§ THE PLAYBOOK SERIES · FREE · NO FORM, NO FOLLOW-UP SEQUENCE

The AI playbook for *Customer services*

How to automate the routine without automating the moment your customer needed a human.

Customer service is where AI headcount stories are made and unmade: Klarna automated the work of 700 agents, then rehired humans within eighteen months. This playbook covers what resolution rates really look like, which conversations should never meet a bot, what your BPO contract is now worth, and who legally owns what the chatbot says.

> *A Canadian tribunal ruled that Air Canada owned its chatbot's invented bereavement policy and had to honour it. Your bot's hallucinations are your promises now. Plan accordingly.*

700 → rehiring

Klarna's arc: agents' work automated in month one, humans hired back within 18 months.

BLOOMBERG / FORBES, 2024-25

41% vs 80%

real enterprise median automated resolution vs vendor headline claims.

ZENDESK BENCHMARKS, 2026

64%

of customers would prefer companies didn't use AI for service at all; 53% would consider switching.

GARTNER SURVEY, 2024

INSIDE THIS PLAYBOOK

PART ONE

What actually *changes.*

Klarna is the whole playbook in one company. February 2024: the AI assistant handles 2.3 million chats in its first month, does 'the work of 700 agents', projects $40m of profit improvement. Eighteen months later the CEO is on record - 'we focused too much on efficiency and cost... the result was lower quality' - and Klarna is rehiring humans. The AI was genuinely good at the routine two-thirds. It was deployed on all of it.

The durable pattern underneath: AI fully handles simple, high-volume queries - order status, passwords, billing lookups - and it is now doing about 30% of service cases across the market. Agent-assist delivers the safest returns. And full automation reliably fails on disputes, fraud and hardship: exactly the conversations that decide whether a customer stays.

01

The routine tier genuinely automates

Order status, password resets, standard billing. Salesforce data has AI at ~30% of cases now, heading for half by 2027. This tier is real, measurable, and where all the credible savings live.

02

Agent-assist is the quiet winner

Summaries, suggested replies, auto-wrap-up, live guidance. Less exciting than a bot with a name, and it is where the CSAT holds: the human stays in the conversation, just faster.

03

Voice arrives properly

Sierra stood up a 70%-resolution voice agent for Singtel in ten weeks and now serves 40% of the Fortune 50. Voice automation is no longer the weak sibling - which makes choosing what it should not handle more urgent, not less.

04

Escalation design becomes the product

Zendesk's data shows the AI-to-human CSAT gap nearly vanishing when handover is clean. The bot is not your experience; the moment it hands over is.

05

The honest number is 41%

Vendor claims run 70-80% resolution; the enterprise median is 41.2%, and deflection is not resolution. Whatever you are quoted, ask how 'resolved' is defined and watch the number change.

06

Complex and emotional stays human

Disputes, fraud, bereavement, hardship. This is what broke Klarna's quality scores and what the 64% of AI-averse customers are really talking about. Route it to people, visibly and quickly.

PART TWO - THE HONEST BIT

What it costs *you.*

The headcount reality is messier than the headlines in both directions. Salesforce really did take support from 9,000 to 5,000. But only about 20% of service leaders have actually reduced staffing, Gartner predicts half of those who cut citing AI will rehire under new titles by 2027, and Commonwealth Bank of Australia had to reverse 45 redundancies and apologise when the union showed call volumes were rising. The cuts that stick are the ones made after the resolution numbers were proven - not before.

The commercial shift is just as sharp. A human interaction costs $6-12; an AI one under $1.50. Hedge funds are shorting the BPO industry on that arithmetic, and vendors are hedging their own model - Zendesk now charges $1.50 per automated resolution. Every service contract you hold was priced for a labour model that is ending.

01

L1 shrinks, escalation hardens

The routine tier gets smaller; everything reaching a human is harder and more emotional. The remaining job is more skilled and more draining - staff it, pay it and support it as such, or churn will do your workforce planning for you.

02

New roles appear whether you plan them or not

Conversation designers, bot-transcript QA, escalation specialists, AI supervisors. Salesforce found 81% of reps saying their roles got more specialised. The org chart moves; better it moves on purpose.

03

Your BPO contract is mispriced

Multi-year, per-FTE pricing against sub-$1.50 AI economics. Reopen it: outcome pricing, automation gainshare, quality floors. Your provider has already run this spreadsheet - the only question is which side of it you are on.

04

You own what the bot says

Air Canada: liable for an invented policy. DPD: bot swearing at customers, viral. Cursor: hallucinated policy, cancelled subscriptions, public apology. Legal, brand and service now share custody of every automated sentence.

05

Customers did not ask for this

64% would rather you didn't; 53% would consider leaving. That is not a reason to stop - it is a reason to automate invisibly well and keep the path to a human short and unhidden.

06

The unit economics can turn on you

Gartner warns AI cost-per-resolution may exceed $3 by 2030 - dearer than offshore humans. Model token and orchestration costs per resolution now, or rediscover them at renewal.

