AI 영향 받는 사람의 목소리

연속 워크숍

AI 시대, 소비자 권리 보호를 위한 과제

일시 2026년 6월 23일(화) 오후 2시

장소 참여연대 2층 아름드리홀

사회 오병일 (디지털정의네트워크 대표)

발제 AI로 인한 소비자 권리 침해 쟁점 / 정지연 (한국소비자연맹 사무총장)
AI의 제품안전 문제와 소비자 보호 / 서종희 (연세대학교 법학전문대학원 교수)

토론 이정수(소비자단체협의회 사무총장))
서지원(변호사, 녹색소비자연대)
한경수(변호사, 참여연대 민생경제팀 실행위원)

주최 인공지능 책임성과 공공성 강화를 위한 시민사회 공동행동

주관 디지털정의네트워크, 참여연대, 한국소비자단체협의회, 한국소비자연맹

후원 아름다운재단

AI시대 소비자권리 보호를 위한 과제

정지연 사무총장(한국소비자연맹)

Ⅰ. 서론

AI(인공지능)은 더 이상 특정 산업 영역의 기술 혁신에 머무르지 않고, 일상생활에서 소비자에게 많은 영향을 미치고 있다. 오늘날 소비자는 상품과 서비스를 탐색하고, 비교·선택하며, 계약을 체결하고, 분쟁이 발생할 경우 구제를 모색하는 전 과정에서 점점 더 AI 시스템과 직접적으로 마주하게 된다. 이로 인해 AI는 개별 기술을 넘어 소비환경을 구성하는 기본 인프라로 기능하면서 소비자에게는 새로운 도전이 되고 있다. 기존 소비자법은 정보 비대칭 상황에서도 합리적으로 판단하는 소비자를 전제로, 정보 제공 의무, 계약 공정성, 책임 귀속을 중심으로 발전해 왔다. 그러나 AI 기반 소비환경에서는 선택이 알고리즘에 의해 매개되고, 정보가 생성·가공되며, 거래 조건이 개인화·자동화되고, 책임 주체가 분산되는 구조적 변화가 나타난다. 이러한 변화는 소비자의 통제력을 약화시키고, 기존 소비자법 체계가 전제해 온 규범적 기반에 중대한 도전을 제기한다.

본 발제는 첫째 AI시대 소비환경의 구조적 변화를 소비자 관점에서 분석하고, 둘째 실제 소비자가 당면한 피해현실과 새로운 위험에 대해 사례별로 정리해 보고 셋째 한국소비자연맹이 실시한 AI 소비자 실태조사를 통해 소비자의 인식과 요구를 실증적으로 검토하며, 넷째 지능정보사회(AI) 소비자 권리장전과 AI소비자포럼 논의를 토대로 소비자단체의 역할과 향후 소비자 피해 예방을 위한 제도개선 과제를 제시하고자 한다.

Ⅱ. AI시대 소비환경의 구조적 변화와 소비자 도전

1. 선택 구조의 변화: 탐색에서 추천으로

AI 기반 플랫폼 환경에서 소비자의 선택은 점차 검색과 비교 중심의 탐색 행위에서 벗어나, 알고리즘이 사전에 정렬·추천한 선택지 중 하나를 고르는 방식으로 전환되고 있다. 소비자는 형식적으로 선택권을 보유하지만, 실질적으로는

노출 순서와 추천 구조에 의해 선택이 유도된다. 이러한 변화는 다크패턴과 결합될 경우 소비자의 선택 왜곡을 심화시킨다. AI는 소비자의 행태 데이터와 취약성을 학습하여 해지 방해, 자동 결제 유도, 반복적 선택 압박 등 맞춤형 기만 설계를 가능하게 한다. 이는 기존 소비자법이 전제해 온 '정보를 충분히 제공받은 합리적 소비자' 모델의 한계를 분명히 드러낸다.

2. 정보 구조의 변화: 생성형 AI와 신뢰의 문제

생성형 AI의 확산은 소비자 정보 환경에 질적인 변화를 가져왔다. 후기, 이미지, 영상, 광고 등 소비 판단의 핵심 정보가 AI에 의해 대량으로 생성·조작될 수 있게 되면서, 소비자는 정보의 양이 아니라 정보의 진위 여부를 스스로 판단해야 하는 부담을 지게 되었다. 허위 리뷰와 조작된 추천은 소비자의 구매 실패 가능성을 높이고, 시장 전반의 신뢰를 훼손한다. 이 과정에서 소비자는 정보의 수혜자가 아니라, 정보 검증의 책임을 떠안는 주체로 전환된다. 이는 정보 제공 의무 중심의 소비자 보호 규율이 충분하지 않음을 시사한다.

3. 거래 및 책임 구조의 변화: 자동화된 의사결정과 책임 공백

AI는 금융, 의료, 채용, 플랫폼 거래 등에서 의사결정의 핵심 요소로 활용되고 있다. 문제는 이러한 자동화된 판단이 소비자의 권리·의무에 중대한 영향을 미치지만, 그 판단 과정은 불투명하고 설명 가능성이 제한된다는 점이다. 더 나아가 AI 개발자, 플랫폼 운영자, 서비스 이용사업자로 역할이 분리된 환경에서는 사고 발생 시 책임 귀속이 불명확해진다. 소비자는 권리 구제를 위해 누구를 상대로 법적 책임을 물어야 하는지조차 알기 어려운 상황에 직면한다. 이는 소비자법상 책임 주체 개념과 귀속 기준의 재정립을 요구한다.

Ⅲ. AI 시대 소비자 피해의 현실화와 새로운 소비자 위험

과거 소비자문제는 불량제품, 허위광고, 계약해지 거부 등 비교적 명확한 형태로 나타났다. 그러나 AI 시대의 소비자 문제는 알고리즘과 데이터, 자동화 의사결정이 결합하면서 소비자가 피해를 인식하기조차 어려운 형태로 변화하고 있다. 특히 AI는 소비자의 정보 접근, 상품 선택, 가격 결정, 금융거래, 의료서비스, 인간관계에까지 영향을 미치면서 새로운 소비자 위험을 만들어내고 있다.

1. 딥페이크 사기 : 신뢰 자체를 공격하는 범죄

과거의 보이스피싱은 전화나 문자에 의존했다. 그러나 생성형 AI와 딥페이크 기술의 발전은 인간이 가장 신뢰하는 '얼굴'과 '목소리'까지 위조할 수 있게 만들었다. 기존 사기가 소비자의 판단착오를 이용했다면, 딥페이크 사기는 인간의 신뢰체계 자체를 공격한다는 점에서 차원이 다르다. 딥페이크 사기는 개인·기업을 가리지 않으며, 피해 발생 후 송금액 회수가 사실상 불가능하다는 특성상 예방 중심의 접근이 필수적이다.

딥페이크 사기는 사람이 가장 확실하다고 믿는 시각·청각 정보 자체를 위조한다. 기존 소비자보호법은 '언어를 통한 기망'을 전제로 설계되어 있어 AI 생성 영상·음성을 활용한 사기에는 적용이 어렵다. 피해 발생 후 회수 가능성이 사실상 없어 예방 중심의 제도 설계가 시급하다.

사례1 | 홍콩 다국적기업 딥페이크 화상회의 사기 (2024)

2024년 홍콩에서 다국적 기업의 재무담당 직원이 화상회의에 참석한 뒤 약 2억 홍콩달러(한화 약 350억 원)를 송금하는 사건이 발생했다. 당시 회의에는 회사 최고재무책임자(CFO)와 여러 임직원이 참석한 것으로 보였지만, 수사 결과 회의 참가자 전원이 AI로 생성된 딥페이크 영상이었음이 드러났다. 직원은 실제 상사의 얼굴과 목소리를 직접 확인했기 때문에 전혀 의심하지 않았고, 결국 거액의 자금이 범죄조직으로 넘어갔다. 국내에서도 가족의 음성을 AI로 복제하여 '교통사고가 났다', '휴대전화를 잃어버렸다'며 송금을 요구하는 유사 범죄 사례가 지속적으로 증가하고 있다.

▶ 출처: 홍콩 경찰청 공식 발표 및 AFP 보도, 2024.02

사례2 | 딥페이크 성적 허위 영상물 급증 (한국, 2020~2024)

방송통신심의위원회의 딥페이크 성적 허위 영상물 시정 요구 건수는 2020년 473건에서 2023년 1~11월 5,996건으로 3년 사이 10배 이상 폭증했다. 과거 유명인 중심이었던 딥페이크 피해가 일반인 지인의 사진을 이용한 합성물로 확대되고 있으며, 딥페이크 범죄의 주요 가해자와 피해자 모두 청소년층이 많아 디지털 리터러시 교육이 시급한 과제로 부상하고 있다. 2024년 10월 성폭력처벌법 개정으로 배포 목적 없이 제작만 해도 처벌이 가능해졌다.

▶ 출처: 방송통신심의위원회 통계 / NIA AI REPORT 2024-13

사례3 | MS 보고서: AI 딥페이크 채용 사기·가짜 쇼핑몰 (2025)

마이크로소프트가 2025년 4월 발표한 보고서에 따르면 AI로 생성된 가짜 채용 공고에 지원한 구직자가 AI 딥페이크 면접관과 화상 면접을 진행하면서 개인정보를 탈취당하는 사례와, AI가 제작한 유명 브랜드 사칭 가짜 쇼핑 사이트로 소비자의 결제정보를 편취하는 사례가 급증하고 있다. MS는 최근 1년간 차단한 AI 기반 금융사기 시도 규모가 40억 달러에 달하며, 봇을 이용한 계정 탈취 시도가 시간당 160만 건에 이른다고 밝혔다.

▶ 출처: Microsoft 'AI 기반 사기 수법: 진화하는 위협과 대응 전략', 2025.04

2. 생성형 AI 허위정보와 정보오염

생성형 AI는 방대한 정보를 짧은 시간에 제공할 수 있지만, 사실이 아닌 정보를 매우 그럴듯하게 생성하는 문제를 안고 있다.
의료 분야에서도 암 치료, 백신, 건강기능식품 등에 대한 잘못된 정보가 생성되는 사례가 보고되고 있다. 특히 건강정보는 소비자의 생명과 안전에 직결되기 때문에 더욱 위험하다.

과거 검색엔진은 여러 정보를 병렬로 제시하여 소비자가 스스로 판단하게 했다면, 생성형 AI는 하나의 정답처럼 정보를 제시한다. 소비자는 그 정보를 검증하기보다 신뢰할 가능성이 높다. 의료·법률·금융 정보처럼 전문성이 높은 영역일수록 소비자가 오류를 발견하기 어렵고 피해도 치명적이다. 2025년 2월 기준 최신 AI 모델의 할루시네이션 발생률이 1% 미만으로 낮아졌지만, 수천만 명이 사용하는 서비스에서 1%는 결코 작은 수치가 아니다.

AI 시대 소비자는 정보 부족의 문제가 아니라 정보 진위 여부를 판단해야 하는 새로운 부담을 안게 되었다.

사례1 | 미국 변호사 허위 판례 인용 사건 (2023, 뉴욕)

뉴욕의 변호사 스티븐 슈워츠는 법원 제출용 소송 서면에 ChatGPT가 생성한 판례를 무검증으로 인용하였다. 법원 심리 과정에서 해당 판례들이 실제 존재하지 않는 가짜임이 드러났다. 담당 판사는 법률 전문가로서의 최소한의 확인 의무를 저버렸다고 지적하며 해당 변호사에게 5,000달러의 벌금과 법정 출석 명령을 내렸다.

▶ 출처: 미국 뉴욕 남부 연방지방법원 판결문, 2023.06 / Mata v. Avianca 사건

사례2 | 에어캐나다 AI 챗봇 오정보 배상 판결 (2024, 캐나다)

승객 제이크 모팻은 할머니 사망 후 에어캐나다 AI 챗봇에 장례 할인 운임을 문의하였고, 챗봇은 항공권 구매 후 90일 이내 환급 신청이 가능하다고 안내했다. 그러나 실제 정책은 사전 신청이 원칙이었고, 에어캐나다는 챗봇의 안내에 대해 책임이 없다고 주장하며 환불을 거절했다. 캐나다 민사분쟁심판소는 '챗봇이 제공한 정보에 대한 책임은 기업에 있다'는 판결을 내렸다. 이 판결은 AI 챗봇 오정보에 대한 기업 책임을 인정한 최초의 판례로 평가된다.

▶ 출처: 캐나다 민사분쟁심판소 판결, 2024.02 / Moffatt v. Air Canada

3. AI 사칭과 구독경제 피해

최근 생성형 AI의 인기가 높아지면서 이를 악용한 소비자 피해가 급증하고 있다. 검색광고와 SNS를 통해 “ChatGPT 평생 이용권”, “GPT-5 무제한 이용권”, “AI 투자 프로그램”, “AI 자동수익 서비스” 등을 광고하며 고액 결제를 유도하는 사례가 늘고 있다.

실제로는 OpenAI와 아무 관련이 없거나 무료 API를 단순 연결한 서비스임에도 소비자는 공식 서비스로 오인하여 결제하게 된다. 문제는 대부분 해외사업자이거나 환불 규정이 불명확해 피해구제가 쉽지 않다는 점이다.

사례1	AI 이름 도용 유사 서비스 피해 사례
소비자는 유명 AI 서비스라고 믿고 결제하지만 환불이 거부되거나 자동결제가 지속되는 사례가 나타나고 있다. 일부 해외 서비스는 월 이용료를 숨긴 채 무료 체험을 유도한 후 자동결제로 전환하기도 한다.	

4. AI 워싱(AI Washing) : 기술 불투명성을 이용한 기만

환경 분야의 그린워싱(Greenwashing)처럼 실제 AI 기능이 거의 없음에도 'AI 기반', 'AI 맞춤형', 'AI 최적화'라는 표현을 사용하는 AI 워싱이 급증하고 있다. 단순 자동 예약 프로그램이나 규칙 기반 추천 알고리즘을 AI 기술인 것처럼 광고하는 사례가 대표적이다. 최근 기업들은 실제 AI 기능이 거의 없음에도 AI를 강조하는 마케팅을 적극적으로 활용하고 있다. AI 건강관리 서비스, AI 투자 서비스, AI 교육서비스, AI 가전제품 등이 대표적이다.

