





# AI기본법 ‘표시의무’, 현장은 무엇을 고민하는가

스타트업얼라이언스 정주연 선임전문위원 <sup>01 02</sup>

01 법학 박사(행정법 전공)

02 이 글은 스타트업얼라이언스의 공식 입장이 아닌, 필자의 주관적인 견해를 밝힙니다.

AI 기본법의 '생성형 AI 표시의무'는 딥페이크와 허위정보 확산에 대응하고, 이용자의 알 권리를 보장함과 동시에 오인·기만 등 잠재적 위험으로부터 이용자를 보호하기 위해 도입되었다. 이러한 제도 취지에 대해서는 산업계 역시 대체로 공감하고 있으며, AI 서비스에 대한 사회적 신뢰를 높여야 한다는 방향성에도 큰 이견은 없다. 다만, 시행을 앞두고 산업 현장에서는 “무엇을 AI 생성물로 볼 것인가”, “표시는 실제 유통 과정에서 유지될 수 있는가”, “책임 주체는 누구인가”와 같은 실질적 쟁점이 제기되고 있다. 특히 AI가 단순 생성이 아니라 편집·보조·변환 형태로 폭넓게 활용되는 현실을 고려할 때, 이를 일률적으로 규율할 경우 고위험 영역 대응이라는 본래 목적에서 벗어나 이용자 보호의 실효성은 제한적인 반면, 사업자 부담만 확대될 수 있다는 우려도 존재한다. 또한 복합적인 서비스 구조와 기술적 한계로 인해 표시의무가 실질적 위험 예방보다는 형식적 준수에 머무를 가능성도 제기된다. 본 이슈페이퍼는 표시의무의 필요성을 전제로 하되, 산업 현장에서 제기되는 구체적 애로를 점검하고 적용 범위, 책임 구조, 기술적 실효성 측면에서 보다 정교한 제도 개선 방향을 모색하고자 한다.

---

## 목차 Table of Contents

1	<b>들어가며: 취지는 분명하다, 그러나 질문은 남는다</b>	04
2	<b>쟁점 1. ‘AI 생성물’의 범위는 어디까지인가</b>	07
3	<b>쟁점 2. 기술적으로 가능한가 - 표시의 실효성 문제</b>	11
4	<b>쟁점 3. 복합 서비스 구조에서 누가 책임을 지는가</b>	16
5	<b>실효성 있는 생성형 AI 표시의무를 위한 제언</b>	19
	5-1 위험도 기반 차등 표시 체계로 전환	
	5-2 ‘눈에 보이는 마크’ 중심 설계에서 기술 중립적 접근으로의 전환	
	5-3 B2B2C·B2B2G 구조에서의 책임 귀속 기준 명확화	
	5-4 중복 규제 정비와 단일 해석 창구 마련	
	5-5 글로벌 정합성 확보와 ‘갈라파고스 규제’ 방지	
6	<b>나가며: 취지를 살리기 위한 정교한 설계가 필요하다</b>	25

# 1

## 들어가며

취지는 분명하다,  
그러나 질문은 남는다

2026년 1월, 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하, “AI 기본법”)이 시행됐다. 법 시행 이후 최소 1년 간의 계도·유예 기간이 적용되면서 기업들이 제도에 적응할 수 있는 시간도 주어졌다. 그럼에도 불구하고 현장의 분위기를 살펴보면, 단순히 준비 기간의 문제를 넘어 제도 설계 자체에 대해 한 번 더 점검할 필요가 있다는 목소리가 이어지고 있다.

AI기본법의 핵심 내용 중 하나는 ‘생성형 AI 표시의무’다. AI가 생성·편집·가공한 콘텐츠임을 이용자에게 명확히 알리도록 함으로써 딥페이크나 허위정보의 확산을 예방하고, 이용자의 알 권리를 보장하겠다는 취지다. 생성형 AI 기술이 빠르게 확산되는 상황에서 이러한 제도적 장치를 마련할 필요가 있다는 점에 대해 산업계 역시 대체로 공감하고 있으며, 신뢰할 수 있는 AI 생태계를 조성해야 한다는 방향성에도 큰 이견은 없다.

다만, 제도가 실제 서비스 환경에서 어떻게 작동하는지에 대해서는 시행 전후를 거치며 보다 구체적인 질문들이 제기되고 있다. 시행 이전부터 산업계에서는 적용 범위와 책임 구조의 불명확성에 대한 우려가 제기된 바 있으며,<sup>03</sup> 이러한 문제의식은 시행 이후에도 해소되지 않은 채 이어지고 있다. 특히 생성형 AI 표시의무가 산업 현실에 비해 과도하게 작동할 수 있다는 우려는 실제 서비스 설계 단계에서 반복적으로 제기되고 있다. 이는 제도에 대한 단순한 이해 부족이라기보다, 법이 실제 서비스 구조와 기술 환경 속에서 얼마나 합리적으로 작동할 수 있는지에 대한 현장의 고민을 보여주는 대목이라 할 수 있다.

