Agentonomics

AI Commerce Readiness Assessment Methodology

How AI Shopping Assistants Evaluate Your Products

Version 2.0 | January 2026

agentonomics.io

Executive Brief

Read this section in 3-5 minutes to understand why AI commerce readiness matters.

The Shift Is Happening Now

AI shopping assistants like ChatGPT and Google Gemini are rapidly becoming primary product discovery channels. Early data suggests these tools influence billions in annual purchases, with adoption accelerating in 2025-2026.

> **The Problem:** AI agents respond very differently than humans. Marketing tactics that work on human shoppers can actively hurt your AI visibility. Scarcity language that creates urgency for humans reduces AI recommendations by 30%.

What Agentonomics Measures

We quantify how AI shopping assistants actually perceive your products using three dimensions:

45%	30%	25%
Agentology How AI interprets your content	Discoverability Can AI find your products?	Commerce Ready Can AI complete purchases?

Research Foundation

Our methodology is built on findings from four peer-reviewed academic papers:

- **Filandrianos et al. (EMNLP 2025)** - Cognitive biases in LLM recommendations
- **Allouah et al. (Columbia/Wharton 2025)** - ACES framework for agent commerce
- **Bansal et al. (Microsoft Research 2025)** - Magentic Marketplace dynamics
- **Cherep et al. (MIT Media Lab 2025)** - AI vs human decision amplification

Key Research Findings

51.4%	60-100%	3-16x
Price equivalent of 0.1 star improvement	First-proposal acceptance rate	AI amplification of biases vs humans

The Opportunity: Businesses that optimize for AI commerce readiness now will have significant competitive advantages as AI shopping grows. Agentonomics helps you measure, understand, and improve your AI visibility.

Part 1: What We Measure

The Three-Category Scoring System

Agentonomics uses a research-backed scoring system that measures three distinct aspects of AI commerce readiness. Each category captures different requirements for AI shopping assistant success.

1. Agentology (45% of total score)

How AI interprets and responds to your content. Based on cognitive bias research showing AI agents have systematic preferences different from humans.

Component	Weight	What It Measures
Attribute Strength	30%	Ratings, reviews, pricing visibility - factors AI weighs heavily
Social Proof	25%	Trust signals, testimonials, review counts
Platform Signals	20%	Badges, certifications, endorsements
Pressure-Free	25%	Absence of scarcity/urgency tactics that hurt AI

2. Discoverability (30% of total score)

Whether AI systems can find and access your product information. If AI can't crawl your site, it can't recommend your products.

Component	Weight	What It Measures
AI Crawler Access	25%	robots.txt allows GPTBot, ClaudeBot, Googlebot
Server Rendering	20%	Content available without JavaScript execution
Structured Data	25%	Schema.org markup completeness
Content Visibility	15%	Content in machine-readable format (not images/PDFs)
llms.txt	15%	AI-specific instruction file at /llms.txt

3. Commerce Readiness (25% of total score)

Whether your site has the data structures and integrations needed for AI-assisted purchases.

Component	Weight	What It Measures
Feed Readiness	25%	ChatGPT Merchant Feed attributes present
Payment Integration	20%	Checkout-capable payment processors detected

Product Identifiers	20%	GTIN, UPC, EAN, MPN for cross-matching
UCP Profile	15%	Google's Universal Commerce Protocol at /.well-known/ucp
Conversational Commerce	20%	Chat widgets, AI shopping assistant integrations

Part 2: The Research Foundation

Every Agentonomics recommendation is backed by peer-reviewed research with specific, measurable effect sizes.

Bias Beware: The Impact of Cognitive Biases on LLM-Driven Product Recommendations

Filandrianos, Dimitriou, Lymperaiou, Thomas, Stamou

EMNLP 2025 (Proceedings of the Conference on Empirical Methods in NLP), pp. 22409-22438

This landmark paper tested how cognitive bias language affects AI recommendations across 6 models (LLaMA 8b/70b/405b, Claude 3.5/3.7, Mistral) with 100 runs per experiment.

-30% **Scarcity language** ("Only 3 left!") reduces recommendation rate by ~30%

-45% **Exclusivity framing** ("Members-only deal") reduces recommendations by ~45%

+42% **Social proof** ("Trusted by 10,000+ customers") increases recommendations

HURTS AI RECOMMENDATIONS

"Only 2 left in stock! Order now before it's gone!"

