Empowered, Confident, and Prepared
Driving Chatbot Product Vision Through User Research

MOLLY MAHAR, *Salesforce*
GREGORY A. BENNETT, *Salesforce*

This case study explores how a series of customer site visits to two international service centers drove design recommendations for a chatbot building platform that could encourage positive agent-chatbot collaboration. The first part of the case focuses on the research undertaken by a team of user experience practitioners at the enterprise software company Salesforce. The team used contextual inquiry and group interviews to better understand the daily experience of customer service agents and service teams in search of ways to responsibly implement automation tools like chatbots within a service center environment. The second part of the case study highlights how the UX team applied these learnings into specific product recommendations and developed a set of principles that could drive the product forward while remaining empathetic and supportive of customer service agents.

INTRODUCTION

In contrast to their job title, customer service agents aren't treated as if they have much agency. Service agents are trained to follow precise scripts and protocols when dealing with problems, and may be quite limited in what power they're granted to actually solve a customer's problem. Their shifts are scheduled to coincide with convenient hours for customers, not necessarily for agents. Agents deal with customers in their worst moments: frustrated, angry, scared, stressed; typically, customers drive the tenor and direction of the conversation. Agents rarely get to hear first-hand about successful service they might provide, because happy customers don't call back to say "thank you." Even if they did, it would be exceedingly unlikely for those customers to get routed back to the same service agent who helped them initially, because of the way automated software routing processes function. Lacking much agency in their day-to-day work, customer service agents can be a vulnerable worker population, treated as low-skill, expendable, replaceable, seasonal workers. The subreddit "Tales From Call Centers" (/r/Talesfromcallcenters) is full of first-hand accounts that highlight these issues for call center employees, demonstrating how this job can be deeply punishing, and only occasionally rewarding.

Larger companies often segment their customer support into different "tiers," or levels of service (Kidd and Hertvik 2019). A Tier 1 agent is less experienced and less knowledgeable than a Tier 3 agent, who has learned the ropes and the products and is expected to be an expert. Tier 1 agents are the most vulnerable population: they are paid the least, have the least power and autonomy in a customer interaction, and are the most replaceable. They are also the most likely to be automated out of a job as companies look to streamline operations and encourage customers to help themselves through self-service. Self-service, for these companies, is considered Tier 0, which includes instances where the customer finds the information themselves, whether through help articles on the website or through a chatbot. The future job outlook for customer service workers is both expanding and contracting, depending on industry (Leopold, Ratcheva and Zahidi 2018), and likely also on level of expertise. On the one hand, even in an automated world, people still like to be

served by other people (that "human touch"); on the other hand, more and more companies are interested in automating Tier 1 tasks down to Tier 0 (self-service) tasks.

There is a very public conversation happening around how the fear that automation will eventually take over everybody's jobs. Automation and self-service can be seen as two perspectives on agency: automation equates to people losing agency, and self-service equates to people gaining agency. Often, it's a balancing act between customer service agents and consumers—and consumers seem to like self-service. There were days when nobody pumped their own gas at the gas station, when every flight check-in involved speaking with a desk agent and printing out a physical ticket, when librarians were the only people with direct access to shelves of books. These days, do most people consider pumping their own gas to be *automation*? For a consumer, self-service (e.g., finding an article online to assist with a problem, using an ATM to get cash, or getting help with a problem from a chatbot) *is* agency. They now have the power to solve their own problems, on their own schedule. For a service agent, companies' efforts to provide self-service to consumers, and therefore agency and convenience, results in automation solutions that have the potential to help or hurt service agents, depending how they're designed and implemented.

It is precisely because of automation's potential for both good and harm in the lives of customer service agents that the User Experience (UX) team for the new chatbot builder product at Salesforce sought to visit customer service centers and observe and learn from agents. Since chatbots for customer service were still relatively new, Salesforce's new product offering would be the first time that many large enterprise customers had ever considered building a chatbot to assist their customer service organization. Because of this unfamiliarity, the UX team wanted to provide best practices and recommendations around how to responsibly and effectively launch a customer service chatbot in a service center. They did not feel confident building a platform without knowing how to build protections and best practices into it that could benefit service agents while also benefiting the companies for which they work. To better understand how automation might be beneficial to service agents, and therefore how to build that into the chatbot building product, the UX team needed a deeper understanding of the experiences of customer service agents.

RESEARCH GOALS

In order to provide recommendations on how to responsibly implement automation, the UX team needed to have a deep understanding of the real-time experience of service agents, as well as which parts of a customer service agent's job were expendable and which parts were enjoyable and satisfying. The team also needed to understand how service centers functioned so that they could provide a product that would successfully automate the parts of a service center agent's job that *were* expendable. In order to implement responsibly, the team needed to better understand how automation might change (for better or worse) the agents' current jobs. Did agents think about how or where automation could take something repetitive off their plate? Did they feel threatened or excited by automation, or feel something else entirely?

