



AI customer service and contact centers

Combining self-service, agent assist, and quality management responsibly.

WORKS WITH THE ASSISTANTS TEAMS ALREADY USE



EDITION 09 / EMBER & ICE

EXECUTIVE WORKING EDITION - 08 OCT 2026

09

AT A GLANCE

Combining self-service, agent assist, and quality management responsibly.



Customer service is where AI is most visible to customers and where mistakes are most costly to trust. Done well, AI resolves simple requests quickly, helps agents find the right answer, summarizes interactions, and makes quality review broader than a small random sample. Done badly, it traps customers in loops, gives wrong policy answers, and burns out agents who inherit the hard cases. This playbook covers the three layers of service AI, how to govern knowledge and escalation, and how to measure resolution rather than deflection.

WHO IT IS FOR

Heads of customer service and support, contact center operations managers, quality and training leads, knowledge managers, and CX leaders deciding how to introduce AI into service channels.

TIME TO USE

60 minutes to read; one working session to apply



What you will have at the end

- A clear map of self-service, agent assist, and quality management uses, with fit and risk for each
- A knowledge and escalation design that keeps customers from being trapped
- A metrics set that tracks resolution, customer effort, and agent experience, not only deflection
- A phased rollout plan with a human fallback and a stop rule



Contents

01	Three layers of service AI and one rule	03
02	Use cases, fit, and where humans must stay	04
03	How service AI produces an answer	06
04	Roles, knowledge ownership, and escalation design	07
05	Accuracy, vulnerability, privacy, and fairness	09
06	Measure resolution, effort, and quality, not only deflection	10
07	A phased rollout with a safety net	12
08	Common failures, an illustrative scenario, and vendor questions	14
09	A first quarter that produces evidence	15
10	Terms used in this publication	16



EXECUTIVE BRIEF

Three layers of service AI and one rule

01



AI in service works in three layers: helping customers directly, helping agents, and helping managers see quality. The rule is that a customer must always have a path to a person.

Customer service has long used automation: phone menus, scripted chat flows, and knowledge bases. Generative AI changes what is possible because it can understand free-form language, draft natural replies, and summarize long conversations. That makes it useful in three distinct places. Customer-facing self-service can answer common questions and complete simple tasks. Agent assist can surface the right article and suggest a reply while a person handles the interaction. Quality and operations tools can summarize, tag, and review interactions at a scale no supervisor can match.



These layers carry very different risk. A summary tool that drafts after-call notes for an agent to confirm is low risk. A bot that tells a customer about refund rights is making a statement on behalf of the company, with real consequences if it is wrong. Teams that succeed usually begin with the agent-facing layer, learn how the model behaves on their own knowledge, and open self-service gradually for narrow, well-understood topics.



- Self-service: answering questions and completing simple, well-defined tasks for customers
- Agent assist: surfacing knowledge, drafting replies, and summarizing for the human agent
- Quality management: reviewing far more interactions, tagging reasons for contact, and spotting coaching needs
- Operations insight: finding recurring causes of contact that product or policy teams can fix

THE RULE

Never design a flow in which a customer cannot reach a person. Make the route to a human easy, quick, and visible, and treat repeated rephrasing, frustration, or a vulnerable customer as automatic triggers for handover.



WORKING QUESTION

Which three reasons for contact would you be comfortable letting a bot handle end to end, and why those three?

WHERE IT FITS

Use cases, fit, and where humans must stay

02

Automate what is frequent, simple, and safe. Keep people on what is emotional, ambiguous, high-value, or regulated.

The reasons customers contact you follow a long tail. A handful of simple requests, such as order status, password help, opening hours, and appointment changes, make up much of the volume. At the other end are complaints, bereavement, billing disputes, fraud concerns, and vulnerable customers, which are fewer in number but heavily weighted in cost, risk, and reputation. A sensible design sends the head of the distribution to automation and keeps the tail with trained humans.

USE CASE	FIT	MAIN RISK AND CONTROL
Order status, simple account tasks via self-service	Strong	Needs secure identity checks and system integration; test the full journey
FAQ and policy answers from approved knowledge	Moderate	Wrong or stale answers. Ground in a governed knowledge base and cite it
Agent assist: suggested articles and replies	Strong	Agent over-trust. Keep agents accountable; show sources
After-call summaries and disposition coding	Strong	Missed or invented details. Agent confirms before saving
Quality review across all interactions	Strong	Unfair scoring of agents. Calibrate against human reviewers; allow appeal

USE CASE	FIT	MAIN RISK AND CONTROL
Complaints, disputes, Weak hardship, bereavement		Needs empathy and discretion. Route to trained humans
Outcomes with legal or Avoid financial effect, such as refunds beyond policy		Human approval; AI may retrieve the rule but not grant exceptions

Signals to hand over to a person

- The customer asks for a person, in any wording
- The customer repeats the same request twice or shows frustration
- Confidence in the answer is low or the knowledge base has no matching article
- The topic is on a defined sensitive list such as complaints, fraud, safety, or vulnerability
- The action would change money, a contract, or access, and policy requires approval

CHECK THE RULES FOR YOUR SECTOR

Regulators in sectors such as financial services, health, telecommunications, and utilities often set expectations for complaint handling, vulnerable customers, and disclosure of automated interaction. Check current rules for your sector and region, and ask counsel what you must tell customers about bots.