실제로 생성형 AI는 텍스트, 이미지, 영상, 음성 등 다양한 형태로 활용되며, 사람과 AI가 협업하는 방식도 빠르게 일반화되고 있다. 또한 AI 서비스는 단일 기업이 모든 과정을 담당하기보다는, 모델 개발사와 이를 활용해 서비스를 구현하는 사업자, 그리고 공공기관 등 다양한 주체가 연결된 복합 구조로 제공되는 경우가 적지 않다. 이러한 현실 속에서 ‘무엇을 AI 생성물로 볼 것인지’, ‘어떻게 표시하는 것이 실효적인지’, ‘누가 책임을 부담해야 하는지’와 같은 질문이 자연스럽게 제기되고 있다.

03 스타트업얼라이언스, AI 기본법과 스타트업: AI 스타트업이 겪는 현실, 스타트업 생태계 동향 리포트, 2025-6, 2025.

이에 본 이슈페이퍼는 생성형 AI 표시의무를 둘러싼 주요 쟁점을 ▲‘AI 생성물’의 범위와 기준, ▲표시의 기술적 실효성과 제도 설계, ▲복합적인 서비스 구조에서의 책임 귀속 문제 등을 중심으로 살펴보고자 한다. 이를 통해 이용자 보호라는 제도의 취지를 훼손하지 않으면서도 산업 현실과 조화를 이루는 보다 합리적인 정책 방향을 모색해보고자 한다.

# 2

## 쟁점 1.

‘AI 생성물’의 범위는 어디까지인가

생성형 AI 표시의무는 “AI가 생성·편집·가공한 콘텐츠”를 이용자가 인지할 수 있도록 하는 것을 기본 전제로 한다. 그러나 실제 서비스 환경에서는 사람이 전 과정을 수행하거나 AI가 전적으로 결과물을 만들어내는 경우보다, 사람과 AI가 협업하는 방식이 훨씬 일반적이다. 이로 인해 현장에서는 “어디까지를 AI 생성물로 보아야 하는가”라는 근본적인 질문이 지속적으로 제기되고 있다.

예를 들어, AI가 추천 기능이나 문장 교정, 요약 등 보조적 역할로 활용되는 경우, 사람이 전체 기획과 구성을 담당하고 AI가 일부 시각 요소나 표현만 보완한 경우, 전문가가 작성한 문서를 AI가 정리하거나 다듬는 경우 등이 모두 표시 대상이 포함되는지에 대해 해석의 여지가 존재한다. 특히 이러한 보조적 활용까지 일률적으로 ‘AI 생성물’로 분류하는 것이 적절한지에 대해서는 현장에서 신중한 검토가 필요하다는 의견이 제기되고 있다.

또 다른 쟁점은 딥페이크와 같은 고위험 활용과 일반적인 AI 활용을 동일한 범주로 다루는 것이 타당한가에 대한 문제다. 기만 목적이 없는 유희적 서비스나 공익적 목적의 AI 활용까지 광범위하게 표시의무 대상에 포함될 경우, 규제 초점이 흐려질 수 있다는 우려가 제기된다. 예컨대 카메라 필터 애플리케이션이나 돌봄 목적의 AI 음성 서비스, 콘텐츠 제작 과정에서의 보조적 AI 활용 등에 대해 일률적인 표시를 요구하는 것이 실제 이용자 보호에 얼마나 기여하는지에 대해서는 추가적인 검토가 필요하다는 시각이 존재한다.

특히 현장에서 혼란이 큰 영역 중 하나는 STT(음성-텍스트 변환), TTS(텍스트-음성 변환), 기계 번역, OCR(광학문자인식) 등 이른바, ‘변환형 기술’이다. 이러한 기능은 기술적으로 LLM이나 딥러닝을 활용하더라도, 특정 인물이나 창작물을 모방하여 새로운 내용을 만들어내는 행위라기보다는 기존 정보를 다른 형식으로 전환하거나 인식·표현하는 데에 가까운 특성을 지닌다. 예를 들어 시각장애인을 위해 책을 읽어주는 TTS 기능이나 영상 속 음성을 자막으로 바꾸는 STT 기능은 정보를 보다 쉽게 전달하기 위한 변환 기술에 해당한다. 이는 특정 인물의 목소리를 합성하여 실제로 하지 않은 발언을 만들어내는 경우와는 위험 수준과 사회적 파급력이 본질적으로 다르다는 점에서, 동일한 규율 체계로 다루는 것이 적절한지에 대한 문제 제기가 이루어지고 있다.

추천 시스템 역시 유사한 맥락에서 살펴볼 수 있다. 온라인 쇼핑몰이 이용자의 구매 이력에 따라 상품을 추천하거나, OTT 서비스가 취향에 맞는 영상을 제안하는 기능은 새로운 콘텐츠를 생성하는 것이 아니라 기존 콘텐츠의 배열과 노출 순서를 조정하는 것에 가깝다. 이러한 기능까지 생성형 AI의 결과물과 동일하게 보아 표시의무를 부과하는 것은 이용자 보호 측면의 실익은 크지 않은 반면, 규제 범위를 과도하게 확장할 수 있다는 지적이 있다.

