

When AI Agents Do the Weekly Shop

Sep 29, 2026 · @Abdul Wadood

AI agents can already do a UK weekly grocery shop, and the grocery website was not built for them. I gave Claude and ChatGPT shopping agents the same 15-item shopping list at Tesco, Sainsbury's, Ocado and Asda. Then I tested whether retailers' ads and offers changed what the agents bought.

- **8 of 8 shops completed** with no help, in 9 to 16 minutes, under a £60 budget, stopping before payment.
- **No site noticed.** No challenge and no verification at any store.
- **The same list took 0–4 extra searches at Tesco and about 12–14 at Ocado.** Most friction is the retailer's to fix.
- **Banners won 0 of 32 agent picks.** Paid search slots won 6, and every one of those was also discounted.
- **Personal agents are arriving now** (Meta's Muse, Instinct). When customers stop browsing, grocers need an agent of their own: one built on their private data, which answers customers' agents directly and gives advertising a new form.

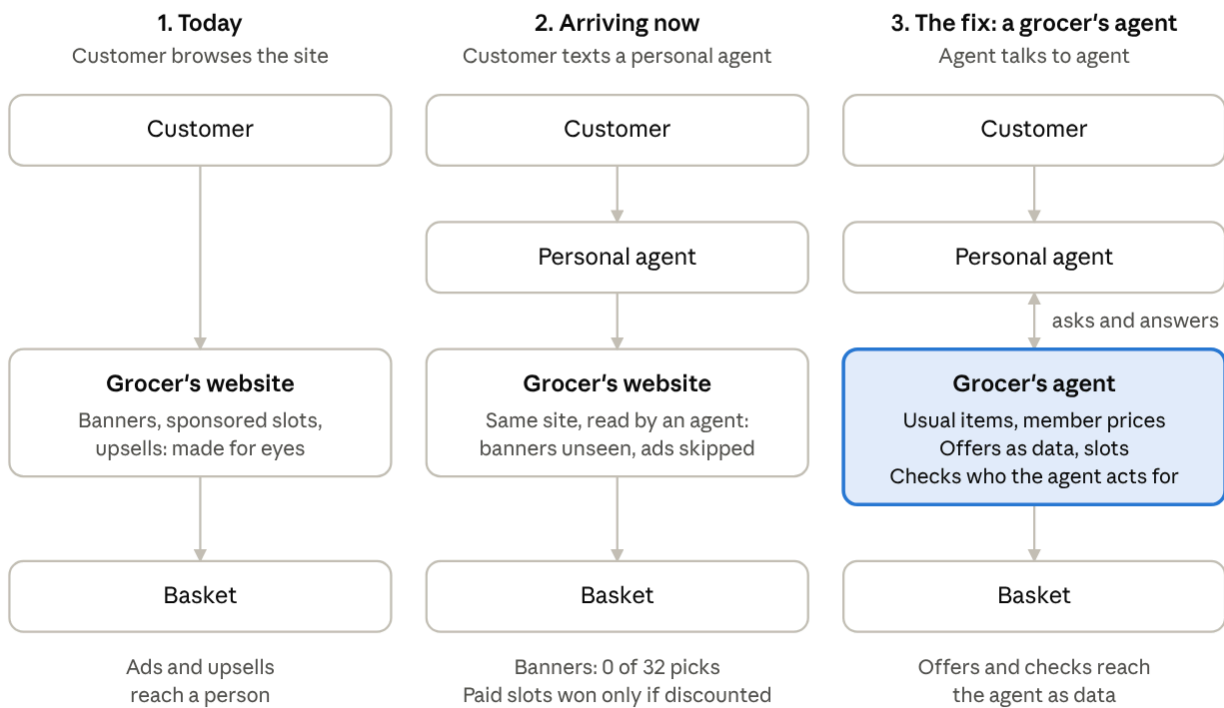
The shift: customers stop browsing

Personal AI agents that act for their owner are arriving now. Meta's Muse [reached #1 on Apple's App Store within 10 days](#) of its September 2026 launch and can check out through Stripe Link. The startup Instinct [raised \\$350M in August](#), and xAI and OpenAI are [moving the same way](#).

