



발행일 2021년 10월 12일
발행처 국회입법조사처
발행인 김만흠 국회입법조사처장
www.nars.go.kr

이슈와 논점

금융분야 AI 가이드라인 도입 추진과 시사점

이수환*, 박소영**

다양한 금융 업무에서 인공지능(AI)이 사용됨에 따라, 금융위원회는 최근 “금융분야 AI 가이드라인”을 발표하였다. 동 가이드라인은 모범규준 형태로, 소비자 보호 측면에서 AI 운영에 대한 최소한의 방향성을 제시한 의의가 있다. 그러나 금융회사 등의 자율규제를 유도하기 위해서는 개인정보보호위원회의 인공지능 자율점검표, EU집행위원회의 인공지능 평가 목록 등을 참고하여 금융시장에 특화된 AI 운영 원칙과 점검항목을 보다 구체적으로 마련하는 논의가 지속적으로 필요하다.

1 들어가며

인공지능(Artificial Intelligence, 이하 ‘AI’)¹⁾이 널리 사용되기 시작하면서 정부와 민간 모두 책임있고 공정한 인공지능 개발 및 활용을 위한 윤리 기준을 마련하기 시작했다. 매년 AI 지수를 발표하는 조사²⁾에 따르면 전 세계적으로 2018년 45개, 2019년 28개, 2020년 23개의 AI 윤리 기준이 마련되었다. 우리나라는 최근 과학기술정보통신부, 개인정보보호위원회, 방송통신위원회 등 여러 부처가 AI 윤리 기준을 구체화하고 있다([표 1] 참고).

시장 분석, 상품 추천, 신용 평가, 부당거래 탐지, 준법 감시 등 다양한 금융 업무의 고도화 과정에 AI가 사용됨에 따라, 금융위원회도 2021년 7월 8일

AI 기반 금융 서비스 개발을 위한 “금융분야 AI 가이드라인”을 발표하였다. 금융은 국민의 경제활동에 직접적인 영향을 미치므로 면밀하게 AI 기술을 접목할 필요가 있다. 이하에서는 금융분야 AI 가이드라인을 살펴보고, 검토가 필요한 사항을 논해보고자 한다.

[표 1] 국내 AI 윤리 정책 및 가이드라인 주요 사례

| 부처 | 정책·가이드라인 | 발표 일자 |
|----------------------|--------------------------------------|--------------|
| 과학기술정보통신부, 정보통신정책연구원 | “인공지능 윤리기준” | '20. 12. 23. |
| 과학기술정보통신부 | “신뢰할 수 있는 인공지능 실현전략” | '21. 5. 13. |
| 개인정보보호위원회 | “인공지능 개인정보보호 자율점검표” | '21. 5. 31. |
| 방송통신위원회 | “인공지능 기반 미디어(매체) 추천 서비스 이용자 보호 기본원칙” | '21. 6. 30. |
| 금융위원회 | “금융분야 인공지능 가이드라인” | '21. 7. 8. |

※ 자료: 각 부처 보도자료

1) 데이터를 해석하여 지식을 추론하거나 정보를 처리하고 특정 목표를 달성하기 위한 결정을 하는 능력이 구현된, 인간이 설계한 소프트웨어 또는 하드웨어 시스템을 말하며, 컴퓨터 처리능력 향상, 우수 알고리즘 등장, 네트워크 발전에 따른 데이터 축적 등으로 급격히 성장하고 있음.

2) Stanford University Human-centered Artificial Intelligence(HAI), 「Artificial Intelligence Index Report 2021」, 2021. 3.

2 금융분야 AI 가이드라인 개요³⁾

(1) 가이드라인 마련 및 향후 계획

금융위원회는 2021년 7월 8일 제1차 ‘디지털 금융협의회 데이터 분과회의’에서 금융분야 AI 가이드라인을 발표하였다.⁴⁾ 금융위원회는 충분한 적용 준비기간을 거쳐 동 가이드라인을 연내에 시행하는 것을 목표로 하고 있다. 이를 위해 올해 3분기 내에 금융업권별·서비스별 실무지침을 마련할 계획이다.⁵⁾

[표 2] 금융분야 AI 가이드라인·지침 마련, 시행 일정

| '21년 7월 | '21년 3분기 | '21년 연내 |
|------------------|---------------------|-------------|
| 가이드라인 최종안 마련, 발표 | 가이드라인 적용 준비 → 전면 시행 | |
| | 실무지침 마련 | 실무지침 발표, 전파 |

※ 자료: 금융위원회(2021)

(2) 가이드라인 규율 형식

금융분야 AI 가이드라인은 AI 활성화를 위하여 법적 규제나 행정지도에 비하여 약한 규제 형식을 취한 것으로 ‘모범규준(Best Practice)’에 해당한다. 금융위원회는 동 가이드라인을 통하여 AI 운영에 대한 최소한의 방향성을 제시하고자 하였다.

