



# CX Leader's Guide to AI

A PRACTICAL FRAMEWORK FOR BUILDING  
AUTONOMOUS CUSTOMER SUPPORT

# Table of Contents

Introduction	<b>3</b>
The Business Case for AI in Support	<b>5</b>
7 Questions Every Leader Should Ask	<b>9</b>
Building Your AI Roadmap	<b>15</b>
Case Study 1: ActiveCampaign	<b>19</b>
Case Study 2: Grammarly	<b>21</b>
Case Study 3: Cotopaxi	<b>23</b>
Case Study 4: YAZIO	<b>25</b>
Final Checklist: How to Use AI the Right Way	<b>27</b>

# Introduction

If you've been evaluating AI vendors, the demos have probably started to blur together. Most of them follow the same script: a chatbot answers customer questions, pulls articles from your knowledge base, and reports impressive deflection rates. The interfaces are clean, the responses are fast, and the pitch decks are full of promises about what's coming next quarter, but there's a problem these demos don't address directly.

95%

of generative AI pilots **never make it to production** because answering a customer's question is not the same as resolving their issue.



[Source](#)

Consider a customer who wants a refund. Traditional AI that's solely generative or RAG-based can explain your refund policy, confirm the customer meets the criteria, and even express sympathy for their frustration. But when that conversation ends, the customer still has to email your support team, wait for a human to review the request, and follow up if nothing happens. In this case, AI didn't really automate anything—it just added a step.

Agentic AI operates differently. Instead of describing how to get a refund, it processes the refund by connecting to your billing system, verifying eligibility, executing the transaction, and confirming completion. And it's how some companies are scaling support without adding headcount, while others are drowning in tickets despite having "implemented AI." This guide is designed to help you make the right choice the first time.

# How to use this guide

Read this guide before evaluating vendors and planning an implementation. It gives you some questions to ask, steps to follow, and benchmarks that tell you what realistic results look like.

The frameworks here come from what we've learned processing over 1.4B customer interactions monthly. [Forethought](#) launched in 2018 as the first AI-native enterprise support platform, built from day one with a multi-agent architecture designed for support at every customer touchpoint.

## **Our AI agents:**

- Autonomously resolve tickets across chat, email, and voice
- Use an agentic reasoning engine that executes workflows
- Completes intelligent routing
- Acts as a copilot for agents
- Identify gaps and generate content

This guide draws from Forethought's experience processing over a billion customer interactions monthly across companies like Grammarly, ActiveCampaign, and Cotopaxi.



# The Business Case for AI in Support

# The Business Case for AI in Support

You've heard the broad claims about AI transforming business, and you've probably felt pressure from leadership to adopt it, but there's actual evidence behind the hype. According to [BCG's research on AI adoption](#), companies that successfully implemented AI achieved 1.5x higher revenue growth over the past three years, 1.6x higher shareholder returns, and 1.4x higher returns on invested capital than companies that haven't.

Looking ahead, successful adopters expect 60% higher AI-driven revenue growth by 2027 and nearly 50% greater cost reductions.

The key word here is "successfully." Most companies that adopt AI don't see these results—BCG found that 74% struggle to achieve and scale value. These results are real, but they're not automatic. The companies that achieved them made specific choices about how to evaluate vendors, sequence their implementation, and prepare their teams. The rest of this guide covers how to make those choices correctly.

## Cost savings

Every support team has a category of tickets that drain agent time without adding value: things like password resets, order status checks, return initiations, and account updates. These are tasks with clear steps and fixed rules, making them perfect candidates for automation.

The question is whether your AI can actually handle them end-to-end. According to [Forethought's 2025 AI in CX Benchmark Report](#), companies using agentic AI achieve 44% deflection compared to 33% for non-agentic solutions, and 72% report actual cost reductions.

Agentic AI makes a significant difference in cost per resolution: \$15 for agentic AI versus \$18-19 for non-agentic.

That gap exists because agentic AI completes the task while non-agentic AI just explains how to complete it. When a customer asks for a password reset, and the AI walks them through the steps but doesn't actually reset the password, the customer either figures it out themselves or comes back later, and the ticket eventually lands in your queue anyway. Real cost savings come from issues that never reach your agents at all.

