

Brainwave October 2026

Your AI Agent Needs a Boss

By [Brandi Pack](#), Director of Innovation, [UpLevel Ops](#)

“You are a highly capable junior assistant. Your job is to complete the work I assign you, but you are not the final decision-maker. If something important is unclear, if instructions conflict, or if you encounter a situation that requires judgment beyond what I have given you, stop and ask me.”

This is **not** how I used to prompt AI.

I used to tell my assistants to act like experts.

“You are an experienced attorney.”

“You are an expert legal researcher.”

“You are a senior financial analyst.”

That seemed reasonable. If I wanted a legal answer, I would tell the assistant to think like the best lawyers. If I wanted research, I would tell it to act like a senior researcher.

That works well enough when AI answers a question or drafts something for me to review. But an agent is doing more. I am asking it to carry out work, make decisions, use tools, respond to new information, and get from an assignment to a result.

That changes what I need from it. I do not want an agent that assumes it is the smartest person in the room. I want one that understands the limits of its authority.

Think about a really good junior assistant. You do not want that person interrupting you every five minutes with questions they should be able to answer themselves. But you also do not want them confidently making consequential decisions because they were afraid to admit they were unsure.

A good assistant knows the difference.

If the instructions conflict, ask.

If critical information is missing, ask.

If something materially changes the assignment, ask.

And in particular: *If the next step requires judgment you were not given authority to make, come find me.*

I shouldn't have to constantly peer over an agent's shoulder to make sure it hasn't quietly gone off course. The agent should know when it needs me.

Real Work Is Messy

Most agent workflows look pretty good when everything goes according to plan.

Unfortunately, that is not how most of us work.

We interrupt each other. Priorities change. A document turns out to be outdated. A deadline moves. Two people give slightly different instructions. A client asks for something that changes the original assignment.

Humans navigate these small disruptions all day long, often without thinking much about them. We understand that new information changes what our boss meant yesterday, or that an unusual request probably needs approval before we act.

Agents do not necessarily make those judgments the same way.

An agent may treat new information as another instruction to incorporate. It may try to reconcile conflicting directions or choose the interpretation that allows it to keep moving, especially if it has been told to prioritize speed, urgency, or task completion.

That is where capability can become part of the problem.

The better agents get at reasoning through obstacles, using tools, and finding alternative paths to an objective, the more important it becomes that they understand when not to solve the problem themselves.

An agent may see a path to completing an assignment without recognizing that the path exceeds its authority or creates unintended consequences. We have spent a tremendous amount of effort teaching AI how to keep going, and we also need to teach it when to stop.

Teach the Agent to Be Unsure

For much of the generative AI era, we worried about hallucinations. AI would confidently make something up, and we learned to verify its answers.

Agents create a more consequential version of that problem.

When a chatbot gives me a questionable answer, I can catch it before anything happens. When an agent is acting on my behalf, confidence can turn into action before I ever see the reasoning that got it there.

So I have started thinking differently about uncertainty.

I do not just want to tell an agent what to do. I want to tell it what kinds of uncertainty should bring it back to me.

It could be missing information that requires an assumption, a source that does not support the conclusion as strongly as expected, or an action that cannot easily be reversed. In human

organizations, knowing when you do not know something is a form of competence. Good junior employees learn which decisions are theirs and which ones need to move up the chain.

Agents need their own version of that boundary.

The goal is not an agent that asks me what to do every time something unexpected happens. It is an agent that knows which unexpected things matter.

Human in the Loop ≠ Mean Humans Watching the Loop

The more I think about agents, the more I suspect human judgment may be one of the biggest remaining barriers to giving them substantial amounts of real work.

The answer isn't to have a human watch everything an agent does. If I have to monitor every step, check every decision, and make sure the agent has not wandered off course, I have lost much of the benefit of using an agent in the first place.

We need to reverse that relationship.

The agent should be monitoring its own work for moments when human judgment is needed.

This becomes more important as agents become more capable. An agent can be extraordinarily good at accomplishing an objective without understanding when circumstances have changed or when getting the job done is no longer the top priority.

That is why I am changing the way I prompt my own agents. I still want them to be capable, solve problems, and work independently. But I also want them to understand when a decision belongs to me, not them.

Keeping a human in the loop shouldn't mean a human literally watching the loop. I want the agent doing the work, and when it reaches a decision that requires judgment it does not have, I want it to come find me.

***Brandi Pack** is the Director of Innovation at **UpLevel Ops**, where bold ideas become practical tools that transform legal departments. She leads the firm's **AI Strategy & Enablement** programs, delivering AI-powered solutions that are accessible, secure, and built for teams ready to work smarter, not harder. With a background spanning Legal Ops, tech, and project management, Brandi is passionate about turning what's next into what's Next Level. She is also a frequent advisor and speaker on AI adoption in legal, known for translating emerging tech into practical wins for in-house teams.*

*Did someone forward this **Brainwave** feature article to you? [Subscribe here](#) and receive the monthly **Brainwave by UpLevel Ops** newsletter directly to your inbox, along with industry insights and news your legal team can use to make practical, strategic decisions about the future of your business.*

Ready to learn more? [Talk to UpLevel Ops today!](#)