

인권과 안전, 민주주의가 보장되는 인공지능 정책을 요구하는 시민사회 선언

오늘 우리 120개 단체들은 인권과 안전, 민주주의가 보장되는 인공지능 정책을 요구하는 선언을 발표합니다.

인공지능은 이미 우리 생활에 널리 사용되고 있습니다. 검색, 자동 번역, 추천 알고리즘, 챗봇처럼 소비자들이 인터넷에서 만나는 수많은 서비스들 뿐 아니라 AI스피커, 가전, 자동차와 같은 제품들에도 인공지능이 탑재되어 있습니다. 일터에서는 인공지능이 업무 성과를 관리 및 평가하기 시작했으며 플랫폼 노동처럼 알고리즘으로 업무를 배치받고 지시받는 사람들이 많아지고 있습니다. 금융 서비스 뿐 아니라 생계에 많은 영향을 미치는 신용 평가나 대출, 보험의 요율이나 지급 여부 결정에도 인공지능이 관여하고 있습니다. 실업 급여 지급 결정이나 부정수급 탐지¹와 같은 사회복지 업무도 인공지능으로 수행될 것으로 예측되고 있습니다. 경찰은 어느 지역이 범죄율이 높은지 예측하는 인공지능 치안활동을 전국적으로 시작하였으며², 과학기술정보통신부와 지방자치단체는 코로나 확진자와 접촉자를 추적하기 위하여 거리 CCTV에 얼굴인식을 도입하고 있습니다³. 이미 수많은 공공기관과 민간기업들이 채용절차에 <AI면접>이나 <AI서류평가> 등 인공지능 도구를 사용하고 있고⁴, 군대도 <AI면접>을 통해 특기병을 결정한다고 합니다⁵. 인공지능 무기 또한 눈앞의 현실로 다가왔습니다⁶.

인공지능의 안전과 인권 침해 문제

이처럼 인공지능은 서비스나 제품을 추천·제공하는 단계부터 노동, 금융, 사회복지, 치안, 군사 영역에서 사람을 평가하고 중대한 의사결정을 내리는 단계까지 나아가고 있습니다. 이는 소비자, 노동자, 시민의 인권과 안전에 중대한 영향을 미칠 것입니다. 특히 안전이 검증되지 않는 인공지능 제품이 일상생활에서 사용된다면 우리의 생명조차 위험에 처할 것입니다.

¹ “머신러닝으로 부정수급 탐지, AI로 강력범죄 전자감독”, 과학기술정보통신부 보도자료(2020. 4. 13).

² “경찰청, 빅데이터 플랫폼 구축하고 인공지능(AI) 범죄 예방 시범운영”, 경찰청 보도자료(2020. 12. 28); “인공지능(AI)으로 범죄예방의 첫걸음 내디딘다”, 경찰청 보도자료(2021. 2. 26).

³ “CCTV 활용 안면인식·동선분석'으로 코로나 감염경로 파악한다”, 뉴스1(2021. 2. 25).

⁴ “공공기관 AI 채용은 ‘안알리즘’”, 한겨레21 제1335호(2020. 10. 23).

⁵ “AI가 軍 특기병 면접본다...7월부터 시범 운영”, 조선일보(2021. 4. 13).

⁶ “인공지능 무기 전쟁”, 한겨레(2021. 4. 10).

미국에서 운행 중인 자율주행차량들이 오작동으로 몇 차례 사망사고를 일으켰지만 그 책임 문제가 여전히 불분명합니다⁷.

또 인공지능이 드러내는 편향성에 대하여 세계적으로 논란이 불거지고 있습니다. 온라인 챗봇이 공개 후 불과 몇시간 만에 인종주의자가 되었고⁸, 많은 기관에서 사용되는 얼굴인식기술이 흑인 여성을 잘 식별하지 못하였습니다⁹. 인공지능은 채용지원자 중 여성을 선호하지 않았고, 흑인들의 재범 위험을 백인보다 2배 높다고 판단하였습니다¹⁰. 지난해 영국에서는 코로나로 인해 취소된 대학입학시험 대신 인공지능으로 성적을 부여하였다가 부유한 지역 학생이 높은 점수를 받고 가난한 지역 학생이 낮은 점수를 받은 것으로 드러나 큰 사회적 논란을 겪기도 했습니다¹¹.

이는 기존의 데이터셋으로 학습하여 미래를 예측하는 인공지능의 본질적인 특성에 따른 문제이기도 합니다. 업계에서는 인간에게도 편견이 있다며, 인공지능은 인간사회의 차별을 학습했을 뿐이라고 주장합니다. 실제로 우리는 인공지능 챗봇 이루다가 어떻게 여성, 장애인, 성소수자, 유색인종에 대한 편견과 혐오를 빠르게 배우고 재현하는지를 직접 목격하였습니다. 문제는 이러한 인공지능의 편향성이 사회적 통제 없이 가정에서, 학교에서, 직장에서 사용되는 제품과 시스템에 바로 적용된다면, 우리 사회에 이러한 편견과 혐오가 항구적이 될 것이라는 점입니다. 그리고 이러한 편견과 혐오에 기초한 차별은 우리 사회에서 더욱 취약한 사람들의 권리를 위협할 수밖에 없습니다.

