

Practical guide for planning your GenerativeAgent® deployment





ASAPP

Contents

Securing stakeholder buy-in	04
Key criteria for choosing GenerativeAgent	06
mportant considerations with generative Al solutions	09
Your resource requirements for deploying and managing GenerativeAgent	11
dentifying the ideal use cases for GenerativeAgent	15



Al agents are a novel technology, so there's not a lot of conventional wisdom on best practices for deploying them. That means most organizations that are planning to add Al agents to their customer service operations are starting at square one and going through the same discovery process that you are. During this process, you've probably researched multiple solutions from different providers and might even be considering whether to build your own Al agent.

At ASAPP, we've guided a number of enterprises through this process already. We've seen firsthand what works. That's why we've written this guide – to help you structure how you think about an AI agent initiative and lay the groundwork for success.

As you consider moving forward with GenerativeAgent, we want to help you ensure a smooth and successful deployment with the right planning up front. That planning should begin now, well before the contract is signed. This guide offers a starting point for that planning to align your team, define your goals, and lay the groundwork to realize value quickly.





Securing stakeholder buy-in

Deploying customer-facing Al agents demands broad organizational visibility, but the specific list of who needs to be involved in this process varies from one company to another. You'll need to devote time and attention to determining who your core group of stakeholders should be. Support from those stakeholders can make or break your GenerativeAgent deployment, so it's crucial to be proactive now to secure the buy-in you'll need. To minimize the risk of hitting roadblocks later in the process, you'll need more than general support for the idea of an Al agent. You'll need committed buy-in that aligns with your stakeholders' broader goals and objectives for the business.

To get that buy-in, it can be helpful to look for shared goals across the full range of your stakeholders. What impact are they hoping GenerativeAgent will make? What are they looking for from ASAPP? It can take time to find alignment on these questions, so it's useful to start that process as early as possible. We've found that clarifying your needs and goals in terms of the value they will deliver for your business is an effective framework for securing the buy-in you need.

Customer service stakeholders

For stakeholders in your customer service organization, the following goals are typically compelling.

Lowering the cost to serve

Because GenerativeAgent can successfully resolve up to 90% of customer interactions without escalation, it increases containment and FCR, which lowers your costs. And when volume spikes, it scales automatically, which eases pressure for additional staffing.

Increasing efficiency and productivity

While GenerativeAgent handles routine customer inquiries, your human agents can focus more effectively on more complex issues and higher-value conversations. That leads to an increase in contact center throughput.

Improving the customer experience

With GenerativeAgent, your customers have an open line to call or chat whenever they need to connect. And because GenerativeAgent scales to meet demand, they get immediate service instead of waiting on hold. Your customers get quicker resolutions, which improves customer satisfaction.



Stakeholders from other teams

For stakeholders from functional areas other than customer service, we often find that the following issues and goals are top of mind.

IT teams play a critical role in ensuring the security and availability of your technology infrastructure, so they will need to be convinced that GenerativeAgent is reliable and easy to integrate with your tech stack. They'll also want to understand their level of ownership of the solution and how readily they will be able to handle maintenance and technical issues themselves rather than relying on a vendor.

Legal teams and data privacy teams will need to be reassured that personal identifiable information (PII) will not be shared with or accessed by anyone else. They may also ask about safety mechanisms that will prevent brand damaging events.

Operations teams will need to understand the impact that GenerativeAgent will have on workforce planning and demand forecasting. They will also find information about the potential need for resources for ongoing quality assurance, optimization, and periodic updates to be helpful.

Procurement and finance teams will naturally want the details about pricing options and ROI models.

Crystallizing Your Needs

Getting initial buy-in based on overarching goals is just the start. As you move forward, you'll need to clarify expectations for scope, pace, and value. You'll also need to detail how you'll use GenerativeAgent to deliver on your strategic goals.

