A joint perspective by
Visa and Trusted Agents Ltd
June 2026

VISA | trusted agents

When AI agents act

Trust frameworks and governance for agentic commerce



Disclaimer

Case studies, comparisons, statistics, research, recommendations and information in this Whitepaper (the "Information") are provided "AS IS" and intended for informational purposes only and should not be relied upon for operational, marketing, legal, technical, tax, financial or other advice. Visa neither makes any warranty or representation as to the completeness or accuracy of the information within this document, nor assumes any liability or responsibility that may result from reliance on such information. Visa is not responsible for your use of the marketing materials, best practice recommendations, or other information, including errors of any kind, contained in this document. Materials and recommendations should be independently evaluated in light of your specific business needs and any applicable laws and regulations. The Information contained herein is not intended as investment or legal advice, and readers are encouraged to seek the advice of a competent professional where such advice is required.

The actual costs, savings and benefits of any recommendations or programs may vary based upon your specific organizational needs and program requirements. By their nature, recommendations are not guarantees of future performance or results and are subject to risks, uncertainties and assumptions that are difficult to predict or quantify. Visa is under no duty to update or revise this Whitepaper in any way.

All intellectual property rights, including but not limited to copyrights, and the ownership and title to this Whitepaper and the information contained therein are owned by and vest exclusively in Visa. Third parties shall not have nor acquire any rights in such intellectual property and shall not copy, distribute or make any other use of the Whitepaper and the Information contained therein.

All brand names and logos are the property of their respective owners, are used for identification purposes only, and DO NOT imply product endorsement or affiliation with Visa.

© 2026 Visa. All rights reserved.

What's inside

04 Introduction

05 When the 'customer' at checkout is an AI agent

07 From AI that talks to AI that acts: what we mean by agentic commerce

09 Why AI agents need trust: verifiable and stable identifiers, mandates and context

12 What this means for merchants and acquirers: staying visible, not just staying safe

13 What this means for financial institutions: From KYC to KYA

14 What this means for ecosystem players and policymakers: building trust frameworks for agents

15 A global perspective: How 'AI agent-forward' are different markets?

16 What's your role in trusted agentic commerce?

17 About Visa and Trusted Agents Ltd

© 2026 Visa. All rights reserved.

Introduction

Commerce has evolved in waves, redefining how people transact, how businesses compete, and how trust is established at the moment of purchase. The shift from in-store to online reshaped checkout around browsers, logins and card details entered on a screen. The rise of mobile and digital wallets made shopping more contextual, tokenised and embedded into everyday user journeys. Through each of these waves, generations of AI have quietly done the heavy lifting of fraud detection, real-time risk scoring and offer personalisation. Each wave brought new participants, new risks, and new infrastructure to keep commerce safe, fast and trusted at scale.

We are now entering the next wave — agentic commerce. Agentic commerce describes a shift where AI agents act on behalf of people and organisations, including carrying out multi-step tasks, making decisions and completing transactions. As these agents move from assistance into meaningful action, our long-standing assumptions about authentication, authority and accountability in commerce need to evolve.

This paper takes a look at what changes when AI agents, rather than humans, become the effective 'customers' at checkout.

We focus on the principles for agentic commerce rather than specific products or protocols: why AI agents require attributable trust, digital identifiers, authentication and delegation; how trust and risk must be redesigned; and what this shift means for financial institutions and policymakers, as agent-driven transactions move from experimentation towards scale.

This paper takes a look at what changes when AI agents, rather than humans, become the effective 'customers' at checkout.

Agentic commerce describes a shift where AI agents act on behalf of people and organisations, including carrying out multi-step tasks, making decisions and completing transactions.

© 2026 Visa. All rights reserved.

When the 'customer' at checkout is an AI agent

Let's start by imagining you are taking a routine, domestic trip for work. Your calendar shows an in-person meeting in another city next month. But this time, you don't open a browser or compare options.

Instead you say:

'Book me a return trip next month. Economy only, train if it's under three hours, flight if not. Hotel near the office. Keep the total under five hundred euros.'

