# 입법의견서

발행일 2025. 10. 01.

# 인공지능기본법 시행령초안에 대한 시민사회 입법의견서

▲고위험 이용사업자 책무 면제 ▲적용 예외 확대 ▲사실조사 및 과태료 유예 등 인공지능 영향 받는 시민의 안전과 인권 보호 장치 부재

일시 | 2025년 10월 1일 (수)

제출 | 민주사회를 위한 변호사 모임 디지털정보위원회 · 정보인권연구소 · 진보네트워크센터 · 참여연대

# 목차

# 들어가며

2025년 9월 25일 이재명 대통령은 유엔 안전보장이사회 공개토의에서 AI(인공지능)와 관련해 "이 무시무시한 도구가 통제력을 상실한다면 허위 정보가 넘쳐나고 테러, 사이버 공격이 급증하는 디스토피아의 미래를 피할 수 없을 것"이라고 밝혔다. 이러한 자신감 있는 발언을 할 수 있었던 배경에는 우리나라가 최소한이나마 인공지능의 위험을 규제하는 '인공지능 발전과 신뢰 기반 조성 등에 관한 기본법'(인공지능기본법)을 보유하고 있다는 점이 작용했을 것이다.

실제로 유엔 인권최고대표가 2025년 인권이사회에서 발표한 <인공지능 관련 활동을 포함한 기술 기업의 활동에 대한 기업 및 인권에 관한 지침 원칙의 실제적 적용(A/HRC/59/32)> 보고서에서는 한국을 유럽연합, 브라질과 함께 위험기반접근을 취하는 인공지능기본법을 가진 국가로 소개하고 있다. 특히 유엔인권최고대표는 인공지능 시대의 권리 보호를 위한 국가의 역할에 대하여 "사람에게 미치는 영향에 초점을 맞춘 인공지능 규제가 필요하다"고 강조하였다.

> *25.* 우선 규제는 보안이나 안전에 중점을 둔 일반적인 위험 모델보다는 사람에게 미치는 영향에 초점을 맞춰야 합니다. 일부 규제 체계는 민간 부문과 공공 부문의 *AI* 기술 사용을 이분법적으로 구분하지만, 양 부문의 상호 연관성을 고려해 보면 모두 인권 보호에 초점을 맞춰야 합니다. 또한, *AI* 규제는 공공 및 민간 부문 모두에서 투명성, 책무성, 그리고 인권에 미치는 영향의 완화를 보장해야 합니다. 투명성과 밀접하게 연관된 데이터 거버넌스, 특히 개인정보 보호 및 데이터 보호와 관련된 사항들이 중요한 고려사항입니다.

우리나라에서도 인공지능에 기반한 딥페이크 성착취물 확산, 생성형 허위조작정보의 유통, AI교과서 논란, 채용AI의 불공정성, 인공지능 도입으로 인한 대량 해고, 배달로봇 사고 등 인공지능 제품과 서비스가 시민의 안전과 인권에 미치는 위험이 이미 가시화되었으며, 앞으로 점점 더 증가할 것으로 예상된다. 최근 개인정보가 인공지능 학습용으로 많이 처리되기 시작하였는데, 통신사 등 시민들의 방대한 개인정보를 보유한 대형기업에서 대규모 개인정보 유출이 끊이지 않아 인공지능 시대를 맞이하는 시민들의 불안이 오히려 커져가는 상황이다.

이에 우리 시민단체는「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(이하, '인공지능기본법')」제정 과정에서 노동자, 장애인, 학생, 환자, 구직자, 소비자 등 인공지능의 영향을 받는 시민을 대신하여 인공지능기본법이 시민의 안전과 인권을 보호하는 최소한의 보호막(가드레일)이 되어야 한다고 요구하여 왔다. 그럼에도 지난 2025년 12월 국회를 통과하고 내년 2026년 1월 22일 시행을 앞두고 있는 인공지능기본법은 비윤리적, 비도덕적 인공지능의 금지를 규정하지 않았고, 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 큰 분야 내지 영역을 상당 부분 누락하고 있다. 국민의 생명, 안전, 자유를 본질적으로 침해할 수 있는 국방 또는 국가안보 목적의 인공지능을 규제대상에서 광범위하게 제외함은 물론이고 인공지능 위험에 상응하는 책임과 의무를 제대로 부과하지 않았다.

이러한 상황에서 인공지능기본법 시행령, 고시, 가이드라인 등 하위법령은 시민의 안전과 기본권을 최소한으로 보호할 수 있는 장치를 마련하여야 할 과제가 있다. 하위법령들이 인공지능 위험이 시민에 미치는 영향을 실질적으로 규율할 수 있어야 시민들이 안심하고 자신이 사용하거나 그 대상이 되는 인공지능을 신뢰할 수 있기 때문이다.

특히 인공지능기본법이 기본법으로서 제대로 기능하기 위해서는 꼭 필요한 위험 규제가 이 법과 하위법령에 규정되어야 한다. 이를 통해서 AI채용 관련 규제 등 추후 우리 사회가 마련할 수 있는

분야별 인공지능 위험 규제가 이 법과 조화롭게 작용할 수 있을 것이기 때문이다. 만약 이 법과 하위법령에 규제를 면제하거나 예외로 취급하는 사항이 많아진다면 이후 개별 영역에서 등장하게 될 인공지능 위험 규제 관련 특별법들과 충돌하거나 원활하게 작용하기 어려워질 우려가 있다.

그러나 지난 2025년 9월 17일 공개된 시행령초안은 그 제정 방향을 “규제보다는 진흥에 무게를 두고, 필요최소한의 규제를 합리적으로 마련하고 유연한 규제체계 도입”이라고 밝히고 있다(제정방향 p3). 이와 같은 방향은 인공지능 위험성에 대한 포괄적 규제체제를 도입한 유럽연합 등 글로벌 규범 동향에 반할 뿐 아니라 “국민의 권익과 존엄성을 보호”한다는 이 법의 목적(제1조)이나 “AI 산업 발전과 안전·신뢰 기반 구축”이라는 두 가치를 균형적으로 표방하겠다는 입법 취지를 심각하게 훼손한다. 시행령 제정방향을 정함에 있어 인공지능으로 인해 피해를 입을 수 있는 영향을 받는 자를 포함하여 시민사회로부터 의견수렴이 부실했다는 점에서 산업 이익에 편향된 결과는 예견된 것이나 다름없다.

특히, 이번 시행령초안은 다음과 같은 점에서 심각한 문제가 있어 이후 의견수렴 과정을 거치면서 반드시 수정, 보완되어야 한다.

- 첫째, 우리나라 인공지능기본법은 유럽연합 인공지능기본법과 달리 공공장소 얼굴인식, 취약성 공격, 직장과 학교의 감정인식 등 인권침해 소지가 큰 인공지능 시스템을 금지하고 있지 않다. 따라서 적어도 시행령의 고영향 인공지능에 관한 정의 및 책무 규정에서 안전과 인권에 위험한 인공지능과 그에 대한 조치를 충분히 규정하여야 마땅하다. 그럼에도 시행령초안에서 고영향 인공지능에 관한 규정은 법률에서 시행령으로 명시적으로 위임한 사항에 대해서조차 아무런 규정을 두지 않고 있다.

- 둘째, 하위법령안은 법령 규정 외적으로 ‘이용사업자’를 협소하게 해석함으로써 인공지능 제품이나 서비스를 업무에 사용하는 사업자를 모두 ‘이용자’로서 보고 일체의 책무를 배제하였다. 이로 인하여 병원, 채용회사, 금융기관 등 업무상 목적으로 인공지능을 이용하는 사업자가 환자, 채용 구직자, 대출 신청자 등 ‘영향받는 자’에 대해서 위험관리, 설명, 사람의 관리·감독 등의 책무를 지지 않아도 된다. “제공받은 제품 또는 서비스를 형태나 내용의 변경 없이 그대로 이용하는” 이용자와 달리, 인공지능을 “업무 목적으로” 이용하여 영향받는 자에 대하여 직접적인 영향을 미치는 병원, 채용회사, 금융기관 등 사업자는 ‘이용사업자’에 해당하는 합당한 책무를 져야 한다.

- 셋째, 인공지능기본법은 국방 또는 국가안보 목적으로만 개발·이용되는 인공지능을 국가정보원장, 국방부장관, 경찰청장의 자체적인 지정만으로 이 법의 적용범위에서 광범위하게 배제하였다. 현재 국방 또는 국가안보 목적 인공지능을 규율할 수 있는 법안이 추진조차 되고 있지 않기 때문에, 가장 인권침해가 심각한 인공지능에 대한 규제 공백이 우려된다. 그럼에도 시행령초안은 국가안보핵심기술 등 ‘국방 또는 국가안보 목적으로만 개발・이용되는 인공지능’ 분야를 폭넓게 인정하고 있다. 이중용도(dual use)로 사용될 수 있는 인공지능이 국방 또는 국가안보 목적 예외로 이 법에서 규정한 최소한의 의무를 배제한 채 은밀하게 개발·운영되는 일이 없도록, 최소한 적용제외 인공지능에 대해서는 국가인공지능위원회가 심의·결정하도록 하여야 한다.

- 넷째, 시행령초안은 안전성 확보 의무가 적용되는 최첨단 인공지능(frontier AI)의 기준으로 ‘학습에 사용된 누적 연산량이 10의26승 이상’으로 매우 좁게 설정하고 있다. 그러나, 현재 이 기준을 충족하는 인공지능 시스템이 과연 몇 개나 있을지 의문이다. 향후 기술 발전에 따라 기준을 변경하더라도 현재 주요 최첨단 인공지능을 포괄할 수 있도록 10의25승 이상으로 규정할 필요가 있다.

- 다섯째, 고영향 사업자 책무에 있어서 시행령초안은 “이행하여야” 하며(법 제34조 제1항), 고시는 “준수하도록 권고할 수 있다”고 정하고 있을 뿐이다(법 제34조 제2항). 시행령과

고시, 가이드라인은 그 법적 지위가 상이하다. 권고에 불과한 사항에 대해서는 특히 비용이 많이 소요되는 조치의 경우 사업자의 준수를 기대하기 어렵고 위반 시 제재도 어려울 수밖에 없다. 그럼에도 시행령초안에는 명시되어야 마땅한 중요한 사항에 대해서 아무런 언급을 하지 않고, 고시나 가이드라인에만 기재되어 있는 경우가 많다. 특히 법률이 위임한 중요 사항과 국민의 권리 의무에 직접적인 영향을 미치는 사항에 대해서는 반드시 시행령에 규정하여야 할 것이다. 따라서 고영향 인공지능 사업자의 책무에 대한 고시 및 가이드라인의 주요 내용은 시행령으로 규정하여야 한다.

- 여섯째, 시행령초안은 법률에서 위임하지 않은 사실조사의 면제를 규정하거나 상당 기간(미상)의 계도기간을 운영하도록 하였다. 인공지능 제품 및 서비스로 인한 안전 사고나 인권 침해가 발생하여도 국가가 최소한의 행정 조사를 포기하거나 사실상 과태료를 미부과하겠다는 방침[1]인 것이다. 사실조사와 과태료에 대해서는 기업들이 많은 민원을 제기한 바 있는 만큼, 그 면제 또는 유예는 시민 안전이나 인권 보호보다 기업 민원을 중시한 결과가 아닌지 의문스럽다. 이러한 전반적인 규제 설계는, 당분간 소비자를 비롯하여 그 영향을 받는 사람을 보호할 수 있는 설명 방안이나 사람의 관리·감독, 문서의 작성·보관 등 책무를 다하지 않고 인공지능 제품이나 서비스를 시장에 출시하여도 된다는 신호를 국가적 차원에서 공식화한 것이나 다르지 않다.

세계적으로 주목받고 있는 한국의 인공지능기본법이 인공지능 위험으로부터 영향받는 시민을 보호하고 국제인권규범에 부합하는 인권기반접근을 달성할 수 있도록, 시민사회는 시행령초안에 대해 다음과 같이 의견을 제시하며, 반드시 반영되기를 촉구한다.

