금융위원회	보도자료

보도시점 2025. 12. 22.(월) 14:00 배포 2025. 12. 22.(월) 09:00

인공지능 대전환(AX), 금융이 선도하겠습니다.

- 권대영 금융위 부위원장, **국정과제 이행**과 **대통령 업무보고 과제**의 신속한 추진을 위해 **금융권 AI 협의회를 개최**하고 현장·전문가의견 청취
- 금융권이 AI 대전환을 선도할 수 있도록 **인프라 및 데이터 지원**, 안전한 AI활용 위한 **규율체계 정비**, 소비자중심의 **포용금융AI 서비스** 등 논의

【관련 국정과제】 46-4. "디지털 금융 및 금융산업의 융복합 촉진"을 위한 금융권 AI 대전환 지원

Ⅰ. 회의개요

금융위원회는 '25.12.22일 **권대영 금융위원회 부위원장** 주재로 금융권 AI 협의회를 개최하여 금융권의 **AI 대전환을 위한 추진과제를 논의**하고 금융권의 AI 활용 관련 **애로사항 및 현장 의견을 청취**하였다. 오늘 협의회는 **국정과제**("금융권의 AI 대전환") **이행**, **대통령 업무보고**('25.12.19일) **과제**의 신속한 후속조치 논의를 위해 개최되었다. 발표된 과제 중 금융회사, 핀테크 기업 등의 AI금융서비스 개발·검증을 지원하는 「**금융권 AI 플랫폼**」은 **오늘부터 이용**이 가능하며, 금융소비자의 AI 학습·활용을 지원하는 「**모두의 금융 AI 러닝 플랫폼**」은 **국민 누구나** 내년 **1월5일부터 이용**이 가능하다.

- **(일시/장소)** '25.12.22.(월) 14:00 / 서울 YWCA회관 4층 대강당
- **(참석)** 금융위 부위원장, 업계, 연구원·학계, 유관기관
  - (업계) 신한·하나·우리은행, KB금융지주, 삼성생명·현대해상, NH투자증권, BC카드, 네이버·카카오페이, 비바리퍼블리카, 뱅크샐러드
  - (전문가) 금융연, 자본연, UNIST 등 - (유관기관) 금감원, 신정원, 금보원, 금결원, 연수원
- **(논의내용)** ①금융권 AI 플랫폼 & 모두의 금융 AI 러닝플랫폼 소개, ②금융데이터 결합·활용 지원, ③금융분야 AI 가이드라인 개정방향, ④AI 기본교육 추진계획 등

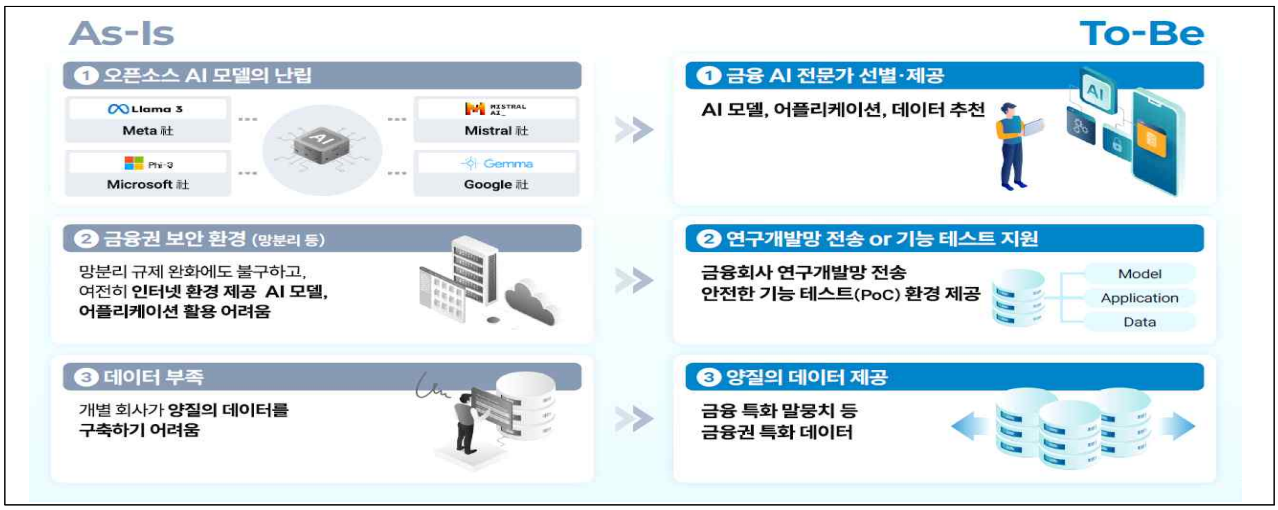
권대영 부위원장은 대한민국이 본격화되고 있는 AI 경쟁을 선도하기 위해서는 정부와 기업이 연구·개발, 교육, 인프라, 정책 및 거버넌스 등 **모든 부문에서 힘을 합친 '총력전'이 필수적**이라고 강조하면서, AI는 방대한 데이터를 활용한 **신용리스크 분석**과 이에 따른 **효율적 자금배분**, 기업들의 **차입비용 축소**, 사기 적발, 금융범죄 예방 등 **금융소비자 보호**에 이르기까지 **AI는 금융의 본질적 역할에 기여**할 수 있다고 강조하였다. 이에 따라, 금융위원회는 **금융권이 AI 대전환을 선도**할 수 있도록 인프라·데이터·교육 지원과 규제 합리화 등 **전방위적인 제도 개선**을 추진할 예정이라고 말했다. 한편, 금융산업에서 인공지능의 도입과 활용이 금융의 본질인 리스크 관리와 소비자 보호를 향할 수 있도록 **AI를 가장 안전하게 잘 활용**할 수 있는 규율체계를 마련하고 금융소비자가 고루 AI 혜택을 누릴 수 있는 **포용적 AI 환경도 조성**할 수 있도록 하겠다고 말했다.

