입법토론회

국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 입법 방향

일시 | **2024.7.11.**(**목**) 오전 **10**시
장소 | 국회의원회관 제5간담회실

주최

과학기술정보방송통신위원회 국회의원 김우영·노종면·박민규·이정헌 ·이훈기·황정아
정무위원회 국회의원 김남근·김용만·이강일·조승래
문화체육관광위원회 국회의원 민형배
외교통일위원회 국회의원 차지호
국토교통위원회 국회의원 윤종군·정준호

건강권실현을위한보건의료단체연합, 광주인권지기 활짝, 문화연대 기술미디어문화위원회, 민주사회를 위한 변호사 모임 디지털정보위원회, 사단법인 정보인권연구소, 서울YMCA 시민중계실, 언론개혁시민연대, 연구공동체 건강과대안, 전북평화와인권연대, 진보네트워크센터, 참여연대, 홈리스행동

순 서

10:00 ~ 10:10	인사	**공동주최** 국회의원
	사회	**한상희** (참여연대 공동대표)
10:10 ~ 10:40	발제	**인공지능 규범의 국제적 흐름과 시사점** | **오병일** (진보네트워크센터 대표)
		22대 국회 인공지능법 입법방향 | **유승익** (한동대학교 교수)
10:40 ~ 11:50	토론	**차지호** (더불어민주당 국회의원)
		김병욱 (민주사회를위한변호사모임 디지털정보위원회 변호사)
		김영규 (한국인터넷기업협회 정책1실 실장)
		과학기술정보통신부 | **남철기** (인공지능기반정책과 과장)
		개인정보보호위원회 | **김직동** (개인정보정책과 과장)
		공정거래위원회 | **이준헌** (시장감시정책과 과장)
		국가인권위원회 | **이진석** (인권정책과 사무관)
11:50 ~ 12:00	**플로어 토론 및 참석자 전체 토론**	

인 사 말

김남근 | 더불어민주당 국회의원 (정무위원회)

반갑습니다. 더불어민주당 성북을 국회의원 김남근입니다.

「국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 입법방향」토론회를 공동주최하게 되어 영광입니다. 토론회의 좌장을 맡아주신 참여연대 한상희 공동대표님과 발제를 맡아주신 오병일 진보네트워크센터 대표님, 유승익 한동대학교 교수님, 토론자로 나서주신 김병욱 민주사회를 위한 변호사모임 디지털정보위원회 위원님, 김영규 한국인터넷기업협회정책실장님, 남철기 과학기술정보통신부 인공지능기반정책과 과장님, 김직동 개인정보보호위원회 개인정보정책과 과장님, 이준헌 공정거래위원회 시장감시정책과 과장님, 이진석 국가인권위원회 인권정책과 사무관님께도 감사를 드립니다.

21대 국회에서 인공지능에 대한 입법논의는 가장 주목을 끌었다 해도 과언이 아닐 것입니다. 하지만 사안의 중요도에 비례해 충분한 숙의의 과정을 거쳤는가에 대해서는

의문이 남습니다. 21대 막판에는 마치 국회의 잘못이 인공지능법안을 통과시키지 못한 데에 있는 마냥 졸속이라도 상관없으니 통과시키라는 분위기마저 형성되었습니다. 많은 사람에게 영향을 미칠 미래의 핵심 기술에 진흥과 촉진이라는 한 방향의 의견만이 영향을 미치는 것을 경계해야 합니다. 시민의 권리가 여전히 문제제기의 차원에 머무르는 사회를, 이제는 벗어나 다음 단계로 나아가야 합니다.

불필요한 규제냐, 위험성에 대한 통제냐 라는 이분법적인 논쟁은 너무 낡았습니다. 유럽은 AI ACT를 통해 규범화에 나서고 있고 미국 또한 행정명령으로 좇아가고 있습니다. 글로벌 스탠다드는 인공지능을 얼마나 폭넓게 정의할 것인가, 위험성을 어떻게 제어 할 것인가, 감독을 위한 거버넌스를 어떻게 구축할 것인가이지 토종 산업을 어떻게 육성할 것인가, 규제에 얼마나 자율성을 부과할 것인가가 아닙니다.

이미 미국 등 해외에서는 인공지능을 둘러싼 소송전이 격화되고 있습니다. 기술을 둘러싸고 정부와 기업, 기업 간, 기업과 시민단체 간의 갈등은 더욱 첨예해질 것이며, 이는 곧 우리나라에서도 펼쳐질 가까운 미래가 될 것입니다. 이러한 상황들이 닥쳐왔을 때 민주당, 나아가 한국 정치는 갈등의 중재자로 당당히 나설 수 있는 역량을 보여줘야 합니다.

토론회를 공동주최하는 과학기술정보방송통신위원회, 정무위원회, 문화체육관광위원회, 외교통일위원회, 국토교통위원회 등 많은 상임위원회 소속의 국회의원들은 인공지능 기술이 우리사회에 얼마나 크게 영향을 미칠지를 증언하는 증인과도 같습니다. 민주당은 대선공약으로 인공지능기본법을 내세웠으며, 차지호 의원, 황정아 의원 등 기술분야의 전문가도 보유하고 있습니다. 충분한 논의의 장이 주어진다면, 이번에야말로 정치가 기술혁신의 발목을 잡는다는 오명을 벗을 수 있을 것입니다. ▢

인 사 말

김용만 | 더불어민주당 국회의원 (정무위원회)

안녕하십니까, 더불어민주당 하남시(을) 국회의원 김용만입니다.

「국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 입법 방향」 입법토론회에 참석해주신 여러분 환영합니다. 함께 토론회를 준비해주신 동료·선배 의원님들과 발제를 맡아주신 오병일 진보네트워크센터 대표님, 유승익 한동대학교 교수님, 좌장을 맡아주신 한상희 참여연대 공동대표님 그리고 각 부처와 시민사회 담당자 분들게 진심으로 감사드립니다.

인공지능이 발전하는 속력은 인식이 따라가지 못할 만큼 빠릅니다. `22년, 한 미술대회에서 인공지능이 만들어낸 그림이 미술대회 1위를 차지하였습니다. 영감의 영역에 인공지능이 도달했음을 시사한 사건입니다. 다음 해, Chat-GPT가 등장하여 평범한 우리 삶에도 인공지능이 가깝게 안착했고, 곧이어 복잡한 코딩마저 인공지능이 척척 해

나가게 되었습니다. 상상이 실현됨에 따라 가까운 미래에 대한 기대감은 더욱 커져만 갑니다.

하지만, 인공지능의 비약적인 발전 이면에는 수많은 정보 약탈이 있었습니다. 모 글로벌 기업은 자사의 인공지능 훈련을 위해 사용자 데이터를 동의 없이 수집하여 소송 중에 있고, 여러 생성형 인공지능이 타인의 창작물을 무단 학습하여 표절 작품을 양산해 저작권 피해 사례가 걷잡을 수 없이 늘어나고 있습니다. 국내 대기업 삼성전자 역시 생성형 인공지능을 이용했다가 소스 코드 데이터가 유출된 사례도 있습니다.

윤리와 철학에 대한 과제도 남아있습니다. '자율주행차량에 탑재된 인공지능은 탑승자와 다수의 보행자 중, 누구의 안전을 우선해야 하는가?'란 문제에 답을 내려야 합니다.모 SNS의 알고리즘은 청소년에게 자극적인 콘텐츠를 시청하지 않게끔 설정할 수 있음에도 회사 이익에 충실하도록 알고리즘이 설정되어 있어, 되려 중독성을 강화하고 있다는 내부 폭로가 있었습니다. 이처럼 아무리 좋은 성능을 가진 인공지능이라도 판단의 기준을 인간이 정한다면 좋은 결과로 이어지지 않을 수 있습니다.

전 세계는 인공지능 무법지대입니다. 그 속에서 인공지능은 초법적인 발전을 이루고 있습니다. 피해 회복이 어려운 데이터 영역의 특성상, 빠른 대처와 예방이 중요합니다. 인공지능에 대한 많은 논의가 진행되어 온 만큼, 22대 국회에서는 결과를 내야 할 때입니다.

동시에 우리의 인공지능 기술이 세계 무대에서 제약으로 작동하지 않게끔 해야 합니다. 우리 법만으로는 탈국가 글로벌 기업을 규제하기 어렵기에, 그들과 속도 경쟁을 할 수 있게끔, 지원과 진흥이라는 추진력도 달아 주어야 합니다. 오늘 논의된 내용을 바탕으로, 인권 윤리와 산업 발전이 함께 어우러진 탄탄한 인공지능 기본법이 만들어지길 바랍니다. 감사합니다. ▢

인 사 말

김우영 | 더불어민주당 국회의원 (과학기술정보방송통신위원회)

안녕하십니까. 더불어민주당 은평구을 국회의원 김우영입니다.

먼저, 「국민의 안전, 인권 및 민주주의와 AI 공존을 위한 제정 방향」 토론회를 공동주최하게 되어 영광입니다.

최근 인공지능 기술이 빠르게 발전함에 따라 미국을 선두로 세계 각국에서도 대규모 투자가 이뤄지고 있습니다. 또한, 각국 정부들은 국민의 삶 전반에 혁신적인 변화가 올 것을 기대하며 다양한 정책과 법률 제정에 앞장서고 있습니다.

대한민국 국회 과학기술정보방송통신위원회에서도 지난 6월 19일, 조인철 의원님께서 「인공지능 산업 육성 및 신뢰 확보에 관한 법률안」을 대표발의 하였습니다. 또한 6월 17일, 국토교통부에서도 「AI · 헬스케어 등 新산업 투자 기반 지원을 도모하기 위한 리츠 활성화 방안」 을 발표했습니다.

이 외에도 관련 법안들이 국회에서도 발의되는 가운데, 우리가 주목해야 할 점은 국민의 안전과 AI 공존이 함께 가는 입법 방향을 찾는 것입니다.

AI 기본법의 입법은 혁신과 기술 발전을 촉진하지만, 공정하고 도덕적이며 사회적 가치, 인권 보호 등도 함께 잘 반영될 수 있도록 올바른 방향을 제시해야 합니다. 개인 자료수집과 활용·공유에 관한 규제, 개인정보 보호법 강화, 데이터의 편향성으로 인한 편견 및 차별 등을 고려한 입법적 방안을 모색하는 구체적인 논의와 체계 마련이 시급합니다.

오늘 입법토론회가 AI 기술 발전도 촉진되는 동시에 국민의 인권, 안전도 함께 보호받을 수 있는 좋은 결과를 도출할 수 있는 생산적이고 의미 있는 토론회가 되길 기대합니다.

뜻깊은 자리를 함께 만들어주신 노종면, 박민규, 이정헌, 이훈기, 황정아, 김남근, 김용만, 이강일, 조승래, 민형배, 차지호, 윤종군, 정준호 국회의원님과 건강권실현을위한보건의료단체연합, 광주인권지기활짝, 문화연대기술미디어문화위원회, 민주사회를위한변호사모임, 디지털정보위원회, 사단법인 정보인권연구소, 서울YMCA 시민중계실, 언론개혁 시민연대, 연구공동체 건강과대안, 전북평화와인권연대, 진보네트워크센터, 참여연대, 홈리스행동 관계자 여러분께 깊은 감사를 드립니다.

저도 과학기술정보방송통신위원회 위원으로서 국민의 안전과 인권, 민주주의가 함께 갈 수 있도록 핵심적인 역할을 하며, 법과 제도적인 개선에 기여할 수 있도록 최선을 다하겠습니다.

참석해주신 모든 분께 다시 한번 진심으로 감사드립니다. ▢

인 사 말

노종면 | 더불어민주당 국회의원 (과학기술정보방송통신위원회)

안녕하세요. 더불어민주당 인천부평구갑 국회의원 노종면입니다.

‘국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 AI기본법 제정 방향 국회 토론회’를 함께 주최하게 되어 기쁩니다. 행사를 준비한 의원실과 시민단체 관계자 여러분 고생 많으셨습니다. 좌장과 발제자, 토론자로 참석해주신 모든 분들께도 감사드립니다.

이제 우리는 AI를 이기는 법이 아닌 AI와 공존하는 법을 찾아야 합니다.

우리나라에 AI기술이 대중적으로 알려진 순간은 조금 특별합니다. 2016년 3월, 이세돌 9단이 알파고를 상대로 한 대국에서 거둔 승리는 여전히 전율로 남아있습니다. 이세돌 9단의 ‘1승’은 우리 국민 뿐만 아니라 모든 인류의 자랑이었습니다. 하지만 아마도 이 ‘1승’은 알파고를 상대로 거둔 마지막 승리가 될 것입니다. 지난 몇 년간 AI기술은 인간이 따라잡지 못할 정도로 급속하게 발전했기 때문입니다.

국제사회는 AI시대 전환에 발 빠르게 대응하고 있습니다.

많은 국가가 AI기술이 필요로 하는 전력 수요를 감당할 수 있는 발전 방식을 찾고 있습니다. AI시대 핵심 부품과 원료를 확보하기 위한 국가 간 수싸움도 치열합니다. AI 글로벌 기업에는 연일 막대한 자본이 흘러 들어가고 있습니다.

AI기술의 부작용을 관리하려는 시도도 눈여겨 봐야합니다.

머신러닝에 기반한 AI기술은 학습 데이터의 양과 질에 따라 오류가 발생할 가능성을 내포하고 있습니다. 최첨단 AI기술이 범죄에 악용될 우려도 상존합니다. AI기술의 오류가 편향과 극단의 결과로 이어지지 않도록 제어해야 합니다. 개인정보 유출, 인권 침해, 딥페이크 등을 막을 합리적인 규제도 필요합니다.

'AI 기본법'에 대한 치열한 토론을 이어가겠습니다.

미래 산업에 대한 국가적 지원을 명시하고, 명확한 기준을 통해 AI 시장의 예측가능성을 확보해야 합니다. AI 기술로부터 국민의 삶을 안전하게 보호하는 대책도 세워야 합니다. 제22대 국회에서 정부와 정치권, 산업계와 시민사회의 의견을 조율한 '잘 만든 AI기본법'이 완성되길 바랍니다.

국회 과학기술정보방송통신위원회 위원으로서 여러분과 함께 머리를 맞대고 최선의 결과를 만들어 가겠습니다. 감사합니다. ▢

인 사 말

박민규 | 더불어민주당 국회의원 (과학기술정보방송통신위원회)

안녕하십니까, 더불어민주당 서울 관악구 갑 국회의원 박민규입니다.

먼저 이 자리를 빛내 주신 모든 분들께 감사의 말씀 드립니다. 특히 오늘의 토론회를 주관한 참여연대 한상희 공동대표님과 발제자로 참여하신 오병일 진보네트워크센터 대표님, 유승익 한동대 교수님, 그리고 토론에 참여해 주신 여러 전문가 분들께 깊은 감사의 인사 드립니다.

오늘 우리는 인공지능과 관련된 중요한 입법 방향을 논의하기 위해 모였습니다. 우리 사회는 기술의 급속한 발전과 함께 인공지능이 우리의 일상과 다양한 분야에 깊숙이 스며들고 있습니다. 인공지능은 경제적, 사회적 발전을 이끄는 중요한 원동력임과 동시에, 우리의 인권과 안전에 중대한 영향을 미칠 수 있는 잠재력을 가지고 있습니다. 따라서 기술의 발전이 우리의 삶을 더욱 윤택하게 만들기 위해서는 그에 상응하는 적절한 법적, 윤리적 기준이 반드시 마련되어야 합니다.

오늘 토론회에서는 인공지능의 국제적 규범 흐름과 우리나라의 입법 방향에 대해 심도 깊은 논의가 이루어질 것입니다. 유럽의 AI Act 와 미국의 행정명령 등 선진국들의 사례는 우리에게 많은 시사점을 제공하고 있습니다. 이들 국가들은 인공지능의 혁신을 촉진하면서도 그로 인한 위험을 관리하기 위해 다양한 법적 장치를 마련하고 있습니다. 이러한 움직임을 보며 우리도 글로벌 스탠다드에 맞추어 나가야 할 것입니다.

특히, 인공지능의 발전이 초래할 수 있는 다양한 윤리적, 사회적 문제들을 사전에 예방하고, 국민의 안전과 인권을 보호할 수 있는 체계를 구축하는 것이 중요합니다. 이는 단순히 기술 발전을 저해하는 것이 아니라, 기술이 인간에게 이로움을 줄 수 있도록 하는 중요한 과정입니다. 오늘 이 자리에서 논의된 내용들이 향후 정책과 법률 제정에 큰 도움이 되길 바랍니다.

우리 더불어민주당은 이번 토론회를 통해 인공지능 관련 법률을 더욱 발전시키고, 국민들이 안심하고 기술의 혜택을 누릴 수 있도록 최선을 다할 것입니다. 또한, 각계각층의 의견을 수렴하여 더욱 포괄적이고 실효성 있는 법안을 마련하기 위해 끊임없이 노력할 것입니다.

오늘 토론회가 인공지능의 안전한 발전을 위한 중요한 밑거름이 되기를 바라며, 참석해 주신 모든 분들께 다시 한 번 감사드립니다. 여러분의 지혜와 통찰이 대한민국의 미래를 밝히는 데 큰 힘이 될 것입니다.

감사합니다. ▢

인 사 말

윤종군 | 더불어민주당 국회의원 (국토교통위원회)

반갑습니다. 더불어민주당 안성시 국회의원 윤종군입니다.

「국민의 안전, 인권 및 민주주의와 AI 공존을 위한 제정 방향」 입법 토론회 개최를 진심으로 환영합니다.

이번 토론회를 함께 준비해 주신 김남근·김용만·김우영·노종면·민형배·박민규·이강일·이정헌·이훈기·정준호·조승래·차지호·황정아 의원님과 건강권 실현을 위한 보건의료단체연합회, 광주인권지기 활짝, 문화연대 기술미디어 문화위원회, 민주사회를 위한 변호사 모임 디지털정보위원회, 사단법인 정보인권연구소, 서울YMCA 시민중계실, 언론개혁시민연대, 연구공동체 건강과대안, 전북 평화와 인권연대, 진보네티워크센터, 참여연대, 홈리스행동 관계자 여러분, 귀중한 시간을 내시어 발제와 토론에 참여해 주시는 많은 분께 깊이 감사드립니다.

미래는 AI의 시대입니다. 증기기관, 대량생산 체계, 컴퓨터와 인터넷으로 대표되는 IT 혁명에 이어 인공지능 기반의 기술혁명은 인류의 미래를 송두리째 바꿀 것입니다.

겨울이 오기 전에 월동 준비를 해야 하는 것처럼 개인정보 유출과 사생활 침해 등 인공지능 발전에 따른 각종 위험을 선제적으로 대비할 때만이 기술 발전에 따른 장밋빛 미래를 누릴 수 있게 될 것입니다.

이번 토론회가 기술 발전에 따른 부작용을 줄이고 인권과 안전 등을 보호하는 가운데 인공지능 기술의 혜택을 온전히 누릴 수 있는 'AI 기본법'을 제정하는 뜻깊은 출발점이 되길 바랍니다.

다시 한번 오늘 토론회 개최를 진심으로 축하합니다. 함께 하신 모든 분께서 건강하고 행복하시길 기원합니다. 감사합니다. ▢

인 사 말

이강일 | 더불어민주당 국회의원 (정무위원회)

반갑습니다. 더불어민주당 국회의원 이강일(청주상당)입니다.

<AI 기본법 제정 방향 토론회>를 공동주최하게 되어 매우 기쁘게 생각합니다.

오늘 뜻깊은 토론회가 마련될 수 있도록 애써주신 참여연대 한상희 공동대표님과 오병일 진보네트워크센터 대표님, 유승익 한동대 교수님, 그리고 토론에 참여해 주신 여러 전문가분께 깊은 감사의 마음을 전합니다.

건강권실현을 위한 보건의료 단체연합, 광주인권지기 활짝, 문화연대 기술미디어문화위원회, 민주사회를 위한 변호사 모임 디지털정보위원회, 사단법인 정보인권연구소, 서울YMCA 시민중계실, 언론개혁시민연대, 연구공동체 건강과대안, 전북평화와인권연대, 진보네트워크센터, 참여연대, 홈리스행동 등 시민사회를 대표하여 함께해주셔서 감사하다는 인사를 드립니다.

유럽연합의 포괄적 규제법인 AI ACT는 2024년 5월 21일 최종 승인 이후 2026년 시행을 앞두고 있으며, 미국 역시 인공지능을 활용하는 정부 기관들에 대하여 인공지능 위험 방지 장치를 의무화하도록 하는 등 인공지능 위험에 대비하기 위한 규제를 마련해 나가고 있습니다.

한국 AI 기본법 제정도 이와 같은 세계적 추세를 참고하여 AI가 가져오는 편익과 효율 이면의 개인정보와 프라이버시 침해, 소비자 권리 침해, 환경문제 등의 사실에 주목하며 적절한 규제를 통해 국민 안전이 모색되어져야 할 것입니다.

인공지능이 야기할 수 있는 위험에 대비할 수 있는 명실상부한 인공지능 기본법이 제정되어 신뢰할 수 있고 안전한 AI 사회가 되도록 오늘 이 자리가 현안을 풀어나가는 계기가 마련되는 뜻깊은 자리가 되길 바랍니다.

다시 한번 이번 간담회가 열릴 수 있게 도와주시고 자리해 주신 모든 분께 깊은 감사의 마음을 전합니다. 감사합니다. ▢

인 사 말

이훈기 | 더불어민주당 국회의원 (과학기술정보방송통신위원회)

안녕하십니까. 더불어민주당 인천 남동을 국회의원 이훈기입니다.

인간중심 인공지능 발전을 위한 <국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 입법 방향> 토론회에 함께해주신 여러분 감사합니다. 발제를 맡아주시는 오병일 진보넷 대표님과 유승익 한동대 교수님 그리고 토론으로 함께해주시는 민변 김병욱 변호사, 인터넷기업협회 김영규 실장, 과학기술정보통신부 남철기 과장, 개인정보위원회 김직동 과장, 공정거래위원회 이준헌 과장, 국가인권위원회 이진석 사무관께도 깊은 감사의 말씀 드립니다.

전 세계의 인공지능 개발과 투자의 파도가 거셉니다. open AI의 챗GPT와 구글의 Gemini(제미나이) 등 미국의 기업들이 AI 산업을 선도하는 가운데, 중국, 영국, 이스라엘, 일본 등 주요국 역시 인공지능 개발에 매진하고 있습니다. 우리나라 역시 문재인

정부 이후, 인공지능 개발과 투자를 위한 정책 방안 논의가 이어졌습니다. 하지만 그와 동시에 필요한 것이 어떻게 인간 중심 인공지능 개발을 할 수 있느냐입니다.

