

“인공지능서비스, 불법·유해정보 대응해야”

- 이용자보호 위한 민관협의회 2차 회의 개최...“사업자 자율조치 방안 필요” -

인공지능서비스 과정에서 불법·유해정보 노출 위험이 커지는 가운데 이용자 보호를 위한 사업자의 자율적 안전조치 방안 등이 논의됐다.

방송미디어통신위원회(위원장 김종철)는 27일 서울에서 인공지능서비스 불법·유해정보 대응을 위한 기술적 조치 동향 등을 논의하기 위해 정부, 산업계, 학계 등이 모인 가운데 ‘인공지능서비스 이용자보호 민관협의회 2차 회의’를 개최했다.

인공지능서비스 확산에 따른 기존 제도의 사각지대를 진단해 안전한 서비스 환경 조성을 위한 다층적인 제도 개선 방안을 모색하기 위한 것으로, 정보통신기술(ICT)·법률·미디어 관련 전문가와 소비자·시민단체, 통신사·플랫폼·인공지능 기업 등 사업자, 정부·공공기관 관계자 등이 참석했다.

이날 2차 회의에서는 김경훈 카카오 인공지능 안전(AI Safety) 리더가 발제를 맡아 인공지능서비스 불법·유해정보 대응을 위한 가드레일 시스템 현황에 대해 발표했다.

국내 최초 한국어 환경에 특화된 오픈소스 기반으로 카카오가 자체 개발한 인공지능 안정성 검증을 위한 가드레일 모델인 ‘카나나 세이프카드(Kanana Safeguard)’를 직접 소개하고, 사업자의 자율적 안전조치 고도화 방향 및 생태계 협력 방안 등에 대해 공유했다.

이어 이성엽 민관협의회 위원이 주재한 종합토론에서는 인공지능서비스 내 불법·유해정보 노출 위험이 커진 가운데 사업자의 자율적 안전조치 활성화를 위한 민관 협력 방안 등에 대해 심도 있는 논의가 이어졌다.

이성엽 민관협의회 위원은 “이번 발표와 종합토론에서 논의된 기술적 동향 등이 사업자의 이용자 보호를 위한 법제 마련 시 제도의 실효성을 제고하는 방향으로 반영될 필요가 있다”고 밝혔다.

방미통위 신영규 방송통신이용자정책국장은 “윤리적으로 신뢰할 수 없는 인공지능 기술은 사회적 불신을 초래해 결국 인공지능 산업 전체의 성장을 가로막을 수 있으므로 기술 진흥과 이용자 보호를 위한 기술적 안전조치는 반드시 동반되어야 한다”고 강조했다.

그러면서 “오늘 논의된 내용을 바탕으로 법과 기술, 제도 등 종합적 지원 방안을 마련해 국민이 안심하고 신뢰할 수 있는 인공지능 이용 환경을 만드는 데 최선을 다하겠다”고 말했다.

한편 방미통위는 지난 2021년부터 ‘지능정보사회* 이용자 보호 민관협의회’를 운영해 왔으며, 2024년부터는 인공지능 환경 변화에 대응하기 위해 ‘인공지능서비스 이용자 보호 민관협의회’로 확대 개편했다.

* 지능정보서비스: 인공지능, 메타버스, 빅데이터, 사물통신, 클라우드 컴퓨팅 등 신기술 기반 서비스

이어 올 5월 정부, 산업계, 법조계, 학계, 시민단체 등 각계 전문가 35명으로 구성된 민관협의회 2기를 출범하고 불법·유해정보로부터 아동·청소년 보호를 위한 기술적 제도적 방안 등을 공유하며 관련 정책 방안 등을 심도 있게 논의하고 있다.

붙임. ‘인공지능서비스 이용자 보호 민관협의회 2차 회의’ 계획안

담당 부서	방송통신이용자정책국 인공지능이용자보호과	책임자	과 장	김혜숙 (02-2110-1570)
		담당자	사무관	김기호 (02-2110-1573)



□ 개 요

- (목적) 인공지능서비스 확산에 따른 기존 제도의 사각지대를 진단하여 안전한 인공지능서비스 기반 조성을 위한 다층적인 제도 개선 방안 모색
- (일시 / 장소) '26. 8. 27.(목) 14:00~16:00 / 엘타워 루비홀
- (참석) 이성엽 민관협의회 위원(고려대 기술경영전문대학원 교수) 포함 **29명**
 - 전문가(ICT·법률·미디어 관련 학계, 법조계 등), 소비자·시민단체, 사업자(통신사·플랫폼·인공지능 기업 등), 정부·공공(방미통위*, KISDI) 등
 - * 방송통신이용자정책국장 신영규, 인공지능이용자보호과장 김혜숙
- (주제) 인공지능서비스 불법·유해정보 대응을 위한 기술적 조치 동향
- (주요 내용) 인공지능서비스 내 불법·유해정보 노출 위험이 커지는 가운데 사업자의 자율적 안전조치 활성화를 위한 정책 방향을 논의

□ 세부 계획(안)

시 간		주요 내용	비고
14:00 ~ 14:05	5'	○ 단체 기념 촬영	
14:05 ~ 14:10	5'	○ 위원 소개	
14:10 ~ 14:15	5'	○ 개회사	· 이성엽 위원(고려대)
14:15 ~ 14:20	5'	○ 인사말씀	· 신영규 국장(방미통위)
14:20 ~ 14:45	25'	○ 발제 - 인공지능서비스 불법·유해정보 대응을 위한 가드레일 시스템 현황	· 김경훈 AI Safety 리더 (카카오)
14:45 ~ 15:55	70'	○ 종합 토론	· 이성엽 위원(고려대)
15:55 ~ 16:00	5'	○ 마무리	