[1] ‘AI기본법 하위법령 제정방향’ 과학기술정보통신부 (2025. 9.), 5쪽

# 인공지능기본법 시행령초안에 대한 검토 및 수정요구사항

## 1. 정의 (법 제2조, 시행령초안 미규정)

- 시행령초안의 내용

- 법 제2조(정의)에서 위임한 내용에 대해 시행령초안에는 아무런 규정이 없지만, 하위법령안 전체적으로 병원, 채용회사, 금융기관 등 인공지능 제품이나 서비스를 업무에 사용하는 사업자를 모두 '이용자'로 보고 일체의 책무를 배제함.

- 문제점

- 우리나라 인공지능산업 가치사슬의 수범자는 인공지능 개발사업자, 이용사업자, 이용자, 영향받는 자로 구성됨. 특히 유럽연합 인공지능기본법에서 '배치자(deployer)'에 해당하는 개념은 '이용사업자'와 '이용자'로 나뉘며, 고위험 인공지능의 경우 이용사업자에게는 책무가 부과되지만, 이용자에게는 책무가 부과되지 않는 차이가 있음.
- 법문상으로 이용사업자는 제공사업자가 "제공한 인공지능을 이용하여 인공지능제품 또는 인공지능서비스를 제공하는 자(법 제2조 제7호)"이며, 이용자는 "인공지능제품 또는 인공지능서비스를 제공받는 자(법 제2조 제8호)"임.

| 한국(인공지능기본법) | EU(AI ACT) |
|---|---|
| 인공지능이용사업자<br>: 인공지능개발사업자가 제공한 인공지능을 이용하여 인공지능제품·서비스를 제공하는 자<br><br>이용자<br>: 인공지능제품·서비스를 제공받는 자 | Deployer[2)]<br>"배포자"란 자신의 권한에 따라 AI 시스템을 이용하는 자연인, 법인, 정부·공공기관, 기타 기관·단체 등을 의미한다. 단, 직업적 활동이 아닌 개인적 활동 과정에서 AI 시스템을 이용하는 경우는 제외된다. |

- 그러나 하위법령안은 법령의 규정 외적으로 '이용사업자'를 협소하게 해석함으로써 인공지능 제품이나 서비스를 업무에 사용하는 사업자를 모두 '이용자'로 보고 일체의 책무를 배제함. 이로 인하여 병원, 채용회사, 금융기관 등 업무상 목적으로 인공지능을 이용하는 사업자가 환자, 채용 구직자, 대출 신청자 등 '영향받는 자'에 대해서 위험관리, 설명, 사람의 관리·감독 등의 책무를 지지 않아도 됨.
- 이러한 규정은 "영향받는 자는 인공지능의 최종결과 도출에 활용된 주요 기준 및 원리 등에 대하여 기술적·합리적으로 가능한 범위에서 명확하고 의미 있는 설명을 제공받을 수 있어야 한다."고 규정한 이 법의 원칙(제3조 제2항)을 형해화함.

※고영향 AI 판단 가이드라인(안)에서 책무가 면제되는 AI 회사들

**보건의료 산업**

| 고영향 AI 사업자 책무 이행 | | 고영향 AI 사업자 책무 이행 불필요 | |
|---|---|---|---|
| 의료 이미지 분석 AI 모델 개발 기업 | 의료영상 진단 AI 시스템 제공 기업 | 병원, 의사 등 의료인, 방사선사 및 임상병리사 등 의료기사 | 환자 |
| 의료영상 인식 및 분석 핵심 AI기술을 개발하여 제공 | 의료진의 진단 정확도 향상을 위한 AI기반 영상 분석 시스템 서비스 제공 | AI 진단 보조 시스템을 활용하여 환자에게 정확한 진단 서비스 제공 | AI 기반 정밀 진단을 통한 치료 |
| AI개발사업자 | AI이용사업자 | 이용자 | 영향받는 자 |

**채용분야**

| 고영향 AI 사업자 책무 이행 | | 고영향 AI 사업자 책무 이행 불필요 | |
|---|---|---|---|
| 영상·음성 분석 AI 모델 개발 기업 | AI 면접 시스템 제공 기업 | 채용기업 | 구직자 |
| 사람의 표정, 음성, 언어 패턴을 분석하는 핵심 AI기술을 개발하여 제공 | 채용 프로세스 효율화를 위한 AI 기반 면접 평가 및 인재 매칭 서비스 제공 | AI 면접 시스템을 활용하여 지원자 선별 및 평가 업무 수행 | AI 면접 시스템을 통한 채용 과정 참여 및 평가 대상 |
| AI개발사업자 | AI이용사업자 | 이용자 | 영향받는 자 |

**대출심사 분야**

| 고영향 AI 사업자 책무 이행 | | 고영향 AI 사업자 책무 이행 불필요 | |
|---|---|---|---|
| 신용평가 예측 AI 모델 개발 기업 | AI 대출심사 시스템 제공 기업 | 은행 및 금융기관 | 대출 신청자 |
| 대출자의 신용도 및 상환능력을 예측하는 핵심 AI기술을 개발하여 제공 | 금융기관의 대출 승인 프로세스 자동화를 위한 AI기반 신용평가 서비스 제공 | AI 대출심사 시스템을 활용하여 대출 신청자에 대한 신용평가 및 승인 업무 수행 | AI 기반 대출심사를 통한 대출 승인 여부 및 조건 결정 |
| AI개발사업자 | AI이용사업자 | 이용자 | 영향받는 자 |

- 시행령초안을 적용했을 때 문제가 되는 구체적 예시

| 구체적 상황 | 시행령초안을 적용했을 때 발생하는 문제 |
|---|---|
| A기업의 AI채용에 응시한 B씨가 좋지 않은 결과를 받았다. 하지만 B씨는 자신의 사투리가 제대로 인식되지 않았을 가능성에 대하여 의심하고 있으며, 이에 대하여 | '영향받는 자'로서 B씨는 "설명을 제공받을 수 있"는 권리가 있고(법 제3조), '이용사업자'는 위험관리, 설명방안, 사람의 관리・감독, 문서의 작성과 보관 등의 책무를 이행하면서(법 제34조), 영향받는 자에 대한 편향적 영향을 확인하고 조치할 수 있다. 그런데 만약 A기업이 '이용자'로만 규정된다면, 자기 업무에 인공지능시스템을 배치하는 모든 사업자는 |

| | |
|---|---|
| 사람이 점검하고 설명해줄 것을 요청하고 싶어한다. | 최종이용자 여부를 불문하고 '배치자(deployer)'로서 규정하고 서로 다른 수준이나마 일정한 책무를 부과하는 유럽연합, 브라질, 미국 콜로라도주 등 해외 인공지능기본법과 규범적으로 상호운용되지 못한다. 우리나라 인공지능기본법에서 제공받은 제품 또는 서비스를 형태나 내용의 변경 없이 그대로 최종이용하는 '이용자'를 별도로 규정하고 있다 하더라도, 인공지능을 "업무 목적으로" 이용하여 영향받는 자에 대하여 직접적인 영향을 미치는 병원, 채용회사, 금융기관 등 사업자는 '이용사업자'에 합당한 책무를 져야 한다. |

- 시행령초안 수정 요구사항

- 법문의 경우 '이용사업자'를 '이용자'에만 대응하여 제공하는 사업자로 한정하지 않았기 때문에 영향받는 자에 대해서도 제공하는 경우가 있을 수 있음. 이 법의 기본원칙 및 국가 등의 책무(제3조)가 영향받는 자의 권리를 규정하고 있고, 위험기반접근에 따라 고영향 인공지능으로부터 영향받는 자를 보호하고자 사업자의 책무 규정을 도입한 입법 취지가 있음. 따라서 "제공받은 제품 또는 서비스를 형태나 내용의 변경 없이 그대로 이용하는" 이용자와 달리, 최소한 인공지능을 "업무 목적으로" 이용하여 영향받는 자에 대하여 직접적인 영향을 미치는 병원, 채용회사, 금융기관 등 사업자는 '이용사업자'에 합당한 책무를 지는 것으로 시행령에 명확히 규정되어야 함.
- 가이드라인안에서 병원, 채용회사, 금융기관 등을 모두 '이용자'로서만 규정하고 일체의 책무를 배제한 도표(p169)는 수정되어야 함. 이용사업자-이용자-영향받는 자가 선형적 관계에 있지 않으며, 도입기업이 모두 '이용자'로만 한정된 것이 아니라 이용사업자로서 '영향받는 자'에 대하여 고영향 제품이나 서비스를 제공하는 경우가 있을 수 있음. 따라서 해당 경우에는 마땅한 책무를 부과하는 것으로 수정되어야 함.

## 2. 영향받는 자의 권리 (법 제3조, 시행령초안 미규정)

- 시행령초안의 내용

- 법 제3조 제2항은 영향받는 자는 인공지능의 최종결과 도출에 활용된 주요 기준 및 원리 등에 대하여 기술적·합리적으로 가능한 범위에서 명확하고 의미 있는 설명을 제공받을 수 있어야 한다고 규정하고 있음. 그러나 이러한 권리를 어떻게 보장할 것인지에 대한 절차를 규정한 조항이 없고, 이에 관련된 시행령 조항도 없음.

- 문제점

- 법 제3조 제2항에서 영향받는 자의 설명요구권을 규정하면서 이에 대한 구체적인 조항을 규정하지 않은 것은 입법 미비라고 할 수 있음. 그러나 법 제3조에서 권리는 명확하게 규정하고 있으므로, 시행령을 통해서라도 이를 이행할 수 있는 방안을 구체화할 필요가 있음.

- 시행령초안 수정 요구사항

- 시행령에서 영향받는 자가 설명요구권을 행사할 수 있는 대상과 세부적인 절차를 규정하고, 인공지능 사업자가 이를 보장하지 않았을 때 권리를 보장받을 수 있는 방안을 규정할 필요가 있음.

## 3. 적용범위 (법 제4조, 시행령초안 제2조)

- 시행령초안 제**2**조의 내용

- 국방 국가안보 목적으로만 이용되는 AI는 AI기본법 적용 대상에서 제외되며, 시행령초안에서 제외 대상 범위를 구체화함. 특히 국정원, 국방부, 경찰청이 수행하는 국방・국가안보 업무를 반영하고 있음.

- 문제점

- 법 제4조 제2항은 하위법령으로 구체화할 것이 아니라, 원칙적으로 법률개정을 통해 삭제되어야 함. 국방 또는 국가안보 목적 인공지능 시스템은 국민의 생명, 안전, 자유를 본질적으로 침해할 수 있어 하위입법으로 유보할 수 없는 사항임. 따라서 국방 또는 국가안보 목적 인공지능은 인공지능 기본법의 적용을 받아야 함.
- 국정원, 국방부, 경찰청이 수행하는 국방・국가안보 업무를 반영한다고 밝혔음. 국정원의 경우, 안보 관련 위협정보 수집, 분석, 인공지능 안보 위협 대응을 추진하고 있으며, 국방부는 AI기반 유무인복합전투체계 개발, 지능형 지휘통제체계 구축 등을 계획하고 있음. 특히 경찰청은 광범위하게 수집, 분석된 치안데이터 및 공공데이터를 활용하여 범죄예측시스템(Pre-CAS) 개발, 운용하고 있으며, 더욱 고도화될 것으로 예상됨.
- 그럼에도 불구하고 하위법령 제정방향은 국정원, 국방부, 경찰청이 수행하는 국방・국가안보 업무를 반영한다고 밝히고 있음. 이는 국민의 생명, 안전, 자유에 직결되는 사무를 담당하는 기관의 인공지능 개발 활용을 방치하는 결과를 낳게 될 것임.
- 또한, 인공지능이 가지고 있는 이중용도(dual use)의 특성상, 오로지 국방 또는 국가안보 목적으로만 개발・이용되는 인공지능과 통상의 목적으로 개발・이용되는 인공지능을 구분할 수 없음(가짜 미디어 탐지 활동이나 집단행동 예측 등). 따라서 “국방 또는 국가안보 목적으로만” 개발・이용되는 인공지능을 법적용에서 배제하는 것은 원천적으로 불가능함.
- 그럼에도 불구하고, 법은 “국방 또는 국가안보 목적으로만 개발・이용되는 인공지능으로서 대통령령으로 정하는 인공지능”에 대한 적용을 제외하고 있으며, 시행령초안 제2조는 각 호의 어느 하나에 해당하는 업무만을 수행하기 위하여 개발・이용되는 인공지능으로 구체화하고 있으며, 각 목에 해당하는 업무로서 국가정보원장, 국방부장관, 경찰청장이 지정하는 업무로 구체화하고 있음.
- 「국가정보원법」 제4조제1항제1호부터 제4호까지에 따른 직무의 경우,
    - 국가정보원장이 지정하는 업무로서 국정원법 제4조제1항제1호부터 제4호까지 업무 중, “국가안보와 국익에 반하는 북한, 외국 및 외국인・외국단체・초국가행위자 또는 이와 연계된 내국인의 활동을 확인・견제・차단하고, 국민의 안전을 보호하기 위하여 취하는 대응조치”(제3호), 중앙행정기관, 지방자치단체 및 공공기관 등에 대한 사이버공격 및 위협에 대한 예방 및 대응(제4호)에 해당하는 정보의 수집・작성・배포가 포함되어 있음.
    - 그러나, 여기에는 내국인의 활동에 대한 확인・견제・차단을 위한 대응조치가 포함되어 있는데, 비록 방첩, 대테러, 국제범죄조직에 관한 정보 및 국가기밀에 해당하는 직무수행이라 할지라도, 이에 활용되는 인공지능은 국방 또는 국가안보 목적으로만 개발・이용되는 인공지능이라 할 수 없으며, 광범위한 목적으로 내국인의 활동을 감시하고 사찰할 수 있는 인공지능으로 활용될 가능성 있음. 또한 국가안보라는 추상적 목적을 내세우며 인공지능을 활용하여 내국인 사찰을 가능하게 하는 규정임.

