

Machine-to-machine payments in the AiFi era

Contents

Executive Summary	3
1 Introduction: The internet has a second customer	4
2 The web's referral bargain has broken down	5
3 The market is already building a third option	6
4 Legacy rails struggle to handle the smallest transactions	8
4.1 The fixed-cost floor	
4.2 The underwriting wall	
4.3 The dispute model mismatch	
4.4 The identity and geography mismatch	
4.5 Comparative summary	
5 Stablecoins are best suited to high-frequency, pay-per-use agentic transactions	13
6 Identity and interoperability decide whether payments scale	16
7 Policymakers should preserve the economics, not force the categories	17
7.1 Preserve micropayment economics	
7.2 Build compliance around identity, authorization, and wallet controls	
7.3 Treat open protocols as competition policy	
7.4 Recognize per-use payment as creator compensation	
7.5 Coordinate internationally before proprietary systems set the defaults	
8 Conclusion: The payments layer decides who participates	19
References	20

Acknowledgements:

The authors thank Christian Catalini, Cesare Fracassi, Omid Malekan, Josh Nickerson, Felix Haynes, Kossi Anyinefa, and Taleen Sample for their reviews and comments. Any errors or omissions are the authors' own.

Executive Summary

Every decade, a new way of using the internet becomes dominant, and each shift changes who participates in the digital economy. The 2010s shift from desktop to mobile brought billions of new people online. The 2020s shift from mobile to agents is different. It does more than change the device. It introduces a second customer: software that retrieves, compares, and increasingly transacts on a human's behalf at a scale and speed the web has never had to serve.

The web's old bargain has broken down. For years, publishers and developers allowed crawlers (automated programs that download, index, and reuse pages at scale) because search engines sent readers back to them. Generative AI changes that exchange: systems crawl to answer rather than to refer, while automated traffic now makes up a large and growing share of web activity. Google search referrals to publishers fell by roughly a third globally from November 2024 to November 2025, widening the gap between the value extracted from content and data and the revenue returned to those who produce it. A functioning per-use market compensates every publisher that sets a price, not only those large enough to negotiate bilateral deals.

A payment layer for the second customer is now emerging. Most sites today face a binary choice: allow automated traffic for free or block it before it reaches the site. Cloudflare and AWS are both moving toward a third option, in which agents pay for access rather than being turned away. x402, an open standard built on the HTTP 402 "Payment Required" response code, provides the mechanism. A server responds to an unpaid request with a price; the agent pays, retries the request, and receives the resource. What matters about x402 is not that it selects one settlement asset or one provider. It is that it standardizes the payment handshake itself, leaving room for competing settlement schemes rather than routing agentic commerce through the same two card networks that intermediate most online payments today.

This paper makes three claims. First, legacy payment rails cannot economically support the high-frequency, low-value core of agentic commerce. Fixed per-transaction costs, underwriting barriers, reversal risk, and identity models built for people rather than machines make many machine-to-machine transactions uneconomic. For the smallest transactions, the cost gap is three to four orders of magnitude. Second, stablecoins are today the strongest fit for this market because they combine low-cost settlement, rapid confirmation, programmability, and global reach. Third, whether this layer develops on open, interoperable standards or inside proprietary marketplaces and closed billing systems will determine whether agentic commerce repeats the fragmentation and weak competition of the systems it replaces.

The policy stakes are global. A long tail of publishers, developers, data providers, and agents will need to settle transactions across borders and around the clock. The amounts will often be too small, and the commercial relationships too fleeting, for legacy systems. Policymakers should therefore focus less on fitting this market into existing categories and more on preserving the conditions it needs to operate: micropayment economics, interoperable identity, open protocols, per-use payments that compensate creators, and room for multiple compliant settlement schemes.

Introduction: The internet has a second customer

The internet was built for one kind of user: a human with eyes, attention, and a payment credential. Its commercial stack of advertising, subscriptions, checkout flows, fraud screening, and chargebacks assumes that the internet activity that matters comes from a human user. But that assumption is weakening. Automated traffic has surpassed human traffic across many areas of the internet, and the fastest-growing category of actor online is the AI agent¹: software that acts on a human principal's behalf with an increasing degree of autonomy.

The internet now has a second customer. This customer does not watch ads, does not browse like a human, and does not fit neatly into the commercial mechanisms the web evolved for people. It needs to retrieve data, invoke services, access content, and make decisions in real time. In many cases, it needs to pay in real time as well. That shift breaks the open web's funding model. Most free content and most public APIs are funded indirectly by attention: a human arrives, sees an ad or converts to a subscription, and the page earns its keep on the visit rather than on the request. An agent generates the request without the attention. It retrieves the content, summarizes it, and returns an answer to its principal. The publisher loses both the human visit and the revenue that visit would have generated. If the value is still being consumed, the charge has to move from the human's attention to the machine's request: per call, per document, per query, per unit of compute. Where the reader is software, direct payment for use is the only mechanism left.

The internet offers many services that are valuable in an agentic economy. The problem is that the payment layer between those services and the agents that use them is poorly suited to the task. Legacy rails were designed for higher-value human transactions between known counterparties, with settlement netted over time and disputes resolved after the fact. Agentic commerce inverts those assumptions. Most transactions are first contact: a server receives a resource request from an agent it has no prior relationship with, no credit history for, and no practical way to underwrite. That single fact drives the rest. Netting is efficient, but it requires some party to carry exposure during the settlement window. That exposure can be priced sensibly only when the counterparty is known and creditworthy. Where the resource is delivered in milliseconds and the buyer is anonymous, the seller either takes settlement at the moment of access or ships value on credit to a stranger. Where no relationship exists, an ex post dispute process has nothing to attach to. Controls must instead operate ex ante by limiting what a payment is authorized to do before it is made.

¹ "Traffic Worldwide: Bot vs. Human," Cloudflare Radar, accessed August 18, 2026, <https://radar.cloudflare.com/traffic#bot-vs-human>. According to Cloudflare's 7-day trailing data for HTML content, bot (automated) traffic accounts for 60.4 percent of web page requests, compared to 39.6 percent for human traffic.

