

HOW READY ARE INDIA'S ECOMMERCE SITES FOR **AI AGENTS?**

October 2026.

Key findings

The major AI agents accessed India's ecommerce sites far more often than a generic bot did and completed 80% of simple shopping tasks. What stopped them was rarely a firewall. It was unclear information, content that needs JavaScript, and location or login walls. We tested 50 sites in two parts in September 2026: a technical crawl, then 450 real tasks run by Claude, GPT and Gemini.

1. **Real agents completed 358 of 450 tasks (80%).** Return windows were easiest (91%). Delivery charges (74%) and finding a product with its price (73%) were harder.

2. **Generic bots get blocked; named AI agents mostly do not.** 21 of 50 sites blocked or failed our generic cloud crawler. Yet on those same sites, the real agents succeeded almost as often as on open sites: 77% of tasks on sites that fully blocked our crawler, against 82% on open sites

3. **robots.txt still matters.** The only site that blocks shoppers' AI agents in robots.txt was also one of the four sites where agents completed a third of tasks or fewer. Claude and Gemini were both refused there.

4. **The most common reason agents failed was unclear information, not blocking.** Of 92 failed tasks, 26 failed because the answer was not clearly stated on the site, mostly delivery charges. 13 failed because content needed JavaScript. Only 11 failed because of a block or error response.

5. **Agents often disagree when policies are complicated.** All three agents gave the same return window on 29 of 50 sites. On 13 sites they gave different answers, usually because the site has several return rules or contradicts itself.

6. **Half of product pages hide the price from simple readers.** On 12 of 24 product pages we could fully test, the price was not in the server HTML, only in JavaScript. Agents that cannot run JavaScript said so when they failed.

7. **Platform defaults decide who is ready for new standards.** All 8 sites showing WebMCP run on Shopify, which switched it on for its stores in August 2026.

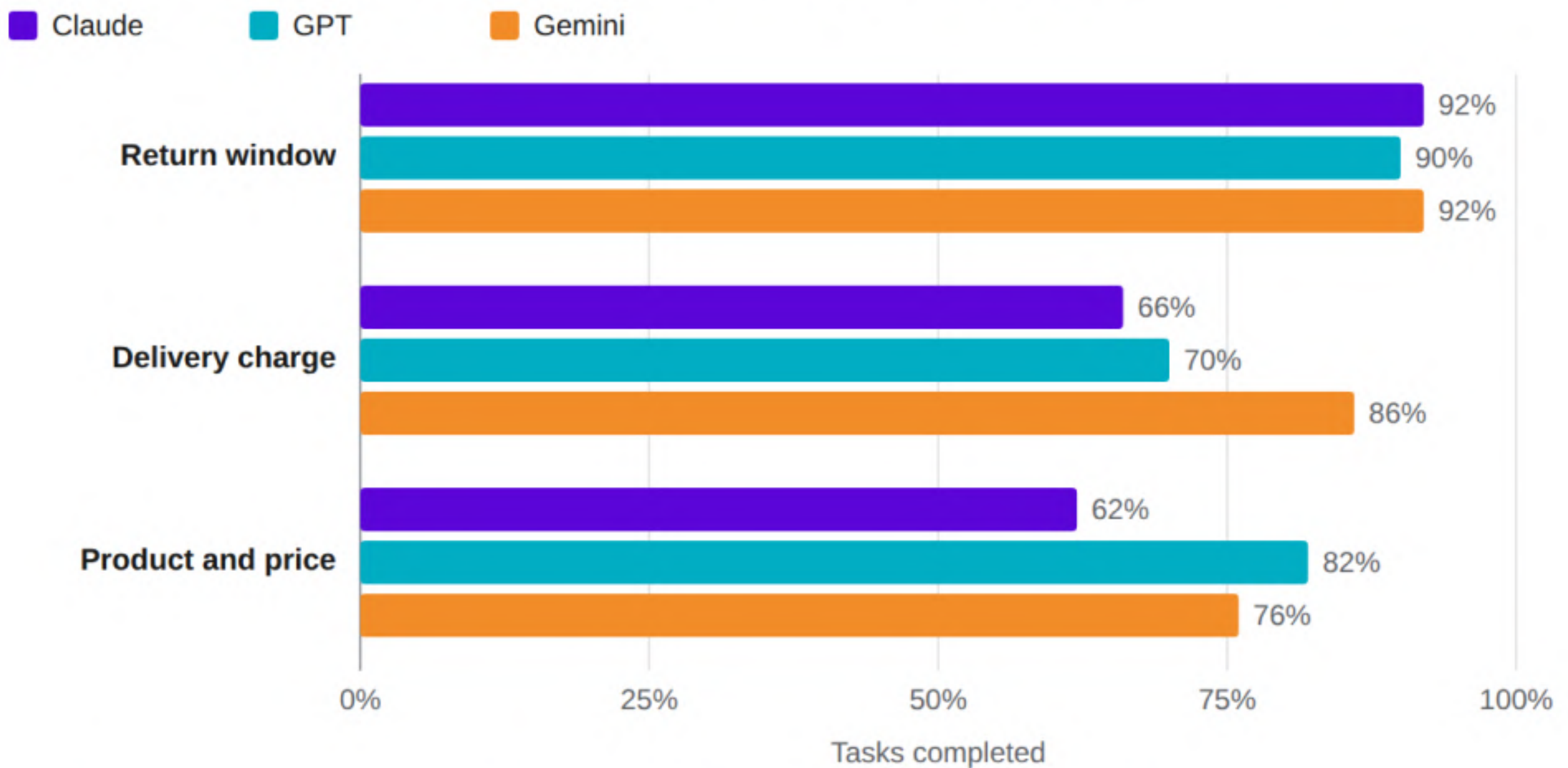
The cross-check between the two parts is the most important result. Without Part 2, we would have reported that four in ten sites block AI agents. With it, we can say what is actually true: most sites block unknown bots, but let the major AI agents in.

