이슈리포트 <정보인권> 2024-12 (통권 제17호)

주요 분야 인공지능 정책 및 이슈 연구
: 공공, 법집행, 교육, 사회복지 분야

정보인권연구소
진보네트워크센터
APC

※ 이 보고서는 진보네트워크센터와 정보인권연구소의 공동연구로 작성되었으며, 진보통신연합(Association for Progressive Communications)의 지원으로 제작되었습니다.

이슈리포트 〈정보인권〉 2024-12 (통권 제17호)

주요 분야 인공지능 정책 및 이슈 연구
: 공공, 법집행, 교육, 사회복지 분야

발행인 : 사단법인 정보인권연구소 (이사장 이영음)
발행일 : 2024년 12월 31일

정보인권연구소
진보네트워크센터

APC

목 차

공공 분야 인공지능 이슈 현황

1. 도입

중앙행정기관, 지방자치단체, 공공기관 등 공공분야에서도 챗봇을 통한 국민대상 서비스, 수사기관의 디지털증거분석시스템 등 업무지원 서비스, 환경 모니터링 AI 등 다양한 AI 시스템이 이미 도입되어 활용되고 있다. 또한, 더 고도화된 AI 시스템이 빠르게 도입되고 있는 중이다. 그러나 도입된 AI 시스템을 통합 관리하지 않기 때문에, 공공분야에서 어떤 목적으로 어떠한 기능을 가진 AI 시스템이 도입되어 있는지 정확하게 파악하기 쉽지 않다. 또한, AI 시스템을 개발하거나 조달할 때 어떤 원칙과 절차를 따라야 하는지에 대한 지침도 미흡하다. 공공분야에 도입되는 AI 시스템의 경우 시민의 권리와 의무에 영향을 미치게 되는 경우가 많으며, 민간 기업의 AI 시스템과 달리 시민은 다른 선택의 여지가 없다. 따라서 공공분야의 AI 시스템을 엄격하게 규율할 수 있는 정책과 감독 시스템이 마련될 필요가 있다.

2. 공공분야의 인공지능 관련 규범

(1) 전자정부법 및 시행령, 지능형 전자정부서비스 도입 및 관리지침

2021년 6월 8일 개정된 전자정부법에서 신설된 제18조의2는 전자정부서비스를 제공할 때 인공지능 등의 기술을 활용할 수 있도록 했으며, 행정안전부 장관이 이를 위해 행정적 · 재정적 · 기술적 지원 등 필요한 지원을 할 수 있도록 했다. 그러나 전자정부서비스에서 AI 기술을 활용할 수 있는 근거를 마련했을 뿐, AI의 특성과 위험성을 고려한 도입 원칙이나 절차 등을 구체적으로 규정하고 있지는 않다. 다만 행정안정부 장관이 지원하는 사업의 선정 및 관리 등에 필요한 세부사항을 고시하도록 하고 있는데, 이것이 〈지능형 전자정부서비스 도입 및 관리지침〉이다. 이 지침의 제7조는 지원사업을 선정할 때 "인공지능 기술 도입에 따른 사업 목적 및 활용 방안의 구체성", "인공지능 학습용 데이터 확보의 지속성 및 품질의 적정성" 등을 고려하도록 하고 있으나, 이는 지원사업을 선정할 때의 고려사항일 뿐 실제 도입 예정인 AI 시스템에 대한 평가 사항은 아니다. 제8조는 행정기관에서 AI 기술을 적용한 서비스를 제공할 때 "의사결정과정에 사전고지 의무를 준수하고 서비스 이용 거부 및 이의제기, 설명요구 권리를 보장"하도록 "노력"할 의무만을 규정하고 있다. 또한 전자정부법령 및 지침은 전자정부서비스에 적용되는 것이지, 공공기관이 업무 목적으로 활용하는 모든 AI 시스템에 적용되는 것은 아니다.

전자정부법

제18조의2(지능형 전자정부서비스의 제공 등) ① 행정기관등의 장은 인공지능 등의 기술을 활용하여 전자정부서비스를 제공할 수 있다.

② 행정안전부장관은 행정기관등의 장이 인공지능 등의 기술을 효율적으로 활용할 수 있도록 행정적 · 재정적 · 기술적 지원 등 필요한 지원을 할 수 있다.

③ 제1항 및 제2항에 따른 인공지능 등의 기술의 종류, 활용 및 지원에 필요한 사항은 국회규칙, 대법원규칙, 헌법재판소규칙, 중앙선거관리위원회규칙 및 대통령령으로 정한다.

전자정부법 시행령

제15조의2(지능형 전자정부서비스의 도입 및 활용) ① 법 제18조의2제1항에 따라 전자정부서비스 제공에 활용할 수 있는 인공지능 등의 기술은 다음 각 호와 같다.

1. 자연어 처리(컴퓨터를 이용해 사람의 언어를 분석하고 처리하는 기술을 말한다)
2. 음성인식
3. 영상인식
4. 그 밖에 전자적 방법으로 학습 · 추론 · 판단 등을 구현하는 기술로서 지능형 전자정부서비스 제공에 필요한 기술

② 행정안전부장관은 법 제18조의2제2항에 따라 다음 각 호의 사업을 지원할 수 있다.

1. 인공지능 등의 기술을 전자정부서비스에 적용 · 실증하는 사업
2. 인공지능 등의 기술을 여러 전자정부서비스에 활용할 수 있도록 공통기반을 구축하는 사업
3. 인공지능 등의 기술을 빅데이터 분석 기법 등 다른 기술이나 서비스와 융합하는 사업
4. 그 밖에 지능형 전자정부서비스의 도입이나 활용에 필요한 사업

③ 제2항에 따른 사업의 선정, 관리 등에 필요한 세부사항은 행정안전부장관이 정하여 고시한다.

지능형 전자정부서비스 도입 및 관리지침

제7조(지능형 전자정부서비스 지원사업의 선정 및 관리)

② 행정안전부장관은 지능형 전자정부서비스 지원사업을 선정할 때에는 다음 각 호의 사항을 고려해야 한다.

1. 사업의 타당성 및 효율성
2. 인공지능 기술 도입에 따른 사업 목적 및 활용 방안의 구체성
3. 인공지능 학습용 데이터 확보의 지속성 및 품질의 적정성
4. 공공업무의 자동화, 새로운 양질의 지능형 대국민 서비스 제공 가능성
5. 그 밖에 행정안전부장관이 정하는 사항

제8조(지능형 전자정부서비스 제공 및 관리) ① 행정기관등의 장은 인공지능 기술을 적용하여 지능형 전자정부서비스로 제공할 때는 의사결정과정에 사전고지 의무를 준수하고 서비스 이용 거부 및 이의제기, 설명요구 권리를 보장하기 위해 노력해야 한다.

② 행정기관등의 장은 지능형 전자정부서비스 도입·운영시 인공지능 활용이 갖는 역기능을 방지하고, 사람 중심의 인공지능 구현을 위해 노력해야 한다.

③ 행정기관등의 장은 안전한 지능형 전자정부서비스를 제공하는 데 필요한 보안 대책을 수립해야 한다.

(2) 지능정보화기본법

지능정보화 기본법에서는 여러 지능정보기술을 나열하고 있는데, 인공지능이라는 용어를 사용하고 있는 것은 아니지만, 제2조 제4호 가목의 기술이 인공지능과 관련된 것으로 볼 수 있다. 공공기관의 AI 활용과 관련해서는 제14조에서 공공지능정보화를 추진하고 이를 효율적으로 추진하기 위해 필요한 방안을 마련하도록 하고 있을 뿐이다. AI 시스템이나 공공기관에 국한된 것은 아니지만, 제46조는 지능정보서비스 제공자가 서비스를 제공할 때 "장애인·고령자 등의 접근과 이용의 편익을 증진하기 위하여 노력"하도록 하고 있다. 또한 "국가기관등은 지능정보제품을 구매할 때 장애인·고령자 등의 정보 접근과 이용 편의를 보장한 지능정보제품의 우선 구매를 촉진하기 위하여 필요한 시책을 마련"해야 한다. 제57조는 국가기관 및 지방자치단체가 지능정보서비스를 제공, 이용할 때 정보보호 시책을 마련하도록 하고 있다. 제60조는 지능정보기술 및 서비스의 안전성 보호조치의 내용과 방법을 정하여 고시하도록 하고 있는데, 아직 고시되지 않은 것으로 보인다.

지능정보화 기본법

제2조(정의)

4. "지능정보기술" 이란 다음 각 목의 어느 하나에 해당하는 기술 또는 그 결합 및 활용 기술을 말한다.

가. 전자적 방법으로 학습 · 추론 · 판단 등을 구현하는 기술
나. 데이터(부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식을 말한다)를 전자적 방법으로 수집 · 분석 · 가공 등 처리하는 기술

제14조(공공지능정보화의 추진) ① 국가기관등은 공공서비스의 지능정보화를 도모하고 국민 편익 증진 등을 위하여 행정, 보건, 사회복지, 교육, 문화, 환경, 교통, 물류, 과학기술, 재난안전, 치안, 국방, 에너지 등 소관 업무에 대한 지능정보화(이하 “공공지능정보화”라 한다)를 추진하여야 한다.

제46조(장애인 · 고령자 등의 지능정보서비스 접근 및 이용 보장)

② 지능정보서비스 제공자는 그 서비스를 제공할 때 장애인 · 고령자 등의 접근과 이용의 편익을 증진하기 위하여 노력하여야 한다.

④ 국가기관등은 지능정보제품을 구매할 때 장애인 · 고령자 등의 정보 접근과 이용 편의를 보장한 지능정보제품의 우선 구매를 촉진하기 위하여 필요한 시책을 마련하여야 한다.

제57조(정보보호 시책의 마련 등) ① 국가기관과 지방자치단체는 정보를 처리하거나 지능정보서비스를 제공 또는 이용하는 모든 과정에서 정보보호를 위한 시책을 마련하여야 한다.

제60조(안전성 보호조치) ① 과학기술정보통신부장관은 행정안전부장관 등 관계 기관의 장과 협의하여 지능정보기술 및 지능정보서비스의 안전성을 확보하기 위하여 다음 각 호와 같은 필요한 최소한도의 보호조치의 내용과 방법을 정하여 고시할 수 있다.

1. 지능정보기술과 지능정보서비스의 오작동 방지에 관한 사항

2. 지능정보기술 및 지능정보서비스에 대한 권한 없는 자의 접근, 조작 등 전자적 침해행위의 방지에 관한 사항

3. 지능정보기술 및 지능정보서비스의 접속기록, 운용 · 활용기록의 저장 · 관리 및 제공 등에 관한 사항

4. 지능정보기술의 동작 및 지능정보서비스 제공을 외부에서 긴급하게 정지하는 것(이하 “비상정지”라 한다)과 비상정지에 필요한 알고리즘의 제공에 관한 사항

5. 기타 지능정보기술 및 지능정보서비스의 안전성 확보를 위해 필요한 사항

(3) 행정기본법 제20조

행정기본법 제20조는 처분에 재량이 있지 않은 경우, AI 시스템을 포함하여 완전히 자동화된 시스템으로 처분을 할 수 있도록 하고 있다. AI 기반의 행정처분이 모든 단계에서 사람의 개입 없이 자동적으로 이뤄지는 경우 민주적 정당성의 문제가 제기될 수 있다. 제20조는 행정의 디지털화를 촉진하고 행정의 효율성과 국민의 편의를 높이기 위해 자동적 처분에 관한 입법적 기준을 마련하기 위한 취지로 입법화되었다. 무분별한 자동적 처분의 도입을 막기 위해 자동적 처분을 도입하려면 법률에 근거를 두도록 하였고, 처분의 재량이 있는 경우는 자동적 처분의 도입을 허용하지 않았다. (행정기본법 해설서)

행정기본법

제20조(자동적 처분) 행정청은 법률로 정하는 바에 따라 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 처분을 할 수 있다. 다만, 처분에 재량이 있는 경우는 그러하지 아니하다.

그런데 자동적 처분을 규정하는 법률에서 처분 대상의 절차적 권리가 약해지거나 배제되어서는 안될 것인데, 이 조항의 규정 만으로는 자동화된 처분을 위해 어떠한 안전조치가 필요하고, 처분 대상이 되는 당사자의 권리에 어떠한 영향을 미칠 것인지 모호한 상황이다. 이 조항은 2021년 3월 시행되었으나, 2024년 말 현재 자동적 처분에 대한 조항을 두고 있는 법률은 「수입식품안전관리 특별법」 제20조의2(수입신고 수리의 자동화)에 불과하다. 이 법률에서도 대상 및 절차 등 세부적인 사항은 시행령에서 정하도록 하고 있다. 한편, 법제처는 질의에 대한 답변에서, 자동적 처분을 도입할 때 행정청이 준수해야 하는 입법 기준 등을 마련하고 있으며, 2024년 중에 중앙행정기관 등에 '자동적 처분 입법 가이드라인'을 안내, 배포할 예정이라고 밝혔다.

한편, 개인정보보호법 제37조의2는 (인공지능 시스템을 포함하여) 완전히 자동화된 시스템으로 개인정보를 처리하여 이루어지는 결정이 자신의

권리 또는 의무에 중대한 영향을 미치는 경우, 정보주체에게 해당 결정을 거부할 권리와 설명을 요구할 권리 등을 부여하고 있는데, 행정기본법 제29조에 따른 자동적 처분은 제외하고 있다. 그 이유는 행정절차법 · 행정소송법 등 처분에 대한 일반 규정에 따라 정보주체의 권리가 보장되고 있는 점을 고려한 것이라고 한다. 또한, 공공분야의 경우에도 행정기본법 제20조의 자동적 처분이 아닌 자동화된 결정에 대하여는 개인정보보호법의 자동화된 결정 규정이 적용되므로 자동화된 결정에 대한 정보주체의 권리를 보장해야 한다.

(4) 공공분야 인공지능 도입을 위한 실무자 안내서(안)

2021년 4월, 행정안전부 디지털정보국 공공지능정책과는 〈공공분야 인공지능 도입을 위한 실무자 안내서(안)〉을 공개하였다. 본 안내서(안)은 "의견수렴을 위해 작성된 초안"이며 내용변경이 가능하다고 명시하고 있으나, 2024년 말까지 공식적인 지침으로 채택되지 않았다. 그 이유가 무엇인지에 대한 질의에 대해 행정안전부는 "공공분야 인공지능 도입을 체계적으로 확산하기 위해 우선 '공공 AI 도입·활용 전략'을 수립 중에 있으며, 이와 관련하여 필요한 지침 등을 마련할 예정"이라고 답했다.

본 안내서(안)은 AI의 특수성을 고려하여 일반 정보화사업과 차별화된 절차를 제시하고 있다. 예를 들어, 기획단계에서는 데이터 수집·조달 계획을 수립하고, AI 윤리 심의를 하도록 하고 있으며, 구축단계에서는 AI 모델 및 알고리즘을 선정하고, 학습 성능을 평가하며, 데이터 관리 계획을 수립하도록 하고 있다. 운영단계에서는 AI 역기능을 모니터링하고, 재학습 계획을 수립할 것을 권고하고 있다.

본 안내서(안)은 공공분야 AI 사업 담당자가 느끼는 어려움으로 ▲AI 성능 평가를 위한 공통 기준이 없어 업체 선정 및 사업 검수가 어려움, ▲알고리즘이 공개되지 않아 내부 구조 접근이 제한되며 돌발상황 대처가 어려움 등이 있음을 인식하고 있는데, 본 안내서(안)이 만들어진 이후에도 공공분야에서 여러 AI 시스템 도입이 추진되고 있음에도 불구하고 공식적이고 체계적인 지침을 왜 발간하지 않고 있는지 의문이다.

(5) 챗GPT 활용 가이드

공공분야에서 AI의 도입, 활용과 관련된 지침이 부재한 가운데, 챗GPT 활용 가이드가 처음으로 발간되었다. 이는 2023년 1월, 윤석열 대통령이 대통령실 비서관과 정부부처에 챗GPT를 업무에 잘 활용하라고 언급했기 때문인 것으로 보인다.

행정안전부는 2023년 5월, 공무원을 위한 〈챗GPT 활용방법 및 주의사항 안내서〉를 발간했다. 이는 8p 정도의 간단한 안내서로서, 공공에서 활용 가능한 분야를 ▲정보탐색능력 활용, ▲언어능력 활용, ▲컴퓨터능력 활용 등 3가지 분야로 나눠, 7가지의 세부적인 활용 방법을 예시와 함께 간단하게 안내하고 있다. 챗GPT의 문제점도 지적하면서 질문에 의사결정이 완료되지 않거나, 외부로 반출이 허용되지 않은 비공개 정보나, 개인정보를 입력하지 않도록 주의하도록 하고, 챗GPT가 내놓은 답변은 반드시 사실여부에 대해 확인을 거치도록 했다.

2023년 6월에는 국가정보원에서 〈챗GPT 등 생성형 AI 활용 보안 가이드라인〉을 발간하였다. 국정원이 본 가이드를 발간한 것은 국정원이 공공부문의 사이버보안을 담당하고 있기 때문이다. 이 가이드는 생성형 AI 기술의 대표적인 보안 위협 7가지 유형을 설명하면서, 공공분야에서 생성형 AI를 사용할 때의 주의사항을 제시하고 있다. 또한 공공기관에서 생성형 AI를 포함한 AI 시스템을 도입할 때의 고려사항과 보안 대책도 제시하고 있다. 그런데 생성형 AI 뿐만 아니라 다른 기능과 특성을 갖는 AI 시스템도 도입되고 있는만큼, 공공기관에서 AI 시스템을 도입할 때의 원칙과 절차를 규정하는 지침이 필요하다. 이 외에 서울디지털재단이 2023년 3월에 발간한 〈ChatGPT 활용사례 및 활용 팁-업무활용편〉이 있다.

공공분야의 AI 도입과 관련한 기본적인 원칙과 절차에 대한 고민없이, 대통령이 특정 AI 애플리케이션을 활용하라고 지시한 것이나, 대통령 한마디에 이에 대한 지침을 내놓는 정부부처의 행태가 적절한 것인지 의문이다.

(6) 공공부문 초거대 AI 도입·활용 가이드라인

대통령직속 디지털플랫폼정부위원회와 한국지능정보사회진흥원은 2024년 4월, 〈공공부문 초거대 AI 도입·활용 가이드라인〉을 발간하였다. 이 가이드라인은 공공부문 기관들이 초거대 AI를 도입·활용하는데 필요한 관련 기준과 절차, 고려사항을 안내하는 것을 목적으로 한다.

본 가이드라인은 초거대 AI 도입 원칙으로 다음과 같은 5가지를 제시하고 있다. ▲민간의 최신 기술을 적기에 도입하고 활용, ▲행정 프로세스와 조직문화 혁신을 함께 수행, ▲부처 간 칸막이를 없애고 하나의 정부를 구현, ▲국가 안보와 국민 권리 보호를 보장, ▲2020년 12월 과기정통부가 발표한 인공지능 윤리기준 준수 등이다.

도입절차는 다음과 같은 6가지 단계로 이루어진다.

1. 데이터보안등급 검토
2. 클라우드 구성방안 : 민간/공공 클라우드
3. 데이터 학습방식 : 범용 LLM, 파인튜닝, 사후학습 등
4. 서비스 도입 방식 : 서비스구매 / 조달절차를 통한 용역발주
5. 서비스 레벨 목표 : 정확도, 응답시간, 가용성 등
6. 유지보수 및 운영

본 가이드라인은 공공분야에서 AI를 도입할 때 참조할 수 있는 공식적인 첫 가이드라인이라고 할 수 있다. 다만, 초거대 AI에 기반한 시스템에 초점을 맞춘 것이다. 일반적인 가이드라인이 만들어진 것은 다행이지만, 아직 개론적인 수준이다. 또한, 도입 원칙에서 국민 권리를 보호하도록 하고 AI 윤리기준을 준수하도록 하고 있으며, 사전 고려사항으로 생성형 AI가 가져올 위험을 이해하고 방지하도록 하고 있음에도 불구하고, 이것이 도입절차의 하나로 포함되어 있지 않다. 또한 구체적으로 어떻게 위험을 식별, 완화할 것인지에 대해서 설명하고 있지 않다.