HELPS AI RECOMMENDATIONS

"Our most popular choice - in stock with fast shipping."

What Is Your AI Agent Buying? Evaluation, Biases, Model Dependence, & Emerging Implications

Allouah, Besbes, Figueroa, Kanoria, Kumar

Columbia Business School / Wharton, arXiv:2508.02630, August 2025

The ACES framework quantifies exactly how different AI models weight product attributes.

Factor	ChatGPT Effect	Gemini Effect	Claude Effect
Rating (per star)	**8.3x** selection odds	5.4x selection odds	6.4x selection odds
Sponsored badge	-22% selection	-23% selection	-13% selection
"Overall Pick" badge	+123% selection	**+567%** selection	+189% selection
Price sensitivity	Moderate	**Highest**	Moderate

Key Insight: A 0.1 star rating improvement provides the equivalent value of a 51.4% price reduction to ChatGPT. Small rating improvements have massive AI impact.

Magentic Marketplace: An Open-Source Environment for Studying Agentic Markets
Bansal et al.
Microsoft Research + Arizona State University, arXiv:2510.25779, October 2025

This research revealed the critical importance of being the first option presented to an AI agent.

First-Proposal Bias

Model	1st Proposal Acceptance	2nd Proposal Acceptance	3rd Proposal Acceptance
ChatGPT	**100%**	~0%	~0%
Gemini	96.7%	~3%	~0%
Claude	93.3%	~6%	~0%

> **Critical Finding:** If your product isn't in the first proposal, it essentially won't be selected. Being first matters 10-30x more than other optimizations.

A Framework for Studying AI Agent Behavior: Evidence from Consumer Choice Experiments
Cherep, Ma, Xu, Shaked, Maes, Singh
MIT Media Lab, arXiv:2509.25609, September 2025

The ABxLab framework compared AI agent decisions to human decisions across 1,500 experimental configurations.

AI Amplifies Human Biases by 3-16x

Bias Type	Human Effect	AI Agent Effect	Amplification
Order/position effects	4 pp	Up to 90 pp	**22x**
Rating sensitivity	5 pp	30-80 pp	**6-16x**
Price sensitivity	9.4 pp	15-93 pp	**1.6-10x**
Authority nudges	9.9 pp	10-60 pp	**1-6x**

pp = percentage points change in selection probability

Opportunity: Because AI amplifies biases, small improvements in AI-favorable factors yield disproportionately large returns. A 5% improvement in rating visibility might translate to 30-80% more AI recommendations.

Part 3: Emerging AI Commerce Standards

New technical standards are emerging specifically for AI commerce. Early adoption provides competitive advantage.

Google's Universal Commerce Protocol (UCP)

A standardized merchant profile format at `/.well-known/ucp` enabling direct integration with Google Shopping and Gemini.

Required Fields	Purpose
name	Business name
url	Primary website URL
contact	Customer service information
policies	Returns, shipping, privacy policies

ChatGPT Merchant Feed

OpenAI's product data format for ChatGPT Shopping requires specific Product schema.org attributes.

Required Attributes:

- name
- price
- availability
- description
- image

Important Attributes:

- gtin (GTIN/UPC/EAN)
- brand
- category
- condition

llms.txt Standard

An emerging standard for AI-specific instructions, similar to robots.txt but for LLMs. Place at `/llms.txt`.

> **Example llms.txt sections:**
>
> - **title:** Site name and purpose
> - **description:** What the site offers
> - **products:** Key product categories
> - **policies:** Returns, shipping information

- **instructions:** How AI should represent the brand

AI Crawler Landscape

Agentonomics tracks 17 AI crawlers. The most critical for shopping:

Crawler	Owner	Purpose
GPTBot	OpenAI	ChatGPT web browsing & shopping
OAI-SearchBot	OpenAI	ChatGPT search features
ClaudeBot	Anthropic	Claude web access
Googlebot	Google	Search + Gemini integration
PerplexityBot	Perplexity	Perplexity AI search

Blocking AI crawlers = invisible to AI shopping. Many sites inadvertently block GPTBot, making their products invisible to ChatGPT Shopping users.

Part 4: Model-Specific Optimization

Different AI models respond differently to the same content. Understanding these differences enables targeted optimization.