What real **AI agents** experienced

Claude, GPT and Gemini completed 80% of the shopping tasks we gave them, including on most of the sites that had blocked our generic crawler. Each agent tried three tasks on all 50 sites, using its own built-in web tools, and had to answer from the official site.

Return windows were easy; delivery charges and prices were harder

Share of tasks completed from the official site, by task and AI agent (50 sites each)



Source: FTA Global Agent Readiness Study 2026, 450 tasks by Claude, GPT and Gemini, 27 Sep 2026

We counted a task as complete only when the agent said it found the answer on the official site, cited an official page, and gave a clear value. Gemini did best on delivery charges, GPT on finding a product and price, and all three were strong on return windows.

The cross-check: blocked to our crawler, but not to real agents

What our generic crawler got in Part 1	Sites	Tasks real agents completed
Open	29	82%
Blocked the plain fetch only	5	82%
Blocked both the plain fetch and the browser	14	77%
Could not load (site error)	2	50%

The sites that fully blocked our crawler still let real agents complete three in four tasks. Five of them let all three agents complete all nine tasks. So the firewalls on these sites are doing what they should: stopping unknown automated traffic, while letting the major AI agents through.

The exceptions were telling. On one large marketplace whose robots.txt disallows shoppers' AI agents, Claude and Gemini were refused, and GPT could open the homepage but not product pages. On a quick commerce site, GPT got "403 Forbidden" on the policy pages. On a large fashion site, two agents got the same "Site Maintenance" page our crawler saw.

Why agents failed

We sorted every failed answer by the reason the agent gave.
Of the 92 tasks agents could not complete:

Reason the agent gave	Failed tasks
The answer was not clearly stated on the site	26
Content needed JavaScript or did not load	13
Blocked, or the site returned an error	11
The site asked for a location or pincode first	5
The site showed a maintenance page	3
The site needed a login	3
Other or unclear	31

Most of the "not clearly stated" failures were delivery charges. Many sites mention free delivery without a clear threshold, or give different figures on different pages.

Do **agents agree** with each other?

Return window answers across the three agents	Sites
All three gave the same answer	29
Different answers, because the site has different rules for different products	9
Contradictory answers, often because the site itself states two different windows	4
At least one agent could not find an answer	8

When a site has one clear rule, agents agree. When it has several, or states the same rule differently on two pages, a shopper's answer depends on which agent they happen to use.

Were the **product answers** real?

Agents gave an official product link and price in 110 of 150 product tasks.
We then opened those links from our own servers. Of the 74 we could check:

What we found at the link	Answers
The page did not exist (404 error)	3, all from Gemini
The page exists and shows the price the agent gave	54 (73%)
The page exists, but the price was not in its source, often because it loads by JavaScript or changed	17

The other 36 links sat on sites that blocked our checker, so we could not verify them. The three dead links show that agents can still invent or misremember product URLs, even when told to use the official site.

