인사이트

AI 답변에 우리 회사가 안 나온다면

크롤러는 화면을 보지 않습니다

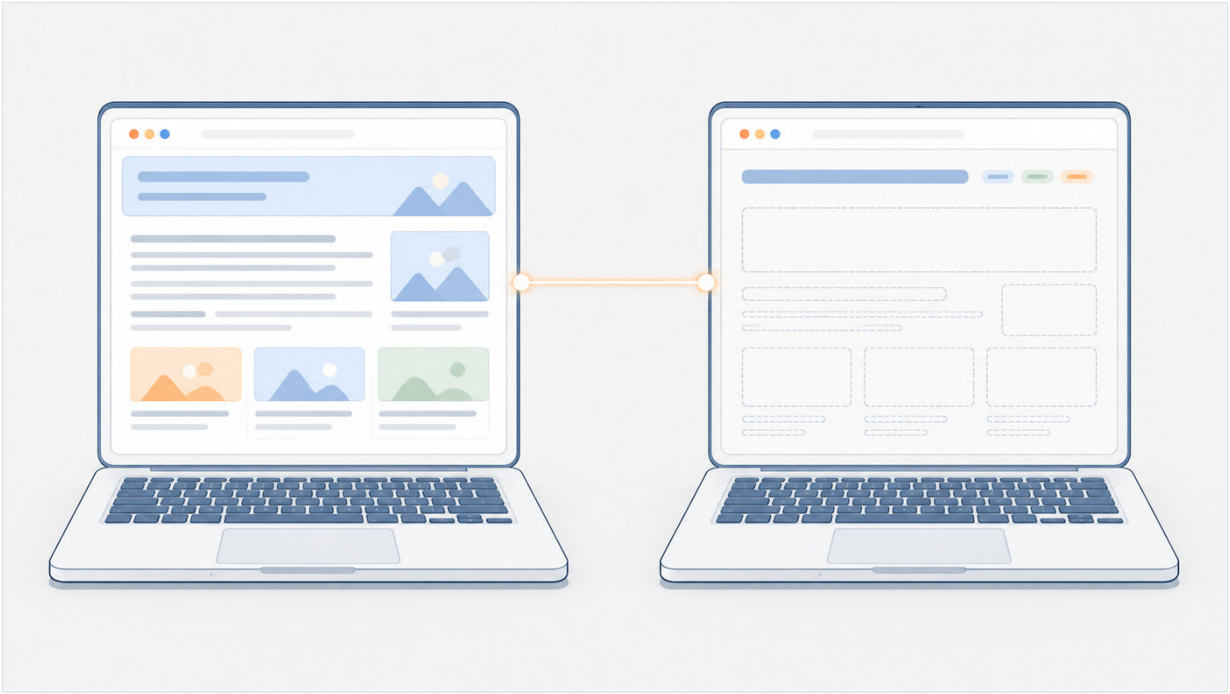
홈페이지에 제품 소개를 올려뒀는데 AI에게 물으면 엉뚱한 답이 돌아온다면, 크롤러가 본문을 아예 못 보고 있을 수 있습니다. 우리 블로그 54편에서 실제로 그랬습니다. 5분 자가점검법과 처방을 정리했습니다.

방동걸 (주)웨이스 대표이사 2026년 8월 20일

최근 우리 블로그를 점검하다 한 가지를 발견했습니다. 글 54편 전체가 크롤러에게는 사실상 빈 페이지였습니다. 제목도 있고 요약도 있고 태그도 붙어 있는데, 정작 본문 문장이 한 글자도 전달되지 않고 있었습니다. 웨이스가 제조 AI 운영체제 VEXPLOR의 블로그를 점검하며 직접 겪은 일에서 나온 처방입니다.

사람이 브라우저로 열면 멀쩡히 읽힙니다. 그래서 한동안 아무도 몰랐습니다.

이 글은 그때 확인한 것과, 같은 문제를 5분 안에 스스로 점검하는 방법을 정리한 것입니다. 홈페이지에 제품 소개와 기술 자료를 올려두었는데 AI에게 회사를 물으면 엉뚱한 답이 돌아오는 경우, 원인이 여기 있을 수 있습니다.



같은 페이지라도 사람이 보는 화면(왼쪽)과 크롤러가 받아오는 문서(오른쪽)는 다를 수 있습니다.

