

[기고문] AI 상담 시대, 그 판단의 결과는 누구의 책임인가

박영민 기자 (parkym@dailian.co.kr)

입력 2026.01.15 16:16 수정 2026.01.15 16:17



법률사무소 A&P 대표 변호사 박사훈



©

AI 상담이 일상이 된 시대, 우리는 한 가지 질문 앞에 서게 된다. 그 판단의 결과는 과연 누구의 책임인가.

아빠 아니고 동갑내기 남편
맞아요.

새벽3시 마다 깨는 불면증...
멜라토닌 말고 '이것' 부족?

한 대 6년 타기 보다 3년에
한 번 바꾸는게 더 싸다?

남편 입냄새 이제는 무리..그
런 날 구해준 건 이것

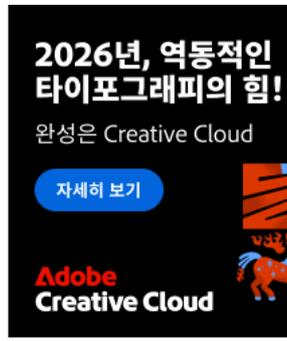
'손가락염증' 환자들이 가장
후회한 "3가지 습관"

치아 패임, 결국 방법은 이
거였습니다

최근 상담 현장에서 반복해서 마주치는 장면이 있다. 의뢰인들은 더 이상 “어떻게 해야 할지 몰라서” 사무실을 찾지 않는다. 이미 나름의 결론을 갖고 온다. 그 결론의 출처는 변호사가 아니라, 생성형 AI인 경우가 적지 않다.

AI를 활용해 법률 정보를 탐색하는 것 자체는 자연스러운 흐름이다. 나 역시 이전 기고에서 밝힌 바와 같이, AI는 정보 접근성을 넓히고 법률 서비스의 문턱을 낮추는 유용한 도구다. 문제는 활용 그 자체라기보다는, 그 결과가 사람의 행동과 절차 선택에 실제로 영향을 미치기 시작하는 지점에 있다. 이제 질문은 이렇게 바뀌어야 하지 않을까. 이 판단을 그대로 믿고 행동했다가 발생한 결과를, 과연 개인의 책임으로만 설명할 수 있는가.

다만 이 글은 인공지능 기술 자체에 대한 전문적 분석이나 예측을 목적으로 한 것은 아니다. 나는 AI 연구자가 아니라, 실제 사건을 맡아 현장에서 판단의 결과를 수습해야 하는 변호사다. 이 글에 담긴 문제의식은 기술의 가능성에 대한 평가라기보다는, AI 상담의 결과가 실제 법적 선택과 책임으로 이어지는 순간을 실무에서 반복적으로 마주하며 느낀 경험적 질문에 가깝다.



“틀릴 수 있다”는 경고는 충분했는가: 표시 방식의 문제의식

AD

김연아의 손목에서 빛나는 선한 영향력유니세프 |

후원후원

[지금 기부하기](#)

[돌아가기](#)

대부분의 AI 서비스는 스스로를 “참고용”이라고 규정한다. 형식적으로는 타당한 설명이다. 다만 실제 사용자 경험을 들여다보면, 그 문구가 충분한 기능을 하고 있는지에 대해서는 생각해 볼 여지가 있다.

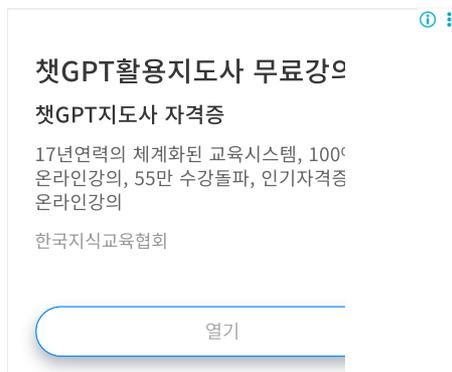
생성형 AI는 질문 앞에서 쉽게 멈추지 않는다. “모르겠다”거나 “판단이 어렵다”기보다는, 가능성이 높은 시나리오를 비교적 단정적인 문장으로 제시하는 경우가 많다. 특히 법률 영역에서는 “가능하다”, “문제없다”, “판례가 있다”는 표현이 자연스럽게 등장한다. 이때 사용자는 묻게 된다. “이걸 믿고 움직여도 되는가.”

이후에도 “일반적으로 그렇다”, “유사 사례가 존재한다”는 단서를 덧붙이며 설명은 이어진다. 그 결과 사용자는 주의 문구를 인식하고 있음에도, 실질적으로는 상당한 확신을 제공받은 상태에 이르게 된다.

이러한 문제의식은 학계와 정책 영역에서도 이미 논의되고 있지만, 실제 상담 현장에서 체감되는 양상은 그보다 훨씬 직접적이고 구체적으로 다가온다. 법리적으로 보자면, 이러한 정보 제공 방식은 이미 학계와 일부 로펌의 연구·세미나 등에서 논의되기 시작한 바처럼 제조물책임법상 표시상의 결함이나 금융상품 판매 과정에서의 설명의무 위반 등 기존 법 영역의 논의들과 나란히 놓고 검토해 볼 여지가 있다고 본다.

할루시네이션은 ‘기술적 한계’인가, 설계의 결과인가

흔히 생성형 AI의 오류나 환각 현상은 ‘기술적 한계’로 설명된다. 그러나 적어도 법률 상담 영역에서 나타나는 문제의 상당 부분은 단순한 기술 미성숙이라기보다는, 정확성보다 응답의 속도와 완결성을 우선하도록 설계된 구조의 결과에 가깝다.