국제표준(ISO) 논의에서도 생성형 AI의 범위를 '새로운 콘텐츠를 생성하는 경우'(generate new contents)로 한정하는 흐름이 확인된다. 이러한 접근은 '생성'과 '변환·배열'을 구분함으로써 규제의 적용 범위를 보다 명확히 하고 실제 오인 가능성이나 기만 위험이 높은 영역에 정책 역량을 집중하자는 취지로 이해된다. 우리 제도 역시 기능의 성격과 사회적 위험도를 종합적으로 고려하여 적용 범위를 정교하게 설정할 필요가 있을 것으로 보인다.

한편, 시행령 제23조 제4항은 일정한 경우 표시의무의 전부 또는 일부를 적용하지 않을 수 있도록 예외를 두고 있다. 예를 들어 서비스명이나 이용자 화면 등을 통해 생성형 AI 활용 사실이 명백한 경우, 내부 업무용에 한정되는 경우, 또는 장관 고시를 통해 추가 예외가 인정되는 경우 등이 이에 해당한다. 그러나 현장에서는 이러한 예외 규정이 실제로 어떻게 적용되고 있는지에 대한 구체적인 사례나 해석 기준을 찾기 어렵다는 의견도 제기된다. 예외가 형식적으로 존재하더라도, 적용범위가 명확하지 않다면 기업은 이를 적극적으로 원용하기보다는 보수적 해석을 택할 수밖에 없다.

그 결과, 위해성과 직접적 관련이 없는 영역까지 표시가 확대되는 이른바, '과잉 표시'로 이어질 가능성도 배제하기 어렵다. 이는 예외 규정이 충분히 작동하지 못할 경우 나타날 수 있는 구조적 문제라는 점에서 'AI 활용 사실이 명백한 경우'의 판단 기준이나 적용 방식에 대한 보다 구체적인 해석이 필요하다는 지적이 이어지고 있다.

결국, 문제의 핵심은 표시의무의 필요성 자체가 아니라, 그 적용 범위를 어떻게 합리적으로 설정할 것인가에 있다. AI의 활용 정도, 기능의 성격(생성·변환·배열), 이용자 오인 가능성, 기만 목적의 존재 여부, 사회적 위험 수준 등을 종합적으로 고려한 보다 명확하고 정교한 기준이 마련될 필요가 있다. 이러한 기준은 제도의 예측 가능성을 높이고 불필요한 규제 부담을 줄이는 방향으로 작동할 수 있을 것이다.

## 💡 여기서 잠깐!

### 핵심 질문 'AI 생성물'의 범위는 어디까지인가

<p><b>Q1.</b> <b>무엇을 'AI 생성물'로 볼 것인가?</b></p> <ul style="list-style-type: none"><li>- 사람이 초안을 작성하고, AI가 일부 문장을 수정·보완한 경우에도 "AI 생성물"로 보아야 하는가?</li><li>- 사람이 전체 기획과 구성을 담당하고, AI가 배경이나 색채 등 일부 요소만 보조한 경우, 결과물 전체에 표시가 필요한가?</li><li>- 전문가가 작성한 문서를 AI가 교정·요약한 경우에도 표시의무가 적용되어야 하는가?</li></ul>	<p><b>Q2.</b> <b>보조적 활용과 핵심적 활용을 어떻게 구분할 것인가?</b></p> <ul style="list-style-type: none"><li>- 생성형 AI가 서비스의 핵심 기능이 아닌 추천·보조 기능으로 사용된 경우에도 동일한 규제가 적용되어야 하는가?</li><li>- 기존의 CG 기술과 본질적으로 큰 차이가 없는 수준의 AI 활용까지 규제 대상에 포함하는 것이 타당한가?</li></ul>
<p><b>Q3.</b> <b>'딥페이크'와 '예술적·창의적 표현물'의 경계는 무엇인가?</b></p> <ul style="list-style-type: none"><li>- 기만 목적이 없는 유희적·공익적 서비스까지 규제 대상으로 분류되어야 하는가?</li><li>- 영화, 광고, 게임, 그래픽 등 창작 분야에서 '예술적·창의적 표현물'의 예외 범위는 어디까지로 설정해야 하는가?</li></ul>	<p><b>Q4.</b> <b>'생성'과 '변환·배열'을 어떻게 구분할 것인가?</b></p> <ul style="list-style-type: none"><li>- STT·TTS·번역·OCR 등 데이터의 형식을 변환하는 기능을 '생성형 AI'로 일률적으로 분류하는 것이 타당한가?</li><li>- 추천시스템과 같이 기존 콘텐츠의 순서를 결정하는 기능이 표시의무 대상에 포함되는가?</li></ul>

# 3

## 쟁점 2.

기술적으로 가능한가 - 표시의 실효성 문제

생성형 AI 표시의무는 이용자가 콘텐츠의 생성 방식을 인지할 수 있도록 하는 데 목적이 있다. 이러한 취지는 이용자의 알 권리를 보장하고, 딥페이크나 허위정보로 인한 피해를 예방한다는 점에서 충분한 공감대를 형성하고 있다. 다만, 실제 서비스 환경에서 이러한 표시가 기술적으로 안정적으로 유지될 수 있는지, 그리고 이용자 보호라는 목적에 실질적으로 기여할 수 있는지에 대해서는 보다 신중한 검토가 필요하다는 의견이 제기되고 있다.