Agents already exist. The commercial future of the open internet now depends on whether the web can charge them per request through a payment system that is economically viable, globally accessible, and interoperable across providers.

2

The web's referral bargain has broken down

The web's original economic arrangement was straightforward. Publishers allowed search engines to crawl and index their pages because search results sent readers back to them. That exchange financed content, discovery, and a great deal of open access.

Generative AI strains that arrangement. AI systems increasingly crawl to answer rather than to refer. Available evidence shows that AI systems crawl thousands of pages for every visitor they refer,² while users click cited sources only a small fraction of the time. Publishers and data providers face pressure from both falling referrals and rising infrastructure costs. Google search referrals fell by roughly a third globally from November 2024 to November 2025,³ while repeated crawling and other automated use increased infrastructure costs.

If automated use grows without replacing lost referral revenue, smaller publishers may have to cut reporting, put more material behind subscriptions, or stop publishing. Creators could become more dependent on large platforms for distribution and licensing income. Readers would have less access to independent information. A practical way to charge for automated use could help sustain the work that those systems rely on.

This is an intellectual-property dispute, but it is also a pricing failure. Large publishers may be able to negotiate bespoke licensing deals. The long tail usually cannot. A local publication, solo developer, or niche data provider rarely has the scale or leverage to negotiate an enterprise contract for each interaction. What they need instead is a way to price access at the moment of use.

That is why the development of pay-per-use access matters. The important shift is the emergence of a third option between open access and total exclusion. Infrastructure providers are beginning to offer pay-per-use access at the protocol layer rather than simply blocking agents.

2 João Tomé, "The Crawl-to-Click Gap: Cloudflare Data on AI Bots, Training, and Referrals," Cloudflare Blog, August 29, 2025, <https://blog.cloudflare.com/crawlers-click-ai-bots-training/>.

3 Charlotte Tobitt, "Global publisher Google traffic dropped by a third in 2025," Press Gazette, January 16, 2026, <https://pressgazette.co.uk/media-audience-and-business-data/google-traffic-down-2025-trends-report-2026/>.

3

The market is already building a third option

Demand for paid access is already visible.

Publishers have asked for payment mechanisms that sit between unrestricted crawling and outright blocking. From September 15, 2026, Cloudflare will block AI training and agentic crawlers⁴ by default on ad-supported pages for new sites and free-tier customers, while continuing to allow search crawlers. But blocking is not what publishers asked for. When Cloudflare designed its Pay Per Crawl system, hundreds of news organizations and platforms gave the company the same message: they wanted to admit AI crawlers and charge them for access rather than leave the door open or wall off their content.⁵ Publisher-side reporting shows what that preference is worth. High-volume sites can earn \$50,000–\$200,000 per month from crawler payments, and nearly 70 percent of publishers expect AI licensing to generate revenue within three years.⁶ Alongside broader industry experiments with attribution and licensing, these point to the same conclusion: content producers want paid access rather than a binary open-or-shut regime.

Pay-per-use access is being introduced by companies that already serve much of the internet. Cloudflare is the reverse proxy in front of roughly a quarter of all websites, and AWS runs 28 percent of the world's cloud infrastructure.⁷ Both are moving toward the same approach: let websites charge agents for access as part of an ordinary web request, rather than block them. When the two firms with the broadest reach across web infrastructure independently adopt the same paywall mechanism, it moves from proposal to deployed infrastructure. That same fact raises the question this paper returns to in Sections 6 and 7: whether paid access is built on open protocols or confined to proprietary marketplaces. A payment mechanism built on an open standard lets any publisher set a price and any agent pay it without joining anyone's

4 "Your Site, Your Rules: New AI Traffic Options for All Customers," Cloudflare Blog, July 1, 2026, <https://blog.cloudflare.com/content-independence-day-ai-options/>; Effective September 15, 2026, Cloudflare sets new defaults across three crawler classifications: on pages that display ads, Training and Agentic crawlers are blocked by default, while Search remains allowed. The new defaults apply to newly onboarded domains, new sites added by existing customers, and existing free-tier customers; other customers may opt out in their Security settings at any point before that date. Multi-purpose crawlers that combine Search with Training are handled according to all of their declared behaviors, so the most restrictive applicable rule governs.

5 "Introducing Pay Per Crawl: Enabling Content Owners to Charge AI Crawlers for Access," Cloudflare Blog, July 1, 2025, <https://blog.cloudflare.com/introducing-pay-per-crawl/>.

6 "Cloudflare's Pay Per Crawl: What It Means for Content Monetization," Eyeful Media, August 25, 2025, <https://www.eyefulmedia.com/blog/pay-per-crawl/>; Luis Rijo, "Media leaders expect 43% more traffic losses as AI reshapes journalism," PPC Land, January 17, 2026, <https://ppc.land/media-leaders-expect-43-more-traffic-losses-as-ai-reshapes-journalism/>. Eyeful Media notes that under new AI-bot monetization models, high-traffic websites can earn between \$50,000 and \$200,000 per month from crawler payments. Furthermore, industry data cited by PPC Land reveals that 69 percent of media leaders and publishers expect AI licensing to become a tangible revenue stream within three years.

7 "Usage Statistics and Market Share of Cloudflare," W3Techs, accessed August 19, 2026, <https://w3techs.com/technologies/details/cn-cloudflare/>; "Usage Statistics and Market Shares of Reverse Proxy Services," W3Techs, accessed August 19, 2026, <https://w3techs.com/technologies/overview/proxy/>; "Chart: Big Three Hold Dominant Lead in Accelerating Cloud Market," Statista, May 5, 2026, <https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/>. W3Techs records Cloudflare as the reverse proxy service for 24.4 percent of all websites, or 84.2 percent of websites whose reverse proxy service is identifiable. Statista, drawing on Synergy Research Group estimates, places AWS at 28 percent of the worldwide cloud infrastructure market in the first quarter of 2026. The two shares measure different layers and are not additive: a single site can sit behind Cloudflare and run on AWS.



venue. A proprietary payment system can restrict that access. If access, billing, and bot verification all run through one company's marketplace, the concentration publishers now oppose simply moves down one layer. The mechanism is settled. The terms are not.

