

Agentic Commerce & PCI DSS

Where the Payment Rails Actually Stand – and What it Means for PCI DSS Scope

Agentic Commerce & PCI DSS

Where the Payment Rails Actually Stand – and What it Means for PCI DSS Scope

Agenda:

1. The Hype vs. the Reality
   Why this field looks more finished than it is
2. The Maturity Curve
   Shopper agent → human-in-the-loop → fully autonomous
3. The Current Rails
   ACP, UCP, AP2, TAP, Agent Pay, MPP, x402 — still competing
4. PCI DSS Impact Today
   Why tokenization is holding scope steady
5. Where Scope Could Creep
   The architecture decisions still being made
6. What Assessors Should Do
   Questions to bring into scoping conversations

To be clear on scope: we're not addressing consumer activities. We're addressing consumer-authorized transactions handled by an agent inside a construct operated by a merchant or their third parties.

Tim Thomas

VP, Sr. Product Manager
PCI Programs, Bank of America



Chris Bucolo

Director Regulatory & Industry Affairs-Strategy Development, ZeroRisk



20 YEARS OF SECURING PAYMENT DATA

PCI Security Standards Council

Why Now?

4,700%

Surge in AI-agent traffic to U.S. retail sites, cited by Visa as the reason it built the Trusted Agent Protocol

169M+

Payments processed by Coinbase's x402 stablecoin protocol in its first year — mostly machine-to-machine, not retail

The activity is real. What's not yet real is a single, settled way for these payments to move — that's the distinction this briefing turns on.

Link to Forbes Article:

Setting Expectations: The Perception vs. The Reality

WHAT PEOPLE ASSUME

- ✗ There's a single, agreed “rail” for agentic payments already in production
- ✗ Full autonomy — agent decides and pays — is the norm today
- ✗ PCI implications are already worked out and standardized
- ✗ One protocol will “win” soon, so we can wait and adopt later

WHAT'S ACTUALLY TRUE (MID-2026)

- ✓ At least six competing protocols are live or in pilot: ACP, UCP, AP2, TAP, Agent Pay, MPP/x402
- ✓ Most production flows today are agent-as-shopper or human-confirms-checkout
- ✓ PCI SSC has issued only high-level AI principles
- ✓ Networks are hedging across card rails and stablecoin rails simultaneously

The Evolution: A Three-Stage Maturity Curve

This is the frame for everything that follows — we are not jumping straight to autonomy.

STAGE 1

Agent as Shopper

Agent researches, compares, and builds a cart. A human then reviews it and pays.

LARGELY HERE TODAY

STAGE 2

Human-in-the-Loop

Agent negotiates and initiates checkout; a human confirms the final payment decision.

EMERGING NOW

STAGE 3

Fully Autonomous

Agent decides, pays, and reconciles within pre-authorized limits — no human in the transaction loop.

EARLY / M2M ONLY

Stage 1 of 3: Agent as Shopper

The agent's job ends at the cart. A human still authenticates, authorizes, and completes the payment step through a familiar checkout.

- Agent searches, compares prices, and reads reviews across merchants
- Agent assembles a cart or shortlist and hands it back to the human
- Consumer completes checkout using an existing wallet, card, or account
- Examples: ChatGPT product discovery or Amazon Rufus recommendations



PCI footprint: unchanged. The agent never touches payment credentials — it's a discovery layer sitting in front of an existing, already-scoped checkout.

Stage 2 of 3: Human-in-the-Loop for Final Decision

The agent moves further into the transaction – selecting the item, initiating checkout, even presenting payment options – but a human still confirms specific payment before funds move.

- Agent initiates checkout and presents a scoped payment token to the merchant
- Consumer approves the specific amount/merchant, often via a push notification or in-app confirmation
- Consumer consent is captured as a signed mandate or authorization record (e.g., AP2 Mandates, Verifiable Intent)
- Examples: OpenAI's Instant Checkout via Shared Payment Tokens, Mastercard/Visa agent tokens through Stripe



PCI footprint: largely unchanged — provided the token issued to the agent is scoped, single-merchant, time-bound, and human confirmation is where cardholder authorization occurs.

Stage 3 of 3: Fully Autonomous Buying & Payment

The agent decides and pays within pre-authorized limits – no human confirms the transaction. This is real today mainly for machine-to-machine, metered, and micro-payments.

- Consumer retail: still rare; would require durable, revocable, spend-capped authorization set up in advance
- Machine-to-machine: already happening — agents paying per API call, per compute unit, per data query
- Settlement splits by rail: card-network tokens for consumer-style flows; stablecoins (x402, MPP) for high-frequency micro-payments
- This is the stage where architecture choices matter most for future PCI scope



PCI footprint: this is where the industry has NOT yet settled the pattern – PCI footprint depends on account data used, whether it be a token or actual PAN.

