

THE AI INSTITUTE / OPERATING INTELLIGENCE

INFLUENCE & PARTNERSHIP RADAR · 25 SEP / 2026



Who Does Your AI Shopping Agent Serve?

Before joining an AI commerce partnership, separate a successful payment from a good decision for the buyer.

Boards, C-suite and senior management

Research period: September 2026 – September 2026

First published: 25 September 2026

Current to 25 September 2026

PUBLICATION RECORD

An Institute thesis and decision framework

An AI commerce partnership should prove whose interests the agent serves before it is rewarded for completing more transactions.

WHAT LEADERS SHOULD TAKE FROM THIS PAPER

- 01 **Separate buyer benefit from merchant conversion.**
- 02 **Ask who controls product visibility and how each partner gets paid.**
- 03 **Test the whole buying journey with local prices, terms and payment options.**
- 04 **Bring an independent evaluator and affected customers into the partnership review.**

Research method

Institute synthesis of one new arXiv preprint, three original supplier announcements and established UK and US public guidance. Research methods and limitations were reviewed separately from commercial claims. The proposed review and pilot are Institute analysis, not validated interventions or legal advice. No private-chat quotation or third-party artwork is reproduced.

First published 25 September 2026 · Updated 25 September 2026 · Research period September 2026 – September 2026 · Version 1.0 · Suggested citation: The AI Institute, *Who Does Your AI Shopping Agent Serve?* (2026).

CONTENTS

Inside this paper

01

A completed purchase is not the whole result
The partnership decision

02

Commercial pressure belongs in the test
The research signal

03

Follow the money and the product list
Who shapes the journey

04

Give the buyer a seat at the review
The missing evidence

05

Measure the purchase the buyer wanted
A bounded pilot

06

Test the actual market, not a global label
Where the promise changes

M

Method and limitations

R

References

How to use this brief

Start with the decision and use the framework to challenge the current plan. Figures retain their original populations and definitions. Institute analysis is labelled, and limitations are recorded before the references.

A completed purchase is not the whole result

Before signing an AI commerce partnership, ask a simple question: would this arrangement still count as successful if the agent persuaded someone to buy the wrong product? A merchant may gain a sale, a platform may earn a fee and a payment provider may process the transaction correctly. The buyer can still be worse off.

This is a leadership issue because different teams will see different evidence. Sales sees conversion. Finance sees transaction cost. Security sees an authorised payment. Procurement sees a supplier contract. Someone must connect those views to the promise made to the customer.

The Institute's recommendation is to add a buyer-interest review before expanding purchasing authority. Define what the agent is meant to achieve for the buyer, who can influence its choice and what would justify stopping a purchase. The review should be part of the partnership agreement, not a late request to the technical team.

Commercial pressure belongs in the test

A preprint submitted on 23 September by researchers at Carnegie Mellon University and Microsoft Research examines shopping agents in controlled marketplaces. Its matched comparison reports buyer-optimal choices falling from 78.6% to 17.3% when commercial steering is enabled. The research covers nine environments; the headline comparison concerns specified model configurations, not consumers or live retail transactions.¹

These are preliminary research results, not a verdict on any named payment network or retailer. The tasks have clear objectives and fixed steering mechanisms. Real buyers may have changing preferences. The authors report improvements from additional verification, but also greater running cost and limited evidence of transfer.¹

For leaders, the useful question is whether the partnership tests buyer benefit separately from task completion. A checkout demonstration alone cannot settle that question. Nor should an early laboratory result justify a blanket ban on useful assistance.

Follow the money and the product list

On 9 September, Mastercard announced Agent Connect and expanded merchant tools, including work with Anthropic’s commerce blueprint. Its stated aims include making products discoverable and supporting authorised transactions. This is a supplier announcement about capabilities and partnerships, not independent evidence that buyers receive better recommendations.²

Visa’s 2 July announcement described live agent-led purchases at participating European merchants, with consumer authorisation and bank involvement. That is evidence of a working transaction route within the programme, not a measure of how often agents selected the best available offer.³

Anthropic’s September commerce blueprint distinguishes agents helping shoppers from agents helping merchants. That distinction should survive the contract. A seller’s assistant can legitimately explain the seller’s catalogue; it should not be presented as an unrestricted comparison service if it cannot examine alternatives.⁴

Ask each partner to state how it earns revenue, which products the agent can see, whether placement is paid and who controls the final recommendation. Commercial incentives are not proof of misconduct. They are facts the buyer needs when deciding how much authority to delegate.