    - 또한 제4호에 따른 중앙행정기관 등에 대한 사이버공격 및 위협에 대한 예방 및 대응을 위한 정보수집 등을 위해서는 기관에 대한 일상적 정보수집, 분석이 필요하며, 이는 국가안보 등의 추상적 목적을 위해 각 기관의 정보에 대한 일상적 수집 및 분석을 전제하여, 국가안보 목적만을 위한 활동이라 할 수 없음.
- 「산업기술의 유출방지 및 보호에 관한 법률」 제2조제2호에 따른 국가핵심기술의 유출 및 침해를 방지하기 위하여 필요한 조사 및 조치의 경우,
    - 국가핵심기술을 "국내외 시장에서 차지하는 기술적·경제적 가치가 높거나 관련 산업의 성장잠재력이 높아 해외로 유출될 경우에 국가의 안전보장 및 국민경제의 발전에 중대한 악영향을 줄 우려가 있는 기술"로서 산업통상자원부장관이 지정하는 기술로 정의하고 있음.
    - 그러나, 국가핵심기술은 안전보장과 경제적 가치의 견지에서 지정되는 기술로서 국방 또는 국가안보만으로 지정되는 기술이라 보기 어려움에도, 그 유출 및 침해를 방지하기 위하여 필요한 조사 및 조치를 위해 개발 이용되는 인공지능을 법적용에서 제외하고 있음.
    - 또한 산업통상자원부장관이 대상기술을 해제하는 경우, 법적용 여부가 국정원장의 지정 해제에 따르는지 불명확함.
- 「국민보호와 공공안전을 위한 테러방지법」 제9조제1항에 따른 정보의 수집 및 같은 조 제4항에 따른 추적의 경우,
    - 테러방지법 제9조에 따르면, 국가정보원장은 테러위험인물에 대하여 출입국·금융거래 및 통신이용 등 관련 정보를 수집할 수 있으며, 대테러조사 및 테러위험인물을 추적할 수 있도록 하고 있음.
    - 테러위험인물에는 일반 국민 및 외국인을 포함할 수 있으므로, 그에 관한 정보수집 및 추적에 활용되는 인공지능을 국방 또는 국가안보 목적으로만 이용되는 인공지능이라 할 수 없으며, 목적 외 활용을 제한하는 규정도 없음.
- 「경제안보를 위한 공급망 안정화 지원 기본법」 제15조제1항에 따른 조기경보시스템 운영·관리의 경우,
    - 공급망 위험에 대한 선제적 파악 및 대응이 "국방 또는 국가안보"와 필연적 관계가 있다고 보기 힘들며, "물자 및 원재료 등의 국내외 수급 동향 및 가격, 생산량의 변화, 외국정부 또는 기업의 정책변경, 물류 또는 지급·결제의 장애 가능성 등을 점검하는 조기경보시스템"은 국가안보 목적 외에도 활용될 수 있는 예측시스템으로 국가안보 목적으로만 포섭되는 시스템이라 할 수 없음.
- 「전자정부법」제56조제3항에 따라 확인된 국가정보원장의 보안조치의 경우,
    - 행정기관의 장은 정보통신망을 이용하여 전자문서를 보관·유통할 때 위조·변조·훼손 또는 유출을 방지하기 위하여 국가정보원장이 안전성을 확인한 보안조치를 하여야 하고, 국가정보원장은 그 이행 여부를 확인할 수 있다고 규정하는데, 정보통신망을 이용한 전자문서 보관·유통 시 위조·변조·훼손 또는 유출 방지를 위한 국정원장의 안전성 확인 보안조치는, 설사 국방 또는 국가안보와 일부 관련성이 있다 하더라도 그 목적으로만 보안조치를 하도록 하는 것은 아님.
- 「국방정보화 기반조성 및 국방정보자원관리에 관한 법률」제2조제7호에 따른 국방정보화사업의 경우,
    - "국방정보화사업"을 국방정보통신망 및 국방정보시스템의 구축·운영 등에 관한 사업으로서 국방부령으로 정하는 것으로 정의하고 있어, 해당 사업이 국방 목적만의 사업인지 법문만으로 예측하기 어려움.
- 「경찰관의 정보수집 및 처리 등 에 관한 규정」제3조제3호·제4호 및 제9호에 따른 정보의 수집·작성 및 배포의 경우,
    - 제9호에 따른 신원조사는 "국방 또는 국가안보 목적"이라 보기 어려우며, 그에 관한 정보의 수집 등을 위한 인공지능도 국가안보 목적으로만 개발·이용되는 인공지능이라 보기 어려움.

- 「국민보호와 공공안전을 위한 테러방지법」 제2조제6호에 따른 대테러활동의 경우,
    - 대테러활동에 포함되는 테러위험인물의 관리, 인원・시설・장비의 보호, 국제행사의 안전확보 등은 일반 국민 또는 외국인, 일반 시설을 대상으로 하는 것으로, 국가안보 목적만으로 활용되지 않으며, 목적 외 활용에 대한 통제장치도 없음.
- 「국가첨단전략산업 경쟁력 강화 및 보호에 관한 특별조치법」제2조제1호에 따른 국가첨단전략기술의 유출 및 침해행위 등의 수사 및 방지 조치의 경우,
    - “국가첨단전략기술”은 공급망 안정화 등 국가・경제 안보에 미치는 영향 및 수출・고용 등 국민경제적 효과가 크고 연관산업에 미치는 파급효과가 현저한 기술로 정의됨. 경제안보는 경제적 파급효과를 의미하는데, 이는 국가안보 목적에 한정되지 않음.
- 결론적으로, 법은 오로지 국방 또는 국가안보 목적으로만 개발・이용되는 인공지능에 대하여 적용을 배제하고자 하였으나, 시행령초안의 다수 규정은 1) 국방 또는 국가안보 목적이라는 추상적 규정을 구체화하지 않고 단순히 국방 또는 국가안보 관련성이 있는 규정을 나열하는데 그치고 있고, 2) 국방 또는 국가안보 목적 외에 개발・이용되는 인공지능도 법적용에 배제하고 있으며, 3) 국정원장, 국방부장관, 경찰청장의 지정이라는 광범위한 재량을 허용하고 있고, 4) 다른 법률에 따라 적용 배제되는 업무를 예정하고 있어, 5) 법문과 시행령초안을 종합한다 하더라도, 국방 또는 국가안보 목적만으로 개발・이용되는 인공지능이 무엇인지 예측가능하지 않음.

● 시행령초안 수정 요구사항

- 원칙적으로 국방 또는 국가안보 목적 인공지능도 적용범위에 포함해야 함.
- 인공지능 기본법의 적용범위에 포함할 수 없다면, 국방 또는 국가안보 목적 인공지능을 규제하는 내용의 다른 법률 개정 또는 특별법을 조속히 제정 추진해야 함.
- 시행령초안에서 적용 예외 대상 범위에 국방 또는 국가안보 목적 외로도 활용될 수 있는 인공지능의 경우 포함되지 않도록 매우 엄격하게 규정하여야 함. 이중용도(dual use)로 사용될 수 있는 인공지능이 국방 또는 국가안보 목적 예외로 이 법에서 규정한 최소한의 의무를 배제한 채 은밀하게 개발·운영되는 일이 없도록, 최소한 적용제외 인공지능에 대해서는 국가인공지능위원회가 심의·결정하도록 해야 함
- 특히,「국민보호와 공공안전을 위한 테러방지법」제9조제1항에 따른 정보의 수집 및 같은 조 제4항에 따른 추적,「경제안보를 위한 공급망 안정화 지원 기본법」제15조제1항에 따른 조기경보시스템 운영・관리,「경찰관의 정보수집 및 처리 등 에 관한 규정」 제3조제3호・제4호 및 제9호에 따른 정보의 수집・작성 및 배포,「국민보호와 공공안전을 위한 테러방지법」제2조제6호에 따른 대테러활동 등은 악용 및 남용가능성이 매우 크므로 반드시 삭제하여야 함.
- 이중용도(dual use)로 사용될 수 있는 인공지능이 국방 또는 국가안보 목적 예외로 이 법에서 규정한 최소한의 의무를 배제한 채 은밀하게 개발·운영되는 일이 없도록, 최소한 적용제외 인공지능에 대해서는 국가인공지능위원회가 심의·결정하도록 해야 함.

## 4. 국가인공지능위원회 (법 제7조~제10조, 시행령초안 제4조~제8조)

- 시행령초안 제**4**조**~8**조의 내용

- 영향받는 자의 참여에 대한 규정이 없음.

- 문제점

- 국가인공지능위원회의 구성, 운영, 개선권고 등, 분과위원회 등을 규정함.
- 그러나 정보주체, 노동자, 소비자 등 국가와 기업의 AI 사업으로부터 '영향받는 자(법 제2조제9호)'의 참여에 대하여 아무런 규정을 하지 않았음.
- '영향받는 자'는 AI 생태계의 중요한 권리주체로서, 국제규범에서 그 참여와 구제 보장을 강조하여 왔음.

> "국가는 권리주체, 특히 가장 큰 영향을 받거나 부정적인 결과를 겪을 가능성이 높은 권리주체가 개발 과정에 효과적으로 참여하고 기여할 수 있는 기회를 창출하고 특정한 신기술의 채택을 촉진해야 한다. 국가는 참여 보장과 포용적 의견수렴을 통해서, 경제적 효율성, 환경적 지속 가능성, 포용성 및 형평성을 갖춘 균형적이고 통합적인 지속 가능 개발 목표에 있어 어떤 기술이 가장 적절하고 효과적인지 결정할 수 있다.(유엔 사무총장, A/HRC/43/29)"

- 이재명 정부가 표방하는 'AI 민주정부'는 AI의 영향을 받는 사람을 '권리주체'로 대우하는 접근으로부터 출발할 수 있음. 특히 우리 사회 인공지능의 건전한 발전과 신뢰 조성을 꾀하는 이 법의 목적(제1조)을 달성하기 위해서는, 인공지능 관련 국가 의사결정에 국가기관 및 기업 뿐 아니라 '영향받는 자'의 참여를 보장하고 의견을 수렴하여야 함.