One approach is x402, an open payment standard introduced by Coinbase in May 2025. It lets a website or API request payment within an ordinary web interaction, using the HTTP 402 "Payment Required" status code. A client requests a resource, receives a 402 response containing the price and payment requirements, and retries with a signed payment authorization. The server verifies the authorization, settles the payment directly or through a facilitator, and returns the resource. That shifts payment away from account creation and manual billing toward a flow native to the web itself. What matters for policy, though, is what the protocol does not do. It does not hard-code a settlement asset, network, or provider. The specification is designed to support multiple payment schemes, including immediate and deferred settlement, rather than a single rail. Actual support depends on which schemes, networks, and assets a given facilitator or server implements. It also means the protocol settles nothing about which rail should carry which transaction. That question is decided by the economics of the payment itself, and it is the subject of Section 4. The ecosystem now spans the largest cloud providers, the largest card networks, major AI labs, and Cloudflare, whose Pay Per Crawl architecture is built around the 402 flow. Optional facilitators can verify payment authorizations and submit approved payments for settlement on behalf of servers. A server can also perform those functions directly. For a creator, this provides a way to earn revenue from an individual use without negotiating a separate agreement with every buyer. A small publication can offer paid access alongside larger providers. Payment for access does not by itself resolve copyright questions or determine how the material may be reused, but it gives creators a practical way to charge for access they authorize.

The 402 response also changes what the agent knows before it acts. Because the price arrives in machine-readable form inside the model's context, cost becomes an input to tool selection rather than a line on a month-end bill the agent never sees. Two properties follow. The first is control. An agent working within a budget can compare the stated price with the value of the task and decline calls outside its mandate. That is easier to oversee than reconciling spending after the fact. The second is competition. When prices are visible before purchase, agents can compare providers of substitutable data and services in real time. That is the basis for genuine price discovery. Both properties depend on discovery as much as on payment. An agent can compare only the endpoints it can find. The standard therefore needs to carry discovery metadata, including price and schema information, alongside the payment flow so an agent can locate a priced endpoint rather than merely pay one. Section 7 returns to what this implies for competition policy.

Nothing in the standard requires stablecoins. That is what makes the market's choice noteworthy. High-frequency, programmatic, and intermediary activity accounts for a substantial share of stablecoin transfer volume and is concentrated in USDC and USDT.⁸ Stablecoins are already the leading dollar-denominated settlement instrument for software operating on public blockchains. Their adoption reflects the combination automated markets require: low transaction costs, continuous availability, and cross-border reach. The standard's authors designed x402 for this use case. They cite near-zero transaction costs, autonomous machine-to-machine payments, and worldwide availability without access to traditional financial rails.⁹ None of this argues for exclusivity. It establishes that the market has already identified what works best today, and Section 4 explains why the arithmetic makes that outcome unsurprising.

4 Legacy rails struggle to handle the smallest transactions

The case against legacy rails in this context is structural, not ideological. Cards, bank transfers, and existing payment intermediaries remain effective for many human transactions and some higher-value agentic transactions. They were not designed, however, for the core of agentic commerce.

4.1

The fixed-cost floor

Payment costs have two components: a percentage of transaction value and a fixed per-transaction fee. The percentage component scales down with the payment and is not an obstacle for agentic commerce. The fixed component is the problem because it does not scale. A \$0.30 fixed fee¹⁰ on a \$60 grocery basket is 50 basis points, a rounding error. The same \$0.30 on a \$0.001 API call is a 30,000 percent transaction cost.

The minimum viable transaction offers a useful comparison across rails. It is the smallest payment for which total processing costs stay at or below 5 percent of the payment's value, a generous ceiling when most merchants consider 3 percent painful. On card-not-present credit, that floor sits somewhere between \$5 and \$15. Card small-ticket programs

⁸ CEX.io, "Stablecoins in Q1 2026," CEX.io Blog, April 1, 2026, <https://blog.cex.io/ecosystem/q1-2026-stablecoin-report-35459>.

⁹ Erik Reppel, Ronnie Caspers, Kevin Leffew, Danny Organ, Dan Kim & Nemil Dalal, "x402: An open standard for internet-native payments," Coinbase Developer Platform, May 6, 2025, <https://www.x402.org/x402-white-paper.pdf>. The whitepaper lists "near-zero transaction costs that enable payments as low as" a tenth of a cent per request and "machine-to-machine transactions that allow IoT devices and AI agents to autonomously pay for resources" among the properties the standard is built to deliver, and states that because it runs on permissionless infrastructure, x402 "is available worldwide without requiring costly conversions or access to traditional financial rails." The protocol originated at Coinbase and has been stewarded by the x402 Foundation, hosted by the Linux Foundation, since April 2026.

¹⁰ Stripe. "Pricing." Accessed August 7, 2026. <https://stripe.com/pricing>. Stripe's published per-transaction fixed component is \$0.30 in the United States, 20 pence in the United Kingdom, and EUR 0.25 in the euro area; the argument in the text holds at any of these values.



bring it to roughly \$1.20. Debit lands near \$4.50 where interchange is capped, and bank transfer between \$5 and \$10. Even bank-operated instant payment systems, the fastest conventional rails available in most markets, sit near \$0.90. A stablecoin transfer on a low-cost chain or rollup carries a network cost measured in hundredths of a cent, which puts its minimum viable transaction between roughly two-tenths of a cent and two cents.¹¹

The gap is three to four orders of magnitude, not an increment. Even instant bank rails carry per-transaction costs roughly 45 to 450 times larger than an onchain stablecoin transfer.¹² Even when instant systems operate as public infrastructure and charge the end user little or nothing, the cost reappears elsewhere. These systems clear only within their own borders, settle between accounts rather than machines, and impose minimum transfer values well above the amounts at issue here. A \$0.001 payment for a single API call, data query, or article crawl is economically impossible on every legacy rail, but trivially cheap onchain.