The savings show up in headcount math. When AI actually resolves issues rather than just acknowledging them, your business volume can double while your team size stays flat.

Final Draft's support team of five, for example, handles their entire customer base because AI resolves 71% of chat and 70% of email volume. Without Forethought, they estimated they would have needed to double the department.

## Customer experience

Customer patience is disappearing faster than most teams realize. The window to meet customer expectations is shrinking, and traditional support models that depend on agent availability can't keep up. According to [Forethought's 2025 Holiday Report](#), 57% of customers now refuse to wait longer than 10 minutes for service, up from 50% in 2024.

20% of those customers won't even wait more than five minutes for service, up from 14% last year.

AI changes the math entirely because there's no queue, wait time, or "we'll get back to you within 24 hours." Customers get responses instantly, but speed alone isn't enough. Forty-four percent of consumers say AI has made service faster, yet satisfaction only improves when that speed leads to actual resolution, not a handoff.

Companies using an agentic AI solution, where the issue is resolved in the same interaction, report a [64% improvement in CSAT](#) because CSAT doesn't just reflect speed; it also reflects whether a customer's problem was actually solved.

## Agent experience

The same repetitive questions that frustrate customers are burning out your agents. That stress comes from answering the same questions hundreds of times a day while toggling between tabs to find information that should be at their fingertips.

According to [research from Cornell](#), 87% of contact center workers report “high” or “very high” stress levels on the job.

AI that actually resolves routine issues removes that burden entirely. [Cotopaxi, for example, saved over 2,000 hours of agent time](#) by letting agentic AI handle the repetitive volume. Tools that surface relevant knowledge articles, past tickets, and macros without agents hunting through systems change the day-to-day reality of the job.

As one team described it, Assist “helps put more information in agents’ hands, which has been incredibly time-saving, useful, and a big morale booster.” New agents ramp faster because the answers are right there, while senior agents avoid burnout because they’re solving problems, not searching for information.



# The 7 Questions Every Leader Should Ask

# The 7 Questions Every Leader Should Ask

Every vendor demo looks impressive in a controlled environment. The real question is which solution will actually work for yours. These seven questions are designed to separate vendors who can deliver on the promise of agentic AI from vendors who can't. Ask them in every evaluation, demand specific answers, and test what you hear.

## 1. What problems are we solving?

The instinct when evaluating AI is to start scheduling demos, comparing features, and reading case studies. But if you don't know exactly what's broken in your operation, you'll end up evaluating vendors against each other rather than against the problem you actually need to solve. Vendors will show you what their product does well, and you'll be impressed, but you won't know whether any of it addresses what's hurting your team today.

Before you talk to anyone, write down the specific metric that has to move for this investment to be worth it. Stay away from initial goals like "improve customer experience," which are likely too broad to measure or optimize against. Something like "reduce average handle time by 30% on password resets" or "deflect 40% of order status inquiries without human touch" will set you up for success. If you can't articulate the problem in one sentence, you're not ready to buy.

### **Questions to clarify your problem:**

- Are agents drowning in repetitive tickets, burning them out?
- Is response time slipping because volume has outpaced hiring?
- Are customers abandoning after waiting too long?
- Is cost per resolution climbing as ticket volume increases?
- What single metric has to improve for this to be a success?

## 2. Is our knowledge base ready?

AI trained on company-specific data performs dramatically better than generic models, but only when that data is accurate, complete, and current. “Garbage in, garbage out” applies here more than anywhere, and outdated articles, aspirational policies that don’t reflect how your team actually operates, or gaps that agents work around with tribal knowledge will produce an AI that confidently gives wrong answers.

Before talking to vendors, walk through your knowledge base honestly. A useful test is whether a new agent could resolve your most common issues using only your documentation. If they need to ask a colleague, your AI will hit the same walls.

If your knowledge base isn’t ready, that’s not a reason to delay indefinitely, but you’ll need a plan to fix your foundation. Some vendors have solutions to generate knowledge articles and policies to quickly close gaps, especially for high-volume issues where documentation is thin.

### **Questions to assess your knowledge base:**

- Are articles current, or are agents working around outdated information?
- Do they reflect actual policies or aspirational ones?
- Do they cover questions customers actually ask, or questions you wish they’d ask?
- Where are agents creating workarounds instead of following documentation?
- Could a new hire resolve common issues using only what’s written down?

