유럽연합 인공지능법안의 개요 및 대응방안

고학수
서울대학교
법학전문대학원
교수

임용
서울대학교
법학전문대학원
부교수

박상철
서울대학교
법학전문대학원
조교수

1) European Commission, COM/2021/206 final, Brussels, 21.4.2021.

2) European Commission, "White Paper on Artificial Intelligence - A European approach to excellence and trust," COM(2020) 65 final, 2020.

3) European Parliament resolution of 20 October 2020 on a framework of ethical aspects of artificial intelligence, robotics and related technologies, 2020/2012(INL). European Parliament resolution of 20 October 2020 on a civil liability regime for artificial intelligence, 2020/2014(INL). European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence technologies, 2020/2015(INI). European Parliament Draft Report, Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters, 2020/2016(INI). European Parliament Draft Report, Artificial intelligence in education, culture and the audiovisual sector, 2020/2017(INI).

4) 우리나라의 경우에도 2021. 7. 31. 현재 7건의 인공지능 관련 법안이 국회에 발의된 상태이다. 그 중 5건('인공지능교육진흥법안,' '인공지능 기술 기본법안,' '인공지능 집적단지의 육성에 관한 특별법안,' '인공지능산업 육성에 관한 법률안,' '인공지능 연구개발 및 산업 진흥, 윤리적 책임 등에 관한 법률안')은 소관 상임위원회에서 심사 중이고, 최근 발의된 '인공지능 육성 및 신뢰 기반 조성 등에 관한 법률안'과 '인공지능에 관한 법률안'은 소관 상임위원회에 접수된 상태이다.

5) AI법안의 설명메모(Explanatory Memorandum) 1.1.항.

6) 단, 군용으로 개발되었거나 군용으로만 활용되는 AI시스템(Art. 2(3)), 제3국의 관청 또는 국제기구가 EU 또는 회원국과의 사법공조를 위한 국제협약의 틀 내에서 AI시스템을 활용할 경우(Art. 2(4))에는 적용이 배제된다.

7) AI법안상의 "risk"는 단순한 위험(danger)뿐만 아니라 불확실성(uncertainty)에 기반한 우려도 내포하는 개념이어서 "위험"이라는 일반적인 역어를 사용하지 않고 "리스크"라고 표기한다. AI법안은 리스크 기반 접근을 취한다고 설명하고 있지만, 개별 AI시스템의 리스크를 어떻게 평가하고 분류할 것인지에 대한 내용을 담고 있지는 않아서, 그에 대한 비판적인 시각도 있다.

유럽집행위원회(European Commission, 이하 'EC')는 2021. 4. 21. 유럽의회(European Parliament)에 「인공지능에 관한 통일규범(인공지능법)의 제정 및 일부 연합제정법들의 개정을 위한 법안(Proposal for a Regulation laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts)」(이하 'AI법안')을 발의하였다.[1] AI법안은 EC의 2020. 2. 19.자 "인공지능 백서 – 수월성과 신뢰성에 대한 유럽식 접근방식" 보고서[2] 및 후속한 유럽의회의 2020~2021년 간의 AI 관련 윤리, 책임, 저작권, 형사, 교육·문화·시청각에 관한 각종 결의들[3]을 계승하고 있으며, 주요국 최초의 AI에 관한 일괄(omnibus) 규제 법안으로 이해된다.[4] AI법안은 공식적으로는 기본권과 유럽연합(이하 'EU')의 가치의 보호, 투자와 혁신 촉진, 기존 법령의 집행권한 및 집행력 강화, EU 단일시장의 발전을 지향한다.[5] EU가 AI의 법적 규율에 있어 미국식 자유시장과 중국식 권위주의 사이의 "제3의 길"을 모색하는 계기가 될 것이라는 기대도 있다. 그러나, 개별 조항들을 살펴보면, 2020. 12. 15. 입안된 디지털서비스법 패키지(The Digital Services Act package)와 마찬가지로 미·중에 AI 기술 내지 플랫폼 산업의 주도권을 빼앗겼다는 위기의식 및 강도 높은 보호 무역적 규제 체계를 통해 이를 돌파하고자 하는 의지도 엿보인다. 주요 내용은 이하와 같다.

1. AI법안의 개요

가. 적용 범위와 체계

AI법안은 통과될 경우 AI시스템을 EU 내에서 출시(placing on the market) 또는 서비스(putting into service)하는 제공자들(providers)이나 EU 내에 위치한 AI시스템의 활용자들(users)에게 적용되고(Art. 2(1)(a), (b)), 제3국에 위치한 AI시스템의 경우에도 그 시스템의 산출물이 EU에서 활용될 경우 그러한 시스템의 제공자들과 활용자들에게도 적용된다(Art. 2(1)(c)).[6] 여기서 활용자란 사적·비전문적 활동 과정에서 AI시스템을 활용하는 경우를 제외하므로(Art. 3(4)), 소비자가 아닌 영업 목적의 활용자를 의미한다는 점에 특히 유의해야 한다.

동 법안에서 AI시스템(artificial intelligence system)은 (i) 인간이 정의한 일련의 목적을 위해, (ii) 기계학습(machine learning), 논리/지식 기반 접근(logic- and knowledge-based approaches), 통계적 접근(statistical approaches), 베이즈 추정법(Bayes estimation), 검색 및 최적화 방법(search and optimization methods) 중 하나 또는 복수의 기술 내지 기법을 활용하여 개발되고, (iii) 해당 시스템이 상호작용하는 환경에 영향을 미치는 콘텐츠·예측·추천·결정 등의 출력을 생성하는 소프트웨어로 정의되고 있다(Art. 3(1), Annex I).

이처럼 폭넓게 정의된 AI시스템에 대해 AI법안은 이를 관련 리스크에 따라 구분하여 달리 취급하는 **리스크 기반 접근(risk-based approach)**을 취하고 있다(Preamble (14)).[7] 즉, AI시스템의 리스크를 수인불가 리스크(unaccept-

able risk), 높은 리스크(high risk), 제한적 리스크(limited risk), 최저의 리스크(minimal risk)로 분류하고, 각 범주에 따라 금지 여부, 출시 전 적합성평가(conformity assessment) 및 기타 규제 여부를 규정한다.[8] AI법안에서 채택한 리스크 단계별 해당 유형(예시 포함)과 규제의 기본 체계는 아래에서 보는 바와 같다.

리스크 단계	해당 유형	규제	
수인불가 리스크	• 잠재의식의 조작 • 아동·장애인의 착취 • 공적인 범용 사회적 평점 시스템 • 실시간 원격 생체정보기반 식별	금지	
높은 리스크	• 제3자 적합성평가를 받아야 하는 제품(기계, 자동차, 전파기기, 의료기기, 완구 등) 또는 그 안전요소	해당 제품에 대한 적합성평가	• AI시스템 요건: 리스크 관리 시스템, 데이터 및 데이터 거버넌스, 기술문서, 기록보존, 투명성 및 활용자 정보제공, 인적 감시, 정확성·견고성·보안. • AI시스템 제공자(신입법체제에 따른 적합성평가 대상 제품의 경우 제조자 포함)의 의무: 위 요건 준수, 품질관리시스템, 기술문서 작성, 로그, 적합성평가, 등록, 시정조치, 신고의무, 당국에 협조, 대표자 선임, 출시 후 모니터링, 중대사고·오작동 신고. • AI시스템 수입자의 의무: 적합성평가 등 확인, 출시 중단 등, 정보제공, 보관·운송 상 유의사항, 당국에 협조 • AI시스템 유통자의 의무: CE마킹 등 확인, 출시 중단 등, 보관·운송 상 유의사항, 당국에 협조, 시정조치 • AI시스템 활용자의 의무: 사용설명서 준수, 시스템 작동 감시, 로그, 제공받은 정보의 활용
	• 원격 생체정보기반 식별 • 핵심기반시설(교통·수도·가스·난방·전력)의 관리·운영	제3자 또는 자율적 적합성평가	
	• 교육·직업훈련 • 채용·인사관리·자영기회 • 필수 서비스의 접근·향유 • 법집행 • 이민·난민·출입국관리 • 사법과 민주적 절차의 집행	자율적 적합성평가	
제한적 리스크 (투명성 의무 대상)	• 사람과 상호작용 • 감정인식, 생체정보 기반 범주화 • 딥페이크		• 고지 또는 공개의무
최저의 리스크			

<표1> AI법안의 리스크 기반 접근의 기본 체계

8) EC는 실제 AI 시스템의 대부분은 최저의 리스크 유형에 해당하고 그 이상의 리스크를 초래하는 AI 시스템들은 단계별로 더 적을 것으로 인식하고 있으며, 홈페이지의 그림에서 피라미드 형태로 이 점을 형상화하고 있다(https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en).

