인공지능 제품안전과 소비자보호
— 인공지능기본법·인공지능책임법안 및 EU AI Act 가치사슬 책임 규정의 시사점을 중심으로 —

I. 들어가는 말

인공지능(Artificial Intelligence, AI)은 의료·금융·교통·교육 등 사회 전반에 걸쳐 급속히 확산되면서, 그로 인한 편익과 함께 새로운 위험을 동시에 낳고 있다. 특히 인공지능시스템이 의도하지 않은 방식으로 작동하거나 해킹 공격을 받을 경우, 에너지 공급 중단, 의료사고, 금융 피해 등 광범위한 손해가 발생할 수 있다.[1)]

그런데 인공지능이 야기하는 법적 문제는 종래의 기술 위험과는 구조적으로 다르다. 인공지능시스템은 데이터·알고리즘·클라우드·플랫폼·응용서비스 등이 결합된 다층적 구조로 개발·운영되며, 이 과정에서 모델 개발자, 데이터 제공자, 배치자, 운영자 등 복수의 행위자가 관여한다. 더욱이 인공지능시스템은 시장 출시 이후에도 업데이트·재학습·미세조정(fine-tuning)을 통해 성능과 위험 특성이 변동되어, 사고 발생 시 원인 규명과 책임 소재 특정이 극히 어려운 이른바 '법적 책임 공백'(legal responsibility gap)의 문제를 야기한다.[2)]

이에 대한 법적 대응으로 우리나라는 2025년 1월 21일 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하 '인공지능기본법')을 제정하여 2026년 1월 21일부터 시행하고 있다.[3)] 그러나 인공지능기본법은 인공지능사업자를 개발사업자와 이용사업자로 이분하는 단순 구도를 취하고 있어, 실제 인공지능 산업의 다층적 가치사슬 구조와 행위자 간 역할 전환 가능성을 충분히 반영하지 못하고 있다는 비판이 제기된다.

반면 EU 인공지능법(Artificial Intelligence Act, 이하 'EU AI Act')[4)]은 인공지능 생애주기와 가치사슬을 반영하여 제공자(provider)와 배치자(deployer)를 구분하고, 특히 제25조에서 '실질적 수정(substantial modification)'이나 '의도된 목적(intended purpose)의 변경'이 있을 경우 책임 주체가 전환되도록 규정함으로써 책임 공백을 방지하고 있다. 나아가 EU는 2026년 5월 7일 '디지털 AI 옴니버스(Digital Omnibus on AI)'에 관한 잠정 합의를 도출하여 AI Act 시행 이후 첫 번째 개정을 단행하였는바, 이는 집행의 현실화와 핵심 규범의 강

---

1) 이소은·서종희, "인공지능 관련 사고와 민사책임 - 불법행위책임을 중심으로", 소비자법연구 제12권 제1호, 2026, 100면; 과학기술정보통신부·한국지능정보사회진흥원, 고영향 인공지능 판단 가이드라인, 2026. 1. 29, 22면.

2) 정해빈, "인공지능과 책임공백에 관한 법경제학적 연구", 법경제학연구 제20권 제1호, 2023, 166-167면; Andreas Matthias, The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata, Ethics and Information Technology, Vol. 6 No. 3, 2004, pp. 175-183.

3) 과학기술정보통신부 보도자료, "AI G3의 기틀인 「인공지능기본법」 시행", 2026. 1. 21; 박상철, "인공지능 기본법의 시행 전 개정 필요성", 정보법학 제28권 제3호, 2024, 7면 참조.

4) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).

화를 동시에 추구하는 의미 있는 동향으로서 우리 법제에도 시사하는 바가 크다. 또한 국내에서는 제21대 국회에 안철수 의원안·황 희 의원안 등 다수의 인공지능책임법안이 제출되었으나[5], 제조물책임법과의 관계 및 소비자 친화적 입증책임 완화 규정 등에서 여전히 미흡한 점이 있다.

본고는 인공지능 제품안전과 소비자보호라는 통합적 시각에서 현행 법제의 문제점을 진단하고, EU AI Act의 가치사슬 책임 규정 및 2026년 디지털 AI 옴니버스 개정 동향이 한국 법제에 주는 시사점을 포함하여 입법적 개선 방향을 제시하는 것을 목적으로 한다.

이하에서는 인공지능기본법의 규제 체계와 그 한계를 검토하고(Ⅱ), 소비자보호 관점에서 민사책임 제도의 역할과 과제를 논하며(Ⅲ), EU AI Act의 가치사슬 책임 규정 및 최신 개정 동향을 검토하여 한국 법제에 대한 시사점을 도출하고(Ⅳ), 바람직한 법적 대응 방향을 제시하며(Ⅴ), 글을 마치고자 한다(Ⅵ).

Ⅱ. 인공지능기본법의 규제 체계와 한계

1. 인공지능기본법의 주요 내용

(1) 규제 대상의 설정

인공지능기본법은 인공지능시스템을 '다양한 수준의 자율성과 적응성을 가지고 주어진 목표를 위하여 실제 및 가상환경에 영향을 미치는 예측, 추천, 결정 등의 결과물을 추론하는 인공지능 기반 시스템'으로 정의한다(제2조 제2호). 핵심 개념인 '자율성'이란 인공지능시스템이 사람의 직접적 개입 없이 독립적으로 데이터를 학습하고 환경을 판단하여 운영될 수 있는 속성을 뜻하고[6], '적응성'이란 배포 전후에 입력되는 데이터와의 상호작용을 통해 스스로 행동 방식을 수정할 수 있는 기계학습 기반 시스템의 특성을 의미한다.

인공지능기본법은 고영향 인공지능(제2조 제4호), 생성형 인공지능(제2조 제5호), 고성능 인공지능(시행령 제24조 제1항)을 주된 규제 대상으로 설정하고 있다. 이 중 고영향 인공지능은 에너지·보건의료·의료기기·원자력·범죄 수사·채용·대출 심사·교통·공공서비스·학생 평가 등의 영역에서 생명·신체 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 인공지능시스템을, 고성능 인공지능은 학습에 사용된 누적 연산량이 10의 26승 부동소수점연산 이상인 시스템을 말한다.[7]

(2) 인공지능사업자의 의무와 책무

---

5) 박신욱, "인공지능이 결합된 소비제품과 사업자의 책임 – 인공지능책임법안을 중심으로", 소비자법연구 제10권 제1호, 2024, 38면.

6) 고영향 인공지능 판단 가이드라인, 12면.

7) 고성능 인공지능의 연산량 기준은 미국 바이든 행정부의 Executive Order 14110(2023. 10. 30.)이 제시한 수치를 참고한 것으로 보인다. 이소은·서종희, 앞의 글, 102면 각주 17 참조.

인공지능기본법은 인공지능사업자에게 투명성 확보 의무(제31조), 안전성 확보 의무(제32조), 고영향 인공지능의 안전성·신뢰성 확보 조치 책무(제34조)를 부과한다. 이러한 규제 체계는 EU AI Act의 위험 기반(risk-based) 접근방식을 따른 것으로 평가된다.[8] 다만 이들 의무는 주로 인공지능사업자에 대한 것으로서, 이용자의 의무나 영향받는 자의 권리에 관한 실질적 규정은 매우 부족한 상황이다.