Customers tend to ask questions or seek help via many channels, depending on availability, context, and customer preference. Common channels are phone, email, web chat, mobile messaging, and social media. Since chatbots are text-based, the highest priority for the team was seeing live chat service agents in action, to most closely reflect the cadence

and issues that companies are likely to use chatbots to assist with. Because different channels require different skill sets, the UX team hypothesized that agents serving different channels might have different needs or desires around automation.

It is important to note that this research was not intended to develop personas around customer service. Salesforce as a company conducts quantitative surveys on a cadence to develop, refine, and modify its user personas based on how its end users interact with the software. The company had already developed and disseminated personas that captured the three major user groups that the UX team would need to interact with: Tier 1 customer service agents (called "case solvers" in the Salesforce parlance), experienced Tier 2 or 3 service agents ("expert agents") and support team supervisor-managers ("team leads").

THE OPPORTUNITY

At Salesforce, user experience practitioners interact regularly with the users of their software. Designers and researchers remotely interview Salesforce administrators (those responsible for configuring the software to align with business processes), sales reps, customer service agents, marketers, business analysts, and others. Customer service agents are notoriously difficult to interview and observe because their time is so tightly controlled and managed by their organizations. Every second counts, and thus it is difficult to send observers into call centers. Responsible data practices, in conjunction with laws around privacy such as GDPR, preclude the sharing or saving of end customer data, making it nearly impossible to observe service agents at work without going onsite and observing in person. Their screens always reflect private information about the customers they're serving, so companies—Salesforce's customers—are rightfully protective of that data and do not share it.

In addition, because the product was just launching, there were very few customers using Salesforce chatbots yet. While it would have been ideal for the UX team to observe agents who were already interacting with chatbots from a support perspective, the immaturity of the technology space meant that the team would likely have to settle with just observing chat agents, and deriving insights and recommendations from their current experiences.

While pursuing opportunities to meet with business customers, the UX team was offered the opportunity to join another product team that had planned a visit to call centers in Manila, the capital of the Philippines. It was an interesting opportunity because the team would visit two call centers that provided outsourced support to the same large fitness technology company, a Salesforce customer. This provided the opportunity to see how multiple service centers functioned to serve the same ultimate client and customer base. The call centers provided English-language support in all major channels: phone, email, web chat, and social media. The visit provided a perfect opportunity to get baseline knowledge of the chat agent experience prior to the implementation of chatbots, and see where chatbots could help or hinder from an agent perspective. In addition, because Manila is an international hub for outsourced customer service, the team expected that these vendors would provide an accurate view into the experiences of high-volume, outsourced service agents in particular—those who are most likely to have their jobs affected in some way by automation, because they are the least visible to the decision-makers at company headquarters.

The UX team traveled to Manila to observe service center agents over the course of five nights. Manila is a hub for international customer service center outsourcing, so the two companies that the UX team visited both supported one large Salesforce customer on a non-exclusive basis (the vendors also had other enterprise customers). The service agents that the team observed were scheduled on the overnight shifts, so that they could support English-speaking customers in the USA and the UK during local business hours.

METHODOLOGY

As noted previously, the primary research goals were to 1) understand the daily experience of service agents at their jobs, 2) understand at a high level how high-volume service centers functioned operationally related to agents and automation, and 3) understand how automation might change a service agent's job from an agent's point of view. This knowledge would then allow the chatbot builder UX team to develop concrete recommendations on how to responsibly implement chatbots within a service center, to benefit end consumers as well as the customer service agents who must work with chatbots in a new kind of human-machine collaboration. To address these three goals, the team planned to shadow service agents while they did their job, conduct brief interviews during or immediately after their shifts or interactions were completed, and interview team leads (customer service managers) about the operation and functions of the service center as a whole.

Contextual Inquiry With Case Solvers

To understand the daily experience of service agents at their jobs, the UX team planned to shadow agents while they worked, observing:
- The general environment of a customer service center
- The general flow and schedule to develop a sense of a "typical" agent workday
- How issues progress up the tiers of service, from Tier 1 to Tier 2 or 3 (typically called "escalations")
- Any ways that agents collaborated with other agents in the course of their jobs
- How and why agents used pre-composed responses in their interactions with customers, and how they maintained and accessed them (pre-composed responses were known to be used by at least some customers because Salesforce offered that functionality in other product features)
- Any differences in the above based on channel used (email, chat, social media, or phone)

During and after these shadow sessions, the UX team planned to conduct short interviews with the service agents to probe more deeply into how agents saw automation potentially affecting their jobs, specifically:
- What portions of the job pleased or satisfied agents, and which portions of the job were displeasing or negative?
- Where did they see automation as being helpful to them? Harmful?
- Were they worried about automation? Did they think about it at all?

