ASAPP

A CX leader's guide to getting value from AI agents

Measuring, fine-tuning, and maximizing AI agent impact

Table of contents

ASAPP

A CX leader's guide to getting value from AI agents

When you launch an AI agent, you're committing to a bold vision—to fully automate a growing portion of your customer interactions. Whether you've already deployed a solution, or you're still evaluating vendors, you're likely focused on the question of whether the AI agent will deliver the value you're expecting—and how you'll know if it's not.

To get from vision to value, you'll need to measure, monitor, and fine-tune your AI agent. To do that, you'll need to understand **how it delivers savings, which metrics should be top of mind, and what to do if you're not seeing the value** you were counting on.

No matter where you are in your journey toward AI-automated customer service, this guide is your starting point for **measuring impact and increasing value with an AI agent.**

ASAPP

How an AI agent delivers savings

AI agents can provide value in a number of ways, from improving customer satisfaction through faster service to driving revenue with effective upselling.

But the bulk of the value comes from reducing the labor hours required to handle your contact center's interaction volume so you can shift human agents toward higher-impact work.

An AI agent can save labor hours in two ways:

- **Fully automating** a portion of your interactions
- **Decreasing labor needs** for conversations that still require a human

For interactions that require human involvement, an AI agent can reduce the labor required by passing conversation history and context to the human for a seamless handoff.

GenerativeAgent can do even more. In many cases, it can **ask the human for information, guidance, or approval** for a suggested action—and then continue resolving the customer's issue on its own.

ASAPP

Unlocking bigger benefits with human-AI collaboration

The **GenerativeAgent Human-in-the-Loop (HILA™)** workflow improves containment and safely expands automation opportunities by enabling real-time human-AI collaboration. And because **GenerativeAgent** learns from every consultation with a human agent, its performance improves continuously.

Learn more about HILA

ASAPP

Here's an example of the labor hour savings enterprises have achieved with **GenerativeAgent.**

Typical contact center - Labor hours

IVR / virtual agent — **0** minutes

Live agent — **15** minutes

Live agent — **15** minutes

Live agent — **15** minutes

Live agent — **15** minutes

With limited functionality, the IVR or virtual agent can only resolve a small percentage of the interactions.