(7) 지방자치단체의 AI 관련 조례

일부 지방자치단체들은 자체적으로 인공지능 조례 혹은 인공지능산업 육성 및 지원에 관한 조례를 제정하고 있다. 예를 들어, 경기도, 부천시가 인공지능 기본조례를 제정하였으며, 경기도, 경상남도, 경상북도, 광주광역시, 대구광역시, 세종특별자치시 등이 인공지능산업 육성 및 지원 조례를 제정하였다. 경기도는 앞서 언급한 두 개의 조례 외에도 인공지능 스타트업 육성 및 지원 조례, 인공지능 윤리기반 조성에 관한 조례를 제정하였다.

그러나 지방자치단체의 조례는 그 적용 범위가 해당 지역으로 제한되고 규범의 내용도 구체적이지 못하다. 예를 들어, 경기도 인공지능 기본조례는 금지된 인공지능과 고위험 인공지능에 대한 정의규정을 두고 있지만, 금지된 인공지능의 개발을 원칙적으로 금지하고 고위험 인공지능은 관계 법령에 따른 규제를 엄격하게 적용하는 범위에서 허용한다는 것만 규정하고 있을 뿐, 어떠한 인공지능이 금지된 인공지능 및 고위험 인공지능에 포함되는지, 규제의 구체적인 내용이 무엇인지는 규정하고 있지 않다. 또한, 대부분의 다른 조례들은 인공지능 산업육성 및 지원에 초점을 맞추고 있다.

경기도 인공지능 기본조례

제2조(정의) 이 조례에서 사용하는 용어의 뜻은 다음과 같다.

1. “인공지능”이란 인간의 지능이 가지는 학습, 추론, 지각, 자연언어 이해 등의 기능을 다양한 수준의 자율성을 가지고 작동하도록 설계된 소프트웨어나 컴퓨터시스템, 그 밖의 장치를 말한다.
2. “금지된 인공지능”이란 인간의 존엄성, 사람의 생명, 자유와 평등에 대한 침해의 위협이 명백하다고 간주되는 등 관계 법령 및 사회적 통념 위반에 해당하는 인공지능을 말한다.
3. “고위험 인공지능”이란 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용되는 인공지능을 말한다.
4. “저위험 인공지능”이란 금지된 인공지능 및 고위험 인공지능 외의 인공지능을 말한다.

제3조(기본원칙) 인공지능 개발 및 이용은 다음 각 호의 내용을 기본원칙으로 추진하여야 한다.

1. 인공지능은 인류의 발전과 편의를 위하여 개발 및 이용되어야 한다.

2. 인공지능의 개발 및 이용은 특정한 개인 또는 단체가 성별, 나이, 인종, 민족, 지역 신체적 조건, 종교, 경제적 사정, 정치적 견해 등을 이유로 차별받지 않도록 이루어져야 하며, 사회적 약자 및 취약계층 등을 위한 접근성을 보장해야 한다.

3. 인공지능의 개발 및 이용은 개인의 개인정보자기결정권을 보장하고 신뢰성과 투명성이 확보될 수 있도록 이루어져야 한다.

4. 금지된 인공지능은 원칙적으로 개발을 금지하고, 고위험 인공지능은 관계 법령에 따른 규제를 엄격하게 적용하는 범위에서 허용하며, 저위험 인공지능은 원칙적으로 허용하도록 노력한다.

3. 공공분야 인공지능 도입 현황

국가기관, 지방자치단체, 공공기관, 공기업 등 공공분야에서 어떠한 AI 시스템을 어느 정도 도입하고 있는지에 대한 공식적인 자료는 없다. 공공분야 AI 도입 현황에 대한 연구보고서를 통해서 짐작할 수 있을 뿐이다.

(1) 인공지능 기반 공공서비스 현황

2023년 10월, 국회입법조사처가 발간한 〈인공지능 기반 공공서비스 실태와 개선과제〉 보고서는 17개 시도에서 운영 중이거나 계획 중인 인공지능 기반 공공서비스를 대상으로 조사하였다. 다양한 인공지능 기반 공공서비스를 운영 중인 지방자치단체에서부터 아직 서비스를 제공하지 않는 곳까지 지방자치단체별로 편차를 보이고 있다.

이 보고서는 인공지능 기반 공공서비스의 유형을 ▲챗봇, ▲ChatGPT 기반 서비스, ▲노인, 장애인 그리고 외국인 대상 인공지능 기반 공공서비스, ▲인공지능 기반 보안관제, ▲하천 시설물 정보 제공 시스템 및 하수관로 결함탐지 시스템 등으로 구분하였다.

인공지능 기반 공공서비스로 가장 많이 활용되고 있는 것은 챗봇인데, 다수의 지방자치단체가 챗봇을 운영하고 있다. 2017년에 시작된 대구시의 '뚜봇', 2019년 서울시 '120상담챗봇', 2023년 경북의 생성형 AI 기반 '챗경북', 부산의 복지를 중심으로 하는 챗봇 '자립 꿀단지' 등이다. 챗경북은 ChatGPT 기반 서비스인데, 보도자료 작성지원 등 반복업무, 정책자료 추천, 정부예산분석, 공모과제 사업제안서 작성 등 사업기획 업무를 지원한다. 노인, 장애인 등을 대상으로 한 서비스는 경남의 인공지능 스피커, 제주시의 AI-IoT 기반 어르신건강관리사업 및 지능형 민원서식 작성 도우미, 대전시 스마트 미러(시·청각 장애인을 위한 민원 안내 시스템) 등이다. 인공지능 기반 보안관제는 CCTV 통합관제센터의 관제요원이 움직임이 있는 영상만을 선별하여 모니터링 할 수 있도록 한다. 대전시는 '하천 시설물 정보 제공 시스템'을, 서울시는 인공지능 기반의 '하수관로 결함탐지시스템'을 구축하고 있다.

가장 다양한 인공지능 기술을 적용한 공공서비스를 운영하고 있는 지방자치단체는 서울시였다. 120상담챗봇 뿐만 아니라 직원용 업무챗봇, 자연어 처리 기술을 적용한 인공지능 회의록 지원시스템, 지능형 영상협업시스템, 규칙 기반의 업무자동화, 보안관제 플랫폼 등을 운영하고 있다.

(2) 공공분야 인공지능 도입 현황

소프트웨어정책연구소(SPRI)는 2022년에 공공부문 408개 기관(중앙행정기관 41개, 지방자치단체 17개, 공공기관 350개)의 인공지능 활용실태를 조사하였다. 각 기관 업무담당자를 대상으로 설문조사를 실시하여, 인공지능 활용 분야, 목적, 사용된 인공지능 기술, 인공지능 도입의 장애요인 등을 분석하였다.

소프트웨어정책연구소는 2024년에 〈공공부문 AI 도입현황 연구〉 보고서를 다시 발간하였는데, 이번에는 이전 10년(2013~2022년)간 조달청 '나라장터'의 입찰정보와 계약정보를 활용하여 분석함으로써, 설문조사 방식을 사용한 선행 연구에 비해 정확하고 객관적인 데이터를 확보하기 위한 것이었다.

소프트웨어정책연구소의 조사에 따르면, 2022년까지 이전 10년간 인공지능 도입 계약 건수는 3,870건이다. 공공부분 420개 기관의 56.7%인 238개 기관이 인공지능을 도입한 것으로 나타났다. 공공기관에서의 인공지능 도입은 2016년 알파고 사건 이후에 매년 빠르게 증가하고 있다. 인공지능 계약이 2015년 107건에서 2022년 922건으로 7배 가까이 증가했고, 금액도 2016년 938억원에서 2022년 1조 831억원으로 늘었다. 2016년 이전에는 대국민 공공서비스를 위한 인공지능의 비중이 높았지만, 그 이후에는 업무 효율화 등 내부역량강화 목적의 인공지능 도입이 증가하다가 2020년 이후에는 챗봇, 자연어처리 등 민원서비스를 위한 기술이 발전하면서 다시 공공서비스 비중이 증가하고 있다.

인공지능이 도입된 분야와 관련해서는 전자정부, 민원 서비스 등에 관련된 시스템 수요가 많은 일반행정 분야가 지속적으로 전체의 20% 이상의 가장 높은 비중을 차지하고 있고, 다음으로 산업/고용(산업육성, 일자리매칭 등, 16.5%), 교통수송/건설(지능형 교통망 등, 11.6%), 기상/재난안전(기상 예측 등, 10.4%) 순으로 많이 도입되었다. 기술적 측면에서는 과거에는 문서 디지털화를 위한 OCR 기술과 장애인 접근성 제고를 위한 TTS(Text-to-Speech) 기술이 주로 활용되었지만, 2016년 이후 다른 기술의 적용이 확대되었다. 챗봇은 2016년 3건을 시작으로 2022년 161건으로 급증하였고, 음성인식과 비정형 데이터 처리 기술이 발전하면서 STT(Speech-to-text)와 자연어처리 기술의 적용도 급속하게 확대되었다.

계약 건수는 국가기관(36.8%), 지자체(23.5%), 준정부기관(19.4%), 기타공공기관(19.4%) 순으로 많았다. 다만, 총 계약금액은 국가기관(56.2%), 준정부기관(27.4%), 지자체(11.7%), 기타공공기관(11%) 순이었는데, 지자체의 계약 규모가 상대적으로 작다는 것을 알 수 있다. 국가기관은 일반행정 분야의 인공지능이 가장 높은 비중을 차지하고 있고, 그 밖에는 기관별 고유 업무에 따른 인공지능을 도입하고 있다. 지자체 역시 일반행정 분야가 가장 많고, 그 다음은 지역의 교통문제 해결을 위해 교통수송/건설 분야의 인공지능 도입이 많았으며, 기상/재난안전 분야의 도입도 증가하고 있다. 지자체를 제외한 다른 공공기관은 공공서비스 보다는 내부역량제고의 목적으로 인공지능을 도입한 사례가 2배 정도 많았다.

이 조사는 공공분야에 도입된 인공지능의 위험도별로 분류하지는 않았다. 인공지능의 위험도에 대한 사회적 기준이 아직 마련된 것은 아니지만, 연구자 나름의 기준을 세워 평가를 할 수도 있었을 것이다. 이는 공공분야에서 안전 및 인권에 미칠 위험이나 영향이 큰 인공지능을 구분하는데 도움이 될 것이다. 예를 들어, 프로젝트의 자세한 내용을 확인한 것은 아니지만, 다음과 같은 인공지능 프로젝트는 안전 및 인권에 미칠 잠재적인 위험이나 영향이 클 것으로 예상된다.

경상남도 AI 기반 지능형 119 신고접수시스템 구축, 경찰청 AI 음성인식 기술을 활용한 성폭력 피해자 조사·지원 시스템 고도화, 대검찰청 빅데이터기반 지능형 디지털증거 통합분석 플랫폼 개발, 제주도 대형 버스 운전자 졸음탐지 및 대응서비스 구축, 사회보장정보원 머신러닝·RPA 기반의 사회서비스 바우처 부정수급 탐지시스템 구축, 인천광역시 데이터 기반 야간 골목길 안전시스템 구축, 관세청 빅데이터 분석모델 (BigFINDER) 개발 등

(3) 공공부문 AI 도입·활용 활성화 방안

2024년 4월 17일, 대통령 직속 디지털플랫폼정부위원회는 전체회의를 개최하여 6개 정책과제를 발표·논의했는데, 그 중 하나의 정책과제가 공공부문 AI 도입·활용 활성화 방안이었다. 정부는 공공부문 AI 도입·활용 활성화를 위한 세부 방안을 다음과 같이 제시했다. 첫째, 공공분야 AI 활용 성공사례 창출과 확산. 특히, 초거대 AI 활용 지원규모를 확대하고, 행정효율화 및 현안해결 등 분야별로 대규모 프로젝트를 발굴해 집중지원한다. 둘째, 공공부문 AI 활용 역량 강화. 〈공공부문 초거대AI 도입·활용 가이드라인〉을 배포하고, AI를 활용하는 실무자 맞춤형 교육프로그램을 제공한다. 셋째, 정부 전용 초거대 AI 기반 구축. 정보화전략계획(ISP) 사업을 통해 중장기 로드맵을 마련하고 시범 적용대상 및 정부 내 학습데이터 선정 등 사전 준비도 진행한다.

2024년 7월 15일 디지털플랫폼정부위원회와 과기정통부는 초거대 인공지능(AI) 공공서비스 개발 사업을 추진하겠다고 밝히면서 8개 사업을 선정

했다. 8대 사업은 ▲초거대 AI 기반의 통합 연구개발 지원 서비스 ▲스마트 소방 안전 서비스 ▲생성형 AI기반 국방시설 건축 행정 지원 ▲AI 근로감독관 지원 서비스 ▶청년 농업인 특화 서비스 ▲초거대 AI 기반 특허심사 지원 서비스 ▲장애인 소통지원 초거대AI 멀티모달 서비스 ▲초거대AI 활용 느린학습자 조기발견 지원 서비스 등이다.

(4) 인공지능 거버넌스

현재 공공분야 인공지능 정책이 어느 부처의 소관인지는 다소 모호하다. 전자정부법은 행정안전부 소관이지만, 지능정보화기본법은 과기정통부 소관이다. 국가적 차원의 인공지능 정책 추진을 소관하고 있는 부처는 과기정통부이고, 2024년 12월 국회를 통과한 AI 기본법 역시 과기정통부 소관이다. 아직 과기정통부나 행정안전부 모두 공공분야에서 인공지능 도입 현황에 대한 조사는 시행하고 있지 않다.

인공지능 기술, 산업, 정책에 대한 주무부처는 과기정통부이지만, 윤석열 정부는 인공지능 주요 정책의 심의·의결을 위한 최고위급 거버넌스 기구로 대통령 산하 국가인공지능위원회를 2024년 9월 26일 설립하였다. AI 기본법이 제정되기 전이지만 대통령령을 통해 일단 출범시킨 것이다. 이 위원회는 대통령과 민간위원이 공동으로 위원장을 맡고, 정부 위원 외에 30명의 민간위원을 두고 있는데, 산업계, 학계, 법조계 인사로만 구성되었을 뿐 시민사회를 대표할 수 있는 인사는 포함하고 있지 않다.

2024년 11월 27일에는 한국 ‘AI 안전연구소’가 출범하였다. AI 안전연구소는 AI의 기술적 한계, 인간의 AI 기술 오용, AI 통제력 상실 등으로 발생하는 다양한 AI 위험에 체계적, 전문적으로 대응하기 위한 목적으로 설립되었다. 이는 2023년 말 영국에서 개최된 AI 안전성 정상회의 이후 영국, 미국 등 주요 국가에서 AI 안전연구소를 설립하고 있는 추세에 발맞추기 위한 것이다.

4. 공공분야 인공지능 활용 문제점

(1) 공공분야 인공지능 등록제도의 필요성

현재 국내에서는 공공기관의 인공지능 시스템 도입시 보고 또는 등록하도록 하고 있지 않기 때문에, 도입 현황을 간접적인 방식으로 확인할 수밖에 없다. 공공기관 인공지능 도입 현황에 대해 언제든 객관적인 정보를 파악할 수 있기 위해서는 공공기관의 인공지능 등록제도를 고려할 필요가 있다. 즉, 공공기관이 인공지능 시스템을 도입할 때, 목적, 활용된 인공지능 기술, 기능, 개발업체 등 주요 정보를 등록하도록 하는 것이다.

민간기업에서 도입하는 인공지능 시스템은 소비자가 선택하지 않을 수 있지만, 시민들이 다른 공공서비스를 선택할 수는 없으므로 공공분야 인공지능 시스템의 영향으로부터 벗어나기 힘들다. 공공기관이 내부 업무목적으로 도입할 경우에도, 시민들의 권리와 의무에 영향을 미치는 정책 결정을 위해 인공지능을 사용할 경우 시민들은 자신도 모르는 사이에 인공지능의 영향을 받게 된다. 따라서 공공분야에 사용되는 인공지능의 위험성을 사전에 파악하고 대응하기 위해서는 어느 기관에서, 어떤 목적으로, 어떠한 기능을 가진 인공지능이 사용되고 있는지에 대한 정보가 필요하다.

또한, 공공분야에서 유사한 목적의 인공지능 시스템을 활용하는 경우가 많을텐데, 이러한 현황을 공유되거나 수요를 파악함으로써 중복투자와 예산낭비를 방지할 수 있을 것이다. 소프트웨어정책연구소의 보고서에서도 “많은 지자체에서 적은 예산으로 체계적인 계획 없이 산발적으로 인공지능 시스템을 도입”하고 있는 문제를 지적하며, “지역 공통의 문제를 발굴해 상위 기관이 주도적으로 인공지능으로 해결한 후 하위 기관에 체계적으로 확산·보급하는 전략이 필요”하다고 제안하고 있다. 이와 관련하여 지능정보화기본법 제7조 제6항은 과기정통부 장관으로 하여금 “국가기관등이 추진하는 지능정보화 사업의 중복투자 방지 등을 위한 방안을 마련”할 수 있도록 하고 있다.

(2) 공공분야 인공지능 도입, 활용 지침의 부재

소프트웨어정책연구소의 조사에 따르면 2022년 이전 10년간 인공지능 도입 계약 건수가 3,870건에 이를 정도로 많은 인공지능 시스템이 이미 도입되었고, 앞으로 더 많은 인공지능 시스템이 도입될 것이다. 그럼에도 불구하고, 공공기관에서 인공지능 시스템을 도입할 때 지켜야할 원칙이나 절차 등을 규정한 지침이 없다는 것은 큰 문제다. 2021년 4월에 행정안전부가 의견수렴을 위해 발표한 〈공공분야 인공지능 도입을 위한 실무자 안내서(안)〉은 아직 공식적인 지침으로 채택된 것이 아니며, 실제 공공기관에서 이 안내서(안)을 어느 정도 활용하고 있는지도 알려져있지 않다. 국회입법조사처는 보고서에서 향후 개선 과제의 하나로 “중앙정부 차원의 가이드라인, 기획 및 조정이 필요”하다고 제안하고 있다. 그나마 2024년에 〈공공부문 초거대 AI 도입·활용 가이드라인〉이 만들어졌다. 그러나 초거대 AI에 기반한 시스템에 초점을 맞추고 있고, 아직 개괄적인 수준이다. 초거대 AI를 포함한 모든 종류의 AI 시스템에 대한, 그리고 AI 위험성에 대한 구체적인 평가 기준 및 절차 등을 포함한 보다 구체적인 지침이 만들어질 필요가 있다.

한편, 영국의 경우 2020년 6월에 〈AI 조달 가이드라인(Guidelines for AI procurement)〉을 발표하였다. 이 가이드라인에서는 공공기관에서 인공지능 시스템을 조달할 때, 10가지 핵심적 고려사항과 함께 조달 절차의 단계별로 고려해야 할 사항을 제시하고 있다. 한국에서도 이와 같이 인공지능에 특화된 조달 지침이 필요하다.

(3) 인공지능 인권영향평가 수행의 필요성

〈공공분야 인공지능 도입을 위한 실무자 안내서(안)〉에는 AI 시스템을 도입할 때 AI 윤리를 검토하도록 하고 있지만, 이 안내서(안) 자체가 공식적인 지침도 아니고 구체적으로 AI 윤리를 어떻게 반영해야 하는지는 자율에 맡겨져 있다. 〈공공부문 초거대 AI 도입·활용 가이드라인〉에서도 어떻게 윤리기준을 준수하고 위험을 식별, 완화할 것인지에 대한 절차를 포함하고 있지 않다.