The **technical crawl** in numbers

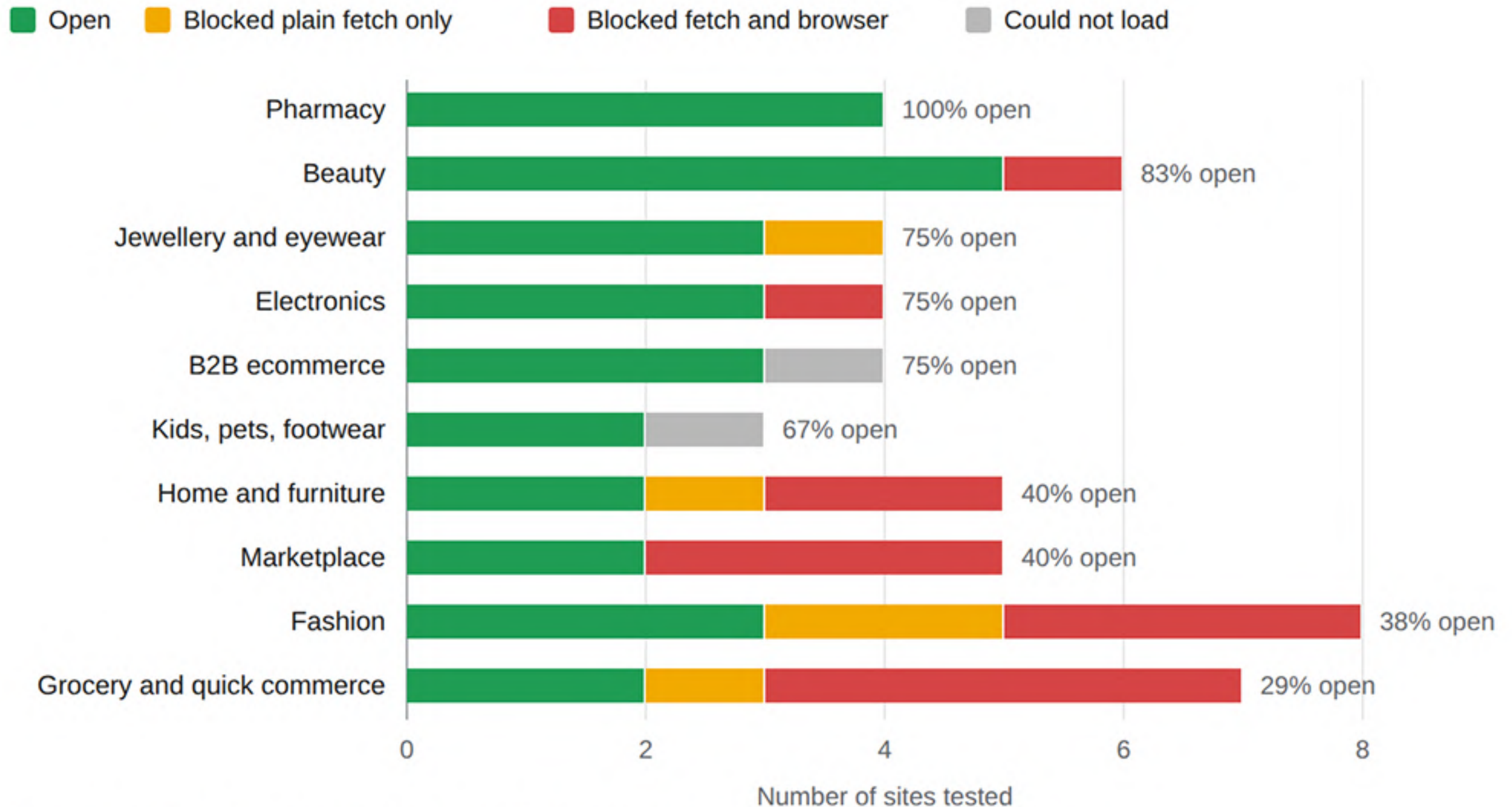
Every number below comes from our own crawl of 50 sites, run on 26 September 2026. The base changes by row, because a site that blocked us could not be tested further.

Pillar	What we measured	Result	Base
Act	Buttons with a readable name	89%	All buttons on 24 product pages
Act	Product pages with a clear "add to cart" or "buy now" control	18 (75%)	24 product pages
Reach	Sites that showed a normal homepage to an automated visit	29 (58%)	50 sites
Reach	Sites that blocked both a plain fetch and a full browser	14 (28%)	50 sites
Reach	Sites whose robots.txt allows shoppers' AI agents (ChatGPT-User, Claude-User, Perplexity-User)	32 (94%)	34 sites with a readable robots.txt
Reach	Sites whose robots.txt allows AI search bots	32 (94%)	34 sites
Read	Median share of homepage text visible without JavaScript	90%	29 open sites
Read	Open sites showing less than half their homepage text without JavaScript	7 (24%)	29 open sites
Read	Product pages with the price in the server HTML	12 (50%)	24 product pages
Read	Product pages with the product name in the server HTML	17 (71%)	24 product pages
Understand	Product pages with Product schema in the server HTML	19 (79%)	24 product pages
Understand	Product pages with return policy in schema	3 (13%)	24 product pages
Understand	Product pages with delivery details in schema	2 (8%)	24 product pages
Act	Form fields with a proper label	40%	58 form fields on 24 product pages
Standards	Sites with an llms.txt file	15 (30%)	50 sites
Standards	Sites with signs of WebMCP	8 (16%)	50 sites
Overall	Average FTA Agent Readiness Score	43 out of 100	42 fully scored sites

Who lets a generic bot in

Grocery, fashion and marketplaces block generic bots most

Number of sites by what our generic crawler got, per category (50 sites, Part 1)



Note: in Part 2, real AI agents still completed 77% of tasks on sites that fully blocked this crawler.

Source: FTA Global Agent Readiness Study 2026, crawl of 50 Indian ecommerce sites, 26 Sep 2026

Every pharmacy site we tested showed a normal page. Grocery and quick commerce did worst: 5 of 7 sites blocked an automated visit in some way. But in Part 2, the major AI agents still completed 63% of tasks on grocery sites, so most of these blocks stop unknown bots rather than ChatGPT, Claude or Gemini.

