The Dual Frontier
A Retailer's Framework for Agentic Commerce

MOLOCO

Introduction

Most retailers' first instinct for onsite AI is an LLM powered chat widget. It's understandable, but it centers on the technology, and not the shopper. Shoppers are unlikely to abandon twenty years of search-and-browse behavior just because a chat bubble appeared in the corner. The harder, more valuable work is starting with shoppers and determining how and when AI can benefit their shopping experience. That's a design problem specific to each retailer's categories, customers, and friction points, and it is solved through iterative exploration, not just adding a chat widget and calling it a day.

Offsite, the infrastructure for agent-driven transactions is forming slower than last September's announcements implied. According to The Information, Away, one of the brands Shopify showcased when OpenAI launched its in-app checkout, still isn't purchasable inside ChatGPT. A DTC brand with roughly $300 million in annual revenue applied for the commerce beta and was told the rollout was proceeding slowly. The bottleneck isn't payments, which work; it's that a merchant's catalog requires hands-on data optimizing before agents can recommend inventory reliably.

These struggles share a root, and the patterns we see in client conversations cluster around four tendencies:

01 **BOTH** / Treating catalog as a checkbox instead of strategy.

Sparse or stale catalog data doesn't just lower conversion onsite; it gives agents less to work with. This no longer just lowers conversion; in the agentic world, it risks exclusion from consideration entirely.

02 **ONSITE** / Leaving first-party data unused

Returns, support tickets, post-purchase failure patterns are where category advantage lives. LLMs will keep improving at general reasoning; they're unlikely to independently learn which of your SKUs actually work for which use cases.

03 **ONSITE** / Chasing agent traffic instead of earning direct habits.

Direct sessions protect margin, generate proprietary data, and don't depend on a platform's ranking algorithm. The goal is building onsite experiences good enough that shoppers skip agents entirely; when they don't, it's ceding the transaction to the LLM provider.

04 **BOTH** / Waiting for maturity instead of learning while the market forms.

By the time agent traffic reaches 10-15% of sessions, retailers who started early will have run dozens of learning cycles on what predicts conversion, what earns trust, and what to protect. Those defaults will be expensive to change.

The best shopping assistance has always been human:
the ex-plumber in the plumbing aisle who asked about your project and level of expertise before recommending parts, the audio specialist who learned about your room before suggesting speakers. That expertise was expensive to staff, so it got cut.
The technology to restore it at scale without the headcount now exists.

This guide provides a framework for winning in agentic commerce across both frontiers.

Part 1 defines the onsite/offsite split and why they're connected.

Part 2 covers building onsite advantage, using ambient intelligence and outcome data to supplement what external agents can offer.

Part 3 addresses how to get discovered by agentic platforms (GEO) and what to protect once you are.

Part 4 reframes retail media economics as agent-driven shopping compresses the funnel.

Part 5 diagnoses where the work typically stalls and how to sequence it.

Appendix B has a 15-minute diagnostic.
It'll tell you whether to start now or fix foundations first.

1 https://www.theinformation.com/articles/openais-shopping-ambitions-hit-messy-data-reality
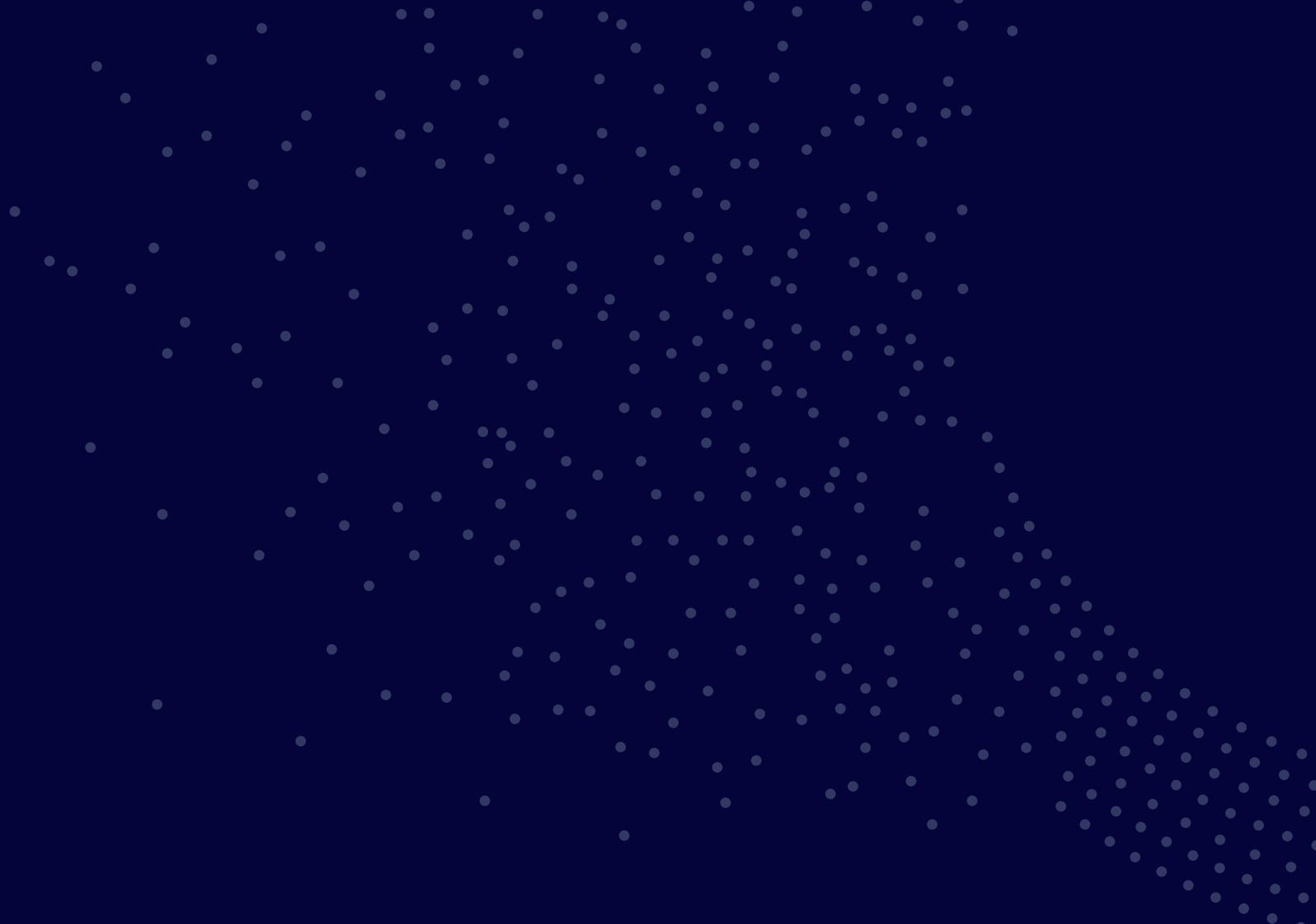
Part 1:
The Expectation Gap and Dual Frontier