01

크롤러는 우리가 보는 화면이 아니라 받아온 문서를 봅니다

웹페이지가 화면에 그려지기까지는 2단계를 거칩니다. 서버가 HTML 문서를 보내주고, 브라우저가 그 안의 자바스크립트를 실행해 나머지를 채웁니다. 요즘 홈페이지는 대부분 두 번째 단계에서 본문을 채웁니다. 화면 전환이 부드럽고 개발이 편하기 때문입니다.

문제는 **크롤러 상당수가 두 번째 단계를 하지 않는다**는 점입니다. 서버가 보낸 HTML을 받아 그 안에 있는 것만 가져가고 끝냅니다. 자바스크립트로 채워질 예정인 본문은, 크롤러에게는 존재하지 않는 글입니다.

클라우드 서비스 업체 Vercel이 **2024년 12월에 공개한 집계**가 이를 보여줍니다. 자사 네트워크를 지나간 직전 한 달치 요청을 종류별로 나눈 것입니다.

크롤러	자바스크립트 실행	월간 요청 수
Googlebot (참고)	함	45억
GPTBot (OpenAI)	하지 않음	5억 6,900만
ClaudeBot (Anthropic)	하지 않음	3억 7,000만
AppleBot	함	3억 1,400만
PerplexityBot	하지 않음	2,440만
Gemini	함	구글 크롤러에 포함 — 별도 수치 없음

출처: Vercel, "The rise of the AI crawler"(2024년 12월 17일 공개). **1년 반 넘은 관측이므로 크롤러 동작은 그사이 달라졌을 수 있습니다.** 중요한 판단에 쓰실 때는 각 사 최신 문서로 다시 확인하시기 바랍니다.

흥미로운 대목이 하나 있습니다. ChatGPT 크롤러는 전체 요청의 11.50%, Claude 크롤러는 23.84%를 자바스크립트 파일을 가져오는 데 사용했습니다. **파일은 내려받지만 실행은 하지 않습니다.** 다운로드 기록만 보고 "잘 읽어가고 있다"고 판단하면 안 되는 이유입니다.

정리하면 이렇습니다. 구글과 애플 계열은 자바스크립트를 실행하므로 늦더라도 본문을 봅니다. 반면 ChatGPT·Claude·Perplexity 계열은 서버가 처음 보낸 HTML에 없는 내용을 영영 보지 못합니다.
검색엔진 크롤러 중에도 자바스크립트를 실행하지 않거나, 실행하더라도 뒤로 미루는 경우가 있습니다. 실행 여부가 보장되지 않으므로 **처음 보내는 HTML에 본문이 들어 있는 편이 안전합니다.**

> **한 줄로 줄이면**
>
> - 크롤러는 화면을 보는 것이 아니라 서버가 보낸 문서를 봅니다

> - 자바스크립트로 그려지는 내용은 상당수 크롤러에게 없는 것과 같습니다
> - 자바스크립트 파일을 내려받는 것과 실행하는 것은 다릅니다

02

우리 블로그 54편이 크롤러에게는 빈 페이지였습니다

우리 경우를 그대로 적겠습니다.

블로그 본문을 편집기 라이브러리로 그리고 있었습니다. 이 편집기는 브라우저 화면이 있어야 동작합니다. 서버에서는 아무것도 만들어내지 못하고 빈 결과를 냅니다. 그래서 서버가 보내는 HTML에는 제목·요약·태그·저자 소개·관련 글·푸터가 전부 들어 있는데 **본문만 빠져 있었습니다.**

