Feedback on the Discussion Draft of the

Artificial Intelligence Access, Gatekeeper Exchange, and Nondiscriminatory Transfer Act of 2026

(AI AGENT Act)

Submitted to the Office of Senator Mark R. Warner

August 28, 2026

Submitted by: Wendy MacKinnon

Title: Founder and Chief Executive Officer, Twinvocate Inc.

Location: Toronto, Ontario, Canada

Contact: wendy@twinvocate.ai

Summary

This submission supports the central construct of the discussion draft and recommends six changes intended to make that construct enforceable in practice. Each recommendation is anchored to a section of the draft.

The Custodial User Agent is the right unit of regulation. Defining a class of software by whose interests it must serve, rather than by what it is technically capable of, is the most important choice in the draft. It is also the choice most likely to be contested by parties whose commercial interest lies in agents that appear to represent the user while answering to someone else. The definition in Section 2 and the duties in Section 3(g) should be defended without dilution, and the non-waivability provision in Section 3(h) is essential to both.

The recommendations that follow address a structural asymmetry. The framework makes the agent legible and accountable to the platform. It does not make the platform legible to the agent. A registered Custodial User Agent bearing the fiduciary-like duties of Section 3(g) is nonetheless unable to verify the party on the other side of a transaction, and therefore unable to discharge those duties at the moment it matters most. Every protocol currently deployed in agentic commerce runs in the same single direction. Without correction, the Act would codify that asymmetry rather than remedy it.

A second asymmetry appears in Section 4(g). The interagency working group is directed to develop proposals preventing harms to businesses and the Government arising from Custodial User Agent fraud, misuse, or genuine mistake. The user, whose interests the Act exists to protect, is absent from that provision. Consumer harm from genuine agent mistake is the most likely category of loss under this framework and it currently has no home in the draft.

Interest of the submitter

I am the founder and chief executive of Twinvocate Inc., a company building a customer-loyal financial agent and a certified-counterparty credentialing layer for agentic commerce. I am a United States citizen, and I write from Toronto, where I also hold Canadian citizenship.

My background spans approximately thirty years in retail technology, payments and digital strategy. I founded SelfPay by Digital Retail Apps in 2011, and hold multiple United States patents in mobile scan-and-pay and purchase verification. I co-founded Paytechs Canada in 2019, now FinTechs Canada. I completed the coursework for a Master of Laws at the University of Toronto Faculty of Law in June 2026, concentrating in innovation and technology law, with the degree to be conferred in October 2026.

I have a commercial interest in this area, which I state plainly. The recommendations below would apply equally to my company and to its competitors, and the principal recommendation asks for an obligation binding parties on both sides of any transaction my company would facilitate.

The cross-border vantage is relevant rather than incidental. Canada has published proposed regulations under its consumer-driven banking framework for comment, the United Kingdom is consulting on the design of its next-generation retail payments infrastructure, and the European Union has already legislated portions of the problem this draft addresses. Where those regimes have resolved questions the draft leaves open, or reached them by a different route, I have noted it.

Part One: Support for the Custodial User Agent construct

Section 2 defines a Custodial User Agent as a software-based agent expressly authorised by a user to interact with a large online platform provider on that user's behalf in a transparent, documented, scope-limited and revocable manner. Four elements of that definition deserve specific endorsement.

- Authorisation is express rather than inferred from use of a service. This forecloses the pattern by which platform terms of service are treated as blanket consent to agent action.
- Scope limitation makes the mandate a bounded instrument rather than a general delegation. This aligns with the long-standing distinction in agency law between actual and apparent authority, and it gives a court something concrete to construe.
- Revocability preserves the principal's control, which is the feature distinguishing an agent from a successor in interest. Section 3(c) sensibly extends this by requiring that revocation be communicable to the platform.
- Documentation makes the mandate evidentiary. A mandate that is not recorded cannot be enforced, disputed or audited.

The duties in Section 3(g) are correctly pitched. The prohibition on actions benefiting the agent to the user's detriment, the ordinarily prudent person standard of care, the real-time records requirement, and the bar on onward delegation without express and revocable consent together establish something close to a fiduciary standard for consumer software. That standard should not be softened to good faith or commercial reasonableness during negotiation.

