

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

THE SHORT VERSION

- 1 AI is giving post-sales teams time back, and retention hasn't moved yet: 81% say it cuts manual work, and NRR has been flat since 2024.
- 2 Three things hold teams back: doubt about what AI means for their jobs, a fuzzy idea of success, and AI dropped in everywhere without anyone deciding where it belongs.
- 3 CRAFT is a five-step way to match AI to the right work: Capture the work, Rank human and machine, Audit the inputs, Fire-test the output, Track success.
- 4 A small team can start with two pieces of it: a baseline of where the week goes, and three lanes for every task, human, assisted and automated.
- 5 Results arrive in order, adoption and hours first, then more time with customers, then retention and revenue, so agree up front what you expect and when.

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. CRAFT is offered as a proposal to test and argue with, and section 5 says plainly how much evidence there is so far.

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. Every figure I quote comes from the full report; the press release linked here carries only a few of them, and I'm happy to share the report itself. 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

customer growth leaders 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. Across all 580 leaders, 49% have no consistent method at all. Among CS leaders alone it's 63%, the highest of any function, 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 across the whole sample 63% still rate their AI maturity as tactical or operational (the same figure as above, for a different question); 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 What holds teams back

I see three reasons, and they feed each other.

Doubt

The public story about AI swings from AGI within a couple of years to a bubble about to burst, and when the loudest voices can't agree, waiting looks like the sensible choice. Inside teams the doubt 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, and they have little reason to help it succeed.

Where impact isn't measured, teams are four times more likely to approve every AI recommendation by hand. Some of that is doubt. Some of it, as John Gleeson argues in the study, is simple prudence: nobody wrote down what a good output looks like, so there's no way to earn the confidence to let go. Both point to the same fix.

A fuzzy idea of success

"AI-ROI" sounds like a target, but it's too broad to steer by. Are we working on cost, revenue or both, and when should we expect to see each? Without that agreement, leaders either expect revenue too early and call a programme a failure at month three, or they settle for hours saved 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. If we stop there, we risk automating the very conversations that keep customers renewing.

Nobody decided where AI belongs

This is the one I see most often. 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. AI has to be matched to the right work, in the right way, and that takes a system. That's why I built CRAFT.

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

CRAFT is how I decide which work goes to AI, which stays with people and which stops altogether, and how to build the AI part well enough that the team chooses to use it. It sits alongside ACE (thecsconsulting.net/ace-framework), the way I run customer success, and works 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.

C	Capture the work Map what your team actually does, stage by stage, from onboarding to expansion. You can't place AI in work you can't see.
R	Rank human and machine Decide task by task what needs a person and what a machine can carry. Protect the conversations that keep customers.
A	Audit the inputs Check the data AI will read, starting with your CRM and call notes. Messy data gives you confident nonsense.
F	Fire-test the output Run it on real accounts and let the CSMs who know them judge it. Trust comes from seeing it work on their own book.
T	Track success Agree a measurable outcome, a set cadence and a plan. Without them you can't tell help from noise.

Starting small

A small team can begin with two pieces of CRAFT and learn by doing. Write down where the week goes and how long the main tasks take, so there's a baseline. Then sort those tasks into three lanes: human, assisted and automated. Switch on one or two tools in the assisted and automated lanes, check the hours after a month, and keep the human lane for the conversations that keep customers. The data checks, the testing with the team and the steering reviews earn their place as the work grows, and the appendix sets out the full programme for teams that want it.

C Capture the work

MONTH 1

Start with the team. Asking people to time every task and mark some of them to stop reads as a headcount exercise unless the leadership says otherwise, out loud and in writing, before the first question: what the inventory is for, where the freed hours go and what it means for their jobs. If roles will change, say how, and what support comes with the change.

Then map the week. With the people who do the work, write out every recurring task from onboarding to renewal and expansion: why it exists and who still uses the output, who does it, how often and how long it

takes. A lot of post-sales work survives on habit, like the weekly report built for a leader who has since moved on. Those tasks stop here, before they reach a lane, and the time they free counts from day one.

What comes out. A 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.