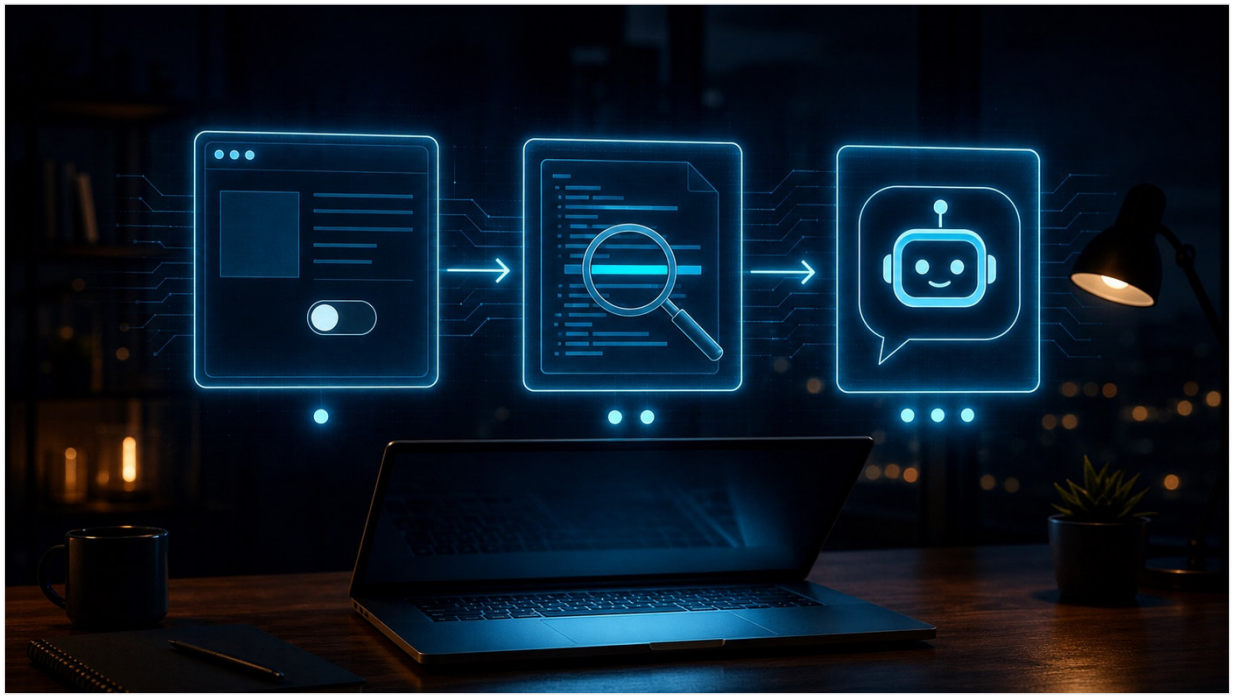
수리 전후를 같은 글 한 편에서 재봤습니다. 세 값 모두 그 한 편에서 나온 것입니다.

측정 항목	수리 전	수리 후
본문 영역의 문단 수	0개	51개
자바스크립트 없이 읽히는 본문	0자	3,443자
전송된 HTML 크기	98KB	107KB

같은 글 한 편을 2026년 8월 20일에 잰 값입니다. 수리 전에도 페이지 전체에는 문단이 10개 있었는데, 전부 저자 소개와 관련 글 같은 주변 요소였습니다. **본문 영역에는 문단이 하나도 없었습니다.**

전송량은 98KB에서 107KB로 약 9KB 늘었습니다. 본문이 통째로 들어갔는데 증가폭이 이 정도인 데는 이유가 있습니다. 종전에도 본문은 화면을 그리기 위한 **데이터 형태로는 이미 실려 나가고 있었기 때문입니다.** 읽히지 않는 데이터가 읽히는 문장으로 바뀐 것이지, 없던 것이 새로 얹힌 것이 아닙니다.

우리는 AI 도구용으로 본문을 평문으로 모아둔 파일을 별도로 운영하고 있었습니다. 그것이 완충 역할을 했을 가능성은 있지만, **얼마나 메꿔줬는지는 우리도 확인하지 못했습니다.** 그 파일을 읽고 인용했는지를 밖에서 알 방법이 없기 때문입니다. 분명한 것은 하나입니다. 크롤러가 먼저 보는 것은 본문 페이지이고, 그게 비어 있으면 나머지는 보조 장치일 뿐입니다.



자가점검 3단계 — 자바스크립트를 끄고 열어보고, 페이지 소스에서 본문 문장을 찾고, AI에게 요약을 시켜봅니다.

03

5분이면 우리 홈페이지도 확인할 수 있습니다

개발자가 아니어도 됩니다. 3가지 방법 중 편한 것을 고르면 됩니다.