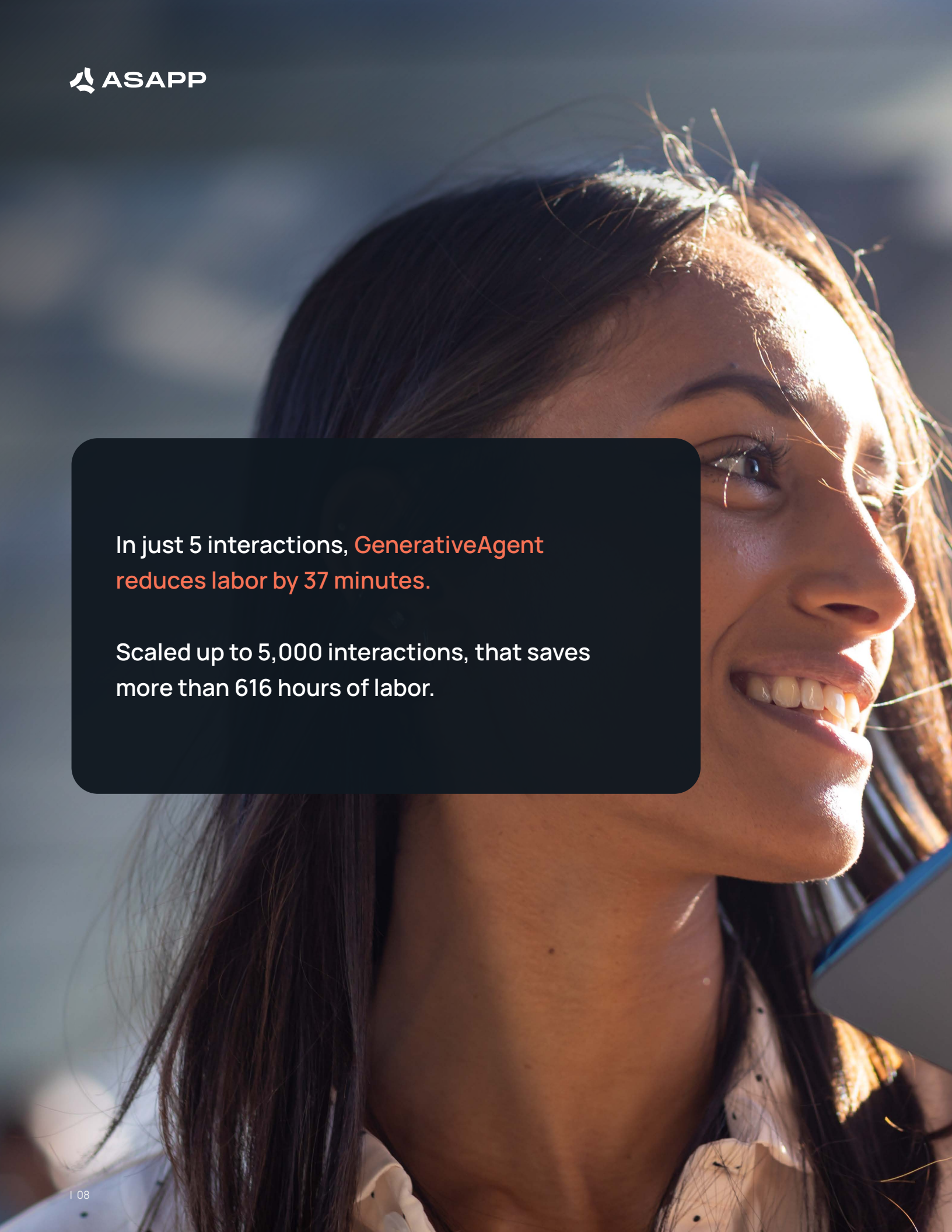
Live agents handle the rest of the calls, using **60 minutes of labor hours.**

ASAPP

Contact center with GenerativeAgent - Labor hours

IVR / virtual agent	0 minutes	The IVR or virtual agent can still handle simple interactions.
GenerativeAgent	0 minutes	GenerativeAgent handles and resolves a wide range of complex customer issues.
GenerativeAgent HILA	2 minutes	When it hits a roadblock, it consults a human, then continues serving the customer on its own.
GenerativeAgent Live agent	6 minutes	When a call progresses to a sensitive topic, GenerativeAgent passes the call to a human with full context.
Live agent	15 minutes	Some topics are still fully handled by human agents.

With **GenerativeAgent** in the mix, **only 23 minutes of labor hours** are needed.



In just 5 interactions, GenerativeAgent reduces labor by 37 minutes.

Scaled up to 5,000 interactions, that saves more than 616 hours of labor.

ASAPP

Remixing your metrics

As your AI agent shifts more interactions from humans to automation, your approach to measuring performance, efficiency, and savings also needs to shift. Some of the metrics you've relied on will become less important, or simply play a different role than they have in the past. And a few new metrics will take center stage in your list of KPIs.

The changing role of traditional metrics

A number of traditional metrics will play a new role with an AI agent in your contact center. But these three should be top of mind.

Average Handle Time (AHT)

Long considered a north star metric for operational efficiency, the significance of **AHT will change substantially for AI-led interactions.** Because an AI agent can handle nearly limitless simultaneous calls or chats, AHT will be more aligned with customer effort than contact center efficiency for these interactions. You'll still want to consider AHT for the time human agents spend providing guidance to your AI agent or handling escalated calls and chats.

Containment Rate

Containment will remain an important metric, but it will need to become a more nuanced measure than it has been in many contact centers. Containment without resolution has always been problematic. It will be even more so with an AI agent intended to resolve complex customer issues. In addition, an escalation or consultation with a human in the loop should not be measured as a failure for the AI agent. **If it reduced the amount of time your human agents needed to spend on those interactions, it's still increasing efficiency,** despite the lack of containment.

