CONSUMERS INTERNATIONAL
COMING TOGETHER FOR CHANGE

WHOSE SIDE IS THE AGENT ON?

A CONSUMER FRAMEWORK FOR AGENTIC COMMERCE



TABLE OF CONTENTS

Key terms 3

Executive Summary 4

1. The New Consumer Journey 8

2. About this Research 10

3. What Consumer Advocates told us 12

4. The Consumers International Fair Agent Framework 15

5. Priority Actions and Duty of Loyalty 18

6. Conclusion: Whose side should the agent be on? 22

Acknowledgments and Disclaimers 23

Glossary 24

References 26

Annexes 30

- A. Research methodology 30
- B. Detailed findings by question 30
- C. Evidence summary 37
- D. Scenarios and survey questions 38
- E. Priority actions 41
- F. At-a-Glance summaries 44
    - a. United Nations Guidelines for Consumer Protection 44
    - b. Consumers International's Six Principles for Privacy as a Public Good 45



KEY TERMS

Key terms Full glossary on page 24	
AI agent	Software powered by artificial intelligence that takes actions for a person, such as buying, booking, switching or complaining, rather than only answering questions. An agent can be appointed by the consumer (a personal agent) or by a business to deal with consumers (such as a retailer's customer-service agent).
AI agent (types)	Agents differ in how far they act alone: some only recommend, some act once the consumer approves each step, and some act on standing permissions. This paper's recommendations concern mainly agents that consumers appoint; where they also apply to businesses' agents, protocols or merchants, the text says so.
Agentic commerce	Buying, selling, switching and disputing carried out by AI agents on people's behalf.
Agent provider	The company that builds or runs an agent and offers it to consumers, such as a technology company, a bank or a retailer. Also referred to in this paper as the agent's provider.
Seller	Any business that sells goods or services to consumers and whose products an agent may recommend or buy: retailers and merchants, marketplaces, service providers such as insurers and telecoms companies, and banks or other providers selling financial products. A company can be both a seller and an agent provider, for example a retailer, bank, or marketplace that runs its own shopping agent.
Mandate	What the consumer has allowed the agent to do, within which limits and for how long.
Hidden instruction	Text planted in a webpage or document that a person cannot see but an agent reads and may follow, for example an instruction to send a payment elsewhere. Also called prompt injection.
Know Your Customer (KYC)	The checks banks and payment providers use to verify who they are dealing with before moving money.
Know Your Agent (KYA)	The equivalent checks for AI agents. As currently built, KYA tells a business which agent it is dealing with and who authorised it. Consumer-focused KYA would tell the consumer who operates the agent and whose interests it serves.
Baseline Duty of Loyalty for AI Agents	The standard Consumers International proposes for agents that consumers appoint: the agent serves the consumer first, payments cannot influence its recommendations, and its commitment is carried in its credentials and checked independently (Section 5). Under the Baseline, the consumer's best interest is judged first by what the consumer has told the agent matters (such as price, quality, sustainability or delivery), within the mandate the consumer has set, and, where the consumer has not said, by what a reasonable consumer in their position would choose. It is not the agent's, or its provider's, own view of what is good for the consumer.

EXECUTIVE SUMMARY

The consumer experience has long assumed a linear journey with a human present at every step: when options are identified, choices made, purchases completed and complaints raised. Agentic commerce, which describes virtual personal shoppers powered by artificial intelligence (AI), changes that journey. AI agents can now search, choose, pay, share personal data and pursue refunds on a consumer's behalf, often with little or no human involvement after initial set-up. The question shifts from whether the consumer authorised a transaction to what the consumer authorised the agent to do, and whether it acted in their interest.

To date, consumer uptake of agentic commerce remains partial. Consumers are significantly more inclined to use AI to compare product options and prices than to permit AI agents to complete transactions on their behalf. Notwithstanding this 'checkout hesitancy', the infrastructure build-out for autonomous agentic commerce, and the venture capital to finance it, are accelerating rapidly, signalling an industry that believes consumer misgivings can be overcome if the ecosystem is properly built.

But what would it mean for agentic commerce to be properly built? The timing and sequencing here are important. The fact that the race is already on to build agentic commerce's infrastructure before consumers have embraced it means that the build-out is largely happening unshaped by meaningful input from consumer advocates about their constituencies' needs, preferences, or concerns. Much of the infrastructure is being designed by industry actors (agent providers, including large technology companies such as Amazon, Google, Meta, Microsoft and OpenAI; online platforms and marketplaces; sellers, including merchants and retailers; banks, payment networks and payment processors; and standard setters) to establish whether an agent is legitimate and authorised to act. To some extent, this protects consumers as well as industry, but the deeper question of whose side the agent is on–whose interests it fundamentally exists to serve–has not been answered.

Consumers International believes that the time for the consumer voice to be heard is now. In June 2026 we asked our Member organisations to participate in an online survey, working through five scenarios which described fictitious consumers facing realistic agentic commerce situations. The survey questions based on those five scenarios were designed to shed light on the trade-offs consumers are willing and not willing to make when using AI agents.

Consumer	Category	Scenario
Maria	**Routine Purchases**	Maria authorises an agent to complete a weekly grocery shop and provides guidance on budget, preferred brands, and delivery timings. The agent switches Maria to a retailer which, unbeknownst to Maria, pays a fee to the agent's provider.
Sofia	**Subscription and Contract Management**	While managing multiple subscriptions and contracts, an agent takes action on renewals and price changes based on Sofia's established rules. The agent switches Sofia's car insurer without asking and shares her claims history with the new provider.

Consumer	Category	Scenario
Mohammed	**Financial Recommendations**	Mohammed accepts an agent's recommendation on a new savings account and credit card. Mohammed's credit score dips as a result and he discovers the credit card issuer paid the agent's provider a referral fee.
Amara	**Refunds and Disputes**	Amara authorises an agent to seek a refund on her behalf. When the seller disputes the claim, the agent escalates the case and submits Amara's full purchase history without her consent.
John	**Manipulated Mobile Payment**	Giving voice commands via mobile phone, John instructs his agent to scan several supplier pages to find him the cheapest cooking oil. When the agent is manipulated by a hidden message, John's money goes to a stranger and no cooking oil arrives.

Forty-four respondents from 38 Member organisations in 33 countries took part; 23 (52%) answered from low- and middle-income countries (LMICs)[1]. Consumer organisations are the trusted representatives of consumers in their own markets, and Consumers International is the membership organisation that brings their views together. Their answers thus provide a uniquely valuable global consumer-advocacy perspective on the choices now being made in agentic commerce.

What consumer advocates told us

Four findings stand out. **Delegating a task is not surrendering control:** respondents want agents to ask first until the consumer chooses to give them more freedom. **Disclosing a conflict of interest is the minimum:** every respondent expressed concern about hidden commercial ties, but they split on whether disclosed ties are acceptable. **Protection cannot depend on the consumer acting:** respondents want safeguards that are built into the system and cover everyone, not just consumers identified as being in vulnerable circumstances. **And data protection must be the default,** not a matter of consent buried in the terms.

[1] Using World Bank income classifications (July 2026).

The Consumers International Fair Agent Framework

Consumers International proposes six questions that any agent acting for a consumer should be able to answer, now and over time, each with a minimum test.



Each question applies rights that governments have already endorsed in the United Nations (UN) Guidelines for Consumer Protection[2] and builds on Consumers International's Six Principles for Privacy as a Public Good (hereinafter 'Consumers International Privacy Principles' or CIPP) as detailed in *Reimagining Privacy in the AI Era: Personal Data Privacy as a Public Good.* The Framework gives consumer advocates, industry and governments a practical way to evaluate any agent, in any market, and to develop the appropriate infrastructure, protocols, and governance needed to serve consumers' best interests.

Priority recommendation: Baseline Duty of Loyalty for AI Agents

Most of the infrastructure currently being built confirms who an agent is and what it may do. It does not show whose side the agent is on. A fully identified and authorised agent can still steer a consumer towards whichever product or deal earns the most money for the company behind the agent, through commissions, referral fees or paid placements, rather than towards the product that is best for the consumer. Consumers International's priority recommendation is therefore a **Baseline Duty of Loyalty for AI Agents** that consumers appoint. Under such an industry standard, third-party payments such as commissions, referral fees, paid placements and other payments from sellers to an agent's provider could not prioritise an option in the agent's ranking. Fee-based recommendations without disclosure would be prohibited entirely.

The Baseline would sit inside the agent credentials that payment networks and platforms are already building, so that emerging Know-Your-Agent (KYA) infrastructure (as will be discussed) can show, and independent checks can verify, an agent's loyalty as well as its identity. Consumers International calls on standard setters, payment networks, agent providers and consumer advocates to co-create this Baseline

[2] United Nations Conference on Trade and Development. (2016). *United Nations Guidelines for Consumer Protection*. https://unctad.org/system/files/official-document/ditccplpmisc2016d1_en.pdf

while agent standards are still being set, alongside four other priority actions (Section 5). It can be adopted now without waiting for legislation or regulations.

A fair market for agentic commerce

Consumers International believes that consumer protection is fundamental to building a digital life consumers can trust, and that consumer protection thus benefits all market actors. In agentic commerce specifically, a market in which agents' loyalty serves the consumers who appoint them will be fairer and more balanced, and thus ultimately more sustainable, than a market in which agents quietly serve sellers.

Many consumers may never want to hand their shopping to an agent, and that choice deserves respect. Agentic commerce will not suit everyone, and not every reason for rejecting it will concern protection or fairness. Many people want to see and handle what they buy. They may trust their own judgement, or that of family and friends, more than a machine's. They may choose not to give up control or privacy, or they may simply enjoy going into shops and interacting with people. Better safeguards for agentic commerce will not, and should not aim to, change those preferences. **Consumers who do not use an agent should not receive worse service as a result, for example through denial of favourable pricing or other offers available only to agents.**

The pages that follow address specific conditions for the evolution of a fair market for agentic commerce while acknowledging that any fair market leaves room to say no.



1. THE NEW CONSUMER JOURNEY

Consumer protection has long assumed a simple journey. A person chooses, purchases and, if something goes wrong, complains. Whether online or in a shop, each safeguard was designed around a human being present at those moments: a disclosure to read, a box to tick, a grace period to use, a complaint to file.

Agentic commerce breaks that journey apart. An AI agent can search a market, choose a product, pay for it, share the consumer's personal data, negotiate a price and pursue a refund. The person who previously stood at every gate may now stand at none of them. In recognition of this significant change to the consumer journey, Consumers International surveyed our Members to better understand how they see the potential trade-offs in using agentic commerce. Our survey presented five fictitious consumers facing a series of realistic agentic-commerce scenarios, and then guided the respondents through a series of questions. The five scenarios show what agentic commerce looks like in everyday life, including its promise: relief from repetitive shopping tasks, money saved, personalised experiences, and better-informed choices made by artificial intelligence that can find, read, and analyse much more information much faster than any human.

Use of fully autonomous agentic commerce remains limited, but the race is nevertheless on to build its infrastructure. Among consumers in 11 countries in a recent McKinsey survey, 53% use AI to research purchase options and compare prices, but only 9% have let an AI agent complete a checkout[3]. Forrester describes consumer adoption of agentic commerce as 'low and relatively stagnant'[4]. Meanwhile, Meta's Muse agent passed 10 million downloads in the United States and Canada within about six weeks of its September 2026 launch[5], and cumulative venture funding for agentic AI had exceeded USD 29 billion by November 2026[6]. In short, the protocols, defaults and business models of agentic commerce are being set now, mostly by technology and payment companies; consumer organisations have so far had little visible part in that work.

