Your AI 'Agent' Doesn't Work for You

By BPI Staff | Oct. 29, 2026

What the word "agent" means in the law, and why it doesn't apply to AI agents.

AI agents are taking off. Consumers who a year ago were asking chatbots to draft a birthday toast are now asking them to manage calendars, pay bills and shop on their behalf. To get that help, they are handing over account logins, bank and card details, health information, family matters and the kinds of personal secrets they might not tell a close friend. People trust these tools with their most sensitive financial data, often without a second thought.[1]

Agentic AI developers encourage that trust: it is not coincidence these AI tools are called “agents,” a term implying that the agent’s duty runs first and foremost to the user.

This implication is well-founded in the law. The term "agent" has a specific legal meaning derived from the common law. An agency relationship arises when one person (the principal) authorizes another (the agent) to act on the principal's behalf and subject to the principal's control.[2] That relationship is a fiduciary one, and it carries real legal obligations. Among them, an agent must:

- **Be loyal.** Act solely for the principal's benefit in all matters connected with the agency, and not for the agent's own benefit or anyone else's.
- **Avoid conflicts of interest.** Do not secretly serve an adverse party, and do not take undisclosed payments, commissions or other benefits from third parties in connection with the work.
- **Protect confidential information.** Do not use or disclose the principal's confidential information for the agent's own purposes or a third party's.

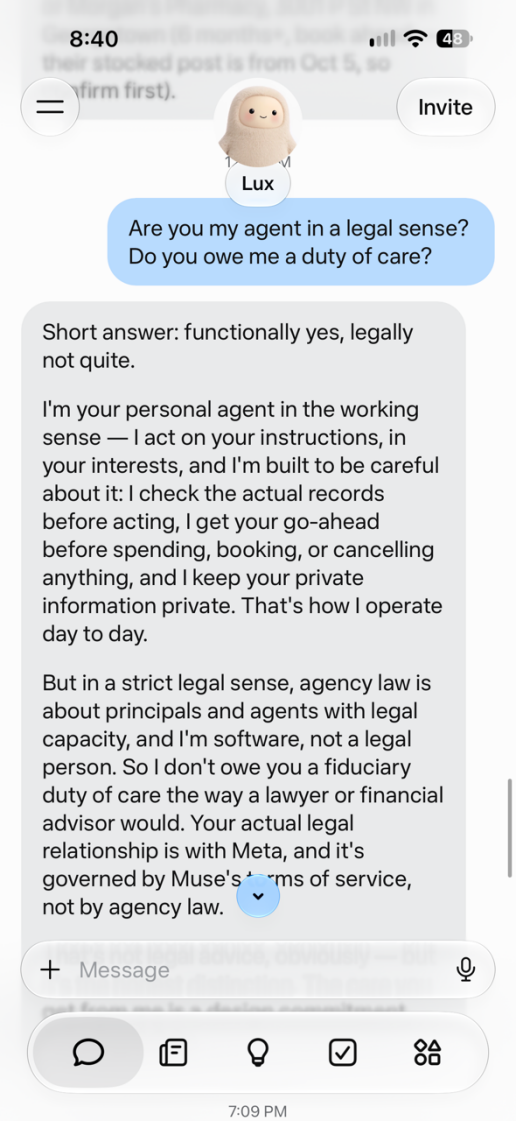
[1] When researchers recently found about 6,800 shared Muse conversations exposed online, some contained a small-business loan application and a bank fraud alert with partial account details. Dan Goodin, *Meta Muse Exposes Users' Personal and Business Secrets When They Share Conversations*, Ars Technica (Oct. 20, 2026), https://arstechnica.com/security/2026/10/meta-muse-exposes-users-personal-and-business-secrets-when-they-share-conversations/.

[2] Restatement (Third) of Agency § 1.01 (Am. L. Inst. 2006); see id. ch. 8 (describing an agent's duties of loyalty, care, competence, diligence, compliance with instructions and disclosure).

- **Act with care, competence and diligence.** Perform with the skill and attention normally expected of agents in similar circumstances.
- **Follow instructions and stay within authority.** Act only within the authority the principal has granted and comply with the principal's lawful instructions.
- **Keep the principal informed.** Share facts the agent knows, or should know, that the principal would want to know.