The service vendor made many of their customer service agents available, such that the UX team was able to spend 1-2 hours with each agent, and observe between two and six agents each evening. Over the course of the five nights, the UX team observed agents that handled cases via chat, email, phone, and social media. During these shadow sessions, the researcher would introduce themselves and then sit beside the service agent while the agent took phone calls, or received and responded to emails, chats, or social media messages.

With the written channels, the researchers would often ask clarifying questions about what they'd just seen on screen, or why an agent did something one way or another. The researchers observed and noted what windows the agent kept on screen, how they arranged them, and what their desks looked like. The UX team found that asking these questions during the course of the agent's workflow was much easier during cases on a written channel, since the customer on the other end of the correspondence didn't know or need to know that there was an observer present. With service agents handling phone calls, all follow-up questions and clarifications needed to happen after the end customer had hung up the phone and the issue was resolved.

To interview agents about automation, the UX team planned to either ask questions during the course of handling customer issues, or to obtain 1:1 time with agents during breaks in their shift and interview them off the floor of the service center, if possible. These interviews were planned to be only a few minutes long. The team ran into a number of issues when attempting to address this portion of the research and was unsuccessful, which will be discussed shortly.

Interviews With Team Leads

To understand how these service centers functioned at a high level, including areas of automation, the UX team planned to interview team leads and supervisors to learn:
- How they would want to change their current setup and workflows
- How they measure current KPIs (key performance indicators) for agents, and how those might change with increased automation
- How supervisors interact with other agents in person on the floor and digitally, during the course of their job

It was unclear prior to the visit what format would be made available for interviewing team leads and supervisors. Upon arrival, the team learned large group sessions had been planned by both vendors. In these sessions, the vendors' participants were a mix of team leads who supervised the teams of agents, and the service account executives who maintained the relationship between the vendor and the fitness technology company, which manages the actual Salesforce implementation and is a Salesforce customer. The Salesforce administrator was also part of the sessions; he traveled with the UX team from the US and is an employee of the fitness technology company.

During these group sessions, team leads discussed areas for improvement in how to implement the Salesforce system, including workarounds that Tier 1 agents at one of the vendors had discovered in order to disassociate themselves from negative customer feedback. Employees discussed common KPIs that were used, including Average Handle Time (AHT), which is a common service center metric. Most of the employee-supervisor

interaction data was actually gathered observationally during shadow sessions with Tier 1 agents, rather than being investigated during the group sessions.

Table 1: Study parameters

Method	Number of Participants	Service Cloud User Persona	Service Channel	Location
Contextual Inquiry	12	Case Solver	Chat	Vendor 1
Contextual Inquiry	5	Case Solver	Email	Vendor 1
Contextual Inquiry	9	Case Solver	Phone	Vendor 1
Contextual Inquiry	2	Case Solver	Social Media	Vendor 1
Contextual Inquiry	10	Case Solver	Chat	Vendor 2
Group Interview	12	Team Leader	Chat, Email, Phone, Social Media	Vendor 1
Group Interview	5	Team Leader	Chat	Vendor 2

Lessons Learned

One interesting development that was unexpected was that the Tier 1 service agents did not seem to understand that the UX team that was observing was unrelated to their clients, the fitness technology company. Because the UX team traveled and arrived along with the Salesforce administrator who represented the fitness technology company, agents universally seemed to assume that the observers were all part of the same group. Even after introductions that the UX team came from the software company that made the software that the agents were using, agents did not seem to grasp nor care that the UX team was not from their client company. However, because the different teams traveled and arrived together, it was very clear to the agents that it was acceptable and expected that they would solve cases somewhat more slowly that day due to having the distraction of answering questions and having an observer present. Having that tacit support, as well as verbal support from the Salesforce administrator (who was, ultimately, the only representative from their client, the fitness technology company), did seem to make agents much more comfortable having the UX team ask questions and dig into their workflows.