방법 1 — 브라우저에서 자바스크립트를 끄고 열어봅니다. 브라우저 설정의 사이트 콘텐츠 항목에서 자바스크립트를 차단한 뒤(메뉴 위치는 브라우저와 버전마다 다릅니다) 우리 홈페이지 제품 소개 페이지를 엽니다. 본문이 그대로 보이면 정상입니다. 빈 화면이거나 로딩 표시만 돌면 크롤러도 같은 것을 보고 있습니다.

방법 2 — 페이지 소스를 열어 본문 문장을 찾아봅니다. 페이지에서 마우스 오른쪽 버튼을 눌러 '페이지 소스 보기'를 선택하고, 본문 중간의 특징적인 문장 하나를 그대로 검색합니다. **제목이나 요약 말고 본문 한가운데 문장이어야 합니다.** 제목·요약은 대개 잘 들어 있어서 그것으로 확인하면 통과한 것처럼 보입니다.

방법 3 — AI에게 직접 물어봅니다. ChatGPT나 Claude에 우리 회사 페이지 주소를 주고 "이 페이지 내용을 요약해 달라"고 합니다. 실제 내용과 다른 이야기를 하거나 "내용을 확인할 수 없다"고 하면 신호입니다.

3가지 모두 결과가 같아야 정상입니다. 방법 2가 가장 정확하고, 방법 3이 가장 현실에 가깝습니다.

우리는 이 점검에서 2번 잘못 판단했습니다

부끄럽지만 그대로 적겠습니다. 같은 함정에 걸리실 가능성이 높아서입니다.

처음에는 글의 요약 문장으로 찾아봤습니다. 소스에 그 문장이 있길래 "잘 들어가 있구나" 하고 넘어갔습니다. 그런데 그 문장은 본문이 아니라 검색 결과용 요약 정보에 들어 있던 것이었습니다. 요약은 대부분의 홈페이지가 정상적으로 내보냅니다. **요약으로 확인하면 거의 항상 통과합니다.**

두 번째로는 본문 한가운데 문장을 골라 찾았습니다. 이번에도 소스에 있었습니다. 그런데 자세히 보니 사람이 읽는 문장 형태가 아니라, 화면을 그리기 위한 데이터 뭉치 안에 들어 있는 것이었습니다. 자바스크립트가 그 데이터를 받아 화면을 그리기 전까지는 문장이 아닙니다. 크롤러가 이걸 문장으로 알아본다는 보장이 없습니다.

그래서 마지막에 확인한 기준은 이것이었습니다. **문장이 문단 태그(`<p>`) 안에 들어 있는가.** 수리 전에는 문단이 10개였고 그중 본문은 하나도 없었습니다. 수리 후에는 51개가 되었습니다.

기술을 몰라도 대체할 수 있는 방법이 있습니다. 소스 창에서 본문 문장을 찾았을 때, **그 문장 주변이 사람이 읽을 수 있는 문장들로 이어져 있는지** 보면 됩니다. 따옴표와 역슬래시가 잔뜩 붙은 데이터 덩어리 안에 있으면 아직 문장이 아닙니다.

> **점검할 때 자주 하는 실수**
>
> - 제목·요약으로 확인한다 — 그건 대개 들어 있습니다. 본문 한가운데 문장으로 확인해야 합니다
> - 문장이 데이터 덩어리 안에 있는데 통과로 본다 — 문단 태그 안에 있어야 문장입니다
> - 첫 화면만 본다 — 스크롤해야 나오는 내용이 따로 채워지는 경우가 있습니다
> - 한 페이지만 본다 — 홈은 정적이고 제품 상세만 자바스크립트인 경우가 흔합니다

04

안 읽히고 있다면 무엇을 고쳐야 하는가

고치는 방향은 하나입니다. **본문을 서버가 만들어 보내게 하는 것**입니다. 업계에서는 서버 사이드 렌더링이라고 부릅니다. 홈페이지를 만든 업체에 "제품 소개 본문이 서버 HTML에 포함되게 해달라"고 요청하면 통합니다.

전면 재구축이 필요한 일은 아닙니다. 우리도 본문을 그리는 부분 하나만 서버 쪽으로 옮겼습니다. 화면도 그대로고 편집 방식도 그대로입니다.

업체에 전달할 때는 이렇게 적으면 오해가 없습니다.