The obvious rebuttal is that small payments can be batched: accumulate a month of sub-cent charges and settle once by card. That works when both parties have a durable billing relationship, which is what this market lacks. An agent performing a research task may touch two hundred services it has never used before and will never use again. Requiring an account or stored credential at each service recreates the friction the agentic model exists to remove. Batching also introduces credit risk, since someone fronts access before collecting, and a gatekeeper, since whoever nets the batch controls the market. Where batch settlement is preferred, as in Cloudflare's deferred-payment scheme for high-volume crawlers, stablecoins serve that mode as well. This asymmetry points to a key advantage of stablecoins in the agentic era: stablecoins settle both instantly and in batches, while cards settle only in batches.

The reason for the gap is structural rather than competitive. Card fees are not high merely because competition is insufficient. They price in chargeback rights, fraud screening, rewards funding, and a four-party settlement chain, and those costs cannot be competed away without removing the features that define the product. Onchain settlement

¹¹ Authors' analysis of published rate schedules. The minimum viable transaction is derived as $V = F / (c - r)$, where F is the fixed per-transaction fee, r the ad valorem rate, and c the 5 percent cost ceiling. Card-not-present credit: 2.0 percent plus \$0.15 under interchange-plus pricing to 2.9 percent plus \$0.30 under blended pricing, yielding \$5.00 to \$14.29. Card small-ticket: 1.65 percent plus \$0.04, yielding \$1.19. Debit: the Regulation II ceiling of 21 cents plus 5 basis points of transaction value, plus the 1 cent fraud-prevention adjustment, yielding \$4.44. Bank transfer: \$0.25 to \$0.50 per entry, yielding \$5.00 to \$10.00. Instant payment: the \$0.045 baseline FedNow credit-transfer fee, yielding \$0.90. Stablecoin transfer on a low-cost chain or rollup: network costs of \$0.0001 to \$0.001, yielding \$0.002 to \$0.02. See Stripe, "Pricing," accessed August 7, 2026, <https://stripe.com/pricing>; Visa, "Visa USA Interchange Reimbursement Fees," accessed August 19, 2026, <https://usa.visa.com/content/dam/VCOM/download/merchants/visa-usa-interchange-reimbursement-fees.pdf>; Debit Card Interchange Fees and Routing, 12 C.F.R. pt. 235 (Regulation II); Federal Reserve Banks, "FedACH Services 2026 Fee Schedule," accessed August 18, 2026, <https://www.frb-services.org/resources/fees/ach-2026>; Federal Reserve Banks, "FedNow Service 2026 Fee Schedule," accessed August 18, 2026, <https://www.frb-services.org/resources/fees/fednow-2026>; The Clearing House, "RTP Participant Fee Schedule," effective January 1, 2026. Figures reflect published U.S. schedules and are directionally conservative: each legacy rail is priced at its most favorable applicable tier.

¹² Federal Reserve Banks, "FedNow Service 2026 Fee Schedule," accessed August 18, 2026, <https://www.frb-services.org/resources/fees/fednow-2026>; "Stablecoin Rails vs Traditional Payment Rails: A Cost and Speed Comparison," Spark, May 15, 2026, <https://www.spark.money/research/stablecoin-payment-rails-vs-traditional>. The Federal Reserve's FedNow network charges a baseline direct fee of \$0.045 per credit transfer. By comparison, Spark's 2026 analysis notes that stablecoin transfers on high-throughput chains like Solana or Layer 2s execute for between \$0.0001 and \$0.001, making the fiat instant payment rail 45 to 450 times more expensive at the network level.



is cheaper because it carries fewer of them: no intermediary chain to compensate, no underwriting or identity check priced into each transaction, and no separate reconciliation step. That is also why the constraint is about size rather than about cards as such. A rail that prices in disputes, rewards, and a four-party settlement chain cannot be discounted enough to serve payments worth a fraction of a cent. The floor comes from the rail's features, not its margin. The same rail remains well suited to the larger, less frequent purchases an agent makes on its principal's behalf. What it cannot do is meter.

4.2

The underwriting wall

Traditional online payments require more than a fee. To accept card payments, a party needs a merchant account. Processors underwrite merchants against fraud and credit risk, typically requiring a legal entity, a processing history, a website, and often a personal guarantee.

Now consider the services that make up the supply side of this market: a solo developer's weather API, a local newspaper's archive endpoint, a sensor network's data feed, or an agent tool published to a registry last Tuesday. These sellers may have no entity, no processing history, and no storefront. The long tail of agent-payable services is not merely expensive to serve on card rails. It is unbankable on them.

Receiving a stablecoin payment requires an address. The agent side still needs provisioning, since a paying agent must hold a wallet and be authorized to spend from it. But the underwriting question moves to where it belongs: the identity and authorization of the paying agent. Verifiable agent-identity frameworks (Section 6) can address that risk without gatekeeping the service on the receiving side.

4.3

The dispute model mismatch

Card-network dispute procedures cover credit-card and many debit-card payments, allowing transactions to be contested after the merchant has been paid. The grounds and time limits vary by network, transaction type, and jurisdiction. Major processors charge at least \$15 in fees, while industry estimates put the merchant's all-in cost near \$128 after internal handling and third-party fees.¹³ That cost is often many times the value of the transaction. For human commerce this is the consumer-protection layer. For sub-cent agentic commerce it is a denial-of-service vector. No rational service sells a \$0.0005 data query if a single dispute can cost thousands of times the sale price, and no platform underwrites billions of daily transactions each carrying open-ended reversal risk.

Onchain confirmation removes card-style chargeback exposure and produces a durable transaction record. It does not by itself guarantee delivery, redemption, or legal finality outside the network. This is a property of the settlement rail, not x402 itself. If an x402 payment scheme were implemented over card rails, it would inherit their chargeback exposure. Whether the principal is worse off for losing that reversal right is a question we address in Section 5.