Your very own AI agent takes it from there. It checks your calendar, your past preferences and the constraints you've set in the request. It then looks at a wide range of rail and flight options, applies fare rules, selects a hotel that meets your location and cancellation preferences, and checks everything against budget. It also knows when to escalate, and if nothing fits cleanly, it will ask for further information.

A few minutes later, your AI agent comes back with a summary:

'Itinerary identified: outbound Tuesday morning, return next day. 2h40m travel time, rail selected, fully refundable fare. Hotel within walking distance, with flexible cancellation. Total €468. Please tap to confirm.'

You check the details and decide that it's good to go. You approve with your face or fingerprint, and move on to your next task of the day.

Now, from your point of view, this was a single decision. But behind the scenes, a very different checkout took place:

- The train company didn't see a person browsing web pages, instead it received a structured request from an AI agent.
- The hotel received a policy-compliant order, not a clickstream.
- The payment arrived with an explicit, machine-readable mandate describing whose money this was, which rules applied, and what should happen if anything fell outside those rules.



Unlike traditional digital commerce, where a person manages their own small and stable set of payment credentials, agentic systems may act through dozens or hundreds of agents operating across merchants, platforms and time horizons.

© 2026 Visa. All rights reserved.

Because here's what really happened. The 'customer' at checkout wasn't you. It was your AI agent, acting on your behalf, using your payment credential.

This shift matters because today, the customer at checkout is the payment account holder, even if they are making the purchase on behalf of another customer. When decisions and transactions are delegated to AI agents, those assumptions break down, making it necessary to surface authorisation, mandate and provenance at the point of sale.

Agentic environments introduce a fundamentally new scale challenge for payments. Unlike traditional digital commerce, where a person manages their own small and stable set of payment credentials, agentic systems may act through dozens or hundreds of agents operating across merchants, platforms and time horizons. This creates an environment where payment credentials are more frequently created and used. Without a way to explicitly identify agentic credentials, the ecosystem cannot reliably distinguish credentials intentionally granted to a trusted AI agent from those obtained illicitly. Addressing this challenge requires the continued evolution of payment credentials. Just as commerce evolved from card-present to e-commerce transactions, agentic commerce requires credentials that are expressly granted to AI agents, tokenised and identifiable as agent-initiated at the point of execution.

Today's digital commerce stack assumes there's a person directly involved. Journeys are designed around humans reading pages, making choices, and typing card numbers. Identity is inferred from logins, devices and behavioural signals. Fraud models track patterns by people.

But agentic commerce breaks those assumptions. Because when AI does the searching, choosing and buying, today's familiar — and human — signals weaken and break. And the critical questions posed by banks, merchants and regulators have to be even sharper:

- Who is this agent?
- Who does it represent?
- What is it allowed to do?
- Who is accountable if something goes wrong?

The core idea is that agents need explicit, attributable trust, anchored in verifiable and stable agent identifiers.

Below, we define what we mean by agentic commerce, explain why trust, identifiers, authentication and delegation are foundational, outline what this means for key players in the market, and highlight how different regions are positioned for what comes next.

Agents need explicit, attributable trust, anchored in verifiable and stable agent identifiers.

© 2026 Visa. All rights reserved.

From AI that talks to AI that acts

What we mean by agentic commerce

Most executives are now familiar with 'AI that talks': chatbots, copilots and virtual assistants that help people find information or complete simple tasks. These systems sit on top of existing customer journeys. They assist, but they don't fundamentally change how data and tasks flow through the business.

But the next wave of AI that's about to sweep across the economy, agentic commerce, is different. It's about AI that acts — with financial, contractual or operational consequences. Critically, this shift becomes increasingly meaningful when AI agents operate across organisational boundaries, interacting with merchants, platforms, and financial institutions rather than remaining within a single domain.

Agentic commerce is where AI agents can browse, compare options, apply your preferences and then pay, all within rules you set. AI that can search and compare is a powerful start, but without payments it is more like window shopping. Agentic commerce becomes truly powerful when those same agents can also complete the transaction securely — turning intent into purchase, not just recommendations.

When commitments are binding, it is important to understand who authorised that action, within what limits and under which conditions. In agentic commerce, three characteristics will matter deeply.

