

Making AI work in post-sales

From hours saved to customers kept

Where post-sales teams really are with AI, what holds them back, and a five-step system that matches AI to the right work and turns the hours it gives back into customers who stay and grow.

Thomas Voigt, CEO and founder of The CS Consulting

This paper is meant as the start of a conversation. It looks at where post-sales teams really are with AI, what I see holding them back, and the system I've built to bring AI into a customer success team on purpose, so that it shows up in retention as well as in the calendar.

1 Where we really are

In October 2026 ChurnZero published its seventh annual Customer Revenue Leadership Study, built on answers from 580 CS and revenue leaders at B2B software and services companies. I expected it to show most teams well into AI by now. The tools are bought, the training has happened, the pilots have names. The numbers tell a more modest story.

81%

say AI has cut the time their teams spend on manual tasks

1 in 2

have no consistent way to measure AI's impact

21%

of CS leaders link what they measure about AI to revenue outcomes

Flat

net revenue retention since 2024

61%

say AI now factors into how they plan headcount

4×

more likely to approve every AI recommendation by hand when impact isn't measured

A word on what "measure" means here, because it's easy to misread. The study asks whether leaders have any consistent way of judging what AI is doing for them, whether that's usage, hours saved or cost. Among CS leaders, 63% have no consistent method at all, and only 21% connect whatever they do measure to revenue outcomes such as retention and expansion. Adoption is one of the things teams measure, but it isn't the whole picture.

The hours are coming back. What hasn't followed is a result a board would recognise. Four in five teams stop AI short of strategic decisions, and 63% of leaders still rate their AI maturity as tactical or operational; 16% call it strategic. The figures are self-reported, and I'd read them as a direction of travel more than a precise measure. The direction is clear enough.

AI has changed our calendars. It has yet to change our customers' minds about staying.

I find that disappointing, and I suspect many CS leaders do too. So I've tried to understand why it's happening. I see three reasons, and the third is the one that matters most.

2 Resistance: belief comes before behaviour

The first obstacle is belief. The public story about AI swings from AGI arriving within a couple of years to an investment bubble about to burst. When the loudest voices can't agree, waiting looks like the sensible choice, and there's little urgency inside teams on a Monday morning.

The doubt also turns personal. With 61% of leaders saying AI now shapes headcount planning, a tool launched with an email about efficiency reads to a CSM as a question about their own job. A team that suspects that has little reason to help the tool succeed. They'll use it when someone is watching, rewrite its drafts by hand and let the pilot prove it wasn't ready. The study shows what this looks like: where impact isn't measured, teams are four times more likely to approve every AI recommendation by hand. That is what doubt looks like in a workflow.

3 The misconception: what success looks like

The second obstacle is how we define success. "AI-ROI" sounds like a target, but it's too broad to steer by. It hides the real decision: are we working on the cost side, the revenue side, or both, and when should we expect to see each?

Without that agreement, one of two things tends to happen. Leaders expect revenue too early and judge a programme a failure at month three. Or they settle for hours saved and stop there, and the hours end up on the headcount line: when budgets shrink, 59% of teams cut people while 27% cut AI spend. Efficiency is the entry ticket. It pays for the work and builds trust. If we stop there, AI becomes a cost-cutting exercise, and we risk automating the very conversations that keep customers renewing.

4 The real problem: nobody decided where AI belongs

Underneath both sits the problem I see most often. AI goes in everywhere at once, matched to nothing in particular. Most of the rollouts I hear about were never run as projects. A tool was bought, there was a launch, the training happened, and then it was pointed at whatever looked easiest, with no owner and no date by which anything had to be true. When budgets grow, 70% of teams put the money into AI tools, and few of them first wrote down what their team actually does all week, or why. You can't apply AI to everything. It has to be matched to the right work, in the right way, and without a system for doing that, the same five gaps open up every time.

