CRO FIELD GUIDE · REVENUE INTELLIGENCE

The Data & AI Revenue Operating System

The data points that decide whether a CRO wins — and the AI layer that turns them from reporting into a control system.

DETECT → DIAGNOSE → DECIDE → DO

AUTHOR	**Guy Karolizki** Revenue & Partnerships Executive · Strategic Advisor · karolizki.com
PURPOSE	Operating blueprint for a data- and AI-led Chief Revenue Officer mandate
SCOPE	Metric architecture · Data model · AI agent stack · Command center · Cadence · 90-day plan
DATE	November 2026

> **How to read this.** Sections 1–3 are the argument and the metrics you personally own. Sections 4–6 are the build: the data model and the AI layer that make the metrics act. Section 7 shows three full loops end to end with the arithmetic. Sections 8–11 are the operating system: dashboard, cadence, governance, and the first 90 days. The appendices are reference material — formulas, an AI-readiness audit, and a prompt library you can use on day one.

CONTENTS

What's in here

01	**The Operating Thesis**	Why the CRO job is decision latency, not reporting
02	**The Metric Architecture**	Four altitudes, and the rule that connects them
03	**The Ten Data Points That Decide the Year**	Definition, formula, failure mode, AI job — plus three nobody tracks
04	**The Data Model Underneath**	What must be true in your systems before any of this works
05	**The AI Layer**	Maturity ladder L0–L4, and where AI genuinely earns its keep
06	**The Agent Catalog**	Nine agents: inputs, triggers, guardrails, success metrics
07	**Three Loops, End to End**	Forecast rescue · silent churn · channel misallocation
08	**The Command Center**	Screen-by-screen blueprint and the morning brief
09	**Operating Cadence**	Daily to quarterly: inputs, AI pre-work, decision rights
10	**Governance & Failure Modes**	Where AI-led revenue goes wrong, and the red lines
11	**The First 90 Days**	Sequenced plan with proof metrics
A	**Appendix — Metric Definitions & Formulas**	Reference table with planning heuristics
B	**Appendix — AI-Readiness Audit**	25 items, scored
C	**Appendix — CRO Prompt Library**	Twelve prompts to run against your revenue data

> THE ONE-LINE VERSION
>
> A CRO does not win by having more dashboards. A CRO wins by shortening the distance between *a revenue event happening* and *a decision being made about it* — and by being right more often when that decision is made. Every metric and every agent in this document is justified only if it moves one of those two things.

SECTION 01

The Operating Thesis

Three claims that determine how you should spend the first quarter.

Claim 1 — Your real KPI is revenue decision latency

Every revenue organisation has a lag between the moment something changes in the market and the moment someone with authority does something about it. In most companies that lag is measured in weeks: the signal appears in a system on day 1, surfaces in a report on day 12, gets debated in a QBR on day 26, and produces an action on day 34. By then the quarter has decided itself.

Everything a data-and-AI CRO builds should be evaluated against a single question: **does this shorten the lag, or does it just make the lag better documented?** Most BI investments do the second. Track your own latency explicitly — see metric 12 in Section 3 — because it is the metric that governs all the others.

Claim 2 — Most revenue orgs measure at the wrong altitude

Two symmetrical failures are common. The first is the executive who watches only ARR, bookings and win rate — metrics so aggregated that by the time they move, the cause is three months old and untraceable. The second is the operator drowning in 140 CRM fields, none of which is connected to a decision anyone is authorised to make.

The fix is a strict hierarchy: five board metrics, roughly fifteen CRO operating metrics, a diagnostic layer owned by function leaders, and a deal-level layer owned by managers. Each level must *decompose* into the one below it, with a named owner at every node. If a metric moves and you cannot name the person whose behaviour changed it, the metric is decoration.