9) 법안 Title II에 관련 조항이 있다.

나. 수인불가 리스크(Unacceptable Risk)를 가진 AI시스템[9]

(1) 해당되는 AI시스템

수인불가 리스크의 AI시스템에 해당하여 그 활용이 금지되는 유형은 다음과 같이 열거되어 있다(Art. 5(1)).

• 잠재의식의 조작(subliminal manipulation)(point (a)): 사람의 행동을 중대하게 왜곡하기 위해 동인 또는 타인에게 신체적 또는 정신적 위해를 가하는(또는 그러한 우려가 있는) 방식으로 동인이 알아챌 수 없는 잠재의식 기법을 활용하는 AI시스템이 이에 해당한다.

• 아동·장애인의 착취(exploitation of children or disabled persons)(point (b)): 아동이나 장애인의 행동을 중대하게 왜곡하기 위해 동인 또는 타인에게 신체적 또는 정신적 위해를 가하는(또는 그러한 우려가 있는) 방식으로 동인의 연령, 신체적 또는 정신적 장애로 인한 취약성을 악용하는 AI시스템이 이에 해당한다.

• 공적인 범용 사회적 평점 시스템(general purpose social scoring)(point (c)): 관청이(또는 관청을 대신하여) 일정 기간에 걸쳐 사람의 신뢰도를 사회적 행동 또는 인지·예측된 개인적·인격적 특성에 기해 평가·분류하기 위한 AI시스템으로서, 그 사회적 평점이 (i) 데이터가 본래 생성·수집된 맥락과 무관한 사회적 맥락에서, 또는 (ii) 일정한 사람이나 그 집단 전체에게 정당화되기 어렵거나 사회적 행동 또는 그 심각성에 상응하지 않는 해롭거나 불리한 처우로 이어지는 경우 이에 해당한다.

• 실시간 원격 생체정보기반 식별(real-time remote biometric identification)(point (d)): 공공장소에서 법집행 목적으로 실시간 원격 생체정보기반 식별 시스템을 활용하는 경우다. 여기서 "실시간"이란 생체정보의 포착·비교·식별 모두가 중대한 지체 없이 이뤄진다는 뜻이다(Art. 3(37)). 단, 그 활용이 (i) 실종아동 등 범죄의 특정 잠재적 피해자를 목표로 한 수색, (ii) 사람의 생명·신체적 안전에 대한 특정한 실질적이고 급박한 위협 또는 테러공격의 방지, (iii) (집행위원회 기본 결의 2002/584/JHA의 2(2)조에서 규정하고 관련 회원국에서 장기 3년 이상의 구금선고 또는 구류명령으로 처벌 가능한) 형사범죄의 실행자 또는 피의자의 인지·소재파악·식별·기소 목적으로 엄격히 필요할 경우에는 예외로 한다.[10]

위 내용 중 공적 범용 사회적 평점 시스템은 중국의 사회신용체계(社会信用体系; Social Credit System)를, 실시간 원격 생체정보기반 식별은 중국의 천망공정(天网工程; Skynet Project)을 구체적으로 염두에 두었던 것으로 알려져 있다.

10) 이러한 예외적인 경우 시스템의 활용이 요구되는 상황의 본질과 그러한 활용이 관련자들의 권리와 자유에 미치는 결과를 고려해야 하고, 이 경우 시기, 지역, 대상자 제한 등 이용에 필요하고 비례적인 안전조치와 조건을 준수해야 하며(Art. 5(2)), 해당 회원국의 사법부 또는 독립행정기관의 사전승인(급박할 경우 사후승인)에 의거해야 한다(Art. 5(3)).

(2) 효과

잠재의식 조작, 아동·장애인 착취, 공적 범용 사회적 평점 시스템의 경우 그 출시·서비스·활용이 모두 금지된다(Art. 5(1)(a)~(c)). 실시간 원격 생체정보기반 식별의 경우 금지되는 것은 법집행 목적으로 공공장소에서 실시간으로 활용하는 것에 한하고(Art. 5(1)(d)), 기타 활용이나 출시·서비스는 다음의 높은 리스크를 가진 AI시스템(원격 생체정보기반 식별)으로 분류되어 기록보존 및 인적 감시와 관련된 추가적인 요건이 부가된다.[11)]

다. 높은 리스크(High Risk)를 가진 AI시스템[12)]

(1) 해당되는 AI시스템

AI법안의 Annex II에 열거된 규정들에 따라 (출시 또는 서비스를 위해) 제3자의 적합성평가를 받아야 하는 제품이거나 그러한 제품의 안전장치(safety component)로의 활용이 의도되어 있는 AI시스템은 (그러한 제품으로부터 독립적으로 출시 또는 서비스되는지 여부와 상관없이) 높은 리스크를 가진 것으로 본다(Art. 6(1)). Annex II에는 안전성 규제를 받는 제품들과 관련한 EU 통합 법(regulation) 및 지침(directive)들이 열거되어 있다. 해당 제품은 기계, 완구, 레저·개인용 선박, 승강기, 폭발성 기체 장치와 보호시스템, 전파기기, 고압기기, 케이블카, 개인보호장구, 기체연료 연소장치, 의료기기, 실험실용 진단기기, 자동차, 민항기, 2륜차·3륜차·4륜자전거, 농림업용 차량, 해상장비, 철도시스템으로서, 모두 원칙적으로 출시 전 적합성평가를 거쳐야 한다.

이에 더해 Annex III에 열거된 아래 8개 목적으로의 활용이 의도된 AI시스템 또한 높은 리스크를 가진 시스템으로 본다(Art. 6(2)).

• 생체인증·범주화
(biometric identification and categorization of natural persons):
– 사람의 실시간 및 사후 원격 생체정보기반 식별

• 핵심기반시설의 관리·운영
(management and operation of critical infrastructure):
– 도로교통 및 수도·가스·난방·전력 공급의 관리·운영에 있어서의 안전요소

• 교육·직업훈련(education and vocational training):
– 사람의 교육·직업훈련기관에 대한 접근권의 결정 및 배정
– 교육·직업훈련기관 학생 또는 교육기관 입학에 통상 필요한 시험의 응시자 평가

11)
실시간 또는 사후 원격 생체정보기반 식별의 경우 특히 로그에 ① 시스템 이용기간 기록, ② 시스템이 입력값에 대해 체크한 참조 데이터베이스, ③ 검색 결과 (참조 데이터베이스에) 부합된 입력값, 결과의 검증에 관여한 사람의 식별에 포함되어야 하고(Art. 12(4)), 최소 2인의 사람이 검증, 확인하기 전에는 인증이 이뤄져서는 아니된다(Art. 14(5)).

12)
법안 Title III에 관련 조항이 있다.