사람을 보호하는 법제도와 보호장치 미비

그럼에도 인공지능으로 영향을 받는 당사자들이나 일반 국민은 인공지능의 도입, 운영, 결정에 대하여 발언과 참여 기회를 전혀 보장받고 있지 못합니다. <AI 채용>에서 불합격된 사람들은 본인이 어떻게 평가받고 채용 탈락으로 결정되었는지 알지 못합니다¹².

공공기관들은 인공지능을 적용한 면접이나 서류평가가 공정한지, 면접대상의 외모나 사투리를 차별하지는 않는지 전혀 검증하지 않고 사용해 왔습니다. 플랫폼 기업의 불투명한 알고리즘은 노동자와 자영업자들에게 고통을 주고 있습니다. 포털 뉴스의 비밀 알고리즘은

⁷ “우버 자율주행차 첫 보행자 사망사고...안전성 논란 증폭”, 한국일보(2018. 3. 20). ; “나무 들어받고 불탄 테슬라...운전석에는 아무도 없었다”, 경향신문(2021. 4. 19).

⁸ “MS 채팅 봇 ‘데이’, 24시간 만에 인종차별주의자로 타락”, 동아사이언스(2016. 3. 27).

⁹ “인공지능도 인종차별?...”얼굴인식 소프트웨어 백인남성 더 정확”, 연합뉴스(2018. 2. 12).

¹⁰ “아마존, '여성차별' 논란 인공지능 채용 프로그램 폐기”, BBC(2018. 10. 11). ; ‘AI 채용’은 남성을 선호해?...성차별 논란에 '자체 폐기', 중앙일보(2018. 10. 12).

¹¹ “사는 곳으로 성적을 결정했다 : 가난한 지역 공립학교 학생들에게 낮은 점수 준 알고리즘, 실패한 실험이 남긴 과제”, 한겨레21 제1329호(2020. 9. 11).

¹² “시민단체, 인공지능 채용 공공기관 13곳에 정보공개 청구 및 결과 발표”, 진보네트워킹센터 보도자료(2020. 10. 27). ; “AI 면접 공정평가” 온라인 토론회서 쏟아진 쟁점”, 뉴스로드(2021. 2. 17).

공론장의 편향성에 영향을 미치고 있습니다¹³. 해외에서는 플랫폼 기업의 알고리즘이 선거와 민주주의에도 영향을 미쳤다는 경각심이 일고 있습니다¹⁴.

세계 여러나라에서 인공지능 알고리즘의 공정성과 투명성에 대한 소송이 제기되었고, 법원은 인공지능의 불투명성에 도전하는 당사자들과 시민사회의 손을 들어 주었습니다. 2017년 폴란드 법원은 정부의 실업자 점수 알고리즘에 대하여 정보공개를 하도록 하였고¹⁵, 미국 법원은 민간기업에서 조달한 교육청의 교사 평가 알고리즘에 대하여 투명성과 적법절차 부족을 이유로 운영을 중단시켰습니다¹⁶. 특히 미국 법원은 민간기업의 영업비밀과 국민의 헌법상 권리인 적법절차를 모두 충족하기 위해서는 공공기관의 중요한 의사결정에 비밀 알고리즘을 사용해서는 안된다고 설시하였습니다. 2020년에는 네덜란드 법원이 사회복지급여 부정수급 탐지 시스템에 대하여 투명성 부족과 개인정보보호법 위반을 이유로 운영을 중단하라는 내용의 판결을 내렸습니다¹⁷.

한국 사회에서 알파고가 던진 충격 이후 인공지능의 놀라운 발전은 언론, 산업계, 학계, 정부와 국회에서 큰 관심을 받았습니다. 그러나 이러한 관심 대부분은 대체로 산업적 가치에만 초점을 맞추어 왔습니다. 기술이 어떻게 사람의 인권과 안전을 보장할 것이며 이 기술을 어떻게 민주적으로 통제할 것인지에 대한 문제는 진지하게 검토되지 않았습니다. 나아가 인공지능 기술이 사람들의 인권과 안전에 영향을 미치는 것이 분명함에도, 정작 영향을 받는 사람들이 의견을 개진할 수 있는 공간은 전무한 상황입니다.

인공지능도 인권과 법률을 준수해야 합니다.

최근 인공지능 챗봇 이루다 개발사의 개인정보 보호법 위반 사건에서 확인하였듯이, 인공지능이라고 하여 인권과 법률 준수 의무에서 예외가 될 수는 없습니다. 모든 인공지능 시스템은 헌법과 관련 법률에 따라 인권 침해와 차별을 금지하고, 소비자의 안전과 권리를 보호하며, 노동자와 노동권을 보호하는 현행 법률을 준수해야 합니다.

인공지능 챗봇 이루다 사건은 우리에게 많은 충격과 분노를 주었습니다. 인공지능 서비스가 우리 사회의 편견과 혐오를 재생산하였을 뿐 아니라 그 개발 과정에서 개인정보를 소홀하게 취급하여 이용자들의 신뢰를 배반하였습니다. 결국 지난 4월 28일 개인정보보호위원회는 이루다 개발사 (주)스캐터랩에 총 1억여 원의 과징금과 과태료를 부과하였습니다. (주)

¹³ “AI가 편집한다지만...포털뉴스 이대로 좋은가?”, MBC 스트레이트(2021. 3. 7).

¹⁴ “무심코 누른 ‘좋아요’가 당신의 투표심리를 조종한다”, 서울경제(2020. 4. 10.)

