

보도	2026.1.16.(금) 조간	배포	2026.1.15.(목)
담당부서	디지털금융총괄국 AI·디지털혁신팀	책임자	국장 이 석 (02-3145-7120)
		담당자	팀장 김은철 (02-3145-7140)

「금융분야 AI 위험관리 프레임워크(AI RMF)」 도입

- 금융회사의 AI 관련 ^①거버넌스, ^②위험평가, ^③위험통제 핵심 프로세스를 제시
- AI 시스템의 도입·활용 **순 주기**에 걸친 위험을 **금융회사 스스로 관리**토록 유도
- 금융권 의견수렴 등을 거쳐 “금융분야 AI 가이드라인”과 함께 '26.1분기 중 **사행** 예정

I 배경

- **(금융권의 AI 활용 동향)** 최근 AI는 기술 진보를 넘어 산업·사회 전반에 패러다임 전환을 촉진하며 경제성장의 핵심 동력으로 부상
 - 금융권 또한 AI 기반의 대고객 서비스를 출시하거나 내부 업무 혁신에 AI를 활용하는 사례가 크게 확대*
- * 세계경제포럼(WEF)은 AI 혜택이 가장 큰 분야로 금융산업을 지목하며 업무의 32~39%를 완전 자동화할 수 있고 추가로 34~37%는 AI로 크게 보조될 수 있다고 평가('25.1월)

국내 금융업권별 AI 활용 현황 (‘25.4월말 기준, 단위 : 개)

구분	은행	증권	보험	카드	합계
대상 회사 수 (a)	20	37	53	8	118
AI 서비스 수 (준비중 포함) (b)	299	143	145	66	653
회사별 평균 AI 서비스 수 (b/a)	15.0	3.9	2.7	8.3	5.5

- **(그간의 감독방향)** 금융당국은 금융권이 AI 대전환을 선도할 수 있도록, 다각도로 지원하며 금융혁신을 유도 중
 - 망분리 규제 합리화, 인프라·데이터 제공 등을 통해 AI의 잠재력이 금융분야에서 현실화될 수 있도록 지원하고 있으며,
 - 다양한 형태의 가이드라인* 등을 마련하여 금융회사가 AI를 도입·활용하는 과정에서 참고할 수 있도록 제공

* ① 운영 가이드라인('21.7.) ② 개발·활용 안내서('22.8.) ③ 보안 가이드라인('23.4.)

II 금융권의 AI 위험관리 현황

- **(금융 AI의 위험)** 다만, AI의 복잡성, 불투명성, 데이터 의존성 등으로 인한 새로운 유형의 위험이 대두되면서 금융소비자의 권익을 침해하거나 금융안정을 훼손할 수 있다는 우려가 점증
 - 특히 금융 분야는 자금 중개라는 본질적 특성상 AI 위험이 산업·사회 전반과 금융소비자에게 큰 영향*을 미칠 수 있는 데다,
 - * (예시) AI 모델이 환각 현상을 보이거나 학습데이터가 편향된 경우 거짓·왜곡된 결과를 산출할 수 있으며, 차별 등 사회적 불평등 및 윤리적 논란이 심화할 소지
 - 실시간 거래가 이루어지는 금융시장에서는 AI 오류가 여타 산업에 비해 빠르고 광범위하게 전파*되어 시스템 위협으로 발현 가능
 - * 심지어 AI 시스템이 정상 작동하더라도 유사 알고리즘이 특정 신호에 동시 반응하여 급격한 시장 변동을 유발하는, 이른바 '플래시 크래시(Flash Crash)' 가능성도 상존
- **(금융회사 위험관리 현황)** 이러한 금융 AI 위험의 차별적 특성 및 파급력에도 불구하고, 거버넌스 및 위험관리는 대체로 미진한 상황
 - 국내 금융회사 점검* 결과, 소수의 회사만이 AI 거버넌스를 구축하거나 위험관리·내부통제 등 관련 내규를 마련**한 가운데,
 - * 은행(20), 카드사(8), 증권사(37), 보험사(53) 등 총 118개사 대상 전수 조사('25.4월)
 - ** 은행 5개(25%), 보험사 4개(7.5%), 증권사 1개(2.7%)만 의사결정기구를 설치하였으며, 약 85%의 금융회사는 AI 윤리원칙, 위험관리 기준 등 미마련
 - AI 서비스를 개발·활용하는 과정에서 감독당국이 제시한 AI 관련 가이드라인을 적용한 비율도 약 60%로 저조
 - 한편, 상당수 금융회사는 AI 거버넌스 및 위험관리 등에 관한 불확실성 해소 및 책임소재 명확화 등을 위해 금융권의 특수성을 반영한 가이드라인 제시를 희망

