

오병일 (디지털정의네트워크 대표)

디지털정의네트워크의 오병일입니다. 지난 12월 16일, 대통령직속 국가인공지능전략위원회는 「대한민국 인공지능 행동계획」(안)을 공개했습니다. 그러나 이 행동계획은 한 국가의 인공지능 전략이라고 보기에는 기본 관점이 지나치게 편향되어 있고, 내용 역시 매우 부실합니다. 이에 대해 시민사회의 입장에서 몇 가지 총론적인 문제를 지적하고자 합니다.

첫째, 이번 AI 행동계획은 마치 한 기업의 AI 사업 전략을 보는 듯, AI의 개발과 활용에만 과도하게 집중되어 있습니다. 심지어 'AI 기본사회' 분야마저도 취약계층도 어떻게 AI를 활용할 수 있을지, 사회적 문제 해결을 위해서 어떻게 AI를 활용할지에 초점을 두고 있습니다. AI 산업 육성과 보편적 접근이 중요하지 않다는 것이 아닙니다. 문제는 AI의 위험성으로부터 시민의 안전과 인권을 보호해야 할 국가의 책무가 거의 보이지 않는다는 점입니다. 국가 정책이라면 아래서는 안 됩니다.

자동차에 비유를 해보자면, 자동차 기술의 개발과 산업 육성도 필요하지만, 자동차를 시장에 출시하기 전에 당연히 안전성 검사를 하도록 하는 체제가 마련되어야 합니다. 신호등 체계도 만들고 과속을 하지 않도록 규제할 필요도 있습니다. 그러나 행동계획은 일단은 자동차를 도입하겠다고 합니다. 이번 행동계획은 '원칙적으로 허용하고 예외적으로 금지하는 네거티브 규제 체계'를 도입하겠다고 밝히고 있는데요, 이는 시민의 안전과 인권을 위협할 수 있는 AI 시스템이라도 일단 도입하겠다는 것과 다르지 않으며, 인공지능 기본법 제정 과정에서 시민사회의 반대로 삭제되었던 '우선도입-사후규제' 원칙을 사실상 되살리는 것입니다.

또한, AI 행동 계획은 마치 AI가 우리 사회의 모든 문제를 해결할 수 있는 것처럼 기술중심주의적 관점에 깊이 매몰되어 있습니다. 이는 자동차가 편리하다는 이유로 초등학교부터 운전 실습을 시켜야 한다고 말하는 것과 크게 다르지 않습니다. 그러나 요새 대학가도 생성형 AI 때문에 난리인 것을 알고 계실 것입니다. AI가 인간의 정신과 역량에 어떠한 영향을 미칠지, 조직 차원에서 AI에 과의존했을 때 발생할 수 있는 문제는 무엇인지에 대해 아직 충분히 밝혀지지 않았습니다. 알고리즘의 오류, 편향, 불투명성은 예외적인 인공지능의 악용 문제가 아니라, 현재 AI 기술에 구조적으로 내재된 문제입니다. 그럼에도 불구하고 이러한 위험을 통제하고 AI 사업자의 책임을 강화하기 위한 장치는 행동계획 전반에서 거의 찾아볼 수 없습니다.

둘째는 AI의 영향을 받는 사람들의 권리의 보호와 피해 구제 방안이 부재하다는 것입니다. AI는 노동자, 소비자, 여성, 장애인을 비롯한 시민들의 삶과 권리에 중대한 영향을 미치게 될 것입니다. 그러나 행동계획에는 AI가 노동자의 근로조건과 노동 환경, 고용 과정에 미치는 영향에 어떻게 대응할 것인지에 대한 종합적인 대책이 없습니다. 소비자 권리에 대한 고려도 부족하고, 성인지적 관점 역시 거의 보이지 않습니다. 지능화된 기술이 권력기관의 감시 수단으로 사용될 가능성에 대해 시민을 어떻게 보호할 것인지에 대한 대책도 찾아보기 어렵습니다.

AI 행동계획은 체계적인 AI 윤리 기준을 마련하겠다고 하지만, 지금 AI 행동계획에 필요한 것은 유엔과 국가인권위원회가 일관되게 강조해왔던 '인권기반접근'입니다.

셋째, AI 산업 육성을 이유로 개인정보 보호의 기본 원칙마저 흔들고 있습니다. 정부는 가명처리조차 하지 않은 개인정보를 AI 학습에 활용할 수 있도록 개인정보 보호법 개정을 추진하겠다고 밝히고 있습니다. 대규모 개인정보 유출 사고가 반복되고 있음에도, 개인정보를 애초의 수집 목적과 무관하게 활용하겠다는 것입니다. 단지 AI 산업을 육성하겠다는 이유만으로, 오랜 시간에 걸쳐 형성된 개인정보 보호의 기본 원칙을 가볍게 뒤집어서는 안 됩니다.

넷째, 이번 AI 행동계획은 내용뿐만 아니라 수립 과정 자체에서도 민주적 거버넌스가 결여되어 있습니다.

행동계획에 이처럼 많은 공백과 문제점이 존재하는 이유는 분명합니다. 시민사회와 영향을 받는 당사자들이 AI 정책 결정 과정에서 철저히 배제되었기 때문입니다. 편향된 데이터로 학습된 AI가 편향된 결정을 내리듯, 편향적으로 구성된 국가 AI 전략위원회에서 공정하고 균형 잡힌 행동계획이 나오기를 기대하는 것은 현실적으로 어렵습니다.