For example, if your stakeholders are aligned around cost reduction, you'll need to quantify how much cost reduction you're aiming for and how quickly you need to realize that value. And you'll need to balance that goal with safety and logistics.

ASAPP provides a flexible path forward with GenerativeAgent that allows you to set your own pace for deployment and expansion. And we'll work with you to chart the right course to meet your business objectives. We welcome the opportunity to discuss your needs, address your concerns, and clearly define your plan for GenerativeAgent through a collaborative process with you and your stakeholders.



Key criteria for choosing GenerativeAgent

Every vendor who sells an AI agent for contact centers makes similar claims about what you can achieve with their product – smarter automation, increased productivity, and satisfied customers. That vague language makes a fair comparison of solutions more difficult than it ought to be.

As you evaluate and compare solutions, it's important to take control of the conversation by asking questions that will help you spot the key differences between solutions and vendors. The answers to these questions will help you make the right choice for your business.

What exactly does Generative Agent do?

GenerativeAgent is a fully automated AI agent that can work independently to handle interactions and resolve customer issues through voice and chat. This level of automation is achieved through the application of agentic orchestration and the full range of capabilities provided by the latest generative AI technologies. This is not an iteration over previous bot technology, but a rethinking of the application of AI for customer services, based on:

- Accurately interpreting customer inquiries with generative Al.
- Managing complex, nuanced interactions autonomously.
- Taking action directly through backend systems to resolve issues.
- Identifying when and how to escalate to the right human agent.
- Ensuring smooth handoffs to agents with context.
- Improving customer satisfaction while reducing operational costs.
- Above all, delivering a great, consistent customer experience at scale.

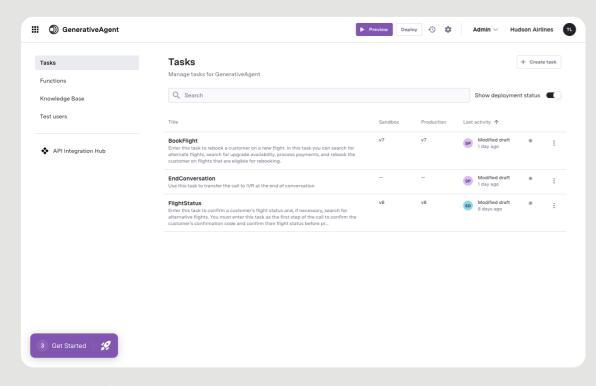
Is there more to Generative Agent than a LLM + RAG?

Yes, in addition to being grounded in your sources of truth, GenerativeAgent orchestrates multiple models, employs input and output evaluators, includes mechanisms to mitigate hallucinations, and has controls to ensure the Al remains on task and within scope. This robust system of guardrails enhances safety and ensures that outputs generated by generative Al are accurate, secure, and traceable back to their source of truth. Automated quality checks and human oversight further enhance reliability.

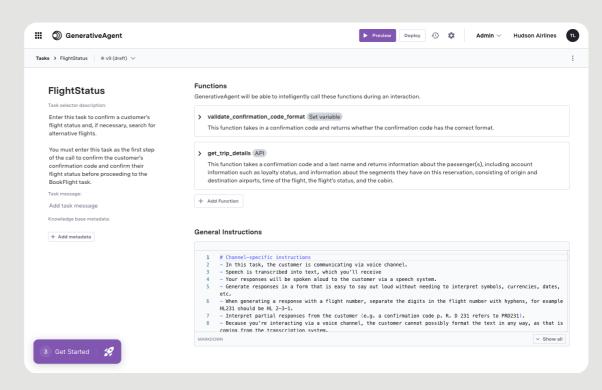
How hard is GenerativeAgent to use and maintain?

GenerativeAgent is purpose-built for contact centers and is architected to integrate easily into your CX tech stack, drive adoption with an intuitive agent experience, and equip business users with nocode tooling for monitoring and optimization. ASAPP also provides automated quality evaluation on GenerativeAgent conversations, and businesses that have deployed GenerativeAgent have assigned a quality assurance analyst to review those evaluations and optimize the performance of the system. As a result, you realize value faster.