가장 먼저 언급되는 부분은 ‘**텍스트 생성물에 대한 표시의 어려움**’이다. 이미지나 영상은 화면에 워터마크를 삽입하는 방식이 가능하지만, 텍스트는 복사·붙여넣기나 재편집 과정에서 표시 정보가 쉽게 사라질 수 있기 때문이다. 예를 들어, 플랫폼에서 ‘AI 생성’이라는 안내 문구를 함께 제공하더라도, 이용자가 해당 내용을 다른 웹사이트나 SNS에 옮겨 게시하는 순간 표시가 유지되지 않을 가능성이 크다. 이미지와 영상 역시, 캡처, 크롭, 재편집, 재업로드 등의 과정을 거치면 워터마크가 제거될 수 있어 표시의 지속성을 완전히 보장하기 어렵다는 지적이 있다.

이와 함께 **표시의무가 실제 이용자 보호로 이어지는지에 대한 점검의 필요성**도 함께 언급되고 있다. 이용자가 모든 AI 관여 콘텐츠에 반복적으로 표시를 접하게 될 경우 이를 일종의 배경 정보로 인식하는 ‘경고 피로’(warning fatigue) 현상이 나타날 수 있으며, 그 결과 정작 위험성이 높은 딥페이크나 허위정보에 대한 경고의 주목도가 낮아질 가능성도 있다. 이는 표시의무의 필요성을 부정하기보다는, 위험도에 따른 차등적 접근이 제도의 실효성을 높이는 데 도움이 될 수 있음을 시사한다.

한편, 글로벌 차원에서는 C2PA(Coalition for Content Provenance and Authenticity)와 같은 메타데이터 기반의 콘텐츠 출처 증명 기술이 점차 확산되고 있다. 우리 법 역시 시행령 제23조 제2항을 통해 ‘기계가 판독할 수 있는 방법’에 의한 표시를 허용하고 있다. 다만, 이용자가 명확히 인식할 수 있는 표시를 함께 요구하고 있어 **메타데이터 기반 방식만으로 표시의무를 충족할 수 있는지에 대해서는 해석상 불확실성이 존재한다는 의견이 있다.** 이러한 불확실성은 기업들이 법적 리스크를 최소화하기 위해 시각적 워터마크나 안내 문구 등을 병행하는 보수적인 방식을 선택하게 만들 수 있으며, 결과적으로 다양한 기술적 표시 방식의 활용이 위축될 가능성도 조심스럽게 제기된다.

그러나 특정 글로벌 표준의 채택을 제도적으로 요구하거나 권장하는 접근 또한 스타트업에게 추가적인 비용 부담과 기술적 진입 장벽으로 작용할 수 있다는 점에서 신중한 검토가 필요하다. 이에 따라 기술적 중립성을 유지하면서 **업계의 자율적인 표시 방식을 폭넓게 인정하는 방향이 보다 바람직하다**는 시각도 존재한다.

### [표시의무] AI기본법 제31조 제2항 및 시행령 제23조 제2항

AI기본법 제31조(인공지능 투명성 확보 의무) ② 인공지능사업자는 생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 경우 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 한다.

AI 기본법 시행령 제23조(인공지능 투명성 확보 의무) ② 인공지능사업자가 법 제31조제2항에 따른 표시를 할 때에는 다음 각 호의 어느 하나에 해당하는 방법으로 할 수 있다.

1. 사람이 인식할 수 있는 방법
2. 기계가 판독할 수 있는 방법. 이 경우 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 1회 이상 안내 문구·음성 등으로 제공해야 한다.

이와 함께 서비스의 인터페이스 자체가 이미 AI 작동 사실을 이용자에게 충분히 전달하고 있는 경우를 어떻게 평가할 것인지에 대한 고민도 제기되고 있다. 예컨대 서비스명에 ‘AI’가 명시되어 있고, 채팅창 형태의 UI를 통해 대화가 이루어지는 생성형 AI 서비스의 경우 일반 이용자 역시 해당 서비스가 AI에 의해 작동한다는 점을 어렵지 않게 인지할 수 있다. 또한 LLM 기반 서비스라 하더라도 특정 목적에 맞게 기능이 제한되고 필터링을 통해 정해진 범위 내에서만 응답을 제공하는 경우에는, 이용자 관점에서 기존의 룰 기반 챗봇과 실질적인 체감 차이가 크지 않을 수 있다는 점도 고려요소로 언급된다.

이러한 맥락에서 단순히 ‘생성형 AI를 사용했다’는 형식적 기준만을 근거로 일률적인 표시의무를 부과하기보다는, 서비스의 이용 맥락과 인터페이스 설계가 이용자에게 제공하는 인지 가능성을 함께 고려하는 유연한 해석이 필요하다는 의견이 제기된다. 이는 불필요한 중복 표시를 줄이면서도 제도의 실효성을 높이고, 다양한 기술적 접근 방식을 포용하는 데 기여할 수 있을 것으로 보인다.