¹³ Mastercard, "What's the True Cost of a Chargeback for Businesses?" accessed August 19, 2026, <https://www.mastercard.com/global/en/news-and-trends/insights/2025/what-s-the-true-cost-of-a-chargeback-in-2025.html> citing Mastercard and Javelin, Chargebacks: The Case for Coordination (2026). Merchants incur an average of \$82 in internal costs and \$46 in third-party fees per chargeback, excluding the cost of goods or services already delivered. Processor dispute fees begin at \$15; see Stripe, "Disputes and Fraud," accessed August 19, 2026, <https://docs.stripe.com/disputes>.

4.4
The identity
and geography
mismatch

Card and bank rails authenticate humans and firms through names, tax identifiers, and addresses. There is no native primitive for “this autonomous agent, acting under this principal’s authority, within these spending limits.” The industry is building new cryptographic layers because the legacy identity model does not extend to machines. These include trusted-agent protocols from the card networks, Web Bot Auth message signatures, and agent-mandate standards. Public-key cryptography is native to blockchain rails and a retrofit everywhere else.

Geography compounds this. Bank transfer and instant payment systems are domestic by design, each operating within one jurisdiction’s banking perimeter. This market is not domestic: the crawler is in Singapore, the model is in Virginia, the publisher is in Ohio, and the sensor network is in Nairobi. Cross-border card transactions add fees and foreign-exchange spreads that widen the fixed-cost floor further, and correspondent banking settles in days rather than seconds. A dollar stablecoin can settle across borders and around the clock at low network cost, while denominating the transaction in U.S. dollars. This capability is not limited to dollars. A compatible wallet can hold stablecoins denominated in different currencies alongside other digital assets. An x402-enabled service can specify which supported assets and networks it accepts, allowing an agent to pay different providers in different currencies through a common payment protocol. Actual use still depends on wallet and payment-service support, liquidity where conversion is needed, and applicable legal requirements.

4.5
Comparative
summary

The table below^{14,15} applies a single test to each option: what it costs to move a very small amount of value between two parties that have no prior relationship. Four properties decide the outcome: transaction pricing, confirmation and settlement timing, reversibility, and geographic reach without local setup on either side. The pattern across the rows is consistent. Rails built for human commerce fail this test on at least one property once the payment is worth a fraction of a cent, and they usually fail first on the fixed fee quantified in Section 4.1.

14 Authors’ analysis synthesizing traditional and digital payment network architectures. Credit card and L1/L2 structural comparisons draw from Yuxin Xiang et al., “SoK: Stablecoins in Retail Payments,” arXiv preprint arXiv:2601.00196 (January 2026), <https://arxiv.org/abs/2601.00196>. Domestic bank rail constraints are based on Bank for International Settlements (BIS) Committee on Payments and Market Infrastructures, Cross-border Linking of Fast Payment Systems (October 2024), <https://www.bis.org/cpmi/publ/d223.htm>. Agentic micropayment capabilities reflect recent developer frameworks, specifically Circle, “How to Build Agentic AI Payments Systems on Arc,” May 8, 2026, <https://www.circle.com/blog/build-agentic-systems-for-high-frequency-sub-cent-transactions>, and Eco, “Why AI Agents Need Stablecoin Payments,” July 8, 2026, <https://eco.com/support/en/articles/14846271-why-ai-agents-need-stablecoin-payments>.

15 Device wallets such as Apple Pay and Google Pay are not a separate rail for this purpose. They use a tokenized card credential and therefore inherit the first row’s cost profile and reversal mechanics. The stablecoin row reports network-level transfer costs and onchain confirmation. End-to-end costs and timing may also depend on custody, compliance, liquidity, foreign exchange, redemption, and conversion infrastructure. For the transaction pattern examined here, however, those costs can be amortized across many payments. Stablecoins therefore retain a substantially lower marginal cost per transaction than conventional payment rails. For low-value transactions, a seller may deliver the resource after an initial confirmation rather than wait for stronger finality. That choice reflects the seller’s tolerance for reversal or reorganization risk; it does not change the network’s finality guarantees.

Table 1: Rail comparisons for small agentic transactions

Payment rail	Typical cost profile	Confirmation/ settlement finality	Reversal/ chargeback risk	Geographic reach	Fit for agentic commerce
Credit cards	High fixed plus percentage fee; often uneconomical for very small payments	Authorization now, settlement later	Yes; reversals can remain open for months	Broad, but intermediary heavy	Good for infrequent, higher-value purchases; poor fit for micropayments
Payment platforms (PayPal, hosted checkout)	Percentage fee plus markup	Mixed; often delayed settlement	Usually yes	Broad, but platform-specific	Better for human checkout than open machine-to-machine commerce
Domestic bank rails / instant payment systems	Lower cost than cards, but not designed for cross-border micropayments	Scheme dependent; often domestic	Lower than cards, but not machine-native	Mostly domestic or regional	Useful for known counterparties and larger repeat flows
Public-chain settlement on L1	Global and open, but often too expensive for routine micropayments	Initial confirmation typically takes seconds to minutes; finality depends on the network and confirmation requirements.	No card-style reversals	Global	Useful in some contexts, but often too costly for high-frequency pay-per-use
Stablecoins over low-cost chains / L2s	Near-zero or nominal network cost	Initial confirmation often arrives within seconds; stronger finality may take longer, depending on the network.	No chargebacks; onchain settlement	Global, 24/7	Best current fit for low-value, cross-border agentic commerce

The point is not that one token or protocol should monopolize agentic payments. The core of agentic commerce requires settlement that is low-cost, programmable, and globally accessible. Today, stablecoins operating on open, interoperable infrastructure best meet that requirement.

5

Stablecoins are best suited to high-frequency, pay-per-use agentic transactions

Stablecoins fit this market for five reasons.