Section 3(h), which prevents waiver or limitation of those duties by any mechanism, is the provision that makes the rest of the framework real. Without it, the duties would be contracted away in terms of service on the day the Act took effect. It should be retained exactly as drafted.

The interoperability obligations in Sections 3(b) and 3(i) address a real and near-term risk. Platforms building native agents have a clear economic incentive to degrade or foreclose third-party agents, and the resulting market would offer consumers a choice among agents that all belong to the counterparty. The functional equivalence requirement, and the requirement that a platform operating its own interoperable agent extend equivalent access to third parties, are the operative provisions and should be preserved.

One provision warrants closer attention. Section 3(j) contemplates affiliate links as a permissible business model so long as the agent has not subordinated the consumer's interests to its own. Affiliate compensation is a structural subordination risk rather than an occasional one, because it makes the agent's revenue depend on which counterparty it selects. If affiliate models are to be permitted, they should carry an affirmative disclosure obligation to the user at the point of recommendation, not merely a general prohibition that will be litigated after the fact.

Part Two: Authorisation runs in only one direction

The gap in Sections 3(c) and 4(c)

Section 3(c) directs the Federal Trade Commission to ensure that the process by which a Custodial User Agent gains access to a large online platform is secure and properly authenticated. Section 4(c) directs the National Institute of Standards and Technology to identify open protocols, or develop model technical standards, to make popular classes of online service accessible to Custodial User Agents, including protocols for verifiable delegation, verification of agent registration, real-time revocation, and auditable action records.

In both provisions the verification runs from the agent outward. The agent proves itself, its registration, and the scope of its mandate to the platform. Nothing in the draft enables the reverse. A Custodial User Agent cannot establish, through any standardised mechanism, who the counterparty is, what that counterparty is authorised to request, or whether its representations to an agent correspond to its representations to a human.

This is not a theoretical gap. It reflects the current state of deployed infrastructure. Bot verification frameworks, the trusted agent protocols developed by card networks with content delivery providers, intent verification schemes, and issuer-operated agent registries all address the same question, which is whether the merchant should trust the agent. No deployed protocol addresses whether the customer's agent should trust the merchant.

Why it defeats the duties in Section 3(g)

Section 3(g) requires that a Custodial User Agent avoid actions resulting in reasonably foreseeable harm to the user and act with the care, skill and diligence of an ordinarily prudent person in like circumstances. Both formulations presuppose that the agent can evaluate the transaction it is entering.

Where the agent cannot verify the counterparty, the standard attaches to a determination the agent has no mechanism to make. Three consequences follow. First, the duty becomes unenforceable in precisely the cases where harm occurs, because the agent cannot be faulted for failing to detect what it had no means of detecting. Second, liability drifts back toward the user, who authorised the agent and now bears the loss. Third, the market receives a signal that counterparty risk is not a regulated concern, which discourages investment in solving it.

The draft already contains the right machinery

Section 3(i) requires a large online platform provider to set privacy and security standards for Custodial User Agent access, to file those standards with the Commission, and provides that the filed standards shall be publicly available. That is a public, platform-side filing obligation. It is the natural home for the reciprocal direction.

Extending that filing to include what a platform represents to agents, and requiring that the resulting file be machine-readable, would create a counterparty register queryable by an agent at transaction time rather than consulted by a compliance function afterwards. No new architecture is required. The obligation, the filing mechanism and the publication requirement all already exist in Section 3(i).

Differential disclosure and Section 3(k)

Section 3(k) provides that nothing in Section 3 confers greater rights of access for a Custodial User Agent than are accessible to a user. That is a sensible ceiling. It is not a floor.

As merchants and platforms restructure product and service data for machine readability, the information presented to an agent and the information presented to a human need not be identical. A party may reasonably conclude that certain terms, conditions or limitations are best omitted from the agent-readable representation. Existing deceptive practice authority reaches the clearest cases. It does not reach the harder ones, where the agent-readable representation is accurate but materially less complete than what a human reader would encounter.

A framework that mandates agent access without addressing what must be disclosed to agents creates an adversarial data surface where none previously existed. Section 3(k) should be clarified to make access parity bidirectional: a platform may not present to a Custodial User Agent information that is materially different from, or materially less complete than, what it presents to a user.