2. 인공지능기본법의 소비자보호 측면에서의 한계

(1) 인공지능 가치사슬을 반영하지 못한 행위자 정의

인공지능기본법의 가장 근본적인 한계 중 하나는, 인공지능사업자를 '인공지능개발사업자'와 '인공지능이용사업자'의 이분법적 구조로 단순화하여 규정한다는 점이다(제2조 제7호). 그러나 실제 인공지능 산업에서는 데이터 제공자, 모델 개발자, 미세조정자(fine-tuner), 통합·배치자, 서비스 운영자 등 다층적 행위자가 가치사슬을 구성한다. 예컨대 Microsoft가 OpenAI의 GPT 모델을 통합하여 Copilot 시스템으로 기업에 제공하는 경우, 가치사슬에는 세 개 이상의 주체가 관여하는데 현행 인공지능기본법의 이분법적 구도는 이를 반영하기에 역부족이다.

또한 '인공지능개발사업자'를 '인공지능을 개발하여 제공하는 자'로 한정하여 정의함으로써, 제3자에게 개발을 위탁하여 제공하는 자가 포함되는지 여부가 불명확하다. 책임법제와의 정합성 측면에서 '개발'이라는 기술적 행위 자체보다는 '거래 시장에의 제공'을 중심으로 규제 대상을 포섭하는 것이 타당하다.[9] 나아가 '인공지능사업자'의 '사업자'라는 용어는 영리적 경제활동을 영위하는 자라는 추상적 범주를 전제하나, 고영향 인공지능의 경우 에너지 공급·먹는물 생산·원자력시설 관리 등 공공기관에 의해 운용되는 경우가 많아 '사업자' 개념이 적합하지 않다는 지적도 있다.[10]

아울러 인공지능기본법상 '이용자'는 '인공지능제품 또는 인공지능서비스를 제공받는 자'로 정의되어(제2조 제8호) 최종적·수동적 수령자를 지칭하고 있으나, 이는 인공지능 특유의 운용적 성격과 위험 관리 구조를 반영하는 인공지능 맥락에서 '이용'과 혼용되어 혼란을 야기한다. EU AI Act가 배치자(deployer)를 독립적 행위자로 설정하고 투명성·위험관리·기록보존 등 규제 의무를 부과하는 것과 대조적이다.[11]

(2) 이용자에 대한 규제 부재와 영향받는 자 보호 미흡

---

8) Amedeo Rizzo & Giorgio Hassan, On the Distinction Between Tort-Based and Risk-Based AI Regulation: A Transatlantic Perspective, Stanford-Vienna TTLF Working Paper No. 131, 2025, at 23.
9) 박상철, 앞의 글(주 3), 33면.
10) 박상철, 앞의 글(주 3), 33면.
11) EU AI Act Article 3(4).

인공지능기본법은 인공지능사업자에게만 의무·책무를 부과하고 이용자에 대해서는 어떠한 규제도 두고 있지 않다. 그러나 이용자와 영향받는 자의 관계를 고려하면 이용자에게도 일정한 책임이 부과될 필요가 있다. 예컨대 학생 평가용 인공지능시스템을 도입한 고등학교(이용자)는 이용 지침 준수, 인간에 의한 관리·감독 주체 지정, 시스템 운영 모니터링 등의 의무를 부담해야 하나, 현행법에는 이러한 규정이 전혀 없다.[12)]

'영향받는 자'(제2조 제9호)에 대한 보호 역시 제3조의 기본원칙에서 설명을 제공받을 권리를 추상적으로 언급하는 데 그치고 있다. 이는 EU AI Act 제26조 제11항이 배포자에게 자연인과 관련된 의사결정 시 해당 자연인에게 인공지능시스템 이용 대상임을 고지할 의무를 부과하는 것과 현저히 대조된다.

(3) '진흥법'으로서의 성격과 규제법으로서의 한계

인공지능기본법은 '진흥법'적 성격이 강하여[13)] 형사처벌 규정 부재, 인공지능 활용 금지 사례 미규정, 의무 위반에 대한 제재가 3,000만 원 이하 과태료(제43조 제1항)에 그치는 등 실질적 준수 유인이 부족하다. EU AI Act가 위반 시 최대 3,500만 유로 또는 전 세계 매출의 7%에 달하는 제재를 규정한 것과 현격한 대비를 이룬다. 인공지능기본법상 규제 공백은 민사책임 제도, 특히 불법행위책임이 보완적으로 더욱 적극적인 역할을 수행하여야 함을 시사한다.

(4) 규제 기술(RegTech)·리걸테크로서 AI의 보완적 역할과 그 한계

인공지능기본법의 집행력 부재를 보완하는 또 하나의 접근으로 AI 기술 자체를 소비자보호 규제 수단으로 활용하는 규제 기술(RegTech)·리걸테크(LegalTech) 방안이 주목받고 있다. 인공지능기본법상 고영향 인공지능 사업자에게 부과되는 안전성·투명성 확보 의무(제31조·제32조·제34조)는 공정거래위원회 등 규제 기관이 한정된 인력으로 방대한 AI 시장 전반을 모니터링하기에 본질적 한계가 있다. 이 지점에서 AI 기술은 집행 비용을 획기적으로 낮추는 규제 기술 수단으로 기능할 수 있다.[14)]

구체적으로 AI는 소비자보호법제 영역에서 다음과 같이 활용될 수 있다. 첫째, 약관법 영역에서 자연어 처리(NLP)와 머신러닝을 활용한 불공정약관 조항의 자동 심사 및 실시간 모니터링이 가능하다. 둘째, 다크패턴(Dark Patterns) 탐지 영역에서는 기계학습 기반 유형 분류, 컴퓨터 비전을 통한 시각적 기만 탐지, 웹 크롤링을 통한 상시 감시 시스템 구축이 가

12) 이소은·서종희, 앞의 글, 108면. EU 인공지능법 제26조(고위험 인공지능 시스템 배포자의 의무) 제1항, 제2항, 제5항 참조.

13) 과학기술정보통신부 보도자료 참조.

14) 김세준, "현행 소비자보호법제에서 인공지능 기술 적용 가능성과 한계", 소비자법연구 제11권 제4호, 2025, 12-13면. AI 기반 불공정약관 심사 시스템 구축 사업의 구체화에 관해서는 공정위 보도자료, "공정위, 불공정 하도급 계약 예방 위해 AI 플랫폼 구축 추진", 2025. 10. 3. 참조.

능하다. 소비자의 인지 편향을 의도적으로 이용하는 다크패턴은 전통적 규율체계로는 효과적으로 대응하기 어려운 영역으로서, 2024년 전자상거래법 개정으로 제21조의2가 신설된 이후에도 AI 기반 모니터링의 필요성이 강조되고 있다.[15] 셋째, 소비자위해감시시스템(CISS)에 AI와 빅데이터 분석 기술을 도입하여 SNS·블로그·온라인 리뷰 등 비정형 데이터로부터 제품 결함 징후를 선제적으로 포착하는 조기 경보 시스템 구축이 가능하다. 넷째, AI 기반 소비자 분쟁해결 지원—지능형 분쟁 접수, 법률 분석 AI를 통한 협상 보조, 자동화된 조정 조서 작성—은 소액·다수 피해라는 소비자 분쟁의 특성을 고려할 때 온라인 분쟁해결(ODR) 절차의 실효성을 획기적으로 제고할 수 있다.

