



Preparing your contact center for the human in the loop



ASAPP Contents

Human in the loop with GenerativeAgent	04
Options for how to use the human in the loop	06
What does the human in the loop do?	07
Equipping the human in the loop	08
How the human in the loop differs from a typical contact center agent	08
Identifying the ideal human-in-the-loop candidates	09
Important skills for all human-in-the-loop personas	12
Staffing and scaling the human in the loop	13
Setting up queues for the humans in the loop	14
Evaluating the performance of the human in the loop	16
The evolving human-AI partnership in your contact center	17

Many AI agent solutions can escalate a customer interaction to a human agent when they get stuck. But GenerativeAgent can often avoid that escalation by asking a human for the help it needs to continue resolving the customer's issue on its own. It's smart enough to know when it needs help, and it knows how to ask for what it needs. Through real-time collaboration with a human in the loop, GenerativeAgent can continue serving the customer without handing off the interaction.

This advanced capability creates new possibilities for how to incorporate GenerativeAgent into your customer service workflows. By creating the right foundation for the human in the loop, you can expand automation opportunities safely, improve the performance of GenerativeAgent through self-learning and continuous optimization, and increase your contact center capacity without shortchanging the customer experience.

This guide provides an overview of the key issues you'll need to consider to get the greatest value out of the relationship between GenerativeAgent and the human in the loop.



Human in the loop with GenerativeAgent

GenerativeAgent enables the human-in-the-loop agent with user-friendly tools in an intuitive interface for an efficient and streamlined workflow. Let's walk through an example.

During the course of a conversation, the customer asks for a payment extension. In this case, GenerativeAgent is prohibited by company policy from making the decision to grant the extension on its own. It needs to request authorization from a human agent.

At this point, GenerativeAgent opens a ticket in the existing contact center platform for a human-in-the-loop agent and asks whether to grant the requested extension.

Human-in-the-loop agents interact with GenerativeAgent through an interface within their existing agent desk.

Through this interface, they can:

- Respond to GenerativeAgent
- Transfer conversations to live agents
- View the interaction thread history
- Access relevant customer information and summarized conversation context

Customer <> GenerativeAgent

GenerativeAgent 11:04:22 AM

Good Afternoon! Thank you for contacting Hudson, how can I help you today?

C Customer 11:04:30 AM

Hello! I just received my last bill, and unfortunately i'm not sure i'll be able to pay it on time

GenerativeAgent 11:05:40 AM

Ok, I can help you with that

I'm pulling up your current bill information

Your current bill is \$150.00, and it is due on March 25.

If you're unable to pay on time, I can help you explore other options.

C Customer 11:04:50 AM

ok, what are my options?

Can I receive an extension?

