



제 439회 국회 (정기회)
제1차 과학기술정보방송통신위원회

인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 일부개정법률안

검 토 보 고

〈국가기관등 사업 수행 데이터의 통합제공시스템 연계 강화〉

■ 김현의원 대표발의(의안번호 제2218002호)

2026. 9.

과학기술정보방송통신위원회

수석전문위원 이지민

목 차

I . 제안경위	1
II . 제안이유 및 주요내용	1
III . 검토의견	3
1. 국가기관등 사업 수행 데이터의 통합제공시스템 연계 강화	3

I. 제안경위

가. 발 의 자 : 김현의원 등 14인

나. 발 의 일 : 2026. 3. 31.

다. 회 부 일 : 2026. 4. 1.

II. 제안이유 및 주요내용

현행법은 인공지능 학습용데이터의 수집·활용 및 관련 시책 등을 규정하면서 과학기술정보통신부장관이 인공지능 학습용데이터를 통합적으로 제공·관리할 수 있는 시스템을 구축·운영할 수 있도록 근거 규정을 두고 있음.

그러나 정부 예산으로 수행되는 사업 과정에서 구축·산출된 인공지능 학습용데이터가 개별 기관 내부에 분산 보관되거나 활용이 제한되는 경우가 많아 국가적 차원의 체계적 관리 및 활용이 어려운 실정임. 특히, 인공지능 학습용데이터에 대한 현장 수요가 급격히 증가하는 상황에서 국가 인공지능 전환(AX) 사업에 대규모 재정이 투입되고 있지만 현행 규정만으로는 정부사업 수행기관 등이 산출한 데이터를 통합제공시스템에 연계·제공토록 하는 법적 의무나 지원 근거가 미비하여 데이터 공급 체계의 효율성이 낮은 상황임.

이에 통합제공시스템의 구축·운영에 관한 사항을 명확화하고 국가기관등의 사업 수행 과정에서 구축·산출된 인공지능 학습용데이터를 통합제공시스템에 연계·제공하도록 의무화하는 한편, 이에 필요한 협

력 요청 및 행정적·재정적·기술적 지원 근거를 마련함으로써 인공 지능 학습용데이터의 체계적 공급과 활용 기반을 강화하려는 것임(안 제15조의2 등).

Ⅲ. 검토의견

1. 국가기관등 사업 수행 데이터의 통합제공시스템 연계 강화

가. 개정안의 내용

- 개정안은 국가기관등의 사업을 수행하는 자에게 사업 수행과정에서 구축·산출된 인공지능 학습용데이터를 통합제공시스템에 연계, 제공하도록 하고, 과학기술정보통신부장관이 국가기관등의 장 또는 국가기관등의 사업을 수행하는 자에게 통합제공시스템 연계, 제공 등에 필요한 협력을 요청할 수 있도록 하며,
- 국가기관등 또는 국가기관등의 사업을 수행하는 자에게 통합제공시스템 연계, 제공에 필요한 행정적·재정적·기술적 지원을 할 수 있는 근거를 마련하고,
- 과학기술정보통신부장관으로 하여금 통합제공시스템에 제공된 학습용데이터의 품질을 관리하도록 하려는 것임(안 제15조의2).

현행	개정안
제15조(인공지능 학습용데이터 관련 시책의 수립 등) ① ~ ③ (생략)	제15조(인공지능 학습용데이터 관련 시책의 수립 등) ① ~ ③ (현행과 같음)
④ <u>과학기술정보통신부장관은 학습용데이터 구축사업의 효율적 수행을 위하여 학습용데이터를 통합적으로 제공·관리할 수 있는 시스템(이하 “통합제공시스템”이라</u>	<u><삭 제></u>

현행	개정안
한다)을 구축·관리하고 민간이 자유롭게 이용할 수 있도록 제공하여야 한다. ⑤ 과학기술정보통신부장관은 통합제공시스템을 이용하는 자에 대하여 비용을 징수할 수 있다. ⑥ (생략) <신설>	<삭제> ⑥ (현행과 같음) 제15조의2(인공지능 학습용데이터 통합제공시스템의 구축 및 운영 등) ① 과학기술정보통신부장관은 학습용데이터 구축사업의 효율적 수행을 위하여 학습용데이터를 통합적으로 제공·관리할 수 있는 시스템(이하 “통합제공시스템”이라 한다)을 구축·관리하고 민간이 자유롭게 이용할 수 있도록 제공하여야 한다. ② 국가기관등의 사업을 수행하는 자는 대통령령으로 정하는 특별한 사유가 없는 한 사업 수행과정에서 구축·산출된 인공지능 학습용데이터의 전부 또는 일부를 제1항의 통합제공시스템에 연계, 제공하여야 한다. ③ 과학기술정보통신부장관은 국가기관등의 장 또는 국가기관등의 사업을 수행하는 자에게 제2항에