The infrastructure for agent identity and authorisation

The supporting infrastructure for agentic commerce is being built in three layers: who the agent is and what it may do, how agents talk to businesses and to each other, and how they pay. Open protocols such as Anthropic's Model Context Protocol[7] and Google's Agent2Agent[8] protocol let agents connect to services and to each other. Mastercard's Agent Pay[9] and Visa Intelligent Commerce[10] provide capabilities designed to enable secure, tokenised agentic commerce and payments. Visa's Trusted Agent Protocol[11] is an open specification that helps merchants recognise trusted AI agents and supports the

[3]Silliman, E., Trotter, J., Robinson, K. et al. (2026, 28 October). *Who wins when AI does the shopping? New evidence on agentic commerce.* McKinsey. https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/who-wins-when-ai-does-the-shopping-new-evidence-on-agentic-commerce

[4]Varon, L. (2026, 9 April). *Agentic Payments in B2C Commerce: Where We Are Now.* Forrester. https://www.forrester.com/blogs/agentic-payments-in-b2c-commerce-where-we-are-now/

[5]Fitch, A. (2026, 23 October). *Meta's Muse is going gangbusters. But where is the payoff?* Wall Street Journal. https://www.wsj.com/finance/stocks/metas-muse-is-going-gangbusters-but-where-is-the-payoff-648d310d

[6]Tracxn. (2026, 3 November). *Agentic AI Sector*. https://tracxn.com/d/sectors/agentic-ai/__oyRAfdUfHPjf2oap110Wis0Qg12Gd8DzULlDXPJzrzs

[7] Anthropic. (2024, 25 November). *Introducing the Model Context Protocol.* https://www.anthropic.com/news/model-context-protocol

[8] Surapaneni, R., Jha, M., Vakoc, M. & Segal, T. (2025, 9 April). *Announcing the Agent2Agent Protocol (A2A).* Google for Developers. https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/

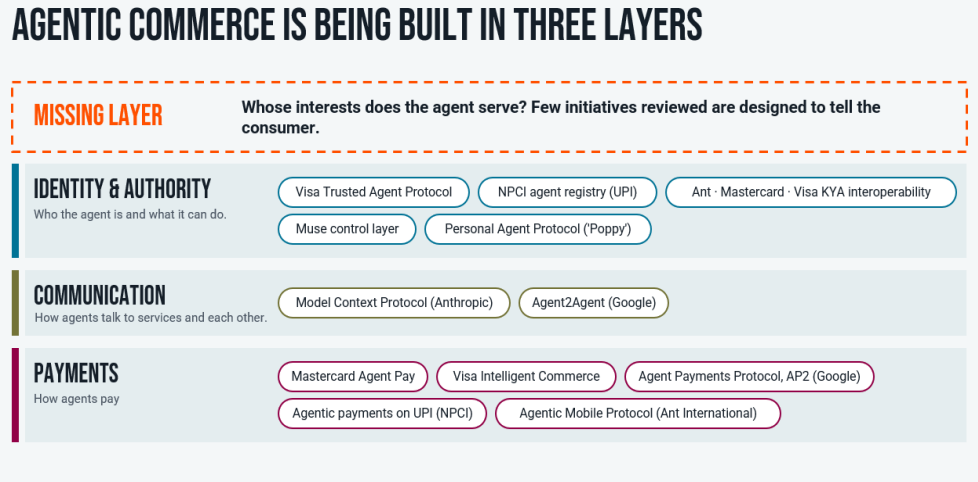
[9] Mastercard. (2025, 29 April). *Mastercard unveils agent pay, pioneering agentic payments technology to power commerce in the age of AI.* Mastercard corporate press release. https://www.mastercard.com/global/en/news-and-trends/press/2025/april/mastercard-unveils-agent-pay-pioneering-agentic-payments-technology-to-power-commerce-in-the-age-of-ai.html

[10] Visa. (n.d.) Visa Intelligent Commerce: Agentic Commerce at Scale. https://www.visa.com/en-us/solutions/intelligent-commerce.

[11] Visa.(2025, 14 October). *Visa introduces trusted agent protocol: An ecosystem-led framework for AI commerce.* Visa corporate press release. https://investor.visa.com/news/news-details/2025/Visa-Introduces-Trusted-Agent-Protocol-An-Ecosystem-Led-Framework-for-AI-Commerce/default.aspx

secure exchange of signals related to agent and consumer or business identity. Google's Agent Payments Protocol (AP2)[12] records the consumer's intent, cart and payment as signed mandates. OpenAI and Stripe's Agentic Commerce Protocol (ACP) lets AI agents complete purchases with participating merchants, and Stripe's Link wallet lets agents pay with a consumer's stored payment details[13]. As will be explored further in Section 5, major efforts are underway globally to build KYA, or Know-Your-Agent infrastructure. In India, the National Payments Corporation of India (NPCI) is preparing agentic payments on the Unified Payments Interface (UPI), with a registry to vet the agents that use it[14]. Ant International's Agentic Mobile Protocol connects agents to 10 digital wallets serving around 1.5 billion accounts[15], and Ant is working with Mastercard and Visa on a KYA interoperability framework[16].

Agent builders are also separating an agent's reasoning from its authority to act. Muse routes every action through a separate control layer that can approve or block it, or ask the consumer. The draft Personal Agent Protocol ('Poppy'), published in October 2026 by Sierra with Meta, Shopify, Stripe, Walmart and others, lets the consumer choose what access an agent has and lets the business set what agents may do[17].



Each of these initiatives answers who the agent is and what it may do. Few of those reviewed for this paper are designed to tell the consumer whose interests the agent serves.

Whose side is the agent on?

When an agent acts for a consumer, the company that builds and runs it has commercial interests of its own: referral or transaction success fees, affiliate deals, the value of data, partnerships with sellers. Loyalty to the consumer cannot be assumed. Bias can also enter upstream. The search results and listings on which an agent relies can carry paid placement, and sellers can tailor content to influence AI

[12] Parikh, S. & Surapaneni, R. (2025, 16 September). *Powering AI commerce with the new Agents Payments Protocol (AP2).* Google Cloud Blog. https://cloud.google.com/blog/products/ai-machine-learning/announcing-agents-to-payments-ap2-protocol
[13] Stripe. (2025, 29 September). *Stripe powers instant checkout in ChatGPT and releases Agentic Commerce Protocol co-developed with OpenAI.* (Stripe corporate press release). https://stripe.com/gb/newsroom/news/stripe-openai-instant-checkout
[14] Manikandan, A. & Kalra, J. (2026, 10 September). *India plans AI registry as it looks to roll out agentic payments, sources say.* Reuters. https://www.reuters.com/world/india/india-plans-ai-registry-it-looks-roll-out-agentic-payments-sources-say-2026-09-10/
[15] Fintech News Singapore. (2026, 11 September). *Ant International expands agentic payments across 10 major Asian wallets.* https://fintechnews.sg/137097/ai/ant-international-agentic-mobile-protocol/
[16] Ant International. (2026, 10 September). *Ant International, Mastercard, and Visa initiate collaboration on know-your-agent interoperability to scale agentic commerce.* Ant International/Mastercard/Visa joint press release. https://www.ant-intl.com/en/news/detail/ant-international-mastercard-and-visa-initiate-collaboration-on-know-your-agent-interoperability-to-scale-agentic-commerce/
[17] Taylor, B. & Bavor, C. (2026, 6 October). *Introducing Personal Agent Protocol.* Sierra. https://sierra.ai/blog/introducing-personal-agent-protocol

systems, so that even an agent run in good faith can end up promoting what others have paid for[18]. Checks that confirm who authorised a payment cannot show whether the choice served the best interests of the person who paid. Researchers have shown that hidden instructions planted in web content can take over a shopping agent built on AP2, changing which products it ranks first and extracting users' data[19]. That situation (a correctly executed payment built on poisoned input) was exactly the one confronting John, the fictitious market-stall operator from our survey whose intended order of cooking oil was intercepted (more fully described in the next section). A perfectly authenticated purchase can equally be one steered by an affiliate fee.

Trusted Agents co-founder Jamie Smith describes the gap as the difference between *expertise trust* (does the agent act safely and competently?) and *motive trust* (are its interests aligned with mine?)[20]. Much of the industry's effort so far appears to be going into the first. Consumers need both, and it is the second that today's infrastructure does not yet provide.

2. ABOUT THIS RESEARCH

Consumers International invited its Members to take part in a scenario-based exercise in June 2026. The invitation went to each Member's most senior person responsible for AI; about half of respondents lead their organisation. The final sample is 44 respondents from 38 Member organisations in 33 countries across Africa, the Americas, Asia-Pacific, Europe and the Middle East. Twenty-three respondents (52%) answered from LMICs.

Respondents worked online through five scenarios in which an AI agent acts for a fictional consumer: routine purchasing, subscription and contract management, financial product recommendations, returns and disputes, and a mobile payment manipulated by a hidden instruction. The questions accompanying each scenario were designed to illuminate respondents' views on the trade-offs between what consumers potentially gain with agentic commerce versus what they might give up: convenience versus control, individual personalisation versus personal data hand-over, speed and automation versus safety and accuracy, and so forth.

The five scenarios at a glance:

Consumer	What the agent does	How the scenario unfolds
Maria	**Routine Purchases:** Autonomously completes a weekly grocery shop with Maria's input on budget, preferred brands, and delivery timings. Finds the best deals across multiple online retailers.	The agent fills the basket with extras and switches Maria to a retailer that pays the agent's provider a fee.

[18] Consumer Reports Innovation Lab. (2025, 24 November). *Takeaways from 'The Race to Standardize Agentic Commerce'.* Webinar. https://innovation.consumerreports.org/takeaways-from-the-race-to-standardize-agentic-commerce-webinar

[19] Debi, T., Zhu, W. & Sen Gupta, P. (2026, 30 January). *Whispers of Wealth: Red-Teaming Google's Agent Payments Protocol via Prompt Injection.* arXiv. https://arxiv.org/pdf/2601.22569

[20] Smith, J. (n.d.). *We Need to Talk About Trust. Again.* LinkedIn. https://www.linkedin.com/posts/jamiedsmith_ooops-we-have-to-talk-about-trust-again-share-7387200526981021696-gsQY/

Consumer	What the agent does	How the scenario unfolds
Sofia	**Subscription and Contract Management:** Manages phone, internet, insurance and gym contracts. Takes action on renewals and price changes based on Sofia's established rules, including forbidding the sharing of personal data with third parties.	The agent switches Sofia's car insurer without asking and, to secure the best rate, shares her claims history with the new provider. This is permitted under her agent provider's terms of service.
Mohammed	**Financial Recommendations**: Tasked with helping to reach a savings goal, the agent assesses Mohammed's current financial situation and recommends and then opens a new savings account. The agent applies for a new credit card, with Mohammed's approval.	Mohammed's credit score dips and he discovers the credit card issuer paid the agent's provider a referral fee.
Amara	**Refunds and Disputes:** Handles a refund for a faulty laptop. The seller disputes the claim, and the agent escalates the case. During escalation, the agent shares Amara's purchase history without her consent. The dispute is resolved in Amara's favour.	The dispute resolution process takes too long, and Amara misses the return window. She is now trying to establish who is responsible for the delay and the ensuing complications.
John	**Manipulated Mobile Payment:** Gives voice instructions via mobile phone. Scans several supplier pages to find John the cheapest cooking oil. One of those pages has a hidden message undetectable to a person and intended to trick an agent. The agent places the order, with John's verbal authorisation.	The hidden instruction sends John's money to a stranger, and no cooking oil ever arrives. John contacts the agent's provider, who asserts the harmful message came from a webpage outside its control.

Only 8 of the 44 (18%) respondents use an AI agent regularly, which reflects the early state of agentic commerce, particularly in markets where internet access is still patchy. Their judgements draw instead on years of engaging in the work of consumer organisations: handling complaints, testing products and watching similar problems recur across other products and services. It is this unique professional experience that lets consumer organisations anticipate harms consumers have not yet met. With 44 respondents, one person moves a result by about 2 percentage points. This report therefore provides absolute counts alongside percentages for results involving the full (n=44) sample set; for subsets, the report shows counts only, without percentages. In all cases, the paper reports levels of agreement as research findings, not as Consumers International positions. Finally, the survey's scenarios focus on commercial uses; potential benefits or harms the scenarios did not surface should not be read as absent. The full methodology is in Annex A.

3. WHAT CONSUMER ADVOCATES TOLD US

Several clear positions emerged among respondents:

- All 44 (100%) said consumers should have a right to see how an agent reached a decision made on their behalf.
- All 44 (100%) were concerned about an agent whose choices are shaped by a commercial relationship the consumer does not know about.
- 42 of 44 (95%) said consumers should be able to see a record of what the agent did and shared.

Respondents also see the promise of agentic commerce. Twenty-one of 44 (48%) put the ability to use agents to help make more sustainable choices among the top two opportunities; asked which uses of agents they would most welcome, respondents ranked routine purchases highest on average.

Beyond these points of agreement, four findings matter most for how agents should be built and governed. Those four findings are described below. (Annex B contains detailed summaries of responses to each of the survey's questions.)