This mistaken assumption that the UX team was in fact working for the vendor's client may have led to more reticence in any answers that would have shown concern or trepidation around the potential for automating the agents' jobs away. Due to the volume of issues these call centers handled, and the nature of the vendors' oversight, the UX team was unable to conduct 1:1 interviews with Tier 1 agents outside of the shadow sessions. The goal for those had been to uncover agent attitudes about automation and, in particular, about

working with chatbots. The UX team also observed that agents were not forthcoming about, or not interested in, revealing any attitudes about automation while working on the floor. This was understandable due to the close physical proximity of other agents and supervisors, who roamed the floor and could easily overhear anything that was said by the agents. Agents were willing to discuss how they currently used automations, and where automation might speed up their workflows, but they did not discuss anything that might be seen or construed as critical to the way the service center functioned or was managed. The UX team thus focused mainly on observations and clarifications after a few failed attempts at digging deeper about automations.

Somewhat similarly, with agents handling phone calls, the observations did not include much time or space for questioning outside of clarifications. Researchers were given a pair of headphones without a microphone and were connected to the live calls to listen in on both sides of the conversation between the agent and the end customer. Agents clearly couldn't answer researchers' questions while the phone line was open, so the UX team had to reserve any questions until after the call was closed out. But that is also the time that agents must complete their "after-call work" (ACW), or case wrap up, which typically involves typing out a summary of the conversation as well as the steps taken to resolve it. At the vendor observed, agents are given roughly 15 seconds to complete this work after every call, before being given a new call. The new call is signaled by a short tone on the phone line before being automatically connected—the agents do not physically pick up a phone nor do they press a button to connect. It happens automatically. There were often moments when a question or an answer between the agent and the UX team went unanswered or was cut off mid-sentence due to an incoming phone call. Those incoming cases happened with the same frequency on the written channels (web chat, social media messaging, email), but the agents had a much easier time multi-tasking, and were able to answer lingering questions while still handling the customer case in front of them.

KEY FINDINGS AND TAKEAWAYS

As part of a software development team, the UX team makes use of broader organizational data-derived personas to help shape and direct their product development efforts. As noted previously, the three relevant ones to this research activity were the "case solver" (a tier 1 service agent), the "expert agent" (an experienced agent), and the "team lead" (a team supervisor). It can be easy to fall into rote acceptance and recitation of these personas to one's product development team if one does not actually interact frequently with the end users of one's product. There is a level of empathy that develops through the richness of small details, the ones that escape the persona and provide real texture to the experience. These minor details often end up being the difference between a development team that truly understands and aligns on why certain product choices are being made, versus one that is simply going with decisions made by others. The UX team was lucky to observe a number of these textural experiences while learning about agents' daily work and the operations of service centers.

Life in the Service Center

There is a surprising amount of security in these multi-tenant call center buildings. Security guards with sign-in sheets and metal detectors were present at both vendors where the UX team performed research. Proximity badges and visitor badges were required for entry into any room. The UX team, as visitors, received special badges that allowed them to bring computers, phones, and note-taking equipment (literally, pens and paper) into the call center. Agents themselves were not allowed to bring phones, pens, pencils, or computers onto the service floor—they had lockers in the hallway where they left their personal belongings. The security risk stems from the customer data that outsourcing vendors have access to in the course of their job; clients do not want agents walking off with it.

Interestingly, the security guards always greeted those entering the premises with "Good morning," even though it was the middle of the night when the team arrived and worked. The teams worked the overnight shift, serving English-language customers in the US and UK. "Good morning" seemed to be the standard greeting for the night shift, since they are just starting their day—it sets the tone that agents will be providing support to people who, in their own time zone, had just begun their day.

The UX team learned, from agents themselves and from their team leads, that most agents in the Manila service center lived many hours away. Some had traveled on three or more modes of transportation to get to work. Since agents often lived far away and weather could be unpredictable, one of the two call center vendors had created sleeping areas where agents could stay and sleep if they were trapped by a monsoon or other inclement weather that affected transportation home.

Entering a call center is much like entering many of the open-plan offices one might see around the world these days. There are groups of agents in pods, formed by a few short rows of desks, and the channel that those pods handle can be identified somewhat by sound. The agents that handle phone calls are always speaking, often quite loudly, leading to a much louder pod. The chat, email, and social media pods are much quieter by comparison, with chat agents being the next loudest. This was due in part to the speed at which they type, the audible alerts that the software puts out when a customer has been waiting too long for a response, and the general chatter that happens as agents speak to one another or ask questions of their supervisors. (The supervisors are always roaming the floor, available for help but also checking over agents' shoulders and keeping tabs on everyone.) Email and social media pods both operated at a much slower pace than phone or chat, and were thus quieter.

