OPINION

연구위원
노성호

인공지능 기본법과 금융투자업의 대응 방향*

『인공지능 발전과 신뢰 기반 조성 등에 관한 기본법』(AI 기본법)과 관련 시행령이 2026년 1월 22일 전면 시행되었다. 이 법은 포괄적 AI 규제 체계로서, '인공지능사업자'를 대상으로 투명성 고지, 안전성 확보, 고영향 AI 책무 등 실무적 의무를 단계적으로 부과하고 있다. 금융투자업에서 AI는 투자자 의사결정과 시장 안정성에 직접적 영향을 미칠 수 있다는 점에서 파급력이 클 것으로 예상된다는 점에서 확산 속도에 부합하는 실무적인 대응이 요구되고 있다.

AI 기본법의 준수는 단순한 규제 의무가 아니라, 업무 표준화를 통한 리스크 감소, 중복규제 회피를 통한 비용 최소화, 적극적 AI 도입을 통한 수익성 극대화라는 세 가지 측면에서 기회로 인식할 필요가 있다. 이를 위해 본 고는 AI 기본법의 핵심 의무인 투명성 고지(제31조), 고성능 AI 안전성(제32조), 고영향 AI 확인·책무·영향평가(제33~35조)를 금융투자업 주요 업무에 적용하여 조항별 적용 가능성을 체계적으로 분석한다.

대응 전략은 크게 세 단계로 구분하여 논의할 수 있다. 단기적으로는 전사적 AI 인벤토리를 구축하고 고영향 AI를 분류하여 고객과 접점이 많은 서비스에 고지 체계를 우선 확립해야 한다. 중기적으로는 신용·담보평가, AML/FDS 등 고객 권리에 중대한 영향을 미치는 업무를 고영향 후보군으로 분류하고, 위험 관리 방안, 설명 방안, 이의제기 절차, 문서 보관 체계를 내재화하여 감독 대응이 가능한 운영 구조를 갖추어야 한다. 장기적으로는 AI 특유의 운영 리스크를 상시 모니터링하고 전사적 내부통제로 정착시킬 필요가 있다. 결론적으로, 규제와 평판 비용을 내재화한 AI 서비스만이 장기적 수익성과 시장 신뢰를 동시에 확보할 것으로 기대한다.

『인공지능 발전과 신뢰 기반 조성 등에 관한 기본법』(이하 'AI 기본법')과 관련 시행령이 2026년 1월 22일부터 전면 시행되었다. 이 법은 '인공지능사업자'[1]를 중심으로 투명성, 안전성, 고영향 AI, 해외사업자 국내대리인 등 실무 차원에서 이행해야 할 의무 조항을 단계적이고 포괄적으로 제시하고 있다. 특히, EU AI Act보다 앞서서 전면 시행되는 AI 관련 규제라는 점에서 그 영향을 면밀하게 살펴볼 필요가 있다.

AI 기본법이 금융투자업에 중요한 이유는 AI가 제공하는 정보 및 이에 기반한 투자 판단이 투자자의 의사결정과 시장 안정성에 직접적인 영향을 미칠 가능성 때문이다. 국제적으로도 AI의 활용이

* 본고의 견해와 주장은 필자 개인의 것이며, 자본시장연구원의 공식적인 견해가 아님을 밝힙니다.
1) '인공지능사업자'는 '개발사업자'와 '이용사업자'를 포함한다.

금융시장에서의 군집행동이나 클라우드 기반 데이터 집중 리스크, 자동화 의사결정의 불투명성 등을 통해 금융시스템의 취약성을 증폭시킬 수 있다는 경고가 지속적으로 제기되고 있다.[2] 국내에서도 많은 금융사에서 AI 서비스를 운영 또는 준비하고 있지만, AI 윤리 원칙이나 위험 관리 기준이 정립되지 않은 상황이며, 의사결정기구를 설치한 곳도 극소수에 불과한 것으로 나타났다.[3] 이는 금융업에서 AI의 확산 속도와 리스크 관리 체계의 성숙도 사이에 격차가 존재함을 시사한다.

규제와 실무 사이의 격차를 메우기 위해서는 규제의 준수를 수동적으로 의무를 이행하는 것에서 나아가 (1) 업무 표준화를 통한 위험 요인 제거, (2) 중복규제의 회피를 통한 비용 최소화, (3) AI 도입의 불확실성 해소로 수익성 극대화의 기회로 전환하기 위한 노력이 필요하다. 더불어 AI 기본법은 산업 전반에 걸친 상위 규범인 반면, 금융권 AI 가이드라인 및 AI RMF(Risk Management Framework) 등은 금융산업에 특화된 이슈인 소비자보호, 시장 신뢰도 제고, 시스템 리스크 통제 등을 내재화하는 규율이라는 점에서 금융투자업계는 두 규범 체계를 포괄하는 접근이 필요하다.