- 시행령초안 수정 요구사항

- 위원회의 의사는 개인과 사회적 차원에서 AI의 영향을 받는 일반 국민이 알 수 있도록 그 내용이 원칙적으로 공개되어야 함. (다만 위원회 의결로서 필요하다고 인정하면 공개하지 않을 수 있도록 예외를 둘 수 있음.)
- 국가인공지능위원회(법 제7조), 분과위원회, 특별위원회 및 자문단(법 제10조)에 '영향받는 자'의 참여를 보장하고 의견을 수렴하여야 함.
    - 각각의 구성원에 영향받는 자 또는 이를 대표할 수 있는 사람으로서, 시민사회단체로부터 추천을 받는 사람을 균형적으로 포함하도록 규정할 필요가 있음.
    - 타 심의・의결 행정위원회의 경우에도 권리 이해관계자 또는 그 대표가 될 수 있는 단체의 참여를 보장해 왔음.

## 5. 표준화 (법 제14조, 시행령초안 미규정)

- 시행령초안의 내용

- 법 제14조 제5항은 '그 밖에 제1항 및 제3항에 따른 표준화 사업의 추진 및 지원 등과 관련하여 필요한 사항'을 대통령령으로 정하도록 하고 있음. 그러나 시행령초안에서 이와 관련된 규정을 두고 있지 않음.

- 문제점

- 인공지능 제품과 서비스가 국제적인 호환성을 갖기 위해서는 국제표준을 준수하는 것이 중요하고, 나아가 우리도 국제표준 제정에 주도적으로 참여할 필요가 있음. 또한 법적 규범이 실질적으로 기술에 내재화되기 위해서는 표준화가 매우 중요함. 이에 인증 등 여러 규범체계와 연계하여 사업자들이 표준을 준수하도록 유도할 필요가 있음. 그러나 시행령초안은 이에 대해 아무런 규정도 하고 있지 않음.

- 시행령 수정 요구사항

- 시행령에서 표준의 고시 절차 및 주체, 표준화 사업의 지원 방식, 표준 준수 권고 및 인증과의 연계 등 표준화 및 표준 준수의 촉진을 위한 구체적인 조치를 포함할 필요가 있음.

## 6. 인공지능 학습용데이터 관련 (법 제15조, 시행령초안 제12조~제14조)

- 시행령초안 제**12**조**~14**조의 내용

- 시행령초안 제12조는 과기정통부장관이 학습용데이터 생산 및 가공 기술 개발 사업 등 지원대상 사업의 유형을 규정하고, 지원대상 사업자 고려사항을 제시하고 있음. 지원대상사업을 선정하기 위한 평가의 세부 기준과 평가 절차 등에 필요한 사항은 과기정통부장관이 정하도록 위임하고 있음. 시행령초안 제13조에서는 학습용데이터 통합제공시스템의 구축 및 관리 운영에 대해 규정하고 있음.

- 문제점

- 정부가 지원하는 학습용데이터의 생산・수집・관리・유통 및 활용은 개인정보보호법, 저작권법 등 관련 법령의 준수가 전제되어야 하고 지원대상 선정시 이 점이 검토되어야 하는데, 이에 대한 언급이 없음.
- 정부의 지원을 받아 생산・수집된 학습용데이터는 사회적 이익에 기여할 수 있어야 하므로, 학습용데이터를 생산한 기업 뿐만 아니라 누구나 사용할 수 있도록 제공해야 함. (물론 학습용데이터 통합제공시스템 운영을 위해 학습데이터 이용에 대한 소정의 대가는 받을 수 있음) 그런데 현재 지원을 받은 학습용데이터가 공공의 자원으로 귀속되는지 여부가 명확하지 않음.
- 세부 기준과 평가 절차 등에 필요한 사항은 과학기술정보통신부장관이 정하도록만 하고 있고, 이를 어떻게 정하는지 규정하고 있지 않음.

- 시행령초안의 수정 요구사항

- 시행령초안 제12조 제2항에 정부 지원 학습용데이터 선정시 고려할 사항으로 ‘개인정보보호법, 저작권법 등 관련 법령의 준수여부’ 포함.
- 제13조 3항에서도 학습용데이터의 최신성, 정확성, 상호 연계성 뿐만 아니라, 관련 법령의 준수도 포함할 것.
- 제12조 제1항의 대상사업에 학습용데이터에 대한 보안 및 폐기와 관련된 사업도 구체적으로 포함해야 함.
- 시행령초안 제12조에서 정부 지원 학습용데이터의 경우, 해당 지원기관에 소유권이 이전되며 누구나 접근하여 사용할 수 있도록 공개된다는 것을 명시함.
- 세부 기준과 평가 절차 등에 필요한 사항은 고시나 가이드라인으로 공개하도록 함.

## 7. 인공지능 기술 도입・활용 지원 (법 제16조, 시행령초안 제15조)

- 시행령초안 제**15**조의 내용

- 시행령초안 제15조에서 기업 및 공공기관의 AI 기술 도입에 필요한 사항, 지원 방안 수립을 위해 관계중앙행정기관 및 지방자치단체와의 협의, 안내자료 제공 등을 규정하고 있음.

- 문제점

- 지원 사항 중 하나로 '2. 이용자 또는 영향받는 자 보호를 위하여 필요한 교육 및 기술 지원'이 포함된 것은 바람직하나 이 뿐만 아니라 AI 시스템의 효과나 영향을 분석하기 위한 도구, 고영향 AI인 경우에는 기본권 영향평가 수행을 위한 도구의 지원이 필요함.
    - 특히, 공공기관의 경우 불필요한 중복 투자를 방지하기 위해 다른 공공기관에서 유사한 서비스나 시스템을 보유하고 있는지에 대해 파악하여 관련 정보를 제공할 필요가 있음. 이를 위해 정부, 지방자치단체, 공공기관이 도입하고 있는 AI 시스템에 대한 등록 의무화가 필요함. 장기적으로는 이와 관련하여 법에서 규정하는 것이 바람직함.

● 시행령초안 수정 요구사항

- 제15조 1항의 각 호 중 하나로 'AI 시스템의 효과나 영향을 분석하기 위한 도구, 고영향 AI인 경우에는 기본권 영향평가 수행을 위한 도구의 지원' 추가.
- 제15조 1항의 각 호 중 하나로 '공공기관이 보유하고 있는 인공지능 시스템에 대한 정보 제공' 추가.

## 8. 인공지능 데이터센터 관련 (법 제25조, 시행령초안 미규정)

● 시행령초안의 내용

- '인공지능 데이터센터 관련 시책의 추진 등에 대해 규정하고 있는 제25조는 시행령 위임 규정은 없으며, 이에 관련 시행령초안에도 없음.

● 문제점

- 인공지능 데이터센터에서 인공지능 학습 및 운영을 위해 물, 전기 등의 에너지를 과도하게 소비하여 기후위기를 심화하고 있고, 때로는 해당 지역의 에너지 부족을 야기한다는 비판이 제기되고 있음. 그러나 본 기본법에서 데이터센터 관련 에너지 정책을 과학적으로 수립하기 위해 필요한 에너지 데이터 사용에 대한 규정을 두고 있지 않음. 이에 제25조에서 인공지능 데이터센터 관련 시책의 추진을 규정하고 있는 바, 이러한 시책 중의 하나로 에너지 관련 데이터 공개에 대한 규정을 시행령에 포함할 필요가 있음.

● 시행령초안 수정 요구사항

- 제25조에서 시행령에 위임하는 조항은 없지만, 제25조에 따른 시책을 추진할 때, 데이터센터 운영이 지역 공동체의 에너지 수급에 미치는 영향을 고려하도록 하고, 기후위기에 미치는 영향을 측정하기 위해 데이터센터의 에너지 사용에 대한 데이터를 공개하도록 하는 등의 규정을 시행령으로 신설할 것을 제안함.

## 9. 인공지능 투명성 확보 의무 (법 제31조, 시행령초안 제22조)

● 시행령초안의 내용

- 생성형 고영향 인공지능 이용자에 대한 사전 고지 및 결과물 표시(워터마크) 의무 부여에 대한 의무 이행 방법과 예외를 시행령초안에 규정함.

- △약관 UI 등을 활용한 사전고지 인정, △비가시적 워터마크 인정, △딥페이크 결과물에 대해선 이용자의 연령 신체적 조건을 고려하여 고지 표시, △사업자 내부 업무용이거나, 생성형 고영향 인공지능 기반이 명백한 경우 투명성 의무 면제 등

● 문제점

- 법 제31조는 인공지능기본법에 대한 국회 법안심사과정에서 딥페이크 성폭력물에 대한 사회적우려와 대책을 반영하는 취지에서 마련된 바 있음. 특히 법 제31조 제2항과 관련이 있는 생성형인공지능의 결과물로 인하여 “실제하지 않는” 허외조작정보, 소비자기만 마케팅, 보이스피싱으로 인한 이용자 오인이나 사기 피해가 발생하고 있음. 법 제31조제3항은 “실제”와 구분하기 어려운 딥페이크 생성물을 특별하게 규제하면서 해당 결과물이 예술적・창의적 표현물에 해당하거나 그 일부를 구성하는 경우에는 전시 또는 향유 등을 저해하지 아니하는 방식으로 고지 또는 표시할 수 있도록 면제하고 있음.
    - 그러나 시행령초안은 생성형 인공지능의 결과물에 대한 표시 방법에 있어 “사람 또는 기계가 판독할 수 있는 형식”을 규정하여 사람이 인지하지 못하는 표시를 허용함(제22조 제2항). 관련 <인공지능 투명성 확보 가이드라인>은 법 제31조 제2항이 제3항과 달리 “명확하게 인식할 수 있는 방식”이라는 “요건이 없음에 따라 비가시적 워터마크 활용 가능”이라고 주장함. 그러나 이는 투명성 의무가 생성형 인공지능 결과물의 대상이 되는 “사람”을 향한 의무라는 입법 취지를 시행령으로 형해화한 것이며, 법에서 위임하지 않은 과도한 예외에 해당함.
    - 더불어 시행령초안은 법 제31조 제1항부터 제3항까지 중 “전부 또는 일부를 적용하지 아니할 수 있다.”고 법 적용의 예외에 대하여 규정함(제22조 제4항). 그러나 하위법령에서 법 적용의 예외를 지나치게 포괄적으로 규정할 경우 위임 범위를 일탈할 뿐더러 명확하지 않은 규정은 투명성 의무가 적용되어야 마땅한 사업자가 탈법적으로 의무를 우회하는 등 악용할 가능성을 낳을 수 있음. ‘전부 또는 일부’라는 규정 또한 적용 범위를 예측하기 어렵게 만듬. 특히 “고영향 인공지능 또는 생성형 인공지능 기반 운용 사실이 명백한 경우”(제22조 제4항 제1호)는 매우 불명확하여 그 대상을 예측하기 어려움. <투명성 가이드라인>에서 제시하고 있는 예시조차 이 조항에 해당하는지 의문스럽고 굳이 예외를 적용해야 하는지를 납득하기 어려움.
    - 이러한 광범위한 예외는 특히 장애, 연령, 심리적인 취약성에 처해 있는 이용자를 투명하지 않은 인공지능 제품 또는 서비스의 위험에 방치할 수 있음. 허위조작정보, 소비자기만 마케팅, 보이스피싱 등에 악용될 수 있는 생성형 인공지능을 개발하거나 이용하는 사업자가 “기계가 판독할 수 있으니 사람이 판독할 수 있는 표시를 생략한다”거나, “결과물이 조악한 수준이기 때문에 누가 보아도 오인할 우려 없이 생성형 인공지능으로 생성되었다는 사실이 명백하므로 표시 의무의 예외에 해당한다”고 주장할 우려도 있음.
- 특히 이 조항에 대해서는 기업협회 등이 “적용범위를 제한”할 것을 요구해 온 만큼, 오인이나 오해로부터 시민을 보호하기 보다 기업 민원을 중시한 결과 과도한 예외 규정을 두게 된 것은 아닌지 의문스러움.