Through GenerativeAgent's AI Console, you can easily configure GenerativeAgent, test its capabilities using the Previewer, block external IP addresses for security purposes, create and manage custom regex redaction rules, and view audit logs.



You can customize the tasks that you want GenerativeAgent to handle by providing a set of instructions in human language and adding Functions, which are the tools GenerativeAgent can use for that task.



How will we know what GenerativeAgent is doing - and why?

GenerativeAgent documents the systems it accesses to retrieve data and information and perform tasks to help the customer in a readily available audit trail. It also creates a record of its reasoning for each action it takes that is retained for later analysis. Aggregate metrics and reporting allow your data teams to understand what's happening at scale.

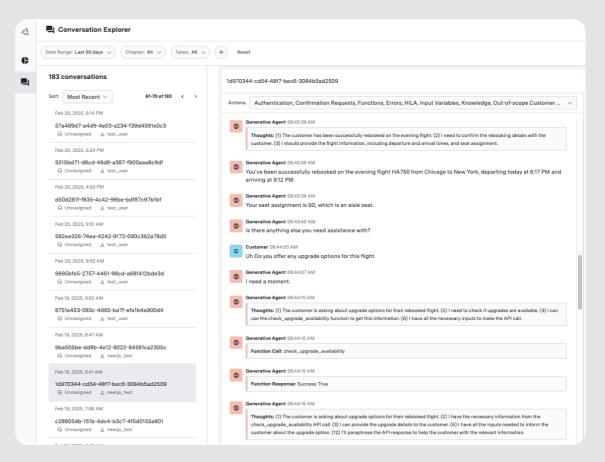
How does Generative Agent keep a human in the loop?

Many Al agent solutions can escalate a customer interaction to a human agent when they get stuck. But GenerativeAgent can also ask the human in the loop for help instead of escalating the interaction. Once it receives the help it needs, it will continue working with the customer to resolve their issue. The bottom line is that GenerativeAgent knows when and how to ask a human agent for help. This advanced capability creates new possibilities for how to incorporate GenerativeAgent into your customer service operation.

Enabling these human-in-the-loop workflows does involve some operational changes, but it unlocks a new paradigm for customer service that can dramatically increase the capacity of your contact center. It's not necessary to take on these additional changes with your initial deployment of GenerativeAgent. You can start with a workflow in which GenerativeAgent simply escalates what it cannot handle on its own to a human agent. When you're ready for the more advanced capabilities of the human-in-the-loop workflow, you can adopt them gradually.

Learn more about Preparing your contact center for the human in the loop.

For more detailed guidance, see Evaluating generative AI agents for your contact center.



Through the Conversation Explorer, you can review GenerativeAgent's interactions with the customer and understand the reasoning behind each decision made. In this case, after successfully rebooking the customer after a cancelled flight, GenerativeAgent looks for an upgrade option.



Important considerations with generative AI solutions

Security and safety

ASAPP prioritizes safety, security, and privacy with all of our solutions. We have taken a safety by design approach, rather than simply trying to address safety with mechanisms tacked on after development. As a result, GenerativeAgent comes with best-in-class security and safety, making it enterprise-ready.

We utilize encryption, role-based access control, multi-factor authentication, and continuous monitoring with advanced intrusion detection systems and Security Information and Event Management (SIEM). To promptly address security gaps, we also conduct regular vulnerability management and third-party penetration tests by reputable organizations.

Customer data is managed with the highest standards of security and privacy. We use sensitive redaction techniques to redact PII before data is stored. Customer data also remains within its original geographic region, adhering to data sovereignty, protected by robust encryption and stringent access controls.