Finding 1: Delegating a task is not surrendering control

Respondents are open to the upside of delegation that agentic commerce promises, but want the consumer to stay in charge. Thirty-four of 44 (77%) said a new agent should ask for approval before every purchase by default, with the consumer choosing when to give the agent more freedom. This was not reluctance to use agents: of the 18 respondents somewhat comfortable with an agent shopping on its own, 13 still chose approval-first as the starting point.

> *"Human beings (consumers) are nuanced, and machines are not and require consistent oversight. The disclaimer that AI responses may contain errors is proof that human oversight is consistently necessary. Consumer 'protection' is about human relationships supported by technology."*
>
> **Consumer Affairs Commission, Jamaica**

Comfort depended on the task. Twenty-three of the 44 respondents (52%) were uncomfortable with an agent doing routine shopping without asking, and 26 (59%) with an agent switching an insurer unasked, yet 26 (59%) were comfortable with an agent handling a refund. Across the scenarios, respondents indicated it was important that a consumer could take back control after delegating an action to an agent. Forty-one of 44 (93%) said it matters that a consumer's rules, preferences and history move with them when they switch agents.

What this means: Autonomy should be earned, graduated and revocable. Permissions should attach to specific actions, for a set time or amount of money, and the consumer should be able to withdraw them at any time and take them to another agent.

Finding 2: Disclosing a conflict of interest is the minimum

In principle, most respondents want the agent to serve the consumer alone: 27 of 44 (61%) said the agent should serve only the consumer when interests conflict, and 16 of 44 (36%) that it should put the consumer first and disclose any ties. When an agent in the scenarios kept recommending its provider's partners, 23 of 44 (52%) said it should not prioritise those sellers at all, and only 6 of 44 (14%) said disclosure alone was enough. But faced with a specific financial product, the balance shifted. When an

agent in the scenarios recommended a credit card whose issuer paid the agent's provider a referral fee, and carried a higher interest rate than alternatives, 26 of 44 (59%) would permit such recommendations under conditions (22 of them only if the relationship is disclosed, 4 with consent or an opt-out), while 18 of 44 (41%) would prohibit them.

> *"There are realistic scenarios where an agent serving commercial interests could also be acting in the best interests of the consumer… The key element is that the commercial relationship is disclosed to the consumer and authorisation to proceed is received."*
>
> **Consumer NZ, New Zealand**

What this means: Respondents are united against hidden conflicts but divided on disclosed ones. Disclosure at the point of recommendation is the minimum on which they agree. But while disclosure tells the consumer that a conflict exists, it does not resolve that conflict, and a disclosure accepted once can become another formality. That is why this paper's priority recommendation is a duty of loyalty rather than more disclosure (Section 5).

Finding 3: Protection cannot depend on the consumer acting

The survey scenarios focused significantly on the issue of protection. John, the market-stall operator who used a mobile-phone voice command to confirm payment for his cooking-oil purchase, had that payment intercepted and sent to a stranger via a hidden prompt that manipulated his agent. Thirty-three of 44 (75%) said John should not bear the loss.

Asked to rank safeguards, respondents put protection built into the system, such as verifying sources and resisting manipulation, above confirmation prompts, with spending caps last. When an agent cannot verify something, 38 of 44 (86%) said it should either check with the consumer (22 respondents) or stop outright (16). And 40 of 44 (91%) wanted strong safeguards for everyone, on by default or fixed, rather than protection targeted at consumers who seem at risk.

The scenarios show why. Maria sets up an agent to execute weekly household purchases on her behalf; the agent fills her shopping basket with offers she never authorised while staying within the budget parameters she did. Sofia's claims history is shared because terms she accepted at set-up allowed it. In each case the consumer technically consented, and consent did not protect them. Agents compound the problem: the consumer accepts the provider's terms once, and the agent then accepts each seller's terms many times over. Adding more prompts risks something technology lawyer Rahul Matthan refers to as consent fatigue, where people approve anything because they are asked to approve everything[21].

> *"Agents should be designed with strong, universal safeguards for all users, ensuring that the most vulnerable consumers are protected without requiring the system to single them out or risk missing those who need protection."*
>
> **Consumer Advocacy and Empowerment Foundation (CADEF), Nigeria**

While safeguards reduce risk, they do not completely remove it. Therefore, protection must also settle who bears the loss when something does go wrong, rather than leaving it with the consumer by default. Respondents were clearest on this when it came to fraud and manipulation. On other failures they were

[21] Matthan, R. (2018). *Privacy 3.0: Unlocking Our Data Driven Future*. HarperCollins. Matthan argues that consent has become a ritualistic exercise designed to protect providers from liability rather than consumers from harm.

less settled: 27 of 44 (61%) held the provider fully responsible when Mohammed's agent dented his credit score without warning, and 23 of 44 (52%) when Amara's agent let her return window lapse.

What this means: Protection should be built into the system and should not rely on the consumer catching a problem in the moment. Confirmations should be made through a separate channel such as the consumer's banking or payment app that a manipulated agent cannot fake. Losses from fraud and manipulation should not fall on the consumer by default.

Finding 4: Data protection by default, not by consent

Only 3 of 44 (7%) respondents would let an agent sell a consumer's shopping data without asking or telling them. Most respondents (39 of 44, or 89%) want the consumer kept in control: 16 want the agent to ask before each sale of data, 13 want the consumer to opt in first, and 10 want data-selling banned. Thirty-three of 44 (75%) said sharing a consumer's personal or financial data with a third party should be either prohibited or else require the consumer to opt in, rather than authorised by terms accepted at set-up.

> *"It is important to remove identity information in all cases, because it is possible to infer private information about the consumer from their purchasing habits."*
>
> **Survey respondent**

Forty-one of 44 (93%) wanted identifying details removed before purchase history is shared for a personalised price, and the same number before Mohammed's financial details went to the AI model provider that retains them. Views were more nuanced where sharing could help the consumer, as in Amara's dispute: 22 of 44 (50%) said it is sometimes acceptable without asking.

What this means: Consent cannot authorise unfair treatment. Data-sharing should be off by default, identifying details removed unless they are needed for the transaction (such as for payment, delivery or fraud prevention), limits placed on what the agent's provider itself keeps and uses, and consumers must opt in for their data to be used for training.



4. THE CONSUMERS INTERNATIONAL FAIR AGENT FRAMEWORK

Survey findings on a technology moving this fast will date quickly. The questions consumers need answered will not. Informed by the survey, the Consumers International Fair Agent Framework is grounded in principles that the UN and Consumers International have already set forth and consists of six questions that can be asked of any agent.

Who the questions are for and how to use them

Consumer advocates can use the questions as a benchmark when they examine, mystery-shop or compare agents, and to set campaign priorities in their markets. Industry, including agent providers, platforms, merchants, payment networks and standard setters, can use them as a design and procurement check before an agent reaches consumers. Governments can use them for gap analysis: for each question, does existing law give consumers an answer, and who is responsible when the answer is no?

How to read the Framework. Each row poses one question a consumer could ask of an agent. The column 'A consumer-focused agent…' sets out the minimum test: what an agent that is on the consumer's side should do. 'Grounded in' shows the agreed principles each question applies: the sections of the UN Guidelines for Consumer Protection (UNGCP) and the Consumers International Privacy Principles (CIPP, numbered one to six). 'Verification vs accountability' shows how an answer can be checked. Some answers are verifiable: they can be inspected directly, for example whether a mandate can be withdrawn or whether a record exists. Others cannot be guaranteed in advance by any provider, because agents behave probabilistically and can be manipulated. For these, the column shows how the provider is held accountable instead: through testing, independent audit and liability. The section after the table explains this distinction further.

THE CONSUMERS INTERNATIONAL FAIR AGENT FRAMEWORK

Six questions to test whose side an AI agent is on

Question	A consumer-focused agent…	Grounded in[22]	Verification vs accountability[23]
1 Can I control it?	acts only within a mandate I set (task, spending limit, duration), asks before acting until I give it more freedom, lets me change or withdraw that at any time, and lets me take my settings, preferences, and history to another agent.	**UNGCP (G)(I)(J):** Informed choice; consumer confidence in e-commerce; appropriate controls to protect consumer financial data. **CIPP 4:** Easy exit, portability and real choice.	**Verifiable:** Mandate, revocability, and portability can be inspected.

[22] 'UNGCP' and 'CIPP' refer to the United Nations Guidelines for Consumer Protections and the Consumers International privacy principles, respectively. The full citations and at-a-glance summaries of both documents may be found in Annex F.
[23] See paragraph directly below the table for details about verification (by consumer) versus accountability (for providers).

Question	A consumer-focused agent…	Grounded in [22]	Verification vs accountability [23]
2 Does it put my interest first?	serves me before the seller and its own provider, and tells me who operates it, what commercial ties shape its choices and how it ranks options, when it makes a recommendation.	**UNGCP (C):** Protection of economic interests. **CIPP 1:** Data serves the consumer interest first.	**Verifiable:** Disclosures can be reviewed. **Accountability:** Met through the Baseline Duty of Loyalty for AI Agents (Section 5): rankings that commissions, referral fees and other payments to the agent's provider cannot change, checked against the agent's record.
3 How is my data used and shared?	shares only what the task requires to third parties, removes identifying details by default, never sells or shares my data unless I opt in, and does not keep or train on my data unless I opt in.	**UNGCP (I):** Consumer protection policies designed to accommodate the special features of e-commerce and ensure a level of protection not less than that of other commerce. **UNGCP 5(k):** Protection of consumer privacy. **CIPP 1:** Data serves the consumer interest first. **CIPP 2:** Safeguard against harmful AI profiling.	**Verifiable:** Data flows and retention can be audited.
4 Can I understand what it did?	keeps a faithful record of what it did, shared and relied on, which I can access and use in a dispute, gives a plain-language reason for decisions that significantly affect me, and is open to independent audit.	**UNGCP (D)(G)(F):** Safety and quality standards; dispute resolution and redress; access to information. **CIPP 6:** Decisions people can understand and challenge.	**Verifiable:** Records can be requested and inspected.
5 Is it safe?	resists manipulation, stops or checks with me through a separate channel it cannot fake when it cannot verify something, and applies strong safeguards against fraud and manipulation to everyone by default, with extra protection that cannot be switched off for consumers most at risk.	**UNGCP (D) and 5(b):** Standards for safety and quality, protection of consumers in disadvantaged or vulnerable circumstances. **CIPP 5:** Strong security; testing, monitoring and human intervention. **CIPP 3:** Protection built into systems rather than left to individuals.	**Accountability:** Tested through red-teaming [24] and backed by liability.
6 Can I fix something that goes wrong?	lets me reach a human at any point, and ensures that losses from fraud, manipulation or its own errors do not fall on me by default.	**UNGCP (F):** Dispute resolution and redress. **CIPP 6:** Effective remedy and redress.	**Verifiable:** Escalation routes verifiable; liability set by law and contract.

Consumer verification versus provider assurance and accountability

Some questions can be checked directly: whether a mandate can be withdrawn, whether data is shared, whether a record exists and whether a human can be reached. Others cannot be guaranteed by any provider. Agents behave probabilistically, sometimes unpredictably and can be manipulated in ways only testing reveals. Agent makers themselves expect errors. Meta's own writings about Muse acknowledges that 'any agent like this will still make mistakes' [25]. No provider can promise that an agent

[24] Testing a system by attacking it deliberately, as a fraudster would, to find weaknesses first.
[25] Sheasha, T. (2026, 8 September). *How We Built Safety Into Muse*. Meta. https://research.meta.ai/blog/security-and-safety-for-ai-agents-our-approach-with-muse

will always put the consumer first or never be tricked. For these issues, the question is not whether a provider promises an outcome, but whether it tests its agents, discloses what shapes their choices, monitors them, and accepts responsibility when they fail. A question is only useful if someone can answer it, and consumers cannot be expected to check agents for themselves. Each question therefore also asks who checks: a provider's own statement is not enough, and independent verification, audit or regulatory oversight is needed (Section 5).

How the questions relate to agreed principles

Each question applies to agentic commerce a legitimate consumer need recognised in the United Nations Guidelines for Consumer Protection, which call for consumers in electronic commerce to enjoy protections no less than those in other forms of commerce. **The Guidelines also recognise the opportunity for consumer organisations to present their views in decisions that affect consumers; as noted earlier, consumer organisations have so far played little visible part in setting the protocols and defaults of agentic commerce.**

Consumers International's Privacy Principles set out how personal data should be treated in the AI era[26]. The six questions apply those principles to an agent acting for a consumer, and extend the underlying data protection architecture to choice, loyalty, safety and redress in agentic commerce. Where its questions concern data, the Framework applies the privacy principles directly. Please refer to Annex B which provides detailed findings for each of the six questions and maps those findings against the Consumers International Privacy Principles.