¹⁵ “Poland: Government to scrap controversial unemployment scoring system”, Algorithm Watch(2019. 4. 16).

¹⁶ “Federal Suit Settlement: End of Value-Added Measures in Houston”, Education News(2017. 10. 10).

¹⁷ “Welfare surveillance system violates human rights, Dutch court rules”, The Guardian(2020. 2. 5).

스캐터랩은 과거에 운영했던 서비스를 통해 수집한 약 60만 명에 달하는 이용자의 내밀한 카카오톡 대화문장 94억여 건을 제대로 된 보호조치를 취하지 않고 자사 다른 서비스인 인공지능 챗봇 서비스 ‘이루다’의 개발과 운영에 이용한 것으로 확인되었습니다. 특히 만14세 미만 아동의 개인정보를 무단으로 이용한 규모가 20만 명에 달했습니다.

그런데 이루다 조사와 조치 결과를 두고 일부 업계는 인공지능 상품과 서비스 개발에 지장을 초래하고 있다고 불멘 소리를 내고 있습니다. 이는 일반 소비자와 이용자들의 생각과 현격한 차이가 있습니다. 우리는 오히려 이루다 사건에서 이용자가 믿고 맡긴 민감한 개인정보를 기업이 이용자가 모르는 다른 상품 개발에 무단으로 마구 사용하거나 판매할 수 있는 현실을 보았습니다. 개인정보 추가이용과 가명정보의 이용을 원하는 산업계의 요구에 따라 2020년 이미 개인정보보호법과 그 시행령이 편향적으로 개정되었음에도, 그조차 지키지 못하겠다는 업계의 엄살은 도저히 납득하기 어렵습니다. 이런 논란을 지켜본 소비자와 이용자들은 인공지능 제품과 서비스를 충분히 믿고 안심하며 이용하기가 어렵습니다. 공공기관과 민간기업의 모든 인공지능 제품과 시스템들이 부족하나마 개인정보 보호법을 준수하고 있다는 신뢰가 바탕이 되어야 시민들 또한 그 혜택을 마음 놓고 누릴 수 있을 것입니다.

무엇보다 인공지능이 우리 사회의 편견과 혐오를 재생산하는 것을 이대로 방치할 수 없습니다. 인공지능 제품과 서비스 역시 고용과 서비스 등에서 성별, 장애, 연령, 인종, 지역 등으로부터 차별을 금지하는 현행 법률을 준수하여야 합니다. 인공지능의 도입이 확산될수록 그로 인한 차별에 대한 우려 또한 커지고 있는 만큼, 효과적이고 포괄적인 차별금지법의 제정도 이루어져야 할 것입니다.

더불어 인공지능 제품과 시스템들은 헌법과 관련 법률에서 보호하고 있는 표현의 자유, 집회시위의 권리 등 인권을 침해해서는 안 됩니다. 제품의 안전과 소비자의 권리를 보호하는 현행 법률들과 노동자와 노동권을 보호하는 현행 법률 또한 준수해야 합니다.

인공지능을 규율하는 법과 제도가 시급히 마련되어야 합니다.

이런 상황에서 네덜란드 암스테르담과 핀란드 헬싱키시¹⁸, 프랑스 앙티브시¹⁹ 등 세계 여러나라 지방자치단체들이 인공지능 알고리즘과 관련한 사항을 시민들에게 모범적으로 공개하고 의견을 수렴하고 있다는 사실은 시사하는 바가 매우 큽니다. 나아가 여러나라 정부가 인공지능을 규율하는 법제도를 도입하였거나 도입을 준비 중입니다. 캐나다 정부는

¹⁸ “Amsterdam and Helsinki launch algorithm registries to bring transparency to public deployments of AI”, VentureBeat(2020. 9. 28).

¹⁹ “Antibes publie l’inventaire de ses algorithmes”, laGazette(2021. 2. 11).

2019년부터 공공기관이 도입·운영하는 인공지능 의사결정 시스템에 대하여 영향평가 후 위험수준별로 정보공개, 인적 개입, 설명, 검사 등의 의무를 부과하는 훈령을 시행 중이며²⁰, 뉴질랜드 정부는 2020년부터 공공기관이 도입·운영하는 인공지능 알고리즘에 대하여 영향평가 후 위험수준별로 규제하고 있습니다²¹. 유럽연합 집행위원회는 최근 유럽시민이 신뢰할 수 있는 인공지능 규제법안을 유럽의회에 발의하였습니다²².

인공지능을 이용한 자동화는 제품과 시스템의 효율성과 예측성을 크게 높일 수 있고 인간이 기존에 수행할 수 없었던 업무 영역을 확장시켰습니다. 그러나 인공지능이 사회적으로 적절히 통제되지 않는다면 그 편향성과 위험성이 사람의 인권과 안전에 중대한 영향을 미칠 것입니다. 이용자와 소비자는 그 의사결정 사유를 설명할 수도 없고, 검증할 수도 없고, 참여할 수도 없는 인공지능을 신뢰하지 않게 될 것입니다. 소비자, 노동자, 시민이 신뢰하지 못하는 인공지능에 대해서는 사회적 수용성 및 효용성 또한 감소할 수밖에 없습니다.