시민사회와 영향을 받는 사람들을 배제한 AI 행동계획은 결코 사회적 공감대를 형성할 수 없습니다. 정부가 계속해서 산업 발전만을 위한 가속 폐달을 밟는다면, 한국의 AI 정책은 그 이행 과정에서 필연적으로 커다란 사회적 반발과 저항에 직면하게 될 것임을 분명히 경고합니다. 정부는 국가 AI 전략위원회를 전면적으로 재구성하고 보다 충분한 숙고와 사회적 공론화를 거쳐서 AI 행동계획을 다시 수립할 것을 촉구합니다.

이지은 (참여연대 공익법센터 간사)

규제혁신이라는 이름으로 프라이버시권, 개인정보자기결정권 등 기본권 침해

- 지난 12월 16일 국가AI전략위원회가 공개한 대한민국 인공지능 행동계획안의 정보인권 관점에서 굉장히 문제가 심각합니다.
- 규제개혁 내지 규제혁신이라는 이름으로 몇몇 특정 AI기업의 AI기술발전을 위해 국민의 개인정보 및 국가기관과 공공기관 등이 보유한 정보들을 원본 그대로 활용할 수 있도록 법제 개편을 기정사실화하고 국가 전 영역에 AI를 도입, 활용하겠다는 것을 전제로 합니다. 특히 규제의 패러다임을 사전통제에서 원칙적으로 허용하고 예외적으로 금지하는 ‘네거티브 규제’ 체계를 AI 분야에 우선 도입하겠다고 선언하고 있는데 이는 검증되지도 않은 AI를 전국가적 차원에서 전국민을 대상으로 ‘실험’을 하겠다는 위험천만한 선언이 아닐 수 없습니다. 현재까지 AI 기술은 알고리즘의 오류 가능성, 불투명성, 개인정보침해, 편향 등의 문제가 드러나고 있고 아직 이에 대해 명확한 해결책도 없는 상황입니다. 또한 AI 시스템이 인간에 미치는 영향, 효과, 안전성은 검증되지 않았습니다. 이와 같은 상황에서 국가의 중요한 역할, 기능 등을 AI로 전향하는 것은 매우 위험하다고 할 것입니다. 마치 온국민을 AI기술의 시험대상으로 삼는 것과 다르지 않습니다.
- 특히 이번 AI행동계획안에서 전제로 하고 있는 AI산업 발전을 위해 사실상 개인정보 원본을 마치 공유재처럼 마음대로 쓸 수 있도록 법개정을 하겠다는 것은 규제 합리화라는 명분이지만 사실상 산업개발을 위해 개인정보자기결정권을 침해하는 것입니다. 헌법으로 보호받는 개인정보자기결정권을 침해하는 것이며, 개인정보보호법의 기본원칙을 무너뜨리는 것입니다. 개인정보는 인간의 존엄성이나 인격의 내적 핵심, 내밀한 사적 영역에 근접하는 민감한 개인정보들을 포함하며, 한 인간의 프라이버시를 구성하는 주요 요소이고, 인간의 정체성과도 밀접한 관련이 있습니다. 인공지능 산업 발전을 위해 개인의 인격권을 포기하라는 것은 60, 70년대 산업성장을 위해 인권을 희생시키던 권위주의 정부에서나 가능한 발상입니다.
- 국가 AI행동계획 중 정보 인권 관련 항목에 대해 구체적으로 보면,

- 04. 국가데이터 통합플랫폼을 통한 민간·공공 데이터 자산화
- 05. AI 데이터 공유 생태계 활성
- 10. AI기반의 사이버 보안 체계 구축, AI 보안 기술 경쟁력 강화 및 보안 전문 인재 양성
- 31. 개인정보·미개방 산업 데이터 활용 규제 과감한 정비
- 32. AI 학습·평가 목적의 저작물 활용 및 유통 생태계 활성화
- 33. 규제 합리화를 통해 혁신을 촉진하는 AI 산업생태계 조성
- 34. 국가 AI 대전환을 뒷받침하는 AI 법제 환경 조성

- 액션플랜4,5: ‘국가데이터 통합플랫폼’ 구축은 사전 위험 대비 분석, 방안 마련 후 신중히 추진해야 함
- 공공데이터는 정부, 공공기관 등이 생산 및 수집, 보유한 국민의 정보임. 이는 목적, 접근권한, 보안수준 등에서 충위가 다양함. 이를 민간데이터화 통합, 공유하겠다는 것은 국가 기밀 보안 차원에서도 위험함. 목적으로 상이하고 접근 권한, 기밀성, 보안등급 등도 다양할텐데 사전 위험 대비 및 영향 평가, 위험 관리 방안 등 마련 후 신중히 추진해야 함
- 액션플랜31: 개인정보보호를 위한 방안은 2026년 1분기 내 수립해야 할 가장 시급한 과제임. : 인공지능 학습을 위해 개인정보 원본 활용이 가능하도록 하는 방향의 법제 추진은 중단할 것, 오히려 인공지능 시대 정보주체 모르게 수집, 프로파일링 등 위험성에 대해 정보주체의 권리를 실질적으로 강화하는 방향으로 개인정보보호법 개정이 이루어져야 함
- 액션플랜 34: AI 법제 환경 조성, AI기본법 하위법령 및 가이드라인 보완은 이용자, 이용사업자, 영향받는자의 정의를 명확히 하고 관련 의무와 권리를 명확히 명시해야 함.
- 인공지능을 “업무 목적으로” 이용하여 영향받는자에 대하여 직접적인 영향을 미치는 병원, 채용회사, 금융기관 등 사업자를 ‘이용사업자’에 포함하여 합당한 책무를 지도록 시행령에 명확히 규정할 것을 포함하여야 할 것임.
- 또한 인공지능법 제3조 제2항에서 영향받는자의 설명요구권을 규정하면서 이에 대한 구체적인 조항을 규정하지 않고 있음. 법에서 권리는 명확하게 규정하고 있으므로, 시행령을 통해서라도 이를 이행할 수 있는 방안을 구체화하는 방향으로 보완되어야 할 것임.

