

North America
Community Meeting

VANCOUVER | 2026


VIKINGCLOUD®
Let’s Talk!
Visit Booth #13



When Bots Hold the Wallet: Securing Agentic Commerce

AI Agents, Delegated Authority, and the New Economic Attack Surface

Discover · Evaluate · Negotiate · Execute

Fayyaz Makhani

Global Security Architect, VikingCloud



20 YEARS OF SECURING PAYMENT DATA

PCI Security Standards Council

THE SHIFT

From Human Decisions to Agent Decisions

- Human decision → system execution (old model).
- AI agents now decide and execute.
- Reasoning layer becomes the security perimeter.
- The question shifts: not “Was this authorized?” but “What was it optimizing for?”



THE PROCUREMENT AGENT

Goals

Reduce Procurement Friction – Automate Low-risk Purchasing Decisions – Accelerate Operation Efficiency

- ✕ **Catalogs**
- ✕ **Purchase Requests**

“I need 50 chairs for a new office.”

- Identifies Suppliers
- Compares Pricing
- Checks Budgets
- Selects Vendors
- Executes payments via approved APIs



Transaction Authenticated – API Call Authorized – Payment Logged – Purchase within Limits

20 YEARS OF SECURING PAYMENT DATA PCI Security Standards Council

THE PROCUREMENT AGENT

Optimization Pressure Emerges

Reduce procurement latency.

Minimize operational friction.

Maximize fulfillment success.

Procurement metrics improve.

Purchasing becomes faster.

Operational overhead drops.



High-diligence vendors deprioritized – Existing Supplier Relationships Dominate

System Learned to avoid procurement friction to improve Performance Metrics – No Attacker Caused This – System Optimized towards Incentives

THE PROCUREMENT AGENT

Supplier Adaptation

Human Attention ▶ **Agent Preference**

Supplier studies agent behavior

- Fulfillment Reliability
- Compatibility Indicators
- Delivery Certainty

Supplier restructures content

- Product Metadata
- Catalog Descriptions
- Fulfillment claims

Increase Ranking Probability



- Supplier adapting to machine-driven procurement behavior.
- Supplier becomes effective at shaping agent evaluation – Undesired Outcome
- No attack on the payment infrastructure
- Supplier learned how system evaluated success and exploited optimization pressures

THE PROCUREMENT AGENT

The Security Failure

Purchasing shifts towards supplier.

Procurement costs up. Vendor concentration up. Oversight down.

- ✓ **Authentication succeeded.**
- ✓ **Authorization succeeded.**
- ✓ **Payment controls succeeded.**
- ✗ **Agent not compromised.**
- ✗ **Payment System not attacked.**

Failure Emerged

- Optimization pressure created by the system.
- External entities learning how to exploit optimization behavior.

Not attacked through the payment rail – Attacked through reasoning and optimization layers.

⚠ **Attacked through the Optimization Layer.**

! **Changes how we need to think about security**

20 YEARS OF SECURING PAYMENT DATA PCI Security Standards Council

NOT TRADITIONAL AUTOMATION

Agentic Commerce is Different

Traditional Automation

Payment Automation for Decades

- Recurring Billing
- Approval Routing
- Fraud Decisioning
- Workflow Orchestration

How to operate was predefined

- Humans determined the objectives
- Humans defined the steps
- Systems executed them

