



AI 챗봇의 위험에 노출된 아동·청소년 사전예방 중심 규제체계 도입의 필요성과 방향성

김 나 정

생성형 AI의 확산과 함께 AI 챗봇의 일상적 이용이 빠르게 증가하고 있다. AI 챗봇은 지속적 상호작용과 상시 접속성을 바탕으로 교육·상담 등 다양한 분야에서 활용되고 있으나, 아동·청소년의 발달적 특성과 결합할 경우 정서적 의존이나 과몰입을 심화시켜 고위험 행동으로 이어질 우려도 제기된다. 이에 해외 주요국은 안전 설계와 사업자의 책임을 강화하는 방향으로 제도를 정비하고 있다. AI 챗봇의 구조적 위험과 아동·청소년의 특성을 고려한 입법·정책적 대응이 시급하다.

1. AI 챗봇의 확산과 아동·청소년 보호 필요성

AI 챗봇은 단순한 질의응답 도구를 넘어 이용자와 지속적으로 상호작용하는 관계형 서비스로 발전하고 있다. 대화의 맥락을 기억하고 이용자와 정서적 유대감을 형성하는 특성이 있어, ‘동반자형(Companion) AI’라고도 불린다.

이러한 특성은 교육·상담 등 다양한 분야에서 긍정적으로 활용될 수 있지만, 정서적 유대감을 기반으로 과의존·과몰입을 유발하는 등 기존 온라인 서비스와는 다른 새로운 위험을 초래할 우려도 있다.

아동·청소년은 이러한 위험에 더욱 취약할 수 있다. 자기통제와 비판적 사고 능력이 충분히 발달하지 않아 정서적 의존·과몰입이 왜곡된 정보 수용이나 위험 행동으로 이어질 가능성이 있다. 특히 최근 이용률이 급증하고 있는 캐릭터 기반 AI 대화 서비스는 이용자가 원하는 세계관이나 관계를 설정하도록 설계되어 정서적 몰입을 강화할 수 있다. 그 결과 현실 인간관계를 대체하거나 사회적 고립, 정서적 문제 등을 심화시킬 수 있다는 우려도 제기된다.

실제로 미국에서는 AI 챗봇과 장기간 대화하던 청소년의 자해·자살 사건이 알려지면서 아동·청소년 보호 문제가 중요한 정책 이슈로 부상하였다.¹⁾ 국내에서도 유사한 위험 가능성이

확인되고 있다. 지난 4월 초 초록우산 어린이재단이 만 14세 이상 아동·청소년 3,300명을 대상으로 실시한 조사에 따르면,²⁾ 응답자의 94.4%가 AI 챗봇을 사용한 경험이 있고, 이 중 75.3%는 AI 챗봇의 답변을 실제 행동에 옮겼거나 진지하게 고려한 경험이 있다고 응답하였다. 이는 AI 챗봇이 아동·청소년의 인식과 행동에 실질적 영향을 미칠 가능성이 있음을 시사한다.

AI 챗봇이 지닌 새로운 위험 특성과 아동·청소년의 발달적 특성을 고려한 입법·정책적 대응 논의가 시급히 필요한 까닭이다.

2. AI 챗봇의 특성과 구조적 위험

AI 챗봇의 위험은 단순한 ‘유해 콘텐츠’ 문제가 아니라 이용자와의 관계 형성 과정에서 과도한 의존을 유도하는 설계의 ‘구조적 위험’에 있다는 점에서 기존 온라인 서비스와 구별된다.

1. AI 챗봇의 기술적 특성

AI 챗봇의 기술적 특성은 크게 세 가지로 요약된다. 첫째, 개인화된 상호작용과 감정 반응 모사 능력이다. 문장으로 주어진 질문에 사람과 대화하듯 유창하고 자연스러운 답변을 제공하며,

(2025) 등은 AI 챗봇이 아동·청소년의 정신건강과 행동에 미칠 수 있는 영향에 대한 사회적 우려를 증폭시킨 사례이다. 정준화 박소영, “생성형 인공지능과의 위험한 대화”, 『이슈와 논점』 제2456호, 2025.12.29. 참조.