그러나 이러한 AI의 규제 기술적 활용은 고유한 법제도적 한계를 내포한다. 첫째, 데이터 편향성 문제이다. AI 모델이 학습하는 과거의 소비자 불만 데이터·피해구제 사례는 보고 편향(Reporting Bias)을 내재하고 있어, 실제 위해의 심각성이 아닌 '신고가 활발한 집단'의 문제에 과도하게 반응하거나, 피드백 루프를 통해 편향이 고착화될 위험이 있다.[16] 둘째, 블랙박스 문제와 절차적 정당성 침해이다. 딥러닝 기반 AI의 의사결정 과정은 인간이 이해하기 어려운 블랙박스 특성을 지니는바, 이는 행정절차법 제23조 제1항의 이유제시의무와 충돌하고 피규제자인 사업자의 방어권을 실질적으로 침해할 위험이 있다.[17] 셋째, 사업자의 지능적 규제 우회이다. 사업자 역시 AI를 활용하여 규제 AI의 취약점을 역공학적으로 파악하거나, 개별 소비자의 심리적 취약성을 분석한 '초개인화된 기만'을 시도할 경우 일반적인 웹 크롤링 기반 규제 AI는 이를 포착하기 어렵다.

따라서 AI 기술을 소비자보호 규제 수단으로 도입하는 경우에는, EU AI Act 제13조·제14조가 고위험 AI에 대해 설명 가능성 확보와 인간 감독(Human Oversight) 의무를 부과하는 것처럼, 설명가능한 인공지능(XAI) 기술 적용 의무화, 알고리즘 영향평가 제도, 정기적 알고리즘 감사(Audit) 등 절차적 정당성 확보 장치를 법적 근거와 함께 마련하여야 한다. 이는 인공지능기본법이 규정하는 '인간에 의한 관리·감독' 원칙(제3조)과도 정합성을 갖는 방향이다.

Ⅲ. 소비자보호 관점에서의 인공지능 민사책임 제도

1. 인공지능 관련 사고의 특수성

인공지능 관련 사고는 일반적인 제조물 사고와 구별되는 세 가지 특수성을 지닌다. 첫째, 인공지능시스템의 오류나 오작동을 인간이 예측하기 어려워 '예견 가능성'을 요소로 하는 과

15) 김세준, 앞의 글, 14-17면; 정신동, "온라인상의 다크패턴과 소비자보호 - 기본규범으로서 전자상거래법 제21조 제1항 제1호의 기능과 한계", 법학논총 제29집 제3호, 조선대학교 법학연구원, 2022, 56-57면.

16) 김세준, 앞의 글, 26-27면.

17) 김세준, 앞의 글, 27-28면; 권은정, "자동적 처분의 행정법적 제 문제 - 처분의 형식·절차 및 구제에 관한 쟁점을 중심으로 -", 법학논총 제38집 제3호, 2021, 100면 이하; 김찬희, "인공지능 기반 자동화 행정행위와 민주적 법치국가 원리", 행정법연구 제72호, 2023, 54면 이하.

실 개념의 적용이 쉽지 않다.[18]

둘째, 인공지능시스템의 자율성으로 인해 인간이 그 기능을 전부 통제하기 어렵다. 고성능 인공지능의 경우 '창발성(emergent abilities)'이나 개발자가 의도한 것처럼 행동하는 척하면서 실제로는 다른 목표를 추구하는 '책략(scheming)'을 보이기도 하며[19], 이는 통제 기제의 작동 실패로 이어질 수 있다.

셋째, 인공지능시스템의 불투명성(opaqueness)으로 인해 사고 발생 시 누구에게 어느 정도의 과실이 있는지, 과실과 사고 사이에 인과관계가 있는지를 증명하는 것이 극히 어렵다.[20] 더욱이 인공지능 가치사슬에서는 개발사업자·이용사업자·이용자 등 다수의 행위자가 관여하므로, 다수의 원인(many things)과 다수의 손(many hands)[21] 문제가 중첩되어 책임 소재 판명이 더욱 복잡해진다.

이러한 인공지능 사고의 특수성은 합리적이고 평균적인 소비자를 전제로 설계된 현행 소비자보호법제의 한계와 직접 연결된다. 특히 인공지능시스템이 소비자의 심리적 취약성을 실시간으로 분석하여 개인화된 기만을 시도하는 경우—예컨대 다크패턴(Dark Patterns)과 AI 광고 기술의 결합—전통적 규율 체계만으로는 대응에 한계가 있으며, 인지 비대칭 상황에서의 소비자 보호를 위한 새로운 법적 수단이 요구된다.[22]

2. 과실책임과 위험책임의 이분구조

(1) 위험 수준에 따른 책임 분화의 필요성

인공지능기술의 위험성은 시스템의 종류와 활용 영역에 따라 현저히 달라, 모든 인공지능 사고에 동일한 책임 법리를 적용하기보다는 위험의 정도에 따라 서로 다른 책임 체계를 적용하는 이원적 구조가 바람직하다. 인공지능기본법이 설정한 고영향·고성능 인공지능과 그 외의 인공지능의 구별이 민사책임 제도 설계의 중요한 기준선이 될 수 있다.

(2) 고영향·고성능 인공지능에 대한 위험책임(무과실책임)

고영향·고성능 인공지능의 경우 인공지능개발사업자에게 위험책임(무과실책임)을 부담시키

---

18) Miriam Buiten, Alexandre de Streel & Martin Peitz, The Law and Economics of AI Liability, 48 Comput. L. & Sec. Rev. 105794, 1 (2023); Jack M. Balkin, The Path of Robotics Law, 6 Calif. L. Rev. Circuit 45, 51 (2015).

19) 이소은·서종희, 앞의 글, 105-106면.

20) 이소은·서종희, 앞의 글, 113면; Buiten, at 6.

21) Andreas Matthias, pp. 175-183 참조.

22) 김세준, 앞의 글, 10-11면. 현행 소비자보호법제가 전제하는 '합리적이고 평균적인 소비자'라는 고전경제학의 인간상은 행동경제학이 제시하는 제한된 합리성(Bounded Rationality)과 인지 편향(Cognitive Bias)의 관점에서 근본적 재검토가 필요하며, 특히 인공지능 제품·서비스 환경에서는 정보비대칭이 인지 비대칭으로 심화된다는 점에서 더욱 그러하다. 서종희, "행동경제학적 관점에서의 소비자법의 개선방향 - 유럽소비자법 등을 중심으로 -", 소비자법연구 제7권 제4호, 2021, 168-169면 참조.