➔ 금융감독원은 금융권이 혁신과 책임의 균형 하에 건전하게 AI를 활용할 수 있도록, 「금융분야 AI 위험관리 프레임워크(이하 'AI RMF)」를 마련하여 금융회사 스스로 위험을 관리할 수 있는 구조적 기반을 제시

※ AI RMF는 '25.12.22. 발표한 「인공지능 대전환(AI), 금융이 선도하겠습니다」(금융위·금감원 등)의 "금융분야 AI 가이드라인 개정 방향"과 관련된 사안으로, 동 가이드라인에 포함된 금융 AI 7대 원칙 중 "1) 거버넌스 원칙"을 구체화

III 금융분야 AI RMF의 주요 내용



※ **(적용 방식)** AI RMF는 법적 강제력이 없는 가이드라인 형태의 자발적 지침으로, 각 회사 사정*에 따라 강화·완화 또는 선택 적용 가능

* 규모, 인적·물적 자원, 영위 업무, AI 활용 범위·수준 및 기술 성숙도 등

1 AI 거버넌스

금융 AI 7대 원칙 반영	정성적 요소 ① 거버넌스 ② 합법성 ③ 보조수단성
-------------------	--------------------------------

◆ 금융회사는 AI 위험관리 등을 위한 **① 의사결정기구** 및 **② 전담조직**을 구성하고, **③ 관련 내규**를 마련하는 등 거버넌스 체계를 구축

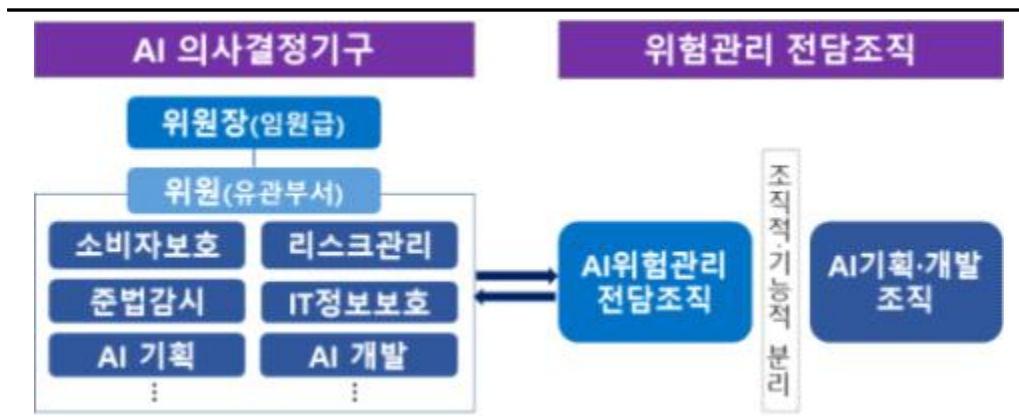
- 명확한 책임소재를 바탕으로 '견제와 균형' 하에 AI 시스템의 기획, 구축, 운영 쏠 단계에 걸친 잠재적 위험을 평가·관리

