



Key AI-first customer service metrics

Analysis, reporting,
and industry benchmarks

Paving the way for AI-first customer service

Today, every company is working on incorporating AI into its ethos. We made this prediction a few years ago, and sure enough, we were right.

For customer service, this means transitioning from an **agent-first** to an **AI-first** model, where the nucleus of the service organization is the AI and leaders build around it. Your customers are expecting this too.

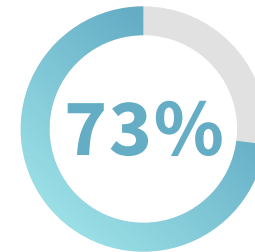
The most transformative effect of an AI-first model is elevating the role of the service organization as a whole.

While previously it may have been seen as a necessary cost center, Customer Service will now have the authority to influence core decisions that affect the growth of the company— decisions like how to evolve the product or service, which third-party vendors to partner with, what the brand should look and sound like...

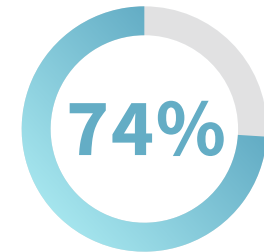
Companies that embrace this shift will see the most growth.

Recommended reading

- [The complete guide to customer service metrics for 2024](#)
- [How data from AI-powered customer service can supercharge the entire business](#)



of surveyed consumers expect more interactions with AI in their daily life



of surveyed consumers believe AI will improve customer service efficiency

Keep in mind...

The benchmarks listed in this report offer a quick glance at how your metrics compare against industry averages. They're intended as guidance, and may vary based on the size and maturity of a company. The benchmarks in this document were compiled in Q2 2024.

Evolving beyond agent-first metrics

In a few years, AI and automation will supplant human agents almost entirely, and human customer service agents will instead be elevated to a more strategic and important role in the company.

This foundational shift is naturally accompanied by a shift in how we measure success. Agent-first metrics will continue to be relevant — albeit adapted to a new paradigm — but there’s a growing need to reliably measure the efficacy of AI Agents.

Historically, metrics like containment rate have come close to measuring a chatbot’s success, but they fall short of providing a complete picture when it comes to AI Agents. The containment rate tells you which proportion of conversations that customers had with your bot ended without being handed off to a human agent. But without context, this metric can only tell you so much. Can you differentiate between a customer who ended a conversation with your bot because they were satisfied, as opposed to one who got frustrated and gave up?

It’s easy to understand why current metrics make it hard to calculate the ROI of AI Agents, let alone the impact of the service organization as a whole.

So, let’s take a step back and think, “*What’s the goal/purpose of an AI Agent?*” The answer at the very core is *to automatically resolve customer inquiries*. This seemingly basic objective has a massive ripple effect.

AI Agents in customer service are only as good as the number of inquiries they can automatically resolve. It’s as simple as that. The north star measurement should be Automated Resolution (AR).

Understanding AR

Automated Resolution (or AR) refers to a fully automated conversation between a customer and a company that is relevant, safe, and accurate and does not escalate to a human agent.

- **Relevant:** the AI Agent effectively understands the inquiry and provides directly related information or assistance
- **Accurate:** the AI Agent provides correct, up to date information
- **Safe:** the AI Agent interacts with customers in a respectful manner and avoids engaging in topics that cause danger and harm
- **No escalation:** the AI Agent resolves the inquiry without any agent support

The AI Agent assesses the conversation to determine whether a successful automated resolution did or didn't happen.

Measuring automated resolutions will revolutionize how you quantify your AI Agent’s performance, helping you focus on what’s most important to your customers.

Recommended reading

- [How we optimize our AI for automated resolutions](#)



How it works

Every hour, Ada takes a sample of engaged conversations that occurred 24 hours ago. A large language model assesses each conversation transcript for accuracy, relevancy, safety, and human involvement, then labels them as Resolved and Unresolved.