장여경 (정보인권연구소 상임이사)

[공공AX 분야 의견] 공공분야 인공지능 전환의 공공성과 책무성 강화해야

인공지능 행동계획(안)은 전략분야 중 하나로 “공공분야 AX(인공지능 전환)”을 제시하고 있습니다. 그러나 국민의 기본권에 직접적인 영향을 미칠 것으로 예상되는 공공분야 AI의 위험성을 통제하고 책임성을 강화하기 위한 대책은 제대로 다뤄지고 있지 않습니다.

때마침 1월 5일 감사원이 공개한 <인공지능 대비실태> 주요 감사 결과는 충격적이었습니다. 감사결과 공공분야 인공지능의 책무성 수준이 심각한 수준인 것으로 드러났기 때문입니다.

현재 공공 AI와 인공지능기본법은 영향받는 자에 대한 피해 대책이 미흡하며, 고영향 인공지능의 검·인증을 담당할 기관 관련 기준조차 정비되어 있지 않았습니다.

특히 공공분야 지능형 CCTV의 성능과 신뢰성 미달 실태는 충격적이었습니다. 지능형 CCTV의 탐지정확도가 낮아 실제 아무런 문제가 없는 개인의 행동을 이상징후로 판단하여 시민에 대한 잘못된 평가나 예측을 하고 공권력을 행사하여 기본권을 심대하게 침해할 우려가 있었습니다. 서울시 11개 자치구의 지능형 CCTV의 침입 등 탐지 정확도는 12~52% 수준에 그쳤고, 한국도로공사의 포트홀 감지시스템의 탐지 정확도 역시 14.7% 수준이었습니다. 근로복지공단의 '직업 복귀 통합시스템'의 추천 적중률도 53.7%로 개발목표에 미달했습니다.

이는 몇년 동안 경쟁적으로 도입된 공공 AI가 기본적으로 갖춰야 할 공공성과 책무성 측면에서 중대한 결함이 있었다는 뜻입니다. 시민들은 공공분야의 책임성과 첨단산업의 전문성을 믿었지만, 사실 공공분야와 산업체가 부패하게 결탁한 것이나 다툼이 없었고 그로 인한 피해가 그동안 시민들에게 고스란히 돌아갔을지 모릅니다.

공공 AI 전환은 반드시 인권에 기초하고 시민의 안전과 신뢰가 보장될 수 있는 방식으로 진행되어야 합니다. 특히 부실한 것으로 드러난 공공 AI에 대한 신뢰성을 회복하고 시민의 인권을 보호할 수 있는 장치를 갖추어야 할 것입니다.

하지만 현재의 국가인공지능행동계획안은, 공공 AI의 신뢰성과 인권 보장 보다 전 공공분야의 인공지능 도입에 급급한 모양입니다. 부실한 공공 AI 실태를 그대로 방치하고 AI-Native한 공무원 협업체계가 구축된다면, 그로 인한 오류는 누가 책임을 지는 것입니까? 공공 AI의 부실햄이 민간클라우드 연계와 감사 면책으로 악화되어 대규모 사고로 이어졌을 때, 그로 인한 시민 피해에 대책이 마련되어 있습니까?

따라서 행동계획에는 공공분야에 설명가능하지 않은 AI, 위험성이 검증되지 않은 AI는 도입하지 않겠다는 명확한 입장 표명이 있어야 합니다. 공공분야 AI의 위험성을 통제하고 사업자 책임성을 강화할 장치를 갖추기 위해서는 <공공분야 AI 윤리기준>으로는 부족하며, 공공 AI가 시민의 권리를 보장할 책임을 질 수 있도록 <법제도>가 마련되어야 합니다.

법제도 측면에서 공공 AI의 영향을 받는 시민들의 권리를 보호하고 구제할 수 있는 방안이 부재한 것은 큰 문제입니다. 경찰, 출입국, 사회복지 등의 분야에서 공공 AI는 시민들을 감시, 차별, 배제하기 위한 목적으로도 사용될 수도 있습니다. 초안에서 강조하는 바와 같이 공공 AI를 위해 "모든 관련 데이터는 표준화되고, 통합되며, 현재화"된다면, 시민에 대한 감시 역시 더욱 용이해질 것이고, 시민들은 자신이 AI의 대상이 되었다는 사실 자체를 인지하지 못할 수도 있습니다. 따라서 공공 AI에 의해 시민의 권리가 어떻게 영향을 받고, 시민의 권리를 보호하고 구제할 수 있는 종합적인 방안이 행동계획으로 포함되어야 할 것입니다.

국가정보원의 역할 확대 및 시민 감시 우려의 문제 또한 짚지 않을 수 없습니다. 국정원은 정보기관으로서의 역할에 맞지 않게 공공 정보통신망에 대한 사이버보안 권한을 갖고 있으며, 행동계획안은 개인정보 보안이나 사이버안보 분야의 AI와 관련한 역할을 국정원에 부여하고 있습니다.

그러나 예산 통제 및 국회 감시를 제대로 받지 않는 초법적인 기관으로서 국정원에 이와 같은 권한을 부여하는 것은 매우 위험하며, 최소한 다른 정부부처 수준의 예산 및 국회 통제를 받도록 해야 합니다.