Ⅱ. 주요 논의사항

❶ 「금융권 AI 플랫폼」 및 「모두의 금융 AI 러닝 플랫폼」 소개

신용정보원은 금융회사 대상 AI 플랫폼과 금융소비자 대상 AI 러닝 플랫폼을 운영한다. **「금융권 AI 플랫폼」**은 금융회사, 핀테크기업 등이 자유롭게 AI 서비스를 개발·실험할 수 있는 **AI 인프라**를 지원하기 위해 구축되었다. AI 플랫폼은 AI 투자여력이 부족한 중소형 금융사, 핀테크기업 등이 쉽고 안전하게 AI에 접근할 수 있도록 AI 전문가가 추천한 **금융권에 가장 적합한 AI 모델, 어플리케이션, 데이터를 선별·제공**한다. 또한, 금융회사, 핀테크기업이 안전하게 **AI 서비스를 검증**할 수 있는 **기능테스트 환경**을 제공하며, 전문가 Q&A 등 **커뮤니티 공간, 금융권 우수사례를 소개·공유**하는 쇼룸도 운영한다.

< 금융권 AI 플랫폼 도입효과 >



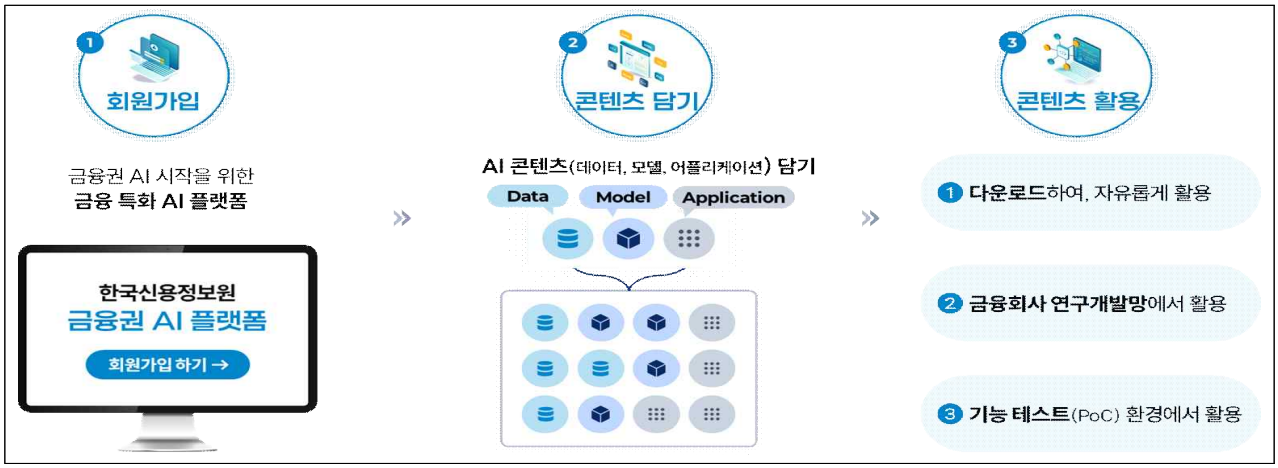
< 금융권 AI 플랫폼 이용 사례(예시) >

✔ **(A저축은행)** 금융권 AI 플랫폼에 추천된 오픈소스 AI 모델과 어플리케이션을 다운로드하여, 저축은행 내 규정 해석 및 보고서 작성을 지원하는 챗봇 구현

✔ **(B카드사)** 금융권 AI 플랫폼에서 제공하는 다양한 데이터셋(고객질의·답변, 법령 설명자료 등)을 업무질의응답, 고객응답을 위한 AI모델 학습·성능평가에 활용