공공분야에서 AI 시스템을 도입할 때 그 위험성에 비례하는 안전조치를 취하기 위해서는 인권영향평가를 수행할 필요가 있다. 국가인권위원회는 2022년 〈인공지능 인권 가이드라인 권고〉에서 정부와 국회에 '인공지능 인권영향평가' 제도의 도입을 촉구한 바 있으며, 2024년 5월 23일에는 자율적인 인권영향평가 수행을 돕기 위해 '인공지능 인권영향평가 도구'를 마련하여 공개하였다. 2023년 10월 30일 미 바이든 행정부의 AI 행정명령의 이행을 위한 미국 관리예산처(OMB)의 이행 가이드에서는 AI 영향평가, 실제 환경 테스트, 독립적인 평가, 이해관계자와의 협의 등 연방 정부 기관이 AI를 도입할 때 수행해야 할 지침을 제공하고 있다. 국내에서도 공공기관이 AI 시스템을 도입할 때 그 위험성을 평가하고 안전조치를 마련하기 위한 인권영향평가를 수행할 필요가 있으며, 이러한 절차가 지침에 반영되어야 한다. 또한, 인권영향평가를 제대로 수행하기 위해서는 해당 AI 시스템에 의해 영향을 받는 사람들, 혹은 이들을 대변할 수 있는 시민단체나 전문가가 영향평가에 참여하도록 보장할 필요가 있다.

국내 공공기관은 인권경영체계를 구축하여 운영해야 하고, 그 일환으로 인권영향평가를 수행하고 있는데, 여기에 AI 시스템 도입시에도 인권영향평가를 수행하도록 할 필요가 있다. 2024년 12월 국회를 통과한 AI 기본법 역시 고영향 AI 사업자로 하여금 인권영향평가를 수행하려고 '노력할 의무'를 부과하고 있다.

(4) 법적 근거의 마련

AI 시스템 도입이 단지 기존 업무를 효율화하는 것에 그치는 것이 아니라, 영향을 받는 사람의 권리나 책임에 큰 영향을 주게되는 경우 추가적인 법적 근거를 마련할 필요가 있다. 예를 들어, 출입국 과정에서 사람들의 신원을 원래 파악해왔다고 하더라도, 신원확인을 위해 사람들의 생체인식정보를 새롭게 수집한다든가 불법 출입국 파악을 위해 AI 시스템을 통해 프로파일링을 하는 경우 이러한 내용이 관련 법에 추가적으로 명시될 필요가 있다. 그래야 법치주의 관점에서도 AI 시스템 사용이 정당화될 수 있으며, 도입의 적정성에 대한 문제제기도 가능해질 것이다.

(5) 공공부문 AI 전문가의 충원 및 교육

국회입법조사처의 보고서는 AI 기술에 대한 공무원과 시민들의 인식 제고의 필요성을 제기하며, 공무원 데이터 리터러시 향상을 위한 교육, AI 기술 인력 강화 등을 제안하고 있다. 소프트웨어정책연구소의 보고서 역시 도메인 지식과 AI 지식을 모두 겸비한 전문가가 필요하기 때문에, 기관 내부의 AI 전문가를 육성할 필요가 있다고 제안하고 있다.

법집행 분야 인공지능 현황과 문제점

1. 도입

국가의 법집행 권한은 국민의 안전을 보호하기 위하여 제한적으로 행사되어야 한다. 특히 법집행 AI는 국가의 최첨단 감시와 권위주의 관행에 복무할수 있다는 점에서 사회적으로 통제될 필요가 있다.

한국 경찰과 출입국기관의 AI는 범죄 예방, 범죄 수사, 출입국 관리 등 공익적 목적으로 개발되어 배치되고 있다. 그러나 관련 AI로부터 중대한 인권 영향을 받게 될 시민들은 사전에 의견을 제출하거나 반영할 수 있는 기회가 없다. 관련 AI의 개발과 배치에 대한 의사결정이 대부분 밀실 속에서 이루어지고 있기 때문이다. 정보주체 시민은 자신의 개인정보가 AI 개발을 위해 수집되고 이용될 때 동의권을 행사하기는 커녕 통지조차 받지 못하고 있다. AI 시스템의 배치에 필요한 적법한 근거들이 갖추어져 있는지, 또 AI 시스템의 개발과 배치 과정에서 개인정보보호법이 철저히 준수되고 있는지도 모호하다. 경찰과 출입국기관의 권한 오남용을 통제하거나 감시하는 거버넌스는 찾아볼 수 없다. 이들 AI로 인한 피해가 발생하였을 때 설명가능성이나 구제가능성이 보장될 수 있는지 여부 역시 불확실하다.

경찰과 법무부는 인권 위험을 방치하고 자기 기관과 사업계의 수요에만

기반하여 AI 기술의 기량을 최대한 발휘하는 데만 급급하다. 법집행 AI의 개발과 배치에 대한 법적 통제가 미비하다는 점도 문제를 악화시키고 있다. 그러나 경찰과 출입국 AI의 개발과 배치는, 모든 수명주기에서 인권과 균형을 이루어야 하며 헌법적으로 통제되어야 한다. 특히 인권에 미치는 영향이 매우 심각한 AI에 대해서는 금지해야 하고, 그렇지 않더라도 고위험에 이르는 AI에 대해서는 엄격히 통제하는 법제도가 마련되어야 한다.

2. 경찰 AI

(1) 도입

한국 경찰은 치안분야에 AI 기술을 적극적으로 배치하고 막대한 예산을 집행하며 관련 산업을 진흥해 왔다.

2017년 11월 발표한 「4차 산업혁명 대응계획」은 경찰청·과학기술정보통신부 합동으로 지능화 기술과 치안인프라 융합 과제를 다음과 같이 포함하고 있다. 첫째, 실종아동·용의자 신원확인 지능형 CCTV(이하 '경찰 지능형 CCTV 사업'), 3D 얼굴인식, AI 기반 범죄분석, 온라인상 음란물 차단, 드론 기반 자율순찰·추적 등을 개발하고 실증한다. 둘째, 2020년까지 범죄 장소 및 종류, 범인 영상 또는 인식자료를 AI 기술로 분석한다. 경찰 지능형 CCTV 사업은 "인공지능 기반 복합인지 기술"로 지칭된다.

경찰은 2019년부터는 5개년 단위 「치안분야 과학기술 진흥 종합계획(이하 '치안과학기술계획')」을 별도로 수립하기 시작하였다. 2019년부터 2023년에 해당하는 제1차 치안과학기술계획 중 "첨단기술을 활용한 범죄 예측·대응 기술 개발" 과제는 앞서 4차 산업혁명 대응계획에서 밝힌 치안 기술 과제와 중첩적이다. 이 과제는 빅데이터와 AI를 이용한 ①범죄 분석·인지 기술, ②범죄 예측·예방 기술, ③민원대응 기술로 구성되는데, ②범죄 예측·예방 기술의 경우 "CCTV 등 카메라를 통하여 범죄를 예측하고, 안면인식을 통하여 범죄자를 추적하고 위험행동을 예측"하는 데 그 목적을 두고 있다.

이러한 사업 계획 하에 추진되어 온 경찰 지능형 CCTV 사업은 복합인지 기능으로 신원확인과 추적, 그리고 범죄와 행동 예측을 목표로 하며, 2023년 개발과 실증 테스트를 마쳤다. 그러나 경찰과 같은 법집행 기관이 공공장소에서 실시간으로 얼굴 또는 동작 등 생체인식을 통해 원격으로 개인을 식별하고 추적하는 것은, 국제규범에서 대단히 위험한 AI로 분류되며 원칙적으로 금지된다는 점은 고려되지 않았다.

2024년부터 2028년에 해당하는 제2차 치안과학기술계획은, GPT 3.5 등 범용 파운데이션 모델이 지구적으로 큰 반향을 미친 후에 마련되었다. 제2차 치안과학기술계획은 현안 대응을 넘어 제1차 치안과학기술계획보다 체계적으로 치안기술을 개발관리하기 위한 로드맵을 담았다. 경찰은 앞서 지능형 CCTV 사업을 한층 고도화하고 전국화하고자 한다. 나아가 경찰 데이터와 유관기관 데이터는 물론 민간이 보유한 외부 데이터까지 통합하여 AI 학습에 광범위하게 이용하고, AI 기반으로 범죄를 예측하여 실시간으로 대응하는 시스템을 구축하겠다고 밝혔다.

이하에서는 제2차 치안과학기술계획을 중심으로 AI 기반 기술을 본격적으로 배치하려는 경찰의 계획을 살펴보고 인권에 미치는 영향에 대하여 검토한다.

(2) 관련 규정

경찰의 치안과학기술 사업은 「국가경찰과 자치경찰의 조직 및 운영에 관한 법률」(제33조) 및 동법 시행령 「치안분야 과학기술 진흥에 관한 규정」(제3조)에 근거하고 있다. 다만 이 규정은 '연구개발' 사업에 대한 근거이다.

국가경찰과 자치경찰의 조직 및 운영에 관한 법률

제33조(치안에 필요한 연구개발의 지원 등) ① 경찰청장은 치안에 필요한 연구 · 실험 · 조사 · 기술개발(이하 "연구개발사업"이라 한다) 및 전문인력 양성 등 치안분야의 과학기술진흥을 위한 시책을 마련하여 추진하여야 한다.

② 경찰청장은 연구개발사업을 효율적으로 추진하기 위하여 다음 각 호의 어느

하나에 해당하는 기관 또는 단체 등과 협약을 맺어 연구개발사업을 실시하게 할 수 있다.

1. 국공립 연구기관

2. 「특정연구기관 육성법」 제2조에 따른 특정연구기관

3. 「과학기술분야 정부출연연구기관 등의 설립 · 운영 및 육성에 관한 법률」에 따라 설립된 과학기술분야 정부출연연구기관

4. 「고등교육법」에 따른 대학 · 산업대학 · 전문대학 및 기술대학

5. 「민법」이나 다른 법률에 따라 설립된 법인으로서 치안분야 연구기관 또는 법인 부설 연구소

6. 「기초연구진흥 및 기술개발지원에 관한 법률」 제14조의2제1항에 따라 인정받은 기업부설연구소 또는 기업의 연구개발전담부서

7. 그 밖에 대통령령으로 정하는 치안분야 관련 연구 · 조사 · 기술개발 등을 수행하는 기관 또는 단체

③ 경찰청장은 제2항 각 호의 기관 또는 단체 등에 대하여 연구개발사업을 실시하는 데 필요한 경비의 전부 또는 일부를 출연하거나 보조할 수 있다.

④ 제2항에 따른 연구개발사업의 실시와 제3항에 따른 출연금의 지급 · 사용 및 관리 등에 필요한 사항은 대통령령으로 정한다.

치안분야 과학기술 진흥에 관한 규정

제3조(치안분야 과학기술 진흥 종합계획 및 시행계획의 수립 등) ① 경찰청장은 「국가경찰과 자치경찰의 조직 및 운영에 관한 법률」(이하 “법”이라 한다) 제33조제1항에 따른 치안분야 과학기술 진흥을 위한 시책의 일환으로 5년마다 치안분야 과학기술 진흥 종합계획(이하 이 조에서 “종합계획”이라 한다)을 수립하여야 한다. 〈개정 2020. 12. 31.〉

② 종합계획에는 다음 각 호의 사항이 포함되어야 한다.

1. 치안분야 과학기술의 현황과 전망

2. 치안분야 과학기술의 발전방향과 목표

3. 치안분야 과학기술의 국내외 환경 분석과 경쟁력 강화 시책

4. 치안분야 과학기술의 중점기술 개발 전략

5. 치안분야 과학기술 전문인력 양성계획

6. 치안분야 과학기술 진흥을 위한 중장기 투자계획

7. 그 밖에 치안분야 과학기술 진흥을 위하여 경찰청장이 필요하다고 인정하는 사항

③ 경찰청장은 종합계획에 따라 연도별 시행계획(이하 이 조에서 “시행계획”이라 한다)을 세우고 추진하여야 한다.

④ 시행계획에는 다음 각 호의 사항이 포함되어야 한다.

1. 해당 연도의 치안분야 과학기술 개발 등에 관한 추진 방향

2. 치안분야 과학기술 진흥에 관한 분야별 세부계획

3. 주요 치안분야 과학기술 연구개발사업별 투자계획

4. 그 밖에 치안분야 과학기술 진흥을 위하여 경찰청장이 필요하다고 인정하는 사항

⑤ 제1항부터 제4항까지에서 규정한 사항 외에 종합계획과 시행계획의 수립 및 시행에 필요한 사항은 경찰청장이 정한다.

연구개발의 기술적 성과를 실제 법집행 현장에서 배치하기 위해서는 경찰 작용을 통제하는 관련 법률들과 개인정보보호법 하에서 적법한 근거를 갖추어야 할 것이다. 그러나 경찰이 첨단기술 시제품의 연구개발을 마치고 실증 테스트 및 시범 설치 등을 거친 후 실제로 배치하는 과정에서 적법한 근거를 갖추었는지는 모호한 상황이다.

(3) 현황

경찰은 다양한 첨단 AI 기술을 개발 및 배치해 왔으며, 2024년부터 적용되는 제2차 치안과학기술계획에서 보다 본격적으로 AI 기술을 배치하기 위하여 기술적, 물적, 제도적, 거버넌스적 체계를 구축하고자 한다.

우선 경찰은 “범죄예방 대응체계” 측면에서 ‘실시간’ 범죄대응 시스템을 개발할 예정이다. 최근 증가한 이상동기범죄(묻지마 범죄) 대응을 위하여 과거의 범죄 데이터를 분석하여 “범죄의 패턴을 분석하고 예측”하기 위한

모델을 개발한다. 또한 “범죄의 선제적 예측과 실시간 식별 및 대응”을 위하여 스토킹, 성범죄자의 행동패턴을 분석하고 “개인의 위험도 평가”를 위한 알고리즘을 개발한다. 또한 순찰차, 로봇, 드론, CCTV 등을 통해 공공장소에서 “실시간”으로 “행동”과 “음성”을 인식하여 비정상 위험을 탐지한다. 경찰은 이와 같은 범죄자 행동 패턴 분석과 비정상 행동 탐지기술을 연계하여 통합적인 ‘지능형 위험징후 분석 시스템’과 ‘지능형 통합관제 시스템’을 설계하는 것을 목표로 하고 있다.

경찰은 궁극적으로 여러 기관에서 연계 또는 통합된 데이터와 실시간 범죄 예측 기능을 경찰이 지휘 또는 지원하는 한국형 RTCC(Real Time Crime Center, 실시간 범죄 대응 센터)를 구축하고자 한다. 경찰은 지능형 실시간 통합관제를 위하여, 자체적으로 보유한 다양한 치안데이터와 다양한 유관기관이 보유한 기관의 데이터를 연계 및 통합한다. 대상이 되는 타기관 데이터에는 오픈소스인텔리전스(OSINT), 지리정보시스템(GIS), CCTV, 국토 공간의 디지털트윈 융합에 기반한 집회 및 군중밀집 위험 시뮬레이션이 포함된다. 이렇게 통합된 데이터는 ‘실시간’으로 분석되어 범죄발생 가능성이나 패턴을 예측하는 기술 개발에 사용된다.

한편, 경찰은 공공장소에서 동선과 객체 인식을 통해 사회자 약자와 범죄자 등 특정 개인을 자동으로 식별하고 추적하는 멀티모달 다중영상 다중객체 인식 및 매칭 기술도 개발할 예정이다. 이는 앞서 경찰 지능형 CCTV 사업을 고도화하는 내용이다.

경찰 지능형 CCTV 사업은 공공장소에서 실시간으로 얼굴 또는 동작 등 생체인식을 통해 원격으로 개인을 식별하고 추적하는 기술로서, 1차 치안과학기술계획 기간 중 개발이 완료되어 2023년 11월 경기도 안양시 통합관제센터에서 실증까지 마쳤다. 해당 사업은 △복합인지 핵심원천 SW기술(과기정통부), △신원확인용 웨어러블 디바이스(산업부), △복합인지 기술 응용 및 인프라(경찰청) 등으로 구성되어 있으며 총 325억원의 예산이 소요되었다. 2024년 사업계획부터는 새로 7.5억의 예산을 투입하여 해당 기술을 고도화하고 경찰청 시스템과 연계하여 전국화하려는 목표로 추진되고 있다.

또한 경찰은 “경찰 장비 첨단화” 측면에서 드론을 도입할 예정이다. 드론의 적용범위는 교통의 관리 및 단속으로부터, 상시 순찰 뿐 아니라 집회와 시위 채증까지 확대해갈 예정이다. 더불어 4족보행 등으로 실외에서 자율적으로 순찰하는 무인 순찰로봇을 개발한다. 이를 위해서 “경찰 순찰로봇 운영 규칙”의 제정을 추진한다. 민원 응대용으로는 대화형 챗봇 및 위험 대응 기능을 갖춘 대면형 로봇을 개발할 계획이다.

나아가 경찰은 “과학기술 기반 고도화”를 추진할 예정이다. AI 기반으로 과학수사 플랫폼을 고도화하여 과학수사 데이터를 통합적으로 분석하겠다는 것이다. 여기에는 경찰이 필요로 하는 데이터를 자동으로 정제하고 정형화하는 기술이 포함되어 있고, 이동통신사와 협력하여 보이스피싱 대화에 ‘실시간’으로 대응하는 기술 또한 포함되어 있다.

경찰은 “인공지능 활용 사이버범죄”에 대한 대응 방안도 계획하고 있다. 여기에는 AI를 배치하여 피싱사이트를 ‘실시간으로’ 탐지하고 가짜뉴스를 자동으로 감지하는 기술이 포함되어 있다. 다양한 인터넷 플랫폼의 대화나 게시물을 ‘실시간’으로 모니터링하는 사이버 순찰 기술도 개발한다. 모바일 금융사기를 탐지하기 위하여 범죄자 ‘음성’을 식별하고 피해자 ‘감정’을 인식하는 기술도 개발한다.

더불어 경찰은 첨단 모빌리티 환경에 부합하는 “교통안전 체계” 구현을 위하여 지능형로봇법을 개정하여 로봇 안전인증 의무화와 사고기록 의무화를 추진한다는 계획도 가지고 있다.

2차 치안과학기술계획은 특히 치안과학으로 경찰과 치안산업의 혁신기반을 조성하겠다고 강조하고 있다. 경찰의 디지털 전환을 위하여 100여 개 업무별로 분리된 정보시스템을 통합하고 데이터의 활용성을 제고하겠다는 계획을 세웠다. 경찰은 경찰청의 정형, 비정형 데이터를 연계하고, 외부데이터는 구매하며, MOU 제휴를 체결하여 데이터를 제공받는 등 과거보다 적극적인 데이터레이크를 구축할 방침이다. 경찰 빅데이터 플랫폼인 데이터레이크의 경우 2024년 현재 경찰청 각 국, 관에 분산된 내부데이터 125종과 외부데이터 172종, 총 297종의 데이터를 연계하여 공동활용의 기반

을 구축한 상황이다. 경찰은 취합된 데이터들을 AI 학습에 이용할 수 있는 데이터센터 또한 신설하여 경찰 AI 모델 개발을 촉진할 계획이다.

(4) 문제점

경찰이 과학적인 범죄 예방 체계를 갖추고 정확한 증거기반으로 수사를 할 수 있으면 공공 안전이 더 강화되는 측면이 있을 것이다. 그러나 경찰이 범죄 예방과 범죄 수사를 명분으로 정당한 개인정보 처리 목적을 넘어 지나치게 많은 개인정보를 처리하고 개인을 감시한다면 이는 헌법에서 보호하는 기본적 인권을 침해할 수 있다. 나아가 공개된 온라인과 오프라인 공간에서 경찰이 정당한 법적 권한 없이 AI 감시 추적 기술을 사용하는 것은, 무고한 시민의 자유를 위축시킬 수 있다는 점에서 인권에 미치는 위험이 크다. 그런 점에서 경찰이 제2차 치안과학기술계획에서 밝힌 AI 치안기술에는 우려스러운 점이 적지 않다.