There was a hierarchy to an agent's job and advancement opportunities at these vendors, as the UX team learned from one of the vendors. Agents typically begin their careers answering emails, which have the most flexibility in response time. Agents will then graduate up to handling chat inquiries, which require faster response times and involve handling more than one chat simultaneously (typically 2-3 conversations). Agents who have been successful on chat might then be upgraded to handling phone calls, *if* they have a good spoken demeanor and high energy. Many of the phone agents used a nickname to introduce themselves to customers, rather than their real names (which in this case were longer or more complex than the names they gave to customers). If an agent succeeds at chat but is not a good fit for handling phone cases, they could be promoted to social media. Social media has a more relaxed timeline, like email, but the added stress that responses are often

very public. For this reason, only the most experienced and talented agents were assigned to handle social media issues at the vendor the team visited. They must know the products, and know how to handle customers well so that issues don't become publicity nightmares. Agents on social media are a much more visible representation of the client company, so they are chosen carefully for their skill and experience.

The Chat Agent Experience

Since chatbots initially will be used on live chat channels, the UX team was primarily interested in observing specifics about how agents on chat channels dealt with customer problems. These observations highlight what the team learned about agents working that channel in particular.

The UX team was able to observe the speed at which chat agents typically responded to customers, above and beyond the SLA. An SLA is a "Service-Level Agreement," which is typically a contractual agreement specifying exactly how long a customer can expect to wait before their issue is resolved (Wikipedia contributors 2019). A client company might promise their consumers that they will solve any problem within 24 hours, for instance. Customer service vendors are thus obligated to also follow that client guideline when dealing with consumer issues. In addition, there are typically internal, procedural SLAs, such as ones that might require a problem to be assigned to an agent within two hours, or closed within eight hours. At these vendors, there are other process requirements for chat, for instance, that customers shouldn't be kept waiting more than two minutes without a response from the agent with whom they're chatting. The UX team also observed processes around how agents could only close out cases (mark them as "resolved") once the customer had confirmed and ended the chat themselves. Otherwise, sometimes agents were left to wait a specified period of time before being able to say that the customer had abandoned the chat. The UX team found that chat agents would typically respond within a few seconds to the customers. This response time was aided by the fact that the service center software allows the agent to see what the customer is typing into the message input field before the customer hits "Enter" to send the text. Thus, by the time the customer finally "sends" their response, the chat agent has already had a chance to see what's coming and start finding an answer and drafting their response outside of the chat window.

During the course of these chats, agents made extensive use of pre-composed responses that they would copy-paste from somewhere else into the chat window, then modify (for instance, with the customer's name) before sending. The UX team observed different workflows around these pre-composed responses, depending on the vendor, leading them to believe that the client itself did not provide or dictate what these responses should be. The agents called these their "spiels" at one vendor. Agents seemed to maintain their pre-composed responses in their own voice and tone, though many noted that they would share their responses with others, or that someone else had helped them get started at the call center by sharing their documents with them. These pre-composed responses were manually maintained, searched, and copy-pasted, making them a significant automation opportunity. Indeed, the client (the fitness technology company) had already programmed a set of pre-composed responses in the Salesforce system as "macros". The UX team observed that agents used a much larger number of pre-composed responses than were available and curated, however.

The UX team was also able to observe a number of chat escalations, whereby a case solver in Tier 1 (the lowest-level agents) passed a customer case up to a Tier 2 or 3 expert agent who was better equipped to handle it or the customer. The team observed that the chat agent who was originally handling the case, upon realizing that they would need assistance, would flag their supervisor, either over chat or by raising their hand or even walking over, provide a brief summary of the issue, and ask for help. The supervisor would then decide who would receive the escalation, and either the original agent or the supervisor would give the Tier 2 agent a quick summary of what was coming. This was a very interesting observation, because the customer service software can automatically escalate from one tier to another tier or to specific agents. Thus, it was a workaround and a clear preference at this vendor to have agents interact directly before handing over a case. This was another area where the UX team saw an opportunity for automation to potentially assist, because the agents clearly found this interaction method useful for both tiers of agents.

Agents + Automation = Teammates

The UX team aims to keep ethical and responsible product development front and center, and although they weren't able to get candid responses to their planned interview questions regarding how agents felt about automation, the team was able to better understand how agents saw themselves and their occupations. Chat agents in particular claimed a satisfaction in solving problems *quickly*. This is perhaps unsurprising, considering they are judged on the speed of issue resolution (average handle time, or AHT). Agents did not have much time during the workday to interact casually with other agents at their tier, but they did regularly communicate via internal chat channels. They made an effort to communicate with other agents to learn from them, to share pre-composed responses, and to provide context to escalations. Thus, agents seemingly found such communications of enough benefit to outweigh any potential negative impact to their resolution time.