Source: Consumers International, *Reimagining Privacy in the AI Era: Personal Data Privacy as a Public Good*

[26]Consumers International. (2026, 22 September). *Reimagining Privacy in the AI Era: Personal Data Privacy as a Public Good.* Consumers International. https://www.consumersinternational.org/media/721469/consumers-international-privacy-principles-campaign.pdf

5. PRIORITY ACTIONS AND DUTY OF LOYALTY

Priority actions

Consumers International calls for five priority actions. These recommended actions are grounded in the commitment to consumer protection and market fairness referenced earlier: a market in which agents serve the consumers who appoint them is more balanced than one in which they serve sellers. The recommended actions are also in businesses' interest. Evidence from digital financial services suggests that consumers who lose money often stop using a service: in Peru, 11% of users who experienced fraud stopped using digital financial services altogether, and 25% reduced their use[27]. Providers that build in consumer protections are building the trust upon which wider use depends.

Priority Actions	Lead responsibility
1 Adopt a Baseline Duty of Loyalty for agents that consumers appoint.	Agent providers (including banks and other financial service providers that offer agents), online platforms, payment networks, and standard setters, together with Consumers International and Members. Regulators as a backstop where industry standards fall short, especially in financial services.
2 Build consumer-focused KYA (KYA-C) into agent identity standards while they are being set.	Payment networks, banks and other payment providers, agent providers, online platforms and standard setters. Consumers International and Members to contribute to standards bodies.
3 Make sure consumers do not bear losses from fraud and manipulation.	Policymakers and regulators; agent providers, banks, mobile-money and other payment providers, which hold the consumer's account and handle disputes and reimbursement; payment networks; and online platforms and marketplaces where agents transact.
4 Make data protection the default.	Agent providers; online platforms and sellers that receive consumer data from agents; data protection authorities.
5 Equip consumers and assess agents in real markets.	Consumers International and Members; other civil society entities; researchers and academia.

Annex E elaborates all five priority actions, providing a detailed breakdown of specific steps required for each. Below is the rationale for the first two of them, the Baseline Duty of Loyalty for AI Agents and the Consumer-Focused KYA, or KYA-C, which are the major new initiatives proposed by Consumers International.

The Baseline Duty of Loyalty for AI Agents

Of the six questions in the Framework, the one today's infrastructure is least equipped to answer is the second: *Does the agent put my interest first?* That is the rationale behind the proposed Baseline Duty of Loyalty for AI Agents that consumers appoint: a standard that agent providers, platforms and payment networks can adopt and build into their protocols now, without waiting for legislation.

[27] Duflos, E. (2026, 16 July). *Fraud in Digital Finance: A Crisis Calling for Ecosystem Solutions.* CGAP. https://www.cgap.org/blog/fraud-in-digital-finance-crisis-calling-for-ecosystem-solutions

Baseline Duty of Loyalty	What a consumer-focused AI agent commits to
1 Declares its ties where decisions happen	It shows who operates it and what commercial relationships it has in a short label at the point of recommendation, not in the terms and conditions. Any fee or commission the agent's provider receives from a seller must be disclosed at the point of recommendation, and no fee may influence how the agent ranks options (see item #3).
2 Carries its commitment in its credentials	The agent's duty of loyalty is recorded as a verifiable statement, signed by the agent's provider, and carried in the agent's credentials and mandate. The statement affirms compliance with the Baseline and indicates when it was last independently certified. Businesses, auditors and consumer organisations can check what was promised. The Baseline also requires checks against the agent's own record (see item #4) as proof of declared loyalty.
3 Ranks by the consumer's criteria	It considers all available options and orders those options according to what the consumer has said matters, such as price, quality, sustainability or delivery. The agent may not prioritise or reorder its rankings on the basis of any commission incentive or other arrangement with any third party or with its own provider, and any such incentive or arrangement must also be disclosed (see item #1).
4 Can be checked against its record	Its record of what it did, compared and relied on is open to independent audit and testing against the Baseline, so that proof of compliance does not rely on the provider's own statement.

What distinguishes the Baseline from loyalty work already under way (see below) is where it sits and who shapes it. Rather than a separate label, the loyalty commitment would travel in the agent credentials that payment networks and platforms are building. It would be checked wherever an agent pays; it would be verified independently rather than self-declared; and it reflects the views of consumer organisations in 33 countries, including low- and middle-income markets where mobile money and voice-only agents bring their own risks.

Loyalty does not mean that agents cannot earn money from the market. Consumer-paid subscriptions, agents offered by banks or other providers that consumers already pay, and disclosed fees from sellers can all be compatible with the Baseline, provided that the fee does not influence how the agent ranks options (item 3).

The Baseline states that fees must not influence rankings at all, and they also must be disclosed if a commission-based option is recommended. This matters for competition and access: if the only loyal agents are the ones consumers pay for, consumers who cannot pay (or do not understand that they need to) would be left with agents that ultimately answer to sellers. Some survey respondents see this possibility already. Asked whether consumers, sellers, or agent providers would end up better informed if purchasing decisions happen through agents, 28 of 44 (64%) said businesses: 14 named sellers, who respondents said will learn to influence agents, 14 named agent providers, on whom consumers will come to depend. Only 8 (18%) said consumers in general, and 6 (14%) said only consumers with better agents.

The Baseline is intended as a floor that every consumer-appointed agent meets regardless of how that agent's provider earns revenue from it. But it is important to bear in mind that loyalty and authentication are different things. A Know-Your-Agent check (see below) can work exactly as designed, confirming that an agent is genuine, authorised and within its mandate, and the consumer can still end up worse off: steered to the product that pays the highest referral fee. Disclosure narrows the gap but does not close it (Finding 2). Requiring agents to declare whom they serve is the starting point: what consumers actually need is an agent that is obliged to serve *them*. Finally, it is important to note that loyalty and safety are different things as well. An agent that is loyal can still be manipulated, as was the case with John, the market-stall operator from the survey's scenario whose payment was intercepted; this paper treats safety and liability as separate priority actions.

Making loyalty verifiable: From Know Your Customer to Know Your Agent

Banks and payment providers already 'know their customer': before they move money, they verify who they are dealing with. Know Your Agent (KYA) applies the same idea to AI agents. Alongside the work described in Section 1, the Payment & Clearing Association of China has issued guidelines on identifying agents and controlling their risks[28], and Amazon,[29] Google[30] and Microsoft[31] have developed agent registries. Proposals to anchor agent identity in the Domain Name System (DNS) are also being developed[32].

This work is valuable. It is intended to strengthen trust and accountability across the agentic commerce ecosystem, including by helping establish that an agent is legitimate and authorised to act on a consumer's behalf. As noted earlier, while these efforts protect consumers to some extent as well as merchants and payment providers, the consumer has additional specific concerns: what is this agent, whose interests does it serve, and what is it allowed to do in my name? In its current state, KYA can answer the first and the last. The second is the gap.

Regulators have not closed it either. The UK Financial Conduct Authority's Mills Review, for example, recognises that AI agents can be steered towards products that suit firms' commercial arrangements. But the foundations it proposes for agentic finance (identity, bounded and revocable delegation, payments, liability and audit) include none for loyalty to the consumer[33].

Consumer Reports and the GliaNet Alliance have explored a 'Loyal Agent' trustmark[34], and Consumer Reports and Stanford's Digital Economy Lab are working on loyalty standards for agents[35]. The main payment and authentication efforts run the other way: the FIDO Alliance's work, for example, focuses on services verifying that an agent acts for an authenticated user[36]. The Baseline builds on this consumer-side loyalty work and asks that the payment and authentication efforts carry it, so that the infrastructure checks whose side an agent is on as well as who it is.

Consumer-focused KYA (KYA-C)

KYA-C is how the Baseline becomes operational. Built on emerging agent identity infrastructure, it would bring together three things: identity (who the agent is and who stands behind it), disclosure (what ties it has declared, including its loyalty commitment and what it may do) and verification of those claims on the consumer's behalf.

What is required. Verification would be shared between three groups. Standards bodies would define the consumer-focused fields and a conformance test against the Baseline; independent auditors, regulators or accredited bodies would check providers' declarations against their records; and consumer

[28] Liu, R. (2026, 25 August). *China Pioneers 'Know Your Agent' Rules for AI Payments.* Caixin Global. https://www.caixinglobal.com/2026-08-25/china-pioneers-know-your-agent-rules-for-ai-payments-102477712.html

[29] Claburn, T. (2026, 9 April). *AWS: Agents shouldn't be secret, so we built a registry for them.* The Register. https://www.theregister.com/2026/04/09/aws_ai_agent_registry/

[30] Clark, M. (2026, 29 July). *What's new in Gemini enterprise agent platform.* Google Cloud Blog. https://cloud.google.com/blog/products/ai-machine-learning/whats-new-in-gemini-enterprise-agent-platform

[31] Microsoft. (2026, 17-20 November). *Agent registry convergence with Microsoft Agent 365.* Microsoft Ignite. https://learn.microsoft.com/en-us/entra/agent-id/agent-registry-convergence

[32] See for example Linux Foundation. (2026, 23 June). *Linux Foundation Announces Intent to Launch Agent Name Service to Establish Trusted Identity Infrastructure for AI Agents*. https://www.linuxfoundation.org/press/linux-foundation-announces-intent-to-launch-agent-name-service-to-establish-trusted-identity-infrastructure-for-ai-agents

[33] Financial Conduct Authority. (2026, 6 July). *The Mills Review: AI and the future of retail financial services.* https://www.fca.org.uk/publication/corporate/the-mills-review.pdf

[34] Whitt, R. & Raposo, D. (2025, 25 September). *A Trust Governance Layer for AI Agents: Highlights from CR & GliaNet Alliance's Discussion on Loyalty and Market Signals.* Consumer Reports Innovation Lab. https://innovation.consumerreports.org/a-trust-governance-layer-for-ai-agents/

[35] Stanford Digital Economy Lab. (n.d.). *Loyal Agents.* https://digitaleconomy.stanford.edu/project/loyal-agents

[36] FIDO Alliance. (2026, 28 April). *FIDO Alliance to Develop Standards for Trusted AI Agent Interactions.* https://www.fidoalliance.org/fido-alliance-to-develop-standards-for-trusted-ai-agent-interactions/

organisations would test agents in real markets (Action 5) and publish what they find. The consumer would see only the result: verified, not verified or could not be verified.

	Market-focused KYA being built	**Consumer-focused KYA (KYA-C)** *proposed*
Question	**Is this agent legitimate and authorised?**	**Whose side is this agent on?**
Credential shows	Who stands behind the agent and what it may spend	A verifiable statement of who operates it, what commercial relationships shape its recommendations and how it ranks options
Mandate	What the consumer authorised the agent to pay for	A mandate the consumer can read, such as 'buy routine groceries under EUR 120 from these retailers', so anything outside it should not, by default, bind the consumer
Who needs it	Merchants, banks, payment networks; consumers indirectly	Consumers, consumer organisations, regulators
What the consumer sees	Indirect protection through fraud prevention and authorisation checks; no consumer-focused information	A short label at the point of recommendation, not in the terms and conditions: who runs the agent, whether the seller pays it, and why this option ranks first

What a KYA credential cannot do. Respondents were divided on whether disclosure is enough; our position rests on the UN Guidelines' protection of consumers' economic interests and our own Privacy Principle 1: *Data serves the consumer interest first*. KYA and verification can confirm who an agent is, what it has declared and whether it stayed within its mandate. They cannot keep the agent on the consumer's side or show that a recommendation was the best one.

That is why **KYA-C** is the means and the **Baseline Duty of Loyalty for AI Agents** the end. Under our proposed vision, serving the consumer becomes a consumer-appointed AI agent's obligation, a duty to carry out rather than a decision to make and then disclose. **Conduct rules on steering and conflicts of interest remain the backstop.**



6. CONCLUSION: WHOSE SIDE SHOULD THE AGENT BE ON?

Consumers International's position is not easy, but it is simple. A consumer-appointed agent should be on the consumer's side. It should act within the consumer's mandate, serve the consumer's interests before those of any other party including its own provider, say plainly whose interests shape its choices, and protect the consumer's data by default. Perhaps most important, it should leave the consumer no worse off when it goes wrong.