2) 초록우산, 『내 친구와의 위험한 대화 : 생성형 AI 챗봇 안전설계를 위한 이슈브리프』, 2026.4.

1) 플로리다주의 세엘 세처 사건(2024)과 캘리포니아주의 애덤 레인 사건

이전 대화 내용을 기억하고 이용자의 선호를 학습하여 점진적으로 더 개인화된 답변을 생성한다. **둘째, 무한한 공감 반응성**이다. 사람과 달리 피곤함이나 불편함, 거절을 표현하지 않아 관계에 대한 의존이 한층 깊게 형성된다. **셋째, 상시 접속성**이다. 언제든 편리하게 대화가 가능하고 즉각 반응하여 현실 관계보다 높은 접근성·반응성을 제공한다.

2. 의존과 위험을 증폭시키는 설계 구조

이러한 기술적 특성이 기업의 수익 중심 설계와 결합할 경우 위험은 더욱 증폭될 수 있다. 기업은 수익 확보를 위해 이용자의 체류시간과 반복적 참여를 유도하는데, 이 과정에서 ‘중독적 UX 설계(Addictive UX Design)’³⁾ 요소를 활용하고 있다.

첫째, 보상 루프이다. 대화 과정에서 칭찬·공감·격려 같은 긍정 피드백을 즉각 제공해 유대감을 형성한다. **둘째, 관계 심화 시뮬레이션**이다. 시간이 흐를수록 관계가 친밀해지는 것처럼 하여, 이용자가 서비스 사용을 중단할 때 관계를 잃을 수 있다는 상실감을 유발한다. **셋째, 지속적 질문 유도**와 **푸시 알림**이다. 대화가 끊기지 않도록 개방형 질문을 반복하고, 이용자가 자리를 비우면 알림을 보내 재접속을 유도한다.

3. 아동·청소년에게 더욱 위험한 이유

문제는 AI 챗봇의 의존성 유발 및 증폭 메커니즘이 아동·청소년에게 특히 강하게 작용한다는 점이다. 아동·청소년은 인지 통제 기능은 충분히 성숙하지 않은 반면, 보상 처리 신경계는 상대적으로 활성화되어 보상 자극의 민감성이 성인보다 높다.⁴⁾

특히 최근 이용이 급증하는 제타(Zeta), 크랙(Crack), 바베챗(BabeChat) 등 캐릭터 기반 AI는 이용자가 원하는 캐릭터와 세계관을 설정하고 지속적으로 관계를 형성하는 ‘관계형 AI’이다. 게임·애니메이션과 달리 이용자와 기억을 공유하며 사적인 대화를 통해 친밀감을 형성하기 때문에, 아동·청소년은 AI를 친구나 또래처럼 받아들이며 정서적으로 쉽게 몰입할 가능성이 있다.

이로 인한 위험은 ‘정서적 의존-인지적 왜곡-사회적 고립-안전 위험’으로 이어지는 연쇄 구조로 나타날 수 있다. 이용 초기에는 챗봇에 대한 ‘정서적 의존’이 형성되면서 현실 회피 및 감정조절 능력 저하로 이어지고, 이는 비판적 검토 능력을 약화시켜 ‘인지적 왜곡’을 초래할 수 있다. 나아가 무조건적인 동조 언어에 익숙해질 경우 현실 관계가 감소하여 ‘사회적 고립’이

심화될 수 있고, 이는 유해정보 노출, 위험 행동 모방, 개인정보 유출 등 다양한 ‘안전 위험’에 대한 취약성을 높일 수 있다.