결국 표시의무가 형식적 준수에 그치지 않고 실질적인 이용자 보호로 이어지기 위해서는 기술적 지속 가능성, 표시 방식의 구체적 기준 등을 보다 명확히 정리할 필요가 있다. 이러한 기준이 보완된다면 제도의 예측 가능성이 높아지고, 산업 현장의 부담을 줄이면서도 표시의무의 본래 취지를 보다 효과적으로 구현할 수 있을 것으로 보인다.

## 💡 여기서 잠깐!

### 핵심 질문 표시의무는 기술적으로 실효적인가?

Q1.

#### 표시의 지속성을 기술적으로 보장할 수 있는가?

- 텍스트 생성물의 경우, 복사·붙여넣기나 재게시 과정에서 표시가 유지될 수 있는가?
- 이미지·영상의 워터마크가 캡처·편집·재업로드 과정에서 쉽게 제거될 수 있는 상황에서, 표시의 실효성을 어떻게 확보할 것인가?

Q2.

#### AI 사업자의 표시의무 이행 범위는 어디까지인가?

- AI 사업자가 콘텐츠를 최초로 제공하는 단계에서 표시의무를 충실히 이행했다면, 이후 이용자의 복사·재게시 등 AI 사업자의 통제범위를 벗어난 활용까지 추가적인 조치를 요구하는 것이 타당한가?
- 표시의무의 이행 여부는 AI 사업자가 실질적으로 통제할 수 있는 '최초 제공 단계'를 기준으로 판단하는 것이 합리적인가, 그리고 그 구체적인 판단 기준은 어떻게 설정되어야 하는가?

Q3.

#### 시각적 표시와 기술적 표시의 관계는 어떻게 설정할 것인가?

- 메타데이터 기반의 기계 판독 방식(C2PA 등)만으로도 법적 의무를 충족했다고 볼 수 있는가?
- 이용자의 인지 가능성을 요구하는 현행 규정이 다양한 기술적 표시 방식의 활용을 제한하고 있는 않은가?
- 특정 글로벌 표준을 제시하기보다는 기술 중립성을 유지하면서 업계의 자율적인 표시 방식을 폭넓게 인정하는 방향이 보다 바람직하지 않은가?

Q4.

#### 인터페이스가 이미 AI 사실을 전달하는 경우는 어떻게 다를 것인가?

- 서비스명·UI·이용 맥락상 AI 작동이 자명한 경우에도 별도의 시각적 고지가 추가로 필요한가?
- 필터를 통해 응답 범위가 제한된 LLM 기반 서비스를 전통적 챗봇과 달리 규율할 합리적 근거가 있는가?

# 4

## 쟁점 3.

복합 서비스 구조에서  
누가 책임을 지는가

오늘날 AI 서비스는 단일 기업이 개발부터 제공까지 전 과정을 담당하는 구조보다는, 여러 주체가 역할을 나누는 다층적 형태로 운영되는 경우가 일반적이다. AI 모델을 개발하는 기업, 이를 API나 솔루션 형태로 제공하는 기업, 이를 활용해 최종 서비스를 구현하는 기업, 그리고 이를 시민이나 이용자에게 제공하는 공공기관이나 파트너 기업까지 이어지는 B2B2C 및 B2B2G 구조가 빠르게 확산되고 있다. 이러한 환경에서는 **생성형 AI 표시의무의 책임 주체를 누구로 보아야 하는지**에 대한 실무적 혼란이 존재한다는 의견이 제기되고 있다.

AI기본법상 표시의무의 주체는 ‘AI 사업자’로 규정되어 있으나, 복합적인 서비스 구조에서는 AI 모델 개발사와 최종적으로 서비스를 구현하여 이용자와 직접 접점을 형성하는 사업자가 서로 다른 경우가 적지 않다. 이 때 표시의무의 1차적 책임을 어느 단계의 AI사업자에게 부과하는 것이 합리적인지에 대한 기준이 보다 명확해질 필요가 있다는 지적이 있다. 특히, 공공서비스 영역에서는 민간 기업이 기술을 제공하고 지자체나 공공기관이 이를 시민에게 제공하는 구조가 일반적이는데, 이 경우 표시 방식과 책임 귀속에 대한 해석 차이가 발생할 수 있다. 이러한 불확실성은 계약 단계에서의 부담 증가뿐 아니라, 서비스 설계와 운영 전반에 영향을 미칠 수 있다.

글로벌 서비스를 제공하는 기업의 경우에는 추가적인 혼란이 발생할 수 있다. **한국 이 이용자에게 제공되는 서비스에 대해 어떠한 방식으로 표시를 이행해야 하는지, 서비스 구조에 따라 구분 적용이 가능한지 등에 대한 구체적인 가이드라인이 충분히 제시되지 않았다는 의견이 있다.** 이처럼 적용 기준이 명확하지 않을 경우 기업은 규제 리스크를 최소화하기 위해 가장 보수적인 방식을 선택하게 되며, 이는 특정 기능의 출시 지연이나 국내 서비스 범위의 제한으로 이어질 가능성도 배제하기 어렵다.

또한 책임 주체가 명확히 정리되지 않은 상태에서는 **기술 제공자와 최종 서비스 제공자 간에 규제 부담이 중첩되거나, 반대로 책임의 공백이 발생할 가능성에 대한 우려도 존재한다.** 이는 제도의 실효성을 저하시킬 수 있을 뿐 아니라, 규제 대응 여력이 제한적인 스타트업에게는 추가적인 불확실성으로 작용할 수 있다.