Goal-driven, not screen-driven

Customers and businesses will set specific objectives, and AI agents will take responsibility for achieving them, often across multiple merchants, channels and payment types, without exposing every step to a human interface. 'Keep these subscriptions within budget,' 'Book compliant travel,' 'Re-order inventory below this threshold.'

Policy restricted, not ad hoc

AI agents may be programmed to operate within explicit parameters: spending limits, risk tolerances, corporate policies and personal preferences. Some will be hard constraints, and others will be high-level principles and preferences. Together, they form a delegation and control framework between the human's intent and an agent's action. For this framework to work across organisations and markets, customer intent and constraints must be representable in ways that merchants, financial institutions, and networks can interpret and rely upon.



The next wave of AI that's about to sweep across the economy, agentic commerce, is different. It's about AI that acts — with financial, contractual or operational consequences.

© 2026 Visa. All rights reserved.

Completing transactions, not just recommending

AI agents do not merely suggest. They will be able to accept terms, initiate payments and update records. Ledgers, contracts, and systems will all look different after AI agents have been involved. Zoom out, and we can already see a progression of AI agent capability emerging:

- Tools and copilots that assist but leave execution to humans
- Task-level agents that complete bounded jobs end-to-end
- Process-level agents that run ongoing activities within defined rules

Our focus is on the second and third levels: where agents can initiate payments, accept terms, and create binding financial or contractual outcomes. These are the critical parts if we want to be able to trust these new, powerful AI agents at scale. Because at this point, identifiers, authentication and authority can no longer be implied. They must be explicit and verifiable.



Ledgers, contracts, and systems will all look different after AI agents have been involved.

© 2026 Visa. All rights reserved.

Why AI agents need trust

Verifiable and stable identifiers, mandates and context

Today's digital commerce model is centred around humans — specifically, on human identity, intent and reputation. A customer logs in, strong authentication confirms who they are, and a card, wallet or account represents how they pay. And today's risk mitigation systems rely on devices, locations and behavioural patterns.

When it's not humans acting, but instead some delegated AI agents, our commercial and operating models start to fray. Gaps emerge not just in fraud detection, but in auditability, dispute resolution and regulatory oversight. As governments increasingly formalise digital authentication and wallet frameworks, these foundations become directly relevant to how AI agents will be recognised and governed.

In practice, three unsatisfactory patterns are already appearing across the economy:

01

AI agents impersonate users

The AI agent logs in 'as' the human and operates in the background. In logs and disputes, there is currently no distinction as to who conducted an action.

02

AI agents hide behind shared service accounts

Multiple AI agents use generic technical identifiers. This works until something goes wrong, at which point it is impossible to tell which AI agent was which, and which instance caused the issue.

03

AI agents disappear inside platforms

From a merchant or issuer's perspective, everything comes from 'Platform XYZ.' The underlying customer, AI agent, its policies and its provider are mostly opaque. All automation looks the same.

Each of these weakens accountability and makes regulators cautious about widening what AI agents are allowed to do. To move beyond this, AI agents are going to need recognised trust credentials of their own.

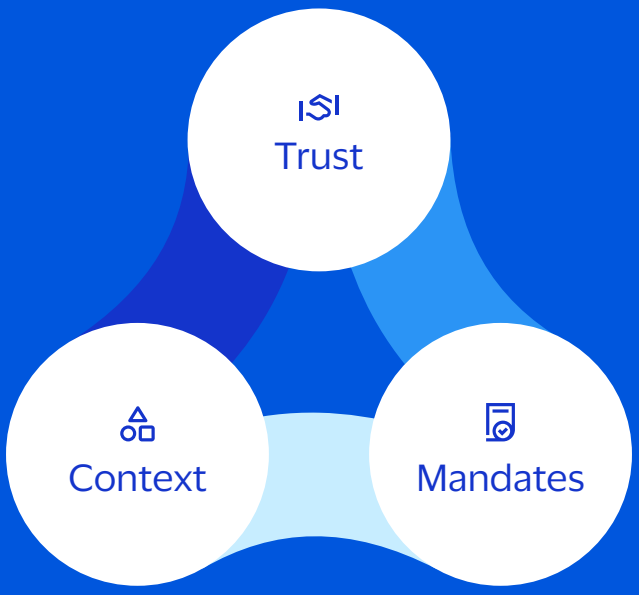
They need attributable, verifiable agency: a reliable way to say, 'This specific AI agent, operated by this provider, at the request of this human or organisation, and acting under these constraints, did this.'



