

The Human/AI Collections Matrix

What To Automate

What To Escalate

What Humans Are Left With



The most important question isn't whether AI can make a call.

It's which calls it should make/take and what happens to your agents when it does.

The obvious automation candidates are easy to identify:

- balance questions
- payment reminders
- basic account information
- routine payment processing
- Common FAQs.

These conversations tend to be predictable, lower risk, and easier to contain within defined rules.

The harder conversations are different.

A consumer says they can't pay because they lost their job. Someone disputes the balance. A complaint starts escalating. An account involves bankruptcy, a deceased consumer, identity uncertainty, or a legal threat. A payment arrangement requires actual negotiation.

Those calls require context, judgment and often empathy.

Deloitte's 2026 research recommends AI-led service for simple, low-risk interactions, human oversight as complexity increases, and human-led handling for high-stakes situations such as disputes, hardship, fraud and complaints. In the same research, 70% of banking executives said they expect AI to shift human agents toward higher-value roles.

That sounds like a sensible division of labor.

But if AI takes the easiest calls, the average human call gets harder.

That changes far more than headcount.

The Human/AI Collections Matrix

There probably shouldn't be one simple rule that says "AI calls go here" and "human calls go there."

A better decision model looks at several characteristics at once:

(Using a 1-5 scale in terms of call difficulty)

Conversation type	Recommended owner	Complexity	Compliance risk	Empathy required	Judgment required
Balance inquiry	AI	1	2	1	1
Due-date reminder	AI	1	2	1	1
Standard account FAQ	AI	1	2	1	1
Simple payment	AI / Hybrid	2	3	1	2
Promise-to-pay arrangement	Hybrid	3	3	2	3
Negotiated payment plan	Hybrid	4	4	3	4
Consumer dispute	Human + AI assist	4	5	3	4
Financial hardship	Human + AI assist	4	4	5	5
Complaint	Human	4	5	5	5
Identity uncertainty	Human	4	5	2	5
Legal threat / attorney	Human	5	5	3	5
Bankruptcy / deceased	Human	5	5	5	5
Vulnerable consumer	Human	5	5	5	5

**These ratings are an Abstrakt operating framework rather than regulatory classifications. Companies should adapt the matrix to their business, clients, policies, and risk tolerance.*

How Much More Call Capacity Does AI Actually Create?

One of AI's most compelling advantages is concurrency.

A human collector can have one phone conversation at a time. An AI voice system can potentially conduct multiple conversations simultaneously, subject to system design, available infrastructure, policies, and regulatory requirements.

Conversation capacity per hour = concurrent conversations × 60 ÷ average handle time

Assume a human collector's increasingly complex queue averages eight minutes per conversation, while the routine conversations assigned to AI average three minutes.

Capacity measure	Human collector	AI: 10 concurrent	AI: 25 concurrent	AI: 50 concurrent
Modeled average handle time	8 min	3 min	3 min	3 min
Concurrent conversations	1	10	25	50
Theoretical conversations/hour	7.5	200	500	1,000
Theoretical conversations/8 hours	60	1,600	4,000	8,000
Best use	Complex	Routine	Routine	High-volume routine

Important: This is capacity, not permissible contact volume

Under the CFPB's Debt Collection Rule, a debt collector is generally presumed to violate the rule's prohibition against repeated or continuous calling if it places more than seven calls to a particular person about a particular debt within seven consecutive days, or calls within seven days after having a telephone conversation about that debt, subject to exceptions and the specific circumstances involved.

State laws and other requirements may add further restrictions.

In other words: AI can remove the labor bottleneck. It does not remove regulatory, contact-strategy, consent, consumer-preference, or right-party-contact constraints.

The more useful question is: How much work can safely move away from humans?

What Happens When AI Takes 44% of the Conversations?

This is where workforce planning gets interesting.

Consider an operation with 100,000 completed monthly conversations.

Call category	Monthly conversations	Share	Modeled AHT
Routine	40,000	40%	3 min
Moderate	35,000	35%	6 min
Complex	25,000	25%	10 min
Total	100,000	100%	5.8 min weighted AHT

Before automation, humans handled all 100,000 conversations.

580,000 handling minutes, or 9,667 hours per month.

Now introduce AI.