The behaviors that the team observed agents take—reaching out and receiving help, providing a heads up to colleagues before escalating a case to them, sharing resources that had been helpful—all seemed designed to help agents feel a sense of preparedness and confidence in the work they were doing and the new problems they were encountering. Since the UX team wanted to maintain empowerment at the core of their experience, the team outlined how agents interacting with a bot should feel: empowered, confident, and prepared. These principles also point to developing bots as *teammates*, rather than as agent replacements. Many large companies seek to develop automation in ways that do not negatively impact their existing agents; the UX team now had a set of design principles that could drive their product design decisions. Would a certain feature make an agent feel more empowered? More confident in their solution for consumers? More prepared to handle new issues? These were the types of features that the UX team wanted to incorporate into the product vision. The behaviors observed around how colleagues interacted also provided insight into how a chatbot could potentially be seen and integrated as part of the team, and be respected as such. Ultimately, it seemed that if a chatbot could help agents continue to solve problems, and do so more quickly, it was likely to be accepted as a teammate.

IMPACT ON THE PRODUCT

The UX team was able to walk away from their weeklong observations with a number of specific design recommendations that could be implemented over time, providing both immediate and long-term value to the bot builder product. Features for customers, like a bot response delay (intentional friction) for more natural conversations, could be implemented immediately. Longer-term recommendations around escalation summaries, a use case for de-escalations, and deeper voice and tone customization have also been adopted to varying degrees. Those long-term features are agent-focused, designed to provide agents with more confidence and make space for more high-value interaction time with customers when handling cases.

Bot Response Delay

In the short term, the observations allowed the UX team to provide best practices on how to adjust the timing of the bot's responses during chat to more closely match expectations that customers would have developed through chatting with human agents. The value proposition of using a chatbot as a frontline Tier 0 resource, which then escalates issues the bot cannot solve to Tier 1 agents, relies on the bot responding quickly to all customer inquiries. However, the UX team had learned in prior research that when companies' bots had been responding instantaneously, it felt unnatural to consumers— especially when multiple messages would arrive at the same time. Observations in the service centers allowed the UX team to provide specific recommendations around timing, and to determine that since human agents at their quickest responded in 1–4 seconds, bots could respond in that timeframe and still be considered a fast response, without the need to respond instantaneously. The chatbot building product was updated in the next release cycle to include a variable "bot response delay" feature that would allow companies to choose a delay time that felt right for their conversation design and customers. The chatbot processing engine would then add this delay to each message, to stagger the arrival of a series of messages sent in quick succession, and to allow consumers a brief chance to read each message before the next one arrives. This feature was designed to benefit end consumers, and does not impact customer service agents, although it was developed through the observations of their chat conversations.

Conversation Summaries

Observations led the UX team to learn that summaries could offer value not only during escalations, but also after the fact, as a way to handle case wrap up and help agents quickly take note of what was done to help the customer. The potential value of providing a summary of a chat conversation seems incredibly obvious in hindsight, but the accepted viewpoint prior to these observations was that agents just read the chat transcript as they received an escalated case, and that that worked fine. Seeing how Tier 1 agents prepped their colleagues in Tier 2 when a case was coming was somewhat revelatory for the UX team. What the team observed was that agents were chatting quickly and handling multiple cases at one time, such that they didn't reliably have time to read over an entire chat before needing to respond and help the customer. Successful summaries, on the other hand, could help keep