따라서 표시의무의 실효성을 높이기 위해서는, 최종 이용자와 직접적인 접점을 형성하여 결과물을 제공하는 사업자를 중심으로 하되, 모델 개발사 등 상위 단계 사업자의 역할과 영향력도 함께 고려하는 유연한 책임 구조가 검토될 필요가 있다. 아울러,

B2B2C·B2B2G 환경에 적용 가능한 가이드라인과 글로벌 서비스에 대한 적용 기준 등이 마련된다면, 제도의 예측 가능성을 높이고 산업 현장의 혼란을 줄이는 데 도움이 될 것으로 보인다.

## 💡 여기서 잠깐!

### 핵심 질문 복합 서비스 구조에서 누가 책임을 지는가

<p><b>Q1.</b> <b>표시의무의 1차적 책임 주체는 누구인가?</b></p> <ul style="list-style-type: none"><li>- AI 모델 개발사와 AI 기술을 활용하여 최종적으로 서비스를 제공하는 사업자 중 누구에게 표시의무를 부과하는 것이 합리적인가?</li><li>- 특히 최종 이용자와 직접적인 접점을 형성하며 결과물을 제공하는 AI사업자를 기준으로 책임을 설정하는 것이 타당한가?</li></ul>	<p><b>Q2.</b> <b>B2B2C·B2B2G 구조에서는 책임을 어떻게 분담할 것인가?</b></p> <ul style="list-style-type: none"><li>- 민간 기업이 개발한 AI 서비스를 지자체나 공공기관이 시민에게 제공하는 경우 표시의무의 귀속 주체는 누구인가?</li><li>- 공공 목적의 AI 서비스에 대해 별도의 기준을 마련할 필요가 있는가?</li></ul>	<p><b>Q3.</b> <b>글로벌 서비스에는 표시의무를 어떻게 적용할 것인가?</b></p> <ul style="list-style-type: none"><li>- 한국 이용자에게만 별도의 표시를 적용해야 하는가?</li><li>- IP 기반으로 이용자를 구분하는 방식이 현실적으로 가능한가?</li><li>- 국가별 규제 차이를 고려한 합리적인 적용 기준은 무엇인가?</li></ul>
---	--	--

# 5

## 실효성 있는 생성형 AI 표시의무를 위한 제언

생성형 AI 표시의무는 이용자의 알 권리를 보장하고 딥페이크나 허위정보로 인한 사회적 위험을 예방하기 위한 제도적 장치이다. 이러한 취지를 실질적으로 구현하기 위해서는 산업 현실과 기술 발전 속도를 반영한 정교한 정책 설계가 필요하다. 다음과 같은 방향을 중심으로 제도 개선을 검토해볼 수 있을 것이다.

### 5-1 위험도 기반 차등 표시 체계로 전환

현재의 논의는 AI의 관여 여부를 기준으로 일괄적인 표시를 요구하는 구조에 가까운 측면이 있다. 그러나 실제 사회적 위험은 모든 AI 활용에서 동일하게 발생하지 않는다. 예를 들어, ① 특정인의 인격권·초상권을 침해할 수 있는 딥페이크 영상, ② 선거·금융 등 공공의 신뢰에 영향을 미칠 수 있는 허위정보 생성, ③ 사기·기만 목적이 명확한 콘텐츠 등은 상대적으로 높은 위험성을 지닌다. 이러한 영역에는 이용자가 직관적으로 인식할 수 있는 명확하고 강력한 표시를 의무화하는 것이 타당할 수 있다.

반면, ① 문법 교정·요약 기능, ② 배경 생성 등 보조적 그래픽 처리, ③ 추천 알고리즘 기반의 콘텐츠 배열, ④ 기만 목적이 없는 유희·공익 서비스 등 저위험 AI 활용까지 동일한 강도로 규율하는 것은 과도하게 작용할 가능성이 있다. EU AI Act가 채택한 리스크 기반 접근과 같이, 규제 강도를 ‘AI 사용 여부’가 아니라 ‘실질적 피해 가능성’에 비례하여 설계하는 방식이 보다 합리적인 대안이 될 수 있다. 모든 영역에 동일한 표시를 요구하기보다는, 고위험 영역에 규제 역량을 집중하는 방향으로의 전환이 필요하다.

구체적으로는 다음과 같은 영역에 차등 적용 또는 예외 인정 방안을 검토해볼 수 있다. 첫째, **결과물의 성격상 위험이 낮은 경우**로, 단어·한 줄 문구·제목·키워드 등 짧은 생성물과 단순 요약·번역·오타 교정 등 기능적 보조 도구, 밝기·색감·해상도 조정이나 노이즈 제거와 같은 단순 보정·후처리 기능 등이 이에 해당한다. 둘째, **사업 및 거래 편의 지원을 목적으로 하는 경우**로, 소상공인·중소기업의 마케팅 문구 보조, 상품 소개문 초안 작성, 고객 응대 문안 보조, 쇼핑 탐색·비교·추천 보조 기능 등이 포함된다. 셋째, **공익 및 취약계층 지원을 목적으로 하는 경우**로, AI 돌봄콜, 배리어프리

(Barrier-free) 지원, 건강정보 탐색 보조, 병원·약국 예약 연계, 날씨·재난 안전 안내 등 정보 접근성 제고가 주된 목적인 서비스가 이에 속한다. 넷째, **개인적·비상업적·오락적 활용 영역**으로, 카메라 필터 효과나 간단한 스타일 변환 등 사회적 위해 가능성이 낮고 이용자의 의사결정에 미치는 영향이 제한적인 기능이 해당한다.