## 3. How well will this integrate with our existing tools?

The difference between AI that resolves issues and AI that just talks about them comes down to system access. If your AI can’t write to your billing platform, it can’t process a refund. If it can’t connect to your order management system, it can’t cancel an order. Listing “Salesforce integration” on a vendor’s website means nothing if your refund logic actually lives in Stripe.

Before evaluating solutions, map every system the AI will need to touch like helpdesk, CRM, billing, order management, shipping, or inventory. Then list the specific actions you want the AI to take, identify which system stores that data, and which API enables changes. Doing this mapping before vendor conversations will save time on both sides. Without clarity on what you need, you’ll end up comparing feature lists instead of evaluating whether a solution can actually work in your environment.

### **Questions to map your integration landscape:**

- Which systems hold the data AI needs to read and write?
- For each action you want automated, which API enables that change?
- Do you have current API documentation for these systems?
- Do you control API access, or does IT need to provision it?
- How long does it typically take your team to set up new API connections?
- Can you show vendors your actual system architecture and ask if they can work with it?

## **4. How does it handle compliance, privacy, and security?**

When AI executes actions autonomously, trust becomes a technical requirement. Audit trails aren't optional if you're operating under GDPR, HIPAA, or SOC 2, and the vendor's compliance posture needs to match your obligations.

Ask vendors to show you their certifications today, including a current SOC 2 Type II, ISO 27001, and a Business Associate Agreement if you need HIPAA compliance.

The same standard applies to data handling and monitoring. You should be able to get specific answers about where customer data lives, who can access it, how long it's retained, and what happens when AI performance drifts from acceptable thresholds. A vendor ready for production should be able to show you an actual audit log from a completed transaction.

### **Questions to evaluate compliance and security:**

- Can you show current SOC 2 Type II, ISO 27001, and BAA certifications?
- Where does customer data live, and who can access it?
- What are your data retention policies?
- What happens when AI performance drifts? Are there automated alerts?
- Can you show me an actual audit log from a completed transaction?

## 5. Will it scale with us?

Most teams don't deploy AI everywhere at once. A typical path is to start with chat deflection, add agent assist tools once that's working, then expand to email and eventually to voice. That sequencing makes sense, but if you buy point solutions for each use case, you end up managing multiple systems with different underlying intelligence.

Scaling with AI means eventually expanding across your entire operation, so you want to look for platforms that cover multiple use cases from the same foundation. When a customer switches channels, the system should retain all information from the previous interaction. If expanding to a new channel means a new contract, a separate implementation, or a different product entirely, that's a signal you'll hit friction as you try to grow.

### Questions to evaluate scalability:

- Does the platform cover chat, email, voice, agent assist, and routing?
- If a customer switches channels, does the system retain full context?
- Can you show me the same issue resolved across three different channels?
- Can we add capabilities after launch, or does that require a separate implementation?
- What does expanding to a new channel actually involve?

## 6. How does it balance automation and human handoffs?

AI that executes autonomously still needs to know its limits. Routine requests should be resolved without human involvement, but complex or ambiguous situations need to reach your team, and the system should make that handoff seamlessly.

The key question is who controls those escalation thresholds. You want the ability to define what counts as routine versus complex based on your own risk tolerance, and to adjust those rules after launch without retraining the entire model. If a vendor describes escalation logic as something "the AI learns over time" without showing you the controls, that's a system you can't govern.

### **Questions to evaluate handoff controls:**

- How do we define what routine is versus what needs human escalation?
- What does the agent see when AI hands off a ticket?
- Does the system retain full context across channels during escalation?
- Can we adjust escalation rules after launch without retraining the model?
- Can you show me the interface for setting these guardrails?

## **7. How do we measure success?**

Define the metrics you care about before talking to vendors, and then ask whether the vendor can actually report on them at the level of granularity you need. Generic dashboards showing overall deflection are fine for executive summaries, but useless for optimization. If password resets deflect at 80% and account upgrades deflect at 20%, you need to see both separately to know where to focus.

Pricing model matters here, too. Outcome-based pricing (paying for resolved tickets rather than seats) aligns the vendor's success with yours. It's also a signal of confidence: vendors who believe their AI works are willing to tie their revenue to results.