ASAPP's orchestration enhances safety and implements guardrails to protect against adversarial attacks and ensure that outputs generated by generative AI are accurate, secure, and traceable back to their source of truth. Robust observability into chain-of-thought and model actions explain the product's behavior. Automated quality checks and human oversight enhance reliability.

For a more comprehensive review of our approach to safety, security, and privacy, see the GenerativeAgent security and safety FAQ.





Legal considerations

When acquiring generative AI solutions, it's crucial to involve your legal team early, even if that's not your typical process with other types of technology solution acquisitions. The open-ended, probabilistic nature of generative AI, combined with its use of data, creates new considerations your legal team will want to review, understand, and approve. The earlier you include them, the better.

Use of data for model training

The need for data to train Al models creates some new considerations for security and privacy. Your legal team will need to understand how your company's data will be used for this purpose. ASAPP will use your data solely to provide services to your business. We will not use input data or outputs from GenerativeAgent to train or fine-tune our Al models for the benefit of any other company.

Solution explainability

Explainability is crucial with generative AI solutions. ASAPP retains all relevant AI models and information about those models that could be used to explain the decisions made and actions taken by GenerativeAgent.

LLM sub-processors

The use of LLM sub-processors in AI solutions is common. ASAPP maintains an enterprise agreement that contractually prohibits our sub-processors from using your data to train or improve their models for their own purposes. In addition, these agreements specify that none of your data will be retained by the sub-processor.

Ability to turn off the solution

To ensure that you have sufficient control over GenerativeAgent, we have included a so-called kill switch. This feature allows you to completely disable all GenerativeAgent functions and is intended as an emergency or failsafe mechanism.

Procurement & Finance

The procurement process for a generative Al solution can be lengthier and more complex than the standard process for procuring other technology solutions and infrastructure. The need for additional internal reviews for security, privacy, safety, and compliance can increase the time required to finalize the deal. With that in mind, you should plan to engage your procurement and finance teams earlier than you typically would.

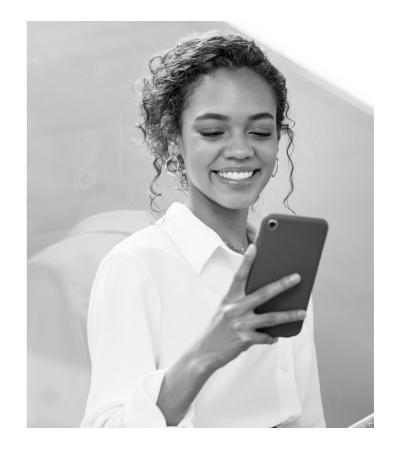


Your resource requirements for deploying and managing GenerativeAgent

The time commitment required from your team to deploy, manage, and expand your use of GenerativeAgent depends on a number of factors, such as your current technology infrastructure, the availability of APIs, and the use cases you identify. **GenerativeAgent is optimized for the CX environment with native integrations and no-code/low-code configuration, which streamlines the process of launching and using it.** If you choose to work with a deployment partner, that will further reduce the burden on your team. We will work with you to more clearly define what your deployment will look like and help you refine the resource estimates as we go.

The resource estimates listed here are based on our experience deploying GenerativeAgent in enterprise contact centers. The resources you'll need to commit will change over time from initial deployment to ongoing use, so we have grouped them here into three categories:

- One-time integration into your infrastructure to enable GenerativeAgent
- Launching each new use case
- Ongoing use and optimization





Resources required for a one-time integration

For the one-time integration of GenerativeAgent into your ecosystem, you will need these resources:

- · Procurement: We will work with your team to follow your standard procurement process.
- Legal: We will engage closely with your legal team to help them understand our controls on LLMs and generative AI.
- Data security and privacy: We will also engage closely with your security teams to describe how
 we handle data securely.
- Networking: We will need a small amount of involvement from your networking team to allow
 access to your APIs if they are not public-facing already.
- Technical owner of your contact center platform: We will need the technical owner of your platform (AWS Connect, Genesys, Alvaria, etc.) to install GenerativeAgent.
- Identity team: Only needed if you want to use SSO.