인공지능 기술은 우리가 인식하지 못한 사이에 이미 생활 속에 자리 잡고 있습니다. 가전제품, 스마트폰, 자동차, 산업로봇, 인터넷과 교육 현장 등 하나하나 열거할 수 없을 정도로 이미 우리와 함께하고 있습니다. 이제는 단순히 인공지능 기술의 발전뿐만 아니라 인간과 공존 그리고 인간에 의한 인공지능 통제 기술도 함께 만들어야 하는 상황에 이르렀습니다.

통제할 수 있는 인공지능 개발과 인권을 보호할 수 있는 기술 개발이 이뤄질 수 있도록 충분하고 다양한 논의가 필요합니다. 아직 우리나라는 초기 단계의 인공지능 기술 개발이 이뤄지고 있지만, 이미 인공지능은 우리 일상과 함께하고 있습니다. 코로나 19 이후 사회 전 분야의 디지털 전환이 이뤄짐에 따라 다양한 인공지능 기술과 서비스를 기대할 수 있게 되었습니다. 면접은 물론, 유튜브 영상 추천 알고리즘, 택시 배차 알고리즘 등 다양한 곳에서 인공지능에 의한 결정이 우리 일상과 맞닿아 있습니다.

어떻게 하면 국민의 안전과 인권을 고려한 인공지능 기술 개발이 이뤄질 수 있을지, 다양한 단위에서 고민과 소통이 필요할 것입니다. 게다가 올해 초 EU에서 인공지능법이 통과되며, 세계 각국 역시 다양한 인공지능 규율에 박차를 가하고 있어, 우리나라만의 인공지능 기본법을 만들기 위한 논의가 이뤄져야 합니다. 그런 점에서 오늘 입법토론회가 민주적인 인공지능 발전을 위한 다양한 논의의 장이 되길 기대합니다.

토론회를 통해 인공지능 기술에 대한 막연한 두려움보다는 인간이 기술의 중심이 되는 인공지능 기술 개발의 밑거름이 만들어지길 바랍니다. 아울러 산업계와 정부는 물론, 시민사회계까지 다양한 단위에서 모인 토론회인 만큼, 우리 국민이 수용할 수 있는 입법 대안을 모색하는 건설적인 시간이길 기대합니다. 다시 한번 오늘 토론회 발제자와 토론자 여러분께 감사드리며, 앞으로도 우리 인공지능 기술의 민주적인 발전과 인권을 지키는 기술 개발을 만드는 데 힘쓰겠습니다.

감사합니다. ▢

인 사 말

조승래 | 더불어민주당 국회의원 (정무위원회)

안녕하십니까?

더불어민주당 대전 유성구갑 국회의원 조승래입니다.

‘국민의 안전, 인권 및 민주주의와 AI의 공존을 위한 AI기본법 제정’입법토론회를 찾아주신 내외 귀빈 여러분을 환영합니다. 오늘 인공지능 시대의 새로운 도약을 위한 첫 걸음에 함께 하게 되어 매우 기쁘게 생각합니다.

22대 국회는 인공지능이라는 새로운 기술의 발전과 함께 찾아온 기회와 도전을 마주하고 있습니다. 인공지능은 우리 사회의 효율성과 편의성을 획기적으로 높여줄 잠재력을 지니고 있지만, 동시에 예상치 못한 위험과 부작용을 야기할 수 있다는 우려 또한 존재합니다.

이러한 상황에서 우리는 인공지능의 긍정적인 면을 극대화하고 부정적인 면을 최소화할 수 있는 균형 잡힌 접근 방식을 모색해야 합니다. 21대 국회에서 논의되었던 인공지능 관련 법안들은 주로 인공지능 산업 육성에 초점을 맞추어져 있어 끝내 매듭을 짓지 못했습니다. 인공지능이 일으킬 수 있는 위험으로부터 국민의 생명과 인권, 그리고 민주주의의 보호가 담보되어야만 인공지능 산업 육성을 위한 법제화도 가능할 것입니다.

유럽연합의 AI ACT, 미국의 행정명령 등 세계 주요 국가들은 이미 인공지능 규제 마련에 적극적으로 나서고 있습니다. 우리도 이러한 국제적인 흐름에 발맞춰 인공지능의 혁신을 지원하면서도, 그로 인한 문제들을 예방하고 해결하는 방안을 종합적으로 검토해야 합니다.

22대 국회에서 인공지능 시대의 새로운 규범을 정립하는 중요한 임무를 수행해야 합니다. 인공지능 기술 발전을 위한 지원 정책과 함께, 인공지능의 윤리적 사용, 투명성 확보, 편향과 차별 방지, 개인정보 보호 등 다양한 측면을 고려한 종합적인 인공지능 기본법을 제정해야 합니다.

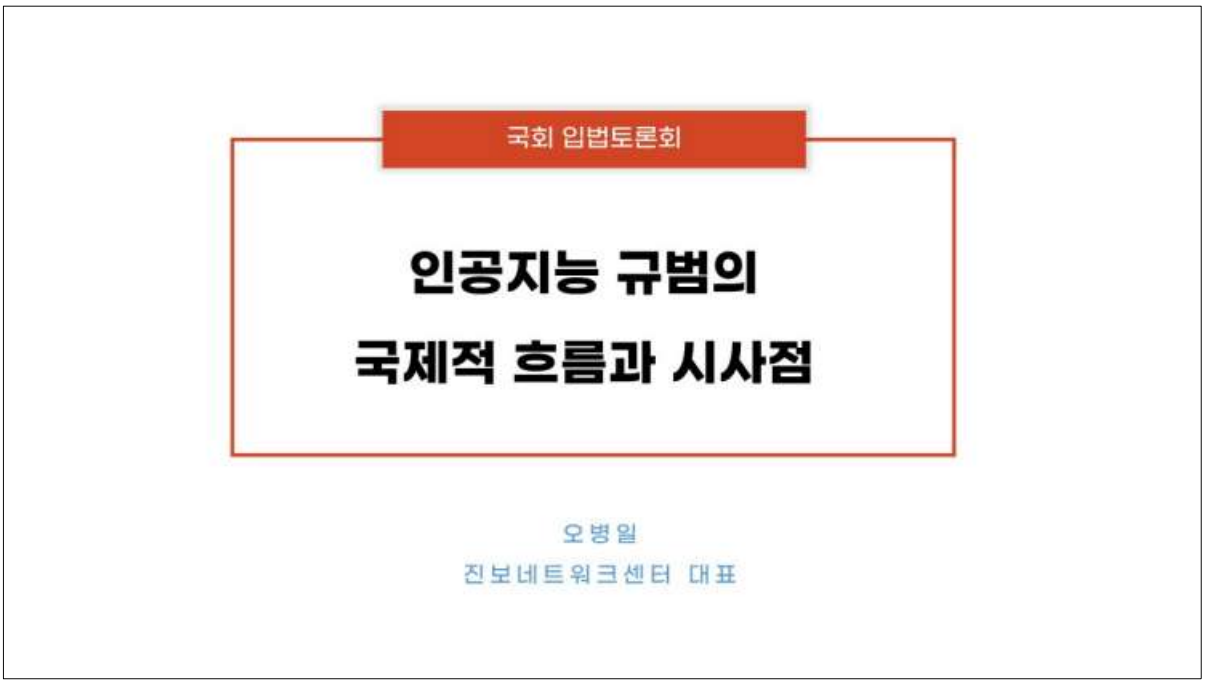
이를 위해서는 정부, 기업, 학계, 시민사회 등 다양한 분야의 전문가들이 머리를 맞대고 함께 고민하는 과정이 필수적입니다. 오늘 이 자리에 모인 여러분들의 적극적인 참여와 활발한 토론은 인공지능 시대의 미래를 밝히는 소중한 밑거름이 될 것입니다.

귀한 시간을 내어 참석해주신 모든 분들께 다시 한번 감사드리며, 오늘 토론회가 인공지능 시대의 새로운 미래를 여는 의미 있는 자리가 되기를 기대합니다.

감사합니다. ▢

인공지능 규범의 국제적 흐름과 시사점

오병일 | 진보네트워크센터 대표



고위험 AI 규제 필요성에 대한 국제적 인식 확산

- 2021.9. 유엔 인권최고대표, 얼굴인식기술을 포함한 고위험 인공지능에 대한 '모라토리엄' 촉구
- 2023.3. 요슈아 벤지오 등 저명한 연구자와 경영자, AI 개발을 6개월 동안 중지할 것을 촉구
- 2023.5. G7 정상회의, 생성형 인공지능에 관한 G7 히로시마 프로세스 수립
- 2023.11. 미국, 인공지능 행정명령 발표
- 2023.11. 영국 블레츨리 파크에서 개최된 AI 안정성 정상회의
- 2024.3. 유엔 총회, 지속가능한 발전을 위한 "안전하고 보안이 되며 신뢰할 수 있는" AI 시스템의 증진에 대한 결의안 채택
- 2024.5. EU AI Act 통과
- 2024.5. AI 서울 정상회의 개최
- 2024.6. 오픈AI 및 구글 딥마인드 전현직 직원, AI의 위험성과 정부 감독의 부족을 고발하며 내부고발자 보호 촉구

AI에 대한 유엔의 대응

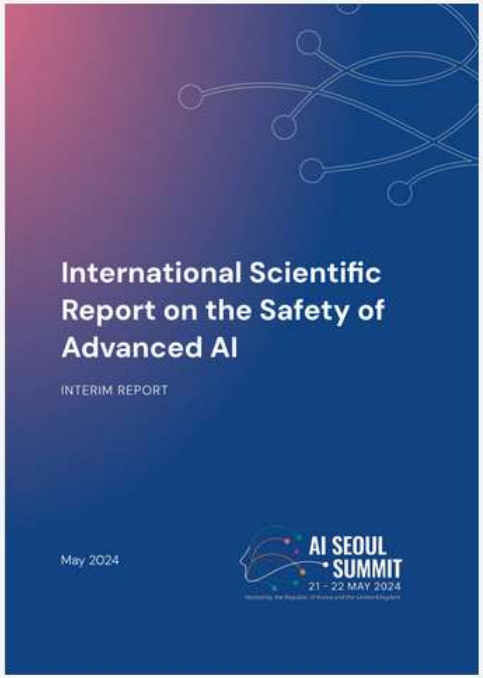
- 2020.3. 유엔 사무총장, 경제,사회,문화적 권리 실현을 위한 신기술의 역할 보고서(A/HRC/43/29)
  - **인공지능 사용에 대한 책임성을 완전하게 보장**하는 적절한 법률 체계와 절차 방법을 마련하고, 감독 체제를 수립하며, 인공지능의 피해에 대한 구제 수단을 구비할 것을 권고
- 2021.9. 유엔 인권최고대표, 디지털 시대 프라이버시권 보고서 (A/HRC/48/31)
  - 위험에 비례하여 법적 요구사항을 엄격하게 적용할 것과 **국제 인권법 하에서 정당화되지 않는** 잠재적 또는 실제적 영향을 초래하는 특정 인공지능 기술, 애플리케이션 또는 사용 사례에 대하여 **금지 요구**.
  - 실시간 원격 얼굴인식기술을 포함한 고위험 인공지능에 대한 '모라토리엄' 촉구
  - 공공과 민간 인공지능 사용의 부정적인 인권 영향을 방지하고 완화하는 **인권 실사와 피해자를 구제하는 규제 체계**의 도입을 요구.
  - 입법 조치는 인공지능에 대한 적정하고 **독립적이고 공정한 감독 체계**를 포함해야 하고, 감독 시스템은 개인정보보호 감독기구, 소비자 보호 기관, 부문별 규제 기관, 차별 방지 기구 및 국가 인권 기구를 포함하도록 권고.
  - 인공지능 시스템의 투명성 확보 강조. 특히 공공부문의 인공지능 의사결정에 대하여 **설명**할 수 있어야 함. 인권에 중대한 위험이 있는 인공지능에 대해서는 **등록제도**를 도입해야 함.

AI에 대한 유엔의 대응

- 2023.10.26. 유엔, 인공지능 고위급 자문위원회 구성
- 2023.12.31. 유엔 인공지능 고위급 자문위원회, 중간보고서 : 인간성을 위한 AI 규율 (Interim Report: Governing AI for Humanity) 발표
- 2024.3. 유엔 총회, 지속가능한 발전을 위한 "안전하고 보안이 되며 신뢰할 수 있는" AI 시스템의 증진에 대한 결의안 채택
  - AI에 초점을 맞춘 최초의 유엔총회 결의
  - AI의 설계, 개발, 배치, 활용에 있어서 인권의 존중, 보호, 증진 강조
  - 안전하고, 보안이 되며, 신뢰할 수 있는 인공지능 활용과 관련된 규제 및 거버넌스 접근 및 프레임워크 지지

첨단 AI의 안전성에 관한 국제 과학 보고서

- 2023.11. 영국 AI 안전 정상회의에서 보고서 작성 결의
- 2024.5 중간보고서 발표
- 요슈야 벤지오 교수가 의장 , 75명의 AI 전문가 참여
- 범용 AI의 기능, 위험, 완화조치에 대한 현재의 과학적 기반을 확인하는 것이 목적



범용 AI의 위험

악의적인 사용 위험

- 가짜 콘텐츠로 인한 개인 피해
- 허위정보 및 여론조작
- 사이버범죄
- 이중 사용 과학 위험

오작동으로 인한 위험

- 제품 기능 문제로 인한 위험
- 편견과 과소 대표로 인한 위험
- 통제력 상실

시스템적 위험

- 노동시장 리스크
- 글로벌 AI 격차
- 시장집중 및 독점으로 인한 위험
- 환경에 대한 위험
- 개인정보보호에 대한 위험
- 저작권 침해

유럽연합 인공지능 법안

주요 경과

- 2019.4.8. AI에 대한 고위급 전문가 그룹 : 신뢰할 수 있는 AI 윤리 가이드라인 발표
- 2020.2.19. EC, AI 백서 발표
- 2020.7.17. AI에 대한 고위급 전문가 그룹 : 신뢰할 수 있는 AI에 대한 최종 평가 목록(ALTAI) 발표
- 2021.4.21. EC, AI Act 제안 발표
- 2022.12.6. 유럽연합 이사회(Council of European Union), AI Act에 대한 공통 입장문(일반적 접근) 채택
- 2023.6.14. 유럽의회 AI Act에 대한 수정안(협상안) 채택
- 2023.12.9. EC, 이사회, 유럽의회 3자 협상을 통해 인공지능 법안에 대해 잠정 합의
- 2023.3.13. 유럽의회에서 공식적으로 통과.
- 2024.5.24. 이사회에서 공식적으로 통과
- 2024.6~7월 공식 저널에 게재 예정. 20일 후 발효.

인공지능의 정의

다양한 수준의 자율성을 가지고 작동하도록 설계되고, 배치 후 적응성을 나타낼 수 있으며, 명시적 또는 암묵적 목적을 위해 수신된 입력으로부터 물리적 또는 가상 환경에 영향을 미칠 수 있는 예측, 내용, 권장 사항 또는 결정과 같은 출력을 생성하는 방법을 추론하는 기계 기반 시스템

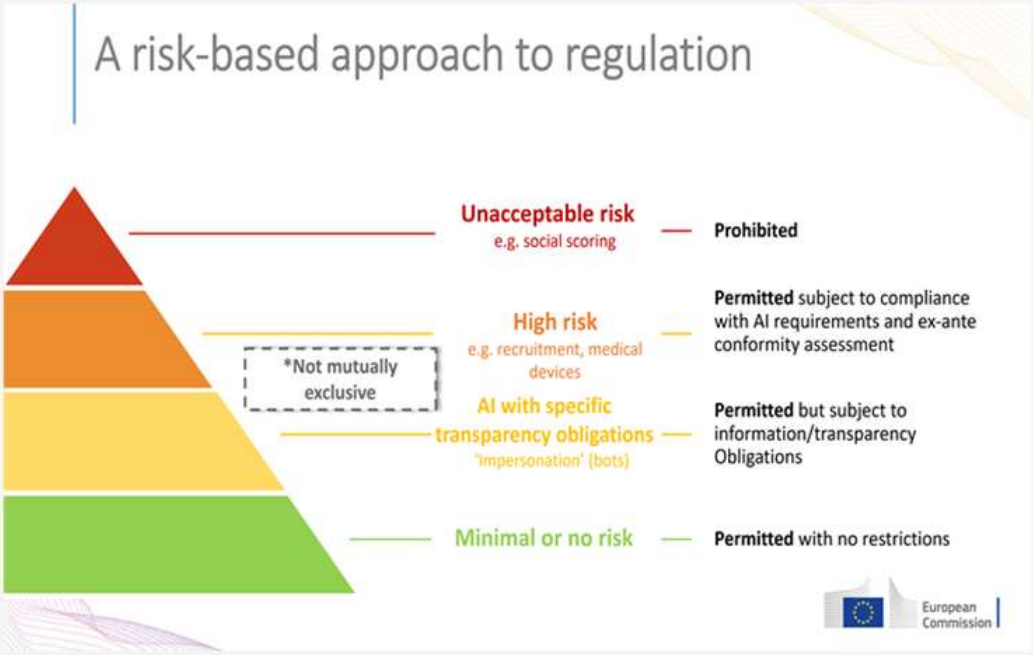
'AI system' means a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments

- 국제적인 조화와 폭넓은 수용을 위해 OECD의 인공지능 정의 차용
  - OECD 정의 : 사람이 정의한 특정 목표 집합에 대해 실제 또는 가상 환경에 영향을 미치는 예측, 권장 사항 또는 결정을 내릴 수 있는 기계 기반 시스템
- 범용 AI 및 생성형 AI 환경 반영

AI Act 적용 범위

- 유럽연합 내에 설립되었는지에 관계없이 **유럽연합 내에 AI 시스템 혹은 범용 AI 모델을 출시**하거나 서비스를 제공하는 제공자
- **유럽연합 내에 설립**되었거나 위치한 AI 시스템 **배치자(Deployer)**
- 시스템에 의해 생산된 **결과물이 유럽연합에서 사용**되는 (제3국에 위치하거나 설립지가 있는) AI 시스템의 제공자 및 배포자;
- 적용 제외
  - 오로지 군사 또는 국방 목적으로만 사용되는 시스템
  - 오로지 과학적 연구 및 개발을 위해 사용하는 AI 시스템 및 모델
  - 업무 목적 외로 AI 시스템을 사용하는 자연인(배치자)

AI Act의 위험 기반 접근법



출처 : https://www.spiceworks.com/tech/artificial-intelligence/articles/ai-regulation-best-approach/

금지되는 인공지능 (Title II)

- 잠재의식 기술이나 조작, 기만적인 방법을 사용하여 행동을 왜곡하고 정보에 입각한 의사 결정을 방해하여 심각한 피해를 초래하는 AI 시스템
- 나이, 장애 또는 사회적 또는 경제적 상황으로 인한 취약점을 악용하여 심각한 피해를 야기하는 AI 시스템
- 인종, 정치적 의견, 노동조합 가입 여부, 종교적 신념, 성생활 또는 성적 지향을 유추하는 생체 인식 분류 시스템(법 집행 기관의 합법적인 라벨링 또는 필터링은 제외) - 신설
- 사회적 행동 또는 개인적 특성을 기반으로 개인 또는 그룹을 평가하거나 분류하여 관련 없는 맥락에서 해롭거나 불균형적인 대우를 하거나 행동과 정당하지 않은 대우를 초래하는 AI 시스템
- 법 집행을 위한 공공장소에서의 '실시간' 원격 생체 인식 (피해자 또는 실종자 수색, 안전 위협 방지, 심각한 범죄 용의자 신원 확인 등 특정 필수 목적 제외)
- 프로파일링 또는 성격 특성만을 기반으로 개인의 범죄 위험성을 평가하는 AI 시스템 (범죄 행위와 관련된 객관적이고 검증 가능한 사실에 근거하여 사람의 평가를 지원하는 경우 제외) - 신설
- 인터넷이나 CCTV 영상에서 비표적 스크래핑을 통해 얼굴 인식 데이터베이스를 생성하는 AI 시스템 - 신설
- 직장이나 교육 기관에서 감정을 추론하는 AI 시스템(의료 또는 안전상의 이유 제외) - 신설

금지되는 인공지능 조항에 대한 평가

- 다양한 요건을 충족해야 하기 때문에 금지되는 범위가 넓은 것은 아님.
- 금지되는 인공지능 목록의 업데이트 메커니즘 부재 (고위험 인공지능은 제7조에서 업데이트 절차를 규정하고 있음)
- 생체인식 분류 시스템, 개인 식별을 전제로 한 생체인식 정보에 기반한 분류 시스템으로 제한 : 의회안의 (개인식별과 무관한) '생체인식 기반 데이터' 개념은 미도입
- 스크래핑을 통한 얼굴인식 DB 금지 : 미국의 클리어뷰(ClearView) AI와 같은 관행 금지
- 감정인식 시스템 : (의회안에서 금지했던) 법집행 및 국경관리 목적의 AI는 제외 + 의료, 안전 목적 감정인식 AI 허용
- 사회적 점수 시스템 : 공공기관(AI Act 초안) 뿐만 아니라 민간 시스템도 포함
- 공공장소에서 실시간 원격 생체인식
  - 의회안에서는 예외없이 금지하였으나 합의안에서는 제한적 허용(법집행 기관의 요구 수용)
  - 사전에 법원의 허가가 필요하며 엄격하게 제한된 범죄에 대해서만 시행 , 기본권 영향평가를 완료하고 데이터베이스에 등록해야 함
  - 사후 원격 생체인식은 금지가 아니라 고위험

고위험 AI 시스템

- 부속서 II의 유럽연합 조화 법률 관할의 제품 혹은 안전요소 + 제3자 적정성 평가를 받은 경우
- 부속서 III의 고위험 인공지능 시스템
  - 사후 원격 생체인식 시스템, 민감 속성에 기반한 생체인식 분류 시스템, 감정인식 시스템
  - 도로, 수도, 가스, 전기 등 중요 인프라
  - 교육 및 직업 훈련 수준 평가 시스템, 시험 중 학생의 금지된 행위 모니터링 시스템
  - 채용, 노동자 성과 관리 시스템
  - 필수적 공공 및 민간 서비스 접근 자격 평가 시스템 (사회복지 수급자격 등)
  - 생명 및 건강보험에서 개인에 대한 위험평가 및 가격책정 시스템
  - 법집행 : 범죄 피해 위험 평가, 거짓말 탐지기, 증거 신뢰성 평가, 재범 위험성 평가, 범죄수사 과정의 프로파일링
  - 이주, 망명 및 국경 통제 관리 시스템 : 거짓말 탐지기, 이주 및 망명 자격 심사
  - 사법 행정 및 민주적 절차 : 사실 확인이나 법적용에 사용되는 시스템, 선거 또는 국민투표 결과나 투표 행위에 영향을 미치기 위한 시스템

고위험 AI 시스템

- 고위험 인공지능 예외 규정
  - AI 시스템이 의사 결정 결과에 중대한 영향을 미치지 않는 등 자연인의 건강, 안전 또는 기본권에 중대한 위해를 가할 위험이 없는 경우
- 부속서 III의 AI 시스템이 자연인 프로파일링을 수행하는 경우 AI 시스템은 항상 고위험으로 간주됨

고위험 AI 시스템의 의무

- AI 시스템 전 주기에 걸쳐 위험관리 시스템 구축
- 데이터 거버넌스 : 데이터셋의 적합성, 대표성, 완전성
- 법률 준수 입증을 위한 기술 문서의 작성
- AI 시스템 전 주기에 걸쳐 (자동화된) 기록 유지
- 이용자(deployer)의 법률 준수를 위한 충분한 사용설명서 제공
- 인간에 의한 감독(human oversight)이 가능하도록 설계
- 적절한 수준의 정확성, 견고성, 사이버보안 제공
- 시스템상에 이름, 상호·상표, 주소 등을 표시 또는 첨부 문서에 표시
- 법률 준수 보장을 위한 품질관리 시스템 구축
- 관련 문서의 보관
- EU 시장 출시 또는 서비스 투입 전 적합성 평가 및 EU 적합성 선언 작성
- 'CE 마크'를 부착하거나 포장 또는 첨부 문서에 적합성 표시
- 필요한 시정조치 및 정보제공 의무
- EU 고위험 데이터베이스에 등록 및 규제기관에 협력
- 접근성 요구사항 준수

범용 AI에 대한 규율

- 용 AI(General Purpose AI, GPAI) 모델 : 대규모 자기지도 학습(self-supervision)을 사용하여 대량의 데이터로 학습된 경우를 포함하여 상당한 일반성을 나타내며 모델이 시장에 출시되는 방식에 관계없이 광범위한 고유 작업을 능숙하게 수행할 수 있고 다양한 다운스트림 시스템 또는 애플리케이션에 통합될 수 있는 AI 모델
- 연구, 개발, 프로토타입 제작 활동을 위해 시장에 출시되기 전에 사용되는 AI 모델에는 적용되지 않음.
- GPAI 모델 의무
  - 기술 문서(훈련 및 테스트 절차 및 결과) 유지
  - 해당 모델을 사용하는 AI 시스템 제공업체에 정보 및 문서 제공
  - 유럽위원회 및 국가 당국과 협력
  - **훈련 콘텐츠의 충분히 상세한 요약본** 공개
  - **저작권법** 준수 : 유럽연합 디지털 단일시장(DSM) 저작권 지침(Directive 2019/790)에 따라 권리자는 과학적 연구 목적이 아닌 한 텍스트 및 데이터 마이닝을 방지하기 위해 저작물 또는 기타 주제에 대한 권리를 유보할 수 있음. 이 경우 범용 AI 모델 제공자는 해당 저작물에 대해 텍스트 및 데이터 마이닝을 수행하려는 경우 권리자의 승인을 받아야 함.