WORKING QUESTION

What share of your contacts is genuinely simple, and how do you know?

HOW IT WORKS

How service AI produces an answer

03

A trustworthy service bot is a language model working inside guardrails: governed knowledge, system access, and clear rules about when to stop.

Most modern service assistants combine a language model with retrieval from a knowledge base. When a customer asks a question, the system finds relevant articles, passes them to the model, and asks it to answer using only that material. For task completion, it also connects to back-end systems through tightly limited actions, such as checking an order or booking a slot. The model is not a database and should not be trusted to recall policy from memory.

01 Understand

The system identifies the customer's intent and any key details, and checks whether the request is in scope.

02 Retrieve

It finds the most relevant approved articles and, where allowed, relevant account data for the verified customer.

03 Answer or act

The model drafts an answer grounded in that material, or performs a limited, permitted action.

04 Check

Guardrails test for out-of-scope topics, low confidence, sensitive content, and policy conflicts, and trigger handover when needed.

05 Record

The interaction, sources used, and outcome are logged for quality review and improvement.

Knowledge is the product

Service AI is only as good as your knowledge. Duplicate articles, contradictory policies, outdated promotions, and internal jargon all leak into answers. Many teams find that the first

real result of an AI project is a knowledge clean-up. Write articles in plain, single-purpose form with an owner, an effective date, and a review cycle. Include what to say, what not to say, and when to escalate. An article that a new agent could follow without help is also an article a model can use well.



PLAIN-LANGUAGE WARNING

A model may give a confident answer when the knowledge base is silent. Configure it to say it does not know and to hand over, and test this specifically with questions it should not be able to answer.

WORKING QUESTION

If you picked ten articles at random from your knowledge base, how many would you trust an AI to quote word for word?



OPERATING MODEL

Roles, knowledge ownership, and escalation design

04



Service AI needs an owner for knowledge, an owner for conversation design, an owner for quality, and a standing route for fixes.

In a contact center, the people who know the most about customers are agents, team leaders, and quality analysts. They should shape the AI, not be surprised by it. Create a small working group that includes frontline staff and gives them real authority to flag unsafe answers and request changes. Make the cycle from flagged problem to fixed knowledge short, ideally days, so that agents see that their feedback matters and keep giving it.



ROLE	OWNS	DECIDES
Head of customer service	Outcomes, customer promise, risk appetite	Which contact types are automated

ROLE	OWNS	DECIDES
Knowledge manager	Article quality, ownership, review cycle	What the AI may cite and when content expires
Conversation designer	Flows, tone, handover paths, fallback wording	How the bot behaves and speaks
Quality and training lead	Scoring, calibration, coaching	How AI quality is measured and how agents are coached
Workforce management	Staffing and scheduling	How volume changes affect rosters
Compliance, legal, and data protection	Disclosure, recording, retention	What customers must be told and what is stored
Technology owner	Integrations, security, monitoring	Access, logging, and incident response

Design for the agent who inherits the hard cases

When AI handles simple contacts, those that reach agents are harder and more emotional on average. Plan for this. Adjust handling time targets, since the remaining mix is not comparable to last year. Give agents a good summary of what the bot already did so customers do not repeat themselves. Invest in coaching on difficult conversations. Treat agent experience as a measured outcome, because attrition among agents who feel they are cleaning up after a bad bot will wipe out any efficiency gain.

- Pass the full conversation and the customer's verified details to the agent at handover
- Show the agent which sources the AI used so they can verify quickly
- Give agents an easy button to mark an AI suggestion as wrong and say why
- Review those flags weekly in a joint meeting with knowledge and design
- Share improvements back with the team so that the loop is visible

WORKING QUESTION

When a bot hands a customer to an agent today, how much of the story does the agent see?

RISKS AND CONTROLS

Accuracy, vulnerability, privacy, and fairness

05



The risks are wrong answers with real consequences, poor handling of vulnerable customers, exposure of personal data, and opaque automated judgment about people.

Service conversations contain personal data, payment details, health information, and emotional distress. They are also recorded, transcribed, and stored. Every layer of AI adds new data flows: transcripts sent to a vendor, summaries stored in a CRM, and scoring data about agents. Map these flows early and decide what is retained, for how long, and who can see it. Mask payment card numbers and other sensitive data before they reach any model.