First, on low-cost networks, stablecoin transfers avoid the fixed per-payment charges that make pay-per-use uneconomic on legacy rails. Second, payment and delivery can occur within the same automated flow, sharply reducing the settlement exposure either party must carry. Third, stablecoins operate on programmable infrastructure, allowing wallets, smart contracts, and authorization schemes to enforce spend caps, allowlists, and signed mandates before payment. Fourth, they can move across borders and around the clock, subject to the availability of compliant wallets, liquidity, and conversion infrastructure. Fifth, the same wallet that spends can also receive, which is what makes an autonomous economic actor possible rather than merely an autonomous spender.

That fifth property is the one most easily overlooked, because it is not about what a payment costs but about who is permitted to charge for anything at all. An agent can be issued a card credential and begin spending in minutes. Accepting card payments is different. It requires a legal entity, merchant underwriting, and a reconciliation process. As a result, someone else's advance credit judgment determines who has the right to sell. The long-tail supply side this paper defends sits on the wrong side of that judgment. The solo developer's API, the local newspaper's archive endpoint, and the sensor network's data feed can all be metered and all be paid for, but none of them can be underwritten as a merchant. A single address that both receives and disburses closes the loop without any counterparty's approval, and closing that loop is what turns a service into a seller. The same problem appears one layer further down the stack in the proprietary agentic marketplaces discussed in the third policy recommendation. Merchant underwriting and marketplace membership both determine who may participate in a market. Neither follows from the economics of the transaction itself. The same wallet could also receive a refund or a token used to demonstrate access rights, where the wallet and service support it. These are possible applications of a wallet's ability to send and receive assets, not capabilities guaranteed by x402 itself.

Stablecoins are not the only settlement mechanism for agentic commerce, and higher-value, lower-frequency purchases between established counterparties will remain well suited to cards and other familiar rails, including the dispute protections those rails provide. The claim is about the core of the market rather than all of it: the high-frequency, low-value transaction between parties meeting for the first time, which is precisely the transaction existing rails cannot price.



The dispute question is worth considering. If an agent buys something defective on its principal's behalf, why would that principal not want the reversal right a card provides? For a discrete purchase of meaningful size, it would, and an open standard lets that purchase run on a card rail. The metered core of this market is different in kind. Processing a single dispute starts at \$15 and costs closer to \$128 all in, as Section 4.3 explains. No remedy process can be run over a \$0.0005 query, and complaints are rarely about non-delivery. They are about quality across thousands of calls. Two substitutes do that work instead. Spend limits and signed authorizations cap the loss before it occurs. Exit is also immediate: an agent can drop an unreliable provider on the next call rather than wait until the end of a billing cycle. When a purchase warrants conditional settlement, the protocol can require escrow without imposing a reversal right on every transaction. For very small purchases, spending controls and the ability to stop future payments may be more practical than resolving each dispute individually.

The current composition of the market also matters. Stablecoin activity is heavily concentrated in dollar-denominated tokens and in regulated issuance. That is an important descriptive fact, not a reason to design policy around exclusivity. Good policy should preserve interoperability so that multiple compliant stablecoins and settlement schemes can develop. It should not prematurely narrow the market to a single issuer, token, or chain.

Example: An illustration of one agentic transaction

The scenario

A logistics operator's routing agent needs local rainfall data for a corridor that no major weather provider covers well. An independent sensor network has the data and would sell a verified reading for a fraction of a cent. The two cannot transact on the rails described in Section 4. The operator has no merchant account, the agent has no billing relationship with it, and creating either would cost more than the sale is worth.

The flow

Under an HTTP-native payment flow, the exchange proceeds as follows.

- The agent requests the record over ordinary HTTP.
- The server answers with a 402 status and a machine-readable payment specification: the amount, the recipient, and the settlement schemes it accepts. The seller sets the price for each endpoint. The protocol sets no price, takes no cut, and provides no haggling step for the two sides to reach a figure.
- The price arrives in the same response that carries the data's terms, before the agent commits. The agent compares the price with the policy its principal set in advance: a per-transaction ceiling, a task budget, and permitted categories of service. It then weighs a fraction of a cent against the value of a more accurate quote. Cost becomes an input to tool selection rather than a line on a bill the agent never sees.
- Judging the purchase worth making, the wallet attaches a signed authorization and retries. The signature replaces enrollment. In the API-key model, identity is server-side state: a row someone provisions in the provider's database. Here, identity is client-side proof that the agent carries with it. That distinction allows a first-contact sale without an account.

- A facilitator verifies the signed payment authorization and submits the approved payment for settlement. Today, the Coinbase Developer Platform operates production endpoints for both functions. The role is not privileged: any party can run a facilitator, other providers build to the same standard, and a server can verify and settle directly without using one at all.
- After receiving the required payment confirmation, the server returns the record. In this example, the complete payment-and-delivery flow takes about two seconds.

Results and implications

No account was created, no credential stored, no invoice issued, no terms negotiated. The parties formed the commercial relationship and completed the transaction inside a single stateless challenge-and-retry flow. Network costs were measured in hundredths of a cent.

Four features bear on the argument made elsewhere in this paper. This is a first-contact sale between parties with no history and no expected future. Section 4.1 shows that aggregation cannot serve this case because batching presumes a durable billing relationship. The seller needed an address rather than underwriting, the inversion described in Section 4.2: merchant acceptance rations who may sell, while one wallet that both receives and disburses removes that gate. Oversight ran before the payment rather than after it, following the ex ante control model described in Section 1 and applied in Section 5. Withdrawing authority means rotating or de-funding a key rather than reclaiming a credential from dozens of provider databases. A price visible at request time also lets agents compare providers as they work, the competition property the third policy recommendation rests on.