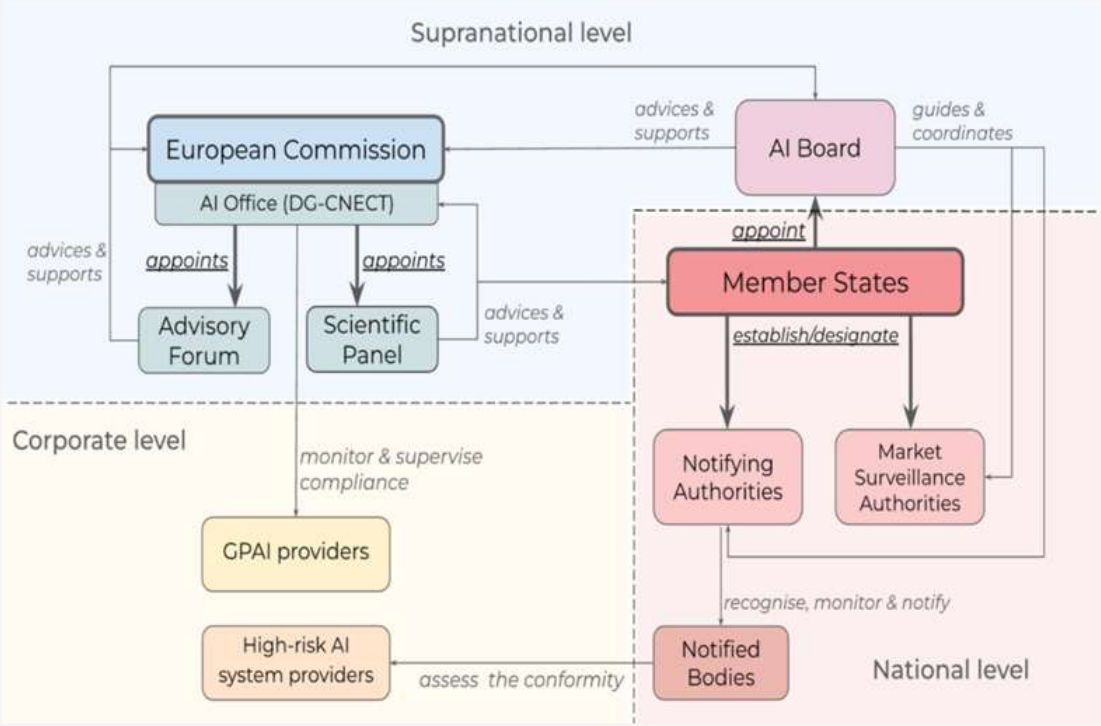
시스템적 위험이 있는 범용 AI에 대한 규율

- 유럽연합 수준에서 시스템적 위험
  - 공공건강, 안전, 보안, 기본권, 사회전체에 실제 혹은 합리적으로 예견된 부정적 영향을 국내 시장에 상당히 크게 미치는 경우.
  - 범용 AI 모델이 영향력이 높은 기능(high-impact capabilities)을 가진 경우
  - 부동 소수점 연산(FLOPs)으로 측정하여, 총 계산 능력이 10^25를 초과하는 경우 시스템적 위험이 있는 것으로 간주
- 제공자는 이러한 기준을 충족하는 경우 EC에 통보. EC는 자체적으로 (과학패널의 고지를 받아) 시스템적 위험을 가진 모델로 결정할 수 있음.
- 제공자의 의무
  - (적대적 테스트를 수행하고 문서화하는 것을 포함하여) 표준화된 모델 평가 수행
  - 시스템적 위험에 대한 평가 및 완화
  - 심각한 사고를 추적, 문서화, 보고 : AI Office 및 관련 당국에 보고
  - 적절한 사이버 보안 보장

AI 거버넌스

- AI 사무국(Office) : EC 산하에 설치. AI Act의 이행 및 준수 감독
- European AI Board (유럽 인공지능 이사회)
  - 회원국당 한 명의 대표로 구성. 유럽개인정보감독관(EDPS) 참관. 이사회 의장은 회원국 대표 중 한명. 일관된 적용을 위한 조언 및 지원 역할
- 자문포럼(58a조)
  - 이사회와 EC에 기술 전문 지식을 조언하고 제공.
  - 업계, 시민사회, 학계 등 이해관계자 대표로 구성. 기본권청, 유럽연합 사이버보안청, 유럽표준화위원회(CEN), 유럽전기기술표준화위원회(CENELEC) 및 유럽전기통신표준협회(ETSI)는 자문 포럼의 영구 회원
- 독립전문가 과학패널(58b조)
  - 인공지능 전문가로 구성. 유럽AI 사무국에 조언 및 지원 : 범용 AI 모델 및 시스템, 시장감시당국업무
- 국가관할당국의 지정 및 단일연락소(59조)
  - 하나의 지정기관(notifying authority) 및 시장감시기관을 국가관할기관으로 지정해야 함.

AI 거버넌스



출처 : Supranational and national bodies involved in the implementation and enforcement of the AIA (Novelli et al., 2024).

AI 거버넌스에 대한 평가

- 초안은 유럽연합 차원의 협력과 조율을 위해 각 국가 감독기관과 유럽개인정보보호감독관(EDPS)으로 구성되는 유럽인공지능이사회 설립 : 의장은 EC가 맡고 자문 역할에 한정
- 유럽의회는 법인격을 갖는 독립적 기구로 유럽 인공지능 사무소(European Artificial Intelligence Office)의 신설 제안 : 사무소에 운영이사회, 사무처, 자문포럼을 두도록 하고 있으며, EC에 대한 자문 및 지원 역할과 함께 훨씬 적극적이고 다양한 역할을 수행.
- 합의안의 AI 사무소는 EC 내 기능일 뿐. AI 이사회는 EC에 대한 자문 역할. 의회안보다는 EC의 주도적인 역할을 인정하는 방향으로 타협.

AI 시스템 배치자(사용자)의 의무

- 초안의 사용자(user)를 배치자(deployer)로 변경
- 고위험 인공지능 배치자의 의무 (26조)
  - 인적 감독 할당 및 필요한 지원 보장
  - 입력데이터의 관련성 및 대표성 확인
  - 사용 지침에 따라 운영 모니터링
  - 자동생성로그 보관
  - 자연인과 관련한 결정을 내리는 경우 당사자에게 고지
- 배치자의 **기본권 영향평가** 신설: 공공기관, 공적 서비스를 제공하는 민간기업, 부속서 III 5조 b,d 운영자(신용평가, 보험평가)
- **고용주의 의무** : 직장에 고위험 AI를 배치하려는 조직의 경우 근로자 대표와 해당 근로자에게 해당 시스템 적용에 대해 고지
- **공공기관** 고위험AI 배치자는 제51조 **등록 의무** 준수 필요
- 고위험 사후-원격생체인식 시스템 제공자는 사전에 법원 허가를 받아야 하며, 각 사용은 특정 범죄 수사를 위해 엄격하게 필요한 경우로 제한됨. 배치자는 모든 사용을 기록하고 연례보고서를 감독기구에 제출해야 함.

투명성 의무

- 자연인과 직접 상호 작용하도록 의도된 AI 시스템의 경우, 자신이 AI 시스템과 상호 작용하고 있음을 알려야 함
- 워터마킹 : 합성 콘텐츠의 경우 해당 콘텐츠가 인위적으로 생성, 조작되었음을 기계가 읽을 수 있는 형식으로 출력물에 표시
  - 법률에서 범죄를 탐지, 예방, 수사 및 기소를 위한 AI 시스템은 예외
- 감정인식 시스템, 생체인식 분류 시스템의 배포자는 이에 노출된 사람에게 시스템 운영에 대해 알려야 함.
- 딥페이크 : 배치자는 해당 콘텐츠가 인위적으로 생성 또는 조작되었다는 사실을 공개해야 함.
  - 콘텐츠가 명백히 예술적, 창작적, 풍자적, 허구적인 저작물인 경우, 저작물의 전시 또는 향유를 방해하지 않는 적절한 방식으로 구현

영향을 받는 사람의 권리

- 인공지능의 **영향을 받는 사람**(affected person) 개념 도입
  - 제공자 - 배치자 - 영향을 받는 사람
- 시장 감시 기관에 **불만(민원)을 제기할 권리**
- 자신에 대한 인공지능의 **의사결정에 대해 설명을 받을 권리**.
  - GDPR이 '오로지 자동화된 처리에만 의존하는 결정'으로 엄격하게 규정하고 있는 것에 비해 폭넓게 권리 인정

벌칙

- 금지된 AI 시스템 의무 위반 : 최대 3500만 유로 또는 총매출액의 최대 7%
- 고위험 시스템 또는 범용 AI 시스템 의무 위반 : 최대 1500만 유로 또는 총매출액의 최대 3%
- 부정확하거나 불완전한 정보 제공 : 최대 750만 유로 또는 총매출액의 최대 1%

다음 단계

- 공식 저널 발표 후 20일 후에 발효
- 발효 후 6개월부터 금지되는 인공지능 관행 적용
- 발효 후 12개월부터 범용 AI 모델 의무 적용
- 다른 조항은 발효 후 24개월 후부터 적용
- 제3자 적합성 평가 대상인 규제 대상 제품의 안전 요소인 경우 발효 후 36개월 후부터 적용
- 실천 강령(code of practice)은 발효 후 9개월 이내에 준비되어야 함.

유럽연합 AI Act 시사점

- 세계 최초로 인공지능을 포괄적으로 규율하는 법
  - → 인공지능 제품과 서비스의 국제적인 성격을 고려할 때 전 세계 다른 국가에 규범화 효과(브뤼셀 효과)를 가질 것
  - → 다만, 실제 적용되는 것은 사실상 2~3년 후이기 때문에 규제가 유예될 우려
  - → 한국의 AI 기업들도 유럽시장에 진출하기 위해서는 EU AI Act 준수 필요
- AI의 혁신을 촉진하면서도 시민의 안전과 인권에 중심을 둔 접근
  - → 다만, 인권보호 측면에서 합의안은 시민사회의 제안 및 유럽의회안에 비해 후퇴함
  - → 프랑스, 독일, 이탈리아 등이 산업계 입장 대변, 각 국의 법집행 당국의 입장 고려
- AI 위험성에 대한 체계화된 접근 : EU AI Act 의제가 제기하는 질문에 대한 고민 필요
- 국내 산업계의 주장 : 유럽의 인공지능 산업이 발전하지 못했기 때문에 강력한 규제 체제를 도입했다?
  - → 유럽 역시 인공지능 혁신 강조 : 샌드박스 조항 도입
  - → 산업계의 입장은 인공지능 산업 발전을 위해 인권 침해를 방치하자는 것?

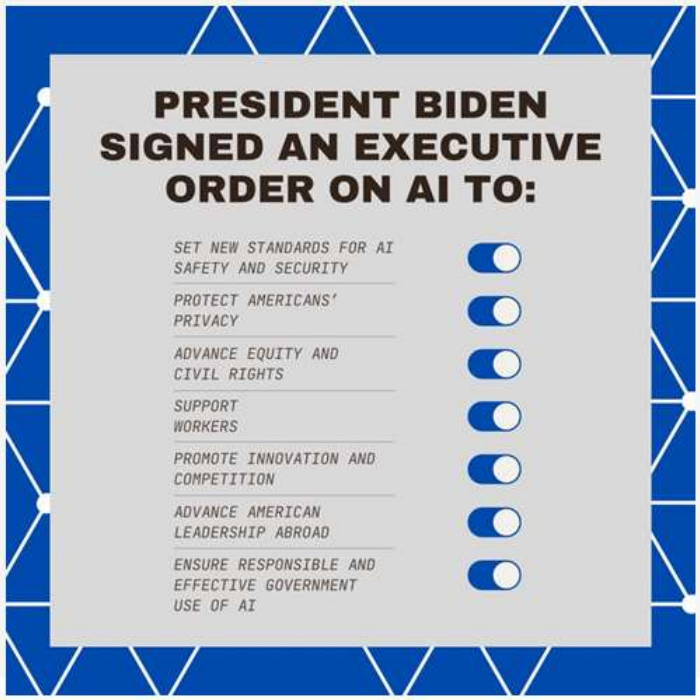
미국 인공지능 행정명령

주요 경과

- 2019.2.11. 트럼프 정부, 인공지능에서 미국의 리더쉽 유지 행정명령 발표 (행정명령 13859)
- 2020.12.3. 트럼프 정부, 연방정부에서 신뢰할 수 있는 인공지능 사용 증진 행정명령 발표(행정명령 13960)
- 2021.1.12. 백악관, 2020년 국가인공지능이니셔티브법에 따라 국가인공지능이니셔티브실 설립
- 2022.10. 백악관 과학기술정책실(OSTP), 인공지능 권리장전을 위한 청사진 발표
- 2023.1.26. 국립표준기술연구소(NIST), AI 위험관리 프레임워크 (AI RMF 1.0) 발표
- 2023.7.21. 바이든 정부, 주요 AI 기업으로부터 AI 위험성을 관리하겠다는 자발적인 약속 받아냄
- 2023.10.30. 바이든 정부, 안전하고 보안이 되며 믿을 수 있는 인공지능의 개발 및 이용에 대한 행정 명령 발표 (행정명령 14110)
- 2023.11.1. 관리예산실(OMB), 인공지능 행정명령에 대한 이행 가이드(초안) 발표 → 2024. 4.9. 관리예산실(OMB), 연방기관의 인공지능 사용을 위한 거버넌스, 혁신, 위험 관리 증진 정책 발표
- 2024.5. 미 상원, 미국 AI 정책 로드맵 발표

인공지능 행정명령 주요 내용

- 목적 : AI의 개발과 사용을 안전하고 책임감 있게 관리하는 것이 가장 시급한 과제이며, 연방정부 차원의 조정된 접근 방식이 필요함.
- 연방기관의 AI 사용 및 강력한 AI(foundation model)에 초점
- 8가지 지도 원칙
  - 안전과 보안
  - 혁신과 경쟁 촉진
  - 근로자 지원
  - 평등과 시민권 보호
  - 소비자 보호
  - 개인정보보호
  - 연방정부의 AI 활용 증진
  - 해외에서 미국의 리더십 강화



인공지능의 안전과 보안

- 국방물자생산법(DPA)에 따라, 강력한 AI 시스템 개발자가 안전 테스트 결과와 기타 중요 정보를 미국 정부와 공유 요구
  - 모든 레드팀 안전 테스트 포함
  - 정수 또는 부동 소수점 연산 기준, 총 계산 능력이 10^26를 초과하는 경우 / 주로 생물학적 서열 데이터를 사용하면서 총 계산 능력이 10^23를 초과하는 경우
- AI 시스템의 안전과 보안, 신뢰성을 보장하기 위한 지침, 표준, 모범사례 개발
  - 생성형 AI를 위한 NIST 위험관리 프레임워크 등
  - 국토안보부 AI 안전 및 보안위원회 설립, CBRN(화학, 생물학, 방사능, 핵) 위협에 대처
- AI 생성 콘텐츠에 대한 인증 및 워터마킹[라벨링] 지침 개발

혁신과 경쟁 촉진

- **인재 유치** : AI에 전문성이 있는 이민자와 비이민자가 미국에서 공부하고 일할 수 있는 기회 확대
- **혁신 촉진** : AI 연구자와 학생들에게 주요 AI 리소스와 데이터에 대한 엑세스를 제공하여 연구 촉진
- **경쟁 촉진** : 소규모 개발자와 중소기업을 지원하여 공정하고 개방적이며 경쟁력있는 AI 생태계 촉진

근로자 지원

- 인공지능은 미국의 일자리와 직장을 변화시키고 있으며, 생산성 향상에 대한 가능성과 함께 직장 내 감시, 편견, 일자리 이동의 위험성 야기
- 이러한 위험을 완화하고 근로자의 단체 교섭 능력을 지원하며, 모두가 접근할 수 있는 인력 교육 및 개발에 투자할 필요
- 일자리 대체, 노동 기준, 직장 내 형평성, 건강 및 안전, 데이터 수집에 관한 원칙과 모범 사례를 개발
- AI의 잠재적 노동시장 영향에 대한 보고서 작성 및 노동 중단 근로자를 위한 연방 정부의 지원 강화 방안 연구

평등과 시민권 보호

- **형사 사법 시스템**에서 AI 및 시민권 강화
  - AI 관련 민권 침해 조사 및 기소를 위한 모범 사례에 대한 교육, 기술 지원, 법무부와 연방 민권 사무소 간의 조정을 통해 알고리즘 차별 해결
  - 선고, 가석방 및 보호 관찰, 재판 전 석방 및 구금, 위험 평가, 감시, 범죄 예측 및 예측 치안, 포렌식 분석에 AI를 사용하여 시민권 보호와 법집행 효율성을 향상할 수 있는 모범 사례 개발
- **정부 혜택** 및 프로그램과 관련된 시민권 보호.
  - AI 알고리즘이 차별을 악화시키지 않도록 임대인, 연방 혜택 프로그램, 연방 계약업체에 명확한 지침 제공
- 더 넓은 경제에서 AI와 시민권 강화
  - **채용, 주택, 금융**에서 AI를 통한 차별 및 편견 방지

소비자 보호

- 사기, 차별, 개인 정보 보호 위협 및 다른 AI 사용 위험으로부터 미국 소비자 보호
- **환자** 보호 : 의료분야에서 AI의 책임감 있는 사용과 생명을 구할 수 있는 저렴한 약품 개발을 촉진
- **승객** 보호 : 운송 부문에서 AI의 안전하고 책임 있는 개발 및 사용 촉진
- **학생** 보호 : AI 기반 교육 도구 등 교육을 혁신할 수 있는 AI의 잠재력을 구체화
- **이용자** 보호 : AI가 통신 네트워크와 소비자에게 미치는 영향 검토

개인정보 보호

- 훈련 데이터의 개인정보 보호 등 개인정보 보호 기술(PET)의 개발과 사용을 가속화하기 위한 연방 정부의 지원
- 암호화 도구와 같은 프라이버시 보호 연구 및 기술을 강화
- 연방기관이 상업적으로 이용 가능한 정보를 처리하는 방법을 평가하고 AI 위험을 고려할 수 있도록 개인정보 보호 지침 강화
  - 개인정보 보호 위험을 완화하는 데 개인정보 영향평가가 어떻게 더 효과적일 수 있는지에 대한 의견 요청
- 개인정보 보호 기술의 효과를 평가할 수 있는 지침 개발
- 초당적인 데이터 개인정보 보호 법안을 통과시킬 것을 의회에 촉구 (Fact sheet)

연방정부의 AI 활용 증진

- AI를 통한 정부 서비스의 효율성 확대, 그러나 차별과 안전하지 않은 결정과 같은 위험 존재
  - 연방기관의 AI 사용에 대한 지침 제정
  - 각 기관은 **최고인공지능책임자(CAIO)** 지정
  - 각 기관 내부에 **인공지능 거버넌스 위원회** 신설
  - AI 권리장전 청사진 및 NIST 위험관리 프레임워크에 기반한 **위험관리 관행 수립** : 공개 의견수렴, 데이터 품질 평가, 서로 다른 영향과 알고리즘 차별 평가 및 완화, AI 사용 통지, 사용하고 있는 AI에 대한 지속적인 모니터링 및 평가, 인간의 고려, AI의 불리한 결정에 대한 구제수단 제공 등
  - 연방 직원의 업무에 생성 AI를 사용하는 방법에 대한 지침 개발
- 보다 효율적인 계약을 통해 특정 AI 제품과 서비스를 효과적으로 획득할 수 있도록 지원
- AI 전문가의 신속한 채용과 직원을 대상으로 한 AI 교육 제공

해외에서 미국의 리더십 강화

- AI 협력을 위한 양자, 다자, 다중 이해관계자 참여 확대
  - AI의 위험을 관리하고 안전을 보장하기 위한 강력한 국제 프레임워크 구축 노력 주도
- 국제 파트너 및 표준 기구와 함께 중요한 AI 표준의 개발 및 구현을 가속화
- 안전하고 책임감 있고 권리를 보장하는 AI의 해외 개발 및 사용을 촉진
  - NIST, AI 위험 관리 프레임워크의 원칙을 통합하는 AI 글로벌 개발 플레이북 발행
- 글로벌 AI 연구 의제를 개발
- 중요 인프라에 대한 국경 간 및 글로벌 AI 위험 해결하기 위해 국제 동맹국 및 파트너와 협력

관리예산실(OMB)의 이행 가이드

- 공공기관은 매년 **모든 AI 사용 사례 목록(인벤토리)**을 OMB에 제출하고 해당 기관 웹사이트에 게시.
  - 안전과 권리에 영향을 미치는 AI의 경우, 사용 방법, 위험, 위험관리 방법 포함
- 안전과 권리에 영향을 미치는 AI에 대해 **위험관리를 위한 최소 관행**을 규정
  - 이를 따르지 못하는 AI의 사용 중단 요구
- AI 공급업체에 일정한 문서 제공 요구
- 안전과 권리에 영향을 미치는 AI에 대한 정의
  - Rights-Impacting AI: 시민의 권리와 자유, 동등한 기회(교육, 주거, 보험, 신용, 고용 등), 정부 서비스(건강, 금융, 공공주거, 사회보장, 교통, 필수서비스 등)에 중대한 영향을 미치는 AI
  - Safety-Impacting AI: 생명 혹은 복지, 기구 및 환경, 핵심 인프라, 전략적 자산의 안전에 중대한 영향을 미치는 AI
- 최고 AI 책임자(CAIO)가 달리 결정하지 않는 한, 안전 혹은 인권에 영향을 미치는 것으로 추정됨.