✔ **(C핀테크기업)** '쇼룸'에 자사 AI 에이전트 서비스와 우수 적용 사례를 소개하여, 여러 금융기관과 협업할 수 있는 기회 마련

< 금융권 AI 플랫폼 이용절차 > ※ **finai.kcredit.or.kr**로 접속



또한, 신용정보원은 AI 접근 및 활용기회가 부족한 금융소비자를 대상으로 하는 **AI 러닝플랫폼**도 소개하였다. AI 러닝 플랫폼은 AI·데이터 입문자가 쉽게 다룰 수 있는 **탐색 데이터셋**과 **활용 안내서**를 함께 제공하여 AI 입문자 및 비전문가도 쉽게 **AI 분석 및 모델링을 할 수 있는 환경을 제공**한다. 이에 따라 금융소비자는 AI 러닝 플랫폼 포털을 통해 서비스 이용신청을 하고 신용정보원이 제공하는 원격분석 환경 내에서 **안전하고 편리하게 AI 기반 모델 개발** 및 **아이디어 검증**이 가능하다. 신정원은 이를 통해 금융소비자 누구나 AI를 잘 다룰 수 있도록 지원하는 **"모두의 금융 AI 환경" 구축**에 기여할 것으로 기대한다고 밝혔다.

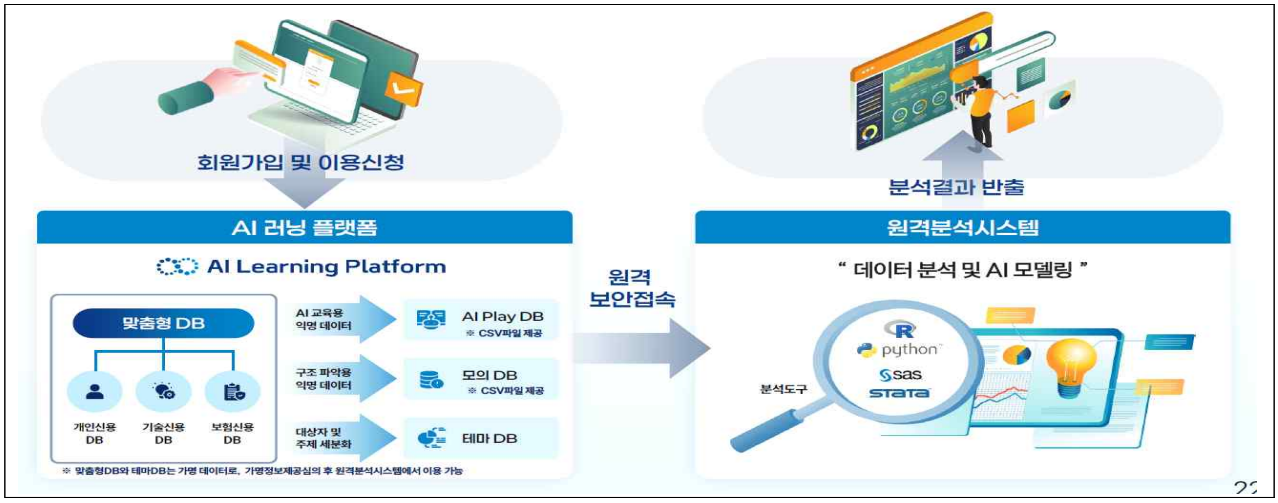
< AI 러닝 플랫폼 이용 사례(예시) >

✔ **(대학생D)** 쉽게 접하기 어려운 대출·연체·보험·카드정보 등 다양한 종류의 금융 데이터셋을 안전한 보안 환경에서 이용하고, 'AI&데이터 활용 안내서'를 기반으로 실습하여 기초 AI 역량 함양 및 AI 전문 인재로 성장 가능

✔ **(금융소비자E)** 'AI Play DB'를 활용하여, 대출·보험 상품의 구조와 신용거래 패턴 등을 면밀히 살펴보고, 대출·보험상품 변경 등 자산관리 능력 제고

✔ **(예비창업자F)** 연령별 신용카드 사용액 수준 등 소비 패턴에 대한 AI 분석 결과를 기반으로 창업 준비 및 사업전략 수립에 활용

< 금융권 AI 러닝플랫폼 개념도 > ※ ailp.kcredit.or.kr로 접속(1.5일~)



❷ 금융데이터 결합·활용 지원방향

< 12.19일 금융위원회 대통령 업무보고 내용 중 >
◇ **금융 AX 확산**을 위해 **결합절차 간소화** 등 **데이터 결합·활용 지원**

대통령 **업무보고 후속조치** 및 **국정과제**("금융권 AI 대전환") 이행을 위한 **금융데이터 결합·활용 지원 세부방향도 공개**되었다. **대량·양질의 데이터 확보는** AI 모델 성능 및 신뢰성에 큰 영향을 미치는데, 금융분야의 고부가가치 데이터는 대부분 개인신용정보를 포함하고 있어 안전한 활용을 위해서는 **가명·익명처리 후 데이터 결합**이 필수적이다. 이에 따라 양질의 AI 학습데이터 확보를 위해 데이터결합·저장방식을 효율화하기 위한 제도개선이 추진될 예정이다.

