



제 439회 국회 (정기회)
제1차 과학기술정보방송통신위원회

인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 일부개정법률안

검 토 보 고

< 고영향 인공지능으로 인한 사이버위협 예방 및 대응 >

■ 김성원의원 대표발의(의안번호 제2220395호)

2026. 9.

과학기술정보방송통신위원회

수석전문위원 이지민

목 차

I . 제안경위	1
II . 제안이유 및 주요내용	1
III . 검토의견	2
1. 고영향 인공지능으로 인한 사이버위협 예방 및 대응	2

I. 제안경위

가. 발 의 자 : 김성원의원 등 10인

나. 발 의 일 : 2026. 8. 3.

다. 회 부 일 : 2026. 8. 4.

II. 제안이유 및 주요내용

최근 고성능 인공지능 모델의 급속한 발전으로 인공지능을 악용한 사이버공격, 대규모 허위정보 생성·유포, 국가 중요 데이터 침해 등 새로운 형태의 국가안보 및 공공안전 위협이 현실화되고 있음.

그러나 현행 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」은 고영향 인공지능에 관한 일반적인 사항만을 규정하고 있을 뿐, 국가안보와 공공안전에 중대한 영향을 미칠 우려가 있는 고위험 인공지능 모델에 대한 별도의 안전관리 및 사이버위협 대응체계는 충분히 마련되어 있지 않은 실정임.

이에 국가가 국가안보 및 공공안전에 중대한 위협을 초래할 우려가 있는 고위험 인공지능 모델을 지정·관리할 수 있는 법적 근거를 마련하고, 범국가적 차원의 고영향 인공지능 사이버위협 대응체계를 구축할 수 있도록 함으로써 국가 핵심 기반시설과 국민의 안전을 보다 두텁게 보호하려는 것임(제33조의2, 제34조제1항제6호 및 제7호 신설).

Ⅲ. 검토의견

1. 고영향 인공지능으로 인한 사이버위협 예방 및 대응

가. 개정안의 내용

- 개정안은 국가가 고영향 인공지능으로 인한 사이버위협의 예방 및 대응체계 구축, 고영향 인공지능의 개발·활용에 이용되는 공공데이터 및 학습용데이터의 보호 방안 등의 시책을 마련하도록 하고(안 제33조의2 신설), 인공지능사업자의 책무에도 고영향 인공지능 개발·활용에 이용되는 공공데이터 및 학습용데이터의 보호·안전관리 조치를 추가하려는 것임(안 제34조제1항 제6호 및 제7호 신설).

현행	개정안
<u><신설></u>	<u>제33조의2(고영향 인공지능과 관련한 국가의 시책) ① 국가는 고영향 인공지능으로 인하여 발생할 수 있는 위험 예방 및 대응을 위하여 시책을 마련하여야 한다.</u> <u>② 제1항에 따른 시책에는 다음 각 호의 사항이 포함되어야 한다.</u> <u>1. 고영향 인공지능으로 인한 사이버위협의 예방 및 대응체계 구축</u> <u>2. 고영향 인공지능의 개발·활용에 이용되는 공공데이터의 보호 방안</u>

현행	개정안
제34조(고영향 인공지능과 관련한 사업자의 책무) ① 인공지능사업자는 고영향 인공지능 또는 이를 이용한 제품·서비스를 제공하는 경우 고영향 인공지능의 안정성·신뢰성을 확보하기 위하여 다음 각 호의 내용을 포함하는 조치를 대통령령으로 정하는 바에 따라 이행하여야 한다. 1. ~ 5. (생략) <u><신설></u> <u><신설></u> 6. (생략) ② (생략)	3. <u>고영향 인공지능의 개발·활용에 이용되는 학습용데이터의 보호 방안</u> 4. <u>그 밖에 고영향 인공지능으로 인한 위험의 예방 및 대응을 위하여 대통령령으로 정하는 사항</u> 제34조(고영향 인공지능과 관련한 사업자의 책무) ① ----- ----- ----- ----- ----- ----- ----- 1. ~ 5. (현행과 같음) 6. <u>고영향 인공지능의 개발·활용에 이용되는 공공데이터의 보호 및 안전관리</u> 7. <u>고영향 인공지능의 개발·활용에 이용되는 학습용데이터의 보호 및 안전관리</u> 8. (현행 제6호와 같음) ② (현행과 같음)
부칙	
이 법은 공포 후 6개월이 경과한 날부터 시행한다.	