Part Three: The European precedent and interoperability

The European Union has legislated the reciprocal direction. Under Article 5b of the amended eIDAS Regulation, a party intending to request credentials from a citizen's digital identity wallet must register in the Member State where it is established. That registration must include the party's identity, contact details, and the intended use, including an indication of the data it will request. Member States must publish that information online in electronically signed or sealed form suitable for automated processing. A relying party may not request attributes outside its registered scope.

Three features are relevant. The register covers the parties asking for information, not only the parties supplying it. The permitted scope of a request is fixed in advance rather than negotiated at the point of transaction. And the register is machine-readable, so it can be queried by software at transaction time.

The United Kingdom has taken a different path so far. Its statutory register under the Data (Use and Access) Act 2025 covers digital verification service providers rather than relying parties. The question is now open there: the Bank of England's current consultation on future retail payments infrastructure asks what identity frameworks and credentials should support retail

payment journeys, and separately what capabilities agentic payments require, and it names stronger payee and merchant identification among the fraud capabilities under consideration.

The competitiveness implication is direct. A United States framework that registers and certifies agents under Section 3(d) but imposes no corresponding obligation on counterparties will interoperate with the European regime in one direction only. American Custodial User Agents operating in Europe will be able to query a public counterparty register that has no domestic equivalent, while European relying parties entering the United States market will encounter no comparable obligation. This is an argument for symmetry that does not require adopting the European approach.

A near-term domestic occasion for the same question

The interoperability argument is not confined to agents operating abroad. In August 2026 The Clearing House said it intends to extend the RTP network across international borders, beginning as a pilot with early adopter banks in the first half of 2027 and with the Bank of New York Mellon among the first participants. The capability that makes this possible is commonly described as one leg out, meaning a payment in which only one side sits within the domestic system and its rules. The Federal Reserve is considering the same step for FedNow, and the two are expected to enable it at different times.

The consequence for this draft is direct. A Custodial User Agent acting for a United States user will, from 2027, be able to initiate payments over domestic real-time infrastructure to counterparties that no United States register describes and no United States regulator has evaluated. Section 3(d) will have registered and certified the agent. Nothing will have been established about the party receiving the money. If the counterparty direction is left unaddressed, the asymmetry the draft creates domestically is exported at the moment the rails open, and the register the agent most needs is the one that does not exist.

Part Four: Section 4(g) omits the user

Section 4(g) directs the Commission to develop proposals to prevent harms to businesses and the Government resulting from Custodial User Agents undertaking actions on behalf of an individual arising from fraud, misuse, or genuine mistake, and establishes an interagency working group spanning the Consumer Financial Protection Bureau, the Federal Deposit Insurance Corporation, the Office of the Comptroller of the Currency, Treasury, Homeland Security, Commerce, Justice and the Securities and Exchange Commission.

The category of harm identified is the right one. Genuine mistake, as distinct from fraud or misuse, is the loss most likely to characterise this market, and convening those agencies is the correct response to it. The omission is the beneficiary. The provision protects businesses and the Government. It does not reach the individual on whose behalf the agent acted and who, under existing machinery, will bear the loss.

That machinery cannot presently recognise the harm. In the payment dispute system there is no reason code corresponding to an agent purchasing something other than what the user instructed. A user in that position must characterise the dispute as a purchase not as described or not authorised. Under current network rules a merchant holding evidence of delivery of the item actually ordered will prevail. The user bears the loss, and the dispute is recorded under a category that does not describe what happened.

The consequences compound. Losses arising from agent mistake are systematically miscoded, so the aggregate is invisible in any dataset a regulator would consult. Rulemaking is conventionally triggered by quantified harm. Harm that cannot be categorised cannot be quantified, and harm that cannot be quantified does not trigger rulemaking. The dispute taxonomy therefore operates as a barrier to the evidence that would justify further intervention.

A related structural point. The parties to a payment dispute are the cardholder, the issuer, the merchant, the acquirer and the network. Custodial User Agent providers are not parties. Liability cannot be allocated to an actor with no standing in the proceeding, and evidentiary artifacts such as intent mandates are of limited use when submitted to a forum with no jurisdiction over the party whose conduct is in question.

The dispute problem worsens as agentic payments move to real-time rails

The account above describes the card dispute system, where the harm is at least miscoded rather than invisible. That is the more forgiving case. Real-time payments over the RTP network and FedNow are credit-push and irrevocable, and neither carries a chargeback mechanism. Where a card dispute forces an agent error into an ill-fitting reason code, a real-time payment offers no code at all, and recovery depends on the goodwill of the receiving institution and the cooperation of the beneficiary.