현행	개정안
	따른 연계, 제공 등에 필요한 협력을 요청할 수 있다. 이 경우 요청을 받은 자는 특별한 사유가 없는 한 이에 따라야 한다. ④ 과학기술정보통신부장관은 제3항에 따른 국가기관등 또는 국가기관등의 사업을 수행하는 자에게 연계, 제공에 필요한 행정적·재정적·기술적 지원을 할 수 있다. ⑤ 과학기술정보통신부장관은 통합제공시스템에 제공된 학습용데이터의 안정적 품질관리 및 적절한 품질수준의 확보를 위하여 필요한 조치를 취하여야 한다. ⑥ 과학기술정보통신부장관은 통합제공시스템을 이용하는 자에 대하여 비용을 징수할 수 있다. ⑦ 그 밖에 통합제공시스템의 구축·관리, 제2항에 따른 연계, 제공의 범위 및 절차와 방법, 제6항에 따른 비용의 징수 등에 필요한 사항은 대통령령으로 정한다.
부칙	
이 법은 공포한 날부터 시행한다.	

다. 검토의견

- 현행법 제15조제4항 및 제5항은 과학기술정보통신부장관이 학습용데이터를 통합적으로 제공·관리할 수 있는 시스템(이하 “통합제공시스템”)을 구축·관리하고, 통합제공시스템을 이용하는 자에 대하여 비용을 징수할 수 있도록 규정하고 있음.
- 이를 근거로, 과학기술정보통신부는 인공지능 학습용데이터를 통합하여 제공하고 관리할 수 있는 시스템인 AI허브를 2017년부터 구축하여 운영하고 있음.
 - AI허브란 AI 학습용 데이터(14개 분야)와 국내외 기관·기업에서 보유한 AI 학습용 데이터를 누구든지 자유롭게 사용할 수 있도록 공개하여 제공하는 시스템으로, 2026년 6월 기준 14개 분야¹⁾에 대한 945종의 데이터셋을 구축하여 제공 중임.
- 최근 여러 분야에서 인공지능 전환(AI)을 추진하는 정부 사업이 수행되면서 다량의 인공지능 학습용데이터가 구축되고 있으나 AI허브에 연계되지 못한 데이터가 일부 있어 학습용데이터로 활용되는데 어려움이 발생하고 있다는 지적이 있음.
- 이에 개정안은 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」 제4조에 따른 공공기관 등(이하 “국가기관등”)의 사업으로 구축·산출된 학습용데이터가 통합제공시스템에 연계·제공되도록 의무화하고,

1) 한국어, 영상이미지, 헬스케어, 교통물류, 재난안전환경, 농축수산, 문화관광, 스포츠, 교육, 로보틱스, 제조, 지식재산, 법률, 금융

이와 관련된 과학기술정보통신부장관의 요청이 있는 경우 국가기관등의 사업을 수행하는 자는 이에 따르도록 함으로써, 공공부문에서 추진하는 사업으로 생긴 학습용데이터가 통합제공시스템에 연계될 수 있도록 적극적으로 지원하는 근거를 마련하려는 것임.

- 국가기관등의 재정이 투입된 사업 수행과정에서 구축·산출된 데이터는 정부 재정의 결과물이므로, 그 공공성을 고려할 때 통합제공시스템을 통해 일반국민이 사용할 수 있도록 공개할 필요성이 인정될 수 있는 것으로 보임.
- 참고로, 기획예산처가 발표한 「2027년도 예산안 및 기금운용계획안 편성지침」에 따르면 각 중앙관서는 AI 학습용데이터의 산출이 예상되는 경우 그 특성에 따라 제공·활용 방안을 사전에 검토하고, 그 방안을 수립할 때 AI허브 등 국가적으로 연계가 용이하고 효과성이 높은 데이터 플랫폼 활용을 우선 검토하도록 하고 있음.
- 다만, 개정안은 다음과 같은 사항을 고려할 필요가 있음.
- 첫째, 안 제15조의2제2항에 따라 인공지능 학습용데이터를 통합제공시스템에 연계·제공할 의무가 부과되는 ‘국가기관등의 사업을 수행하는 자’의 범위가 불분명하여 수범자의 예측가능성이 떨어질 우려가 있고, 국가기관등의 사업을 통해 구축·산출된 학습용데이터를 최종적으로 관리하는 것은 국가기관등의 역할이므로 통합제공시스템 연계·제공 의무대상을 ‘국가기관등’으로 규정하는 방안이 타당한 것으로 보임.