A shopper asks ChatGPT for standing desk recommendations, gets clear on weight capacity, cable management, and height range for their setup, then lands on a PDP with a spec sheet and 47 reviews to parse. That gap between the conversation they just had, and the experience offered is where they're lost.

This is the onsite problem: shoppers are forming expectations about what intelligent shopping feels like, and most retail sites are not meeting them. Not because the technology is missing, but because it's implemented wrong...a chat widget that waits to be discovered instead of surfacing when behavior signals confusion.

The offsite problem is different but related. A growing share of shopping journeys will start in agent surfaces rather than search queries. ChatGPT, Gemini, Perplexity, and whatever comes next will shape consideration sets before shoppers reach any retailer's site, and increasingly close the transaction without them reaching it at all. Models can reason over unstructured content, but they won't infer claims for you that could break trust in them. If your catalog requires guessing on allergens, compatibility, or fit, agents will recommend someone else.

These problems share a root: frontier models are getting good enough at product reasoning that shoppers will use them.

Some retailers will benefit (at least initially), others will get routed around. This splits Agentic commerce across two battlegrounds:

01 Onsite (your properties):

Where you control the experience and hold outcome data no one else has. Where you can build habits that make external agents less necessary. This is the offense: earning direct relationships through intelligence that frontier agents can't replicate.

02 Offsite (agent ecosystems):

Where consideration sets form and where you need to be discoverable without giving away your differentiation. Open protocols are arriving fast: Google's UCP is already live, OpenAI's commerce integration ACP is onboarding merchants now. These will standardize transactions but not whether agents recommend you in the first place. What you can control: catalog depth, outcome signals, and what you share versus protect is covered in Parts 2 and 3. The goal is staying in the game without becoming a commodity.

Retailers who treat onsite and offsite as competing priorities will fail at both. Retailers who see them as two expressions of the same capability will compound advantages in both directions.

AGENT ECOSYSTEM

ONSITE AGENTIC EXPERIENCES
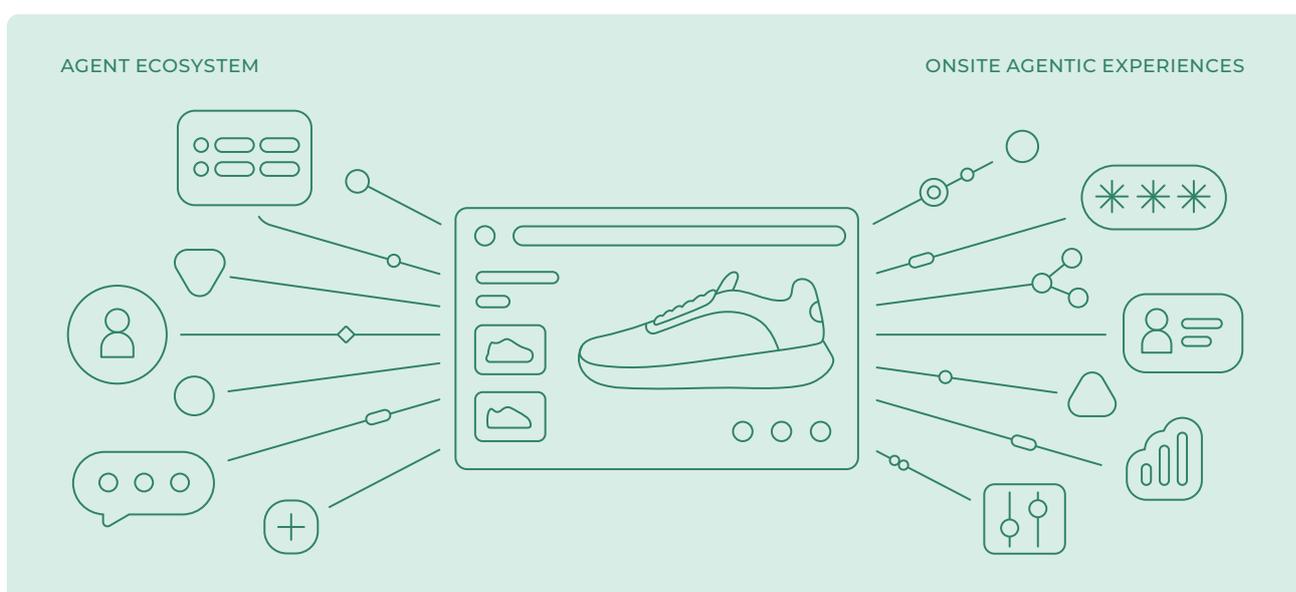
Part 2:
The Onsite Advantage