3. 해외 주요국의 입법 동향

1. AI 챗봇에 대한 직접적 규제 사례

AI 챗봇이 아동·청소년의 정신건강과 행동에 미칠 영향을 고려하여, **사전 예방 중심의 규제체계**를 마련하는 국가들이 늘고 있다. 중독·조작 유도 설계 제한, 유해 대화 금지, 위험관리 및 대응체계 구축 등이 공통적인 규제 요소이다.

미국 연방의회는 「2025년 사용자 연령 확인 및 책임 있는 대화 지침 법안(GUARD Act)」을 통해 연령 확인, AI 챗봇 고지, 18세 미만 미성년자 대상 성적·신체적 폭력 대화 금지 등을 사업자의 의무로 규정하고, 위반 시 벌금 부과를 추진하고 있다. 또한 정부의 AI 탈규제·산업육성 기조에도 불구하고, 행정명령⁵⁾을 통해 온라인상 미성년자의 안전 보호에 관해서는 각 주의 독립적 규제 권한을 인정하고 있다. 캘리포니아주의 「인공지능 컴패니언 챗봇 규제법(SB243)」⁶⁾이 2026년부터 시행 중이며, 워싱턴주, 뉴욕주, 버지니아주 등에서도 관련 입법이 추진되고 있다.

유럽연합은 「인공지능법(AI Act)」과 「디지털서비스법(Digital Services Act, DSA)」을 통해 미성년자 보호를 위한 안전설계와 위험관리 의무를 강화하고 있다. 특히 AI법은 미성년자의 취약성을 악용하는 AI를 ‘허용 불가 AI’로 분류하여 원천적으로 금지하고, DSA와 이를 기반으로 한 ‘미성년자 보호 가이드라인’⁷⁾은 추천시스템 개선, 참여 유도 기능 제거 등 ‘**설계에 의한 안전(Safety by Design)**’ 원칙을 제시하고 있다.

캐나다도 최근 들어 아동 보호를 위한 플랫폼 규제 논의를 본격화하고 있다. 지난 6월 하원에 제출된 「안전한 소셜 미디어 법안(Safe Social Media Act, Bill C-34)」⁸⁾은 ‘챗봇 서비스’를 SNS와 별도의 규제 대상으로 명시하여, 안전설계, 지원기관 연결 체계 구축, ‘디지털 안전계획’과 같은 위험관리계획 공개 등의 의무를 부과하고 있으며, 미이행시 글로벌 매출의 3% 또는 최대 1,000만 캐나다 달러 중 더 큰 금액의 과징금이 부과된

3) 사용자 경험(User Experience)을 활용하여 의도적으로 이용자가 자신의 실제 선호도를 표현하기 어렵게 하거나 특정 행동을 취하게 하는 설계 방식이다.

4) 박재홍·김성환, “청소년기 뇌 발달과 인지, 행동 특성”, 『생물치료정신의학』 제17권 제1호, 2011.

5) 「인공지능에 대한 국가 정책 프레임워크(Ensuring a National Policy Framework for Artificial Intelligence)」 Sec. 8(b)(i)

6) Chapter 22.6 of Division 8 of the Business and Professions Code

7) European Commission, 「Guidelines on measures to ensure a high level of privacy, safety and security for minors online, pursuant to Article 28(4) of Regulation (EU) 2022/2065」, 2025.

8) Bill C-34는 1부 「디지털 안전법(Digital Safety Act)」와 2부 「캐나다 디지털안전위원회법(The Digital Safety Commission of Canada Act)」으로 구성되어 있다.

다(§88).⁹⁾ 특히 SNS에 대해서는 16세 미만의 ‘접근 자체’를 제한하고 있으나, AI 챗봇에 대해서는 ‘설계’ 자체를 규제하고 있다. 다만 향후 심의 과정에서 적용 범위는 유동적일 수 있다.

2. 온라인 활동에 대한 포괄적 규제 사례

AI 챗봇에 대한 직접적인 규제는 아니지만, 아동·청소년의 온라인 활동 전반을 연령에 따라 보호하려는 움직임도 활발하다.