Resolved conversations indicate that:

- The conversation was accurate, relevant, and safe
- No human was involved

If your current AI chatbot solution does not do this automatically, you can still measure AR manually following this method:

1. Take a statistically significant sample of your chatbot's conversations
2. Review the transcripts and mark each conversation as **resolved** or **unresolved**, and **contained** or **uncontained**
3. Divide the number of **resolved and contained** conversations by the total sample number to get the Automated Resolution Rate (AR%)
4. Use this formula to calculate AR across your entire conversation volume:

$$\text{AR} = \text{conversation_volume} \times \text{AR\%}$$

Measure ROI using AR

AR makes it very easy to get a quick idea of the ROI on your AI Agent. Here are formulas for common customer service KPIs:

Agent hours saved

$$\text{AR} \times \text{average_handle_time} / 60$$

Costs saved (agent hours)

$$\text{hours_saved} \times \text{agent_cost_per_hour}$$

Cost savings (conversations)

$$\text{AR} \times \text{conversation_volume} \times \text{cost_per_conversation}$$

Projected headcount saved

$$\text{AR} / \text{average_tickets_per_agent}$$

Automated Resolution

Containment vs. AR

When you first start measuring AR, you'll naturally notice that the number is lower than your containment rate. This is normal.

Remember that containment tells you how many conversations didn't reach a human, regardless of whether or not the inquiry was resolved.

On the other hand, AR only takes into account **successful** conversations. AR is a more accurate indication of how helpful your AI Agent really is by measuring the success of those conversations and the content they contain. It gives you more insight, with the help of machine learning, to tell you if the customer got the help they needed.

AR also gives you insight into why a conversation was resolved or not, so you can more easily identify gaps and improve performance. With containment, it's almost impossible to identify why a customer exited the conversation (were they happy or frustrated?) so you think you're performing well, but you actually have problems you can't see and can't fix their root cause.

The industry-standard metric of containment is only one piece of the puzzle. Automated resolution rate is more holistic, and is a better indicator of your bot's performance.

Automated Resolution

Status ⓘ
Not Resolved

Reason ⓘ
The bot provided irrelevant links and did not address the customer's issue.

✓ CONTAINED

✗ NOT RESOLVED

9:41

FinTech Bot

I accidentally transferred funds to the wrong account

I'm sorry to hear you're having trouble.

To learn more about how to transfer funds from one account to another, check out the help article below.

[Learn more about transferring funds](#)

• The customer has exited the conversation

From agent-centric to customer-centric

Currently, most of the KPIs that CS leaders track are agent-first and rely quite heavily on the performance of teams of human agents. The thing is, these KPIs are mostly influenced by one lever: budget. If you can hire more agents, have around-the-clock shifts, offer multi-lingual support, of course stats like CSAT, customer wait times, employee engagement, and more, will improve. But as we all know, budget is always a limiting factor.

AI-first customer service takes a more customer-centric approach. As more companies make the shift, CS leaders will need to introduce new KPIs for their teams that focus less on measuring human agents and more on measuring the quality of the service as a whole.

That doesn't mean that the current KPIs will disappear altogether, but they will evolve to suit the new way of working. Here are some examples.

AR will become the North Star

Increasing AR and AR% will become one of the biggest objectives of the CS org. CS leaders will design strategies for their teams to improve AR and increase conversation volumes.

AI Managers and AI Specialists will have new KPIs that directly influence AR.

Human agent metrics will become AI Agent metrics

When companies reach 100% AR (meaning the AI Agent is resolving 100% of customer inquiries):

- CSAT becomes exclusively an AI Agent metric
- Average Handle Time will matter less from a cost perspective and more from an optimization perspective
- Cost per Resolution will more-or-less stabilize and will not factor in agent headcount
- First Contact Resolution will be a good indication of whether the AI Agent is able to accurately understand and resolve a customer's inquiry from the first try or if it needed more clarification
- CS leaders will likely use Time in Bot Dashboard as an indicator of AI Manager and AI Specialist engagement

Customer Experience agents will have revenue KPIs

The Customer Experience team (previously support agents) has a mandate to focus on long-term relationships with customers. To measure that, they'll have KPIs associated with revenue, such as:

- Return customers
- Customer loyalty
- Expansion
- Net Promoter Score



Customer service data benchmarks

Automated Resolution Rate

Definition: the percentage of customer conversations that have been fully automated, that were relevant, safe, and accurate and not escalated to a human agent.

Some ways to learn from AR%:

- This is the best metric to determine how well your AI Agent is doing.
- Use this number to calculate the ROI of AI-powered customer service automation.
- Use this number while calculating headcount projections — either seasonal, full time, or expansion to new markets — by determining the volume of conversations that agents would have to handle.
- Dig into transcripts of unresolved conversations to see how you can optimize the AI Agent.
- During conversation volume spikes, check if AR% remained stable. If not, then you might need to plan ahead to create knowledge base content for seasonal spikes, major campaigns, product launches, etc...