• 채용·인사관리·자영기회
(employment, workers management and access to self-employment):
– 구인 광고, 원서의 심사·선별, 면접·시험 과정에서의 지원자 평가 등 채용·선발
– 승진·해고 결정, 직무 배분, 근로관계에 따른 성과·행동의 감시·평가

• 필수적 공공·민간서비스의 접근·향유
(access to and enjoyment of essential private/public services and benefits):
– 관청에 의한(또는 관청을 대신한) 사람의 공공부조·서비스 수급자격평가·부여·감축·철회·환수
– 사람의 신용도의 평가 또는 신용평점의 부여(단, 소규모 공급자가 자가 목적으로 서비스에 활용하는 경우 제외)
– 소방, 구급 등 긴급출동서비스의 배치 및 배치 우선권의 설정

• 법집행(law enforcement):
– 법집행기관의 범죄 또는 재범 위험 또는 형사범죄의 잠재적 피해자의 위험을 평가하기 위한 사람에 대한 개별적인 위험 평가
– 법집행기관의 거짓말탐지기(유사한 도구 포함)로의 활용, 사람의 감정상태 탐지 위한 활용
– 법집행기관의 딥페이크 적발
– 형사범죄의 수사·기소 과정에서 증거의 신빙성 평가
– 법집행기관의 사람에 대한 프로파일링 또는 사람의 인격적 특징·성격·범죄전력의 평가에 기해 실제·잠재적 형사범죄의 발생·재발 예측
– 법집행기관의 형사범죄 인지·조사·기소 과정에서의 사람에 대한 프로파일링
– 사람에 관한 범죄분석(데이터 내 미지의 패턴을 식별하거나 숨은 관계를 발견하기 위해 상이한 데이터 소스·포맷의 복잡·대량의 데이터셋 검색)

• 이민·난민·출입국관리
(migration, asylum and border control management):
– 당국의 거짓말탐지기나 유사한 도구로의 활용 또는 사람의 감정상태의 감지
– 회원국의 영토로 입경하려 하거나 입경한 사람의 안보·불법이민·보건 등의 관련 리스크 평가
– 당국의 사람에 대한 출입국 서류나 보조 서식 검토 또는 위조서류의 적발
– 망명·비자·거주허가 신청 및 입국 지위를 신청하는 사람의 자격 관련 민원에 대한 검토 보조

• 사법과 민주적 절차의 집행
(administration of justice and democratic process):
– 사법당국이 사실관계와 법을 연구, 해석하고 법을 구체적인 사실관계에 적용하는 것의 보조

EC는 (i) 위에 열거된 8개 영역에 활용되는 것을 목적으로 하면서, (ii) 인간의 건강과 안전에 위해를 가하거나 기본권에 부정적 영향을 미칠 리스크를 가지고 있는 AI시스템을 추가하는 방식으로 Annex III의 리스트를 업데이트할 권한을 부여 받고 있다(Art. 7).

(2) 효과

AI법안에 의하면 높은 리스크를 가진 AI시스템은 아래 표에 정리한 요건을 충족해야 한다(Art. 8(1)). 높은 리스크를 가진 AI시스템은 데이터 거버넌스, 투명성, 통제성, 정확성, 견고성, 보안성 등의 요건들을 갖춰야 하고, 이를 담보하기 위해 리스크 관리 시스템, 기술문서, 기록보존 등 프로세스를 마련하여 추적·검증해야 한다는 취지이다.

항목	시스템 요건	각 항목의 세부 구성요소
리스크 관리 시스템 (risk management system) (Art. 9)	- 이러한 시스템을 전 생애주기에 걸쳐 지속적·반복적 프로세스로 수립·실시·기록·유지(정기적·체계적 현행화 포함)(paras. (1)~(4)) - 개발과정, 출시 또는 서비스 전 테스트(paras. (5)~(7)) - 아동에 미치는 영향을 특별히 고려(para. (8))	- 각 시스템의 알려지거나 예상가능한 리스크의 식별·분석 - 시스템이 의도된 목적 하에, 합리적으로 예상가능한 오용 조건 하에 각 활용될 때 야기될 수 있는 리스크의 추산·검토 - 출시 후 모니터링 시스템으로부터 수집되는 데이터의 분석에 기한 다른 발생가능한 리스크의 평가 - 적합한 리스크 관리 조치의 채택 - (para. (2))
데이터와 데이터 거버넌스 (data and data governance) (Art. 10)	- 훈련·검증·시험데이터의 관련성·대표성·무오류성·완전성(para. (3)) - 시스템이 활용될 곳의 지리·행태·기능적 상황에 특유한 특성·요소 고려(para. (4)) - 편향의 모니터링을 위해 꼭 필요한 경우 고유식별정보, 민감정보, 전과를 처리할 수 있으나 이 경우 합당한 보안조치(para. (5))	- 설계상 선택 - 데이터 수집 - 데이터 전처리(애노테이션·레이블링·클리닝·보강·집합화 등) - (특히 데이터가 측정·대표하는 정보에 관한) 가정 설정 - 필요한 데이터셋의 가용성, 규모, 적합성의 사전 평가 - 편향을 고려한 검토 - 데이터 갭 또는 흠결의 식별 및 해결방안 모색 - (paras. (1), (2))
기술문서 (technical documentation) (Art. 11)	- 기술문서를 출시·서비스 전 작성하여 지속적으로 현행화(para. (1)) - 기술문서는 높은 리스크를 가진 AI시스템 요건 준수 여부의 검증이 가능하고 당국에 준수 여부의 조사를 위한 필요한 정보를 전부 제공하도록 작성(para. (1))	- 개요 - 시스템의 요소와 개발 과정의 상세 - 감독·기능·제어의 상세 - 리스크 관리 시스템의 상세 - 생애주기 중 시스템 변경 - 적용되는 EU 통합 표준 - EU의 적합성확인 사본 - 출시 후 성능 평가 시스템의 상세 - (para. (1), Annex IV)
기록보존 (record-keeping) (Art. 12)	- 자동화된 이벤트 기록(로그) 기능을 갖춰 설계·개발(para. (1))	- 시스템의 의도된 목적에 합당한 생애주기에 걸친 기능의 추적 (para. (2)) - 건강·안전·기본권보호 리스크 야기 또는 실질적 변경의 모니터링 기능 및 출시 후 모니터링 기능(para. (3))
투명성 및 활용자에 대한 정보제공 (transparency and provision of information to users) (Art. 13)	- 활용자들이 시스템의 출력을 해석하고 적합하게 활용하기에 충분할 정도로 투명하게 작동되도록 설계·개발(para. (1)) - 설명서(instruction)를 디지털 포맷 등으로 제공(para. (2))	설명서 포함 사항(para. (3)): - 제공자(대표자 포함)의 신원·연락처 - 시스템 성능의 특성·능력·한계 - 최초 적합성평가 당시의 시스템 및 성능의 변경 - 인적 감시 조치 - 예상 내용연수 및 유지·보호조치
인적 감독 (human oversight) (Art. 14)	- 인간-기계 간 인터페이스(HMI) 등 활용 중 사람이 건강, 안전, 기본권에 대한 리스크 여부를 효과적으로 감시할 수 있도록 설계·개발(paras. (1), (2)) - 인적 감독은 출시 또는 서비스 전 제공자가 식별하여 시스템에 (기술적으로 가능하면) 탑재하거나, 활용자가 이행하기에 적합하도록 보장(para. (3))	- 시스템이 할 수 있는 것과 한계를 완전히 이해하고 작동을 적정하게 모니터링하여 이상, 오작동, 예상 밖의 작동을 최대한 빨리 포착하고 해결 - 시스템(특히 사람의 결정을 위해 정보를 제공하거나 추천하는 시스템)이 생성하는 출력을 자동으로 신뢰(과신)하는 경향(automation bias)을 인지 - 시스템의 특성과 툴·방법의 해석을 고려하여 시스템의 출력을 정확히 이해할 수 있어야 함 - 특정한 상황에서 시스템 활용 중단 또는 시스템 출력의 무시·대체·번복 가능 - 중지 버튼 등을 통해 시스템 작동 개입, 중단 가능 - (para. (4))
정확성·견고성·보안 (accuracy, robustness and cybersecurity) (Art. 15)	- 합당한 수준(설명서에 고지)의 정확성·견고성·보안을 달성하고 생애주기 간 이러한 측면에서 일관되게 작동하도록 설계·개발(paras. (1), (2))	- 정확성·견고성의 경우(para. (3)): (특히 사람 또는 다른 시스템과의 상호작용에 있어서의) 시스템·환경에서 발생하는 오차, 오류, 비일관성에 대해 복원력(resilience) 필요(지속학습 시스템은 출력값이 향후의 작동에 입력값으로 투입됨으로써 발생하는 출력값 편향(feedback loop)을 합당한 경감조치로써 해결하도록 개발) - 보안의 경우(para. (4)): 시스템 취약성을 악용하는 제3자의 미인가된 활용 또는 성능의 변경 시도에 대한 복원력 필요(특히 훈련 데이터셋을 조작하려는 공격(data poisoning)의 방지·제어, 모델의 실수를 유발하도록 설계된 입력값(adversarial examples), 모델상의 결함(model flaws) 등 AI에 특유한 취약성을 해결)

<표2> 높은 리스크를 가진 AI시스템의 요건

이러한 시스템의 제공자, 제조자, 수입자, 유통자, 활용자들은 아래 표에 요약된 의무를 각각 부담한다. 앞서 강조하였듯이 활용자는 소비자가 아닌 사업목적 이용자에 한한다. 유통자, 수입자, 활용자 또는 다른 제3자는 자기 명의·상표로 시스템을 출시·서비스하거나, 이미 출시·서비스된 시스템의 의도된 목적을 변경하거나, 시스템에 실질적 변경을 가할 경우 (변경 전 제공자를 갈음하여) 제공자로 간주된다(Art. 28).