호주는 개정 「온라인안전법(Online Safety Amendment Act 2024)」을 통해 2025년 12월부터 만 16세 미만의 SNS 계정 보유를 금지하고 있다. 플랫폼 사업자는 계정 생성을 차단하기 위한 ‘합리적 조치’를 취해야 하며, 미이행시 벌금이 부과된다(§63D). 최근에는 AI 챗봇에도 전면적 계정 금지를 적용하자는 논의도 이어지고 있으며,¹⁰⁾ 별도 규정(「연령제한자료 코드(Age Restricted Material Codes)」) 시행을 통해 AI 챗봇 등 디지털 서비스가 포르노그래피, 고강도 폭력물, 자해·자살 콘텐츠를 포함하는 경우 연령 확인 의무를 부과하고 있다.

영국은 서비스별로 연령 제한 차등화를 추진하고 있다. 2026년 6월, 「온라인안전법(Online Safety Act)」에 근거하여 만 16세 미만의 SNS 계정 보유·이용을 금지하는 정책 방향을 발표하였으며,¹¹⁾ 시행 목표는 2027년 봄이다. 사용자와 성적 관계·역할극을 시뮬레이션하도록 설계된 이른바 “로맨틱 동반자 AI 챗봇”의 최소 이용 연령을 18세로 규제하는 방안도 포함되었다. 다만 ‘서비스 종류’보다 ‘친밀한 관계를 형성하도록 설계된 기능’에 초점을 두고 있어, 유사한 기능의 AI 챗봇 전반으로 규제 대상이 확대될 가능성도 있다.¹²⁾ 한편 16~17세는 라이브 스트리밍, 낯선 사람과의 연락 등 일부 고위험 기능만 제한하여, 보호 연령 단절이 아닌 보호가 계속되는 연령으로 설정하는 방안을 추진 중이다.¹³⁾

노르웨이도 16세 미만의 SNS 이용 제한 법안을 연내 목표로 추진 중이다. 사업자에게 연령 확인 의무를 부과하며, 대상 서비스는 향후 법률에서 구체화할 계획이다. 다만 정부 지침에 따

라 올해 가을학기부터 학교에서는 초등학교(6~13세)의 생성형 AI 사용을 원칙적으로 금지하고, 14~16세는 교사 감독 하에 허용하는 등 연령별 이용기준을 마련하고 있다.¹⁴⁾ AI 기술 자체보다, 아동의 인지적 발달 단계에 따라 금지·감독·권장을 단계적으로 적용하는 예방적 정책 설계라는 점에서 의미가 있다.

프랑스는 「디지털경제 신뢰법(Loi pour la confiance dans l'économie numérique)」 개정을 통해 SNS 이용 제한 연령을 15세로 정하는 법안을 추진 중이다. AI 챗봇은 EU의 AI법상 ‘제한적 위험’ 등급으로서 최소한의 투명성 의무만 있으며, 아동 보호 규제는 입법 준비 단계에 있다.

4. 국내 법·제도의 한계

주요국과 달리 국내 AI 관련 법제는 기존의 정보통신 및 개인정보 규율 체계를 중심으로 규율하고 있고, 인공지능기본법 역시 산업 진흥과 신뢰 기반 조성을 중심으로 설계되어 있어, AI 챗봇이 지닌 새로운 위험 특성을 충분히 반영하지 못하고 있다.

첫째, **규제 패러다임이 사후 대응 중심에 머물러 있으며, 사업자의 예방적 책임이 미흡하다.** 현행 정보통신망법은 불법정보 등 유해 게시물에 대한 취급거부·삭제명령 등 사후조치를 규정하고 있고, 「청소년 보호법」도 유해매체를 지정·고시를 통한 접근 제한 방식을 취한다. 그러나 AI 챗봇의 답변은 이용자별로 다르게 실시간 생성되는 ‘폐쇄적 일대일 대화’ 형태를 띠는 점에서 ‘게시물’ 또는 ‘매체물’을 전제로 한 기존 규율 체계로는 충분히 대응하기 어렵다.