는 것이 타당하다. 이에 대한 근거는 두 가지이다. 첫째, 효율적 위험 분산이다. 인공지능개발사업자는 책임보험 가입 등을 통해 적은 비용으로 위험을 분산시킬 수 있고, 그 비용은 인공지능제품·서비스의 가격에 반영되어 이용자 집단 전체에 분산될 수 있다.[23]
둘째, 이익·위험의 귀속 일치이다. 상당한 위험이 수반되는 활동으로부터 이익을 얻는 자가 그 위험이 현실화되어 발생한 손해도 배상하는 것이 형평의 원칙에 부합한다.[24] 다만 위험책임 도입이 기술 혁신을 저해하지 않도록 책임보험 가입 의무화와 공적 보상기금 제도의 도입을 병행하여야 한다. 이와 관련하여 뉴질랜드 사고보상공사(ACC) 제도와 독일의 자동차 사고 피해 보상기금이 참고 모델로 제시된다.[25]

(3) 그 외 인공지능에 대한 과실책임

고영향·고성능 인공지능에 해당하지 않는 인공지능시스템의 경우에는 과실책임을 적용함이 타당하다. 과실의 판단 기준은 관련 법령의 의무 규정을 일응의 기준으로 삼되[26], 인공지능 기술이 활용되는 영역별로 특례 조항을 두어 사고 원인을 규명하는 방안도 고려할 만하다. EU AI Act가 제공자에게 기술문서와 기록 유지를 의무화한 것은 바로 이 입증 곤란성 문제를 제도적으로 해결하기 위한 것으로 이해된다.[27]

IV. EU AI Act의 가치사슬 책임 규정과 한국 법제에 대한 시사점

1. EU AI Act의 가치사슬 책임 구조

(1) 제공자와 배치자의 구분

EU AI Act는 제공자(provider)를 '인공지능시스템이나 범용인공지능모델을 개발하여, 혹은 개발된 시스템 또는 범용인공지능모델을 시장에 출시하거나 자신의 이름이나 상표로 서비스로 제공한 자'로 정의한다(제3조 제3항). 단순히 연구 목적으로 개발만 하고 외부에 유통시키지 않은 경우는 '시장 출시 행위'가 없으므로 제공자에 해당하지 않는다.[28] 즉, 개발행위와 시장 출시 행위가 결합되어야 제공자 지위가 발생한다.
배치자(deployer)는 '개인적·비전문적 활동에 사용되는 경우를 제외하고, 자신의 권한 하에 인공지능시스템을 사용하는 자연인 또는 법인, 공공기관, 기관 또는 기타 단체'로 정의된다(제3조 제4항). EU AI Act는 배치자를 단순히 수동적 소비자가 아닌 독립적 법적 행위자

23) 편집대표 김용덕, 주석민법 채권각칙 6권 제5판, 한국사법행정학회, 2022, 66-67면(박동진 집필부분).
24) 편집대표 김용덕, 앞의 책, 67면(박동진 집필부분); 이소은·서종희, 앞의 글, 115면.
25) 뉴질랜드 Accident Compensation Act 2001 제20조 제1항; Simon Connell, Community Insurance versus Compulsory Insurance, 39 Legal Stud. 499 (2019).
26) 대법원 2001. 1. 19. 선고 2000다12532 판결; 이소은·서종희, 앞의 글, 119면.
27) EU AI Act Article 11, 12.
28) EU AI Act Article 3(3).

로 설정하여, 투명성·위험관리·기록 보존·모니터링 및 보고 의무를 부과한다.[29]

(2) 인공지능 가치사슬의 특징과 EU AI Act의 대응

인공지능 가치사슬은 전통적 제조업과 달리 선형적·고정적이지 않고 순환적·동태적 성격을 지닌다. 또한 데이터·알고리즘이라는 비물질적 자원의 비경합성·비배재성으로 인해 가치망은 네트워크형을 띠게 된다.[30]

EU AI Act는 이러한 인공지능 고유의 산업 구조를 반영하여, 제공자·수입자·유통자·배치자 등 가치사슬 상 전 주체를 운영자(operator)로 포괄하고[31], 특히 전문 88항과 제25조의 규정을 통하여 다수의 행위자가 상호 얽히는 가치사슬 구조를 전제로 책임을 분배하고 있다.

(3) 제25조의 제공자 지위 전환 규정

EU AI Act 제25조는 고위험 인공지능시스템에 관하여 본래 제공자가 아닌 주체라도 일정한 조건 하에서 '제공자'로 간주되어 제16조의 제공자 의무를 부담하도록 규정한다. 제공자 지위 전환의 세 가지 요건은 다음과 같다.

첫째, 자신의 이름이나 상표를 표시한 경우이다(제25조 제1항 (a)). 이미 출시된 고위험 인공지능시스템에 자신의 이름이나 상표를 부착한 자를 제공자로 간주하는 것으로, 구 EU 제조물책임지침 제3조(1)에서 상표를 부착한 자를 제조자로 보는 법리[32]를 인공지능 맥락에 적용한 것이다.

둘째, 실질적 수정(substantial modification)이 가해진 경우이다(제25조 제1항 (b)). '실질적 수정'이란 공급자가 수행한 초기 적합성 평가에서 예상하거나 계획하지 않은, 시장 출시 또는 사용 개시 후 이루어진 인공지능시스템의 변경으로서, 그 결과로 고위험 인공지능 요건에 따른 해당 시스템의 적합성이 영향을 받거나 의도된 목적이 변경되는 경우를 말한다(제3조 제23항).[33] 이는 미세조정(fine-tuning) 등 추가 학습 행위가 실질적 수정에 해당할 수 있음을 시사한다.

셋째, 의도된 목적(intended purpose)의 변경이다(제25조 제1항 (c)). 유통업자·수입업자·배치자 또는 제3자가 고위험 인공지능시스템이 아닌 인공지능시스템 또는 범용인공지능모델의 의도된 목적을 변경하여 해당 시스템이 처음으로 제6조 요건에 따른 고위험 인공지능시스템에 해당하게 되는 경우, 변경을 가한 자가 새로운 제공자로 간주된다.[34]

한편, 제공자 지위가 전환된 경우에도 최초 제공자는 새로운 제공자가 적합성 평가 등 의무

---

29) EU AI Act Article 3(4), Article 26.
30) 한국지능정보사회진흥원, "글로벌 AI 가치사슬 분석을 통한 AI 경쟁력 강화 제언", The AI Report, 2024, 9면.
31) EU AI Act Article 3(8), Article 2.1.
32) EU Directive 85/374 Article 3(1); CJEU Case C-264/21 (Keskinäinen Vakuutusyhtiö Fennia v Koninklijke Philips NV).
33) EU AI Act Article 25.1(b), Article 3(23).
34) EU AI Act Article 25.1(c), Article 3(12).

이행에 필요한 기술적 접근 및 지원을 받을 수 있도록 협력해야 한다(제25조 제2항). 다만 최초 제공자가 사전에 고위험으로 변경하여서는 아니 된다고 명확히 특정한 경우 이 협력 의무가 면제될 수 있다.[35]

(4) 범용인공지능모델에 대한 특별 규율

EU AI Act는 범용인공지능모델(General Purpose AI Model) 제공자에게 투명성 의무로서 ① 훈련·평가 결과를 포함한 기술문서의 작성·유지·제공 의무, ② 인공지능시스템 제공자에 대한 모델의 기능과 한계 관련 정보 제공 의무, ③ 훈련에 사용된 콘텐츠에 대한 공개 의무를 부과한다.[36] 범용AI 모델 실천 강령(Code of Practice)은 2025년 9월 최종 확정되어 현재 적용 중이다.

2. EU AI Act 최신 동향: 2026년 디지털 AI 옴니버스(Digital Omnibus on AI)

(1) 디지털 AI 옴니버스의 배경과 의의

EU AI Act는 2024년 6월 채택 이후 단계적 시행을 거쳐왔으나, 각 회원국의 국가감독기관 지정 지연, 조화 표준 미확정, 기술 지원 도구 부재 등 현실적 집행 여건 미비로 인해 업계의 준비 부담이 가중된다는 지적이 지속되었다. 이에 EU 집행위원회는 2025년 11월 19일 '디지털 AI 옴니버스'를 제안하였고, 2026년 5월 7일 EU 이사회·유럽의회·집행위원회 삼자 간 잠정 합의가 성립되어 AI Act 시행 이후 최초의 체계적 개정이 이루어지게 되었다.[37]
이번 개정은 전면적 재검토가 아니라 집행 현실화(실행 기한 조정)와 핵심 규범 강화(딥페이크 금지 확대)를 동시에 추구하는 것이 특징이다. 아래 [표 2]는 Digital Omnibus에 따른 EU AI Act 주요 이행 기한을 정리한 것이다.