(3) 가이드라인 주요 내용

금융분야 AI 가이드라인은 금융분야 AI 시스템의 개발, 사업화, 활용 등 전 과정에서 “신뢰성”을 제고하여 AI 활성화에 기여하는 것을 목적으로 한다. 동 가이드라인은 금융서비스 및 금융상품 제공을 위한 업무에 AI 시스템을 직·간접적으로 활용하거나 활용하고자 하는 금융회사, 그리고 금융연관

서비스 제공을 위한 업무에 AI 시스템을 직·간접적으로 활용하거나 활용하고자 하는 비금융회사(이하, 금융회사 등)에 적용된다.

금융회사 등은 선제적 위험관리체계를 구축하여야 한다. 조직 내 AI 윤리원칙과 기준을 수립하고, 위험관리정책을 마련하는 한편, AI 시스템의 위험평가·관리를 위한 구성원의 역할·책임·권한을 구체적으로 정의하여야 한다.

금융회사 등은 AI 시스템의 목적, 고객의 특성 등을 고려하여 탄력적으로 AI 시스템을 활용할 수 있으나, AI가 인간의 의사결정 과정을 대체하는 경우에는 AI 시스템을 효과적으로 감독·통제하고 “책임성”을 유지할 수 있도록 AI 시스템을 설계하여야 한다.

개발 단계에서는 AI 학습 데이터의 품질을 조사하고 개선하는 노력을 지속하여야 하고, 개인신용정보 오·남용을 방지하고 안전한 정보 활용을 위한 충분한 조치를 거쳐야 한다. 설명의무가 있는 금융서비스에 AI 시스템을 활용하는 경우에는 “설명 가능한 인공지능 기술” 도입을 위해 노력을 기울여야 한다.

AI 시스템의 평가·검증 단계에서는 서비스 특성에 맞게 “성과와 공정성의 목표 수준과 판단 지표”를 선정하여야 한다. 서비스에 따라 우선시되어야 하는 성과와 공정성 기준이 달라질 수 있다는 점을 유념하여⁶⁾ 위험요인을 통제하여야 한다.

도입·운영·모니터링 단계에서는 “적절한 권리구제 방안”을 고지하고 성능을 주기적으로 모니터링 하며 “최선의 보안 시스템을 구축”하기 위하여 노력해야 한다.

3 자율규제를 위한 점검표 도입

(1) 금융분야 AI 자율점검표 도입 필요성

6) 예를 들어 소비자에게 금융거래 기회를 제공하는 서비스는 적격자의 금융거래가 거절될 수 있는 위험(False Negative 오류)을 최소화하여야 하는 반면, 불법 거래를 탐지하는 서비스는 불법 거래를 탐지하지 못하는 위험(False Positive 오류)을 최소화하여야 함.

3) 금융위원회, 「금융분야 AI 가이드라인 및 주요 검토 필요사항」, 2021. 7. 8.
 4) 금융위원회, 「금융분야 AI 가이드라인 및 주요 검토 필요사항」, 2021. 7. 8.; 금융위원회·금융감독원 보도자료, 「금융분야 인공지능(AI) 가이드라인이 시행됩니다. - 금융권 AI 활용을 활성화하고 AI 기반 금융서비스에 대한 신뢰를 제고하기 위한 모범규준 마련·발표 -」, 2021. 7. 8.
 5) 2021년 10월 5일 기준으로, 발표된 실무지침은 없음.

금융분야 AI 가이드라인은 AI를 운영함에 있어서 소비자 보호를 위한 방향성을 제시한다는 측면에서 의의가 있다. 그러나 동 가이드라인 상의 “신뢰성”, “책임성”, “공정성” 등 용어는 다소 추상적이고, “최선의”, “적절한” 방안이 무엇인지 불분명하여 동 가이드라인의 적용 대상인 금융회사 등에 부담으로 작용할 우려가 있다.

금융위원회는 ‘모범규준’과 ‘업권별 자율규제’ 형식의 2단계 규율을 계획하고 있지만,⁷⁾ 금융회사 등의 자율규제를 유도하기 위해서는 보다 명확한 기준과 원칙이 제시될 필요가 있다. 이를 기초로 금융회사 등이 자체적으로 의무·권장사항을 점검해 볼 수 있도록 국내외 사례를 참고하여 자율점검표를 마련하는 방안을 검토해 볼 수 있다.

