

# Salesforce Agentforce: The \$2-Per-Conversation Bet

Your Contact Center Is About to Make This Deployment Decision.

01

# Background

How autonomous agents arrived in the enterprise contact center.

## **\$2 per conversation. 85% deflection. What the benchmarks don't show.**

- Salesforce Agentforce reached GA in October 2024. It is designed for autonomous, customer-facing operation — resolving inbound inquiries without human intervention, escalating only when the agent determines it cannot complete the task.
- Pricing: \$2 per conversation for the Flex tier. Enterprise/Unlimited editions include 1,000 conversations/month; every conversation beyond that costs \$2.
- Best published result: Wiley reports 85% case deflection and \$2.7M in annual savings. Salesforce internal deployment reports 84% autonomous resolution in the first reporting period.
- Agentforce performs best in bounded, high-frequency, clear-resolution query categories. Performance degrades predictably when query scope widens or resolution requires exception handling.
- Human agent cost benchmark: \$2–4 per contact at average handling time. At high deflection rates, Agentforce is cost-neutral or better. The model breaks when deflection rate is lower than assumed.

02

## Decision Required

The question before any production deployment.

## Have you modeled your actual conversation volume at \$2/conversation — using your interaction mix, not Salesforce's benchmarks?

At 100,000 monthly interactions, full Agentforce scope costs \$200,000/month — before escalations add human agent cost on top. The economic case depends entirely on your deflection rate in your query mix.

Salesforce's published benchmarks come from best-case deployments with high-volume, bounded query categories. Your interaction profile is different. The deflection rate you will achieve is also different.

Autonomous agent failures in customer-facing roles are brand events, not silent errors. An incorrect resolution, a dead-end loop, or a miscommunicated account status reaches a customer — and may reach their social feed.

## Three deployment postures.

### Option A

#### Full deployment across all query categories

Launch Agentforce for the majority of contact center volume simultaneously. Maximum speed to cost savings if deflection rates match benchmarks. Maximum brand exposure if autonomous failure rates are higher than expected.

### Option B

**Recommended**

#### Scoped pilot — 2–3 bounded categories, 90 days

Deploy autonomously only in your highest-volume, clearest-resolution query categories. Measure actual deflection rate, CSAT delta, and per-conversation cost against your model. Expand only when your own data, not Salesforce's benchmarks, supports it.

### Option C

#### Human-in-the-loop only — agent assist, not autonomous

Deploy Agentforce as a human agent assist tool: AI drafts responses, humans send them. Eliminates autonomous failure modes. Captures a fraction of the cost savings. The right posture for organizations with high CSAT requirements or regulatory constraints on automated customer communication.

## Scoped pilot first. Model your economics before you commit.

Identify top 3 query categories by volume and resolution simplicity. Pilot Agentforce scoped to those categories for 90 days.

Measure deflection rate, CSAT delta, escalation rate, and escalated-contact handling time — not just resolution rate.

Model your actual \$2/conversation exposure at full deployment scale using pilot deflection data — not Salesforce benchmarks.

Define failure criteria before go-live: CSAT floor, escalation rate ceiling, autonomous-error rate that triggers rollback.

Audit data residency and processor obligations before any customer interaction flows through Agentforce infrastructure.

Expand scope only when your own pilot data shows stable performance — not based on published case studies from different interaction mixes.

## Four risks to model before deployment.

1.

### Volume cost exposure at scale

100,000 contacts/month at \$2 = \$200,000/month before escalations add human agent cost. Model your numbers against your actual volume and realistic deflection rate before committing.

2.

### Autonomous failure is a brand event

An autonomous agent that provides incorrect account information or loops a customer through failed resolution steps produces a complaint — and a social media post. CSAT impact from autonomous errors persists beyond the individual interaction.

3.

### Escalation path degradation

Escalated contacts arrive frustrated. Without explicit handoff design — structured context transfer, queue priority routing — human agents spend more time on escalated contacts than on baseline contacts, offsetting deflection savings.

4.

### Salesforce platform concentration risk

Agentforce means your autonomous agents, CRM data, AI interaction logs, and escalation routing are all on Salesforce infrastructure. A platform outage is a contact center outage. A pricing change directly affects your operational cost structure.

## Six questions before you deploy.

1. What is your current monthly inbound contact volume, broken out by query category? Agentforce economics only model correctly against actual volume segmented by query type.
2. Which query categories have resolution paths that can be fully executed without human judgment? Those are your Agentforce candidates. Complex exception handling and high-stakes account decisions are not.
3. What is your current blended cost per contact, including agent salary, benefits, management, and tooling? If your effective cost per contact is below \$2, Agentforce autonomous resolution may not save money at scale.
4. Who owns the CSAT outcome for autonomous agent interactions — IT, contact center operations, or the AI platform team? If unclear, the accountability gap will surface after your first autonomous agent failure.
5. Does your current customer service data handling require specific residency, retention limits, or processor agreements under GDPR, CCPA, HIPAA, or sector regulations?
6. What does your current escalation handoff look like when a digital channel escalates to a human agent? A poor escalation handoff today becomes a larger problem at autonomous agent scale.

AI INSIGHT LAB

# The Deployment Memo

One enterprise AI deployment, dissected every Tuesday.  
Written for executives who have to decide, not just read.

Subscribe at [aiinsightlab.cloud](https://aiinsightlab.cloud) — free during beta.