Finding 1: Firewalls block generic bots, but mostly let the major AI agents in

Almost every site says "yes" to AI agents in robots.txt, and many firewalls say "no" to unknown automated visitors. Part 2 showed that most of those firewalls still let the major AI agents through, so the real risk is for smaller or newer agents, not for ChatGPT, Claude and Gemini.

What robots.txt says

Type of AI bot	Allowed everywhere we checked	Blocked on some pages	Blocked	Base
Shoppers' AI agents (ChatGPT-User, Claude-User, Perplexity-User)	32	1	1	34 sites
AI search bots (OAI-SearchBot, Claude-SearchBot, PerplexityBot)	32	1	1	34 sites
AI training crawlers (GPTBot, ClaudeBot, Google-Extended and others)	30	4	0	34 sites

Only one site, a large marketplace, blocks shoppers' AI agents in robots.txt on purpose. One fashion site blocks the product pages we tested for every bot, AI or not. Four sites limit training crawlers: two block only Common Crawl or ByteDance, one blocks almost all of them, and one is the fashion site above. None of the 50 sites used Cloudflare's new Content-Signal line to state how AI may use their content.

What happened when our generic crawler visited

The firewall told a different story. 14 sites refused both a plain fetch and a full headless browser. They returned "Access Denied" pages, HTTP 403 errors, "Just a moment" challenge pages or empty 202 responses. 5 more sites blocked the plain fetch but let the browser through after a JavaScript challenge. One large fashion site served a "Site Maintenance" page to both.

We could not read robots.txt at all for 16 sites, mostly because the firewall blocked that file too. So for those sites, even the site's own instructions to bots were unreachable.

Why this matters

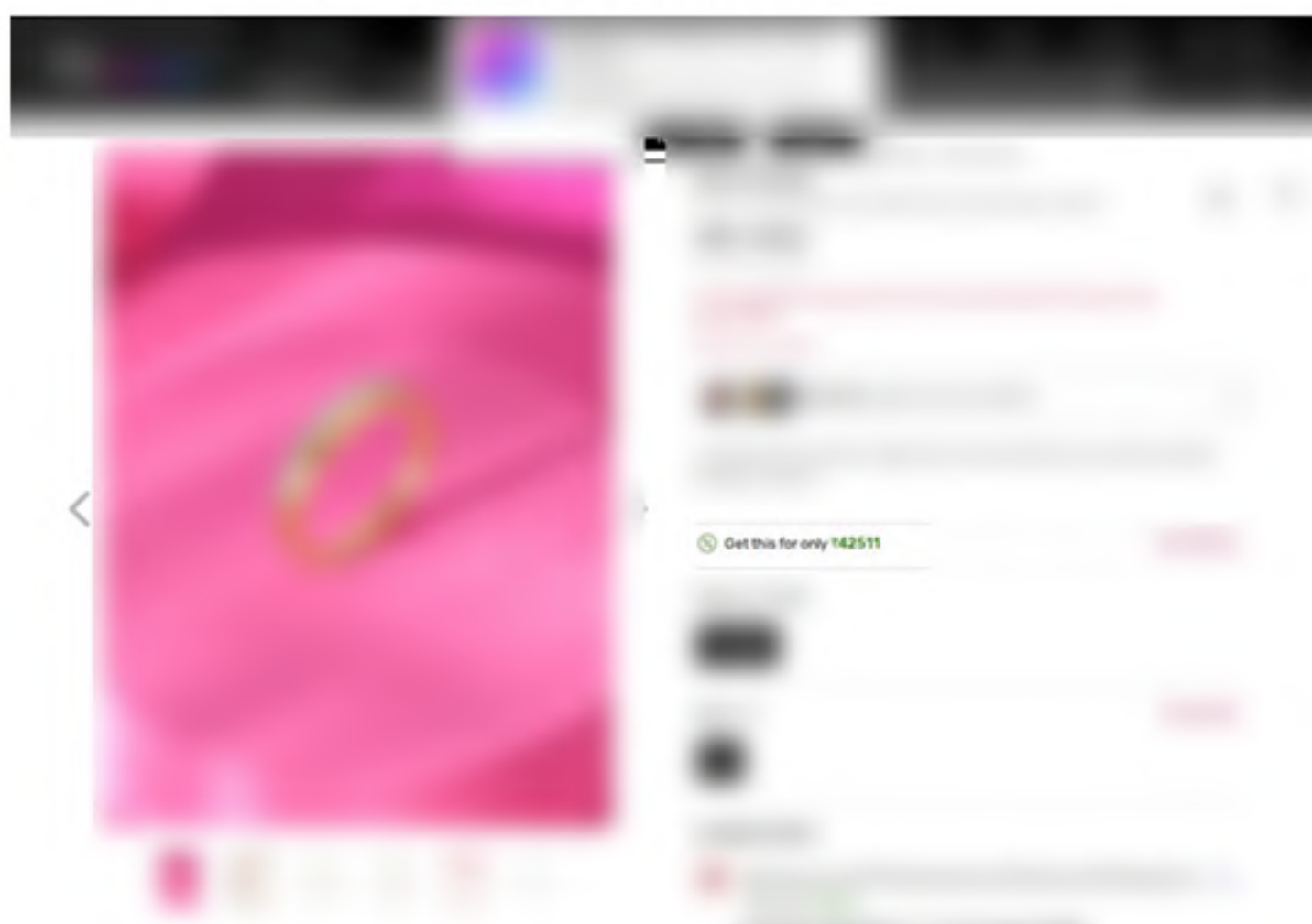
In Part 2, real agents completed 77% of tasks on the 14 sites that fully blocked our crawler, close to the 82% on open sites. So these firewalls tell known AI agents apart from unknown traffic. That is good for today's big assistants, but it means any new agent, a smaller AI company or a brand's own procurement bot may still hit the wall. The one clear exception was robots.txt: the marketplace that disallows shoppers' agents there was one of the hardest sites for agents to use.