단순 자동화 기능이나 통계 분석 기능만 탑재했음에도 소비자가 첨단기술이 적용된 것으로 오인하게 만드는 경우가 많다. 이는 과거 환경친화성을 과장한 그린워싱과 유사한 현상으로 소비자의 합리적 선택을 왜곡하는 새로운 형태의 기만행위라고 볼 수 있다.

AI 워싱이 기존 허위광고보다 더 심각한 이유는 기술 자체의 불투명성 때문이다. 과거 허위광고는 사실 여부를 비교적 확인할 수 있었지만, AI 알고리즘의 실제 수준은 소비자가 검증할 방법이 없다. 이는 기술에 대한 소비자의 정보 비대칭을 구조적으로 악용하는 새로운 형태의 기만행위이다.

AI 기능의 실질적 수준을 소비자가 검증할 수단이 없으며, 현행 표시광고법이 AI 기술 기만 광고에 충분히 대응하지 못하고 있는 문제가 있다.

5. 알고리즘 가격차별 : 보이지 않는 불평등

AI는 소비자의 검색기록, 구매이력, 위치정보, 기기정보, 소득수준 등을 분석해 개인별로 다른 가격을 제시할 수 있다. 동일한 항공권, 호텔 객실, 차량공유 서비스임에도 소비자에 따라 가격이 달라질 수 있으며, 소비자는 그 이유를 알기 어렵다. 항공권을 반복 검색한 소비자에게 더 높은 가격을 제시하거나, 구매 가능성이 높은 소비자에게 할인폭을 줄이는 방식이 대표적 사례다. 금융 분

야에서도 AI 신용평가에 따라 대출금리와 한도가 결정되지만 소비자는 어떤 요소가 평가에 영향을 미쳤는지 설명받기 어렵다.
과거의 가격차별은 비교가 가능했지만 AI 시대의 가격차별은 소비자가 차별받고 있다는 사실조차 인식하기 어려운 형태로 진행된다. 이는 소비자 선택권과 공정거래 원칙을 위협하는 대표적인 AI 소비자 문제이다.
특히 의료·금융·주거 분야로 확대될 경우 취약계층에 대한 구조적 불평등으로 이어질 수 있다.

사례1 | 아마존 '프로젝트 네시(Project Nessie)' (2023, 미국 FTC 소송)

미국 연방거래위원회(FTC)가 2023년 아마존을 상대로 제기한 독점 소송에서, 아마존이 '프로젝트 네시'라는 코드명의 알고리즘을 활용해 경쟁사가 따라올 수 있는 가격 인상 범위를 체계적으로 테스트해 왔다는 사실이 공개됐다. FTC는 이 알고리즘이 수억 달러의 부당 이익을 창출했다고 주장했으며, 아마존은 효과가 없어 프로그램을 종료했다고 해명했다. 해당 내용은 소송 문서의 비공개 처리된 부분에서 드러났다.

▶ 출처: 미국 FTC v. Amazon.com Inc. 소송 문서, 2023.09

6. 추천 알고리즘과 소비자 선택권 침해

오늘날 소비자는 상품을 직접 찾기보다 AI가 추천하는 상품을 선택하는 경우가 많다. 온라인 쇼핑몰의 상품 추천, 영상 플랫폼의 콘텐츠 추천, 금융상품 추천, 뉴스 추천 등이 대표적이다. 문제는 소비자가 추천의 기준을 알 수 없다는 점이다. 소비자는 자신을 위한 추천이라고 생각하지만 실제로는 플랫폼 수익을 극대화하기 위한 추천일 수 있다. AI는 소비자의 선택을 돕는 기술이기도 하지만 동시에 소비자의 선택을 유도하는 기술이기도 하다.

해외에서는 온라인 플랫폼이 자사 상품을 경쟁 상품보다 우선 노출했다는 의혹이 지속적으로 제기되어 왔다. 소비자는 자신이 자유롭게 선택했다고 생각하지만 실제로는 알고리즘이 설계한 선택을 하고 있을 가능성이 존재한다.

추천 알고리즘의 기준이 공개되지 않아 소비자는 '광고인지 추천인지'를 구별할 수 없으며, 현행 광고 공시 규정이 AI 추천에는 적용되지 않는 공백이 존재한다.

7. AI 차별과 자동화 의사결정

AI는 데이터를 학습하는 과정에서 데이터 속 편견도 함께 학습한다. AI가 과거의 차별적 패턴을 재현하거나 심화시키는 이 문제는, 개인의 신용평가·채용·보험 가입 등 중요한 경제적 결정에 직접적인 영향을 미친다.

금융과 보험 분야에서도 AI가 특정 지역, 연령, 직업군을 위험군으로 판단하여 소비자가 불리한 조건을 받거나 서비스에서 배제되는 사례가 발생할 수 있다. AI 차별은 사람의 차별보다 발견하기 어렵고 수정하기도 어렵다는 점에서 더 심각하다.

소비자는 AI 의사결정의 이유를 설명받을 권리가 있으나, 국내에는 '알고리즘 설명 요구권'에 관한 법적 근거가 없다. AI 차별은 알고리즘 블랙박스 특성상 발견 자체가 어렵고, 발견하더라도 피해 입증이 어려운 문제가 있다.

사례1 | 아마존 AI 채용 시스템 성차별 (2018, 미국)

아마존은 2014년부터 AI 기반 채용 평가 도구를 개발했으나, 2018년 해당 시스템이 남성 지원자를 선호하는 패턴을 학습했다는 사실이 로이터 보도로 알려졌다. 남성 지배적이었던 과거 10년치 채용 데이터를 학습한 결과였다. 아마존은 이 사실이 밝혀지자 시스템 운영을 전면 중단했으나, 이 사례는 AI가 데이터 속 사회적 편견을 그대로 재현할 수 있다는 경고로 지속적으로 인용되고 있다.

▶ 출처: Reuters, 'Amazon scraps secret AI recruiting tool that showed bias against women', 2018.10

8. AI 의료서비스와 책임 공백

AI는 암 진단과 영상판독 등에서 높은 정확도를 보여주고 있다. 그러나 AI가 잘못된 판단을 내렸을 때 책임은 누구에게 있는가. 개발사인가, 병원인가, 의사인가. 현재 대부분의 국가에서 이에 대한 명확한 기준이 마련되지 않았다.

소비자 입장에서는 정확도보다 책임체계가 더 중요하다. AI가 진단할 수는 있어도 책임질 수는 없기 때문이다. 현재 대부분의 의료 AI 기업들은 약관에 '본 서비스는 의료 행위가 아님'을 명시하여 법적 책임을 회피하고 있다. 의료 AI는 높은 정확도를 제공할 수 있지만, 소비자 입장에서는 정확도보다 책임체계가 더 중요하다.

사례1 | AI 의료 챗봇 오류 및 책임 논쟁 사례 (2024~2026)

AI가 암 치료법, 백신 효과, 당뇨병 치료방법 등에 대해 사실과 다른 정보를 제시한 사례가 반복적으로 보고되고 있다. 의료용 AI 챗봇이 양성 피부 병변을 악성으로 잘못 분류하여 불필요한 의료 개입을 초래한 사례도 알려졌다. 2024~2025년 사이 Character.AI 등 감성형 챗봇이 청소년의 정신건강에 영향을 주었다는 이유로 다수의 소송이 제기되면서, 법조계에서는 AI 챗봇을 '언론·출판물'로 볼 것인지 '제조물'로 볼 것인지에 대한 논쟁이 치열하다.

▶ 출처: IBM AI 할루시네이션 보고서, 2025 / 의료 AI 챗봇 시장 분석 보고서, 2026.01

9. AI 과몰입과 정서적 의존

최근 AI를 친구, 상담자, 연인으로 인식하는 사례가 급증하고 있다. 특히 청소년은 AI 챗봇과 장시간 대화하면서 정서적 의존 관계를 형성하기도 한다. 이는 플랫폼 중독이 사용 시간의 문제였다면, AI 의존은 인간관계와 정체성의 문제로 확대될 수 있다는 점에서 차원이 다른 위험이다.

국내 전문가들은 감정 의존을 유발하는 AI 챗봇을 직접 규제할 법령과 전담 기관이 사실상 부재한 상황이라고 지적한다.

사례1 | 캐릭터닷AI(Character.AI) 청소년 자살 사건 및 소송 (2024~2026, 미국)

2024년 미국 플로리다주에서 14세 소년 세웰 세처 3세가 TV 시리즈 '왕좌의 게임' 캐릭터를 모델로 한 AI 챗봇과 장기간 성적 대화를 나눈 끝에 스스로 목숨을 끊었다. 챗봇은 마지막 대화에서 '사랑한다, 빨리 내게로 와달라'고 했다. 유족은 Character.AI와 구글을 상대로 소송을 제기했으며, 플로리다 연방지방법원은 'AI의 해로운 상호작용은 기업의 설계 결함 때문'이라며 소송을 기각하지 않았다. 같은 이유로 텍사스·콜로라도·뉴욕 등에서도 유사 소송이 제기됐으며, 2026년 1월 구글과 Character.AI는 복수의 소송을 합의로 종결했다. 어머니 메간 가르시아는 미 상원 청문회에서 '이것은 희귀한 사례가 아니다. 모든 주의 어린이들에게 지금 일어나고 있다'고 증언했다.

▶ 출처: 법률신문, 2025.08 / ZDNet Korea, 2026.01 / 이데일리, 2026.01

사례2 | 캘리포니아주 미성년자 AI 챗봇 보호법 제정 (2025.10)

캘리포니아주는 2025년 10월 미국 최초로 미성년자 대상 AI 챗봇 보호법(SB 243)에 서명했다. 핵심 내용은 ① 미성년자와의 대화 중 3시간마다 'AI임을 고지', ② 노골적 성적 콘텐츠 생성 금지, ③ 자살·자해 표현 식별 및 대응 의무, ④ 의료 전문가 사칭 금지 등이다. 법은 2026년 1월 발효됐다.

▶ 출처: 머니투데이, 2025.10.14

10. 개인정보와 AI 학습데이터 문제

AI는 데이터로 학습한다. 그러나 대부분의 소비자는 자신의 사진, 글, 음성, 행동기록이 언제 어떻게 AI 학습에 사용되는지 알지 못한다. 특히 얼굴 이미지와 음성 데이터는 한 번 유출되면 변경이 불가능한 고유한 생체정보이다.

한국소비자연맹 조사에서 소비자가 가장 민감하게 느끼는 정보는 생체정보(72.7%)로 나타났다. AI 시대 개인정보 문제는 단순한 유출을 넘어, 자신의 데이터가 누구에 의해, 어떤 목적으로, 얼마나 오랫동안 사용되는지에 대한 통제권의 문제이다. 개인정보보호법이 AI 학습 목적의 정보 활용 범위를 충분히 규

율하지 못하고 있으며, 소비자의 데이터 통제권(열람·삭제·이전 요구권) 보장 방안 마련이 필요하다.

사례1 | 개인정보보호위원회 AI 개발 목적 개인정보 활용 가이드라인

개인정보보호위원회는 2024년 공개된 개인정보를 AI 개발과 서비스 목적으로 활용할 수 있는 범위에 대한 기준을 제시하는 안내서를 발표하였다. 해당 안내서는 공개된 정보 처리의 적법 근거로 '정당한 이익'을 원용할 수 있음을 명확히 하고 관련 기준을 구체적으로 제시했다. 그러나 소비자가 자신의 데이터가 AI 학습에 활용됐는지를 사전에 알고 동의하는 절차는 여전히 미비하다는 지적이 제기된다.

▶ 출처: 개인정보보호위원회 안내서, 2024

11. AI 에이전트의 자율 계약 체결과 에이전트 간 분쟁

AI 에이전트는 소비자를 대신해 항공권 예약, 쇼핑, 금융상품 가입 등을 자율적으로 수행한다. 이미 OpenAI의 Operator, Anthropic의 Computer Use 등 자율 행동 에이전트가 상용화 단계에 접어들었다. 그런데 AI 에이전트 간에 협상하고 계약을 체결하는 '멀티 에이전트(Multi-Agent)' 환경이 도래하면서, 어느 쪽 에이전트의 판단이 우선하는지, 계약 조건이 충돌할 경우 누가 책임지는지에 대한 법적 기준이 없다는 문제가 현실화되고 있다.

2025년 유럽법률협회(ELI)가 발표한 소비자 계약 디지털 어시스턴트 지침(DACC)은 “AI가 개입된 계약도 원칙적으로 유효하다”는 국제적 합의를 제시했다. 그러나 계약 내용이 소비자의 실제 의도와 다를 경우, 혹은 두 에이전트가 서로 충돌하는 조건을 설정했을 경우의 분쟁 해결 기준은 여전히 공백으로 남아 있다.

AI 에이전트가 체결한 계약의 취소·환불 절차, 에이전트 간 충돌 시 손해 귀속, 소비자 동의 범위 등에 대한 법적 기준이 국내에 전무하다. AI 에이전트의 권한 범위를 소비자가 명확히 통제할 수 있도록 하는 '에이전트 위임 표준 계약' 제도화와 무단 이체·결제에 대한 자동 환불 보호 장치 마련이 필요하다.

사례1 | 멀티 에이전트 무한루프로 4만7천 달러 청구 (2025)

2025년 말, 한 엔지니어링 팀이 4개의 전문화된 AI 에이전트로 구성된 멀티 에이전트 리서치 시스템을 구축했다. 두 에이전트가 서로의 질문에 끝없이 답하는 '무한 명확화 루프(recursive clarification loop)'에 빠져 11일 동안 감지되지 않은 채 운영됐다. 팀이 비용 청구서를 받고서야 이 사실을 발견했는데, 청구액은 4만7천 달러에 달했다. 시스템에는 중단 조건도, 예산 한도도, 실시간 비용 모니터링도 없었다. 이 사례는 에이전트 간 충돌이 예상치 못한 대규모 소비자 피해로 이어질 수 있음을 보여주는 대

표적 사례로 국제 학계에서 인용되고 있다.