> **홈페이지 담당 업체에 보낼 요청 문구**
>
> - 제품 소개·기술 자료 페이지의 본문이 **서버가 보내는 초기 HTML에 포함**되게 해주세요
> - 확인 기준은 **자바스크립트를 끈 상태에서 본문이 보이는지**입니다
> - 본문 문장이 문단 태그 안에 들어 있어야 하며, 데이터 형태로만 실려 있으면 안 됩니다
> - 대상 페이지 목록: (우리 회사에서 중요한 순서로 5~10개를 적어 전달)

받는 쪽에서 "서버 사이드 렌더링으로 바꿔달라는 말씀이시죠"라고 되물으면 정확히 통한 것입니다.

당장 손대기 어렵다면 완충 장치를 먼저 둘 수 있습니다. 회사·제품 핵심 정보를 평문으로 정리한 파일을 따로 두는 방법입니다. 우리도 그렇게 해두고 있습니다.

다만 이건 보완재이지 대체재가 아닙니다. **효과를 과신하지 않는 편이 좋습니다.** 이런 파일을 읽는 도구와 읽지 않는 도구가 갈리고, 실제로 인용에 쓰였는지는 밖에서 확인하기 어렵습니다. 본문 페이지를 고치는 것이 먼저이고, 이건 그 다음입니다.

05

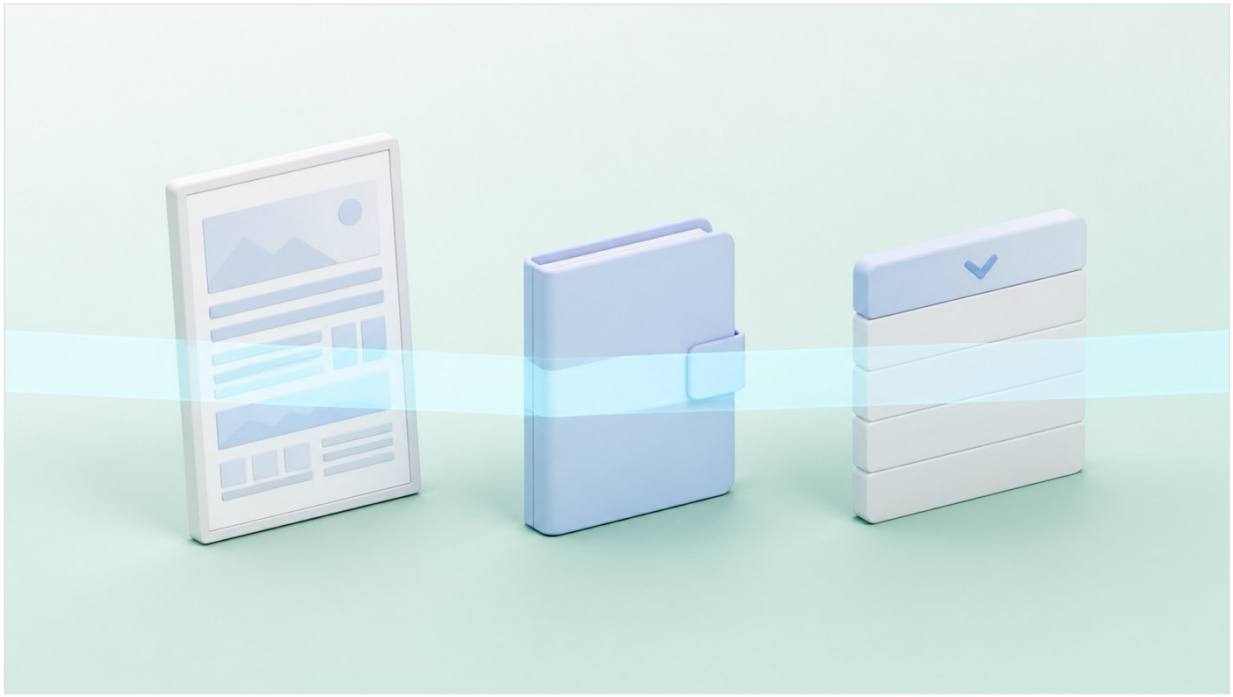
자바스크립트만 문제인 것은 아닙니다

같은 이유로 크롤러가 놓치는 것이 몇 가지 더 있습니다.

이미지 안의 글자. 사양표나 비교표를 이미지로 만들어 올린 경우, 그 안의 숫자와 항목명은 검색되지도 인용되지도 않습니다. 보기 좋게 만든 표일수록 손해가 큽니다. 값을 보여줄 때는 이미지가 아니라 표로 적는 편이 낫습니다.