시행일	적용 대상	비고 (Digital Omnibus 반영)
2025. 2. 2.	AI 활용 금지 조항 (소셜 스코어링, 서브리미널 조작 등)	원안 그대로 시행
2025. 8. 2.	범용AI 모델(GPAI) 의무 거버넌스 조항	원안 그대로 시행
2026. 12. 2. (원안: 2026. 8. 2.)	AI 생성 친밀한 콘텐츠 딥페이크 금지 조항 추가 워터마킹 등 투명성 의무	신규 금지: 딥페이크 성적 콘텐츠 투명성 기한 단축
2027. 12. 2. (원안: 2026. 8. 2.)	부속서 III 고위험 AI 시스템 (에너지, 의료, 고용, 교육 등)	Digital Omnibus로 약 16개월 연장

35) EU AI Act Article 25.2.
36) EU AI Act Article 53(1) a, b, d.
37) European Commission / Council of the EU / European Parliament, Provisional Agreement on Digital Omnibus on AI (May 7, 2026); Inside Privacy, "EU AI Act Update: Timeline Relief, Targeted Simplification, and New Prohibitions", May 18, 2026.

2028. 8. 2. (원안: 2027. 8. 2.)	부속서 I 고위험 AI 시스템 (제품 내장 안전 부품 등)	Digital Omnibus로 약 12개월 연장

(2) 고위험 AI 시스템 의무 이행 기한 연장

Digital Omnibus의 핵심 변경은 고위험 AI 시스템 의무 이행 기한의 단계적 연장이다. 부속서 III에 해당하는 용도 기반 고위험 AI 시스템의 의무 준수 기한은 당초 2026년 8월 2일에서 2027년 12월 2일로 약 16개월 연장되었다. 또한 부속서 I에 해당하는 제품 내장형 고위험 AI 시스템에 대해서는 2028년 8월 2일까지 추가 연장이 이루어졌다.

(3) 딥페이크 성적 콘텐츠 금지 규정 신설

Digital Omnibus는 2026년 12월 2일부터 식별 가능한 자연인의 친밀한 신체 부위나 성적 노골적 행위를 현실적으로 묘사·조작하는 AI 시스템, 즉 이른바 '누디파이어(nudifier)' 애플리케이션의 시장 출시 및 배치를 금지한다. 위반 시 최대 3,500만 유로 또는 전 세계 매출의 7% 제재가 부과될 수 있다.

(4) AI 리터러시 의무의 조정

Digital Omnibus는 AI 리터러시 의무의 부담 방식도 조정하였다. 종전에는 제공자·배치자가 임직원의 AI 리터러시를 '보장(ensure)'하도록 규정하고 있었으나, 개정 합의에서는 AI 리터러시 '발전을 지원하는 조치를 취할(take measures to support the development of AI literacy)' 의무로 완화되었다.

(5) 한국 법제에 대한 추가적 시사점

디지털 AI 옴니버스의 동향은 우리 인공지능기본법 운용에 세 가지 시사점을 제공한다. 첫째, 규범의 단계적 시행과 현실적 집행 여건 마련의 중요성이다. 둘째, 딥페이크 등 신종 위험에 대한 선제적 규범화이다. EU가 딥페이크 성적 콘텐츠 금지를 AI Act 개정으로 신속히 대응한 것처럼, 우리 법제도 생성형 AI의 특수한 위험에 대해 명문 금지 조항을 마련할 필요가 있다. 셋째, 중소기업·스타트업 배려 규정의 도입이다.

3. EU AI Act 가치사슬 책임 규정이 한국 법제에 주는 시사점

(1) 행위 유형 중심의 행위자 정의 필요

EU AI Act가 제공자와 배치자를 구체적 행위를 기준으로 정의하는 것과 달리, 인공지능기본법은 추상적 범주를 중심으로 사업자를 정의하고 이를 다시 개발사업자와 이용사업자로 이분한다. 인공지능의 위험을 관리·예방하고 피해 발생 시 책임 귀속을 명확히 하기 위해서는, 행위자가 개발·시장 출시·유통·배치·운영·재학습 등 어떠한 구체적 행위를 수행하였는지를 기준으로 법적 지위를 정의하여야 한다.

(2) 실질적 수정·의도된 목적 변경에 따른 책임 전환 규정의 도입

EU AI Act 제25조와 같이, 인공지능이용사업자가 고위험 인공지능시스템에 실질적 수정을 가하거나 의도된 목적을 변경하는 경우 개발사업자를 대체하여 새로운 법적 책임 주체가 되도록 하는 '제공자 지위 전환' 규정을 도입하는 것이 필요하다.

(3) 글로벌 공급망에 대한 역외 적용 규정 정비

인공지능기본법 제36조는 국내에 주소 또는 영업소가 없는 인공지능사업자에 대해 국내대리인 지정을 규정하고 있으나, 이는 단순한 행정적 집행력 확보에 그친다.[38] EU AI Act는 EU 역내에서 AI 시스템을 제공하거나 서비스를 이용 가능하게 하는 경우, 제공자가 제3국에 소재하더라도 EU 소재 수입자나 EU 역내 대리인에게 제공자 의무를 부과하는 방식으로 역외 적용을 실효적으로 달성한다. 한국도 이를 실질화할 필요가 있다.

(4) 범용인공지능모델에 대한 별도 규율 체계 마련

ChatGPT, Gemini, Claude 등 범용인공지능모델은 다양한 산업에 통합되어 활용되고 있으나, 인공지능기본법은 이에 대한 별도의 규율 체계를 마련하고 있지 않다. EU AI Act가 범용인공지능모델 제공자에 대해 별도 의무를 부과하고 2025년 9월 범용AI 모델 실천 강령을 최종 확정한 것처럼[39], 우리 법제에서도 범용인공지능모델의 특수성을 반영한 별도 규율을 도입할 필요가 있다.

4. 인공지능에 의한 가격차별화에 대한 재고

가격차별이 성립하려면 ① 시장지배력, ② 소비자별 수요 차이 식별, ③ 차익거래 불가라는 세 조건이 필요하다. 전통적으로 이 중 ②가 가장 큰 현실적 장벽이었는데, AI가 이 장벽을 사실상 허물어 버린다.

사회후생(social welfare)은 특정 거래를 통해 발생하는 경제적 이익의 총합이다. 일반적으

38) 인공지능기본법 제36조.

39) EU AI Act Article 53(1) a-d; 가치사슬 논문, 10-11면; EU AI Office, Code of Practice for General-Purpose AI Models (최종본 2025. 9.).