As governments increasingly formalise digital authentication and wallet frameworks, these foundations become directly relevant to how AI agents will be recognised and governed.

© 2026 Visa. All rights reserved.

We describe this with a simple framework: Trust, Mandates and Context.



Establishing trusted AI interactions requires three foundations — verifiable trust, clear mandates and operational context.

Trust

Can another party recognise a legitimate AI agent and its operator? Can it distinguish trusted AI agents from hostile automation? Can we assess this AI agent's behaviour over time and understand its reputation?

Mandates

Can counterparties understand what authority an AI agent has been granted, by whom, within which limits, and for how long, with that information made available at the time of the transaction or post-transaction?

Context

Can systems understand the preferences, history, constraints and objectives the AI agent is optimising against at the moment of execution?

When these three elements are explicit and combined, AI agents can be made accountable. Ensuring accountability at scale requires shared ways for agents to declare a verifiable identity, delegated authority, and execution context so counterparties can interpret and validate them consistently.

Visa's Trusted Agent Protocol provides that identity layer for agents, users, and merchants to evaluate, trust across use cases. Many early agents will rely on the trust and security provided by payment wallets and platforms, while protocols like the Trusted Agent Protocol provide a scalable path for agents that require higher degrees of independent trust or autonomy. With these shared signals in place, disciplines like Know Your Agent become practicable rather than abstract when interacting with merchant domains. We can reconstruct what happened, determine whether an AI agent acted within its mandate, and allocate liability fairly when something goes wrong. Critically, this record of instructions, mandates, identity assurance, and execution history forms compelling evidence for dispute resolution, investigations, and regulatory review.

This is where 'Know Your Agent' (KYA), an emerging framework for handling AI agent liabilities and risk, emerges as a practical complement to 'Know Your Customer' (KYC).

KYA is not a single system or registry, but rather a risk discipline spanning digital identity, delegation and evidence, ensuring that agents are identifiable, authorised and auditable. KYA does not replace KYC, but extends compliance and trust to include both the consumer and the agent in a world where non-human actors initiate transactions.

Establishing trust in agentic payments requires a high level of assurance at two critical moments: when agentic credentials are created, and when a mandate is established for an AI agent to act on behalf of a cardholder. In practice, this level of assurance is only achievable through modern means of digital identity authentication. Lower-assurance mechanisms that rely on inferred identity or behavioural signals do not scale to agentic environments where credentials may be created and exercised programmatically.

Trust frameworks like these are already emerging globally. In Europe, eIDAS2 and digital wallet initiatives (particularly those focused on business wallets) are establishing high-assurance authentication, credential standards and delegation approaches. While not designed specifically for AI agents, these frameworks provide shared trust rails that agentic commerce can build upon and align with.

By anchoring agents in recognised identifiers, authentication and mandate systems, markets can accelerate safe adoption while containing systemic risk.

Establishing trust in agentic payments requires a high level of assurance at two critical moments:

01
When agentic credentials are created

02
When a mandate is established for an AI agent to act on behalf of a cardholder

© 2026 Visa. All rights reserved.

What this means for merchants and acquirers

Staying visible, not just staying safe

For merchants and acquirers, agentic commerce requires two things at once: how to handle non-human traffic when it shows up, and how to interpret new market demand from AI agents.

AI agents will be able to screen inventory, read terms and compare offers on behalf of customers. If products, prices and policies are only visible in HTML pages or PDFs, the customer's AI agent will approximate what's possible and available or simply look elsewhere. And it's not just about being accessible and legible. It's also about making the criteria for making the transactions explicitly defined, not being open to interpretation.

And we must remember that not all automation will be benign. Hostile bots will continue to exist and are already attempting to exploit agentic commerce capabilities. Our biggest challenge now is to distinguish between the two and handle the risks they will increasingly pose.