This is not a distant concern. The RTP network has operated since 2017, reaches more than eleven hundred financial institutions, processed approximately 1.3 trillion dollars in 2025, and raised its per-transaction limit to ten million dollars in February 2026. Both it and FedNow are built on ISO 20022, which is the messaging environment in which any attribution of agent initiation or reference to a mandate would have to be carried. The rails that agentic commerce is most likely to migrate toward are therefore the rails on which the consumer has the least recourse and on which the technical means of recording what the agent was authorised to do already exists.

Two implications follow for Section 4(g). The working group should be directed to consider agent error across payment rails rather than only where dispute machinery happens to exist, since a remedy available on cards and unavailable on real-time payments will shape which rail agent providers choose. And because Treasury, the Federal Deposit Insurance Corporation and the Office of the Comptroller of the Currency are already named in the provision, the agencies with standing over those rails are already at the table.

The interagency working group contemplated by Section 4(g) is precisely the right body to address this, and it already includes every agency whose participation would be required. The provision needs its scope widened rather than its machinery redesigned.

Part Five: Instruction integrity

Section 3(d) permits the Commission to authorise independent certification bodies. Section 3(g) sets the duties those bodies would assess. Neither presently addresses the integrity of the agent’s instructions.

Prompt injection is conventionally treated as a security vulnerability. In the context of this draft it is more accurately a loyalty failure. An agent that reads a counterparty’s content and thereby has its instructions altered has, at that moment, ceased to act for the user. It will nonetheless satisfy every element of registration under Section 3(d) and identity binding, because those are established at onboarding while the compromise occurs at runtime.

This matters for the standard of care in Section 3(g). A duty of loyalty defeasible by content the agent was directed to read is not a duty the agent can reliably perform. Certification criteria should require assessment of resistance to instruction manipulation from counterparty-supplied content. The real-time records already required by Section 3(g) should be sufficient to reconstruct the instruction chain, not only the actions taken, since an action log without the instructions that produced it cannot distinguish an agent that erred from an agent that was captured.

Recent empirical work indicates that the problem is broader than manipulation. In a study published in August 2026, researchers at the Wharton School exposed six frontier models to an identical shopping task two hundred times per condition, varying only the material the agent encountered before it reached the purchase page (Kumar, Meincke, Shapiro, Mollick, Puntoni and Mollick, Agentic Shopping is Complicated and Contingent, Wharton Generative AI Labs Research Report 6, August 2026, SSRN abstract 7355899). A single prior source shifted product selection by as much as ninety-nine percentage points. Presenting the same sources in a different order produced a different selection. Most significantly for this draft, returning the same three sources in one tool call rather than in three sequential calls changed the product selected by up to forty-seven percentage points. Nothing about the information differed. Only the format of its delivery did.

Two implications follow for Sections 3(g) and 4(c). First, a real-time record of the agent's actions, and even of the content it read, would not have captured the variable that determined the outcome, because that variable was an implementation detail of the retrieval process rather than an item of content. Records sufficient to reconstruct an instruction chain should therefore extend to the prior context the agent retrieved, the sequence in which it was received, and the format in which it was delivered. Second, in the same study a product that was cheapest, highest rated and most reviewed was passed over by five of the six models in the control condition, with no prior content and no injected instruction present at all. The ordinarily prudent person standard in Section 3(g) presumes a baseline of user-aligned selection that this evidence does not support. Consistency and loyalty are distinct properties, and it is the second that Section 3(g) requires.

It is worth acknowledging what commentators have already noted, which is that reliable explanation of model reasoning is not currently available. That is a reason to specify the requirement now so that the standards work under Section 4(c) is directed toward it, rather than a reason to omit it.

Part Six: The counterfactual

It is worth stating what the market produces absent this framework, because a working example already exists.

In April 2026 American Express released a developer kit for agentic commerce together with a commitment to protect cardholders from charges arising from registered agent error. The protection is conditioned on the agent being registered with American Express, on American Express receiving the cardholder's authenticated purchase intent before the transaction, and on the error consisting of a deviation from that intent.