The Rails are Not Settled: Six Competing Protocols, Mid-2026

Protocol	Backed By	Layer	Settlement Rail
ACP	OpenAI + Stripe	Checkout protocol	Card (Shared Payment Token)
UCP	Google + Shopify	Merchant/commerce protocol	Multiple (Google Pay, PayPal, cards)
AP2	Google → FIDO Alliance	Consent / mandate layer	Rail-agnostic (cards, stablecoins)
TAP	Visa + Cloudflare	Agent identity / bot detection	Card
Agent Pay	Mastercard	Network token layer	Card (Agentic Tokens) + stablecoin
MPP / x402	Stripe/Tempo, Coinbase/Cloudflare	Machine-to-machine payments	Stablecoin

The common design pattern is not that every protocol uses the same token model. The common pattern is that none of them are built around exposing raw PANs to the agent layer. That's the real through-line for the PCI conversation.

The Connective Tissue: Keep PAN Out of the Agent Layer

Consumer

Grants authorization

AI Agent

Never receives the PAN,
only a scoped token

Token Issuer

Network or PSP issues a
Single merchant, time bound token
(SPT / Agentic Token / TAP credential)

Merchant / Network

Detokenizes within its own
existing PCI-scoped environment

The PAN never reaches the agent, the LLM, or the orchestration layer in any of the six protocols. That’s why we see no material PCI DSS scope impact from agentic commerce as currently architected.

PCI DSS Scope Today: No Material Change

Tokens, not PANs

Across both card and non-card approaches, the agent layer should not handle raw PANs

Detokenization stays put

Detokenization still happens inside the network or processor's existing, already-assessed CDE — not inside the agent or LLM layer.

Consent ≠ cardholder data

Mandates and Verifiable Intent records capture authorization, not account data. They're closer to Requirement 12 evidence than Requirement 3 data.

Familiar TPSP model

Token issuers and agent-identity providers slot into the existing third-party service provider oversight model — the mechanism isn't new, just the counterparties are.

PCI SSC’s AI Principles Still Apply Directly

The PCI SSCs September 2025 AI guidance principles map cleanly onto agent-initiated transactions:

Principle	What It Means for Agentic Flows
No exemption	Using AI anywhere in a payment flow does not exempt that system from any applicable PCI DSS requirement.
Least privilege (Req. 7)	Agents should be provisioned only the access and data needed for their specific function — not broad account access “just in case.”
Data protection (Req. 3 & 4)	Cardholder data touched by an AI system must be secured at rest and in transit exactly as with any other system.
Accountability & logging	Agent actions must be logged and traceable to a responsible human individual — an agent cannot itself be the accountable party.

Looking Ahead: Where PCI Scope Could Still Creep In

Orchestration layers touching raw data	An LLM or agent that logs, caches, or reasons over unmasked PAN e.g., for troubleshooting or context — pulls that system into scope.
Agent credential vaulting	If an agent (or its developer platform) stores long-lived card credentials instead of ephemeral tokens, the agent's infrastructure becomes a CDE component.
New TPSP categories	New Service Provider types, for example, a universal autonomous checkout provider, will need the same AOC diligence as any processor.
Multi-party consent ambiguity	Autonomous, standing-authorization flows blur who “authorized” a transaction — weak consent records could shift liability and scoping questions during disputes.

Design Principles to Keep Scope Minimal

Tokenize at the edge, every time	Never let raw PAN cross into an agent, LLM context window, or orchestration log — tokenize before the agent ever sees it.
Scope tokens tightly	Single-merchant, single-use or short-lived, amount-capped tokens by default; broad standing tokens should be the exception.
Separate consent from payment data	Keep mandate/authorization records (who approved what, when, within what limits) in systems distinct from account data storage.
Treat agent-identity providers as TPSPs from day one	Bring Know-Your-Agent and mandate-verification vendors into the same due-diligence and AOC process as any other service provider.
Design for auditability, not just function	Every autonomous transaction should produce a traceable record linking agent action → authorization → settlement → accountable human.

Questions to Bring Into Scoping Conversations

1. Does any agent, LLM, or orchestration layer ever receive, log, or cache an unmasked PAN or sensitive authentication data — even transiently?
2. Are agent payment credentials true scoped tokens (merchant-bound, time/amount-limited, revocable), or long-lived stored credentials?
3. Who are the new third parties in this flow — agent-identity verifiers, mandate issuers, wallet providers — and is there an AOC or equivalent for each?
4. Can the client produce an auditable chain from agent action to human accountability for every autonomous transaction?

Key Takeaways

It's an evolution, not a leap

Shopper-agent and human-confirmed flows dominate today; full autonomy is still early and mostly machine-to-machine.

The rails are still being negotiated

Six-plus protocols are competing — don't let a client's vendor pitch imply the industry has settled.

Tokenization is holding the line on scope

No protocol reviewed exposes raw PANs to the agent layer — that discipline is the reason scope hasn't expanded, and it needs to hold.

The next architecture decisions will decide future scope

Credential storage, consent design, and new TPSP relationships are where scope creep would actually originate.


20
YEARS OF SECURING
PAYMENT DATA
PCI
Security
Standards Council

Thank You