금융당국은 ①이미지, 영상, 음성 등 **비정형데이터** 및 **합성데이터*** 등 보다 다양한 데이터의 활용기준을 제시하기 위한 **가명·익명처리시 유의사항, 체크리스트** 등을 구체적으로 제시할 계획이다. ②데이터 결합절차 효율화를 위해 주기적·반복적 정보결합의 경우 **데이터 결합 패스트트랙** 등을 통해 데이터 결합시간 단축을 지원할 예정이다. 한편, 데이터전문기관이 결합정보를 안전하게 관리할 수 있는 환경을 갖춘 경우 ③**결합정보를 파기하지 않고 재사용**할 수 있도록 허용하여 결합데이터의 지속적 활용과 데이터의 양적 성장도 도모할 계획이다.

* 실제 데이터의 통계적 특성을 모방하여 만든 가상데이터

오늘 발표한 데이터 결합·활용 지원은 신용정보법 **시행령 및 감독규정 개정**, 금융분야 **가명·익명처리 안내서 개정** 등을 통해 **내년 1분기중 시행**될 예정이다.

❸ 금융분야 AI 가이드라인 개정방향

금융연구원은 금융분야 AI 가이드라인(안) 개정방향에 대한 연구결과를 발표하였다. 그간 금융당국은 「금융분야 AI 운영 가이드라인('21.7월)」, 「금융분야 AI 개발·활용 안내서('22.8월)」, 「금융분야 AI 보안 가이드라인('23.4월)」 등을 마련하여 운영해왔다. 그러나 최근 생성형 AI 등 **새로운 AI 기술의 도입·확산, 인공지능기본법 제정**('25.1월, '26년 1월 시행) 등 기술발전 및 규제환경 변화를 반영한 가이드라인 개정 필요성이 커짐에 따라 **기존 가이드라인을 통합·개정**하고 업무 전반에 걸친 **AI 위험관리의 방향과 원칙을 제시**하고자 한다.

통합 가이드라인(안)은 AI 활용의 7대원칙으로 **①거버넌스, ②합법성, ③보조수단성, ④신뢰성, ⑤금융안정성, ⑥신의성실, ⑦보안성**을 제시하고, 이에 대한 세부이행 사항 등을 제안한다. 가이드라인(안)은 AI 기술의 빠른 발전속도, 금융분야의 AI 수용도, 관련 법·제도 환경변화 등을 고려하여 기존 가이드라인과 마찬가지로 **모범규준(Best Practice), 업권별 자율규제 형식으로 규율**하면서 금융권 의견을 지속 수렴하여 **상시적으로 개선·보완**해나갈 예정이다.

< 금융분야 AI 가이드라인 개정안 주요내용 > ※ 상세내용은 별첨 참고

구분	주요 내용
거버넌스 원칙	①AI 위험관리, 윤리원칙 수립을 위한 최고 의사결정기구 설치, ②독립적 위험관리 전담조직 구축, ③AI 위험평가체계 구축 등
합법성 원칙	AI기본법, 금소법, 신정법·개보법 등 적용법규를 사전파악하여 업무절차에 반영하고 주기적 점검·개선
보조수단성 원칙	①AI 활용에 대한 임직원 책임소재 명확화, ②임직원 개입이 필요한 상황을 차등화 규정하고 개입방법 선택·운영, ③정기적 교육 실시
신뢰성 원칙	①AI모델의 정기적 성능관리, ②AI 학습데이터 품질관리, ③AI의 공정성·편향성 점검, ④이해관계자에 대한 AI의 설명가능성 확보
금융안정성 원칙	①금융안정위험 평가, ②비상정지장치, 백업모형 등 안전장치 마련, ③제3자 IT리스크 관리, ④감독당국 보고절차 마련
신의성실 원칙	①금융소비자 이익 침해 등 이해상충 방지, ②AI 활용시 사전고지 및 서비스 오류 신고, 선택권 보장 등 소비자 보호대책 마련
보안성 원칙	①AI 보안위협 식별·관리, ②AI 특화공격 탐지·대응, ③AI자산 보호·관리, ④외부모델 및 데이터 검증, ⑤보안성 검증·운영 관리

오늘 공개되는 금융분야 통합 AI 가이드라인(안)은 향후 금융권의 의견을 충분히 수렴하고 인공지능기본법 하위법규 및 가이드라인 논의동향을 반영하여 **내년 1분기 중 시행**될 예정이다.

❹ AI 기본교육 추진계획

금융연수원은 금융소비자, 금융회사 종사자를 위한 AI 기본교육 추진계획을 발표하였다. 금융연수원은 **AI 기본사회 구현** 및 **금융소비자 권익 보호**와 금융 분야의 **위험 예방 및 정보격차 해소**를 위해 금융산업 종사자뿐 아니라 금융소비자 대상 AI 교육이 중요해지고 있다고 강조하면서, 금융소비자 대상 AI 기본교육계획안을 발표했다.