최근 인공지능 규범과 관련한 여러 국제문헌들은 인공지능을 규율하는 새로운 법률이 필요하다고 강조하고 있습니다. 인공지능 제품과 시스템을 국가적으로 감독하고, 인공지능 시스템에 대한 정보공개와 참여로써 그 투명성과 민주성을 보장하며, 위험한 인공지능을 제도적으로 통제하는 한편, 피해를 입거나 인권을 침해당한 사람에 대하여 권리구제를 보장해야 한다는 것입니다.

또한 인공지능의 무기화가 빠르게 진행되고 있는 현실에서 “인공지능이 인간의 통제없이 개별 표적을 선별하고, 공격하는 자율살상무기(LAWS)”를 규율 할 수 있는 법적, 제도적 장치가 선제적으로 마련되어야 합니다. 이는 전쟁의 비인간화와 비인간성이 초래하는 인류 존엄성의 훼손을 막기 위해 필요한 조치입니다.

이에 우리 시민사회는 인공지능 규율법의 마련을 촉구하며, 이 법률과 제도로써 ①인공지능 국가 감독 체계 마련 ②정보 공개와 참여 ③인공지능 평가 및 위험성 통제 ④권리구제 절차 보장을 요구합니다. 인공지능 역시 다른 모든 제품 및 기술 시스템과 마찬가지로 법률을 준수하고, 인권을 침해하지 않으며, 안전을 보장하여야 하고, 국가와 지방자치단체, 공공기관은 물론 기업은 이를 위한 조치를 취해야 할 의무가 있습니다. 딥러닝 등 특정한 기술적 특성이나 기업자율적인 윤리가 이런 의무를 회피하는 이유가 될 수는 없습니다.

²⁰ The Government of Canada, “Directive on Automated Decision-Making”.

²¹ New Zealand Government, “Algorithm charter for Aotearoa New Zealand”.

²² European Commission, “Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)”.

【참고】 유엔 사무총장(2020)²³

(b) 국가는 민간부문 활동에 관한 조치를 포함한 입법 조치를 취해야 할 의무를 상기하고 이를 준수해야 한다. 이를 통해 신기술은 모든 사람이 경제·사회·문화적 권리를 비롯한 인권을 완전히 향유하고 인권에 미치는 부작용을 방지하는 데 기여할 수 있을 것이다.

(h) 신기술이 사용되는 상황에서 책임을 완전하게 보장하는 적절한 법률체계와 절차방법이 마련되어야 한다. 이를 위해 국내 법제도의 공백을 검토 및 평가하고, 필요한 감독 체제를 수립하고, 신기술로 인한 피해에 대해 사람이 이용할 수 있는 구제 수단이 구비되어야 한다.

【참고】 유럽 기본권청(2020)²⁴

인공지능에 대해 새로운 정책이나 새로운 입법을 도입할 때 입법자와 국가들은 ... 모든 기본권이 존중될 수 있도록 보장해야 한다. 특정한 기본권 보호조치는 관련 정책과 법률을 수반할 필요가 있다 ... 인공지능 개발자와 사용자 모두를 기본권에 대한 자의적인 간섭으로부터 효과적으로 보호하고 법적 확실성을 갖추기 위해서 관련 안전조치가 법으로 규정될 필요가 있다. 인공지능의 개발과 사용에 있어 기본권을 지키고 보호하기 위한 자율적인 제도들도 추가적으로 권리 침해 완화를 도우는데 도움이 될 수 있다. 입법자가 인공지능 법의 범위를 설정할 때는, 법적 명확성의 최소 요건, 즉 법치주의 원칙 및 기본권 보장의 전제 조건에 충분한 주의를 기울여야 한다.

인공지능 제품과 시스템은 국가적으로 감독되어야 합니다.

최근 정부는 과학기술정보통신부/4차산업혁명위원회 등 산업진흥 부처/기구가 인공지능 정책을 주도하고 있으며 관련 입법 또한 빠른 속도로 추진되고 있습니다²⁵. 그러나 인공지능으로 사람에 대한 의사결정에까지 이르도록 하면서 그로부터 사람을 보호하기 위한 조건이나 절차에 대한 논의는 충분히 이루어지지 못하였으며, 이는 과학기술정보통신부 등이 주도하는 인공지능 정책이 산업계의 요구로 점철되어 있는 데 따른 결과입니다.

반면 다른 나라의 경우 산업관련 부처가 인공지능 정책에 관여하더라도 부분적이거나 기술과 관련한 역할에 그치고 있습니다. 이들 국가에서 국민을 위해 인공지능을 감독하는 역할은 주로 시장감독기구, 국가인권기구 또는 평등기구, 개인정보보호감독기관에서 수행하고 있습니다. 이 기관들은 인공지능 제품과 시스템으로부터 소비자 보호, 인권 침해 구제, 차별 금지, 개인정보 보호 등을 소관하고 있습니다.

²³ “Question of the realization of economic, social and cultural rights in all countries: the role of new technologies for the realization of economic, social and cultural rights”, 유엔문서번호 A/HRC/43/29(2020. 3. 4).
²⁴ European Union Agency for Fundamental Rights(2020), “Getting the Future Rights : Artificial Intelligence and Fundamental Rights”.
²⁵ “인공지능 시대를 준비하는 법·제도·규제 정비 로드맵 마련”, 과학기술정보통신부 보도자료(2020. 12. 24). ; “4차산업혁명특 간담회 - AI에 지나친 쏠림 줄이기...자율규제 방식이 바람직”, 한국경제(2021. 3. 30).