PART THREE

What good looks *like.*

Good customer service is a deliberate split: machines resolving the routine tier at proven rates, humans owning the conversations where trust is won or lost, and a handover between them so clean customers barely feel the seam. Quality is measured in CSAT and retention, deflection is a cost line, and no contract - vendor or BPO - is priced on the old labour model.

01

Automation with an honest scorecard

Resolution measured end-to-end, CSAT tracked separately for bot and human contacts, and the routine tier automated only as fast as quality holds. The board sees the same numbers the floor does.

02

Humans where humans are the point

Disputes, fraud, hardship and anger routed to skilled people fast, with the bot's context attached. The escalation team is senior, supported and staffed for difficulty, not volume.

03

Guardrails on every automated sentence

Grounded answers, tested policies, no invented refunds, and an owner for what the bot is allowed to promise. The Air Canada ruling taped to the wall, metaphorically or otherwise.

04

Contracts priced for the new economics

Vendor deals on outcome pricing with resolution honestly defined; BPO agreements rebuilt around automation gainshare and quality. Nobody pays human rates for machine work - in either direction.

THE CHANGE, ACROSS ALL FIVE PILLARS

PEOPLE	PROCESSES	DATA	PLATFORMS	PARTNERS
Escalation roles upgraded - senior pay and real support for the hardest conversations, all day, every day.	Handover designed as the product, and the bot's boundary decided with legal and brand in the room.	Grounded knowledge and transcript QA behind every automated answer the company is now liable for.	Vendor deals on outcome pricing, with 'resolved' defined by you rather than the sales deck.	BPO contracts rebuilt around automation gainshare and quality floors before the margin quietly moves.

PART FOUR

The first *90 days.*

The order matters more here than anywhere: prove resolution quality first, touch headcount last. Klarna and CBA ran the other sequence so you don't have to.

I

WEEKS 1-2

Map contacts by what they really are

Volume, cost and emotional stakes per contact type. Draw the line between routine and human-required - with legal and brand in the room for the bot-boundary decision.

II

WEEKS 3-6

Automate one routine contact type

Grounded knowledge, tested answers, clean escalation, approval gates on anything the bot can promise. Agent-assist for everyone else at the same time - it is the safer half of the value.

III

WEEKS 7-10

Hold it against real measures

End-to-end resolution, bot CSAT vs human CSAT, escalation quality, cost per resolution including tokens. If quality holds, widen. If not, you just saved yourself a public reversal.

IV

WEEKS 11-13

Then - and only then - the structure

Resize the routine tier on proven numbers, upgrade the escalation roles, open the BPO renegotiation, and tell the team the honest plan including what you will not automate.

THE DAY-90 SCORECARD · WHAT YOU SHOULD BE ABLE TO SHOW

→ End-to-end resolution for automated contacts - your definition, not the vendor's

→ Bot CSAT versus human CSAT, with transcripts actually reviewed

→ Time for a frustrated customer to reach a human

→ Cost per resolution including tokens and orchestration, not just licences

→ The never-automate list agreed in writing with legal and brand

PART FIVE

The traps - and six questions worth an *honest hour.*

01

Cutting on the projection

CBA cut 45 roles on a bot's promise while call volumes rose, then reversed and apologised. Headcount moves last, on measured resolution, or it moves back - expensively and publicly.

02

Worshipping deflection

A deflected customer is not a resolved one; sometimes they are just a customer who gave up. Measure whether the problem ended, not whether the ticket did.

03

Letting the bot near policy

Bereavement fares, refunds, account restrictions - anywhere an invented answer becomes a legal promise. Ground it, gate it, and give legal a veto they can actually use.

04

Hiding the humans

Customers who want a person and cannot find one become the 53% who leave. A short, visible path to a human is the cheapest retention programme you will ever run.

05

Leaving the BPO contract sacred

Your provider's costs are collapsing inside a contract priced on FTEs. Every quarter you do not reopen it is margin you donated.

06

Ignoring what the job becomes

The humans who remain get the hardest conversations all day, every day. Without senior pay, real support and career paths, your quality plan quietly walks out the door.

THE SELF-ASSESSMENT · MORE THAN TWO UNCOMFORTABLE ANSWERS IS AN AGENDA, NOT A CRISIS

→ What is your true end-to-end resolution rate for automated contacts - by your definition or the vendor's?

→ Which contact types have you formally decided a bot must never handle?

→ How does bot CSAT compare with human CSAT - and who reviews the transcripts?

→ What would the Air Canada ruling cost you if it were applied to your bot yesterday?

→ When does your BPO contract next open, and what is your automation position for it?

→ Could a frustrated customer reach a human in under two minutes if they needed to?

START WITH THE WORKSHOP

Book your *Catalyst Workshop.*

1-2 executive days to turn this playbook into your first move, the case for it, and a 90-day mobilisation plan. hello@veriteer.com

HELLO@VERITEER.COM →