- 둘째, 안 제15조의2제2항에 따라 국가기관등의 사업을 수행하는 자에게 이미 학습용데이터 연계·제공 의무가 부과되고 있는데, 같은 조 제3항에서 국가기관등의 사업을 수행하는 자에게 연계, 제공 등에 필요한 협력 요청을 수인할 의무를 다시 부과하고 있어 유사한 의무가 중복되어 부과되는 문제가 있으므로, 이에 대한 검토가 필요함.
- 셋째, 국가기관등의 사업을 통해 구축된 데이터가 인공지능 학습용데이터에 해당하더라도 개인정보 보호, 저작권 보호, 경영상 영업비밀 유지 등 데이터의 제공이 적절하지 않은 경우가 있을 수 있음.
- 안 제15조의2제2항은 대통령령으로 정하는 특별한 사유가 있는 경우 통합제공시스템에 대한 연계·제공 의무 대상에서 제외하고 있으나 그 사유 일체가 대통령령으로 위임되고 있어 예외 해당 여부가 모호한 문제가 있음.
- 따라서, 국가기관등의 사업으로 구축된 데이터가 통합제공시스템에 연계·제공되지 않을 수 있는 예외 사유로 명백히 규정되어야 할 사항은 법률에 명시하고 그 외에 추가될 수 있는 사유는 대통령령으로 위임하여 규정하는 방안을 검토할 필요가 있음.
- 넷째, 국가기관등의 사업을 통해 구축된 데이터가 통합제공시스템에 연계·제공되어야 할 인공지능 학습용데이터에 해당하는지 사업수행주체가 명확하게 판단할 수 있는지 불분명한 측면이 있음.
- 현행법 제2조제12호는 학습용데이터를 인공지능의 개발·활용 등에 사

용되는 데이터로 정의²⁾하고는 있으나 이것만으로 국가기관등의 사업을 수행하는 자가 가진 데이터가 안 제15조의2에 따라 연계·제공의무가 있는 학습용데이터인지 구분하기 어려울 것으로 보임.

- 따라서, 통합제공시스템에 연계·제공되어야 하는 학습용데이터의 기준을 명확하게 마련하고 사업수행주체가 사업을 착수하기 전에 미리 학습용데이터로 연계·제공해야 하는 대상임을 고지하는 등 수범자가 미리 파악할 수 있도록 하위 법령에 세부적인 기준, 절차 등을 마련할 필요가 있음.
- 다섯째, 개정안은 현행법 제15조제4항에 따른 통합제공시스템 구축·관리 및 제5항에 따른 통합제공시스템 이용 비용 징수 근거를 안 제15조의2제1항과 제6항으로 각각 이동시키므로, 현행법 제15조제4항 및 제5항에 관한 사항을 대통령령으로 위임하도록 한 현행법 제15조제6항³⁾에 대한 수정이 필요함.
- 여섯째, 개정안의 시행일은 공포 후 즉시로 규정되어 있는데 안 제15조의2제7항에 따라 통합제공시스템 관련 사항이 대통령령에 위임될 것인바, 위임되는 사항을 반영하는 시행령 개정에 필요한 유예기간을 고려하여 시행일을 수정하는 방안을 검토할 필요가 있음.

2) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.

12. "학습용데이터"란 인공지능의 개발·활용 등에 사용되는 데이터를 말한다.

3) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

제15조(인공지능 학습용데이터 관련 시책의 수립 등) ⑥ 그 밖에 제2항에 따른 지원대상사업의 선정 및 지원, 학습용데이터 구축사업의 시행, 통합제공시스템의 구축·관리 및 제5항에 따른 비용의 징수 등에 필요한 사항은 대통령령으로 정한다.

다. 정부 및 관련 단체 의견

1) 과학기술정보통신부: 수용

- 실제 공공·민간부문 현장의 인공지능 학습용데이터 부족 문제에 대응하여 국가 차원에서 인공지능 학습용데이터 관리를 효율화하려는 입법 취지에 동의함.

2) 산업통상부: 수정 수용

- 국가기관등의 사업으로 산출된 학습용데이터가 영업비밀 등에 해당할 수 있어 의무화보다는 데이터 제공 근거를 마련하는 것으로 완화하거나, 의무화는 유지하되 사업수행자가 영업비밀 등의 사유가 있는 경우 데이터를 제공하지 않을 수 있도록 수정이 필요하며, 행정적·기술적·재정적 지원의 주체를 정부로 변경할 필요가 있음.

3) 국가데이터처: 수정 수용

- 통합제공시스템은 민간뿐 아니라 공공에도 제공될 필요가 있으므로 공공부문도 제공 대상에 포함되도록 문구를 수정하고, 개정안과 같이 학습용데이터 미제공 사유를 대통령령으로 일괄 위임하기보다는 민간 정보나 영업비밀 등에 대해서는 구체적으로 명시할 필요가 있음.

문 의 처
02)6788-5230