나. 검토의견

- 최근 인공지능기술의 발달로 인공지능의 성능이 고도화되면서 인공지능을 이용한 사이버위협¹⁾의 심각성이 대두되고 있음.
- 2026년 4월, 엔트로픽이 공개한 Mythos 모델이 기존 보안체계의 취약점을 상당수 탐지할 뿐만 아니라 이를 공격할 수 있는 코드까지 스스로 작성하고 이를 검증하는 능력까지 보여준 바 있음.
- 그러나, 현행법은 인공지능을 기반으로 하는 사이버위협에 대한 예방·대응에 관한 사항을 명시적으로 규정하고 있지 않으며, 인공지능 기반 위협을 포함한 사이버위협에 관한 사항은 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」, 「정보통신기반 보호법」, 「국가정보원법」 등 정보보호와 관련된 개별법에서 규정하고 있음.
- 정보통신망을 대상으로 하는 사이버위협은 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」에 따른 침해사고¹⁾에 해당하여 정보통신망 침해 등 금지행위(제48조), 침해사고 대응조치(제48조의2), 침해사고 신고(제48조의3), 원인분석(제48조의4) 등의 규정이 적용됨.
- 그리고, 정보통신기반시설을 대상으로 하는 사이버위협은 「정보통신기반

1) 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

7. "침해사고"란 다음 각 목의 방법으로 정보통신망 또는 이와 관련된 정보시스템을 공격하는 행위로 인하여 발생한 사태를 말한다.

가. 해킹, 컴퓨터바이러스, 논리폭탄, 메일폭탄, 서비스거부 또는 고출력 전자기파 등의 방법

나. 정보통신망의 정상적인 보호·인증 절차를 우회하여 정보통신망에 접근할 수 있도록 하는 프로그램이나 기술적 장치 등을 정보통신망 또는 이와 관련된 정보시스템에 설치하는 방법

보호법」에 따른 침해사고에 해당²⁾하여 정부는 기반시설 침해행위 등 금지(제12조), 침해사고 통지(제13조), 복구조치(제14조), 대책본부의 구성(제15조) 등에 기반하여 대응 중임.

- 또한, 중앙행정기관, 공공기관 등 공공부문을 대상으로 하는 사이버위협은 국가정보원이 「국가정보원법」³⁾ 및 「사이버안보 업무규정」⁴⁾에 따라 예방·대응 업무를 수행 중임.

○ 개정안은 국가의 시책과 인공지능사업자의 책무에 고영향 인공지능 기반 사이버위협 대응과 고영향 인공지능의 개발·활용을 위해 투입되는 데이터의 보호에 관한 사항을 추가하려는 것으로,

국가가 고영향 인공지능으로 인하여 발생할 수 있는 사이버위협 문제에

2) 「정보통신기반 보호법」

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

2. "전자적 침해행위"란 다음 각 목의 방법으로 정보통신기반시설을 공격하는 행위를 말한다.

가. 해킹, 컴퓨터바이러스, 논리·메일폭탄, 서비스거부 또는 고출력 전자기파 등의 방법

나. 정상적인 보호·인증 절차를 우회하여 정보통신기반시설에 접근할 수 있도록 하는 프로그램이나 기술적 장치 등을 정보통신기반시설에 설치하는 방법

3. "침해사고"란 전자적 침해행위로 인하여 발생한 사태를 말한다.

3) 「국가정보원법」

제4조(직무) ① 국정원은 다음 각 호의 직무를 수행한다.

4. 다음 각 목의 기관 대상 사이버공격 및 위협에 대한 예방 및 대응

가. 중앙행정기관(대통령 소속기관과 국무총리 소속기관을 포함한다) 및 그 소속기관과 국가인권위원회, 고위공직자범죄수사처 및 「행정기관 소속 위원회의 설치·운영에 관한 법률」에 따른 위원회

나. 지방자치단체와 그 소속기관

다. 그 밖에 대통령령으로 정하는 공공기관

5. 정보 및 보안 업무의 기획·조정

4) 「사이버안보 업무규정」

제3조(사이버안보 업무의 수행) 국가정보원은 사이버안보를 위하여 다음 각 호의 업무(이하 "사이버안보 업무"라 한다)를 수행한다.

2. 사이버보안 업무

가. 법 제4조제1항제4호 각 목의 기관(이하 "중앙행정기관등"이라 한다)을 대상으로 하는 사이버 공격·위협에 대한 예방 및 대응 업무

나. 법 제4조제1항제5호에 따라 수행하는 사이버공격·위협에 대한 예방 및 대응 관련 기획·조정 업무

체계적으로 대비할 수 있도록 하고 고영향 인공지능에 대한 인공지능사업자의 책임을 강화하려는 취지로 보임.