● 시행령초안을 적용했을 때 문제가 되는 구체적 예시

| 구체적 상황 | 시행령초안을 적용했을 때 발생하는 문제 |
|---|---|
| 장애가 있는 인터넷 이용자 C씨는 본인이 동영상 사이트에서 여러 차례 접했던 건강 관련 정보가 실제로는 존재하지 않는 허위조작정보이고 AI가 | 생성형 인공지능의 투명성 의무는 사람에 대한 투명성이어야 하며, 이는 특히 취약성을 가지고 있는 시민이 허위조작정보나 사기판매 등 후속적인 피해를 보지 않을 수 있도록 보장하는 첫 |

| | |
|---|---|
| 생성하였다는 사실을 최근에서야 타인에게 전해듣고 깜짝 놀랐다. 한 노부부는 인공지능으로 생성한 가짜 영상을 믿고, 실제로 존재하지 않는 관광지를 찾아 3시간을 운전해 갔다가 사실이 아니라는 걸 알고 충격을 받았다. | 단계이다. 그러나 시행령초안은 기계가 판독할 수 있는 형식으로 표시되면 사람이 판독할 수 있는 표시를 병행하지 않아도 되도록 허용하였으며, “명확하게 인식할 수 있는 방식”이라는 주관적인 요건으로 투명성 의무 전부 또는 일부를 면제하였다. |

- 시행령초안 수정 요구사항

- 시행령초안 제22조 제2항 기계가 판독할 수 있는 표시 형식이라 하더라도 사람이 판독할 수 있는 표시 형식을 반드시 병행하도록 하여야 함
- 시행령초안 제22조 제4항 제1호의 예외 규정을 삭제하여야 함

## 10. 인공지능 안전성 확보 의무 (법 제32조, 시행령초안 제23조)

- 시행령초안의 내용

- 시행령초안 제23조는 법 제32조 제1항 “학습에 사용된 누적 연산량이 대통령령이 정하는 기준 이상인 인공지능시스템”을 ‘학습에 사용된 누적 연산량이 10의 26승 부동소수점 연산 이상인 인공지능시스템’으로서 과학기술정보통신부장관이 인공지능기술 발전 수준, 위험도 등을 고려하여 고시하는 기준에 해당하는 인공지능시스템으로 정하고 있음.

- 문제점

- 시행령초안은 “대통령령으로 정하는 기준 이상인 인공지능시스템”을 학습에 사용된 누적 연산량이 10의 26승 부동소수점 연산 이상으로 규정하고 있는데, 기업들이 정확한 수치를 공개하지 않기 때문에 추정치일 수밖에 없지만, 현재 이 기준을 충족하는 최첨단 인공지능 시스템은 거의 없는 것으로 알려져 있음. 적용대상이 거의 없을 정도로 기준을 높게 설정하는 것은 사실상 규제를 하지 않겠다는 것이나 마찬가지임. 더구나 학습에 사용된 누적 연산량이 10의 26승 부동소수점 연산 이상이어야 한다는 요건에 더하여, 인공지능기술 발전 수준, 위험도 등을 고려하여 고시하는 기준에도 해당해야 하므로, 이 조항의 적용 대상은 더욱 제한적일 수 밖에 없음. 현재 전 세계적으로 최첨단 AI 로 인정되는 주요 서비스를 포괄할 수 있도록 10의 25승 정도로 설정할 필요가 있음.
- 법 제32조 제3항은 "과학기술정보통신부장관은 제1항 각 호에 따른 사항의 구체적인 이행 방식 및 제2항에 따른 결과 제출 등에 필요한 사항을 정하여 고시"하도록 하고 있음. 고시에 포함되어야 할 안전조치의 구체적인 이행방식 및 결과 제출에 필요한 사항을 시행령이 아닌 고시에 위임한 것은 문제임. 전문적·기술적 사항 또는 경미하고 자주 변경되는 부수적 사항에 대하여 고시에 규정할 수는 있겠지만, 안전성 확보를 위한 구체적인 이행방식의 주요 내용은 시행령에 규정했어야 함.

- 시행령초안 수정 요구사항

- 시행령초안 제23조 제1항에서 10의 26승을 10의 25승으로 변경할 것을 제안함.
- 법 개정 이전에라도 법 제32조 제1항의 구체적인 이행방식, 즉 ① 인공지능 수명주기 전반에 걸친 위험의 식별・평가 및 완화 ②인공지능 관련 안전사고를 모니터링하고 대응하는 위험관리체계 구축에 대한 정의, 목적, 내용 등이 무엇인지 시행령에서

대략적으로 제시할 필요가 있음. 이를 통해 인공지능 사업자 뿐 아니라 이용자, 영향받는 자 등 국민의 예측가능성을 높이고 법적 안정성을 확보하여야 함. 이를 위해 현재 <안전성 확보 고시(안)>의 주요 내용을 시행령으로 올려야 함.

## 11. 고영향 인공지능 관련

### 1) 고영향 인공지능 목록 (법 제2조, 시행령초안 미규정)

- 시행령초안의 내용

- 법에서 위임한 고영향 인공지능 범위와 관련하여 추가하거나, 구체화한 부분이 전혀 없음. 특히 제2조 제4호 카목에서는 “그 밖에 사람의 생명·신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역으로서 대통령령으로 정하는 영역”을 위임하고 있으나, 시행령에서는 이에 대하여 아무런 구체적 규정을 두지 않음.

- 문제점

- 기본권 침해 위험이 크거나, 중대한 영향을 미칠 우려가 큰 경우 누락없이 고영향 인공지능으로 규정하여 사업자 등에게 적정한 의무를 부과하고, 책임과 권리구제의 근거를 마련하여야 함.
- 법 제2조 제4호에서 고영향 인공지능 목록을 제시하고 있는데, 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 큰 분야 내지 영역이 상당 부분 누락되어 있음. 특히 법상의 목록은 ‘안전’에 미치는 고영향이 다수이며, ‘인권’에 미치는 영향에 대한 규정은 매우 일부만 열거되어 있으며, 특히 “채용, 대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가”라는 규정은 예시 규정에 불과함.
- 따라서 법 제2조 제4호 카항에서 법상 명시된 고영향 인공지능 목록 외에 “사람의 생명, 신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역”을 대통령령으로 구체화할 수 있도록 위임 규정을 마련하고 있으므로 적어도 시행령에서 누락된 영역을 추가하거나 구체화하여야 함. 그러나, 시행령초안에 추가되어야 마땅한 고영향 영역에 대해 규정이 이루어지지 않음.
- 특히 우리나라 인공지능기본법은 유럽연합 인공지능기본법과 달리 공공장소 얼굴인식, 취약성 공격, 직장과 학교의 감정인식 등 인권침해소지가 큰 인공지능 시스템을 금지하고 있지 않음. 2025년 9월 발표된 프랑스 AI안전연구소 등의 <AI 레드라인(Red lines)> 성명에서는 각국 정부에 대하여 △핵무기를 비롯한 대량살상 내지는 치명적인 피해로 이어질 수 있는 무기 관련 지시 및 제어, △인간 사칭, 대중 감시, 인프라 공격 등 사이버 공간에서의 악용, △폐기 원칙 위반 및 자동화된 자기복제를 금지할 것을 제안하기도 하였음. 따라서 적어도 시행령의 고영향 인공지능에 관한 목록 정의에서 안전과 인권에 위험한 인공지능을 충분히 규정하여야 마땅함. 그럼에도 시행령초안에는 법률에서 시행령으로 명시적으로 위임한 고위험 인공지능 목록 추가 사항에 대해서 아무런 규정을 하지 않았음.
- 이로 인하여 고영향 AI 판단 가이드라인(안)에서는 유럽연합의 금지 인공지능에 해당하기 때문에 법원의 허가가 필요한 공공장소 원격 생체인식 AI나 항공기 승객 감시AI가 고영향 분야에 해당하지도 않는 것으로 설명하고 있음.

<고영향 AI 판단 가이드라인(안) 중 고영향 불인정 사례>

비해당 사례

C 항공사는 D 인공지능시스템을 통해 운항 중인 객실 내에서 승객의 행동을 분석하여 비상 상황에 대비하여 신속한 대처방안을 승객에게 제시하고 있다.

비해당 사례

시민 A는 자녀 B와 한강 공원으로 나들이를 가서 주말을 보내고 있던 중, 자녀 B가 없어진 것을 뒤늦게 알게 되어 3일 넘게 B를 찾았으나 결국 행방을 알지 못해 경찰에 신고하였다. 관할 지구대에서 출동한 경찰관 C는 시민 A의 최근 동선을 탐문하여 인근 CCTV 정보를 확인한 뒤 CCTV 관제센터와 연계된 인공지능시스템을 활용하기 위해 자녀 B의 키, 성별, 안면 사진 등 정보를 전산에 입력하였다. 해당 인공지능시스템은 3일 전부터 현재까지 CCTV 영상 정보를 실시간으로 분석하여 자녀 B의 현재 위치를 추적하여 자녀 B의 행방을 찾을 수 있었다.

- 시행령초안을 적용했을 때 문제가 되는 구체적 예시

| 구체적 상황 | 시행령초안을 적용했을 때 발생하는 문제 |
|---|---|
| 중등학교 교사 D씨는 수업시간에 학생들을 대상으로 감정인식도구를 사용한다.<br>학부모 E씨는 학교 폭력을 예방한다는 이유로 일부 지방자치단체 학교가 공공장소에서 학생의 얼굴을 실시간으로 인식하는 CCTV를 도입한 것이 학생의 인권을 중대하게 침해한다고 생각한다.<br>콜센터 상담직원 F씨는 최근 직장에 도입된 인공지능도구가 자신의 업무를 모니터링하는 것에 대하여 회사가 아무런 공지나 동의를 구하지 않는 데 대하여 문제의식이 있다. 동료들과 이 불투명한 인공지능도구의 기능에 대하여 다양한 추측을 해보면서, 향후 해고 등 고용상 불이익한 결과로 이어지지 않을지 불안해하고 있다.<br>이 콜센터 인공지능은 고객의 감정을 실시간으로 분석하려는 목표를 가지고 있다.<br>G씨는 경찰이 지리정보를 기반으로 범죄예측프로그램을 운영한다는 소식을 들었다. G씨가 현재 거주하고 있는 지역은 범죄율이 높은 편인데 자신 또한 해당 지역에 거주하고 있다는 이유만으로 낙인찍히는 건 아닌지 우려가 된다. | 유럽연합은 인간과 사회에 용납할 수 없는 위험을 미치는 여러 인공지능을 금지하며, 안전 뿐 아니라 분야별 인권에 고위험을 미치는 여러 인공지능을 고위험으로 규제한다. 미국 콜로라도주 인공지능기본법은 차별로부터 보호되는 집단에게 불리한 '영향' 또는 '불법적 차별 대우'를 야기하는 알고리즘 차별을 금지하며, 교육 등록 및 교육 기회, 고용 또는 고용 기회, 금융·대출 서비스, 필수 정부서비스, 건강관리, 주택, 보험 및 법률서비스 영역에서 중대한 결정을 내리는 인공지능을 고위험으로 규제한다. 그러나 한국 인공지능기본법은 금지 인공지능을 규정하지 않았으며 고영향 인공지능 목록에는 인권에 위험한 영향을 미치는 인공지능을 매우 일부만 열거하고 있다. 따라서 시행령초안은 중대한 고영향 인공지능을 추가적으로 규정하여야 한다. 특히 감정인식이나 고용에 위험을 미치는 결정을 내리는 데 활용되는 인공지능은 유럽연합 뿐 아니라 미국 콜로라도주 인공지능기본법에서도 고위험으로 규정되어 있으나, 시행령초안은 이에 대하여 아무런 규정을 두지 않았기 때문에 개인과 사회에 매우 위험한 활용이 우려된다. |

- 시행령초안 수정 요구사항

- ▲(범죄 수사나 체포 업무 외) 생체인식정보를 분석·활용하는 경우, ▲(채용결정 외) 근로조건, 근로계약이나 종료에 영향을 미치는 결정, ▲(대출결정 외) 개인에 대한 신용 또는 위험평가, ▲(학생평가 외) 교육기관 입학 결정,▲직장이나 학교에서 사람에 대한 모니터링, ▲수사 및 기소, 재판, 형집행 등 기본권을 침해할 수 있는 국가기관의 권한행사에 이용하는 경우, ▲외국인의 출입국, 난민인정, 귀화 등 출입국 사무에 활용되는 경우, 나아가 ▲위험한 기계, 기구 설비 및 ▲어린이들이 사용하는 장난감에 사용하는 인공지능, ▲감정인식, ▲정보통신망의 운영, ▲선거 및 투표행위, 투표결과에 영향을 미치기 위해 사용되는 등 기본권 침해 위험이 크거나, 중대한 영향을 미치는 영역을 빠짐없이 고영향 인공지능으로 규정해야함.
- 아울러 법이 개정되어 금지된 인공지능이 규정되기 전까지는 수용불가한 위험도를 지닌 금지된 인공지능의 경우, 예를 들어 ▲공개된 장소에서 실시간 원격 생체인식에 활용하는 인공지능, ▲인간의 심리, 사고, 행동 등을 현저하게 왜곡하기 위하여 잠재의식에 영향을 미치는 인공지능, ▲직장과 학교에서 감정 분석 또는 ▲사회적 행동이나 인격적 특성에 기반한 사회적 점수를 이용하여 특정인의 신뢰도를 평가하거나 분류하는 인공지능, ▲인적 개입에 의한 조치없이 무기를 운용할 목적으로 개발되거나, ▲범죄를 저지를 가능성을 평가하거나 예측할 목적으로 자연인에 대한 위험평가를 실시하는 인공지능 등에 대하여 개발이나 활용자체를 금지하지 않더라도 최소한 고영향 인공지능으로 포섭하여 사업자에게 적정한 의무를 부과할 필요가 있음.