Caution: we ran this crawl from Google Cloud servers, as an unknown bot. The blocks in this finding describe what unknown automated traffic gets. Part 2 describes what the major AI agents get, and the two are very different.

Finding 2: JavaScript hides key content, including prices

Most homepages are readable without JavaScript, but product pages are not: half of them hide the price from simple agents. And the price is the one fact every shopping agent needs.

What a shopper sees (JavaScript on)



What a simple AI agent gets (JavaScript off)



The same product page from one site in our study, loaded with JavaScript on (left) and off (right). With JavaScript off, the page shows one line: "You need to enable JavaScript to run this app." Brand details are blurred.

Homepages

Across the 29 sites that showed a normal homepage, the median site showed 90% of its homepage text without JavaScript. But 7 of the 29 showed less than half. Two, a large marketplace and a dairy delivery brand, showed nothing at all.

Product pages

We could fully test 24 product pages. The results:

What a simple agent can see in the server HTML	Product pages	Share
The product name	17 of 24	71%
The price	12 of 24	50%
Product schema, which often includes the price	19 of 24	79%

So on some pages the price sits inside schema but not in the visible text. Whether AI agents read schema is still debated, which makes these pages a gamble.

Why this matters

Most AI crawlers and many AI assistants read only the HTML the server sends, without running JavaScript. When the price is missing from that HTML, the agent cannot confirm the price. It may skip the product, guess, or pick a rival whose price it can read. Part 2 bore this out: 13 failed tasks were put down to content that needed JavaScript, 11 of them by Claude, whose page reader does not run JavaScript.

Finding 3: Product schema is common, but return and delivery data is missing

Indian ecommerce sites have done the basic schema work, but not the parts that answer a shopper's real questions: can I return it, and when will it arrive?

What the product schema included	Product pages	Share
Product schema in the server HTML	19 of 24	79%
Any Product schema (in server HTML or after JavaScript)	21 of 24	88%
Brand	19 of 24	79%
An offer with price	18 of 24	75%
Stock availability (in server HTML)	16 of 24	67%
SKU, GTIN or model number	15 of 24	63%
Ratings	11 of 24	46%
Return policy	3 of 24	13%
Delivery details	2 of 24	8%

Two more patterns stood out:

- **Schema that only appears after JavaScript.** 2 sites added Product schema through JavaScript. Agents that read only the server HTML never see it.
- **Price in schema but not on the page.** 7 sites had the price in schema but not in the visible server HTML. If an agent ignores schema, it sees no price at all.

At site level, 23 of 29 open homepages carried Organization schema, and 17 of 50 sites described their site search in schema, which helps agents search the catalogue directly.

Why this matters

Returns and delivery are the two questions shoppers ask most before buying. When they are missing from structured data, an agent must dig through policy pages, often with vague wording. Sites that add return and delivery schema give agents a direct, trusted answer.

Finding 4: Buttons are labelled, forms are not, and pop-ups get in the way

Most buttons tell an agent what they do, but most form fields do not. That matters most for the fields agents need to fill: pincode, phone number and search.

What we checked on 24 product pages	Result
Buttons with a readable name	89% of all buttons
Form fields with a proper label	23 of 58 fields (40%)
Form fields labelled only by placeholder text	31 of 58 fields (53%)
Pages with a clearly named "add to cart" or "buy now" control	18 of 24
Pages where that control was a real button or link	17 of 24
Sites where a pop-up covered the homepage or product page on arrival	3 of 24

What the **pop-ups** were

The three full-screen pop-ups were a login prompt on a marketplace homepage, a "call back" request form on a mattress product page, and a "you are visiting from a different country" prompt on an electronics site. Each one stands between an agent and the product.

What we **could not** find

On 6 of 24 product pages we found no clearly named add to cart control. Some of these sites ask for a delivery location first, and others use unlabelled icons. Either way, an agent has to guess.

Why this **matters**

Browser agents read the names and labels on a page, much like screen readers do. Placeholder text disappears as soon as someone types, and is often ignored by assistive tools. A proper label on the pincode field is one of the cheapest AXO fixes there is.

Finding 5: Platform defaults decide who is ready for new standards

Almost every Indian ecommerce site that supports the newest agent standards got them from its platform, not from its own team. That is good news for brands on those platforms, and a warning for everyone else.