첫째, 대시민 투명성이 매우 부족하다. 어떤 치안기술 솔루션은 종합계획 수립 이후 일사천리로 실행까지 이르는데, 이 종합계획에 대해 경찰 외부 시민이 의견을 제출하거나 반영될 수 있는 절차가 거의 없다. 종합계획은 경찰청 내부 총괄기획위원회에서 수립한 후, 전문가 검토, 국민 설문조사, 공청회를 실시하며, 국가과학기술자문회의 심의의결을 거쳐 수립된다고 한다. 연도별 시행계획은 경찰청 내부기구에서 인권영향평가를 시행하고 인권위원회 보고를 거쳐 국가경찰위원회의 의결로 수립된다. 그러나 설문조사 결과는 종합계획을 정당화하는 근거로만 인용되고 있을 뿐 비판 의견을 포함하고 있지 않다. 공청회는 사전적으로나 사후적으로 개최 정보조차 거의 공개되어 있지 않다. 데이터의 처리나 알고리즘의 설계 방향에 대한 정보도 전혀 공개되어 있지 않다. 결국 해당 사업으로 인권에 영향을 받게 될 일반 시민이나 이들을 대표하는 인권시민단체가 해당 사업에 대한 정보를 접하고 의견을 제출할 수 있는 경로가 사실상 부재하다.

둘째, 치안과학기술계획 전반적으로 경찰이 개인정보 데이터를 다루는 방식이 우려스럽다. 경찰은 전국 경찰이 보유한 다양한 치안데이터를 막연한 목적으로 최대한 통합할 뿐 아니라 타기관 또는 민간이 보유한 외부 데

이터와도 연계하고 이를 AI 학습에 쓰겠다는 취지를 밝히고 있기 때문이다.

예를 들어 경찰은 한국형 RTCC 시스템 구축을 위하여 “경찰이 보유한 치안데이터와 유관기관이 보유한 다양한 데이터를 통합하는 체계”를 마련하겠다고 밝혔다. 나아가 경찰은 RTCC를 넘어서는 막연한 경찰 목적으로 업무 목적에 따라 분리된 시스템과 데이터를 통합하고 방대한 데이터레이크를 구축하겠다는 계획을 가지고 있다.

이때 경찰이 보유한 치안데이터가 무엇인지는 구체적으로 알 수 없다. 다만 “법령상 활용가능한 데이터”로서 “국민안전과 범죄예방 등 제한된 목적으로만 활용”한다고 밝히고 있을 뿐이다.

그러나 데이터에는 개인정보가 포함되어 있으며, 특히 경찰이 보유하고 있는 개인정보는 범죄경력자료 뿐 아니라 사상 · 신념, 노동조합 · 정당의 가입 · 탈퇴, 정치적 견해 등 민감정보를 아우른다. 그런데 경찰은 민감할 수도 있는 개인정보를 이처럼 방대하게 이용하며 나아가 통합하겠다는 계획을 가지고 있으면서도 개인정보를 어떻게 보호하겠다는 것인지 그 내용은 밝히고 있지 않다. 다만 종합계획 일부에서 웨어러블 장비와 드론 운용과 관련하여 사생활 침해 우려에 대하여 관련 법령을 정비하고 교육을 실시한다는 정도를 언급하고 있을 뿐이다.

이는 전반적으로 경찰의 인공지능 사업이 개인정보를 수집하거나 이용함에 있어 개인정보보호법 준수 계획을 갖추고 이를 집행하였는지 의심스럽게 만든다. 특히 개인정보 수집과 이용, 처리 등에 적법성 분석과 감독 거버넌스가 매우 취약해 보인다. 경찰이 혹여나 가명처리를 했다는 이유로 개인정보보호법을 우회하거나, 연구개발이라는 이유로 별다른 제약 없이 방대한 데이터 학습을 마친 후 그 결과물인 AI 알고리즘을 그대로 경찰 작용에 배치하려는 것은 아닌지 우려스럽다.

2023년 2월 16일 독일연방헌법재판소는 예측치안을 비롯해 경찰의 자동화된 개인정보 분석 또는 평가가 법익의 균형성을 위반하여 위헌이라고 결정하였다(1 BvR 1547/19, 1 BvR 2634/20). 우선 경찰에 보관된 개인정

보가 개인에 대한 분석 또는 평가를 위해 자동적으로 처리되는 경우, 이 과정에서 개인정보가 사용되는 모든 사람의 개인정보 자기결정권이 원칙적으로 제한된다고 보았다. 경찰이 이전에 확보한 정보를 수사 단서로 활용하는 일이 드문 것은 아니지만, 자동화된 분석 또는 평가처럼 대량의 복잡한 정보를 처리하는 방법은 기본권에 미치는 영향이 크다는 것이다. 따라서 이러한 경찰의 추가적인 개인정보 처리가 헌법적으로 정당하려면 목적 '변경'의 원칙상으로 추가적인 근거가 갖추어져야 한다. 그런데 헤센주와 함부르크주 예측치안에 관한 법률은 "범죄행위를 예방하기 위하여"라는 요건 외에는 별다른 제한을 두고 있지 않아 법적으로 무제한의 데이터 세트를 무제한의 방법으로 처리할 수 있도록 허용했다는 점이 위헌 결정의 이유가 되었다. 경찰이 보관하고 있는 개인정보라 하더라도 추가적인 분석이나 평가에 사용하기 위해서는 그만큼 구체화된 위험이 있다는 사실을 증명해야 한다는 것이다. 특히 주거지 감시, 온라인 수색, 통신 감청, 교통정보 조회 등은 이러한 데이터 분석에 사용되어서는 안된다.

그런데 한국의 경찰은 민감할 수도 있는 개인정보를 경찰 자기관이 보유했다는 이유로 막연한 경찰 목적으로 대규모로 통합하고 광범위한 경찰 목적으로 이를 이용하고 나아가 타기관 또는 민간 데이터와 통합하려고 계획하고 있다. 구체적인 법률적 근거나 이를 위한 계획도 전혀 갖추지 않은 상태이다. 이는 헌법이 보호하는 개인정보에 대한 권리를 침해할 수 있다.

셋째, '실시간' 탐지와 분석, 그리고 '예측치안'을 강조한 AI 알고리즘의 목표가 인권 침해적이다. 경찰은 여러 사업에서 '실시간' 탐지와 분석을 목표로 제시하고 있는데 이는 고도의 감시라는 점에서 인권에 미치는 위험이 크다. 오프라인과 온라인 공간에서 위치나 대화 등에 대한 실시간 탐지는 통신비밀을 침해할 가능성 또한 높다. 또한 확률적으로 도출되는 경찰 AI의 '예측'은 우리 사회에 용인할 수 없는 차별과 인권 위험을 초래할 수 있다. 편향적인 집단 분석에 기반하여 개인을 분석 또는 평가할 경우 특정 인종 등 일부 집단에 차별적인 결과를 초래할 수 있고 그 결과가 다시 학습에 사용되어 '피드백 루프' 문제를 야기할 수 있다.

특히 공공장소에서 실시간으로 얼굴이나 동작 등 생체정보를 이용하여

사람을 원격으로 식별하는 경찰 활동은 국제 인권규범에서 금지하고 있다. 이에 대하여 2024년 8월 발효된 유럽연합 AI법은 “많은 사람들의 사생활에 영향을 미치고, 지속적으로 감시받고 있다는 생각이 들게 하며 집회의 자유 및 그 밖의 기본권 행사를 간접적으로 침해할 수 있다”는 이유에서 원칙적으로 금지하였다. 다만 피해자 식별, 테러 대응 등 극히 예외적인 사유에 대하여 법원의 사전 허가를 받아 제한적으로 허용할 뿐이다. 유엔 인권최고대표는 이보다 앞선 2021년 법적 통제가 결여된 얼굴인식을 유예할 것을 각국 정부에 요구한 바 있다. 한편 유럽연합 AI법은 개인의 특성에만 기반하는 예측치안에 대해서도 금지하였다. AI의 범죄 예측 역시 범죄 활동과 직접적으로 관련된 객관적이고 검증가능한 사실에 기초해야 한다는 것이다.

넷째, 경찰 치안과학기술계획이 치안산업 진흥을 주요 목표로 제시하고 있다는 점이 우려스럽다.

경찰은 1차 치안과학기술계획이 기술 이전에 따른 4억 2,850만원에 달하는 경제적 성과를 달성했으며, 2023년 산업박람회 등을 통해 총 398만 달러(51. 7억 원) 규모의 수출계약을 체결했다고 밝혔다. 2차 치안과학기술계획에서는 보다 본격적으로 치안산업 진흥에 기여하겠다는 계획을 담고 있다. 경찰은 치안과학기술 연구개발을 전담하는 법정기관(과학치안진흥원)과 경찰이 설립을 지원하는 민간 사업자단체(치안산업진흥협회)를 축으로 하여 민-관-경-산-학-연 파트너십을 추진한다. 또한 장차 치안산업의 국내적 적용 뿐 아니라 해외 수출도 국가적으로 지원하고 관리하기 위하여 ‘치안산업 진흥에 관한 법률’을 제정할 방침이다.

그러나 치안기술은 경찰 또는 공공기관이 보유한 시민의 데이터를 직접 자기 목적의 AI 학습에 쓰는 것이다. 이러한 데이터를 적법한 근거를 갖춘 경찰 목적 외로 민간기업의 치안상품 개발을 지원하는 데 이용하는 것은 정당하지도 비례적이지도 않은 개인정보 자기결정권의 침해가 될 수 있다. 나아가 치안기술 개발 자체의 동기와 방향이 공공성보다 상업적 이윤 동기에 치우칠 경우, 비례적인 국가 공권력의 발동으로 시민의 안전을 지키기보다 인권침해적이고 과도한 탐지력을 중시하는 기술을 생산할 수 있다.

시민의 인권에 중대한 영향을 미치는 경찰 AI의 목표 설정이 시민의 감시의 시선이 닿지 않은 채, 경찰과 산업계의 밀실 속에서 이루어진다면 민주주의에 대한 큰 위협이 될 수 있다. 냉전 시기 군산학복합체가 오늘날까지 미국의 민주주의와 평화 체제를 위협하였다는 비판을 상기할 필요가 있다.

다섯째, 인권에 부정적 영향을 미칠 수 있는 경찰 AI의 개발과 배치에 대한 독립적인 인권 감독이 현재 공백 상태이다. 경찰 수사 기법에 대하여 일정한 비밀 보호가 요구된다 하더라도 경찰 작용에 대한 전문적이고 독립적인 인권 감독 거버넌스를 마련할 수 있다. 현재 경찰청 내부에서 자체적으로 시행하는 인권영향평가는 이러한 조건을 충족하고 있지 못하다. 독립적인 인권 감독 거버넌스는 경찰 AI의 적법성, 데이터, 알고리즘을 독립적으로 검토하고, 헌법과 개인정보보호법을 준수하는지 여부도 감독할 수 있어야 한다. 치안과학기술에 대한 초기 계획 단계에서부터 참여하여 경찰의 기술 개발과 배치를 견제하고 인권침해를 구제할 수 있어야 함은 물론이다. 무엇보다 시민의 인권에 중대한 영향을 미치는 경찰의 AI 개발과 배치에 대하여 전반적으로 감독하는 거버넌스 및 통제 절차가 경찰 작용법 등에 법률적으로 규정될 필요가 있다.

3. 출입국 AI

(1) 개요

법무부는 2019년 4월부터 공항 출입국장에 배치하기 위한 목적으로 “인공지능 식별추적 시스템(이하 ‘출입국 AI’)”을 개발해 왔다. 이 사업에서 법무부는 얼굴데이터를 비롯하여 출입국 관리 목적으로 수집한 내국인과 외국인의 개인정보 1억 7천만 건을 복수의 민간기업에 AI 학습용으로 제공하였다.

사업주체인 과학기술정보통신부, 법무부는 물론 개인정보보호위원회도 이 사업이 개인정보보호법을 비롯한 관련 법을 위반하지 않았고 아무런 문제가 없다고 주장한다. 그럼에도 법무부는 해당 학습 데이터와 실증랩을 조기 폐기하여 자신의 개인정보가 사용되었는지 확인해 달라는 내국인과

외국인 정보주체의 요청을 거부하였다. 2024년 12월 현재 이 사건은 헌법소원 심판 중이다.

(2) 관련 규정

법무부는 출입국 AI의 적법성 근거로 출입국관리법의 다음 조항을 제시하였다. 개인정보보호위원회는 이들 조항이 출입국 AI의 적법한 근거가 된다고 추인하였다.

출입국관리법

제1조(목적) 이 법은 대한민국에 입국하거나 대한민국에서 출국하는 모든 국민 및 외국인의 출입국관리를 통한 안전한 국경관리, 대한민국에 체류하는 외국인의 체류관리와 사회통합 등에 관한 사항을 규정함을 목적으로 한다.

제3조(국민의 출국) ⑤ 출입국관리공무원은 제3항에 따라 수집하거나 제출받은 생체정보를 출국심사에 활용할 수 있다.

제6조(국민의 입국) ⑥ 출입국관리공무원은 제4항에 따라 수집하거나 제출받은 생체정보를 입국심사에 활용할 수 있다.

제12조의2(입국 시 생체정보의 제공 등) ⑤ 출입국관리공무원은 제1항 또는 제3항에 따라 제공 또는 제출받은 생체정보를 입국심사에 활용할 수 있다.

제28조(출국심사) ⑥ 출입국관리공무원은 제12조의2제1항 또는 제3항에 따라 제공 또는 제출받은 생체정보를 출국심사에 활용할 수 있다.

그러나 출입국관리법은 “출국심사에” 또는 “입국심사에” 생체정보를 활용할 수 있다고만 규정하고 있다. 이는 출입국자의 신원을 확인하기 위해 생체정보를 1:1로 대조할 수 있다는 의미일 뿐이다. 따라서 출입국관리법은 얼굴정보를 비롯해 내국인과 외국인 출입국자의 개인정보를 1:n으로 자동 식별추적되는 AI 시스템의 개발에 이용할수 있는 정당하고 적법한 근거가 되기 어렵다.

(3) 현황

출입국 AI는 법무부와 과학기술정보통신부가 공동으로 추진하는 사업으

로 주요 목적을 다음과 같이 밝혔다. 첫째, 기존의 “지문인식 방식의 공항 출입국 관리시스템을 데이터인공지능 기반으로 고도화”한다. 둘째, “AI 기업들에게 공공 분야 실증 및 시장수요를 제공함으로써 컴퓨터 비전 기술력의 조기 확보와 국내 AI 산업 활성화에 기여”한다. 이러한 목적 하에 국가는 복수의 신청 기업에게 출입국 데이터를 학습용으로 제공하고, 민간 기업은 국가가 마련한 실증랩에서 자사가 보유한 AI 알고리즘을 학습시켰다.

즉 이 사업은 처음부터, 법무부가 제공한 얼굴이미지 정보 등을 학습자료로 하여 민간 기업 얼굴인식 AI의 국내외 경쟁력을 향상시키는 것을 주요 목표로 삼고 있었다. 출입국 AI 사업을 주무한 과학기술정보통신부 산하 공공기관 정보통신산업진흥원(NIPA)은 “법무부가 보유한 얼굴데이터의 가치가 5천억~1조원의 가치를 지녔다”고 밝혔다. 이 사업으로 본래 “학습용 얼굴데이터 확보를 위해서는 개인의 동의가 필요하며, 1인당 2~10만원 이상의 수집비용 및 가공비용과 관리 부담이 발생”하는 문제를 해결할 수 있다는 것이다. NIPA는 이 사업이 “대규모 공공데이터를 확보”함으로써 AI 기업의 “현안 해결”을 도모하여 성장의 발판을 마련할 수 있을 것이라고 기대하였다. 더불어 이렇게 개발된 얼굴인식 기술은 “은행거래, 쇼핑, 미아찾기, 불법 이민자 식별 등 다양한 분야로 사업영역을 확대할 것”이라고 전망하였다.

2019년 4월부터 2021년 10월 사업의 문제가 공론화되고 중단되기 직전까지 출입국 목적으로 보관된 개인정보가 이 사업을 위해 제공되었다. 총 12곳의 신청 기업이 자사 얼굴인식 AI 학습을 위하여 내국인 5,760만 건 및 외국인 1억 2천만 건의 개인정보를 이용하였다. 약 1.7억 건에 달하는 이 데이터는 본래 출입국 데이터 원본 총 3.2억 건 중 안면사진 크기, 용량, 파일 훼손 여부를 ‘필터링’하여 선별된 것이다. 기업에 제공된 개인정보는 얼굴이미지 정보 외에도 여권번호, 국적, 생년, 성별 등이 있었다. 제공된 얼굴이미지 정보는 ‘전처리’되어 기계판독이 가능한 얼굴인식 데이터로 변환되었으며 출신 대륙, 연령대별로 분류되었다. 신청 기업들은 국가와 명확한 용역 계약을 맺지도 않은 상태였다. 국가는 추후 우수한 업체를 선정하겠다는 이유로 계약을 체결하지 않은 다수 기업에 AI 학습 데이터를 우선 제공하였다.

한편 법무부는 위 출입국 데이터 외에도 공항에 동작인식 CCTV를 수백 대 설치하고 실제 공항 이용객의 이른바 '리얼데이터'를 별도로 수집하고 있었다. '이상행동' 감지 시스템의 개발을 위해서이다. 그러나 출입국 AI 사업이 공론화되자마자 법무부는 공항내 리얼데이터 수집 CCTV를 철거하였다. 개인정보보호위원회는 CCTV 설치와 운영 자체는 범죄 예방 및 시설안전 목적으로 볼 수 있지만 얼굴이나 동작 등 생체인식을 위해서는 별도의 법적 근거가 필요하다고 보았다. 다만 사업이 중지되어 CCTV 영상이 사용된 사실이 없다는 이유로 위법성 판단을 하지 않았다.

2021년 10월 언론보도와 국정감사를 통하여 이 문제가 공론화된 후 시민사회는 출입국 AI의 인권침해를 비판하고 피해자를 구제하기 위해 활동했다.

그러나 이 사업의 개인정보보호법 위반 여부를 조사하고 심의한 개인정보보호위원회는 출입국 AI가 출입국관리법에 근거를 두고 있기 때문에 개인정보보호법을 위반하지 않았다고 결정하였다. 다만 개인정보처리 위탁 사실에 대한 공지를 지체하였다는 이유로 법무부에 1백만원이라는 소액의 과태료 처분을 내렸을 뿐이다.

내국인과 외국인 정보주체 20여 명은 시민사회 도움으로 자신의 개인정보가 이 사업 학습 데이터로 사용되었는지 여부에 대하여 열람을 요청하였다. 그러나 법무부는 얼굴만으로는 개인을 "알아볼 수 없다"는 이유를 들어 이 요청을 거부하였다.

이에 청구인들은 개인정보분쟁조정위원회에 열람권 거부에 대한 분쟁조정을 신청하였다. 그러나 법무부가 AI 학습에 이용된 모든 데이터를 파기하고 실증랩을 폐쇄하는 바람에 이 요청은 기각되었다.

이어서 시민사회는 감사원에 이 사업의 위법성에 대한 공익감사를 청구하였다. 그러나 감사원 역시 감사청구 이유가 부족하다는 이유를 들어 청구를 종결시켰다.

시민사회는 마지막으로 2022년 7월 헌법재판소에 헌법소원을 제기하였

다. 2024년 12월 현재 이 사건은 헌법심사 중에 있다.