To thrive in agentic commerce, merchants and acquirers need to stay visible and stay safe.

Staying visible means becoming accessible to AI agents:

- Exposing products, availability, pricing and key policies in structured, machine-readable and precise forms

Staying safe means evolving today's automated controls:

- Preparing for a world where automated digital traffic will represent high-intent, policy-compliant demand from trustworthy AI agents
- Reviewing fraud and bot management tools that currently treat all non-human traffic as hostile — because we must be careful not to screen out legitimate customer AI agents too. Standardised agent trust signals, whether conveyed through protocols or carried via commerce and payment networks, will be essential to make that distinction reliably at scale.

Merchants and acquirers should also plan for agent-driven disputes. When a customer says, 'my AI agent booked this, not me,' resolving these cases will depend on whether the AI agent's reputation and mandate information is not only available, but accurate and independently verifiable.

At this point, it's important to say that the adoption of AI agents by merchants and acquirers can be gradual. Managed, well-scoped use cases like subscription management can be delivered by AI agents operating within clear and cryptographically signed limits, allowing the business to learn without excessive risk.



To not only survive but thrive in agentic commerce, merchants and acquirers need to stay visible and stay safe.

© 2026 Visa. All rights reserved.

What this means for financial institutions

From KYC to KYA

For financial institutions, AI agents turn a traditional two-party relationship into a new, governed triangle: customer, AI agent and the financial institution.

This shift in commercial relationships is subtle but important. We move from single authentication events ('this person authenticated and then paid') to new instructions ('this person authorised this AI agent to act within these rules until revoked'). This is the structural change we need to pay attention to. If we only treat AI agents as 'invisible scripts' behind each customer, we can't understand the crucial context behind each transaction, reducing the ability to handle risk coherently.

Instead, treating AI agents as delegated execution entities operating under explicit customer authority clearly defines the role the agent is playing in this new commerce flow, enabling financial institutions and issuers to manage liability in a new way.

'KYA' becomes a new discipline, involving three areas:

01

Agent onboarding, verification and certification

Decide which types of AI agents the business will work with, classify them by risk (internal AI agents vs. customer-owned AI agents vs. third-party AI agents), and establish appropriate verification and certification processes. Just as with KYC or KYB, checks are required to ensure that the AI agent operates a secure and trusted commerce experience.

02

Mandate lifecycle management

Give customers clear ways to grant, review and revoke authority. Enforce limits in real time and clearly map each agentic transaction to a mandate approved by the customer. Recognise that mandates will change continually.

03

Monitoring and supervision

Track AI agent behaviour over time. Set thresholds for intervention when patterns suggest misconfiguration, misuse or compromise.

For payment networks, the design challenge is to carry AI agents and new mandate and trust signals in ways that are interoperable, privacy-preserving and usable at authorisation time. Network-level capabilities such as Visa's Trusted Agent Protocol illustrate one direction of travel: carrying agent identifiers, the token that contains mandate references, and transaction context alongside payments, so AI agents can be recognised as governed actors without fragmenting existing rails — building on existing trust and risk coordination tools rather than bypassing them.

For payment networks, the design challenge is to carry AI agents and new mandate and trust signals in ways that are interoperable, privacy-preserving and usable at authorisation time.

© 2026 Visa. All rights reserved.

What this means for ecosystem players and policymakers

Building trust frameworks for agents

Around the world, governments have an increasingly central role in developing digital identity infrastructures. These infrastructures could be used to accelerate agentic commerce. Specifically, the implementation of new trust frameworks could become the 'rules of the road' for everyone participating in agentic commerce, including who can act, on what terms and with what consequences. As AI agents increasingly appear in payments and commerce, not to mention customer services and loyalty flows, these AI agent trust frameworks will need to rapidly evolve.