GPT-6 will be smarter than GPT-5. It still won't know that your customers return the slim-fit oxfords from Brand A at three times the rate of the classic-fit.

Why Retailers Should Win Onsite

Frontier models won't have all the retailers outcome data. They won't know which of 47 mattress SKUs actually works for side sleepers with back pain based on return rates. They can't tell shoppers that customers who buy the mid-tier air purifier for rooms over 400 square feet return it at twice the rate of the premium, because that pattern lives in your returns data, not in any spec sheet. They won't discover that furniture shoppers care about assembly time while ignoring weight capacity because they've never seen how those customers actually behave.

Better chat interfaces won't save anyone; faster responses won't either. Public reviews and ratings give frontier models a partial picture, but the fuller one lives in data only retailers hold: return reasons by SKU attribute, support ticket patterns that surface defects before they're public, behavioral signals showing which product comparisons actually lead to kept purchases. That accumulated knowledge of what matters in specific categories is what no outside model can synthesize. Onsite agent interactions add a new layer of signal: when shoppers share budget, timeline, and use-case context directly, that's intent, data richer than any clickstream and it's now yours.

Agent platforms will eventually know more about individual shoppers across their full purchase history; retailers will know more about what actually works in their categories. Depth beats breadth when the question is specific. And as protocols mature, the onsite agent that resolves a shopper's uncertainty and the offsite agent that sent them will share context, making the outcome data you're capturing now more valuable, not less.

Ambient Intelligence, Not Just Chatbots

Most retailers default to chat widgets because that's what "AI assistance" looks like. That is the wrong model.

Shoppers don't abandon twenty years of search-and-browse muscle memory just because a chat bubble appeared in a corner. When assistance requires the shopper to seek it out, most won't, and teams might conclude the demand isn't there. This is the wrong lesson from the right observation.

The better model is ambient intelligence: assistance that surfaces at moments of friction and disappears when not needed. The interface is still conversational; what changes is when it appears. Silent by default, helpful when behavior signals uncertainty.

Google figured this out with AI Overviews. They embedded intelligence directly into search results. Users didn't have to change behavior, the experience just got smarter. That's ambient. It's already generating billions of monthly chats.

A good store clerk doesn't wait for you to ask. They see you standing in the aisle, holding two boxes, eyes flicking between spec sheets and they walk over with exactly the context you need. Ambient intelligence works the same way. A shopper lands on a product detail page for a monitor arm, scrolls down, scrolls back up, toggles between variants, pauses on the specs. The system recognizes the pattern and surfaces a small decision window: "This arm works with desks 0.5 to 2 inches thick. Your monitor up to 27 inches is supported. Most customers install in 10 minutes." They can go deeper if they want, or add to cart and move on.

The trigger signals are readable: dwell time on specs, toggling between variants, pogo-sticking between product pages, scroll patterns that suggest confusion. These are signals only the retailer sees; no offsite agent has access to how a shopper is behaving on your site. The triggers differ by category, but the principle is the same: surface help that matches how they're already thinking, not help that interrupts it. Albertsons found that when assistance is contextually right, even in grocery where shoppers resist anything that slows a routine trip, their Ask AI capability delivers a 10% increase in basket size. The triggering logic is something you test and learn over time.

But the harder problem isn't when to surface, it's what to say when you do. Outcome data tells you 'runs large, size down.' User context tells you 'you've bought medium in similar brands; this one runs small.' Both resolve uncertainty a sizing chart can't.

Four Moments Where Intelligence Creates Value

Not every moment in a shopping journey benefits from intervention.
Knowing where to surface help and where to stay silent is itself a form of specialization.

01 Explore (vague intent, needs structure):

A shopper arrives with "running shoes for trails" but needs help narrowing. The choice is between surfacing meaningful starting points based on what actually differentiates options or dumping 200 results sorted by popularity. The intervention here is a guided entry: "Rocky terrain or groomed paths? Under or over 5 miles?" Two questions that cut the consideration set by 80%.

02 Evaluate (clear options, needs comparison clarity):

They've identified 3-4 candidates. The choice is between making tradeoffs explicit or forcing them to open four tabs and decode spec sheets. The intervention is comparison synthesis: "Option A has better grip but runs warm. Option B drains faster but fits narrow. Option C is the budget pick with no major tradeoffs." This is where outcome data shines. The retailer knows Option A triggers "runs hot" returns; a frontier agent doesn't.