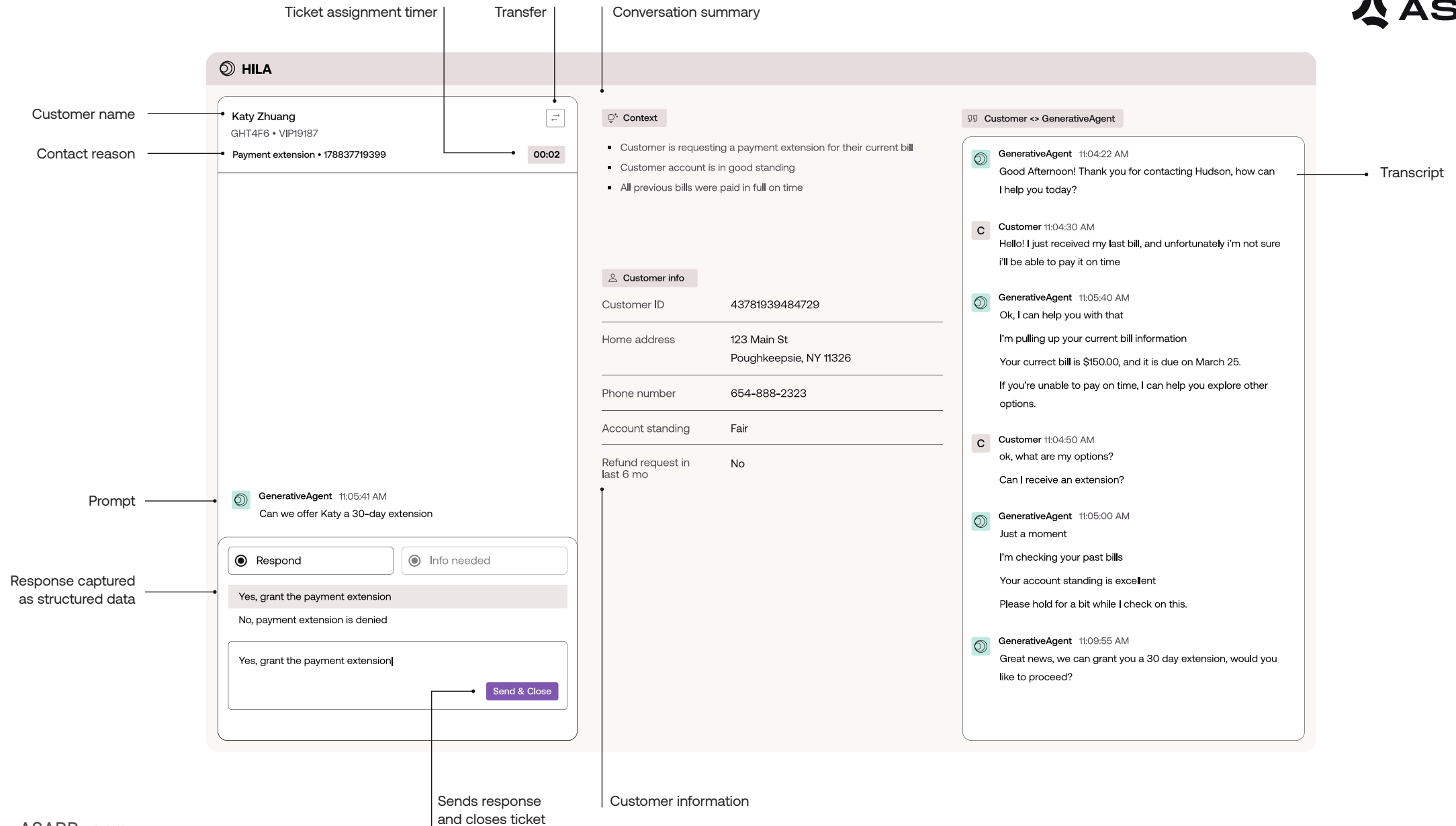
GenerativeAgent 11:05:00 AM

Just a moment

I'm checking your past bills

Your account standing is excellent

Please hold for a bit while I check on this.




The human-in-the-loop agent can review the information the customer has provided and grant or deny the extension with a simple click. The agent can also add free text to explain their response to GenerativeAgent if additional information would be helpful.

After submitting the response, the agent can close the ticket. GenerativeAgent uses the response from the human-in-the-loop agent to continue the customer conversation. In this case, the customer's request for an extension is granted, and the interaction is resolved without a hard escalation.

The human-in-the-loop interface includes a range of helpful fields, including:

- **Transfer:** Allows the agent to transfer the conversation from the GenerativeAgent to a live agent.
- **Ticket assignment timer:** Tracks the time elapsed since the ticket was assigned to the agent.
- **Prompt:** Indicates the specific assistance the GenerativeAgent needs to unblock the customer. This is generated by the GenerativeAgent itself.
- **Response:** The human-in-the-loop agent can respond to GenerativeAgent through an open text field or structured options, depending on the configuration.
- **Send:** After selecting a response, the agent can click Send to submit the response and close the ticket simultaneously.
- **Context:** Provides a summarized context of the conversation between GenerativeAgent and the customer.
- **Transcript:** Displays the complete customer interaction thread prior to the human-in-the-loop ticket being opened.
- **Customer:** Shows the customer's details, including company and specific account information for authenticated customers.

Customer <-> GenerativeAgent



GenerativeAgent

11:04:22 AM


Good Afternoon! Thank you for contacting Hudson, how can I help you today?

C

Customer

11:04:30 AM

Hello! I just received my last bill, and unfortunately I'm not sure i'll be able to pay it on time



GenerativeAgent

11:05:40 AM

Ok, I can help you with that

I'm pulling up your current bill information

Your current bill is \$150.00, and it is due on March 25.

If you're unable to pay on time, I can help you explore other options.


C

Customer

11:04:50 AM

ok, what are my options?

Can I receive an extension?



GenerativeAgent

11:05:00 AM

Just a moment

I'm checking your past bills

Your account standing is excellent

Please hold for a bit while I check on this.