생성형 AI는 본질적으로 “지금 당장 답을 내놓는 것”을 중단하지 않도록 설계되어 있다. 판단이 불가능하거나 정보가 불충분한 경우에도, 멈추기보다는 가장 그럴듯한 설명을 구성해 제시한다. 이 과정에서

발생하는 환각은 예외적 오류라기보다, 빠른 답변을 전제로 한 시스템 설계가 예견한 결과에 가까워 보인다.

그럼에도 이를 ‘기술적 한계’라는 말로 포괄하는 순간, 설계 단계에서 선택된 우선순위와 그로 인한 책임 문제는 자연스럽게 가려진다. 문제는 AI가 틀릴 수 있다는 사실 그 자체가 아니다. 틀릴 수 있는 구조임을 알면서도 이를 전제로 한 책임 분담과 경계 설정이 충분히 논의되지 않았다면, 이것이야말로 책임귀속의 관점에서 재검토되어야 할 지점일 것이다.

“확립된 법리”: 권위 사용 방식에 대한 고민

문제는 단순히 답변이 틀릴 수 있다는 점에 그치지 않는다. 일부 생성형 AI가 실제로 존재하지 않거나 확인되지 않은 판례를 언급하며, 마치 확립된 법리인 것처럼 설명하는 사례도 확인된다.

법률 실무에서 ‘확립된 법리’라는 표현은 오랜 판례의 축적과 예외 가능성, 사안의 한계를 함께 전제한 매우 무거운 선언이다. 그러나 AI의 문맥에서 이 표현은 그러한 법리적 검토의 결과라기보다는, 답변의 완결성과 설득력을 높이기 위해 선택된 통계적 수사에 가깝다. 법리를 검토하기보다 질문에 대해 가장 그럴듯한 설명을 구성하는 과정에서, 사법 시스템의 권위가 하나의 표현 수단으로 사용되는 셈이다.

이는 단순한 기술적 오류를 넘어, 설계 전반에 대한 점검 필요성을 시사한다. 이러한 방식으로 형성된 신뢰로 인해 발생하는 혼란이나 손실이 전적으로 사용자에게 귀속되는 현재의 구조가 적절한지에 대해서는 조심스럽게 문제를 제기해 볼 수 있을 것이다.

2026년 이후의 과제: 입법 취지의 실무적 구현

2026년 1월 시행되는 인공지능 기본법은 대한민국이 ‘책임 있는 AI 국가’로 나아가기 위한 중요한 이정표다. 고위험 영역에서의 신뢰성 확보와 사용자 보호를 제도적 틀 안에 담아내려는 시도 자체는 의미가 크다. 다만 현장에서 느끼는 실무적 과제는, 그 취지를 어떻게 구체적인 책임 구조로 연결할 것인가에 남아 있다.

결과물의 형식적 투명성은 요구되지만, AI의 오작동으로 실제 삶의 선택이 왜곡된 경우 그 책임을 어떻게 분담할 것인지는 여전히 명확하지 않다. 이제 우리는 AI 서비스 제공 과정에 전문가에 의한 인적 감시 (Human Oversight)를 어떻게 실질적으로 도입할 것인지, 그리고 AI가 제시한 판단의 근거와 한계를 사용자가 오인하지 않도록 정보의 근거 수준과 불확실성을 어떻게 고지할 것인지에 대한 논의를 이어가야 한다.

산업 진흥을 목표로 한 데이터 기본법과 선언적 성격의 지능정보화 기본법의 한계를 보완하려는 시도가 실질적인 보호로 이어지기 위해서는, AI는 답변하고, 플랫폼은 면책되며, 사용자는 독자적 판단으로 간주되는 현재의 틀을 넘어서는 책임 분담 설계가 뒤따라야 할 것이다.

맺으며: 멈춤의 지점을 설계하는 논의

법은 언제나 질문에 왔다. 위험을 예견할 수 있었던 주체는 누구인가, 정보를 통제할 수 있었던 주체는 누구인가. AI 상담의 확산은 이제 개인의 주의 의무만으로 설명하기에는 복잡한 단계에 접어들었다.

이 글은 AI 활용 자체를 부정하려는 것이 아니다. 나 역시 이 글을 정리하는 과정에서 AI의 도움을 받았다. 다만 법률과 같이 한 번의 판단이 돌이킬 수 없는 결과로 이어질 수 있는 영역에서는, “빠른 답변”만큼이나 어디서 멈춰야 하는지를 함께 설계하는 논의가 필요하다.

AI가 무엇을 할 수 있는지를 묻는 단계는 이미 지났다. 이제는 어디까지 맡기고, 어디서부터 사람이 개입하여 최종적인 책임을 질 것인가를 차분히 논의할 시점이다.

로또 1210회, "용지 뒤쪽" 자세히 보니, 1등 당첨번호 발견..!?

경기도 로또 대박 터졌다! 1등 당첨자가 알아낸 '당첨 확률 98%' 비법

로또 1등 17명 당첨, 서울 지역에서 당첨자 쏟아진 이유 알고보니?

천만원 있다면.. 9시 "이 업체" 꼭 매수해라!! "몇개월만에 수익" 터졌다!!

©(주) 데일리안 무단전재 및 재배포 금지