(4) 문제점

출입국 AI 사업의 첫번째 문제점은 전반적으로 법적 근거가 모호하다는 것이다. 법무부와 개인정보보호위원회는 출입국관리법이 국가가 보유한 방대한 개인정보를 기업들에 학습용으로 제공하고 출입국 AI를 개발할 수 있는 근거가 된다고 주장한다. 그러나 출입국관리법의 조항은 출입국절차에서 1:1 신원확인 목적으로 개인의 생체정보를 처리할 수 있다고만 규정하고 있다. 이것만으로는 출입국 정보를 다른 목적, 즉 이동하는 정보주체를 자동으로 추적하고 얼굴과 동작을 식별하는 기능을 가진 AI 개발에 이용할 수 있는 근거가 될 수 있다고 보기 어렵다.

둘째, 출입국 AI 사업은 인공지능 학습에 방대한 개인정보를 이용하면서도, 정보주체의 동의를 받기는 커녕 최소한의 공지조차 없었다. 정보주체 보호나 권리 행사를 보장하기 위한 조치도 충분히 이루어지지 않았다. 정보주체는 자신의 개인정보가 학습에 사용되는 데 대하여 반대할 수 있는 기회도 갖지 못했다.

특히 얼굴인식정보는 생체인식정보로서 한층 더 두텁게 보호되는 민감정보에 해당한다. 그럼에도 법무부는 출입국관리법의 무관한 조항을 근거를 제시하며 기업들에게 민감정보를 대규모로 제공하였다. 반면 정보주체들의 열람 요청에는 얼굴이미지 정보만으로 개인을 식별할 수 없다는 이유를 들어 거부하였고, 이후에는 데이터를 아예 폐기하여 정보주체가 피해사실을 확인하고 권리구제를 위해 민사소송 등으로 대응할 수 있는 길을 차단해 버렸다. 이는 국내외 개인들의 방대한 개인정보를 기반으로 대규모 국가 사업을 시행한 국가기관으로서 매우 책임성이 결여된 행태이다.

셋째, 출입국 목적으로 공공장소에서 얼굴인식을 통해 사람을 식별하고 추적하려는 출입국 AI의 목표와 기법이 인권 침해적이다. 사람의 얼굴을 자동으로 식별하여 자동으로 추적하는 AI 기능은 과거에 상상하지 못했던 새로운 기술로 인권에 심각한 위험을 미칠 수 있다. 유럽 AI법은 공공장소

에서 실시간으로 얼굴이나 동작 등 생체정보에 기반하여 사람을 원격으로 식별하는 법집행 기관의 활동을 원칙적으로 금지하였다.

넷째, 국가가 출입국 목적으로 보유하고 있는 방대한 개인정보를 국내 얼굴인식 기업 다수에 제공한 이 사업의 목적과 방법이 중대하게 잘못되었다. 신청 기업들이 이 사업의 명목으로 학습시킨 알고리즘 각각은 출입국 사업 목적이나 계약 관계가 불분명하다. 따라서 이 사업의 주된 목적이 출입국 목적보다 국가가 보유한 방대한 데이터를 관련 기업에 제공하여 산업을 진흥하는 데 치중되어 있다는 비판이 타당하다.

출입국 AI는 공공장소인 공항 출입국장에 실시간 원격 얼굴인식이라는 고위험 AI를 개발하는 국가 사업이었다. 그럼에도 그 개발 과정은, AI 학습에 사용된 개인정보의 정보주체들은 물론 출입국 AI로부터 영향을 받을 국내외 시민들에게 매우 불투명하고 인권침해적이었다. 결론적으로 출입국 AI 개발사업은 국가 기관과 사업계의 수요에만 기반하여 추진되었고, 인권에 미치는 위험이 매우 높았으며 책임성도 갖추지 못했다. ◘

교육 분야 인공지능 이슈 현황

1. 도입

교육분야에서 인공지능(AI)을 어떤 목적으로 어떻게 활용하고 있는지에 대해 실태 조사가 이루어진 바는 아직 없다. 우선 사교육시장을 중심으로 AI 기능을 포함한 학습 애플리케이션(예를 들어, 영어 학습이나 문제 풀이 지원 앱)을 개발하면 개개인이 구매해 학습용으로 활용하고 있으며, 교사들이 수업 준비를 위해 생성형 AI나 시험문제 출제 AI 서비스를 사용하는 등 개별적으로 사용하는 경우가 있을 것이다. 학교에서 공개된 AI 서비스를 수업 중에 활용하거나, 특히 일부 대학에서는 AI 강의 시스템을 실험적으로 개발, 도입하는 경우도 있는 것으로 보인다. 2024년 12월 현재 교육부 차원에서 도입하여 사용하고 있는 AI 시스템은 초등학생의 영어 말하기 학습용으로 개발된 'AI 펭톡' 서비스 정도다. 그러나 2023년부터 교육부가 AI 디지털교과서 도입 계획을 발표하였고 2025년 실제 도입을 목표로 사업을 추진하면서, 교사 및 학부모 등 교육관계자들이 이에 크게 반발하였다. 결국 2024년 12월 26일, AI 디지털교과서의 법적 지위를 교과서가 아니라 교육자료로 정하고, 학교별로 활용 여부를 선택하도록 하는 법이 야당 주도로 국회를 통과하였다. 사실상 AI 디지털교과서 사업이 좌초될 상황에 처하자 교육부는 정부에 재의요구권 행사를 요구하겠다며 끝까지 반발하고 있는 형국이다. 정부가 재의요구권을 행사하여 법안이 사실상

부결되더라도 거대한 반발 여론이 확인된 만큼 AI디지털교과서 사업 추진은 어려움을 겪을 것으로 보인다.

2. 교육분야 인공지능 규범

(1) 교육분야 인공지능 윤리원칙

2022년 8월 11일, 교육부는 「교육분야 인공지능 윤리원칙」을 발표하였다. 2020년 12월에는 범정부 차원의 「인공지능 윤리기준」이 발표되었고, 2022년 5월에는 국가인권위원회가 「인공지능 개발과 활용에 관한 인권 가이드라인」을 발표한 바 있다. 교육부는 "기존의 인공지능 윤리기준(원칙)은 대부분 범용적이고 개발자(공급자) 중심으로 규범화되어, 교육현장에 바로 적용하기에는 제한적"이기 때문에, 교육에 특화된 윤리원칙을 마련한 것이라고 밝혔다.

교육분야 인공지능 윤리원칙은 '사람의 성장을 지원하는 인공지능'이라는 대원칙 하에, 다음과 같은 10개의 세부원칙을 제시하였다 : ▲인간성장의 잠재성을 이끌어낸다 ▲학습자의 주도성과 다양성을 보장한다 ▲교수자의 전문성을 존중한다 ▲교육당사자 간의 관계를 공고히 유지한다 ▲교육의 기회균등과 공정성을 보장한다 ▲교육공동체의 연대와 협력을 강화한다 ▲사회 공공성 증진에 기여한다 ▲교육당사자의 안전을 보장한다 ▲데이터 처리의 투명성을 보장하고 설명 가능해야 한다 ▲데이터를 합목적적으로 활용하고 프라이버시를 보호한다.

그러나 추상적인 원칙만을 제시하고 있을 뿐, 실제 교육분야에서 활용되는 AI의 개발이나 운영에 활용될 수 있는 구체적인 절차나 지침을 제공하지는 않고 있다. 또한, 10대 원칙의 설명에는 AI 산업계를 의식하여 원칙을 완화하는 것으로 보이는 내용도 존재한다. 예를 들어, "교육당사자의 안전을 보장한다"는 원칙의 제정취지에서 "다만, 의도치 않은 개인이나 집단이 희생되지 않도록 주의하고, 안전을 위한 조치가 기술적 발전을 저해하지 않도록 유의하는 자세도 필요할 것이다"라고 설명하고 있다. 또한, "데이터 처리의 투명성을 보장하고 설명 가능해야 한다"는 원칙의 제정취지

에서는 “다만, 데이터 처리의 투명성을 우선적인 원칙으로 하되, 투명성을 달성하는 과정에서 데이터 처리의 효율성 등의 관점도 함께 고려될 수 있을 것이다”라고 설명하고 있다.

이후 구체적으로 살펴볼 AI 디지털교과서 사업이 이러한 원칙을 존중하고 있는지도 의문이다. 교사, 학부모, 야당의 반대에도 불구하고 AI 디지털교과서 사업을 밀어붙인 것은 “교육당사자 간의 관계를 공고히 유지한다”는 원칙에 어긋난다. 충분한 사전 준비와 시범사업도 없이 AI 디지털교과서를 성급하게 도입한 것은 “교육당사자의 안전을 보장한다”는 원칙과 충돌한다.

(2) 인공지능(AI) 공공성 확보를 위한 현장 가이드라인

교육분야 인공지능 윤리원칙에 앞서, 서울시 교육청은 2021년 8월, 「인공지능(AI) 공공성 확보를 위한 현장 가이드라인」을 발표하였다. 이 가이드라인은 단위학교에서 웹 기반 머신러닝 도구나 AI 튜터, AI 스피커, AI 챗봇, 학교 시설관리에 활용되는 AI 등 AI 시스템을 도입할 때 우선 AI의 등급을 평가하고 세부적인 영향평가를 하도록 하고 있다.

AI의 등급은 의사결정 영향의 정도와 AI에서 처리하는 개인정보의 민감성 정도에 따라 4개의 등급으로 구분되고, 위험성이 높을수록 보다 엄격한 절차와 심의를 거치도록 하고 있다. 등급 평가를 위한 구체적인 기준도 제시하고 있다. 또한 AI 영향평가를 수행할 때 참고할 수 있는 체크리스트도 제시하고 있다.

본 가이드라인은 교육분야 인공지능 윤리원칙보다 현장에서 활용할 수 있는 구체적인 지침을 제시하고 있고, 개략적인 수준이기는 하지만 AI 영향평가를 위한 체크리스트를 제시하고 있다는 점에서 의미가 있다. 다만, 실제 현장에서 어느 정도 활용되고 있는지에 대한 조사가 필요하다.

【그림1】 인공지능 등급평가 매트릭스

3. 인공지능(AI) 디지털교과서 개요

교육부에 따르면 AI 디지털교과서는 “학생 개인의 능력과 수준에 맞는 다양한 맞춤형 학습 기회를 지원할 수 있도록 인공지능을 포함한 지능정보화기술을 활용하여 다양한 학습자료 및 학습지원 기능 등을 탑재한 교과서”를 의미한다. 윤석열 정부의 교육부는 AI 디지털교과서 도입을 추진했다. 그러나, AI 기반 코스웨어 형태로 제공되는 AI디지털교과서에 대해 교육 전문가들은 ‘반복 학습을 통해 지식을 주입하는 형태’로 수업의 가이드인 교과서가 주입식이면 교사의 수업도 주입식이 될 수밖에 없다는 점을 들어 우려를 나타냈으며, 전국교직원노동조합, 학부모단체를 비롯한 시민사회단체, 야당 등의 반대에 부딪혔다.

(1) AI 디지털교과서 사업의 경과

교육부는 2023년 2월, '모두를 위한 맞춤 교육의 실현'을 슬로건으로 하는 〈디지털 기반 교육혁신 방안〉을 발표했다. "개별 학생의 역량 및 선호·학습 속도에 최적화된 맞춤 교육 체제"를 실현하겠다는 것이다. "AI 등 첨단 기술을 활용하여 교육의 질 제고가 가능"하고 디지털 대전환 시대에 "교육 내용·방식의 근본적 변화가 요구된다"고 진단하였다. AI 디지털교과서는 이러한 교육 혁신의 핵심적인 수단으로 제안되었다.

2023년 6월 8일, 교육부는 〈인공지능 디지털교과서 추진방안〉을 확정하여 발표하였다. 2025년에 초등 3, 4학년, 중등 1학년, 고등 1학년을 대상으로 수학, 영어, 정보, 국어(특수교육)를 우선 도입하고, 이후 국어, 사회, 과학 등 전과목 도입을 목표로 2028년까지 단계적으로 확대 추진한다는 계획이다. 다만, 발달단계와 과목의 특성 등을 고려하여 초1, 2학년, 고등학교 선택과목, 예체능, 도덕 교과는 제외한다.

추진방안에 따르면, 정부와 공공기관은 통합학습기록저장소(통합로그인, 대시보드 등 포함)를 구축하고, 과목별 디지털교과서는 민간기업이 개발하도록 했다. 기존에는 교과서 출원 자격이 발행사로 제한되었는데, AI 디지털교과서는 에듀테크 기업이 발행사와 컨소시엄을 구성하여 참여 가능하도록 했다. 컨소시엄을 구성하여 AI 디지털교과서 개발에 참여한 에듀테크 기업은 29년 이후에는 단독으로 출원할 수 있는 자격을 부여할 계획이라고 한다.

교육부는 AI 디지털교과서의 교과서로서의 법적 지위를 마련하기 위해 2023년 10월 17일 〈교과용도서에 관한 규정〉을 개정하였다. 개정된 시행령에는 디지털교과서의 정의 및 검정 절차별 필요사항이 포함되었다. 디지털교과서는 지능정보화 기술을 활용한 학습지원 소프트웨어로 정의(시행령 제2조 제2호)했고, 디지털교과서를 검정할 때에는 기술결함 및 기술·서비스 적합성 여부 등에 대해서도 검정하도록 했다.

2023년 2월 계획 당시에는 2024년 상반기로 예정되었던 AI 디지털교과

서에 대한 검정심사는 2024년 9월에야 진행되었다. 2024년 9월 24일 본심사결과가 발표되었고, 이어 이의신청 검사, 수정본 검토를 거쳐 11월 29일 최종 합격결과가 발표되었다. 검정 심사 결과 12개 출원사에서 제작한 76종의 교과서가 합격하였다.

(2) AI 디지털교과서 개발 가이드라인의 주요 내용

교육부는 2023년 8월 30일, 〈인공지능(AI) 디지털교과서 개발 가이드라인〉(이하 가이드라인)을 발표하였다. 가이드라인에 따르면, AI 디지털교과서 제공을 위해 포털이 구축되며, 하나의 계정으로 AI 디지털교과서 포털과 각 발행사의 디지털교과서를 이용할 수 있도록 통합 인증 체계가 제공된다.

【그림2】 AI 디지털교과서 서비스 구성도

*출처: 가이드라인 21쪽

AI 디지털교과서에서는 개별 학생별로 강·약점, 학습 태도 등을 다각도로 진단하고 분석 결과가 제공되며, 학생들의 학습 이해도와 특성 분석을 바탕으로 개인의 능력과 목표에 맞춘 맞춤 학습 경로 및 콘텐츠를 제시한다. 또한, AI 기술을 활용한 학습 데이터 수집·분석을 통해 학생 개개인의 특성에 맞는 개별화된 학습을 지원하는 소위 'AI 튜터' 서비스가 제공된다.

학생에게 맞춤형 콘텐츠를 제공하기 위하여 학생의 흥미, 수준, 학습 상황 등을 분석하는데, 이는 학생들의 민감한 개인정보를 기반으로 할 수 있다. 가이드라인에 따르면, "말하기를 좋아하는 학생의 영어 학습을 위해 음성인식 기반의 대화 시뮬레이션 콘텐츠를 제공하여 학생이 지속적으로 학습에 몰입할 수 있도록 개인의 학습 패턴을 고려한 학습 콘텐츠를 추천"하는 방식을 예시로 들고 있다.

가이드라인에서 예시로 제시하고 있는 학생 대시보드 사례를 보면, ▲학생들의 기본 개인정보부터, ▲학습 참여도(로그인/로그아웃 일시, 로그인 횟수, 학습한 콘텐츠의 수, 학습 시간, 게시글 등록건수), ▲학습 성취도(평가 결과, 과제 제출 여부, 문제 풀이 건수 및 결과, 오답노트 유무 및 건수), ▲학습 이력(최근 학습한 단원, 최근 학습한 문제, 풀어본 문제, 클릭한 콘텐츠), ▲학습 분석(어려워하는 단원, 학습 맵 추천, 학습 역량, 로그인 간격 또는 학습 시간 등 학습패턴) 등 방대한 개인정보를 수집, 분석하는 것을 예정하고 있다.

【그림3】 학습 진단 및 추천 개요

*출처: 가이드라인 74쪽

대시보드에서는 AI 디지털교과서가 비단 수업시간 뿐만 아니라 방과 후에도 학생의 모든 학습 과정을 모니터링할 뿐만 아니라, 심지어 학생의 '기분'까지 파악하는 기능을 포함하고 있다. 이러한 기능이 '감정인식' 기술을 통해 파악되는 것인지는 실제 AI 디지털교과서의 기능을 검토할 필요가 있다. 학부모의 대시보드를 통해서는 자녀의 데이터를 열람하고 자녀의 기분까지 파악할 수 있도록 정보를 제공하고 있다.

【그림4】 학생 대시보드 사례

*출처: 가이드라인 86쪽

‘AI 튜터’는 질의응답, 추가 학습자료 제공, 학습전략 제안, 학습진도 모니터링, 피드백 및 성취평가, 오답노트 등의 기능을 제공한다. 학생이 궁금한 점이 있을 때 질문하면 답변하는 기능을 포함하고 있는데, 챗봇형, 음성인식형 등 개발사에 따라 다양한 형태로 구현할 수 있다.

‘AI 보조교사’ 서비스는 교사에게 학생별 학습활동 정보를 제공하고, 맞춤 수업설계를 지원한다. 교사는 학생의 데이터를 바탕으로 개별적인 학습계획을 수립하고, 학습활동을 모니터링한 결과를 바탕으로 개별 학생 맞춤형 학습경로를 제공할 수 있다.

AI 디지털교과서는 클라우드 기반(SaaS) 웹 서비스 형태로 제공되는데, 이를 위해 클라우드 보안인증(CSAP) ‘중’ 등급 이상의 인프라(IaaS)와 SW(SaaS) 사용을 의무화하고 있다. 학생의 학습데이터의 수집 및 저장은 발행사에 의해 이루어진다. 발행사는 AI 디지털교과서에서 발생하는 학습데이터를 자체적으로 활용하여 교과에 특화된 학습 분석 정보를 제공하게 된다.

가이드라인은 학생들의 개인정보를 포함한 데이터 관리 정책을 수립할 때 다음과 같은 사항을 포함하도록 하고 있다.

- 신뢰성 확보 : AI 디지털교과서에서 수집된 데이터는 정확하고 신뢰할 수 있어야 한다. 학습데이터에 오류가 있거나 왜곡된 정보가 포함되어 있을 경우, 학습 알고리즘의 성능이 저하되거나 부정확한 결과를 도출할 수 있으므로, 데이터의 정확성을 검증하고, 필요한 경우 데이터 정제 과정을 거쳐야 한다. 또한, 학습데이터 수집과 활용의 전 과정에서 데이터 편향성이 최소화되도록 데이터 품질과 위험을 관리해야 한다.
- 데이터 활용 동의 : 학생의 학습데이터를 사용하기 위해서는 데이터 수집 시점, 데이터 활용 목적, 범위, 기간 등에 대해 동의서를 통해 학생(학부모)의 명시적인 동의를 얻어야 한다.
- 데이터 보안 : AI 디지털교과서 활용 과정에서 발생하는 학습데이터는 개발사가 엄격한 보안과 안전 규정을 준수하여 관리하여야 한다. 이를 위해 관리적, 기술적 보안 사항에 대한 지침 준수가 필요하다.

