고성능 AI 보안위협 대응
관련 금융권 간담회 안건

고성능 AI 관련
금융권 보안 위협 대응 방안

2026. 5.

금 융 위 원 회
금 융 감 독 원
금 융 보 안 원

목 차

Ⅰ. 현 황

1. 배 경

□ 최근, **고성능 AI 발전***으로 **사이버위협 환경**이 **급격**하게 **변화**

* '26.4월 미국 앤트로픽社의 Mythos(미토스), '26.5월 미국 OpenAI GPT-5.5 등

○ **미토스 선공개버전**(4월 발표)이 보안에서 탁월한 성능*을 보인다고 알려지면서, AI가 **해킹에 악용될 경우**에 대한 **불안감** 확산

* 예: 보안성이 높은 OpenBSD 운영체제에서 27년간 발견되지 않은 취약점, 자동화 도구가 16년간 500만회를 검사할 동안 발견되지 않은 취약점 탐지 등

- 특히, **스스로 탐색·판단·실행**하는 **에이전트**(Agent) AI는 공격측의 **해킹역량**을 **비약적으로 향상**시키는 도구가 될 수 있음

○ 인적 개입이 전제가 된 **기존 보안체계**의 대응 속도와 범위를 월등히 넘어서는 **보안 취약점 탐지** 등이 가속화될 우려

□ **고성능 AI**에 대한 **위험성**이 대두되는 한편, **AI가 보안역량**을 **강화**할 수 있는 **핵심수단**이 될 수 있다고 평가

○ 침해위협 탐지, 취약점 분석·관리 등 **정교한 방어체계 구축**을 위해 **AI**가 **충분한 역할**을 수행 가능

□ 한편, 급변하는 **AI 환경 속**에서도 **금융분야**는 **망분리 규제**로 인해 **AI 활용**이 **제한**되어 **보안역량 강화**에 **한계**

○ 망분리 규제로 외부와의 접점이 **최소화**되는 **측면**은 있으나

○ 이로 인해 **AI를 활용한 방어체계**(취약점 탐지·보안시스템 구축 등)는 **원천적으로 제한**되는 한계점("양날의 검")

2. 대응 현황

□ **(정부)** 금융위는 미토스 이슈 제기 즉시(4月 중순)부터 **업계·전문가와 긴밀한 소통**을 통해 **체계적 대응방안 모색** 중

* 금융위 주관 민간 전문가 회의(4.14일, 4.15일, 부위원장 주재)
「고성능 AI 보안 위협 상황대응반」 운영(4.28일, 5.7일, 디지털국장)

〈※참고〉 보안전문가 및 금융회사 주요의견

· **(보안전문가)** ▲미토스가 금융회사 시스템의 취약점을 효과적으로 공격할 수 있음, ▲취약점 점검, 모의해킹 등에도 AI를 사용하여 보안 수준을 높일 필요

· **(금융회사)** ▲AI기반 사이버공격 위협의 현실화로 인한 파급효과 高, ▲기존 사람 중심의 공격 대비 공격의 속도·범위가 확대될 수 있어 위험할 수 있음

□ **(금융사)** 금융사·인프라기관* 등은 **자체적인 보안관리 강화 조치****를 취하고 있으나, **아래와 같은 애로사항 토로**

* 금융결제원, 한국거래소, 예탁결제원, 신용정보원, 보험개발원 등
** 자체 AI모델 통한 취약점 점검·모의해킹, 공급망 영향분석 등

❶ AI 보안 위협에 대한 신속한 민·관 정보 공유 필요

* 민간과 정부가 합심하여 앤트로픽社 등의 논의 동향을 파악할 필요
** AI 보안 위협의 수준이나 대응요령 등과 관련하여 정확한 정보의 신속 공유 필요

❷ "AI는 AI로 방어"하기 위한 과감한 규제 개선

* 미토스를 필두로 AI가 보안부문 공격·방어 수단에 조만간 활용될 것으로 예상
** 기존 「로드맵」보다 **적극적인 망분리 완화** 조치 필요

❸ 적극적 보안패치 등 독려하는 가이드라인 등 필요

* **보안패치**는 **전산시스템상 오류를 유발**할 수 있어 테스트 등 여러 절차 필요
** 미토스 보고서 공개시기(**7월경**)에 참여기관(글래스윙)의 **보안패치가 다수 배포**될 수 있어, 패치에 따른 보안성 확보 vs 시스템 안정성 등 부담감
↳ **우선순위** 설정, 보안을 위한 패치조치시 **일정범위내 면책조치** 등 필요