RISK	EXAMPLE	CONTROL
Wrong policy answer	Bot states an incorrect refund or warranty rule	Governed knowledge, citations, test set, rapid correction path
Customer trapped	No exit to a human; repeated loops	Always-available handover; frustration triggers
Vulnerable customers	Distress or confusion not recognized	Sensitive topic list, trained human routes, monitoring
Personal data	Card numbers or health details sent to a model	Redaction, minimization, vendor and retention controls
Disclosure	Customer believes they are talking to a person	State clearly that it is an automated assistant; check rules
Unfair agent scoring	Automated quality scores penalize accents or styles	Calibrate with humans, test for bias, allow appeals
Prompt injection	Customer text tricks the bot into revealing or doing something it should not	Limit permissions, validate actions, never expose secrets to the model



Controls to confirm before any customer-facing launch

- There is an obvious, working route to a human at every step

- The bot says that it is automated and what it can and cannot do
- Allowed actions are narrow, authenticated, and logged
- Sensitive topics and vulnerable-customer signals are defined and tested
- Recording, transcript retention, and vendor processing have been reviewed by data protection and legal
- A kill switch can disable the bot immediately and route all traffic to staff

Sector and national rules about customer treatment, recording, automated decision-making, and disclosure differ and change. The NIST AI Risk Management Framework and ISO/IEC 42001 offer general structures for managing AI risk, and data protection law and the EU AI Act may apply depending on where you operate and what the system does. Check current requirements with your counsel and regulators.

WORKING QUESTION

What is the worst thing your bot could say with confidence today, and how soon would you find out?

MEASURING VALUE

Measure resolution, effort, and quality, not only deflection

06

Deflection tells you how many people did not reach an agent. It does not tell you whether their problem was solved.

A bot that frustrates customers into giving up has a high deflection rate and a terrible service. Measure whether the customer got what they needed. Track contained and resolved interactions separately, and check for repeat contacts on the same issue within a few days, which is a strong sign that the first answer failed. Pair operational metrics with customer-reported measures and with agent experience.

METRIC	HOW TO MEASURE	WATCH FOR
Resolution rate	Share of contacts where the issue is solved, with no repeat within a set window	Counting abandonment as success
Repeat contact rate	Same customer and topic within 30 days	But answers that sound right but are not
Customer effort and satisfaction	Post-contact survey, split by bot, assisted, and human	Low response from unhappy customers
Handover quality	Share of handovers where the agent had full context	Customers repeating themselves
Average handling time and after-call work	Before and after, by contact mix	Mix change making comparisons unfair
Answer accuracy	Sampled interactions graded against policy	Spot checks that miss sensitive topics
Agent experience and attrition	Pulse surveys, turnover, sick days	Gains in efficiency offset by burnout

Build a test set from real contacts

Take 100 to 200 recent, anonymized contacts across your top reasons for contact, including difficult and ambiguous ones. For each, write the correct answer or correct handover decision. Run your assistant against the set before launch and after every change in model, prompt, or knowledge. Grade for accuracy, completeness, tone, and correct escalation. Make sure the set includes questions the bot should refuse or hand over, since those cases matter most. Pass marks are your choice; an illustrative one is no policy errors on the sensitive subset and correct handover in all cases that require it.

QUALITY MANAGEMENT CAUTION

AI can review every call, not just a sample, which is powerful. But it can also produce surveillance without insight. Use it to find coaching themes and process failures first, and treat individual scoring with care, human calibration, and transparency to agents.

WORKING QUESTION

How would you know today if your first-contact resolution got worse while handling time improved?

IMPLEMENTATION GUIDE

A phased rollout with a safety net

07

Start with agent assist and summaries, fix the knowledge, then open customer-facing automation on narrow topics with strong handover.

01 Understand contact demand

Analyze a few thousand contacts by reason, effort, and outcome. Identify the top reasons that are frequent, simple, and well documented.

02 Repair the knowledge

Consolidate duplicates, assign owners, and rewrite the articles behind your top reasons for contact in clear, single-purpose form.

03 Launch agent assist first

Give a pilot team article suggestions and after-call summaries. Gather feedback and measure time and accuracy.

04 Open one narrow self-service topic

Choose a safe topic, such as order status, with strong handover. Run it in a limited audience and compare against the human path.

05 Expand by evidence

Add topics only when accuracy, resolution, and satisfaction hold. Extend quality analytics once the team trusts the approach.

06 Review and rebalance

Revisit staffing, handling targets, and coaching with the new contact mix.