## 2) 고영향 인공지능 확인 (법 제33조, 시행령초안 제24조~제25조)

● 시행령초안의 내용

- 시행령초안 제24조에서 고영향 인공지능 확인절차를 구체화하고 있음. 제25조에서는 고영향 인공지능 전문위원회에 대하여 정하고 있고, 제24조 제5항에서는 재확인 요청을 받은 경우 위 전문위원회의 자문을 받아 고영향 해당여부를 재확인하도록 정하고 있음.
- 디지털의료기기 관련 법률 등 타법상 의무를 “영역별 소관부처와 협업하에 특수성을 고려한 고영향AI 판단기준(제정방향 p10)”으로 인정하고 이를 시행령초안 [별표 1]에 ‘이행조치 인정 기준’으로 열거함.

● 문제점

- 과학기술정보통신부는 과학기술정책 소관부처로서 인공지능시스템이 기본권에 미칠 영향이나 위험에 대하여 충분한 전문성이나 경험을 보유하고 있지 못함. 따라서 위험성 판단과 관련하여 각 분야의 소관부처에 의견을 조회하거나 협의하는 등의 내용을 시행령에서 구체화하였어야 하나 이러한 고려사항이 시행령초안에 전혀 반영되어 있지 않음.
- 한편, 고영향 인공지능 여부의 객관적 확인을 위하여 과학기술정보통신부 장관이 인공지능 사업자에게 관련한 자료의 제출을 요구할 수 있다는 규정이 포함될 필요가 있으나, 현재 공개된 시행령초안에는 관련한 언급이 없음.
- 시행령초안 제24조 제1항에서, 고영향 인공지능 해당 여부의 확인을 요청하려는 경우에는 별지 서식의 확인 요청서를 과학기술정보통신부장관에게 제출해야 한다고 정하고 있어 첨부서류로 일정한 자료의 제출을 요구한다는 내용이 포함될 가능성은 있으나, 아직 별지 서식은 공개되지 않고 있어 확인할 수 없음.
- [별표1]에 열거된 법률들은 대부분 안전에 대한 고영향에 속함. 반면 인권에 고영향을 미치지만 구체적인 규제입법에 이르지 못하고 다만 조직법적으로 소관하는 사안에 대해서는 타법소관 또는 기관협업이 규정되지 않음. 이로 인하여 채용AI, 공권력 침해,

차별 등 특별한 법률 규정을 가지고 있지 않은 인권 관련 고영향 여부를 모두 과학기술정보통신부가 판단하는 결과를 낳음.

● 시행령초안 수정 요구사항

- 고영향 인공지능 확인 절차와 관련하여 과학기술정보통신부 장관이 각 분야의 소관부처에 의견을 조회하거나 협의하도록 의무를 부과하는 내용을 시행령에 포함시켜야 함. 특히 기본권에 초래할 수 있는 위험에 대해서는 반드시 국가인권위원회의 의견을 청취하여야 함.
- 또한 과학기술정보통신부장관이 인공지능사업자에게 고영향 인공지능의 기준을 충족하는지에 대한 구체적인 자료의 제출을 요구할 수 있다는 내용을 시행령으로 포함시켜야 함.
- 고영향 인공지능에 해당하는지 여부의 확인을 요청하는 경우 고영향에 해당하지 않는다는 판단을 하더라도, 그 목록을 공개하여 심사의 공정성, 객관성에 대한 통제가 가능할 수 있도록 하여야 함.
- 시행령초안 제25조에서는 단순히 '④ 전문위원회는 고영향 인공지능 확인 자문 업무를 효율적으로 수행하기 위하여 5명의 위원으로 구성하는 회의를 개최할 수 있다.'고 하고 있는데, 인권 관련 전문가나 영향을 받는자의 이해관계를 대변할 수 있는 자를 포함하여야 한다는 내용이 규정될 필요가 있음.
- 고영향 인공지능의 목록은 등록제도를 운영할 필요도 있음.
- 특별한 법률 규정을 가지고 있지 않더라도 조직법적으로 인권 고영향을 미치는 인공지능에 대해서도 관련 기관과 협업하도록 규정하여야 함. 특히 인권 관련 고영향에 대해서는 반드시 국가인권위원회의 의견을 청취하여야 함.

### 3) 고영향 인공지능 책무 (법 제34조, 시행령초안 제26조)

● 시행령초안의 내용

- 법률에서는 "다음 각 호의 내용을 포함하는 조치를 대통령령으로 정하는 바에 따라 이행하여야 한다"고 시행령에 위임했으나, 시행령초안은 이에 대하여 구체적인 규정을 하지 않았고, 대부분의 사항을 고시 또는 가이드라인에 규정하고 있음. 그런데 고영향 사업자 책무에 있어서 시행령초안은 "이행하여야" 하며(법 제34조 제1항), 고시는 "준수하도록 권고할 수 있다"고 정하고 있을 뿐임(법 제34조 제2항).
- 다만 시행령초안 제26조는 ① 법 제34조 제1항의 조치 중 일부에 대한 홈페이지 게시의무, ② 조치의무를 이행한 것으로 간주되는 경우, ③ 인공지능이용사업자(이하 '이용사업자'라 함)에 대한 인공지능개발사업자(이하 '개발사업자'라 함)의 협력의무, ④ 5년간 보관 의무, ⑤ 법 제34조 제3항에 따라 조치의무를 이행한 것으로 간주되는 경우를 정함.

● 문제점

- 법 제34조는 시행령에 고영향 인공지능의 안전성·신뢰성을 확보하기 위해 이행해야 하는 6가지 조치의 구체적인 '이행 방법과 절차'를 위임하였음. 그런데 시행령초안 제1항의 각호는 법 제1항 각호에서 정한 조치의 일부에 대한 게시의무만 구체적으로 정하고 있어 구체적인 이행 방법 및 절차에 대한 규정이 미비하고 공백이 존재함. 책무 이행과 관련된 구체적인 사항은 준수 대상인 시행령초안에 명시되어 있지 않고, 권고 대상인 고시와 가이드라인에만 규정되어 있음. 따라서 시행령초안으로는 법상 책무 위반을 판단할 수 있는 근거가 부족함.
- 시행령초안은 기본적인 고영향 인공지능사업자의 책무를 이행한 것으로 간주하는 경우에 더 치중하고 있음.

- 시행령초안 제1항의 단서는「부정경쟁방지 및 영업비밀보호에 관한 법률」상 영업비밀에 해당하는 사항의 게시의무를 면제하고 있음. '영업비밀'의 개념이 추상적이고, 사업자가 자의적으로 판단할 수 있는 부분으로 실제 게시의무를 해태하는 것을 정당화하는 이론적 근거로 작용할 수 있음.
- 법에 나열한 6가지 조치 중 아래와 같이 4가지 조치에 대한 일부 내용에 대한 게시의무를 규정하고 있는데, 결국 게시 대상을 축소하는 것이 골자임.

  1) 위험관리방안의 수립·운영 ➜ 위험관리방안의 주요 내용

  2) 인공지능이 도출한 최종결과, 최종결과 도출에 활용된 주요 기준, 개발·활용에 사용된 학습데이터 개요 등에 대한 설명 방안 수립·운영 ➜ 기준 및 설명 방안의 주요 내용

  3) 이용자 보호방안의 수립·운영 ➜ 이용자 보호방안

  4) 고영향 인공지능에 대한 사람의 관리·감독 ➜ 고영향 인공지능을 관리·감독하는 사람의 성명 및 연락처
- 설상가상으로 제2항은 개발사업자로부터 법 제1항 제1호부터 제3호 사항 조치의 모두 또는 일부를 이행한 인공지능시스템을 제공받은 이용사업자가 '중대한 기능 변경을 초래하지 않은 경우' 이용사업자의 조치의무를 면제하고 있음. 인공지능기술을 이용한 제품·서비스를 제공할 때 개발사업자와 이용사업자의 역할과 개별 단계에서 요구되는 조치의 내용이 다름.
  따라서 원칙적으로 모든 인공지능사업자가 개별 단계별로 법 제34조의 내용을 준수하여야 함에도 시행령초안은 일방적으로 이용사업자의 조치의무를 면제하고 있고, 그 요건도 개발사업자가 법 제34조 제1항 제1호부터 제6호의 모든 조치를 이행한 경우가 아니라 제1호부터 제3호를 이행한 경우로 축소하였으며, 이마저도 '전부 또는 일부'라고 정하고 있음. 심지어 '중대한 기능 변경을 초래하지 않은 경우'라고 추상적인 개념을 사용하여 조치의무 면제 범위를 더 넓힘 ➜ 사실상 이용사업자의 조치의무를 전부 면제하는 내용임.
- 제3항은 이용사업자가 개발사업자에게 필요한 자료의 제공을 요청할 수 있고, 개발사업자가 이에 협력하도록 노력하여야 한다고만 정하고 있을 뿐 협조의무를 명문화 하지 않음. 가장 낮은 수준의 '노력'을 규정하고 있어 '개발사업자'와 '이용사업자' 간의 자료제공이 원활히 이루어질 수 있을지 의문임.
- 제4항에서 법 제1항 제5호에서 정한 문서의 명칭을 적시하지 않았고, 보관의무 5년은 짧음. 채무불이행에 따른 손해배상책임의 소멸시효가 10년이므로 5년으로 제한할 경우 분쟁이 발생한 경우 자료 확보가 어려울 수 있음.
- [별표1]에 따른 조치를 해당 법령에 따라 이행한 경우에는 제1항에 따른 조치를 이행한 것으로 간주하고 있는데, 이는 중복 규제를 방지하기 위한 취지이지만, 문제는 [별표1]에 규정된 법령이 인공지능 기본법의 취지를 반영하여 개정되지 않을 경우, 인공지능 기본법상의 규율이 무력화될 위험이 있음. 따라서 별표 1에 포함된 해당 법령이 인공지능 기본법상의 취지를 반영하여 개정될 필요가 있음.