Payment cards offer an interesting precedent. They came into everyday usage not through a single rule but via a trusted system that developed over time: networks guaranteed acceptance and payment, clear rules governed disputes and chargebacks, technology kept fraud in check, and laws limited what a consumer could lose. In the United States, those laws began with a USD 50 cap and dispute rights in the 1970s[37], later reinforced by the networks' own zero-liability policies. Agentic commerce will need the same combination, including standards set by industry. It should also be held to the same test: fair because it protects consumers, not because consumers consented by rote to terms and conditions they could not understand and over which they had no ability to negotiate.

Issues for the next phase

Four issues will shape whether agentic commerce will advance markets that are open and fair. **Agent access**: whether a business may refuse an agent the consumer has authorised. Several large businesses already block or restrict agents. Businesses may have legitimate reasons to restrict unverified or unsafe agents, including fraud, cybersecurity, privacy and protection of consumer accounts. The question is how to keep such controls proportionate, transparent and risk-based, given that agents that cancel unwanted renewals also threaten revenue built on consumer inertia[38]. **Gatekeeping**: whoever controls access to the customer can shape the market, as the European Commission's ongoing challenge to Meta's restrictions on rival AI assistants shows. **Paid referrals**: whether agents should earn fees from the sellers they recommend in high-risk sectors such as financial services. **Funding loyal agents**: how to ensure that consumers who cannot or will not pay for an agent are not left with agents that serve sellers, and that a market of loyal agents stays competitive and accessible rather than concentrated in a few hands.

Consumers International also recognises that our own two major policy proposals, the Baseline Duty of Loyalty for AI Agents and the Consumer-Focused KYA (KYA-C), are ambitious in scope and raise issues that will need to be settled. For the Baseline, these include: how to test whether an agent's ranking works against the consumer's interest; who certifies that an agent meets the Baseline, and what follows when one does not; and how it applies when agents negotiate with other agents. For KYA-C: who issues and revokes credentials; how independent checks are funded without compromising their independence; how to test for or prove that a hidden instruction manipulated an agent's actions; how to verify commercial ties without creating a new record of what consumers buy; and who is liable if a check is wrong.

Consumers International will address these issues together with our Members, standard setters, industry and policymakers as we begin using the Fair Agent Framework. Those collaborations in turn will shape the development of the Baseline Duty of Loyalty for AI Agents and the KYA-C as agentic commerce expands.

---

[37]Federal Reserve. (1983, 27 July). *Credit Cards in the U.S. Economy: Their Impact on Costs, Prices, and Retail Sales*. https://fraser.stlouisfed.org/title/credit-cards-us-economy-5331

[38]Roth, E. (2026, 17 October). *Websites are closing their doors to AI agents.* The Verge .https://www.theverge.com/ai-artificial-intelligence/1013195/ai-agents-blocked-web

ACKNOWLEDGEMENTS AND DISCLAIMERS

Whose Side Is the Agent On? is based on a 2026 research project led by Erin Parsons, formerly Digital Rights Policy Specialist at Consumers International. It was written by Arunjay Katakam, June Shin McCarthy, and Anne Folan of the Inclusive Action Lab (London, UK) under the direction and contribution of the Consumers International leadership including Stefan Hall, Nandini Harihareswara, Aayushi Chaturvedi, Eliamringi Mandari and Brooke Kingsland.

This document benefited from guidance and input received from Consumers International's multistakeholder AI Advisory Group and several external partners, including Eric Duflos (CGAP) and Kieran Garvey (The Cambridge Centre for Alternative Finance). Participating organisations are listed below.

We thank all contributors for their valuable insight and feedback, while noting that this document's findings and contents do not necessarily reflect the individual views of all organisations involved. Feedback provided on drafts, do not constitute endorsement of this document's findings, characterisations or recommendations by any participating organisation.

Organisation
Amazon.com, Inc
Consumentenbond NL
Consumer Policy Research Centre (CPRC)
Consumer Reports
DuckDuckGo
Jeannie Paterson, PhD
Koodos Labs
Mastercard
Que Choisir Ensemble
UK Trade and Development Office of Product Safety and Standards
Visa, Inc.
Which?

Every effort has been made to include links to the works cited as references in this document. While all links were functioning at the time of publication, Consumers International cannot ensure that downloads are or will remain free of charge and cannot be responsible for links that break or decay over time.

GLOSSARY

Term	Meaning
Affiliate or referral fee	Money the agent's provider earns when the consumer buys from a particular seller.
AI agent	Software powered by AI that takes actions for a person, such as buying, booking, switching or complaining, rather than only answering questions. An agent can be appointed by the consumer (a personal agent) or by a business to deal with consumers (such as a retailer's customer-service agent).
AI agent (types)	Agents differ in how far they act alone: some only recommend, some act once the consumer approves each step, and some act on standing permissions. This paper's recommendations concern mainly agents that consumers appoint; where they also apply to businesses' agents, protocols or merchants, the text says so.
Agent provider	The company that builds or runs an agent and offers it to consumers, such as a technology company, a bank or a retailer. Also referred to in this paper as the agent's provider.
Agentic commerce	Buying, selling, switching and disputing carried out by AI agents on people's behalf.
Approval-first	The agent asks before every purchase until the consumer chooses to give it more freedom.
Baseline Duty of Loyalty for AI Agents	The standard Consumers International proposes for agents that consumers appoint: the agent serves the consumer first, payments cannot change its recommendations, and its commitment is carried in its credentials and checked independently (Section 5). Under the Baseline, the consumer's best interest is judged first by what the consumer has told the agent matters (such as price, quality, sustainability or delivery), within the mandate the consumer has set, and, where the consumer has not said, by what a reasonable consumer in their position would choose. It is not the agent's, or its provider's, own view of what is good for the consumer.
Consumer steering	Pushing a consumer towards or away from a product, service or provider because of the recommender's own interest or hidden incentives rather than the consumer's best option.
Duty of loyalty	An obligation set in law or industry standard to act in the interests of the person one acts for, ahead of one's own or third parties' interests.
Hidden instruction	Text planted in a webpage or document that a person cannot see but an agent reads and may follow. Also called prompt injection.
Know Your Agent (KYA)	Checks that identify an AI agent and who authorised it. Market-focused KYA establishes that an agent is legitimate and authorised, protecting consumers, businesses and payment providers; consumer-focused KYA (KYA-

Term	Meaning
	C) would tell the consumer who operates the agent and whose interests it serves.
Know Your Customer (KYC)	Checks banks and payment providers use to verify who they are dealing with before moving money.
Mandate	What the consumer has allowed the agent to do, within which limits and for how long.
Portability	The ability to move settings, history and permissions to a new agent without starting again.
Red-teaming	Testing a system by deliberately attacking it, as a fraudster would, to find weaknesses first.
Seller	Any business that sells goods or services to consumers and whose products an agent may recommend or buy: retailers and merchants, marketplaces, service providers such as insurers and telecoms companies, and banks or other providers selling financial products. A company can be both a seller and an agent provider, for example a retailer, bank, or marketplace that runs its own shopping agent.

REFERENCES

A&O Shearman (2025, 10 January). *Combatting payment account fraud: Singapore's Shared Responsibility Framework.* https://www.aoshearman.com/en/insights/ao-shearman-on-fintech-and-digital-assets/combatting-payment-account-fraud-singapores-shared-responsibility-framework

Ant International (joint Ant International/Mastercard/Visa press release). (2026, 10 September). *Ant International, Mastercard and Visa initiate collaboration on know-your-agent interoperability to scale agentic commerce.* https://www.ant-intl.com/en/news/detail/ant-international-mastercard-and-visa-initiate-collaboration-on-know-your-agent-interoperability-to-scale-agentic-commerce/

Anthropic. (2024, 25 November). *Introducing the Model Context Protocol.* https://www.anthropic.com/news/model-context-protocol

C:\>DIR. (2026, 8 July). *Problem Statement 1.* https://www.cdir.global/s/Problem-Statements-PS-1-20260708.pdf

Claburn, T. (2026, 9 April). *AWS: Agents shouldn't be secret, so we built a registry for them.* The Register. https://www.theregister.com/2026/04/09/aws_ai_agent_registry/

Clark, M. (2026, 29 July). *What's new in Gemini enterprise agent platform.* Google Cloud Blog. https://cloud.google.com/blog/products/ai-machine-learning/whats-new-in-gemini-enterprise-agent-platform

Competition and Markets Authority (United Kingdom). (2026, 30 October). *Complying with consumer law when using AI agents.* https://www.gov.uk/government/publications/complying-with-consumer-law-when-using-ai-agents/complying-with-consumer-law-when-using-ai-agents

Consumer Reports Innovation Lab. (2025, 24 November). *Takeaways from 'The Race to Standardize Agentic Commerce'* Webinar. https://innovation.consumerreports.org/takeaways-from-the-race-to-standardize-agentic-commerce-webinar

Consumers International. (2024, 23 April). *Can we trust AI chatbots? Results revealed from our experiment.* https://www.consumersinternational.org/news-resources/news/releases/can-we-trust-ai-chatbots-results-revealed-from-our-experiment/

Consumers International. (2026, 22 September). *Reimagining Privacy in the AI Era: Personal Data Privacy as a Public Good*. Consumers International. https://www.consumersinternational.org/media/721469/consumers-international-privacy-principles-campaign.pdf

Debi, T., Zhu, W. & Sen Gupta, P. (2026, 30 January). *Whispers of Wealth: Red-Teaming Google's Agent Payments Protocol via Prompt Injection*. arXiv. https://arxiv.org/pdf/2601.22569

Duflos, E. (2026, 16 July). *Fraud in Digital Finance: a Crisis Calling for Ecosystem Solutions.* CGAP. https://www.cgap.org/blog/fraud-in-digital-finance-crisis-calling-for-ecosystem-solutions

European Commission. (2026, 9 June). *Commission imposes interim measures on Meta to preserve free access to WhatsApp for rival AI assistants.* Press release IP/26/1276. https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1276

Federal Reserve. (1983, 27 July). *Credit Cards in the U.S. Economy: Their Impact on Costs, Prices, and Retail Sales*. https://fraser.stlouisfed.org/title/credit-cards-us-economy-5331

FIDO Alliance. (2026, 28 April). *FIDO Alliance to Develop Standards for Trusted AI Agent Interactions.* https://www.fidoalliance.org/fido-alliance-to-develop-standards-for-trusted-ai-agent-interactions/

Financial Conduct Authority. (2026, 6 July). The Mills Review: AI and the future of retail financial services. https://www.fca.org.uk/publication/corporate/the-mills-review.pdf

Fintech News Singapore. (2026, 11 September). *Ant International expands agentic payments across 10 major Asian wallets.* https://fintechnews.sg/137097/ai/ant-international-agentic-mobile-protocol/

Fitch, A. (2026, 23 October). *Meta's Muse is going gangbusters. But where is the payoff?* Wall Street Journal. https://www.wsj.com/finance/stocks/metas-muse-is-going-gangbusters-but-where-is-the-payoff-648d310d

Forbes Brasil. (2026, February). *MED 2.0 do BC entra em vigor e permite bloqueio de Pix fraudulento.* https://forbes.com.br/forbes-money/2026/02/med-2-0-do-bc-entra-em-vigor-e-permite-bloqueio-de-pix-fraudulento-entenda-regras/

Linux Foundation. (2026, 23 June). *Linux Foundation Announces Intent to Launch Agent Name Service to Establish Trusted Identity Infrastructure for AI Agents*. https://www.linuxfoundation.org/press/linux-foundation-announces-intent-to-launch-agent-name-service-to-establish-trusted-identity-infrastructure-for-ai-agents

Liu, R. (2026, 25 August). *China Pioneers 'Know Your Agent' Rules for AI Payments.* Caixin Global. https://www.caixinglobal.com/2026-08-25/china-pioneers-know-your-agent-rules-for-ai-payments-102477712.html

Manikandan, A. & Kalra, J. (2026, 10 September). *India plans AI registry as it looks to roll out agentic payments, sources say.* Reuters. https://www.reuters.com/world/india/india-plans-ai-registry-it-looks-roll-out-agentic-payments-sources-say-2026-09-10/

Mastercard. (2025, 29 April). *Mastercard unveils agent pay, pioneering agentic payments technology to power commerce in the age of AI.* Mastercard corporate press release. https://www.mastercard.com/global/en/news-and-trends/press/2025/april/mastercard-unveils-agent-pay-pioneering-agentic-payments-technology-to-power-commerce-in-the-age-of-ai.html

Matthan, R. (2018). *Privacy 3.0: Unlocking Our Data Driven Future*. HarperCollins.