사람의 인권과 안전을 보장하고 권리구제의 책무가 있는 국가는 인공지능 국가 감독 체계를 수립하고, 여기에 공정거래위원회, 국가인권위원회, 개인정보보호위원회가 반드시 참여하도록 해야 합니다.

인공지능 국가 감독 체계는 인공지능에 대하여 사전적으로나 사후적으로 감독 권한을 행사하고, 진정 사건을 조사하고 피해자의 권리를 구제할 수 있는 충분한 권한과 자원을 구비해야 하며, 감독대상이 되는 인공지능 제품과 서비스의 경우 자료와 데이터를 보관하도록 의무화하여 그 설명과 조사가 가능하도록 하여야 합니다.

【참고】 유럽평의회 인권위원장(2019)²⁶

각국은 공공기관과 민간기업이 개발, 도입, 사용하는 인공지능 시스템의 인권 준수에 대한 독립적이고 효과적인 감독을 위해 입법 체계를 수립해야 한다. 이 입법 체계는 상호협력하는 행정적, 사법적, 준사법적 또는 의회의 감독기구들의 조합으로 구성된 구조를 포함할 수 있다. 각국은 적절한 경우 기존 국가인권기구에 인공지능 시스템의 인권 준수에 대해 독립적이고 효과적인 감독 역할을 수행하는 권한의 부여를 검토해야 한다.

감독기구는 인공지능 시스템을 개발, 도입 및 기타 사용하는 공공기관과 민간기업들로부터 독립적이어야 하며 감독기능을 수행하기에 적절하고 충분한 분야간 전문지식, 역량 및 자원을 구비해야 한다.

독립적인 감독기구는 인공지능 시스템의 인권 준수를 능동적으로 조사 및 모니터링하고 영향을 받은 개인들로부터 진정을 접수, 처리하며 인공지능 시스템의 성과와 기술 개발에 대한 주기적인 검토를 보다 일반적으로 실시해야 한다. 감독기구는 인권 침해 발생의 위험성을 포착한 상황에 개입할 수 있는 권한을 보유해야 한다. 감독기구는 또한 정기적으로 의회에 보고하고 활동에 대한 보고서를 발표해야 한다.

공공기관과 민간기업들은 감독기구가 인공지능 시스템의 효과적인 감독에 필요한 모든 정보를 요청할 때 이를 제공해야 하고 이들에 정기적으로 보고해야 한다. 또한 인공지능 시스템이 인권에 미친 영향과 관련해 감독기구의 권고를 이행 해야 한다. 또한 감독 절차는 투명해야 하며 적절한 공공 조사의 대상이 될 수 있어야 하고 감독기구의 결정은 이의제기 또는 독립 심사의 대상이 될 수 있어야 한다.

【참고】 유럽 기본권청(2020)

국가는 인공지능 시스템이 기본권에 미치는 부정적인 영향을 감독하고, 필요한 경우 이를 실질적으로 해결할 수 있는 효과적인 책임 시스템을 구축해야 한다 ... 인공지능 시스템이 기본권에 미치는 영향을 평가할 수 있도록 충분한 정보를 제공하도록 요구하는 법적 의무도 포함될 수 있다 ... 국가는 인공지능 사용 시 기본권을 보호하기 위해 개인정보 보호 기관, 평등 기구, 국가 인권 기구, 옴부즈 기관 및 소비자 보호기관 등 기존의 감독 전문 조직을 더 잘 활용해야 한다. ... 이들 기관들이 충분한 자원, 권한 및 중요한 전문지식을 갖추어 기본권 침해를 예방 및 평가하고 자신의 기본권이 인공지능에 의해 영향을 받는 사람들을 실질적으로 도울 수 있어야 한다.

²⁶ Council of Europe Commissioner for Human Rights(2019), “Unboxing artificial intelligence: 10 steps to protect human rights”.

인공지능 정책은 소비자, 노동자, 시민에게 투명하게 공개되고 그 참여가 보장되어야 합니다.

인공지능 제품과 시스템의 사용은 그 소비자, 노동자, 시민에게 위험하거나 중대한 영향을 미칠 수 있습니다. 따라서 이 제품과 시스템을 개발, 도입하고 의사결정에 사용하는 공공기관과 민간기업은 그 소비자나 이용자에게 모든 인공지능 기능과 그 의미에 대하여 이해할 수 있는 형태로 공개하고 가능한 한 당사자의 선택권을 보장해야 합니다.

특히 정보공개 의무를 지고 있는 공공기관의 경우 그 인공지능 시스템 운영의 담당자와 책임자에 대한 정보뿐 아니라, 훈련 데이터셋과 처리 방법에 대한 정보, 개인정보 보호와 차별 방지를 위한 조치에 대한 정보, 인적 감독에 대한 정보, 그 기능의 영향력과 위험성에 대한 정보를 공개하여야 마땅합니다.

이 권리는 공공기관이 민간에서 조달한 시스템이나 서비스를 사용하는 경우에도 똑같이 보장되어야 합니다. 기업의 영업비밀 보호가 공공기관의 투명성 의무를 회피하는 조건이 되어서는 안되며, 국민 앞에 투명하게 공개할 수 없는 민간 시스템을 공공기관이 도입해서는 안 될 것입니다. 더불어 인공지능 시스템의 공공조달에 대해서는 위험성을 평가하고 통제하는 절차가 마련되어야 하며 그 평가 내용 또한 공개되어야 합니다.