Assume the organization automates:

- 85% of routine conversations
- 30% of moderate conversations
- 0% of complex conversations

Call category	Original volume	AI handled	Human handled
Routine	40,000	34,000	6,000
Moderate	35,000	10,500	24,500
Complex	25,000	0	25,000
Total	100,000	44,500	55,500

AI now handles 44.5% of conversations.

At first glance, someone looking only at call volume might assume human workload should decline by roughly the same amount.

It doesn't.

The remaining human handling time is approximately:

- Routine: 18,000 minutes
- Moderate: 147,000 minutes
- Complex: 250,000 minutes

Total: 415,000 minutes, or 6,917 hours.

So while AI removed 44.5% of the conversations, it removed only about 28% of the human handling time.

Why?

Because AI took short calls.

The humans kept the long ones.

AI Can Make Your Human KPIs Look Worse

Before automation, the modeled human AHT in our example was: **5.8 minutes**.

After automation: **7.5 minutes**.

That's roughly a 29% increase in human AHT.

If those numbers showed up on a dashboard with no context, someone might conclude that collector productivity had deteriorated badly.

In reality, the opposite may be happening.

Collectors are no longer spending as much time answering balance questions or handling straightforward payments. They're spending more of their day solving disputes, working through hardship situations, negotiating arrangements, and managing exceptions.

That means organizations adopting AI need to be careful with traditional performance metrics.

Traditional KPI	What AI may do to it	What it could actually mean
Calls per hour	Decrease	Humans are receiving more complex conversations
Average handle time	Increase	Easy calls have been removed
Escalations per collector	Increase	Humans are becoming escalation specialists
Supervisor assistance	Increase	Remaining situations require greater judgment
First-call resolution	Could change	Humans receive problems AI couldn't resolve
Dollars collected per conversation	Could increase	Collectors focus on higher-value interactions
QA risk per call	Could increase	Remaining calls contain more judgment and compliance exposure

A company could therefore automate successfully while watching several traditional human productivity metrics move in the "wrong" direction.

The mistake would be measuring the post-AI workforce as though it were still doing the pre-AI job.

QA Becomes More Important, Not Less

If the human queue becomes more complex, the downside of reviewing only a tiny sample of conversations becomes harder to ignore.

Imagine that your collectors increasingly spend their day handling:

- Disputes
- Negotiations
- Hardship
- Complaints
- Unusual account situations
- Escalations
- Vulnerable consumers

Those are precisely the conversations where organizations want visibility.

At the same time, AI-generated conversations need oversight too.

Deloitte predicts that quality assurance could move from manual sampling toward continuous monitoring of both human and AI interactions, including AI-assisted script adherence, complaint detection, and compliance monitoring.

That creates two parallel QA responsibilities

Human QA	AI QA
Did the collector follow policy?	Did AI remain inside its guardrails?
Were disclosures handled correctly?	Were required disclosures produced correctly?
Was the consumer treated appropriately?	Did AI identify consumer sentiment or vulnerability?
Was negotiation appropriate?	Did AI exceed an authorized negotiation boundary?
Did the collector recognize risk?	Did AI recognize an escalation trigger?
Was the issue resolved correctly?	Was the AI response accurate?
Was escalation necessary?	Should AI have escalated sooner?

NIST's AI Risk Management Framework similarly emphasizes regular evaluation of deployed AI systems, real-time monitoring where appropriate, and mechanisms for tracking existing and emerging AI risks over time.

In other words, deploying the AI isn't the end of the implementation.

Someone still needs to determine whether it's making the right decisions.

A New KPI: Escalation Accuracy

One of the most useful measures in an AI-enabled contact center may have nothing to do with traditional collector productivity.

It is: Did AI know when to stop?

Organizations can measure this using two concepts common in classification systems.

ESCALATION PRECISION

Of all conversations AI escalated to a human, how many actually needed human involvement?

High precision means AI isn't constantly handing routine issues back to people.

ESCALATION RECALL

Of all conversations that should have gone to a human, how many did AI successfully identify?

High recall means AI is catching the situations organizations don't want automated.

A missed escalation could mean AI continued a conversation involving a dispute, complaint, vulnerable consumer, legal issue, or other circumstance the organization intended to route to a person.

Automation rate tells you how often AI stayed on the call.

Escalation accuracy tells you whether it should have.

Those are very different measures.