안전에 영향을 미치는 AI

- 댐, 응급서비스 전력망, 에너지 전송, 화재안전 시스템, 식품안전, 교통통제, 상하수도, 원자력 시스템 등 안전 기능 통제
- 선거 또는 투표 인프라
- 직장, 학교, 주택, 교통, 의료 또는 법 집행 환경 내 로봇 사용
- 운동력, 생물학적 또는 화학적 작용의 전달, 잠재적으로 유해한 전자기 자극의 전달
- 육상, 지하, 해상, 공중, 우주 등 차량의 이동
- 유해 화학물질이나 생물학적 매개체의 전달, 안전, 설계, 개발의 통제
- 산업 장비, 시스템 또는 구조물의 설계, 시공 또는 테스트
- 의료 기기, 의료 진단 및 치료, 보험 / 약물 중독 / 자살 등의 위험평가 등
- 위험한 무기 또는 폭력 행위의 존재 감지
- 응급 상황에 대한 응급 구조대원 소환 선택
- 정부 시설에 대한 접근 또는 보안 통제
- 수출, 투자 또는 운송에 대한 제재, 무역 제한 등에 따른 집행 조치의 결정이나 이행

인권에 영향을 미치는 AI

- 보호되는 표현을 제약하는 행위
- 법 집행 맥락에서 개인에 대한 위험평가, 예측, 식별, 차량 추적, 생체인식, SNS 모니터링, 선고 및 가석방 등 결정
- 이민, 망명, 구금 상태와 관련되어 위험 평가 결정, 개인에 대한 모니터링 및 예측
- (공공 생체인식 감시) 공공장소에서 일대다 신원 확인을 위한 생체 인식 인증 실시
- (감정인식) 사람의 감정, 사고, 장애 또는 속임수를 감지하거나 측정하는 경우
- 명시적인 동의 없이 사람의 모습이나 목소리를 복제하는 행위
- 교육 맥락에서 부정행위 탐지, 입학 절차, 학생 진도 등
- 세입자 심사, 모니터링, 주택 가치 평가 등
- 채용, 급여 또는 승진, 성과 관리, 징계 권고, 직장 감시 등
- 의료기기, 의료 진단, 치료, 보험 건강위험 평가 등
- 신용 평가, 보험 평가 등 금융 분야
- 정부 서비스에 대한 접근
- 법적 구속력이 있는 개인과의 공식적 의사소통을 목적으로 하는 언어 간 번역 등
- 입양 매칭, 아동보호 조치, 아동 양육권 추천, 노인 및 장애인 보호조치에 대한 권고 등

안전과 권리에 영향을 미치는 AI의 위험 관리를 위한 최소관행

- 도입 전
  - AI 영향평가 완수
  - 실제 환경에서 AI 성능 테스트
  - AI에 대하여 독립적으로 평가
- 도입하는 동안
  - 지속적인 모니터링 수행
  - AI 사용 위험에 대한 정기적인 평가
  - 권리와 안전에 미치는 신규 위험 완화 -- 완화 조치 실행이 불가능한 경우 AI 사용 중단
  - 인력에 대한 적절한 교육훈련 및 평가 보장
  - 권리나 안전에 고위험을 초래하는 결정의 일부로서 적절한 인적 검토 제공
  - AI 사용 사례 인벤토리를 통해 공중에 정보 공개

권리에 영향을 미치는 AI에 대한 추가적인 최소 관행

- 도입 전
  - AI가 형평성, 공정성에 미치는 영향을 파악, 평가하고 알고리즘 차별이 존재하는 경우 이를 완화
  - **영향을 받는 집단의 피드백을 수렴**하고 반영
- 도입하는 동안
  - AI 기반 차별에 대하여 지속적으로 모니터링하고 완화 실시
  - 부정적인 영향을 받은 개인에 대해 통지
  - **인적 검토와 구제절차** 관리
  - AI 기반 의사결정에 대한 **거부(opt-out) 옵션** 제공

미 상원 AI 정책 로드맵

- 미 상원 척 슈머 위원장을 중심으로 초당적인 AI 워킹그룹 보고서 발표
- 2023년에 9차례의 AI 인사이트포럼 개최를 통해 주요 AI 이슈 제안
  - 창립 포럼
  - 미국의 AI 혁신 지원
  - AI와 인력
  - AI의 영향력 있는 활용
  - 선거와 민주주의
  - 개인정보 보호 및 책임
  - 투명성, 설명 가능성, 지적 재산권 및 저작권
  - AI 위험으로부터 보호
  - 국가 안보
- 이 로드맵이 초당적인 AI 법안 검토를 위한 기반이 될 것으로 보임
- 시민사회는 AI 법안이 AI의 위험성을 실질적으로 통제할 수 있어야 함을 주장하며, Shadow report 발간

인공지능 행정명령의 시사점

- 트럼프 정부에서의 행정명령이 AI의 활용과 윤리적 원칙의 표명에 그쳤다면, 바이든 정부의 행정명령은 AI의 부정적 영향에 대한 인식과 AI의 책임감 있는 사용을 위한 대책에 초점
- 글로벌 AI 규범 형성에 미국의 리더쉽 회복의 의지 표명
  - 개인정보보호(GDPR), 디지털 시장 독점 규제(DMA) 등 분야에서는 유럽연합이 리더쉽 발휘 & 유럽 AI Act 타결
  - 영국 AI 안전 정상회의 직전에 행정명령 발표
- 행정명령은 인권과 안전에 영향을 미치는 AI를 활용하는 공공기관에 강력한 의무를 부과하고 있지만, 이는 결국 민간에서 개발하는 AI에 영향을 미치게 됨.
- 법률이 아닌 행정명령의 한계
  - 대규모 AI에 대한 보고 의무 : 국방물자생산법(DPA) 활용 논란 - 현재 진정으로 국가안보적 의미가 있는지, 국방물자생산법 취지에 맞는지.
- 미국은 자율 규제를 위해 AI 법안 제정을 주저하고 있는가?
  - 행정명령에서 바이든 행정부는 의회와 협력하여 미국이 책임 있는 혁신을 선도할 수 있도록 초당적 입법을 추진하겠다는 의지 표명

국제동향 시사점

고위험 AI의 규제 필요성

- 고위험 AI의 위험성과 이를 통제할 필요성에 대해서 국제적인 공감대 형성
  - 2023.11. 영국 블레츨리 파크에서 개최된 AI 안정성 정상회의
  - 2024.3. 유엔 총회, 지속가능한 발전을 위한 "안전하고 보안이 되며 신뢰할 수 있는" AI 시스템의 증진에 대한 결의안 채택
- 유럽, 미국을 비롯하여 AI의 위험성에 따라 차별적인 규제를 적용하는 '위험 기반 접근 ' 에 대한 공감대 형성

> 우리는 안전하고, 혁신적이고 포용적인 AI 생태계들을 육성하는, 위험 기반 접근법들을 포함한
> 정책·거버넌스 체계들을 지지한다
>
> - AI 서울 정상회의 선언문

인공지능 규제 방식

- 인공지능에 대한 포괄적 규율 vs 부문별 규율
  - 유럽 AI Act 는 AI의 위험성을 통제하기 위한 포괄적 법안이고, 미국의 행정명령은 AI 환경에 대응한 부문별 정책 패키지임. (위상이 다름)
  - AI의 위험성 규율과 관련해서는 포괄적인 프레임워크와 특수성에 대한 고려가 모두 필요함. 예를 들어, 유럽 AI Act도 범용 AI, AI의 위험성, 범죄수사 목적 등 특수성을 고려한 규정을 두고 있음. 반면, 미국 NIST의 위험관리 프레임워크 역시 포괄적 프레임워크임.
  - 미국 역시 AI에 대한 포괄적인 규제 법안 추진
- 법적 규제 vs 자율규제
  - 자율규제는 보완재, 최소한의 책임성을 위한 법적 규제는 필요함

"윤리 강령은 인권 책무를 중요하게 보완할 수는 있지만 대체물은 아니다."
"민간이 주력하고 공공이 추구하는 인공지능 윤리는 인권 기반 규제에 대한 반발을 내포한 경우가 많다. 인공지능 분야에서 윤리는 특정 과제를 해결하는 중요 체계를 제공하지만, 윤리가 모든 국가에서 법률로 구속되는 인권을 대체하는 것은 아니다. 기업과 정부는 윤리 강령과 지침을 개발할 때에도 인공지능 운영의 모든 측면에 인권 고려사항과 책임을 확실하게 통합시켜야 한다"

(유엔 의사 표현의 자유 특별보고관, 2018. 8. 29. 보고서)

AI 산업 육성 vs 위험성 규제

- 육성과 규제는 배치되는 것이 아님. AI의 안전성에 대한 신뢰 없이는 AI 산업 발전도 불가능
- 규제의 목적은 기업에게 부담을 주고자 하는 것이 아니라, 안전과 인권의 보호임
- AI 규제에 대한 산업 위축론은 논점 왜곡 → 산업이 위축되면 안전과 인권을 보호하지 말자는 것인가?
- AI의 위험성을 식별하고 이에 대한 적절한 규제 체제가 무엇인지에 대한 논의에 초점을 두어야 함

범용 AI의 사회적 위험요소

"제품을 빠르게 출시하는 것이 중요한 역동적인 시장에서 시장 점유율을 놓고 경쟁하는 범용 AI 개발자는 위험 완화를 위해 투자할 인센티브가 제한적일 수 있다.
이로 인해 안전과 윤리를 보장하는 조치에 대한 투자는 소홀히 하면서 범용 AI 모델을 최대한 빨리 개발하기 위해 경쟁하는 '바닥을 향한 경주' 시나리오에 대한 우려가 제기되고 있다
범용 AI 규제에 관한 국제적 공조가 이루어지지 않는다면, 규제 '바닥을 향한 경쟁'으로 인해 각국이 국내외 안전 보장에 불충분할 수 있는 느슨한 규제를 통해 AI 기업을 유치하려고 시도할 수 있다."

첨단 AI의 안전성에 관한 국제 과학 보고서

인권과 안전에 기반한 AI 거버넌스

- 인권, 안전, 민주주의에 기반한 AI 거버넌스가 필요함
  - 언제까지 인권과 안전의 규범을 산업발전의 후순위로 놓을 것인가
  - 안전과 인권을 침해할 수 있는 인공지능에 왜 국가적인 지원을 해야 하는가 → 산업 정책도 공공성과 인권 보호를 위한 AI를 중심으로 지원하도록 변화 필요
- 인공지능에 대한 국제적인 규율 필요
  - AI 제품 및 서비스의 보급, 그리고 AI가 야기하는 위험성은 본질적으로 국제적 성격
  - 실제 미국, 유럽 등 주요 국가의 AI 윤리 원칙 및 위험 관리 관행의 내용은 상당히 유사함
  - 유럽 및 미국의 규제 사례는 한국에 참조가 될 수 있음

> 우리는 인간 중심적인 AI를 활용하여 국제 난제를 해결하고, 민주주의적 가치·법치주의 및 인권·기본적 자유와 프라이버시를 보호 및 증진하고, 국가 간의 그리고 국내적인 AI 및 디지털 격차를 해소함으로써, 인간의 복지를 향상하고, 유엔 지속가능발전목표 진전을 포함하여 AI를 실용적으로 활용하도록 지원하기 위해, AI 안전·혁신·포용성을 향상시키는 국제 협력 강화를 촉구한다.
>
> - AI 서울 정상회의 선언문

감사합니다.

22대 국회 인공지능법 입법방향

유승익 | 한동대학교 교수

Ⅰ. 들어가며

2024년 5월 31일 개원한 제22대 국회에서 인공지능 및 관련 법제에 대한 관심이 고조되고 있다. 지난 6월 26일, 초당파적 모임인 국회 AI포럼이 창립총회 및 기념 세미나를 열고 국내 AI 산업 지원 필요성을 역설하며 인공지능 기본법 통과를 촉구했다.

현재까지(2024. 7. 3. 기준) 제22대 국회에서 총 5건의 인공지능법이 발의되었다. 안철수 의원이 대표발의한 「인공지능 산업 육성 및 신뢰 확보에 관한 법률안」(의안번호: 2200053)은 개원 첫 날 발의되었고, 이후 여야를 막론하고 발의가 이어지고 있다. 특히 여당은 6. 17. 당선인 108인 전원이 「인공지능 발전과 신뢰 기반 조성 등에 관한 법률안」(의안번호: 2200543)을 공동발의하였다. 앞으로 인공지능법 제정안은 계속해서 발의될 것으로 보인다.

하지만 22대 국회에서 발의되고 있는 제정 법률안은 21대 국회에서 발의되어 논의되었던 법률안의 취지와 내용을 답습하면서, 주로 인공지능을 산업 진흥의 관점에서 접근함으로써, 인공지능이 갖는 위험성을 간과하고 있다는 지적도 계속되고 있다.

이 발제에서는 이번 국회에서 지금까지 발의된 총 5개의 인공지능법의 내용을 살펴보고, 22대 국회의 인공지능법 입법방향을 제시하고자 한다.

Ⅱ. 제22대 국회 인공지능법 개관 및 주요 내용

1. 법안 개요

현재까지 발의된 인공지능법은 총 5건으로 그 내용은 다음과 같다.

의안번호	제안일자	제안자	법안명
2200053	2024-05-31	안철수의원 등 12인	인공지능 산업 육성 및 신뢰 확보에 관한 법률안
2200543	2024-06-17	정점식의원 등 108인	인공지능 발전과 신뢰 기반 조성 등에 관한 법률안
2200673	2024-06-19	조인철의원 등 19인	인공지능산업 육성 및 신뢰 확보에 관한 법률안
2200675	2024-06-19	김성원의원 등 11인	인공지능산업 육성 및 신뢰 확보에 관한 법률안
2201158	2024-06-28	민형배의원 등 13인	인공지능기술 기본법안

법안의 대체적인 입법방향은 법안명에서 알 수 있듯이 인공지능 “산업”의 “발전과 육성”과 “신뢰 확보”라 할 수 있다. 산업의 관점에서 인공지능에 접근하고 있으므로 자연스럽게 자국 산업의 육성에 초점이 모아지고, 신뢰 기반 조성 등의 기조는 산업진흥에 보조적·종속적 지위에 머무르게 된다. 이러한 취지는 정점식의원안의 제안이유에서 엿볼 수 있다. “우리나라 인공지능 산업의 발전을 지원하기 위해 국내 산업과 사회·문화적 맥락을 고려하여 인공지능 산업의 혁신을 저해하지 않으면서도 인공지능의 부작용과 위험을 최소화하기 위한 법적 기반 마련이 필요”하다는 것이다. 다른 법률안도 대동소이한 제안이유를 부기하고 있다.

2. 법안의 체계

법안들은 대체로 “인공지능 발전과 신뢰 기반 조성 등에 관한 법률”, “인공지능산업 육성 및 신뢰 확보에 관한 법률” 등의 제명 하에 규정되어 있고, 정점식 의원안에서 보는 바와 같이 대개 5장, 30여개의 조문으로 구성되어 있다. 법안들은 서로 대동소이한 내용으로 구성되어 있다. 다만, 조인철 의원안의 경우, 인공지능 실증 규제특례(안 제14조), 인공지능제품의 비상정지(안 제28조) 등을 규정하고 있으며, 민형배 의원안의 경우, 지역인공지능위원회의 구성 등을 포함하였다. 김성원 의원안의 경우, 여전히 “우선허용·사후규제 원칙”을 규정하고 있다(안 제11조).

인공지능 발전과 신뢰 기반 조성 등에 관한 법률안(정점식 의원안의 경우)

제1장 총칙

제1조(목적)
제2조(정의)
제3조(기본원칙)
제4조(국외행위에 대한 적용)
제5조(다른 법률과의 관계)

제2장 인공지능의 건전한 발전과 신뢰 기반 조성을 위한 추진체계

제6조(인공지능 기본계획의 수립)
제7조(국가인공지능위원회)
제8조(위원회의 기능)
제9조(위원의 제척 · 기피 및 회피)
제10조(분과위원회 등)
제11조(국가인공지능센터)
제12조(인공지능안전연구소)

제3장 인공지능기술 개발 및 산업 육성

제1절 인공지능산업 기반 조성

제13조(인공지능기술 개발 및 안전한 이용 지원)
제14조(인공지능기술의 표준화)
제15조(인공지능 학습용데이터 관련 시책의 수립 등)

제2절 인공지능기술 개발 및 인공지능산업 활성화

제16조(기업의 인공지능기술 도입 · 활용 지원)
[제14조(인공지능 실증 규제특례)]
제17조(창업의 활성화)
제18조(인공지능 융합의 촉진)
제19조(제도개선 등)
제20조(전문인력의 확보)
제21조(국제협력 및 해외시장 진출의 지원)
제22조(인공지능집적단지 지정 등)
[제21조(대한인공지능협회의 설립)]

제4장 인공지능윤리 및 신뢰성 확보

제23조(인공지능 윤리원칙 등)
제24조(신뢰할 수 있는 인공지능)
제25조(인공지능 신뢰성 검 · 인증 지원 등)

제26조(고위험영역 인공지능 고지 의무)
제27조(고위험영역 인공지능의 확인)
제28조(고위험영역 인공지능과 관련한 사업자의 책무)
[제28조(인공지능제품의 비상정지)]
제29조(생성형 인공지능 고지 및 표시)
제30조(생성형 인공지능 안전 확보 의무)
제31조(민간자율인공지능윤리위원회의 설치 등)

제5장 보칙

제32조(인공지능산업의 진흥을 위한 재원의 확충 등)
제33조(실태조사, 통계 및 지표의 작성)
제34조(권한의 위임 및 업무의 위탁)
제35조(벌칙 적용에서 공무원 의제)
제36조(벌칙)

부칙

3. 법안의 주요내용

법안은 인공지능의 건전한 발전을 지원하고 인공지능사회의 신뢰 기반 조성에 필요한 기본적인 사항을 규정함으로써 국민의 권익과 존엄성을 보호하고 국민의 삶의 질 향상과 국가경쟁력을 강화하는 데 이바지함을 목적으로 한다고 밝히고 있다(정점식의원안 제1조). 요약하자면, 인공지능산업 육성 도모와 신뢰 기반 조성이다.

법안들의 주요 내용은 다음과 같다.

① 인공지능산업 육성 및 신뢰 확보를 위한 추진체계로, 과학기술정보통신부 장관은 3년마다 인공지능기술 및 인공지능산업의 진흥과 국가경쟁력 강화를 위하여 인공지능 기본계획을 수립·시행, 대통령 또는 국무총리 소속으로 심의·의결기관인 국가인공지능위원회와 그 산하에 전문위원회(인공지능 신뢰성 전문위원회 포함) 설치, (지능정보사회진흥원 산하) 국가인공지능센터 설치가 포함된다.

② 인공지능 기술개발 및 산업 육성을 위하여, 인공지능기술 개발 및 안전한 이용 지원 사업 실시, 인공지능기술의 표준화, 인공지능 학습용데이터 관련 시책의 수립, 기업의 인공지능기술 도입·활용 지원, 창업 활성화, 인공지능 융합의 촉진, 과기정통부장관의 법령정비 등 제도개선 노력의무, 전문인력의 확보 시책 추진, 국제협력 및 해외시장 진출의

지원, 인공지능집적단지 지정, 대한인공지능협회의 설립 등이 포함된다.

③ 인공지능윤리 및 신뢰성 확보를 위하여, 정부는 인공지능사업자 및 이용자가 인공지능의 개발 · 이용과정에서 지켜야 할 인공지능 윤리원칙 제정·공표, 신뢰할 수 있는 인공지능 기반조성을 위한 시책 마련(과기정통부장관), 인공지능 신뢰성 검 · 인증 지원 사업 추진(과기정통부장관), 고위험 영역 인공지능의 확인, (제품 또는 서비스 제공자의 이용자에 대한) 고위험 영역 인공지능 고지 의무, 고위험 영역 인공지능과 관련한 사업자의 책무, 생성형 인공지능 고지 및 표시 의무 및 안전 확보 의무 부과, 민간자율인공지능윤리위원회의 설치등을 규정하였다. 벌칙규정은 업무상 비밀누설, 직무상 목적 외 용도 사용(징역 3년 이하 또는 3천만원 이하 벌금)에 한정된다.

④ 기타 보칙으로 인공지능산업의 진흥을 위한 재원의 확충 방안 마련, 실태조사, 통계 및 지표의 작성·관리 및 공표, 비밀누설 등에 대한 벌칙 조항, 국가인공지능센터 또는 이와 유사한 명칭 사용한 자에 대한 과태료 부과 조항을 규정하고 있다.

Ⅲ. 22대 발의 법안의 문제점과 제정방향

1. 인공지능 정의 규정의 문제점과 제정방향

발의된 법안들은 인공지능을 다음과 같이 정의하고 있다.

“‘인공지능”이란 학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것을 말한다.”

안철수의원안의 경우, “‘인공지능”이란 자율성을 가지고 외부의 환경 또는 입력에 적응하여 학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것을 말한다.”라고 다르게 규정하고 있다.[1)]

유럽연합 AI Act의 “인공지능시스템"에 대한 정의는 다음과 같다.