Ⅱ. 대응 방향

1 미토스 등 고성능 AI 보안 위협 선제 파악

· 과기부·외교부 등 관계부처와 **AI 보안 위협 적기 파악**

· 금보원·주요 금융사 등 **금융권 차원**에서도 **보안위협 실체 파악**을 위해 노력

2 "AI는 AI로 방어하는" 보안체계 신속히 구축

▸ **[망분리 규제 완화]**

- **보안강화 목적 AI 활용**시 **망분리규제 예외** 신속 추진
- **선별된 금융사**(AI·보안역량) 대상으로 **망분리규제 전면 해제 검토**

▸ **[조직·기능 강화]**

- **민간 기술자문단**, **금융권 AI 보안 협의체** 상시 운영
- 금융보안원 內 「**AI 연구소**」 및 「**AI 지원센터**」 신설 추진

3 금융회사의 AI 보안 위협 체계적 대응 지원

- AI 보안 위협 대응 관련 **가이드라인 제시·자율점검 지도**
- **긴급 보안패치** 시 전산불안 등과 관련하여 **일정 범위 內 면책 검토**
- **소규모·영세 핀테크 기업** AI 보안 전환 **지원**

1 고성능 AI 보안 위협 선제 파악

□ 과기부·외교부 등 관계부처와 **적극 협업***하여 **고성능 AI로 인한 보안위협**, 보안 **리스크 요인** 등 적기 파악

* 정부(과기부·금융위·외교부 등)와 글로벌 AI 기업(앤트로픽(5.11), OpenAI(5.18)) 간 간담회 실시

□ **금보원·주요 금융회사** 등 **금융권 차원**에서도 **다양한 채널**을 통해 AI 보안 위협의 **명확한 실체파악** 등을 위해 노력중

2 "AI는 AI로 방어하는" 보안체계 신속히 구축

1. 망분리 규제 완화

1 보안목적 AI 활용시 망분리 규제 긴급 완화 조치

(비조치의견서 : [1회차] 6~7月 [2회차] 8~9月 [3회차] 4분기중 (잠정))

○ **(대상) 일정한 보안역량** 갖춘 금융회사 중 **신청社**

* ▲총자산 10조원 이상, ▲상시 종업원수 1,000명 이상(**총 49개**)
↳ CISO의 정보기술부문 外 업무 겸직금지 적용 기준(「전자금융거래법」§21의2③)

※ 신청社 중 보안역량·AI 활용 준비 부족 시 **탈락 가능**

○ **(조치) 보안목적***에 한해 정보처리시스템 **망분리 규제 한시**** **완화**

* ① AI 활용한 내부 취약점 확인, ② 보안 SaaS 솔루션을 통한 방어시스템 구축
** 테스트 소요시간·보완조치 등을 감안하여 1년간 유예

○ **(조건) ▲보안성이 검증된 AI 등만 사용, 추가 보안조치 준수**
▲테스트 결과 확인된 사항* **당국보고 + 全 금융권 공유 등**

* (예) ▲고성능 AI 보안 위험성 특징, ▲공격용도로 악용할 경우 예상되는 수법, ▲효과적인 방어를 위한 대응 요령 등

○ **(절차) 금융사 신청*** **→ 평가·선정**** **→ 금융위 보고 → 비조치의견서**

* 금융회사 준비상황·테스트 여력 등을 감안하여 2~3회차로 나누어 신청 접수
↳ **1회차**(6月) : **10개 이내**, 2회차(8月) : 10~20개사, 3회차(4분기) : 미정

** 민간 기술자문단의 보안역량 평가 등을 감안하여 회차별 참여사 선정

※ 미신청 금융회사 등에 대해서는 망분리 규제완화 조치가 필요없는 **"외부 공격표면 대상" AI 취약점 점검**(Blackbox 방식)도 실시
(금융보안원, 7월까지 최대 17개사)

