

금융권의 AI 활용을 적극 지원하겠습니다.

- GPT-4 등 인터넷 환경에서 제공되는 상용 AI 활용을 허용하는 「금융분야 망분리 개선 로드맵」 발표에 이어, 금융회사 내부망에 직접 설치하는 오픈소스 AI 활용 지원으로 금융권 AI 이원(Two-track) 활용 체계 마련
- 금융회사 내부망에 오픈소스 AI 설치·활용을 지원하는 플랫폼 구축, 금융 분야 한글 말뭉치 구축 등 「금융권 생성형 AI 활용 지원 방안」 발표

'24년 12월 12일(목), 금융위원회는 김소영 부위원장 주재로 「금융권 AI 협의회」를 개최하였다. GPT-4 등 상용 AI 활용을 허용하는 「금융분야 망분리 개선 로드맵」('24.8.13.) 발표에 이어, 이번 협의회에서는 금융회사의 내부망에 직접 설치하는 오픈소스 AI 활용 지원을 포함한 「금융권 생성형 AI 활용 지원 방안」을 발표하였다.

【「금융권 AI 협의회」 개요】

- 일시·장소 : '24.12.12(목) 10:00, 은행회관 14층 세미나실
- 참석자 : (금융위) 부위원장(주재), 디지털금융정책관
(유관기관) 금융감독원, 신용정보원, 금융보안원, 금융결제원,
금융연수원, 보험연수원
(학계) 금융연구원, 자본시장연구원, UNIST
(업계) 하나은행, 신한은행, KB증권, 미래에셋자산운용,
삼성생명, 현대해상, BC카드 등

김소영 부위원장은 이날 모두발언에서 “국내 금융회사들은 AI 인프라와 데이터 부족, 생성형 AI 활용에 대한 명확한 거버넌스 부재 등 애로사항을 제기해 왔다”며, “이러한 의견을 종합하여 금융권 AI 활용 인프라, 금융권 특화 데이터 지원, 금융분야 AI 가이드라인 개정 등 금융회사에 실질적으로 필요한 지원 체계를 갖출 것”이라고 밝혔다.

< 추진 배경 >

GPT 등 새로운 콘텐츠를 생성할 수 있는 생성형 AI 등장에 따라 금융권에서도 AI에 대한 관심이 지속적으로 증가하고 있다. 생성형 AI는 인터넷망에서 제공되는 **상용 AI**와 회사 내부 시스템에 설치하는 **오픈소스 AI**로 구분되며, 국내 금융회사들은 AI 활용 목적, 비용 효율성 등을 고려하여 두 가지 모두를 사용하는 방안을 고민하고 있다.

예를 들어, **상용 AI**는 다양한 목적에 대해 높은 성능을 제공하나 주로 금융회사 통제권 밖의 클라우드 환경에서 제공되므로 개인신용정보를 활용하지 않는 일반적인 고객상담이나 투자정보 분석 등에 활용할 수 있다.

한편 **오픈소스 AI**는 작은 컴퓨팅 자원에서도 동작이 가능하고 금융회사 내부망에 직접 설치하여 보안성과 비용 효율성에 장점이 있으므로 내부 정보 등을 활용하는 내부규정 챗봇이나 홍보물 등 콘텐츠 제작에 활용할 수 있다.



이에, 금융위원회는 금융회사들이 서비스 목적에 따라 다양한 방식으로 생성형 AI를 활용할 수 있도록 금융권 AI 이원(Two-track) 활용 체계를 마련하고자 한다. **상용 AI**는 지난 8월 13일 발표한 「금융분야 망분리 개선 로드맵」에 따라 규제 샌드박스를 폭넓게 허용하는 한편, 이번 협의회에서는 **오픈소스 AI**를 금융회사 내부망에 손쉽게 설치하고 활용할 수 있도록 종합적인 지원 방안을 마련하였다.

< 주요 내용 >

첫째, 금융권 생성형 AI 이원(Two-track) 활용 체계를 구축한다.

현재 국내 금융회사들은 망분리, 자체 보안규정 등으로 인해 AI 서비스 개발을 위한 오픈소스 AI 모델, 데이터 등을 다운로드 받아 내부망에 설치하는 데 제약이 있다. 또한 오픈소스 AI 모델은 누구나 수정 및 재배포가 가능한 특성으로 인해 수많은 변형 모델*이 난립하여, 금융회사가 AI 모델의 성능 및 안전성을 검증하고 사용하는 데 어려움이 있다.

* 세계 최대 AI 플랫폼 '허깅페이스'에 등록된 오픈소스 AI 모델은 약 110만 개

또한 생성형 AI 서비스 개발은 초기 비용이 매우 많이 들어 실제 개발에 착수하기 전에 기술의 실현 가능성을 미리 테스트(PoC*)하는 절차가 필수적임에도, 테스트 환경과 자원이 부족하다는 애로사항도 제기되었다.