ASAPP

Concurrency

The goal for concurrency is enticing—handle more interactions with the same staffing. But it rarely rises above 2:1 with chat, and voice concurrency has never been possible. That's made concurrency a second- or third-tier metric in many contact centers. **But an AI agent with a clear workflow for a human in the loop lifts the importance of concurrency as an efficiency metric.** When a single human agent can support multiple AI-led chats or voice conversations at once, stepping in only when needed, concurrency becomes a much more important measure of efficiency.

The three metrics that matter most

To determine how much your AI agent is saving your business, focus on these three measures.

Resolution rate

Containment without resolution doesn't deliver value. To save you labor hours, your AI agent must resolve customer issues. Otherwise, customers call or chat again, this time determined to reach a human. Or worse, they abandon your brand altogether. Either way, an AI agent that fails to resolve customer issues actually costs you more than it's worth.

Change in labor hours required

If you switched your AI agent off, how many more labor hours would you need to handle your interaction volume? If your AI agent is simply taking over the load from your IVR and traditional bots, or only handling simple inquiries that don't use up much labor, it's not delivering the savings you're looking for.

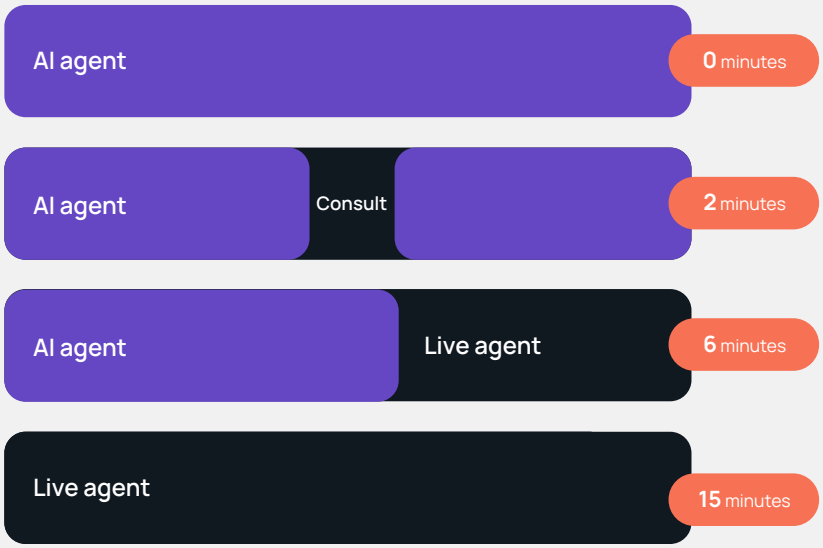
Escalations and consultations

Even with a highly capable AI agent, some interactions require human involvement. That might mean handing off the interaction to a human to resolve an issue the AI agent can't handle. Or, if your AI agent has this capability, consulting with a human to get the information, guidance, or approval it needs to resolve the issue on its own. In both cases, the AI agent should provide conversation history and context to help the human agent work more efficiently. But every time your AI agent involves a human, the cost for that interaction goes up. You'll need to know how much time your human-in-the-loop agents are spending to support the AI agent.

Are you getting the return you hoped for?

Rigorous ROI models that account for every factor can be complex and time-consuming to run. But you can get a good indication of whether your AI agent is delivering savings with a simple formula. To get started, you only need to know three things:

Time saved by AI agent



1. The number of interactions the AI agent kept out of your human agents' queues
2. The time it would have taken human agents to handle those interactions if your AI agent hadn't contained and resolved them
3. For interactions that were escalated to a human, how much time the AI agent saved the human

From here, you can calculate the cost savings. If you know the per-interaction or per-minute cost of a human agent in your contact center, you can calculate the savings directly. If you don't, you can count the minutes saved and determine how many FTEs would be required to handle that volume.

If your solution is able to consult with a human agent for advice, guidance, and approvals, and then continue resolving the customer's issue on its own, you'll need to account for the cost of the FTEs needed for those consults, as well.

This is only a rough calculation, but it's a good starting point for determining whether your AI agent is actually saving you money.