- 목적 외 활용 금지 : AI 디지털교과서를 통해 수집된 데이터는 AI 디지털교과서 서비스 고도화를 위한 목적으로 사용하여야 하며, 발행사의 자체 서비스를 위한 목적으로 사용해서는 안된다. 또한, 이를 위해 자체 서비스 인프라와 AI 디지털교과서의 인프라를 분리하여 관리해야 한다.
- 데이터 제공 : 데이터 전송요구권에 따라 학생(학부모)의 데이터 전송요구 등의 요청에 적절히 대응해야 한다.
- 개인정보 비식별화 : 국가수준 학습분석 등을 위한 학습데이터는 비식별(익명화, 가명화) 조치하여야 한다.
- 데이터 보존 및 파기 : 개발사는 AI 디지털교과서에서 수집된 학생의 학습데이터를 개인정보 활용 동의서에 따라 관리·보존하여야 하며, 데이터 보존 기간이 지난 학생의 학습데이터는 사용자 동의하에 보관 주기를 연장하거나 안전한 방법으로 파기하여야 한다.

가이드라인은 AI 디지털교과서를 개발할 때 AI 리스크 관리를 위한 프로세스를 마련하고 준수하도록 하고 있다.

- 안정성 및 포괄성 관리 : AI 디지털교과서는 선정성, 폭력성을 띄는 내용 또는 사회적 가치관을 훼손 및 왜곡하는 내용 등을 포함하지 않도록 해야 한다.
- 공정성 관리 : AI 알고리즘이 학습데이터를 통해 편향된 결과를 야기하지 않도록 해야 한다.
- 알고리즘의 투명성 : AI 디지털교과서는 개인정보 활용 방식과 특정 의사결정 또는 행동을 수행한 방식에 대해 설명 가능하도록 해야 한다.
- 책임성 관리 : AI의 특정 의사 결정 또는 행동에 대한 책임 주체가 명확히 설정되어야 한다.

그러나 이상의 내용은 AI 디지털교과서 제작업체가 고려해야 할 가이드라인일 뿐이다. 실제 교육부의 검정을 통과한 AI 디지털교과서가 가이드라인에서 제시하고 있는 기능을 어떠한 기술적 방식을 통해서 구현했는지, 그리고 개인정보 보호와 AI 리스크 관리를 위한 원칙들을 어떤 절차를 통해서 이해하였는지에 대한 조사가 필요하다.

(3) AI 디지털교과서의 검정 과정

AI 디지털교과서의 검정 절차는 크게 내용심사와 기술심사로 구분된다. 내용심사는 교과별 검정기관에서 추진하는데, 수학은 한국과학창의재단, 영어 및 정보는 한국교육과정평가원이 담당한다. 기술심사는 기술심사지원기관인 한국교육학술정보원과 협력하여 추진되는데, 개발사가 제출한 '자체기술검증결과서'를 기초 자료로 활용한다.

【그림5】 AI 디지털교과서 검정 기술심사 기준

심사 영역	심사 항목	심사 요소	검정 심사	현장적 합성
사용성 검사 (10)	• AI 디지털교과서가 기술적 결함이나 오류 없이 작동하는가?	• 디지털교과서의 기능 오류 확인 • AI 성능 테스트 • 부하 테스트	○	-
	• AI 디지털교과서가 웹 접근성 및 상호 호환성을 확보하고 있는가?	• 기기 및 브라우저 간 호환성 • 장애 학생 접근의 수월성 • 자동 번역을 통한 다국어 지원	○	-
기술 표준 준수 (10)	• AI 디지털교과서에 포함된 기술이 관련 규격이나 표준을 준수하였는가?	• 인프라 환경 요건 충족 • 관련 표준 준수	○	-
	• AI 디지털교과서 개발 준수사항을 반영하였는가?	• 목적 외 활용 금지 조치 준수 • 선행학습 금지 조치 준수 • 인공지능 윤리 준수 및 조치 • 저작권 확보	○	-
적합성 검사 (40)	• AI 디지털교과서의 연계 기능이 적절하게 구성되었는가?	• 인증체계 • 교육과정 표준체계 적용 • 시작 화면 구성	○	○
	• AI 기반 맞춤형 학습 지원 기능이 적절하게 작동하는가?	• 학습 진단 및 추천 • 대시보드 및 데이터 시각화 • 학습 안내 및 지원을 위한 AI 튜터 • 수업 설계 및 처방을 위한 AI 보조교사 • 교사 재구성 기능 지원	○	○
	• AI 디지털교과서의 UI/UX 설계와 상호작용이 사용자 관점에서 편리하게 구성되었는가?	• UI/UX의 편의성 • 상호작용의 적절성	○	○
신뢰성 검사 (40)	• 데이터를 적절하게 수집하고 안전하게 관리하는가?	• 데이터 수집 · 저장 • 데이터 전송	○	○
	• 개인정보 및 정보보안 체계가 신뢰성 있게 운영되는가?	• 개인정보 보호, 예방, 대처방안 • 정보 보안, 예방, 대처방안 • 관련된 물리적/기술적 체계 운영	○	○
	• 사용자 지원 및 서비스 관리가 신뢰성 있고 안정적으로 운영되는가?	• 사용자 지원 및 대응 관리 체계 운영 • 장애관리 체계 운영 • 서비스 품질관리체계 구성	○	○

*출처: 가이드라인 162쪽

자체기술검증결과서는 개발사가 기술심사 기준에 따라 외부 기관에 의뢰하거나 자체 사용성 점검 결과와 규격 검사 결과 등을 구체적인 문서로 작성하여 제출한다. 검정심사에 합격한 AI 디지털교과서는 학생, 교사 등 실제 사용자가 참여하는 현장적합성 검토를 진행하며, 필요할 경우 수정하도록 명령할 수 있다. 기술심사의 영역과 기준은, 크게 사용성 검사, 기술표준 준수 검사, 적합성 검사, 신뢰성 검사로 구분되는데, 세부적인 심사항목 및 심사요소는 그림5와 같다.

4. AI 디지털교과서의 문제점

(1) 이해관계자와의 협의 및 사전 준비 미흡

우선 AI 디지털교과서의 가장 큰 문제점 중 하나는 이해관계자와의 충분한 협의없이 교육부가 밀어붙이고 있다는 점이다. AI 디지털교과서에 대해서 교사, 학부모, 인권단체 등은 다양한 우려를 제기해왔다. 2024년 6월 〈교육부의 2025 AI디지털교과서 도입 유보〉를 요구하는 국회 국민동의청원이 53,884명의 동의를 얻어 국회 교육위원회에 회부되었다. 청원의 주요 내용은 △코로나19 장기화에 따른 원격수업으로 디지털 기기 활용 수업의 명암 확인 △학생 스마트 기기 과의존 및 유해성 문제 △교과서 개발, 검정 일정 지연과 교사들의 반대 여론 등에 대한 우려를 들어 AI디지털교과서 도입이 효과적 교육방식이 맞는지 검증 과정을 거치라는 것이다. 이어 전국교직원노동조합, 참교육을위한전국학부모회, 정치하는엄마들 등 교사단체, 교육단체, 학부모단체를 포함한 127개 단체는 2024년 8월 28일, ‘AI 디지털교과서 중단 공동대책위원회’를 구성하였다. 이들은 교육부가 세계 최초로 국가 단위 AI 디지털교과서를 도입하기 위해 밀어붙이고 있다며, 이는 우리 아이들이 아무도 겪어보지 못한 부작용을 가장 먼저 겪게 될 수 있다는 것과 다름없다고 비판하였다. 공대위는 2024년 9월 6일부터 10월 1일까지 AI 디지털교과서 도입 중단과 정책 타당성 검토를 위한 공론화위원회 구성을 요구하는 서명운동을 전개하였는데, 그 결과 한 달 만에 10만 명이 넘는 국민이 서명에 참여했다고 한다.

(2) AI 디지털교과서의 교육적 효과에 대한 우려

AI 디지털교과서의 교육적 효과가 명확하다면 교사와 학부모가 반대할 이유가 없을 것이다. 그러나 이에 반대하는 단체들은 AI 디지털교과서가 디지털 기기의 사용 시간을 늘려 스마트기기 중독을 야기할 수 있고, 디지털 방식의 학습이 전통적인 방식에 비해 문해력을 저하시킨다는 연구 결과도 있으므로, 오히려 청소년에게 부정적 영향을 미칠 수 있다고 우려한다. '학생 맞춤형 교육'을 강조하지만 실제 AI디지털교과서는 반복적으로 과제를 제시하는 방식으로 구조화되어 있어 '학생 맞춤형 주입식 교육'이 될 가능성도 큰 만큼 학생의 성장발달을 지원하는 적합한 수업 도구인지 보다 면밀한 검토가 필요하다.

하지만 교육부는 AI디지털교과서의 효과성 검증을 요구하는 교육계의 요구는 무시한 채 제도 도입을 밀어붙이고 있다. 시범사업의 형태로 부분적으로 시행하고, 그 결과를 평가하면서 단계적으로 도입하자는 요구에도 교육부는 충분한 사전 준비 단계 없이 2025년 3월 초등학교 3~4학년, 중학교 1학년, 고등학교 1학년 전체 학생을 대상으로 수학, 영어, 정보 교과를, 국어의 경우 특수교육 대상 학생에 한하여 도입 입장을 밝혔다.

(3) AI 디지털교과서 시행의 법적 근거

교육부는 시행령인 〈교과용도서에 관한 규정〉을 개정하여 AI디지털교과서 도입을 위한 법적 근거를 마련하였다고 밝히고 있다. 하지만 국회 입법조사처는 시행령 개정을 통한 AI디지털교과서 도입은 헌법에 명시된 교육제도 법률주의에 위배된다는 보고서를 발간하였다. 야당은 교육계의 반발을 고려할 때 AI디지털교과서를 학교 현장에 이를 단계적으로 도입할 필요가 있다며 법적 지위를 '교과용도서'가 아닌 '교육자료'로 규정하고 학교에서 필요에 따라 사용할 수 있도록 하는 초중등교육법 개정안을 발의하였고, 2024년 12월 26일 국회 본회의를 통과하였다. 하지만 교육부는 AI디지털교과서의 법적 지위가 '교과용 도서'일 때 정책 추진에 의미가 있다며 정부에 '거부권' 행사를 요청하였고, 국회는 2025년 1월 17일, 〈AI디지털교과서 검증 청문회〉를 진행하는 등 논란은 계속되고 있다.

(4) 개인정보보호 측면의 문제점

가. AI 디지털교과서의 개인정보 수집, 이용의 근거

AI 디지털교과서는 학생 개인마다의 특성이나 성취도에 기반하여 학습의 방향을 제시하기 때문에 AI 디지털교과서를 통한 학습 과정에서 다양한 학생 개인정보를 수집할 수밖에 없다. 이러한 개인정보 수집, 이용은 적법한 근거에 따라 이루어져야 한다. 통상 학교에서 인적사항, 학적사항, 출결상황 등 학생 개인정보를 수집하는 것은 초중등교육법(제25조)에 근거한다.

AI 디지털교과서 운용을 위한 개인정보 처리의 주체, 즉 개인정보처리자는 AI 디지털교과서 발행사이다. 그런데 발행사의 경우 교육 관련 법에 개인정보 처리를 위한 법적 근거가 없으므로 정보주체의 동의를 받아 처리할 수밖에 없다. 이때 정보주체의 동의는 자유로운 동의여야 하며(개인정보보호법 시행령 제17조), 따라서 정보주체가 원하지 않을 경우 동의하지 않을 수 있어야 한다. 만일 학교에서 동의가 사실상 강제된다면 적법한 동의라고 할 수 없다. 교육부의 계획처럼 AI 디지털교과서가 '교과서'로 사용될 경우 특정 학교의 전체 학생에 적용될 수밖에 없는데, 이는 자유로운 동의의 원칙과 충돌할 수밖에 없다. 따라서 AI 교과서가 원활하게 도입되기 위해서는 AI 교과서 활용과정에서의 개인정보 처리에 대해 사회적 합의를 통해 법적 근거를 마련하는 것이 우선될 필요가 있다.

나. AI 디지털교과서 활용 과정에서의 개인정보 흐름 및 처리자 명확화 필요

AI 디지털교과서 활용과정에서 학교(교사), 교육행정정보시스템(NEIS), AI 디지털교과서 서비스, 통합 게이트웨이, 학습데이터 허브 등 다양한 기관과 시스템이 관여한다. 이 과정에서 개인정보를 처리하는 처리자(책임주체)가 어떠한 개인정보를 어떠한 적법 근거 하에 처리되는지 명확하지 않다. 서로 다른 기관 간의 개인정보의 제공 흐름, 구체적인 개인정보 항목, 적법 근거 등을 명확히 할 필요가 있다. 그래야 개인정보 침해 예방 및 문제가 발생할 경우 책임 소재를 명확하게 따질 수가 있다.

다. 학생에 대한 과도한 개인정보 수집 및 감시

AI 디지털교과서는 학생의 기본적인 정보 뿐만 아니라, 교과서 시스템에의 로그인/로그아웃 일시, 횟수 및 간격, 학습한 콘텐츠의 수, 학습 시간 및 패턴, 게시글 등록 건수, 평가 결과, 문제 풀이 현황, 과제 현황, 학습 이력 등 매우 세세한 정보를 수집한다. 비단 학교 내에서의 학습 과정 뿐만 아니라, 방과 후의 학습 과정까지 모니터링하고 있는데, 이는 학생의 전체 삶에 대한 모니터링으로서 학생의 프라이버시와 자율성을 심각하게 침해할 우려가 있다. 이러한 감시는 학생들에게 압박감을 줄 수 있어 교육적 측면에서도 바람직하지 않다.

(5) 인공지능 기능 관련 우려

유럽연합 AI Act의 경우 “직장이나 교육 기관에서 감정을 추론하는 AI 시스템”의 활용을 금지하고 있다. 또한, 교육 수준을 평가하거나 시험 중 학생의 금지된 행위를 모니터링하는 시스템 등을 ‘고위험’으로 규정하고 엄격한 의무를 부과하고 있다. 이러한 의무에는 위험관리체계의 마련, 훈련 데이터 평가 및 관리체계, 기술 문서의 작성 및 보관, 이용자(배치자)에 대한 상세한 정보 제공, 도입 이후 모니터링 방안, 로그기록 보관, 보안조치, 접근성 조치 등이 포함된다. 한국의 경우 2024년 12월 26일, AI 디지털교과서를 무효화 한 초중등교육법 개정안과 함께, AI 기본법 제정안도 국회를 통과하였다. 이에 따르면, “유아교육 · 초등교육 및 중등교육에서의 학생을 평가하는 AI”도 고영향 AI로 규정하고 있다. 그러나 유럽연합 AI Act 에 비해 고영향 AI 제공자의 의무는 법에서 구체적으로 규정하고 있지 않다.

AI 디지털교과서의 AI 관련 기능이 갖춰야할 요건에 대한 법적 규범은 없지만, 앞서 살펴보았듯이 〈AI 디지털교과서 개발 가이드라인〉에서 개인정보를 포함한 데이터 관리 정책을 수립할 때 고려사항과 AI 리스크 관리를 위한 원칙을 제시하고 있다. 그리고 검정 과정에서 개발사가 제출한 ‘자체기술검증결과서’를 기초 자료로 활용하여 기술심사를 하도록 하고 있다. 이와 관련하여 몇 가지 문제를 제기할 수 있다.

첫째, 가이드라인이 AI 디지털교과서 개발 및 운영과 관련한 고려사항을 충분히 포함하고 있는지 의문이다. 예를들어, 개발 및 운영 과정과 관련한 기술문서와 로그기록의 보관, 모니터링 등 AI 디지털교과서 운영 과정 상의 요건, 이에 대한 교육부의 감독 메커니즘에 대해서는 다루지 않고 있다.

둘째, 실제 활용될 AI 디지털교과서가 어떠한 AI 기능을 포함하고 있는지, 어떠한 AI 기술이 적용되었는지, 가이드라인의 원칙들을 어떻게 반영하였는지 등이 불투명하다. 예를 들어, AI 튜터 기능에 생성형 AI 기술이 활용되었는지, 편향성이나 환각 문제는 어떻게 통제하고 있는지 등이 문제가 될 수 있다. 또한 학생들의 감정을 파악하는 경우 '감정인식 AI'를 사용하는지 여부, 감정인식 AI의 정확성을 신뢰할 수 있는지 여부 등도 문제가 될 수 있다. 교과서의 내용 자체는 누구나 볼 수 있도록 드러나기 때문에 검정이 제대로 이루어졌는지에 대한 사회적 평가가 가능하지만, AI 디지털교과서의 기술적 측면은 공개된 AI 디지털교과서를 통해서 파악하기 힘들다. 이러한 내용들이 공개되어 있지 않다면 가이드라인에서 제시한 규범이 제대로 준수되고 있는지에 대한 사회적 신뢰를 갖기 힘들다. 이에 대해 교육부에 정보공개를 청구하였으나 교육부는 "가이드라인에 따른 요구사항을 구현하기 위해 각 개발사에서 자율적으로 구성하여 교과서를 개발하는 사항"이라는 원론적인 답변만 보내왔다.

(6) 지방교육재정의 악화

교육부는 AI 디지털교과서를 기존의 서책형교과서와 함께 사용할 것을 강제하고 있다. 따라서 각 학교는 AI 디지털교과서 구매 비용을 추가적으로 지출하여야 한다. 한 번 구매해 1년 이상을 사용하는 서책형 교과서와 달리 AI 디지털교과서는 매달 일정 금액의 구독료를 발행사에 지급하여야 한다. 국회 입법조사처가 초중고교 AI 디지털교과서 구독료를 추계한 결과 2025년에는 연 4,067억원의 예산이 필요할 것이라는 결과가 나왔다. 모든 과목, 모든 학년에 AI 디지털교과서가 도입되는 2028년에는 연 1조 7,343억원이라는 천문학적 예산이 소요된다고 전망하였다. 하지만 교육부는 AI 디지털교과서 도입을 위한 별도의 재원 조달 방안도 마련하지 않고 있어, 시도교육청이 지방교육재정으로 이 모든 부담을 떠안아야 한다. 이토록 막

대한 재정을 AI 디지털교과서 사업에 투입하면, 학교 노후시설 개선, 다문화 교육, 기초학력 증진 등 꼭 필요한 사업 예산이 축소될 수밖에 없다. AI 디지털교과서 중단 공동대책위원회는 효과가 검증되지 않은 AI 디지털교과서에 막대한 혈세를 쏟아붓는 것은 지나친 공교육 재정 낭비일 뿐만 아니라, 에듀테크 산업계만 배 불리는 것이라고 비판한다.

(7) 세계최초 AI 디지털교과서에 대한 국제 사회의 우려

국제교육연맹(Education International, EI)은 2024년 7월 29일부터 8월 2일까지 아르헨티나 부에노스아이레스에서 제10차 총회를 진행하였다. EI는 전 세계 178개국 383개의 교원 단체 3,200만 명의 교사들을 대표하는 세계 최대 교원조직이다. 제10차 총회에서 EI 집행위원회는 「기술, AI 그리고 교직의 미래(Technology, Artificial Intelligence and the Future of the Teaching Profession)」 결의문을 채택하였다. 결의문은 AI와 디지털 기술의 무차별적인 교육 현장 투입에 대한 강한 우려를 표명하고 있다. 결의문 제안자인 EI 집행위원 랜디 와인가튼(Randi Weingarten) 미국 교원노조 AFT 위원장은 “코로나 이후 AI 기술이 교육 분야에 무차별적으로 들어오고 있다.”면서 “이로 인한 ▲학생 간 디지털 격차(digital divide) 문제 ▲불평등의 심화 ▲공교육 재정이 사기업으로 흘러 들어가는 문제”를 지적하고, EI 차원의 더 강력한 국제적 대응을 촉구했다. 비슷한 시기 이주호 교육부 장관이 국회 교육위원회 ‘교육현안 질의’ 과정에서 “국제기구에서도, 심지어 EI 같은 교사들의 단체에서도 ‘(AI 디지털교과서가) 굉장히 효과가 있다’는 의견을 냈다”라고 답변한 내용과는 상반된 결과이다.