무엇보다 공공 AI에서 민주적이고 효과적인 거버넌스 체제가 부재하다는 점은 매우 심각한 문제입니다. 공공 AI 정책에서 노동자·시민 등 영향받는 사람과 그 대표자가 참여할 수 있어야 하고, AI에 대하여 규제권한을 가지고 있는 규제기관들을 포함하는 효과적인 감독 거버넌스 체제가 수립되어야 합니다. 이는 AI시대 민주주의의 과제이기 때문입니다.

김은진 (민주사회를 위한변호사모임 디지털정보위원회)

AI기본사회 분야 발언

안녕하세요 민변 디지털정보위원회 김은진 변호사입니다.

국가인공지능전략위원회의 2025. 12. 『대한민국 인공지능 행동계획』(안) 중 '글로벌 AI 기본 사회 기여'부분의 내용과 한계, 수정 방향에 대하여 크게 세가지 말씀드리겠습니다.

첫째, 『대한민국 인공지능 행동계획』(안)은 'AI기본사회' 실현과 관련하여, 모든 국민이 인공지능의 혜택을 포용적으로 누리는 'AI 기본 사회' 실현하기 위해 「모두를 위한 AI기본 사회 추진계획」 수립을 정하고 있습니다. 이러한 추진계획은 AI 기술 발전의 혜택을 모두에게 분배하겠다는 방향성 자체는 긍정적이지만, 인권 기반 관점에서 통제·참여·구제 장치가 매우 추상적이라는 문제가 있습니다. 이에 「모두를 위한 AI기본 사회 추진계획」에 국제인권규범 및 헌법상 기본권을 기준으로 한 '인권 기반 원칙(예: 차별금지, 데이터 권리, 절차적 권리)'를 명시하고, 고위험 AI 규제·인권영향평가·취약계층 보호에 관한 내용을 추가 보완할 필요가 있습니다.

둘째, 위 행동계획(안)은 기존 2020년 「인공지능 윤리기준」만으로는 새로운 AI 기술의 위험 요인과 기술 환경을 충분히 반영하기 어렵기에, 과기정통부를 중심으로 AI 개발 사업자 및 이용 사업자, 이용 사업자, 이용자, 영향을 받는 자 등 사회 각계의 의견을 수렴하는 'AI 공론장' 거버넌스 등을 구축한다고 정하고 있습니다. 하지만 '국민 모두의 공론장'이라는 수사적 표현 이면의 한계를 지적할 수밖에 없습니다. 이러한 시스템의 실제 설계·운영 권한은 과기정통부, 국가AI전략위, 산업계 등에게 집중될 가능성이 높고, 구조적 약자인 시민사회 단체와 인권 단체는 들려리로 소외될 위험이 큽니다. 또 제안된 구상은 참여의 깊이(의사결정 권한·비토권·동의권 등)를 전혀 구체화하지 않고 있어 문제가 있습니다. 따라서 공론장 참여 구조, 의제 설정, 권한 배분, 책임·구제 장치에 AI 영향을 받는 집단과 인권·시민단체의 실질적 참여권을 구체화하여야 합니다. 공론화 결과를 정책에 우선 반영한다는 추상적인 내용이 아니라, "AI 활용 전·중·후 단계의 인권영향평가", "독립적 감독", "정보 접근권", "이의제기·중단 요구·손해배상" 등 효과적 구제수단 등 정책 반영 내용을 구체화할 필요가 있습니다.

셋째, 위 행동계획(안)은 AI교육 정책의 기획-연구-평가-백서발간-정책환류를 통합적으로 시행할 수 있는 민관 협력형 AI교육정책 연구체계를 구축하고, 지역·계층 간 AI 교육격차 해소를 위한 지원체계 강화, 교육 AI서비스 품질 및 안정성을 위한 거버넌스를 구축한다고 정하고 있습니다. 그런데 AI 연구 생태계를 운영하는 "국가 AI 교육 연구센터"는 민간 에듀테크 기업와의 협력하는 모델로 삼고 있어

교육당사자의 의견과 사회적 공론화보다 사기업의 이익이 더 우선적으로 반영될 우려가 있습니다. 또 “국민 AI 기본역량 교육 확대”에서 기본 역량 교육의 방향이 기본적인 기술 이해와 활용에 치우쳐 있는 한계가 있습니다. 위 행동계획(안)은 2027년 전까지 교육 AI서비스 시행(교육 AI 규제샌드박스 시범운용 등)을 계획하고 있는데, 이는 교육당사자들의 실질적 참여와 공론화 없이 교육 AI서비스 도입을 전제하여 충분한 사회적 논의와 준비없이 1년~2년 내 실질적 도입을 예정한다는 점에서 문제가 있습니다. 그러므로 ▲민관 협력형 AI 교육정책 연구에 교육당사자들의 실질적 참여와 사회적 공론화 반영의 거버넌스 구성이 보강되어야 할 것하고, ▲ 단순한 도구 활용이 아닌 기술에서의 인권 보호 및 권리 행사 등에 관한 내용이 AI 기본역량 교육의 내용 및 방안에 반드시 포함될 수 있도록 정책 권고안에 명시되어야 합니다. 아울러 ▲ 교육당사자들의 실질적 참여와 공론화 없이 교육 AI서비스 도입을 전제하여 충분한 사회적 논의와 준비없이 1년~2년 내 실질적 도입을 계획하는 것은 재검토가 필요할 것입니다.

결론적으로, 국가인공지능전략위원회의 이번 행동계획(안)은 ‘AI 기본 사회’라는 비전을 제시했음에도 불구하고 구체적인 인권 보호 장치와 실질적인 시민 참여가 결여되어 있다는 점에서 근본적인 보완이 필요합니다. 정부와 산업계 중심의 추상적 선언에서 벗어나 국제인권규범에 입각한 구체적인 통제·구제 장치를 명문화하고, 시민사회와 교육 당사자가 단순 의견 개진을 넘어 의사결정에 실질적으로 참여할 수 있는 권한 있는 거버넌스를 구축하며, 사기업 이익과 기술 활용에 편중된 성급한 교육 정책 대신 인권 중심의 AI 역량 교육과 충분한 사회적 합의를 우선하는 방향으로 전면 수정되어야 합니다.

