금융 AI 협의회 발족식 발제자료 (2024.3.28.)

금융미래를 열어가는 금융보안파트너

생성형 AI의 소개 및 활용시 규제 관련 논의사항

금융보안원
FINANCIAL SECURITY INSTITUTE

AI혁신실 실장 김성웅(swkim@fsec.or.kr)

Table of Contents

금융보안원
FINANCIAL SECURITY INSTITUTE

1. 생성형 AI 개요

1. 생성형 AI 개요

금융보안원 FINANCIAL SECURITY INSTITUTE

인공지능(AI)

일반적으로 인간 지능이 필요한 작업을 기계가 수행할 수 있도록 하는 컴퓨터 과학의 모든 분야를 포괄하는 광범위한 용어. 기계 학습(ML), 생성형 AI는 AI의 하위 범주

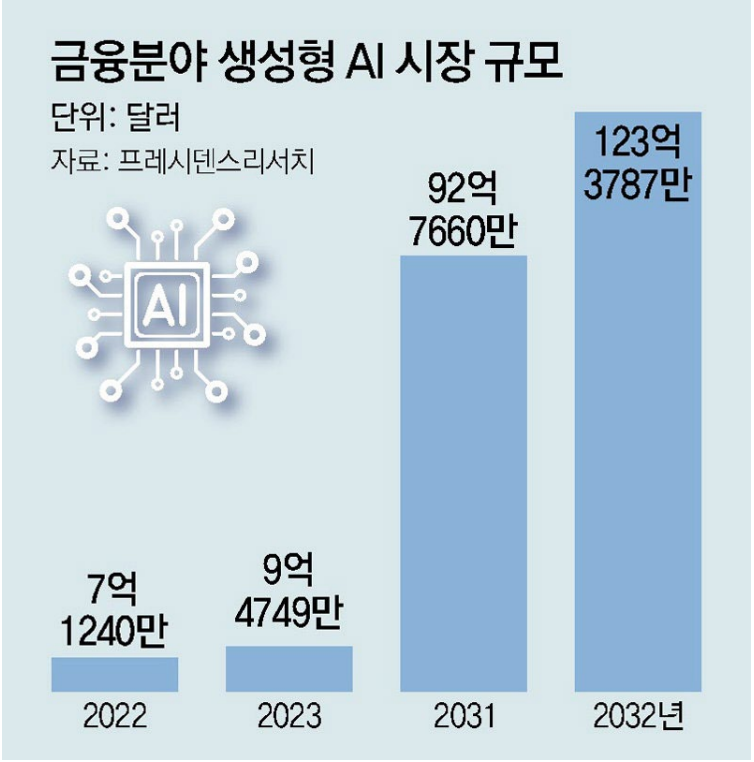
판단형 AI

의사결정 등에 필요한 데이터에 대해 학습한 후 이를 사용하여 향후 발생할 수 있는 상황에 대한 판단을 하는데 중점을 둔 AI의 유형

생성형 AI

텍스트, 이미지 등 다양한 데이터에 대해 학습한 후 기존에 없던 새로운 데이터를 생성하는데 중점을 둔 AI의 유형

Chat GPT 등장으로 금융권에 촉발된 생성형 AI 도입 및 활용 열풍



* 동아일보, "AI는 생산성 높여줄 조력자… 인구감소 韓에 해결책 될 것"('23.11.10.)

1. 생성형 AI 개요

금융보안원
FINANCIAL SECURITY INSTITUTE

파운데이션 모델

일반상식, 과학, 문학 등 방대한 양의 데이터를 기반으로 학습된 다양한 작업에 적용가능한 기초 생성형 AI 모델로, 추가 기법을 이용하여 사용자의 요구사항에 맞게 변형하는 게 가능

상용 모델

일반 AI 기업에서 상업적으로 개발 및 유지 관리하는 모델로, 특정 서비스나 애플리케이션에 사용될 목적으로 만들어지며, 대개 사용 비용을 지불해야 함. GPT-4와 같은 모델이 있음