(2) 국내외 AI 관련 자율점검표 도입 사례

가. 국내 사례

개인정보보호위원회는 2021년 5월 31일 “인공지능 자율점검표”를 발표하였다([표 1]).⁸⁾ 이는 개인정보 처리 특성(복잡성·불투명성, 자동화·불확실성, 대규모 데이터 처리 등)을 고려하여 「개인정보 보호법」, 과학기술정보통신부의 “인공지능 윤리 기준”, 국제적으로 통용되는 개인정보보호 중심설계 원칙 등을 반영한 결과이다.

동 점검표는 AI 관련 개인정보 보호 6대 원칙(적법성, 안정성, 투명성, 참여성, 책임성, 공정성)을 기반으로 점검해야 할 항목을 기술하여 보다 명확한 개인정보 보호 기준을 제시하고 있다.⁹⁾

7) 금융위원회, 「금융분야 AI 가이드라인 및 주요 검토 필요사항」, 2021. 7. 8.

8) 개인정보보호위원회 보도자료, 「개인정보위, 인공지능(AI) 자율점검표 발표 - 인공지능 개발·운영 시 자율점검이 가능한 개인정보보호 기준 제시 -」, 2021. 6. 1.

9) 개인정보보호위원회 보도자료, 「개인정보위, 인공지능(AI) 자율점검표 발표 - 인공지능 개발·운영 시 자율점검이 가능한 개인정보보호 기준 제시 -」, 2021. 6. 1.

[표 3] AI 관련 개인정보보호 6대 원칙

- ① 적법성 : 개인정보의 수집·이용·제공 등 처리의 근거는 적법·명확해야 한다.
- ② 안정성 : 개인정보를 안전하게 처리하고 관리한다.
- ③ 투명성 : 개인정보 처리 내역을 정보주체가 알기 쉽게 공개한다.
- ④ 참여성 : 개인정보 처리에 대한 소통체계를 갖추고 정보주체의 권리를 보장한다.
- ⑤ 책임성 : 개인정보 처리에 대한 관리 책임을 명확히 한다.
- ⑥ 공정성 : 개인정보를 수집 목적에 맞게 처리하여 사회적 차별·편향 등 발생을 최소화한다.

※ 자료: 개인정보보호위원회(2021)

또한 동 점검표는 상기 6대 원칙에 따라 업무처리 단계(기획·설계, 개인정보 수집, 이용·제공 등 8 단계)에서의 16개 점검항목과 54개 확인사항을 제시하고 있다. 특히 점검항목은 “의무사항”과 “권장사항”으로 구별되어 있음을 참고할 수 있다. 예를 들어, “AI 서비스 운영 과정에서 개인정보 유출 시 정보주체 통지, 유출 신고, 피해구제 지원 등에 관한 대응 절차를 마련하여 이행하는가?”는 의무사항으로, “AI 개발·운영 과정에서 자율적인 개인정보 보호활동을 적극적으로 수행하는가?”는 권장사항으로 명시하였다.¹⁰⁾

나. 국외 사례

EU 집행위원회는 2020년 7월 17일 스스로 평가(self-assessment)해 볼 수 있는 “신뢰할 수 있는 인공지능을 위한 평가 목록(Assessment List for Trustworthy Artificial Intelligence, 이하, 평가 목록)”을 발표하였다. 동 평가 목록은 사용자가 불필요한 위험에 노출되지 않고 AI의 이점을 누릴 수 있도록 자체 평가 항목을 제시하였다.¹¹⁾ 평가 목록은 AI 윤리 요구사항(① 인간 행위자와 감독, ② 기술적 견고성과 안정성, ③ 프라이버시·데이터 거버

10) 개인정보보호위원회, 「인공지능(AI) 개인정보보호 자율점검표」, 2021. 5. 31.

11) AI HLEG(High-Level Expert Group on AI), 「The Assessment List For Trustworthy Artificial Intelligence for self assessment」, 2020.

년스, ④ 투명성, ⑤ 다양성, 차별금지, 공정성, ⑥ 사회·환경적 웰빙, ⑦ 책임성)에 대한 구체적인 질문들로 구성되어 있다(예컨대, 책임성과 관련하여 “독립적인 제3자가 AI 시스템을 감사할 수 있는지 확인했습니까?” 등).