나아가 공공기관의 경우 인공지능 시스템의 운영으로 영향을 받는 소비자, 노동자, 시민 등 모든 이해당사자들이 참여할 수 있고 그 의견을 수렴하는 공청회를 사전에 실시하는 것이 필요합니다.

【참고】 유엔 사무총장(2020)

신기술의 개발 및 도입에 관한 의사결정에 모든 관련 이해당사자의 참여를 보장하고, 특히 공공부문에서 인공지능이 지원하는 의사결정에 대하여 적절한 설명가능성이 보장될 필요가 있다.

【참고】 캐나다 정부 훈령 <자동화된 의사결정에 대한 지침>(2019)

6.2. 투명성

(의사결정 전 공지)

6.2.1. 해당 의사결정이 부록 C에 규정된 바대로 자동화된 의사결정 시스템에 의해 전체 또는 부분적으로 수행된다는 내용을 관련 웹사이트에 공지한다.

6.2.2. <Canada.ca 콘텐츠 스타일 가이드>에 부합하는 뚜렷하고 쉬운 용어를 이용하여 공지한다.

(의사결정 후 설명)

6.2.3. 부록 C에 규정된 대로 결정이 내려진 방법과 이유에 대해 영향을 받는 개인들에게 이해가능하게 설명한다.

인공지능의 위험성이 제도적으로 평가되고 통제되어야 합니다.

대다수 인공지능 제품과 시스템은 소비자, 노동자, 시민에게 위험하지 않을 수 있습니다. 그러나 어떤 인공지능은 개인이나 공동체의 권리나 안전에 영향을 미칠 수 있습니다. 영향이 미치는 범위는 협소하고 미미한 수준에 그칠 수 있지만 때로는 광범위하거나 심각한 데 이를 수 있습니다. 그 빈도 또한 일시적일 수 있지만 준영구적으로 발생할 수도 있습니다.

모든 인공지능 제품과 시스템은 개인정보 보호법, 제조물책임법, 소비자관련법, 노동관련법 등 현행 법률을 준수해야 합니다. 여기에 더하여 위험한 인공지능은 제도적으로 관리되어야 하며, 인공지능의 위험성은 생애주기별로 평가되어 그 위험성을 완화하기 위한 조치가 취해져야 합니다.

의료, 운송, 에너지, 금융 및 공공부문에서 인권에 법적 또는 상당한 영향을 미치는 경우, 사람의 부상·사망이나 상당한 물질적·비물질적 손상을 초래하는 경우, 채용과 업무관리 등 노동권과 입점업체 배정에 영향을 미치는 경우, 소비자권에 영향을 미치는 경우, 원격으로 생체를 인식하거나 침입적으로 감시하는 인공지능은 고위험으로 보아야 할 것입니다. 이러한 고위험 인공지능에 해당할 경우 그 훈련 데이터셋은 편향성 방지와 안전성을 보장하는 조치를 사전과 사후에 취해야 하고, 감독기관의 조사나 피해자 권리구제시 검증이 가능하도록 인공지능 시스템 생애주기별 데이터와 방법론을 문서화, 기록, 보존해야 하며, 필요한 경우 인적 개입을 의무화해야 합니다. 공공기관의 모든 인공지능 시스템과 민간기업의 고위험 인공지능 시스템은 국가기관에 등록하여 그 의무를 준수하는지 국가적으로 관리해야 합니다.

사람의 안전과 생계, 인권에 대하여 명백한 위험을 야기하는 어떤 인공지능 시스템은 금지되어야 합니다. 실시간으로 거리에서 불특정다수의 얼굴을 인식하여 추적하는 시스템은 원칙적으로 금지되어야 합니다. 경찰 등 국가기관이 전국민의 얼굴을 인식하는 일은 결코 있어서는 안 될 것입니다.

【참고】 유럽 기본권청(2020)

입법자는 기본권의 모든 범주를 포괄하는 의무적 영향 평가의 실시를 고려해야 한다. 이는 민간부문과 공공부문 모두에 적용되어야 하며, 인공지능 시스템이 사용되기 전에 이루어져야 한다 ... 영향 평가는 ... 정기적이고 반복적으로 실시되어야 한다. 평가는 투명하게 실시되어야 한다. 그 결과와 권고사항은 가능한 범위까지 공개되어야 한다. 기업과 공공기관들은 기본권에 미치는 영향을 완전하게 평가하는 데 필요한 정보를 수집할 수 있어야 한다.

【참고】 유럽 집행위원회(2021²⁷)

고위험 인공지능 시스템은 시장에 출시되기 전에 엄격한 의무가 부과된다.

- 적절한 위험 평가 및 완화 시스템
- 위험과 차별적 결과를 최소화하기 위해 시스템에 공급하는 데이터 세트의 고품질 보장
- 결과의 추적성을 보장하기 위한 활동 기록
- 시스템에 대하여 필요한 모든 정보 및 그 준수 여부를 평가하는 기관의 목적 수행에 필요한 모든 정보를 제공할 수 있는 상세한 문서화
- 사용자에게 대해 명확하고 적절한 정보 제공
- 리스크를 최소화하기 위해 적절한 인적 감독 조치
- 높은 수준의 견고성, 보안성 및 정확성

인공지능으로 피해를 입거나 인권을 침해당한 사람에 대하여 권리구제가 보장되어야 합니다.