03 Commit (ready to buy, last hesitations):

Near Add to Cart with unresolved questions. Will this work with their existing setup? Is this the right size? The intervention is uncertainty resolution: "Fits curved monitors up to 32 inches" or "Runs large, most customers size down." They either convert or leave to validate elsewhere.

04 Complete (in cart, checking for gaps):

Uncertain about completeness. Did they forget something? Will they regret not adding the accessory? The intervention is domain-informed completion: "Most customers add the wall anchors for this shelf" beats "frequently bought together" because it's based on what actually gets returned with "fell off wall" in the notes.

The Goal Is Habit Formation

Efficiency is measurable in months. Habit formation takes longer but matters more.

This is the bet: when a site resolves uncertainty faster than competitors, shoppers will notice. Not consciously at first, but after three or four sessions where questions were anticipated, the behavior shifts: they will start coming directly. Direct session rates climb, branded search referrals grow, and agent-referred traffic becomes a smaller share of conversions. They skip the Google search, skip the ChatGPT query, trust that they'll find what they need.

This is the real prize: not capturing the session but earning the default. Retailers who build this don't need to worry as much about offsite agents because their shoppers still see a reason to come direct.

Part 3:
Offsite Visibility Without Being Vulnerable

A shopper asks ChatGPT for a monitor arm recommendation. If your catalog isn't structured for that query, you're not in the answer. If it is structured but your prices are stale, you're in the answer once, then skipped next time.

The goal of onsite intelligence is earning direct relationships, but not every shopper will form that habit. Some will start with ChatGPT because they're early in research, some will use Perplexity to compare across retailers, some will ask Gemini because it's integrated into their phone.

Invisibility to these agents means exclusion from the consideration set. Visibility with an inaccurate catalog means one recommendation, one disappointment, then deprioritization. Visibility and accuracy but with full differentiation exposed means becoming one of five equivalent options competing on price.

Protocols Are Plumbing

The agent commerce stack is converging on standards: ways for assistants to discover what you sell, verify what's true, and complete transactions. Google's UCP, OpenAI's ACP commerce integrations, and whatever follows will reduce integration friction. They also create a trap: teams mistake "connected" for "competitive."

Protocols tell an agent how to ask and how you respond. They don't improve the truthfulness or usability of what you return. If your catalog is sparse, inconsistent, or stale, protocols transmit that weakness at machine speed. And agents are optimized for user trust: when confidence is low, missing attributes, messy variants, unreliable availability, the agent doesn't try anyway. In the old world, poor data meant lower conversion. In the agentic world, poor data means you may never enter consideration.

Optimize for Reasoning Engines

SEO earned you a ranking.
GEO earns you a recommendation.

When a shopper asks Perplexity for a standing desk that fits a small apartment with good cable management, the engine isn't matching keywords. It's reasoning over structured attributes: dimensions, cable routing features, weight capacity, assembly complexity. If those attributes exist in your catalog as machine-readable data rather than buried in marketing copy, you're more likely to be a candidate. If the engine also finds third-party reviews, consistent pricing across sources, and recent availability updates, it gains the confidence to cite you. If not, it recommends someone else.

This is the discipline emerging as generative engine optimization: structuring content and product data so reasoning engines can understand, trust, and surface your products in AI-generated answers. The tactics differ from SEO: schema markup over keyword density, explicit use-case attributes over inferred relevance, conversational FAQ content over landing-page prose. The principle is the same: if you're not optimized for how the system decides, you're invisible to it.

The catalog work described in Part 2 serves double duty here. Structured attributes that power onsite intelligence also make your products legible to external agents. Outcome-derived insights, "runs large, size down" or "doesn't work with curved screens" become the contextual signals that give reasoning engines confidence to recommend you over competitors with sparser data. The investment compounds across both frontiers.

Signal Governance

Visibility requires giving agent platforms enough to work with. The trap is giving them everything. The same data that gets you recommended can teach platforms (and competitors) exactly what makes you valuable.

01 Share freely (table stakes):

Product specs, pricing, availability, variants, compatibility, ratings. The factual truth that gets you into consideration sets. Every competitor shares this, so opting out isn't viable.

02 Share selectively (stand out):

Curated insights derived from outcome data. The conclusion without the underlying signal. Enough for agents to recommend you with more confidence than competitors who only share tier one.

03 Protect (your learning engine):

Raw outcome streams. Return rates by attribute. Support ticket patterns. Behavioral sequences. Which interventions worked. The system that generates insights, not the insights themselves.

04 Capture (new signal):

Agent interactions surface intent data, budget, timeline, use-case context, brand preferences that clickstreams don't reveal. This feeds user profiles and ML models, improving personalization across all placements, not just agent surfaces.

