

ZOOM  
-IN해외 금융당국의 원칙 기반 AI 규율체계  
현황과 시사점

- 영국 FCA는 금융서비스 분야의 AI 도입 활성화를 위해 특정 기술을 금지하기보다는 위험을 식별하고 완화하는 데 중점을 두는 규제 접근방식을 채택
- 그리하여 AI 활용에 대한 5가지 기본원칙 준수와 효과적인 책임 및 거버넌스 구현에 초점을 두고 이를 위한 기존 고위경영자 인증제도(SM&CR)에 따른 고위경영자의 책무를 강조
- 싱가포르 MAS는 AI 위험관리 프레임워크 구축을 위한 다양한 이니셔티브 및 기본원칙(FEAT) 제시 등을 통해 금융기관의 AI 도입을 촉진
- 대만 FSC는 금융기관의 AI 기술 활용을 지원함과 동시에 위험을 효과적으로 관리할 수 있도록 최근 행정지침 성격의 가이드라인을 출시
- 국내 금융당국은 금융권의 AI 활용이 가속화됨에 따라 관련 규율체계 개선을 논의 중으로, 향후 광범위하게 이용될 수 있는 AI 활용을 위해 위험관리 체계 마련이 선행될 필요

- 영국 금융감독청(Financial Services Authority: FCA)은 금융서비스 부문의 AI에 대한 강력한 규제를 도입하기보다는 큰 틀에서 AI가 금융시장 내 안전하고 책임성 있게 개발 및 사용되고, 이로인 혁신(beneficial innovation)을 주도하는 방식으로 활용할 계획을 발표
  - 영국 정부는 AI 활용에 대한 지침을 제공하기 위해 2023년 3월 5가지 기본원칙에 기반한 AI 규제 통합 프레임워크를 제한하였으며, 이는 AI의 혜택과 잠재적 위험 간 균형을 고려한 혁신 친화적 접근방식을 취함
    - 재무부(HM Treasury)는 해당 원칙에 따라 각 규제기관에게 AI에 대한 전략적 접근방식을 설명하는 업데이트를 발표할 것을 요청하였으며, FCA는 이를 위해 금융서비스 부문에 대한 검토를 수행
    - FCA는 AI 활용을 위한 5가지 기본원칙에 비추어, 금융서비스 부문 내 AI의 고유 위험이 크지 않으므로 새로운 규제를 도입하기보다는 기존의 다양한 법률 체계로 규율이 가능하다고 판단
  - 그리하여 FCA는 영국 정부의 AI 혁신 친화적 접근방식<sup>1)</sup>을 지지하고, 금융서비스 부문의 AI에 대한 규제 접근방식을 특정 기술을 금지하기보다는 위험을 식별하고 완화하는 데 중점을 두고 AI가 이로인 혁신을 촉진하도록 유도한다는 입장

1) Gov.UK, 2024. 2. 6, A pro-innovation approach to AI regulation: Government response, Consultation Outcome.

- 기업이 AI 기술을 안전하고 책임감 있게 도입할 수 있는 방법과 AI 혁신이 금융소비자와 시장에 미치는 영향을 이해하는 데 초점
  - 특정 기술을 의무화하거나 금지하지 않는 기술 중립적이고 원칙 중심적이며 성과 중심적인 규제기관으로서의 일반적인 역할을 강조
- 이는 최근 EU가 도입한 AI 시스템의 위험수준에 따른 금지사항과 요구사항을 제시한 세계 최초의 법적 구속력을 갖는 AI 규제법(AI Act)과는 상반되는 정부 후원의 자발적 자율규제(voluntary self regulation) 형태

□ FCA의 이러한 접근방식은 금융서비스 부문에서 AI 활용에 대한 5가지 기본원칙을 준수하기 위해 기존 규정 및 지침과의 관련성을 설명한 문서<sup>2)</sup>에 기반

- FCA는 AI가 초래할 수 있는 금융시장의 잠재적 위험을 경계하고 있는 만큼 위험관리 조치로서 5가지 기본원칙 준수를 위해 기존 규정 및 지침과 어떻게 매핑되는지를 설명
- 해당 설명 문서는 개별 규제기관이 AI의 영향과 위험을 고려하고 AI 거버넌스에 대한 접근방식을 결정할 때 지침을 제공하는 역할
- FCA는 금융시장 전반의 AI 도입을 면밀히 모니터링하고, 필요한 경우 향후 규제 마련을 고려할 수 있음을 시사