### **Questions to evaluate measurement and pricing:**

- Can you show me reporting broken down by topic, channel, and issue type?
- How granular can the dashboards get?
- What's your pricing model? Seats, usage, outcomes, or something else?
- If outcome-based, how do you define a resolved ticket?
- Can I see a sample report from a current customer?



# Building Your AI Roadmap

# Building Your AI Roadmap

Teams that succeed with AI prove value fast, then scale systematically. The three-phase approach that follows balances speed with sustainability. Each phase builds on the last and compounds returns over time.

## Phase 1: Quick Wins

Start with AI that answers questions intelligently, even before it can take actions. For most companies, email or chat are the highest-volume channels and the right place to begin. Generative AI can help customers self-serve on “how do I...” questions, reducing ticket volume even when the AI isn’t executing tasks like processing returns or updating accounts. This is also a lower-risk way to start because if the AI surfaces a wrong article, the customer simply asks for an agent, and there’s no failed transaction or broken workflow to recover from.

Forethought Solve is an AI agent that understands customer intent and resolves issues across several channels. In Phase 1, you could deploy it in chat without autoflows. It would start by handling customer questions and, later, could be connected to systems that could take action.

The goal here is to prove that AI understands your content and that customers trust its answers to build the confidence to move into more complex automation.

## Phase 2: Scale

This phase runs two parallel tracks: making your agents more efficient now, while building the infrastructure that enables autonomous AI later. The sequence matters because agent efficiency tools deliver immediate ROI, giving your organization time and breathing room, while you do the integration work that Phase 3 requires.

### Agent efficiency

The first track focuses on reducing the friction your agents face today using tools that surface relevant information automatically so they don’t have to hunt through systems while customers wait. It also means implementing intelligent routing that classifies and prioritizes tickets without manual tagging. Lastly, it means gaining access to analytics that identify gaps in your knowledge base before they turn into ticket spikes.

Forethought's products for this track include Assist, an AI copilot that works inside your helpdesk and surfaces context the moment an agent opens a case; Triage, which automatically classifies tickets and routes them based on historical patterns; and Discover, which analyzes support interactions to find knowledge gaps and auto-generate article suggestions.

## Infrastructure

The second track is building the connections that enable AI to take action. This means working with IT and security to connect your CRM, billing platform, order management, and other systems by setting up API keys, testing connections, and defining what actions AI will be allowed to take. This creates a foundation so that when Phase 3 begins, the pipes are in place.

## Phase 3: Agentic AI

With infrastructure in place from Phase 2, AI can move from answering questions to resolving issues. Now, if a customer requests a password reset, AI can reset it. The work is actually done and creates an accurate deflection because the ticket never reaches an agent.

### Autonomous workflows

The core capability in this phase is agentic AI that executes step-by-step processes inside your connected systems. Forethought Autoflows is the reasoning engine that powers within Solve: it determines what the customer needs, identifies the required steps, calls the appropriate APIs, and completes the transaction. The same Autoflows power every channel, so when a customer switches from chat to email to phone, context is maintained throughout.

### Channel expansion

Now, with Autoflows connected, you can expand to additional channels knowing they'll have the same resolution capability from day one. Forethought Solve works across chat, email, and voice, powered by the same Autoflows, so customers get consistent experiences regardless of how they reach you, and context is maintained when they switch mid-issue.

## Continuous improvement

As AI takes on more responsibility, visibility becomes even more critical. Tools that automatically evaluate conversations against your own metrics for accuracy, empathy, and resolution quality create a feedback loop that catches problems early. Forethought's Discover includes Agent QA for this purpose, evaluating 100% of conversations rather than a manual sample.

Most teams reach this phase six to twelve months after Phase 1. That timeline reflects the work of proving AI can handle your content, building organizational trust, and establishing the integration infrastructure that autonomous execution requires.



ActiveCampaign >

# Case Study 1: ActiveCampaign

# Case Study 1: ActiveCampaign



**“A lot of what we saw felt like signing onto a roadmap instead of a real product. Once we started using Autoflows, the way we thought about automation shifted. We weren’t just clearing tickets—we were building flows that looked more like product features.”**

**— Phil Lynch, Senior AI Program Manager, ActiveCampaign**

## What they prioritized

ActiveCampaign’s previous chatbot relied on rigid decision trees that required constant manual updates. When they evaluated replacements, they focused on finding a working product rather than a vendor selling a roadmap. They also wanted AI that could serve internal teams by meeting employees where they already worked.