Options for how to use the human in the loop

When deploying GenerativeAgent, you have a number of options for how it fits into your workflows and what role the human in the loop will play. The bottom line is that you get to decide which intents GenerativeAgent will handle and what it should do when it hits a roadblock.

Typical scenarios in which GenerativeAgent needs human intervention include:

- **Lack of knowledge base or API:** GenerativeAgent (or any AI chatbot) capabilities are only as good as the knowledge and APIs provided to it. In cases where GenerativeAgent lacks access to a system or cannot find the information it needs, the human in the loop can close this gap by stepping in to provide the necessary data or information or execute tasks in systems GenerativeAgent cannot access. It is your decision whether GenerativeAgent escalates these interactions to a human or simply asks for the help it needs to continue working with the customer itself.
- **Lack of authorization:** For some scenarios, you might not want GenerativeAgent to make certain decisions on its own. For these cases, you can require human decision-making or authorization by policy. Common examples include offering discounts, closing an account, or granting payment extensions. You can choose to require GenerativeAgent to ask the human in the loop for the decision or authorization it needs when it encounters these situations. Explicit customer request: No matter how effective your automated service is, sometimes a customer will specifically request to speak to a human. You can choose to have GenerativeAgent transfer the customer or offer to consult the human in the loop itself as a means of helping the customer reach a resolution faster. This option can be helpful and even appealing to customers when agent queues have long wait times.
- **System initiated:** In the case of an API error, safety risk, or ambiguous data, GenerativeAgent can seek human assistance for resolution.

In each of these situations, the human in the loop plays an important role by assisting GenerativeAgent in real time, ensuring that customers receive accurate and swift resolutions to their most complex problems. This human collaboration with GenerativeAgent accelerates adoption and expansion to cover additional use cases. The flexibility in deciding how you define the human-in-the-loop role enables your organization to deploy and expand the use of GenerativeAgent on your own terms and at your own pace.



What does the human in the loop do?

The role of a human-in-the-loop agent varies somewhat from the role of an agent that handles customer interactions directly. However, the workflows are similar, and the tools provided by GenerativeAgent will feel familiar.

The specific tasks the human in the loop handles will depend on your policies and implementation decisions. But in general, the human-in-the-loop agents will perform these tasks:

- **Triage tickets in real-time**
 - Monitor incoming consults by GenerativeAgent in the customer support platform.
 - Use the purpose-built human-in-the-loop interface to access relevant transcript, context and customer information.
 - Prioritize multiple tickets based on priority.
- **Resolve complex issues promptly**
 - Perform investigations and troubleshooting, provide approvals and authorization as required by policy, and access systems and databases that GenerativeAgent cannot access because the necessary API does not exist.
 - Maintain constant communication with other support teams, product managers, and engineers to expedite the resolution process.
- **Utilize and update knowledge articles and other designated sources of truth**
 - Reference and update the knowledge sources (including tools, such as your CRM or other systems with information critical to resolving customer issues) to provide accurate information.
 - Suggest improvements or updates to the knowledge sources based on recurring queries or new information.
- **Escalate customers to live agents when required**
 - Assess tickets for particularly complex or sensitive issues where a live agent is necessary and escalate to the live agent immediately. Note: The human-in-the-loop agent does not handle customer interactions directly. These need to be escalated specifically to a live agent queue.

Equipping the human in the loop

A critical component in enabling any agent is providing access to the right tools. Here's what the human in the loop will require.

- **Existing messaging system:** The human in the loop will use your contact center's existing ticket/messaging system to monitor, track, and manage GenerativeAgent consults instantly. GenerativeAgent raises tickets via the existing messaging system. These tickets get routed to dedicated human-in-the-loop queues. The GenerativeAgent human-in-the-loop interface integrates seamlessly as an iframe application in the agent workspace so it is visible and at the agent's fingertips every time a ticket is created.
- **Knowledge base:** It will be necessary for the human in the loop to have access to the knowledge base to reference and update information promptly as needed.
- **Software applications and sources of data:** The human in the loop must have access to all applications and data sources needed to resolve customer issues, including CRMs, databases, other systems of record, and troubleshooting software.
- **Internal communication platforms (helpful but not necessary):** Chat apps and video calls enable the human in the loop to collaborate in real time with other team members and departments as they troubleshoot and escalate issues.