A simple way to think about an agentic commerce trust framework is across three phases:

01	**Before the transaction**	Allocating identifiers, credentials, mandates
02	**During the transaction**	Standards for evidence, how agents signal their role and authority, how counterparties verify those signals, and how risk engines use them
03	**After the transaction**	Supervisory access, what evidence is retained, how disputes are categorised, and how systemic risks are monitored

Across these phases, different layers of the ecosystem play distinct but complementary roles. Government-led digital identity and wallet initiatives provide an early blueprint for how such trust layers can be structured at scale, while network level capabilities such as Visa Intelligent Commerce can operationalise these foundations by carrying identity, mandate, and risk signals at transaction time. In parallel, Visa's Trusted Agent Protocol provides the construct in which agents are able to display verifiable intent and authority to merchants and platforms in a consistent way — together allowing identity, mandates and accountability to function reliably in day-to-day transactions.

Importantly, AI agents should not be treated as loopholes. If rules already exist for human-initiated transactions, there should be clear positions on how they apply when AI agents act instead — and what additional safeguards are justified.

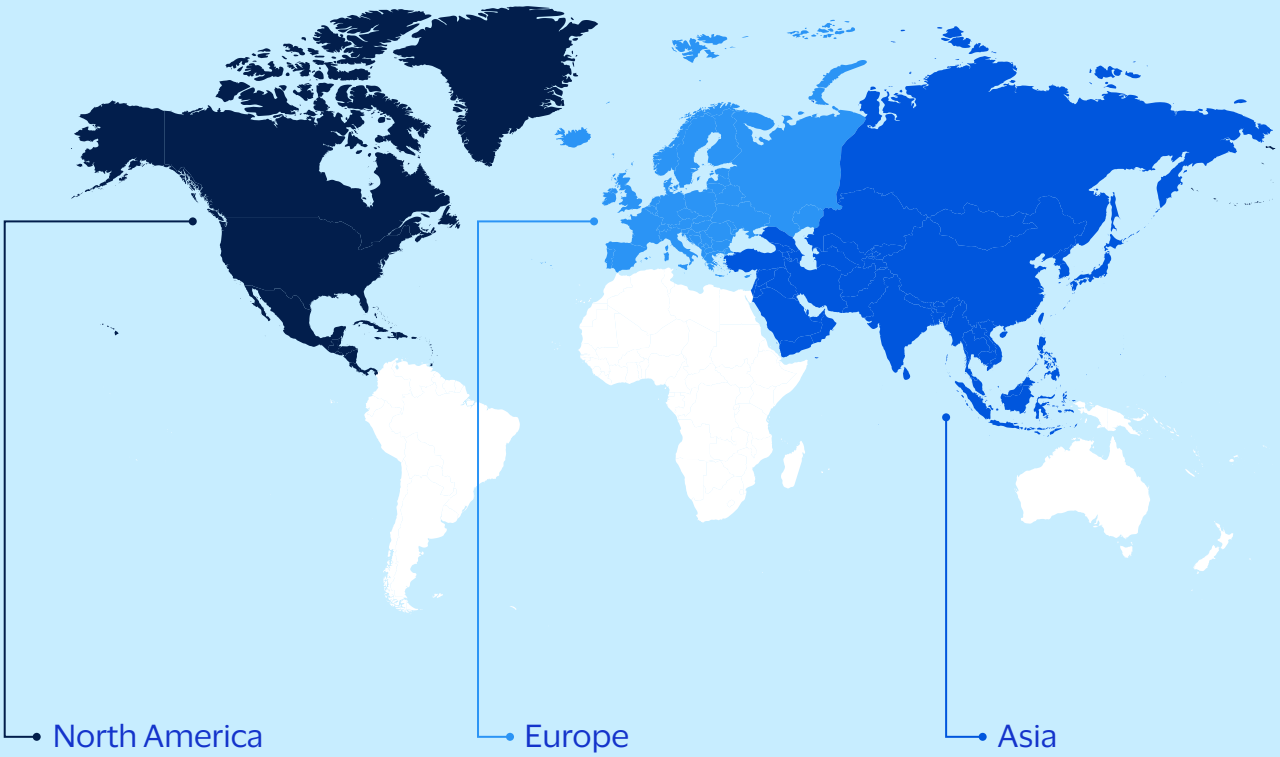
If rules already exist for human-initiated transactions, there should be clear positions on how they apply when AI agents act instead — and what additional safeguards are justified.