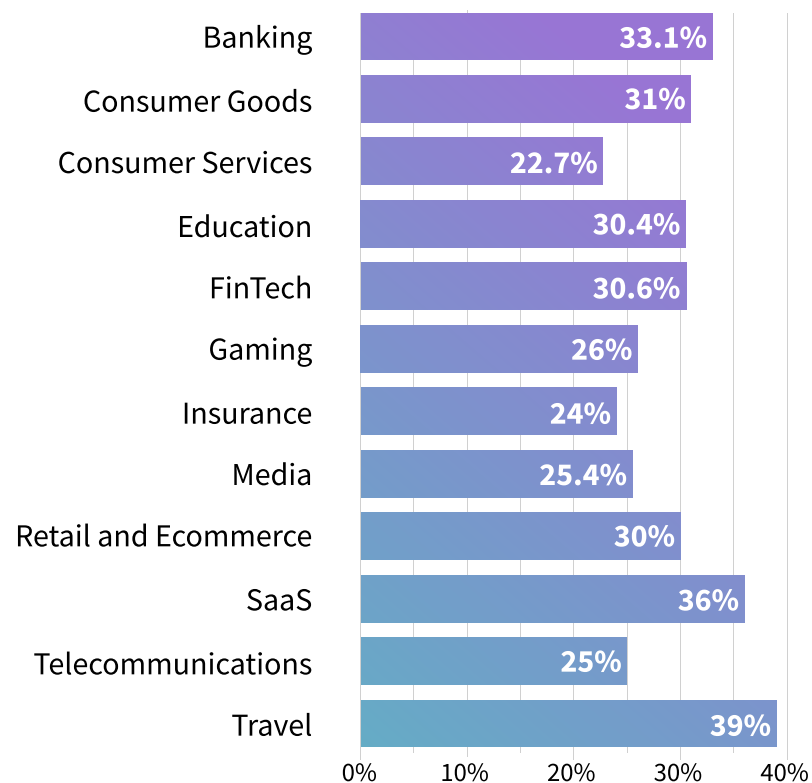
Check how AR% might influence these other metrics:

- CSAT Score: CSAT Score should increase as AR% increases.
- Average customer wait times for live agent: since the AI Agent is resolving more inquiries, wait times should decrease.

Average Automated Resolution Rate across all industries

32.5%

Average Automated Resolution Rate by vertical



Customer service data benchmarks

Containment Rate

Definition: the percentage of customer conversations that were not escalated to a human agent, regardless of whether or not they were resolved.

Some ways to learn from Containment Rate:

- CR% gives you a very quick glance at how many conversations are not being escalated to a human agent so you can quickly spot irregularities between specific time periods (day-to-day, week-over-week, etc...)
- Dig into conversations that were “contained” but not “resolved” to uncover opportunities for training and improving the AI Agent.

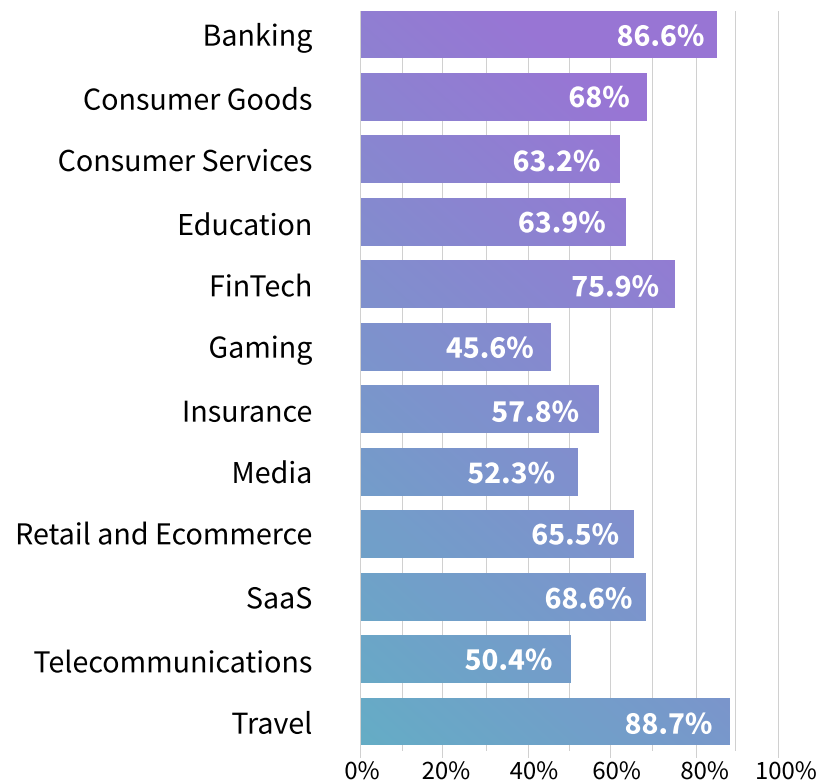
Check how CR% might influence these other metrics:

- AR%: keep an eye out on the discrepancy between AR% and CR%. While AR% will naturally be lower at the start, it should steadily increase and the discrepancy should decrease.
- Conversation Volume: as CV increases, CR% should remain stable. If it's not, you can find opportunities to optimize automation by looking for patterns in uncontained conversations.

Average Containment Rate across all industries

63.9%

Average Containment Rate by vertical



Customer service data benchmarks

CSAT Score

Definition: the percent of conversations customers reviewed positively, out of all conversations they reviewed.

Some ways to learn from CSAT Score Rate:

- Dig into the transcripts of conversations marked positively or negatively to uncover patterns that made the customers happy or unhappy.
- Be mindful of this metric specifically during big campaigns or seasonal sales, it can be a good indicator of whether the big brand moments were received positively or negatively by customers.

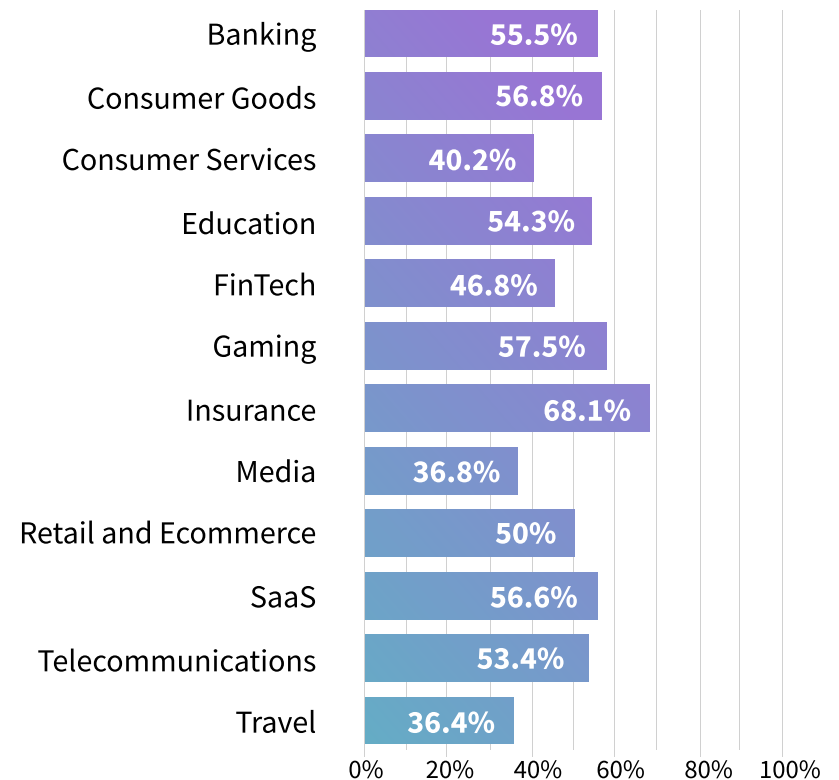
Check how CSAT Score might influence these metrics:

- AHT: CSAT Score typically increases as AHT decreases.
- AR%: CSAT Score typically increases as AR% increases.
- CR%: Comparing CSAT Score with CR% can indicate whether or not customers are happier when their inquiry gets escalated to a human agent.

CSAT Score across all industries

51.6%

Average CSAT Score by vertical



Customer service data benchmarks

Average Handle Time

Definition: the amount of time it takes to resolve an inquiry. This is measured separately for AI Agents and human agents.