responses within the designated SLA time period, and keep customers happy. This in turn could help agents feel confident and prepared when they address cases that may have been escalated to them by a bot. The content of the summary is also important. The team observed that agents were not telling expert agents *what they'd said*, but rather *what they'd done* to help the customer already, and what had, or, more frequently, had not worked. This meant that summaries should ideally be action-oriented: what actions had the chatbot taken already, and what were those outcomes? That information could be quite useful to a Tier 1 or 2 agent, who could then hop into a chat with an acknowledgement of what had been tried already, and an immediate plan for next steps. The concept of adding a summary has been added to the product roadmap for multiple automation-related products at Salesforce since being introduced by the chatbot UX team.
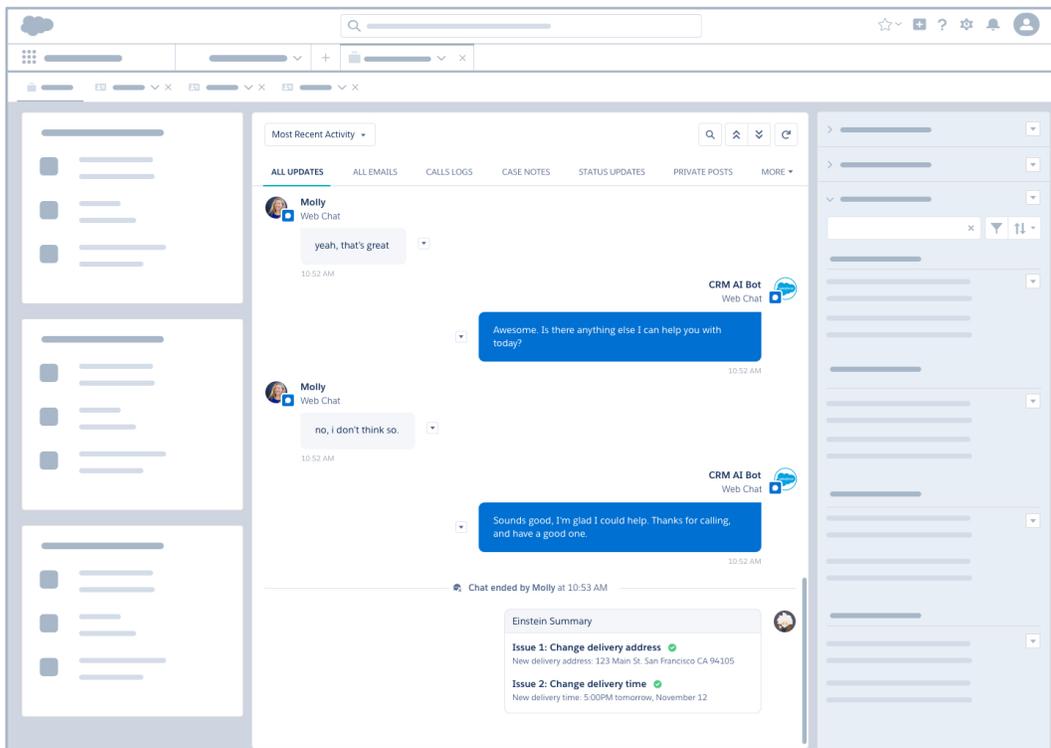

Figure 1: Wireframe of a chat transcript within the Salesforce Service Cloud agent console that contains the suggested summary component in context at the end of the conversation. Image © Salesforce, used with permission.
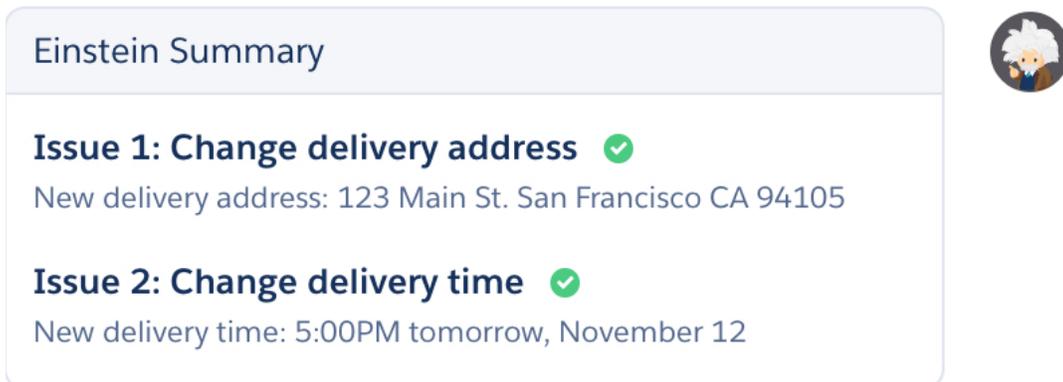
Figure 2: Wireframe of the summary component. Image © Salesforce, used with permission.

Agent-to-Bot Handoffs (De-escalations)

Observations also revealed a need for de-escalations: when an issue goes from a higher tier to a lower tier of service. In this case, the UX team saw value for agents to be able to pass conversations back to bots, who could then handle simple interaction flows for them. The UX team observed many agents waiting for the SLA to run out when a customer didn't respond, before they could close out a case. This wasted precious moments for the vendor's client company—time during which the agent couldn't help another customer, but also wasn't helping their current customer—as well as appearing quite boring to the agents themselves. The UX team hypothesized that being able to hand a conversation back to a bot that could "close out" the conversation and ensure that the customer had, indeed, left the chat, could open up the agent's time and either give them more breathing room between chat conversations, or allow them to accept a new chat if they wanted. A pared-down version of this feature, allowing a bot to handoff a conversation to another bot, has been implemented in the chatbot builder product already, and agent-to-bot handoffs are now an acknowledged opportunity area by the chatbot builder product team.