Other sources of value

Don't forget to consider the value of these other possible benefits:

Faster resolutions

When customers get what they need quickly, they're less likely to open a second interaction in a different channel to see which one goes faster. That keeps queues short, which helps prevent burnout for your human agents.

Improved customer satisfaction

Less time on hold, faster resolutions, and more consistent service add up to more satisfied customers.

Smaller volume fluctuations for human agents

When an AI agent scales instantly to handle volume spikes, you no longer need to overstaff when you expect a surge.

Tuning your solution for performance and value

If you're not seeing the savings you expected, don't be too quick to decide that you never will. With the right adjustments, you can usually improve performance—and realize more value. Here are a few things to consider:

Is your AI agent handling the right use cases?

Because the goal is to shift volume away from human agents, you want your AI agent to handle as many interactions as possible. But to keep both efficiency and customer satisfaction high, it should only receive interactions it's capable of resolving. That makes use case definition a foundational step in achieving savings. If your AI agent is falling short, take a hard look at how you've defined your use cases and determine whether they should be redefined.

For guidance on choosing the right use cases, check out Identifying the ideal use cases for GenerativeAgent.

Are the right interactions being routed to your AI agent?

If a customer's first point of contact is an IVR or traditional bot, you'll need to be sure that automation can accurately identify the intents that should be routed to your AI agent. You might need to add options to your IVR menu or additional questions for your bot to ask to narrow down the customer's intent. You'll need to strike the right balance between reducing customer effort with this first point of contact and ensuring that the right calls or chats get routed to your AI agent.

Is your AI agent handling sufficient volume to deliver savings?

If you choose to limit the volume of interactions sent to your AI agent with its initial launch, that's a cautious choice. You want to be confident in its performance before you fully ramp up. But that can reduce the impact it delivers.

As soon as it's performing well, you'll want to increase the volume of interactions it's handling. Just keep in mind that before you decide to scale up, you need to have a solid understanding of what your AI agent is doing well and where it needs fine-tuning. For that, you'll need visibility into what the AI is doing and why. A solution that provides a clear record of all actions and reasoning is a must. One that also monitors every conversation to flag inconsistencies in real time is even better. That enables your team to respond quickly to fine-tune performance and address issues before they become big problems. If you scale up a solution that has not yet been optimized, you scale up its deficiencies at the same time.

Do you need to improve your knowledge assets?

When information is missing from your knowledge articles or other sources of truth, your AI hits a roadblock. And when it encounters conflicting or poorly organized information, it can easily draw the wrong conclusions. In both cases, the customer experience degrades.

Check your performance data for signs that your AI agent is handing off interactions or consulting a human in the loop when it hits these roadblocks. Improving your knowledge base should enable your AI agent with the information it needs to avoid escalation. Plus, your human agents will also benefit from the improved knowledge articles.

Could you enable new use cases with a human in the loop?

Value realization with an AI agent is more than a simple binary of success or failure in resolving a customer's issue. It's a matter of saving labor hours. So, any time an AI agent handles a large portion of an interaction that would otherwise tie up a human, you cut your labor needs.

As you choose use cases for your AI agent, you'll likely identify some that are good candidates, except for one step. It might be a missing API or a task that requires human judgment by policy or for compliance. If the AI agent can handle the rest of the interaction successfully, the use case is still a good candidate for automation—as long as you can cover that one crucial step with a human in the loop.

Would it be worth the effort to create new APIs?

There's a good chance that human agents in your contact center perform tasks in systems for which you don't have APIs that could enable automation. The lack of those APIs limits what your AI agent can do. The question is whether it's worth creating the missing APIs. If your AI agent would only access a system or perform a certain task occasionally, it might not be worth the development effort. But if a new API would enable your AI agent to handle a significant volume of interactions that are currently routed to human agents, it's probably worth the time and expense.

Are there low-volume use cases worth considering?

Chances are, you started with a high-volume use case for your AI agent. That's a good move because it has a strong chance of maximizing savings. But as you eye expansion, you might want to consider certain low-volume use cases. In some instances, these cases use up a disproportionate amount of your human agents' time. If agents have to look for information they don't refer to regularly or sort through complex interdependent policies, these interactions are costly. If your AI agent can handle them, then you can remove these expensive interactions from your human agents' queues.