AI를 이해하고, 신종 금융사기로부터 **스스로를 보호**하며, AI의 추천·조언을 맹신하지 않고 **최종 결정은 주체적**으로 할 수 있도록 교육하는 금융소비자 대상 AI 기본교육은 동영상 강의 형태로 제작될 예정이며, 내년 2분기 중 **유튜브, 공공기관 채널 등 다양한 채널을 통해 무료 배포**하여 학교 등에서 이루어지는 **청소년의 AI교육에도 활용**할 수 있도록 할 예정이다.

❺ AI를 활용한 포용금융 확대

한편, **UNIST 이용재 교수**는 포용금융을 위해 활용될 수 있는 AI서비스 관련 연구를 소개하였다. 이 교수는 금융사기 등 다양한 사회문제에 대해 AI를 활용하여 해결방안을 모색할 수 있는데, 이를 위해서는 먼저 **여러 부처·기관에 흩어진 데이터를 안전하게 결합하여 분석**할 수 있는 환경을 구축해 나가야 한다고 강조하였다.

금융위원회는 오늘 논의된 전문가 의견을 적극 반영하여 **발표 과제를 신속히 이행**하고, AI를 활용한 대안신용평가, 금융사기 방지 등 금융소비자가 체감할 수 있는 **금융분야 AI 혁신을 지원**하는 한편, **안전한 AI 규율체계가 확립**될 수 있도록 관련 제도를 지속 정비해나갈 계획이다.

[별첨1] 금융위원회 부위원장 모두발언
[별첨2] 금융권 AI 플랫폼 및 모두의 금융 AI 러닝플랫폼 소개(신용정보원)
[별첨3] 금융데이터 결합·활용 지원방향(금융감독원)
[별첨4] 금융분야 AI 가이드라인 개정방향(금융연구원)
[별첨5] AI 기본교육 추진계획(금융연수원)

담당 부서	금융위원회 금융데이터정책과	책임자	과 장	서나윤 (02-2100-2620)
		담당자	사무관	권진웅 (02-2100-2625)
			사무관	김예빈 (02-2100-2622)
<공동>	금융감독원 디지털금융총괄국	책임자	국 장	위충기 (02-3145-7120)
		담당자	팀 장	이우람 (02-3145-7135)
			팀 장	김은철 (02-3145-7141)
<공동>	한국신용정보원 금융AI데이터센터	책임자	상 무	임충태 (02-3705-5810)
		담당자	센터장	정권수 (02-3705-5433)
<공동>	금융연수원 AI디지털연수부	책임자	본부장	문영성 (02-3700-1504)
		담당자	부 장	유승엽 (02-3700-6851)

대한민국 정책브리핑
OPEN 출처표시 공공누리 공공저작물 자유이용허락

참고 금융분야 AI 가이드라인 개정안 주요내용

1. 적용대상 · 방식

□ **(적용대상)** 업종·업무에 관계없이 **AI를 도입·활용하는** 모든 금융회사 및 **비금융회사**(예: 핀테크기업)가 **가이드라인 적용대상**에 해당 가능

○ **고영향·고성능 AI 사업자**를 주로 규율하는 AI기본법과 달리 **규율 대상을 포괄적으로 규정**

□ **(적용방식)** 개별 회사가 **AI 활용수준, AI 활용에 따른 영향, 인적·물적 자원** 등을 **종합 고려**하여 자율적으로 가이드라인 적용 가능

○ **고영향 AI**의 경우 AI기본법상 **고영향 AI 사업자**의 **책무**를 **이행**하고, 추가적으로 **금융분야 AI 통합 가이드라인**의 내용을 **적용 가능**

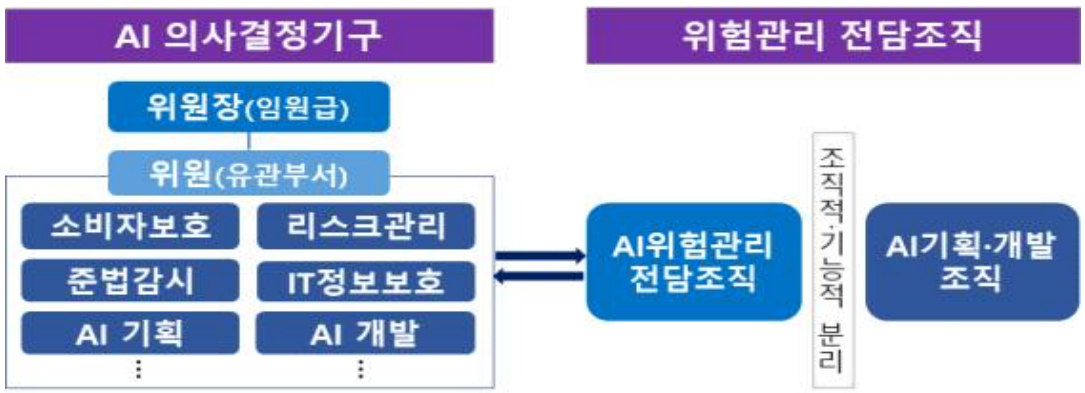
○ **고영향 AI 이외**의 경우에도 **소비자 보호** 및 **금융안정성 강화** 등을 위해 **금융 AI 7대 원칙**에 기반한 동 **가이드라인**의 내용을 적용 가능