Model Market Share (B2C AI Shopping)

~70%	~25%	~5%
ChatGPT	Gemini	Claude + Others

Model-Specific Sensitivities

Factor	ChatGPT	Gemini	Claude
Rating weight	**8.3x/star**	5.4x/star	6.4x/star
Badge response	Moderate	**HIGHEST**	Moderate
Price sensitivity	Moderate	**HIGHEST**	Moderate
First-proposal acceptance	**100%**	96.7%	93.3%
Ad aversion	-22%	-23%	-13%

Optimization Priorities by Model

ChatGPT (70% market share):

- Maximize visible ratings and review counts
- Optimize page speed to be in first proposal
- Avoid sponsored/ad badges where possible

Gemini (25% market share):

- Pursue platform endorsement badges (567% boost)
- Emphasize competitive pricing
- Ensure UCP profile is present

Claude (5% market share):

- Balanced approach across all factors
- Lower ad aversion - less penalized for promotions
- Quality descriptions matter more

Key Trade-offs

Some optimizations benefit one model while slightly hurting another:

- **Heavy badge pursuit:** Huge Gemini benefit (+567%), moderate ChatGPT benefit (+123%)
- **Price reduction:** Helps Gemini most, ChatGPT prefers rating improvements
- **Promotional messaging:** Claude tolerates better than ChatGPT/Gemini

Part 5: The Agentonomics Advantage

What Makes Agentonomics Different

Capability	What It Provides
Quantified AI Visibility	Not guesswork - actual extraction tests show what AI sees
Model-Specific Scores	Know how ChatGPT, Gemini, and Claude each perceive you
Research-Backed Recommendations	Every suggestion includes effect sizes from peer-reviewed papers
Competitive Benchmarking	Compare your scores against industry peers
Actionable Implementation	Developer-ready code snippets for each fix

ROI Context

The research provides concrete value quantification:

51.4%
Price equivalent of 0.1 star improvement (to ChatGPT)

60-100%
Selection probability for first-position products

0%
Selection if AI crawlers blocked

Grade Scale

Grade	Score	Interpretation
A	70-100	Elite - optimized for AI commerce with complete schema and positive signals
B	55-69	Good - minor issues like image-based ratings or incomplete schema
C	40-54	Average - typical e-commerce with common AI visibility problems
D	25-39	Needs Work - significant gaps in AI accessibility
F	<25	Critical - AI cannot reliably understand or recommend this site

Get Started: Run your free AI commerce readiness assessment at **agentonomics.io**

Appendix: Full Research Citations

Bias Beware: The Impact of Cognitive Biases on LLM-Driven Product Recommendations

Giorgos Filandrianos, Angeliki Dimitriou, Maria Lymperaiou, Konstantinos Thomas, Giorgos Stamou

EMNLP 2025 (Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing), pp. 22409-22438, November 2025

Code: github.com/geofila/BiasBeware

What Is Your AI Agent Buying? Evaluation, Biases, Model Dependence, & Emerging Implications for Agentic E-Commerce

Amine Allouah, Omar Besbes, Josue D Figueroa, Yash Kanoria, Akshit Kumar

Columbia Business School / Wharton School, arXiv:2508.02630, August 2025

Magentic Marketplace: An Open-Source Environment for Studying Agentic Markets

Bansal et al.

Microsoft Research + Arizona State University, arXiv:2510.25779, October 2025

Code: github.com/microsoft/multi-agent-marketplace

A Framework for Studying AI Agent Behavior: Evidence from Consumer Choice Experiments

Manuel Cherep, Chengtian Ma, Abigail Xu, Maya Shaked, Pattie Maes, Nikhil Singh

MIT Media Lab, arXiv:2509.25609, September 2025

Additional Resources

- **Schema.org Product specification:** schema.org/Product
- **Google UCP documentation:** developers.google.com/commerce
- **OpenAI ChatGPT Plugins/Shopping:** platform.openai.com/docs
- **llms.txt standard:** llmstxt.org

Methodology Version History

Version	Date	Changes
2.0	January 2026	3-category scoring system (Agentology, Discoverability, Commerce)
1.0	December 2025	Initial 2-category system (Machine Readability, Agentology)

Agentonomics

agentonomics.io

Document generated January 2026