The transaction illustrates the paper's narrower claim. Had the agent renewed an annual software license, conventional rails and their dispute and rewards machinery would serve that purchase well. The sensor reading is the class of exchange with no incumbent option: too small to price on legacy infrastructure, too fleeting to justify onboarding, too dispersed for domestic instant payment systems. Nothing in the flow is specific to one asset or chain. Stablecoins carry most of this traffic because they satisfy its cost, speed, and reach requirements at once, not because the standard requires them.

x402 payment flow

Logistics routing agent  Weather sensor network

HTTP-native payments



x402 · HTTP-native payment handshake for APIs & agents

6

Identity and interoperability decide whether payments scale

Payments alone do not solve the problem. The internet also needs to identify which agent is acting, whose authority it carries, and what limits apply. That need grows as agents delegate and transact through new protocols such as Agent-to-Agent (A2A), Agent Payments Protocol (AP2), and x402. Know-Your-Agent frameworks provide that missing identity layer.

Agent identity should bind cryptographic activity to accountable principals in proportion to the risk and authority involved, without recreating full human onboarding for every microtransaction. A low-risk research agent spending a small monthly budget should not face the same compliance burden as an agent moving significant value or taking consequential actions. Policymakers should therefore require levels of assurance that rise with the risk of harm rather than impose one-size-fits-all treatment.

Interoperability requires portable credentials. An agent verified once should be able to present that verification to any service, wallet, or marketplace it transacts with, rather than repeating onboarding at each one. Identity needs the same portability as payments. If verification must be repeated for every counterparty, the identity layer restores the setup costs that open payment rails remove. A fragmented identity stack also recreates the gatekeeping those rails are meant to avoid.

Portable does not have to mean universally identifying. A credential presented everywhere becomes a tracking identifier. That outcome carries the same risks as concentration. The workable design uses attestation rather than identification. A server learns that an agent acts for an accountable principal within stated limits, but nothing beyond what the transaction requires. Agent-credential standards should be judged by whether they combine portability with minimum disclosure, not whether they produce a single registry.

This is also why open standards matter beyond payments alone. Interest in open-weight AI models reflects concern about concentration in the AI stack. Open payment and identity protocols apply the same principle to the stack's financial and commercial layers.

7

Policymakers should preserve the economics, not force the categories

This market needs surprisingly little from policymakers, but it does need the right kind of restraint. Policymakers should preserve the economics of digital micropayments while ensuring accountability, competition, and interoperability. Five recommendations follow.

7.1 Preserve micropayment economics

This recommendation addresses regulatory obligations, not commercial pricing. Rules requiring per-transaction reporting, recordkeeping, disclosure, identity verification, or screening are calibrated for remittances, payroll, and consumer card purchases. For those payments, a few cents of compliance cost per transaction is immaterial. Applied unchanged to a payment worth a fraction of a cent, the same duty costs more than the transaction it governs and operates as a prohibition rather than a safeguard. Where an obligation must apply, it should attach to the account, wallet, or relationship that can bear it, not to each transfer. Its cost should scale with the risk addressed rather than the number of payments made.

7.2 Build compliance around identity, authorization, and wallet controls

Oversight should target the systems and controls maintained by issuers, wallet providers, and account providers rather than manual review of each individual payment. If an obligation is too costly to impose on each transaction, it can instead be met at the account, wallet, or customer-relationship level. Automated monitoring can clear low-risk transactions without friction while flagging suspicious or higher-risk payments for review, which is increasingly how existing payments regimes already operate. Assurance should also scale with authority. An agent making sub-cent data purchases within a \$5 monthly budget does not warrant the same diligence as one moving five figures. A regime that cannot distinguish between them will overburden the first or underpolice the second. Agentic commerce requires this supervision to be programmable. A principal's limits and a provider's controls should be enforced at the moment of payment, not reconstructed afterward.

7.3 Treat open protocols as competition policy

The main risk is not that this market grows too quickly. It is that it hardens into a small number of proprietary marketplaces that meter access, identity, and payment on closed terms, restricting competition, access, and cross-compatibility. Open, interoperable standards supporting a range of asset types, issuers, and settlement rails are the best safeguard against that outcome. Two asks follow from that. No rule should require agentic payments to pass through licensed marketplaces rather than licensed issuers. Supervising who issues money is not the same as deciding who may sell. Public engagement in standards bodies should favor open protocols over single-vendor designs, including on discovery. A standard that lets an agent pay an endpoint but not find one keeps payment open while concentrating the market at the directory layer. A framework that can carry compliant digital settlement across jurisdictions and use cases will hold up even if dollar stablecoins remain the leading instruments in the near term.

7.4

Recognize per-use payment as creator compensation

Section 2 described the collapse of the referral bargain as a pricing failure and not just an intellectual-property dispute. That framing has a policy consequence. Legislatures face mounting pressure to resolve the AI content standoff through compulsory licensing, statutory royalties, or expanded liability. These instruments take years to litigate and tend to reward parties large enough to litigate them. A functioning per-use payment market reaches further and sooner. It lets publishers earn revenue when agents purchase access, including publishers without the leverage to negotiate an enterprise contract. A local newspaper's archive can participate on the same terms as a national brand. Policymakers who want AI systems to compensate journalists, authors, and developers should treat the payment layer as the mechanism.

7.5

Coordinate internationally before proprietary systems set the defaults

The market for the second customer is cross-border by nature, and its norms are being set now, in venues that already exist. Work is already underway through the G20 cross-border payments agenda, the Committee on Payments and Market Infrastructures and IOSCO, and the IETF's web-authentication track. A consistent position is available across all three: open protocols over single-vendor designs, compliance discharged at the identity and wallet layer, portable credentials with minimum disclosure, and room for multiple compliant settlement schemes rather than incompatible national or platform-specific silos. Delay is not neutral here. Each quarter without a common position gives proprietary systems more time to become entrenched. Once adopted, payment infrastructure is difficult to displace.

8

Conclusion: The payments layer decides who participates

The internet's next commercial chapter will be written by humans and by software acting on their behalf. That software must be able to pay. The web and the agent face the same problem from opposite sides. A publisher cannot charge for a request unless the requester can pay the asking price. A price produces no revenue if no rail can carry the payment. This paper has concentrated on the payment side because that is the side that does not work. Demand is already visible in the breakdown of the old referral bargain, in the spread of pay-per-use experiments, and in the emergence of HTTP-native payment standards. The obstacle is a payment layer whose fixed costs, onboarding requirements, and reversal machinery were built for a human buying something substantial from a merchant who has been underwritten.

