PROGEOLAB

PROGEOLAB RESEARCH

Pharma AI Visibility

How 14 pharmaceutical giants appear to AI answer engines · April 2026

PROGEOLAB
April 2026
progeolab.ai/research

Contents

CHAPTER 1

Johnson & Johnson Serves Chrome 64/64 But Blocks ChatGPT 0/64: Pharma AI Visibility Report

1. Pharma's AI Visibility Challenge

The pharmaceutical industry occupies a unique position in the AI visibility landscape. Drug information is literally life-or-death content — prescribing information, contraindications, dosing guidelines, and recall notices must be accurate and current. Yet the industry's regulatory caution around content distribution creates a tension with AI-era information access.

Our Fortune Global 500 dataset contains 14 pharmaceutical companies spanning 8 countries. The pharma sector shows a moderate GEO gap: 2 of 14 companies block ChatGPT-User while serving Chrome, but both are among the sector's most consequential content producers.

2. The J&J Pattern: Browser-OK, AI-Blocked

Johnson & Johnson presents the single cleanest example of AI-specific blocking in the entire Fortune 500 dataset. J&J's probe results:

User Agent	Successful Probes	Rate
Research UA	—	—
Googlebot	—	—
Chrome	**64 / 64**	**100%**
ChatGPT-User	**0 / 64**	**0%**

Every single one of our 64 HTTP endpoints returns content to a Chrome browser. Every single one blocks ChatGPT-User. This is not partial blocking or inconsistent WAF behavior — it is a comprehensive, deliberate policy of serving browsers while rejecting AI.

J&J does not have JSON-LD on its homepage, no llms.txt, and no AI-specific robots.txt directives. The company has made no investment in AI discoverability and has actively invested in AI blocking. AI Readiness Score: 1/11.

The healthcare consequence: when a patient asks ChatGPT about Tylenol dosing, Motrin interactions, or Janssen immunology products, the model cannot access jnj.com. It answers from training data that may predate the latest FDA label update. J&J's own regulatory-reviewed content is replaced by whatever the model remembers.

3. AstraZeneca: Googlebot Trusted, ChatGPT Rejected

AstraZeneca shows a different but equally instructive pattern:

User Agent	Successful Probes
Research UA	—
Googlebot	—
Chrome	**45 / 64**
ChatGPT-User	**0 / 64**

AstraZeneca is accessible to Chrome (45/64) but blocks ChatGPT-User entirely. AstraZeneca also has JSON-LD on its homepage (Organization type) and is behind Cloudflare's WAF — suggesting sophisticated bot management is in place.

AstraZeneca's Cloudflare deployment likely uses Bot Management with AI-specific UA rules. The JSON-LD investment shows awareness of structured data for search engines, but the AI blocking means this structured data is inaccessible to AI crawlers.

4. The Open Pharma Companies

The majority of pharma companies — 11 of 14 — are accessible to ChatGPT-User. The sector's gap rate (2/14 = 14.3%) is higher than banking (5.3%) but lower than telecom (26.7%).

The most AI-accessible pharma companies:

Company	Country	Chrome	ChatGPT	JSON-LD	Score
GSK	Britain	63	63	Yes	5
Merck	U.S.	10	10	Yes	5
Novartis	Switzerland	9	9	Yes	5
Bristol-Myers Squibb	U.S.	10	10	Yes	5
Roche	Switzerland	10	10	—	4
Pfizer	U.S.	10	10	—	4
Novo Nordisk	Denmark	6	6	—	4
Sanofi	France	7	7	—	4
Eli Lilly	U.S.	2	2	—	4

GSK leads with 63/64 probes accessible to both Chrome and ChatGPT — nearly full transparency to AI systems. GSK also has JSON-LD. Merck, Novartis, and Bristol-Myers Squibb combine JSON-LD with ChatGPT accessibility for the highest pharma AI readiness scores.

Notably, no pharma company has an llms.txt file, no pharma company has explicit AI bot rules in robots.txt, and no pharma company links to Wikidata. The sector's AI optimization is zero — accessibility is

accidental, not deliberate.

5. Regulatory Context: Why Pharma Blocks AI

Pharmaceutical content distribution is regulated by the FDA (U.S.), EMA (EU), MHRA (UK), and national agencies. Key regulatory concerns around AI include:

Off-label promotion. If AI generates answers that describe uses beyond a drug's approved indications, the pharma company could face regulatory scrutiny — even if the content originated from their website and was redistribution by AI without context.

Adverse event reporting. FDA requires pharma companies to monitor and report adverse events mentioned in connection with their products. If AI redistributes drug information and users report side effects to ChatGPT rather than to the company's safety database, the adverse event reporting chain breaks.