2 선별된 금융회사 대상 망분리 규제 전면 해제 검토

(기획형 혁신금융서비스 : 연내 선정 추진)

○ **고도의 보안역량·AI 활용능력** 등을 갖춘 금융회사에 대해서는 **망분리 규제를 전면 해제**하는 방안 검토·추진

- 민간 기술자문단, 혁신위 등을 통해 **▲보안역량, ▲AI 활용능력, ▲망분리 대체 보안조치** 등을 **꼼꼼히 심사·선별**

- 선별된 금융사는 **▲전면적·체계적 AI 보안체계 구축***, **▲혁신적인 대고객 AI 서비스 마련** 등**에 AI를 폭넓게 활용

* 일회성 취약점 탐지·외부 Saas 솔루션 外, **AI기반 보안체계로 전면 전환** 가능
↳ ▲AI활용 자체 보안프로그램 마련 ▲AI기반 패치 자동화 등 활용 가능

** 금융회사의 생산성 향상을 위한 ▲챗봇 상담·자산관리, ▲여신심사, ▲기업금융, ▲내부통제 등 폭넓은 분야에서 활용 가능

2. 조직 · 기능 강화

1 AI·보안분야 전문가·금융업권 등과 긴밀한 협의채널 구축

(민간 기술자문단 출범 : '26.5月~, 금융권 협의체(상황대응반 등) : 운영 中)

○ **(민간 기술자문단) AI·보안·정보보호** 등 분야에 정통한 **학계·보안업계·법조계** 등 인사로 **자문단**을 **구성**(5月)

* 과기부, KISA, 금융보안원 등에서 추천한 전문가 포함

- **▲망분리 완화** 관련 **자문·평가 ▲**고성능 AI 보안위협 **전망 및 대응 조치, ▲국내외 최신 동향** 등 관련 심도 있는 자문

○ **(금융권 협의체)** 旣 구축한 「**고성능 AI 보안위협 금융권 상황대응반**」 등 **협의채널**을 상시적·수시 개최(5月~)

* [1차 회의] 은행 3개사·증권 2개사, [2차 회의] 인프라 기관 5개사

- AI 보안 위협 관련 **금융권 대응상황·애로사항**을 **청취**하고, 최근 **국내외 논의 동향, 주요 대응요령·정책** 등 신속히 전파

2 「AI 연구소」 및 「AI 지원센터」 신설 추진

(금융보안원 조직 개편 : ~'26.6月)

○ **(AI 연구소)** 고성능AI 보안위협 분석, 대응기법·수단 개발, 침해대응 지원 등을 수행하는 **「금융AI보안연구소」 신설**

- 금보원 內 AI 침해대응 등 제한된 기능을 하던 조직(AI 혁신부)을 **AI 보안과 관련된 전 부문**(기술개발·관제·침해대응·인력양성 등)을 **포괄하여 추진**할 수 있는 조직(AI연구소)으로 **확대 개편**

- **국내·외 AI 보안 연구소** 등과도 적극적으로 **교류**하여 **고성능 AI 보안 위협 최신 동향**을 신속히 **파악·전파**

<※참고 「금융 AI 보안연구소」 신설(안)>

AS-IS (1부)	TO-BE (1연구소 3부 1실)
AI혁신부 (4팀) AI전략팀 AI기술팀 AI RED팀 ASAP팀	금융AI보안연구소 (3부 1실) AI보안총괄부 (4팀 1센터) AI 위협대응 총괄 및 데이터 사업 AI보안총괄팀 AI보안연구팀 데이터사업팀 데이터결합평가팀 금융AI보안 지원센터 AI안전평가부 (3팀) AI 서비스 안전성 및 신뢰성 평가 AI안전평가팀 AI기술팀 AI RED팀 ASAP실 (2팀) AI 기반 보이스피싱 대응 ASAP기획팀 ASAP운영팀 클라우드평가부 (3팀) 클라우드 및 AI플랫폼 평가 클라우드기획팀 클라우드평가1팀 클라우드평가2팀

○ **(지원센터)** 고성능AI 적극 대응이 어려운 **중소 금융사** 등의 **애로사항 접수, 대응요령 안내** 등을 지원하는 **전담조직 신설**