---

1) 안철수의원안은 알고리즘에 대한 별도의 정의 규정을 두고 있다. “‘알고리즘”이란 문제의 해결, 업무의 수행 또는 장비 · 장치 · 기기 등의 운용을 위하여 기술(記述)된 연산, 규칙과 절차명령 또는 논리의 집합으로 이루어진 체계를 말한다.”(안 제2조 제1호)

Article 3 Definitions For the purposes of this Regulation, the following definitions apply: (1) 'AI system' means a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;	제3조(정의) 이 규정의 목적상 사용하는 용어의 뜻은 다음과 같다. (1) "인공지능시스템"이란 배포 이후에 적응 성을 보이고 명시적 또는 묵시적 목적을 위하여 물리 환경이나 가상 환경에 영향을 미칠 수 있는 예측, 콘텐츠, 권고[추천]나 결정 등의 산출물을 생성하는 방법을 입력을 통하여 추론할 수 있는 다양한 수준의 자율성을 가지고 작동하도록 설계된 기계 기반 시스템을 말한다.

특히, 인공지능시스템에 대한 정의에서 “다양한 수준의 자율성"을 부가함으로써, 인공지능의 자율성의 변화가능성을 고려하였고, 완전한 자율성뿐만 아니라 일정 정도의 자율성을 갖는 기계 기반 시스템도 인공지능에 포함시키고 있다. 또한 인공지능은 인간이 설정한 명시적 또는 묵시적 목적을 위해 수신된 입력으로부터 결정을 내리고 추천하는 시스템임을 선언하고 있다.

인공지능에 대한 정의는 인공지능법에서 주의를 기울여 규정해야 하는 사항이다. 국회 발의 법안들은 지능정보화 기본법의 “지능정보기술"의 정의 조항 일부에 ‘지각', ‘언어의 이해' 등을 덧대는 방식으로 규정하는 데 그치고 있다.[2] 이는 인공지능의 고유한 특성을 포착하지 못 하고 있으며, 인공지능시스템의 다양한 수준의 자율성과 그 변화 양상을 포괄하지 못 하고 있다.

유럽연합 AI Act의 입법례에 참고하여, 시민사회가 제안하는 인공지능의 정의는 다음과 같다.

> “인공지능”이란 다양한 수준의 자율성을 가지고 작동하도록 설계되어, 특정 목적들을 위해 실제 또는 가상 환경에 영향을 미치는 콘텐츠를 생산하거나 예측, 권고 또는 결정을 내릴 수 있는 시스템을 말한다.

---

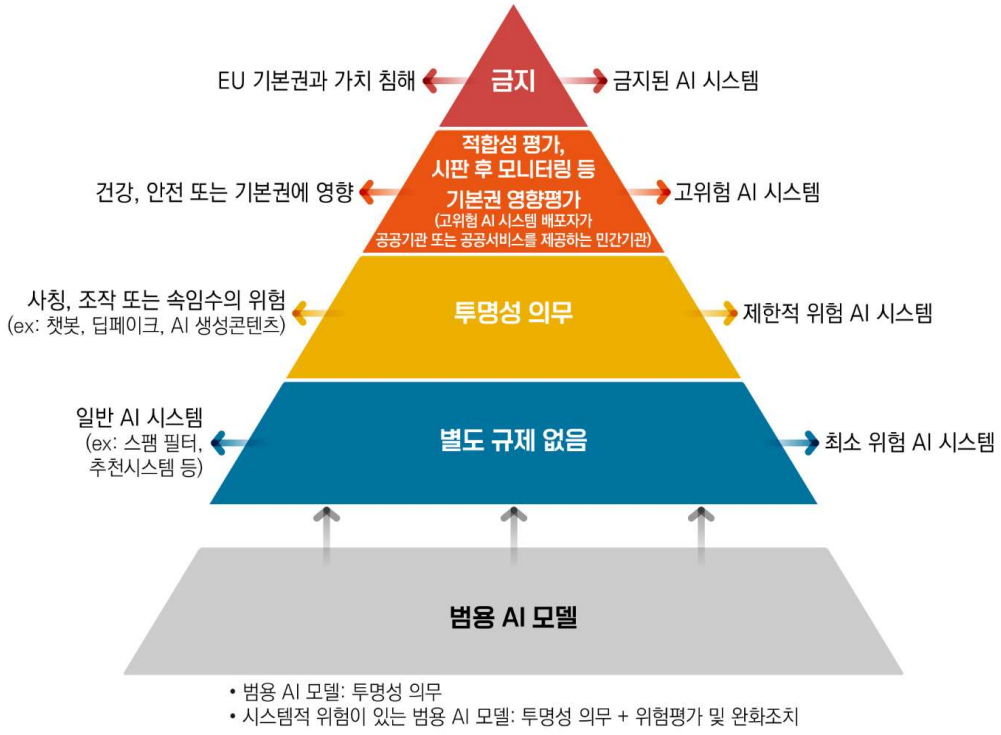
2) 지능정보화 기본법 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.
4. “지능정보기술” 이란 다음 각 목의 어느 하나에 해당하는 기술 또는 그 결합 및 활용 기술을 말한다.
가. 전자적 방법으로 학습 · 추론 · 판단 등을 구현하는 기술

2. 금지 인공지능 규율의 부재

(1) 금지 인공지능 규율의 필요성

수용불가한 위험도를 가진 인공지능에 대한 명시적 금지는 인공지능법이 필수적으로 포함해야 하는 사항이다. 유럽연합의 AI ACT는 용인할 수 없는 인공지능 시스템에 대한 개발 및 활용을 금지하고 있다.

① 인간의 잠재의식 또는 특정 집단의 취약점을 악용하여 피해를 유발할 수 있는 시스템, ② 사회적 행동이나 개인의 특성에 기반하여 생성 · 수집된 정보로 개인이나 집단의 '사회적 점수(social score)를 도출하여 불리한 대우를 유발할 수 있는 시스템, ③ 프로파일링이나 성격 특성만을 토대로 개인의 범죄 가능성을 예측 · 평가하는 시스템, ④ 직장 및 교육기관에서 개인의 감정 추론을 위해 사용하는 시스템, ⑤ 공공장소에서의 실시간 원격생체 인식 시스템(납치 · 인신매매 · 성적 착취 피해자 및 실종자 수색, 테러 위협 방지, 범죄자 및 범죄 용의자 신원 및 위치 파악을 위해 예외적으로 허용됨) 등이다.[3)]



(그림) 유럽연합 인공지능법상 AI 시스템의 분류 및 규제 체계[4)]

3) 심소연, 규제중심의 유럽연합 인공지능법(EU AI Act), 최신 외국입법정보 통권 제242호, 2024, 4-5면.

(2) 법안의 문제점

22대 국회에서 발의된 인공지능법안들은 금지 인공지능을 규율하는 규정을 찾아볼 수 없다. 한국 사회에서 어떠한 인공지능 시스템이 용인될 수 없는지에 대한 숙의와 합의가 필요하다. 현재 법안들은 금지에 대한 규율에 침묵함으로써 국민의 권리를 침탈하거나 민주적 기본질서에 위협이 되는 금지 인공지능을 사실상 용인하고 있다. 21대에 유력하게 논의되던 법안들(특히 2023. 2. 14. 과방위 법안심사소위를 통과한 것으로 알려진 소위 '소위원장 대안')에서 특히 문제되었던 '우선허용 사후규제' 조항은 22대 국회의 법안에서는 대부분 삭제되었지만(김성원 의원안은 포함), 금지 인공지능 규율이 부재하다는 관점에서 보면 이 '포괄적 네거티브 규제원칙'은 법안의 지도원리로 여전히 작동하고 있다 해도 과언이 아니다.

(3) 금지 인공지능 규율의 제정방향

한편, 금지된 인공지능에 대한 규율이 21대 국회에서 법안으로 제안된 바 있었다. 안철수 의원이 대표발의한 「인공지능 책임 및 규제법안」(의안번호: 2123709)은 인공지능의 유형을 금지된 인공지능·고위험 인공지능·저위험 인공지능으로 구분하면서, 금지된 인공지능 이외의 인공지능 개발 및 이용에 대하여 우선허용·사후규제 원칙을 적용하도록 하고 있다(안 제5조).[5)]

인공지능 책임 및 규제법안(안철수의원안)

제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.
5. "금지된 인공지능"이란 다음 각 목의 어느 하나에 해당하는 것으로서 인류평화, 인간의 존엄성, 자유와 평등, 민주주의, 기본권에 대한 침해의 위협이 명백하다고 간주되는 등 대통령령으로 정하는 기준에 해당하는 인공지능을 말한다.

---

4) 심소연, 규제중심의 유럽연합 인공지능법(EU AI Act), 최신 외국입법정보 통권 제242호, 2024, 3면.

5) 해당 법안의 검토보고서에 따르면, 금지된 인공지능의 개발·이용을 금지하면서, 제6조 각 호에 한해 예외적으로 허용하고 있음을 지적하면서, 이러한 예외적 허용에 대한 인·허가나 관리·감독 수단을 별도로 규정하고 있지 않으므로 관련 규정 마련을 검토할 필요가 있다고 지적하고 있다. 또한 국가정보원은 금지된 인공지능의 범위에 국가안보 업무를 수행하는 기관이 안보침해세력의 활동을 확인·견제·차단하기 위한 경우도 포함되는 문제가 있으므로 해당 부분을 제외할 필요가 있다는 의견을 제출한 바 있으며, 이에 대해 과학기술정보통신부는 동법에 따른 '금지된 인공지능'은 인류평화, 인간의 존엄성, 자유와 평등 등에 대한 침해의 위협이 명백하다고 간주되는 인공지능을 의미하는 것으로, 국가안보 침해 세력의 활동에 대한 확인·견제·차단 목적의 인공지능은 금지된 인공지능에 해당한다고 보기 어렵다는 입장을 밝혔다. 인공지능 책임 및 규제법안 검토보고, 과학기술정보방송통신위원회, 2023, 22-23면.

가. 사람의 신체적 또는 정신적 피해를 유발하거나 유발할 가능성이 있는 방식으로 특정인의 생각과 행동을 현저하게 왜곡하기 위하여 잠재의식을 활용하는 인공지능

나. 사람의 신체적 또는 정신적 피해를 유발하거나 유발할 가능성이 있는 방식으로 특정인의 생각과 행동을 현저하게 왜곡하기 위하여 그가 속한 집단의 취약성을 활용하는 인공지능

다. 본래 데이터가 생성 또는 수집된 상황과 관련이 없는 특정인이나 그 집단에 대하여 해롭거나 불리한 처우 또는 사회적 행동이나 그 심각성에 비추어 부당하거나 불균형적이며 불리한 처우를 초래하는 사회적 점수를 산정하여 특정 기간 동안 특정인이나 그 집단의 신뢰성을 평가 또는 분류할 목적으로 사용하는 인공지능

라. 법 집행을 위하여 공개적으로 접근 가능한 공간에서 실시간 원격 생체인식을 활용하는 인공지능

제6조(금지된 인공지능의 원칙적 금지) 누구든지 금지된 인공지능을 개발하거나 이용하여서는 아니 된다. 다만, 대통령령으로 정하는 바에 따라 다음 각 목의 어느 하나에 해당하는 경우에는 그러하지 아니하다.

1. 범죄피해자 및「실종아동등의 보호 및 지원에 관한 법률」제2조제2호에 따른 실종아동등에 대한 수색
2. 특정인의 생명이나 신체적 안전 또는 테러 공격에 대한 실질적이고 임박한 위협의 예방
3. 범인 추적 등 범인 검거를 위한 검거활동에 반드시 필요한 경우

유럽연합 AI Act는 다음과 같이 자세하게 용인할 수 없는 인공지능을 금지하고 있다.

CHAPTER II PROHIBITED ARTIFICIAL INTELLIGENCE PRACTICES	제II장 금지된 인공지능 관행
Article 5 Prohibited AI Practices	제5조(금지된 인공지능 관행)
1. The following AI practices shall be prohibited:	1. 다음과 같은 인공지능 관행은 금지된다.
(a) the placing on the market, the putting into service or the use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of, materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing a person to take a decision that that person would not have	(a) 개인이나 집단이 정보에 입각한 결정을 할 수 있는 능력을 눈에 띄게 침해하여 해당 개인이나 집단의 행동을 실질적으로 왜곡하고 그에 따라 개인이 해당 개인, 다른 개인 또는 집단에 상당한 피해를 유발하거나 유발할 가능성 있는 방식으로 결정을 하지 아니하였을 결정을 하게 할 목적으로 사람의 의식을 초월한 잠재의식에 영향을 미치는 기술 또는 조작이나 기만을 위한 기술을 사용하는 인공지능시스템의 시장 출시, 서비스 제공 또는 사용

otherwise taken in a manner that causes or is likely to cause that person, another person or group of persons significant harm;	
(b) the placing on the market, the putting into service or the use of an AI system that exploits any of the vulnerabilities of a person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm;	(b) 개인이나 다른 사람에게 상당한 피해를 유발하거나 합리적으로 유발할 가능성이 있는 방식으로 개인이나 특정한 집단에 속하는 개인의 행동을 실질적으로 왜곡할 목적으로 해당 개인이나 집단의 연령, 장애나 특정한 사회적 또는 경제적 상황으로 인한 취약성을 악용하는 인공지능시스템의 시장 출시, 서비스 제공 또는 사용
(c) the placing on the market, the putting into service or the use of AI systems for the purpose of the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following:	(c) 다음 중 어느 하나 또는 모두를 초래하는 사회적 점수를 통하여 자연인이나 집단의 사회적 행동 또는 알려지거나 추론되거나 예상되는 성격이나 특성에 기초하여 특정한 기간 동안 자연인이나 집단을 평가하거나 분류하기 위한 인공지능시스템의 시장 출시, 서비스 제공 또는 사용
(i) detrimental or unfavourable treatment of certain natural persons or whole groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;	(i) 데이터가 최초로 생성되거나 수집된 상황과 관계가 없는 사회적 상황에서 특정한 자연인 또는 전체 집단에 유해하거나 불리한 처우
(ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity;	(ii) 특정한 자연인 또는 전체 집단의 사회적 행동이나 그 심각성에 비하여 부당하거나 불균형한 해당 자연인 또는 전체 집단에 유해하거나 불리한 처우
(d) the placing on the market, the	(d) 자연인의 자료수집분석 또는 그 성격

putting into service for this specific purpose, or the use of an AI system for making risk assessments of natural persons in order to assess or predict the likelihood of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity;	및 특징에 대한 평가만에 기초하여 자연인이 범죄를 저지를 가능성을 평가하거나 예상하기 위하여 자연인의 위험평가를 실시하기 위한 인공지능시스템의 시장 출시, 특정 목적을 위한 서비스 제공 또는 사용. 이 금지는 범죄활동과 직접적으로 관련된 객관적이고 검증 가능한 사실에 기초한 사람의 범죄활동 연루에 대한 인간의 평가를 지원하기 위하여 사용된 인공지능시스템에는 적용되지 아니한다.
(e) the placing on the market, the putting into service for this specific purpose, or use of AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage;	(e) 인터넷이나 폐쇄회로텔레비전 영상의 얼굴 이미지를 임의로 수집하여 안면 인식 데이터베이스를 생성하거나 확장하는 인공지능시스템의 시장 출시, 특정 목적을 위한 서비스 제공 또는 사용
(f) the placing on the market, the putting into service for this specific purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons.	(f) 의료 또는 안전상의 이유로 시장에 인공지능시스템을 사용하거나 시장에 출시하는 경우를 제외하고, 직장 및 교육기관에서 자연인의 감정을 추론하기 위한 인공지능시스템의 시장 출시, 특정 목적을 위한 서비스 제공 또는 사용
(g) the placing on the market, the putting into service for this specific purpose, or the use of biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not	(g) 인종, 정치적 의견, 노동조합 가입, 종교적 또는 철학적 신념, 성생활 또는 성적지향을 추론하기 위하여 생체인식데이터를 기반으로 자연인을 개별적으로 분류하는 생체인식분류시스템의 시장 출시, 특정 목적을 위한 서비스 제공 또는 사용. 이 금지사항은 법집행 분야에서의 생체인식데이터 또는 생체인식데이터 분류에 기반한 이미지 등 적법하게 취득한 생체인식데이터의 라벨링

cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement;	또는 필터링에는 적용되지 아니한다.
(h) the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives:	(h) 법집행을 목적으로 하는 공개적으로 접근 가능한 공간에서의 실시간 원격생채식별시스템 사용. 다만 그 사용이 다음의 목적 중 어느 하나에 엄격하게 필요한 경우에는 그러하지 아니하다.
(i) the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as searching for missing persons;	(i) 특정한 유괴, 인신매매 또는 성 착취 피해자의 표적 수색과 실종자 수색
(ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack;	(ii) 자연인의 생명이나 신체적 안전에 대한 구체적이고 실질적이며 임박한 위협 또는 진정하고 현존하거나 진정하고 예측 가능한 테러공격 위협의 방지
(iii) the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation, prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years;	(iii) 부속서 II에서 규정하고 관련 회원국에서 4년 이하의 징역형 또는 구금명령에 따라 처벌될 수 있는 범죄에 대한 수사, 기소나 형사처벌의 집행을 목적으로 용의자나 범죄자의 소재 파악 또는 식별
Point (h) of the first subparagraph is without prejudice to Article 9 of Regulation (EU) 2016/679 for the processing of biometric data for purposes other than law enforcement.	이 항 제(h)호의 규정은 법집행 이외의 목적을 위한 생체인식데이터의 처리와 관련하여 규정(EU) 제2016/679호 제9조를 침해하지 아니한다.

2. The use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement for any of the objectives referred to in paragraph 1, point (h), shall be deployed only for the purposes set out in paragraph 1, point (h), to confirm the identity of the specifically targeted individual, and it shall take into account the following elements:	2. 제(1)항(h)에서 규정하는 목적 중 어느 하나에 대한 법집행을 목적으로 한 '실시간' 원격생체식별시스템은 구체적으로 대상이 된 개인의 신원을 확인하기 위하여 제(1)항(h)에 언급된 목적을 위해서만 사용하여야 하고 다음의 요소를 고려한다.
(a) the nature of the situation giving rise to the possible use, in particular the seriousness, probability and scale of the harm that would be caused if the system were not used;	(a) 사용 가능성을 유발하는 상황의 성질 특히 해당 시스템을 사용하지 아니할 경우 초래될 피해의 심각성, 확률 및 규모
(b) the consequences of the use of the system for the rights and freedoms of all persons concerned, in particular the seriousness, probability and scale of those consequences.	(b) 관련된 모든 자의 권리와 자유에 대한 해당 시스템의 사용 결과 특히 해당 결과의 심각성, 확률 및 규모
In addition, the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement for any of the objectives referred to in paragraph 1, point (h), of this Article shall comply with necessary and proportionate safeguards and conditions in relation to the use in accordance with national law authorising the use thereof, in particular as regards the temporal, geographic and personal limitations. The use of the 'real-time' remote biometric identification system in publicly accessible spaces shall be authorised only if the law enforcement authority has completed a fundamental rights impact assessment as provided for in Article 27 and has registered the system in the EU database	또한 이 조 제(1)항(h)에서 규정하는 목적 중 어느 하나에 대한 법집행을 목적으로 한 '실시간' 원격생체식별시스템의 사용은 특히 시간적, 지리적 및 개인적 제한과 관련하여 그 사용을 승인하는 국내법에 따라 사용과 관련된 필요하고 비례적인 안전장치 및 조건을 준수하여야 한다. 공개적으로 접근 가능한 공간에서의 '실시간' 원격생체식별시스템 사용은 법집행기관이 제27조에 언급된 기본권 영향평가를 완료하고 제49조에 따른 유럽연합 데이터베이스에 해당 시스템을 등록한 경우에만 승인된다. 다만 합리적인 사유가 있는 긴급한 경우에는 유럽연합 데이터베이스 등록하지 아니하고 해당 시스템의 사용을 개시할 수 있다. 이 경우 지체 없이 해당 등록을 완료한다.

according to Article 49. However, in duly justified cases of urgency, the use of such systems may be commenced without the registration in the EU database, provided that such registration is completed without undue delay.	
3. For the purposes of paragraph 1, point (h) and paragraph 2, each use for the purposes of law enforcement of a 'real-time' remote biometric identification system in publicly accessible spaces shall be subject to a prior authorisation granted by a judicial authority or an independent administrative authority whose decision is binding of the Member State in which the use is to take place, issued upon a reasoned request and in accordance with the detailed rules of national law referred to in paragraph 5. However, in a duly justified situation of urgency, the use of such system may be commenced without an authorisation provided that such authorisation is requested without undue delay, at the latest within 24 hours. If such authorisation is rejected, the use shall be stopped with immediate effect and all the data, as well as the results and outputs of that use shall be immediately discarded and deleted.	3. 제(1)항(h) 및 제(2)항의 목적상 법집행을 목적으로 공개적으로 접근 가능한 공간에서 '실시간' 원격생체식별시스템을 사용하려면 합리적이 요청이 있는 때에 제(5)항에서 정하는 국내법의 상세규칙에 따른 사법당국 또는 독립적 행정당국의 사전승인을 받아야 한다. 그 결정은 해당 시스템이 사용되는 회원국을 구속한다. 다만 합리적인 사유가 있는 긴급한 경우에는 승인 없이 해당 시스템의 사용을 개시할 수 있다. 이 경우 늦어도 24시간 이내에 부당한 지체 없이 승인을 요청한다. 승인이 거절되는 경우에는 즉시 사용을 중지하고 모든 데이터와 해당 사용의 결과 및 산출물을 폐기 및 삭제한다.
The competent judicial authority or an independent administrative authority whose decision is binding shall grant the authorisation only where it is satisfied, on the basis of objective evidence or clear indications presented to it, that the use of the 'real-time' remote biometric identification system concerned is necessary for, and proportionate to, achieving one of the	그 결정이 구속력을 갖는 관할 사법당국 또는 독립적 행정당국은 제출된 객관적 증거 또는 명확한 지표에 기초하여 관련된 '실시간' 원격생체식별시스템이 제(1)항(h)에서 규정하는 목적 중 어느 하나를 달성하는 데 필요하고 비례적이며 해당 요청에서 확인되고 특히 기간과 지리적 및 개인적 범위에 관하여 엄격하게 필요한 것으로 제한된다고 판단하는 경우에만 승인을 하여야 한다. 요청에 대한

objectives specified in paragraph 1, point (h), as identified in the request and, in particular, remains limited to what is strictly necessary concerning the period of time as well as the geographic and personal scope. In deciding on the request, that authority shall take into account the elements referred to in paragraph 2. No decision that produces an adverse legal effect on a person may be taken based solely on the output of the 'real- time' remote biometric identification system.	결정을 할 때에는 해당 당국은 제(2)항에 언급된 요소를 고려한다. 사람에게 부정적인 법적 효력을 미치는 결정은 '실시간' 원격생체식별시스템의 산출물만을 근거로 할 수 없다.
4. Without prejudice to paragraph 3, each use of a 'real-time' remote biometric identification system in publicly accessible spaces for law enforcement purposes shall be notified to the relevant market surveillance authority and the national data protection authority in accordance with the national rules referred to in paragraph 5. The notification shall, as a minimum, contain the information specified under paragraph 6 and shall not include sensitive operational data.	4. 제(3)항을 침해하지 아니하는 범위에서 법집행을 목적으로 공개적으로 접근 가능한 공간에서 '실시간' 원격생체식별시스템을 사용하려면 제(5)항에 언급된 국내규칙에 따라 관련 시장감시기관 및 국가데이터보호기관에 통지한다. 이 통지에는 최소한 제(6)항에 따라 규정하는 정보를 포함하고 민감한 운용정보를 포함해서는 아니 된다.
5. A Member State may decide to provide for the possibility to fully or partially authorise the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement within the limits and under the conditions listed in paragraph 1, point (h), and paragraphs 2 and 3. Member States concerned shall lay down in their national law the necessary detailed rules for the request, issuance and exercise of, as well as supervision and reporting relating to, the authorisations referred to in paragraph	5. 회원국은 한계 내에서 제(1)항(h), 제(2)항 및 제(3)항에 언급된 조건에 따라 법집행을 목적으로 공개적으로 접근 가능한 공간에서 '실시간' 원격생체식별시스템을 사용하는 것을 전부 또는 일부 승인할 가능성을 규정하기로 결정할 수 있다. 관련 회원국은 그 국내법에서 해당 요청, 제(3)항에 언급된 승인의 부여 및 실행과 해당 승인에 관한 감독 및 보고에 관하여 필요한 세부규칙을 정한다. 이 규칙은 또한 제(1)항(h)(iii)에 언급된 범죄를 포함한 같은 항 제(h)호에 언급된 목적과 관련하여 관할기관이 법집행을 목적으로 해당 시스템을 사용하도록 승인할 수 있음을 명시한다. 회원국은 채택 후 늦어도 30일 이내에 해당 규칙을 유럽연합집행위원회에