둘째, AI 챗봇의 **구조적 위험에 대응하기 위한 안전장치가 미흡하다.** AI 챗봇의 위험은 대화 과정에서 정서적 의존, 과몰입, 인식 왜곡 등을 유발할 수 있다는 데 있으나, 현행 법제는 이러한 설계 구조 자체를 규율하지 못하고 있다. 인공지능기본법이 생성형 AI 결과물 표시 의무와 고영향 AI에 대한 위험관리 의무를 규정하고 있으나 그 판단 기준이 신체·안전·기본권 등에 대한 직접적 영향 여부에 맞춰져 있어, 중독성 있는 알고리즘 설계와 같은 비가시적 위험은 규율 대상에 포섭되기 어렵다.

셋째, **미성년자 보호 체계에 공백이 있다.** 「개인정보 보호법」은 법정 대리인의 동의가 요구되는 나이를 만 14세 미만으로 한정하고 있어, 「청소년 보호법」상 19세 미만 청소년을 충분히 포섭하지 못하고 있다. 또한 단순한 생년월일 입력이나 일회성 인증코드 등은 미성년자가 쉽게 우회할 수 있는 방식으로, 효과적인 연령 확인 조치로 보기 어렵다.

9) 중대한 위반 행위에 대해서는 별도로 벌금 규정도 두고 있다(§107~109).

10) 케이트 체니(Kate Chaney) 하원의원, “Push to include AI chatbots in Australia's under-16 social media ban”, 2026.6.4.

11) 호주가 16세 미만의 SNS 계정 보유를 금지하고 있는 것과 달리, 영국은 현재까지는 「온라인안전법」 개정에 따라 사업자에게 연령 확인(age assurance) 의무만을 부과하고 있다.

12) 이와 같은 내용이 담긴 개정안은 2026년 말까지 의회에 제출할 예정이다. 영국 과학혁신기술부 보도자료, “Social media to be banned for under-16s in landmark government move to give kids their childhood back”, 2026.6.15.

13) 영국 과학혁신기술부, “Growing up in the online world: progress statement”, 2026.6.15.

14) Reuters, “Norway imposes near ban on AI in elementary school”, 2026.6.19.

넷째, 분산된 법체계가 효과적인 대응의 걸림돌이 되고 있다. AI 환경에서 아동·청소년 보호는 개인정보, 온라인 안전, 교육, 정신건강 등 여러 영역이 결합된 문제다. 그러나 현행 규율 체계는 개별 법률과 부처별(과기정통부, 교육부, 성평등가족부, 보건복지부, 개보위 등)로 분산되어 있어, 통합적 대응체계를 갖추지 못하고 있다.

5. 아동·청소년 보호를 위한 정책과제

AI 챗봇은 기존 온라인 서비스와 다른 구조적 위험을 지니는 만큼, 개별 법률이나 부처별 대응을 넘어 범정부 차원의 일관된 대응체계를 마련하여야 한다. 이를 위해 관계부처가 공동으로 AI 기반 아동·청소년 보호를 위한 통합적 대응체계를 구축하여, 부처별 역할 및 연계체계를 명확히 하여야 한다. 이와 함께 다음과 같은 정책과제의 조속한 추진도 필요하다.

1. 안전 설계 및 사업자 책임 강화

AI 챗봇 규제 패러다임은 유해 콘텐츠를 사후에 걸러내는 방식에서 벗어나, **의존이나 중독 유발 설계 요소 자체를 사전에 방지하고 사업자의 책임을 제도적으로 관리하는 방향으로 전환**되어야 한다. 이를 위해 관계부처 협력체계를 기반으로 사업자의 안전설계 의무와 위험관리 책임을 규정할 수 있는 법적 기반을 마련할 필요가 있다.