© 2026 Visa. All rights reserved.

A global perspective

How 'AI agent-forward' are different markets?

Agentic commerce is already emerging globally, but the starting points will differ based on markets and jurisdictions.



North America

Here, rapid digital innovation exists within a fragmented identity and authentication market and contested data-sharing rules. Early AI agent experiences are most likely to emerge within large, existing payment ecosystems, with payment networks and financial institutions playing a stabilising role through shared payment-level signals

Europe

Europe has already scaled identity, authentication and payments infrastructure. Setting a high bar for digital trust creates a foundation for formally governing AI agents, albeit with some complexity. AI Agents operating in Europe will benefit from anchoring in the EU's emerging authentication and mandate frameworks, including the eIDAS2 digital wallet initiatives.

Asia

With the dominance of super-app models and platforms, agent-initiated payment flows operating under user mandates are already familiar. Combined with growing attention to AI governance, parts of Asia are natural testbeds for production-grade agentic commerce. Markets that anchor AI agents in shared authentication and delegation frameworks will find it easier to allow agentic commerce to scale without uncontrolled risk.

© 2026 Visa. All rights reserved.

What's your role in trusted agentic commerce?

Agentic commerce only becomes trusted, scalable and viable when the ecosystem can confidently answer four questions:

01	02	03	04
Which AI agent acted?	On whose behalf?	Within what limits?	What exactly did it do?

For business leaders, the next step is to apply that lens internally.

For merchants and acquirers

Are we accessible to AI agents today, or only to humans with browsers?

Do our risk and operational controls distinguish trusted agents from hostile automation?

Where could AI agents bring us strong customer intent and policy-compliant market demand?

For financial institutions

Where are agents already close to our business processes and financial infrastructure?

Should we help issue and manage AI agent mandates for any specific customer segment?

Can our systems treat AI agents and customer mandates as first-class entities across the business?

For ecosystem players and policymakers

Do our existing definitions clearly cover AI agents acting on a user's behalf?

What minimum Trust Frameworks will be needed, that all legitimate AI agents will respect?

How will we avoid either the over-centralisation or fragmentation of AI agent trust?

This paper has been deliberately less focused on AI agent technology, and more on the principles for agentic commerce. We believe that over the next few years, AI agents will be acting on behalf of customers at scale. The opportunity is to prepare deliberately. With coordinated action between businesses, technology providers and governments, agentic commerce can uplift human agency, grow markets and build trust from the start. And by aligning agent capabilities with emerging authentication layers and trust frameworks, we can ensure that accountability is designed in.

We believe that over the next few years, AI agents will be acting on behalf of customers at scale. The opportunity is to prepare deliberately.

© 2026 Visa. All rights reserved.

Acknowledgements

For Visa

Authors of this paper are Milja Radosavljevic and Ranjiva Prasad

The authors would like to thank the following colleagues for their input and review:
Rob Walls, Parth Awasthi, Gilles Verstraeten, Steven Sacia, Helena Hu, Tod Frincke

For Trusted Agents

Authors of this paper are Jamie Smith and Gam Dias

About Visa

Visa (NYSE: V) is a world leader in digital payments, facilitating transactions between consumers, merchants, financial institutions and government entities across more than 200 countries and territories. Our mission is to connect the world through the most innovative, convenient, reliable and secure payments network, enabling individuals, businesses and economies to thrive. Visa has been an innovator for decades in the application of AI to payments and money movement, including fraud detection, risk decisioning, tokenisation, dispute management and programme controls. In 2025, Fortune recognised Visa with a #2 ranking on its AIQ 50 list of global AI leaders. Learn more at Visa.com.

About Trusted Agents Ltd

Trusted Agents is a specialist advisory firm focused on the business, governance, identity and trust architectures required for Agentic Commerce. It works with enterprises, financial institutions and digital innovation leaders to prepare for a world in which AI agents act on behalf of individuals and businesses. Its work centres on the practical implications of agentic identity, delegation, customer context and digital trust in agent-mediated transactions. You can find out more at thetrustedagents.ai/trust/.

© 2026 Visa. All rights reserved.

© 2026 Visa. All rights reserved.