2. 주요내용

1 **(거버넌스)** 최고 경영자를 포함한 경영진은 **AI 개발·활용에 대한 관심**을 갖고 **역할과 책임 분담** 필요

① (AI 거버넌스 체계) **AI 위험 통제·관리, 윤리원칙 수립** 등을 위한 **의사결정기구를 설치**하고, **독립적 위험관리 전담조직을 구축**하며, AI 도입·활용 등 **전체 프로세스 관리**를 위한 **내규·매뉴얼 마련**

예시 AI 거버넌스 체계 구성



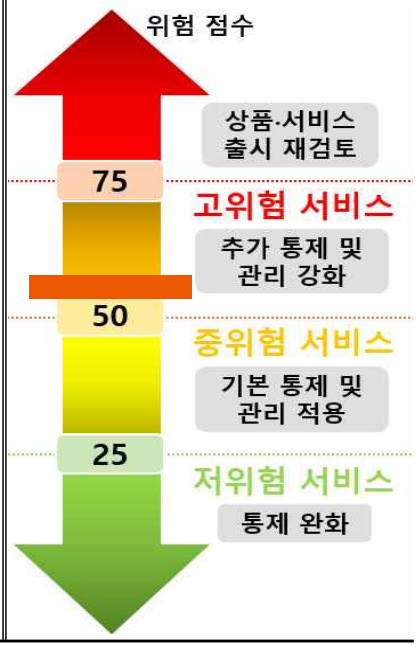
② (위험평가체계) AI 원칙의 **위반 가능성**을 점수화하여 **AI 위험을 평가·경감 후 위험등급을 산정**하는 등 **종합적 위험평가체계*** 구축

* AI 서비스별 위험등급을 기준으로 차등화된 통제·관리를 수행하고, 모니터링 및 사후관리, 문서화·교육, 감독당국 정보공유 등 위험통제를 위한 제반절차 이행

예시 **AI 위험평가 항목 및 등급 분류 예시(회사별 조정 가능)**

부문	평가항목 및 배점		위험경감	잔여 위험		
합법성 원칙 (20%)	금융소비자보호법 위반 가능성	8	(4)	4	∑ 9	총합 54점
	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	∑18	
	편향성	6	(2)	4		
	공정성	6	(2)	4		
	설명가능성	6	(1)	5		
	성능	6	(3)	3		
신의성실원칙 (20%)	계약 권리 침해	6	(3)	3	∑10	
	책임 투명성	6	(3)	3		
	소비자 보호방안	8	(4)	4		
보안성 원칙 (30%)	보안	8	(3)	5	∑17	
	안정성	8	(4)	4		
	위탁/관리	8	(3)	5		
	프라이버시	6	(3)	3		

위험등급의 확정



※ 거버넌스, 위험평가 및 위험통제 등의 구체적인 내용은 향후 금융감독원이 공개 예정인 「금융분야 AI 위험관리 프레임워크」(AI Risk Management Framework in Financial Sector)를 참고 가능

2 **(합법성)** AI 개발·이용시 **적용 법규를 사전파악**하여 내부정책 및 **업무절차에 반영**하고 **주기적인 점검·개선** 필요

예시 **AI 개발 · 이용시 검토 필요 법규**

① **(AI기본법) 사전 고지의무, 고영향 AI 사업자 책무*, AI 영향평가** 등 **준수** 필요

* ① **위험관리방안** 수립·운영, ② 인공지능 도출결과에 대한 **설명 방안** 수립·시행, ③ **이용자 보호방안** 수립·운영, ④ 고영향 인공지능에 대한 **사람의 관리·감독**, ⑤ **안전성·신뢰성 확보**를 위한 조치내용 **확인문서의 작성과 보관**

② **(금소법)** AI 활용과정에서 설명의무, 적합성·적정성원칙, 부당권유금지, 불공정영업행위 금지, 광고규제 등 **금융소비자 보호의무 준수** 필요

③ **(기타 금융법령) 정보보호 의무**(개보법·신정법), **정보보안 의무**(전금법), **로보어드바이저 의무**(자본시장법) 등 준수 필요

