일반에 대한 규제에 적합하지 않다고 보이며, 인공지능 생성물의 유통 과정에서 발생할 수 있는 부작용을 방지하기 위한 규제는 건전한 정보통신 이용환경 조성을 목적으로 하는 정보통신망법을 통해 마련되는 것이 적절함.

3) 2026년 8월 말 기준, 의안번호 2213740호(박정훈의원 대표발의), 의안번호 2216186호(조인철의원 대표발의), 의안번호 2216369호(서영석의원 대표발의), 의안번호 2216849호(이주희의원 대표발의), 의안번호 2219336호(박충권의원 대표발의) 등

3) 한국인터넷기업협회: 신중검토

- 인공지능으로 생성된 결과물이 업로드되는 과정에서 자동 압축되거나 편집 과정에서 일부가 삭제되는 등 비의도적인 표시 변경 행위까지 제재 대상으로 포함될 여지가 있어 금지·처벌 대상의 범위를 고의적이고 적극적인 표시 조작 행위 및 중대한 소비자 오인 유발 행위 등으로 한정할 필요가 있고,

현행법에서는 예술적·창의적 표현물에 대해서는 별도의 고지·표시를 허용하는 만큼 개정안과 같이 표시의 훼손·위작·변작을 금지할 때도 예술적·창의적 표현 목적이 있는 경우 별개의 기준을 적용할지 논의가 필요함.

문 의 처
02)6788-5230