No map of the work	The team's week was never written down or measured, so there's nothing to prove the gain against and the tool goes looking for easy tasks, including ones nobody needs any more.
No line between human and machine	The question of which work should stay human was never settled, so CSMs fear the tool will reach the conversations they care about most.
No check on the data	AI reads whatever it's given. Half-finished call notes and a health score that counts logins produce confident nonsense.
No test with the people who'll use it	Outputs reach the team before they're good enough, trust never arrives, and every recommendation gets checked by hand.
No agreed measure, and no plan for the hours	Success was never defined, so the answer to "is AI working?" stays vague, and the freed hours drift into email, meetings or the headcount line.

Each of those gaps feeds the resistance and the misconception above. That's why I sat down and built CRAFT: a five-step process that closes them in order, that puts AI only where it belongs, and that a CS leader can run as a project with a clear owner, a plan and a review.

5 CRAFT: a system for bringing AI into post-sales

CRAFT is the method I use to decide how much AI a post-sales team should use, where and when, and then to build it well enough that the team actually chooses to use it. Some work should go to AI, some should stay with people, and some should stop altogether. CRAFT sits alongside ACE, the way I run customer success, and is applied stage by stage across Activate, Cultivate and Expand. The name is on purpose: craft means taste and judgement, and the willingness to send something back when it isn't good enough.



Run like a project

An owner and a sponsor. The CS leader owns the programme day to day, an executive, usually the CRO, sponsors it, and each use case has its own named owner.

A plan with dates. Six months in four phases, then a quarterly rhythm through month twelve.

A baseline before anything changes, so we can prove what the AI did.

Gates between the steps. Each letter has a short list of what must be true before the next one starts.

A cadence. A weekly working session, a fortnightly check-in with the CS leader, a monthly steering review with the sponsor.

A decision log. Every lane change, killed idea, retired task and data problem is written down with the reason.

C Capture the work

MONTH 1

With the leaders. I start with one-to-one conversations with the CRO, the CS leader and two or three of the people doing the work, because the picture has to be human before it's operational. I want to hear where the week goes, what frustrates them and which customer moments they're proudest of. Then we sit down together and write out every recurring task across the lifecycle, from onboarding through adoption, value realisation, renewal and expansion: preparing a kick-off, writing the success plan, chasing usage, building the QBR deck, logging the call, updating the health score, spotting the expansion signal, drafting the renewal note.

The questions I ask. Where did your last week actually go? Which task would you happily never do again, and which one would you fight to keep? Why does this task exist, and who still uses what it produces? Which customer conversation do you wish you had more time for? Where do you wait on someone else, or someone on you? The answers show the work behind the job description, and they tell me early where trust lives in the team.

What we record. For each task: why it exists and who still uses the output, who does it, how often, how long it takes, which systems and sources it draws on, and where it touches Sales, RevOps or Product. A lot of the value sits in those handovers, and they rarely show up on anyone's plan.

Why the why matters. A lot of post-sales work survives on habit, and reporting more than anything. A weekly report built for a leader who has since moved on, a spreadsheet updated every Friday that nobody opens:

these keep going because no one ever asked what they're for. There's no point teaching AI to produce something nobody needs. Those tasks leave the inventory before they ever reach a lane, and the time they free up counts towards the baseline gain from day one.

What comes out. A complete task inventory, a short list of tasks to stop, and a baseline measured before anything changes: hours per task, accounts per CSM, the share of the week spent with customers, time to value, GRR and NRR. For many leaders it's the first time they see where their team's week actually goes, and the answer is rarely "with customers".

Before we move on. The inventory is signed off by the CS leader and checked with at least two CSMs, every task has a recorded reason to exist or is marked to stop, and the baseline is recorded.

R Rank human and machine

MONTH 1

With the leaders. Once the tasks we're stopping are gone, we go through what's left task by task and place each one in a lane. Human is where trust is built or lost: the difficult renewal conversation, the executive relationship, the judgement call on a wobbling account. Assisted is where AI prepares and a person decides: the pre-call brief, the first draft of a follow-up, the QBR narrative. Automated is work a person shouldn't be doing by hand any more, with a person spot-checking: CRM updates, meeting notes, task creation, routine data refreshes.