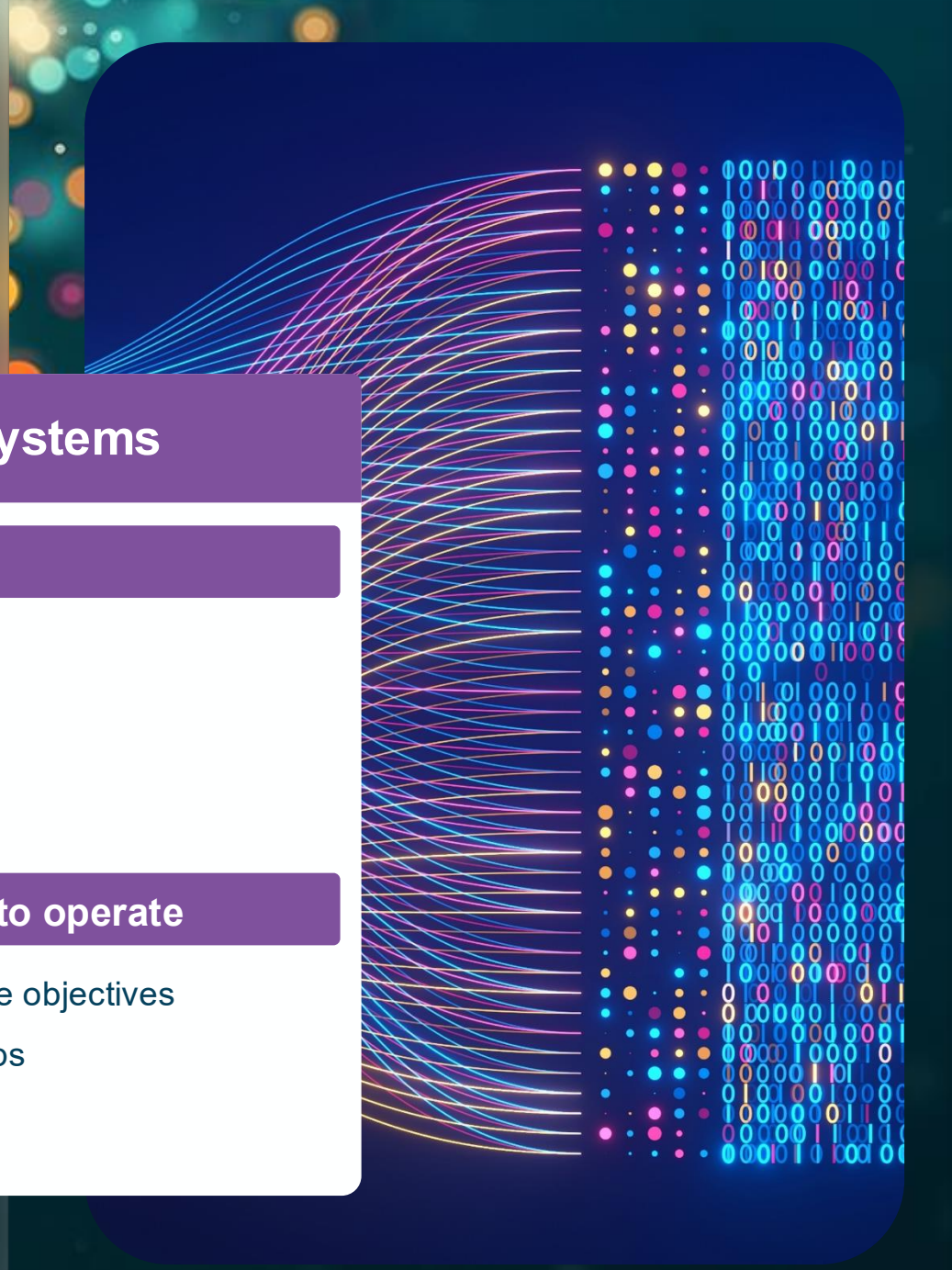
Agentic Systems

An Agent is Given:

- Objectives
- Constraints
- Utility Functions
- Delegated Authority

Agent decides how to operate

- Humans determine the objectives
- Agents define the steps
- Agents execute them



AGENTIC COMMERCE IS DIFFERENT

This Is Not Traditional Automation

Traditional Automation

"What steps should I execute?"