이와 같은 영역 구분은 규제 자원을 실제 고위험 분야에 집중할 수 있도록 하는 한편, 산업 현장에 구체적인 예측 가능성을 제공하는 기반이 될 수 있다. 또한 불필요한 규제 부담을 완화하면서도 딥페이크, 허위 정보, 기만 콘텐츠와 같은 핵심 위험에 대한 대응력을 높인다는 점에서, 투명성 의무의 입법 취지를 보다 효과적으로 구현하는 방향이 될 것이다.

## 5-2 ‘눈에 보이는 마크’ 중심 설계에서 기술 중립적 접근으로의 전환

현행 논의는 주로 시각적 표시(워터마크, 팝업 등) 방식에 초점을 맞추고 있다. 그러나 이러한 방식은 콘텐츠가 외부로 반출되거나 편집·재업로드되는 과정에서 쉽게 제거될 수 있어 표시의 지속성을 보장하기 어렵다는 한계가 있다.

한편, 글로벌 차원에서는 C2PA 등과 같은 메타데이터 기반의 콘텐츠 출처 증명 체계가 확산되고 있으며, 이는 콘텐츠의 생성·편집 이력을 기술적으로 추적할 수 있도록 설계된 국제적 논의의 일환으로 이해되고 있다. 우리 AI기본법 시행령 제23조 제2항 역시 ‘기계가 판독할 수 있는 방법’에 의한 표시를 허용하고 있어, 비가시적 워터마킹 등 다양한 기술적 표시 방식의 활용 가능성을 제도적으로 열어두고 있다.

다만, 특정 국제 표준의 채택을 제도적으로 요구하거나 권장하는 접근은 스타트업과 중소기업에게 추가적인 비용 부담과 기술적 진입 장벽으로 작용할 수 있으며, 결과적으로는 일부 대형 사업자에게 유리한 환경을 조성할 가능성도 있다는 점에서 신중한 검토가 필요하다. 또한 향후 제도 개선 과정에서 시각적 표시와 기술적 표시를 모두 의무화하는 방향으로 규제가 강화될 경우, 산업계 전반의 부담이 증가할 수 있다는 우려도 제기된다.

이러한 점을 고려할 때, 표시 방식은 **특정 기술이나 표준을 중심으로 일률적으로 설계하기보다는 기술 중립성 원칙을 바탕으로 다양한 방식의 표시를 폭넓게 인정하는 접근이 보다 바람직할 수 있다.** 특히, 스타트업의 경우 단순한 시각적 마크 부착이 규제 준수를 보다 용이하게 하는 현실적인 수단이 될 수 있으므로, 업계가 서비스의 특성과 기술적 여건에 맞추어 자율적으로 표시 방식을 선택할 수 있도록 하는 유연한 제도 설계가 필요하다.

결국, 표시의무의 실효성을 높이기 위해서는 비가시적 워터마킹과 같은 기술적 방식의 활용 가능성을 열어두되, 이를 의무화하기보다는 선택 가능한 수단으로 인정하고, 업계의 자율적인 표시 방식을 폭넓게 수용하는 방향으로 정책적 균형을 모색할 필요가 있다. 이러한 접근은 이용자 보호라는 제도의 취지를 유지하면서도 산업 현장의 부담을 최소화하고, 제도의 예측가능성과 수용성을 높이는 데 기여할 것으로 기대된다.

### 5-3 B2B2C·B2B2G 구조에서의 책임 귀속 기준 명확화

오늘날 AI 서비스는 AI 모델 개발사, 플랫폼사업자, 솔루션 제공 기업, 그리고 지자체·공공기관 등 다양한 주체가 참여하는 다층적 구조로 운영되는 경우가 많다. 이러한 환경에서는 표시의무의 책임주체가 누구인지에 대한 실무적 혼란이 발생할 수 있다.

예를 들어, AI 모델 개발사와 최종 서비스 제공자가 다른 경우, 또는 민간 기업이 기술을 제공하고 공공기관이 이를 시민에게 제공하는 경우, 표시의무의 귀속 주체와 책임 범위를 법령 또는 시행령에서 명확히 정리할 필요가 있다. 공공 목적의 AI 서비스에 대해서는 별도의 기준을 마련하는 방안도 검토할 수 있다. 책임이 불명확할수록 기업은 규제 리스크를 최소화하기 위해 가장 보수적인 설계를 선택하게 되며, 이는 서비스 기획 단계에서의 위축으로 이어질 가능성이 있다.