How the human in the loop differs from a typical contact center agent

The primary role of an agent is to engage with customers through voice, chat, or other channels and resolve their issues in an efficient and empathetic manner. In contrast, the primary role of the human in the loop is to unblock and guide GenerativeAgent as it handles calls or chats and resolves the customers' issues.

Because GenerativeAgent can successfully resolve many customer issues on its own, the role of the human in the loop is geared toward more complex, high-value tasks rather than routine interactions. This shift in the balance of responsibilities empowers human-in-the-loop agents to operate with greater efficiency and impact.

With GenerativeAgent handling the vast majority of an interaction, the human-in-the-loop agent is not responsible for:

- **Engaging in pleasantries:** GenerativeAgent handles routine greetings and pleasantries with customers, which allows human-in-the-loop agents to dive straight into resolving issues.
- **Documenting interactions:** The GenerativeAgent human-in-the-loop functionality automatically logs and reports interactions, freeing the human-in-the-loop agent from this routine task.
- **Providing routine status updates:** Automated systems provide routine status updates to customers, which relieves the human-in-the-loop agent of this burden.
- **Owning complete resolution:** Human-in-the-loop agents assist with specific hurdles faced by GenerativeAgent rather than all issues faced by the customer. Their role is only to help GenerativeAgent when it gets stuck because it encounters ambiguous information, cannot access a system it needs, or requires human authorization to perform a specific task.

Identifying the ideal human-in-the-loop candidates

Because the primary responsibilities of a human-in-the-loop agent differ from those of a typical frontline agent, you'll need to fill these roles with a different kind of candidate with a slightly different skillset.

It's helpful to think of the human-in-the-loop agent as a new paradigm that depends on a particular set of personas, depending on the type of issue GenerativeAgent needs help to resolve. For example, resolving a technical issue, such as an API error, would require different skills than a situation in which an angry customer demands to speak to a human. In other words, these two scenarios each require a different human-in-the-loop agent persona. Keep in mind that the skills for these personas overlap, and a single employee might effectively function as more than one human-in-the-loop agent persona.

This list of personas describes the skills and responsibilities that will be important for the human-in-the-loop agent during different scenarios. Your contact center management will ultimately decide the optimal operational mix for their human-in-the-loop agents given your company's interaction volume, staffing models, and specific domain requirements.

Human-in-the-loop personas

Knowledge base advisor

This persona is needed when GenerativeAgent is unable to find the information it needs in the knowledge base or cannot access a system that's required to resolve a customer's issue.

Entry-level

Skills required

Research, knowledge management, problem-solving

Responsibilities

Updates and enhances the knowledge base, provides guidance to GenerativeAgent on new or unclear queries.

Decision authority

For some types of interactions, GenerativeAgent will need to perform a task that requires human authorization by company policy. In these cases, the human-in-the-loop agent will provide that decision authority.

Mid-level

Skills required

Decision-making, attention to detail, customer service

Responsibilities

Provides instructions on routine authorizations, such as promotional offers and basic account changes.

Senior-level

Skills required

Policy understanding, high-stakes decision-making

Responsibilities

Provides instructions for GenerativeAgent on handling sensitive decisions and sets policies for critical scenarios, such as account closure and bulk discounts.

Customer trust advisor

Sometimes a customer who is impatient or dissatisfied with their interaction with GenerativeAgent will specifically request to speak to a human. In these cases, a human agent will need to take over with a focus on easing the customer's frustration.

Mid-level

Skills required

Communication, empathy, conflict resolution

Responsibilities

Guides GenerativeAgent on how to rebuild trust with dissatisfied customers and resolve concerns based on established protocols.

Senior-level

Skills required

Advanced negotiation, dispute resolution, customer advocacy

Responsibilities

Provides strategic input to GenerativeAgent on complex or escalated customer scenarios, managing VIP cases indirectly.

Technical support advisor

Technical issues, such as API errors and ambiguous data, can prevent GenerativeAgent from completing tasks that are required to resolve a customer issue. In these cases, the human-in-the-loop agent steps in to provide GenerativeAgent with the technical support that it needs.

Entry-level

Skills required

Troubleshooting, basic technical skills, analytical thinking

Responsibilities

Offers solutions to GenerativeAgent when technical errors occur and provides direction on error handling and data ambiguities.