PDF로만 있는 카탈로그. 제품 상세를 PDF 한 장으로만 제공하면 본문 페이지에는 내용이 없습니다. 최소한 핵심 사양은 페이지에 텍스트로 함께 적어두는 편이 낫습니다.

클릭해야 펼쳐지는 내용. 자주 묻는 질문처럼 접혀 있는 영역은 펼치기 전 상태의 HTML에 내용이 들어 있는지 확인이 필요합니다. 들어 있으면 문제없고, 클릭할 때 불러오는 방식이면 크롤러는 보지 못합니다.



자바스크립트 말고도 크롤러가 지나치는 것들 — 이미지 안의 글자, PDF로만 있는 자료, 클릭해야 펼쳐지는 영역입니다.

06

본문이 읽혀도 이름이 안 나오는 경우가 있습니다

여기까지 고쳤다면 최소 조건은 맞춘 것입니다. 다만 본문이 읽히는 것과 AI 답변에 회사 이름이 나오는 것은 다른 문제입니다.

2026년 8월 ChatGPT와 Gemini에 제조 관련 질문 4개를 직접 넣어보고 답변이 무엇을 인용하는지 봤습니다. **4건은 통계가 아니므로 비율로 말할 수 없습니다.** 다만 질문 유형에 따라 답이 갈리는 방향은 뚜렷했습니다.

- **"~ 업체를 추천해달라"는 질문**에는 회사 이름이 나옵니다
- **"~를 어떻게 하나"는 방법론 질문**에는 표준 문서와 기관 자료만 나오고 **업체 이름은 아예 등장하지 않습니다**
- **"국내 도입 사례"를 물으면** 도입한 제조사 이름은 나오는데, 그 시스템을 구축한 공급사는 나오지 않습니다

이게 뜻하는 바가 있습니다. 방법론 글을 아무리 많이 만들어도 그 자체로는 회사 호명으로 이어지지 않습니다. 인용된 출처는 대체로 언론과 공공 기관 자료였고, 업체 자사 홈페이지가 인용된 것은 이미 이름이 널리 알려진 곳뿐이었습니다.

그래서 순서가 이렇게 됩니다. 본문이 읽히게 만드는 것이 먼저이고(안 그러면 출발선에도 못 섭니다), 그 다음이 우리 이름과 함께 인용될 만한 근거를 밖에 만드는 일입니다. 첫 번째는 며칠이면 되고, 두 번째는 오래 걸립니다.

07

왜 지금 이걸 봐야 하는가

예전에는 검색 결과에서 우리 순위가 밀리는 정도의 문제였습니다. 지금은 다릅니다. 고객이 업체를 고를 때 검색창이 아니라 AI에게 먼저 묻는 경우가 늘고 있습니다. 그 답에 우리 회사가 등장하지 않으면, 검토 대상 목록에 처음부터 들어가지 못합니다.

우리가 겪은 일이 그 예입니다. 글을 54편 쌓는 동안 본문은 크롤러에게 전달되지 않고 있었습니다. 콘텐츠를 더 만드는 것보다 이미 만든 것이 읽히고 있는지 확인하는 편이 먼저였습니다.

홈페이지를 새로 만든 지 얼마 되지 않았다면 특히 한 번 보시기를 권합니다. 최근 방식일수록 자바스크립트로 화면을 그리는 경우가 많습니다. 5분이면 끝납니다.

WACE(웨이스)는 공장이 스스로 판단하고 멈추지 않게 하는 제조 AI 운영체제 VEXPLOR를 만드는 회사다.

원문 blog.vexplor.com/post/crawler-cannot-read-your-site
발행 2026년 8월 20일 · 방동걸 (주)웨이스 대표이사

회사 주식회사 웨이스 (WACE Inc.) · 대표이사 방동걸 · 사업자등록번호 439-87-02852
주소 (18462) 경기 화성시 동탄구 동탄영천로 150, 에이동(A동) 1702-1708호 (영천동, 실리콘앨리동탄)
연락처 전화 070-4730-0629 · contact@wace.me
홈페이지 wace.me · linkedin.com/company/wace-inc

원문 홈페이지 링크드인