AI 기본법 및 시행령의 핵심 조항: 투명성, 안정성, 고영향 AI

이 절에서는 AI 기본법 제31~35조에 명시된 핵심 의무를 세 가지 영역으로 분류하고, 금융투자업에서 AI 기반 서비스의 도입 과정에서 각 영역이 실무적으로 미치게 될 영향을 평가한다.

첫 번째 항목은 투명성 확보 의무이다(법 제31조 및 시행령 제23조). AI 기본법 제31조는 인공지능사업자에게 크게 세 가지 의무를 부과하고 있는데, 이는 고영향 · 생성형 AI 기반 서비스라는 사실의 사전 고지(제1항), 생성형 AI의 결과물이라는 표시(제2항), 현실과 구분이 어려운 가상 콘텐츠에 대한 별도 고지 및 표시(제3항)이다. 이와 관련하여 시행령 제23조는 사전 고지를 계약서에 기재하거나 화면 · 단말기에 표시하는 등으로 구체화하고, 그 방법을 제한하고 있다. 금융투자업에서는 이와 같은 조항이 고객이 접하는 모든 AI 기반 소통 방식에 적용될 가능성이 있다. 특히, AI 기반 PB · 챗봇은 서비스 진입 시점(예: 앱 시작 화면)에서의 고지와 대화 결과의 AI 활용 여부 표시 체계가 필요하다. 예를 들어, 종목에 대한 분석을 자동으로 생성하는 서비스는 AI 고시 의무를 위반할 경우 투자자의 오인을 유발할 수 있어 사용자가 AI의 결과물임을 명확히 인식할 수 있는 표준 고시안을 마련할 필요가 있다.

두 번째는 고성능 AI의 안전성 확보 의무이다(법 제32조 및 시행령 제24조). 누적 연산량 10^{26} FLOPs[4] 이상인 AI 시스템[5]에 대해 AI 기본법 제32조는 시스템의 "수명주기 전반에 걸친 위험의

2) 금융업에서 AI 활용에 따른 위험 요인에 대한 논의는 OECD, 2023, *Generative Artificial Intelligence in Finance*, FSB, 2024, *The Financial Stability Implications of Artificial Intellgence* 등을 참고할 수 있다.

3) 금융감독원, 2026. 1. 16, 「금융분야 AI 위험관리 프레임워크(AI RMF)」 도입, 보도자료.

4) Floating point Operations Per second의 약자로 1초 동안 수행할 수 있는 부동소수점연산의 횟수를 의미한다.

5) 해당 기준을 넘는 AI 시스템의 예로 GPT-4.5, Grok3 등이 있다.

식별 · 평가 및 완화"를 요구하고 있다. 금융투자업계에서 이와 같은 조건을 충족하는 '초거대' 모형을 직접 개발하는 사례는 드물 것으로 예상된다. 그러나 대형사를 중심으로 자체 파운데이션 모형을 구축하거나 미세조정(fine-tuning)을 거친 고성능 모형의 연산을 직접 수행하는 경우 해당 의무의 적용 가능성을 배제할 수 없다. 이를 대비하여 안전성 의무를 수동적인 보안 점검으로 제한하지 않고 위험 모니터링 및 (차단 · 롤백 등) 리스크 완화 체계를 선제적으로 갖추는 것이 필요하다.

세 번째로 고영향 AI[6] 관련 의무는 여러 조항(법 제33~35조)에 걸쳐서 다루어지고 있다. 제33조(확인)는 고영향 해당 여부를 사전 검토하고 장관에게 확인을 요청할 수 있게 하며, 이는 분쟁 시 주의의무 입증을 통한 방어 수단이 된다. 제34조(책무)는 위험관리방안, 인공지능 개발 결과 및 과정 설명 방안, 이용자 보호 방안 등에 대한 조치를 요구한다. 제35조(영향평가)는 권고형이나, 감독 대응 관점에서 실무에 AI 기반 서비스를 도입하기 전 최종 확인 단계로 내재화될 가능성이 크다. 금융투자업의 경우 고영향 AI 사례 중 하나로 '채용 · 대출 심사'가 명시되어 있다는 점에서 증권사 신용공여(를 위한 담보 · 신용평가), 내부 인사 평가 자동화 시스템 등에 AI가 활용될 경우 해당 항목의 적용을 받을 가능성이 있다.