### 5-4 중복 규제 정비와 단일 해석 창구 마련

현재 기업들은 시기본법 외에도 정보통신망법 개정 논의, 부처별 가이드라인 등 복수의 규제를 동시에 고려해야 하는 상황에 놓여있다. 문제는 규제의 엄격성 자체보다도 기준의 중첩과 해석의 불일치에서 비롯되는 불확실성이다.

① 어느 규정을 우선 적용해야 하는지, ② 부처별 가이드라인이 충돌할 경우 어떻게 판단해야 하는지, ③ 시행령과 행정해석의 범위는 어디까지인지에 대한 명확한 기준이 부족한 상황이다. 이러한 불확실성은 기업에게 상당한 부담으로 작용할 수 있다. 따라서 규제 간 정합성을 정비하고, 기업이 참고할 수 있는 ‘단일 해석 창구(One-Stop 체계)’를 마련하는 것이 필요하다. 명확한 가이드라인과 공식·질의 회신 체계가 구축된다면 법적 예측 가능성을 높이는 데 기여할 수 있을 것이다.

## 5-5 글로벌 정합성 확보와 ‘갈라파고스 규제’ 방지

AI 산업은 본질적으로 글로벌 시장을 전제로 성장한다. 국내에만 과도하거나 독자적인 규제가 적용될 경우, ① 국내 기업이 해외 출시를 우선 고려하게 되거나, ② 국내 투자 위축, ③ 글로벌 자본의 한국 시장 회피로 이어질 가능성도 배제할 수 없다.

EU AI Act, 미국 투명성 관련 입법 등 주요 국가의 규제 동향과의 정합성을 사전에 충분히 검토할 필요가 있다. 국내 기업이 국제 기준과 다른 별도의 체계를 이중으로 준수해야 하는 상황은 곧 경쟁력 약화로 이어질 수 있기 때문이다. 입법 및 시행 과정에서 글로벌 사례에 대한 재검토와 함께 산업계의 의견을 충분히 수렴하는 절차가 중요하다.

# 6

## 나가며

취지를 살리기 위한  
정교한 설계가 필요하다

AI기본법의 생성형 AI 표시의무는 기술 발전에 따른 새로운 사회적 위험에 대응하기 위해 도입된 제도이다. 이용자의 알 권리를 보장하고 딥페이크와 허위정보로부터 사회를 보호하겠다는 취지 자체는 충분히 공감할 수 있다.

그러나 제도는 의도만으로 평가되기보다는 실제 현장에서 어떻게 작동하는지와 그 목적을 얼마나 효과적으로 달성하는지에 따라 평가될 필요가 있다. 산업 현장에서 제기되는 우려는 규제를 축소하거나 배제하자는 주장이라기보다, 제도의 취지를 살리기 위해 보다 정교한 설계가 필요하다는 제안에 가깝다.

위험도에 비례하지 않는 일괄적 표시 요구, 기술적으로 지속되기 어려운 형식적 워터마크 중심의 설계, 다층적 서비스 구조에서의 책임 불명확성, 중복 규제로 인한 해석상의 혼란, 그리고 글로벌 기준과의 정합성 부족 등은 표시 의무가 이용자 보호 장치로 기능하기보다는 산업 부담 요인으로 인식될 가능성을 높일 수 있다. 이러한 상황이 지속될 경우 제도에 대한 사회적 수용성 또한 약화될 수 있다.

특히, **스타트업에게 ‘시간’과 ‘예측 가능성’은 생존과 직결되는 요소**이다. 규제의 존재 자체보다도 적용 범위를 예측하기 어려운 불확실성이 더 큰 리스크로 작용할 수 있다. 이러한 불확실성은 혁신을 ‘도전’의 영역에서 ‘회피’의 영역으로 전환시키는 요인이 될 수 있다.

AI 산업은 아직 성장 과정에 있으며, 기술과 서비스 모델은 빠르게 변화하고 있다. 이러한 환경에서 규제는 완성된 답안이라기보다 지속적으로 보완되고 조정되는 설계도에 가까울 필요가 있다. 이용자 보호와 산업 경쟁력은 상충되는 목표가 아니라, 정교한 제도 설계를 통해 함께 달성해야 할 과제이다.

따라서 지금 필요한 것은 규제의 확대나 축소라는 단순한 선택이 아니라, 목적에 부합하는 작동 방식으로서의 재설계이다. 규제는 수단으로서 기능해야 하며, 그 수단이 산업의 성장 가능성을 제약하는 장벽이 되지 않도록 구조를 다시 점검할 시점이다. 이러한 균형 잡힌 접근이 이루어질 때, 생성형 AI 표시의무는 사회적 신뢰를 축적하는 동시에 산업의 지속가능한 성장을 뒷받침하는 제도로 자리매김할 수 있을 것이다.





**STARTUP  
ALLIANCE**

발행인 스타트업얼라이언스

기획/제작 스타트업얼라이언스

발행처 스타트업얼라이언스

서울특별시 마포구 마포대로 122,

4층(프론트원)

<https://startupall.kr>

발행일 2026.4.28

ISSN 2982-4834

이 책의 저작권은  
스타트업얼라이언스에 있으며  
무단복제와 전재를 금합니다.



