한국정보통신기술협회 Telecommunications Technology Association http://www.tta.or.kr	**보도자료**	*ICT 표준화와 시험인증의 글로벌 리더*
일시	**2025. 2. 11(화) 배포 / 배포시부터 보도해 주시기 바랍니다.**	
담당	TTA AI융합시험연구소 AI신뢰성센터 곽준호 팀장(010-5110-2693), 신예진 책임(010-5110-6093)	

TTA, 범용 AI 안전성 강화를 위한 위험 관리 프레임워크 발표

- 범용 AI 위험 요소 체계화 및 국제표준 기반 관리 체계 제시 -

한국정보통신기술협회(회장 손승현, 이하 TTA)는 2025년 2월 10일, AI신뢰성센터를 통해 '범용 인공지능 (GPAI) 위험 관리 프레임워크' 연구 보고서를 발표했다. 이 보고서는 범용 AI (General-Purpose AI, GPAI) 기술의 잠재적 위험을 체계적으로 식별하고 분석하여, 이를 효과적으로 관리할 수 있는 국제표준 기반의 위험 관리 프레임워크를 제시한다.

TTA는 2021년부터 과학기술정보통신부 및 관계 기관과 협력하여 인공지능 신뢰성 인식 제고와 확산을 위한 다양한 노력을 기울여왔다. 이번 연구는 급격히 발전하는 범용 AI 기술이 사회적, 윤리적 영향을 미칠 가능성을 고려하여 범용 AI의 안전하고 지속 가능한 발전을 지원하기 위한 기초적인 방향성을 제시하고자 추진되었다.

보고서는 기존 연구에서 도출된 위험 요소를 분석하고, 이를 기반으로 범용 AI에 특화된 8가지 핵심 위험 요소를 정의하였다. 이를 통해 범용 AI 기술과 관련된 이해관계자들이 위험을 보다 명확히 파악하고, 체계적으로 대비할 수 있도록 돕는다.

또한, 국제표준을 참고하여 범용 AI 기술의 특수성을 반영한 위험 관리 프레임워크를 설계하였으며, 관련 조직에서 활용 가능한 가이드와 함께 사례 분석을 통해 프레임워크의 효과를 검증하였다.

본 보고서의 AI 위험관리 프레임워크는 국내의 인공지능 사업자들이 안전하고 신뢰할 수 있는 인공지능 서비스를 제공하는데 실질적인 도움을 줄 수 있도록 국내 산·학·연 전문가 20명으로 구성된 연구자문위원회의 검토를 거쳐 완성도를 높였다.

또한 AI 분야 세계적인 석학인 요슈아 벤지오 캐나다 몬트리올대 교수와 스튜어트 러셀 UC버클리대 교수의 의견수렴을 통해 국제적 공신력도 확보할 수 있었다.

TTA 손승현 회장은 “최근 중국의 딥시크(DeepSeek) 공개 이후, 개인정보 보호, 저작권, 국가 안보 등 다양한 측면에서의 안전성 이슈가 부각되며 AI 위험에 대한 인식이 높아지고 있다” 면서 “이번 연구 보고서가 AI 모델 및 서비스 개발 기업들이 인공지능을 안전하게 관리하는 데 중요한 참고자료로 활용되기를 기대한다” 라고 밝혔다.