3. Those rules shall also specify in respect of which of the objectives listed in paragraph 1, point (h), including which of the criminal offences referred to in point (h)(iii) thereof, the competent authorities may be authorised to use those systems for the purposes of law enforcement. Member States shall notify those rules to the Commission at the latest 30 days following the adoption thereof. Member States may introduce, in accordance with Union law, more restrictive laws on the use of remote biometric identification systems.	통지한다. 회원국은 유럽연합 법령에 따라 원격생체식별시스템의 사용에 관한 보다 엄격한 법률을 도입할 수 있다.
6. National market surveillance authorities and the national data protection authorities of Member States that have been notified of the use of 'real-time' remote biometric identification systems in publicly accessible spaces for law enforcement purposes pursuant to paragraph 4 shall submit to the Commission annual reports on such use. For that purpose, the Commission shall provide Member States and national market surveillance and data protection authorities with a template, including information on the number of the decisions taken by competent judicial authorities or an independent administrative authority whose decision is binding upon requests for authorisations in accordance with paragraph 3 and their result.	6. 제(4)항에 따른 법집행을 목적으로 한 공개적으로 접근 가능한 공간에서의 '실시간' 원격생체식별시스템 사용을 통지받은 회원국의 국가 시장감시기관 및 국가 데이터보호기관은 그 사용에 관한 연례보고서를 유럽연합집행위원회에 제출한다. 이 목적을 위하여 유럽연합집행위원회는 그 결정이 제(3)항에 따른 승인 요청 및 그 결과에 대한 구속력을 갖는 관할 사법당국 또는 독립적 행정당국이 한 결정의 횟수에 관한 정보를 포함한 서식을 회원국과 국가 시장감시기관 및 데이터보호기관에 제공한다.
7. The Commission shall publish annual reports on the use of real-time remote biometric identification systems in publicly accessible spaces for law enforcement purposes, based on aggregated data in Member States on the basis of the annual reports	7. 유럽연합집행위원회는 제(6)항에서 규정하는 연례보고서를 기초로 회원국에서 집계된 데이터에 기반하여 법집행을 목적으로 한 공개적으로 접근 가능한 공간에서의 '실시간' 원격생체식별시스템 사용에 관한 연례보고서를 공표한다.

referred to in paragraph 6. Those annual reports shall not include sensitive operational data of the related law enforcement activities. 8. This Article shall not affect the prohibitions that apply where an AI practice infringes other Union law.	 8. 이 조는 인공지능 관행이 다른 유럽연합 법령을 위반하는 경우에 적용되는 금지에는 영향을 미치지 아니한다.

다음은 금지 인공지능에 대한 규율을 대한 조문 형태의 제안이다. 앞서 언급한대로, 무엇을 금지 인공지능으로 분류할 것인가의 문제는 광범위한 사회적 합의와 검토가 필요하다. 따라서 해당 제안은 잠정적이다.

제ㅇㅇ조 (인공지능 개발과 활용의 금지)
① 누구든지 다음 각 호의 사항에 해당하는 인공지능을 개발하거나 활용할 수 없다.
1. 성별·연령·장애·지역·인종·종교·국가 등 오로지 개인의 특성에 따라 합리적 이유없이 차별하거나 차별하려는 인공지능
2. 사회적 약자 및 취약 계층에 속한 사람의 행동을 중대하게 왜곡하려는 의도로 이들의 생명, 신체, 재산 등을 침해하는 인공지능
3. 인간의 심리, 사고, 행동 등을 왜곡하기 위한 의도로 개인의 잠재의식에 영향을 미치는 인공지능
4. 특정된 개인 또는 집단을 차별하거나 불이익을 줄 의도로 인간의 사회적 행동, 인격적 특성에 기초하여 개인의 신뢰도를 평가하거나 분류하는 인공지능
5. 공공장소에서 실시간 원격 신원확인을 위한 목적의 인공지능
6. 인간의 감정을 인식하거나 예측하여 개인의 권리와 의무에 중대한 영향을 미치기 위한 목적의 인공지능
7. 인간의 실질적 통제 없이 무기를 운용할 목적의 인공지능
8. 인간의 실질적 통제 없이 운용되는 예측 치안 및 경찰직무집행을 목적으로 하는 인공지능
9. 기타 인간의 존엄을 침해하는 인공지능으로서 대통령령으로 정하는 인공지능
② 인공지능 제공자는 제1항 각호에 해당하지 않는 인공지능을 개발 또는 제공하는 과정에서 해당 인공지능이 제1항 각호에 해당하게 되는 경우 즉시 해당 인공지능의 개발 또는 제공을 중단하여야 한다.
③ 인공지능 활용자는 인공지능 시스템이 제1항 각호에 해당하는 인공지능에 해당함을 알게 된 경우에는 즉시 활용을 중단하고 관계기관에 신고하여야 한다.

3. 고위험 인공지능 규율의 불완전성

고위험영역 인공지능 규율에서 중요한 점은 ① 고위험영역 인공지능의 규율 객체에 대한 구체적 규정 및 객체별 의무부과(제공자, 활용자, 최종이용자 등), ② 안전 그리고 인권에 미치는 위험에 대한 체계적 규정, ③ 필수적으로 포함되어야 할 고위험영역의 법률적 규정이다. 특히 경찰을 포함한 수사기관의 수사, 재판, 선거 등 주요 공공영역의 인공지능, 출입국 관리, 산업안전, 고용관계, 학교교육, 신용평가 영역 등은 반드시 법률적으로 규율해야 한다. 하위입법으로 위임하더라도 법률에서 고위험영역이 무엇인지 예측가능해야 한다.

"고위험영역 인공지능을 시장에 출시하는 제공자는 물론 이를 업무용으로 도입하는 활용자 모두, 사전에 인공지능의 위험을 방지하거나 완화하는 조치를 취하도록 강력한 의무를 부과하여야 한다. 이러한 고위험영역의 의무는 사후에 안전과 인권을 침해하는 사고가 발생할 경우 그 책임 문제에 대하여 조사하고 구제하는 조치를 뒷받침할 수 있어야 한다. 공공기관에 도입되는 인공지능의 경우 고위험영역에 준하는 위험 방지 및 완화 의무가 부과되는 것이 바람직하다.
구체적으로 고위험영역 인공지능 제공자의 경우, 위험 관리, 데이터셋 관리, 기술문서와 로그기록 작성, 정보를 제공하는 투명성, 사람의 관리감독, 견고성, 정확성, 사이버보안을 보장하도록 하고, 시장 출시 전에 적합성 평가와 인증 절차를 이행하여야 한다. 시장출시 후에도 제대로 작동하고 있는지 모니터링하고, 중대한 문제가 발생할 경우 당국에 신고해야 한다. 고위험영역 및 공공기관 활용자에 대해서는 인권영향평가를 의무화하고 그 주요사항을 공공적으로 등록하여 일반에 공개하도록 해야 한다.
유럽연합 AI ACT는 물론 미국 OMB 규칙의 경우에도, 고위험을 안전과 인권 영역 별로 체계적으로 구분하여 규정하였다. 유럽연합과 미국 모두 고위험영역 인공지능에 대하여 영향평가를 사전에 실시하고 식별되는 위험을 방지하거나 완화하는 조치를 의무화하였다. 일정 수준의 위험 완화 수준에 도달하지 못하거나 도입 전 엄격한 테스트를 통과하지 못하는 고위험영역 인공지능은, 유럽의 경우 시장에 출시되지 못하고 미국의 경우 연방정부에 조달되지 못한다. 또한 고위험영역 인공지능의 개발과정에 대한 문서를 작성보관하고, 데이터 평가 결과 드러난 편향에 대하여 조치하며, 사람이 관리감독하도록 하는 것 역시 공통적인 의무로 규정되어 있다.
국가인권위원회는 고위험 영역 인공지능의 경우, 21대의 과방위 법안보다 범위를 확대하여 재정의하고, 알고리즘의 투명성과 설명가능성, 인권침해·차별 예방 조치 여부 등을 사전에 엄격히 점검할 것을 권고하였다. 더불어 인권영향평가 제도를 도입하여 인공지능 개발·출시 전 인권영향평가를 실시하고, 출시 후 기능 수정 및 활용 범위 변경 시 재평가를 하도록 요구하였다."[6)]

6) 22대 인공지능법 제정에 대한 시민사회 의견서

22대 국회에서 발의되어 있는 법안들은 모두 “고위험영역 인공지능" 또는 “고위험 영역에서 활용되는 인공지능"(김성원의원안)을 정의하고 있다.

김성원의원안의 경우 고위험 영역에서 “활용”하는 인공지능으로 한정함으로써 개발 단계 고위험 인공지능의 규율을 회피하고 있다.

법안 모두 고위험영역을 정의하고 있기는 하지만, 안전 및 인권에 미치는 위험을 체계적으로 규율하고 있다기보다 산발적·제한적으로 규정하고 있다. 에너지, 먹는물 등의 공급, 보건의료의 제공 및 의료체계, 의료기기, 핵물질과 원자력시설 등에서 사용하는 인공지능을 고위험영역으로 분류하고 있는데, 수사기관 수사 일반, 재판, 선거, 출입국 관리 등 주요 공공영역은 물론 산업안전, 일반 고용관계, 학교 교육, 신용평가 등은 규율하고 있지 않다.

특히, 인권 위험과 관련하여, ‘채용, 대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능’을 고위험영역으로 분류하고 있는데, ‘채용, 대출 심사 등’이라는 제한적 예시를 통해 ‘개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능’이 무엇인지 예측가능한지 의문이다. 또한 ‘그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능’이라 규정함으로써 기본권 보호에 관한 사항을 광범위하게 하위입법에 포괄적으로 위임하고 있는데 이 또한 대통령령에 어떤 사항이 포함될지 예측가능하지 않다.

> “고위험영역 인공지능 제공자에게는 이것이 고위험이라는 고지 의무만이 부과되어 있다. 또한 위험관리방안, 기술문서 작성보관, 인공지능결과물 설명, 이용자 보호, 사람의 관리감독 등의 일부 조치조차 제공자가 자율적인 ‘방안’을 마련하는 책무에 그쳐 있고 책무 위반에 대한 처벌규정을 두지 않아 아무런 실효성을 갖추지 못했다.
> 더불어 인공지능 제품이나 서비스를 개발하거나 이를 시장에 제공하는 사업자를 넘어, 이를 제공받아 업무에 활용하는 인공지능 활용자에 대해서는 명시적인 의무나 책무를 규정하지 않았다. 예를 들어 금융기관이 대출심사 AI를 공급받아 금융소비자를 대상으로 활용하거나 의료기관이 의료진단 AI를 공급받아 환자를 대상으로 활용할 때 이들 기관의 의무와 책임이 규정되어 있지 않은 것이다. 더불어 제공자와 활용자 모두 위험 평가, 데이터 평가, 로그기록 보관, 사전적합성 평가 또는 인권영향평가, 이용자 또는 영향을 받는 사람에 대한 설명, 출시후 모니터링, 중대한 사고 보고 등 인공지능의 위험을 사전에 평가하고 완화하며, 도입 이후 작동을 모니터링하기 위한 조치 의무를 지고 있지 않았다.”[7]

7) 22대 인공지능법 제정에 대한 시민사회 의견서

고위험영역 인공지능 관련 정의 규정은 다음과 같다.

정점식 의원안	조인철 의원안	김성원 의원안	안철수 의원안	민형배 의원안
제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 3. “**고위험영역 인공지능**”이란 다음 각 목의 어느 하나에 해당하는 인공지능으로서 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용되는 인공지능을 말한다. 가. 「에너지법」 제2조제1호에 따른 에너지, 「먹는물관리법」 제3조제1호에 따른 먹는물 등의 공급을 위하여 사용되는 인공지능 나. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계 등에 사용되는 인공지능 다. 「의료기기법」 제2조제1항에 따른 의료기기에 사용되는 인공지능 라. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항 제1호 및 제2호에 따른 핵물질과 원자력시설의 안전한 관리 및 운영을 위하여 사용되는 인공지능 마. 범죄 수사나 체포 업무에 있어 생체정보(얼굴 · 지문 · 홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적 · 생리적 · 행동적 특징에 관한 개인정보를	제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 3. “고위험영역 인공지능”이란 다음 각 목의 어느 하나에 해당하는 인공지능으로서 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 인공지능을 말한다. 가. 「에너지법」 제2조제1호에 따른 에너지, 「먹는물관리법」 제3조제1호에 따른 먹는물 등의 공급을 위하여 사용되는 인공지능 나. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계 등에 사용되는 인공지능 다. 「의료기기법」 제2조제1항에 따른 의료기기에 사용되는 인공지능 라. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항에 따른 핵물질과 원자력시설의 안전한 관리 및 운영을 위하여 사용되는 인공지능 마. 범죄 수사나 체포 업무에 있어 생체정보(행정기관이 보유하고 있는 얼굴 · 지문 · 홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적 특징에 관한 개인정보를	제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 3. “**고위험 영역에서 활용되는 인공지능**”이란 다음 각 목의 어느 하나에 해당하는 인공지능으로서 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 인공지능을 말한다. 가. 「에너지법」 제2조제1호에 따른 에너지, 「먹는물관리법」 제3조제1호에 따른 먹는물 등의 공급을 위하여 사용되는 인공지능 나. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계 등에 사용되는 인공지능 다. 「의료기기법」 제2조제1항에 따른 의료기기에 사용되는 인공지능 라. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항에 따른 핵물질과 원자력시설의 안전한 관리 및 운영을 위하여 사용되는 인공지능 마. 범죄 수사나 체포 업무에 있어 생체정보(행정기관이 보유하고 있는 얼굴 · 지문 · 홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적 특징에 관한 개인정보를	제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 4. “고위험영역 인공지능”이란 다음 각 목의 어느 하나에 해당하는 인공지능으로서 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용되는 인공지능을 말한다. 가. 「에너지법」 제2조제1호에 따른 에너지, 「먹는물관리법」 제3조제1호에 따른 먹는물 등의 공급을 위하여 사용되는 인공지능 나. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계 등에 사용되는 인공지능 다. 「의료기기법」 제2조제1항에 따른 의료기기, 「디지털의료제품법」 제2조제2호에 따른 디지털의료기기에 사용되는 인공지능 라. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항 제1호 및 제2호에 따른 핵물질과 원자력시설의 안전한 관리 및 운영을 위하여 사용되는 인공지능 마. 생체정보(얼굴 · 지문 · 홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적 · 생리적 · 행동적 특징에 관한 개인정보를 말한다)를 처리하는 데 사용되는	제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 3. “고위험영역 인공지능”이란 다음 각 목의 어느 하나에 해당하는 인공지능으로서 사람의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용되는 인공지능을 말한다. 가. 「에너지법」 제2조제1호에 따른 에너지, 「먹는물관리법」 제3조제1호에 따른 먹는물 등의 공급을 위하여 사용되는 인공지능 나. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공 및 이용체계 등에 사용되는 인공지능 다. 「의료기기법」 제2조제1항에 따른 의료기기에 사용되는 인공지능 라. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항 제1호 및 제2호에 따른 핵물질과 원자력시설의 안전한 관리 및 운영을 위하여 사용되는 인공지능 마. 범죄 수사나 체포 업무에 있어 생체정보(얼굴 · 지문 · 홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적 · 생리적 · 행동적 특징에 관한 개인정보를

말한다)를 분석 · 활용하는 데 사용되는 인공지능 바. 채용, 대출 심사 등 개인의 권리 · 의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능 사. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영에 사용되는 인공지능 아. 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」에 따른 공공기관 등(이하 "국가기관등"이라 한다)이 공공서비스 제공에 필요한 자격 확인, 결정 또는 비용징수 등에 사용하는 인공지능으로서 국민에게 영향을 미치는 의사결정을 위하여 사용되는 인공지능 자. 그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능	말한다)를 분석 · 활용하는 데 사용되는 인공지능 바. 채용, 대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능 사. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영에 사용되는 인공지능 아. 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」에 따른 공공기관 등(이하 "국가기관등"이라 한다)이 사용하는 인공지능으로서 국민에게 영향을 미치는 의사결정을 위하여 사용되는 인공지능 자. 그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능	말한다)를 분석 · 활용하는 데 사용되는 인공지능 바. 채용, 대출 심사 등 개인의 권리 · 의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능 사. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영에 사용되는 인공지능 아. 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」에 따른 공공기관 등(이하 "국가기관등"이라 한다)이 사용하는 인공지능으로서 국민에게 영향을 미치는 의사결정을 위하여 사용되는 인공지능 자. 그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능	인공지능 바. 채용, 신용등급평가, 대출 심사 등 개인의 권리 · 의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능 사. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설 및 교통체계의 주요한 작동 · 운영에 사용되는 인공지능 아. 공공기관 등이 공공서비스 제공에 필요한 자격 확인, 결정 또는 비용징수 등에 사용하는 인공지능으로서 국민에게 영향을 미치는 의사결정을 위하여 사용되는 인공지능 자. 그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능	말한다)를 분석 · 활용하는 데 사용되는 인공지능 바. 채용, 대출 심사 등 개인의 권리 · 의무 관계에 중대한 영향을 미치는 판단 또는 평가 목적의 인공지능 사. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영에 사용되는 인공지능 아. 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」에 따른 공공기관 등(이하 "국가기관등"이라 한다)이 사용하는 인공지능으로서 국민에게 영향을 미치는 의사결정을 위하여 사용되는 인공지능 자. 그 밖에 국민의 안전 · 건강 및 기본권 보호에 중대한 영향을 미치는 인공지능으로서 대통령령으로 정하는 인공지능

4. 제재 규정의 미비 문제

인공지능법은 금지 인공지능, 고위험 인공지능 등을 정의하고 이를 금지하고 또는 차등화된 의무 부과하여야 한다. 또한 실효성을 확보하기 위하여 다양한 수준의 제재 규정을 포함하여야 한다.

유럽연합 AI Act의 경우, 용인할 수 없는 인공지능을 금지하고 이를 위반하는 경우 전세계 연간 매출액의 최대 7%의 과징금을 부과한다. 고위험 인공지능 시스템에 대한 의무를 준수하지 않는 경우에는 전세계 연간 매출액의 최대 3%의 과징금이 부과된다.

미국의 AI 행정명령의 경우, 명시적 금지나 제재 규정이 없지만 연방정부 AI가 차별금지명령에서 금지하는 불법적인 차별이나 유해한 편견을 초래하는 경우 해당정보의 사용을 중단해야 한다.

국회 발의 법안들의 경우, 금지 인공지능에 대해서는 앞서 언급한대로 정의 규정 자체가 없어 그에 대한 제재 규정도 없다. 고위험영역 인공지능의 경우, 해당 인공지능 제공자의 사전 고지 의무만이 부과되어 있을 뿐이고, 위험관리방안, 기술문서 작성보관, 인공지능결과물 설명, 이용자 보호, 사람의 관리감독 등의 일부 조치는 제공자가 자율적인 '방안'을 마련하는 책무를 부담하는 데 그치고 있으며, 활용자(사업자에 의해 개발되거나 제공된 인공지능을 받아 업무에 활용하는 경우를 말함)에 대해서는 어떠한 의무도 부과하지 않고 있다. 부과된 의무에 대한 실효성을 담보할 수 있는 제재 규정은 존재하지 않는다.

법안은 국가인권위원회 위원의 직무상 비밀 위반, 직무상 목적 외의 용도 사용한 자 등에 대한 벌칙 조항만을 두고 있다.

한편, 발의된 모든 법안들은 소위 "윤리 프레임"에 갇혀 있는 것으로 보인다. 과기정통부와 산하 연구원 등이 제시하였거나 추진하고 있는 인공지능 윤리 가이드라인 등을 법률적으로 정당화하는 법제로 보인다. 그러나 이러한 인공지능 윤리 프레임은 역으로 인공지능이 인간과 인류에 미치는 위험성을 윤리적으로 과소평가하고 있으며, 안전과 인권에 위험을 실질적으로 예방하고 완화할 수 있는 제도적 장치들을 "자율적 윤리"의 이름으로 우회하고 잠탈할 수 있다. 인공지능의 개발과 활용에서 필요한 규범은 윤리적 가이드라인이기도 하여야 하지만 실효성을 실질적으로 보장하는 제도적·법적 장치이기도 하여야 한다.

5. 범용 인공지능에 대한 규율 문제

'범용 인공지능'(General Purpose AI, GPAI) 모델이란, 대규모 자기지도학습(self-supervision)을 사용하여 대량의 데이터로 학습된 경우를 비롯하여 상당한 일반성을 나타내며, 모델이 시장에 출시되는 방식에 관계없이 광범위한 고유 작업을 능숙하게 수행할 수 있고, 다양한 다운스트림 시스템 또는 애플리케이션에 통합될 수 있는 AI 모델을 의미한다.[8] chat-gpt(특히, GPT-4o)가 등장하면서 생성형 인공지능뿐만 아니라 범용 인

8) 22대 인공지능법 제정에 대한 시민사회 의견서

공지능에 대한 사회적 관심과 법적 규제의 필요성도 커지고 있다.

인공지능법은 범용 인공지능에 대비하여 이에 대한 정의와 위험 방지 대책을 포함해야 한다. 특히 위험성이 높은 범용 인공지능에 대한 레드티밍과 관리를 의무화해야 한다.