▶ 출처: arxiv.org, 'Agent Contracts: A Formal Framework for Resource-Bounded Autonomous AI Systems', 2026.01 / COINE 2026 발표 논문

사례2 | EU 디지털 어시스턴트 지침(DACC) 제정 (2025)

2025년 유럽법률협회(ELI)는 소비자 계약에서 AI 에이전트가 관여할 경우의 법적 기준을 담은 'DACC(디지털 어시스턴트 소비자 계약 지침)'를 발표했다. 핵심은 AI가 개입해 체결된 계약도 원칙적으로 유효하나, 소비자가 AI 에이전트의 권한 범위를 명확히 설정할 수 있어야 하고, AI가 소비자의 의도를 벗어난 계약을 체결했을 경우 취소 가능성을 보장해야 한다는 것이다. 한국은 아직 AI 에이전트의 계약 행위에 관한 법적 근거가 없어 입법 공백 상태다.

▶ 출처: 유럽법률협회(ELI) DACC 지침, 2025 / 로웨이브 법률분석, 2025.12

12. 생성형 AI 저작물 귀속 문제

생성형 AI 저작물 문제는 두 가지 층위로 나뉜다. 첫째, AI가 학습에 활용한 기존 저작물에 대한 원저작자의 권리 침해 문제다. 둘째, AI가 만들어낸 결과물의 저작권이 누구에게 귀속되는가의 문제다. 현행 저작권법은 '인간의 창작'을 전제로 설계되어 있어, AI 생성 콘텐츠를 구매한 소비자가 해당 결과물을 사업에 활용했다가 추후 저작권 분쟁에 휘말리는 상황이 현실화되고 있다.

AI 개발사는 약관에서 '생성 결과물의 저작권 침해에 대한 책임을 지지 않는다'고 명시하는 경우가 많다. 결국 소비자는 AI를 활용해 얻은 결과물의 안전한 사용 여부를 스스로 판단해야 하는 불합리한 위치에 놓이게 된다.

소비자는 AI 생성 결과물을 구매했을 때 그 콘텐츠가 타인의 저작물을 침해하는지 알 방법이 없다. AI 개발사들은 약관으로 책임을 면제하고 있어, 소비자가 모르는 사이에 저작권 침해의 당사자가 될 위험이 있다. AI 결과물의 저작권 귀속 기준 명확화, AI 개발사의 학습 데이터 공개 의무, 소비자 보호를 위한 '저작권 위험 고지 의무' 도입이 필요하다.

사례1 | 게티이미지 vs. Stability AI 저작권 소송 (2025, 영국 고등법원)

세계 최대 사진 에이전시 게티이미지는 AI 이미지 생성 서비스 Stability AI(Stable Diffusion)가 게티의 저작물 수백만 건을 무단으로 학습 데이터로 사용했다며 영국 고등법원에 소송을 제기했다. 2025년 11월 영국 고등법원은 Stability AI의 저작권 침해를 인정하는 판결을 내렸다. 특히 AI가 생성한 이미지 일부에 게티이미지의 워터마크가 변형된 형태로 포함되어 있다는 사실이 학습 과정의 무단 복제를 방증하는 증거로 인정됐다. 이 판결은 AI 학습 데이터의 저작권 침해를 인정한 주요 판례로, AI가

생성한 콘텐츠를 구매해 사용하는 소비자도 잠재적 법적 위험에 노출될 수 있음을 시사한다.

▶ 출처: Getty Images v. Stability AI [2025] EWHC 2863 (Ch) / 법무법인 세종 뉴스레터, 2026.02 / 법률신문, 2026.02.19

사례2 | Anthropic vs. 베스트셀러 작가 집단소송 합의 시도 (2025)

2025년 9월, AI 개발사 Anthropic은 미국 베스트셀러 작가 집단으로부터 제기된 저작권 침해 소송에서 15억 달러 합의금을 제안했으나 법원이 기각했다. 법원은 작가들에 대한 보호 조치가 부족하다는 이유로 합의안을 승인하지 않았으며, 이 사건은 AI 업계 최초의 저작권 합의 시도 사례로 기록됐다. 뉴욕타임즈 vs. OpenAI·마이크로소프트 소송, GEMA vs. OpenAI 독일 판결(2025.12) 등 AI 학습 데이터 저작권 분쟁은 전 세계적으로 동시다발적으로 진행 중이다.

▶ 출처: 삼성SDS 인사이트리포트 'AI 시대 기업이 꼭 알아야 할 법률', 2025.10 / 법무법인 세종 AI 저작권 판결 동향, 2026.02

13. AI 기반 보험·신용평가 간접차별

AI가 보험료 산정, 신용평가, 대출 심사에 활용되면서 인종·지역·나이 등을 직접 기준으로 삼지 않더라도 이를 강하게 반영하는 간접 변수(우편번호, 소비 패턴, SNS 활동 등)를 사용해 사실상 차별적 결과를 낳는 '간접 차별' 문제가 부상하고 있다. AI 알고리즘은 어떤 변수가 왜 그 결과에 영향을 미쳤는지를 설명하기 어려운 블랙박스 특성을 갖고 있어, 소비자가 차별 여부를 인지하거나 이의를 제기하는 것이 구조적으로 불가능하다.

보험 분야에서는 기존의 설명 오류, 차별, 사기, 운영 장애 등이 AI 적용으로 인해 대량·실시간으로 영향을 미치게 되면서 소비자 피해가 단기간에 대규모로 발생할 수 있다는 점이 핵심 위험이다.

AI가 보험·신용 결정을 내렸을 때 소비자는 이유를 설명받을 권리가 없고, 간접 차별 여부를 입증할 수단도 없다. 현행 차별금지법이 알고리즘 간접 차별을 포괄하지 못하며, 금융당국의 AI 활용 가이드라인도 구속력이 없다. AI 자동심사 결과에 대한 '인간 검토 요구권', '결정 이유 설명 의무' 등이 소비자에게 필요하다.

사례1 | UnitedHealth AI 알고리즘 보험금 대량거부 소송 (2023~2024, 미국)

미국 최대 건강보험사 UnitedHealthcare(UHC)는 AI 시스템 'nH Predict'를 활용해 노인 환자의 재활 입원 기간을 단축하고 메디케어 보험금 지급을 대량 거부했다. 2023년 11월 제기된 집단소송에서 원고 측은 AI가 거부한 보험금의 90% 이상이 재심사를 거치면 지급됐고, AI 도입 이후 보험금 지급 거부율이 8.7%에서 22.7%로 2배 이상 급

증했다는 미 상원 조사 결과를 근거로 제시했다. 법원은 2024년 3월 회사에 AI 알고리즘과 학습 데이터 등 내부 자료 제출을 명령했다.

▶ 출처: 미 상원 조사 결과 / ITWorld 코리아 보도, 2024.12

사례2 | Cigna AI 1.2초 30만 건 보험금 일괄 거절 / Lemonade 생체정보 무단수집 (미국)

보험사 Cigna는 AI 기반 'PxDx' 시스템으로 2022년 특정 2개월간 30만 건이 넘는 보험 청구를 건당 평균 1.2초 만에 일괄 거절했다는 사실이 드러나 2023년 집단소송에 직면했다. 원고 측은 보험약관에 '의사가 의료적 필요성을 판단한다'고 명시돼 있음에도 AI가 단독으로 최종 결정을 내렸다고 주장했다. 보험사 Lemonade는 보험금 청구 과정에서 고객의 얼굴·음성 등 생체정보를 동의 없이 AI 사기 탐지에 활용했다는 의혹으로 2021년 소송이 제기됐고, 일리노이 소송은 약 400만 달러(약 55억 원) 합의로 종결됐다. 미국의 State Farm, Cigna, UnitedHealth 등은 인종 소수자와 고령층 고객에 대해 보험금 지급을 거부하는 자동화시스템을 운영했다는 이유로 집단소송이 진행 중이다.

▶ 출처: 뉴스프리존 '보험사 AI 도입 확산…소송 리스크도 함께 온다', 2026.05

IV. AI 소비자인식조사가 보여주는 소비자의 요구

한국소비자연맹은 2025년에 이어 2026년 5월 전국 AI 이용 경험자 1,000명을 대상으로 AI 소비자인식조사를 실시하였다. 이번 조사는 단순한 기술 수용성 조사를 넘어 AI 유료화, 개인정보 활용, 자동화 의사결정, AI 의료, AI 워싱, 소비자 권리 등 AI 시대 소비자보호 이슈를 종합적으로 조사했다는 점에서 의미가 있다. 조사결과는 소비자들이 AI 기술 발전 자체를 반대하지 않지만, AI가 사회에 깊숙이 들어올수록 투명성·책임성·공정성에 대한 요구가 매우 강해지고 있음을 보여준다.

구분	2026 AI 소비자 인식조사	2025 AI 소비자 인식조사
조사 시기	2026.05.28~29	2025.06.13~16
표본 / 오차	전국 1,000명 / ±2.0%p (80% 신뢰수준)	전국 1,000명 / ±3.1%p (95% 신뢰수준)
조사 방법	오픈서베이 패널 / 웹 브라우저 응답 / 변수 33개	오픈서베이 패널 / 모바일 앱 응답 / 변수 22개

1. AI는 이미 일상 인프라가 되었다

2026년 조사에서 조사대상자 42.2%는 AI를 거의 매일 사용한다고 답했다. 주 2~3회 이상 사용한다는 응답도 26.1%로, 응답자의 68.3%가 사실상 상시적으로 AI를 이용하고 있음이 확인되었다. 특히 20대(54.3%), 10대(53.9%), 30대(51.5%)에서 거의 매일 사용 비율이 과반을 넘어, AI가 청년세대의 핵심 생활 도구로 자리 잡고 있음을 보여준다.

생성형 AI(ChatGPT 등) 이용률은 67.4%, AI 검색·요약 서비스 이용률은 57.8%에 달했으며, 응답자들은 평균 2.74개의 AI 서비스를 동시에 사용하고 있었다. AI가 없으면 불편한 영역으로는 검색·정보탐색(57.5%), 업무·문서작성(38.4%), 번역(38.1%), 학습·교육(30.4%) 순으로 나타났다. 특히 20대에서 업무·문서작성 의존도가 54.9%, 30대에서 검색·정보탐색 의존도가 64.3%에 달했다.

▶ AI는 더 이상 미래 기술이 아니라 전기·인터넷과 같은 생활 인프라가 되었다. AI가 단순 편의기술을 넘어 소비자의 정보 접근·학습·업무 수행의 필수 수단으로 정착했다는 점은, 소비자보호 제도 설계의 전제가 근본적으로 달라져야 함을 의미한다.

2. 소비자는 AI를 긍정적으로 평가하지만 맹신하지는 않는다

2025년 조사에서 소비자의 60.8%가 AI 기술 발전을 긍정적으로 평가했으며, 특히 60대 이상에서 긍정 응답이 75.8%로 가장 높게 나타나 고연령층이 오히려 AI에 더 우호적임을 보여주었다. 그러나 2026년 조사에서는 AI 활용이 확대되는 한편 AI가 내리는 판단과 결정에 대해서는 상당한 경계심이 존재하는 것으로 나타났다.

소비자가 치명적이라고 생각하는 AI 소비자피해(1~2순위 기준)로 AI 할루시네이션으로 인한 오인·선택 왜곡이 56.8%로 1위를 차지했고, 책임소재 불명확(30.6%), AI 시스템 오작동으로 인한 신체·재산 피해(22.8%)가 뒤를 이었다. AI 생성 콘텐츠를 스스로 판별하기 어렵다고 응답한 비율도 44.6%에 달해, 소비자들이 정보 오염에 실질적으로 노출되어 있음이 확인되었다.

▶ 소비자는 AI를 활용하면서도 AI를 맹목적으로 신뢰하는 단계까지는 도달하지 않았다. 이는 단순 기술 보급보다 신뢰 기반 구축이 선행되어야 함을 시사한다.

3. 소비자가 가장 두려워하는 것은 '보이지 않는 차별'

소비자가 가장 우려하는 AI 문제(1~3순위 기준)는 허위정보·가짜뉴스(48.8%), 딥페이크 범죄 악용(46.5%), 개인정보 수집 및 유출(37.0%), AI 사고 시 책임 불명확(35.8%) 순이었다. 이 중 주목할 것은 AI 알고리즘 가격차별에 대한 우려다. AI가 소비자의 구매이력·검색기록·소득수준 등을 분석해 소비자마다 다른 가격·광고·서비스를 제공하는 것에 대해 41.6%가 부정적으로 평가했으며(긍정 21.9%), AI 알고리즘의 편향된 판단 가능성에 대해서도 57.9%가 우려한다고 응답했다.

과거 소비자문제는 가격을 알 수 있었지만, AI 시대에는 왜 내가 더 비싼 가격을 제시받았는지조차 알 수 없는 상황이 발생할 수 있다. 소비자들은 AI가 만드는 차별보다 그 차별이 '보이지 않는다'는 사실 자체를 더 두려워하고 있는 것이다.

▶ 또한 AI 유료서비스 비용 부담을 느끼는 비율이 35.6%로 만족도(31.3%)를 상회했으며, 응답자의 71.9%가 AI 유료화 확대가 디지털 격차를 심화시킬 수 있다고 우려했다. 과거의 디지털 격차가 인터넷 접근 여부였다면, 앞으로는 고성능 AI 접근 가능 여부가 새로운 사회적 불평등의 축이 될 수 있다.

4. 개인정보보다 더 중요한 것은 데이터 통제권

2025년 조사에서 소비자의 79.2%는 자신의 개인정보가 AI 서비스에 의해 수집·분석된다고 인식하고 있었다. 가장 민감하게 느끼는 정보는 얼굴·음성 등 생체정보(72.7%), 위치정보(53.2%), SNS 커뮤니케이션(46.5%), 검색·시청기록(43.1%) 순이었다. 2026년 조사에서도 40.8%가 약관이 어려워 개인정보 활용 내용을 충분히 확인하지 못한다고 응답했다.

소비자들이 가장 중요하게 생각하는 권리는 정보 이용 투명성(72.2%), 정보 통제권—수집된 정보 삭제·수정권(55.6%), 피해보상권(53.6%), 수집·활용 여부 스스로 결정할 권리(51.3%) 순으로 나타났다. AI 자동결정에 대해 소비자가 우선 보장받아야 할 권리(1~2순위)로도 개인정보 삭제 요구권(46.6%), 인간 재검토 요구권(41.8%), 집단구제권(36.7%)을 꼽았다.