Mid-level

Skills required

Advanced troubleshooting, technical expertise, system analysis

Responsibilities

Manages complex technical instructions for GenerativeAgent, ensuring accurate guidance for API and system errors.

Quality assurance advisor

In some instances, GenerativeAgent will encounter ambiguous or non-standard scenarios or conflicting information in designated sources of truth. A quality assurance advisor can ensure the correct information reaches the customer and identify opportunities to improve internal systems, content, and flows.

Mid-level

Skills required

Analytical skills, attention to detail, AI system understanding

Responsibilities

Reviews GenerativeAgent interventions, identifies recurring issues, and updates GenerativeAgent protocols to minimize human handover.

Senior-level

Skills required

Strategic analysis, QA leadership, data-driven decision-making

Responsibilities

Leads the refinement of GenerativeAgent strategies, provides high-level recommendations for continuous improvement.

Compliance and risk advisor

Some customer interactions involve information and actions that are compliance-sensitive. In these cases, the human-in-the-loop agent provides guidance and authorization based on deep knowledge of legal and regulatory requirements.

Senior-level only

Skills required

Regulatory knowledge, risk assessment, ethical decision-making

Responsibilities

Instructs the bot on compliance-sensitive actions, ensuring adherence to legal and regulatory standards.

Important skills for all human-in-the-loop personas

Each human-in-the-loop persona requires its own set of skills, some of which are highly differentiated to your business, industry, and the persona's specific responsibilities. But all human-in-the-loop personas need some foundational skills to be effective in working with GenerativeAgent. You'll want to prioritize these skills in the hiring process and in ongoing professional development.

Software and tool usage

- **Live chat tools:** Because GenerativeAgent communicates with the human in the loop via chat, experience with tools like LiveChat, ChatBot, and others that are used for real-time collaboration in customer support can help in getting GenerativeAgent the assistance it needs to resolve a customer issue quickly.
- **Ticketing tools:** Agents can use the existing ticket or messaging system to monitor, track, and manage escalated tickets instantly. GenerativeAgent raises tickets via the existing messaging system.

Data entry and management

- **Data handling:** Familiarity with data protection and privacy regulations is necessary to ensure secure handling of customer data.

Knowledge of performance metrics

- **KPI understanding:** Familiarity with key performance indicators, such as response time, resolution time, and customer satisfaction scores will help the human in the loop track and improve performance.
- **Analytical skills:** The ability to analyze one's own performance metrics and reports is a key component in improving service quality.

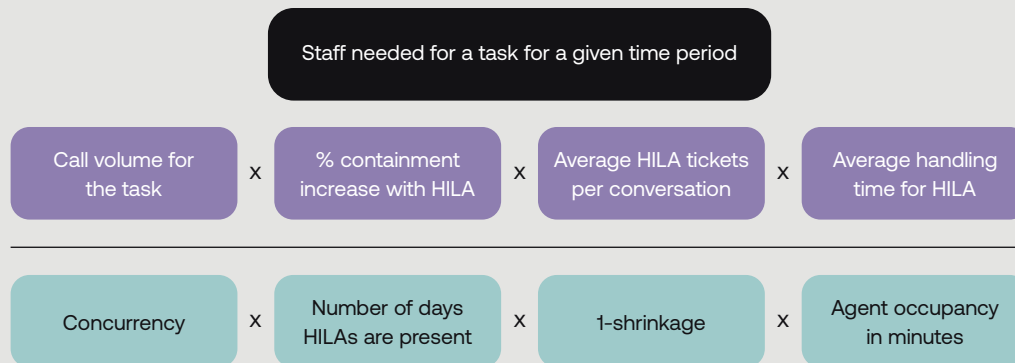
Documentation skills

- **Knowledge base maintenance:** Because GenerativeAgent relies on the knowledge base to stay grounded in your policies, the ability to update and manage knowledge articles or FAQs is a critical skill for equipping GenerativeAgent to handle customer inquiries.

Staffing and scaling the human in the loop

When determining staffing levels for the human in the loop, there are several key factors to consider.

At a very high level, you should consider this set of parameters for each intent that GenerativeAgent will handle:



It's important to keep in mind that some of the factors that determine staffing needs for the human in the loop depend on policy choices that your organization makes. For example, you might choose to limit concurrency for your human-in-the-loop agents initially until your contact center leadership feels confident in their ability to simultaneously oversee more interactions.