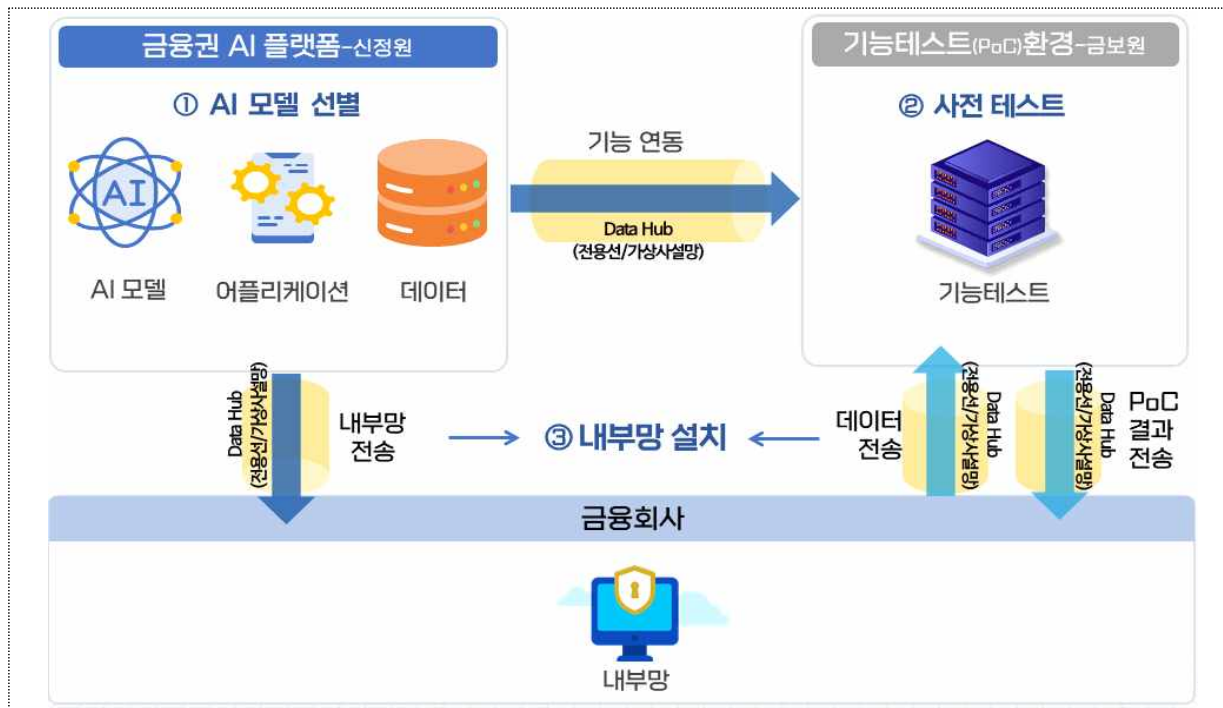
* PoC(Proof of Concept) : 실제 서비스 개발을 시작하기 전, 서비스·기술 등의 실현 가능성을 입증하는 절차

이러한 문제점을 해소하기 위해 금융권 오픈소스 AI 서비스 개발 및 활용을 통합 지원하는 「금융권 AI 플랫폼」을 구축한다. 플랫폼에서는 금융분야에 적합한 성능과 안전성을 지닌 오픈소스 AI 모델, 데이터 등을 전문가 그룹이 선별하여 제공한다.

이와 함께, 금융회사들이 다양한 오픈소스 AI 모델, 어플리케이션, 데이터를 활용하여 최적의 조합을 탐색하고 혁신적인 AI 서비스 아이디어를 실험해 볼 수 있는 기능테스트(PoC) 환경과, AI 모델, 데이터 등을 금융사 내부망에 쉽게 설치할 수 있도록 지원하는 인프라도 구축할 예정이다.

「금융권 AI 플랫폼」 구축으로 금융권 AI 이원(Two-track) 활용 체계가 마련됨으로써 금융회사들이 상용 AI와 오픈소스 AI를 전략적으로 선택하여 활용하고, 이를 통해 보다 전문화되고 특화된 AI 서비스를 개발할 수 있을 것으로 기대된다.

< 금융권 AI 플랫폼 구축 개요도 >



둘째, 금융분야 AI 학습 등을 위한 금융권 특화 데이터를 구축한다.

현재 Llama3(Meta社) 등 주요 오픈소스 AI 모델들은 주로 영미권 언어와 일반적인 데이터를 학습하여 한국어 능력이 부족하고 금융분야에 대해서는 전문성이 결여된 답변을 제공하는 한계가 있다.