로 사회후생은 소비자잉여와 생산자잉여의 합으로 정의된다.[40]

[SW(사회후생) = CS(소비자잉여) + PS(생산자잉여)]

소비자잉여(consumer surplus)란 소비자가 어떤 재화나 서비스에 대해 실제로 지불한 가격보다 더 높은 가격을 지불할 의사가 있었을 때 그에게 발생하는 이익이다.[41] 한편, 생산자잉여(producer surplus)란 생산자가 어떤 재화나 서비스를 공급하기 위해 최소한 받아야 한다고 생각한 금액보다 실제로 더 높은 가격을 받을 때 그에게 발생하는 이익이다.[42] 완전경쟁시장에서는 소비자잉여와 생산자잉여의 합도 최대화된다(사회후생의 극대화). 반대로 독과점시장에서는 경쟁이 제한되어 거래량이 감소하거나 가격이 왜곡되어 일부 상호이익적 거래가 실현되지 못한다. 그 결과 완전경쟁시장이었다면 확보할 수 있었던 사회후생을 잃게 된다. 이를 사중손실(deadweight loss)이라 한다(그림 1-5 참조).[43]

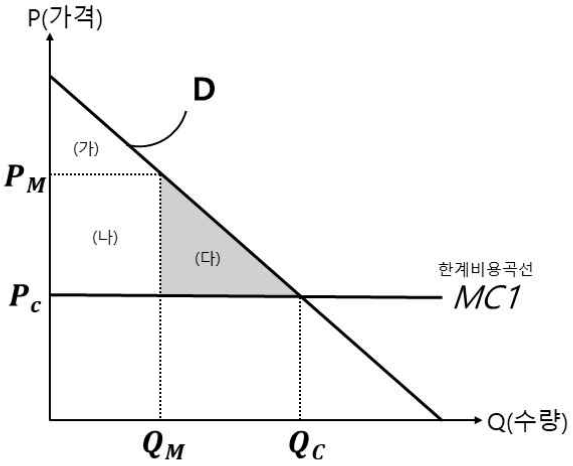

그림 1-5: 독점시장의 사회적 후생

(1) 전통적 제3종 → AI형 제1종으로의 진화

가격차별(price discrimination)이란 동일한 재화 또는 서비스를 서로 다른 소비자에게 서로 다른 가격으로 판매하는 행위이다. 이는 독과점시장의 기업이 소비자잉여의 일부를 생산자잉여로 이전시켜 이윤을 증가시키는 일종의 전략으로 볼 수 있다[44]. 일반적으로 가격차별

40) 서승환, 『미시경제학』, 제4판, 법문사, 2011, 316-317면.
41) 따라서 예컨대 어떤 소비자가 어떤 상품에 관하여 최대 10만 원까지 지불할 의사가 있었는데, 그가 실제로 지불한 시장가격이 7만 원이라면, 그 소비자는 3만 원의 소비자잉여를 얻는다.
42) 예컨대 생산자가 어떤 상품을 5만 원 이상이면 판매할 의사가 있었는데 실제 시장가격이 7만 원이라면, 생산자는 2만 원의 생산자잉여를 얻는다.
43) 서승환, 앞의 책, 347-348면. 그림 1-5 설명> 완전경쟁시장의 사회 후생은 (가), (나), (다) 영역의 합이나, 독점시장의 사회 후생은 (가), (나) 영역의 합이다. (다) 영역이 사중손실에 해당한다.

은 가격 설정 방식과 소비자 정보의 정도에 따라 아래와 같이 제1종, 제2종, 제3종 가격차별로 구분된다.

제1종 가격차별(first-degree price discrimination)은 기업이 각 소비자의 최대 지불의사를 완전히 파악하여 소비자마다 서로 다른 가격을 부과하는 형태이다. 이 경우 기업은 소비자잉여를 모두 흡수하며, 따라서 생산자잉여가 극대화된다.[45]

제2종 가격차별(second-degree price discrimination)은 소비자의 유형을 직접 구별하지 않고 구매수량이나 선택한 메뉴에 따라 서로 다른 가격을 적용하는 방식이다. 대표적인 예로는 대량구매 할인, 정액제와 종량제의 결합, 이부가격제(two-part tariff), 멤버십 요금제 등이 있다. 이 경우 소비자는 자신의 소비 성향에 따라 스스로 가격체계를 선택한다.[46]

제3종 가격차별(third-degree price discrimination)은 기업이 소비자 집단의 특성을 구분하여 서로 다른 가격을 적용하는 방식이다. 학생 할인, 경로우대 할인, 지역별 가격 차이, 시간대별 요금 차별 등이 대표적 사례이다. 일반적으로 가격탄력성이 낮은 집단에는 높은 가격이, 가격탄력성이 높은 집단에는 낮은 가격이 부과된다.[47]

즉, 기존 제3종 가격차별은 학생·노인 등 집단 단위로만 가능했다. 그러나 AI는 개별 소비자의 구매이력, 검색 패턴, 체류 시간, 위치 정보, 심지어 사용 기기(아이폰 vs 안드로이드)까지 실시간 분석하여 사실상 소비자별 최대 지불의사를 추정한다. 이는 제1종 가격차별에 근접한 형태로, 소비자잉여가 거의 전부 생산자잉여로 전환될 수 있다.

(2) Damaged Quality와 AI 개인화 광고의 결합

'Damaged Quality' 개념[48]과 AI가 결합하면 더 정교한 착취가 가능하다. AI는 특정 소비자가 광고형 저가 요금제에 만족할 확률이 높다고 판단하면 처음부터 그 요금제만 노출시켜, 소비자가 상위 요금제의 존재 자체를 인식하지 못하게 할 수 있다. 메뉴형 이부가격제에서 소비자의 자발적 선택을 전제로 유인합치 제약이 설계되지만, AI 개인화 광고는 애초에 선택지 자체를 왜곡한다.

(3) 다크패턴 + AI 가격차별

'초개인화된 기만' 문제가 정확히 여기서 발현된다. 항공권 예약 사이트가 재방문 소비자에

---

44) 가격차별이 성립하기 위해서는 일반적으로 다음과 같은 조건이 필요하다. ① 기업이 일정한 시장지배력 또는 가격결정력을 갖고 있어야 한다. 따라서 개별 기업이 가격을 결정할 수 없는 완전경쟁시장에서는 가격차별이 어렵다. ② 기업이 소비자별 수요 차이를 식별할 수 있어야 한다. ③ 낮은 가격으로 구매한 소비자가 높은 가격 시장에 재판매하는 차익거래가 불가해야 한다.

45) 서승환, 앞의 책, 342-244면

46) 서승환, 앞의 책, 342면

47) 서승환, 앞의 책, 336-340면

48) Damaged Quality란, 기업이 의도적으로 낮은 품질의 상품이나 서비스를 만들어 저가 상품으로 제공하는 전략을 말한다. 이 역시 일상에서 자주 접할 수 있는 제2종 가격차별의 중요한 사례이다. 여기서 핵심은 높은 품질에 기꺼이 비싼 가격을 지불할 의사가 있는 소비자가 비교적 품질이 낮은 저가 상품을 선택하는 것을 방지하기 위해 기업이 일부 기능이나 품질을 의도적으로 제한하는 데 있다.

게 가격을 올려 표시하거나, 동일 좌석을 기기·위치에 따라 다른 가격으로 제시하는 행위는 다크패턴과 AI 가격차별의 교집합이다. 전자상거래법 제21조의2가 다크패턴을 명시적으로 규율하지만, 가격 자체의 개인화는 현행법의 사각지대에 있다.