R Rank human and machine

MONTH 1

Three lanes. 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, routine data refreshes.

How we decide. A task stays out of the automated lane if it carries emotional weight for the customer, if the customer would care whether a person or a system did it, or if automating it would lose information or trust. Between assisted and automated, the questions are how much judgement it needs and how easily a mistake can be undone. 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.

Segment changes the budget. In a low-touch or digital segment there's far less human time to go round, so more of the work can run automated end to end, as long as there are written conditions for handing an account back to a person. As Ejieme Eromosele puts it in the ChurnZero study, segment sets how much human attention an account can justify, and the risk of each task decides how that attention is spent.

What comes out. A lane map and two or three pilots, each with a one-page card: the outcome, a named owner, the data it uses, today's baseline and the number that will tell us it worked. If the card can't be filled in, the idea isn't ready.

A Audit the inputs

MONTH 2

Check what AI will read. For each pilot, open up the data the AI would use, starting with the CRM and the call notes. Is it correct, current, complete and consistent across systems, and is it ours to use this way under our contracts and data protection rules? Does the health score measure value or only logins? Is the renewal date in the CRM the real one?

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

F Fire-test the output

MONTHS 3 TO 4

Break it before the team sees it. Before a jet engine is certified it gets broken on purpose. For each pilot we write down what good looks like in the words of the people who'll use it (for a pre-call brief, perhaps readable in two minutes, the customer's goal in their own words, one risk and one opportunity), then run it on real accounts and send it back until it meets that picture.

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. People trust what they've helped shape, so this is also where adoption is won.

T Track success

FROM MONTH 1, THEN QUARTERLY

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. Count what the programme itself costs too, so the hours given back are a net number.

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.

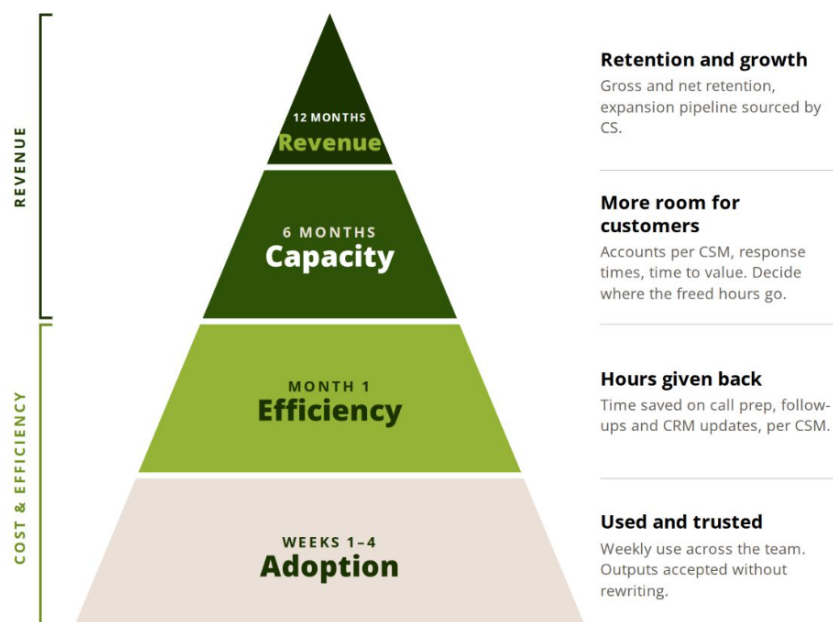
4 Measuring success: the AI success pyramid

CRAFT tells a team what to do and in which order. It doesn't tell the board when to expect a result, and that's where most programmes come unstuck. If success isn't defined before the first pilot, everyone fills the gap with their own picture: the CFO hopes for lower costs, the CRO for higher NRR, the team fears for its jobs. So before anything is built, I ask the leadership team to agree what we'll measure at each stage, when we expect to see it, how often we look and who hears about it. The pyramid is the tool I use for that conversation.

The AI success pyramid



What to expect from AI, and when.