공공기관 또는 민간기업이 개발, 도입, 사용하는 인공지능 시스템으로 인하여 피해를 입거나 인권을 침해당한 사람에게는 국가기관의 구제수단에 대한 접근이 보장되어야 합니다. 피해를 입거나 인권을 침해당한 당사자가 인공지능 제품과 시스템에 대응할 수 있으려면, 인공지능이 언제 어떻게 사용되었는지, 어디서 어떻게 이에 대한 이의를 제기할 수 있는지 알 수 있어야 합니다. 또한 인공지능 제품과 서비스는 우리 헌법과 인권규범이 보호하고 있는 적법절차의 권리를 보장해야 합니다.

따라서 인공지능 사용 기관은 일반적으로 소비자와 이용자에게 인공지능 사용 여부 및 그 기능, 인공지능을 사용한 의사결정에 대한 사항과 이의제기 방법에 대하여 설명하여야 합니다. 특히 공공기관이나 민간기업이 인공지능을 사용하여 당사자에 대해 법적 또는 상당한 영향을 미치는 의사결정을 내리는데 그 결정에 대하여 설명하지 못한다면 이 시스템을 사용해서는 안 될 것입니다.

금융서비스나 채용 등 개인을 평가하거나, 거부하는 등 사람에 대한 중대한 의사결정이 인공지능을 사용하여 자동적으로 내려졌을 경우, 그 대상이 된 정보주체는 그 사실을 통지받고, 인적 개입, 의사 표현, 그 결정의 이유 설명, 이의 제기, 안전조치 등을 보장받을 권리가 있습니다. 국가가 인공지능을 이용하여 국민에게 불이익한 행정 또는 형사상 처분이나 의사결정을 내리는 경우 국민은 그에 대한 사전적인 청문과 의견 진술, 문서 열람, 결정 이유 요구권 등 헌법상 적법절차를 보장받아야 합니다. 이 권리는 공공기관이 민간에서 조달한 시스템이나 서비스를 사용하는 경우에도 똑같이 보장되어야 합니다.

²⁷ “Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence”, European Commission(2021. 4. 21).

【참고】 유엔 의사표현의자유 특별보고관(2018)²⁸

개인 이용자들은 인공지능 시스템의 반인권적 영향에 대한 구제 수단에 접근 할 수 있어야 한다. 기업들은 인공지능 처리시스템에 대해 표출되는 모든 이용자의 불만사항 및 이의제기에 적시에 대응하기 위해 사람에 의한 검사 및 구제 시스템을 두어야 한다. 인공지능 시스템에 불만이 접수되고 구제가 요청된 빈도에 대한 데이터 뿐 아니라, 실제 이용되는 구제수단의 종류와 효과성에 대해서도 주기적으로 공개 되어야 한다.

【참고】 유럽평의회 인권위원장(2019)

인공지능 시스템이 머신러닝이나 이와 유사한 기법으로 특정한 인적 개입에서 독자적으로 의사결정을 할 수 있는 환경이라 하더라도, 인공지능 시스템은 항상 인적 통제를 유지해야 한다. 각국은 인공지능 시스템의 다양한 생애주기 단계에서 발생할 수 있는 인권 침해에 대한 명확한 책임의 선을 수립해야 한다. 인권을 침해하는 조치가 인간 책임자 또는 운영자에 의해 직접 지시되지 않은 경우라 하더라도, 인공지능 시스템의 개발, 도입 또는 사용으로 발생하는 인권 침해에 대한 책임과 책무는 항상 자연인 또는 법인 측에 있다.

공공기관 혹은 민간기업 인공지능 시스템의 개발, 도입, 혹은 사용으로부터 인권 침해 피해를 주장하는 누구라도 소관 국가기관에서 효과적인 구제수단을 제공 받을 수 있어야 한다. 나아가 각국은 불투명한 방식으로 당사자의 인지 없이 오로지 혹은 상당히 인공지능 시스템에 의해 산출된 결과물의 대상이 되었다고 의심하는 사람들에게 효과적인 구제수단에 대한 접근을 제공해야 한다.

효과적인 구제수단에는 인공지능 시스템의 개발, 도입 또는 사용으로 인한 피해에 대하여 신속하고 적절한 배상 및 보상이 포함되어 있어야 하며, 민사적, 행정법적 또는 적절한 경우 형사적 조치를 포함할 수 있다.

국가인권기구는 각각의 권한에 따른 자신의 결정을 제시함으로써 이러한 보상의 근거가 될 수 있다.

각국은 충분한 인적 개입 없이 이루어진 자동화된 의사결정에 기반하여 중대한 영향을 미칠 수 있는 의사결정의 대상이 되지 않을 권리를 개인에게 부여해야 한다. 적어도 개인은 이러한 자동화된 의사결정에서 인적 개입을 확보할 수 있어야 하며, 그러한 결정이 실시되기 전에 자신의 의견을 피력할 기회를 부여받아야 한다.