**<표 1> AI 기본원칙에 기반한 FCA 프레임워크 주요 내용**

기본원칙	주요 예시
안전, 보안, 견고성	사업원칙: 기업은 적절한 기술, 주의, 근면성을 가지고 사업을 수행 적절한 위험관리 시스템을 통해 책임감 있고 효과적으로 업무를 조직하고 통제하기 위한 합리적인 주의를 기울여야 함
적절한 투명성과 설명 가능성	소비자 의무: 정직하고, 소매 고객의 공정하고 개방적인 거래를 보장 사업원칙: 고객의 정보 요구를 적절히 고려하고 명확하고 공정하며 오해의 소지가 없는 방식으로 고객과 의사소통을 수행
공정성	소비자 의무: 소매 고객에게 좋은 결과를 제공하고 선의로 행동하며 예측가능한 피해를 방지 사업원칙: 고객의 이익을 존중하고 공정하게 대우하며, 명확하고 공정하며 오해의 소지가 없는 방식으로 정보를 전달
책임 및 거버넌스	고위경영자 대응: 최고 위험책임자 기능(Chief Risk function)이 위험노출 설정 및 관리를 포함하여 회사의 위험통제에 대한 전반적인 관리 책임을 갖도록 보장 소비자 의무: 소매 고객에게 좋은 결과를 제공할 의무를 전략, 거버넌스 및 리더십에 반영하도록 보장
이의신청 및 구제	분쟁해결(DISP) : 기업의 분쟁처리 방법을 자세히 설명하는 규칙과 지침, 보상 계획

자료: FCA

2) FCA, 2024, AI Update.

- FCA는 AI에 대한 규범적 규칙 적용보다는 금융기관의 효과적인 책임 및 거버넌스의 구현에 초점을 맞추고 있으며, 이를 위해 기존 고위경영자 인증제도(Senior Managers and Certification Regime: SM&CR)에 따른 고위경영자의 책임과 의무가 핵심적 역할을 한다고 강조
  - AI 원칙은 AI 수명주기 전반에 걸쳐 명확한 책임라인을 확립하여, AI 시스템의 공급 및 사용에 대한 효과적인 감독을 보장하도록 강력한 거버넌스 조치가 있어야 함을 요구
  - SM&CR 제도에 기반하여 고위경영자는 해당 영역/기능에서 AI를 사용할 책임이 있음을 분명히 하고 기업의 위험통제 전반을 관리할 책임이 있으며, AI의 안전하고 책임감 있는 사용을 위해 합리적인 조치를 취해야 할 의무를 가진다고 명시
    - AI 배포가 회사의 위험 감수성, 규제 요구사항 및 공정한 고객 대우 원칙과 일치하는지 확인할 책임
  - 이에 따라 기업은 AI 기반 시스템의 감독 및 거버넌스를 담당하는 고위경영자를 식별해야 하고, AI 사용을 위한 인증제도(certification regime)를 확대하는 것을 하나의 해결책으로 제안
    - AI 시스템의 기술적 복잡성을 감안할 때, 이를 개발하거나 배포하는 직원이 역량을 갖추는 것이 중요하므로, AI 시스템을 관리하는 직원에게 인증제도를 확장 적용하는 것을 고려
  
- 싱가포르 통화청(Monetary Authority of Singapore: MAS)은 AI 활용에 대한 공적 가이드라인으로서 공정성, 윤리, 책임 및 투명성(Fairness, Ethics, Accountability and Transparency: FEAT)을 촉진하는 14가지 원칙을 제시<sup>3)</sup>하였으며, 해당 원칙을 기반으로 금융부문의 책임 있는 AI 사용을 지원
  - 이는 금융부문에서 AI 기술을 적용할 때 고려해야 할 기본원칙으로, 해당 원칙을 기반으로 금융기관 내부 거버넌스를 구축, 강화함으로써 AI 활용에 대한 일반 대중의 신뢰를 높이도록 지원
    - 데이터 분석 및 AI(AI and Data Analysis: AIDA) 사용에 있어 공정성, 윤리, 책임성, 투명성 4가지 측면에서 준수해야 하는 14가지 원칙과 구체적 사례를 제시하고 있어 금융기관이 AI 서비스를 설계 및 도입할 때 체크리스트로서 기능
  - 최근 발표한 MAS의 MindForge 프로젝트의 위험관리 프레임워크에서는 생성형 AI(generative AI) 활용 시 금융기관이 직면한 특정 문제를 해결하기 위한 FEAT 원칙의 확대 적용을 제안<sup>4)</sup>
    - 금융기관은 AI 및 생성형 AI 시스템의 수명주기 전반에 걸쳐 위험을 식별하고 완화하는 것이 중요
    - 생성형 AI 활용 증가는 제3자 지적재산권을 침해하는 콘텐츠의 생성 가능성 등 관련 위험을 모니터링 하기 위한 새로운 거버넌스의 수립이 요구되므로, 생성형 AI 활용 시 금융기관이 직면한 특정 문제를 해결하기 위해 FEAT 원칙의 확장을 고려

---

3) MAS, 2018, *Principles to Promote Fairness, Ethics, Accountability and Transparency(FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector*, Monographs or Information Paper.