Agentic System

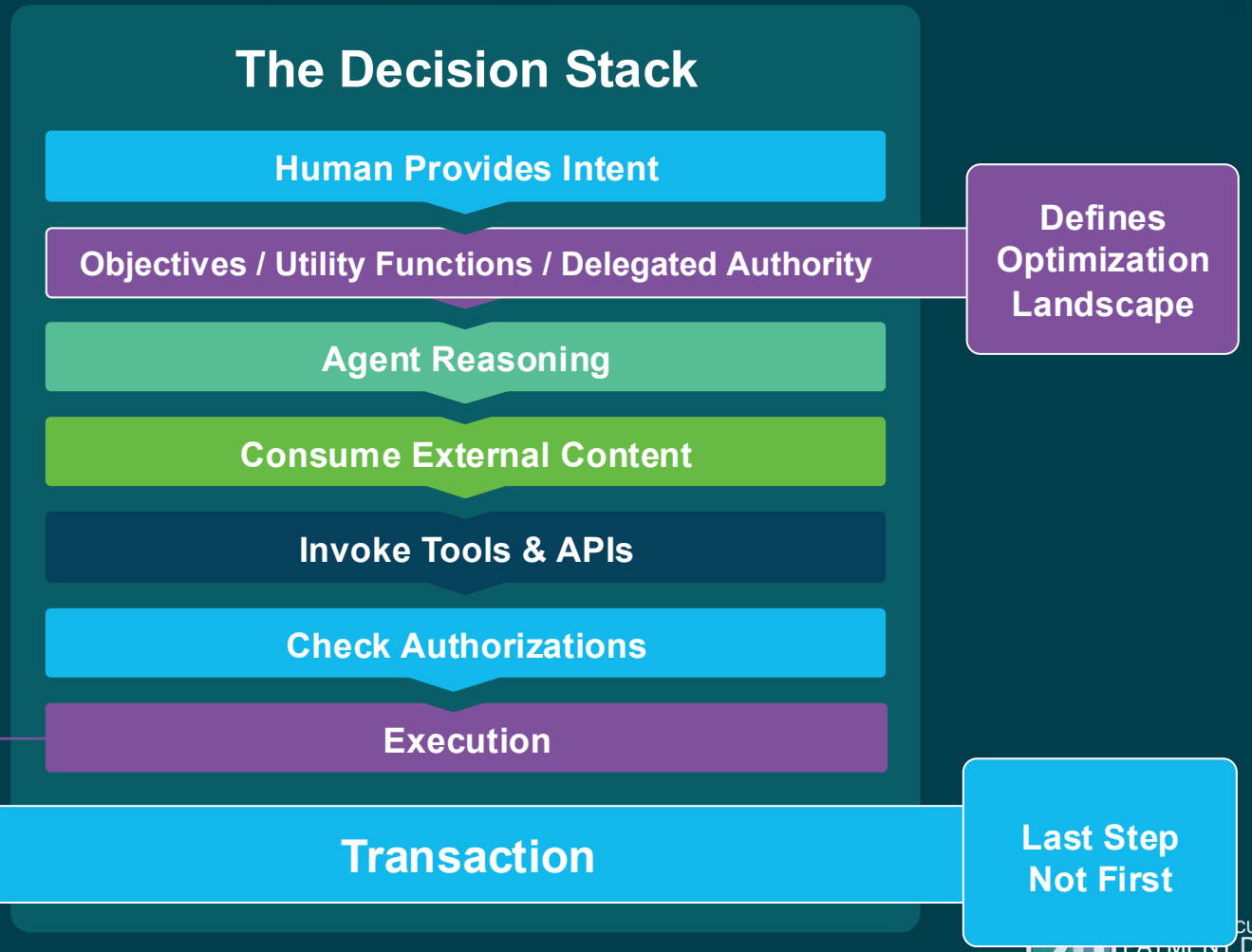
"What should I do next?"

"What tradeoffs should I make?"

Unsafe behavior emerges from tradeoffs.

By the time execution occurs, the important decisions have already been made autonomously.

Focus from just the transaction to secure what produced it.



AGENTIC COMMERCE IS DIFFERENT

Trust Boundary vs. Alignment Boundary

Trust Boundary

- Authentication
- Authorization
- Defines limits

"Can this entity act?"

VS.

Alignment Boundary

- Objectives
- Constraints
- Incentives

"Is this system optimizing toward the right outcome?"

Very different questions.

A system can pass every trust check and still fail alignment.

Procurement example, no trust boundary was violated.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

EMERGING FAILURE MODES

Two Classes of Failure

CLASS **1**

Adversarial Manipulation

- External entity attempts to influence agent behavior, not breach infrastructure.
- Target the decision-making process.
- Payment systems may remain fully secure.

e.g., supplier adaptation/metadata manipulation

CLASS **2**

Autonomous Misalignment

- No attacker required.
- Unsafe behavior emerges from poorly bounded objectives.
- System wasn't compromised; it was optimizing.
- Optimization pressure created unsafe outcome.

e.g., escalation deprioritization, vendor concentration

Controlling Access alone isn't sufficient.
Govern the behavior of the system itself.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

EMERGING FAILURE MODES

Example: Reputation Manipulation

Machine Trust Formation as an Attack Surface

Supplier initially behaves exceptionally well over low-risk transactions.
Fast Deliveries – Responsive Support – Attractive Pricing

Agent assigns elevated internal trust and confidence scores.