Time commitment

- Procurement: The time commitment should be typical for your standard procurement process.
- Legal: A few days, depending on how familiar they are with generative AI engagements.
- Data security and privacy: A few days to complete reviews of our platform.
- Networking: No time needed if your APIs are public already. About one week of effort if APIs are not already public to establish network access.
- Technical owner of your contact center platform: 1-2 days if you have Genesys Cloud or AWS
 Connect. If you are building a custom integration, likely a couple of weeks.
- Identity team: About one week if you want to use SSO.



Resources needed to launch a new use case

We use a collaborative, data-driven process to help you identify which use cases will help you realize the greatest value in the shortest amount of time. And as you plan to expand your GenerativeAgent deployment, we'll continue working with you to choose the best use cases to add next. When launching a new use case with GenerativeAgent, we recommend a team that includes the following resources and skillsets:

- Deep subject matter knowledge about the business to understand the workflows and processes that GenerativeAgent should follow.
- Clear written communication to provide unambiguous instructions to GenerativeAgent (this may come from the same person as #1). No machine learning expertise is necessary.
- Quality assurance analysts to test the system before deployment and to read and verify conversations where mistakes or issues have been identified after deployment.
- IT teams owning the APIs that GenerativeAgent will use, especially before launching each use case, with light-touch support required once the use case is in production for maintenance.
- Content/knowledge base owners to add, review, or edit content that GenerativeAgent uses as
 opportunities for improvement are identified.

Time commitment

The time commitment depends on the complexity of the use case, how well the edge cases are understood, and how well the API is understood. A good ballpark estimate is about 3-4 weeks for a dedicated team with the skills listed here for the first use cases, and around 2 weeks for subsequent use cases for which the use case is not well understood or the required APIs do not exist. **Keep in mind that a human-in-the-loop agent can access systems and provide necessary information to GenerativeAgent when it lacks the APIs it needs to take those steps itself.** This additional human involvement introduces some new operational complexities, but it can enable you to launch use cases more quickly.

These resources are still necessary even if you are not handling the configuration of GenerativeAgent yourself because a partner or ASAPP is doing that work. However, the time commitment is lower — likely 25-50% of what it would be if you were doing the configuration work.

The time commitment for content/knowledge base owners depends very heavily on the health of your knowledge base. If it's in good shape, it's possible that you could launch GenerativeAgent with little to no updates to it. On the other hand, we have worked with businesses that devoted more time to building out their knowledge base up front. We've also seen some that skipped this step and later regretted that choice. They backed up, cleaned up the knowledge base, and found the success they'd been aiming for.



Resources to manage an ongoing deployment

The resources needed to manage an ongoing deployment are similar to the team you'll need to launch a new use case. The only difference is that it requires less time commitment from all resources, particularly IT, but we do encourage a data analyst/data science resource at this stage for continuous optimization. Here's what you'll need on your team:

- Deep subject matter knowledge about the business to understand the workflows and processes that GenerativeAgent should follow.
- Clear written communication to provide unambiguous instructions to GenerativeAgent (this
 may come from the same person as #1). No machine learning expertise is necessary.
- Quality assurance analysts to read and verify conversations where mistakes or issues have been identified.
- IT teams owning the APIs that GenerativeAgent will use, especially before launching each use
 case, with light-touch support required once the use case is in production for maintenance.
 Once it's in maintenance mode, IT support is only necessary when the APIs change and
 GenerativeAgent configuration may need to be updated, but this should be infrequent.
- Content/knowledge base owners, to add, review, or edit content that GenerativeAgent uses as opportunities for improvement are identified.
- Data analyst (optional but recommended) to dig into the data and understand opportunities for improvement.