EXHIBIT / FOUR VIEWS OF THE SAME PURCHASE

BUYER	AGENT	SELLER	PAYMENT
Needs met Did the choice meet the agreed budget, features and delivery needs? Institute review framework	Task completed Completion alone does not show whether better alternatives were visible. Institute review framework	Sale converted A sale alone does not show whether the buyer's priorities were preserved. Institute review framework	Authority verified An authorised transaction alone does not establish a suitable product choice. Institute review framework

Give the buyer a seat at the review

The review needs more than the platform's demonstration. Include a commercial owner who can explain the revenue model, a payment specialist who can explain authorisation and refunds, and an evaluator who is paid to find failures rather than increase sales. Bring in customers or customer representatives who can challenge what the business calls a good outcome.

An independent university group could contribute controlled tests; a merchant could supply accurate product and service information; a consumer organisation could examine whether the promise is understandable. This is a proposed partnership design, not an announced Institute relationship or endorsement.

The UK Competition and Markets Authority's evidence review explains how digital presentation can affect consumer choice. US Federal Trade Commission guidance addresses the recognisability of advertising. Neither establishes that a particular AI shopping service has broken the law. They provide established questions for a local review of how commercial influence is presented.⁵⁶

Keep the tests and reporting rights separate from the sales target. Agree who can publish an unfavourable result, who must correct misleading product information and who may halt the pilot. Without those rights, an impressive advisory panel may have little practical influence.

Measure the purchase the buyer wanted

Start with one product category and a spending limit. Use test transactions or a controlled environment, not unsuspecting customers. Write down the buyer's requirements before browsing: the total budget, delivery deadline, essential features and acceptable trade-offs. Include cases where the right answer is to ask a question or buy nothing.

Change the order and prominence of offers while keeping the buyer's requirements fixed. Compare the chosen item with the alternatives the evaluator can verify. Check the final payable price, delivery terms and return conditions. Record whether the agent asks before changing the brief.

Measure cost as well as quality. Additional checking can consume time and money; a low-value purchase may not justify an elaborate process. Compare the agent with a simpler search-and-confirm journey. Expand only when the result is worth the extra operating effort.

Do not make staff responsible for correcting choices they cannot inspect. Customer-service teams need enough of the decision record to resolve a complaint, while access and retention should be limited to what the service genuinely needs.

Test the actual market, not a global label

A European card-payment pilot does not establish coverage for every payment method or jurisdiction. The US and UK sources above address different forms of commercial presentation. Local advice is needed on the applicable consumer, advertising, privacy and payment duties; this brief does not propose one worldwide legal checklist.

For a cross-border launch, test the intended language, currency, delivery area, payment method and route for refunds. In markets where the agent sees only a small share of sellers, describe that limit. A restricted catalogue can still be useful if the customer understands what is excluded.

Approve the partnership against a clear promise: whose interests the agent serves, what it can compare and when it must ask. Then assign a business owner to the failures that cross organisational boundaries. The next useful conversation is with the people who can demonstrate those conditions, not just the people announcing the partnership.

RESEARCH RECORD

Method and limitations

Method

Institute synthesis of one new arXiv preprint, three original supplier announcements and established UK and US public guidance. Research methods and limitations were reviewed separately from commercial claims. The proposed review and pilot are Institute analysis, not validated interventions or legal advice. No private-chat quotation or third-party artwork is reproduced.

Limitations

No global agent-shopping adoption rate or independent live-market buyer-benefit estimate was established. The preprint is not peer-reviewed in this record and does not evaluate the named commerce partnerships. Regional examples do not establish legal equivalence or availability in other markets. No affiliation, guest acceptance or partnership with the Institute is implied.



Sources and limitations

Source facts link to the original publication. Institute analysis reconciles definitions and turns the material into a management proposition. Frameworks are decision aids, not claims of universal causality. Read every important statistic with its population, date, definition and caveat.

Research cutoff: 25 September 2026. Review cycle: six months, or earlier after a material change in the underlying facts, standards or policy.

REFERENCES

Evidence behind the thesis

01 Li, Epperson, Deng and Huang: CAVEAT

New v1 submitted 23 September 2026; listed 24 September. Preprint. Controlled marketplaces, not a live-commerce audit. Full methods and limitations: <https://arxiv.org/html/2609.27273v1>

02 Mastercard: Agent Connect and merchant capabilities

9 September 2026. Supplier announcement; no independent buyer-outcome or adoption claim.

03 Visa: live agent-led transactions in Europe

2 July 2026. Participating European issuers and merchants; provider account, not a global market census.

04 Anthropic: Building commerce agents with Claude

2 September 2026. Supplier blueprint. Merchant conversion statistics are not used as evidence of buyer benefit.

05 CMA: Online choice architecture evidence review

5 April 2022. UK evidence review, not a new AI rule or a finding against any named partner.

06 FTC: Native advertising guidance for businesses

December 2015 US business guidance. Background on advertising presentation, not a universal or AI-specific statute.

THE AI INSTITUTE

Make the decision.
Then measure what changes.

Research · Readiness · Capability

theai.institute