협의회에서도 금융 전문성을 갖춘 AI 서비스를 개발하기 위해 금융 특화 한글 빅데이터 지원을 요청하는 목소리가 많았다. 또한 이상거래 탐지(FDS), 신용평가, 금융보안 등 공익 목적의 AI 데이터 활용을 위한 주기적인 공급 체계도 부재하다는 의견이 제기되었다.

이에, 금융위원회는 업권별 협회, 금융연수원, 보험연수원 등 유관기관과 협력하여 생성형 AI 모델을 학습 및 검증하는데 필요한 ‘금융권 특화 한글 말뭉치*’를 구축하고 「금융권 AI 플랫폼」을 통해 제공할 예정이다.

* 생성형 AI의 언어 학습 등 특정한 목적을 위해 구축한 대규모 텍스트 집합

금융 법규 및 가이드라인, 업권별 보도자료, 금융연수원·보험연수원의 연수 자료 등을 기반으로 금융권이 공동으로 활용할 수 있는 양질의 데이터를 확보하고, 생성형 AI 개발과 활용에 필요한 다양한 형태로 제공한다.

특히 생성형 AI의 정확한 답변을 유도할 수 있는 검색 증강 생성(RAG*)용 데이터, AI의 성능 및 윤리 평가에 활용할 수 있는 평가지원 데이터에 대한 수요가 큰 만큼, 우선 확보되는 공개 데이터를 기반으로 신속하게 구축해 나갈 계획이다.

* RAG(Retrieval-Augmented Generation) : AI 모델이 외부 특화 데이터, 최신 정보 등을 검색한 후 결과물을 생성하도록 하여 정확도와 신뢰도를 향상시키는 방법

아울러, 현재 금융사기방지(금융결제원), 신용평가(신용정보원), 금융보안(금융보안원) 데이터 등 유관기관에서 제공하는 공익 목적 데이터의 제공 채널을 「금융권 AI 플랫폼」으로 일원화한다. 이를 통해 금융회사들이 이상거래 탐지(FDS) 등 특성화 AI 개발을 위한 데이터를 AI 플랫폼을 통해 정기적으로 편리하게 확보할 수 있을 것으로 기대된다.

셋째, 「금융분야 AI 가이드라인」 개정을 추진한다.

금융위원회는 금융권 AI 활용을 활성화하기 위하여 그간 가이드라인, 안내서 등을 제시*해왔으나, 생성형 AI 등 급격한 기술 발전과 금융권 내부통제 강화 등 제도 변화에 따라 개정이 필요한 상황이다. 특히, 많은 금융회사들이 AI 활용에 대한 명확한 거버넌스를 요청하였다.

* 「금융분야 AI 운영 가이드라인」(‘21.7월), 「금융분야 AI 개발·활용 안내서」(‘22.8월), 「금융분야 AI 보안 가이드라인」(‘23.4월)

금융위원회는 이러한 의견들을 반영하여 수차례 협의회 논의를 통해 금융권 AI 개발·활용의 주요 원칙을 마련하였다.

< 금융 AI 7대 원칙(안) >

분야	7대 원칙
거버넌스	① 최고경영자를 포함한 경영진은 AI 개발·활용에 대한 관심을 갖고 역할과 책임을 분담해야 함
	② AI 활용 전단계에서 금융·AI 등 관련 법규를 준수해야 함
	③ 현 단계에서 AI는 업무의 보조 수단이므로 최종 의사결정과 그에 따른 책임은 임직원이 수행함
AI 개발 단계	④ AI 개발 과정에서 신뢰할 수 있는 데이터와 모델을 사용해야 함
	⑤ AI 설계·학습 등 전과정에서 금융 안정성 위험을 최소화해야 함
AI 활용 단계	⑥ AI 활용 시 금융소비자의 이익을 최우선으로 해야 함
	⑦ AI 활용 시 보안성 기준 및 점검·개선 체계를 마련해야 함

금융분야 AI의 거버넌스와 개발·활용에 관한 명확한 원칙을 제시하고, 원칙에 대한 구체적인 내용들은 향후 유관기관, 업계 전문가 등과 논의를 거쳐 안내서로 제시할 계획이다. 안내서에는 금융 AI 7대 원칙의 적용 기준, 생성형 AI 관련 윤리 등을 반영한 자세한 설명과 사례를 담을 예정이다. 이번 「금융분야 AI 가이드라인」 개정으로 금융회사들이 실제 업무에 AI를 활용할 때 불확실성이 상당 부분 해소될 것으로 기대된다.

발제 이후 이어진 자유토론에서 신한은행 임은택 상무는 “금융권 AI플랫폼이 구축된다면 검증된 AI모델과 학습데이터를 활용할 수 있게 되어 큰 도움이 될 것으로 기대가 된다”고 말했다.