Operational readiness

- Agents are trained, have been consulted, and know how AI use will and will not be used in their evaluation
- Handover, kill switch, and incident procedures are tested end to end
- Contact tagging is consistent enough to measure before and after
- Peak-period behavior has been tested, including outages of connected systems
- Customer-facing wording about automation has been reviewed
- Weekly review of flagged answers and failed handovers is scheduled and owned
- Run the new flow in shadow mode first, comparing its answers with what agents actually said
- Roll out by channel and by customer segment, not all at once
- Keep a complete human path for customers who prefer it
- Communicate with staff early, since job security concerns are real and rumors travel fast
- Report results openly, including failures and what was changed as a result

WORKING QUESTION

Which single contact type would be the safest first test of customer-facing automation in your operation?

PITFALLS AND BUYING

Common failures, an illustrative scenario, and vendor questions

08

The usual causes of failure are launching before the knowledge is ready, hiding the human option, and celebrating deflection.

- Treating a bot as a headcount reduction project and cutting staff before quality is proven
- Pointing the model at an unmanaged knowledge base and discovering the contradictions through customers
- Making the human route hard to find in order to lift containment numbers
- Ignoring peak events such as outages, promotions, and billing runs, when the bot is most needed
- Measuring average handling time on a changed mix and drawing the wrong conclusions

ILLUSTRATIVE SCENARIO, NOT A REAL COMPANY

A subscription business launches a chat assistant with a deflection target. Deflection looks excellent, but repeat contacts and cancellations rise. A review shows the bot kept answering billing disputes with generic policy text. The firm adds a billing handover, rewrites the dispute article, and changes the target to resolution without repeat contact.

Questions to ask any vendor

Vendor diligence

- Which parts of a conversation are sent to which models and sub-processors, and where
- How are transcripts, recordings, and personal data retained, redacted, and deleted
- Can the assistant be restricted to our approved knowledge, and does it cite sources
- How do handover, confidence thresholds, and sensitive topic rules work and who can change them
- What logging, analytics, and audit functions are available, and can we export them
- How does the vendor handle prompt injection, abuse, and outages

- What happens to our data, flows, and knowledge structures if we leave

Insist on a trial using your own anonymized contacts, including the difficult ones. Watch how the tool behaves when the answer is not in the knowledge base, when the customer is angry, and when the question tries to push it out of scope. These moments, not the happy path in the demo, are what your customers and agents will remember.

WORKING QUESTION

Which of your current service metrics would reward a bot for behaving badly?

30 / 60 / 90

A first quarter that produces evidence

09

Fix the knowledge, help the agents first, and open customer-facing automation only where evidence and handover are strong.

First 30 days: ground it

- Analyze contact demand and pick the top simple, frequent reasons
- Name owners for knowledge, design, and quality, and form a frontline working group
- Capture baseline resolution, repeat contact, handling time, and satisfaction
- Define the sensitive topic list and handover rules

Days 31 to 60: test it

- Pilot agent assist and after-call summaries with one team
- Rewrite and own the articles behind the top contact reasons
- Build a test set of 100 to 200 anonymized contacts and run it on every change
- Brief and consult staff on how AI use affects roles and evaluation

Days 61 to 90: decide

- Run one narrow self-service topic in a limited audience with a kill switch
- Compare resolution and repeat contact against the human path
- Decide which topics to add, change, or stop
- Adjust staffing, targets, and coaching to the new contact mix

GLOSSARY

Terms used in this publication

10

TERM	MEANING
Agent assist	Tools that support a human agent during an interaction with suggested articles, draft replies, and summaries.
Containment	The share of contacts that finish in self-service without reaching a person, which is not the same as resolution.
Deflection	A reduction in contacts that reach agents, which may or may not reflect solved problems.
Disposition code	A label an agent or system assigns at the end of a contact to record its reason and outcome.
Escalation	Passing a contact to a more skilled person or a different process when the first level cannot or should not resolve it.
First contact resolution	The share of issues solved in the first interaction without the customer needing to return.
Guardrail	A rule or technical check that limits what an AI assistant may say or do.
Handover	The transfer of a customer from an automated assistant to a person, ideally with full context.
Knowledge base	The managed set of articles that agents and assistants use to answer customers.
Prompt injection	An attempt to make an AI system ignore its instructions by placing manipulative text in its input.

TERM	MEANING
Quality management	The practice of reviewing interactions against standards to coach agents and improve the process.
Shadow mode	Running a new system alongside the live process without affecting customers, in order to compare its outputs.



A note on use

Free assessments are available at halden.company/assessments, and you can reach the team at halden.company/contact. Figures here are illustrative, and nothing in this playbook is legal advice.

TAKE THE NEXT STEP

Take the free Halden assessment at halden.company/assessments, or talk to us at halden.company/contact about turning this into a plan for your organization.