Platforms will keep asking for more data "to improve recommendations." Some of that is true; richer data does enable better recommendations. Some of it is how platforms learn to need you less. The retailers who navigate this well track which signals actually improve their recommendation rates versus which ones just train the platform.
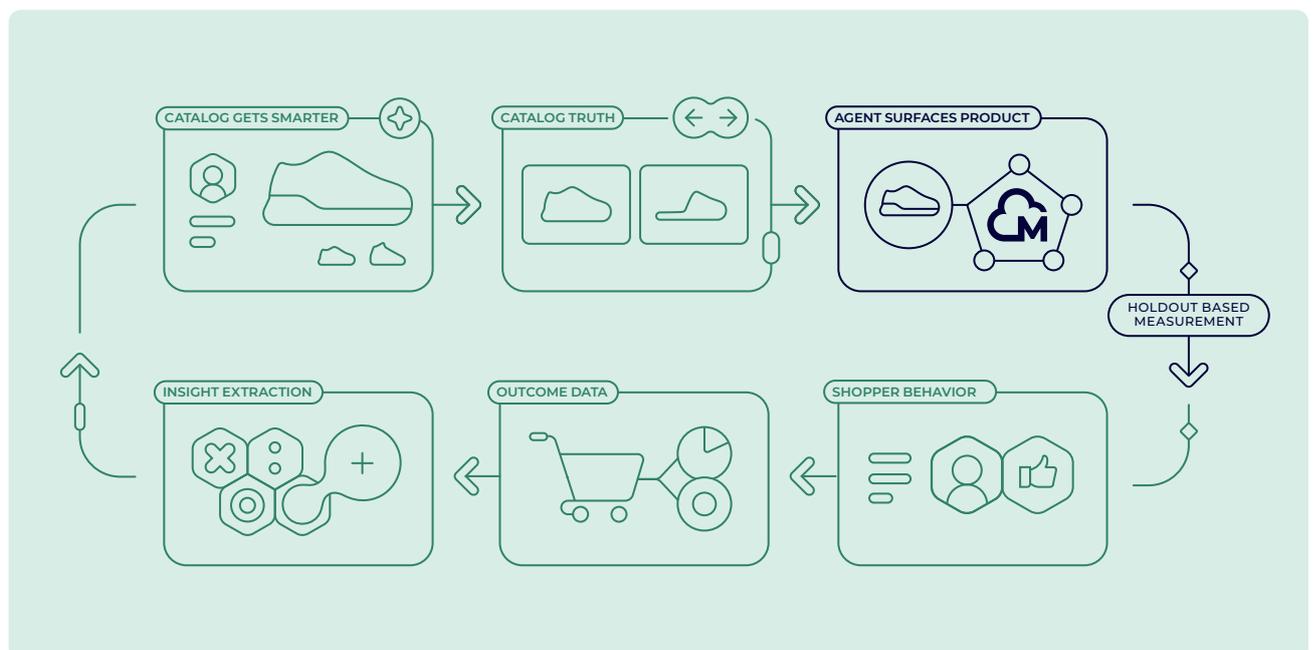
What This Looks Like by Category

Fashion - **Share freely:** sizes, materials, care instructions, variant images.
Share selectively: "slim-fit styles from Brand A run a full size small" (the insight your return data revealed). Protect the return rates by style and brand that taught you which fits actually work.

Electronics - **Share freely:** specs, compatibility lists, warranty terms.
Share selectively: "known connectivity issues with Ring Gen 2 doorbells" (the insight your support tickets revealed). Protect the ticket volume and resolution patterns that surface these issues before they're public.

Grocery - **Share freely:** ingredients, allergens, nutrition, pricing. **Share selectively:** "store-brand pasta is an accepted substitute; store-brand peanut butter is not" (the insight your fulfillment data revealed). Protect the behavioral data showing which substitutions get accepted, rejected, or trigger churn.

Furniture - **Share freely:** dimensions, materials, weight capacity. **Share selectively:** "assembly takes 45 minutes for most customers" (the insight your reviews and returns revealed). Protect, the analytical method that identifies assembly time matters more than weight capacity to your shoppers.

Part 4:
The Retail Media Yield Management Shift

Traditional e-commerce economics reward engagement: more page views, more time on site, more ad impressions during browsing. Revenue scales with attention captured.

Agent commerce inverts this. Shoppers converting in three touches instead of twelve is success; specialization resolved their uncertainty faster. But the dashboard looks worse: page views down, time on site down, display impressions down. The inversion is already underway. Retailers who adapt their economics and ad servers will capture it; retailers who protect old dashboards will blame the market.

Why the Math Still Works

The concern is obvious: fewer page views means fewer ad impressions means less RMN revenue. But this assumes the only monetizable moment is passive browsing. Amazon's Rufus offers early evidence. Customers using the assistant are 60% more likely to convert; the 250 million users who engaged in 2025 generated an incremental $10 billion in GMV. Fewer browsing sessions, more decisive ones. And decisive sessions should command higher prices: brands will pay more per outcome when the intent is already there.

The advertiser used to pay for 10 clicks at $1 each to get one purchase. Now they pay $10 for one conversion with much higher efficiency. Their economics stay the same; the path just compresses. The opportunity is selling outcomes against moments that didn't exist before, provided you can prove the outcomes are incremental.

That proof requires your ad server and agent surface sharing the same infrastructure. Separate systems mean high-intent moments you can't price and conversions you can't attribute.

Agent commerce creates new outcome opportunities across the shopping journey:

01   Discovery:

The shopper is building a consideration set. An agent surfaces your product as a recommendation. That's an outcome opportunity, a conversion you can drive and attribute, that didn't exist when discovery meant banner ads against browse traffic.

02   Comparison:

The shopper is qualifying options. An agent synthesizes tradeoffs and highlights a winner. That's an outcome opportunity, influencing the decision at the moment it's made, that didn't exist when comparison meant hoping they clicked your sponsored listing.