The bottom two levels are the cost side, and they come quickly: adoption in the first four weeks, then hours given back from month one. The top two are the revenue side. Capacity shows around month six, when we

decide on purpose where the freed hours go, and the first revenue read comes at month twelve. The months mark when I'd expect to see each level first. The levels overlap in time, and each one rests on the one below.

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. My own reading of ChurnZero's numbers, and the study doesn't ask this directly, is that most teams have reached efficiency, measured it loosely and skipped capacity.

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	<p>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.</p> <p>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.</p> <p>The win to celebrate. The week everyone opens the brief before every call.</p> <p>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.</p>
MONTH 1 Efficiency Hours given back	<p>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.</p> <p>What we measure. Hours per task against the baseline from Capture, reviewed monthly with the CS leader and the sponsor.</p> <p>The win to celebrate. The month the hours given back reach a number you can name.</p> <p>The warning sign. Savings claimed without a baseline. 81% of leaders say AI saves time, and half can't measure what it's worth.</p>
MONTH 6 Capacity More room for customers	<p>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.</p> <p>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.</p> <p>The win to celebrate. The renewal conversation that now happens at 120 days, or internal meetings or the headcount line. 61% of the customer a CSM finally has time to go and see.</p> <p>The warning sign. Hours that drift into email, leaders now plan headcount around AI.</p>
MONTH 12 Revenue Retention and growth	<p>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.</p> <p>What we measure. GRR, NRR and CS-sourced pipeline, reported quarterly to the board against the baseline. With annual contracts, month twelve gives an early read rather than the full picture. To get closer to cause, compare accounts where the team spent the freed time on purpose with a similar group where it didn't.</p> <p>The win to celebrate. A renewal or expansion the board can trace back to time the team got back.</p> <p>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.</p>

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.

5 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, against 39% of the least mature, a gap the study calls statistically reliable. It's self-reported, and it doesn't prove that AI made the difference. The arrow could just as easily point the other way, because companies whose customers stay have more money to spend on AI. Still, the teams furthest along are also the ones whose customers are staying, and that's worth watching.

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.

The same honesty applies to CRAFT. It's a proposal for discussion, built from what I've seen work and fail in post-sales teams. I've run it with teams a couple of times, and it hasn't yet earned a revenue number of its own. I've written it down in enough detail that a team can test it, measure it against its own baseline and tell me where it held and where it broke. That's the conversation I'd like this paper to start.



CRAFT in practice

The simplest place to start is your own team. Put ten of its recurring tasks into the three lanes, and tell me where it held and where it broke. If your team then 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](#)

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

Source: ChurnZero, 2026 Customer Revenue Leadership Study: The State of AI in SaaS Customer Growth, October 2026 (580 CS and revenue leaders at B2B software and services companies; self-reported). All figures are from the full report. A summary is in ChurnZero's press release, "ChurnZero study finds AI investment is outpacing measurement in customer success", 6 October 2026, churnzero.com/press-release/ai-research-customer-growth.

Appendix: running CRAFT as a programme

For larger teams, or anyone who wants the full version, this is how I run CRAFT as a project: six months in four phases, then a quarterly rhythm through month twelve.

Run it 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, set out below.

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.

Rules for customer data. What each pilot can see, where that data goes and who can use the output, agreed with whoever owns security and data protection before anything goes live.

Keeping the first month manageable. A full lifecycle inventory, a baseline and a lane map is a lot for four weeks. With a large team, or more than one segment, I'd start with one segment or one stage of the lifecycle, run CRAFT on that and widen it in the quarterly lane review.

The questions I ask in Capture. 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? Which AI tools are you already using, including the ones nobody signed off?

Before we move on: the gates

Capture	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, the baseline is recorded, and the team has heard from the sponsor what the programme means for their jobs.
Rank	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.
Audit	Each pilot's data sources pass the checks, or there's a dated fix in the plan.
Fire-test	There's a written quality standard for each pilot, at least three CSMs say they'd use it as it is, nothing from one customer can surface in another customer's output, and the owner and whoever owns data protection sign it off for launch.
Track	A monthly steering review against the baseline, a decision log the board can follow, and a quarterly lane review where tasks move from human to assisted to automated once they've earned it, and occasionally back.