- **(의사결정기구) AI 위험관리** 등을 위한 **의사결정기구**(예: AI 윤리 위원회, AI 위험관리위원회)를 설치하여 AI 도입·활용에 적극 관여*
 - * AI 윤리원칙 등 AI 관련 내규 제·개정, 위험관리·소비자보호정책 수립, 고위험 AI 서비스 승인 등 AI 관련 중요사항을 심의·의결
 - 특히, 위원장은 CEO가 AI와 관련된 사업계획, 전략, 위험 등을 명확히 인식할 수 있도록 CEO에게 정기적으로 보고*
 - * **7대 원칙 ① 거버넌스 원칙** 최고경영자를 포함한 경영진은 AI 개발·활용에 대한 관심을 갖고 역할과 책임을 분담해야 함

※ **책무구조도**를 도입한 금융회사는 AI 관련 **내부통제** 및 **위험관리** 등에 관한 **책무**를 반영하여 AI 도입·활용에 따른 **책임소재**를 **명확화**하는 방안을 고려

- **(위험관리 전담 조직) AI 도입·개발 추진 조직과 독립된 위험관리 전담 조직**을 설치하여 AI 관련 업무 전반을 통제·관리
 - AI 위험의 인식·측정·평가 등 위험관리 전반을 통할하고, 「AI 기본법」 등 관련 법규상 각종 의무의 준수 여부를 관리·감독

예시 AI 거버넌스 체계 구성



※ 회사의 규모, 자원, 영위업무, AI 활용 범위 등을 고려하여 조직을 구성하되, 이해상충 방지를 위한 독립성 확보 방안을 수립할 필요

□ **(내규)** AI 윤리기준을 근간으로 AI 위험관리규정 및 지침 등 AI 관련 내규*를 수립하고, 세부적인 업무매뉴얼을 마련

* AI 거버넌스 구조, 위험의 인식·측정·관리, 고위험 AI 통제, AI시스템 기획·개발·운영·활용 절차, 보안 업무, 소비자보호 업무, 관련 법규 위반 시 처리 등을 포함

○ 특히, 금융 및 AI 관련 법규 준수 의무*를 부과하고, 최종 의사 결정 책임을 임직원이 지도록 내규화**

* **7대 원칙 ② 합법성 원칙** AI 활용 전 단계에서 금융·AI 등 관련 법규를 준수해야 함

** **7대 원칙 ③ 보조수단성 원칙** 현 단계에서 AI는 업무의 보조 수단이므로 최종 의사 결정과 그에 따른 책임은 임직원이 수행함

○ 한편, 표준화된 매뉴얼을 통해 AI 도입·활용 쏠 프로세스를 관리하고, 부서간 역할 및 책임을 명확화하여 잠재 위험에 대응

예시 AI 관련 내규 체계



2 AI 위험평가

금융 AI	정량적 요소
7대 원칙 반영	② 합법성 ④ 신뢰성 ⑥ 신의성실 ⑦ 보안성

◆ 금융회사는 AI 위험의 체계적인 인식·측정·관리 등을 위해서 **위험기반 접근방법** (Risk-based approach)의 **종합 평가체계**를 구성*
 * 단, 규칙 기반 접근방법(Rule-based approach) 등 금융회사의 AI 도입 범위, 규모, 인력 등 자체 실정에 맞는 위험등급 분류도 허용

□ **(평가체계)** “금융 AI 7대 원칙” 중 정량적 요소를 토대로 위험평가 체계를 설계하고, 원칙 준수를 위한 기본요소를 평가 항목 등에 반영*

* **7대 원칙 ② 합법성 원칙** AI 활용 전 단계에서 금융·AI 등 관련 법규를 준수해야 함

7대 원칙 ④ 신뢰성 원칙 AI 개발 과정에서 신뢰할 수 있는 데이터와 모델을 사용해야 함

7대 원칙 ⑥ 신의성실의 원칙 AI 활용 시 금융소비자의 이익을 최우선으로 해야 함

7대 원칙 ⑦ 보안성 원칙 AI 활용 시 보안성 기준 및 점검·개선 체계를 마련해야 함

○ ① **위험 인식·측정** → ② **위험 경감** → ③ **잔여위험 평가** → ④ **위험등급 산정** 등 종합적·체계적 평가를 통해 AI 서비스별 위험을 분류