(4) 사회후생 측면에서의 평가

제3종 가격차별은 총생산량 증가 여부에 따라 사회후생 효과가 달라진다. AI 가격차별도 마찬가지인데, 문제는 AI가 총생산량 증가 없이 소비자잉여만 흡수하는 방향으로 최적화될 가능성이 높다는 점이다. 특히 AI는 기업의 이윤 극대화 목표로 훈련되므로, 사회후생보다 생산자잉여 극대화에 편향될 구조적 유인이 있다.

(5) 입법적 공백

현재 인공지능기본법에는 AI를 이용한 가격차별에 관한 직접적 규율이 없고, 표시광고법은 가격의 개인화 자체보다 '허위·과장'에 초점을 맞추고 있어 적용에 한계가 있다. EU AI Act 부속서 III는 신용평가·고용 분야의 AI를 고위험으로 분류하지만 소매 가격 설정 AI는 포함되지 않는다.
결국 이 문제는 앞서 보완한 논문의 VI장 입법 제언에 "AI 기반 동적 가격차별(dynamic pricing)에 대한 투명성 의무 및 소비자 고지 규정 신설" 항목을 추가하는 방향으로 연결할 수 있다.

(1) 가격 알고리즘 투명성 의무

가장 핵심적인 규제 수단은 AI 가격결정 알고리즘의 투명성 확보이다. 블랙박스 문제의 해법으로 제시한 설명가능한 AI(XAI) 의무와 같은 맥락이다. 구체적으로 사업자에게 가격 결정에 사용된 변수의 종류와 가중치, 동일 상품에 대해 소비자별로 다른 가격이 제시될 수 있다는 사실 자체, 소비자 본인에게 적용된 가격 결정 요인에 대한 개인별 설명을 제공하도록 의무화하는 방향을 생각해 볼 수 있다. EU AI Act 제13조가 고위험 AI에 대해 사용자에게 투명한 정보 제공을 의무화한 것처럼, 소비자 가격에 직접 영향을 미치는 AI 시스템을 고위험 AI로 분류하고 동일한 투명성 의무를 부과하는 것이 하나의 방향이다.

(2) 가격 개인화 고지 의무

투명성보다 한 단계 낮은 수준의 규제로, 알고리즘의 내용까지는 공개하지 않더라도 소비자에게 개인화 가격이 적용되고 있다는 사실 자체를 고지하도록 의무화하는 방안이다. 소비자는 자신이 최대 지불의사액에 근접한 가격을 제시받고 있다는 사실조차 알지 못하는 경우가

대부분이다. 이는 단순한 정보비대칭을 넘어 인지 비대칭의 문제이다. 표시광고법상 기만적 광고의 '부작위에 의한 기만' 규정을 확대 해석하거나, 전자상거래법에 개인화 가격 고지 조항을 신설하는 방법을 검토할 수 있다.

(3) 기준가격 공시 의무

소비자가 자신에게 제시된 가격이 개인화된 것인지 판단하려면 비교 기준이 필요하다. 이를 위해 동일 상품·서비스에 대한 기준가격(reference price)을 공개적으로 공시하도록 의무화하는 방안을 고려할 수 있다. 소비자가 자신에게 제시된 가격과 기준가격을 비교함으로써 개인화 가격차별 여부를 스스로 확인할 수 있게 된다. 이는 레포트가 설명한 가격차별의 전제조건인 '차익거래 불가' 요건을 간접적으로 약화시키는 효과도 있어서, 소비자들이 정보를 공유하여 최저가를 추적하는 행동을 촉진시킨다.

(4) 알고리즘 영향평가와 감사 의무

RegTech의 한계로 지적한 피드백 루프와 편향 고착화 문제는 AI 가격차별 알고리즘에도 동일하게 적용된다. AI 가격차별 알고리즘이 특정 집단—저소득층, 고령층, 정보 취약 계층—에 대해 체계적으로 높은 가격을 부과하는 방향으로 편향될 위험이 있다. 이를 방지하기 위해 독립적 제3자 기관에 의한 정기적 알고리즘 감사(Algorithm Audit) 를 의무화하고, 집단별 가격 분포를 분석하여 차별적 결과가 발생하는지 점검하는 절차를 도입할 수 있다.

(5) 동적 가격차별의 금지 또는 상한 규제

보다 강력한 규제로, 특정 영역에서는 AI 기반 동적 가격차별(dynamic pricing) 자체를 금지하거나 가격 변동 폭에 상한을 설정하는 방안도 검토할 수 있다. 의료·교육·필수 공공재 등 접근성이 사회적으로 중요한 영역에서 AI가 지불의사가 높은 소비자에게 과도하게 높은 가격을 부과하면, 레포트가 설명한 사중손실 감소 효과 없이 분배적 불평등만 심화될 수 있기 때문이다. 다만 이러한 강한 규제는 사업자의 혁신 유인을 저해할 수 있으므로, 영역별로 차등 적용하는 것이 바람직하다.

(6) 한계와 균형점

규제에는 필연적으로 한계가 따른다. 제1종 가격차별은 사중손실을 완전히 제거하고 사회후생을 완전경쟁시장 수준으로 회복시킨다. AI 가격차별도 이론적으로는 거래량을 늘려 사회후생을 개선할 여지가 있다. 따라서 AI 가격차별을 전면 금지하는 것은 오히려 사회후생을 감소시킬 수 있으며, 핵심은 분배적 정의와 효율성 사이의 균형을 어떻게 설정하느냐이다.

이에 대한 향후의 노력이 필요하다.

V. 바람직한 법적 대응 방향

1. 인공지능기본법의 보완

(1) 행위 유형 중심의 행위자 정의 체계 재편

인공지능기본법상 인공지능사업자 정의를 EU AI Act를 참고하여 개발사업자·이용사업자의 이분법에서 벗어나, 시장 출시자·배치자·유통자·수입자 등 구체적 행위 유형을 기준으로 재편하여야 한다. 특히 인공지능시스템을 실제 운용하는 배치자를 독립적 법적 행위자로 규율하고, 이들에게 투명성·위험관리·모니터링·보고 의무를 차별적으로 부과하는 것이 필요하다.

(2) 제공자 지위 전환 및 이용자 의무 규정 신설

인공지능시스템의 순환적 생애주기를 반영하여, 이용사업자가 개발사업자 제공 시스템에 실질적 수정을 가하거나 의도된 목적을 변경하는 경우 이용사업자가 새로운 법적 책임 주체로 전환되는 규정을 신설하여야 한다. 아울러 EU AI Act 제26조를 참고하여 고위험 인공지능을 운용하는 이용자(배치자)에게 이용 지침 준수, 관리·감독 주체 지정, 시스템 이용 대상자 고지 등의 의무를 명시적으로 부과하여야 한다.

(3) 영향받는 자의 권리 강화

인공지능제품·서비스에 의하여 중대한 영향을 받는 자에게 ① 인공지능시스템 이용 대상이 되고 있다는 사실을 통지받을 권리, ② 결정에 활용된 주요 기준·원리에 대한 설명요구권, ③ 이의제기권, ④ 자동화된 결정에만 의존하지 않을 권리를 명시적 권리로 확립하여야 한다.

(4) 딥페이크 등 신종 위험에 대한 명문 금지 규정 도입

EU AI Act Digital Omnibus가 딥페이크 성적 콘텐츠 생성 AI에 대한 금지 규정을 신설한 것에 상응하여, 인공지능기본법 또는 관련 개별법에 동의 없는 딥페이크 생성·유포 AI의 활용 금지, 생성형 AI 서비스 사업자의 워터마킹 의무 등을 명문화할 필요가 있다. 특히 인공지능기본법에 '인공지능의 금지 행위' 규정이 없다는 점은 EU AI Act와 비교하여 중대한 입법 공백으로, 조속한 보완이 요구된다.