금융투자업 주요 업무별 적용 가능성

금융감독원에 따르면 국내에서도 118개 금융사가 653개 AI 서비스를 운영(준비 중 포함)하고 있으며,[7] 해외의 경우 챗봇, 로보어드바이저, 리서치 자동화 등의 업무에 AI를 적극적으로 도입하여 업무 효율성과 비용 절감을 추구하고 있다.[8]

〈표 1〉에서는 금융투자업 업무 유형별로 AI 기본법 제31~35조의 적용 가능성을 요약하였다. 우선, 로보어드바이저의 경우, 구체적인 적용 방식에 따라 고영향 AI로 분류되지는 않을 수 있으나, 고객에게 제공되는 설명 또는 리포트가 생성형으로 작성되면 제31조 결과물 표시 의무의 적용을 받을 수 있다. 이는 생성형 AI에 기반한 PB 상담에도 동일하게 적용될 것으로 판단된다.

6) 고영향 AI의 정의는 AI 기본법 제2조 제4항에서 "사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는" AI로 규정하고 있다. 과학기술정보통신부는 「고영향 인공지능 판단 가이드라인」을 발표하여 고영향 AI를 식별하는 실무적인 기준을 제시하였는데 해당 가이드라인에서 금융업으로 분류되는 분야별 사례로는 대출 심사가 유일하게 포함되었다.

7) 금융감독원, 2026. 1. 16, 「금융분야 AI 위험관리 프레임워크(AI RMF)」 도입, 보도자료.

8) 김진영·노성호, 2026, 『특허 분석을 통하여 살펴본 금융투자업의 AI 활용과 시사점』, 자본시장연구원 연구보고서 26-03.

〈표 1〉 업무 유형별 AI 기본법 적용 가능성

	투명성 (제31조)	안전성 (제32조)	고영향 AI		
			확인 (제33조)	책무 (제34조)	영향평가 (제35조)
로보어드바이저	조건부	조건부	낮음	조건부	조건부
AI PB	직접	조건부	조건부	조건부	조건부
시장 리서치	직접	낮음/조건부	통상 낮음	통상 낮음	통상 낮음
신용 · 담보평가	조건부	조건부	직접 가능성 높음	직접 가능성 높음	권고
AML/FDS	제한적	조건부	조건부	조건부	조건부
알고리즘 트레이딩	통상 낮음	조건부	낮음	낮음	낮음
시장감시	제한적	조건부	조건부	조건부	조건부
내부업무 자동화	제한적 (내부 한정)	조건부	조건부	조건부	조건부

주: '직접'은 즉시 의무 발생 가능한 영역, '조건부'는 요건 충족 시 적용 가능한 경우, '제한적'은 특정 상황에 한하여 적용 가능한 경우를 의미한다.

한편, 신용 · 담보평가의 경우 고영향 AI 사례에 '대출 심사'가 명시되어 있어, AI가 승인여부, 한도, 금리 등 핵심적인 거래 조건을 결정할 경우 고영향 사례로 분류될 가능성이 높다. 이 경우, 제34조에 따른 책무를 부여받게 되면 서비스 제공 사업자는 위험관리 및 설명 방안, 인간 직원을 통한 관리 절차 등을 마련할 필요가 있다. 이상거래 및 자금세탁 방지 체계(AML/FDS)의 경우는 대출 심사와는 다르지만 자동화된 의심 거래 차단 및 계좌 제한 절차는 고객의 권리에 중대한 영향을 미칠 수 있다는 점에서 적용 여부에 대한 검토가 필요하다.

마지막으로 내부 업무 자동화에 활용되는 AI의 경우 고객에게 결과물을 적용하지 않는다면 투명성(제31조), 안정성(제32조) 등의 의무의 적용 가능성은 제한적일 것이다. 하지만 내부 절차에 따른 결정이 고객의 권리, 거래 조건 등에 실질적인 영향을 미칠 경우라면 고영향 AI로 분류될 가능성을 검토할 수 있다.

금융투자업의 대응 전략

AI 기본법 및 시행령은 안전한 AI 활용을 우선적인 경영 과제로 인식할 것을 요구하고 있다. 특히 생성형 AI와 고영향 인공지능에 대해 사전 고지, 위험 관리, 설명가능성 제고, 인간의 감독, 문서 보관 등을 요구하고 있어, 고객과 접점이 많은 금융 서비스는 직접적인 영향을 받을 수 있다. 이에 금융투자업자는 전사적 AI 인벤토리를 구축하고, 도입하려는 AI 모형의 고영향 여부를 분류한 뒤, 서비스 진입

단계에서 고지 · 표시를 사용자가 접하는 화면에 내장하는 등의 최소 준수 체계를 우선 확립할 필요가 있다.