3 **(보조수단성)** 現단계에서 **AI는 업무의 보조수단**이므로 최종 의사결정과 그에 따른 책임은 **임직원이 수행** 필요

① (책임 명확화) AI 시스템에 대한 **내부통제 체계를 구축**하여 인공지능 활용에 대한 **책임소재 명확화**

② (인적개입 원칙) AI 시스템 운영 전 단계에 걸쳐 임직원 개입이 필요한 **상황을 차등화하여 사전에 규정**하고 **개입방법**을 **선택·운영**

<예시 : 인공지능 의사결정에 대한 인적 개입 방법>

구분	주요 내용 및 설계 방법	고위험 인공지능
승인 및 기록 보존	○ (내용) 사람의 검토 또는 승인을 거쳐 실행 ○ (방법) 전자승인·제3자검토·윤리위원회 심의 등 ○ (기타) 사유·근거 기록 의무화	필수 적용
인공지능 권고 재정의 (Override)	○ (내용) 인공지능 권고를 무시·수정·반대 결정을 할 수 있도록 관리자에게 권한 부여 및 이행 수단 확보 ○ (방법) 오버라이드 실행가능한 시스템 구축 ○ (기타) 사유·증빙 등을 표준화하여 기록 보존	필수 적용
긴급 정지 기능 및 격리	○ (내용) 이상 또는 긴급 상황 발생시 해당 인공지능 시스템을 즉시 중단하고 문제 요소 격리 ○ (방법) 긴급 정지. 안전모드 전환 등의 기능을 인공지능 운영 시스템에 반영 (직관적 UI 제공 필수)	필수 적용
실시간 모니터링 대시보드	○ (내용) 성능, 공정성, 신뢰도, 규제 위반 등의 신호를 실시간으로 시각화 ○ (방법) 임계치 초과 시 자동 알림·보고	권고

③ (정기적 교육) 보조수단성 원칙이 효과적으로 준수되도록 **업무담당자 및 감독자** 등을 대상으로 **정기적인 교육 실시** 필요

4 **(신뢰성)** AI 개발 과정에서 **신뢰할 수 있는 데이터·모델 사용** 필요

① (모델 성능관리) AI 모델의 성능을 측정할 수 있는 **명확한 지표를 설정**하고 **정기적 점검·개선**

② **(데이터 품질관리)** AI **학습 및 참조**에 사용하는 **데이터**, AI **시스템에 입력**되는 **데이터 품질*** **검증·확인**

* 데이터 품질관리 5대 기준 : 정확성, 완전성, 일관성, 대표성, 적시성

③ **(공정성·편향성 점검)** AI 서비스가 모든 집단에 대해 **차별없이 공정하게 작동**하도록 **데이터·모델을 분석·개선**

④ (설명가능성 확보) AI 의사결정 과정과 결과에 대해 **이해관계자가 합리적으로 이해**할 수 있도록 **설명 가능한 형태**로 **서비스 제공**

< 예시 : 대출심사시 설명가능한 AI 서비스 제공사례 >

설명가능한 AI 시스템 생성 자료

- 결과: 거절
- 이용된 기초정보(해당 고객의 주요 원본 정보):

항목명	설명	값
제1금융권 대출 건수	제1금융권 대출 건수	18
리볼빙 잔액 비율	신용한도 대비 리볼빙 잔액 비율	63
거래 연체 비율	연체되지 않은 거래 비율	94
상환 비율	과거 정상 상환된 신용 거래 비율	96
... (이하 생략) ...		

- 주요 판단 사유: 제1금융권 대출 건수(부정적), 리볼빙 잔액 비율(부정적), 거래 연체 비율(긍정적), 상환 비율(긍정적)

자동 설명 생성 ↓ 고객 대응 업무 담당자가 해석하여 설명 생성

고객 제공 설명

고객님은 연체되지 않은 거래 비율 94% 및
과거 정상 상환된 신용 거래 비율 96%가
긍정적인 요인이었으나,

제1금융권 대출 건수 18건 및
신용한도 대비 리볼빙 잔액 비율 63%가
부정적으로 작용하여 **대출이 거절**되었습니다.

5 **(금융안정성)** AI 개발·학습 등 全 과정에서 **금융 안정성 위험 최소화**

① (금융안정위험 평가) AI위험 평가시 전체 금융시스템이나 금융시장 전반에 미칠 수 있는 **위험 평가** 및 **관리방안 마련**

② (안전장치 마련) AI모형 오작동시 **백업모형 활용**, 사후 개입이 가능한 **비상정지장치**(Kill Switch) 등 시스템 위험관리장치 마련

③ (제3자 IT리스크 관리) AI 모델을 **외주개발**하거나 **오픈소스** 기반 AI를 활용하여 내부에 **AI 시스템**을 **구축**하는 경우 별도 **관리방안 마련**