03   Decision:

The shopper is ready to buy but hesitating. An agent resolves uncertainty or completes the basket. That's an outcome opportunity, closing the sale at the point of friction, that didn't exist when decision-stage meant retargeting ads over days.

What This Requires

Measurement becomes non-negotiable. Without holdouts, there's no way to distinguish efficiency gains from disengagement. If sessions-to-convert drops from 12 to 3, is that because the agent surface helped or because those shoppers were going to buy anyway? Retailers who can answer this question will justify continued investment; retailers who can't will cut the wrong things.

Pricing models follow. CPM assumes volume. Agent surfaces have lower volume but higher intent. CPA or hybrid models that reward conversion become more attractive to brands who want efficiency, and more defensible for retailers who can prove incrementality.
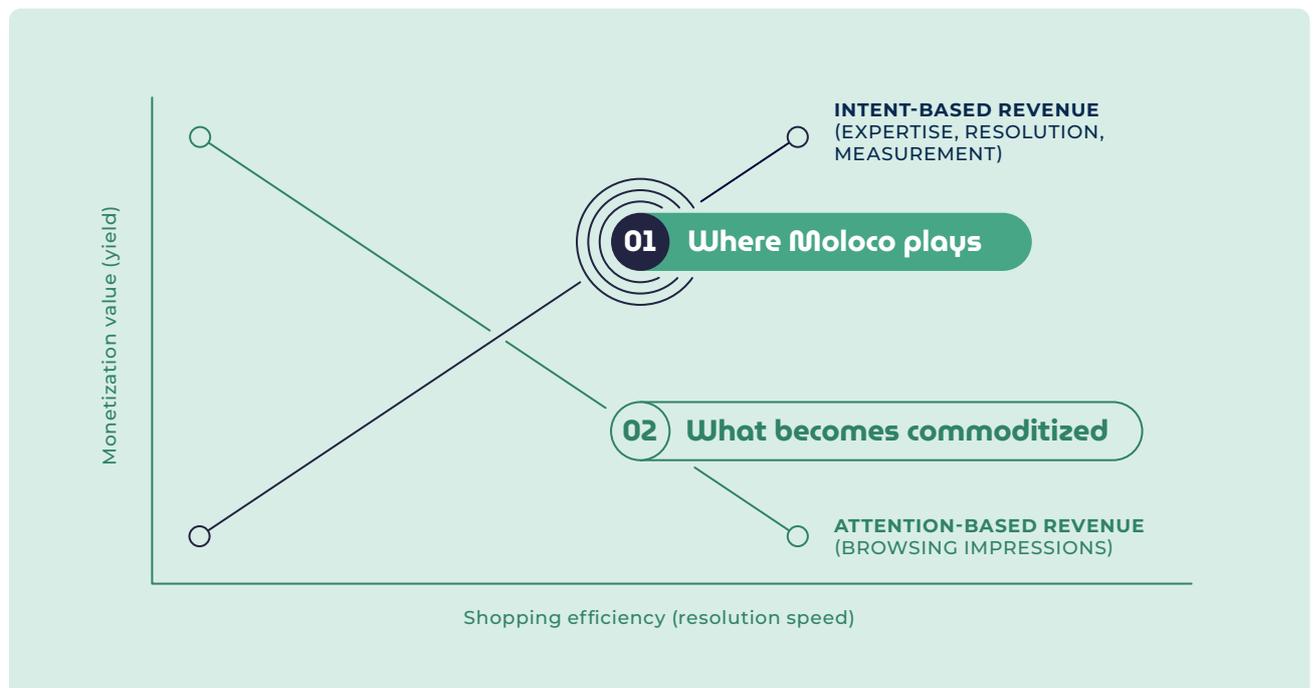
> But proving incrementality requires connected infrastructure. If your agent surfaces and ad server operate as separate systems, you create high-intent moments you can't price and conversions you can't attribute. The economic model only works when agent intelligence and monetization share the same foundation.

The Margin Trap to Avoid

Barnes & Noble had the distribution and data to dominate online books but protected brick-and-mortar retail instead because the margins looked so much better than having to figure out shipping logistics.

Agent commerce makes shopping more efficient. Some retailers will see declining engagement metrics and conclude it's destructive. They'll throttle investment to protect page views and ad impressions while competitors who adapt measurement grow margin per transaction, capturing shoppers who simply want buying to be easier.

Protecting old metrics is how you guarantee the disruption happens to you instead of through you.

Y-axis: Monetization value (yield)
X-axis: Shopping efficiency (resolution speed)

INTENT-BASED REVENUE
(EXPERTISE, RESOLUTION, MEASUREMENT)

01 Where Moloco plays

02 What becomes commoditized

ATTENTION-BASED REVENUE
(BROWSING IMPRESSIONS)

Part 5:
Where the
Work Stalls

Catalog teams own product data but don't see support tickets. UX teams build surfaces but don't see return rates. RMN teams optimize for impressions while agent success requires fewer of them. Data science builds models but waits months for engineering to deploy them.

This is the real reason most retailers will fail at agentic commerce: not missing technology but missing coordination.