● 시행령초안 수정 요구사항

- 법 제34조 제1항에서 위임한 정한 6가지 조치의 구체적인 '이행 방법과 절차'를 시행령에 규정 하여야 함. 특히 국민의 권리 의무에 직접적인 영향을 미치는 사항에 대해서는 반드시 시행령에 규정하여야 함. 이에 <사업자 책무 고시(안)>의 내용을 고시가 아니라 시행령에 반영해야 함.(고시안의 상향 규정) 특히,
  - 법 제34조 제1항 제1호의 위험관리방안의 수립・운영과 관련하여 위험관리방안에 반드시 포함되어야 하는 내용과 주기 등에 대한 것을 명문으로 정하여야 함.
  - 제2호의 '기술적으로 가능한 범위에서의 인공지능이 도출한 최종결과, 인공지능의 최종결과 도출에 활용된 주요 기준, 인공지능의 개발・활용에 사용된 학습용데이터의 개요 등에 대한 설명 방안의 수립・시행'은 개별 기술 관련 문서에 대한 보관의무를 명확히 정하고 이용자 및 영향을 받는 사람에게 전달할 수 있는

방안을 포함할 필요가 있음.
  - 제3호의 이용자 보호방안과 관련하여 인공지능의 개발 뿐만 아니라, 운영 과정에서 이용자를 보호할 수 있는 방안을 포함해야 하며, 특히 '개인의 선택, 판단과 결정에 영향을 미치는 방식으로 작동하는 경우 이를 명확히 고지'하는 고지의무를 명문화할 필요가 있음. 또한, 법에서 이용자 보호방안으로 규정한 한계가 있음에도 불구하고, 인공지능 시스템이 이용자가 아닌 당사자에게도 영향을 미칠 수 있는 바, 인공지능 시스템으로부터 중대한 영향을 받는 사람에 대한 보호방안 역시 포괄할 수 있어야 함.
  - 제4호는 사람의 관리・감독을 규정하고 있으며, 이는 특히 고영향AI가 배치될 현장에서 시민의 안전과 인권에 미치는 부정적 영향을 완화하기 위해 매우 중요한 부분임. 고영향 인공지능에 대한 사람의 관리・감독은 '자격'을 갖춘자에 의한 관리감독이 필요하므로 자격요건 보완이 필요함. 시스템 중단조치까지 포함하는 구체적인 조치를 규정하여야 함(고시안의 상향 규정).
  - 위와 같은 내용 뿐만 아니라 문제가 발생한 경우 관련 조치를 이행할 의무 등을 명문으로 정하여야 함.
- 시행령초안 제1항은 게시의무의 대상이 지나치게 협소하여 이를 확대하여야 하며, 제1항의 단서는 반드시 삭제해야 함.
- 제2항의 조치의무 면제 요건을 두는 것은 삭제하여야 함. 설령 이를 유지하더라도 이용사업자의 면제 요건을 '제34조 제1항 제1호부터 제3호까지 사항의 조치를 모두 또는 일부 이행한 인공지능 시스템을 제공받은 이용사업자라 인공지능 시스템의 중대한 기능변경을 초래하지 않은 경우'로 정하고 있는데, 이는 1) 제1호부터 제3호까지 사항의 조치 ➜ 제1호부터 제6호의 조치, 2) 모두 또는 일부 ➜ 모두, 3) 중대한 기능변경을 초래하지 않은 경우 ➜ 기능변경을 초래하지 않은 경우로 수정하여야 함.
- 제3항에서 이용사업자가 개발사업자에게 자료제공을 요청할 경우 성실의무를 명문화해야 함.
- 제4항의 문서 보관의무 5년은 10년으로 수정하여야 함. 나아가 문서는 안전사고나 인권침해 발생시 사실조사를 하거나 피해구제를 위하여 매우 중요한 사항이므로 해당 문서의 명칭을 시행령에서 명확히 기재하여 제도의 통일성을 기하여야 함.
- [별표1] 관련 타 법이 본 법의 취지를 반영하여 개정될 수 있도록 관련 부처와 협력해야 한다는 내용을 규정할 필요가 있음.

## 4) 고영향 인공지능 영향평가 (법 제35조, 시행령초안 제27조)

- 시행령초안 제**27**조의 내용

- AI제품·서비스 제공 시 사람의 기본권에 미치는 영향평가를 실시할 수 있으며, 구체적 내용·방법을 시행령에서 규정
- 사업자가 자율적으로 실시하되, 고영향AI에 대해선 영향평가 실시 노력의무 규정
- 영향받는 대상 식별, 영향대상 기본권 식별, 영향의 내용·범위, AI사용행태, 위험의 예방·손실의 복구, 개선 방안 등을 반영하여 평가를 수행하며, 제3자에게 위탁 수행 가능
- AI 영향평가 가이드라인 수록 사항
  - 영향평가 수행주체, 평가 대상 AI시스템 단위, 평가 주기
  - 개인정보 영향평가 등 타 영향평가와의 관계
  - AI활용 영역별 영향평가 수행을 위한 프로파일 작성 방법
  - EU 영향평가 안내 및 관련 지침(상호운용성 확보를 위함)

- 문제점

- 법상 영향평가에 대한 노력의무만을 부과하여 수행의 실효성을 담보하기 어렵고, 국가기관등이 고영향 인공지능을 이용하는 경우 영향평가 실시 제품 서비스의 우선적 고려는 권고적 효력만을 가짐.
- 비록 법상 기본권 영향평가의 노력의무와 시기, 실효성 확보방안(우선적 고려)을 제한적으로 규정하였다 하더라도, 시행령을 통해 영향평가의 실효성을 제고하고, 기본권 영향평가의 구체적인 내용과 방법 등을 규정해야 함.
- 그럼에도 불구하고, 시행령초안은 영향평가에 포함될 사항 7가지와 평가주체(직접 또는 제3자)만을 규정하는데 그치고 있으며, 그 외 사항은 과기정통부장관이 수립, 보급할 수 있다고 정하고 있음
- 시행령초안은 영향평가에 포함되어야 하는 사항으로, “해당 고영향 인공지능으로 인하여 발생할 수 있는 사람의 기본권에 대한 사회적·경제적 영향의 내용 및 범위”(제3호)를 규정하고 있음. 그러나 인공지능 시스템은 사람의 생명, 신체, 기본권에 전인적으로 영향을 미칠 수 있다는 점에서, 사회적·경제적 영향에 국한되지 않음. 특히 인공지능이 민주주의, 법치주의 등 헌법적 가치에 미치는 부정적 영향이 광범위하게 보고되고 있음. 그럼에도 불구하고, 시행령초안은 사회적·경제적 영향에 국한하고 있음.
- 영향평가는 그 형식화와 부실화를 방지하기 위한 “공개 및 점검” 사항을 포함하여야 하나, 시행령초안은 영향평가의 공개 등에 관한 사항을 규정하지 않음.
- 평가 주체와 관련하여 시행령초안 제27조 제2항은 인공지능사업자가 "직접 또는 제3자에 의뢰하여" 영향평가를 실시하도록 하고 있는데, 평가 주체를 유연하게 할 필요는 있지만 영향평가가 인공지능사업자로부터 독립적으로 이루어질 수 있도록 보장할 필요가 있음.
- 기본권 영향평가의 활성화를 위해 정부는 기본권 영향평가를 수행할 수 있는 기관의 자격요건을 규정하고, 이에 부합하는 기관을 영향평가 전문기관으로 지정할 필요가 있음. 기본권 영향평가 수행기관은 인공지능에 대한 전문성 뿐만 아니라, 해당 인공지능이 사용되는 분야의 전문성, 인권에 대한 전문성을 갖춰야 함. 따라서 과기정통부 뿐만 아니라 인공지능 활용 분야에 따라 해당 정부기관에서 지정할 수 있을 것임. 이러한 기본권 영향평가 전문기관 제도를 통해 기본권 영향평가에 전문성을 가진 사람들을 양성하고, 영향평가에 대한 경험을 축적해나갈 수 있을 것임.
- 법 제35조는 기본권 영향평가를 '사전에' 하도록 하고 있지만, 원래 인권영향평가는 제품의 출시 전 뿐만 아니라 정기적으로 또는 상당한 기능의 변경이 있는 경우 다시 수행하도록 하고 있음. 인공지능 기본권 영향평가도 이를 권고할 필요가 있음.
- 2011년 유엔의 <기업과 인권에 관한 이행지침>은 기업의 인권실사 의무를 요구하고 있는데, 인권영향평가는 인권실사의 핵심도구임. 인공지능에 대한 인권영향평가는 인공지능 기술 분야에서 유엔 규범을 이행하는 것으로 볼 수 있음. 다른 영향평가와 달리 인권영향평가의 경우 이해관계자와의 소통과 협의가 특히 강조됨. 시행령초안은 제27조 제1항 제1호에서 "해당 고영향 인공지능을 이용한 제품 또는 서비스에 의하여 사람의 생명, 신체의 안전 및 기본권에 영향받을 수 있는 가능성이 있는 대상의 식별"을 포함하도록 하고 있지만, 이에 그치고 있음. 나아가 영향평가 과정에 영향받는 자와 협의하도록 할 필요가 있음.
- 기본권 영향평가가 실효성을 갖고 수행되기 위해서는 정부의 지원이 필요함. 인공지능 기본법은 인공지능 기술 개발이나 산업 육성에는 여러 지원을 제공하면서도 영향평가와 같이 위험을 최소화할 수 있는 제도나 관행에 대한 지원은 미약함. 특히 중소기업도 영향평가를 수행할 수 있기 위해서는 이들에 대한 컨설팅, 교육, 재정적 지원이 필요함.
- 시행령초안 제27조 제3항은 과학기술정보통신부장관이 "영향평가의 구체적인 내용, 방법을 수립하여 보급할 수 있"도록 하고 있음. 그런데 국내에서 인권 및 인권영향평가에 전문성을 가지고 있는 기관은 국가인권위원회이므로, 인공지능 기본권 영향평가와 관련해서도 국가인권위원회와 협의하는 것이 바람직함.
- 법 제35조 제2항은 "국가기관등이 고영향 인공지능을 이용한 제품 또는 서비스를 이용하려는 경우에는 영향평가를 실시한 제품 또는 서비스를 우선적으로 고려하도록 하고

있는데, 이를 위해서는 국가계약법(제7조제1항 등) 및 그 시행령(제21조 등), 지방계약법, 조달사업법(제14조, 제18조 등) 등에 따라 참가자의 자격을 제한하거나 달리 정하는 계약방법으로 계약을 체결하게 하거나, 납품검사, 조달물품 지원 및 지정(혁신제품 지원 포함) 등을 하도록 할 필요가 있음.

- 시행령초안 수정 요구사항

- “해당 고영향 인공지능으로 인하여 발생할 수 있는 사람의 기본권에 대한 사회적·경제적 영향의 내용 및 범위”에서 ‘사회적·경제적 영향’을 ‘전반적 영향’으로 개정해야 함(제27조 제1항 제3호).
- “해당 고영향 인공지능으로 인한 위험의 예방 및 손실 복구 등에 관한 사항”을 ‘해당 고영향 인공지능으로 인한 위험의 예방, 완화, 구제 등에 관한 사항’으로 개정(제27조 제1항 제6호)
- 제27조 제2항에 '인공지능사업자는 영향평가가 독립적으로 이루어질 수 있도록 보장한다'는 내용을 추가해야 함.
- 시행령에 기본권 영향평가 전문기관의 자격 요건, 전문기관의 지정 및 지정취소 등의 요건과 절차에 대한 규정이 포함되어야 함.
- 영향평가를 사전에 뿐만 아니라 정기적으로, 또는 상당한 기능의 변경이 있는 경우 수행하도록 권고해야 함.
- 해당 인공지능 시스템으로부터 영향받는 자(또는 제27조 제1항 제1호에서 식별된 대상) 또는 이들을 대리할 수 있는 사람과 협의하도록 권고해야 함.
- 형식적이거나 부실한 영향평가를 방지하기 위해서 영향평가의 내용을 공개할 필요가 있음. 특히, 공공기관이나 공공분야에 조달되는 인공지능 시스템에 대한 영향평가 결과보고서는 필수적으로 공개하도록 해야 함.
- 정부는 기본권 영향평가를 수행하려는 사업자(특히 중소기업)에게 컨설팅, 교육, 재정적 지원을 할 수 있음을 규정함.
- 과학기술정보통신부는 시행령초안 제27조 제3항의 영향평가의 구체적인 내용, 방법을 수립하여 보급할 때 국가인권위원회와 협의하도록 함.
- AI 영향평가 가이드라인은 다른 사회적 영향평가와 구별하고 EU 기본권 영향평가와의 상호호환성을 고려하여, <AI 기본권 영향평가 가이드라인>으로 이름을 변경할 것을 제안함.

## 12. 사실조사 등 (법 제40조, 시행령초안 제31조)

- 시행령초안의 내용

- 법 제40조(사실조사)에서 조사를 하도록 하였음에도 조사를 실시하지 않을 수 있는 면제 사유를 규정하였으며, 특히 “부당한 목적의 신고·민원”에 대해서 사실조사를 하지 않을 수 있도록 근거를 마련함.