Consider this hypothetical scenario. Note that this is a simple linear model. It does not take into account service levels and is not predictive.

Let's say the AHT for a live agent is 8 minutes per interaction and that of a human-in-the-loop agent is 3 minutes per interaction.

Human-in-the-loop agent	Live agent
AHT 3 minutes per interaction	AHT 8 minutes per interaction
Throughput $60 \text{ mins} / 3 \text{ mins} = 20 \text{ interactions per hour}$	Throughput $60 \text{ mins} / 8 \text{ mins} = \sim 7 \text{ interactions per hour}$

Now that we have throughput, some common parameters would be:

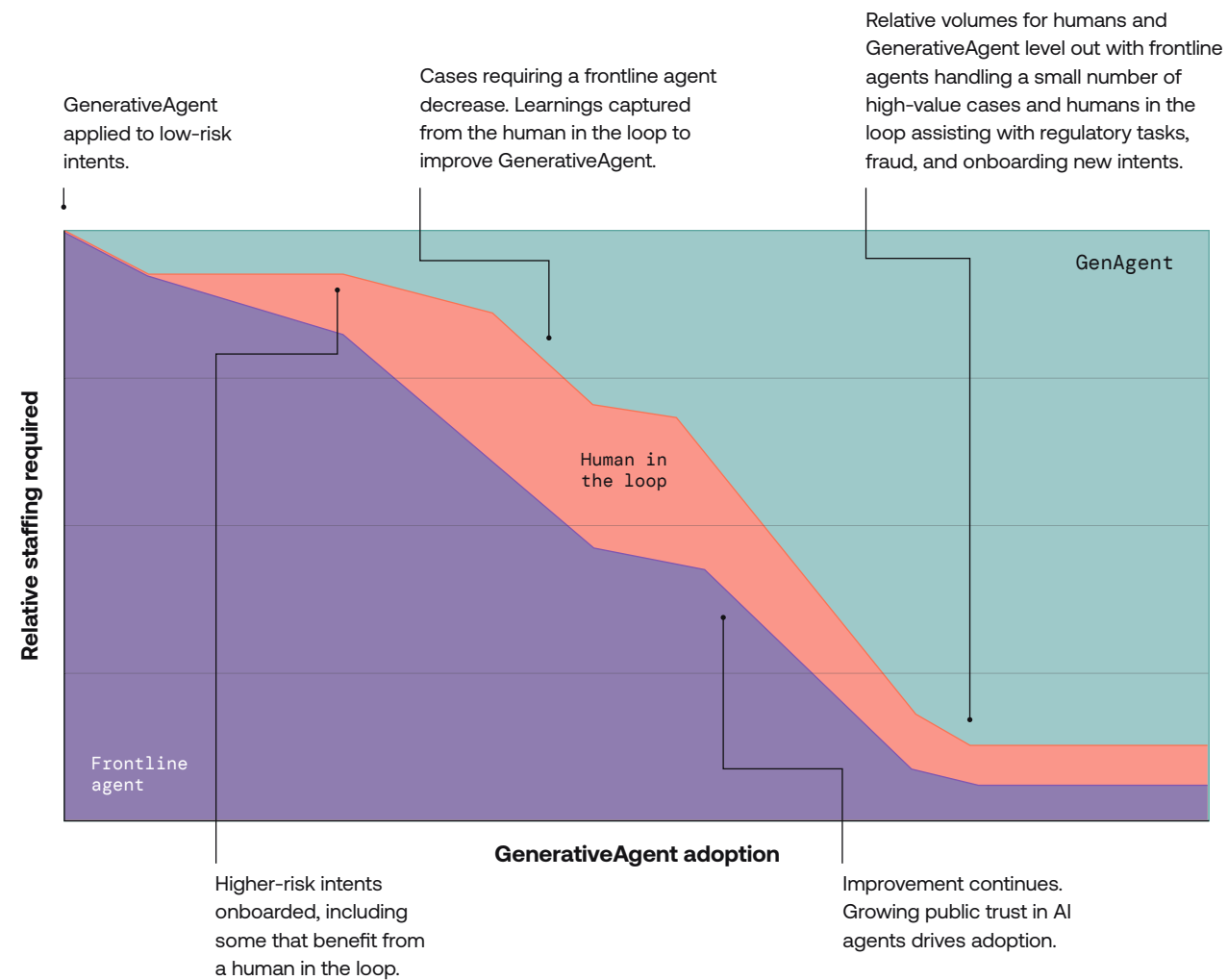
- **Operational hours:** 8 hours per day
- **Occupancy rate:** 85%

Knowing that the shrinkage rate is 20%, we can now estimate that for 3,000 interactions a day, required agent staffing would be as follows for equivalent volume.