4) MAS, 2024. 5. 27, Project MindForge, Schemes and Initiatives.

**〈표 2〉 싱가포르 금융부문의 AI 사용에 대한 FEAT 원칙 주요 내용**

원칙	주요 내용
공정성 (Fairness)	<ul style="list-style-type: none"> <li>· 정당성: AIDA 주도 의사결정이 정당한 이유 없이 개인/집단에게 불이익을 주지 않아야 함</li> <li>· AIDA 주도 의사결정을 위한 입력 요소로 개인 속성(예: 연령, 소득수준 등)을 정당하게 사용</li> <li>· 정확성과 편향: AIDA 주도 의사결정에 사용되는 데이터, 모델들은 정기적으로 검토 및 검증되어야 하며, 의도되지 않은 편향을 최소화해야 함</li> <li>· AIDA 기반 의사결정은 모델이 설계 및 의도한 대로 작동하도록 정기적으로 검토해야 함</li> </ul>
윤리 (Ethics)	<ul style="list-style-type: none"> <li>· AIDA의 사용은 회사의 윤리적 기준, 가치 및 행동강령과 일치해야 함</li> <li>· AIDA의 결정은 최소한 인간중심의 결정과 동일한 윤리적 기준을 준수해야 함</li> </ul>
책임 (Accountability)	<ul style="list-style-type: none"> <li>· 내부적 책임: AIDA 주도 의사결정에 AIDA를 사용하는 것은 내부기관의 승인이 필요</li> <li>· AIDA 사용 기업은 내부개발 및 외부소싱 모두에 대해 책임을 가지며, AIDA 사용에 대한 경영진과 이사회에 인식을 적극적으로 제고</li> <li>· 외부적 책임: AIDA 사용 기업은 데이터 주체(소비자 및 이해관계자)에게 AIDA 주도 의사결정에 대한 질문, 이의신청 등의 채널을 제공해야 함</li> <li>· AIDA 주도 의사결정을 검토할 때 데이터 주체가 제공한 관련 보안적 데이터가 고려되어야 함</li> </ul>
투명성 (Transparency)	<ul style="list-style-type: none"> <li>· AIDA 사용은 일반적 의사소통의 일부로서 데이터 주체에게 적극 공개되어야 함</li> <li>· 데이터 주체의 요청 시, AIDA 주도 결정을 내리는 데 사용된 데이터, 해당 데이터가 미치는 영향에 대해 명확한 설명을 제공해야 함</li> <li>· 또한 AIDA 주도 결정이 미칠 수 있는 결과에 대해 명확한 설명을 제공해야 함</li> </ul>

자료: MAS

□ 싱가포르 금융당국은 다양한 금융분야에서 활용할 AI 기술 개발에 적극적으로 나서고 있으며, 위험 관리를 위한 다양한 이니셔티브를 통해 책임 있는 AI의 사용을 장려함으로써 AI 생태계 육성을 주도

— MAS는 2023년 8월 금융부문 기술 및 혁신 지원제도(Financial Sector Technology and Innovation Scheme: FSTI) 3.0에 따라 금융부문의 AI 기술 개발 및 도입을 위해 3년간 1억 5천만 싱가포르달러를 투자할 것이라고 발표

— MAS는 금융기관의 데이터 분석 및 AI의 책임 있는 채택을 촉진하기 위해 Veritas 이니셔티브를 운영하고 FEAT 원칙에 대한 평가 방법론을 제공<sup>5)</sup>

- Veritas는 MAS와 금융기관이 협력하여 AI의 책임 있는 사용을 위한 프레임워크를 구축하는 이니셔티브로, 1단계에서 은행 사용 사례에 대한 공정성 측정 기준과 평가 방법론을 제시
- 2022년 2월 발표된 Veritas 2단계는 금융기관을 대상으로 FEAT 원칙의 평가 방법론과 평가 사례를 연구한 백서를 개발하고, 금융기관의 책임 있는 AI 사용을 촉진하기 위해 Veritas Toolkit Version 1을 출시하여 공정성 원칙 준수를 위한 평가 방법론에 초점
- 2023년 6월 Veritas 3단계를 통해 윤리, 책임성, 투명성 원칙에 대한 평가 방법론으로 확장한 Veritas Toolkit Version 2를 출시하였으며, 금융기관의 기존 거버넌스 프레임워크에 평가 방법론을 통합하고 원칙 평가 사례를 연구한 백서를 공개

5) MAS, 2023. 10. 26, Veritas initiative, Schemes and Initiatives.

