radware

Infographic

3 Questions to Guide Agent Trust Decisions

AI agents act on behalf of human users online, creating new opportunities, but their activity can also introduce security, fraud, and business risks.



Can you distinguish valuable AI agent activity from risky activity?

Agent activity that can create value

- Searches for information
- Discovers and compares products
- Completes user-requested tasks
- Initiates actions or transactions
- Interacts with digital services

Agent activity that may signal risk

- Misrepresents its identity
- Deviates from its declared intent
- Attempts unexpected or harmful actions
- Tries to access restricted workflows
- Exhibits anomalous behavior

Three questions to ask:

1. "Can I identify the agent?"

Determine what is interacting with your application:

- Is the traffic generated by an AI agent?
- Which agent or platform is behind it?
- Can its identity be cryptographically verified?
- Do technical and behavioral signals support identity claims?

Why it matters:

Knowing an agent's identity does not guarantee trust, but it provides essential context for evaluating risk and trust.

2. "Can I understand its intent?"

Identification tells you *who*. Intent tells you what it is trying to accomplish, and that's essential for assessing trust.

Understand the agent's intent and assess whether its actions remain aligned with it. For example, is the agent trying to:

- Search for information
- Browse products
- Access an account
- Complete a transaction

Why it matters:

Even a verified agent can behave unexpectedly, become misdirected, perform harmful actions and more. Trust must be continuously assessed.

3. "Can I govern what the agent can do?"

Once you understand identity and intent, decide which actions and workflows it should be permitted to access.

Move beyond allowing or blocking entirely to instead evaluate agent trust to:

- Allow access to approved activities
- Apply controls where neededRestrict access to sensitive workflows
- Restrict sensitive activitiesBlock actions that present unacceptable risk
- Block unacceptable risk

Why it matters:

An agent may be permitted to browse product pages while being restricted from sensitive workflows such as login or checkout. Granular governance enables valuable activity without giving the agent unrestricted access.

The value of Agent Trust:

- Enable trusted agent engagement
- Contain agent-driven risk
- Make informed agent-access decisions
- Retain control over agent-driven journeys
- Participate confidently in the agentic economy

Radware Agent Trust Management

Identify, continuously assess trust, and govern AI agent activity.

Learn More



Source: *Building Trust in the Age of AI Agents blogpost*

© 2026 Radware Ltd. All rights reserved. The Radware products and solutions mentioned in this report are protected by trademarks, patents and pending patent applications of Radware in the U.S. and other countries. For more details, please see https://www.radware.com/LegalNotice/. All other trademarks and names are the property of their respective owners.

RW86 465 | 09.25.2026

radware