WebMCP

8 of 50 sites showed signs of WebMCP, the new browser standard that lets a site tell agents which actions it supports. All 8 run on Shopify.

[Shopify switched on WebMCP tools for its standard storefronts in August 2026](#), with no work needed from store owners. No site on any other platform showed WebMCP.

llms.txt

15 of 50 sites had an llms.txt file. But look closer:

Type of llms.txt file	Sites
The same "Agent Instructions" template, about 4.4 KB each, typical of a platform-generated file	8
A custom file written about the brand and its catalogue	6
A generic file that does not describe the site at all	1

Why this matters

Platform choice is becoming an AXO decision. A brand on a platform that ships agent features by default gets them for free. A brand on a custom stack must build them. Neither WebMCP nor llms.txt is proven to change results yet, so we did not include them in the readiness score. But brands on custom stacks should watch these standards closely, because their competitors on big platforms may get them overnight.

Finding 6: Agents come from abroad, and some sites treat them as foreign shoppers

AI agents usually run on servers outside India, so sites that change their content by country may show agents the wrong store. Our crawl ran from cloud servers outside India, and two open sites reacted to that.

- An eyewear brand redirected us to its US store. Every page we reached, including the return policy, was the US version, with US prices and US rules.
- An electronics brand covered every page with a full-screen "you are visiting from a different country" prompt. A human can click "continue". Many agents will not, or will read the prompt instead of the product.

This is a quiet but real risk. [Vercel and MERJ found](#) that the major AI crawlers they measured operate from US data centres. When an Indian shopper asks an AI assistant about a product, the assistant may fetch the page from the US. If the site then shows US prices or a country pop-up, the answer the Indian shopper gets is wrong.

What to do about it

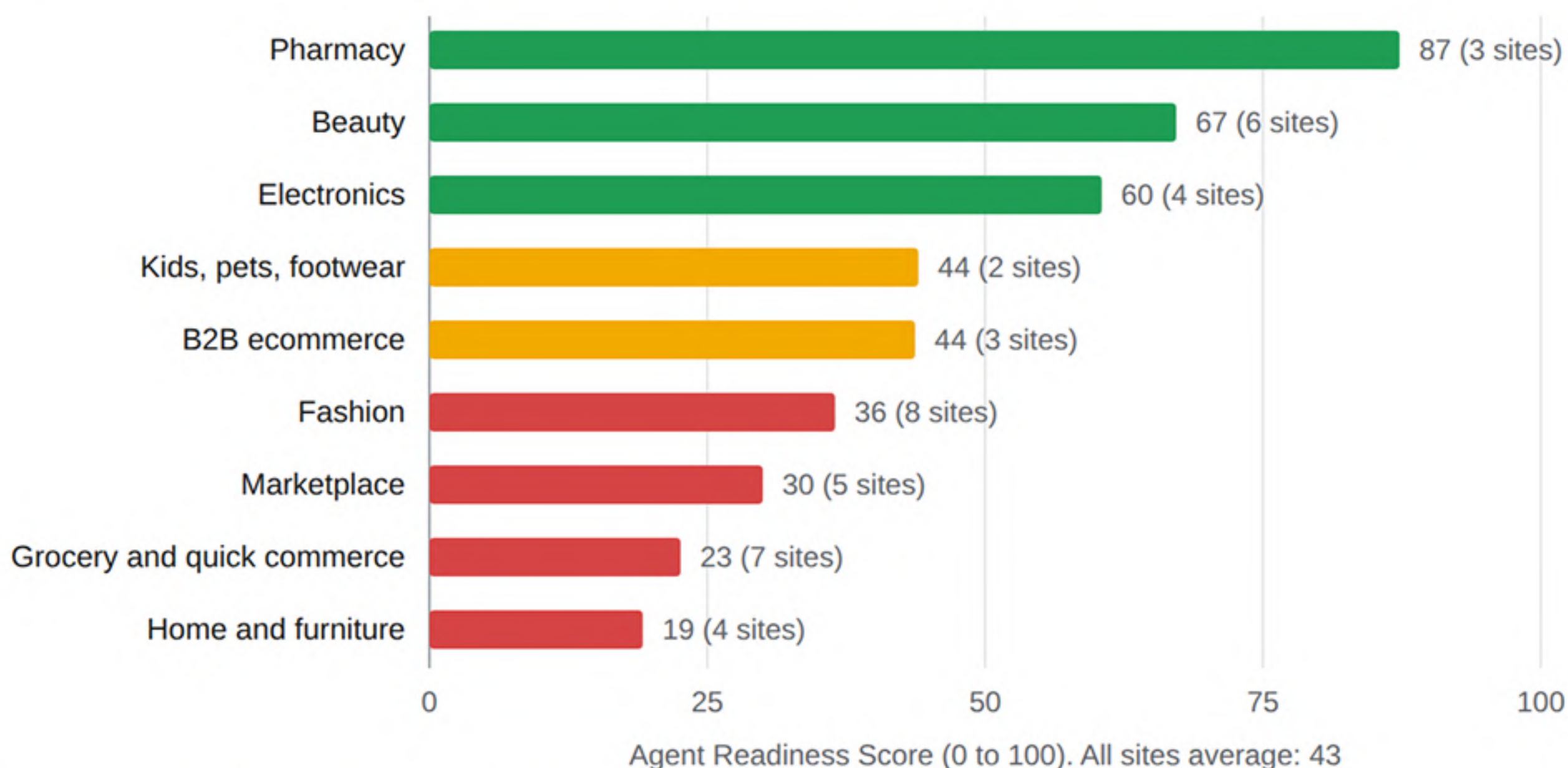
Serve the Indian store by default on Indian domains, whatever the visitor's location, and offer other regions as a choice rather than a forced redirect. Never cover the page with a country prompt. A small banner does the same job without blocking agents.