유럽연합의 경우 범용 인공지능 일반에 대하여 기술문서 작성보관, 정보를 제공하는 투명성, 당국에 대한 협력 의무, 훈련 콘텐츠의 요약본 공개, 저작권법 준수 등을 의무화하였다. 범용 인공지능 중 부동소수점 연산 10^25를 초과하는 등 일정 수준 이상으로 시스템적 위험이 높은 경우 적대적 테스트 등을 의무화하고 사고에 대한 국가 보고 및 사이버 보안을 의무화하였다. 미국 AI 행정명령의 경우에도 범용 인공지능 등 강력한 인공지능 시스템에 대하여 안전 평가 결과와 중요 정보를 정부와 공유하도록 의무화하였다.[9]

유럽연합법에서 범용 인공지능의 정의는 다음과 같다.

Article 51 Classification of general-purpose AI models as general-purpose AI models with systemic risk	제51조(범용인공지능모델의 시스템적 위험이 있는 범용인공지능모델로의 분류)
1. A general-purpose AI model shall be classified as a general-purpose AI model with systemic risk if it meets any of the following requirements:	1. 범용인공지능모델이 다음의 요건 중 어느 하나를 충족하는 경우에는 시스템적 위험이 있는 범용인공지능모델로 분류된다.
(a) it has high impact capabilities evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks;	(a) 지표 및 척도를 포함하여 적절한 기술적 도구 및 방법론에 기반하여 평가한 고영향력이 있을 것
(b) based on a decision of the Commission, ex officio or following a qualified alert from the scientific panel, it has capabilities or an impact equivalent to those set out in point (a) having regard to the criteria set out in Annex XIII.	(b) 유럽연합집행위원회의 결정에 기초하여 직권으로 또는 과학심사단의 적격한 경고 이후 부속서 XIII에서 정하는 기준을 고려할 때 제(a)호에서 정하는 사항과 동등한 능력 또는 영향력이 있을 것
2. A general-purpose AI model shall be presumed to have high impact	2. 범용인공지능모델은 부동소수점연산에서 측정된 그 훈련에 사용되는 누적 계산량이

9) 22대 인공지능법 제정에 대한 시민사회 의견서

capabilities pursuant to paragraph 1, point (a), when the cumulative amount of computation used for its training measured in FLOPs is greater than 10^25.	10^25보다 클 때에는 제(1)항(a)에 따른 고 영향력이 있는 것으로 추정된다.
3. The Commission shall adopt delegated acts in accordance with Article 97 to amend the thresholds listed in paragraphs 2 and 3 of this Article, as well as to supplement benchmarks and indicators in light of evolving technological developments, such as algorithmic improvements or increased hardware efficiency, when necessary, for these thresholds to reflect the state of the art.	3. 유럽연합집행위원회는 이 조 제(2)항 및 제(3)항에서 정하는 기준치를 변경하고 필요한 경우에는 알고리즘 향상 또는 하드웨어 효율성 증대 등의 발전하는 기술 개발을 고려하여 최신기술을 반영하는 이러한 기준치에 대한 척도 및 지표를 보완하기 위하여 제97조에 따른 위임입법을 채택한다.

일부 법안(정점식, 안철수, 민형배의원안)의 경우 '생성형 인공지능'을 정의하고, 이에 대해 고지 및 표시 의무, 안전 확보 의무를 부과하고 있다. 다만 위반시 제재 규정은 없어 실효성을 담보할 수 없다. 그나마 고지 및 표시 의무 부과 규정은 정점식의원안에서만 두고 있다.

모든 법안에서 범용 인공지능(또는 파운데이션 모델)에 대한 규정은 찾아볼 수 없다.

정점식의원안

제29조(생성형 인공지능 고지 및 표시) ① 생성형 인공지능을 이용하여 제품 또는 서비스를 제공하려는 자는 해당 제품 또는 서비스가 생성형 인공지능에 기반하여 운용된다는 사실을 이용자에게 사전에 고지하고, 해당 제품 또는 서비스의 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 한다.

② 제1항에 따른 고지 및 표시에 필요한 사항은 대통령령으로 정한다.

제30조(생성형 인공지능 안전 확보 의무) ① 학습에 사용된 누적 연산량이 대통령령으로 정하는 기준 이상인 생성형 인공지능을 이용하여 제품 또는 서비스를 제공하려는 자는 인공지능안전을 확보하기 위하여 다음 각 호의 사항을 이행하여야 한다.

1. 인공지능 수명주기 전반에 걸친 위험 식별, 평가, 완화
2. 인공지능 관련 안전사고를 모니터링하고 대응하는 위험관리체계 구축

② 제1항에 따른 생성형 인공지능을 개발하려는 자는 제1항 각 호에 따른 사항의 이행 결과를 과학기술정보통신부장관에 제출하여야 한다.

③ 과학기술정보통신부장관은 제1항 각 호에 따른 사항의 구체적인 이행 방식 및 제2항에 따른 결과 제출 등에 필요한 지침을 정하여 고시하여야 한다.

6. 인공지능 국가-지역 거버넌스 체계

인공지능법에서 중요하게 다루어야 하는 사항은 인공지능 감독 및 통제에 관한 국가-지역 거버넌스 체계이다.

이와 관련하여 국가인권위원회는 "국가는 인공지능을 독립적이고 효과적으로 감독할 수 있는 체계를 수립하여 개인의 인권과 안전을 보장하고, 인공지능 때문에 피해를 입은 사람이 진정을 제기하는 등의 방법으로 권리구제를 받을 수 있는 기회를 제공하여야 한다"는 가이드라인을 제시한 바 있다.[10)]

구글(Google)은 AI백서('The AI Opportunity Agenda')를 통해, 인공지능의 범용적 성격을 고려하여 AI 규제의 획일성을 경계하면서, 규제기관을 기관 간 기구(interagency apparatus)로 구축할 것을 권고하고 있다.

The AI Opportunity Agenda의 일부[11)]

"We believe there are four major universal policies that policymakers should consider to ensure AI researchers and innovators can convert ideas and data into new discoveries, products, and services.
First, as a general principle, given the cross-cutting nature of AI, it is essential that governments avoid siloed approaches to AI regulation. While we need case-specific answers for the unique issues of each sector, it will often be true that a regulatory debate on an issue like data will implicate multiple equities and interests within a government - agencies responsible for privacy, cybersecurity, economic growth, trade, law enforcement, health, and finance all may have a reason to weigh in on the issue.
Governments need to build an interagency apparatus that can effectively represent and balance these competing equities - leaving a critical element of AI policy to one agency, without weighing trade-offs, risks an overall AI strategy that is misaligned with the public's broader interests."

AI 연구자와 혁신가가 아이디어와 데이터를 새로운 발견, 제품, 서비스로 전환할 수 있도록 정책결정자가 고려해야 할 4개의 주요 보편적 정책이 있습니다.
첫째, 원칙적으로, AI의 범용적인(cross-cutting) 성격을 고려하면, 정부가 AI 규제에 대해 획일적인 접근방식을 취하지 않는 것이 중요합니다. 각 부문 고유의 사안의 경우 사례별로 방안이 필요하겠으나, 데이터와 같은 사안에 대한 규제 관련 논쟁에서는 정부 내 다양한 이해관계가 개입되어 있는 경우가 많습니다. 개인정보 보호, 사이버보안, 경제성장, 통상, 법집행, 보건, 금융 등을 소관하는 기관들은 각자 해당 사안에 관여할

10) 국가인권위원회, <인공지능 개발과 활용에 관한 인권 가이드라인>(2022. 5. 11. 권고)

이유가 있을 것입니다. 정부는 이처럼 서로 대립되는 이해관계를 효과적으로 대변하고 그 균형을 유지할 수 있는 **기관 간 기구(interagency apparatus)**를 구축할 필요가 있습니다. 상충관계를 고려하지 않고 AI 정책의 핵심 요소를 하나의 기관에 맡긴다면 전체적인 AI 전략이 국민의 광범위한 이익에 부합하지 않는 위험이 초래될 것입니다.

인공지능 감독을 위한 국가-지역 거버넌스 구축에서 독립성과 통합성이 고려되어야 한다.

심의·의결기구로서 (국가)인공지능위원회를 대통령 소속(정점식, 조인철, 안철수의원안) 또는 국무총리 소속(김성원, 민형배의원안)으로 두고 있다. 예를 들어 정점식의원안의 경우, 해당 위원회는 정부위원, 민간위원, 대통령비서실의 과학기술 수석비서관으로 구성된다(위원장: 대통령, 간사: 대통령실 수석비서관, 40명 이내). 대통령실 중심의 위원회가 독립성을 확보할 수 있을지 의문이다.

발의된 법안들은 대부분의 인공지능 통제에 관한 주요 내용은 과학기술정보통신부가 주무부처로 규정되어 있다. 과기정통부가 독점하다시피 하는 이러한 감독 거버넌스는 (구글 AI백서가 권고한 바와 같이) 상충하는 이해관계를 조정하고 범부처가 관여할 수 있는 체계가 아니다. 새로운 국가 독립 감독기관을 구축할 필요가 있다.

한편, 민형배의원안의 경우 지역인공지능위원회의 구성을 고려하고 있다. 중앙뿐만 아니라 지역 거버너스를 고려하고 있다는 점에서 전향적인 입법으로 평가된다. 앞으로 인공지능에 관한 국가-지역 거버넌스도 함께 고려할 필요가 있다.

11) Google, The AI Opportunity Agenda, https://storage.googleapis.com/gweb-uniblog-publish-prod/documents/AI_Opportunity_Agenda.pdf

"인공지능 문제는 사회 여러 분야에 걸쳐 있고, 분야별 전문성이 상호 협력할 수 있는 거버넌스 체계가 마련되어야 한다. 과학기술정보통신부 단독으로는 자율주행자동차의 책임 문제나 경찰AI의 인권 위험을 모두 포괄하는 전문성을 가지고 있지 않다. 카카오택시, 네이버쇼핑, 쿠팡 등 국민 소비생활에 큰 피해를 끼쳤던 플랫폼기업의 알고리즘 조작 사건에 대해서는 공정거래위원회가 전문성을 발휘하였으며, AI챗봇 이루다의 개인정보 침해 문제에 대해서는 개인정보보호위원회가 가장 전문적이었다. 따라서 각 부처가 전문적인 기존의 소관을 인공지능 분야에도 적용하는 집행을 계속할 수 있을 것이다. 다만 인공지능법에 대한 고유한 집행은 경제부처보다는 규제기관 중에서 한 곳을 지정하여 담당하도록 하거나 해외에서 논의되듯 새로운 국가 독립 감독 기관을 신설하는 것을 고려해봄 직하다.
마지막으로 우리 시민사회는 22대 국회가 특정 부처, 특정 상임위원회, 산업계의 조급한 이해관계를 벗어나, 범국회적이고 미래지향적인 논의를 통하여 인공지능의 위험을 실효적으로 통제할 수 있는 사회적 합의 도출에 나설 것을 요구한다. 이를 위해서는 인공지능 제공자, 활용자는 물론 영향을 받는 사람을 포함하는 각계의 다양한 의견을 충분히 수렴하는 절차를 마련해야 할 것이다. 방송통신융합 문제가 우리 사회에 처음 등장하였던 17대 국회 당시 〈방송통신융합 특별위원회〉를 만들어 범상임위적으로 이 문제를 논의하고 신규 규제기관인 방송통신위원회 설립 및 운용에 관한 법안을 공동으로 심의하고 합의에 도달하였던 사례를 참고해볼 수 있을 것이다."[12)]

Ⅳ. 나오며

22대 국회는 개원한지 불과 한 달여만에 5개의 법안을 발의했다. 해당 법안들의 전반적인 기조는 '토종 AI산업 육성'과 '자율규제'이다. 그러나 이는 유럽연합, 미국, 유럽 각국 등의 인공지능법제에서 흐르는 글로벌 스탠다드에 한참 미달한다. 인공지능 산업은 역내뿐만 아니라 역외 시장에서 표준화되고 유통되어야 한다. 따라서 인공지능법제와 규범 또한 국제적 표준을 준용해야 한다. 현재 논의되고 있는 법안들은 근시적 자국산업 보호라는 미명하에 국내 인공지능 산업의 갈라파고스화를 부추길 뿐이다.

현재 22대 국회는 과기정통부, 과방위, 일부 산업계 등의 이해관계에 매몰되어 법안 통과에만 주력하고 있다. 그러나 인공지능은 그 특성상 ICT 산업과는 달리 범용성을 띠며 인간과 인류성 그 자체를 위협하는 침습적 기술이다. 국회가 인공지능의 산업적, 사회적, 국가적 파급효과와 중요성을 진정으로 고려한다면, 산업 육성과 자율규제, 인공지능 윤리와 같은 연성규범과 함께 인공지능의 위험성을 실질적으로 통제할 수 있는 법규범, 거버넌스, 통제방안을 마련해야 한다. ▢

12) 22대 인공지능법 제정에 대한 시민사회 의견서

토론문

차지호 | 더불어민주당 국회의원

토론문

김병욱 | 민주사회를위한변호사모임 디지털정보위원회, 법무법인 두율 변호사

1. 들어가며

인공지능 기술의 영향력이 급속도로 확산되는 가운데, 인공지능기술이 제공할 편익과 효용의 관점에 치우쳐 인공지능 기술이 초래할 새로운 형태의 기본권 침해나 차별 위험에 대해서 등한시되는 경향이 있음.

21대 국회에서 발의되었던 다수의 법안과 더불어 22대 국회에서 발의된 법안의 경우에도 인공지능의 위험성에 대한 실효적인 규제에 대한 고민은 부족해보임. 새롭게 제정될 인공지능 기본법에 인공지능의 안전한 활용과 신뢰성 확보를 담보하고, 위험성에 대한 실효적인 규제를 가능하게 하는 법률적 근거가 반드시 담겨야할 것임.

인공지능 기술의 영향력은 일부 영역과 분야에 한정되지 않고, 정도는 갈수록 심화할 것임. 인공지능이 가진 영향력에 비추어 단순히 기술 발전이나 산업진흥의 관점에서 접근하는 것은 적절하지 않음. 일개 부처나 기관이 아니라 여러 분야의 관점, 전문성을 포괄할 수 있는 통합적, 융합적 접근이 필요함.

다만 현 상태에서도 각 기관의 설립 목적과 취지에 맞게 각자의 규율 영역에서 본래의 기능이 추진되고, 적절히 수행되어야할 필요가 있음. 이는 향후 제정될 인공지능 기본법

과 상충되는 것이 아니며, 기본법의 추상적인 규율을 개별 영역에서 구체화시키는 것으로 볼 수 있고, 같은 맥락에서 상호 보완관계로 볼 수 있음.

2. 인공지능 위험성 규제에 대한 접근

가. 개인정보보호의 관점

1) 인공지능 기술 개발의 원료가 되는 학습데이터의 가치와 중요성이 커지고 있음. 그러나 인공지능에 의한 데이터의 활용 방식이나 활용 범위는 불명확하고, 불투명함. 학습데이터는 특정 개인의 식별 가능성이라는 위험성을 내재되어 있기 때문에, 개인정보의 활용범위에 대한 통제권이라 할 수 있는 정보주체의 개인정보자기결정권은 매우 심각한 위기에 처해 있다고 할 것임.

2) 인공지능의 위험성에 대한 규제는 위험기반접근으로 수렴해가는 경향을 보이고 있으므로, 개인정보보호법제 역시 이러한 위험기반 접근방식을 도입할 필요가 있어보임. 즉, 고위험 영역에서 개인정보를 수집, 활용하는 경우 동의의 요건을 강화하거나 영향평가의 적용 대상이나 엄격성을 강화할 필요가 있을 것임.

3) 나아가 인공지능기술에 의해 개인정보가 수집되고, 활용되는 경우 정보주체의 개인정보에 대한 통제권의 실효적인 보장을 위한 방안이 강구될 필요가 있음.

자동화된 의사결정에 대한 정보주체의 대응권으로서 개인정보보호법 제37조의 2가 신설되어 2024. 3. 15.부터 시행되고 있고, 동 규정에 의하면 인공지능기술에 의한 개인정보처리에 대하여 설명요구권, 거부권이 보장됨.

> **개인정보보호법 제37조의2(자동화된 결정에 대한 정보주체의 권리 등)** ① 정보주체는 **완전히 자동화된** 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 개인정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외하며, 이하 이 조에서 "자동화된 결정"이라 한다)이 자신의 권리 또는 의무에 **중대한 영향을 미치는 경우**에는 해당 개인정보처리자에 대하여 해당 결정을 **거부할 수 있는 권리**를 가진다. 다만, 자동화된 결정이 제15조제1항제1호·제2호 및 제4호에 따라 이루어지는 경우에는 그러하지 아니하다.
>
> ② 정보주체는 개인정보처리자가 자동화된 결정을 한 경우에는 그 결정에 대하여 **설명 등을 요구할 수 있다.**

③ 개인정보처리자는 제1항 또는 제2항에 따라 정보주체가 자동화된 결정을 거부하거나 이에 대한 설명 등을 요구한 경우에는 정당한 사유가 없는 한 자동화된 결정을 적용하지 아니하거나 인적 개입에 의한 재처리·설명 등 필요한 조치를 하여야 한다.

④ 개인정보처리자는 자동화된 결정의 기준과 절차, 개인정보가 처리되는 방식 등을 정보주체가 쉽게 확인할 수 있도록 공개하여야 한다.

⑤ 제1항부터 제4항까지에서 규정한 사항 외에 자동화된 결정의 거부·설명 등을 요구하는 절차 및 방법, 거부·설명 등의 요구에 따른 필요한 조치, 자동화된 결정의 기준·절차 및 개인정보가 처리되는 방식의 공개 등에 필요한 사항은 대통령령으로 정한다.

다만 GDPR(유럽연합 일반 데이터보호규칙) 제22조가 계약의 이행, 정보주체의 동의 등 예외적인 경우를 제외하고 완전 자동화된 의사결정 자체를 금지하는 것과 달리, 개인정보보호법 제37조의 2는 정보주체의 권리 또는 의무에 중대한 영향을 미치는 경우에 한하여 이를 거부하고, 설명, 인적개입에 의한 재처리를 요구할 수 있는 권리를 부여하고 있어 그 체계를 달리하고 있음.

입법취지와 정보주체의 권리 보장 측면을 고려하여 거부권 행사의 요건이 되는 '중대한 영향'을 완화해석하여 정보주체의 인공지능에 의한 자동화된 의사결정에 대한 거부권을 폭넓게 보장해야 할 것임.

GDPR 제22조 프로파일링을 포함한 자동화된 개별 의사결정

1. 정보 주체는 프로파일링을 포함해 자동화된 처리만을 바탕으로 한 결정으로서 자신과 관련한 법적 영향 또는 이와 유사하게 중대한 영향을 미치는 결정의 대상이 되지 않을 권리를 가진다.

2. 결정이 다음에 해당되는 경우에는 제1항이 적용되지 않는다.

(a) 정보 주체와 정보 관리자 간의 계약 체결 또는 이행에 필요한 경우

(b) 정보 관리자가 준수해야 하며 정보 주체의 권리 및 자유와 정당한 이익을 보호하기 위한 적절한 조치

(c) 정보 주체의 명시적 동의에 기초한 경우

설명요구권의 경우 완전히 자동화된 시스템을 전제로 하고 있어 그 범위가 협소해질

우려가 있음. 이를 엄격하게 해석해서는 안될 것임.

4) 개인정보가 개인정보처리자에 의해 임의로 인공지능 기술개발의 학습데이터로 제공되고, 활용되지 않도록, 수집한 개인정보를 인공지능 학습에 활용하거나 이를 목적으로 제3자에게 제공하는 경우 그 목적을 별도 항목으로 구체적이고 명확하게 안내하도록 하여, 정보주체의 실질적인 동의권을 보장하여야 함(개인정보보호법 제22조 관련). 개인정보 유출과 위법한 데이터 재식별조치 등 불법적인 행위를 막기 위해 보다 강한 규제가 요구됨.

관련하여, 최근 네이버 개인정보처리방침에서는 개인정보 수집목적에 신규서비스 요소의 발굴 및 기존서비스 개선을 언급하면서, 정보검색, 다른 이용자와의 커뮤니케이션, 콘텐츠 생성·제공·추천, 상품 쇼핑 등에서의 인공지능 기술적용이 포함된다고 설명하고 있는데, 네이버 서비스 이용자의 정보수집 목적에 인공지능 서비스 개발을 포함하는 것이 개인정보 최소수집의 원칙이나 목적 명확성의 원칙에 반하지 않는지 의문임.

나. 독점 규제 및 시장기능 보호의 관점

1) 카카오택시 알고리즘 조작, 쿠팡 알고리즘 조작 사례 등을 통하여, 플랫폼에서 인공지능 알고리즘을 통한 경쟁저해, 시장지배력 남용이 가능하다는 점은 이미 확인되었음. 플랫폼은 중개자로서 특정 사업자를 우대하거나 차별하는 것이 가능하고, 독점력을 이용한 시장지배력 남용이 이루어질 가능성이 있음. 인공지능 기반 플랫폼에서 이러한 특성은 더욱 심화할 가능성이 큼.

이러한 피해는 소비자 피해와 더불어 경쟁사업자들을 배제함으로써 시장에서의 경쟁을 저해하여 혁신에도 악영향을 미칠 수 있음.

2) 위험기반 접근법은 공정거래 관련 규제에도 적용될 수 있음. 검색엔진, SNS서비스 등 규모가 크고, 일반 국민의 일상적인 생활을 영위하는 데 필수적인 재화 및 서비스의 제공과 관련한 플랫폼 등 사업자 및 소비자에 대하여 해악을 끼칠 가능성이 더 크다고 볼 수 있는 경우, 그에 상응하는 기술문서 작성, 영향평가 등 위험기반에 따른 의무를 부과할 수 있을 것임.

또한 위험도에 따라 더 강화된 공정성, 투명성, 책임성의 원칙을 관철할 수 있을 것임. 플랫폼의 알고리즘에 대한 접근권을 보장하거나, 설명을 요구할 수 있어야 함.

3) EU에서는 거대 플랫폼 사업자의 시장지배력 남용을 방지할 목적으로 일정한 규모의 플랫폼 사업자를 '게이트 키퍼'로 지정하여, 특별한 규율을 부과하는 디지털시장법(DMA)이 제정되어 2022. 11. 1.부터 발효되었는데, 이는 같은 취지에 입각한 것으로 볼 수 있음.

본격적인 규제를 위해서는 별도 법령의 근거가 필요하겠지만, 현행 법체계하에서도 적극적인 법적용과 행정작용을 통해 플랫폼에서의 시장지배력 남용행위와 불공정거래행위를 규율할 수 있음.