How we decide. I start with three questions. Does this moment carry emotional weight for the customer? Would the customer notice or care whether it came from a person or a system? Does it create information or trust we'd lose if it were automated? A yes to any of them keeps a task out of the automated lane. Then two questions of my own separate assisted from automated: how much judgement does it need, and how easily can a mistake be undone? We start where a lot of time comes back and the risk to the relationship is small.

A worked example. Take the QBR. Pulling usage and support data into the deck can be automated. Drafting the value story can be assisted, with AI preparing and the CSM deciding what's true and what matters. Running the conversation with the executive sponsor stays human. One meeting, three lanes, and a CSM who walks in prepared rather than exhausted.

What comes out. A lane map of the team's work and a shortlist of two or three pilots. Each pilot gets a one-page use-case card: the business outcome, a named owner (often outside CS, which is the point), the systems and data it draws on, the workflow and next best action it triggers, today's baseline and the KPI that will tell us it worked. If the card can't be filled in, the idea isn't ready, and it's far cheaper to find that out on paper.

Before we move on. Every task sits in a lane, the sponsor has agreed the priorities, and each pilot has a complete card with a baseline and a target.

A Audit the inputs

MONTH 2

With the leaders. For every pilot we open up what the AI would actually read, starting with the CRM and the call notes. We check four things. Is the data correct, or has it drifted through manual updates and old definitions? Is it current, fed by systems rather than by whoever remembered to update it? Is it complete? Is it consistent, so the CRM, the CS platform and finance tell the same story?

The questions that matter. Are the call notes complete or half-finished? Does the health score measure value or activity? Is the renewal date in the CRM the real one? We also look at what the data is organised around. Customer milestones, such as goals agreed, the right stakeholders engaged and the product built into daily workflows, give AI real signals to work with. Journey stages mostly tell it where the customer sits in our own process.

Typical fixes. Agreeing one definition of an active user across the CRM and the CS platform. Turning the renewal date into a field a system maintains rather than a note someone updates. A short, shared structure for call notes, so a summary has something real to summarise. Moving the health score from logins towards signs of value, such as goals reached and the right stakeholders engaged. None of these is glamorous, and each one makes every AI use case built on top of it better.

What comes out. A clear verdict on each data source and, where it falls short, a dated fix with an owner. AI won't fix bad data; it will only help us find it faster. Where the inputs are weak, the task stays with a person until they're fixed.

Before we move on. Each pilot's data sources pass those four checks, or there's a dated fix in the plan.

F Fire-test the output

MONTHS 3 TO 4

With the team. Before a jet engine is certified it gets broken on purpose, and nothing should reach a team until it has been pushed the same way. For each pilot we write down what good looks like, in the words of the people who'll use it. Then we run it on real accounts and send it back until it meets that picture. The first version is usually recognisable, the second pretty good, and it often takes a third before anything holds. Lauren Hughes at Justworks described exactly this with her team's personas: two flat attempts, a third that held.

What good looks like. For a pre-call brief, for example, the standard might be: readable in two minutes, names the customer's own goal in their words, flags what has changed since the last conversation, points to one risk and one opportunity, and suggests the question worth asking.

Who decides. The CSMs do. We ask the person with twelve minutes before a call whether they'd trust the brief. If they wouldn't, it isn't finished. That rejection loop is taste, the human skill at the centre of the method, and it's also where adoption is won, because people trust what they've helped shape.

Before we move on. There's a written quality standard for each pilot, at least three CSMs say they'd use it as it is, and the owner signs it off for launch.

T Track success

MONTHS 5 TO 6, THEN QUARTERLY

With the leaders. We agree up front which level of the AI success pyramid we're aiming for, then measure at three rhythms. Adoption weekly: are the briefs opened and the drafts sent? Capacity monthly: hours given back, accounts per CSM, the share of the week spent with customers. Outcomes quarterly: time to value, GRR and NRR.

A plan for the hours. If AI gives a CSM back hours every week and the business doesn't decide where they go, they disappear into email and internal meetings, or onto the headcount line. So the human lane gets a plan too: which accounts get more time, which conversations happen earlier, such as the renewal conversation at 120 days rather than 30, and which customers we go and see in person.