의무자	의무	상세
	위 시스템 요건 준수(Art. 16(a))	
	품질관리시스템(quality management system)(Arts. 16(b), 17)	- 이 법의 준수를 담보하는 품질관리시스템을 갖출 의무(상세는 Art. 17(1) 각호에서 규정)
	기술문서 작성(obligation to draw up technical documentation)(Arts. 16(c), 18)	- 기술문서 작성 의무
	자동 생성 로그(automatically generated logs)(Arts. 16(d), 20)	- 시스템이 자동으로 생성하는 로그를 (그러한 로그에 대한 활용자와의 계약 또는 법상 관리권이 있는 한) 보존 의무
	적합성평가(conformity assessment)(Arts. 16(e), 19)	- 출시·서비스 전 적합성평가를 받아 적합성선언 후 CE마크 부착 의무
제공자 (provider)	등록(registration)(Arts. 16(f), 51, 60)	- 출시·서비스 전 시스템(타 시스템 일부인 경우 제외)의 EU데이터베이스 등록
	시정조치(corrective actions)(Arts. 16(g), 21)	- 출시·서비스한 시스템이 이 법에 적합하지 않다고 판단되면(또는 그 이유가 있으면) 적합하게 만들거나, 철회 또는 회수 위해 필요한 시정조치를 취할 의무
	신고의무(duty of information)(Art. 22)	- 시스템에 건강·안전·기본권보호 관련 리스크가 있음을 알게 된 경우, 즉시 시스템이 활용된 회원국의 당국 또는 인증서를 발부한 인증기관에 신고
	당국에 협조(cooperation with competent authorities)(Art. 23)	- 회원국 당국의 요청 시 적합성을 입증하는 데 필요한 모든 정보와 자료(계약이나 법상 로그 관리권이 있으면 로그 포함) 제공
	대표자(authorised representatives)(Art. 25)	- EU 내 출시 전 수입자가 확인이 안되면, EU 역외 제공자는 EU 내 대표자를 선임 - 대표자는 적합성선언과 기술문서 사본 보관, 당국에 정보·자료제공, 협조 의무
	중대사고·오작동 신고(reporting of serious incidents and of malfunctioning)(Art. 62)	- 기본권을 보호하는 EU법상 의무를 위반하는 중대 사고와 오작동 시 후관리기관(MSA)에 신고하고, 사후관리기관은 이를 회원국의 당국에 전달
제조자 (manufacturer) (Art. 24)	준수의무	- 신입법체제에 따른 적합성평가 대상 제품[13] 과 관련된 AI시스템이 출시·서비스되면 자기 명의로 제조한 제조자도 이 법의 준수에 대한 책임을 지고, 제공자와 마찬가지의 의무를 부담
수입자 (importer) (Art. 26)	제공자의 적합성평가 등 확인(para. (1))	- 출시 전 제공자의 적합성평가, 기술문서, CE마킹 부착, 관련 문서와 사용설명서 첨부 확인
	출시 중단 등(para. (2))	- 시스템이 이 법에 적합하지 않다고 판단되면(또는 그 이유가 있으면) 출시 중단 - 건강·안전·기본권보호 관련 리스크 있을 경우 제공자에 고지하고 사후관리기관에 신고
	정보 제공(para. (3))	- 명칭, 등록상호·상표, 주소를 시스템 상(불가능하면 포장 또는 관련 문서에) 표기
	보관·운송 상 유의사항(para. (4))	- 시스템을 책임지는 동안 보관, 운송 조건이 요건 준수를 위협하지 않도록 조치
	당국에 협조(para. (5))	- 회원국 당국의 요청 시 모든 정보와 자료(로그 포함) 제공, 협조
유통자	CE마킹 등 확인(para. (1))	- 출시 전 CE마킹 부착, 관련 문서와 사용설명서 첨부, 제공자와 수입자의 이 법상 의무 준수 여부 확인
	(distributor)	(위 수입자의 경우와 동일)
	(Art. 27)	- 출시한 시스템이 이 법에 적합하지 않다고 판단되면(또는 그 이유가 있으면) 적합하게 만들거나, 철회 또는 회수 위해 필요한 시정조치를 하거나, 제조자·수입자·관련자가 시정조치를 하게 할 의무
활용자 (user) (Art. 29)	사용설명서 준수(paras. (1)~(3))	- 사용설명서에 따라 활용(인적 감시 목적의 활용은 예외), 입력값은 시스템의 의도된 목적에 비추어볼 때 관련성 필요
	시스템 작동 감시(para. (4))	- 사용설명서에 따라 모니터링 - 사용설명서에 따른 활용이 건강·안전·기본권보호 관련 리스크 발생시킨다고 볼 이유가 있거나 이와 관련된 중대사고·오작동 발견 시 제공자·유통자 고지
	로그 기록(para. (5))	- 시스템이 자동으로 생성하는 로그를 (그 로그가 관리 하에 있는 한에서) 기록
	제공받은 정보의 활용(para. (6))	- 투명성 원칙(Art. 13)에 따라 제공받은 정보를 일반정보보호법(GDPR) 등에 따른 정보보호영향평가(DPIA)를 위해 활용

<표3> 높은 리스크를 가진 AI시스템 관련자의 의무

13) 기계, 완구, 레저·개인용 선박, 승강기, 폭발성 기체 장치와 보호시스템, 전파기기, 고압기기, 케이블카, 개인보호장구, 기체연료 연소장치, 의료기기, 실험실용 진단기기가 이에 해당한다.

높은 리스크를 가진 AI시스템에 대한 나머지 조항들은 적합성평가 절차를 상세하고 규정하고 있는데(Chapters 4~5, Arts. 30~51), 이 중 핵심적인 부분을 요약하면 다음과 같다.

• 인증기관(notified bodies): 적합성평가기관(conformity assessment bodies)은 각 회원국들에 설치된 통보기관(notifying authorities)에 의해 인증기관으로 지정되어 감독을 받게 된다(Arts. 30, 31). 이러한 인증기관은 높은 리스크를 가진 AI시스템의 적합성을 검증한다(Art. 33(1)).

• 적합성평가(conformity assessment): 제공자는 원격 생체정보기반 식별, 핵심기반시설의 관리·운영 시스템의 경우에는 (i) EU통합기준을 적용하여 준수를 입증하는 경우 자율적합성평가(conformity assessment based on internal control) 또는 제3의 인증기관의 품질관리시스템 및 기술문서에 기한 적합성평가를 받으면 되나, (ii) EU통합기준을 미적용 또는 부분적용하거나 통합기준이 없는 경우 제3의 인증기관의 적합성평가를 받아야 한다(Art. 43(1)). 교육·직업훈련, 채용·인사관리, 필수서비스 접근권, 법집행, 이민·난민·출입국관리, 사법과 민주절차의 집행 시스템의 경우에는 제공자가 자율적합성평가를 받으면 된다(Art. 43(2)). 자율적합성평가는 제공자가 품질관리시스템, 기술문서 상의 정보, 시스템의 설계개발 ·과정 및 출시 후 모니터링이 이 법을 준수한다는 점을 스스로 확인하는 방식으로 이루어진다(Annex VI). 인증기관은 적합성평가 통과 시 확인서(certificate)를 발급한다(Art. 44).

• 적합성선언(declaration of conformity): 제공자는 각 시스템에 대해 적합성선언을 작성하여 출시 또는 서비스 개시 후 10년간 보관해야 한다(Art. 48).

• CE마킹(CE marking): 제공자는 CE마킹을 시스템 상에 (불가능하거나 어려우면 포장 또는 부속문서에) 부착해야 한다(Art. 49). CE는 유럽공동체(EC)를 일컫는 프랑스어 Communauté européenne의 약자로서, 회원국마다 난립하던 인증제도를 1993. 7. 22.자 EC 결의(Council Decision 93/465/EEC)로 CE마킹으로 통합하여 오늘에 이르고 있으며, 우리나라의 국가통합인증마크(KC)에 상응한다.