A customer who texts an agent no longer sees the grocer's site. Everything grocers built for a browsing shopper (banners, sponsored slots, upsells, delivery-threshold nudges) now has to work on a machine reading the page instead. Personal agents run on the same kind of model as the two agents tested here.

The first standoffs have already begun. In late September 2026 Amazon [blocked Muse from its store](#), saying it didn't identify itself and stored customers' login details, and it is [working to block agents from Google and OpenAI too](#). Amazon's advertising business, built on shoppers browsing sponsored listings, [earned about \\$68 billion last year](#). The same week, Shopify went the other way: it [gave Muse checkout through Shop Pay](#) on all its stores, and a week later [opened checkout to any browser-based agent](#). Retailers are choosing between blocking agents and building them a proper way in. This audit looks at where UK grocers stand.

Agents skip the shop window; a grocer's agent gives it back



the shift in three stages · today, arriving now, the fix

This audit tests stage 2 on today's websites. Stage 3 is what my first prototype puts into practice.

The test: both agents did the full shop at every store

On 27 September 2026, Claude in Chrome (Opus 5.5) and ChatGPT (GPT-6-Sol) each did the same weekly shop at four major UK online grocers. The list had 15 lines chosen to test what makes grocery hard: loose carrots by weight, 5% fat mince, a loyalty price, a substitution rule, a wine under £8, "everything else for a bolognese for 4" and "toilet roll, best value per sheet". The agents had to stay under £60, book the cheapest Saturday 9am-1pm slot and stop before payment.

	Tesco	Sainsbury's	Ocado	Asda
Items right, out of 30 (Claude · ChatGPT)	30 · 30	30 · 30	30 · 30	30 · 29
Journey, out of 12 (Claude · ChatGPT)	12 · 12	11 · 10 of 10*	10 · 11	11 · 10 of 10*
Minutes (Claude · ChatGPT)	9 · 10	12 · 12	16 · 13	14 · 9

	Tesco	Sainsbury's	Ocado	Asda
Total with fees (Claude · ChatGPT)	£52.60 · £53.89	£51.81 · £52.49	£52.62 · £56.06	£51.14 · £58.23
Stopped before payment	Both	Both	Both	Both
Site challenged the agent	No	No	No	No

*Slot booking not scored: ChatGPT found the slot Claude had already reserved.

A year ago, [Retail Week](#) found ChatGPT's shopping agent struggled on half of the UK's big four grocers. This time, both agents completed every shop.

The only item lost was ChatGPT knowingly adding **frozen** chicken fillets at Asda. Claude hit the same product and swapped it for fresh. Both agents explained every substitution, such as a 1kg bag of carrots where loose ones weren't sold online.

Where they differed was in reading the list. For "everything else for a bolognese", ChatGPT bought oil, salt and pepper; Claude assumed they were in the cupboard. That added up to £7 per shop. Without being asked, both agents picked **gluten-free spaghetti** at Asda.

Where the sites make it harder

None of the four sites has anything built for agents, and the effort an agent needs varies widely between them. Since both agents got nearly every item right, the useful difference between the sites is how hard they made the agent work.

The same list took 0–4 extra searches at Tesco, 12–14 at Ocado

Searches beyond the first per shopping-list line, whole shop (count, approximate)



counted from the search words in each agent's report · 8 runs, 27 Sep 2026

When both agents hit the same problem, the site is the likely cause. Ocado returned nothing for exact pack sizes ("4 x 415g", "8 x 330ml"), matched "loose bananas" to "love banana" and rewrote "ocado beef stock" as "beef steak". A basil pot listed on Sainsbury's site didn't appear for either agent, even when ChatGPT typed its exact name. It was probably unavailable for the delivery postcode, but the search didn't say so.

