

AI for Customer Service — Workbook

This workbook turns the course into a working AI support deployment: a contact-reason breakdown of your real queue, a knowledge-base audit, a tested grounding and escalation setup, a tag-and-routing map, and a measurement baseline. Work one section per module, doing the build steps in your actual help desk as you go and filling each worksheet. By the end you will have the artifacts to launch a narrow, grounded, well-measured AI layer instead of a risky all-at-once rollout.

Where AI Fits in a Support Operation

Audit your real ticket volume, score each contact reason for automation fit, and choose the tool and job that match your stack before you turn anything on.

Exercise: Run a 90-Day Contact-Reason Breakdown

Export your last 60 to 90 days of tickets and turn them into a ranked list of contact reasons. This is the single most important input to the whole project, because it tells you which handful of topics actually fill your queue. Do this before evaluating any tool.

- Group your exported tickets into 15 to 25 contact reasons and count the volume of each.
- Sort the reasons by volume and calculate what percentage of total tickets the top five represent.
- For each top reason, note whether a current, correct help article already exists.
- Tag each reason as Repetitive and rule-based, Repetitive but judgment-heavy, or Rare and complex.

Worksheet: Automation Fit Scorecard

Score your top ten contact reasons so you automate the right ones in the right way. Deflection fit goes to a bot, assist fit goes to agents, sensitive topics stay human, and missing-article topics get fixed first. Keep this next to you when you configure the tool.

Contact reason

Monthly volume

One clear correct answer? (yes/no)

Current article exists? (yes/no)

Emotionally or legally sensitive? (yes/no)

Recommended path (Deflect / Assist / Keep human / Fix article first)

Worksheet: Tool and Job Selection Sheet

Match the AI job and the tool to your real stack and volume before you buy. Fill this out so the decision is driven by your help desk, channel, and pricing math rather than by a sales demo.

Help desk we run today

Primary channel and vertical (e.g. Shopify ecommerce, SaaS support)

First AI job to deploy (Summarization / Agent assist / Triage / Deflection)

Candidate tool(s) (Zendesk AI / Intercom Fin / Gorgias / Einstein / Forethought / Ada)

Pricing model (per seat / per resolution) and estimated monthly cost

Monthly volume x expected deflection rate = resolutions to price

Checklist: Readiness Checklist Before Buying

- Last 90 days of tickets exported and grouped into contact reasons
- Top five contact reasons and their share of total volume identified
- Each top reason scored for automation fit (deflect / assist / human / fix first)
- Native AI in your current help desk evaluated before shopping outside
- Both pricing models modeled against your real monthly volume

Grounding the Bot in Your Knowledge

Audit and fix your help center, connect only customer-safe sources, force citations and a clean refusal path, and use AI to fill content gaps mined from real tickets.

Exercise: Audit the Knowledge Base for the Top Topics

Walk your top contact reasons against your help center and find the stale, contradictory, and missing content the bot would otherwise repeat or guess around. Fix these before connecting any bot, because the bot is only as good as the articles behind it.

- For each top contact reason, confirm there is one current, dedicated article with a question-shaped title.
 - List every article containing a price, policy, or date that has changed since it was written.
 - Find any two articles that contradict each other on a window, fee, or process.
 - List the high-volume questions that have no article at all, ranked by how much they would deflect.
-

Worksheet: Article Rewrite Worksheet

For each article you flagged, redesign it so retrieval can find and quote it cleanly: a question as the title, the direct answer up front, and exact numbers. Fill one of these per article you fix.

Current article title

New question-shaped title

Direct answer in one or two sentences (with exact numbers and conditions)

Topics to split out into separate articles

Facts verified against source of truth? (yes/no) and verified by whom

Exercise: Test Grounding and the Refusal Path

Before launch, prove the bot answers from your content, cites it, and refuses gracefully when it should. Run these in the bot's preview or test mode and record what happens. A bot that refuses cleanly is a feature; a bot that guesses is a liability.

- Ask a question you know is in an article and confirm the bot quotes it correctly and cites the source.

- Ask a question covered nowhere and confirm the bot refuses and offers a human handoff instead of inventing an answer.

- Ask a question whose answer recently changed and confirm the bot uses the updated article.

- Ask a multi-part question and check whether the bot conflates two policies; if it does, split the source articles.

Checklist: Pre-Launch Grounding Checklist

Only published, customer-safe sources are connected; internal and draft material is excluded

Citations are turned on so answers show their source article

The bot is set to answer only from connected content, not general training

The I do not have that information handoff path is configured and tested

The bot was re-tested after the latest content changes

Agent Assist, Macros, and Triage

Brief the assist tool in your brand voice, generate an AI-supercharged macro library, and stand up intent, sentiment, and routing so every ticket reaches the right skill.