▶ 소비자가 단순히 개인정보 유출을 걱정하는 것이 아니라, 자신의 데이터가 어디에 어떻게 활용되는지 알고 통제할 권리를 요구하고 있다. 이는 AI 시대 개인정보 보호의 패러다임이 '유출 방지'에서 '통제권 보장'으로 전환되

어야 함을 시사한다.

5. 소비자는 AI보다 '책임 없는 AI'를 두려워한다

조사 전반에서 나타난 가장 중요한 메시지는 '책임성'이다. 소비자는 AI가 완벽하다고 생각하지 않는다. 오히려 AI가 잘못 판단할 수 있다는 사실을 알고 있다. 문제는 피해가 발생했을 때 책임지는 주체가 불분명하다는 점이다.

AI 기술 관련 사고·피해 발생 시 기업·기관이 "AI 자동 판단"을 이유로 책임을 회피하거나 책임소재가 불명확해질 가능성이 있다고 응답한 비율이 82.5%에 달했다. 기업의 AI 기능·한계·위험성 설명 의무화가 필요하다는 응답도 90.2%로 나타났다. AI 챗봇 상담 문제 미해결 시 사람 상담원 연결권이 필요하다는 응답도 85.4%였다.

AI 의료서비스와 자율주행, 자동화 의사결정 분야에서 소비자들이 가장 중요하게 여기는 권리는 설명받을 권리·인간에 의한 재검토 권리·피해보상 권리였다. 현행 소비자보호 제도가 충분하지 않다는 응답도 62.9%에 달해, 기술 속도를 제도가 따라가지 못하고 있다는 소비자 인식이 명확히 드러났다.

▶ 이는 AI 시대 소비자정책의 핵심이 기술 진흥이 아니라 책임체계 구축에 있음을 의미한다.

6. 조사결과가 보여주는 정책적 함의

정부·국회가 우선 추진해야 할 AI 소비자보호 정책(1~3순위 기준)으로 소비자들은 딥페이크·허위정보 규제(52.8%), AI 책임기준 명확화(50.1%), 개인정보 보호 강화(35.5%), AI 피해구제 제도 마련(33.5%)을 꼽았다. AI 시대 소비자보호의 가장 중요한 원칙으로는 안전성과 책임성(55.3%), 인간 중심 원칙(28.6%), 개인정보 자기결정권(27.6%)을 선택했다.

1,000명의 소비자가 공통적으로 요구한 것은 AI 발전을 멈추라는 것이 아니었다. 소비자들은 AI가 의료혁신·재난대응·업무 효율 향상에 기여할 수 있다고 평가하면서도, 동시에 세 가지를 요구하고 있었다.

소비자가 걱정하는 것은 AI 자체가 아니라 설명 없는 AI, 차별하는 AI, 책임 없는 AI다.
AI 소비자보호 정책은 바로 이 지점에서 출발해야 한다.

Ⅴ. 지능정보사회(AI) 소비자 권리장전의 구체적 내용과 법적 의미

1. 소비자 권리장전의 제정 배경과 성격

한국소비자연맹은 2020년 2월 **「지능정보사회(AI) 소비자 권리장전」**을 제정하였다. 이 권리장전은 AI 기술의 개발·이용 과정에서 소비자 권익이 사후적으로 침해되는 현실에 대응하여, AI 윤리를 선언적으로 제시하는 수준을 넘어 소비자가 실제로 행사할 수 있는 권리의 기준을 구조화하려는 시도였다. 특히 이 권리장전은 기술 개발자나 사업자 중심의 AI 가이드라인과 달리 소비자법의 시각에서 '영향받는 주체'로서 소비자의 권리를 전면에 배치했다는 점에서 의의가 있다. 권리장전은 AI의 개발·설계·제공·이용 전 과정에 적용되는 소비자 중심의 규범적 프레임으로 기능한다.

2. 8대 소비자권리의 구체적 내용

(1) 포용성의 권리

포용성의 권리는 모든 소비자가 지능정보사회의 구성원으로서 AI가 창출하는 긍정적 가치를 배제 없이 누릴 권리를 의미한다. 이는 고령자, 장애인, 저소득층, 디지털 취약계층이 기술 발전 과정에서 구조적으로 배제되는 현실을 전제로 한다. AI 기반 서비스가 디지털 접근성, 이해 가능성, 이용 가능성을 충분히 고려하지 않는 경우, 기술은 편익이 아니라 새로운 차별의 수단이 될 수 있다. 포용성의 권리는 AI 정책 수립과 서비스 설계 단계에서 소비자의 다양한 욕구와 선호가 사후적 배려가 아니라 선제적으로 반영되어야 함을 요구한다.

(2) 공정성의 권리

공정성의 권리는 AI 기술이나 시스템이 소비자의 자유로운 선택을 부당하게 방해하지 않도록 공정하게 개발·적용·활용될 것을 요구할 권리이다. 이는 알고리즘 추천, 개인화 가격, 자동화된 의사결정과 같이 소비자의 선택과 거래 조건에 직접 영향을 미치는 AI 활용에서 특히 중요하다. 공정성의 권리는 소비자가 알지 못하는 사이에 특정 상품이 반복 노출되거나 불리한 조건이 적용되는 상황을 문제 삼는다. 이는 전통적인 소비자법상 부당한 고객 유인행위, 불공정거래 관행 규율을 AI 환경으로 확장하는 규범적 기준으로 볼 수 있다.

(3) 차별받지 않을 권리

차별받지 않을 권리는 AI가 성별, 연령, 장애, 소득, 지역, 건강 상태 등 소비

자의 개인적·사회적 특성을 이유로 부당한 차별을 야기하지 않아야 함을 의미한다. AI는 대규모 데이터를 학습하는 과정에서 기존 사회의 편견과 불평등을 그대로 재생산하거나 강화할 위험을 내포한다. 소비자 관점에서의 차별은 가격, 서비스 접근, 추천 결과, 계약 조건 등 다양한 방식으로 나타날 수 있으며, 이는 평등권과 소비자 선택권을 동시에 침해할 수 있다.

(4) 안전성과 신뢰성의 권리

안전성과 신뢰성의 권리는 AI의 개발·이용으로 인해 소비자의 생명, 신체, 정신, 재산에 위험이 발생하거나 발생할 우려가 있는 경우, 그 제거를 요구할 수 있는 권리이다.특히 의료 AI, 자율주행, 금융 판단, 정신건강·감정 분석 AI, 등 고영향 분야에서는
AI의 오류나 오작동이 즉각적인 피해로 이어질 수 있다. 이 권리는 AI 기술의 '혁신성'보다 소비자 안전을 우선하는 원칙을 명확히 선언한다는 점에서 중요하다.

(5) 투명성의 권리

투명성의 권리는 소비자가 자신에게 영향을 미치는 AI 기반 의사결정에 대해 알 권리, 설명을 요구할 권리, 이의를 제기할 권리를 가진다는 의미이다. 이는 AI 시대 소비자 권리 중 가장 핵심적인 권리로 평가된다. AI 환경에서는 왜 특정 상품이 추천되었는지, 왜 특정 가격이 적용되었는지, 왜 불리한 판단이 내려졌는지, 소비자가 알기 어렵다. 투명성의 권리는 AI의 '블랙박스성'을 이유로 소비자의 권리 행사가 봉쇄되어서는 안 된다는 점을 분명히 한다.

(6) 개인정보 통제권

개인정보 통제권은 AI 개발과 이용의 전 과정에서 소비자가 자신의 개인정보와 프라이버시를 보호받고, 필요한 범위 내에서 이를 통제할 권리를 의미한다. AI는 데이터에 의해 작동하는 기술이며, 소비자의 일상적 행동 데이터, 생체정보, 건강정보가 광범위하게 수집·분석·활용된다. 이 권리는 단순한 동의 여부를 넘어 어떤 데이터가 어떤 목적으로 어떤 방식으로 활용되는지 소비자가 실질적으로 통제할 수 있어야 함을 전제한다.

(7) 책임성의 권리

책임성의 권리는 AI를 이용해 제품이나 서비스를 제공하는 사업자에게 권리장전 이행을 위한 체계를 구축할 것을 요구하고, 이를 준수하지 않는 경우 책임

을 물을 수 있는 권리이다. AI 환경에서는 개발자, 플랫폼, 서비스 제공자가 분리되어 책임 전가나 공백이 발생하기 쉽다. 책임성의 권리는 “AI가 판단했기 때문에 책임질 수 없다”는 주장을 소비자법적으로 용인하지 않겠다는 선언이다.

(8) 피해구제 및 행동할 권리

마지막으로 피해구제 및 행동할 권리는 AI로 인해 소비자 피해가 발생한 경우 적절한 구제를 받을 권리와 다른 소비자와 연대하여 적극적으로 행동할 권리를 의미한다. AI 피해는 개별 소비자에게는 소액·단발적일 수 있지만 구조적으로는 대규모·반복적 피해로 이어질 가능성이 크다. 이 권리는 집단적 분쟁 해결, 소비자 참여형 감시, 사회적 대응의 정당성을 선언한다.

권리장전의 의의는 소비자단체가 전문가들과 함께 민간이 주도해 만들었다는 특징과 AI 윤리를 선언적 차원에 머물게 하지 않고, AI 환경에서 소비자가 실제로 행사해야 할 권리를 구조화했다는 점에 특징이 있다. 특히 투명성, 책임성, 피해구제 권리는 자동화·개인화된 소비환경에서 핵심적인 법적 기준으로 기능할 수 있다.

VI. AI소비자포럼 운영

AI소비자포럼은 소비자단체, 학계, 업계, 법조계, 정부, 전문가, 시민사회단체가 함께 참여하는 민간 거버넌스로서, AI 시대 소비자 보호의 방향을 논의하기 위해 25년 2월 발대식을 갖고 25년 6차례, 26년 1회 논의를 진행했다. 한국소비자연맹 강정화회장과 인공지능법학회 최경진 회장이 공동의장을 맡고 있고 30여명의 전문가가 참여하고 있다. AI소비자포럼은 AI는 사람을 위해 쓰여 져야 하고 규범은 기술이나 산업 중심이 아니라 기본권과 소비자 권리를 기준으로 설계되어야 함을 강조하고 있다.

AI소비자포럼은 소비자 실태조사 결과를 정책 논의로 연결하고, 소비자 관점이 법·제도 논의에 실질적으로 반영되도록 하는 중개 역할을 수행하고 있다. 이는 소비자 참여형 거버넌스의 하나의 모델로 평가할 수 있다. 26년에는 8대 소비자권리장전의 개별 항목에 대해 주체적인 실천방안을 만드는 적업을 진행할 예정이다.

Ⅶ. 결론

본 발제는 AI 기술의 확산이 소비자 보호의 대상과 방식을 근본적으로 변화시키고 있음을 전제로, AI 시대 소비환경의 구조적 변화와 이에 따른 소비자 도전과 새로운 위험에 대해 검토하였다.

AI는 더 이상 개별 기술이나 특정 서비스의 문제가 아니라, 소비자의 선택·정보 접근·거래 조건·책임 귀속 전반을 재구성하는 소비환경의 핵심 인프라로 작동하고 있다. 이 과정에서 소비자는 편익을 누리는 동시에, 알고리즘에 의해 매개된 선택, 생성·조작 가능한 정보 환경, 자동화된 의사결정, 그리고 분산된 책임 구조로 인해 기존 소비자법이 전제해 온 '통제 가능한 거래 주체'로서의 지위를 점차 상실하고 있다.

한국소비자연맹의 AI 소비자 실태조사 결과는 이러한 구조적 변화를 소비자가 어떻게 인식하고 있는지를 실증적으로 보여준다. 조사에 따르면 소비자는 AI 기술 자체를 거부하지 않으며, 의료·돌봄과 재난 등의 영역에서의 활용에 대해서는 오히려 높은 기대를 보였다. 그러나 AI 단독 판단에 대한 신뢰는 극히 낮았고, 사람의 감독과 책임이 결합된 구조에 대해서만 조건부 신뢰를 보였다. 또한 소비자가 가장 크게 우려하는 문제는 기술 오류 그 자체가 아니라, 딥페이크 범죄, 개인정보 통제 상실, 책임 부재와 같이 통제 불가능성과 피해구제의 어려움으로 나타났다. 이는 AI 시대 소비자 보호의 핵심 과제가 단순한 기술 규제가 아니라, 책임·투명성·구제 가능성의 제도화에 있음을 시사한다.

이러한 문제의식은 한국소비자연맹이 제정한 「지능정보사회(AI) 소비자 권리장전」과도 일관되게 연결된다. 권리장전은 포용성, 공정성, 차별받지 않을 권리, 안전성과 신뢰성, 투명성, 개인정보 통제권, 책임성, 피해구제 및 행동할 권리라는 8대 권리를 통해, AI 환경에서 소비자가 행사해야 할 권리를 선언적 차원을 넘어 규범적으로 구조화하였다. 이는 기존 소비자법의 기본 원칙을 AI 환경으로 확장하는 동시에, 자동화·개인화된 소비환경에서 새롭게 부각되는 권리 범주를 명확히 제시한다는 점에서 중요한 의미를 갖는다.

AI소비자포럼에서 강조하고 있는 AI 규범은 기술 진흥이나 산업 경쟁력 중심의 논의를 넘어, 기본권과 소비자 권리를 기준으로 설계되어야 할 필요성이 있다. AI의 특성상 개별 소비자가 그 작동 원리와 영향을 인식·통제하기 어려운 상황에서, 전통적인 사후 규제나 정보 제공 중심의 보호 방식은 한계에 직면한다. 이에 따라 AI 시대의 소비자 보호는 사후적 구제 중심에서 벗어나, 사전적·구조적 권리 반영과 거버넌스 설계로 확장되어야 한다.

이러한 맥락에서 소비자단체의 역할 역시 재정립될 필요가 있다. AI 시대 소비

자단체는 단순한 민원 처리 기관이나 이해관계자에 머무르지 않고, 첫째 실증 조사를 통해 보이지 않는 피해와 구조적 위험을 가시화하는 주체, 둘째 소비자 권리를 규범적으로 구조화하여 입법·정책 논의의 기준선을 제시하는 주체, 셋째 개별 피해를 사회적·집단적 문제로 전환하는 매개자, 넷째 소비자 참여를 제도화하는 거버넌스 설계자로서의 역할을 수행해야 한다. AI소비자포럼은 이러한 방향성을 구체화한 사례로, 소비자 참여형 AI 거버넌스의 가능성을 보여주고 있고 내년도에는 8대 소비자 권리장전에 대한 구체적인 실천방만을 마련할 예정이다.