## Mistakes they avoided

Rather than limiting AI to external support, ActiveCampaign deployed internally from the start, using Solve in Slack to handle employee questions. They also treated AI as an ongoing program rather than a one-time implementation, building a dedicated three-person team to optimize performance continuously.

## How Forethought helped

ActiveCampaign deployed Solve for chat in their help center and implemented Autoflows to resolve issues end-to-end. They rolled out Assist across Zendesk and Slack, which handled over 1,000 internal chats per week, so agents had immediate access to context and history.

## Results

- 60%+ chat deflection rate
- 40% fewer tickets reaching agents despite business growth
- 1,300 internal chats per week, most resolved under 2 minutes
- 5+ full workdays saved weekly



# Case Study 2: Grammarly

# Case Study 2: Grammarly



**“When you buy Forethought, you’re not just buying a tool—you’re getting real support behind it. No matter how powerful the chatbot is, if your internal systems can’t take API calls, you’re very limited. If you want the bot to actually perform actions, your documentation and systems need to support it. That was something we had to overcome, and it made a huge difference.”**

**— Tanya Budaieva & Ira Smyk, Grammarly Support Team**

## What they prioritized

Grammarly built a 70-criteria evaluation matrix before talking to vendors. Security certifications were non-negotiable, but beyond that, they needed a tool that understood complex questions, responded with empathy, and integrated with existing systems without a heavy engineering lift. They also prioritized implementation speed and a responsive vendor team.

## Mistakes they avoided

Grammarly rejected chatbots that couldn’t maintain context across follow-up questions and refused solutions requiring massive manual decision trees. They avoided vendors promising months-long implementations, and they didn’t deploy until their knowledge base was strong enough to support AI.

## How Forethought helped

Forethought met every requirement in Grammarly’s evaluation matrix. They deployed Solve for chat and email, went live in a week and a half, made possible because their knowledge base was optimized. API integrations enabled system actions, which drove a 5-10% deflection gain.

## Results

- CSAT tripled to 4.2 out of 5
- Deflection started at 60%, climbed to 87%, never dropped below 80%
- QA confirms accuracy of resolutions
- Eliminated manual maintenance overhead from previous chatbot



cotopaxi®

# Case Study 3: Cotopaxi

# Case Study 3: Cotopaxi



**“When implementing Forethought and going from our old chatbot to the new one, the biggest thing was the ability to route customer questions automatically through specific intents. We didn’t have to force customers to pick specific categories. That just made things so much easier to build.”**

**— Alexia Bench, Consumer Insights Manager, Cotopaxi**

## What they prioritized

Cotopaxi needed to increase deflection without sacrificing the 4.5 CSAT their agents had built. They wanted clear control over when AI engaged and what it said, deep reporting to monitor specific conversations and intent flows, and a tool that wouldn’t add work to their agents’ plates. Plus, integration with Kustomer and Shopify was essential for full customer context.

## Mistakes they avoided

Cotopaxi ran a proof-of-concept before committing to deployment and refused to sacrifice service quality for automation. They also avoided treating their knowledge base as “set it and forget it,” assigning one team member to lead ongoing audits and updates.

## How Forethought helped

Cotopaxi deployed Solve for chat with defined customer intents that didn’t force customers into rigid categories. Integration with Kustomer and Shopify gave AI full context on each interaction, while Discover’s Knowledge Gap tool identified missing articles. Its article templates proved especially useful during the holiday rush when gaps needed to be closed quickly.

## Results

- Saved \$76,000 in first six months of 2024
- Increased deflection rate by 28%
- Maintained agents’ 4.5 CSAT score
- Kept headcount steady through holiday season despite company growth



**YAZIO**

# Case Study 4: YAZIO

# Case Study 4: YAZIO



**“We would drown in tickets if we didn’t have Solve’s 80% deflection. We’d have to hire three full-time agents focused on customer support to maintain our team’s current workload if we were to try to replace it with human support.”**

**— Michael Trojan, User Success Specialist (AI), YAZIO**

## What they prioritized

YAZIO was growing rapidly and couldn’t hire fast enough to keep up. They needed 24/7 support across six languages and wanted agentic AI that wouldn’t require the constant maintenance of rule-based decision trees. Given the sensitive nature of nutrition and eating topics, they also prioritized careful language handling and high security standards for customer data.

