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)