- 또한 금융부문의 생성형 AI 활용을 위한 이니셔티브인 Project MindForge를 운영 중으로, 생성형 AI의 책임 있는 사용을 위한 명확한 프레임워크를 개발하고 위험관리를 강화하는 것이 목표
  - MindForge의 1단계 프로젝트로 2023년 11월 은행 대상 생성형 AI 위험관리 프레임워크를 개발한 가운데 여기에는 기존 FEAT 원칙과 Veritas 방법론을 평가하고, 위험관리를 위한 추가 원칙과 권장사항을 제안
  - 다음 단계에서는 보험 및 자산관리 업계의 금융기관을 포함하도록 범위를 확장할 계획이며, AI 거버넌스 구현에 중점을 두고 금융산업을 위한 AI 거버넌스 핸드북 개발을 목표

□ 대만 금융감독위원회(Financial Supervisory Commission: FSC)는 2024년 6월 금융기관의 AI 기술 활용을 지원함과 동시에 위험을 효과적으로 관리할 수 있도록 행정지침 성격의 ‘금융산업을 위한 AI 활용 가이드라인’<sup>6)</sup>을 발표

- FSC는 금융산업에서 책임 있는 AI 혁신을 촉진하기 위해 2023년 10월 ‘금융산업의 AI 적용을 위한 핵심원칙 및 추진정책’<sup>7)</sup>을 발표하고, 6가지 핵심원칙과 8가지 지원정책을 개괄적으로 설명
- 이러한 핵심원칙을 바탕으로 금융기관의 AI 기술 활용을 장려하고, 책임 있는 혁신에 중점을 두며 신뢰도 높은 AI를 도입하기 위해 올해 6월 금융산업을 위한 가이드라인을 발표
- 가이드라인은 크게 AI 시스템의 위험 평가요소와 위험 수준, 제3자 감독 등 일반적인 문제를 다루는 일반원칙과, AI 수명주기에 걸쳐 평가된 위험을 바탕으로 6가지 핵심원칙을 이행할 때의 고려사항과 예방조치 등을 제시한 6개의 장으로 구성
  - 가이드라인에는 생성형 AI에 대한 위험관리 및 제어 방법을 별도로 제시하고 있어 최근 확산되고 있는 AI 기술 영역을 함께 포괄
- FSC에 따르면, 해당 가이드라인은 금융기관의 AI 도입과 활용, 관리 시 주의해야 할 사항을 주로 제시하는 구속력 없는 행정지침 성격을 갖고 있으며, 해당 가이드라인을 기반으로 금융협회의 자율규제 규범 또는 금융기관별 내부 거버넌스에 핵심원칙과 조치를 마련하도록 권장
  - 금융기관은 AI 사용에 따른 위험을 완화하는 메커니즘과 핵심원칙 구현 방법을 자체적으로 수립할 수 있으며, 보다 비용 효율적인 방법을 채택하도록 강조

---

6) FSC, 2024. 6. 20, 金管會發布「金融業運用人工智慧(AI)指引」.

7) FSC, 2023. 10. 17, The FSC publishes core principles and policies for AI applications in the financial industry, Press Release.

**<표 3> 대만 금융산업을 위한 AI 활용 가이드라인 주요 내용**

핵심원칙	주요 내용
1. 거버넌스 및 책임 메커니즘 수립	<ul style="list-style-type: none"> <li>· AI 시스템 관리, 목적, 적용업무 또는 운영에 대한 명확한 구조와 위험관리 정책 수립</li> <li>· 오류나 예상치 못한 사건을 처리하기 위한 포괄적인 절차 구축</li> <li>· 고위험 AI 시스템의 경우 위험, 전문성 등을 평가한 후 독립적인 제3자 AI 전문가의 검토 및 평가를 위한 메커니즘을 구축</li> </ul>
2. 공정성과 인간중심적 가치	<ul style="list-style-type: none"> <li>· AI 시스템 사용 시 공정성을 평가하고, 편견을 지양하며 차별을 방지</li> </ul>
3. 개인정보보호 및 고객권리보호	<ul style="list-style-type: none"> <li>· 고객 개인정보보호, 데이터 유출 위험 방지 및 민감한 정보를 수집하지 않도록 데이터 최소화 원칙 준수</li> <li>· 고객의 AI 서비스 이용 여부 선택권 존중</li> <li>· 고객과 기관 모두의 위험과 대안 솔루션의 타당성, 비용 등을 고려한 후 대안 솔루션 제공 여부 결정</li> </ul>
4. 시스템 견고성과 보안 보장	<ul style="list-style-type: none"> <li>· 시스템 견고성과 보안 보장, 사이버 보안 조치를 수립 및 구현</li> <li>· 제3자 개발 AI 시스템 활용 시 해당 공급업체에 대한 위험관리 및 감독 필요</li> </ul>
5. 투명성과 설명 가능성 구현	<ul style="list-style-type: none"> <li>· AI 시스템 운영의 투명성 보장</li> <li>· 자체 개발 또는 아웃소싱 여부와 관계없이 AI 시스템이 어떻게 작동하는지 예측이나 결정 논리의 명확한 설명 필요</li> </ul>
6. 지속가능개발 촉진	<ul style="list-style-type: none"> <li>· AI 시스템을 사용할 때 지속가능개발 원칙에 맞춰 개발 전략 및 구현을 조정</li> <li>· 또한 디지털 전환 과정에서 직원들에게 적절한 교육과 훈련을 제공</li> </ul>