4) 한편 독일연방카르텔청은 2019. 2. 7. 페이스북이 실질적인 동의를 받지 않은 상황에서 제3자로부터 이용자들의 개인정보를 수집하고 이를 해당 이용자의 페이스북에 연계·통합시켜 온 것에 대하여 착취적 거래조건 강제행위로서 독일 경쟁법에 위반한다는 결정을 하였는데, 이를 참고하여 데이터의 과도한 수집이나 처리(독점)로 인하여 소비자 착취가 발생하는 경우 경쟁법의 적용을 적극적으로 검토할 필요가 있음.

다. 인권 보장과 차별 구제의 관점

1) 국가인권위원회는 인권의 보호와 향상을 위한 업무를 독립적으로 수행하는 기관으로 기본권 침해 문제와 차별에 대한 조사, 구제 업무에 대하여 축적된 전문성과 역량을 보유하고 있음.

인공지능기술이 새롭게 야기하는 기본권 침해 및 차별의 진정 및 구제절차에 있어, 인권에 대한 전문성과 역량을 확보하고 있는 국가인권위원회가 적극적으로 개입하고, 역할을 담당할 필요가 큼.

2) 국가인권위원회는 인권 보호와 향상과 관련하여 국가 기관등에 대한 시정권고, 의견표명 권한이 있음. 국가인권위원회는 2022. 5.경 인공지능개발과 활용에 관한 인권 가이드라인을 마련하였고, 해당 가이드라인에 기초하여 인공지능 관련 정책 수립, 관계법령 제·개정할 것을 과학기술정보통신부장관, 개인정보보호위원회 위원장 등에게 권고하였고, 2023. 1. 25. 얼굴인식기술의 도입 및 활용에 있어 인권을 보호하고, 인권영향평가를 실시할 것을 요구하는 내용의 권고를 하였음.

아울러 국가인권위원회는 21대 국회에서 논의 중이던 인공지능 법률안에 대하여 2023. 8. 24. 인권침해, 차별문제를 예방, 규제할 수 있는 규정을 마련할 것을 의견표명하고, 인

공지능 인권영향평가 도구안을 개발하여 2024. 4.경 상임위 의결로 확정하였고, 2024. 7. 8. 과기정통부 장관에게 활용해줄 것을 의견표명하였음. 향후에도 이러한 역할을 더욱 적극적으로 수행할 필요가 있음.

3) 사전적으로 인공지능이 야기하는 기본권 침해 및 차별 위험에 대해서 식별하고 방지 및 완화조치를 도입하도록 하는 인공지능 인권영향평가는 독립성과 객관성을 담보한 상태에서 수행되어야 함. 이를 위해 기술을 직접 개발한 당사자가 아니라, 인권에 대한 전문성과 역량을 갖춘 제3의 기관 또는 기관 내의 별도 부서가 인권영향평가를 수행하여야 함. 국가인권위원회가 인권영향평가에 있어 주도적인 역할을 수행해야 할 것임.

3. 나오며

인공지능기술의 영향력이 날이 갈수록 심화하고 있는 상황에서, 인공지능기술이 인권에 미치는 영향이 불확정적이긴 하나, 점차 뚜렷해지고 있는 양상임.

인공지능의 영향력이 사회 전반에 걸쳐있는 만큼, 개인정보보호, 독점규제 및 시장기능 보호, 인권보장과 차별구제를 포함하여 다양한 관점(산업 안전, 금융소비자 보호 등)과 전문성을 포괄하는 통합적인 규제 거버넌스가 마련될 필요가 있음.

이와 관련하여 활발한 논의와 토론를 거친 사회적 합의를 도출해야할 필요성이 긴절하고, 국회에서의 논의도 과방위 내에서가 아니라 여러 상임위를 아우르는 통합적인 논의가 이루어져야할 필요가 있어보임. 인공지능 기본법에는 이러한 합의의 결과가 반영되어야 할 것임. ▢

토론문

김영규 | 한국인터넷기업협회 정책1실 실장

■ 인공지능 기술 발전과 그에 따른 규제 및 산업 육성의 필요성

- 인공지능(AI) 기술의 급속한 발전은 현대 사회의 경제, 안보, 그리고 일상생활에 지대한 영향을 미치고 있음
- 편향성, 판단오류, 윤리적 문제 등 문제점이 나오고 있어 업계는 인공지능의 역기능 방지를 위한 제도적 장치가 근본적으로 필요함을 인식하고 있음
- 우리나라는 ICT 산업 성장을 통해 디지털 산업 강국으로 도약하였으며, ICT 기술은 우리나라 경제 전반에 스며들어 국가 경제의 성장을 견인하였음
- 전 세계의 인공지능 열풍 속에서 우리 기업들은 꾸준한 연구·개발을 통해 미국, 중국에 이어 세계 세 번째로 초거대 인공지능을 상용화함
- 국내 주요 ICT 기업은 AI를 미래 수출 먹거리로 추진하고 있어 이는 수출 경제에도 큰 기여가 가능할 것으로 예상됨
- 특히 생성형 인공지능뿐만 아니라 온디바이스 인공지능의 확대에 따라 휴대폰, TV, CCTV 등 정보통신방송기기에 인공지능 HW·SW가 접목되고 있어 산업의 핵심 차별

화 요소로 작용하고 있음

- 최근 국내 ICT 기업들은 사우디, 동남아시아 국가 등과 거대언어모델(LLM) 공동 개발 합의, 생성형 인공지능 서비스 개발 지원 합의, 기술 제공 등을 통해 수출 경제에 기여하고 있음

- 미·중이 AI 패권 경쟁을 하고 있는 가운데 중위권 국가인 우리나라, 영국, 캐나다, 인도, 이스라엘 등의 치열한 경쟁에서 글로벌 경쟁력을 확보하기 위해서는 내수 중심을 넘어 글로벌을 겨냥한 전략적 투자와 연구개발이 필요

■ 인공지능 산업의 발전을 저해할 수 있는 법안의 문제점과 개선 필요성

1. 인공지능 법안의 문제점

- 제22대 국회가 발의한 6개 법안을 보면 산업의 발전과 육성보다는 규제로 작용하여 인공지능 산업을 혁신을 저해할 우려가 있음

- 먼저 법안에서 규정하고 있는 인공지능의 정의는 '인간의 지적 능력' 등 **추상적인 개념을 기반으로 하고 있어 인공지능의 의미를 불명확**하게 하고, 이에 따라 **적용 대상이 모호해질 우려**가 있음

- 추상적으로 규정하게 되면 입법 취지와 다른 대상이 본 법안의 수범 대상이 될 가능성이 커 입법 취지를 실현하지 못할 가능성이 크기에 인공지능의 법적 개념 명확화가 필요

- 또한 수범자의 정의는 향후 다양한 개별법, 특별법 등에서 준용될 가능성이 상당히 크기에 특히 신중한 검토가 필요

- 인공지능을 법적으로 정의한 해외 사례를 보면 **'인공지능 기술 및 실제 작동 방식'을 기반으로 구체적으로 범위를 규정**하고 있음

- 또한 인공지능의'연구'와'개발'을 구분하고 있지 않아 **인공지능을 연구 목적으로 개발하는 사업자에 대해서도 규제 대상으로 포섭하고 있어 인공지능에 대한 연구개발 및 산업 발전을 저해할 가능성이 있음**

- '인공지능 기술을 활용한' 제품과 '인공지능 기술 또는 인공지능제품과 관련한 서비

스’를 모두 인공지능제품과 인공지능서비스로 정의하고 있어 인공지능 기술 발전에 따라 대부분의 제품 및 서비스가 제정안의 적용 대상에 포섭될 우려가 있음

- 이 경우 인공지능의 발전 내지 혁신에 지대한 영향을 미치는 오픈소스 소프트웨어들에 관여하는 개발자 또는 사업자들조차 고위험영역 인공지능 개발사업자나 이용사업자에 해당할 여지가 있어 이러한 오픈소스 소프트웨어 또는 연구개발에 대해서도 **불필요한 규제가 이뤄지면 한국의 인공지능 산업만 뒤처지게 되는 결과가 초래될 수 있음**

※ EU의 인공지능법에서는 인공지능 연구개발을 촉진하려는 목적으로 연구개발을 규제 대상에서 배제하고 있음 (EU AI Act 제3조 제63호)[1] 또한 “제공자(provider)”의 개념에 시장 출시 여부를 포함하는 등 적용 범위를 명확히 제한하고 있음 (EU AI Act 제3조 제3호)[2]

2. 인공지능 산업 발전을 위한 규제 개선 필요성

- **본 법안에서 기본 원칙 또는 규제의 원칙으로 ‘우선 허용, 사후 규제’를 규정하지 않았다는 문제점이 있음**

- 신기술 서비스·제품인 인공지능 기술에 대해서는 「행정규제기본법」이 규정하고 있는 바와 같이 우선허용·사후규제 원칙을 적용하여 산업 육성과 진흥을 도모할 필요가 있음

「행정규제기본법」

제5조의2(우선허용·사후규제 원칙) ① 국가나 지방자치단체가 **신기술을 활용한 새로운 서비스 또는 제품(이하 “신기술 서비스·제품”이라 한다)과 관련된 규제를 법령 등이나 조례·규칙에 규정할 때는 다음 각호의 어느 하나의 규정 방식을 우선적으로 고려하여야 한다.**

---

1) EU AI Act 제3조 제63호 : ‘범용 AI 모델’은 AI 모델이 대규모 자체 감독을 통해 대량의 데이터로 훈련되고, 상당한 일반성을 발휘하며, 모델 출시 방식과 관계없이 광범위한 개별 작업을 능숙하게 수행할 수 있고, 다양한 다운 스트링 시스템이나 애플리케이션에 통합될 수 있는 경우를 포함한 AI 모델을 의미한다. **출시되기 전 연구, 개발 또는 프로토타입 제작 활동에 사용되는 AI 모델은 제외한다.**

2) EU AI Act 제3조 제3호 : **‘제공자’**란 AI 시스템 또는 범용 AI 모델을 개발하거나 자신의 이름 또는 상표 하에 유료 또는 무료로 AI 시스템 또는 범용 **AI 모델을 개발하여 출시하거나 AI 시스템을 서비스 개시하는 자연인이나 법인, 공공 기관, 관청 또는 기구**를 말한다.

1. 규제로 인하여 제한되는 권리나 부과되는 의무는 한정적으로 열거하고 그 밖의 사항은 원칙적으로 허용하는 규정 방식
2. 서비스와 제품의 인정 요건 · 개념 등을 장래의 신기술 발전에 따른 새로운 서비스와 제품도 포섭될 수 있도록 하는 규정 방식
3. 서비스와 제품에 관한 분류기준을 장래의 신기술 발전에 따른 서비스와 제품도 포섭될 수 있도록 유연하게 정하는 규정 방식
4. **그 밖에 신기술 서비스 · 제품과 관련하여 출시 전에 권리를 제한하거나 의무를 부과하지 아니하고 필요에 따라 출시 후에 권리를 제한하거나 의무를 부과하는 규정 방식**

- 인공지능 기술은 기술의 변화와 발전 속도가 빨라 사전에 모든 것을 규정해 사전에 규제하는 방식보다는 **다양한 실험과 시도를 통해 기술을 발전시키면서, 그와 관련해 발생하는 문제점 및 한계점에 대해서 적절한 사후 조치를 하는 것이 현실적인 대응 방안임**

- 지난 2023년 2월 27일 메타버스 발전을 지원하기 위해 제정된 「가상융합산업진흥법」은 제4조에서 우선허용 및 사후규제 원칙을 규정하여 변화와 발전 속도가 빠른 메타버스 기술 영역의 건전한 발전을 도모하고 있음

■ 인공지능 기본법 제정 방향

1. 인공지능 규제의 적용 범위와 예외 규정을 통한 신뢰성 확보

- EU 인공지능법의 경우 생성형 인공지능을 포함한 범용 AI 제공자(provider)에게 적용되는 의무를 명확히 규정하여 적용 범위를 제한하고 있음을 고려하여 **인공지능의 결과물이 실제로 이용자에게 피해를 줄 우려가 있는 경우만으로 적용 범위를 한정하는 방안도 대안으로 검토할 필요가 있음**

- (전문 제97조) 해당 의무는 인공지능 모델이 **시장에 출시된 경우에만 적용**되며 **사업자 내부적으로만 사용되는 경우는 적용 범위에서 제외**됨을 명확히 하고 있음
  ※ 범용 AI의 경우도 제3조 제63호를 통해 시장에 출시되기 전인 모델은 제외되도록 규정하고 있음
- (제50조 제1항) **실제 문제가 될 수 있는 경우(이용자의 오인 가능성이 있는 경우)에만**

제한적으로 의무가 적용될 수 있도록 하며, 이용자가 인공지능임을 명확하게 인지할 수 있는 경우에는 의무의 예외를 두고 있음

- (제50조 제2항) 범용 AI 시스템으로 생성된 콘텐츠에 표시함에 있어 다양한 유형의 콘텐츠의 특수성 및 한계를 고려하도록 하고, **AI 시스템이 일반적인 편집을 위한 보조 기능을 수행하는 등의 범위에서는 표시 의무 적용을 배제**하고 있음
- (제50조 제4항) **범용 AI 이용 배포자에 적용되는 공개 의무는 딥페이크에 해당하는 콘텐츠 생성 AI 시스템에 한해 적용**되며, **콘텐츠가 명백히 예술적, 창의적, 풍자적, 허구적 유사 저작물 또는 프로그램 일부를 구성하는 경우에는 별도의 고려**가 이루어지고 있음

▶ **인공지능 시스템 개발 이후 특정 애플리케이션을 개발하는 사업자**가 '해당 인공지능 시스템을 활용하는 구조'를 설계하는 과정에서 관련 신뢰성 확보 조치를 마련하는 것이 합리적

2. 인공지능 산업 경쟁력 강화를 위한 방안

- 우리나라 ICT 위상에 비해 인공지능 경쟁력이 상대적으로 낮은 주된 이유는 글로벌 빅테크 대비 민간투자가 위축되어 있다는 것
- 국내 5개 사의 연간 합산 투자액이 3.2조 원인 데 반해, 오픈AI는 모델 고도화에만 약 133조 원 규모를 투입할 예정으로 정부의 지원 없이는 경쟁할 수 없는 상황
- 정부가 앞장서 반도체 산업을 국가 주력산업으로 육성했듯이 정부의 집중 지원과 민간투자를 더 끌어낼 수 있는 조세 혜택과 같은 제도적 장치 필요
- 만약 기본법에서 육성이 어렵다면 산업진흥법안과 같은 별도의 법안으로 진흥에 대한 부분을 빠르게 추진하되 여러 논의가 되는 인공지능의 역기능에 대해서는 더 세밀한 논의를 진행하면서 제도를 만들어가는 방향을 제안함
- 또한 정부는 인공지능에서 소외되는 직업군, 사회적 약자들을 위한 인공지능 리터러시 교육을 진행해 주길 바람
- 앞으로 오늘과 같은 자리가 자주 개최되어 인공지능이 초래할 수 있는 위험성에 적절히 대응하면서 인공지능 산업 발전을 통한 사회적 편익이 저해하지 않는 잘 균형 잡힌 규제방안이 모색되기를 기대해 봄 ▢

과학기술정보통신부 토론

남철기 | 인공지능기반정책과 과장

1. 논의 배경

□ 전 세계적 **챗GPT 돌풍**을 계기로, AI에 대한 **막연한 기대감**이 **높은 효용성**으로 **증명**되고, 누구나 쉽게 AI를 활용하는 **'AI 일상화'** 촉발

○ **글로벌 빅테크** 기업은 압도적 **컴퓨팅 파워·대규모 자본** 등을 토대로 **초거대AI 플랫폼을 선점**하기 위한 **속도전 치열**

○ 우리도 **독자적 초거대AI 생태계**(AI반도체·클라우드 – AI모델 – 응용서비스)를 바탕으로 **글로벌 경쟁**에 참여하고 있음

□ 우리나라는 그간 **데이터 축적, AI R&D 및 인력양성과 더불어 국산 AI반도체**(NPU, PIM) **출시** 등 **AI 생태계 기반 조성**에 집중

○ 이와 같은 **정책 성과**와 **민간의 한발 앞선 도전**으로, **우리 기업**은 **독자적 초거대AI 모델*을 개발**하고 **세계를 향해 도전 시작**

* 네이버, LG AI, 카카오, SKT, 코난테크놀로지, 엔씨소프트 등

□ 한편, **주요국·국제기구**는 **'19년 OECD 'AI 원칙' 채택을 시작**으로 AI **윤리·신뢰성 확보**를 위한 **원칙**과 **민간 자율 실천 방안**을 **제시**

※ **OECD 'AI 원칙'**('19.5), **UNESCO 'AI 윤리권고'**('21.11), **UN 'AI 결의안'**('24.3) 등

○ 특히, **생성형 AI 등장 이후** 각 국은 AI **위험성 예방**을 위하여 **규범의 제도화를 추진**하고 있으나, **내용 측면에서는 차이** 존재

※ 美·英·EU·G7·UN 등 국제사회는 **AI 위험성에 대한 대응 필요성을 공감**하고 있으며, **규범체계 정립 및 안전성 강화 논의**를 본격 진행 중

□ **우리나라는 사람 중심의 AI 구현**을 위한 OECD AI **권고안**('19.5월) **수립** 및 이에 **기반**한 G20 **정상선언문**('19.6월) **반영을 주도**하였으며,

○ 이후 OECD, UNESCO AI **권고안 내용을 반영**하여 **3대원칙· 10대 요건**으로 구성한 「AI **윤리기준**」 발표('20.12)하고,

○ 이러한 **정책적 노력**에 **호응**하여 **민간**에서도 AI **윤리·신뢰성 전담조직 신설*** 등 AI **윤리·신뢰성 준수**를 위한 **자발적 노력 확산** 중

* 네이버 '퓨처 AI 센터'('24.1.~), SKT 'AI 거버넌스 TF'('24.1.~) 등

2. 글로벌 규범 동향 및 국내 규범 정립 방향

□ **생성형 AI 기술의 급격한 확산**과 함께 가짜뉴스·편향성·개인정보 침해·저작권 권리 침해 등 **부작용**에 대한 **우려**도 대두되는 상황

○ 美·英·EU·G7·UN 등 국제사회는 AI **위험·부작용 관련 다양한 이슈에 대응**하기 위해 **규범 및 다양한 제도적 기반 마련** 중

〈 참고 : 주요국의 AI 규범 마련 현황 〉

◇ **(美)** 안전성·보안성·신뢰성 3대 원칙에 따라 안전하고 투명하게 AI 기술을 개발하겠다는 7개 **AI 기업의 자발적 서약 확보**('23.7월), **AI 행정명령 발표**('23.10월)

〈 美 AI 행정명령 개요 및 주요내용 〉

■ **(개요)** AI의 무책임한 사용으로 인한 **위험을 완화**하고 AI를 활용한 **생산성·혁신성 향상을 극대화**하기 위해 준수해야 할 **원칙 및 부처별 조치 필요 사항**

■ **(주요내용)** 새로운 **AI 안전 및 보안 기준** 제시, 개인정보 보호, 평등 및 시민권 증진, 소비자·환자·학생 지원, 글로벌 리더십 확보 등 **8개 주요 행동지침** 포함

◇ **(EU) AI**에 대한 **EU의 글로벌 리더십을 확보**하기 위해 **AI 시스템을 위험도에 따라** 4가지로 분류·규제하는세계 최초의 **AI 규제법인 「AI Act」 제정**('24.5월)

- 챗GPT 등장 후 **범용 AI에 대한 정의** 및 **워터마크 표시** 등 **의무를 부과**를 추가하였으며, **의무 위반 시** 최대 전년도 **전 세계 매출액 7%의 과징금** 부과

◇ **(日) G7 히로시마 AI 프로세스** 주도(국제 AI 행동강령 발표, '23.10월)

- AI **수명 주기 전반**에 걸친 **위험 식별·평가 및 완화하기 위한 조치**, AI 시스템의 기능·제한 사항·사용 및 **오용에 관한 공개 보고서 게재 등 11개 항목으로 구성**

□ 우리나라는 **파리 이니셔티브**('23)를 통해 **AI·디지털 국제기구 신설**을 제안하였으며, **디지털 권리장전** 수립, **AI 서울 정상회의 개최** 등 추진

○ 특히, AI 안전에만 국한되지 않고 **혁신·포용**까지 포함하는 **균형적인 AI 거버넌스 발전 방향**을 제시한 **「서울 선언」의 가치를 바탕으로,**

○ **AI 발전 및 신뢰 기반 조성**을 **균형 있게 고려**한 **국내 규범인 AI 기본법 제정**을 통해 **국가 AI 경쟁력 강화의 제도적 기반 마련** 필요

3. 22대 국회 AI기본법 추진현황 및 주요 내용

□ **22대 국회**에서 **AI 기본법안(7.8 현재) 6건 발의**

※ AI 기본법 발의 현황(6건): 정점식·안철수·김성원(국민의힘), 조인철·민형배·권칠승(민주당)

○ **6건의 의원안** 모두 **21대 국회** 시 **여야가 함께 논의한 법안**을 기초로 발의되는 등 **AI 발전과 신뢰 기반 조성의 균형**을 갖추고자 함

□ 여당에서 당론으로 발의한 정점식 의원안의 경우

○ AI 산업 육성을 지원하기 위한 AI 기본계획 수립, 국가AI위원회 신설, R&D·표준화·전문인력 확보 등에 대한 근거 마련과

○ 신뢰 기반 조성을 위해 고위험영역AI 및 생성형AI 정의 및 사업자에 대한 의무 부과, AI안전연구소 운영 근거 신설 등이 포함됨

□ 21대 국회 과방위 대안과 비교할 때 AI 안전성 확보를 위해 '우선허용·사후규제 원칙' 삭제, 생성형 AI 고지 및 표시 등이 반영되었으며,

○ AI에 대한 범정부 리더십을 강화하기 위해 국가AI위원회의 지위를 격상(위원장: 국무총리·민간 공동 → 대통령)과 AI 위험에 대한 국가 차원의 체계적 대응을 위한 전담기관인 AI안전연구소 운영 근거를 마련하고,

○ 美·EU에서 선제적으로 도입한 고도의 성능(누적 연산량 기준으로 판단)을 가진 생성형 AI에 대한 안전 확보 의무도 반영되는 등 안전성 확보를 위한 규정을 강화

□ 다만, 사업자 의무 위반에 대한 제재 규정 도입은 AI 산업 수준에 대한 고려와 함께 사회적 합의가 필요한 사항으로 생각되며,

○ 강력한 규제를 포함한 AI법을 제정한 EU도 고위험AI에 대한 규제의 시행에 시간을 두고 있다는 점을 고려하여 우리도 단계적인 보완 입법을 통해 글로벌 규범 정합성에 부합 필요 ▢

개인정보보호위원회 토론

김직동 | 개인정보보호정책과 과장

공정거래위원회 토론

이준헌 | 시장감시정책과 과장

국가인권위원회 토론

이진석 | 인권정책과 사무관

메 모