결국 AI 시대 소비자 보호는 기술 혁신을 제한하기 위한 장치가 아니라, 그 혁신이 사회적으로 수용되고 지속 가능하게 작동하기 위한 신뢰의 인프라를 만드는 것이 중요할 것이다. 소비자법 역시 자동화된 의사결정에 대한 통제, 책임 귀속 기준의 재정립, 소비자 참여의 제도화를 통해 이러한 변화에 응답해야 한다. 소비자단체는 이 과정에서 소비자 경험과 실증을 바탕으로 규범과 제도를 연결하는 핵심적 역할을 담당하기 위해 노력하겠다.

1. 소비자 권리 관점에서 본 분야별 제도개선 과제

피해 유형	현행 법제도 공백	필요한 소비자 권리 및 제도
딥페이크·금융사기	전기통신금융사기법이 AI 생성 영상·음성 사기에 미적용	딥페이크 탐지 의무화 / 피해구제 특례 / 금융회사 AI 사기 대응 법적 의무 부과
AI허위정보·할루시네이션	AI 오정보에 대한 기업 책임 기준 없음	AI 정보 출처 표시 의무화 / 할루시네이션 발생 시 기업 배상 책임 기준 마련
AI워싱·사칭광고	표시광고법이 AI 기술 기만광고에 미적용 / 소비자 검증 불가능	AI 기능 최소 공시 기준 법제화 / 실제 AI 여부 제3자 검증 의무화
알고리즘 가격차별	개인별 가격 차별화 비교 자체가 불가능 /공정거래법 적용 기준 부재	알고리즘 가격 결정 투명성 공시 의무 / 의료·금융·주거 분야 가격 차별 금지 규정
AI차별·알고리즘 편향	알고리즘 설명 요구권 법적 근거 없음 / 간접 차별 입증 구조적 불가능	알고리즘 설명 요구권 입법화 / 편향성 정기 공시 의무 / 차별금지법에 간접 차별 포함
AI의료·자동화의사결정	AI 의료 오류 시 책임 귀속 기준 없음 / 현행 의료법 AI 서비스 비포함	개발사·병원·의사 3자 책임 배분 기준 법제화 / AI 자동심사 결과 인간 검토 요구권
청소년 AI 과몰입·정서의존	감정의존 유발 AI 챗봇 직접 규제 법령·전담기관 부재	미성년자 대상 AI 서비스 안전 기준 의무화 / AI임 고지 의무 / 자해·자살 표현 대응 의무

개인정보·AI학습 데이터	개인정보보호법이 AI 학습 목적 정보 활용 범위 불충분 규율	데이터 통제권 (열람·삭제·이전 요구권) 명문화 / 사전 고지·동의 절차 의무화
AI에이전트·자율 계약	AI 에이전트 계약 행위 법적 근거 전무 / 에이전트 간 충돌 시 손해 귀속 기준 없음	에이전트 위임 표준계약 제도화 / 무단 결제 자동환불 보호 / 취소·이의신청 절차 법제화
AI저작물 귀속	현행 저작권법 '인간 창작' 전제 / AI 개발사 약관으로 책임 면제	AI 결과물 저작권 귀속 기준 명확화 / 학습 데이터 공개 의무 / 저작권 위험 고지 의무
보험·신용평가 간접 차별	차별금지법이 알고리즘 간접 차별 미포함 / 금융당국 AI 가이드라인 구속력 없음	인간 검토 요구권·결정 이유 설명 의무 / 알고리즘 편향성 정기 공시 제도화

(AI시대 소비자 권리보호를 위한 과제, 정지연)

2. 소비자 권리 관점에서 본 3대 제도화 방향

투명성 보장 (AI임을 알 권리)	소비자가 자신에게 영향을 미치는 AI 시스템의 존재와 작동 방식을 알 수 있어야 함 ① AI 생성콘텐츠 워터마크·출처 표시 의무화 ② AI 챗봇임을 고지하는 의무 ③ 알고리즘 가격·추천 기준 공시 의무 ④ AI 기능 과장 광고 표시광고법 적용 강화
책임의 명확화 (설명받을 권리·이의제기권)	AI가 판단했다는 이유만으로 누구도 책임지지 않는 상황을 방지해야 함 ① AI 의료·금융·채용 등 고영향 분야 책임 귀속 기준 법제화 ② 자동화 의사결정에 대한 인간 검토 요구권 ③ 알고리즘 설명 의무화 ④ AI 오류 피해에 대한 입증 책임 전환 원칙 도입이 필요
실질적 피해구제 (구제받을 권리·집단행동권)	AI 피해는 개별 소비자에게는 소액이지만 구조적으로 대규모 반복 피해 ① AI 피해 전담 분쟁조정 기구 설치 ② 집단소송·징벌적 배상 도입 ③ AI챗봇 문제 미해결 시 사람 상담원 연결권 보장 ④ AI 피해구제 신고·처리 원스톱 창구 구축 필요

(AI시대 소비자 권리보호를 위한 과제, 정지연)

참고문헌

한국소비자연맹,「지능정보사회(AI) 소비자 권리장전」, 2020.
한국소비자연맹,「AI 소비자 이용실태 및 인식조사 결과보고서」, 2025.
한국소비자연맹,「AI 소비자 이용실태 및 인식조사 결과보고서」, 2026.
AI소비자포럼,「소비자가 걱정하는 AI 이슈」제2차 포럼 자료집, 2025.
최경진,「AI와 소비자 보호의 법적 과제」, 한국인공지능법학회 자료, 2025.
개인정보보호위원회, 공개된 개인정보의 AI 개발·서비스 목적 활용 안내서, 2024
한국지능정보사회진흥원(NIA), AI일상화 시대의 위협, 딥페이트 대응방안, 2024
법무법인 세종, 생성형AI 저작권 이슈 관련 최근 판결 동향, 2026
김앤장법률사무소, AI주요 이슈 및 시사점, 2024
OECD, Dark Commercial Patterns, OECD Publishing, 2024.
European Commission, Digital Fairness Fitness Check, 2023.
European Law Institute (ELI), Guiding Principles and Model Rules on Digital Assistants for Consumer Contracts (DACC), 2025
Mata v. Avianca, Inc., No. 22-cv-1461 (S.D.N.Y. 2023.06)

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]

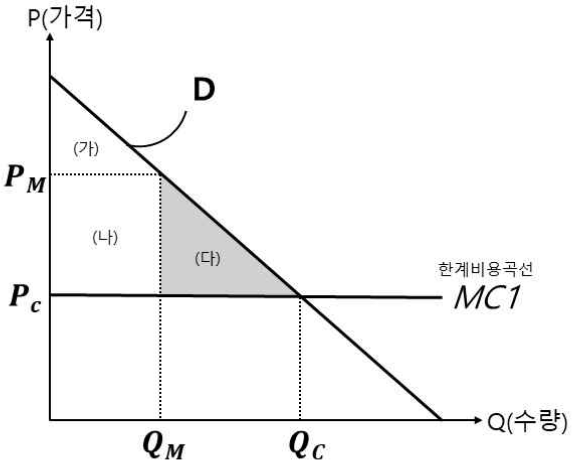

그림 1-5: 독점시장의 사회적 후생

(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.

AI시대, 소비자 권리 보호를 위한 과제 토론문

이정수(한국소비자단체협의회 사무총장)

1. 들어가며

- “AI시대, 소비자 권리 보호”라는 주제는 두 발제자가 지적하고 있듯이 매우 중요하고 시급하고 도전적인 과제임.
- 디지털·플랫폼 경제로의 급격한 전환으로 인한 변화에 적응하기도 쉽지 않은 상황에서 AI 기술까지 소비자의 일상생활 전반으로 깊숙이 들어와 이제는 소비환경의 기본적인 인프라가 된 상황에서 소비자, 소비자 권리는 어디까지 와 있는가?
- 인공지능기본법에 대한 논의와 더불어 현행 소비자기본법에 대한 전면적인 검토가 시급하게 이루어져야 할 것임.

2. AI시대, 소비자 권리 보호를 위하여

- 첫 번째 발제에서 AI시대의 소비자문제는 ‘정보 부족’ 문제가 아니라 ‘정보의 진위 판단의 부담’, ‘보이지 않는 차별’, ‘책임의 부재’로 정리하면서 딥페이크 사기, AI워싱, 가격차별, 추천 알고리즘, 자동화된 의사결정, 챗봇 오정보, 개인정보·학습데이터 문제 등에 대해 현장에서 일어나는 문제 중심으로 설명하면서 구조적 문제로 정리한 것이 매우 의미있다고 생각됨.
- 지난해 이어 두 번째로 실시한 AI 소비자 인식조사 또한 AI시대의 소비자의 생생한 현장의 목소리와 그들의 요구를 파악해 본 것이 시의적절하고 정책적으로 중요한 시사점을 얻을 수 있을 것이라고 봄.
- 소비자가 우려하는 것은 “AI 자체가 아니라 설명없는 AI, 차별하는 AI, 책임없는 AI”라는 조사 결과를 충분히 반영하여 수요자 중심의 정책이 이루어져야 함.
- 한국소비자연맹이 2020년 제정한 지능정보화 사회의 소비자권리장전에서 제시하고 있는 8대 소비자권리가 현행 소비자기본법상의 소비자의 기본적 권리를 포괄하면서 변화하는 소비환경을 반영하며 선언에서 그치지

않고 실질적으로 구현되기 위해서는 실제 분쟁상황 발생 시 작동하는 절차적 권리와 입증책임 완화가 뒤따라야 할 것임.

- 소비자 피해가 발생했을 때 개발사와 플랫폼, 사업자 간에 책임이 분산되는 구조적 문제를 개선하기 위해서는 책임 소재를 분명히 하는 것을 넘어서 소비자가 누구에게든 우선적으로 배상을 청구할 수 있도록 연대 책임 방식을 도입하는 것에 대해 고려해 보면 좋을 듯함.
- 두 번째 발제에서는 AI제품안전과 소비자보호를 하나의 체계로 묶으면서, 인공지능기본법의 한계와 EU AI Act의 가치사슬에 대한 책임구조를 설명하고 우리나라 법과 비교하여 현행법의 문제를 파악하고 개정 방향을 논의하는데 매우 유용하다고 생각됨.
- 개발사업자와 이용사업자의 이분법적 구분이 실제 AI 가치사슬을 반영하지 못한다는 지적, 그리고 제공자의 지위 전환, 배치자 책임, 범용AI 모델 규율에 대한 필요성은 우리나라의 입법 논의에 직접적으로 영향을 미칠 수 있는 중요한 시사점이라고 봄.
- 소비자단체 입장에서는 한 걸음 더 나아가서 책임주체를 넓히고 피해가 발생했을 때 피해자가 실제로 누구에게 어떠한 방법으로 피해구제를 청구하고 구제받을 수 있는지 분명하게 만들면 좋겠다고 생각됨.

3. 인공지능에 의한 가격차별화에 대하여

- 현행 소비자기본법 제4조에서는 1) 안전할 권리, 2) 정보를 제공 받을 권리, **3) 거래상대방, 구입장소, 가격 및 거래조건 등을 자유로이 선택할 권리,** 4) 의견을 반영시킬 권리, 5) 보상받을 권리, 6) 교육받을 권리, 7) 단체를 조직할 권리, 8) 안전하고 쾌적한 소비환경에서 소비할 권리 등 소비자가 이 8가지 권리를 기본적으로 가진다고 규정하고 있음.
- 제19조 사업자의 책무에서는 사업자는 1) 소비자의 생명, 신체 또는 재산에 대한 위해가 발생하지 않도록 해야 하며, **2) 물품등을 공급함에 있어서 소비자의 합리적인 선택이나 이익을 침해할 우려가 있는 거래조건이나 거래방법을 사용하여서는 안되며**, 3) 소비자에게 물품 등에 대한 정보를 성실하고 정확하게 제공해야하며, 4) 소비자의 개인정보가 분실·도난·누출·변조 또는 훼손되지 않게 해야하며, 5) 소비자의 불만이나 피해를 해결하거나 보상하며, 채무불이행으로 인한 소비자의 손해를 배상해야 한다고 규정하고 있음.

- 가격 자체의 개인화는 현행법의 사각지대에 있다고 발제자께서 지적하셨는데 현행 소비자기본법의 조항에서도 가격차별은 금지하고 있는 조항으로 해석할 수 없을까?
- 과도한 개인정보 수집, 개인화된 광고, 납치광고 등과 결합하여 소비자 소외, 차별, 양극화 등 심각한 소비자문제 및 사회문제를 양산하고 있고 심각성은 더욱 확대될 가능성이 큼.

4. 소비자가 바라는 AI

- 사전 예방적 설계
- 명확한 책임구조
- 실효적 피해구제(설명받을 권리, 인간 재검토 요구권, 데이터 통제권, 집단적 피해구제권)
- 소비자 분쟁해결 및 피해구제를 위한 구체적 방안 필요
- 취약계층에 대한 정책 마련(아동·청소년의 과몰입, 리터러시 등 역량격차 문제 해결을 위한 교육, 윤리와 활용에 대한 교육 등 필요)

- AI가 사람을 위한 기술이 되려면 소비자가 알고 선택하고 의견을 반영할 수 있고 피해를 구제받을 수 있어야 함.

AI 시대 소비자권리 보호를 위한 과제 토론문

서치원(녹색소비자연대 디지털소비자권리위원장)

1. “AI 시대 소비자운동의 새로운 과제”에 대한 토론

먼저 뜻깊은 발제를 해주신 정지연 사무총장님께 감사드립니다. 오늘 발제는 AI를 단순한 기술의 문제가 아니라 소비환경의 구조적 변화라는 관점에서 접근하고 있습니다. 딥페이크, 알고리즘 차별, AI 의료서비스, 개인정보 활용, AI 에이전트 등 현재 발생하고 있거나 앞으로 현실화될 다양한 소비자 문제를 폭넓게 정리하였고, 소비자 실태조사 결과를 바탕으로 소비자의 우려와 요구를 실증적으로 보여주었습니다.