〈표 2〉 AI 기본법 대응 목표와 과제

	목표	위험 관리	비용 최소화	수익 극대화
단기	최소 준수 체계 구축	√ 전사적 AI 인벤토리 구축(내 · 외부용 구분) √ 고영향 AI 분류 √ AI PB, 챗봇 등에 투명성 고지	√ 고지 · 표시 문구 표준화 √ 사전검토 체크리스트 템플릿 마련 √ 기존 규제와 연계성 검토	√ AI 상담 · 리서치 자동화 PoC 확대 √ 고객 응대 리드타임 단축 √ 내부업무 자동화로 즉각적 생산성 확보
중기	고영향 업무 통제 체계 정착	√ 위험관리, 설명방안, 인간 기반 관리 체계 구축 √ 문서보관, 이의제기 프로세스 내재화	√ 벤더 실사 절차 표준화 √ 해외 벤더 계약에 국내대리인, 감사권, 재학습 제한	√ 설명가능성 기반 시스템 신뢰도 제고 √ 고영향 AI 검증을 서비스 경쟁력으로 전환 √ 리서치 정확성 관리로 품질 우위 확보
장기	모니터링 및 운영 리스크 내재화	√ 모델 드리프트, 데이터 오염 상시 감시 √ 사고 대응 지침 마련 √ 전사적 시스템 리스크 관리	√ 로그 통합 보관 시스템 구축 √ 감사 대응 자료 데이터베이스 구축 √ 규제 대응 프로세스 자동화	√ 규제 정합성이 높은 AI 서비스 상용화 √ 개인화 기반 추천 시스템 확대 √ 신뢰성 높은 서비스 확대

중기적으로는 신용 · 담보평가, 자동 한도 조정, AML/FDS와 같이 고객 권리에 중대한 영향을 미칠 수 있는 업무를 고영향 후보군으로 분류하고, 사전검토 절차와 인간의 개입(human-in-the-loop) 절차를 시스템에 내재화해야 한다. 위험 관리 방안, 모형 설명 방안, 이의제기 절차, 5년간의 문서 보관 체계를 수립하여 감독 대응이 가능한 운영 구조를 갖추는 것이 핵심이다. 고영향 AI의 영향 평가 역시 권고적 성격이지만 사실상 서비스 출시 전 최종 절차로 기능할 가능성이 높아 표준화된 검사 항목을 기반으로 시범 적용하는 것이 바람직할 것이다.

장기적으로는 모델 드리프트(model drift),[9] 환각 현상(hallucination)[10] 등 AI 특유의 운영 리스크를 상시 모니터링하고, 문제 발생 시 격리와 롤백에 이어 고객에게 빠르게 공지하는 사고 대응 체계를

9) 배포된 머신러닝(machine learning) 모형이 시간이 지나면서 실제 업무 환경 및 입력 데이터의 변화로 인하여 예측 성능이 저하되는 일종의 노후화 현상을 의미한다.

10) 생성형 AI가 거짓 또는 편향된 결과물을 만들어내는 현상을 의미하며 이는 확률적인 과정을 통해 생성되는 출력물에서 일반적으로 발생할 수 있다. 대형언어모형(Large Language Model)에서 흔히 발생하는데, 이에 대한 구체적인 논의와 대응 방안은 노성호, 2024, 『증권업 경쟁력 강화 시리즈 2: 대형언어모형의 발전과 금융정보분석에의 활용 방안』, 자본시장연구원 이슈보고서 24-02.를 참고할 수 있다.

마련하고 이를 전사적인 내부통제 절차로 정착시켜야 한다. 동시에 비용 측면에서는 고지 문구, 사전 검토 항목, 벤더 실사 양식을 표준화하고, 기존 금융소비자보호 및 개인정보보호 관련 규제를 고영향 AI에 적용되는 책무 사항과 비교 분석하여 중복되는 업무로 인한 비용을 최소화할 수 있다. 나아가 해외 업체에서 개발한 AI 모형 도입 시에는 국내대리인 지정, 감사권, 재학습 제한 등을 표준 계약으로 정립하여 벤더 위험을 구조적으로 통제하려는 노력이 필요하다.

수익 측면에서는 규제 부담이 상대적으로 낮은 생성형 AI 기반 상담, 투자 리서치 영역에서 생산성 개선 효과를 조기에 확인할 필요가 있다. 한편, 고영향 영역에서는 설명가능성(explainability)과 인간 관리자의 체계적인 감독을 AI 기반 서비스의 경쟁력으로 전환해야 한다. 결국 규제 비용과 긍정적인 평판으로 발생하는 기대 수익을 내재화하여 적합한 AI 서비스만이 장기적으로 수익성과 시장 신뢰를 동시에 확보할 것으로 기대한다.