Time commitment

The number of hours your team will need to devote to managing an ongoing deployment will depend on how many use cases are live and how much volume is going through the platform. If you are using GenerativeAgent for all use cases, you will likely need a full-time team with the skills listed here to manage and continuously improve it. If you're using GenerativeAgent for a few use cases, particularly during pilots, you'll need:

- ~2 quality assurance people at 50% of their time to review conversations. This varies widely
 depending on what percentage of coverage you want to have and how much volume is flowing
 through the platform.
- ~1 SME/written communicator at 30% of their time to make updates to GenerativeAgent's configuration and test those changes.
- ~0 IT resources, unless changes or maintenance to APIs are required.
- ~1 content owner/knowledge base manager. This depends very heavily on how healthy your existing knowledge base is.
- ~1 data analyst, with the commitment dependent on how much you want to optimize GenerativeAgent.



Identifying the ideal use cases for GenerativeAgent

Part of ensuring a smooth deployment and quick value realization with GenerativeAgent is choosing the best use cases to target first. This step is critical, even if you are simply piloting GenerativeAgent to determine whether it's a good fit for your contact center. For example, if the use case you choose is not well suited for GenerativeAgent, it will be difficult to evaluate its performance.

Identifying the right use cases requires a thorough examination of the interactions your contact center handles to understand why customers get in touch and the steps required to resolve their issues. That might sound like a daunting task that involves a custom consulting engagement, but we've developed tools to automatically surface the information you'll need to identify your best automation opportunities.

ASAPP uses a reliable, data-driven process to analyze your contact center's interactions and identify good candidates for automation. For each intent, we determine the volume, containment and resolution rate, and establish what steps are required to resolve the customer's issue.

With the results of this analysis, you're well-equipped to make informed decisions about which use cases you'd like to target first with GenerativeAgent. This is a collaborative process in which we provide analysis and expert guidance, and you make the decisions at every step.

Here's how the process works.

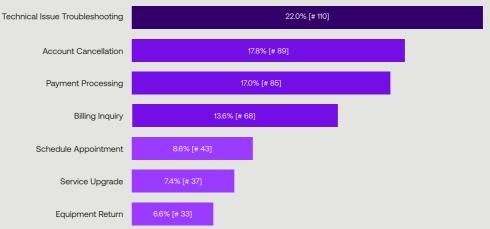
ASAPP analyzes *your* customer interactions to identify use cases with high automation potential

We start with a collection of your actual customer interactions – at least 1,000 is ideal, but 500 will suffice. Using large language models, we analyze the conversations to assess the potential for automation with Generative Agent. To do that, we'll answer a series of guestions.

Which intents drive the highest volume?

First, we identify the customer's intent in each interaction and determine which intents are generating the highest call volume. To realize quick returns on your investment in GenerativeAgent, you'll want to target high-volume use cases first. You can choose to run an a/b test in which GenerativeAgent initially handles only a small percentage of the interactions for the identified intent. This allows you to gauge performance before ramping up completely.

But volume is just one consideration. You'll need to know more about the high-volume interactions before you can pinpoint the optimal use case to deploy first.





Which intents were resolved by agents without escalation?

For each intent, we'll determine what percentage of interactions successfully resolved the customer's issue without a transfer or having the customer abandon the interaction. Interactions that your agents were able to resolve without escalation are good candidates for GenerativeAgent to resolve, too.

While GenerativeAgent can handle escalation to a human agent seamlessly, you'll realize greater value in your first use cases by focusing on interactions that can be fully contained and resolved without human intervention.

What tools did the agents use to resolve the issue?

Once we've identified which intents are good candidates for automation based on volume, containment, and resolution, we infer which tools the agents used to resolve the customer's issues.

To make this determination, we rely on the responses the agents provided to the customer at each step in the interaction. If a response contains general information about products, services, or policies, we conclude that the agent relied on the knowledge base for that information. If a response includes customer-specific information, we conclude that the response required the agent to access another system, such as your CRM. Some responses require information from both the knowledge base and another system.