또한 동 평가 목록은 추상적인 용어를 쉽게 이해할 수 있도록 용어사전(Glossary)을 첨부하였다는 점도 참고할 수 있다. 예컨대, “AI 신뢰성(AI reliability)”에 대하여 “AI 시스템이 이전에 훈련되거나 테스트 되지 않은 새로운 입력에 대해서도 예상대로 작동하면 신뢰할 수 있다고 한다.”고 정의하고 있으며, “책임성(Accountability)”은 “자신의 행동과 그에 따른 결과에 대하여 책임이 있으며, 목적과 동기, 이유를 설명할 수 있어야 한다.”고 하였다. 그리고 “공정성(Fairness)”에 대하여는 “개인들 사이, 그리고 개인과 그룹 사이의 평등한 대우라는 이상을 구현하는 것이며, 개인의 권리와 자유가 침해되었을 때 구제받을 수 있는 절차적인 관점도 포함된다.”고 설명하고 있다.

(3) 시사점

개인정보보호위원회의 “인공지능 자율점검표” 상의 6대 원칙은 「개인정보 보호법」의 개인정보 보호 원칙¹²⁾을 기본으로 하여 도출된 것이다.¹³⁾ 따라서 금융분야 AI 자율점검표를 도입할 경우에도 「금융소비자 보호에 관한 법률」, 「자본시장과 금융투자업에 관한 법률」 등을 기반으로 하여 금융시장에 특화된 AI 운영 원칙과 점검항목을 마련할 필요가 있을 것이다. 예를 들어, 「금융소비자 보호에 관한 법률」은 금융소비자의 기본적 권리¹⁴⁾와 영업규

제¹⁵⁾ 등을 규정하고 있는데, 이와 연계하여 AI 관련 금융소비자 보호 원칙과 평가 목록을 마련할 수 있을 것이다. 그리고 평가 목록을 작성함에 있어서 “의무사항”과 “권장사항”으로 명확히 구별하여 금융회사 등의 부담을 완화하는 방안을 생각해 볼 수 있다. 또한 EU 집행위원회의 평가 목록과 같이 금융분야 AI 가이드라인 상의 추상적인 용어들의 개념을 명확히 할 필요가 있다.

4 나가며

금융분야 AI 가이드라인은 소비자 보호 측면에서 인공지능 거버넌스 및 시스템 구축에 대한 방향성을 제시하였다는 의미가 있으나, 타 부처의 AI 가이드라인과의 정합성 및 기존 금융시장 관리체계를 고려하여 자율점검표를 마련하는 방안 등을 더 구체적으로 논의할 필요가 있다.

과학기술정보통신부는 학계, 기업, 시민단체 등 다양한 사회구성원이 참여하여 AI 윤리에 대해 의견을 수렴하는 “윤리 정책 플랫폼”을 마련한다고 발표한 바 있다.¹⁶⁾ 금융분야에서도 합리적인 AI 윤리 기준과 정책에 대한 활발한 논의의 장이 계속 이루어지길 바란다.

『이슈와 논점』은 국회의원의 입법활동을 지원하기 위해 최신 국내외 동향 및 현안에 대해 수시로 발간하는 정보 소식지입니다.

12) 「개인정보 보호법」 제3조(개인정보 보호 원칙) ① 개인정보처리자는 개인정보의 처리 목적을 명확하게 하여야 하고 그 목적에 필요한 범위에서 최소한의 개인정보만을 적법하고 정당하게 수집하여야 한다. (후략)

13) 개인정보보호위원회, 「인공지능(AI) 개인정보보호 자율점검표」, 2021. 5. 31.

14) 「금융소비자 보호에 관한 법률」 제7조(금융소비자의 기본적 권리) 금융소비자는 다음 각 호의 기본적 권리를 가진다.

1. 금융상품판매업자등의 위법한 영업행위로 인한 재산상 손해로부터 보호받을 권리

2. 금융상품을 선택하고 소비하는 과정에서 필요한 지식 및 정보를 제공받을 권리

3. 금융소비생활에 영향을 주는 국가 및 지방자치단체의 정책에 대하여 의견을 반영시킬 권리

4. 금융상품의 소비로 인하여 입은 피해에 대하여 신속·공정한 절차에 따라 적절한 보상을 받을 권리

5. 합리적인 금융소비생활을 위하여 필요한 교육을 받을 권리

6. 금융소비자 스스로의 권익을 증진하기 위하여 단체를 조직하고 이를 통하여 활동할 수 있는 권리

15) 「금융소비자 보호에 관한 법률」 제13조 내지 제28조

16) 과학기술정보통신부, 「과기정통부, 신뢰할 수 있는 인공지능 실현전략 발표」, 2021. 5. 14.