예시 AI 관련 제3자 계약 체결 시 포함 사항

① 제3자가 제공하는 서비스 등에 대해 기능, 위험 등에 대한 명확한 설명
② 사이버 보안, 데이터 및 프라이버시 보호에 관한 내용
③ 서비스를 제공하는 지역, 국가 장소
④ 제3자의 파산 및 사업 운영 중단 또는 계약 종료 시 해당 서비스, 데이터 등에 대한 금융회사 등의 접근 및 복구, 반환의 보장에 관한 내용
⑤ 사고 발생 시 정보 공유 및 사고처리, 배상책임과 관련한 내용
⑥ AI 관련 피해 발생 시 책임 소재
⑦ AI 관련 법규 명령 및 정책 등의 준수 의무
⑧ 안전하고 신뢰할 수 있는 AI 개발·활용과 관련 금융회사 등의 행동강령, 정책 등 준수

④ (감독당국 정보공유 및 보고) 시스템위험을 초래할 수 있는 **AI 사고** 발생 시 **감독당국에 보고**하고, **AI 활용구조**를 **감독당국**이 사전에 **파악**할 수 있도록 **정보 공유** 필요

6 **(신의성실)** AI 활용시 **금융소비자 이익을 최우선 목표**로 관리

① (이해상충 방지) 대고객 서비스에서 AI 모형 활용시 금융소비자 이익 침해 등 **이해상충 문제가 발생**하지 않도록 **점검** 및 **예방조치 마련**

② **(소비자 보호대책 마련)** AI가 허위광고, 불완전판매 등 금융소비자의 신뢰를 저해하지 않도록 AI 활용시 금융소비자에게 **사전고지** 및 **서비스 오류 신고절차 마련, 소비자 선택권 보장 절차** 마련

예시 인공지능 관련 금융소비자 보호대책

- 인공지능 서비스 제공 시 해당 서비스가 인공지능을 활용하고 있음을 사전 고지
- 인공지능의 성능을 부풀리거나 인공지능을 활용하지 않음에도 인공지능을 활용하는 것처럼 속이는 거짓·과장 광고 금지
- 금융소비자가 인공지능 서비스 오류 등 위험을 신고할 수 있는 창구 마련
- 금융소비자가 인공지능 서비스에 대해 이의를 제기하고 인공지능이 아닌 기존 서비스 이용을 요구하는 경우 제공이 가능한 체계 마련 등

7 **(보안성)** AI 활용시 **보안성 기준** 및 **점검·개선체계 마련**

① (보안위협 식별·관리) 전통적 보안위협과 별개로 AI 시스템에 **특화된 보안위협**을 체계적으로 **식별·관리**

예시 AI 활용에 따른 주요 보안 위협 요인

▶ **(고도화된 피싱 공격)** 생성형 AI로 제작한 가짜 영상, 이미지 및 오디오로 피해자가 신뢰하는 지인들을 흉내내어 더욱 식별하기 어려운 보이스피싱 등 수행, 생성형 AI를 활용하여 피해자의 SNS 내용 등을 통해 맞춤형 피싱 메일 제작 자동화, 대규모 피싱 공격 수행

▶ **(악성코드 생성 증가 및 고도화)** 특수 제작된 생성형 AI를 악용하여 악성코드를 대량 생산 가능, 생성형 AI를 사용하여 기존 악성코드 탐지 시스템을 우회하는 등 고도화 가능

▶ **(중요 정보 유출)** 금융회사 내부에서 상용 생성형 AI를 사용할 때 고객 정보나 기밀 정보를 활용할 경우 데이터 유출 위험에 노출, 공격자가 금융회사가 제공하는 생성형 AI 서비스를 대상으로 적대적 공격을 통해 내부 데이터 추출 가능

▶ **(공급망 보안 이슈)** 상용 또는 오픈 생성형 AI 모델을 사용할 때, 해당 상용 모델 공급자 및 오픈소스 관련 제3자 위협에 노출 가능

▶ **(생성형 AI 모델 및 출력 조작)** 모델의 학습데이터를 조작하여, 잘못된 정보를 바탕으로 투자를 추천하는 등 모델이 부정확하고 공격적이거나 해로운 출력 가능

② (AI특화공격 탐지·대응) 식별된 AI 특화 보안위협과 관련된 공격에 대해 **탐지, 차단할 수 있는 대응체계** 구축

③ (AI자산 보호·관리) 데이터, 모델 파라미터 등 **핵심 자산이 무단 접근·유출·변조**되지 않도록 **암호화, 무결성 검증, 접근통제** 등 적용

④ (외부모델 및 데이터 검증) 외부에서 도입하는 모델·데이터에 대해 **보안 및 신뢰성 검증을 수행**하여 **공급망 위험 최소화**

⑤ (기존 보안관리 확장) 출력조작, 고도화된 피싱위험 등 AI 시스템의 특성에 맞게 **접근통제 및 권한을 관리**하고, **네트워크 보안 및 데이터 보호** 등 기존 IT 보안체계를 **확장적용**

⑥ (보안성 검증·운영 관리) AI시스템의 보안성을 **개발 단계부터 체계적으로 검증**하고 **운영과정**에서 **지속 관리**