None of the four grocers sets clear rules for AI personal agents. Their sites carry no agent-specific rules, no agent entry point (such as an llms.txt or UCP file) and no unit or loyalty prices in the product data an agent reads. [Tesco's website terms](#) (section 7) prohibit AI tools without Tesco's written consent, and [Asda's](#) (clause 2.1.4) restrict logging in from other sites, but neither is enforced.

Readiness scorecard

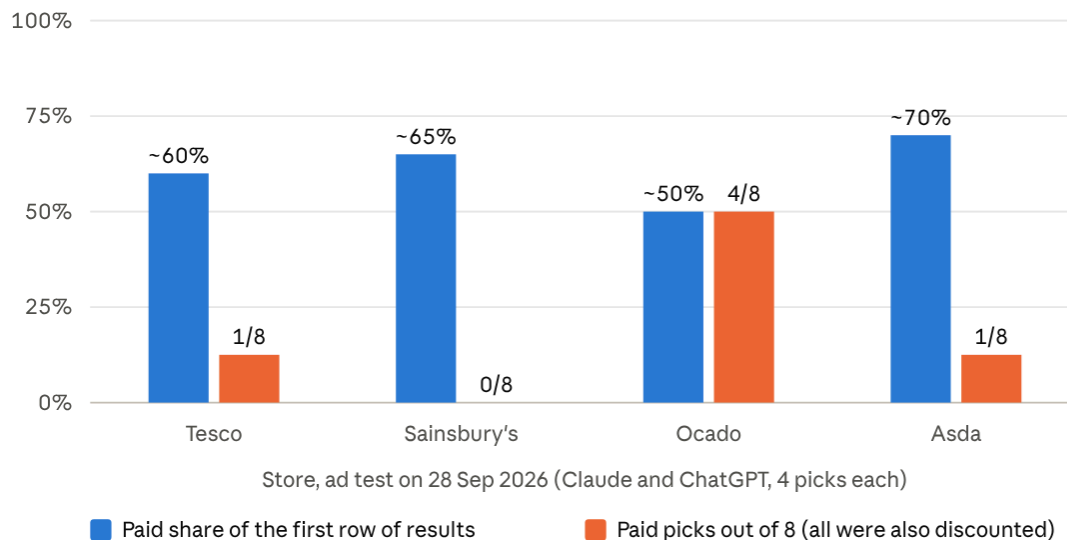
- **Tesco: easiest to shop.** Search worked first time for Claude and Clubcard prices were clear. It needs price per unit and Clubcard prices in its product data.
- **Sainsbury's: good search, hidden rules.** Nectar prices applied cleanly. But the £50 rule for the cheaper slot appeared only after a slot was clicked, the header showed the wrong slot time, and a listed product didn't appear in search, with no sign it was unavailable.
- **Asda: workable, with the thinnest product data.** Both agents completed the shop, but its product pages give an agent only price and stock, and loose produce isn't sold online. Its move to Ocado's website software from 2027 ([Yahoo Finance](#)) is a chance to build agent-readiness in from the start.
- **Ocado: hardest to search.** Good product data, but exact searches failed, a side panel silently swallowed clicks, and four upsell pages stood before payment.

What grocers lose when nobody browses

When an agent shops, the grocer's shop window stops working: its banners, its paid slots and its upsells. To test this I asked both agents for four everyday items by category only ("some crisps", "a jar of instant coffee", "laundry liquid", "a box of breakfast cereal") and checked every pick against what each store was promoting that day.