Learning velocity requires these teams working off the same feedback loop with compatible incentives. A support ticket about curved-monitor compatibility should reach the catalog within days, not quarters. A/B test results should change what surfaces in weeks, not fiscal years. The retailers who fix this will compound advantages while competitors debate ownership. A partner like Moloco Commerce Media already operating your retail media can serve as connective tissue: the catalog normalization is done, the experimentation primitives exist, the outcome data is already flowing. Agent commerce becomes an extension, not a second build.

The Learning Window

Agent-referred traffic is small today, which is precisely why now is the time to start.

Learning in agent commerce can't be compressed with budget or headcount. Each cycle teaches you something: which catalog attributes actually predict conversion, which onsite interventions earn trust versus annoy, which signals to share offsite and which to protect. Target 4-6 experiments a quarter. This knowledge accumulates through iteration.

The question that predicts success: how fast can you close the loop? A shopper struggles with monitor arm compatibility. That shows up in support tickets Tuesday. How quickly does "doesn't work with curved screens over 32 inches" appear in the catalog and agent surfaces? Days? Weeks? Quarters?

That cycle time, from signal to encoded insight to live intervention, is measurable. **Retailers running 30 cycles per year will learn things retailers running 3 cycles never discover.**

Two windows are closing simultaneously. Onsite, shoppers are forming habits about what intelligent shopping feels like; a site that feels static while a competitor's anticipated needs is training shoppers to start somewhere else next time. Offsite, agents are building defaults about which retailers to trust, which catalogs to rely on, which signals to weight. By the time agent traffic reaches 10-15% of sessions, those defaults will be expensive to change.

The Cold-Start Problem

Behavioral learning requires behavior. New categories, new SKUs, and long-tail products don't have outcome data yet. This is where catalog structure earns its keep. Rich attributes, explicit compatibility relationships, and structured tradeoffs let agent surfaces perform reasonably before behavioral signals accumulate. The catalog does the work until the data catches up.

Retailers who treat catalog structure as "checkbox" and behavioral learning as "the real intelligence" will have dead zones across half their assortment. The two systems bootstrap each other.

Conclusion

This is winnable. Retailers navigated shifts before: online, mobile, social, retail media. Each one looked disruptive from the outside and felt like execution from the inside. Agentic commerce is the same. The capabilities are familiar: catalog structure, behavioral learning, signal governance, measurement. What's new is how they combine.

You have the outcome data, the customer relationships, the category knowledge. The advantage is yours to build, or yours to lose while waiting for someone else to prove the model.What's hard is coordinating the work across teams that have never shared a feedback loop. But that's an org problem, not a technology problem. Retailers who solve it will compound advantages quietly, cycle after cycle, until the gap becomes obvious to everyone else.

You have what the Chat platforms want but can't build themselves: real knowledge about what your customers actually need. The question is whether you'll use it.

Next step: Run the diagnostic in Appendix B with your team. Fifteen minutes gives you a score; if you're below 18, the learning window is shorter than your build time. We'll benchmark your results and recommend a 12-week pilot plan.

Appendix A:
Technical Glossary

ACP (Agentic Commerce Protocol):
OpenAI's emerging interface for commerce transactions within ChatGPT and other OpenAI surfaces.

Agent:
AI system that can perceive, reason, and act on a user's behalf. In commerce: systems that help shoppers discover, evaluate, and purchase products through conversation.

Ambient Intelligence:
AI assistance that appears at moments of uncertainty in existing user flows, not as a separate chat interface.

Holdout:
Control group in experiment that doesn't receive treatment; used to measure true incrementality.

Incrementality:
Revenue or conversions that wouldn't have happened without the intervention, as opposed to redistribution from other channels.

Mismatch Rate:
Percentage of agent-surfaced claims (price, availability, etc.) that don't match catalog reality. Must stay below 0.5% for trust.

PDP (Product Detail Page):
Individual product page where shoppers view specs, reviews, and make purchase decisions.

Pogo-sticking:
Behavior where shopper rapidly clicks back and forth between multiple PDPs, indicating comparison friction.

RMN (Retail Media Network):
Advertising business where retailers sell placements on their properties to brands and suppliers.

Truth Rails:
System ensuring every agent-generated claim traces back to verified catalog data, preventing hallucination.

UCP (Universal Commerce Protocol):
Google's open-source standard for agent-driven commerce, announced January 2026. Covers discovery, transaction, and monetization.

Variant:
Different versions of the same product (colors, sizes, materials). Must be structured as parent-child, not orphaned SKUs.

Appendix B:
Readiness Diagnostic

Before committing resources, assess where you stand. This diagnostic helps determine whether to start now or build foundations first, what realistic timelines look like, and whether to build in-house or partner.

Part 1: Catalog Readiness (0-12 points)

Score each 0-2 (0 = not true, 1 = partially, 2 = clearly true):

01 For top 100 SKUs across key categories, more than 70% have structured attributes describing use case, compatibility, and fit

02 Products with color/size/material options are modeled as parent-child relationships, not orphaned SKUs

03 Price and availability feeds update at least hourly with lag under 2 hours

04 Promotional pricing can be queried programmatically: discount amount, end date, trigger conditions

05 More than 30% of products have explicit relationships (compatibility, substitution, completion)

06 Mismatch rate between catalog data and reality is tracked and stays under 0.5%

Catalog Score: ___ / 12

Part 2: Technical Infrastructure (0-10 points)

Score each 0-2:

01 Infrastructure can serve sub-100ms API responses with live catalog lookups

02 Experimentation platform supports A/B tests with eligibility logging and holdout groups

03 Event instrumentation tracks user journeys from query through interaction to outcome

04 Frontend can render dynamic modules on search/PLP/PDP without full rebuilds

05 Third-party API integrations have shipped successfully in the past 12 months

Infrastructure Score: ___ / 10

Part 3: Organizational Readiness (0-5 points)

Answer 0 = no, 1 = yes:

01 Data engineering resources available for 3+ months of catalog transformation work

04 Product resources who own agent commerce as a strategic initiative, not a side project

02 Frontend/UX resources available for building and iterating on new surfaces

05 Leadership supports investment with 12+ month payback horizons

03 Data science resources who can design experiments and build optimization models

Organizational Score: ___ / 5

Part 4: Strategic Urgency (0-5 points)

Score 0 = not urgent, 1 = urgent:

01 Competitors already visible in agent responses when testing queries in your category

04 Retail media business exceeds $10M and could expand into agent surfaces

02 Category is high-consideration (appliances, furniture, electronics) where assistance matters more

05 Margin compression risk if you become interchangeable in agent shortlists

03 Agent-referred traffic (referrers like chat.openai.com) growing month-over-month

Urgency Score: ___ / 5

Interpreting Your Score

Total: ___ / 32

26-32: Foundations in place. Start learning loops immediately; catalog to behavior to insight to catalog. Partner accelerates but isn't required. 12-16 weeks to meaningful results.

18-25: Foundations first. Invest 8-12 weeks in catalog structure for top categories while piloting one onsite surface in parallel. Partner recommended for catalog work. 20-24 weeks to meaningful results.

10-17: Narrow and prove. Pick one category, one surface, one feedback loop. Prove the model before expanding. Partner strongly recommended. 24-32 weeks to results; this is a 2026-2027 initiative.

0-9: Build the case. Use the next 6-12 months to improve catalog hygiene and monitor agent traffic trends. Consider this a 2027 initiative unless urgency changes.

Appendix C:
Infrastructure Partner Profile

Building catalog infrastructure, experimentation platforms, and measurement systems in-house requires 12-18 months and specialized teams: data engineers for catalog normalization, data scientists for holdout frameworks, ML engineers for real-time decisioning, commerce specialists who understand retail constraints.

Large retailers with deep engineering teams and patient capital can develop these capabilities internally. Most can't.

What to Look For

Operates RMN substrate today.
The partner should already run catalog normalization, ad serving, and incrementality measurement for retail media. These shouldn't be roadmap items; they should be battle-tested with live retailers. Agent commerce uses the same foundations extended to new surfaces.

Extends to agent contexts without rebuilding.
Offsite protocol integration, onsite intelligence, and intent-based monetization should leverage existing substrate. If they're starting from zero on agent commerce, you're funding their R&D.

Has truth rails and holdout frameworks at scale.
Truth rails (every claim traces to verified data) and holdout-based incrementality are prerequisites for trust and economic validation.

Enables learning velocity.
You should start iterating in weeks, not waiting 12-18 months for infrastructure. The partner should help run experiments, capture behavioral signals, and feed learnings back into catalog structure.

Where Moloco Fits

Moloco Commerce Media operates catalog structuring, real-time truth, ad serving, and incrementality measurement for retail media today. Agent commerce extends that substrate:

01 Catalog normalization for retail media becomes agent-readable catalog for offsite discovery

02 Real-time decisioning for ad serving becomes ambient intelligence for onsite surfaces

03 Holdout-based incrementality for RMN becomes the measurement framework for agent surfaces

For retailers already partnering with Moloco on retail media, agent commerce is infrastructure extension. For others, it's partnering with a platform that already operates the foundation layer agent commerce requires.

Appendix D:
Open Questions

Some questions aren't settled yet by any retailer, platform, or model provider. These are the problems we're actively working on.

What's Clear

01 High-consideration categories see stronger results than low-consideration

02 Behavior-triggered assistance provides benefits always-on can't

03 Catalog structure is non-negotiable; mismatch kills trust faster than bad UI

04 Monetization works when contextually relevant, backfires when forced

What We're Working to Prove

Incrementality:
Do agent-assisted sessions create net-new revenue or redistribute existing demand? Early data is mixed. We're building holdout frameworks and longer time horizons to answer definitively.

Speed vs. depth:
Is sub-300ms response necessary even if it means simpler recommendations? We're running head-to-head tests to find the real threshold.

Attribution:
How do you credit conversions that start offsite and finish onsite? We're developing measurement approaches for journeys that cross boundaries.

What No One Knows Yet

01 Will agent ranking algorithms converge or fragment across platforms?

02 Does shopper engagement with agent surfaces sustain, or does novelty wear off?

03 Where's the trust threshold for sponsored recommendations, and is it binary or gradual?

04 Will platforms vertically integrate in ways that close the window for retailers?

Our approach:
Build capabilities that create value under multiple scenarios. The answers will emerge from live experimentation, and we intend to be generating that data.

Note on Moloco Commerce Media

The infrastructure this document describes (catalog truth, behavioral learning loops, holdout-based measurement, real-time decisioning) is what Moloco Commerce Media operates today for retail media. Agent commerce extends that foundation to new surfaces, not a rebuild from scratch.

For retailers evaluating partners: we're not building toward these capabilities. We're extending what's already live.

MOLOCO