What comes out. A monthly steering review against the baseline, a decision log that becomes the story for the board, and a quarterly lane review where tasks move from human to assisted to automated once they've earned it, and occasionally back.

6 A framework still needs a finish line

CRAFT tells a team what to do and in which order. On its own, it doesn't tell the board when to expect a result, or what that result should be, and that's where most programmes come unstuck.

If success isn't defined before the first pilot, every stakeholder fills the gap with their own picture. The CFO hopes for lower costs, the CRO for higher NRR, the team fears for its jobs, and the CS leader ends up defending one programme against several scorecards at once. A programme measured against everyone's hopes will disappoint all of them.

So before anything is built, I ask the leadership team to agree four things: what we'll measure at each stage, from adoption to revenue; when we expect to see it, so a month-three review isn't held to a month-twelve number; how often we look, weekly, monthly or quarterly depending on the level; and who hears about it, from the team that needs to see its own wins to the board that needs to see the trend. Measured this way, success also builds belief. Each level reached is a visible win, shared early, that earns the next one. The pyramid is the tool I use for that conversation.

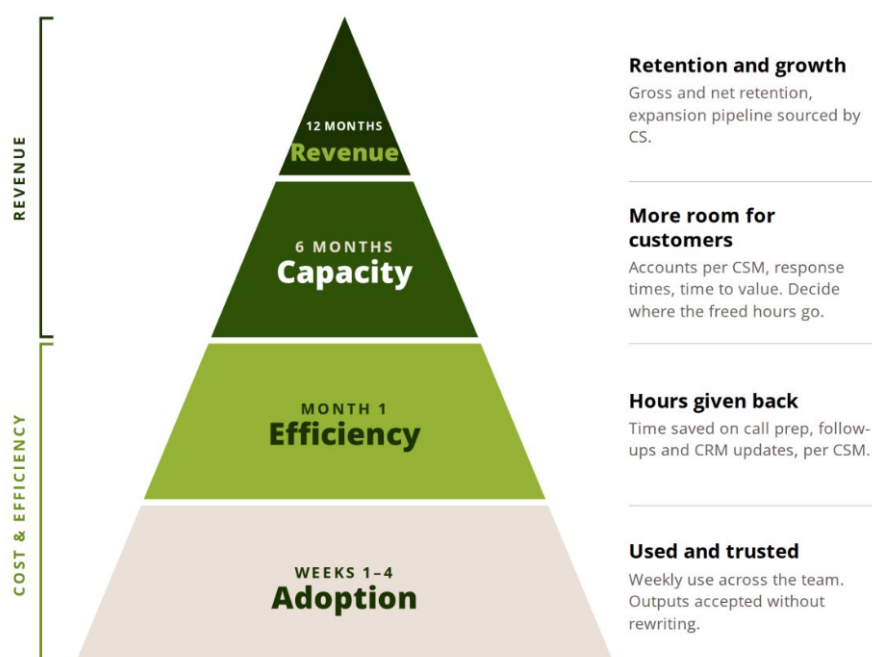
7 Measuring success: the AI success pyramid

AI doesn't skip the queue. The value arrives in a sequence, and agreeing that sequence up front is what turns a vague "AI-ROI" into something a team and a board can steer by. The pyramid answers the four questions from the section before: what we measure, when we expect to see it, how often we look, and who hears about it. It sits in the T of CRAFT.

The AI success pyramid



What to expect from AI, and when.



The bottom two levels are the cost and efficiency side. They come first, they come quickly, and they're where trust is built: adoption in the first four weeks, then efficiency, the hours given back, from month one. The top two levels are the revenue side. Capacity shows around month six, when we decide on purpose where the freed hours go, and revenue shows at month twelve.

Each level earns the next. A team that skips adoption builds on outputs the CSMs don't trust; a team that stops at efficiency turns AI into a cost-cutting exercise; a team that expects revenue at month three cancels a good programme too early. Put ChurnZero's numbers on the pyramid and most teams have reached efficiency, measured it loosely, and skipped capacity. That's the gap this paper is about.