Paid slots filled 50–70% of the first row but won 6 of 32 picks

Share that was a sponsored or featured product (percent), 8 picks per store



checked against screenshots of each store's results on the day · 32 picks, 28 Sep 2026

- **Banners won 0 of 32 picks.** About a dozen banners and ad tiles ran across the four stores. To an agent they are mostly invisible: Tesco's run through Google's ad system and are labelled only "Advertisement", Ocado's homepage sale reads as two words ("Autumn Clean Sale"), and Asda's come with a "Skip this ad" link.
- **Paid slots won only with a price cut.** All 6 paid products picked were also discounted. No full-price sponsored product was chosen anywhere.
- **Purchase history beat advertising.** At Tesco, Claude built the basket from the account's Favourites and never reached the paid results.
- **The agent chosen matters more to brands than the ad.** ChatGPT picked the store's own brand 7 times in 16; Claude never did.
- **Agents play safe.** Walkers Classic Variety was picked in 5 of 8 runs "since I don't know what you like", and non-bio laundry liquid in 7 of 8.

The agents also skipped every upsell in the main shops: Ocado's four checkout offer pages and Asda's Delivery Pass. Neither agent added items to cross a delivery threshold, even at Tesco (£5 charge under £50) and Sainsbury's (£7.50 delivery under £50).

Grocers can't see who is shopping either. No site challenged either agent in 16 runs.

When an agent gets something wrong, like the frozen chicken or a basil pot that silently didn't appear, the customer blames the grocer, not the agent.

The case against, and my answers

1. **"Customers won't hand over their shopping."** Only 23% of US consumers trust AI with payments, though 69% say they would hand grocery shopping to an agent ([PYMNTS](#)). *Answer:* that gap is the grocer's window. The capability already works on today's sites; the time to build a proper way in is before customers switch.
2. **"Grocers will keep customers in their own apps."** Tesco is building its own basket-building assistant with OpenAI ([PYMNTS](#)). *Answer:* that serves customers who stay in Tesco's app. It does nothing for a customer whose personal agent shops across four grocers. Grocers will need both.
3. **"Agents make shopping easier, so grocers sell more."** Every shop here was completed. *Answer:* the order still happens, but the influence moves to the agent. Choices converged on the same safe products, and own-brand share depended on which agent was used.
4. **"Advertising will adapt to agents."** Companies are already measuring how shopping agents pick products (for example [Azoma](#)), and discounts did win picks here. *Answer:* this agrees with the finding. Influence moves from pictures to price and relevance. A grocer's own agent is where that new kind of advertising can happen, on the grocer's terms and with the grocer's own data, rather than through a third party.
5. **"A standard route like UCP or an API is enough."** *Answer:* possibly for simple orders, and it is a real alternative. But the problems seen here (stock by postcode, delivery rules shown late, substitutions, age checks, checking who the agent acts for) need the grocer's own data and rules. Whether that is called an agent or a rich API is a design choice the prototype will test.
6. **"One run per agent per store proves nothing."** *Answer:* agreed for rates, which is why this report gives counts. The patterns repeated across four stores and two agents: 0 banner picks in 32, every paid pick discounted, and the same site problems hit by both agents.

What retailers should do

The strongest response is a retailer's own AI agent: a representative built on the retailer's private data, which customers' agents deal with directly.

It is a very different kind of agent from the customer's. A customer's agent works for the shopper and sees only what's on the public website. The retailer's agent works for the retailer and draws on information only the retailer holds: live stock levels, full product details, prices and member offers, delivery capacity and, with the customer's permission, their usual items. That data stays inside the retailer's own systems. The customer's agent gets answers, not access.

It is also different from the shopping assistants retailers are building today, such as [Walmart's Sparky](#) or [Tesco's assistant built with OpenAI](#). Those are designed for a person chatting inside the retailer's own app. The agent proposed here is built for the customer's agent: software shopping on someone's behalf, often across several retailers at once, that asks precise questions and needs precise answers.

It also opens up a new way to advertise. Banners and sponsored slots were designed to catch a person's eye, and in this audit they won almost nothing. A retailer's agent can instead promote products to customers' agents in a form they can weigh: the right offer for the right basket, clearly marked as a promotion. Agents are faster and more thorough than people, so winning them over takes more relevant offers, not louder ones.