<그림 1> CE마킹

14) 법안 Title IV에 관련 조항이 있다.

라. 제한적 리스크(Limited Risk)를 가진 AI시스템 — 투명성 의무 대상[14)]

(1) 해당되는 AI시스템

투명성 의무 대상이 되는 AI시스템에는 다음 세 유형이 해당된다(Art. 52).

• 사람과 상호작용하도록 의도된 AI시스템(AI systems intended to interact with natural persons)(para. (1)): 챗봇 등 대화형 에이전트의 경우가 대표적이다. 단, (i) AI시스템과의 상호작용 사실이 상황 및 활용 맥락상 명백한 경우, (ii) 법에 따라 형사범죄를 인지·예방·수사·기소하는 시스템의 경우(공중이 이를 범죄신고에 이용하는 경우 제외)는 예외다.

• 감정인식 시스템(emotion recognition system), 생체정보 기반 범주화 시스템(biometric categorization system)(para. (2)): 법상 형사범죄의 인지·예방·수사를 위한 활용이 허용된 생체정보 기반 범주화 시스템은 제외된다.

• 실존하는 사람·대상·장소 또는 다른 주체·사건에 눈에 띄게 닮았고 거짓으로 진정·진실된 것처럼 보이는 화상·시각·청각 콘텐츠를 생성·조작하는 시스템(AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful)(para. (3)): 소위 딥페이크(deep fake)를 생성하는 시스템을 의미한다.

유의할 점은 투명성 의무는 높은 리스크를 가진 AI시스템의 경우에도 부과될 수 있다는 것이다. 즉, 특정한 AI시스템이 법안 Title IV에 따라 투명성 의무 대상이 된다고 하여 높은 리스크를 가진 AI시스템의 유형에서 바로 배제되는 것은 아니다(Art. 52(4)).

(2) 효과

제한적 리스크를 가진 AI시스템의 경우 제공자나 활용자에게 고지 또는 공개 의무가 부과된다. 사람과 상호작용하는 시스템의 경우 제공자는 그 사람이 AI시스템과 상호작용 중이라는 사실을 고지받는 방식으로 시스템을 설계·개발해야 한다(para. (1)). 감정인식 또는 생체정보 기반 범주화 시스템의 경우 활용자는 그러한 시스템의 작동을 그에 노출된 사람에게 고지해야 한다(para. (2)). 딥페이크의 경우 활용자는 그 콘텐츠가 인공적으로 생성·조작되었다는 점을 공개해야 한다(para. (3)).

마. 최저의 리스크(Minimal Risk)를 가진 AI시스템

최저의 리스크를 가진 AI시스템의 경우 법안상의 규제가 적용되지 않는다. 이러한 시스템에 대해서도 EC와 회원국들이 행동강령(code of conduct)을 정해서 권고할 수는 있으나(Art. 69), 자발적 준수 사항에 불과하다.

바. 기타

AI법안의 나머지 부분은 규제샌드박스(Title V, Arts. 53~55), 유럽AI위원회와 회원국별 기관의 설치(Title VI, Arts. 56~59), 집행절차(Title VIII, Chapter 3, Arts. 63~68), 기밀유지와 벌칙(Title X, Arts. 70~71), 권한위임(Title XI, Arts. 73~74), 부칙(Title XII, Arts. 75~85)으로 구성되어 있는데, 우리나라 입장에서 특기할 만한 사항은 아래와 같다.

• 벌칙(Art. 71): 수인불가 리스크를 가진 AI시스템의 금지(Art. 5) 및 높은 리스크를 가진 AI시스템의 데이터와 데이터 거버넌스 요건(Art. 10) 위반은 3천만 유로, 회사의 경우 직전사업연도 전세계 총매출액 6% 중 높은 액수 이하의 과징금에 처하며(Art. 71(3)), 기타 법 위반은 2천만 유로, 회사의 경우 직전사업연도 전세계 총매출액 4% 중 높은 액수 이하의 과징금에 각각 처한다(Art. 71(4)). 최근 EU의 디지털 및 데이터산업 관련 여타 규제와 마찬가지로 제재의 가능 수위가 상당히 중하다.

• 유럽AI위원회(European Artificial Intelligence Board)(Arts. 56~58): 새로이 창설될 유럽AI위원회는 각 회원국의 감독당국의 장 및 동등한 고위공무원 및 유럽정보보호관(EDPS)으로 구성되어 EC를 위한 자문과 보조를 수행한다.

2. AI법안에 대한 평가와 시사점

AI법안을 통해 EU는 데이터 프라이버시에 이어 AI 거버넌스에 대해서도 적극적인 목소리를 내겠다는 의지를 보여주고 있다. AI법안이 입법되면 2016년의 일반정보보호법(GDPR)과 더불어 EU가 디지털 경제의 혈류와 신경이라 할 수 있는 데이터와 알고리즘 모두에 대해 규제의 선발주자로서 국경이 사라져가고 있는 글로벌 경제의 규제 방향에 대해 상당한 영향을 미칠 것으로 예상된다.[15)]

우선 AI법안은 EU가 표방해 온 신뢰(trust)와 수월성(excellence)에 기반한 에코시스템을 구축하기 위해,[16)] EU 경내에서의 안전과 기본권의 존중을 강조하면서 동시에 기술혁신의 조장과 규제의 예측가능성 제고를 염두에 두고 마련된 것으로 설명되고 있다.[17)] 특히 AI 기술이 시장에 적용되기 이전부터(적합성평가) 그 이후까지(출시 후 모니터링) 전(全) 단계를 아우르는 규제 프레임워크를 구축하려고 노력한 점은 주목할 만하다. 또한 학습을 통해 계속 개량되는 AI

15) EU의 선도적인 규제가 글로벌 경제와 규제에 미치는 영향을 잘 보여주는 예로 Anu Bradford, "The Brussels Effect - How the European Union Rules the World" (Oxford 2020) 참조. 저자는 이러한 영향을 "브뤼셀 효과(Brussels Effect)"라 부르고 있다. EP의 연구 보고서에서도 AI에 대한 글로벌 거버넌스 모델이 부재한 상황에서 EU가 선제적으로 규제를 입안함으로써 GDPR의 사례처럼 규제의 선발주자로서의 이득('first mover' advantage)을 취하고, 이를 미·중을 따라잡는 수단으로 활용하는 방안에 대해 언급하고 있다(European Parliamentary Research Service Study, "European framework on ethical aspects of artificial intelligence, robotics and related technologies - European added value assessment" (Sep. 2020)(이하 "EPRS 연구보고서"), para. 4.2.2.).

16) EU AI백서 등 참조.

17) 설명메모 1.1.항.

18) 설명메모 5.2.1.항.

19) EPRS 연구보고서, 2.4.2. 및 4.2.2.항.

20) GDPR의 경우 법안의 공식 발표는 2012. 1.에 이루어졌고, 매우 복잡한 논의 및 법안 수정의 과정을 거쳐 최종 법안은 2016. 4.에 통과되었다. 그리고 2018. 5.부터 집행되기 시작하였다. 법안 발표에서 집행 단계에 이르기까지 6년이 넘는 오랜 기간이 소요되었고, 그 과정에서 법안의 내용도 적지 않게 수정되었다.

21) European Parliament resolution of 20 October 2020 on a framework of ethical aspects of artificial intelligence, robotics and related technologies, 2020/2012(INL).

기술의 특성을 고려하여 장래에 있어서도 규제의 실효성을 담보할 수 있는 법기술적인 유연성(future proof)도 추구하고 있다.[18)] 한편 AI법안의 기저에도 AI 기술의 개발과 활용에 대한 민주적 통제(democratic oversight)의 필요성과 이를 통한 정당성(legitimacy)의 확보가 규제의 중요한 목표라는 EU의 시각이 흐르고 있다.[19)]

AI법안은 법(regulation) 형식이므로 지침(directive)과 달리 유럽의회에서 가결되어 시행되면 회원국들의 개별 입법을 요하지 않고 바로 회원국 내에서 효력이 발생한다. 다만 AI법안의 내용을 살펴봄에 있어 상당한 기간 동안 이루어질 심의 과정에서 유럽의회와 회원국들의 의견에 따라 그 내용이 달라질 수 있음을 주지할 필요가 있다.[20)] 또 AI법안이 가결되는 경우에도 그 시점이 현재로서는 확실치 않다. 하지만 2019년 EC 의장 당선 과정에서 AI 법의 시급한 제정을 공약한 폰데어라이엔(Ursula von der Leyen) 의장의 임기가 2024년까지 남아 있고, 동 법안이 유럽의회가 2020. 10. 20. EC에 대해 AI 관련 윤리적 원칙의 보호 등을 위한 입법절차를 촉구한 것[21)] 에 대한 응답도 겸하고 있다는 점에서 입법에 대한 상당한 추동력이 있다는 점은 부인할 수 없다.