This is a functioning implementation of agent registration, intent capture and error protection. It works because American Express is issuer, acquirer and network simultaneously and therefore need not negotiate rules with any other party. It demonstrates that the mechanism is technically and commercially achievable, and it supplies the consumer remedy that Section 4(g) currently omits.

It also demonstrates the competition problem the draft exists to address. The same company has stated that a strategic priority is making its own inventory and offers discoverable across AI platforms. The party operating the agent registry is also a party selling to the agents it registers. Consumers within that ecosystem receive protection, which is valuable. They do not receive representation, which is different.

The predictable equilibrium without federal action is that agent trust becomes available only inside closed loops operated by vertically integrated firms with a commercial interest on both sides. That outcome is worse for competition, worse for smaller merchants, and worse for consumers transacting outside any single perimeter. It is a direct argument for the access obligations in Section 3(i), and for extending the framework to counterparty verification.

Recommendations

1. **Extend authentication to the counterparty. (Sections 3(c) and 4(c))**

   Section 3(c) should require that the access process be authenticated in both directions, not only as to the Custodial User Agent. The protocols and model standards identified by NIST under Section 4(c) should expressly include counterparty authentication and counterparty credential formats alongside verifiable delegation, and certification criteria under Section 3(d) should include the capacity to consume counterparty credentials. No technology need be specified. The direction needs to be named so the standards work does not proceed on the assumption that only one side requires a standard.

2. **Make the Section 3(i) filings a machine-readable counterparty register. (Section 3(i))**

   Section 3(i) already requires large online platform providers to file privacy and security standards with the Commission and provides that those filings be publicly available. Extend the filing to include the representations a provider makes to Custodial User Agents, and require the resulting register to be published in a signed, machine-readable format suitable for automated query at transaction time. This uses machinery the draft already contains and would place the United States on a comparable footing with the European relying-party register.

3. **Include harm to users in the Section 4(g) working group. (Section 4(g))**

   Section 4(g) directs proposals to prevent harms to businesses and the Government from agent fraud, misuse or genuine mistake. Extend it to harms to users arising from the same causes, and direct the working group specifically to address dispute categorisation for agent error, the standing of Custodial User Agent providers in dispute resolution, and the absence of any recovery mechanism on irrevocable real-time rails, which is where agent-initiated payments are most likely to migrate. The agencies required are already named in the provision.

4. **Add instruction integrity to the duties and certification criteria. (Sections 3(d) and 3(g))**

   Require that certification bodies authorised under Section 3(d) assess resistance to instruction manipulation from counterparty-supplied content, and clarify that the real-time records required by Section 3(g) must be sufficient to reconstruct the instruction chain, not only the actions taken. Those records should extend to the prior context the agent retrieved, the order in which it was received, and the format in which it was delivered, since recent empirical work shows that delivery format alone can change which product an agent selects.

5. **Make access parity bidirectional. (Section 3(k))**

   Clarify that Section 3(k) sets a ceiling on agent access rights rather than permitting differential disclosure. A large online platform provider should not present to a Custodial User Agent information that is materially different from, or materially less complete than, what it presents to a user in the equivalent interaction.

6. **Require disclosure of affiliate compensation, and retain Sections 3(g) and 3(h) intact. (Sections 3(g), 3(h) and 3(j))**

   If affiliate models are permitted under Section 3(j), require affirmative disclosure to the user at the point of recommendation rather than relying on after-the-fact assessment of whether the consumer's interests were subordinated. Retain the ordinarily prudent person standard and the prohibition on secondary commercial use of user data in Section 3(g), and retain the non-waivability provision in Section 3(h) without qualification.

Closing

The construct at the centre of this draft is correct and, so far as I am aware, without precedent in United States law. Defining a class of software by whose interests it must serve is the durable contribution here, and it will outlast any particular technical standard. The non-waivability provision in Section 3(h) is what will keep it from being contracted away.

The recommendations above are directed at making that construct operable. An agent can be registered under Section 3(d), identity-bound, certified, and legally obliged under Section 3(g) to act in a user's best interests, and still be structurally incapable of doing so if it cannot see who is on the other side of the transaction. The draft has the harder half right. The remaining half is a matter of symmetry.

I would welcome the opportunity to discuss any part of this submission with your staff.

Respectfully submitted,

Wendy MacKinnon
Founder and Chief Executive Officer, Twinvocate Inc.
Toronto, Ontario