Agent Readiness Scores

The average site scored 43 out of 100, but the spread is huge: 12 sites scored 80 or more, while 18 scored under 20. Almost all of the low scores come from sites that blocked our generic crawler at the door.

On the technical score, pharmacy leads; grocery and home trail

Average Part 1 Agent Readiness Score by category (out of 100, 42 fully scored sites)



Source: FTA Global Agent Readiness Study 2026, crawl of 50 Indian ecommerce sites, 26 Sep 2026

Jewellery and eyewear has no category score, because the product pages we captured for those four sites were listing pages, not single products. We left them out rather than mark them down for our mistake.

Read these scores with Part 2 in mind. The Reach pillar reflects what our generic crawler got, so sites that blocked it scored near zero, even though real agents did well on most of them. Home and furniture, for example, averaged 19 on this score, yet real agents completed 89% of tasks there. Among the 29 open sites, where the score measures the pages themselves, it tracks real agent success reasonably well, with a correlation of 0.63. In the next edition we will rebuild the Reach pillar around what real agents get.

The top 5 sites

Rank	Site	Category	Reach (of 25)	Read (of 25)	Understand (of 20)	Act (of 20)	Score (of 100)
1	boAt	Electronics	25	25	17	19	95
2	Flipkart	Marketplace	25	25	20	13	92
3	Mamaearth	Beauty	25	24	16	17	92
4	Moglix	B2B ecommerce	25	25	17	16	92
5	Truemeds	Pharmacy	25	25	15	17	91

Pillar scores are rounded. The total is scaled to 100 from the 90 points we could measure, because Trust was not scored in Part 1.

What the leaders share: they let agents in, they send their key content in the server HTML, and their product pages carry schema with price and stock. None of them is perfect. Most of them lose points on unlabelled form fields, and only two of the five describe returns and delivery in schema.

How the 42 scores spread

Score band	Sites
80 to 100	12
60 to 79	5
40 to 59	4
20 to 39	3
0 to 19	18

The middle is thin. Most sites are either broadly ready or not reachable at all, which means the single biggest lever for most brands is access, not fine-tuning.

What **Indian ecommerce** brands should do now

Make your delivery and return rules unmistakable, make the price readable without JavaScript, and keep your bot rules open to shoppers' agents. In Part 2, unclear information caused more failed tasks than blocking did, so clarity comes first. Firewall rules still matter for smaller and newer agents.

Priority	Action	Who owns it	Finding it fixes
1	State one clear delivery charge rule and one clear return window, each on one public page, and repeat them on product pages	Ecommerce and legal	Part 2 failure reasons and agent disagreement
2	Put product name, price, stock and variants in the server HTML, not only in JavaScript	Engineering	Finding 2
3	Keep robots.txt open to shoppers' AI agents and AI search bots, even if you block training crawlers	SEO	Finding 1 and Part 2
4	Review firewall, CDN and bot-management rules, so verified AI agents are allowed and not just the ones your vendor already knows	Security or IT, with SEO	Finding 1
5	Add return policy and delivery details to product schema	SEO and engineering	Finding 3
6	Show products and prices without forcing a location, pincode or login first	Product and engineering	Part 2 failure reasons
7	Give every form field a real label, starting with pincode, phone and search	Engineering or UX	Finding 4
8	Remove full-screen pop-ups on first load, including login and country prompts, and serve the Indian store by default	Product and engineering	Findings 4 and 6
9	If you are not on a platform that ships agent features, plan your own path to WebMCP and agent-friendly APIs	Product and engineering	Finding 5

How to check your own site in 15 minutes

1. Open a product page, turn off JavaScript in your browser and reload. If the price disappears, you have a Read problem.
2. Open `yoursite.com/robots.txt` from a phone on mobile data and from a laptop on office Wi-Fi. If either shows an error or "Access Denied", your firewall is blocking even that file.
3. Ask ChatGPT, Claude or Perplexity: "What is the price of [one of your products] on [your site]?" If it says it cannot access your site, or quotes a different site, you have a Reach problem.

For the full list of practices, see the FTA guide to Agent Experience Optimization.

Method

We crawled 50 Indian ecommerce sites on 26 September 2026 with an open, repeatable script and scored each one on the FTA AXO Framework, then ran 450 real agent tasks on the same sites on 27 September.

