

## KISO, 딥페이크 기반 허위조작정보 대응 강화

### 딥페이크·생성형 인공지능 등 신기술 자율규제 방안 모색 위법 게시물 관련 관계기관과 긴밀히 협력

네이버, 카카오, SK커뮤니케이션즈(이하 SK컴즈) 등 국내 포털과 인공지능(AI) 챗봇 기업들이 회원사로 가입해 있는 한국인터넷자율정책기구(KISO)가 최근 사회적 이슈로 떠오르고 있는 ‘딥페이크<sup>1)</sup>에 의해 생성된 허위조작정보’에 대응하기 위한 논의에 착수했다.

KISO 산하 신기술위원회(위원장 이재신 중앙대 미디어커뮤니케이션학부 교수)는 최근 열린 회의에서 딥페이크 등 인공지능 기술 오남용 문제와 관련해 회원사들의 대응 현황을 공유하고 향후 대책을 논의했다. 학계 전문가들로 구성된 신기술위원회에는 국내 대표 포털인 네이버와 카카오를 비롯해 심심이, 스캐터랩, 튜넵 등 챗봇서비스 업체가 참여하고 있다.

신기술위원회는 특히 4월 국회의원 선거를 앞두고 딥페이크 기술을 이용한 허위조작정보가 확산할 가능성을 예의주시하며 구체적인 사례를 수집·분석하기로 했다. 또한 설문조사 등 이용자의 의견을 수렴한 후 자율규제 측면에서 적절한 대응책을 마련해 나가기로 했다.

KISO는 선거의 공정성을 확보하기 위해 중앙선거관리위원회와 협력해 딥페이크 영상을 이용한 허위조작정보에 대해 명확히 확인된 경우, 빠르게 조치할 계획이다.

1) 딥 러닝(Deep learning)과 가짜(Fake)의 합성어인 딥페이크(Deepfake)는 인공지능을 활용하여 진위 여부를 구별하기 어려운 이미지, 음성, 동영상 등 다양한 콘텐츠를 제작할 수 있는 기술을 의미한다.

관련하여 네이버는 신고센터 메인 페이지에 △선거 관련 허위정보 △답페이지 생성물에 신고 영역을 신설해 이용자들이 최대한 거르고 대처할 수 있는 체계를 마련했다. 또한 블로그, 카페 등에서 동영상을 게시하거나 답페이지 관련 검색어 입력시 답페이지 활용에 관한 주의 및 안내 문구를 제공해 이용자를 보호하고 있다. 이 밖에도 뉴스 서비스에서는 AI 로봇이 생성한 기사를 이용자가 인지하도록 기사 본문에 해당 사실의 표기를 강조하고 있다.

카카오는 이용자들이 스스로 쉽게 AI 저작물을 파악하는데 도움이 되는 방향으로 정책과 기술적 조치 방안을 마련했다. 카카오는 언론사에서 AI를 이용해 생성한 기사 상단에 “AI가 생성한 기사입니다”라고 노출해 이용자가 쉽게 인지할 수 있도록 했다. 생성형 AI 이미지 생성 모델 ‘달로’에는 비가시성 워터마크를 부착하는 기술을 도입했다. 또한 키워드 검색, 배너 등을 활용한 캠페인을 통해 이용자의 주의를 환기하고, 답페이지 신고 요령 등을 자세히 안내할 예정이다.

SK컴즈는 ‘AI챗’ 서비스를 이용한 이미지 생성 시 주요 정치인의 성명에 관한 ‘검색어 제한 조치’를 통해 선거 관련 답페이지 제작을 방지하는 방안을 마련하고, 선거 관련 허위 정보 신고 채널 운영과 서비스 내 답페이지 주의 안내 문구를 표시하여 이용자의 주의를 환기할 예정이다.

앞서 KISO는 네이버, 카카오, SK컴즈 등 포털 3사와 함께 구축한 ‘청소년 보호 검색어 데이터베이스(DB)’에 답페이지 관련 유해 검색어들을 새롭게 추가하는 등 청소년 이용자 보호에도 이미 적극 나서고 있다.

또한 KISO의 신기술위원회는 지난해 9월 챗봇서비스의 긍정적·부정적 측면을 고려하고 인간과 교감하며 대화하는 AI 챗봇서비스가 인간에게 도움이 되는 방식으로 활용될 수 있도록 하는 ‘챗봇 윤리 가이드라인’을 발표한 바 있다.

‘챗봇 윤리 가이드라인’은 △인간의 존엄성 및 권리 존중 원칙 △프라이버시 보호 및 정보보안 원칙 △다양성 존중 원칙 △투명성 원칙 △책임 원칙으로 구성된 5대 기본원칙을 제안했다. 이에 회원사들은 약관이나 챗봇 대화 화면에서 이 가이드라인을 준수하기 위한 필요한 조치를 이행하고 있다.

이재신 KISO 신기술위원회 위원장은 “온라인 환경에서 사실로 오인 가능한 허위조작정보가 확산되면 사회 신뢰를 저하시키고, 수많은 콘텐츠의 진위를 판별

하는 것이 더욱 어려워질 수 있다”며 “인공지능 기술의 오남용으로 인한 사회적 문제에 대응하기 위한 연구를 활성화 해나갈 것”이라고 밝혔다.

※ 문의: 박엘리 정책팀 팀장([ellee@kiso.or.kr](mailto:ellee@kiso.or.kr), 02-563-6196)