- **▲고성능AI 보안위협·대응요령** 등 상세 안내, **▲AI 취약점 점검, ▲공격 탐지룰** 개발·제공 등 **다각도 지원**

3 금융회사의 AI 보안 위협 체계적 대응 지원

1 AI 보안 가이드라인 배포 및 정보보안 대비태세 점검

(가이드라인 배포 : ~'26.6月, 정보보안 대비태세 점검 : ~'26.6月)

○ **(가이드라인 배포) 고성능 AI 보안위협**에 대비하여 **全 금융회사**가 **준수**해야 할 **세부 대응요령*** 등을 마련·배포(금감원·금보원)

* ▲전산자원 현황 전수점검, ▲패치 우선순위, ▲불필요한 외부접점(노후화된 PC 등) 폐기, ▲계정관리·접근권한 업데이트, ▲공급망 관리 강화 등

※ 고성능 AI 보안테스트 등으로 추가 대응요령 고지가 필요한 경우 수시로 가이드라인 추가 업데이트·배포

○ **(정보보안 대비 태세 점검) 금감원**에서 체크리스트를 배포하여 금융회사의 **적시 대응, 신속한 사고 복구** 등 **대비태세 점검***

* ❶ IT 위험 관리체계, ❷ IT 자산 식별·관리, ❸ 보안 취약점 관리, ❹ 사이버 위협 대응, ❺ 사고 대응(금융소비자 피해보상 포함) 등 5개 분야로 구성

- **미흡 금융회사**에 대해서는 **적시 개선 및 보완**을 지도하고, 그 중 **일부 회사**에 대해 **현장점검**을 **실시**하여 점검의 실효성 제고

2 긴급보안패치 등과 관련한 임·직원 면책

(금융위·금감원 지도공문 : '26.6月~7月)

○ **보안 패치** 등 과정에서 발생한 **경미한 전산시스템 장애**에 대해 **신속한 복구·소비자 보호조치 전제**로 제재조치 감경·면책

3 중·소형 핀테크기업의 AI 보안 전환 지원

○ **핀테크기업**의 **보안점검 비용 부담** 등을 **지원**(핀테크지원센터)

○ 금보원을 통해 ▲핀테크기업의 **AI 위험 관리체계 수립** 지원, ▲**취약점 점검도구 제공** 등을 수행

Ⅲ. 향후 추진계획

1 금융회사에 대한 보안목적 망분리 규제완화 조치 실시

① 망분리 규제 완화 대상 **금융회사 1차 선정***(비조치의견서, 6월 중순)

* ①금융회사 신청접수(5.22~29일) → ②민간전문가 등을 통한 심사·선정(6월 초) → ③ 고성능 AI 기반 취약점 테스트 실시(6~7월)

② **2회차**(8~9월), **3회차**(4분기 이후) **테스트** 순차 시행

※ 고도의 보안역량·AI 활용능력 등을 갖춘 금융회사에 대해서는 망분리 규제를 전면 해제하는 방안 검토·추진

2 금융권 AI 보안역량 강화를 위한 지원방안 마련

① 금보원 내 **「AI연구소」 및 「AI지원센터」 신설 추진**(6월중)

② 고성능 **AI 보안위협 대비 가이드라인 마련·배포***(6월중)

* 보안패치 관련 임직원 면책 관련 지도공문 포함

③ 금융권 **정보보안** 체크리스트 배포 및 **대비태세 점검**(6월~)

구분	주요내용	추진일정
1 망분리 규제완화	▫ **보안목적 AI 활용시 망분리 규제 긴급 완화 조치**	
	① 금융회사 신청접수	5.22~29일
	② 민간전문가 등을 통한 심사·선정	6월 초
	③ 금융위원회 보고	6.17
	④ 선정대상 금융회사에 대한 비조치의견서 발부	6월 중
	⑤ 고성능 AI 기반 취약점 테스트 실시	6~7월
	⑥ 2차 선정 절차 수행	8~9월
	⑦ 3차 선정 절차 수행	'26.말
2 지원방안 등	▫ **「AI연구소」 및 「AI지원센터」 신설**	6월
	▫ **AI 보안 가이드라인 마련·배포**	6월
	▫ **체크리스트 배포 및 점검**	6월~