The sites

50 leading Indian ecommerce sites across 10 categories: marketplaces (5), fashion (8), beauty (6), electronics (4), grocery and quick commerce (7), pharmacy (4), home and furniture (5), jewellery and eyewear (4), kids, pets and footwear (3), and B2B ecommerce (4). This is a selection of leading sites across categories, not a ranking of the top 50 by revenue.

The pages

For each site, the script visited four pages: the homepage, one category page, one product page and the return policy page. It found these pages on its own by following the site's links and sitemap.

The visits

Each page was visited three ways:

1. A plain fetch with a normal browser identity, the way many AI assistants read a page.
2. A plain fetch with a clearly named research bot identity, to see how sites treat declared bots. We never pretended to be any AI company's crawler.
3. A full headless Chrome browser with JavaScript on, plus a second load with JavaScript off for screenshots

The script also read each site's robots.txt and checked it against 13 AI user agents, and looked for llms.txt, WebMCP and Content-Signal.

The checks

Pillar	What the script checked
Reach	HTTP status, block and challenge pages, robots.txt rules for shoppers' agents, AI search bots and training crawlers
Read	Share of page text present without JavaScript, and whether the product name and price appear in the server HTML
Understand	Product, Offer, return policy, delivery and Organization schema, in the server HTML and after JavaScript
Act	Named buttons, labelled form fields, a clear add to cart control, and full-screen pop-ups on arrival
Standards	llms.txt, WebMCP signals and Content-Signal (reported, not scored)

The score

Reach 25 points, Read 25, Understand 20, Act 20. Trust (10 points in the full framework) needs a human review of policies, so it was not scored here. We scaled the 90 measured points to 100. A site that blocked both the plain fetch and the browser scored zero on Read, Understand and Act, because an agent would see nothing.

Quality checks

We reviewed the page data for every site and checked screenshots wherever a result looked unusual. On 8 sites the script captured a listing page instead of a single product. We excluded those 8 from product-level results and from scoring, rather than mark them down for our error. Our pop-up check also flagged invisible full-screen page elements on two sites; we counted only pop-ups that showed visible text.

The real agent test (Part 2)

On 27 September 2026, three AI models each received the same three tasks for every site: Claude (claude-haiku-4-5), GPT (gpt-5-mini) and Gemini (gemini-3.8-flash). Each ran through its company's API with its own web search and page-reading tools switched on, and was told to use only the official website and to say so if it could not. That made 450 tasks in total.

We counted a task as complete only when the agent said it found the answer on the official site, cited an official page, and gave a clear value. We then opened every official product link the agents gave, from our own servers, to check that the page existed and showed the stated price. The test cost about USD 25 in API fees.

These are the faster, lower-cost models from each company, used through their APIs. Consumer agents such as ChatGPT agent mode or Claude in Chrome run full browsers with larger models, and may do better on JavaScript-heavy pages.

Limitations, and what comes next

- **Part 1** measures what an unknown automated visitor gets.
- **Part 2** measures what three major AI agents get. Both are snapshots, and both have limits worth knowing.

What to keep in mind

- **Cloud location.** We ran the Part 1 crawl from Google Cloud servers outside India, as an unknown bot. Part 2 confirmed that many of those blocks do not apply to the major AI agents, so Part 1's access results describe unknown traffic, not ChatGPT, Claude or Gemini.
- **One visit per page.** Sites change often, and a block or pop-up can depend on time of day, traffic or A/B tests. These results are a snapshot of 26 September 2026.
- **One product page per site.** A site's other product templates may do better or worse.
- **Automated checks have limits.** Our checks can miss a well-hidden add to cart button or misread a complex page. Where results looked odd, we checked screenshots by hand.
- **Trust not scored.** Policy clarity and consistency need human review, so the Trust pillar is not in these scores.
- **Standards not scored.** WebMCP and llms.txt are reported but not scored, because their benefit is not yet proven.
- **Agent answers are not all checked by hand.** We verified product links and prices automatically, and compared return windows across agents. We did not hand-check every delivery answer against the site.
- **Model choice.** Part 2 used each company's faster, lower-cost model. Larger models and full browser agents may do better.

What comes next

The next edition will rebuild the Reach pillar around real agent access, fix the product pages for the 8 sites we could not score, add a hand-checked accuracy review of delivery answers, and test full browser agents such as ChatGPT agent mode and Claude in Chrome on a sample of sites.

Sources

All study data comes from FTA's own crawl of 26 September 2026. Outside sources used for context:

- [The rise of the AI crawler](#), Vercel and MERJ, December 2024: AI crawlers and JavaScript, and where they run from
- [WebMCP for Liquid storefronts and Hydrogen](#), Shopify developer changelog, August 2026
- [Overview of OpenAI crawlers](#), OpenAI: the difference between training, search and user-triggered agents
- [Join the WebMCP origin trial](#), Chrome for Developers, June 2026



//fta.

Stay tuned for more FTA Global research on how AI agents
are reshaping search, discovery and visibility.



@FTA-Global



ftaglobal



www.ftaglobal.in