- 다만, 개정안은 다음과 같은 사항을 고려할 필요가 있음.
- 첫째, 현행법상 고영향 인공지능 해당 여부가 인공지능의 성능이 아니라 인공지능이 활용되는 영역을 기준으로 판단된다는 점을 고려할 때, 개정안이 고영향 인공지능을 대상으로 규율하는 것이 인공지능 성능 향상에 따라 증가하는 인공지능 기반 사이버위협에 대응한다는 개정안의 취지에 부합하는지 검토할 필요가 있음.
- 현행법 제2조제4호에 따르면 ‘고영향 인공지능’이란 ‘사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 인공지능시스템’으로 각 목에 나열된 영역에서 활용되는 인공지능⁵⁾임.

5) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

- 4. “고영향 인공지능”이란 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 인공지능시스템으로서 다음 각 목의 어느 하나의 영역에서 활용되는 것을 말한다.
 - 가. 「에너지법」 제2조제1호에 따른 에너지의 공급
 - 나. 「먹는물관리법」 제3조제1호에 따른 먹는물의 생산 공정
 - 다. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계의 구축·운영
 - 라. 「의료기기법」 제2조제1항에 따른 의료기기 및 「디지털의료제품법」 제2조제2호에 따른 디지털의료기기의 개발 및 이용
 - 마. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항제1호에 따른 핵물질과 같은 항 제2호에 따른 원자력시설의 안전한 관리 및 운영
 - 바. 범죄 수사나 체포 업무를 위한 생체인식정보(얼굴·지문·홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적·생리적·행동적 특징에 관한 개인정보를 말한다)의 분석·활용
 - 사. 채용, 대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가
 - 아. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영
 - 자. 공공서비스 제공에 필요한 자격 확인 및 결정 또는 비용징수 등 국민에게 영향을 미치는 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」 제4조에 따른 공공기관 등(이하 “국가기관등”이라 한다)의 의사결정
 - 차. 「교육기본법」 제9조제1항에 따른 유아교육·초등교육 및 중등교육에서의 학생 평가
 - 카. 그 밖에 사람의 생명·신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역으로서 대통령령으로 정하는 영역

- 한편, 현행법은 고성능 인공지능을 정의하고 있지는 않으나, 제32조에서 ‘학습에 사용된 누적 연산량이 대통령령으로 정하는 기준 이상인 인공지능시스템’에 대해서 인공지능사업자에게 안전성 확보 의무를 부과⁶⁾하여 일정 수준 이상의 성능을 갖춘 인공지능에 대해서 별도로 규제하고 있음.
- 둘째, 현재 정부는 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」, 「정보통신기반 보호법」, 「국가정보원법」 등의 개별법에 따른 사이버 위협 예방·대응 체계를 통해 이미 인공지능 사이버위협을 관리하고 있는 바, 안 제33조의2와 같이 인공지능 사이버위협에 대하여 국가가 별도로 시책을 마련하도록 할 실익이 크지 않을 것으로 보임.
- 참고로, 과학기술정보통신부는 인공지능으로부터의 사이버위협에 체계적으로 대응하고자 2026년 5월 과학기술관계장관회의에서 「AI 기반 사이버 위협에 대응하기 위한 민간 정보보호 추진계획」을 발표하였음.
- 셋째, 현행법 제34조제1항제1호에 따라 인공지능사업자가 수립·운영하는 ‘위험관리방안’에 고영향 인공지능의 개발을 위하여 수집된 데이터 관리에 관한 사항이 포함되어 있는 것으로 보이는바,
- 안 제34조제1항제6호 및 제7호에 따라 추가되는 ‘고영향 인공지능의 개

6) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

제32조(인공지능 안전성 확보 의무) ① 인공지능사업자는 학습에 사용된 누적 연산량이 대통령령으로 정하는 기준 이상인 인공지능시스템의 안전성을 확보하기 위하여 다음 각 호의 사항을 이행하여야 한다.

1. 인공지능 수명주기 전반에 걸친 위협의 식별·평가 및 완화
 2. 인공지능 관련 안전사고를 모니터링하고 대응하는 위험관리체계 구축
- ② 인공지능사업자는 제1항 각 호에 따른 사항의 이행 결과를 과학기술정보통신부장관에게 제출하여야 한다.
- ③ 과학기술정보통신부장관은 제1항 각 호에 따른 사항의 구체적인 이행 방식 및 제2항에 따른 결과 제출 등에 필요한 사항을 정하여 고시하여야 한다.

발·활용에 이용되는 공공데이터의 보호 및 안전관리’ 및 ‘고영향 인공지능의 개발·활용에 이용되는 학습용데이터의 보호 및 안전관리’ 책무가 현행법에 따른 고영향 인공지능과 관련한 사업자 책무와 중복될 우려가 있어, 이에 대한 검토가 필요함.