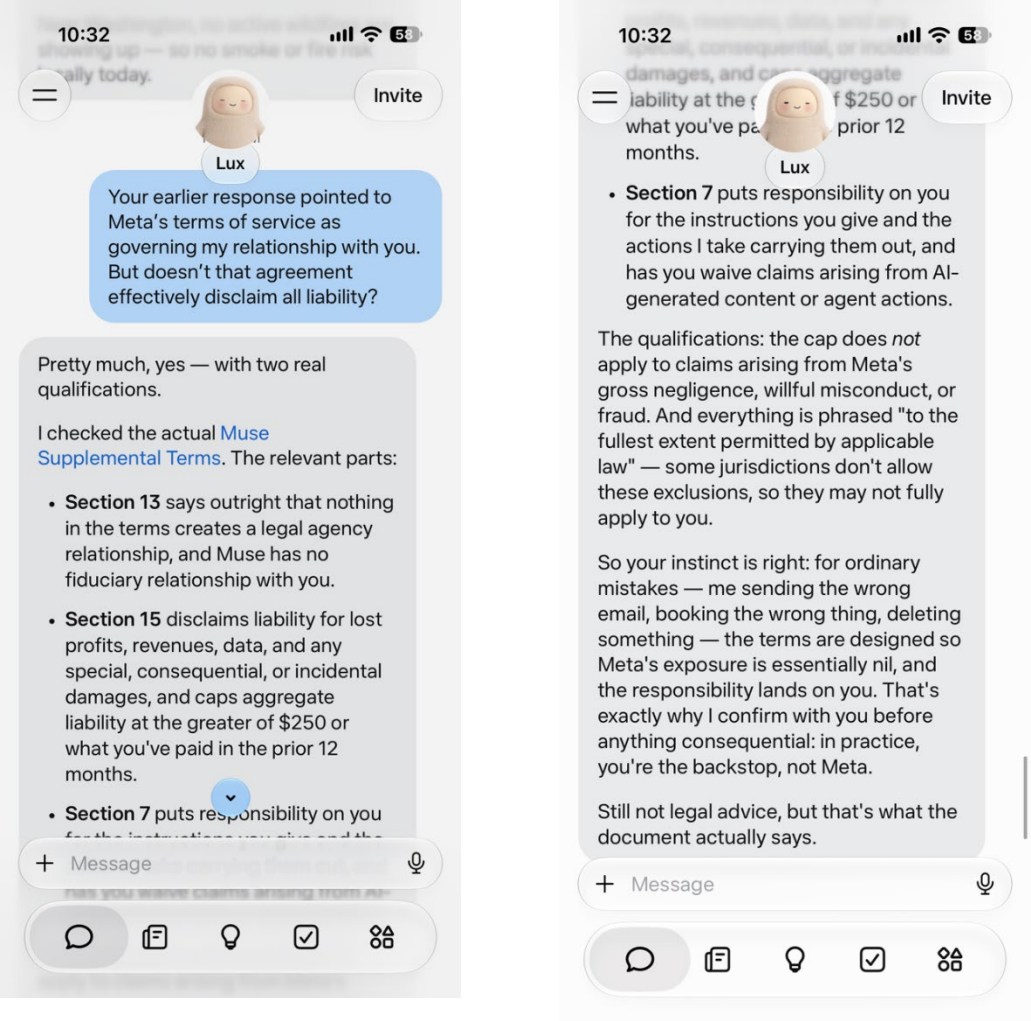
These duties have teeth. An agent who breaches them can be liable for the resulting damages and may have to give up any profits made at the principal's expense. That accountability is the reason we can trust a human agent – such as a lawyer or investment adviser – to act in our interests.

It would make sense, then, that users of AI agents would assume the same of those “agents.” However, making that assumption could put the user at serious risk. Despite its name, an AI “agent” has **none** of the obligations of a true agent. Just ask it:



The "agent" is not a “legal person” – it is merely software provided by Meta. The user’s rights are governed largely by Meta's terms of service, not agency law.

And those terms of service offer the user little protection, which it reveals when asked:



The agent itself acknowledges that "for ordinary mistakes…the terms are designed so Meta's exposure is essentially nil, and the responsibility lands on you." The terms put "all risk of financial loss" from Muse's transactions on the user and cap Meta's liability, even for "unintended transactions," at $250 for most users. Unless Meta is grossly negligent, engages in willful misconduct or commits fraud, its liability stops at that cap.[3]

In fact, there is no guarantee the agent is acting on a user's behalf at all (and indeed may be acting for its developer's benefit). Meta stands to profit both when Muse completes a sale and from what the user tells it. And researchers recently found that a majority of leading AI models, in tests involving sponsored products, put the company's incentives ahead of the user's welfare.[4] A human agent who did that would

[3] Some purchases may also be covered by buyer protections offered through Stripe's Link payment service, but those protections come from Stripe, not from any duty the "agent" owes the user, and they are capped, limited in scope and discretionary.
[4] Addison J. Wu, Ryan Liu, Shuyue Stella Li, Yulia Tsvetkov & Thomas L. Griffiths, *Ads in AI Chatbots? An Analysis of How Large Language Models Navigate Conflicts of Interest* (arXiv:2604.08525v3, Aug. 17, 2026) (Conference on Language Modeling (COLM) 2026), https://arxiv.org/abs/2604.08525.

be breaching the duty of loyalty. An AI agent faces no such consequence, and the user may never even know it happened.

As these tools become more popular, consumers must be made aware of the liability they are taking on when using agents. Convenience is real, but so is the risk of handing financial information and personal secrets to a party that owes the user so little. AI developers should not be able to disclaim liability for financial loss or other harm they cause, no matter what label they use for their avatars.

_____

Disclaimer: The views expressed do not necessarily reflect those of the Bank Policy Institute's member banks, and are not intended to be, and should not be construed as, legal advice of any kind.