Some ways to learn from AHT:

- For AI Agent AHT:
 - The AHT of the AI Agent should be proportional to the complexity of the resolution it's trying to automate — dig into the longest and shortest conversations to see if you notice patterns.
 - If you notice that AI Agent AHT for simple FAQ interactions is high, then you probably need to optimize your knowledge base content.
- For human agent AHT:
 - Review transcripts for conversations that are way over the average human agent AHT to see if agents need to be better enabled or if you can use the AI Agent to collect key information from customers before they get passed on.
 - Review the conversations with the shortest Agent AHT to see if they are better suited for the AI Agent.

Check how AHT might influence these metrics:

- CSAT Score: CSAT Score typically increases as AHT decreases.

Average Handle Times across all industries (min:sec)

AI AGENT:

01:58

HUMAN AGENT:

14:00

Average Handle Time by vertical (min:sec)

	AI	HUMAN
Banking	03:34	14:25
Consumer Goods	02:06	15:42
Consumer Services	02:15	15:07
Education	01:55	15:42
FinTech	02:25	15:31
Gaming	02:05	15:49
Insurance	01:28	12:14
Media	02:05	15:49
Retail and Ecommerce	02:24	13:14
SaaS	01:45	14:45
Telecommunications	02:03	14:58
Travel	02:25	15:12

Customer service data benchmarks

Conversation Volume

Definition: the number of conversations that customers have with a company per set length of time (day, week, month, quarter, year).

Some ways to learn from CV:

- Review yearly CV numbers to uncover patterns for volume spikes and be better prepared for them.
- Keep an eye on CV during big brand moments such as product launches or major campaigns to measure customer engagement.
- Unusual spikes in CV can signal outages or issues in other areas of the company, such as logistics or operations, and you can often notify the relevant teams before they are even aware of the issue.

Check how CV might influence these metrics:

- CSAT Score: comparing CV and CSAT Score during volume spikes can signal positive or negative brand sentiment.
- AR%: AR% should either increase or remain stable as CV increases. If AR% decreases, it means there are opportunities to improve the automated experience.

Average Conversation Volume across all industries

864 per day

Average Conversation Volume per day by vertical

Banking	2,084
Consumer Goods	522
Consumer Services	555
Education	236
FinTech	497
Gaming	1,883
Insurance	129
Media	1,192
Retail and Ecommerce	673
SaaS	564
Telecommunications	572
Travel	4,630

Example weekly AI Agent report

Weekly AI Agent metrics

Be sure to include week-over-week charts for these metrics to track trends

AR

How many conversations did the AI Agent automatically resolve?

XX

AR%

What is the latest AR%?

XX%

CR%

What is the latest CR%?

XX%

AHT

What is the latest AHT?

XX

CSAT Score

Surveys collected

XX

CSAT Score

XX%

Weekly AI Agent insights

Top 3 customer inquiries that were **automatically resolved**

Inquiry volume

AR%

Reasons for success

Weekly AI Agent insights

Top 3 customer inquiries that were **escalated to a human agent**

Inquiry volume

Hand-off %

Opportunities for improvement

Other noteworthy insights

For example, any issues with the AI Agent or any knowledge base gaps or opportunities that the AI Agent identified.

Example weekly AI Agent report

Updates and amendments

Such as:

- *Any improvements made to the knowledge base, product, or any other areas of the business that was informed by AI Agent data*
- *Any improvements made to the AI Agent itself – new integrations, new training, etc...*
- *Any issues from the previous weeks that have been resolved*

Weekly AI Agent successes

Such as positive customer reviews, early issue identification, anecdotes of very successful generations, or the AI Agent handling an unexpected CV spike.


Partner with the pros

We've helped over 350 innovative customer service leaders deliver transformative results using AI. If you're looking for guidance, our consultants are ready to chat about where this technology can fit into your company and how it can help you transform your customer service organization.

Schedule a call

Ada is an AI-powered customer service automation platform on a mission to make customer service extraordinary for everyone. Ada envisions a world where every customer interaction is resolved by AI, and makes it easy for businesses to automatically resolve the greatest number of customer service conversations — across channels and languages — with the least amount of effort.

Since 2016, Ada has powered more than 4 billion automated customer interactions for brands like Meta, Verizon, AirAsia, Yeti, and Square. Born in Toronto, Ada serves companies and their customers worldwide.

-  +1 (855) 232-7593
-  hello@ada.support
-  [@Ada_CX](https://twitter.com/Ada_CX)
-  facebook.com/getadasupport
-  linkedin.com/company/ada-cx
-  youtube.com/@ada_cx

ada.cx