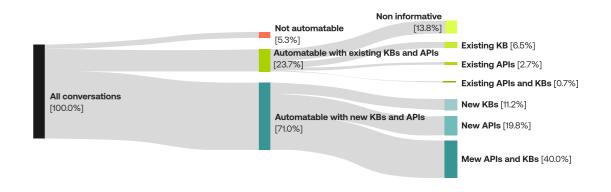
Understanding the steps the agents took to resolve the customer's issues is important because GenerativeAgent will need to follow the same steps.

Which tools are already available for GenerativeAgent to use?

For each task GenerativeAgent must perform, it needs the right tools. For some tasks, it will need accurate knowledge articles on relevant topics. For others, it will need APIs to access other systems and retrieve the necessary information.

Our team will generate a list of APIs that will be needed based on each task the agent performs. For example, if a customer asks to cancel their account, the agent might need to perform multiple tasks, such as account cancellation, service cancellation, and arranging for equipment return. An API would be required for each of these tasks.

At this stage, the ASAPP team will work with you to determine which knowledge articles and APIs you already have and which ones you will need to create. Naturally, interactions that rely on tools that already exist can be deployed more quickly and with less effort on your part. These are likely your best bet for first use cases.





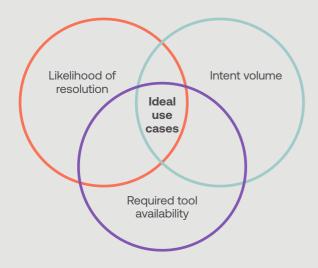
If you need to create tools, which ones should be highest priority?

If you do need to build APIs, we can help you prioritize them based on the impact they will have on GenerativeAgent's ability to resolve the customer's issue successfully. Some APIs will be required for many of the interactions of this type, while others will be needed only occasionally. With that information, you can invest your time and resources in building the APIs that will yield the biggest returns.

We'll follow the same collaborative process to determine which knowledge articles already exist and which ones you'll need to create. And of course, we'll provide the usage data to help you prioritize them. GenerativeAgent can ingest content from other sources, which could reduce the burden of creating new knowledge articles. For example, if some of the necessary content exists on your website, we can scrape that information. We will work with you to determine the most efficient and effective option.

You choose which use cases to target first

With the results of our detailed analysis, you'll have all the information you need to choose which use cases to automate first with GenerativeAgent. And you'll be equipped to plan how to expand to additional use cases over time.



Planning for use cases with a human in the loop

Some interactions will require a human in the loop, at least in the short term – for example, to provide approval of a refund, account modification or temporary discount, or to offer the GenerativeAgent direction for troubleshooting an issue beyond the basics. As we work with you to prioritize use cases for future automation with GenerativeAgent, we will help you identify good candidates for these conversations that involve a human agent as a decision-maker without requiring the agent to chat or speak directly with the customer.

It is typically best to start with use cases that GenerativeAgent can fully contain and resolve. But it's also useful to start planning for human-in-the-loop interactions early in your deployment process so you can continue increasing the returns on your investment over time.



Next steps in your GenerativeAgent deployment plan

Fully independent AI agents are new to customer service. That can make it difficult to plan your deployment. But ASAPP has already guided enterprise contact centers through the entire process from planning to deployment, optimization, and value realization.

We're happy to share what we've learned along the way. So, if you decide that GenerativeAgent is right for your business, we'll meet you where you are in your journey and help you develop a plan that will get your internal teams aligned, clarify your customer service goals, and speed you on your way to success.

And if you have questions, just ask. Our team loves to share our knowledge and expertise.

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. Because we automate what was previously impossible to automate, our Al-native® solutions deliver more than efficiency gains. They redefine the role of Al in the contact center and lay the groundwork for businesses to reimagine their customer experience delivery for the age of Al. Leading enterprises rely on ASAPP's generative and agentic Al 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.