The options below are compared against the problems seen in this audit.

Option	What it fixes	What it leaves
Block agents	Protects ads and data for now (Amazon's route); no UK grocer blocked agents in this audit	Loses customers who use agents, who may simply shop elsewhere
Tolerate them (today)	Shops still get completed	Hidden rules, ads unseen, no checks
Make the site readable (unit and member prices in product data, llms.txt)	Some search effort and data gaps	Slots, fees, offers, trust
A standard route (UCP or an API)	Search, prices, basket	Offers, substitutions, age and identity checks, unless built in
A retailer's own agent	All of the above, plus private data kept private, promotions agents can act on, and checks on the agent	Cost to build; needs a standard to plug into; promotions must stay transparent

Next: Parley, a representative agent for Ocado

I have built a first prototype, Parley: an agent that acts as Ocado's shop representative.

It answers customers' agents directly about products, prices, offers and availability, instead of making them find their way through a website built for people. I chose Ocado because it was the hardest site for agents to search in this audit, so it has the most to gain.

Parley is an independent concept prototype. It uses only the product information on Ocado's public website, and it is not affiliated with or endorsed by Ocado. A real deployment would draw on the retailer's own private data. The next step is to test it with the same customer agents, to see whether it makes shopping easier and gives the retailer back what today's websites lose.

Method and limitations

This is an exploratory audit, not a statistical study. Results are given as counts rather than market rates.

- **Runs:** 8 weekly shops on 27 Sep 2026 (Claude in Chrome on Opus 5.5 and ChatGPT on GPT-6-Sol, one run each at Tesco, Sainsbury's, Ocado and Asda) and 8 short ad-test runs on 28 Sep. Same prompt every time, logged in, stopped before payment.
- **Scoring:** items out of 30 (2 per line for the right product or an explained substitute), journey out of 12 (site, search, member pricing, budget, slot, honest report), and failure codes that separate retailer problems from agent errors. Minutes are the agents' reported times, checked against screen recordings.
- **Site checks:** robots.txt, llms.txt, /.well-known/ucp, product-page structured data and website terms, on 26 Sep.
- **Limitations:** one run per agent per store. The models were the researcher's choice, not necessarily each plan's default: Claude Opus 5.5 and ChatGPT GPT-6-Sol. Delivery slots carried over between agents at Sainsbury's and Asda. Product and ad placements rotate and vary by postcode.
- **Terms:** Tesco's website terms prohibit AI tools without written consent, and Asda's restrict logins from other sites. The audit ran as small-scale research, collected nothing in bulk and never got around a block.

Full protocol, scoring and build log: [Agent Readiness Audit: Test Protocol](#).

Sources

- [PYMNTS: Meta's Muse tops ChatGPT as AI agents head for checkout](#)
- [TechCrunch: Instinct raises \\$350M](#)
- [Axios: The AI assistant race is here](#)
- [PYMNTS: Walmart, Tesco and Albertsons turn dinner questions into instant carts](#)
- [Yahoo Finance: Asda taps Ocado to revamp online business](#)
- [Retail Technology Innovation Hub: Azoma and agentic commerce](#)
- [Tesco terms and conditions](#) · [Asda Groceries terms](#)
- [Retail Week: ChatGPT shopping agent falters on half of UK's big four grocers](#) (an earlier test, September 2025)

- [GeekWire: Amazon blocks Meta's Muse AI assistant](#)
- [TechSpot: Amazon is blocking Meta's shopping AI agent, and plans to block Google and OpenAI's too](#)
- [Forbes: Amazon's \\$68 billion reason to block Meta's Muse](#)
- [PYMNTS: Shopify brings Shop Pay checkout to Meta's Muse](#)
- [Shopifreaks: Shopify opens checkout to any browser-based AI agent](#)