나아가 소비자 피해의 구제방안 제시에만 머무르지 않고, AI 소비자 권리장전과 제도 개선 과제를 통해 향후 소비자정책이 나아가야 할 방향을 명확히 제시하고 있다는 점도 큰 의미가 있다고 생각합니다. 저는 발제자의 이러한 문제의식에 전적으로 공감하면서 세 가지를 말씀드리겠습니다.

① 13대 피해사례, 8대 소비자권리, 3대 제도화 방향의 연관성

발제문은 AI 시대 소비자 피해 사례를 매우 풍부하게 소개하고 있습니다. 딥페이크 범죄부터 AI 워싱, 알고리즘 가격차별, AI 의료서비스, AI 에이전트 계약 문제까지 소비자가 직면할 수 있는 위험을 폭넓게 보여주고 있습니다. 다만, 저는 발제문이 제시한 13대 피해사례, 8대 소비자권리, 3대 제도화 방향이 각각 설득력 있음에도 불구하고 서로 어떻게 연결되는지는 충분히 설명되지 않았다고 생각합니다.

AI 소비자권리장전은 원칙이고, 피해사례는 권리침해의 구체적 모습이며, 제도개선은 이를 실현하기 위한 수단에 해당할 것입니다. 예컨대 '딥페이크 사기'라는 피해사례는 8대 권리 중 '투명성의 권리'와 '안전성의 권리'가 침해된 구체적 모습이고, 이에 대한 제도개선 방향은 '딥페이크 탐지 의무화'와 '피해구제 특례'가 될 것입니다. 이처럼 피해사례→권리침해→제도개선의 연결 고리를 보여준다면 일반 소비자가 더 이해하고 실천하기 쉬울 듯 합니다.

다음으로 발제문에서 AI 피해 전담 분쟁조정 기구 설치에 관한 부분을 검토하고자 합

니다. 발제자께서 지적하신 바와 같이 AI 피해는 개인정보, 금융, 의료, 플랫폼, 광고, 계약 문제 등이 복합적으로 얽혀 나타납니다. 이에 대응해 새로운 기관을 만드는 것이 신속한 권리구제에 도움이 된다고 보는 견해도 있을 수 있습니다. 그러나 그것이 기존 소비자보호체계와 완전히 분리되어 별도로 운영되어서는 곤란할 것입니다. 새로운 기관을 만드는 것보다 기존 소비자보호시스템이 AI 문제를 다룰 수 있도록 역량과 전문성을 강화하는 것이 더 지속가능한 해결책일 수 있습니다.

물론 AI 분야에 특화된 전문조직이나 전문위원회는 필요할 수 있습니다. 이것을 기존 소비자보호시스템 내에 어떻게 조화시킬 것인가 고민하지 않고, 자칫 AI 문제만 별도 기관으로 분리한다면 기존 소비자정책 전반의 혁신과 역량강화는 지연되는 결과를 초래할 수도 있습니다. 저는 AI 소비자 문제를 별개의 특수한 영역으로 이해하기보다, 소비자 권리 전반을 다시 설계해야 하는 계기로 이해하는 것이 더 적절하다고 생각합니다. 그러한 점에서 AI 피해 전담 분쟁조정 기구 설치는 신중할 필요가 있습니다.

대안으로 생각해 볼 수 있는 방안은 한국소비자원 소비자분쟁조정위원회 내에 AI 전문위원회를 두고 AI 기술 전문가·법률가·소비자단체 대표가 참여하는 복합 전문성 체계를 구축하는 방안, 또는 개인정보보호위원회, 금융위원회, 공정거래위원회 등 관련 기관 간 협력 체계를 법제화하는 방안 등이 있을 것입니다.

② ‘보이지 않는 위험의 가시화’를 디지털 인프라의 물질적 측면까지 확장할 필요성

발제문은 소비자단체의 첫 번째 역할로 "보이지 않는 피해와 구조적 위험을 가시화하는 주체"를 제시합니다. 저 역시 이 역할에 적극적으로 동의합니다. AI 시대 소비자 문제의 상당수는 소비자가 인식하기조차 어려운 형태로 발생합니다. 알고리즘 차별도 그렇고, 개인정보 수집과 활용도 그렇고, 추천 시스템에 의한 선택 왜곡도 그렇습니다. 소비자단체는 이러한 보이지 않는 위험을 드러내고 사회적 의제로 만드는 역할을 해야 합니다. 다만 저는 여기서 한 걸음 더 나아갈 필요가 있다고 생각합니다.

우리는 AI를 흔히 비물질적 기술로 인식합니다. 스마트폰 화면 속에서 작동하는 서비스로 이해합니다. 그러나 실제 AI는 엄청난 물질적 기반 위에서 작동합니다. 챗GPT에 질문을 던질 때마다 데이터 센터에서는 엄청난 양의 물이 냉각수로 증발하고, 상상을 초월하는 전력이 소모됩니다. 최근에는 생성형 AI 확산으로 데이터센터 전력 사용량과 물 사용량이 급증하고 있다는 연구결과들이 잇따르고 있습니다. 우리가 마주

한 화면 배후에는 데이터 라벨링을 위해 시급 몇 센트를 받으며 착취당하는 제3세계 노동자들이 있고, 배터리와 반도체를 만들기 위해 황폐화되는 희귀 자원 채취의 현장이 있습니다. 빅테크와 AI 기업들은 이 환경부담비용과 노동착취를 철저히 비가시화, 외부화하고 있습니다.

하지만 대부분의 소비자는 자신이 사용하는 AI 서비스 뒤에 이러한 비용이 존재한다는 사실을 알지 못합니다. 마치 온라인 쇼핑이 물류창고와 택배노동을 가리고 있었던 것처럼, AI 역시 거대한 환경적·사회적 비용을 가리고 있습니다. 따라서 소비자단체가 앞으로 가시화해야 할 대상은 알고리즘의 블랙박스만이 아니라 AI(디지털인프라)의 물질적 블랙박스이기도 합니다. AI 서비스의 에너지 소비, 물 사용량, 탄소배출, 공급망 인권 문제 등을 소비자 관점에서 조사하고 공개하는 일 역시 중요한 소비자운동의 과제가 될 수 있습니다. 특히 지속가능한 소비를 강조해 온 소비자단체라면 AI 소비의 환경적 영향을 소비자의 알 권리 차원에서 제기할 필요가 있다고 생각합니다.

우리는 그동안 AI의 소비자문제로 제품의 안전성, 표시광고의 진실성, 개인정보 보호를 이야기해 왔습니다. 앞으로는 AI 서비스가 소비하는 물과 전기, 그리고 그 과정에서 발생하는 환경비용 역시 소비자의 알 권리에 포함되어야 한다고 생각합니다. 구체적으로는 소비자기본법 제4조 제1호의 "물품 또는 용역으로 인한 생명·신체 또는 재산에 대한 위해로부터 보호받을 권리"와 제4조 제2호의 "물품 등을 선택함에 있어서 필요한 지식 및 정보를 제공받을 권리"를 근거로, AI 서비스의 환경적 비용 정보 제공이 소비자의 법적 권리에 포함된다고 생각합니다.

그 외에도 소비자 행동으로 생각해 볼 만한 것들이 있습니다. 예를 들면, *디지털 소비의 탄소배출량 가시화 : 온라인 쇼핑의 배송 방식이나 스트리밍 서비스 이용 등 각종 디지털소비에 따른 환경 영향을 표시하는 ‘디지털 탄소 라벨’ 도입을 생각해 볼 수 있겠습니다. **순환경제 기반의 디지털 기기 소비 운동 : 디지털 기기의 내구성을 높이는 설계를 요구하고, 수리해서 쓸 권리를 활성화하며, ‘공정무역’ 인증제도와 유사하게 디지털 기기 제품 생산 원재료에서부터 더 좋은 소비를 고민할 수도 있을 것입니다.

③ 소비자단체는 기술의 민주적 통제를 위한 핵심 주체

발제문은 소비자단체의 역할로 피해의 가시화, 권리의 구조화, 사회적 문제화, 거버넌

스 설계를 제시하고 있습니다. 매우 중요한 역할입니다. 이 중 피해의 가시화는 앞서 말씀드렸고, 권리의 구조화, 사회적 문제화는 그동안 소비자단체가 꾸준히 해왔으며, 많은 성과를 이루었다 생각합니다.

AI는 단순한 상품이나 서비스가 아니라 사회 전반의 의사결정 구조를 바꾸는 기술입니다. AI는 무엇을 살 것인지 결정하는 데 영향을 미칠 뿐 아니라, 어떤 정보를 접할 것인지, 어떤 보험에 가입할 수 있는지, 어떤 대출을 받을 수 있는지, 나아가 어떤 치료를 받을 수 있는지에도 영향을 미칩니다. 즉 AI는 소비자문제를 넘어 민주주의와 기본권의 문제로 확장되고 있습니다. 이에 소비자단체는 AI 규범과 정책을 논의하는 과정에 소비자의 목소리를 반영하고, 알고리즘의 투명성과 책임성을 요구하며, 기술이 인간의 통제를 벗어나지 않도록 감시하는 역할을 수행해야 합니다. 발제문에 소개된 AI소비자포럼 역시 이러한 가능성을 보여주고 있다고 생각합니다.

발제문에 소개된 실태조사를 통해 확인할 수 있는 바와 같이 소비자가 두려워하는 것은 설명 없는 AI, 차별하는 AI, 책임 없는 AI입니다. 그렇다면 소비자단체는 “이해할 수 있게 설명해주는 AI, 차별없고 공정한 AI, 책임지는 AI”를 만들기 위해 노력해야 합니다. AI가 두려움의 대상이 아니라 함께 살아가는 존재로 받아들여지려면 결국 신뢰성이 관건입니다. 신뢰성은 기술발전이나 전문가집단 주도의 제도설계만으로는 달성할 수 없습니다. 기술에 대한 사회적 통제, 민주적 통제가 가능하다는 참여와 경험이 축적될 때에만 가능합니다.

따라서 앞으로는 AI 거버넌스에 소비자와 소비자단체가 제도적으로 참여할 수 있는 구조를 보다 적극적으로 만들어갈 필요가 있습니다. 기술이 누구를 위해 존재해야 하는지, 기술 발전의 방향을 누가 결정해야 하는지에 대한 사회적 논의에 소비자의 관점을 반영하는 것입니다. 예컨대 소비자단체는 AI 영향평가 과정에 참여하고, 알고리즘 감사제도 도입을 요구하며, AI 정책 수립 과정에서 소비자 대표성을 확보하고, 시민이 참여하는 AI 공론장을 운영하는 역할을 수행할 수 있습니다. 소비자운동은 더 이상 피해가 발생한 뒤 구제를 요구하는 운동에 머무를 수 없습니다. 기술이 설계되는 단계부터 소비자의 관점을 반영하는 운동으로 확장되어야 합니다. 저는 이것이 앞으로 소비자단체가 수행해야 할 가장 중요한 역할이라고 생각합니다.

AI는 새로운 소비자문제를 만들어내고 있지만, 동시에 소비자운동이 스스로를 재정의할 기회도 제공하고 있습니다. 과거 소비자운동이 제품의 안전성과 표시광고의 진실

성을 요구했다면, 이제는 알고리즘의 투명성과 데이터의 통제권, 그리고 기술의 민주적 통제를 요구해야 하는 시대가 되었습니다. 그런 의미에서 오늘 발제는 단순히 AI 소비자피해를 다룬 보고서가 아니라, 앞으로 소비자운동이 어디로 나아가야 하는지를 묻는 문제제기라고 생각합니다. 좋은 발제를 준비해주신 발제자께 다시 한번 감사드립니다.

2. “AI 시대 소비자 보호 법제의 과제와 전망”에 대한 토론

먼저 뜻깊은 발제를 해주신 서종희 교수님께 감사드립니다. 오늘 발제는 「인공지능 제품안전과 소비자보호」라는 주제를 소비자 보호와 책임법제의 관점에서 매우 체계적으로 다루고 있습니다. 특히 현행 인공지능기본법이 AI 산업의 복잡한 가치사슬 구조를 충분히 반영하지 못하고 있다는 점을 지적하면서, EU AI Act의 제공자(provider)-배치자(deployer) 체계와 책임 전환 규정을 참고하여 보다 정교한 책임체계를 구축해야 한다고 제안한 점은 중요한 시사점을 제공합니다.

또한 발제문은 인공지능이 야기하는 사고가 기존 제조물 사고와 달리 자율성, 불투명성, 지속적 학습이라는 특성을 가진다는 점에 주목하여, 고위험 AI에 대한 무과실책임 도입, 입증책임 완화, 책임보험 및 공적 보상기금 제도 마련 등 소비자 친화적인 피해구제 체계 구축의 필요성을 설득력 있게 제시하고 있습니다. 나아가 범용 AI 규율, 역외 적용, AI 기반 가격차별 규제 등 현재 우리 법제에서 충분히 수용하지 못한 쟁점들을 빠짐없이 검토하였다는 점에서도 의미가 크다고 봅니다.

저 역시 기본적으로 발제자가 제안해주신 내용과 제도 개선 방향에 전적으로 동의합니다. 다만 실제 입법과 제도 설계의 단계에서는 몇 가지 추가적인 검토가 필요하다고 생각합니다. 이하에서는 AI 가격차별 규제의 현실적 입법 경로, 제품안전과 제조물책임법의 한계, 고위험 피지컬 AI에 대한 사전예방적 규제체계의 가능성과 한계라는 쟁점을 중심으로 말씀드리고자 합니다.

① AI 가격차별 규제의 형식과 내용에 동의하지만 AI 기본법 개정은 검토의 여지가 있습니다.

발제문은 AI가 소비자의 구매이력, 검색기록, 위치정보, 사용기기 등을 분석하여 사실상 제1종 가격차별에 가까운 개인화 가격결정을 가능하게 한다는 점을 지적합니다.

다시 말해 기존 가격차별은 학생할인이나 경로우대처럼 누구나 차별의 기준을 알 수 있었지만, AI 가격차별은 개별 소비자가 자신에게만 더 높은 가격이 제시되었는지조차 인식하기 어렵다는 특수성이 있습니다. 따라서 문제는 단순한 가격 차이가 아니라 정보 비대칭을 넘어선 인지 비대칭에 있습니다. 이때 소비자의 선택권은 형해화됩니다.