자료: FSC

□ 국내 금융당국은 최근 생성형 AI 기술에 따른 금융권의 AI 활용이 가속화됨에 따라 망분리 규제 완화 등 관련 규율체계 개선을 논의 중으로, 향후 광범위하게 이용될 수 있는 AI 활용을 위해 위험관리 체계 마련이 선행되어야 할 것으로 판단

- 금융당국은 금융분야에서 AI 활용이 안전하게 확대되고 정착할 수 있도록 2021년 「금융분야 AI 가이드라인」을 마련하였으며, 2022년 8월에는 ‘금융분야 AI 활용 활성화 및 신뢰확보 방안’을 통해 로드맵을 제시하고 이에 따른 다양한 후속조치를 수행
- 다만 현재 망분리 및 데이터 결합 규제로 인해 금융권 내 생성형 AI 활용에 제약이 있어 금융권의 AI 활용도 제고를 위해서는 관련 규제를 완화시켜야 한다는 의견이 지속됨에 따라, 금융당국은 망분리 적용 업무, 시스템 등을 세분화하여 전자금융거래와 무관한 영역에서는 규제를 완화, 해제하는 방안을 고려하는 등 금융권 망분리 제도 개편 계획을 발표<sup>8)</sup>
- 향후 AI 기술이 금융서비스에서 광범위하게 사용될 가능성이 높아지고 있는 만큼 AI 신뢰도를 높이고, 위험을 체계적으로 관리할 수 있는 장치 마련이 뒷받침될 필요

8) 금융위원회, 2024. 7. 8, 김소영 금융위 부위원장, 금융연 주최 「미래금융세미나」 기조연설, 보도자료.

- 앞서 논의된 해외 국가들은 AI 도입을 촉진하기 위해 AI 윤리 원칙 및 가이드라인, 위험관리 프레임워크 등을 수립하고 금융기관이 자발적으로 채택할 수 있는 접근방식을 취하고 있는 것을 참고하여 국내 금융당국은 금융기관의 책임 있는 AI 도입을 위한 제도적 지원을 지속할 필요
- 영국 SM&CR 제도에 따른 AI 사용에 대한 고위경영자의 책임과 역할 강조 사례를 참고하여, 국내에서도 AI 활용의 효과적 거버넌스 구축과 자율적 통제를 위해 고위경영자의 내부통제 수립 책임을 강화하고, 적절한 내부통제를 수립한 경우 책임을 경감하는 인센티브가 부여되도록 책무구조도상의 면책 범위 확대를 고려<sup>9)</sup>
- 기존 금융분야 AI 가이드라인의 추상적인 내용, 모호함 등으로 실제 적용에 어려움을 겪고 있다는 의견<sup>10)</sup>을 고려하여, 보다 구체적이고 모범사례를 포함하여 실무적으로 유용하도록 가이드라인을 개선하는 것을 검토해 볼 수 있음

선임연구원 안유미

---

9) 이효섭, 2024, 생성형 AI 확산에 따른 금융산업 지형변화 및 대응 과제, 자본시장연구원 『자본시장포커스』 2024-13호.

10) 최근 대한상공회의소가 실시한 ‘금융권 AI 활용현황과 정책개선 과제 조사’(2024. 6. 20) 결과에 따르면 2021년 금융위원회가 발표한 ‘금융분야 AI 가이드라인’에 대해 ‘내용이 추상적이어서 활용하기 어렵다’(53.4%)거나 ‘잘 모르겠다’(37.8%)라는 평가가 대다수로 나타남