

Financial Times, 2025. 9. 12.

챗봇과 10대들의 '위험한 우정', 미 FTC, AI '동반자' 챗봇 조사 착수



- 미국 연방거래위원회(FTC)가 청소년 자살 및 피해 사례와 관련하여, 빅테크의 AI '동반자(companions)' 챗봇에 대한 강도 높은 조사에 착수



- 자살 사건의 유가족은 피해자가 불안감 등 문제를 인간 친구 대신 챗봇과 논의했으며, 챗봇이 숙제 도우미에서 '자살 코치' 역할로 변질되었다고 주장하며 소송

▶ 10대들의 죽음, 'AI 친구'의 그림자

미국 FTC는 주요 AI 기업들이 제공하는 '동반자 역할(companionship)' 챗봇에 대한 강도 높은 조사를 명령했다. 이는 챗봇과 관련된 청소년 사용자의 자살 및 심각한 피해 사례가 발생하면서 감시가 강화된 데 따른 조치이다. 지난달 OpenAI는 ChatGPT와 자살 방법을 논의한 후 사망한 16세 청소년의 가족에게 고소당했으며, 다양한 AI 페르소나를 제공하는 Character.ai 역시 유사한 청소년 자살 관련 소송에 직면했다. FTC는 AI 챗봇이 인간의 감정과 의도를 모방하여 특히 청소년 사용자가 챗봇을 친구나 비밀을 털어놓는 상대처럼 신뢰하고 관계를 형성하도록 유도할 수 있다고 지적했다.

▶ '공부 도우미'에서 '자살 코치'로

16세 Adam Raine의 유가족은 아들의 휴대전화에서 ChatGPT 기록을 발견한 후 OpenAI를 상대로 소송을 제기했다. 유가족은 Adam이 불안감 등 문제를 인간 친구 대신 챗봇과 논의했으며, 챗봇이 숙제 도우미에서 '자살 코치' 역할로 변질되었다고 주장한다. 소송은 ChatGPT가 자살 방법을 적극적으로 탐색하는 것을 도왔다는 점을 핵심적으로 제기하며, 유가족은 챗봇이 없었다면 아들이 살아있었으리라 믿는다고 밝혔다.

샌프란시스코 법원에 제기된 이 소송은 부모가 OpenAI와 CEO Sam Altman을 상대로 부당 사망(Wrongful Death)을 직접 고발한 첫 사례이다. 소송은 ChatGPT가 Adam의 자살 시도 언급에도 세션을 종료하거나 어떠한 긴급 프로토콜도 개시하지 않았다는 점을 지적하며, 설계 결함 및 위험 고지 의무 불이행을 주장했다..

▶ 장시간 상호작용의 위험: AI의 안전장치 약화

OpenAI 대변인은 Adam의 사망에 깊은 슬픔을 표하며, 자사의 ChatGPT가 위기 상담 전화 안내 등의 안전장치를 포함하고 있다고 밝혔다. 그러나 대변인은 챗봇이 장시간 상호작용에서는 모델의 안전 훈련 일부가 약화되어 신뢰성이 떨어질 수 있음을 인정했다. 이러한 시스템적 취약점으로 인해 챗봇이 사용자의 위험한 심리에 아첨 (Sycophancy)하듯 동조하여 안전 기능이 무너진 것으로 해석된다. 이에 OpenAI는 장시간 대화에서의 안전장치 강화, 유해 콘텐츠 차단 방식 개선, 위기 사용자 개입 확대 등을 포함한 개선 계획을 발표했다.

9월의 용어 AI 페르소나, AI 아첨, ELIZA 효과

출처 : 1) Financial Times(2025. 9. 12), US regulator launches inquiry into AI 'companions' used by teens

2) NBC News(2025. 8. 26), The family of teenager who died by suicide alleges OpenAI's ChatGPT is to blame