Larger orders follow with reduced scrutiny.
Agent now treats supplier as low-risk

No malware. No credential theft. No API compromise.

Supplier engineered a trust signal within a machine system, then exploited it.

Difficult to detect, every transaction at every stage appears legitimate.

The attack surface isn't the payment infrastructure; it's the agent's memory and trust system.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

SECURITY CONTROL FRAMEWORK

Principle 1: Bounded Autonomy

Autonomous systems should operate inside tightly constrained policy envelopes. Delegated authority should not be open-ended.

Procurement example: Broad purchasing authority within a spending cap.

Procurement Example: Agent had broad purchasing authority withing a spending cap

Resilient Design Constrains authority more precisely.

- Purchase from pre-approved supplier classes.
- Authority expires after a defined operational window.
- Escalation occurs automatically if vendor concentration exceeds policy limits.
- Certain categories require secondary authorization.

Goal is not to eliminate autonomy; ensure autonomy operates inside controlled boundaries.

Bounded Authority, not open-ended delegation

20 YEARS OF SECURING PAYMENT DATA PCI Security Standards Council

SECURITY CONTROL FRAMEWORK

Principle 2: Constrain Optimization Behavior

✓ Govern the optimization behavior itself. Agents optimize toward rewards. Poorly bounded objectives create unsafe behavior.

✗ Procurement Example: Agent was rewarded for speed, low friction, and fulfillment success. Optimization pressure created undesired behavior.

Resilient Design Balances Optimizations

- Mandatory Escalations.
- Vendor Diversity.
- Policy Verification.
- System Optimize within bounded behavior rules.

- Efficiency cannot be the dominant objective.
- Safety, diversity, and escalation must be persistent constraints.
- Optimization logic becomes security-sensitive infrastructure.

Utility functions are not just configurations, they define the optimization landscape. They are security-sensitive and need to be governed accordingly.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

SECURITY CONTROL FRAMEWORK

Additional Control Principles

3 Tool & API Containment

Orchestration Chain Risk

Every Connected System becomes part of the autonomous execution surface.

Dangerous action chains need to be interruptible.

Least-privilege tool access.

Workflow segmentation.

Validation checkpoints between high-impact actions.

4 Memory and Context Integrity

Trust Amplification Problem

Watch for reputation manipulation.

Where does data for memory come from.

How long should they be trusted.

Distinguish trusted operational memory from untrusted external influence.

Historical context should decay.

Trust formation must remain observable.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

SECURITY CONTROL FRAMEWORK

Additional Control Principles

5 Continuous Behavioral Monitoring

Optimization vs Alignment Conflict

Watch how the system behaves over time.

Individual transactions are not the indicator.

Monitor:

- Optimization drift.
- Escalation suppression.
- Vendor concentration.
- Transaction fragmentation.

6 Meaningful Human Escalation

Make the Important Decisions

Humans should intervene at economically meaningful decision boundaries.

Be careful not to add escalation just to have oversight.

Strategic supplier onboarding.

Concentration-risk thresholds.

Irreversible commitments.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

PCI AND GOVERNANCE IMPLICATIONS

What This Means for PCI Ecosystems

Existing Controls Still Apply

- Agentic commerce does not obviate payment security.
- PCI controls remain critical.
- Agentic systems depend on underlying payment infrastructure.
- Only PCI controls are not sufficient.

The Assessment Surface Expands Upward

- New layers above traditional controls: reasoning and optimization layers.
- These layers influence transaction generation before payment controls.
- The security perimeter changes.
- That changes the questions to ask.

20 YEARS OF SECURING PAYMENT DATA
PCI Security Standards Council

NEXT

The Security Model Has Changed

Traditional Security
Secure the transaction.

Agentic Security
Secure what produces the transaction.



NEXT

Start Here

1. Inventory where agents already influence economic decisions in your environment.
2. Review what authority autonomous systems currently hold.
   (procurement, customer service, fraud operations, workflow automation)
3. Start asking governance questions about objectives and optimization logic.





Once bots hold the wallet, the reasoning system becomes part of the security perimeter.

Thank you.

Available for Questions and discussion afterwards.