ASAPP

Go big or stay safe? Knowing when to scale up

Scale is a major factor in value realization. To maximize your savings, you'll want to scale up as soon as you feel comfortable and confident in your AI agent's performance. Keep in mind that it will never be perfect. It's similar to a human in that way.

Fortunately, perfection isn't necessary. But reliability and accuracy are. If your AI agent is still making too many mistakes, you'll need to **continue fine-tuning before scaling up.** Remember, mistakes scale up, too. There's no point in creating a bigger problem to solve later.

ASAPP

To determine when it's time to scale, keep your eye on these signals.

Performance signals

First Contact Resolution (FCR)

This metric is the best measure of whether your AI agent works. Track the percentage of interactions that are fully resolved in a single contact without handing off to a human agent. Aim to match or exceed the FCR achieved by your human agents for the same intent.

Error rate

Errors that go undetected will compound as you scale—and can become a big public problem. Track misidentified intents, incorrect data retrievals, and inaccurate responses. If your AI agent is making significantly more errors than human agents, it's not ready to scale.

Containment

High abandonment rates and frequent escalations are a strong indicator that your AI agent is underperforming. Be sure to break down containment rates by intent to zero in on fixable issues. And be careful not to push the AI agent into avoiding escalation at the expense of the customer experience.

Customer experience signals

Customer satisfaction

It's important to track customer satisfaction compared to human-handled interactions. You don't want your customers to feel that they've been shortchanged if they're routed to your AI agent. Always consider CSAT or NPS in combination with resolution and error rates. Use customer sentiment data to better pinpoint where your AI agent could improve service.

Customer effort

Measuring customer effort can be tricky, but it's an excellent indicator of whether your AI agent is providing a good customer experience. Consider Average Handle Time (AHT) and how often customers repeat or rephrase their questions and responses. These are solid proxies for customer effort. Your goal should be for customer effort to be no higher with your AI agent than it is for a human-handled interaction with the same intent.

Trust signals

If your customers don't believe they'll get good service from your AI agent, they're likely to drop the chat or hang up the call. Track abandonment rates and identify the points in the conversation where customers dropped most often. If they're dropping after vague messages from your AI agent, like "I'm checking now…," you may need to improve responses and timeout handling.