EI 소속 교원노조인 전교조 대표이자 집행위원 자격으로 총회에 참석한 전희영 전교조 위원장은 해당 결의문에 적극 지지 입장을 밝혔다. 또한 한국의 AI 디지털교과서 추진사업 과정에서 발생하는 ▲법률상 교과서 개념 충돌 문제 ▲정보 인권 침해 및 개인정보 유출 문제 ▲학생들에게 미칠 영향성 검증 미흡 등을 지적하였다. 이와 함께 EI에 한국 AI 디지털교과서 사업을 검증하고 대응할, 관련 전문가로 구성된 국제조사단 파견을 긴급하게 요청하였고, EI는 긍정적으로 검토하겠다는 입장을 밝혔다.

사회복지 분야 인공지능 현황

1. 도입

다양한 국가에서 복지 분야에 디지털 기술을 접목하고 최근에는 AI를 활용한 복지 서비스로 발전하고 있다. 정부는 AI와 빅데이터를 이용해 복지 사각지대를 발굴하고 사람의 손길 없이도 맞춤형 서비스를 제공하는 등 데이터 복지를 표방한다. 그러나 데이터를 이용한 복지 서비스는 사생활 침해 및 통제, 부당한 결정, 책임 회피 등 우려점이 존재한다. 또한 복지 분야는 무엇보다 사람을 돌보는 것이 가장 중요한데, 이러한 일을 AI에 맡기면서 취약계층이 오히려 보호대상에서 밀려나거나 심적 불안감을 더욱 증폭시키는 등의 위험이 크다.

2. 사회복지 분야의 AI 관련 규범

사회복지 분야 AI는 대부분 공공분야에서 활용되고 있음에도 이에 대해 따로 규정하고 있는 법은 없으며, 각 지방자치단체가 자의적으로 제정한 인공지능 조례 혹은 AI 산업 육성 및 지원 조례안 등이 관련 규범에 가장 가깝다고 볼 수 있다. 그러나 지방자치단체의 조례는 그 적용 범위나 논의 대상이 좁고, AI에 대한 규범이라기보다 AI 개발 혹은 AI 산업 육성에 보다 치우쳐 있다는 점에서 한계가 있다.

3. 취약계층 돌봄 목적의 AI 시스템

사회복지 AI 역시 어디서 어떤 AI를 어느 정도로 도입하고 있는지에 대한 공식적인 자료는 없다. 다만 지방자치단체에서 적극적으로 도입하고 있는 것으로 보이고, 이를 언론홍보자료 등을 통해 짐작할 뿐이다. 사회복지 AI 중 가장 가시적으로 활용되고 있는 취약계층 돌봄 목적 및 사회복지 목적의 AI 시스템 사례를 살펴보고자 한다.

(1) 현황

노인 등 건강 취약계층 돌봄 사업에 AI 시스템이 도입되기 시작했다. 2023년 보건복지부와 한국사회보장정보원, 한국건강증진개발원이 함께 “AI, IoT 기반 어르신 건강관리 사업”을 발표했다. 그동안 방문건강관리사업을 담당해왔던 보건소와 기술 기반 건강증진서비스를 하겠다는 것이 사업의 골자다. 방문에서 비대면으로 전환하는 것이다. 사업대상은 “허약·관리 및 건강관리 행태 개선이 필요한 만 65세 이상 어르신”이며 이들에게 손목 활동량계, 블루투스 체중계, 혈압계, 혈당계, 일반형 AI 스피커, 화면형 AI 스피커 등 디바이스를 제공하고 이를 통해 여러 건강지수들을 모니터링 및 스크리닝하는 것이 사업 내용이다. 아울러 사기업인 네이버도 AI 시스템이 전화를 걸어 건강 상태나 식사 여부 등을 확인하는 ‘클로바 케어콜’ 서비스를 출시하고 확장해나가기 시작했다.

AI, IoT 기반 어르신 건강관리 사업에 참여하게 되면 건강관리 등을 목적으로 손목밴드형 생체정보 수집기, 블루투스 체중계, 혈압계, 혈당계 등을 지급하고 이를 통해 신체정보를 측정한다. 보통 디바이스를 통해 수집된 데이터는 디바이스와 연동된 모바일 앱을 통해 기록된다. “AI, IoT 기반 어르신 건강관리 사업”의 경우 이상 수치가 나타나거나 또는 1주일 이상 측정하지 않을 시 유선으로 확인하도록 되어 있다. 사용자의 정보는 담당 간호사, 영양사, 운동전문가 등에게 공유된다.

【그림1】 'AI스피커(일반형)' 제품 목록

종류	사진	회사/모델명	주요기능	신청 연도
AI스피커 (일반형)		원더풀플랫폼 (다솜이)	말벗기능	2020
			알림기능	
			정보제공	
			음악제공	
			치매예방	
			SOS기능	
		행복커넥트 (NUGU)	말벗기능	2020
			알림기능	
			정보제공	
			음악제공	
			치매예방	
			SOS기능	
		효돌	감성케어	2020
			알림기능	
			정보제공	
			음악제공	
			치매예방	
			SOS기능	
		KT (GIGA-Genie LTE2)	말벗기능	2021
			알림기능	
			정보제공	
			음악제공	
			치매예방	
			SOS기능	

*출처: 2023년 AI, IoT 기반 어르신 건강관리 사업 안내서 67쪽

만일 스마트폰이 없거나 측정용 디바이스 사용이 어려운 경우 AI 스피커(비화면/화면)가 지급된다. AI 시스템이 전화를 걸거나 스피커를 통해 대화를 나누며 사용자의 상태를 체크하는 방식이다. 네이버의 클로바 케어

콜 서비스도 비슷한 메커니즘으로, 사용자와 대화를 나누며 전체적인 건강 상태를 모니터링한다. 이때 사용되는 AI는 이전에 대화를 나누었던 복약지도 내용, 식사 여부 등을 기억해 연속적인 대화를 이어가도록 한다.

(2) 문제점

가. 민감한 데이터 유출 및 프라이버시 침해

AI, IoT 기반 어르신 건강관리 사업은 인적정보 외에도 기기를 통해 혈압, 심박수, 혈당 등의 개인정보가 필수로 저장되고 동의 여부에 따라 약물 복용 여부 및 질환 진단여부, 입원 횟수 등 건강 상태와 관련된 정보도 저장 및 전송된다. 이러한 정보는 개인의 병력 등을 파악할 수 있는 매우 민감한 정보이며 일반 개인정보보다 한층 더 두텁게 보호해야 한다. 그렇기에 목적에 필요한 최소한도로만 수집하고 보호해야 하는데, 취약계층 돌봄이라는 사업의 목적에 맞도록 최소한도로 수집하고 있는지 점검할 필요가 있다.

정부가 2024년 4월에 발표한 초거대 AI 도입 활용 가이드라인을 보면 “의료급여 수급자 관리를 위해 수급자의 연령, 지역, 병명, 직업 등의 데이터를 분석하여 시사점 도출”을 향후 AI 활용 영역으로 제시하고 있다. AI 개발을 위해 민감정보의 활용이 제안되고 있는 것이다. 이러한 생체정보 기반의 프로파일링은 고위험 AI에 해당할 수 있다. 유럽연합 AI법(AI Act)에 따르면, 인권에 중대한 위험을 초래할 수 있는 AI 시스템의 경우 고위험으로 분류되어 엄격한 사업자 의무가 부과되어야 한다. 하지만 아직 우리나라에는 이러한 의무 규정이 없는 상황이다. 또한 알고리즘이 투명하지 않고 과도한 정보를 수집하는 경향이 있으며, 수집된 건강정보가 AI 시스템과 어떻게 연계되고 의료진에게 어떤 방식으로 전달되는지 등 정보의 활용 목적과 처리 과정이 명확하지 않다. 특히 사회복지 관련 AI 규범이 부재한 상태에서 시스템이 도입되고 있어, 고위험 AI의 안전한 활용을 위한 원칙과 절차가 마련되지 않은 것이 문제점으로 지적된다.

나. 정보 동의 방식

서비스 제공 대상인 취약계층, 특히 노인들의 기기 사용에 대한 명확한 이해와 동의 여부가 문제가 될 수 있다. 디지털 취약계층은 기술에 대한 이해도가 낮아 자신이 제공하는 데이터의 의미와 영향을 제대로 파악하지 못한 채 동의할 수 있다. 따라서 약관을 충분히 설명하고 대상자의 이해도를 확인하는 추가적인 절차를 도입하는 등 안전장치가 필요하다.

또한 클로바 케어콜은 건강(민감)정보 수집·이용과 제3자 제공에 동의하지 않으면 서비스 이용이 불가능하다. 구청 등 지방자치단체가 '클로바 케어콜'의 개인정보처리자이고, 클로바 케이콜 사업을 운영하는 네이버는 또 다른 개인정보처리자로서 개인정보를 제공받도록 되어 있다. 그런데 만일 클로바 케어콜 서비스의 운영 주체가 지방자치단체라면, 설사 네이버가 운영 업무에 관여하더라도 이는 '위탁' 처리일 뿐이며, 지방지차단체로부터 위탁받은 목적으로만 개인정보를 처리해야 한다. 그런데 네이버에 대한 개인정보 제공을 위탁이 아니라 제3자 제공으로 규정한 이유는 〈건강(민감)정보 수집이용, 제공 동의서〉에 명시된 바와 같이 "AI학습을 통한 통화품질 및 AI서비스 품질향상"을 위해 네이버가 독자적으로 민감정보를 활용하기 위한 것으로 보인다. 그런데 개인정보보호법은 정보주체의 동의를 받을 때 개인정보의 수집과 제3자 제공 동의를 구분하여 각각 동의를 받아야 하며, 목적 외 이용에 동의하지 않는다는 이유로 재화 또는 서비스의 제공을 거부하여서는 안된다.(22조) 그런데 클로바 케이콜 동의서는 서비스 제공에 필요한 개인정보와 "AI학습을 통한 통화품질 및 AI서비스 품질향상"을 통합하여 동의를 받고 있고 이에 동의하지 않으면 서비스를 제공할 수 없다고 하고 있으므로, 개인정보보호법 위반 소지가 있다. 전문기관에 데이터를 위탁하는 것과 달리, 클로바 케어콜의 제3자 제공은 서비스 제공 외의 목적으로도 정보가 활용될 수 있다는 점에서 근본적으로 다르다. 이러한 강제적 동의 방식은 지양해야 한다.

이렇게 정부기관 외 다른 업체가 참여한 사업인 경우에는 수집된 데이터의 제3자 제공 동의를 통해 업체 역시 데이터를 공유받는 경우가 대부분이다. 이러한 경우 수집된 데이터가 상업적 목적으로 활용되거나 보험사, 제

약회사와 같은 또다른 제3자에게 제공될 경우 보험료가 상승하는 등 문제가 발생할 수도 있다. 아울러 데이터가 해킹되거나 보안을 제대로 하지 않는 경우, 민감한 정보가 유출돼 프라이버시 침해가 발생할 수도 있다.

다. 취약계층 돌봄목적 AI 시스템의 사회적 영향에 대한 고려

만일 AI, IoT 기반 어르신 건강관리 사업에 동의하지 않는다거나 디지털 격차로 인해 이용할 수 없는 경우, 이전과 마찬가지로 방문 돌봄을 받게 된다. 이 경우 따로 편성된 인력에 의해 돌봄을 받는 것이 아니라 AI 사업에 동원되고 남은 인력이 배정될 가능성이 높아 전통적인 방식으로 돌봄을 받고자 하는 사람들이 더 취약해지는 결과를 낳을 수 있다. 또한 돌봄 업무를 하는 공무원들의 입장에서 보면 AI으로 인해 기존의 일이 줄어드는 것이 아니라 다른 일로 대체되고 또 인력을 감축하는 일을 겪을 수 있다.

또한 AI 시스템 도입은 필연적으로 사람 간의 상호소통 기회를 줄인다. 더 많은 기술에 의존하는 건강 관리는 가족과 간병인 간의 상호작용을 감소시킬 수 있다. 측정용 디바이스 외 지급되는 AI스피커(대면/비대면)는 주로 대화를 통해 건강정보를 측정해 돌봄을 수행하는데, 이 과정에서 다양한 상호작용을 경험한다. 그런데 AI 시스템을 이용한 돌봄은 장점도 있는 반면 노인의 정서적 안정감을 약화시키고, 기술로 대체할 수 없는 인간적 유대감을 훼손하는 결과를 가져올 수 있다. 연구 결과에 따르면 노인들은 돌봄인력을 대체하는 로봇에게 상호작용을 시도하고 사람처럼 인식하며 동료애를 느낄 수 있다. 그러나 반대로 같은 연구에서는 우울과 외로움이 심한 경우 로봇과의 관계가 집착으로 변해 스스로를 고립시키는 등 사회적 관계에 부정적 영향을 가질 수도 있다고 짚기도 한다. 따라서 이러한 시스템 도입에는 신중한 접근이 필요하며, 노동자와 사용자에게 미치는 영향에 대한 충분한 검토가 선행되어야 한다.

4. 사회복지 체제

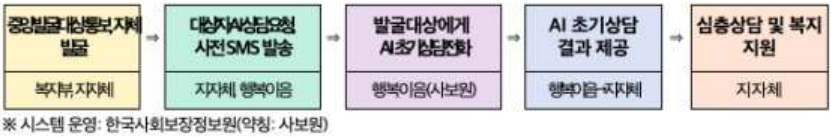
(1) 현황

가. 복지 대상자 파악 및 관리

보건복지부는 위기 의심 가구에 지방자치단체 사회복지공무원이 전화로 초기상담을 진행한 뒤 심층상담과 가구방문을 통해 사회보장급여나 민간 서비스 연계 등 복지서비스를 지원하고 있었다. 그런데 지난 2024년 7월, '제4차 복지사각지대 발굴' 사업의 일환으로 초기상담 단계에 AI 시스템을 도입했다. 이 사업은 같은해 11월 전국으로 확대되었다. 도입 취지는 AI에게 초기상담을 맡기고 공무원은 도움이 필요한 위기가구를 집중 상담해 빠르게 복지위기를 발굴하고 지원한다는 것이다. 시범사업 참여 지자체는 주민에게 상담전화를 하기 전 초기상담을 진행한다는 문자메시지를 시스템으로 발송한다. 이후 AI 시스템이 사전에 파악된 위기가구에 전화해 복지 도움이 필요한 상황인지 파악하는 초기상담을 진행하게 된다. AI가 진행한 초기 상담내용은 시스템을 통해 지자체 공무원에게 자동으로 제공되고 심층상담, 가구방문 등을 위한 자료로 활용된다.

지방자치단체 외 공공기관에서도 앞서 언급한 '클로바 케어콜' 서비스를 이용한 AI 안부전화 서비스가 도입되었다. 사전에 신청한 고령의 연금 수급자와 교직원 부모를 대상으로 AI가 전화를 걸어 건강·식사·수면·운동 등 안부를 묻고 일상적인 대화를 나누는 것으로, 클로바 케어콜 측은 대상자와의 지난 대화를 기억해 개인 맞춤형 대화 서비스를 제공함으로써 섬세한 정서 관리가 가능하다고 밝혔다. 또한 전화 통화 중 건강 관련 위기 상황에 대한 징후를 포착해 119 호출, 병원 연계 등 돌봄 관제 서비스도 함께 제공된다고 한다.

【그림2】 AI 복지상담 운영체계



*출처: 보건복지부 보도자료 24. 7. 21. <인공지능(AI) 초기 복지상담 전화로 위기가구 지원에 나선다>.

나. 부정수급 탐지

국가지원금은 약자를 보호하고 돕기 위해 반드시 필요한 제도지만, 부정수급에 관한 논란이 빠지지 않는다. 이 부정수급 문제를 데이터 추적으로 해결하기 위해 AI를 도입하는 움직임이 숨가쁘다. 한국에서는 기획재정부가 세금계산서 발행 후 취소, 가족 간 거래 등 의심되는 패턴을 알고리즘을 이용, 탐색해 부정수급을 찾아내는 시스템을 도입했고, 한국전력도 취약층의 전기요금 할인 혜택을 부당하게 이용하는 사례를 발견하기 위해 AI 기술을 접목했다. 매달 할인 혜택을 받는 가구의 자격 요건을 수작업으로 판별해왔는데, 이를 AI 애플리케이션에 맡긴 것이다. 또한 사회보장정보원도 사회서비스 전자바우처 부정수급탐지시스템을 도입했다. 대상사업 별로 일괄결제, 중복결제, 심야결제, 연속결제 등 주요 이상결제 유형을 탐지하도록 하고 부정사용이 의심되는 경우 지급을 보류한 뒤 실제 여부를 확인하도록 했다.

(2) 문제점

가. 데이터 편향과 검증 부족

국제앰네스티의 보고서에 따르면 세르비아는 2022년 사회카드 등록 시스템을 도입하며 지원 자격을 결정하는 과정에 데이터를 기반으로 하는 자동화 시스템을 도입했다가 오히려 소외된 계층에 피해를 입히는 결과를 맞이했다. 이 등록 시스템은 기존 정부 데이터베이스에서 소득, 연령, 가구

구성, 건강 상태, 고용 상태 등의 데이터를 가져와 개인의 사회경제적 프로필을 구축하는 방식이었으며 사회복지사가 데이터베이스를 한번 더 분류하도록 해 완전한 자동화는 아니었다. 그러나 이 시스템으로 정부 데이터베이스 기록이 최신 상태가 아닌 경우가 많은 소외계층의 개인은 특히 더 불이익을 받았다. 앰네스티는 “이 시스템은 비공식적 일자리에 있으며 개인 사정이 매우 다양한 사람들의 경제적 현실을 오래된 데이터로 축소했다”고 지적했다.

프랑스 역시 2010년부터 사회복지 시스템에 부정수급을 저지를 가능성이 있는 위험점수를 부여하는 알고리즘을 도입한 바 있다. 이 알고리즘은 개인 데이터와 가족의 데이터를 주기적으로 업데이트해 위험점수를 평가하는데, 위험 점수를 높이는 기준에 저소득, 실업, 불우한 지역에 거주, 소득의 상당 부분을 임대료에 지출, 장애가 있는 상태에서 일하는 등 취약계층을 차별하는 매개변수가 포함돼 있다는 사실이 밝혀졌다.

이 두 가지 사례를 보면 어떤 데이터가 입력되는지가 AI에게 얼마나 중요한 일인지 알 수 있다. AI는 훈련 데이터에 기반해 학습하고 결정을 내리며 이 데이터에는 우리 사회의 편향이 반영될 수 있다. 이미 특정 지역, 소득 수준, 인종 등과 관련된 편향이 데이터에 포함되어 있을 경우 AI는 이러한 편향을 학습해 부당한 판단을 내릴 가능성이 크다. 즉 불평등을 더욱 심화하고 사회적 약자에 불이익을 가할 수 있다. 또한 세르비아의 사례를 보면 알 수 있듯 인간의 개입이 있더라도, AI 도입으로 인해 이미 인력이 감축되고 사람이 AI가 내린 결론에 대한 검증만을 하게 되는 경우 이러한 불공정함을 잡아내기 어렵다. 아울러 성능 검증이 제대로 이루어지지 않은 알고리즘을 사회복지 분야에 도입해 당장의 지원금이 없으면 삶을 이어가기 힘든 사회적 소외계층에 큰 타격을 입힐 수 있는 만큼 제대로 된 모니터링과 감독이 필요하다. 또한 영향을 받는 사람들에게 이러한 AI가 도입되고 있음이 제대로 고지 될 필요가 있다.

나. 결정에 대한 설명 요구와 이의제기 권리

수급자 관련 알고리즘은 인권영향평가와 규제가 필요한 고위험 AI에 해

당한다. AI는 데이터 기반 알고리즘으로 판단을 내리기 때문에 개개인의 특수한 상황과 맥락을 이해하지 못할 수 있다. 세르비아의 사례를 보면, 한 여성이 갑작스럽게 사망한 딸의 장례식 비용으로 약 24만 원을 기부받았는데, 이 기부금이 소득으로 집계되어 수급자 자격을 박탈당했다. 이처럼 AI는 수급자의 일시적 소득 증가나 가족 상황 변화와 같은 복잡한 맥락을 제대로 반영하지 못한다.