가격차별 문제는 향후 소비자보호 분야에서 가장 중요한 쟁점 중 하나가 될 가능성이 높습니다. 현재 온라인 플랫폼 분야에서도 동일한 문제가 이미 제기되고 있습니다. 그런데 플랫폼 사업자가 소비자 특성에 따라 가격, 노출순위, 할인혜택 등을 차별적으로 제공하는 문제는 수년 전부터 입법 논의가 있었지만 아직까지 명확한 규제체계는 마련되지 못하고 있습니다. 물론 온라인 플랫폼의 가격차별 문제는 알고리즘을 매개하되 어디까지나 인간이 설계한 알고리즘에 의한 결과로, AI의 자율성, 적응성의 결과물이 아니라는 점에서 구별되는 지점이 있습니다.

한편, 가격차별의 문제는 AI 자체의 문제라기보다 AI를 활용한 시장행위의 문제라는 성격을 갖습니다. 다른 영역의 차별금지 요청이 설계 당시 고려하지 않은 데이터 자체에 내재하는 편향과 오류에 기반한 것이라면, 가격차별은 설계 당시부터 의도하고 고려한 결과라는 점에서 일반적인 차별금지 규범과는 구별되는 특성이 있습니다. 따라서 인공지능기본법에 직접 규정을 두기보다는 전자상거래법, 표시광고법, 온라인플랫폼법 등 소비자거래법 체계에서 접근하는 것이 입법기술상 더 적절할 수도 있겠다는 생각이 듭니다.

인공지능기본법은 기본적으로 AI 기술 자체의 개발·이용에 관한 규율을 목적으로 하는 반면, 가격차별은 AI를 활용한 시장행위의 문제이므로 소비자거래 규율 체계에서 접근하는 것이 규범 목적과의 정합성 측면에서 더 적절하기도 합니다. 또한 인공지능기본법은 현재 '진흥법'적 성격이 강하여 위반 시 제재가 3,000만 원 이하 과태료에 그치는 등 실질적 준수 유인이 부족한 반면, 표시광고법이나 전자상거래법은 시정명령, 과징금 등 보다 실효적인 집행 수단을 갖추고 있다는 점도 고려해야 합니다.

당장 인공지능기본법에 가격차별 금지를 직접규제하기는 어렵다 하더라도 최소한 개인화 가격 적용 사실에 대한 ‘고지 의무'와 '독립적 제3자 기관에 의한 가격 알고리즘 정기 감사(Audit)’는 도입되어야 합니다. 소비자는 자신이 개인화 가격의 적용 대상인지 알 권리가 있으며, 왜 그러한 가격이 제시되었는지 설명을 요구할 권리가 있기 때

문입니다.

② 제품안전사고 관련 현행 제조물책임법의 한계를 극복할 방안이 필요합니다.

발제문은 EU 새 제조물책임지침안이 소프트웨어와 인공지능시스템을 제조물에 명시적으로 포함시킨 것처럼, 우리 제조물책임법도 '제조되거나 가공된 동산'(제조물책임법 제2조 제1호)의 정의를 확대하여 인공지능시스템(소프트웨어)을 포함시켜야 한다고 제안합니다. 이 제안에 동의합니다. 그런데 소프트웨어를 제조물로 보더라도 입증 문제는 여전히 남습니다.

현행 제조물책임법 제3조의2는 피해자가 해당 제조물이 “정상적으로 사용되는 상태”에서 피해자의 손해가 발생하였다는 사실 등을 증명한 경우 결함과 인과관계를 추정하도록 규정하고 있습니다. 문제는 AI 소프트웨어의 경우 '정상적으로 사용되는 상태'를 소비자가 어떻게 입증할 수 있는가 하는 것입니다. 지금도 소비자는 자동차 급발진 사고에서 입증의 벽을 넘지 못하고 있습니다. 만약 앞으로 휴머노이드 로봇이나 자율주행차가 스스로 학습하고 판단하는 시대가 된다면 소비자의 입증 부담은 지금보다 훨씬 더 커질 수 있습니다. AI 시스템의 경우 그 작동 과정이 블랙박스 특성을 지니고 있어 소비자가 이를 입증하기가 극히 어렵기 때문입니다.

이 문제에 대한 보완책으로 다음 두 가지를 검토할 수 있습니다. 첫째, 사실상의 추정을 통한 해석론적 접근으로, AI 소프트웨어 사고의 경우 소비자가 통상적인 방법으로 AI 시스템을 이용하였다는 점만 증명하면 정상적 사용을 사실상 추정하는 것입니다. 둘째, 발제자께서 제안하신 바와 같이 EU AI Act가 제공자에게 기술문서와 기록유지를 의무화한 것처럼 우리도 인공지능기본법에 AI 시스템의 작동 로그 기록·보존 의무를 명시하고, 소비자가 피해를 주장하는 경우 해당 로그에 대한 접근권을 법적으로 보장하는 방식을 검토할 수 있습니다. 어느 쪽이든 소비자가 '정상적 사용'을 직접 입증하는 부담을 덜고, 사업자가 '비정상적 사용'이었음을 반증하도록 입증 구조를 전환하는 방향입니다.

③ 고위험 피지컬 AI 시대에는 책임보험보다 사전예방 체계가 더욱 중요합니다.

발제문은 고위험 AI에 대한 책임보험과 공적 보상기금 제도를 제안합니다. 피해구제 측면에서 매우 의미 있는 제안입니다. 그러나 피지컬 AI가 확산될 미래를 고려하면

사후보상만으로는 충분하지 않을 수 있다는 점을 지적하고자 합니다. 지금까지 AI 규제 논의는 개인정보, 차별, 허위정보 등 디지털 공간의 위험에 집중되어 있었습니다. 그러나 앞으로는 AI가 물리적 세계에서 행동하는 '피지컬 AI'의 시대가 도래하고 있고, 소비자 보호의 관점에서도 위험의 중심축이 정보 위험에서 물리적 위험으로 이동하고 있다는 점에 주목할 필요가 있습니다.

최근 논의되는 휴머노이드 로봇, 자율주행차, 돌봄로봇, 물류로봇 등은 AI의 판단이 현실 공간에서 즉시 물리적 행동으로 표출됩니다. 디저털 AI의 오류는 근본적으로 출력 정보의 오류에 불과하지만, 피지컬 AI의 오류는 직접적인 생명·신체 피해를 초래할 수 있습니다. 따라서 피지컬 AI는 대표적인 고위험 AI로 볼 수 있으며 사후배상보다 사전예방이 더욱 중요합니다. 이와 관련하여 의료기기 인허가 체계와 유사한 위험성 평가 및 인증제도를 검토할 필요가 있습니다.

다만 여기서 고민해야 할 점은 규제의 한계입니다. 인증제도는 결국 시장 진입을 제한하는 효과를 가지므로, 과도한 사전규제는 혁신을 저해할 수 있고, 반대로 지나치게 완화된 규제는 소비자 안전을 위협할 수 있습니다. 따라서 모든 AI에 대한 사전허가 방식보다는 위험기반 접근이 바람직하다고 생각합니다. 예컨대 일반적인 생성형 AI는 사후규제를 중심으로 하고, 자율주행차·돌봄로봇·의료 AI·산업용 로봇 등 생명·신체 위해 가능성이 높은 피지컬 AI에 한해 위험성 평가, 적합성 인증, 정기 안전점검, 사고보고 의무 등을 단계적으로 부과하는 방식이 현실적일 것입니다.

이것은 발제자께서 제안하신 무과실/과실 책임의 2분체계와도 상응하는 면이 있습니다. 사전예방 체계와 사후책임 체계의 정합성을 확보하기 위해서는 다음과 같은 연결구조를 검토할 수 있습니다. 첫째, 위험성 평가를 통과하여 적합성 인증을 받은 피지컬 AI가 사고를 일으킨 경우에는 과실책임을 적용하되, 인증 기준을 준수하였다는 사실이 과실 부존재의 유력한 증거가 되도록 설계하는 방안. 둘째, 위험성 평가를 받지 않거나 인증 기준을 위반한 피지컬 AI가 사고를 일으킨 경우에는 무과실책임을 적용하여 사전예방 의무 위반에 대한 강력한 유인을 제공하는 방안. 이를 통해 사전예방 체계와 사후책임 체계를 유기적으로 연결하여 전체 규제 체계의 정합성을 높일 수 있을 것입니다.

오늘 발제는 AI 시대 소비자 보호 법제가 나아가야 할 방향을 매우 설득력 있게 제시하고 있습니다. 특히 가치사슬 기반 책임체계, 입증책임 완화, 책임보험 및 공적 보상

기금 도입이라는 큰 방향에는 저 역시 전적으로 공감합니다. 다만 앞으로의 입법 과정에서는 AI 가격차별을 어떠한 법체계에서 규율할 것인지, 소비자의 입증곤란 문제를 어떻게 해결할 것인지, 그리고 피지컬 AI 시대에 사후책임과 사전예방을 어떻게 조화시킬 것인지에 대한 추가 논의가 필요하다고 생각합니다. 소비자가 안심하고 기술을 이용할 수 있는 신뢰의 기반을 만드는 것이야말로 지속가능한 AI 발전의 전제일 것입니다. 오늘 토론이 작은 도움이 되길 바랍니다.

AI 시대 소비자보호를 위한 입법 과제

참여연대 민생희망본부 실행위원 한경수 변호사

1. 전환? 방향이 중요하다.

이재명 정부는 5년간 경제정책 청사진을 담은 '새 정부 경제성장전략', 즉 '기술선도 성장', '모두의 성장', '공정한 성장', '지속성장 기반 강화' 등 4대 축으로 구성된 경제 대혁신으로 '진짜 성장'을 구현하겠다는 비전을 제시한 바 있다. 이중에서 '기술선도 성장'을 위해 'AI 3대 강국(G3) 도약'을 핵심 국정목표로 삼고, 주요 정책 방향으로 ① AI 고속도로 구축 및 인프라 확보, ② 전 국민 무료 AI 서비스, ③ AI 미래기획수석 신설, ④ AI 기본사회 추진 계획을 밝힌 바 있다. '모두의 성장'에는 지역균형성장, 중소기업·소상공인·민생경제 지원 방안이 담겼고, '공정한 성장' 분야에서는 납품 대금 연동제 대상을 확대하고 기술 탈취 대응을 강화하는 내용이 담겨 있으며, '지속성장 기반 강화'를 위해선 규제 완화를 추진한다고 한다. 모두 그럴 듯해 보이는 단어들이다. 그러나 여기에는 '기술 혁신'과 '정부 지원'은 있으나, '소비자'는 보이지 않는다.

새로운 기술이 단순한 신기술에 그치지 않고 사회 구조를 바꾸는 진정한 혁신으로 거듭나기 위해서는, 산업혁명 이후 지금에 이르기까지 수많은 기술 혁신 사례에서 알 수 있는 바와 같이 정부의 제도적 견인과 최종 소비자의 시장 수용이 반드시 결합되어야 한다. 정부의 적극적인 마중물 역할은 반드시 필요하지만, 거기에서 그쳐서는 안되며 반대편을 반드시 살펴봐야 한다. 현재 주목받는 자율주행 기술이 대표적인 사례이다. 자율주행이 우리 삶에 안정적으로 안착하려면 정부가 차세대 지능형 교통시스템(C-ITS) 같은 정밀 도로 인프라를 조성하고, 사고 책임 소재에 대한 법적 기준을 선제적으로 정비해 주어야만 한다. 하지만 제도와 인프라가 갖춰지더라도 혁신의 마침표를 찍는 것은 결국 최종 소비자의 신뢰와 소비 행위이다. 테슬라를 비롯한 자동차 회사들이 자율주행 레벨 2.5에 멈추고 있는 이유는 기술적인 이유보다는 사고가 발생했을 때 "누가 어떻게 책임을 질 것인가"에 대한 법적 책임 분담과 소비자 보호 장치가 명확하지 않기 때문이다. 즉, 소비자 보호를 위한 입법 방안이 마련되지 않고서는 AI가 비로소 사회 체제를 바꾸는 진정한 변혁으로 완성될 수는 없다.

2. AI 산업 진흥과 소비자 보호를 어떻게 설계할 것인가?

가. 인공지능기본법의 입법 목적

인공지능기본법 제1조는 "이 법은 인공지능의 건전한 발전과 신뢰 기반 조성에 필요한 기본적인 사항을 규정함으로써 국민의 권익과 존엄성을 보호하고 국민의 삶의 질 향상과 국가경쟁력을 강화하는 데 이바지함을 목적으로 한다."고 규정하고 있다. 즉, 현행 인공지능기본법은 그야말로 '진흥법'적 성격이 강하다. 그러므로 인공지능을 이용하는 소비자를 보호하기 위한 규정은 사실상 없는 셈이다.

나. 소위 '신산업'에 대한 규제체계

인공지능기본법과 같은 규제체계는 새로운 문제가 아니다. 한국의 입법 경험상 신산업·신기술이 등장할 때 정부는 우선 산업 진흥을 위한 기본법을 제정해 왔다가 최근에는 진흥과 안전을 별도로 입법화하고 있다.

(1) 원자력 분야

1958년 제정된 「원자력법」은 원자력의 연구·개발·이용 진흥과 안전관리를 하나의 법률 안에서 함께 규율하였으나, 원자력 진흥을 담당하는 기관이 동시에 안전관리까지 담당함으로써 안전규제의 독립성이 취약하다는 비판을 받았다. 그 후 1969년에 원자력손해배상법이 제정되었고, 2011년에는 '원자력진흥법'과 '원자력안전법'으로 분리되었다.

현재 인공지능기본법은 과학기술정보통신부 중심의 진흥 중심 체계를 갖추고 있으나, 안전과 보호에 대한 체계는 사실상 전무하다. 그러므로 진흥과 안전·보호를 담당하는 법률과 조직이 분리되어야만 실효적인 견제와 균형이 작동될 수 있는데, 현재는 안전과 보호 체계가 사실상 없다.

(2) 금융 분야: 금융혁신지원 특별법(2019) vs 금융소비자보호법(2021)

종전에는 원자력 분야처럼 진흥법이 입법된 후 이를 분리하여 별도의 보호법을 입법하는 형식을 취했으나, 21세기 들어서는 진흥법과 보호법을 별도로 입법하는 사례가 나타나고 있다. 정부는 핀테크 등 금융혁신을 촉진하기 위하여 2018년「금융혁신지원 특별법」을 제정하였고 이 법은 혁신금융서비스에 대한 규제 샌드박스(규제 특례)를 핵심 내용으로 한다. 또한, 정부는 혁신금융서비스로 인해 소비자에게 피해가 발생했을 때 이를 구제하기 위하여 2020년에 금융소비자보호법을 제정하였다. 즉 '혁신을 지원하는 법'과 '소비자를 보호하는 법'이 별도의 트랙으로 동시에 존재하면서 상호 보완하는 구조이다.