오픈 모델

누구나 다운로드를 받아 사용할 수 있는 공개 AI 모델로, 사용자가 모델을 구동할 수 있는 컴퓨팅 인프라만 갖추고 있다면 활용 가능함. LLaMA2 같은 모델이 있음

1. 생성형 AI 개요

금융보안원 FINANCIAL SECURITY INSTITUTE

생성형 AI를 실무에 활용하기 위해서는 별도의 기법들을 추가 적용하는 작업이 필요함

검색증강생성(Retrieval-Augmented Generation)

- 사용자의 질의와 관련된 데이터베이스나 문서 집합 등에서 검색하고, 그 정보를 AI가 참고할 수 있도록 전달하는 기법
- 직장인이 보고서를 쓰기 전에 관련 자료를 찾아보고 획득한 정보를 바탕으로 글을 쓰는 것과 유사
- 생성형 AI가 인터넷, 내부 저장소 등에 있는 법률 문서를 참고하여 법률 분야에 특화된 응답을 하도록 지원

미세조정(Fine-Tuning)

- 범용적으로 활용될 수 있도록 이미 훈련된 AI를 특정 작업/분야에 맞게 추가적으로 조정하여 특화된 AI로 개선하는 기법
- 법학을 전공한 학생이 반도체 회사의 법무팀에 입사 후 특허법 관련 교육을 듣는 것과 유사
- 법률 분야에 대한 응답에 특화된 생성형 AI를 만들기 위해 법률 문서를 생성형 AI에 추가 학습

금융보안원 FINANCIAL SECURITY INSTITUTE

상용 모델과 오픈 모델 비교

	상용 모델	오픈 모델
모델 예시	▪ GPT4 (사이즈 : 약 1.76T으로 추정)	▪ LLaMA2 (사이즈 : 70B 기준)
초기 인프라 투자 비용*	▪ 약 7,500억 원 (엔비디아 A100 80GB 기준 (약 3천만 원) x 25K대) ※ MS는 GPT 기반 서비스 운영에 약 5조 8,000억 원을 투자한 것으로 추산-경향신문, `23.11.)	▪ 최소 약 5억 7천만 원 ~ (엔비디아 A100 80GB 기준 (약 3천만 원) x 19대)
운영 비용 **	▪ 입력 기준 영단어 약 100만 개당 약 1만3천 원($10) ▪ 출력 기준 영단어 약 100만 개당 약 4만 원($30)	▪ 입출력 기준 영단어 약 100만 개당 약 1,300~2,600원($1~2 사이) ***
성능 비교****	▪ 86.4점	▪ 68.9점

* 서비스 운영을 제외한 모델 학습 시에만 필요한 초기 GPU 개수 기준으로 추정

** OpenAI 공식사이트 참고

*** 오픈 모델 학습 및 추론 시 가장 많이 활용하는 엔비디아 A100 80GB 기준으로 추정

**** (메타 LLaMA2 논문) MMLU 점수, LLM의 언어 이해력을 나타내는 점수로서, 과학, 문학 등 다양한 분야에 대한 이해력을 0~100점 사이로 표현

AI 기술의 급격한 발전으로 설비 교체 주기가 짧아지면서 금융회사는
오픈 모델을 이용한 직접 설치 방식을 선택하기보다는
사용량에 따라 이용료를 지급하고 대규모 초기 투자 없이 최신 AI 기술을
빠르게 이용할 수 있는 클라우드를 활용한 소프트웨어 서비스 방식의 AI 모델을 선호