또한, 미래에셋자산운용 최용민 이사는 “AI 인력과 인프라가 여전히 부족한 금융회사들이 많은 가운데, 도메인 특화 데이터를 기반으로 한 금융권 AI 플랫폼이 제공하는 정보와 테스트 인프라는 금융권 전반에 효과적인 지원이 될 것”이라고 기대감을 나타냈다.

보험연수원의 배병한 본부장은 “금융권 AI의 성능은 데이터가 좌우한다고 해도 과언이 아니며, 금융권 특화 데이터 구축을 위해 보험연수원도 연수자료 공유 등에 있어 적극 협조하겠다”고 말했다.

〈 향후 추진 계획 〉

금융위원회는 이번 「금융권 생성형 AI 활용 지원 방안」의 신속한 추진이 가장 중요하다는 금융회사들의 요청에 따라 과제들을 속도감 있게 진행해 나갈 예정이다.

금융권 오픈소스 AI의 설치·활용을 지원하는 「금융권 AI 플랫폼」은 '25년 상반기 구축을 목표로 추진하고, 금융권 특화 데이터 구축은 금융 법규 등 공개 데이터를 기반으로 '25년 1분기에 지원을 시작하여 업권별 보도자료, 연수자료 등으로 단계적으로 확대해 나간다.

더불어, 금융분야 AI의 저변 확대를 위하여 금융회사 임직원뿐만 아니라 소비자, 금융 취약계층 등에 대한 AI 리터러시 교육 지원 방안도 금융연수원, 보험연수원, 업권별 협회 등과 지속 소통할 계획이다.

[참고] 인공지능 관련 용어 설명

[별첨] 1. 금융위원회 김소영 부위원장 모두 발언

2. 「금융권 생성형 AI 활용 지원 방안」

담당 부서	금융위원회 금융데이터정책과	책임자	과 장	신상록 (02-2100-2620)
		담당자	사무관	김하영 (02-2100-2696)
<공동>	금융감독원 디지털혁신국	책임자	국 장	곽범준 (02-3145-7160)
		담당자	팀 장	심은섭 (02-3145-7162)
<공동>	금융연수원 종합기획부	책임자	부원장	김영태 (02-3700-1503)
		담당자	부 장	문영성 (02-3700-1511)
<공동>	보험연수원 전략사업부	책임자	본부장	배병한 (02-920-0803)
		담당자	부 장	이우호 (02-920-0860)
<공동>	한국신용정보원 금융AI데이터센터	책임자	전 무	김평섭 (02-3705-5804)
		담당자	센터장	이철흠 (02-3705-5433)
<공동>	금융보안원 AI혁신실	책임자	상 무	박진석 (02-3495-9004)
		담당자	실 장	김성웅 (02-3495-9960)
<공동>	금융결제원 금융데이터융합센터	책임자	전 무	류재수 (02-531-1002)
		담당자	센터장	김옥선 (02-531-1650)



- (인공지능의 개념) 컴퓨터 등 기계를 활용하여 인간의 인지, 학습, 추론 능력 등을 인공적으로 구현하는 기술
 - (분석형 AI) 머신러닝* 등을 활용하여 데이터를 학습한 후 향후 발생할 수 있는 상황을 분류·판단하는데 중점을 둔 AI 유형
 - * 사람이 직접 프로그래밍하지 않아도 컴퓨터가 데이터를 통해 스스로 학습하고, 그 학습을 기반으로 새로운 작업이나 문제를 해결하는 기술
 - (생성형 AI) 텍스트, 이미지, 미디어 등 기존 데이터를 학습한 후 새로운 콘텐츠를 생성·구축하는데 중점을 둔 AI 유형
- 가장 널리 쓰이고 있는 생성형 AI 기술은 '거대언어모델'(LLM, Large Language Model)이며 상용 LLM과 오픈소스 LLM으로 분류
- (거대언어모델) 인간의 언어를 이해·학습하며 생성할 수 있도록 대규모의 텍스트 데이터로 훈련된 생성형 AI 모델
 - (상용 모델) AI 전문기업(OpenAI, Google 등)에서 최신 기술과 대규모 데이터 셋을 활용하여 상업적으로 개발한 모델
 - 대형 모델로 기본 성능이 우수하나, 사용자 필요에 맞는 모델 수정이 어렵고 서비스 활용 시 입력 정보가 제공기업으로 전송
 - * 예) GPT-4(오픈AI社), HyperClova(네이버社), Claude(앤티트로픽社), Gemini(구글社)
 - (오픈소스 모델) 누구나 다운로드 받아 사용할 수 있는 공개 모델로 금융회사 내부망에 설치·활용할 수 있으며, 상대적으로 작은 컴퓨팅 자원으로도 활용이 가능한 모델
 - 상용 모델 대비 모델의 크기가 상대적으로 작으나, 수정이 용이하여 최적화에 유리하고 외부로 정보 전송 없이 활용 가능
 - * 예) LLama(메타社), Grok(xAI社), Gemma(구글社)