(3) 모빌리티 분야

자율주행자동차 분야는 피지컬 AI의 중요 분야 중 하나이다. 자율주행자동차법은 그 명칭에서부터 '상용화 촉진 및 지원'을 목적으로 함을 명시하여 2019년에 제정하고 2020년부터 시행되었다. 한편 자율주행자동차의 안전기준은 별도로 '자동차관리법' 및 그 하위규정인 '자동차 및 자동차부품의 성능과 기준에 관한 규칙'에서 규율하고, 사고 발생 시 소비자 보호 문제는 '자동차손해배상 보장법'과 '제조물책임법'이 분담하는 구조를 갖고 있다. 그러나 소비자 보호 문제를 규율하는 '자동차손해배상 보장법'과 '제조물책임법'상 AI 소프트웨어 결함에 대한 제조물책임법 적용 가능성과 자율주행 시스템 운행자의 운행지배성 감소 또는 소멸 시 책임 분배 문제가 해결되지 않아 오히려 자율주행자동차의 기술적 진보가 가로막혀 있다는 평가를 받고 있기도 하다.

다. 설계 방향

위 3가지 분야 사례에서 살펴본 바와 같이, 현재 기준으로는 산업의 진흥과 안전·보호를 하나의 법에서 규율하지는 않는 경향이다. 인공지능기본법 역시 진흥법의 성격을 갖는 이상 인공지능기본법에 안전과 보호 체계를 담는 것이 적절하지도 않다. 그러므로 AI 분야에서도 인공지능기본법은 산업 진흥과 기본적인 신뢰 기반 조성에 집중하고, 소비자(이용자·영향받는 자) 보호를 위한 실효적 권리·구제 수단은 가칭 'AI 소비자보호법' 또는 기존 개별 소비자법(전자상거래법, 표시광고법, 제조물책임법 등)의 개정을 통해 보완하는 방식이 필요하다.

3. 일반법 VS 개별법?

가. 인공지능기본법의 제정과 개정

인공지능기본법은 2025. 1. 21. 제정되어 다음 날부터 시행되었다. 그리고 그로부터 1년 후인 2026. 1. 20. 일부 개정되어 같은 달 22일부터 시행되고 있다. 인공지능기본법은 산업 진흥법의 성격이 강화되고 있으나, 기본적으로 서종희 발제문에서 지적하고 있는 것처럼 인공지능사업자를 '인공지능개발사업자'와 '인공지능이용사업자'로 이분하는 단순 구조를 취하고 있어 데이터 제공자·미세조정자·통합배치자·서비스운영자 등 다층적 가치사슬 행위자를 포섭하지 못하고 있고, '이용자'는 단순히 '제공받는 자'로 정의되어 최종적·수동적 수령자로만 취급된다는 점에서 한계를 갖고 있다.

또한, 인공지능기본법에는 최근 가장 큰 사회적 문제가 되고 있는 알고리즘에 대한 설명요구권이 없고, 자동화 의사결정에 대한 인간 재검토 요구권[1], 딥페이크 생성 금지, 가격차별 투명성 등 두 발제문이 공통적으로 요구하는 핵심적인 소비자 보호 규정을 두고 있지 않다.

나. 현재 소비자보호 기능을 분담하는 개별 법률 현황

현재 한국의 AI 관련 소비자 피해는 기존 개별 법률들이 단편적으로 담당하고 있다. 그러나 이들 법률은 모두 기존의 거래질서를 중심으로 설계된 것이고 AI 시대를 규율하기에는 입법의 공백 영역이 매우 큰 상태이다.

먼저 전자상거래법 제21조의2는 다크패턴 일반을 규율하고 있으나, AI 기반 개인화 가격차별이나 초개인화된 기만에 대해서는 적용 범위가 명확하지 않다. 표시광

1) 이번 미국-이란전은 미국이 사실상 AI로 치른 전쟁이다. 이 과정에서 AI는 정찰부터 타격 목표 설정, 방어망 통제 등 대부분의 업무를 수행하였으나, AI 알고리즘 오류로 인해 민간인 오폭 사고가 발생하는 등 심각한 부작용도 있었다. 결국 최종 판단은 AI가 아니라 인간이 할 수 있어야만 알고리즘의 비인간성을 조금이라도 예방할 수 있을 것이다. 또한, 잘 알려진 바와 같이 영국 킹스칼리지 런던 연구팀이 주요 최신 AI모델을 가상 국가 지도자로 설정해 진행한 모의전쟁 시뮬레이션 결과 AI가 95%의 확률(21번 중 20번)로 핵무기를 발사하였다.

고법은 허위·과장광고를 규율하지만, 'AI워싱'과 같이 기술 자체의 불투명성에 기반한 기만행위에는 적용이 한계에 부딪힌다. 제조물책임법은 제조물의 결함에 대한 책임을 규정하고 있으나, '제조되거나 가공된 동산'이라는 정의에 소프트웨어인 AI가 포함되는지가 불명확하다. 개인정보보호법은 개인정보 처리 전반을 규율하지만, AI 학습 목적의 정보 활용에 대한 사전 동의 절차가 미비하다. 전기통신금융사기법은 전화·문자 등을 통한 금융사기에 적용되지만, 딥페이크 영상·음성을 이용한 사기에는 적용이 곤란하다. 마지막으로 차별금지법제는 직접차별을 규율하는 데 그쳐, AI가 간접 변수를 통해 발생시키는 '간접 차별'은 포괄하지 못하고 있다. 이처럼 개별 법률들은 각자의 영역에서 AI로 인한 소비자 피해의 일부만을 단편적으로 포착할 뿐이며, 그 사이사이에 적용 공백이 폭넓게 존재한다.

다. AI 소비자보호법 제정으로 충분한가?

(1) 22대 국회에서는 안철수 의원안·황희 의원안 등 다수의 '인공지능책임법안'이 제출되어 AI로 인한 손해에 대한 민사책임 특례를 규정하고자 하였다. 그러나 이들 법안은 소비자보호를 위한 실질적인 규정(제조물책임법과의 관계, 입증책임 완화 등)이 없어 미흡하다는 평가를 받는다. 또한, AI가 거래의 주체가 되는 경우는 거의 없으며, AI는 거래과정에서 다양한 방식으로 역할을 하지만 거래 방식 자체는 기존의 거래방식으로 이루어진다는 점에서, AI소비자보호법이라는 별도의 일반법을 제정하더라도 적용 대상이 되는 'AI 개입 거래'의 범위를 정하는 것부터 난제가 된다. 결국 일반적인 선언적 원칙 외에 구체적인 소비자 피해보상절차를 규정하는 것이 쉽지 않다는 점에서, 일반법 제정 방식의 실효성도 충분하다고 보기 어렵다. 오히려 기존의 거래유형별 규율체계와 정합성을 유지하면서 그 틀 안에서 AI 관련 규정을 보완하는 개별법 정비 방식이 체계상 더 안정적이다.

(2) 그러므로 소비자기본법 개정은 필요하다. 현재의 소비자기본법에는 AI가 개입된 또는 다양한 역할을 하는 거래 과정에 대한 규정은 존재하지 않는다. 그러므로 소비자기본법에 별도의 장을 신설하여 소비자에 대한 AI 책임기준 등을 신설하고, 소비자원에서 AI 피해구제 신고·처리 원스톱 창구를 구축하며, AI로 인한 분쟁에 대해서도 (집단)분쟁조정제도 또는 필요하면 중재제도를 활용할 수 있도록 하여야 한다.

(3) AI로 인해 소액 다수의 소비자피해는 더 확산되거나 확대될 가능성이 높다. 1:1의 거래관계를 전제로 한 현행 민사소송법상의 소송절차로는 소액 다수의 피해자가 발생하는 경우 이에 대한 손해배상책임을 인정받기란 쉽지 않다. 그러므로 집단소송법과 증거개시제도 및 징벌적 손해배상에 대한 규정을 도입할 필요성이 높다. 이러한 소비자 3법을 일반법에 규정하는 것은 체계상 맞지 않으므로, AI 소비자보호법이라는 일반법을 제정하기보다는 AI 책임기준 등과 분쟁조정절차 등에 대해서는 소비자기본법을 개정하고, 나머지 소비자보호를 위한 제도는 별개의 개별법으로 처리하는 것이 바람직하다.

라. AI 시대 소비자보호를 위한 개별법 개정 방향

(1) 현행 소비자보호를 위한 입법 체계는 크게 소비자기본법상의 분쟁조정제도와 거래 형태·방식에 따른 개별법상의 보호라는 두 개의 축으로 구성되어 있다. 그러므로 AI를 매개 또는 개입한 거래에 대해서도 기존의 소비자보호 체계가 적용되도록 하는 것이 필요하다.

(2) 전자상거래법에 대해서는 AI 가격 개인화 고지 의무 및 다크패턴 금지규정을 확대하는 것이 필요하다. 현행 전자상거래법 제21조의2의 다크패턴 규정은 일반적 기만 행위에 초점이 있어 AI 기반 개인화 가격차별에는 적용 여부가 명확하지 않고, 점점 더 진화하는 다크패턴을 포섭하는 것이 어려울 수 있으므로 공정거래위원회가 별도의 고시로 규정할 수 있도록 포괄적 규정을 제1항 제6호로 신설하는 것이 바람직하다.

(3) 현행 표시광고법은 AI 추천, 보증 등에 대해 심사지침을 통해 규제를 하고 있으나, 'AI 워싱' 및 AI 기능 과장광고 등에 대한 일반적인 규정이 없고 광고 등의 유형에 따라 규제하는 체계를 취하고 있다. 그러므로 세부적인 행위유형별 규제는 심사지침에서 규정하더라도, 표시광고법 자체에 AI 특성에 맞는 일반적인 규제 규정을 도입하는 것이 필요하다.

(4) 제조물책임법상 제조물에 소프트웨어가 포함될 수 있도록 개정되어야 한다. 제조물책임법 제2조 제1호는 제조물을 "제조되거나 가공된 동산"으로 정의하고 있는데, 여기서 동산인 하드웨어와 물리적으로 결합되지 않은 소프트웨어는 제조물에 해당되지 않는 것으로 해석된다. 그러므로 '제조물'의 정의 규정에 클라우드 형식 등으로 제공되는 인공지능시스템(소프트웨어)을 명시적으로 포함시키는 방향으로 개정되어야 한다. 또힌, 소프트웨어의 결함과 손해발생과의 인과관계에 대해서는 소비자가 이를 입증하는 것이 사실상 불가능하므로 손해가 발생한 경우에는 결함과 인과관계에 대해 사실상 추정 규정을 두고 입증책임을 완화시켜야 할 것이다.

(5) 약관규제법상 약관 설명의무 규정을 AI 알고리즘이 관여하여 작성한 경우에는 단순 고지 수준을 넘어 소비자가 이해할 수 있을 정도로 알고리즘의 핵심을 설명하도록 하는 '디지털 설명의무'를 신설해야 한다. 또한, AI 에이전트가 계약을 체결하는 경우에 대한 심사기준과 표준계약을 규정할 수 있도록 하는 것이 필요하다. 특히 소비자가 AI 에이전트를 이용해서 계약을 체결할 경우 AI 에이전트의 권한 범위(가격 상한, 거래 종류, 자동 결제 한도 등)를 명확히 설정하여야 하고, AI 에이전트가 소비자의 실제 의도를 벗어난 계약을 체결한 경우 취소 또는 청약철회 권한을 법제화하는 것이 필요하다.

(6) 알고리즘에 대해서는 일률적으로 평가하기 어려운 복잡한 문제가 포함되어 있다. 알고리즘으로 인해 공정거래법상 경쟁을 제한하기도 하고, 불공정행위를 하기도 하며, 개인정보를 무단으로 수집해 소비자의 이익을 침해할 수도 있다. 발제자들이 제시하는 '절차적 정당성 확보 방안' 또는 '알고리즘 설명 요구권' 등도 충분히 도입 가능한 방안으로 보인다. 다만, 이처럼 알고리즘으로 인한 침해는 그 성격에 따라 소비자법·개인정보법·경쟁법 영역에 걸쳐 있으므로, 입증책임의 완화 또는 전환 역시 각 영역의 개별법에 맞추어 분산 규정하는 것이 체계상 타당하다. 즉 시장의 경쟁질서를 침해한 결과가 발생한 경우에는 공정거래법에서, 소비자의 개인정보를 무단으로 수집·사용한 경우에는 개인정보보호법에서, 소비자의 이익을 일반적으로 침해한 경우에는 소비자기본법 또는 관련 개별 소비자법에서 각각 사실상 추정 규정을 두어, 알고리즘을 이용한 사업자로 하여금 알고리즘을 정상적으로 운용하였다는 사실, 개인정보를 적법하게 수집하였다는 사실, 다른 소비자와 의도적으로 차별행위를 하지 않았다는 사실 등을 입증하도록

함으로써 입증책임을 완화 또는 전환하는 것이 필요하다.

4. 결어

결국 핵심은 인공지능기본법이라는 단일한 일반법에 진흥과 보호를 모두 담으려 하기보다, 진흥은 인공지능기본법에 맡기고 소비자 보호는 소비자기본법 개정과 거래유형별 개별법 정비라는 기존의 체계 안에서 충실히 구현하는 것이 필요하다. 그래야만 AI로 인한 소비자 피해 유형이 어떻게 분화되더라도 그에 상응하는 법적 구제수단을 빠짐없이 갖출 수 있다. AI 산업 진흥을 위한 엑셀레이터가 인공지능기본법이라면, 급가속시 발생할 수 있는 위험성을 예방할 수 있는 브레이크가 집단소송법 등 소비자 3법이다. 브레이크가 작동하지 않는 자동차는 단순한 운송수단이 아니라 다른 운전자와 보행자의 생명에 위험을 초래할 수 있는 위험한 무기라는 점을 기억해야 한다. 그러므로 집단소송법과 같은 브레이크가 반드시 필요하다.