Voice and Tone Customization Tools

Finally, seeing how each agent customized and curated their pre-composed responses, the UX team recommended adding features addressing voice and tone customization in the future. As a customer calling a help desk, one might feel that the agent is simply following a script—and in some ways and for some questions, they are—but the team observed that those agents actually spent a great deal of effort trying to optimize their response time while adding their own personal touch to each of their communications. Bots should do the same, and if agents had helper bots that they could de-escalate to, those bots should be customized to fit the voice and tone of the agent with whom they're working.

COMMUNICATING THE RESULTS

After every research engagement, the UX team posts a research report to an internal website so that it is accessible to the rest of the UX organization and product stakeholders. In this case, because the findings impacted multiple products, the team also gave a presentation for the entire UX organization that highlighted the research done and the guiding outcomes that now drive the product—Empowered, Confident, and Prepared—after the service center observations. Salesforce UX is very user-driven, and agent agency in particular is a hot topic for the organization, so designers were very engaged. The goal with that presentation was to drive more empathy amongst designers by providing a very visceral description of the call center life, and bring more detail into the persona of a service center agent, a "case solver."

Chat summarization, in particular, has been presented numerous times in internal company executive summits, because it impacts a number of chat-related products that incorporate intelligence. Summarization is not only relevant to bot interactions, but can be applied to wrap up activities as well as analysis on cases. The designs for summarization, originally created to be used in chat escalations from bots to agents, have thus seen more life and are currently being incorporated into three different products.

In addition, this research has seen a long lifespan due to its first-hand nature. It provides a wealth of anecdotes that can be drawn on by the UX team during discussions with product management, engineering, and other stakeholders. Learnings from the research have even been incorporated into best practices that are recommended to customers worldwide who are using the Salesforce Einstein Bots chatbot building product.

CONCLUSION

This case study reveals how informative an ethnographic observation can be, even when key research questions aren't answered. The UX research team never was able to get first-hand responses to how agents felt about automation, beyond immediate ways that agents could be helped by minor automations in their workflows. And yet, the observations yielded a wealth of information that led to a richer, deeper understanding of the end users that the bots UX team was designing for. Such is the value of ethnography, to provide insight even while withholding concrete answers.

Customer service agents, like most other employees, find satisfaction from doing their jobs well. They seek to solve customer problems. The challenge is how to provide agency without autonomy, because it is unlikely that at any time in the near future, companies will give customer service agents complete autonomy over their schedule, what questions they answer, or even their time. The nature of a service agent is to be ready at a moment's notice to respond to nearly any inquiry. Empowerment and agency in this context, then, means providing resources to allow agents to do this efficiently and to allow them to move up the ranks and gain recognition and skills from learning to address new problems.

The chat center observations that the UX team undertook in Manila also allowed the team to better understand what allowed agents to be confident and prepared in how they handled conversations with customers: they had a library of communal knowledge that was regularly curated and updated, they communicated with others when they needed their help, and they took advantage of every opportunity to provide better service to their customers.

Doing this allowed agents to feel some agency in their activities, because they could personalize responses to their liking, keeping their personality. The UX team learned firsthand how important it is to design within this framework so that human-AI collaboration doesn't lose those elements that provide agency and satisfaction to service agents.

Molly Mahar is a Senior Product Designer at Salesforce and design lead for Einstein Bots, a product offering that leverages AI and chatbots to help support centers manage and grow their operations at scale. Her work focuses on making AI and machine learning understandable and usable by non-technical audiences. *m.mahar@salesforce.com*

Gregory A. Bennett is Lead User Researcher at Salesforce for Service Cloud Einstein, a suite of products that leverage conversational AI to benefit support centers and their agents. As a formally trained linguist, his research focuses on empowering businesses to design conversations that feel natural and helpful, build user trust, and meet user expectations for linguistic behavior. *gbennett@salesforce.com*

REFERENCES CITED

"r/Talesfromcallcenters." reddit. Accessed October 1, 2019. https://www.reddit.com/r/talesfromcallcenters/.

Kidd, Chrissy, and Joe Hertvik. 2019. "IT Support Levels Clearly Explained: L1, L2, L3, and More." BMC Blogs, April 25. Accessed October 1, 2019. https://www.bmc.com/blogs/support-levels-level-1-level-2-level-3/.

Leopold, Till Alexander, Vesselina Stefanova Ratcheva, and Saadia Zahidi. 2018. "The Future of Jobs Report 2018." World Economic Forum, September 17, 2018. Accessed October 1, 2019. https://www.weforum.org/reports/the-future-of-jobs-report-2018.

Wikipedia contributors. 2019. "Service-level agreement." Wikipedia website, July 11. Accessed October 1, 2019. https://en.wikipedia.org/w/index.php?title=Service-level_agreement&oldid=905761301.