Content approval processes. Pharma marketing content undergoes medical-legal-regulatory (MLR) review. AI-generated answers that paraphrase or recombine approved content may produce statements that have not undergone MLR review.

These are legitimate concerns. However, blocking AI does not address them — it makes them worse. When AI cannot access current, approved content from the company's website, it substitutes stale or third-party content that has undergone no regulatory review at all.

6. What Patients See When AI Cannot Read Drug Info

The practical impact of pharma AI blocking is testable. When a patient asks an AI about a J&J medication:

What the AI should cite: Current prescribing information from jnj.com, reviewed by J&J's regulatory affairs team, approved by the FDA, updated with the latest safety data.

What the AI actually cites: Training data from months ago, supplemented by third-party sources (WebMD, Drugs.com, Wikipedia) that may not reflect the latest label changes. The AI answer carries no regulatory provenance and may contradict the current label.

The irony: J&J's blocking is likely motivated by regulatory caution, but the blocking produces exactly the uncontrolled content redistribution that regulatory teams fear.

7. The Compliance-Visibility Tradeoff

The pharma sector needs a framework that reconciles regulatory compliance with AI visibility:

Allow retrieval, block training. Allow ChatGPT-User and PerplexityBot (which fetch content for real-time answers) while blocking GPTBot and Google-Extended (which ingest content for model training). This ensures patients get current prescribing information while limiting training data contribution.

Implement llms.txt with regulatory content. Create an llms.txt file that points specifically to approved prescribing information, patient education materials, and safety data. This curates what AI can access rather than leaving it to crawl everything.

Add JSON-LD with medical schema types. Use MedicalWebPage, Drug, and MedicalCondition schema types to provide structured metadata that helps AI systems understand the regulatory context of the content.

8. The Pharma AI Scorecard

Company	Country	Chrome	ChatGPT	JSON-LD	WAF	Score	Status
Merck	U.S.	10	10	Yes	—	5	Accessible
Novartis	Switzerland	9	9	Yes	F5 Bigip	5	Accessible
Bristol-Myers Squibb	U.S.	10	10	Yes	Imperva	5	Accessible
GSK	Britain	63	63	Yes	F5 Bigip	5	Accessible
Roche Group	Switzerland	10	10	—	F5 Bigip	4	Accessible
Pfizer	U.S.	10	10	—	Cloudflare	4	Accessible
AbbVie	U.S.	1	1	—	Cloudflare	4	Accessible
Sanofi	France	7	7	—	F5 Bigip	4	Accessible
Guangzhou Pharmaceutical Holdings	China	2	2	—	—	4	Accessible
Eli Lilly	U.S.	2	2	—	Cloudflare	4	Accessible
Novo Nordisk	Denmark	6	6	—	F5 Bigip	4	Accessible
AstraZeneca	Britain	45	0	Yes	Cloudflare	2	GEO Gap
Johnson & Johnson	U.S.	64	0	—	—	1	GEO Gap
Bayer	Germany	0	0	—	F5 Bigip	0	Unreachable

No pharma company scores above 5/11. The sector has the lowest maximum score of any industry vertical we've analyzed. The opportunity for the first pharma company to implement llms.txt and explicit AI bot policies is significant — it would immediately become the most AI-visible pharmaceutical company in the Fortune 500.

Source data: *Cross-source analysis from the Fortune Global 500 AI Accessibility audit (April 2026). Domain mapping from the audit database. All 14 companies classified as "Pharmaceuticals." UA probe data from four runs. JSON-LD from homepage extraction. WAF from response body signatures. Regulatory context from publicly available FDA, EMA, and MHRA guidance documents.*

About PROGEOLAB

PROGEOLAB is an AI-native visibility intelligence platform.

This report is part of the PROGEOLAB Fortune 500 AI Accessibility Audit — a series of research studies on how large enterprises appear to (or disappear from) AI answer engines. All measurements are from live HTTP probes across four user agents: a research bot, Googlebot, Chrome, and ChatGPT-User. No estimates, no third-party data sources.

Methodology in brief

500 companies · 67 probes each · 4 user agents · 134,000 probe requests. Data collected April 16–19, 2026. Response bodies stored and re-validated with MD5-hash soft-404 detection to eliminate the ~25× inflation that status-code-only scans produce.

Contact & next steps

Visit **progeolab.ai/research** or request a demo for a complimentary AI visibility analysis of your organisation.

For press enquiries, data requests, or syndication, write to **research@progeolab.ai**.

PROGEOLAB