이러한 경우 시민은 왜 수급자격이 박탈당했는지에 대해 명확한 설명을 요구해 그 판단이 어떻게 이루어졌는지에 대해 알 권리와 잘못된 판단에 대해 이의를 제기할 권리가 있다. 그러나 AI가 결정을 내린 경우 제대로 된 설명을 듣는 것부터 매우 어려워질 수 있다. 특히 취약계층의 경우 복잡한 행정절차나 기술적 용어를 이해하기 어려울 수 있어, 자신의 권리를 보호하는 데 더 큰 어려움을 겪을 수 있다. 실제 세르비아의 사람들은 사회복지사로부터 수급자격이 박탈당한 이유에 대한 설명을 들을 수 없었고, 이의를 제기하고 이를 해결하기 위해 관공서를 찾아다니며 복잡한 절차 속을 헤매야 했다. 설명을 받을 권리와 이의제기할 권리는 시민이 자신의 권리를 보호할 수 있는 기반이며, 이는 특히 취약계층에게 더욱 중요하다. 또한, 부분자동화 시스템에 대해서도 주의가 필요하다. 미국의 COMPAS 사례에서도 판사가 결과를 검토했음에도 결국 AI의 판단에 의존하게 되었다. 이처럼 자동화 편향과 의존성이 발생할 수 있는 만큼 당사자들이 충분한 이의제기 기회를 가질 수 있도록 하는 것이 필요하다. 한국에서도 복지 대상자 파악 또는 부정수급자 파악을 목적으로 AI 시스템을 사용할 경우, 해외에서 발생한 문제와 유사한 문제가 발생할 수 있기 때문에, 이러한 위험을 파악하고 통제하기 위한 정책과 제도가 마련될 필요가 있다. ◘

한국의 인공지능법 현황

1. 도입

2024년 12월 26일, 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안(이하 AI 기본법)」이 국회를 통과했다. 윤석열 대통령의 갑작스러운 비상계엄 선포와 해제, 국회의 탄핵에 이르기까지 정치적으로 혼란스러운 와중에도 AI 기본법은 '민생법안'으로 치장되어 정당간 이견없이 국회를 통과하였다. 시민사회는 AI의 위험성을 통제할 수 있는 규제가 AI 기본법에 포함되어야 한다고 요구하였지만, 정부, 여야정당, 주요 언론은 현재 한국의 AI 산업 경쟁력을 높이기 위한 지원 정책이 더 필요하다는 것에 합의하였다. 〈첨단 AI의 안전성에 관한 국제 과학 보고서〉 중간보고서에서는 AI 개발을 위한 전 세계적 경쟁이 각 국의 규제 완화를 야기할 위험을 지적하였는데, 한국은 그러한 우려의 대표적인 사례이다. AI 기본법의 국회 통과로 법안을 둘러싼 논의는 일단락될 것으로 보인다. 그러나 AI 기술의 빠른 발전과 도입으로 여러가지 문제가 발생할 가능성이 높기 때문에, 이러한 위험을 보완하기 위한 AI 기본법 개정 논의가 조만간 시작될 수도 있다. 이에 AI 기본법의 주요 내용과 경과, 문제점을 정리하고자 한다.

2. AI 기본법 주요 내용

국회를 통과한 AI 기본법에서 설명하고 있는 법안의 주요 내용은 아래와 같다.

- 가. 인공지능의 건전한 발전과 신뢰 기반 조성에 필요한 사항을 규정함으로써 국민의 권익과 존엄성을 보호하고 국민의 삶의 질 향상과 국가경쟁력을 강화하는 데 이바지함을 목적으로 함(안 제1조).
- 나. 인공지능, 고영향 인공지능, 생성형 인공지능, 인공지능윤리 및 인공지능사업자 등에 대하여 정의함(안 제2조).
- 다. 과학기술정보통신부장관은 3년마다 인공지능기술 및 인공지능산업의 진흥과 국가경쟁력 강화를 위하여 인공지능 기본계획을 국가인공지능위원회의 심의·의결을 거쳐 수립·시행하고, 기본계획에는 인공지능 정책의 기본방향, 전문인력 양성, 신뢰 기반 조성 등에 관한 사항이 포함되어야 함(안 제6조).
- 라. 인공지능산업의 진흥 및 인공지능 신뢰 기반 조성을 위한 주요 정책 등에 관한 사항을 심의·의결하기 위하여 대통령 소속의 국가인공지능위원회를 두고, 국가인공지능위원회는 기본계획의 수립, 인공지능 활용 촉진, 고영향 인공지능 규율 등에 관한 사항을 심의·의결함(안 제7조 및 제8조).
- 마. 과학기술정보통신부장관은 인공지능 관련 정책의 개발과 국제규범 정립·확산을 위하여 인공지능정책센터를 지정할 수 있고, 인공지능안전을 확보하기 위하여 인공지능안전연구소를 운영할 수 있음(안 제11조 및 제12조).
- 바. 정부는 인공지능기술의 개발 활성화와 안전하고 편리한 이용을 위하여 국내·외 동향 및 관련 제도의 조사, 기술의 실용화, 연구개발 등의 사업을 지원할 수 있고, 과학기술정보통신부장관은 인공지능기술과 관련된 표준화를 위하여 표준 제정 등의 사업을 추진할 수 있음(안 제13조 및 제14조).
- 사. 과학기술정보통신부장관은 인공지능기술의 개발 및 인공지능산업의 진흥을 위하여 관련 전문인력을 양성하고, 해외 전문인력의 확보를 위한 각종 시책을 추진할 수 있음(안 제21조).
- 아. 국가 및 지방자치단체는 인공지능산업의 진흥과 인공지능 개발·활용의 경쟁력 강화를 위하여 인공지능 및 인공지능기술의 연구·개발을 수행하는

기업, 기관이나 단체의 기능적 · 물리적 · 지역적 집적화를 추진할 수 있음(안 제23조).

- 자. 정부는 인공지능윤리의 확산을 위하여 안전성 · 신뢰성, 접근성, 사람의 삶과 번영에의 공헌 등의 사항을 포함하는 인공지능 윤리원칙을 제정 · 공표할 수 있고, 과학기술정보통신부장관은 인공지능 윤리원칙의 실천방안을 수립하고 이를 공개 및 홍보 · 교육하여야 함(안 제27조).
- 차. 과학기술정보통신부장관은 법인 · 기관 · 단체 등이 인공지능 안전성 · 신뢰성 확보를 위하여 자율적으로 추진하는 검증 · 인증 활동을 지원하기 위한 사업을 추진할 수 있음(안 제30조).
- 카. 고영향 인공지능 또는 생성형 인공지능을 이용한 제품 또는 서비스를 제공하는 인공지능사업자는 해당 사실을 이용자에게 사전에 고지하여야 하며, 생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 경우 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 하고, 인공지능시스템을 이용하여 실제와 구분하기 어려운 가상의 결과물을 제공하는 경우 그 사실을 이용자가 명확하게 알 수 있도록 고지 또는 표시하여야 함(안 제31조).
- 타. 인공지능사업자는 학습에 사용된 누적 연산량이 대통령령으로 정하는 기준 이상인 인공지능시스템의 안전성을 확보하기 위하여 위험 식별 · 평가 · 완화 등의 사항을 이행하여야 함(안 제32조).
- 파. 인공지능사업자는 고영향 인공지능 또는 이를 이용한 제품 · 서비스를 제공하는 경우 안전성 · 신뢰성을 확보하기 위한 조치를 이행하여야 함(안 제34조).
- 하. 과학기술정보통신부장관은 이 법에 위반되는 사항을 발견하거나 혐의가 있음을 알게 된 경우 인공지능사업자에 대하여 자료를 제출하게 하거나 소속 공무원으로 하여금 필요한 조사를 하게 할 수 있고, 위반 사실이 있다고 인정되면 위반행위의 중지나 시정을 위하여 필요한 조치를 명할 수 있음(안 제40조).

3. AI 기본법에 대한 시민사회의 대응

21대 국회(2020.5.30~2024.5.29)부터 이상민, 정필모, 윤영찬 의원등 여러 의원들이 인공지능 법안을 발의하였다. 2023년 2월 14일, 21대 국회 과방위 법안심사소위는 인공지능 법안을 갑작스럽게 통과시켰다. 인공지능 법안은 제정법이었음에도 불구하고 인공지능 위험에 대하여 공청회나 사회적인 논의가 충분히 이루어지지 않았다. 심사소위 회의록을 보아도 주요 쟁점에 대한 실질적 토론없이 졸속으로 통과되었음을 확인할 수 있다. 더구나 이 법안은 인공지능에 대한 기본법을 표방하면서도 '우선허용·사후규제 원칙'을 포함하는 등 산업육성에만 초점을 맞추고 있고, 인공지능의 위험성을 통제하고 사업자에 책임을 부과할 수 있는 내용을 포함하고 있지 않았다. 시민사회는 법안심사소위 통과안을 비판하며 과방위에서 처리하지 않을 것을 요구하였다.

과학기술정보통신부는 국회 인공지능 법안의 내용을 실질적으로 주도하고 있었다.

2023년 4월 26일, 과기정통부는 시민사회가 국회 과방위에 제출한 의견서에 대한 답변서를 보내왔으며, 시민사회는 이에 대한 반박 의견서를 다시 제출하였다.

인공지능 법안에 대한 국가인권위원회의 의견도 국회 논의 과정에서 전혀 반영되지 않았다. 국가인권위원회는 2023년 8월 21일, 국회의장에게 인공지능 법안에서 '우선허용·사후규제' 원칙을 삭제하고, 인권영향평가 등 인권침해·차별 문제를 예방·규제할 수 있는 규정을 마련할 것을 권고하였다.

전 세계적으로 AI 위험성을 통제해야 한다는 문제의식이 높아지고 있었고, 각국은 인공지능을 규제하기 위한 행정적, 입법적 규율체계를 마련하기 시작했다. 2023년 10월 30일, 바이든 행정부는 '안전하고 보안이 보장되며 신뢰할 수 있는 AI의 개발과 사용에 대한 행정명령'을 발표하였다. 유럽연합은 2023년 12월 8일, 세계 최초로 고위험 인공지능을 포괄적으로

규제하는 ‘인공지능법안(AI Act)’에 대해 합의하였다.

시민사회는 한국에서도 인공지능의 위험성을 규율하기 위한 법이 필요하다는 것에 동의하였지만, 국회에 계류되어 있는 인공지능 법안은 산업육성에 치우쳐있기 때문에 이를 폐기하고 22대 국회에서 실질적으로 논의할 것을 촉구하였다. 과기정통부는 시민사회의 의견을 수렴하여 인공지능 법안을 수정했다고 밝혔지만, ‘우선허용·사후규제’ 조항을 삭제했을 뿐 사업자의 책임을 강화하라는 요구를 반영한 것은 아니었다. 오히려 과기정통부는 기업들과의 비공개 간담회에서 “시민단체 반대 의견을 최소한으로 수렴했다”는 입장을 밝혔다. 과기정통부는 AI 정상회의 개최를 핑계로 인공지능 법안의 통과를 압박하였지만, 결국 21대 국회에서 인공지능 법안은 통과되지 못하고 임기만료로 폐기되었다.

22대 국회가 개원하자마자 인공지능 법안의 발의가 이어졌다. 시민사회는 인공지능법의 제정 방향을 제시하는 의견서를 국회에 제출하고 토론회를 개최하였다. 또한, 인공지능이 전 사회적으로 영향을 미치고 있는 만큼, ‘국회 인공지능 특별위원회’를 구성하여 범상임위적으로 심사할 것을 촉구하였다.

정점식 의원안은 국민의힘 의원 전원이 서명한 사실상 정부여당안이었다. 과기정통부가 21대 국회 말미에 공언한대로 ‘우선허용·사후규제’ 조항은 포함하고 있지 않았지만, 그 외의 내용은 21대 국회 병합안(과방위 심사소위 통과안)과 크게 다르지 않았다. 시민사회 의견서에서 지적하고 있는 바와 같이, ▲금지해야 할 인공지능에 대해 규정하지 않고 있고, ▲고위험 인공지능 분야를 협소하게 규정하고 있을 뿐더러, ▲고위험 인공지능 사업자(개발자 및 운영자)의 의무에 대한 규정이 미흡할 뿐만 아니라 처벌 조항이 없어 실효성에 한계가 있었다. 또한 ▲인공지능의 영향을 받는 사람의 권리 및 구제 조항도 없었고, ▲학습 데이터 공개 등 범용 인공지능 사업자의 의무 조항도 배제되었다. 시민사회는 산업육성에 편향된 과기정통부가 인공지능의 주무부처로서 적합한지에 대해 의문을 제기하였다.

2024년 9월 24일, 22대 국회의 과방위는 인공지능 법안에 대한 상임위

차원의 공청회를 개최하였다. 시민사회는 국회에 발의된 인공지능 법안의 주요 이슈별로 상세한 의견서를 작성하여 제출하였다. 그러나 국회 과방위는 두 번의 심사소위만에 인공지능 법안(AI 기본법)을 통과시켰다. 22대 국회가 개원하자마자 규제완화에 초점을 둔, 비슷한 내용의 인공지능 법안이 앞다투어 발의되었고, 그중 19개 법안이 병합되어 처리된 것이다.

정부여당안에 비해 일부 진전된 내용이 있음에도 불구하고, 과방위 통과안 역시 21대 국회 병합안과 크게 다르지 않았다. 여전히 ▲금지해야 할 인공지능에 대한 규정이 없고, ▲고위험 인공지능 사업자의 책무 위반에 대한 처벌 규정도 미흡하다. 책무 위반에 대한 과기정통부의 시정명령을 이행하지 않을 때에야 과태료를 부과하고 있을 뿐이다. ▲인공지능의 영향받는자의 정의가 포함된 것은 다행이지만, 정작 영향받는 자의 권리 및 구제에 대해서 규정하고 있지 않다. ▲학습 데이터 공개 등 범용 인공지능 사업자의 의무 조항 역시 포함되어 있지 않다. 더구나 독소조항을 새롭게 포함하였는데, ▲국방 또는 국가안보 목적의 인공지능을 이 법의 적용에서 배제한 것이다.

시민사회는 AI로 야기될 위험을 예방하고 문제발생 시 권리 구제절차 등을 규정하여 국민의 안전과 인권 및 민주주의를 보장할 수 있는 '시민사회 인공지능 기본법안'을 별도로 준비하고 제안하였으나 발의가 이루어지지 못했다. 인공지능에 대한 제정법이자 기본법은 인공지능 위험 규제와 권리구제에 대한 시민사회 제안도 함께 검토하고 충분한 사회적인 공론화를 거쳤어야 했다. 시민사회는 향후 AI 기본법의 지속적인 개선을 위한 디딤돌로 삼기 위해, 준비한 법안을 2024년 12월 3일 더불어민주당 김남근 의원 소개로 입법 청원하였다.

4. AI 기본법의 문제점

시민사회는 국회를 통과한 AI 기본법이 산업육성에 치중해있고 인권을 외면하고 있다고 비판한다. 그럼에도 불구하고, 애초에 발의된 정부여당안보다는 일부 진전된 부분도 있는데, 다음과 같은 점들이다.

▲일부 발의안에 포함되었던 ‘우선허용·사후규제 원칙’ 조항이 최종 AI 기본법에 포함되지 않았다. ▲국제적인 정합성을 위해 OECD 정의를 차용하여 ‘인공지능 시스템’을 정의하였고, ‘영향받는 자’의 개념도 정의에 포함되었다. ▲고영향[고위험] 인공지능의 영역에 “유아교육 · 초등교육 및 중등교육에서의 학생 평가”가 포함되었다. ▲기본원칙(제3조)에서나마 “인공지능이 사람의 생명 · 신체의 안전 및 기본권에 중대한 영향을 미치는 경우”, 그 결과의 이유 및 원리 등에 대하여 기술적 · 합리적으로 가능한 범위에서 명확하고 의미 있는 설명을 제공받을 수 있는 ‘영향받는 자의 권리’를 규정하였다. ▲영향받는 자를 포함한 이해당사자가 신고 및 민원을 제기할 수 있도록 하고, 사실조사를 통해 과기정통부 장관이 법 위반사항을 조사하고 시정명령을 할 수 있는 권한을 신설하였다. ▲딥페이크 등 생성형 AI의 고지 및 표시 의무 등 투명성 의무가 강화되었다. ▲미약하지만, 고영향 AI 사업자에 대해 영향평가를 위해 “노력할 의무”를 부과하였다.

그러나 대체적으로는 시민사회가 요구해왔던 핵심적인 규제를 배제하였는데, 주요 문제점은 다음과 같다.

첫째, 금지해야 할 인공지능에 대한 규정이 없다. 정부는 유럽연합 외에 아직 인공지능 금지에 대한 규정을 두고 있는 나라가 없고, 산업 발전을 저해할 수 있다는 이유로 이에 반대하였다.

둘째, 고영향 인공지능의 범위가 EU AI Act에 비해서 여전히 협소하다. 예를 들어, △범죄 수사나 체포 업무 외의 영역에서 생체인식정보를 분석 · 활용하는 데 사용되는 인공지능, △수사 및 기소 등 기본권을 침해할 수 있는 국가기관의 권한 행사에 이용되는 인공지능, △사람의 감정인식에 사용되는 인공지능, △사법부 또는 행정부에서 판결, 결정, 심판 등의 업무에 사용되는 인공지능, △정보통신망의 운영에 사용되는 인공지능, △선거 및 투표행위, 투표결과에 영향을 미치기 위하여 사용되는 인공지능, △제품 안전에 영향을 미칠 수 있는 인공지능 등이 포함되는지 명확하지 않다

셋째, 고영향 인공지능 사업자의 책무 위반에 대한 처벌 규정도 미흡하

다. 사업자의 책무도 구체적이지 않을 뿐더러 책무 위반 자체에 대해서 처벌하는 것이 아니라, 단지 책무 위반에 대한 과기정통부의 시정명령을 이행하지 않을 때에야 과태료를 부과하고 있을 뿐이다.

넷째, 인공지능에 영향받는 자의 정의가 포함된 것은 다행이지만, 정작 영향받는 자의 권리 및 구제에 대한 조항은 두고 있지 않다.

다섯째, 학습 데이터 공개 등 범용 인공지능 사업자의 의무 조항 역시 포함되어 있지 않다. 이는 개인정보 보호 및 저작권 보호와 충돌할 수 있다.

여섯째, 인공지능 인권영향평가도 "노력할 의무"만을 부여하고 있어 실효성이 있을지 의문이다. 이와 관련하여 2024년 5월에 국가인권위원회가 '인공지능 인권영향평가 도구'를 발표한 것은 긍정적인 상황이지만, 정작 주무부처인 과기정통부는 이를 무시하고 있다.

일곱째, 독소조항을 새롭게 포함하였는데, 국방 또는 국가안보 목적의 인공지능을 이 법의 적용에서 배제한 것이다. 정부와 국회는 이를 별도의 법률로 규제해야 한다고 하지만, 이러한 법률이 언제 만들어질지는 미지수이다. 국방 또는 국가안보 목적의 인공지능 역시 AI 기본법의 적용을 받도록 하되, 필요하다면 국방 또는 국가안보 목적의 AI에게만 적용되는 특별법을 제정할 수도 있었다. 이 조항이 포함된 것은 국가정보원의 요구 때문인데, 인권 침해와 정치 개입의 오랜 역사를 가지고 있는 국가정보원과 최근 계엄령에 동원되었던 군이 AI 기술을 남용할 경우 어떻게 통제할 수 있을지 우려된다.