

보도시점 2026. 6. 17.(수) 10:00 배포 2026. 6. 16.(화) 16:00
(2026. 6. 17.(수) 석간)

AI 안전 협력 강화를 위한

(한) AI안전연구소 - (미) 오픈AI 간 업무협약(MoU) 체결

- 고위험 분야에서 최첨단 AI 모델의 안전 확보를 위한 협력 체계 구축

【관련 국정과제】 23. 국민의 안전과 보편적 삶의 질 제고를 위한 ‘AI기본사회’ 실현

과학기술정보통신부(부총리 겸 과기정통부 장관 배경훈, 이하 ‘과기정통부’)는 6월 17일(수) 한국전자통신연구원(ETRI) 서울사무소에서 최첨단 AI의 안전성 확보를 위해 인공지능안전연구소(소장 김명주, 이하 ‘AI안전연구소’)와 오픈AI 간 고위험 분야 AI 안전 협력 강화를 위한 업무협약(MoU)을 체결했다고 밝혔다.

이번 업무협약 체결은 지난해 10월 과기정통부와 오픈AI 간 MoU 체결을 통해 구축한 AI 협력기반을 AI 안전 분야까지 확대하기 위한 것으로, 올해 두 차례 과기정통부 제2차관과 오픈AI 고위관계자와의 면담 등에서 논의된 AI 안전 협력 방안을 토대로 AI안전연구소와 오픈AI 간 실질적인 협력을 본격화하기 위해 이루어졌다.

AI안전연구소와 오픈AI는 이번 업무협약을 통해 고위험 분야별 안전 평가 방법론과 벤치마크에 관한 지식과 모범사례를 공유하고, 한국어와 한국의 사회·맥락을 반영한 평가 체계 개발을 위한 기술 정보를 교환하며, 국제적으로 적용 가능한 AI 안전 평가 체계 마련을 위해 지속적으로 협력해 나가기로 하였다.

오픈AI가 각국의 AI안전연구소와 업무협약을 체결한 것은 미국, 영국, 일본에 이어 이번이 네 번째로, 한국은 이번 협약을 계기로 최첨단 AI의 위험 검증과 평가 기준 마련을 위한 국제 협력 네트워크에서의 역할을 한층 강화할 수 있게 되었다.

과학기술정보통신부 이진수 인공지능정책기획관은 “고성능 AI, 자율형 에이전트 AI 등 최첨단 AI 모델의 안전성 확보를 위해, 글로벌 선도기업과의 협력을 강화하고 AI 안전성 평가 체계를 고도화해 나갈 시점이다”면서, “이번 협약을 계기로 양 기관이 급변하는 글로벌 AI 기술과 이용 환경에 대응하여 최첨단 AI의 안전 확보에 적극 협력하길 기대한다”고 밝혔다.

AI안전연구소 김명주 소장은 “AI가 국가 핵심 인프라와 안보에 미치는 영향이 커질수록 고위험 분야에 대한 엄밀한 안전 평가는 선택이 아닌 필수가 되고 있다.”라며, “이번 협약을 통해 AI안전연구소는 오픈AI와 함께 최첨단 AI의 위험을 과학적으로 검증하고, 국제적으로 통용되는 평가 체계를 만드는 데 기여하겠다.”고 밝혔다.

오픈AI 이상현 아시아태평양 지역 정책 총괄은 “한국은 AI기술의 활용과 혁신이 빠르게 이루어지고 있는 중요한 나라이며, AI안전연구소는 책임있는 AI 발전을 위한 국제 논의에 중요한 역할을 하고 있다.”라며, “이번 협력을 통해 고위험 분야 AI 안전성 평가에 대한 지식과 모범 사례를 공유하고, 신뢰할 수 있는 AI 개발과 안전한 이용 환경 조성에 함께 기여하게 되어 뜻깊게 생각한다.”고 말했다.

AI안전연구소와 오픈AI는 업무협약 체결 이후에 실무 협의를 거쳐 구체적인 협력 과제와 일정을 확정해 나갈 예정이다.

담당 부서	과학기술정보통신부 인공지능안전신뢰정책과	책임자	과 장	장두원 (044-202-6290)
		담당자	사무관	김현희 (044-202-6295)
유관 기관	인공지능안전연구소 AI안전정책 및 대외협력실	책임자	실 장	최민석 (031-739-7654)