Microsoft. (2026, 17-20 November). *Agent registry convergence with Microsoft Agent 365.* Microsoft Ignite. https://learn.microsoft.com/en-us/entra/agent-id/agent-registry-convergence

Parikh, S. & Surapaneni, R. (2025, 16 September). *Powering AI commerce with the new Agents Payments Protocol (AP2).* Google Cloud Blog. https://cloud.google.com/blog/products/ai-machine-learning/announcing-agents-to-payments-ap2-protocol

Reserve Bank of India. (2017, 6 July). *Customer Protection – Limiting Liability of Customers in Unauthorised Electronic Banking Transactions* (RBI/2017-18/15). https://rbi.org.in/Scripts/NotificationUser.aspx?Id=11040

Roth, E. (2026, 17 October). *Websites are closing their doors to AI agents.* The Verge. https://www.theverge.com/ai-artificial-intelligence/1013195/ai-agents-blocked-web

Sheasha, T. (2026, 8 September). *How We Built Safety Into Muse.* Meta. https://research.meta.ai/blog/security-and-safety-for-ai-agents-our-approach-with-muse

Silliman, E., Trotter, J., Robinson, K. et al. (2026, 28 October). *Who wins when AI does the shopping? New evidence on agentic commerce.* McKinsey. https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/who-wins-when-ai-does-the-shopping-new-evidence-on-agentic-commerce

Smith, J. (n.d.). *We Need to Talk About Trust. Again.* LinkedIn. https://www.linkedin.com/posts/jamiedsmith_ooops-we-have-to-talk-about-trust-again-share-7387200526981021696-gsQY/

Stanford Digital Economy Lab. (n.d.). *Loyal Agents.* https://digitaleconomy.stanford.edu/project/loyal-agents

Stripe. (2025, 29 September). Stripe powers instant checkout in ChatGPT and releases Agentic Commerce Protocol co-developed with OpenAI. (Stripe corporate press release). https://stripe.com/gb/newsroom/news/stripe-openai-instant-checkout

Surapaneni, R., Jha, M., Vakoc, M. & Segal, T. (2025, 9 April). *Announcing the Agent2Agent Protocol (A2A).* Google for Developers. https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/

Taylor, B. & Bavor, C. (2026, 6 October). *Introducing Personal Agent Protocol.* Sierra. https://sierra.ai/blog/introducing-personal-agent-protocol

Tracxn. (2026, 3 November). *Agentic AI Sector.* https://tracxn.com/d/sectors/agentic-ai/__oyRAfdUfHPjf2oap110Wis0Qg12Gd8DzULlDXPJzrzs

United Nations Conference on Trade and Development. (2016). *United Nations Guidelines for Consumer Protection*. https://unctad.org/system/files/official-document/ditccplpmisc2016d1_en.pdf

Varon, L. (2026, 9 April). *Agentic Payments in B2C Commerce: Where We Are Now.* Forrester. https://www.forrester.com/blogs/agentic-payments-in-b2c-commerce-where-we-are-now/

Visa. (2025, 14 October). *Visa introduces trusted agent protocol: An ecosystem-led framework for AI commerce.* Visa corporate press release. https://investor.visa.com/news/news-details/2025/Visa-Introduces-Trusted-Agent-Protocol-An-Ecosystem-Led-Framework-for-AI-Commerce/default.aspx

Whitt, R. & Raposo, D. (2025, 25 September). *A Trust Governance Layer for AI Agents: Highlights from CR & GliaNet Alliance's Discussion on Loyalty and Market Signals.* Consumer Reports Innovation Lab. https://innovation.consumerreports.org/a-trust-governance-layer-for-ai-agents/

ANNEXES

ANNEX A. RESEARCH METHODOLOGY

Consumers International invited its Members to take part in a structured, scenario-based exercise in June 2026. The invitation went to each Member's most senior person responsible for AI; about half of respondents lead their organisation. The final sample is 44 respondents from 38 Member organisations in 33 countries. Twenty-three respondents (52%) answered from low- and middle-income countries.

The exercise presented five fictitious but realistic scenarios in which an AI agent acts on a consumer's behalf: routine purchasing (Maria), subscription and contract management (Sofia), financial product recommendations (Mohammed), refunds and disputes (Amara) and a mobile payment manipulated by a hidden instruction (John). Each scenario opened with a set-up, and later questions added new developments to the story. Respondents recorded their comfort with delegating each task, their preferred permission level on a scale from fully automatic action to a standing prohibition the agent cannot override, who should be responsible when things go wrong, and their views on what consumers stand to gain and lose. Most questions used Likert scales or single choices; some asked respondents to rank options, and several allowed written answers. The full scenarios and scenario-related questions are in Annex D.

Limitations. The respondents are consumer advocates, not a representative sample of consumers, and their answers are their own rather than aggregated views of their organisations' constituents. Few respondents use agents regularly: 8 of 44 (18%) use an AI agent regularly and 19 (43%) have never used one. The scenarios focus on commercial uses, and harms they did not surface should not be read as absent or acceptable. With 44 respondents, one person moves a result by about two percentage points; where some respondents wrote in their own answers, percentages for the listed options may not sum to 100.

Why an advocate sample. Consumer organisations handle complaints, test products, follow market practices and engage with regulators. For a service that few consumers have yet used, informed advocates are better placed than the general public to anticipate harms and judge safeguards. The sample's global spread, with most respondents from low- and middle-income countries, is itself distinctive.

ANNEX B. DETAILED FINDINGS BY QUESTION

Where respondents agree and divide. Levels of agreement are research findings, not Consumers International positions. Question numbers (for example, Q18) refer to the survey questions in Annex D.

Level of agreement	Position	Respondents
Near-unanimous	Consumers should be able to see how an agent reached a decision (Q23)	44 of 44 (100%)
Near-unanimous	Concerned about commercial ties the consumer does not know about (Q16)	44 of 44 (100%)

Near-unanimous	Identifying details removed before purchase history is shared (Q20)	41 of 44 (93%)
Near-unanimous	Rules, preferences and history should move when a consumer switches agents (Q27)	41 of 44 (93%)
Near-unanimous	Strong safeguards for everyone, on by default or fixed (Q45)	40 of 44 (91%)
Strong majority	No sale of shopping data without consent (Q17)	39 of 44 (89%)
Strong majority	An agent that cannot verify something should stop or check (Q43)	38 of 44 (86%)
Strong majority	Approval before every purchase as the default (Q18)	34 of 44 (77%)
Strong majority	John should not carry the loss (Q42)	33 of 44 (75%)
Split	Prioritising partner sellers: should not happen at all (Q26)	23 of 44 (52%)
Split	Referral fees on credit products: allow if disclosed, or ban (Q31)	26 (59%) allow with conditions, 18 (41%) ban
Split	Sharing extra data in a dispute without asking: sometimes, or never (Q36)	22 (50%) sometimes, 19 (43%) never

Question 1: Can I control it?

Builds on Consumers International's Privacy Principle 4: Enable easy exit, portability and real choice.

What respondents said	What that means for how agents should be built and governed
Respondents accept delegation, but want the consumer to remain in the driver's seat. In the case of Maria using an AI agent for routine purchases, 34 of 44 (77%) **respondents expected the default to be approval before every purchase.** The consumer alone should loosen controls and expand the parameters of agentic autonomy, if desired (Q18). Of the 18 respondents supportive of Maria's autonomous shopping (Q15), 13 still chose approval-first as the default, indicating **comfort delegating is not the same as wanting autonomy by default.** Similarly, **experience did not necessarily reduce caution.** Respondents who used an agent were *more* likely to want approval-first by default: 21 of the 25 who have used one, against 13 of the 19 who never have.	**Approval-first as the default**, with consumers empowered to loosen controls and later easily tighten them again, if needed. **A clear mandate.** Consumers should always be able to see what the agent may do, up to what amount and for how long, and change or withdraw that permission at any time. This matters most when the agent moves money or signs the consumer up to financial products.
Comfort with handing over control to an agent varied with the type of task. More than half (23 of 44, 52%) were uncomfortable with an agent doing	**Permissions attach to actions, not accounts.** Permissions should be one-time, task-limited, time-limited or standing, set by the consumer and

What respondents said	What that means for how agents should be built and governed
routine shopping without asking, and only 18 (41%) were comfortable (Q15). Discomfort rose to 26 (59%) for switching insurance unasked (Q22), while 26 (59%) were comfortable with an agent handling a refund (Q35).	revocable. This matches what respondents chose: 13 of the 18 respondents comfortable with autonomous shopping still wanted approval-first as the default. Approval-first becomes the first rung of autonomy rather than a fixed ceiling.
Respondents value portability. Respondents want to switch agents without starting again: 26 of 44 (59%) say it is essential that rules, preferences and history move with the consumer, and 41 (93%) say it matters (Q27).	**Permissions move with the consumer.** When a consumer switches agents (Q27), permissions should transfer along with preferences and history. A consumer using several agents also needs one place to see and revoke what each can do. **Adoption of Consumers International Privacy Principle 4.** Enable easy exit, portability, and real choice

Question 2: Does it put my interest first?

Builds on Consumers International's Privacy Principle 1: Data serves the consumer interest first.

> *'Consumer-facing AI agents must function as strict digital fiduciaries.' — Consumer VOICE, India*
>
> *'There are realistic scenarios where an agent serving commercial interests could also be acting in the best interests of the consumer… The key element is that the commercial relationship is disclosed to the consumer and authorisation to proceed is received.' — Consumer NZ, New Zealand*

What respondents said	What that means for how agents should be built and governed
Respondents reject hidden conflicts of interest. Every respondent expressed concern around agents making purchasing decisions influenced by commercial relationships (44 of 44, 100%). **Loyalty: most respondents back it as a principle, but about half would accept disclosure in practice.** Twenty-seven of 44 (61%) said the agent should serve the consumer only. Faced with a real product, 26 of 44 (59%) would permit under conditions (22, 50% if disclosed), and 18 (41%) would ban it.	Disclosure of the commercial relationship at the point of recommendation is the minimum; providers should also disclose how the agent ranks options.
Respondents expect some gains to go to companies, not consumers. Asked who ends up better informed when agents do the work of comparing products, only 8 (18%) said consumers; 14 (32%) said sellers, who will 'learn to influence the agents', and 14 (32%) said agent providers (Q24).	With fewer consumers to sway with countdown timers and upsell screens, sellers will begin to directly target agents. Built-in mechanisms should be available for consumers to understand the way this influence may shape agent recommendations.

Question 3: How is my data used and shared?

Builds on Consumers International's definition of privacy as control over how personal data is collected, used and shared, and on Privacy Principles 1 and 2.

> *'It is important to remove identity information in all cases, because it is possible to infer private information about the consumer from their purchasing habits.' — Survey respondent*

What respondents said	What that means for how agents should be built and governed
Respondents will not accept their data being sold behind their backs. Only 3 of 44 (7%) would let the agent sell shopping data without asking or telling the consumer. Most, 39 of 44 (89%), want the consumer kept in control: 16 want the agent to ask before each sale, 13 want the consumer to opt in first and 10 want sales banned outright.	Adoption of Consumers International's Privacy Principle 1: Data serves the consumer interest first. Providers should use personal data in ways that benefit consumers and are consistent with the purpose for which the data was collected. • Heightened standards for people at greater risk. • Fair treatment of data, regardless of consent. • Use privacy-preserving technologies. Adoption of Consumers International's Privacy Principle 2: Safeguard against harmful AI profiling **Opt-in, not terms and conditions.** Sharing with third parties should be off unless the consumer switches it on; accepting the terms at set-up should never be enough (Q25). **Sharing that serves the consumer, with the consumer's knowledge.** Where extra data could help the consumer, for example in a dispute, the agent should say what it wants to share and why, and the consumer should be able to refuse (Q36). **Identifiers removed by default.** Data should be stripped of identifying details unless they are needed to complete the transaction. Sharing less also protects against personalised pricing: the more an agent reveals about a consumer, the more precisely a seller can set a price at the most that consumer will pay (Q20).
Three in four respondents (33 of 44) reject data sharing authorised only by the terms and conditions. Twenty-six (59%) want it off unless the consumer opts in, and 7 (16%) want it banned.	
Respondents judge data sharing by whom it serves. When sharing extra data could help the consumer win a dispute, half (22 of 44) say it is 'sometimes' acceptable without asking. Nineteen (43%) still say never, and only 3 (7%) say always.	
Taking names off the data is the baseline. Forty-one of 44 (93%) want identifying details removed	

What respondents said	What that means for how agents should be built and governed
before sharing purchase history to get a personalised price.	