Human-in-the-loop agent	Live agent
Number of agents = $3,000 \text{ interactions} / (20 \text{ interactions per hour} \times 8 \text{ hours per day}) = 18.75 \text{ agents}$	Number of agents = $3,000 \text{ interactions} / (7 \text{ interactions per hour} \times 8 \text{ hours per day}) = 53.57 \text{ agents}$
Adjusted for shrinkage and occupancy = $18.75 / (0.85 \times 0.8) = \sim 28 \text{ agents}$	Adjusted for shrinkage and occupancy = $53.57 / (0.85 \times 0.8) = \sim 79 \text{ agents}$

As you calculate your staffing needs, you'll want to keep in mind that the relative volume of work handled by human agents, GenerativeAgent, and human-in-the-loop agents will shift over time for several key reasons:

- The burden on human-in-the-loop agents will be greatest in the initial weeks after deploying a new intent.
- The consultative role of human-in-the-loop agents will be larger for some intents than for others.
- Over time, GenerativeAgent will learn from its interactions with human-in-the-loop agents, which will improve its ability to resolve customer interactions on its own.
- As your customers grow more comfortable interacting with GenerativeAgent, they will become less likely to specifically request a human.



Setting up queues for the humans in the loop

Once you've calculated your human-in-the-loop staffing needs based on the formula on the previous page, you'll need to establish queues for them. These should be dedicated queues for agents who engage only as the human in the loop and do not handle customer chats or calls themselves. This ensures that these agents are fully available to engage with GenerativeAgent to keep the interactions it handles moving toward resolution.

Setting up queues in which agents serve as the human in the loop and handle customer interactions directly can have some unintended negative consequences, including:

- Increased handle times for tickets due to frequent task switching and the challenge of prioritizing GenerativeAgent interactions vs. live chats and calls
- Missed human-in-the-loop consults while the agent is focused on interacting directly with a customer
- Skewed performance reporting for both GenerativeAgent and the human in the loop

With your initial deployment, you'll likely limit GenerativeAgent to a small number of intents, which will keep things simpler for the humans in the loop, too. As you expand to additional intents, you will want to consider how best to organize the human-in-the-loop queues. The principles you rely on to structure your current agent queues can inform how you structure these new queues, as well. For example, you might create queues based on groups of related intents.

Evaluating the performance of the human in the loop

When evaluating your human-in-the-loop agents, many of the key performance metrics you'll rely on are the same ones you already use with typical agents. However, in nearly every case, the interpretation, performance targets, and relative importance of these metrics differs from how they are used in evaluating agents who handle customer interactions directly. For example, average handle time will still be an important metric, but evaluating a human-in-the-loop agent based on the total handle time of GenerativeAgent interactions would not be a fair measure. Instead, the average handle time in this case would include only the time spent on the GenerativeAgent-human agent consult.

A baseline set of performance evaluation metrics for human-in-the-loop agents might include:

- Average handle time (human-in-the-loop consult only)
- Average first response time (human-in-the-loop consult only)
- Escalation rate (human-in-the-loop initiated)
- Total consults handled per hour
- Customer satisfaction metrics (end customer interaction)
- Resolution rate (end customer interaction)
- First contact resolution (end customer interaction)

It is important to note that human-in-the-loop agents would not want to be penalized in cases where the CSATs or resolution rates are low because the GenerativeAgent hallucinated or did not handle the conversation with care. Satisfaction scores should therefore be scrutinized to understand whether performance issues originated with the human or the technology.

The evolving human-AI partnership in your contact center

As you incorporate more advanced AI solutions like GenerativeAgent into your contact center, the need for human expertise that complements the new technologies will become increasingly important. That means the role of the contact center agent will continue to evolve. They won't need to be experts in AI, but they will need to adapt and acquire some new skills.

This shift will require some changes in recruiting, onboarding, and training of agents who will work alongside GenerativeAgent as the human in the loop. The good news is that the flexibility of GenerativeAgent enables you to adapt and move forward with these changes at your own pace.

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 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 age of AI. 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.