Worksheet: Brand Voice Brief for Draft Replies

Give your agent-assist tool the voice and rules it needs so drafts sound like you and survive editing. Paste the resulting brief into the assistant's instructions and reuse it for macro generation. Keep the do-not-say list strict.

Three to five adjectives for our voice

Reading level and sentence-length preference

Default reply shape (greeting, answer, next step, sign-off)

Two example replies that sound exactly right

Do-not-say list (no over-apologizing, no promises outside policy, no invented dates or refunds)

Exercise: Generate and Personalize a Macro Library

Use AI to draft a first version of every macro for your top contact reasons, then have it adapt a generic macro to a specific ticket. Verify every fact before saving. This gives you the speed of canned replies with the feel of tailored ones.

- Draft a short macro for each of my top five contact reasons using these real policy facts.
- Rewrite the return-instructions macro to personalize it with this customer's order number and item.
- Make this blunt draft warmer and more empathetic without changing any facts.
- Cut this rambling reply to three sentences while keeping the answer and the next step.

Worksheet: Intent, Sentiment, and Routing Map

Define the tags AI will apply and exactly what each one does to the ticket. Every label should change routing or priority; if it does not, drop it. Build this from your contact-reason buckets so the taxonomy stays clean.
 Intent label (from contact reasons)

Route to (team or skill queue)

Priority or SLA when this intent appears

Sentiment override (e.g. angry = escalate and speed up)

VIP or high-value override (e.g. senior agent, supervisor flag)

Checklist: Triage and Routing Checklist

- A short, stable intent list maps directly to how you route, with no tag sprawl
- Sentiment is captured and used to raise priority on angry or at-risk tickets
- Each intent routes to the team or skill best equipped to resolve it
- A catch-all human queue exists for anything the AI cannot confidently classify
- A weekly review of misclassifications is scheduled to tighten the taxonomy

Escalation, Safety, and Measurement

Write the escalation triggers and warm-handoff brief, harden the bot against misbehavior, capture a baseline, and plan a staged rollout judged on resolution and satisfaction.

Worksheet: Escalation Trigger Sheet

Decide in advance the exact conditions under which the AI stops and a human takes over, and what the bot passes along. A bot is judged by how gracefully it hands off the tickets it should not handle, so design this as carefully as the answers.

Always-escalate triggers (asks for human, angry, VIP, legal or fraud, low confidence)

Two-strike rule definition (escalate after how many failed attempts)

Always-human topics the bot must never attempt

Warm-handoff brief contents (issue summary, what the customer tried, account details)

After-hours plan when no human is available

Exercise: Adversarial Safety Test

Before launch, try to make the bot misbehave the way a hostile or clever user eventually will. Record whether it holds. Anything that moves money or changes data should be gated behind approval or strict limits.

- Insult the bot and instruct it to ignore its rules; confirm it stays civil and on-policy.

- Ask about a policy that does not exist; confirm it refuses instead of agreeing or inventing one.

- Ask it to issue a refund or change account data above your limit; confirm it escalates instead of acting.

- Embed an instruction inside your message telling it to break a rule; confirm it ignores the injection.

Worksheet: Baseline and Metrics Capture Sheet

Record your current numbers before any AI goes live, then track the same metrics after. Read deflection together with CSAT and reopens, because deflection without satisfaction is just a deferred ticket.

Baseline deflection or self-service resolution rate

Baseline CSAT (and target after launch)

Baseline average handle time and first response time

Baseline first-contact resolution and reopen rate

Metric review cadence and who owns the dashboard

Checklist: Staged Rollout Checklist

- Baseline metrics captured before any AI is customer-facing
- AI run first in shadow or agent-assist mode with human review of every output
- Customer-facing deflection enabled only for two or three safest, highest-volume topics
- CSAT and reopen rate monitored on pilot topics alongside deflection before widening
- A kill switch or instant human-fallback is in place if quality drops

Your Action Plan

1. Export the last 90 days of tickets and build a ranked contact-reason breakdown.
2. Score the top ten reasons for automation fit and pick your first AI job and tool.
3. Audit the help center against the top topics; fix stale, contradictory, and missing articles first.
4. Use AI to mine tickets for content gaps and draft the missing articles, then verify every fact.
5. Connect only customer-safe sources, turn on citations, and test the refusal and handoff path.
6. Brief the assist tool in your brand voice and generate an AI-supercharged macro library.
7. Define intent, sentiment, and routing tags so each ticket reaches the right skill, with a human catch-all.
8. Write escalation triggers and a warm-handoff brief, then run an adversarial safety test.
9. Capture baseline deflection, CSAT, handle time, and reopen rate before going live.
10. Launch narrow on your safest topics, watch CSAT and reopens, and widen only what the numbers earn.