## Mistakes they avoided

YAZIO rejected rule-based chatbots and refused app-based solutions that would expose customer data. They worked closely with Forethought on language translations before launch rather than rushing to deploy, and they didn’t measure CSAT immediately, which gave customers time to adjust to AI interactions before evaluating satisfaction.

## How Forethought helped

Forethought deployed Solve across six languages, including German, English, Portuguese, Italian, French, and Spanish. The system automatically redacted sensitive data and met security requirements like SOC 2, ISO 27001, and GDPR compliance.

## Results

- First month: 60% deflection at 3.1 CSAT
- Six months later: 80% deflection at 4.0 CSAT
- 40% growth in tickets while team handled same volume as previous year
- Would need three additional full-time agents to replace Solve



# **Final Checklist: How to Implement AI the Right Way**

# Final Checklist: How to Implement AI the Right Way

Before you deploy, review these five areas where implementations typically succeed or fail. If you can't check a box, that's worth addressing before you move forward.

## 1. Align AI to your business goals

- Found the specific metric that must move for this to be worth it  
**Metric:** \_\_\_\_\_
- Defined today's baseline for that metric  
**Baseline:** \_\_\_\_\_
- Set a target improvement with a timeline (e.g., reduce CPR by 25% in 90 days)  
**Target:** \_\_\_\_\_
- Identified which problems AI will solve (e.g., repetitive tickets, agent burnout)  
**Problems:** \_\_\_\_\_
- Confirmed AI addresses root cause, not just symptoms
- Built a business case showing ROI across cost, CX, and agent experience
- Leadership understands what success looks like

## 2. Involve CX leaders, engineers, and agents early

- CX leader owns the implementation  
**Owner:** \_\_\_\_\_
- Agents consulted on which tasks drain their time most  
**Tasks:** \_\_\_\_\_
- Team understands what AI will handle and what stays with humans
- A training plan exists for agents to work with AI tools (like Assist)
- Feedback loop established so agents can report when AI fails  
**Channel:** \_\_\_\_\_

### 3. Keep humans in the loop for oversight

- Defined clear escalation thresholds (e.g., complexity, risk, ambiguity)
- Agents receive full context when AI hands off tickets
- Guardrails prevent AI from making decisions outside your framework
- Audit logging is enabled for every AI action and decision
- QA process reviews AI interactions regularly (manually or via Agent QA)

**Frequency:** \_\_\_\_\_

- Can adjust AI behavior without retraining the entire system
- Humans can override AI decisions when needed

### 4. Track the right mix of metrics

- Measuring deflection rate (e.g., percentage resolved without human)

**Target:** \_\_\_\_\_

- Tracking cost per resolution

**Target:** \_\_\_\_\_

- Monitoring CSAT

**Target:** \_\_\_\_\_

- Watching agent efficiency (tickets per hour, time to resolution)

**Target:** \_\_\_\_\_

- Granular reporting by topic, channel, and complexity

- Can see which issues deflect well and which need improvement

- Metrics reviewed weekly during the first 90 days, monthly after

**Frequency:** \_\_\_\_\_

- The vendor provides dashboards showing these breakdowns

## 5. Choose a partner, not just a tool

Vendor has production customers you can reference

**References:** \_\_\_\_\_

QA process reviews AI interactions regularly (manually or via Agent QA)

**Model:** \_\_\_\_\_

Platform scales across use cases (chat, email, voice, agent tools, routing)

The implementation team will be responsive during deployment

Security certifications match your requirements (SOC 2, ISO 27001, HIPAA if needed)

**Required:** \_\_\_\_\_

Can pilot before full commitment

Contract terms allow you to expand gradually (Phase 1 → Phase 2 → Phase 3)

## See what this looks like for your team

Forethought has helped companies like Grammarly, ActiveCampaign, Cotopaxi, and YAZIO move from evaluation to production with measurable results. If you're ready to start the process, our team can walk you through what each phase would look like in your environment.

[Request A Demo](#)