이러한 불확실성에도 불구하고 AI법안의 광범위한 적용 범위, 글로벌 규제에 대한 잠재적 영향, 수출 시장으로서 EU의 중요성을 고려할 때, AI법안의 내용을 살펴보고 우리나라에 미칠 영향에 대해 미리 생각해 볼 필요가 있어 아래에 몇 가지 시사점을 논하고자 한다.

가. 리스크 기반 접근의 수단으로서의 인증(적합성평가) — 무분별한 수용의 위험

EU는 AI법안에서 리스크가 높은 AI를 CE인증, 즉 적합성평가의 틀로 규제하고자 한다. 우리나라도 일반 재화의 경우 이에 상응하는 KC인증 시스템이 있고,[22)] KC인증의 틀로 AI를 규율할 수 있는 기반 작업이 이미 국내법 체계에서도 어느 정도 이루어져 있다고 볼 수 있다. 구체적으로 전파법 제58조의2 제1항은 방송통신기자재등의 적합성평가 시 기술기준, 전자파흡수율(specific absorption rate) 등 전자파인체보호기준, 전자파적합성(electromagnetic compatibility) 뿐 아니라 "다른 법률에서 방송통신기자재등과 관련하여 과학기술정보통신부장관이 정하도록 한 기술기준이나 표준"(제7호)에 따르도록 하고 있다. 그런데, 지능정보화 기본법 제21조 제1항은 과학기술정보통신부장관이 지능정보기술의 안정성 · 신뢰성 · 상호운용성 등을 확보하기 위하여 필요한 기술기준을 정하여 고시할 수 있다고 규정하고, 제2항은 대통령령으로 정하는 국민의 생명 또는 신체안전 등에 밀접한 지능정보기술에 관련된 사업자는 과학기술정보통신부장관이 정하여 고시하는 기준에 적합하도록 지능정보기술을 개발 · 관리 · 활용하여야 한다고 규정하고 있다. 이에 따라 지능정보화 기본법 시행령 제16조 제2항은 "대통령령으로 정하는 국민의 생명 또는 신체안전 등에 밀접한 지능정보기술"을 (i) 군

22) KC인증 시스템의 유관 기관으로는 전파연구원(전파법에 따른 방송통신기자재등 적합성평가 담당, 과학기술정보통신부 소속), 국가기술표준원("전기용품 및 생활용품 안전관리법"에 따른 전기용품 및 생활용품의 안전인증·안전확인·공급자적합성확인 담당, 산업통상자원부 소속), 한국교통안전공단(자동차관리법에 따른 자동차 및 자동차부품 자기인증 담당, 국토교통부 소속) 등이 있다.

사적 목적으로 개발 · 관리 · 활용하려는 지능정보기술, (ii)「의료법」 제24조의2 제1항에 따른 수술등의 의료행위에 직접 이용되어 사람의 신체에 영향을 미칠 수 있는 지능정보기술, (iii) 지능정보기술이 오작동(誤作動)될 경우 사람에게 중대한 위해(危害)를 끼칠 우려가 있는 지능정보기술로 정하고 있다. 결국 과학기술정보통신부가 향후 이러한 기준을 고시하면 사전규제로 작용하게 되고, 전파법상 적합성평가의 기준도 될 수 있어 EU의 AI법안과 비슷한 체제가 될 수 있는 것이다.

그러나 AI법안이 시도하고 있는 AI시스템에 대한 인증은 리스크에 대한 사전적인 대처 등의 장점은 있지만 다음과 같은 문제점도 함께 가지고 있어 그 무분별한 수용은 위험하다고 생각된다.

(1) AI법안은 아직 제대로 출시되거나 시험되지 않은 AI를 포함한 다양한 유형의 AI를 리스크에 대한 경험적 증거에 입각한 비용편익분석(cost-benefit analysis)보다는 사전주의 원칙(precautionary principle)[23] 에 기하여 강도 높게 규제하려는 입장을 취하고 있는 것으로 보인다. 수용불가 또는 높은 리스크를 가진 것으로 열거된 AI시스템 유형의 대다수는 아직 제대로 활용되거나 사업화조차 시도하지 못한 개발 초기 상태에 머물러 있다. 이러한 AI시스템이 활용 용도 별로 인간 및 사회와 어떻게 상호작용하여 어떠한 위해(harm)를 야기하고 어떠한 시장실패(market failure)로 이어지며 규제가 어느 정도까지 효과적인 역할을 할 수 있을지에 대한 실증적 근거가 매우 부족한 실정이다. 되돌릴 수 없는 손해(irreversible damage)가 분명히 예견되는 상황도 아니어서 이러한 사전주의적 접근 방식 자체의 적절성에 대해 의문이 들 뿐만 아니라, 현실에 아직 존재한 적도 없는 기술을 틀어막듯 이루어지는 규제 방식은 연구개발 과정 자체를 과도하게 위축시키고 기존 사업자는 물론 스타트업들의 경쟁력을 손상시킬 우려가 있다. 규제 대상인 AI시스템에 대해서는 데이터의 관련성·대표성·무오류성·완전성을 기하려 하면서 정작 법률안 자체는 선험적(a priori) 가정에 입각할 뿐 리스크의 정확한 평가를 위한 제대로 된 데이터의 뒷받침이 없다는 것 또한 아이러니하다. EU의 입장에서는 디지털 기술의 개발과 적용이 뒤쳐져 있고 국제적인 경쟁력을 갖춘 자국 플랫폼의 성장을 당분간 기대하기 어려워 소비자 보호나 통상의 관점에서 접근하면서 규제를 도구로 삼아 글로벌 운동장의 평탄화를 꾀하는 것이 이해못할 바는 아니다. 그러나 자국 플랫폼과 정보기술 기업들이 국내는 물론 해외 정보기술 시장(예를 들어 일본의 모바일 인스턴트 메시징(MIM) 시장 등)에 활발히 진입하고 성장하고 있는 우리나라로서는 EU와는 입장이 다르므로 규제 방식이나 내용에 있어서도 우리 실정에 맞는 보다 전략적인 접근을 할 필요가 있다.

(2) AI법안은 각 AI시스템의 다양한 활용 방식과 그에 대한 수요 등의 개별적 특성을 고려하여 세분화된 부문별(sectoral) 규제를 하기 보다는 이질적 기술과 용도를 한데 묶어 일괄(omnibus) 규제를 하려는 경향이 강하고 일정한 유형의 AI시스템에 대해 인증 절차를 일반적으로 적용하는 것도 일괄 규제의 성격이 있는 것으로 보인다. 그러나 개별 영역에 따라 다른 판단이 필요한 경우가 적

23) "사전배려 원칙" 또는 "사전예방 원칙"으로 번역되기도 한다.

지 않다. 예를 들어, 의료진단기기의 경우 정확성(accuracy)(특히 민감도(sensitivity))이 핵심적으로 중요하고, 자동차의 경우 공중장소(public space)에서의 안전성(safety)과 책무성(accountability)이 중요하다. 한편 공공부문의 경우 공정성(fairness)(특히 채용·사회보장·기반시설접근권의 경우 분리성(separation), 사법·출입국관리의 경우 충분성(sufficiency)) 확보가 중요하고, 생체인증의 경우 프라이버시가 특히 강조될 필요가 있다. 이처럼 동일하거나 유사한 AI시스템이라 할지라도 구체적으로 활용되는 분야 및 용처에 따라 제기되는 문제의 본질이 상이할 수 있다. 이러한 상황에서 현 AI법안은 각 AI시스템별로 나타나는 특유의 시장실패 요인에 집중하여 정확하게 규제하는 것이 아니라, 광범위한 유형의 AI에 대해 잠재적인 문제까지 포함하여 발생가능한 모든 이슈를 상정한 후 관련 규제들을 한꺼번에 획일적으로 규제하는 방식을 채택하고 있다. 상이한 부문에 걸쳐 다른 법체계가 진화해 온 것은 각 부문의 특수한 시장실패에 맞춤형으로 대응하는 과정에서 발생하는 자연스러운 결과인데, 단지 일정한 유형의 AI를 (추가로) 적용했다는 이유만으로 일괄적인 규제를 부과하는 것은 또 다른 부작용을 불러올 가능성이 높다.