Stablecoins provide the largest part of the answer available today, though other settlement mechanisms still have a role. Stablecoins make it possible to price content, data, and services at the moment of use, across borders, for buyers who are software. Whether this layer develops on open, interoperable standards or inside a small number of proprietary marketplaces will shape agentic commerce. That choice will determine whether the market avoids or inherits the fragmentation and weak competition of the systems it replaces. Policymakers who preserve room for open standards, portable credentials, and multiple compliant settlement schemes will get a more competitive and more durable commercial foundation for the open internet. Those who do not will get the same gatekeepers, metering a larger market.

References

Bank for International Settlements Committee on Payments and Market Infrastructures. Cross-border Linking of Fast Payment Systems. Basel: BIS, October 2024. <https://www.bis.org/cpmi/publ/d223.htm>.

CEX.io. "Stablecoins in Q1 2026: Rising Similarities With the 2022 Bear Market." CEX.io Blog, April 1, 2026. <https://blog.cex.io/ecosystem/q1-2026-stablecoin-report-35459>.

Circle. "How to Build Agentic AI Payments Systems on Arc." May 8, 2026. <https://www.circle.com/blog/build-agentic-systems-for-high-frequency-sub-cent-transactions>.

The Clearing House. "RTP Participant Fee Schedule." Effective January 1, 2026.

Cloudflare Blog. "Introducing Pay Per Crawl: Enabling Content Owners to Charge AI Crawlers for Access." July 1, 2025. <https://blog.cloudflare.com/introducing-pay-per-crawl/>.

Cloudflare Blog. "Your Site, Your Rules: New AI Traffic Options for All Customers." July 1, 2026. <https://blog.cloudflare.com/content-independence-day-ai-options/>.

Cloudflare Radar. "Traffic Worldwide: Bot vs. Human." Accessed August 18, 2026. <https://radar.cloudflare.com/traffic#bot-vs-human>.

Debit Card Interchange Fees and Routing. 12 C.F.R. pt. 235 (Regulation II).

Eco. "Why AI Agents Need Stablecoin Payments." July 8, 2026. <https://eco.com/support/en/articles/14846271-why-ai-agents-need-stablecoin-payments>.

Eyeiful Media. "Cloudflare's Pay Per Crawl: What It Means for Content Monetization." August 25, 2025. <https://www.eyefulmedia.com/blog/pay-per-crawl>.

Federal Reserve Banks. "FedACH Services 2026 Fee Schedule." Accessed August 18, 2026. <https://www.frb-services.org/resources/fees/ach-2026>.

Federal Reserve Banks. "FedNow Service 2026 Fee Schedule." Accessed August 18, 2026. <https://www.frb-services.org/resources/fees/fednow-2026>.

Mastercard. "What's the True Cost of a Chargeback for Businesses?" Accessed August 19, 2026. Citing Mastercard and Javelin, Chargebacks: The Case for Coordination, 2026. <https://www.mastercard.com/global/en/news-and-trends/insights/2025/what-s-the-true-cost-of-a-chargeback-in-2025.html>.

Newman, Nic. Journalism, Media, and Technology Trends and Predictions 2026. Oxford: Reuters Institute for the Study of Journalism, January 2026.

Reppel, Erik, Ronnie Caspers, Kevin Leffew, Danny Organ, Dan Kim, and Nemil Dalal. "x402: An Open Standard for Internet-Native Payments." Coinbase Developer Platform, May 6, 2025. <https://www.x402.org/x402-whitepaper.pdf>.

Rijo, Luis. "Media Leaders Expect 43% More Traffic Losses as AI Reshapes Journalism." PPC Land, January 17, 2026. <https://ppc.land/media-leaders-expect-43-more-traffic-losses-as-ai-reshapes-journalism/>.

Spark. "Stablecoin Rails vs Traditional Payment Rails: A Cost and Speed Comparison." May 15, 2026.

<https://www.spark.money/research/stablecoin-payment-rails-vs-traditional>.

Statista. "Big Three Hold Dominant Lead in Accelerating Cloud Market." May 5, 2026. <https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/>.

Stripe. "Disputes and Fraud." Accessed August 19, 2026. <https://docs.stripe.com/disputes>.

Stripe. "Pricing." Accessed August 7, 2026. <https://stripe.com/pricing>.

Tobitt, Charlotte. "Global Publisher Google Traffic Dropped by a Third in 2025." Press Gazette, January 16, 2026. <https://pressgazette.co.uk/media-audience-and-business-data/google-traffic-down-2025-trends-report-2026/>.

Tomé, João. "The Crawl-to-Click Gap: Cloudflare Data on AI Bots, Training, and Referrals." Cloudflare Blog, August 29, 2025. <https://blog.cloudflare.com/crawlers-click-ai-bots-training/>.

Visa. "Stablecoins and the Future of Onchain Finance." Accessed August 18, 2026. <https://corporate.visa.com/en/solutions/crypto/stablecoins/stablecoins-and-the-future-of-onchain-finance.html>.

Visa. "Visa USA Interchange Reimbursement Fees." Accessed August 19, 2026. <https://usa.visa.com/content/dam/VCOM/download/merchants/visa-usa-interchange-reimbursement-fees.pdf>.

W3Techs. "Usage Statistics and Market Share of Cloudflare." Accessed August 19, 2026. <https://w3techs.com/technologies/details/cn-cloudflare>.

W3Techs. "Usage Statistics and Market Shares of Reverse Proxy Services." Accessed August 19, 2026. <https://w3techs.com/technologies/overview/proxy>.

Xiang, Yuexin, et al. "SoK: Stablecoins in Retail Payments." arXiv preprint arXiv:2601.00196 (January 2026). <https://arxiv.org/abs/2601.00196>.