홍지욱 (전국민주노동조합총연맹 부위원장)

인공지능기본법이 곧 시행됩니다. 시행령에 여러 지침과 가이드라인도 만들어졌습니다. 그리고 이제 행동계획이라는 이름으로 국가 정책방향도 곧 내놓으려 하고 있습니다. 숨이 가쁠 정도로 이렇게 달려가야 하는 이유는 무엇일까요?

국가 경쟁력, 토종 AI 산업 발전과 기업 육성 등의 이름으로 인공지능, AI라는 말이 국가 산업정책의 곳곳에 박히고 있습니다. 뒤떨어지면 안된다, 선두에 서야 한다 강변하며 모든 것을 희생해서라도 가야 할 대한민국 전체의 길인 것처럼 유파지릅니다.

인공지능을 왜 만들고 사용합니까? 사람들이 살아가는데 도움을 주기 위해서 아닌가요? 그러나 사람들의 권리나 안전은 내팽개친 채 인공지능 산업 발전, 기업 성장을 전면에 부각시키고 있는 것이 한국 정부의 인공지능 산업정책인 것 같습니다.

이미 현장 곳곳에서 인공지능이 사용되고 있고, 인공지능을 통해 현장 노동자들의 노동 자체를 데이터화하고 자산화하고 있으며, 심지어 안전과 고용마저도 위협합니다. 이러한 인공지능의 개발과 운용을 규율하는 법제도가 제대로 만들어져야 합니다. 그러나 안타깝게도, 민주노총이나 시민사회의 여러 우려와 비판에도 불구하고, 인공지능기본법도, 시행령도 많은 한계와 문제점을 그대로 유지하고 있습니다. 그리고 또 이제 행동계획까지 그런 양태를 고스란히 갖고 가고 있습니다.

인공지능 산업의 발전만큼이나 중요한 것은 인공지능으로부터 영향받는 자들의 권리와 안전이 어떻게 충분히 지켜질 수 있는가입니다. 사람이 살아가는데 도움을 주기 위해 만들어지고 사용되는 인공지능으로부터 우리의 권리와 안전이 위협받을 수는 없지 않겠습니까?

인공지능을 사용하는 기업이 자기들도 이른바 이용자랍시고 그 인공지능으로부터 영향을 받는 노동자에 대한 책임을 지지 않아도 되는 얼토당토 않은 법제도 속에서 노동자의 권리와 안전은 어떻게 지켜질 수 있습니까?

노동자의 채용부터, 현장 통제와 업무 배치 및 인사 평가, 급기야는 고용 자체까지 소위 알고리즘화된 평가 기준으로 인공지능이 좌우하게 되는 현실이, 인공지능 산업 발전으로 한국정부가 그리는 산업 현장입니까? 노동자에게 재교육과 직무 전환 훈련을 제공한다고요? AI 도입과 운용으로 구조적이고 상시적인 위험과 불안에 처할 수 있는 노동자들에게 개인적인 대책이면 되는 걸까요?

노동자들은 정의로운 전환을 이야기합니다. 그 누구도 배제되지 않아야 한다는 것, 그리고 그 전환 과정 전반에 노동자들의 주도적인 참여가 있어야 한다는 것이 정의로운 전환의 핵심입니다. 인공지능이 우리 현장에 필요한지, 어떻게 운용되는지, 어떤 영향을 주는지 노동자들이 알고 함께 판단할 수 있어야 합니다. 산업 현장부터 산업별 협의, 교섭까지, 국가 산업정책 거버넌스까지, 영향받는 자로서 노동자의 참여가 반드시 필요합니다. 그것이야말로 인공지능 산업의 정의로운 전환으로 가는 길일 것입니다.

전은경 (참여연대 사회복지위원회 팀장)

대한민국 인공지능 행동계획에 대한 의견(보건복지분야)

참여연대 사회복지위원회에서 활동하는 전은경입니다.

이재명정부는 기본사회를 새로운 대한민국의 비전으로 제시하고 있습니다. 소득, 주거, 돌봄, 교육 등 모든 영역에 사회안전망을 갖추고, 국민의 기본적 삶을 국가가 책임지겠다고 합니다. 그러나 여전히 기본사회에 대한 개념이 불분명한 상태에서 진행되는 모두를 위한 AI 기본사회 계획은 어떤 의미가 있을까요. AI를 너무 강조한 나머지 AI를 잘 활용하는 기본사회, AI가 사회의 기본적인 가치들 앞에 서는 AI우선 사회가 되지 않을까 하는 우려의 말씀을 먼저 드립니다.

우선, 복지 영역과 관련해서 이번 행동계획에는 AI 기반의 정밀 복지와 돌봄의 접근성 확대 계획이 포함되어 있습니다. 이를 위해 복지부는 올해 3분기까지 AI복지 돌봄 혁신 로드맵을 수립하고, 4분기까지는 AI기반 사회위험 예측 모델 및 현장 연계 프로토콜을 개발하고, 27년 2분기까지 피지컬 AI 기반 돌봄 인프라 및 인간-로봇 공존 마을형 실증사업을 위한 계획을 마련한 다음, 30년 1분기까지 돌봄 로봇 등 핵심 기술의 실증과 상용화를 추진한다는 야심찬 계획을 가지고 있습니다.