2. 생성형 AI 활용사례 및 서비스

2. 생성형AI 활용사례 및 서비스

금융권 외 생성형 AI를 활용한 사례

상용모델 활용사례

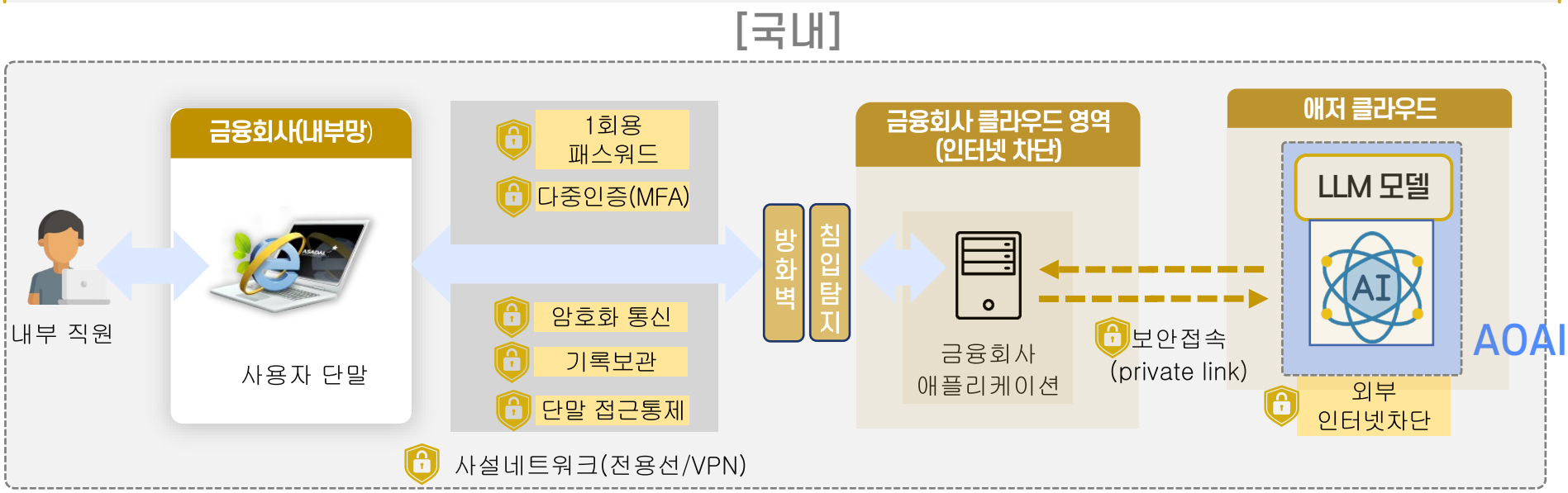
한국남부발전㈜	▪ 한국남부발전은 내부 직원이 활용할 수 있는 사내 정보 정보검색 서비스인 『통합검색 플러스 GPT』를 공개
동아일보	▪ 동아일보는 일반 독자를 대상으로 경제경영 뉴스 전문 챗봇 서비스인 『AskBiz 서비스』를 출시
SK telecom	▪ SK텔레콤은 AI 비서 '에이닷'에 ChatGPT 기반 서비스인 『챗T』를 추가하여 사용자 친화적인 질의응답이 가능하도록 AI 비서를 고도화

오픈모델 활용사례

LG CNS	▪ LG CNS는 LLaMA2 기반 코드 생성 AI 서비스인 『AI 코딩』를 출시하여 다수의 사내 시스템 운영 뿐만 아니라 고객사의 개발 프로젝트에도 활용
POLARIS OFFICE	▪ 문서작업 솔루션 업체 폴라리스 오피스는 오픈모델 중 하나인 솔라(Solar)를 탑재하여 실시간 문서 번역 및 요약 등 AI 기반 문서작업 솔루션을 개발

클라우드 기반의 AI 서비스 구조 예시

MS의 클라우드 애저(Azure)가 제공하는 Azure OpenAI(AOAI)



- 금융회사는 인터넷이 차단된 클라우드 영역에서 애플리케이션 운영
- AOAI 서비스는 ① 국내 전산센터 내 위치하고 외부 인터넷과 차단
  ② 금융회사 클라우드와 보안접속(private link)으로 연결
  ③ 각 금융회사에게 별도로 격리된 자원으로 할당 가능 (PTU 단위로 제공)