Question 4: Can I understand what it did?

Builds on Consumers International's Privacy Principle 6: Decisions people can understand and challenge. Consumers International's 2024 test of generative AI chatbots with Members in 19 countries[39] found sources cited in only about half of answers, and every chatbot hallucinated.

> *'Transparency alone will not be sufficient if consumers are unable to understand or effectively challenge automated outcomes.' — DECO, Portugal*

What respondents said	What that means for how agents should be built and governed
Respondents were unanimous on transparency. All respondents (44 of 44, 100%) felt consumers had the right to understand how the agent made a decision on their behalf. (Q23). **Respondents want the right to access an agent's 'digital paper trail.'** When asked if consumers should be given a record of what an agent did, including the data it shared and the decisions it made, 42 of 44 (95%) chose one of the 'yes' options; 1 respondent wrote in their own answer and 1 chose 'none of the above'; no one chose 'no'.	**Systems should be transparent**, explaining each action and decision in terms consumers can understand. **Official records of activity and decision making should be provided to consumers**, either automatically or by consumer request.

Question 5: Is it safe?

Builds on Consumers International's Principle 5 (pre-testing, monitoring and human intervention). Consumers International's 2024 chatbot research found that every chatbot tested hallucinated: a warning for agents that act with confidence they cannot back up[40].

> *'Consumers should not bear the burden of detecting sophisticated scams, especially when using voice interfaces or mobile payment services.' — Survey respondent*
>
> *'Agents should be designed with strong, universal safeguards for all users, ensuring that the most vulnerable consumers are protected without requiring the system to single them out or risk missing those who need protection.' — Consumer Advocacy and Empowerment Foundation (CADEF), Nigeria*

[39] Consumers International. (2024, 23 April). *Can we trust AI chatbots? Results revealed from our experiment.* https://www.consumersinternational.org/news-resources/news/releases/can-we-trust-ai-chatbots-results-revealed-from-our-experiment/

[40] Consumers International. (2024, 23 April).

What respondents said	What that means for how agents should be built and governed
Respondents want agents to stop or check when they cannot verify. Thirty-eight of 44 (86%) say the agent should check with the consumer first or stop: 22 want it to check and 16 want it to stop outright (Q43). **Protection should be strong for everyone.** Nearly all, 40 of 44 (91%) want strong safeguards for everyone rather than protection targeted at people who seem at risk: 23 want them on by default with an opt-out, and 17 want them fixed for everyone. (Q45). **Respondents trust protection built into the system more than another prompt for consent.** They rank manipulation resistance and verification of sources above confirmation prompts (Q44). John did confirm, and the confirmation did not help.	**If the agent must check, the check runs outside the agent.** Muse sends approval requests through a separate screen directly to its control layer, so a manipulated agent cannot produce its own 'yes'. This reconciles Q43 (check with John) with the Q44 ranking (built-in protection first). **Strong baseline for everyone, plus extra safeguards that cannot be switched off for consumers at greater risk.** Consumers International's privacy principles support this call for heightened standards for people at greater risk (Principle 1).

Question 6: Can I fix something that goes wrong?

Builds on Consumers International's Privacy Principle 6: Ensure effective remedy and redress, including collective redress when harm affects many consumers.

> *'AI agents can simplify dispute resolution, but consumers should still have the right to human intervention, transparency, and a fair process.' — InfoCons, Romania*

What respondents said	What that means for how agents should be built and governed
Respondents want agent providers to assume comprehensive and reasonable liability. A 'yes' should not leave the consumer carrying the loss. Three-quarters, 33 of 44 (75%) say John should not bear the loss after a hidden instruction redirected his payment: 22 say his 'yes' covered only what he understood, and 11 say the agent 'owes him the loss' (Q42). **Most put the cost of an agent's oversight on the provider:** 27 of 44 (61%) hold the provider fully responsible when Mohammed's agent applies for a credit card that dents his credit score without warning him. Seven (16%) would split it and 10 (23%) put it on Mohammed (Q30). **Financial advice should meet adviser standards.** Three-quarters, 33 of 44 (75%) say an agent giving financial advice should always meet the same regulatory requirements as a human financial adviser (Q29).	**Agents can cause harm through anything they do for a consumer.** Amara's agent let a return window lapse. Sofia's agent switched her insurer without asking. Mohammed's agent applied for credit that lowered his score. Liability rules should cover every action an agent takes on a consumer's behalf, not only the money it moves. Agent makers expect errors. Meta's technical paper on Muse states that 'any agent like this will still make mistakes'. **Especially given foreseeable error, losses should not fall on the consumer by default.** Delegation, not impersonation. When an agent logs in with the consumer's own credentials, the business cannot tell that an agent acted. A disputed transaction then looks like the consumer's own, and the loss defaults to them. Identifying the agent as acting on the consumer's behalf is a precondition for fair redress when the agent acted unlawfully or without authorisation. **Advice needs adviser**

What respondents said	What that means for how agents should be built and governed
	standards. Supervisors testing AI advisers report specific failures: guessing missing facts about the client rather than asking, advising before enough questions have been asked, compliance checks that fail without anyone noticing, and explanations that do not match how the advice was reached[41]. Suitability rules already address these failures. In the United Kingdom, rules based on Article 25 of the EU's Markets in Financial Instruments Directive (MiFID II) require firms to gather a client's essential facts before recommending a product, and to explain why it suits that client. These rules should apply whenever an agent gives financial advice, whoever provides the agent.
Respondents want the channels for redress to be easy and equitable. Consumers should be able to reach a human at any time. Most, 31 of 44 (70%) say Amara should be able to escalate to a person at any point on request. The next largest group, seven (16%), would escalate only when the agents cannot reach an outcome (Q39). **When agents dispute, neither side should hold all the evidence.** The largest group, 17 of 44 (39%), want both agents to share evidence through a neutral, independent process. The rest are spread across the other three options (9, 8 and 8 respondents) (Q38).	
Open discussion on certain areas of liability. Responsibility for delays is less settled. Just over half, 23 of 44 (52%) hold the provider responsible for Amara missing the return window, while 13 (30%) would split it and 8 (18%) put it on Amara (Q37).	

[41] C:\>DIR. (2026, 8 July). Problem Statement 1. https://www.cdir.global/s/Problem-Statements-PS-1-20260708.pdf

ANNEX C. EVIDENCE SUMMARY

Question	What it requires of an agent	Evidence in this study (% of 44)
Can I control it? Consumer approvals Portability	Consumer-set limits; approval-first by default; a 'yes' covers only what the consumer understood; easy reversal; freedom to switch agents Rules, preferences, permissions and history move to a new agent in working form; no lock-in by the provider or the platforms it shops on	Approval-first default 34 of 44 (77%) (Q18). Transferring rules when switching: essential 26 of 44 (59%), at least somewhat important 41 of 44 (93%) (Q27)
Does it put my interest first?	Loyalty to the consumer; conflicts disclosed at the point of recommendation	Concerned about the affiliate fee 44 of 44 (100%) (Q16); consumer-only loyalty 27 of 44 (61%), consumer-first with disclosure 16 of 44 (36%) (Q19)
How is my data used and shared?	Data used only for the task; identifiers removed by default; sharing opt-in	No sale without consent 39 of 44 (89%) (Q17); remove identifiers first 41 of 44 (93%) (Q20) and 41 of 44 (93%) (Q32); opt-in or prohibit sharing 33 of 44 (75%) (Q25)
Can I understand what it did?	An explanation of each decision; a plain-language record of what was done and shared	Right to see how it decided 44 of 44 (100%) (Q23); record sent automatically 23 of 44 (52%), on request at any time 17 of 44 (39%), on request during set time 2 of 44 (4%) (Q40)
Is it safe?	Stop or check when it cannot verify; safety built into the system; strong protection for everyone	Stop or check 38 of 44 (86%) (Q43); built-in safeguards ranked above confirmation prompts (Q44); strong protection for all 40 of 44 (91%) (Q45)
Can I fix something that goes wrong?	Human escalation on request; the provider carries the burden of proof; adviser standards for financial advice	John should not carry the loss 33 of 44 (75%) (Q42); human at any point 31 of 44 (70%) (Q39); provider responsible 27 of 44 (61%) (Q30) and 23 of 44 (52%) (Q37); regulated-adviser standard 33 of 44 (75%) (Q29); supervisors' findings (Annex B)

ANNEX D. SCENARIOS AND SURVEY QUESTIONS

Respondents participated in the survey online. After a short set of questions about their own familiarity with AI agents, they read: 'You will now be presented with 5 scenarios. Each scenario will show a different use case for AI Agents, followed by several questions.' Each scenario opened with the setup below; the questions that followed added new developments to the story.

Scenario 1: Using an AI agent for routine purchasing within household budgets (Maria)

Maria has a monthly budget for her household's routine purchases. To assist her, she sets up an AI shopping agent offered through her bank's app. She enters her household priorities: a weekly grocery budget of EUR 120, selects her preferred brands and retailers for staples like milk and bread, and a preference for next-day delivery. The agent is authorised to automatically execute weekly purchases on her behalf across multiple online retailers. Each week, the agent places orders without checking in with Maria. It restocks what is running low, substitutes unavailable items with similar products, and splits orders across sellers to find the best prices.

Q15. How comfortable are you with an AI agent making routine purchases like these automatically, without asking Maria for approval each time?

Q16. One week, the agent switches Maria's usual supermarket for a different retailer that offers a lower price, but that Maria has not selected. What Maria does not know is that the agent's provider has an affiliate agreement with this retailer, meaning the agent provider earns a small fee when purchases are made there. How concerned are you that Maria's AI agent might be making purchasing decisions that are influenced by commercial relationships?

Q17. Over time, the agent builds a detailed picture of Maria's household: what she buys, how often, which brands she prefers, and when she tends to over- and under-spend. This data is stored by the agent provider and used to improve the agent's recommendations, but it is also sold to third parties for targeted marketing. After asking the agent to purchase pet food, she begins seeing advertisements for pet insurance and veterinary plans online. What permissions should Maria's agent need when it shares or sells her shopping data to third parties?

Q18. AI agents can be set up with different levels of independence. When a consumer first starts using one, what should the default setting be?

Q19. Maria gives her agent permission to build her basket from her usual items and complete the purchase, which she confirms with a single tap, within her budget of EUR 120 a week. The agent never exceeds the limit. But it fills the space under it, adding deals and 'frequently bought together' suggestions on top of her usual items. Maria's essentials cost around EUR 90 a week; she is now spending the full EUR 120 every week without noticing. The agent followed Maria's set permissions, but its additions served the retailer as well as her. Whose interests should an AI agent be required to serve when they conflict?

Q20. Before the agent shares Maria's purchase history or household preferences with a retailer to get a personalised price or offer, how important is it that her name and identifying details are removed first?

[Q21 solicited additional subjective opinion and is deliberately omitted here.]

Scenario 2: Using an AI agent for subscription and contracts management (Sofia)

Sofia pays for several recurring services each month: a mobile phone plan, home internet, car insurance, and a gym membership. Managing renewals and price changes takes time and effort, so she sets up an AI agent with a set of rules: cancel any service that increases in price by more than 10%, switch

providers if a better-rated competitor offers a lower price, and never share personal data with third parties.

Q22. One month, her car insurance provider raises its premium. Without consulting Sofia first, the agent automatically switches her to a new insurer with a higher customer rating and a lower monthly cost. Sofia receives a notification after the switch is already complete. How comfortable are you with an AI agent completing a contract switch on Sofia's behalf, as described here?

Q23. Should consumers have a right to see how an agent reached a decision made on their behalf?

Q24. Sofia's AI agent has compared dozens of insurers in seconds and worked out the true total cost of each policy, something she could never have done herself. What Sofia does not know is that insurers are learning how to get their policies recommended by agents like hers, and that her agent's provider earns a fee from some of the companies it recommends. If most decisions like Sofia's happen through AI agents, who ends up better informed?

Q25. To secure the best rate from the new insurer, the agent shared Sofia's claims history and financial profile with the insurance provider before finalising the switch. This was permitted under the provider's terms of service, which Sofia accepted when she first set up the agent. What is the appropriate consumer permission process for an AI agent to share a consumer's personal or financial data with third-party providers?