AI 기반 예측 모형이 복지와 돌봄에 대한 접근성 확대에 일부 기여할 수는 있습니다. 하지만 현재 한국 사회의 복지 문제가 정밀한 예측의 부재로 인한 것일까요? 아닙니다. 지금 우리에게 시급한 것은 정교한 알고리즘이 아니라, 촘촘한 복지 인프라 구축과 국가의 실질적인 책임 이행입니다. 복지 예산을 늘리고 인프라를 확충하는 데 '속도전'을 펼친다는 이야기는 들어본 적이 없습니다. 그런데 왜 유독 돌봄에 AI를 들이는 일에는 이토록 유례없는 속도를 내는 것입니까?

보건의료 분야와 관련해서 제시된 액션 플랜도 상당히 우려스럽습니다.

AI 학습을 위해 민간의료기관의 의료데이터를 확보하고, 공공데이터와 민간데이터를 결합하여 활용할 길을 열어주겠다는 계획이 있습니다. AI 학습을 위해 재식별 위험이 매우 높은 의료 데이터를 결합하여 쓰면, 국가 자산이 된다는 주장입니다. 하지만, 국가의 이익과 기업의 이익은 다릅니다. 결국 국민의 정보를 재료로 삼아 AI 기업만 돈을 버는 데이터 약탈 구조를 정당화하는 결과로 이어지지 않을까요?

AI 기술로 기존의 문제를 해결할 수 있다는 환상은 보건의료 분야에서도 어김없이 반복되고 있습니다. 계획은 보건의료 특화 파운데이션 모델을 구축해 의료격차를 해소하고 공공의료를 구현하겠다고 주장합니다. 그러나 이는 실제 개발 주체인 민간 대기업의 모델에 보건의료 시스템이 종속되는 결과를 초래할 위험이 큽니다. 국가 보건의료 시스템의 안정적 운영을 기업의 영업 전략이나 서비스 유지 여부에 맡기는 것은 국가적 위험을 자초하는 일이며, 공공인프라의 운영권을 사기업에 고스란히 양도하는 것과 다름없습니다.

AI 기반 글로벌 바이오·헬스 5대 강국 도약 계획은 경제적 수치에 매몰되어, 임상적 안전성과 의료윤리를 산업 발전의 걸림돌로 취급하고 있습니다. 국민의 건강권을 지탱해야 할 재정과 데이터가 특정 기업의 기술 실증을 위한 도구로 여겨져서는 곤란합니다. 산업 논리가 의료의 공공성을 침식하지 못하도록 법적, 제도적 절연 장치를 마련하는 것이 최우선입니다. 바이오·헬스 강국의 평가지표 역시 산업 성과가 아니라 ‘의료 접근성 개선율’, ‘가계 의료비 부담 절감 기여도’, ‘AI 도입에 따른 안전 사고율’ 등 국민들이 누리는 실질적 이익을 중심으로 재편되어야 하며, 이를 위해 산업계 위주의 거버넌스를 시민 중심의 공공 거버넌스로 전환하는 일이 시급합니다.

AI 기본의료 구현 계획도 문제의 진단과 해법이 겉돌고 있습니다. 정부는 지역 간 의료 접근 격차 해소 등 의료에서 발생하는 문제를 AI로 해결하겠다고 하지만, 이는 국가가 마땅히 수행해야 할 공공의료 인력 확충 의무를 회피하고, 기술 뒤에 숨는 무책임한 태도입니다. AI는 의료인의 일을 돋는 ‘보조 도구’일 뿐이며, 대면 진료 우선의 원칙을 흔들어서는 안 됩니다. 지금 우리에게 필요한 것은 AI 솔루션 도입 예산이 아니라 사람이 중심이 되는 견고한 의료안전망입니다.

이영아 (참여연대 평화군축센터 팀장)

AI 행동계획 기자회견 발언문 (국방)

<대한민국 인공지능 행동계획(안)>은 ‘AI 기반 국방강국’이라는 장밋빛 미래를 그리고 있습니다. 효율성 및 군사력 강화를 위해 국방 AX 전환 및 가속화를 목표로 삼고 이를 방산수출 목표에까지 반영하고 있습니다. 그러나 현재 우리가 직면한 ‘인간의 생명과 안전을 좌우하는 복잡한 윤리적 판단과 무력 사용을 AI와 로봇에 맡길 것인가’라는 질문에 대한 성찰은 행동계획(안) 어디서도 찾아볼 수 없습니다.

그동안 국제사회는 무력 충돌 상황에서도 인간의 생명과 존엄성을 지키고, 민간인을 보호하고 무력 충돌의 영향력을 최소화하기 위해 국제인도법을 비롯한 다양한 원칙을 확립해 왔습니다. 그러나 인간의 판단이나 통제 없이 작동하는 자율무기체계(Lethal Autonomous Weapons Systems, LAWS)가 개발·사용되는 상황은 이러한 원칙을 뒤흔들고 있습니다.

군사분야 AI는 인간의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 수 있는 고위험 AI로 규정되고 있습니다. 이에 현재 대두된 여러 쟁점에 대한 검증과 토론을 비롯해 규제와 통제 방안에 대한 충분한 숙의와 사회적 합의가 수반되어야 합니다. 이에 국방 AI 행동계획에 대한 문제와 몇 가지 수정 제안을 하고자 합니다.

첫째, 국방AI 거버넌스에 국제법, 인권 및 인도법, 평화학 등 다양한 전문가와 시민사회 의견 수렴 보장해야 합니다. 국방부 내 국방AI위원회, 국방AI자문위원회에는 과학기술과 군사 분야 전문가 뿐 아니라 인권, 평화, 군축, 국제법, 국제정치 전문가 등의 참여가 반드시 보장되어야 합니다.