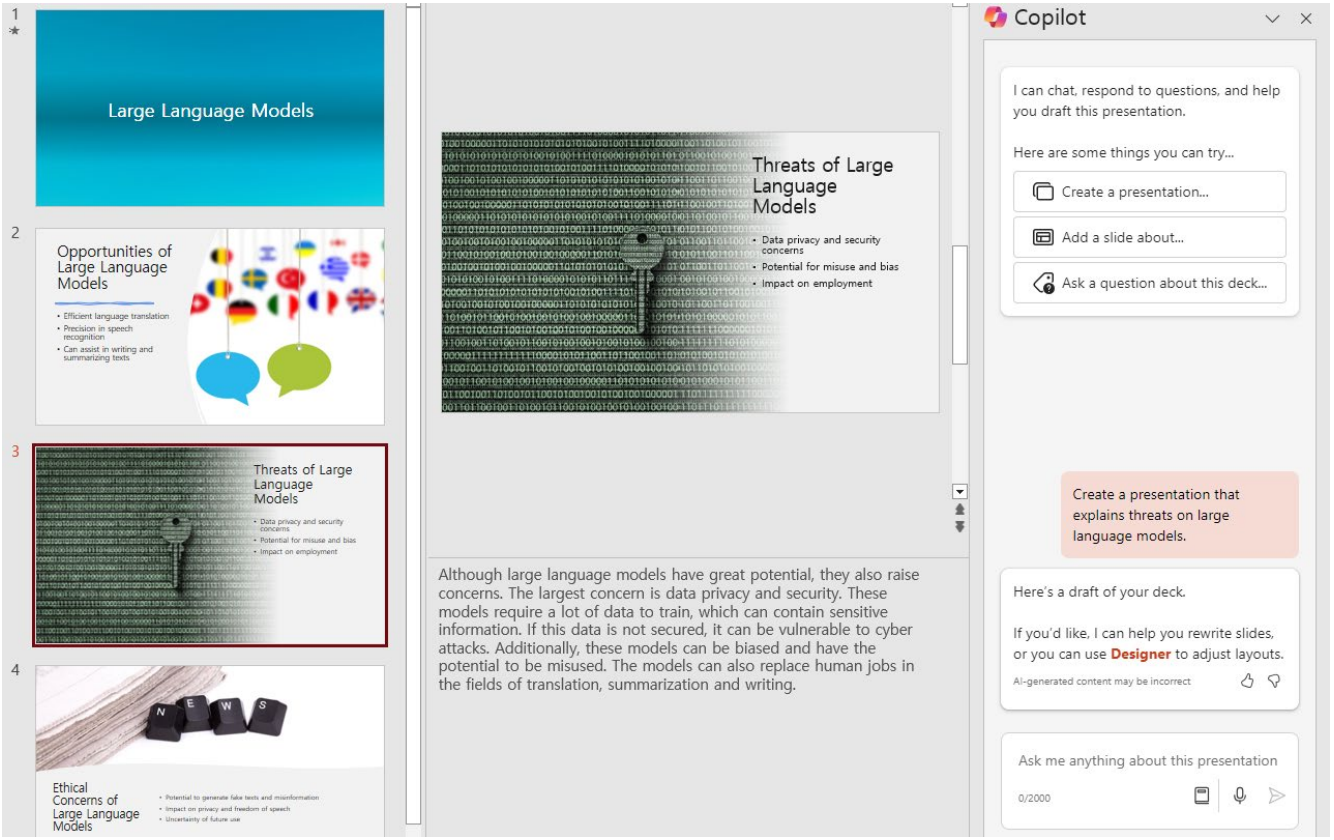
※ 프로비전된 처리량(Provisioned Throughput Units, PTU): 사용자가 사전 구매하고 이용할 수 있는 애저 클라우드 컴퓨팅 자원의 용량 단위