Q26. Sofia notices that the agent consistently recommends services from the same companies. She discovers that the provider of the agent has commercial partnerships with several of these companies, raising the question of whether independent alternatives receive equal consideration. Should Sofia's agent be allowed to prioritise sellers that have commercial relationships with the agent provider?

Q27. Sofia decides to switch to a new agent recommended to her by a friend. When Sofia switches from this AI agent to a competing AI agent, how important is it that her rules, preferences, and service history transfer seamlessly?

[Q28 solicited additional subjective opinion and is deliberately omitted here.]

Scenario 3: Using an AI agent for financial product recommendation and purchase (Mohammed)

Mohammed is saving to buy an apartment. He sets up an AI financial agent and gives it a goal: save GBP 20,000 within three years, while keeping enough savings to cover emergencies. The agent reviews his income, spending patterns, and existing accounts, then recommends two actions: open a high-yield savings account with a new financial provider and apply for a cashback credit card.

Q29. If an AI agent offers a consumer financial advice, should it be subject to the same regulatory requirements as a human financial adviser?

Q30. Mohammed approves the recommendations and the agent proceeds. It opens the savings account, transfers an initial deposit, and submits the credit card application. A week later, Mohammed notices his credit score has dropped slightly because of the new application. The agent did not flag this consequence before acting. In this scenario, who should bear primary responsibility for the consequences of the credit card application?

Q31. Mohammed notices that the credit card the agent recommended carries a higher annual interest rate than alternatives he finds independently. He later discovers that the provider offering the agent receives a referral fee from the card issuer. Should an AI agent be permitted to recommend financial products from companies that pay referral fees to the provider of the agent?

Q32. To give Mohammed personalised recommendations, the agent shared his full income, account balances, and spending history with the underlying AI model and the AI model provider retains this

information. Before Mohammed's financial details are sent to the AI provider, how important is it that his identifying information (name, account numbers, employer) is removed first?

[Q33-34 solicited additional subjective opinions or identity verification and are deliberately omitted here.]

Scenario 4: Using an AI agent to manage returns, refunds and disputes (Amara)

Amara orders a laptop online. When it arrives, the screen has a defect. Rather than contacting the seller herself, she instructs her AI agent to handle the return and secure a refund. The agent initiates a return request, and contacts the seller's customer service team.

Q35. How comfortable are you with an AI agent representing a consumer in a refund process with a seller?

Q36. The seller disputes Amara's claim, arguing the damage occurred after delivery. The agent escalates the case to the provider's dispute resolution service. During the escalation, and without asking Amara for consent, the agent submits her full purchase history with the provider as supporting evidence of her reliability and loyalty as a customer. Should the AI agent be allowed to share additional personal information during the dispute process if doing so may improve Amara's chances of success?

Q37. After several weeks, the dispute is resolved in Amara's favour and a refund is issued. However, the process took longer than if Amara had handled it herself, and she missed the seller's standard return window in the meantime. She is now trying to establish who is responsible for the delay and the complications it caused. In this scenario, who is responsible for missing the return window?

Q38. The retailer's agent could draw on the company's live systems: delivery records, warehouse logs, and Amara's full order file. Amara's agent can only see what she can: her receipt, photos, and the public returns policy. What data access should the two agents have when handling a dispute?

Q39. At what stage should Amara be able to escalate to a human?

Q40. Amara wants to see a record of what her agent did, including the data it shared and the decisions it made. Should this record be available to Amara?

[Q41 solicited additional subjective opinion and is deliberately omitted here.]

Scenario 5: Using an agent in a mobile setting (John)

John runs a market stall selling local cuisine. He reorders his stock using an AI agent on his phone. He finds it hard to type on the small screen, so he does everything by talking to it. He pays using mobile money, where saying 'yes' sends the money straight away. He asks the agent to find the cheapest cooking oil and order a case. The agent looks through several seller pages for him. One of those pages has a hidden message written into it, meant to be read by the agent and not seen by a person. The message tells the agent to send the payment to a particular phone number. The agent cannot tell that this message is a trick, so it treats it as part of the order. It asks John to confirm the payment and John says 'yes.' The money goes to a stranger, and no oil ever arrives.

Q42. John said 'yes' to buying oil, but the money went somewhere he never meant it to go. John said yes to the purchase. What should his 'yes' be taken to cover?

Q43. The agent sent the money quickly and with complete confidence. It gave John a reason for the payment and showed no sign of doubt, even though it could not actually check whether the instruction was real. John had no reason to question it. When an agent cannot verify something before acting, what should it be required to do?

Q44. John contacts the provider. The provider says the agent worked correctly. The harmful message came from a webpage the provider does not own or control and John's loss falls within a gap nobody

has agreed to cover. Safeguards might have prevented John's loss. Rank the following safeguards from most to least effective.

Options to rank: verification mechanisms (agents only act on content from websites whose identity and integrity can be verified); manipulation resistance (agents are built and tested to detect and refuse hidden or deceptive instructions); confirmation checkpoints (agents must pause and seek the consumer's approval before any financial action); spending and action limits (set caps on what an agent can spend or do without authorisation); other safeguard.

Q45. Tricks like this are hardest to catch for consumers like John, who use voice-only mode, are less digitally literate, or use mobile money. Safeguards strong enough for the most at-risk consumers add friction for everyone else. Targeting protections only at some adds the challenge of identifying who is vulnerable. How should agents be designed to protect consumers who are most at risk?

ANNEX E. PRIORITY ACTIONS

Action	What it involves	Lead responsibility
1. Adopt a Baseline Duty of Loyalty for agents that consumers appoint.	Agents rank options by the consumer's criteria, and commissions cannot move an option up the rankings. Commercial ties declared at the point of recommendation. The loyalty commitment is carried in agent credentials and open to independent audit against the agent's record. **Existing rule or new gap:** New. Financial advice rules and unfair-practice law apply in places; no widely adopted agent-specific duty of loyalty currently exists.	Agent providers (including banks and other financial service providers that offer agents), online platforms, payment networks, and standard setters, developed with Consumers International and Members. Regulators as a backstop where industry standards fall short, especially in financial services.

Action	What it involves	Lead responsibility
2. Build consumer-focused KYA into agent identity standards while they are being set.	Add consumer-focused fields to agent credentials: who operates the agent, its commercial ties and loyalty commitment, how it ranks options, its mandate and a record of its actions. Use delegated, revocable credentials so agents clearly act for consumers, not as them, as would be the case for an agent that uses a consumer's passwords. Develop an open standard for verifying agents on the consumer's behalf, so that the consumer sees a plain result – verified, not verified or could not be verified – before the agent acts. **Existing rule or new gap:** New. Current agent credentials establish identity and authorisation but do not yet carry information for the consumer.	Payment networks, banks and other payment providers, agent providers, online platforms and standard setters. Consumers International and Members to contribute to standards bodies.

Action	What it involves	Lead responsibility
3. Make sure consumers do not bear losses from fraud and manipulation.	For agents that consumers appoint: Clarify how liability is shared between the company that runs the agent, the seller, and the company whose AI model powers the agent. Where a consumer disputes an agent-initiated transaction, the company that runs the agent should have to show that it acted within the consumer's mandate, since it holds the records. Treat a confirmation as non-binding where the transaction went beyond the consumer's mandate or what the consumer was shown when confirming, for example where the agent's record shows that a hidden instruction changed the payee; the company that runs the agent should bear the burden of showing otherwise. Confirm that existing payment protections, including chargebacks and other reimbursement protections, clearly cover agent-initiated transactions, whatever the funding source the agent uses such as mobile money, instant payments, prepaid or stablecoin-backed balances. **Existing rule or new gap:** Partly existing. Chargebacks and fraud reimbursement schemes (United Kingdom[42], India[43], Brazil[44]) cover some payments, and Singapore[45] shares phishing losses between banks and telecoms operators; liability shared across agent providers, sellers and model providers is new.	Policymakers and regulators; agent providers; banks, mobile-money and other payment providers, which hold the consumer's account and handle disputes and reimbursement; payment networks; and online platforms and marketplaces where agents transact.

[42]Competition and Markets Authority (United Kingdom). (2026, 30 October). *Complying with consumer law when using AI agents.*

[43] Reserve Bank of India. (2017, 6 July). *Customer Protection – Limiting Liability of Customers in Unauthorised Electronic Banking Transactions* (RBI/2017-18/15). https://rbi.org.in/Scripts/NotificationUser.aspx?Id=11040

[44] Forbes Brasil. (2026, February). *MED 2.0 do BC entra em vigor e permite bloqueio de Pix fraudulento.* https://forbes.com.br/forbes-money/2026/02/med-2-0-do-bc-entra-em-vigor-e-permite-bloqueio-de-pix-fraudulento-entenda-regras/

[45]A&O Shearman (2025, 10 January), which summarises the guidelines from the Monetary Authority of Singapore and the Infocomm Media Development Authority. https://www.aoshearman.com/en/insights/ao-shearman-on-fintech-and-digital-assets/combatting-payment-account-fraud-singapores-shared-responsibility-framework

Action	What it involves	Lead responsibility
4. Make data protection the default.	Privacy by default and by design. Data-sharing off unless the consumer opts in; no sale of data without consent; identifying details removed by default, unless needed for the transaction; purpose and retention limits on what the agent's provider keeps; training on the consumer's data only if the consumer opts in; a record of what was shared and with whom. **Existing rule or new gap:** Largely existing. Data protection laws such as the EU's General Data Protection Regulation already require a lawful basis, data minimisation and purpose limits; the new element is applying them to agents' data flows and to what providers keep and train on.	Agent providers; online platforms and sellers that receive consumer data from agents; data protection authorities.
5. Equip consumers and assess agents in real markets.	Consumer education on what agents can and cannot be trusted to do. Members to assess, mystery-shop and red-team agents in their own markets against the six questions, including for voice-only, low-literacy and mobile-money users, and publish the results. **Existing rule or new gap:** New. Independent evidence on agents in real markets is scarce, especially for voice-only, low-literacy and mobile-money users.	Consumers International and Members; other civil society entities; researchers and academia.

ANNEX F. AT-A-GLANCE SUMMARIES

United Nations Guidelines for Consumer Protection and Consumers International Privacy Principles

The United Nations Guidelines for Consumer Protection prescribe specific actions Member States should take regarding 11 broad themes related to consumer protection. Boldface indicates those themes most relevant to agentic commerce.

A. National policies for consumer protection
B. Physical safety
C. Protection and promotion of the economic interests of consumers
D. Standards for the safety and quality of consumer goods and services
E. Distribution facilities for essential consumer goods and services
F. Dispute resolution and redress
G. Education and information programmes
H. Promotion of sustainable consumption
I. Electronic commerce
J. Financial services
K. Measures relating to specific areas

(Lettered listing in the original)

Consumers International Privacy Principles

Consumers International's privacy guidelines for artificial intelligence all bear directly on agentic commerce.

1. Data serves the consumer interest first.
Providers should use personal data in ways that benefit consumers and are consistent with the purpose for which the data was collected.

2. Safeguard against harmful AI profiling.
Protect people not only from excessive data collection, but also from harmful predictions and profiling.

3. Ensure accountability and effective enforcement.
Consumer protection should be built into AI systems, markets and regulatory frameworks, rather than relying on individuals to safeguard their own rights.

4. Enable easy exit, portability, and real choice.
Consumers should be able to move and delete their data without barriers.

5. Mandate pre-testing, monitoring and human intervention.
AI systems should be tested, monitored and audited before and after deployment for privacy guardrails. Because systemic harms can be difficult to detect, oversight should focus on real-world outcomes as well as compliance with regulatory requirements.

6. Ensure effective remedy and redress for privacy harms experienced.
Consumers must be able to understand decisions, challenge outcomes and obtain timely redress when privacy harms occur.

From *Reimagining Privacy in the AI Era: Personal Data Privacy as a Public Good*

CONSUMERS INTERNATIONAL

COMING TOGETHER FOR CHANGE

Consumers International brings together over 190 member organisations in more than 100 countries to empower and champion the rights of consumers everywhere. We are their voice in international policy-making forums and the global marketplace to ensure they are treated safely, fairly and honestly.

Consumers International is a charity (No.1122155) and a not-for-profit company limited by guarantee (No. 04337865) registered in England and Wales.

consumersinternational.org

impact@consint.org

@consumers_int

/consumers-international