(5) AI 규제 기술(RegTech) 도입 시 절차적 정당성 확보 체계 마련

인공지능기본법의 집행력 공백을 보완하기 위해 AI 기술을 소비자보호 규제 수단으로 활용하는 경우, 그 절차적 정당성 확보가 병행되어야 한다. 구체적으로 다음 세 가지 제도적 장치가 필요하다.

첫째, 설명가능한 AI(XAI) 적용 의무화이다. AI 규제 시스템이 불공정약관 조항이나 다크패턴을 위법으로 판단하는 경우, 그 판단 근거를 인간이 이해할 수 있는 형태로 제공하도록 설명가능한 AI 기술 적용을 인증 기준으로 삼아야 한다. 이는 행정절차법 제23조 제1항의 이유제시의무와의 정합성을 확보하고 피규제자의 방어권을 실질적으로 보장하는 토대가 된다.

둘째, 알고리즘 영향평가(Algorithm Impact Assessment) 의무화이다. 소비자 권익에 직접적 영향을 미치는 AI 규제 시스템을 도입하기 전에, 데이터 편향성 검증, 차별적 결과 가능성 분석, 피드백 루프 위험 점검 등을 포함한 사전 영향평가를 의무화하여야 한다. EU AI Act 제9조의 위험관리 체계가 이에 관한 참고 모델을 제공한다.

셋째, 정기적 알고리즘 감사(Algorithm Audit) 및 인간 감독 체계이다. AI 규제 시스템 도입 이후에도 사업자가 AI의 취약점을 역공학적으로 파악하여 규제를 우회하는 기술 경쟁이 발생할 수 있다. 이에 대응하기 위해 독립적 전문기관에 의한 정기 감사와 함께, EU AI Act 제14조의 '인간 감독(Human Oversight)' 의무에 상응하는 규정을 인공지능기본법에 명시할 필요가 있다.

2. 인공지능책임 법제의 정비

(1) 위험책임 입법 및 소비자 친화적 입증책임 완화

고영향·고성능 인공지능에 대해 개발사업자에게 무과실책임을 부과하는 명문의 입법이 필요하다.[49] 동시에 인공지능 가치사슬의 복잡성으로 인한 피해자의 입증 곤란 문제를 해소하기 위해, 2026년 2월 개정된 상생협력법(제40조의5·제40조의6·제40조의11)의 증거개시 제도와 자료보전명령 제도를 인공지능 피해 구제에도 도입하여야 한다.[50]

(2) 제조물책임법의 적용 범위 확대

EU 새 제조물책임지침안(COM/2022/495 final)이 소프트웨어와 인공지능시스템을 제조물에 명시적으로 포함시킨 것처럼[51], 우리 제조물책임법도 '제조되거나 가공된 동산'(제2조 제1호)의 정의를 확대하여 인공지능시스템(소프트웨어)을 포함시켜야 한다.

---

49) 이소은·서종희, 앞의 글, 124면.

50) 대·중소기업 상생협력 촉진에 관한 법률(2026. 2. 19. 법률 제21377호로 일부개정) 제40조의5, 제40조의6, 제40조의11; 이소은·서종희, 앞의 글, 122-123면.

51) Vorschlag für eine RICHTLINIE DES EUROPÄISCHEN PARLAMENTS UND DES RATES über die Haftung für fehlerhafte Produkte (COM/2022/495 final) Art. 4(1).

(3) 보험 및 기금 제도의 정비

고영향·고성능 인공지능에 대한 무과실책임 도입과 연계하여 책임보험 가입 의무를 법률로 규정하고, 책임보험으로 피해가 충분히 전보되지 못하는 경우를 대비한 공적 보상기금 제도를 도입하여야 한다.[52]

VI. 맺음말

인공지능기술이 소비자의 일상에 깊숙이 침투하면서, 인공지능 제품안전과 소비자보호는 현재의 긴급한 법적 과제가 되었다. 2026년 1월 시행된 인공지능기본법은 진흥과 규율의 포괄적 틀을 제시하였으나, 소비자보호 관점에서 세 가지 명확한 한계를 지닌다. 첫째, 인공지능 가치사슬의 다층적 구조를 반영하지 못한 단순한 이분법적 행위자 정의로 인해 법적 책임 공백이 발생할 우려가 있다. 둘째, 이용자의 의무와 영향받는 자의 권리에 관한 실질적 규정이 불충분하다. 셋째, '진흥법'적 성격이 강하여 피해 구제 수단으로서의 실효성이 미흡하고 딥페이크 등 신종 AI 위험에 대한 금지 규정이 없다.

EU AI Act는 제공자와 배치자를 구체적 행위를 기준으로 구분하고, 제25조에서 실질적 수정·의도된 목적 변경 시 책임 주체가 전환되도록 규정함으로써 인공지능 가치사슬의 순환적·동태적 특성을 법리에 반영하고 있다. 나아가 2026년 5월의 디지털 AI 옴니버스 개정을 통해 고위험 AI 의무 이행 기한을 현실에 맞게 조정하면서도 딥페이크 성적 콘텐츠 금지 등 핵심 보호 규범은 강화하는 방향을 취하고 있다.

바람직한 입법 방향은 다음과 같이 정리된다. ① 시장 출시자·배치자·유통자 등 구체적 행위 유형 중심의 행위자 정의 체계를 도입하고, 이용사업자의 실질적 수정·목적 변경 시 제공자 지위 전환 규정을 신설하여야 한다. ② 고영향·고성능 인공지능에 대해서는 개발사업자에게 위험책임(무과실책임)을 명문화하되, 책임보험과 공적 보상기금을 통해 사회적으로 위험을 분산하여야 한다. ③ 제조물책임법상 제조물 개념을 인공지능(소프트웨어)을 포함하도록 확대하고, 증거개시·자료보전명령 제도를 통해 소비자의 입증 부담을 완화하여야 한다. ④ 글로벌 공급망에 대한 역외 적용 조항을 실질화하여야 한다. ⑤ 인공지능기본법에 금지 행위 규정을 신설하여 딥페이크 성적 콘텐츠 등 생성형 AI의 위험에 선제적으로 대응하여야 한다. ⑥ AI 규제 기술(RegTech)을 도입하는 경우, 설명가능한 AI(XAI) 적용 의무화, 알고리즘 영향평가, 정기적 알고리즘 감사 및 인간 감독 체계를 법적 근거와 함께 마련하여 절차적 정당성을 확보하여야 한다.[53]

민사책임법의 목표는 인공지능기술에 수반되는 위험을 완전히 제거하는 것이 아니라, 사회적으로 수용 가능한 수준으로 합리적으로 감소시키면서 피해자인 소비자를 두텁게 보호하는 것이어야 한다. 기술 혁신과 소비자보호라는 두 가지 요청 사이의 균형을 어떻게 달성할 것

52) 이소은·서종희, 앞의 글, 117면; 권영준·이소은, "자율주행자동차 사고와 민사책임", 민사법학 제75호, 2016, 48면.

53) 이소은·서종희, 앞의 글, 125면.

인지는 결국 집단적이고 민주적인 의사결정 과정을 통해 결정되어야 할 사회적 선택의 문제임을 잊지 말아야 한다.