조금 더 일반적으로는 AI법안 하에서 적용되는 규제와 의무의 사항들이 지나치게 세세한 점도 문제다. 법안이 채택한 AI시스템의 정의를 보면 회귀분석 등 AI 이전의 전통적인 통계분석을 망라하고 있어 데이터나 실증에 기반한 거의 모든 산업활동이 포섭될 수 있다. 그에 따라, 의도한 바는 아닐지 라도 AI법안이 실질적으로는 전산업적 규제로 귀착될 가능성이 있다. 그런 성격에 비해 규제 항목들은 과하게 세세하거나 형식적인 조치를 요구하여 준수비용이 상당하고,[24] 어느 수준까지 의무를 달성해야 하는지 구체적인 기준을 제시하고 있지 못하다. 일례로 훈련·검증·시험데이터의 대표성, 무오류성, 완전성을 법적 의무로 명시하고 있는데(Art. 10), 이는 현실성이 떨어진다. 주어진 데이터 가용성 등의 근본적인 한계 내에서 가급적 합리적인 샘플링과 예측정확도를 추구해야 할 뿐인 AI 개발의 현실이 적절히 고려되지 않은 것으로 보인다. AI로 훈련된 모델이 몇 퍼센트의 예측정확도(predictive accuracy)를 달성해야 법적 의무를 이행한 것인지 하는 등의 구체적인 기준이 마련되지 않는 한 불확실성으로 인한 비용이 상당할 것으로 보인다. 다른 한편, 이러한 기준을 구체화하여 제시하는 것 자체가 매우 어렵다는 또 다른 한계가 있다. 이와 별도로, 다양한 형태로 나타날 수 있는 잠재적·실제적 상충(tradeoff) 관계를 어떻게 조화롭게 조율할 것인지에 대한 문제도 있다. 예를 들어, 높은 리스크를 가진 AI의 의무 중 정확성과 데이터 거버넌스(특히 공정성) 사이에 그리고 인적 감시와 프라이버시 사이에 상충 관계가 나타날 수 있는데, 이와 같이 상충하는 의무를 어떻게 조화롭게 해결 또는 충족할 수 있는지에 대한 기준도 결여되어 있다. 또한, 자동화편향(automation bias), 데이터오염(data positioning), 적대공격(adversarial attack), 모델결함(model flaw) 등 최근에 학계에서 연구되고 있는 일부 개념이나 현상들이 그대로 의무사항의 일부로 언급되고 있기도 한데(Arts. 14, 15), 제재가 수반되는 법적 의무 사항에 이러

24) 입안자들도 이러한 문제를 인식하여 중소사업자 및 스타트업들의 부담을 경감시키기 위한 노력을 기울였다고 설명하고 있으나(AI법안 Title V 및 설명메모 1.1., 3.3., 5.2.5.항 등), 법안이 의도하는 효율적인 규제 수준에 도달하기까지 규제자들의 전문성 축적 및 역량 강화 과정에서도 적지 않은 비용이 초래될 것으로 예상된다.

한 용어들이 명시적으로 언급되는 것이 어떤 기술적, 정책적 함의를 가질 것인지에 대해서도 보다 많은 검토가 필요하다.

AI법안은 특히 기술중립성(technological neutrality)을 충분히 지키지 못하고, 규제를 통해 정부가 승자(winner) 기술을 현실적으로 선택하는 왜곡이 발생할 수 있다는 문제가 있다.[25] 수범자들의 예측가능성을 높이기 위한 목적일 수는 있지만, AI시스템을 정의하면서 기계학습, 베이스 추정법 등 특정한 기술들을 Annex I에 열거(개정 가능)하는 방식으로 정의를 내리고 있는 것이다. 과거 국내 공인인증서 의무화와 관련한 시행착오에서 경험하였듯이 기술특유적 규제의 가장 큰 문제는 시장에서 선택할 수 있는 다양한 형태의 기술 중 특정 형태의 기술의 채택을 강요하는 결과가 초래될 수 있고, 이는 다양한 기술의 실험·경쟁·혁신을 저해하는 효과를 가져온다는 것이다.[26] AI법안에 담긴 AI시스템의 개념은, EC의 AI 전문가그룹이 법안 발표에 앞서 AI를 행동과정 중심(즉, 인식(percept), 판단(reasoning), 동작(action))으로 정의했던 것[27] 에 비해 도리어 퇴보한 면이 있는 것이어서 아쉬움을 준다.

AI법안의 이러한 난점들을 고려하면 동 법안의 접근과 내용을 추종하기보다는 일단 각 분야에 이미 존재하던 법령들이 AI시스템으로 인한 시장실패를 효과적으로 규율할 수 있는지,[28] 그렇지 않다면 어떻게 개선될 수 있는지를 개별적으로 검토하면서 그 과정에서 부분적으로 참고하는 방향이 타당하다고 생각한다. AI법안을 참고하는 경우에도 공공부문과 민간부문을 명확히 구분하여 투트랙(two track)으로 접근할 필요가 있다. AI법안에 높은 리스크를 가진 AI시스템으로 열거된 항목 중 상당수가 공공부문에서 적용되는 것이다(핵심기반시설, 공공부조·사회서비스, 소방·구급, 법집행, 이민·난민·출입국관리, 사법·민주적 절차의 집행). 그런데, 최근의 인천공항과 한국공항공사의 AI채용과 관련된 논란과 정보공개 청구 관련 분쟁에서 볼 수 있듯이 국가와 공공기관은 헌법적 원리를 직접적으로 적용 받는 관계로 충분한 투명성의 확보, 공정성에 대한 검증 및 그에 관한 사회적 합의 없이는 AI를 적극적으로 활용하기 어려울 수 있다. 따라서 전략적으로 인증제를 공공부문에서의 적극적인 AI활용을 가능케 해 주는 디딤돌로 고려해 볼 수도 있을 것이다. 지금도 정부 조달 정보보호시스템은 정보보호제품 평가인증(CC인증)을 받아야 하고(국가정보화 기본법 제38조, 시행령 제35조), 정부 조달 암호제품은 암호모듈검증(K-CMVP)을 받아야 하므로(전자정부법 제56조 제3항, 시행령 제69조 제1항 제1호), 이러한 인증제의 연장선상에서 접근하는 것을 구상할 수도 있다. 이와 달리 민간영역의 경우 인증제를 무차별적으로 시행하고 강력한 사후규제까지 부과하면 신기술의 시장진입 자체가 어려워지고 정부실패의 폐해가 커질 가능성이 있다.

25) 설명메모 5.2.1.항에서는 법안이 기술중립적이 되도록 노력했다고 언급하고 있다.

26) 미국 클린턴 행정부의 글로벌 전자상거래 프레임워크(Framework for Global Electronic Commerce of 1997)도 "정부의 규제 시도는, 특히 그러한 규제가 기술특유적인 한에서, 입안될 당시 이미 낡아 있을 가능성이 높다"고 지적한 바 있다(President William J. Clinton & Vice President Albert Gore, Jr., "A Framework for Global Electronic Commerce" (1997), https://clintonwhitehouse4.archives.gov/WH/New/Commerce/).

27) "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions." (High-Level Expert Group on Artificial Intelligence, European Commission, A Definition of AI: Main Capabilities and Disciplines (2019), p. 6).

28) EC도 AI법안이 디지털 서비스법(Digital Services Act)을 포함한 EU의 타 법규와 조화롭게 입안되었음을 강조하고 있으나(설명메모 1.2.항 등), 위에서 언급한 것과 같은 (i) 기존 법규를 통한 규율가능성에 대한 진지한 검토와, 그에 기반한 (ii) 규율범위의 최소화 노력은 부족한 것으로 보인다.

29) 이와 동일하지는 않지만 AI법안에서도 서로 입장이 다를 수 있는 회원국간의 협조를 이끌어내고 조율하기 위한 기구로 유럽인공지능위원회(European Artificial Intelligence Board)의 설립을 예정하고 있다는 점은 참고할 만하다(Title IV, Chapter 1).