- 고영향 인공지능 관련 사업자의 책무 이행에 관한 세부적인 기준을 제시하는 「고영향 인공지능 사업자 책무 가이드라인」에 따르면 고영향 인공지능에 대한 위험관리방안은 인공지능 개발·활용을 위한 데이터 수집, 모델 학습 과정을 포함한 인공지능 수명주기⁷⁾ 전체를 대상으로 하고 있으며, 인공지능사업자가 수집한 데이터의 안전한 관리를 위한 사항을 포함하고 있음.

- 구체적으로, 「고영향 인공지능 사업자 책무 가이드라인」은 인공지능사업자가 데이터 수집 시 이용자 보호를 위해 적법한 절차대로 데이터를 수집 및 관리하고 데이터 유출 및 오용 방지를 위한 기술을 적용하도록 하는 것을 목표로 설정하고 있으며,

자가점검의 내용으로도 ‘데이터의 보유 기간 설정과기 절차 및 방법 등 데이터 관리 계획을 수립하고 준수하였는지’, ‘데이터 유출·오용·손실 방지를 위해 암호화, 접근 권한 관리 등 적절한 기술적·관리적·물리적 보안 방안을 적용하고 있는지’ 등을 명시하고 있음.

- 넷째, 현행법 제2조제12호에 따르면 ‘학습용데이터’란 ‘인공지능의 개발·

7) 「고영향 인공지능 사업자 책무 가이드라인」은 ‘인공지능 수명주기’를 ‘인공지능의 개발 배포 운영 및 폐기에 이르기까지의 전 과정’으로 정의함.

활용 등에 사용되는 데이터’로 민간데이터, 공공데이터를 구분하지 않고 인공지능의 개발·활용 등에 사용되면 모두 학습용데이터에 해당하므로, 안 제33조의2제2항제2호는 제3호와, 안 제34조제1항제6호는 제7호와 그 내용이 중복되는 것으로 보여, 이에 대한 검토가 필요함.

- 즉, 안 제33조의2제2항제3호에 따른 ‘고영향 인공지능의 개발·활용에 이용되는 학습용데이터의 보호 방안’에는 같은 항 제2호에 따른 ‘고영향 인공지능의 개발·활용에 이용되는 공공데이터의 보호 방안’이 포함되고, 안 제34조제1항제7호에 따른 ‘고영향 인공지능의 개발·활용에 이용되는 학습용데이터의 보호 및 안전관리’에는 같은 항 제6호에 따른 ‘고영향 인공지능의 개발·활용에 이용되는 공공데이터의 보호 및 안전관리’가 포함되는 관계임.

다. 정부 및 관련 단체 의견

1) 과학기술정보통신부: 신중검토

- 고영향 인공지능으로 인한 사이버위협과 관련하여 국가의 시책을 마련하고자 하는 취지에는 공감하나, 개정안에 따른 국가 시책의 내용을 구체화할 필요가 있고, 구체화된 내용에 따라 소관 부처 및 관련 법률에 대한 검토가 필요함.
- 그리고, 사이버 위협 및 대응체계 구축 등에 관한 사항은 현재 「국가정보원법」, 「정보통신기반 보호법」, 「정보통신망 이용촉진 및 정보

보호 등에 관한 법률」 등 개별 법률을 통해 규율·운영되고 있어 신규 조항 신설 시 이중규제 및 규제 강화 효과를 야기할 우려가 있음.

- 또한, 고영향 인공지능 사업자의 책무에 데이터 보호·안전관리 의무를 추가하더라도 현행법에서 이미 위험관리방안의 수립·운영(제34조제1항제1호) 등을 통해 데이터 관리를 포함한 전반적인 위험 관리를 포함하여 관리하고 있어 별도 규정을 마련하는 것은 실익이 낮음.

2) 국가정보원: 신중검토

- 인공지능 데이터 보호를 포함한 사이버 위협 및 대응체계 구축 등에 관한 사항은 「국가정보원법」, 「정보통신기반 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 등 각 법률에 따라 이미 규율·운영되고 있는 상황으로, 국가·인공지능사업자에 대한 고영향 인공지능 관련 별도 의무를 신설할 경우, 법 적용·해석의 혼란, 이중 규제 문제 등을 야기할 수 있어 신중한 검토가 필요함.
- 또한, 현행법의 입법 목적은 인공지능산업 활성화이므로 인공지능 기반 사이버 위협 및 대응체계 구축을 현행법에 별도로 규율하는 것은 오히려 규제 강화 효과로 작용하여 입법 취지에 부합하지 않음

문 의 처
02)6788-5230