클라우드 기반의 AI 서비스 예시

M365의 Copilot

※ 온라인 구독 기반 오피스 프로그램인 M365에 탑재된 생성형 AI 기반 기능으로, 사용자의 문서 작업에 대한 편리함 제공

- 사용자의 질의에 맞춰 적절한 PPT 장표와 스크립트를 제작



2. 생성형AI 활용사례 및 서비스

금융보안원 FINANCIAL SECURITY INSTITUTE

금융회사가 클라우드 기반의 AI 서비스를 사용해야 하는 이유 및 필요성

	클라우드 기반 상용 모델	직접설치한 오픈 모델
비용 측면	▪ 클라우드 사용량에 따라 비용이 결정되는 방식이므로 초기 인프라 투자 비용이 발생하지 않음 ▪ 입출력 기준 단어처리 비용에 대하여 이용한만큼만 지불하면 됨 ※ 영단어 약 100만 개당 약 $10~30	▪ 오픈 모델 이용 시 대규모 초기 인프라 투자 비용 발생 ▪ 최소 5억 7천만 원 소요 (학습 비용) ▪ 실제 업무를 위해서는 더 많은 서비스 운영 비용이 추가적으로 필요 ※ 영단어 약 100만 개당 약 $1~2
기술 측면	▪ 클라우드서비스제공업자가 지속적으로 서비스를 업데이트함에 따라 최신 기술 도입이 용이	▪ AI 모델을 직접 운영·관리하는 비용이 발생 ▪ 최신 기술을 직접 개발하는 데 막대한 시간 소요 →유연한 신기술 도입 및 대응 어려움

3. 생성형 AI 활용 시 규제 관련 이슈

3. 생성형 AI 활용 시 규제 관련 이슈

AI 활용의 3가지 유형

클라우드 기반

① 해외 서버

✓ 복잡한 설치 과정 없이, 웹 브라우저 등으로 쉽게 이용 가능

✓ 유지보수 및 업데이트가 편리

✓ 외부 시스템과 데이터 송수신 과정에서 프라이버시 노출 위험 상존

✓ 데이터 현지화* 이슈가 존재

② 국내 서버

✓ 데이터 현지화 이슈 없음

③ 직접 설치형 (오픈 모델)

✓ 금융회사가 데이터 처리 과정을 완전히 통제할 수 있음

✓ 외부망과 연결 불필요하므로 데이터 보호 강화

✓ HW/SW 구매, 서버 운영 등 높은 초기 비용이 소요

* 데이터가 발생한 국가 내에서 해당 데이터를 저장 및 처리할 것으로 요구하는 정책으로, EU, 중국 등 많은 국가에서 이를 적용하고 있음
— 「주요국 국경간 데이터 이동 규제 현황 및 시사점 (한국무역협회 국제무역통상연구원, `22.12.14.)」

3. 생성형 AI 활용 시 규제 관련 이슈

3가지 유형별 세부 특징

	클라우드 기반		③ 직접 설치형 (오픈 모델)
	① 해외 서버	② 국내 서버	
신기술 활용 가능여부	✓ CSP 업체의 지속적인 서비스 업데이트 → 신기술 도입 용이		✓ 기존 인프라로 인한 제약으로 신기술 도입 어려움
비용	✓ 사용량에 따라 비용이 결정됨으로 초기 투자 비용이 적음		✓ 상당한 수준의 설비 투자 필요
편의성	✓ 복잡한 설치 과정 없이 활용 가능		✓ 시스템 별도 설치 및 운영절차 마련 필요
규제	✓ 데이터 현지화 이슈	✓ 망분리 규제 준수에 관한 금융당국의 확인	✓ 중요/비중요업무 관계없이 활용가능

3. 생성형 AI 활용 시 규제 관련 이슈

국내 금융산업의 경쟁력 강화를 위해 개인신용정보 등을 취급하는 중요업무에 대해 ❶ 클라우드 방식의 AI 활용에 대한 금융당국의 확인 또는 ❷ 금융회사가 직접 설치하는 방식으로 인공지능 기술을 활용하도록 유도할 것인지 논의할 필요

❶-1 AI 활용이 필수적인 경우에 한해 망분리 규제 준수 등 금융당국의 확인을 받거나,

❶-2 데이터센터의 위치(국내), AI서비스의 보안성 등이 확인된 경우 모든 업무에 대해 금융당국의 확인 및 검토 필요

※ 클라우드 서비스 제공자 등 제3자가 금융회사 내부 데이터에 접근 불가한지 확인 및 데이터 격리 · 차단 방식 등 확인

❷ 국내 금융산업 특화 AI 모델 · 환경 개발 등 금융분야 AI 기술력 및 경쟁력 확보 방안 등 논의 필요

금융미래를 열어가는
금융보안파트너

금융보안원
FINANCIAL SECURITY INSTITUTE

감사합니다.

금융보안원 홈페이지
www.fsec.or.kr