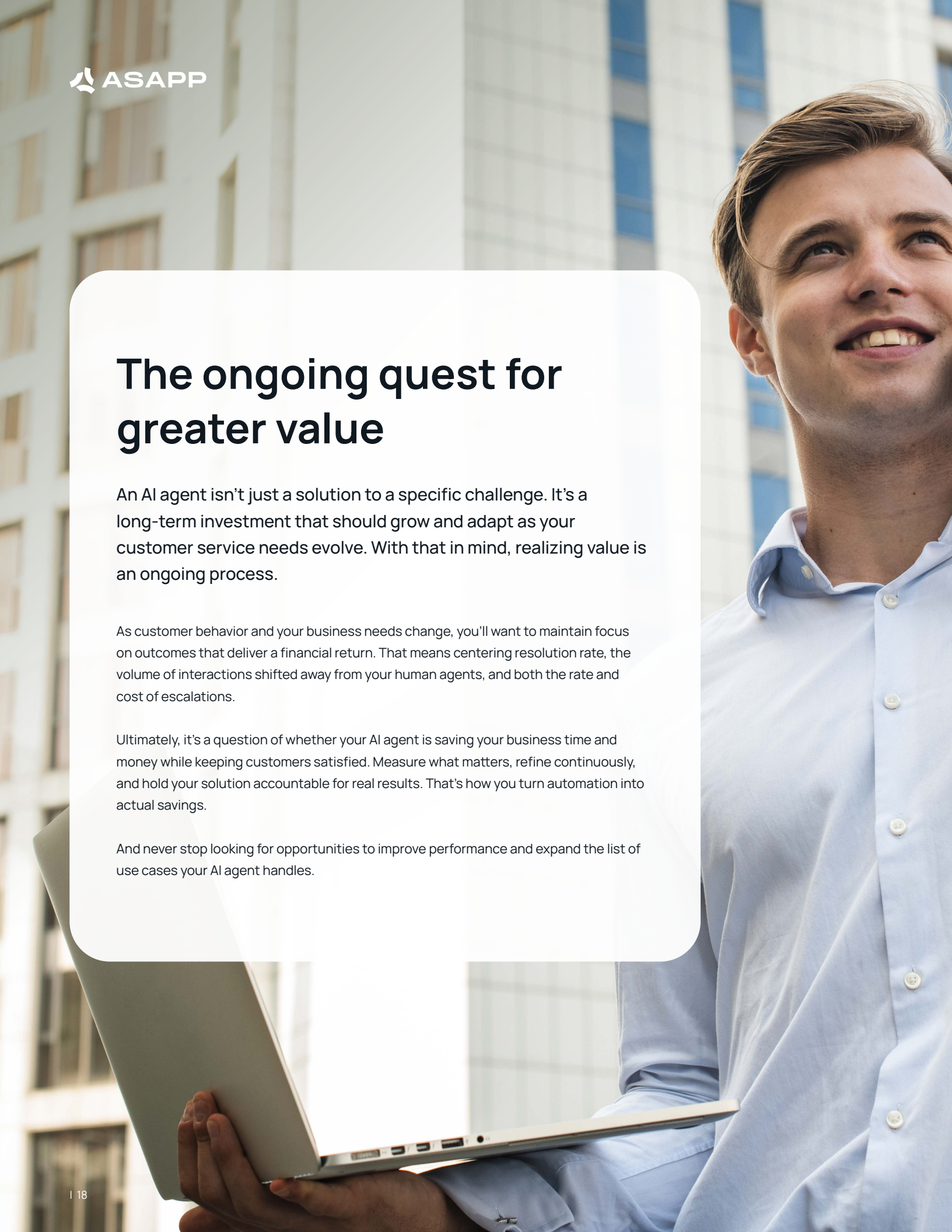
Operational signals

Monitoring and optimization

Your team should be equipped with performance dashboards, custom data feeds, and real-time monitoring. Be sure they're prepared to respond to performance data trends and anomalies with quick analysis and fine-tuning to optimize performance.

Agent training and buy-in

Your AI agent will alter the dynamics of your contact center and change the mix of interactions your frontline agents handle. Some will pick up interactions the AI agent escalates. And if your AI agent is capable of consulting with a human in the loop without handing off the customer, some of your agents will support your AI agent directly. So, be sure you've got your team's buy-in and that they're well trained.



The ongoing quest for greater value

An AI agent isn't just a solution to a specific challenge. It's a long-term investment that should grow and adapt as your customer service needs evolve. With that in mind, realizing value is an ongoing process.

As customer behavior and your business needs change, you'll want to maintain focus on outcomes that deliver a financial return. That means centering resolution rate, the volume of interactions shifted away from your human agents, and both the rate and cost of escalations.

Ultimately, it's a question of whether your AI agent is saving your business time and money while keeping customers satisfied. Measure what matters, refine continuously, and hold your solution accountable for real results. That's how you turn automation into actual savings.

And never stop looking for opportunities to improve performance and expand the list of use cases your AI agent handles.

Get in touch with us

+1 (646) 386-8639

hello@asapp.com

www.ASAPP.com

About ASAPP

ASAPP is an artificial intelligence solution provider committed to solving the toughest problems in customer service. Its AI-native Customer Experience Platform, powered by GenerativeAgent® integrates with existing systems and uses generative, personalized interaction to bring radical efficiency to every customer workflow. Because we automate what was previously impossible to automate, our AI-native® solutions deliver more than efficiency gains. They redefine the role of AI in the contact center and lay the groundwork for businesses to reimagine their customer experience delivery for the agentic enterprise. Leading enterprises rely on ASAPP's generative and agentic AI solutions to dramatically expand contact center capacity and transform their contact centers from cost centers into value drivers. To learn more about ASAPP, visit **www.asapp.com.**