예시 AI 서비스 위험평가 절차(세부 항목 조정 가능)

부문	① 위험 인식·측정	② 위험 경감	③ 잔여 위험	④ 위험등급 산정
합법성 원칙 (20%)	금융소비자보호법 위반 가능성	8	(4)	4
	AI기본법 위반 가능성	4	(3)	1
	데이터 관련법 위반 가능성	4	(2)	2
	개별 업권법 위반 가능성	4	(2)	2
신뢰성 원칙 (30%)	품질	6	(4)	2
	편향성	6	(2)	4
	공정성	6	(2)	4
	설명가능성	6	(1)	5
신의성실 원칙 (20%)	성능	6	(3)	3
	계약 권리 침해	6	(3)	3
	책임 투명성	6	(3)	3
보안성 원칙 (30%)	소비자 보호방안	8	(4)	4
	보안	8	(3)	5
	안정성	8	(4)	4
	위탁/관리	8	(3)	5
	프라이버시	6	(3)	3
총합 54점				<p>위험 점수</p> <p>75 고위험 서비스 추가 통제 및 관리 강화</p> <p>50 중위험 서비스 기본 통제 및 관리 적용</p> <p>25 저위험 서비스 통제 완화</p>

- **(고영향 AI)** 법체계와 정합성 확보를 위해 「AI기본법」상 고영향 AI*에 대해서는 위험점수에 따른 등급 분류와 관계없이 고위험 서비스로 분류

* AI 기본법(§2 iv)은 고영향 AI의 예시로 '대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가'를 포함

3 AI 위험통제

금융 AI 7대 원칙 반영	정성적 요소 ② 합법성 ⑤ 금융안정성
-------------------	-------------------------

◆ 금융회사는 위험 수준별 차등화된 통제·관리를 수행하고, 초고위험 AI에 대해 출시 여부 재검토 등 위험통제를 위한 제반 절차 이행

- **(위험기반 통제)** 위험평가 세부 절차를 따라 산출한 AI 서비스의 위험등급을 기준으로, 위험별로 차등화된 통제·관리 수행

예시 서비스 위험 수준에 따른 차등화된 통제·관리 수행방안

저위험: 통제 완화	기본 통제·관리	고위험: 통제 강화
✓ 승인절차 및 작성 문서 등 축소	✓ 상품·서비스 출시 前 경감조치 검증 ✓ 상품·서비스 운영단계 모니터링 기준 적용보고 ✓ 상품·서비스 위험 변경시 위험 수준 재평가 ✓ AI 세부 업무 방법에 따른 관리 (AI 위험관리 업무매뉴얼, 검증매뉴얼 등)	✓ AI윤리위원회 사전 승인/사후 검증 ✓ 제3자에 의한 평가검증 ✓ 운영 단계 모니터링 강화

- **(초고위험 AI 출시 재검토)** 초고위험* AI 시스템 등에 대해서는 AI 의사결정기구를 통해 출시 여부를 재검토

* 금융안정성 훼손 또는 소비자 권익 침해 우려가 있다고 AI 의사결정기구가 판단하는 경우 등

- **(제반 내부통제 절차 이행)** 모니터링·사후관리, 문서화 및 교육, 감독당국 정보공유 등 위험통제를 위한 제반 절차 이행

IV 향후 계획

- 금융감독원은 업권별 협회* 등을 통해 금융권에 AI RMF(안)을 배포하고, 설명회·간담회 등을 통해 충분한 의견수렴을 거친 후 금년 1분기 중 최종안을 확정·시행할 예정

* '26.1.16.(금)부터 금융권 AI 플랫폼(<https://finai.kcredit.or.kr>)의 공지사항 메뉴를 통해서도 확인 가능

- 아울러, 모범사례 전파, 도입회사 실태점검 등을 통해 AI RMF의 거버넌스, 위험관리 및 위험통제 프로세스가 금융회사의 내부통제 체계에 안착할 수 있도록 충실히 지원할 계획