30) 동 협정 제2장 부록 2-나-1, 1-가, 전파법 제58조의8, 「한·EU FTA 체결에 따른 방송통신기자재등의 적합성평가 상호인정에 관한 고시」(과학기술정보통신부고시 제2017-7호).

나. AI 규제 정책에 대한 조율 메커니즘의 정립

AI법안의 CE인증에 상응하는 우리나라의 KC인증은 현실적으로는 각 부처간의 치열한 권한 충돌의 장으로 작동하는 면이 있고, AI법안상 높은 리스크를 가진 AI시스템은 우리나라의 경우 여러 부처에 걸쳐 규제 권한이 분산되어 있는 것이 사실이다. 최근 온라인 플랫폼에 관한 구체적인 규제 권한을 어느 기관에 귀속시킬지 논란이 되고 있는 것에서 보듯 이러한 문제는 비단 인증제에만 국한된 것은 아니다. AI법안의 경우에도 특정 부처가 적극적으로 국내 계수를 추진하면 다른 부처들이 이를 가만히 지켜볼 수만은 없어 누구도 원치 않은 '규제 레이스'가 촉발될 우려도 있다. 디지털 경제의 규제 전반에 관하여 제기되는 것이지만 AI와 관련해서도 개별 규제들의 중첩적·누적적 집행으로 인한 사회적 비용이 발생하지 않도록 여러 부처 사이의 조정 기능이 활성화되어야 하며, 산업계와 시민사회 모두의 목소리를 경청하여 기술혁신을 촉진하면서 동시에 그로 인한 혜택에서 소외되는 집단이나 구성원이 생기지 않도록 정책을 조율하는 메커니즘의 정립이 절실하게 요구된다.[29]

다. EU와의 상호인정협정(Mutual Recognition Agreement)의 선제적 준비

AI법안은 EU가 유럽적 가치를 관철하고 자국 시장을 보호하기 위해 AI 분야에서는 물론이고 그 외의 영역에서 인증제의 활용도를 높이려는 움직임의 신호탄으로 보인다. AI법안은 제3국법에 의해 설치된 제3국의 적합성평가기관이 EU와 협정 체결 시 인증기관(NB)의 활용을 수행할 권한이 부여될 수 있다고 규정하고 있다(Art. 39). 여기서 EU와의 협정이란 상호인정협정을 의미한다. 한·EU FTA에 따라 우리나라의 경우에는 전파법상 방송통신기자재, EU의 경우에는 전자파적합성 또는 전기안전 기준 적용범위에 해당하는 기자재에 대해 각각의 시험성적기관의 시험성적서를 인정해 주고 있다.[30] 즉, 적합성평가 자체가 서로 인정되는 2단계 상호인정은 아니지만, 그 전단계인 시험성적서를 상호 인정해 주는 1단계 상호인정의 단계에 있는 것이다. 국내 정보통신기업 입장에서는 1단계 상호인정 하에 시험성적서를 국내 시험기관에서 받을 수만 있어도 그만큼 해외에 민감한 영업비밀이 유출될 우려가 경감될 것이다. 그러나, 이에 머무르지 않고 가능하다면 캐나다와 같이 2단계 상호인정으로 나아갈 필요가 있고, 적용 범위도 현재의 전자파적합성(EMC)에서 보다 넓은 영역으로 확대할 필요가 있다. 따라서 AI법안의 통과, 기타 인증제의 확대 움직임에 대비하여 한·EU FTA의 개정 또는 상호인정협정의 개별적 체결을 통해 AI 관련 상호인정협정을 미리 준비할 필요가 있다. 이러한 준비가 결실을 맺는다면, AI법안이 통과됨에 따른 국내 수출기업 및 디지털 산업의 충격도 완화될 수 있을 뿐 아니라, 이러한 대비를 제대로 하지 않은 제3국들에 비해 EU 시장에서 경쟁력을 강화하는 기회가 될 것이다.

라. 미국과의 조율 및 공조

AI법안의 발표 후 미국과 EU간 이견을 조율하기 위한 노력들이 시작되었으나,[31] 법안 통과 및 시행 과정에서 미국과의 갈등 소지가 적지 않아 보인다. AI법안이 현 모습대로 통과되면 당장 AI의 개발과 사업화를 선도 중인 미국의 플랫폼 및 테크 기업들이 광범위한 규제 리스크를 부담하게 되고, 최대 전세계 총매출액의 6%까지의 막중한 제재의 위험에 노출된다. 더욱이 제3의 인증기관(NB)의 적합성평가 과정에서 동 기관에 AI 알고리즘을 공개해야 할 가능성이 열리게 된다. 그 이유는 적합성평가 시 심사 대상인 기술문서(TD)에 시스템 디자인 시방서(design specifications of the system)를 포함해야 하는데, 여기에는 (i) AI시스템과 알고리즘의 일반적 로직, (ii) 핵심 디자인 방식, (iii) 주요 분류 방식, (iv) 시스템의 최적화 대상 및 상이한 모수(parameter)의 관련성, (v) 법적 요건을 준수하기 위해 채택된 기술 솔루션과 관련된 상충관계와 관련한 결정이 들어가야 하기 때문이다(Art. 11(1), Annex IV, 2(b)). 미국은 북미자유무역협정(US-Mexico-Canada Agreement (구 NAFTA)) 체결 이후 이를 모델 삼아 우방국들에게 디지털무역협정 체결을 요청하고 있고, 알고리즘 공개를 요구하는 규제의 금지가 그 핵심적 내용 중 하나를 이루고 있다. 일례로 미국이 일본과 2019. 10.7.자로 체결한 미일디지털무역협정(US-Japan Digital Trade Agreement; 日米デジタル貿易協定) 제17조 제1항도 "어떠한 일방의 체약국도, 타방의 체약국의 사람이 소유한 소프트웨어 또는 해당 소프트웨어를 포함한 제품의 일방의 체약국의 영역으로의 수입·유통·판매 또는 사용의 조건으로, 해당 소프트웨어의 소스코드의 이전, 해당 소스코드로의 액세스 또는 해당 소스코드에 의해 표현된 알고리즘의 이전 또는 해당 알고리즘으로의 액세스를 요구할 수 없다"[32] 고 규정한다. 디지털서비스법 패키지와 마찬가지로 AI법안이 규정하는 내용은 미국이 구상하고 있는 디지털무역질서의 방향과 정면으로 충돌하기 때문에, 미국으로부터의 이의제기가 지속적으로 이루어질 것이고 그 과정에서 안전지대(safe harbor)의 논의 등 어떤 형태로든 절충 과정은 불가피할 것으로 예상된다.

이러한 상황에서 우리나라가 미국의 이러한 구상을 무시하고 유럽식 접근을 조율 없이 수용하면, 향후 미국이 디지털무역협정 체결을 요구할 경우 난감한 상황에 처할 수 있다. 미국과의 협정 체결 과정에서 기존에 계수해 놓은 규제체계를 번복하는 결과마저 초래될 수 있고, 산업과 시장에 커다란 혼란이 나타날 가능성도 있다. 따라서 우리나라 산업의 상황을 고려하여 국가 전략을 면밀하게 검토할 필요가 있다.

31) 미국과 EU간에 2021. 6. 15. 발족된 무역기술위원회(Trade and Technology Council)가 대표적인 예다(EC, "EU-US Launch Trade and Technology Council to Lead Values-Based Global Digital Transformation" (15 June 2021), https://ec.europa.eu/commission/presscorner/detail/en/IP_21_2990).

32) "Neither Party shall require the transfer of, or access to, source code of software owned by a person of the other Party, or the transfer of, or access to, an algorithm expressed in that source code, as a condition for the import, distribution, sale, or use of that software, or of products containing that software, in its territory."

*　　*　　*

이상 EU의 AI법안의 기본적인 구성과 개요를 검토하고, 법안의 발표와 EU의 움직임이 우리나라에 주는 몇 가지 주요한 시사점에 대해 살펴보았다. 국경이 없는 디지털 시장에서 제품만큼이나 규제간 경쟁도 치열해지고 있다. 우리 AI 정책과 법이 실증적인 검토에 입각하여 합리적인 방향으로 형성될 수 있도록 민간, 공공영역이 지혜를 모아 노력해야 할 것이다.