각국은 개인이 인공지능 시스템에 의해 야기된 인권 침해의 피해자라는 자신의 주장을 입증하는 것과 관련된 피고 또는 제3자 소유의 정보에 접근할 수 있도록 보장해야 한다. 여기에는 적절한 경우 훈련 및 테스트 데이터, 인공지능 시스템이 어떻게 사용되었는지에 대한 정보, 인공지능 시스템이 특정 권고, 의사결정 또는 예측에 어떻게 이르렀는지에 대해 유의미하고 이해가능한 정보도 포함되며, 인공지능 시스템의 결과물이 어떻게 해석되고 실행되었는지에 대한 상세정보도 포함된다.

소관 국가기관이 인공지능 시스템의 개발, 도입, 혹은 사용으로 인해 야기된 인권 침해 문제를 검토할 때, 인공지능 시스템이 현출하는 ‘객관성의 유혹(allure of objectivity)’에 적절한 회의적인 태도를 견지해야 하며, 인권 유린으로 난관에 처한 개인이 문제의 조치에 대한 책임자들에 비해 높은 수준의 증거에 구속받지 않도록 보장해야 한다.

오늘 선언에 참여한 시민사회, 노동, 인권단체들은 인공지능이 미치는 영향과 위험성으로부터 사람들의 인권과 안전, 민주주의를 보장하기 위하여 정부와 국회가 빠른 시일내 위와 같은 조치를 보장하는 입법 절차에 착수할 것을 다시 한번 촉구합니다. (끝)

2021. 5. 24.

²⁸ “Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression”. 유엔문서번호 A/73/348(2018. 8. 29).

가짜뉴스체크센터, 건강권 실현을 위한 보건의료단체연합(건강권 실현을 위한 행동하는 간호사회, 건강사회를 위한 약사회, 건강사회를 위한 치과의사회, 노동건강연대, 인도주의실천의사협의회, 참의료실현청년한의사회), 경제정의실천시민연합, 공익변호사와 함께하는 동행, 공익인권법재단 공감, 공익인권변호사모임 희망을만드는법, 군인권센터, 김포장애인야학, 김포장애인자립생활센터, 대구참여연대, 대전참여자치시민연대, 데이터민주주의포럼, 라이더유니온, 매체비평우리 스스로, 무상의료운동본부, 미디어기독연대, 민주노총 공공운수노조, 민주사회를 위한 변호사모임 (교육청소년위원회, 노동위원회, 디지털정보위원회, 소수자인권위원회, 아동인권위원회, 여성인권위원회), 민주언론시민연합, 부산참여연대, 빈곤사회연대, 사단법인 경기장애인자립생활센터협의회, 사단법인 두루, 사단법인 여수시민협, 사단법인 정보인권연구소, 사단법인 제주참여환경연대, 사회적협동조합 빠띠, 상록수장애인자립생활센터, 생명안전시민넷, 서울YMCA 시청자시민운동본부, 성남참여자치시민연대, 소비자시민모임, 손잡고, 신대승네트워크, 언론개혁시민연대, 연구공동체 건강과 대안, 울산시민연대, 익산참여연대, 인권운동공간 활, 인천평화복지연대, 장애인권법센터, 장애해방열사_단, 전국교직원노동조합, 전국민주노동조합총연맹, 전쟁없는세상, 정태수열사추모사업회, 진보네트워크센터, 진보3.0, 차별금지법제정연대, 참여연대, 참여와자치를 위한 춘천시민연대, 참여자치21(광주), 충남참여자치지역운동연대, 충북참여자치시민연대, 투명사회를 위한 정보공개센터, 팔레스타인평화연대, 평등과 연대로! 인권운동더하기(거창평화인권예술제위원회, 광주인권지기 활짝, 구속노동자후원회, 국제민주연대, 노동인권실현을위한노무사모임, 다산인권센터, 대학성소수자모임연대 QUV, 대한불교조계종 사회노동위원회, 문화연대, 민족민주열사.희생자추모(기념)단체연대회의, 민주사회를 위한 변호사 모임, 민주주의법학연구회, 민주화실천가족운동협의회, 반도체노동자의 건강과 인권지킴이 반올림, 불교인권위원회, 빈곤과 차별에 저항하는 인권운동연대, 사회진보연대, 삼성노동인권지킴이, 새사회연대, 서교인문사회연구실, 서울인권영화제, 성적소수문화인권연대 연분홍치마, 실천불교전국승가회, 어린이책시민연대, 예수회 인권연대연구센터, 외국인이주.노동운동협의회, 울산인권운동연대, 원불교인권위원회, 인권교육센터 들, 인권교육온다, 인권연극제, 인권운동네트워크 바람, 인권운동사랑방, 장애물없는생활환경시민연대, 장애여성공감, 장애와인권발바닥행동, 전국불안정노동철폐연대, 전국장애인차별철폐연대, 전북평화와인권연대, 제주평화인권연구소 왓, 진보네트워크센터, 청소년인권행동 아수나로, 천주교인권위원회, 한국게이인권운동단체 친구사이, 한국기독교교회협의회(NCCK) 인권센터, 한국비정규노동센터, 한국성적소수자문화인권센터, 한국성폭력상담소, 한국청소년청년감염인커뮤니티 알, 행동하는성소수자인권연대, HIV/AIDS인권연대나누리+, 표현의자유와언론탄압공동대책위원회, 피스모모, 한국방송기술인연합회, 한국여성단체연합, 한국여성민우회, 한국장애인자립생활센터협의회, 형명재단 (총 120개 단체)