The pyramid, level by level

This is how I use each level with a leadership team: what it means, what we measure and how often, the win worth celebrating, and the warning sign that tells us we're not there yet.

WEEKS 1 TO 4 Adoption Used and trusted	What it means. The team uses AI every week and accepts its outputs without rewriting them. This is the foundation: if people don't trust it, nothing above this level happens. What we measure. Weekly use across the team, the share of outputs accepted as they are, and how many drafts still get rewritten by hand. We look at it every week in the working session and share it with the team. The win to celebrate. The week everyone opens the brief before every call. The warning sign. Outputs rewritten by hand and every recommendation approved line by line. Teams that can't measure impact are four times more likely to work this way.
MONTH 1 Efficiency Hours given back	What it means. Time saved on the tasks we picked, such as call prep, follow-ups and CRM updates, per CSM, plus the time freed by tasks we stopped altogether. This is the cost side, and usually where the first business case sits. What we measure. Hours per task against the baseline from Capture, reviewed monthly with the CS leader and the sponsor. The win to celebrate. The month the hours given back reach a number you can name. The warning sign. Savings claimed without a baseline. 81% of leaders say AI saves time, and half can't measure what it's worth.
MONTH 6 Capacity More room for customers	What it means. The freed hours turn into customer time, by decision rather than by accident. This is the hinge between the cost side and the revenue side. What we measure. Accounts per CSM, response times, time to value for new customers and the share of the week spent with customers. Monthly with the sponsor, quarterly with the leadership team. The win to celebrate. The renewal conversation that now happens at 120 days, or the customer a CSM finally has time to go and see. The warning sign. Hours that drift into email, internal meetings or the headcount line. 61% of leaders now plan headcount around AI.
MONTH 12 Revenue Retention and growth	What it means. Gross and net retention, and expansion pipeline sourced by CS. This is where AI earns its place in the boardroom, and it only shows up when the three levels below hold. What we measure. GRR, NRR and CS-sourced pipeline, reported quarterly to the board against the baseline. The win to celebrate. A renewal or expansion the board can trace back to time the team got back. The warning sign. Revenue expected at month three, or never measured at all. Only 21% of CS leaders link their AI measurement to revenue outcomes, and NRR has been flat since 2024.

Efficiency is the entry ticket. The bigger prize sits at the top, in customers who stay and grow, and it needs a twelve-month plan where each level earns the next.

8 What success looks like so far

The question I hear most once people have seen the pyramid is whether anyone has reached the top yet. The honest answer is that clean, revenue-level proof is still rare, and that's a big part of why I wrote this paper. There is one encouraging number in the ChurnZero study. While NRR across the market has been flat since 2024, 65% of the most AI-mature teams report NRR of 100% or more. It's self-reported, and it doesn't prove that AI made the difference, but the teams furthest along are also the ones whose customers are staying.

Where I see real progress today, it starts lower down the pyramid, in teams that agreed a baseline before they switched anything on. They knew how many hours a CSM spent preparing a QBR, or writing the summary email after every client call, so when AI took over part of that work they could name the hours that came back and decide where they went. It's a modest kind of success. It's also a measurable one, and it's the kind that earns the next level.

If we really want to get our heads around AI, and what we can do and control today, there are a lot of variables. Two of them matter most: being intentional about where AI goes, so it's matched to the right work in the right way, and having a shared understanding of what success could look like.

Thomas Voigt

CEO and founder, The CS Consulting



CRAFT as a workshop

If your team feels it needs a facilitator on board to put CRAFT into practice, please reach out. I've done this with teams a couple of times to fast-track a more deliberate, intentional use of AI and get to results quicker.

[Book a meeting with Thomas](#)

[Send Thomas an email](#)

thecsconsulting.net · linkedin.com/in/thomas-voigt · London

Source: ChurnZero, "ChurnZero study finds AI investment is outpacing measurement in customer success", 6 October 2026 (580 CS and revenue leaders; self-reported), churnzero.com/press-release/ai-research-customer-growth.