- 문제점

- 인공지능 제품 및 서비스로 인한 안전 사고나 인권 침해가 발생하더라도 법 제40조의 사실조사에 이르기까지도 여러 단계를 거쳐야 함. 피해 신고가 있거나 부처가 인지하여야 하고, 그 이후에도 사실조사는 강행 규정이 아니라 “필요한 경우... 조사를 하게 할 수 있다”는 훈시 규정에 따라 이루어질 뿐임.
- 그럼에도 시행령초안에서는 더 나아가 법률에서 위임하지 않은 사실조사 면제사유를 규정함. “위반 사항 또는 혐의에 대하여 충분한 자료나 증거가 이미 확보되어 있는 경우”는 매우 모호하며 “신고나 민원이 민원인의 사적이익을 위하여 제기된 경우나 공무를 방해하기 위하여 제기된 경우 등 목적이 정당하지 않은 경우”는 매우 주관적임. 이처럼 모호하거나 주관적인 사유로 사실조사 면제사유를 규정하는 것은 사실조사 착수조차 어렵게 만들 수 있음.
- 또한 본법에서 훈시규정으로 두고 있음에도 시행령에 단서조항으로서 예외 규정을 별도로 두는 것이 법체계상 어색한 측면이 있음. 시행령에서 규정하는 예외 사유에 해당하는 경우 본법에 따라 장관의 재량에 의하여 사실조사를 시행하지 않으면 해결되는 문제라고 보여짐.
- 사실조사에 대해서는 기업들이 집중적인 민원을 제기한 바 있는 만큼, 최소한의 행정 조사 권한조차 시민의 안전이나 인권 보호 보다 기업의 민원을 중시한 결과로 보임.

- 시행령초안의 수정 요구사항

- 사실조사 면제 규정을 삭제하여야 함.

## 13. 과태료 (법 제43조, 시행령초안 제32조)

- 시행령초안의 내용

- 인공지능기본법 시행 초기 기업들의 혼란을 최소화할 수 있도록 실질적으로 규제 유예와 동일한 효과 달성 가능한 과태료 계도기간 운영 추진
- 이해관계자 수렴을 통해 구체적인 계도기간 등을 확정 예정

- 문제점

- 인공지능사업자에 대한 최소한의 규제(투명성 사전고지의무/국내대리인 미지정/중지명령·시정명령 미이행)에 대한 과태료 계도기간 운영은 실질적으로 인공지능기본법 상의 규제를 유예하는 신호를 줄 수 있음.
- 제43조 제1항 제3호는 사실조사에 따른 중지명령이나 시정명령을 이행하지 아니한 자에 대해 과태료를 부과하고 있으므로, 이미 사업자에게 개선 및 시정의 기회를

부여하였음에도 이조차 위반한 경우에 과태료를 부과하는 것이 도저히 과도하거나 불확실성을 야기할 조항으로 보이지 않음. <계도기간 적용 프로세스>를 보면 위반행위 중지 시정명령을 받았음에도 미이행 시 오히려 중지 시정명령보다 약한 행정지도만을 받도록 하는 것이 논리적 체계인지 알 수 없음.

- 제3호를 제외하면 사실상 법 위반시 곧바로 과태료가 부과되는 경우는 1) 투명성 사전고지의무 및 2) 국내 대리인 미지정 의무를 위반하였을 때인데, 이에 대하여는 사업자에 대한 최소한의 규제로 보일 뿐만 아니라 불확실성으로 인한 혼란 가능성도 높지 않은 의무규정이므로, 현행 과태료 규정이 계도기간을 두어야 할 정도로 사업자에게 불리하게 적용된다고 보이지 않음.
- 법에서 과태료 부과규정에 대하여 별도의 기한 유예를 두지 않았음에도 시행령을 통하여 과태료 계도기간을 운영하고자 함은 법률의 위임 범위를 넘어선 것이라고 보여짐 .
- 본래 법률의 경과규정은 법률 준수를 예비하기 위한 것이지, 미준수를 예비하는 것은 국민의 안전과 인권을 위험에 방치하는 것이나 다름이 없음.

● 시행령초안의 수정 요구사항

- 과태료 계도기간 운영에 반대함.

# 나가며 : 16대 수정 요구사항 요약

국제사회는 인공지능의 위험에 대해 규제마련에 나서고 있다. 2025년 9월 30일 미국 캘리포니아주는 대형 AI기업에 투명성과 안전 프로토콜을 의무화하는 AI 안전법 'SB 53'을 통과시켰고, 학습데이터 공개 등 인공지능의 투명성과 책무성을 의무적으로 요구하는 입법을 마쳤다. 유럽연합은 이에 앞선 올 2월 2일부터 인공지능기본법(EU AI Act)의 금지인공지능규제조항을 시행하였고, 8월 2일부터는 범용AI등에 대한 규제조항도 시행하기 시작했다. 또한 노벨상 수상자 등 국제적 석학들이 인류에게 용납할 수 없는 위험을 초래하는 시스템의 개발, 배포 및 사용을 명확하게 금지하기 위한 '사용한계선(레드라인 RED LINE)'설정에 대해 2026년 말까지 국제적 합의를 도출하고, 강력한 집행 메커니즘을 통해 이러한 한계선이 실제로 작동하도록 보장할 것을 촉구하는 서명을 벌이고 있기도 하다.

그러나 시행령초안은 이와 같은 국제적 흐름을 간과하지 않고 시행령에서나마 국민의 생명, 안전 등을 최소한이나마 보호할 수 있는 세부적인 내용을 규정하는 것이 마땅하다. 그러나 앞서 지적한 바와 같이 지난 9월 17일 공개된 시행령초안은 이에 부합한다고 보기 어렵다.

이에 시민사회는, 과기부가 공식 입법예고를 위한 시행령안 마련에 앞서 의견수렴과정에서 영향을 받는 자 등 시민사회의 의견을 충분히 수렴하여 아래와 같이 시행령초안을 수정, 반영할 것을 제안한다.

1) 법 제2조의 정의규정에서 혼란을 주고 있는 이용자 개념을 시행령에서 명확히 해야 함.

2) 법 제3조의 영향받는 자의 권리를 시행령에서 구체적으로 명시해야 함.

3) 국방 또는 국가안보 목적 외로도 활용될 수 있는 인공지능의 경우 포함되지 않도록 매우 엄격하게 규정해야 함.

   특히, ▲「국민보호와 공공안전을 위한 테러방지법」제9조제1항에 따른 정보의 수집 및 같은 조 제4항에 따른 추적, ▲「경제안보를 위한 공급망 안정화 지원 기본법」 제15조제1항에 따른 조기경보시스템 운영・관리, ▲「경찰관의 정보수집 및 처리 등 에 관한 규정」 제3조제3호・제4호 및 제9호에 따른 정보의 수집・작성 및 배포, ▲「 국민보호와 공공안전을 위한 테러방지법」제2조제6호에 따른 대테러활동 등은 악용 및 남용가능성이 매우 크므로 반드시 삭제해야 함. 최소한 적용제외 인공지능에 대해서는 국가인공지능위원회가 심의·결정해야 함.

4) 국가인공지능위원회의 의사는 원칙적 공개, 영향받는 자의 참여 보장 및 의견 수렴 절차 마련을 시행령에 규정해야 함.

5) 법 제14조 제5항에서 위임한 표준화 및 표준 준수의 촉진을 관련해 표준의 고시 절차 및 주체, 표준화 사업의 지원 방식, 표준 준수 권고 및 인증과의 연계 등 표준화 및 표준 준수의 촉진을 위한 구체적인 조치를 규정해야 함.

6) 정부 지원 인공지능 학습용 데이터는 선정에 개인정보보호법, 저작권법 등 준수 여부 반드시 포함시키고, 지원 대상 사업에 보안, 폐기 규정도 구체적으로 포함해야 함.

7) 기업 및 공공기관 AI기술 도입에 필요한 사항, 지원방안으로 AI 시스템의 효과나 영향 분석도구, 고영향 AI인 경우 기본권 영향평가 수행 도구 지원을 추가하고, 구체적인 영향 평가의 내용, 방법을 수립하여 보급할 때 국가인권위원회와 협의하도록 함.

8) 데이터센터 관련 시책의 추진에 지역공동체의 에너지 수급에 미치는 영향고려, 기휘위기에 미치는 영향을 측정할 수 있도록 에너지 사용 데이터 공개 규정 신설해야 함.

9) 고영향인공지능, 생성형AI의 투명성 의무에서 사람이 판독할 수 있는 표시 형식 병행 및 광범위한 예외규정을 삭제해야 함.

10) 안전성 확보 의무 대상 인공지능을 학습에 사용한 누적연산량 10의 25승 부동소수점 연산 이상으로 수정하고, 인공지능 안전성 확보 의무의 실효성을 담보하고 사업자 뿐 아니라 이용자, 영향을 받는 자 등 국민의 예측가능성을 높이고 법적 안정성을 확보하기 위해 고시에 위임하고 있는 안전성확보 고시는 시행령에서 규정해야 함.

11) 고영향인공지능 목록에 생체인식정보 분석, 활용하거나 근로조건, 계약이나 종료에 영향을 미치는 경우, 수사 및 기소, 선거 및 투표 등 기본권 침해 위험이 크거나, 중대한 영향을 미치는 영역을 빠짐없이 정의해야 함.

12) 공개된 장소의 실시간 원격생체인식 활용 인공지능, 인간의 심리, 사고, 행동을 현저히 왜곡하기 위해 잠재의식에 영향을 미치는 인공지능, 사회적 점수를 매겨 신뢰도 평가 및 분류하는 인공지능, 대량살상무기 개발 등 금지인공지능은 법개정 전 보호공백을 최소화하기 위해 고영향 인공지능으로 포섭해서 적정한 의무를 부과해야 함.

13) 고영향인공지능을 과기부 장관이 확인할 때 각분야의 소관부처에 의견 조회 및 협의하는 의무 부과하고, 기본권 침해 위험이 있는 경우 반드시 국가인권위원회 의견 청취하도록 해야 함.

14) 고영향인공지능 책무에서 안전성·신뢰성 확보하기 위한 6가지 이행 조치는 고시가 아니라 시행령에 구체적인 방법 및 절차에 대한 규정을 담아야 함. 게시의무 대상을 확대해야 하고, ‘부정경쟁 방지 및 영업비밀보호에 관한 법률’상 영업비밀에 해당하는 사항 면제 규정 단서 삭제 및 조치의무 면제 요건을 삭제해야 함.

15) 영향평가의 목적을 인공지능으로 인한 위험의 예방, 완화, 구제 등에 관한 사항으로 개정하고 독립적 수행을 보장, 사전 뿐 아니라 사후 등 정기성 담보 및 상당한 기능 변경이 있는 경우도 수행하도록 하며, 공공기관의 경우 영향평가 내용 공개 필수로 함. 과기부가 영향평가 구체적 방법 등을 수립하여 보급 시 국가인권위원회와 협의해야 함.

16) 사실조사 면제규정 삭제 및 최소한의 규제라고 할 과태료 조항의 계도기간 운영에 반대함.

과기부는 시행령초안의 공식적인 입법예고 전에도 영향을 받는 자 등 광범위한 이해 관계자의 의견수렴을 거쳐 바람직한 방향으로 수정할 것을 약속한 만큼, 이와 같은 시민사회의 의견을 제대로 수렴하여 반영할 것을 촉구한다. 끝.

인공지능기본법 시행령초안에 대한 시민사회 의견서

발행일 2025. 10. 01.
발행처 민주사회를 위한 변호사모임 디지털정보위원회
·정보인권연구소 ·진보네트워크센터· 참여연대
담당
참여연대 공익법센터 이지은 선임간사 02-723-0666
정보인권연구소 장여경 상임이사 02-701-7687
진보네트워크센터 오병일 대표 02-774-4551
민주사회를 위한 변호사모임 디지털정보위원회
02-522-7283

※ 본 자료는 각 단체 홈페이지에서 다시 볼 수 있습니다.