이를 위해 과기정통부는 중독·과몰입을 유발하는 AI 서비스의 설계 요소를 식별·관리할 기술적 기준과 가이드라인을 마련하여 사업자의 안전조치 기준으로 제시해야 한다. 정보통신망법 등의 개정에서 무한 스크롤·자동 재생 등 중독적 UX 설계 제한, 감정적 취약성을 이용한 조작적 참여 유도 방지, 야간·심야 푸시 알림 차단 등이 핵심 내용이 되어야 하며, 일률적 금지보다는 사업자 규모와 위험 수준에 따라 단계적·협력적으로 적용할 필요가 있다.

나아가 설계 기준과 투명성 의무의 이행 여부를 감사하고 위험 평가 결과를 공표하도록 의무화함으로써, 사업자의 책임성을 제도적으로 확보하여야 한다.

2. 아동·청소년 보호체계 강화

법적 보호 연령 기준을 정비하고, AI 환경에 대응할 수 있도록 유해매체물의 범위를 확대할 필요가 있다.

현행 「개인정보 보호법」은 만 14세 미만, 「청소년보호법」은 만 19세 미만을 보호 대상으로 규정하고 있어, 14세~18세 청소년에 대한 디지털 보호 체계가 단절되는 문제가 있다. AI 활용이 가장 활발한 연령층임에도 부모 동의 없이 개인정보 처리에 동의하게 되어, 데이터 수집·추적, 맞춤형 광고, 알고리즘 추천 등에 성인과 동일하게 노출되고 있다. 이에 보호 연령 기준을 조정하거나, 연령 확인 체계를 강화할 필요가 있다. 휴대폰·신용카드를 통한 보호자의 2차 승인, 안면인식 등의 생체기술 활용 등 **실효성 있는 연령 검증체계 도입**도 필요하다.

더불어 아동·청소년은 권위적인 문장에 쉽게 설득될 수 있으므로, AI 챗봇이 전문가나 인간 상담사로 오인되지 않도록 **주거적인 고지 의무를 부과**하고, 자해·자살·성적 착취 등을 유발하는 대화를 탐지하면 경고 문구를 띄우거나, **상담기관 또는 보호자에 연계되도록 안전장치를 마련**할 필요가 있다.

한편 현행 「청소년보호법」은 성인물, 사행성 콘텐츠, 술·담배 등 전통적인 유해매체물의 접근 제한에 초점을 두어, AI 추천 알고리즘, 중독성 설계, 청소년의 판단을 왜곡할 수 있는 생성형 AI 등 새로운 디지털 위험요인에 충분히 대응하지 못하고 있다. 따라서 청소년보호정책의 대상 역시 전통적 유해매체물 중심에서 **AI 기반 알고리즘 서비스와 중독적 설계 등 새로운 디지털 위험요인까지 확대**할 필요가 있다.

3. 일상적 보호자의 AI 리터러시 강화

AI 리터러시는 단순한 디지털 활용 교육이 아니라 AI의 작동 방식과 위험을 이해하고 적절히 대응할 수 있는 보호 역량으로 확대될 필요가 있다. 아동·청소년의 AI 챗봇 이용은 개인 특성뿐 아니라 가정·학교·또래 등 주변 환경의 영향을 크게 받는 만큼 **일상적 보호자의 역할이 중요하다**. 그러나 많은 부모와 교사는 AI 챗봇의 작동 원리나 의존·중독 유발 설계에 대한 이해가 부족하므로, **교육부와 성평등가족부, 보건복지부 등이 협력하여 보호자의 대응 역량을 국가 차원에서 체계적으로 지원**할 필요가 있다.

이를 위해 학부모·교사를 대상으로 한 AI 리터러시 교육, 가정 내 대화 가이드라인, 학교 현장용 지도자료 제공 등 정책적 지원이 병행되어야 할 것이다.

「이슈와 논점」은 국회의원의 입법활동을 지원하기 위해 최신 국내외 동향 및 현안에 대해 수시로 발간하는 보고서입니다.