둘째, 국방 AI 기본법은 제정시한을 두지 말고 충분한 논의와 의견 수렴 절차를 거쳐 진행되어야 합니다. 행동계획(안)은 2026년 1분기 국방 AI기본법 발의, 2분기 제정을 권고하고 있습니다. 그러나 고위험 AI 영역인 국방 AI기본법을 제정하는 데 있어 국제사회에서 진행되고 있는 논의를 신중히 검토하고, 국방 AI에 대한 규제 방안 및 윤리적 기준을 기본법에 포함해야합니다. 이 과정에서 민사회를 포함한 다양한 이해관계자들과의 간담회 등 공론장을 마련하고 의견수렴 절차를 거쳐야 합니다.

셋째, ‘방산 AI 육성을 통한 방산수출 4대 강국 진입’ 목표를 삭제해야 합니다. 행동계획은 ‘방산 시장의 경쟁력은 AI 기술에 있다’며, 국내 AI 스타트업을 글로벌 시장에서 경쟁력을 갖춘 ‘K-팔란티어’ 수준으로 성장할 수 있도록 육성하는 전략 등을 제안하고 있습니다. 그러나 롤모델로 삼고 있는 팔란티어는 이스라엘의 이란 선제공격, 팔레스타인 집단학살을 뒷받침한 첨단기술을 판매해온 기업이며, 이에 유엔 팔레스타인 특별보고관은 가자지구 집단학살 지원 기업으로 팔란티어를 지목하기도 했습니다.

자율살상무기 및 AI 기반 의사결정지원체계는 인간의 생사 여부를 기계가 결정하거나 의도하지 않은 피해를 줄 수 있다는 위험이 존재합니다. AI 무기에 의한 오폭이나 민간인 피해 발생 시, 그 법적, 윤리적 책임이 어디에 있는지 명확하지 않아 전쟁 범죄에 대한 처벌을 어렵게 만드는 문제도 있습니다. AI 체계가 복잡하고 기술이 발달할수록 예측이 어렵고, 인간 통제를 벗어날 위험이 있습니다. 이러한 문제가 제기되고 있고, 이에 대한 국제사회의 우려가 지속적으로 증가하고 있는 상황에서 방산AI 확대를 국가 목표로 삼는 것은 적절치 않습니다.

마지막으로 무엇보다 지금 시급히 추진해야 하는 것은 국방부, 방사청, 외교부, 과기정통부 등 범부처가 협력하여 자율무기를 포함한 군사 AI 규제에 대한 사회적 합의를 위한 공론장을 설계 구축하는 것이 필요합니다. 공론장을 통해 ▷자율무기의 정의 ▷자율성의 허용 범위와 통제 방안 ▷책임 주체 ▷투명성과 설명 가능성 확보 방안 ▷민간인 보호와 국제법 준수 방안 ▷무기체계의 신뢰성과 안전성 평가 방안 ▷변화된 전쟁에 대한 전군 교육 방안 ▷보안 대책 ▷군에 대한 민주적 통제 방안 ▷그 밖의 광범위한 윤리적·법적 쟁점 등을 충분히 검증, 토론, 숙의해야 합니다. 그리고 이를 바탕으로 군사 AI 규제를 통한 통합적인 로드맵을 수립하는 것이 필요합니다. 이 과정에 국방 AI 추진체계 재설계나 국방 AI 기본법 추진 등의 계획보다 선행되어야 합니다.

온다 (한국여성민우회 성평등미디어팀)

국가인공지능행동계획 성평등 분야 의견

시민사회공동의견서의 여성 부문 의견에 한국여성민우회에서 제출한 의견을 덧붙여서 말씀드립니다. 앞서 발표해주신 바와 같이, 이번 행동계획은 AI의 무조건적 도입과 산업 증진의 측면에만 집중할 뿐 AI가 불러오는 위험과 기본권 침해에 대응할 방안은 거의 가지고 있지 않습니다. 특히 성평등 관점으로 검토하였을 때, 성평등·인권과 여성에 관한 정책이 매우 부족하고 일부 성평등가족부의 역할도 분절적으로 제시되어 있습니다.

따라서 무엇보다 먼저 제시하고자 하는 전반적인 수정 의견은 AI 정책에서 통합적인 성평등 실현 과제를 설정하고, 이에 따른 액션플랜을 수립하여야 한다는 것입니다. AI 개발과 활용 분야는 여성 참여와 대표성이 낮고, AI는 성차별과 여성혐오 및 젠더폭력의 문제를 일으키고 있으며 여성은 노동 등 영역에서 AI 전환에 따른 영향을 크게 받는 집단이기도 합니다. 이는 개별적인 문제가 아니라, 구조적 성차별에서 비롯된 문제로서 이에 대한 통합적 대책이 필요합니다. 따라서 국가는 AI 분야에서의 성평등 증진을 국가의 책무로 인식하여 목표를 설정하고 그에 따른 성평등 정책의 추진 계획을 제시해야 합니다. AI 분야 전반의 성별 영향 분석과 평가가 필요합니다.

다음으로 세부 액션플랜에 대한 수정 의견을 발표하겠습니다. AI 관련 국가 정책 결정 구조에 성평등 관점이 반영되어야 하고, AI 분야에 성평등한 참여가 보장되어야 합니다. 54번, 55번, 81번, 82번, 83번 플랜과 관련되어 있는데요. AI 정책 수립 과정에서의 성평등 관점을 위해 성평등 분야 전문가와 단체가 거버넌스에 참여하고, 정책 결정 기구와 회의체계 전반에 특정 성별이 60%를 넘지 않도록 하는 등 여성 참여를 제도적으로 보장해야 합니다. 또한, 26번 플랜에서 AI 분야 인력의 경력단절 대책을 제시하고 있는데, AI 분야의 성별 격차는 단순히 경력단절뿐 아닌 기술 분야 교육, 고용, 승진 등 전 과정에서의 구조적 성차별의 결과이므로 이에 대한 문제 인식과 대책이 필요합니다.

다음으로 평등과 인권 가치에 기반한 AI 윤리기준 마련과 정책 반영이 필요합니다. 관련 액션플랜은 55번과 80번, 82번입니다. 공공부문 윤리기준에서 성차별을 비롯한 AI의 차별 문제를 다루어야 하고, 「모두를 위한 AI 기본 사회 추진 계획」수립에 있어서도 평등권을 비롯한 인권 보장을 기본 원칙으로 삼아야 합니다. 인공지능 윤리기준을 점검하여 고도화하는 계획에서도 자율적 준수가 아닌 공적 감독과 책임에 연동되는 의무적 평가도구로 재설계할 필요가 있습니다. 그리고 92번의 교육 분야 액션플랜인데요. 전 세대와 집단 대상 AI 기본 역량 교육에 성차별을 비롯한 차별 혐오 예방과 인권 보장 내용을 포함해야 합니다.

AI 시대 경제와 일자리 분야 정책에 젠더 관점 도입이 필요합니다. AI 도입으로 영향받는 돌봄 등 직군과 비정형, 플랫폼노동 등 고용 형태에 여성 비율이 높은 문제를 인식하고 대응해야 합니다. 관련 액션플랜은 86번과 87번인데요. 관련 대책에 성평등가족부의 참여와 성평등 관점 도입이 필요합니다.

돌봄 분야에서 AI의 선제적 도입과 사후적 문제 해결을 전제하고 있는 점도 이번 행동계획의 문제입니다. 88번과 89번 액션플랜입니다. AI 돌봄 정책에 돌봄 정의 관점의 개입이 있어야 합니다. 돌봄 현장에 대한 실질적 이해에 기반한 AI 도입 여부 판단이 이루어질 수 있도록 돌봄 당사자가 AI 돌봄 정책 수립 과정에 참여해야 합니다.

마지막으로 AI를 악용하는 젠더폭력에 대한 개발 및 활용, 유통 주체의 책임이 강화되어야 합니다. 관련 액션플랜은 90번입니다. 현재 딥페이크 범죄에 한정한 사후 대책 위주로 서술되어 있습니다. 플랫폼 및 AI 개발 주체가 딥페이크와 사이버불링, 혐오발화 등 젠더폭력 확산 방지를 위한 책임을 질 수 있도록 정책적 개입이 필요하겠습니다.

하장호 (문화연대 정책위원장)

정부의 인공지능 행동계획안은 AI 기술 도입으로 인해 기대되는 사회적 효용과 이로인해 발생할 사회적 갈등과 문제들을 균형있게 살피지 못하고 있음.

AI 기술이 가져올 긍정적 변화는 그 구체적 기대값이 명확하지 않은데 반해 AI 기술의 전면화에 따른 우리사회 각 부문의 문제와 예상되는 피해는 구체적이며 불가역적인 변화를 예고하고 있음.

정부의 계획안은 이러한 기대와 우려 사이에서 새로운 기술환경이 가져올 사회변화를 예측하고 공공의 이익을 보장하면서도 사회적 갈등과 피해는 조율하기 위한 구체적인 행동 방향이 담겨 있어야 함에도 불구하고 근거없는 기술적 낙관에 기반한 산업성장 지원책과 규제 완화 등에 대한 장미빛 계획만을 제시하고 있음.

문화분야는 AI 기술이 특별한 제약없이 가장 적극적으로 활용되고 있는 분야이며, 계획에서도 지적하고 있는 바와 같이 기술 도입에 따른 생태계 변화와 혼란이 충분히 예상되고 있으나 실제 제출된 계획안에서는 이러한 우려들을 불식시킬만한 구체적인 방안은 제시되고 있지 못함.

특히 문화산업 분야에서의 '워크슬롭' 현상과 이로 인한 소모적인 과잉노동, 창작자의 번아웃 문제는 이미 현실화 되고 있으며, AI 도입으로 인한 일자리 감소 문제나 노동환경 악화도 예상되고 있고, AI 도입에 따른 저작권 문제나 저작인격권 문제는 이미 갈등의 현장이 되고 있음.

문화 부문에서의 AI 도입은 단순히 기술적 환경 변화라는 측면을 넘어 예술의 정체성을 뒤흔드는 변화를 가져올 수 있으며, 창작과 산업 생태계 전반에 걸친 대대적인 변화와 문화적 전환을 예고하고 있음.

이는 계획에서 제시하고 있는 기술 개발, 기술지원, 예산지원, 인력 육성과 같은 개별화 된 지원책만으로 대비할 수 있는 변화가 아니며 생태계 전반을 아우르는 거시적인 공론장과 거버넌스의 구축을 통해서 민관이 같이 협력하며 준비할 때 그 피해는 최소화하면서도 새로운 미래를 담아내는 생태계 구축이 가능할 것임.

특히 문화예술분야 경우 4개 정도 액션플랜이 제시되어 있는데, 콘텐츠산업 증진이 주 내용임. 문화정책의 영역은 문화산업, 콘텐츠 뿐만 아니라 개인의 일상 생활 문화, 문화소양교육, 게임정책 등도 있음. 그런데 해당 내용은 다 빠져있고 실제로 생산자나 기업, 유통기업 등에 대한 증진정책 중심으로 제시되어있음.