Claim 3 — AI's value is the loop, not the chart

The instinct is to use AI to make dashboards prettier or to write narrative summaries of charts. That is a marginal gain. The structural gain is collapsing the four-step loop that currently costs you weeks:

STEP	TODAY, WITHOUT AI	WITH AN AI REVENUE LAYER	LATENCY SHIFT
Detect	A human notices an anomaly in a report, usually late and usually only in the metrics they personally watch	Continuous anomaly detection across every segment, channel, cohort and rep — including the slices no one thought to look at	Weeks → hours
Diagnose	Analyst pulls data, builds a cut, presents in a meeting; three competing narratives emerge	Automated decomposition of the variance into contributing factors with the evidence chain attached	Days → minutes
Decide	Debate, opinion, seniority; the loudest hypothesis usually wins	Ranked options with expected value, confidence and cost — the human still decides, but on arithmetic	Weeks → one meeting
Do	Action items decay in a spreadsheet; nobody measures whether the fix worked	Agent executes the mechanical parts (lists, briefs, playbook updates, alerts) and instruments the outcome	Partial → tracked

THE SHIFT IN ONE SENTENCE

The old model is **Data → Dashboard → Human analysis → Meeting → Decision → Action**. The new model is **Data → AI detects → AI diagnoses → AI ranks options → Human decides → Agent executes → Outcome measured**. Note that the human moves *later* in the chain and their judgement gets applied to a better-framed choice. That is the entire prize.

The 60-second test

Before building anything, benchmark yourself. A data-and-AI CRO should be able to answer these five questions in under a minute, from one surface, with confidence intervals — not by opening seven dashboards and calling two analysts.

#	QUESTION	WHAT A GOOD ANSWER SOUNDS LIKE
1	Are we going to make the quarter, and how sure are you?	"$11.2M against $12.0M at 84% confidence; $1.4M of the gap is concentrated in 14 late-stage enterprise deals."
2	If we miss, what will have caused it?	"Enterprise pipeline creation has run 27% below plan for six weeks; SMB is at plan. Cause is the SDR territory re-cut, not demand."
3	Where is the next dollar of growth cheapest?	"Partner-sourced mid-market in DACH: 34% win rate, 11-month CAC payback, 118% NRR — versus 23 months payback in the comparable direct motion."
4	Which customers are quietly leaving?	"Eleven accounts, $2.4M ARR, each showing three or more of the five leading churn signals; renewal dates 60–120 days out."
5	Which three decisions should *I personally* make this week?	A ranked list with expected value attached to each, not a list of everything that is broken.

If you cannot answer all five today, the gap is almost never a shortage of data. It is that the data is not modelled around decisions. Section 4 fixes that.

SECTION 02

The Metric Architecture

Four altitudes, one rule: every metric must decompose into named drivers with named owners.

ALTITUDE 0 — BOARD	ALTITUDE 1 — CRO	ALTITUDE 2 — FUNCTION	ALTITUDE 3 — DEAL / REP
5	**15**	**~40**	**n/a**
Metrics · Monthly	Metrics · Weekly	Diagnostics · Weekly	Records · Continuous

ALTITUDE	AUDIENCE	METRICS	THE AI JOB AT THIS LEVEL
0 · Outcome	Board, CEO	ARR / NNARR growth · Net Revenue Retention · Gross margin-adjusted CAC payback · Rule of 40 contribution · Forecast accuracy (trailing 4 quarters)	Explain variance to plan in plain language, with attribution to Altitude 1 drivers
1 · Control	CRO	Qualified pipeline creation vs. requirement · Coverage ratio · Weighted forecast + confidence · Stage conversion · Sales cycle · Win rate by segment · Slippage & no-decision rate · ASP · Logo & revenue churn · Expansion ARR · Partner-sourced share · Rep attainment distribution · Ramp time · Cost of revenue motion · Decision latency	Detect anomalies, rank what deserves your attention this week, quantify the gap in dollars
2 · Diagnostic	VP Sales, VP Marketing, VP CS, VP Partners, RevOps	Channel-level CAC and conversion · Lead-to-MQL-to-SQL by source · Meeting set rate · Demo-to-proposal · Proposal-to-close · Discount depth · Loss reasons · Support ticket volume by account · Product adoption depth · Partner activation and productivity · Territory coverage · Quota capacity model	Decompose the Altitude 1 anomaly into contributing factors; separate mix effects from performance effects
3 · Action	Managers, AEs, CSMs	Individual deal health · Buying-group coverage · Days since last meaningful buyer activity · Next-step quality · Account health composite · Renewal readiness	Prioritise the list, draft the intervention, flag the risk before the human sees it

THE DECOMPOSITION RULE

No Altitude 0 metric may exist without a documented decomposition into Altitude 1 drivers, each with a single named owner. When NRR drops 4 points, you should be able to walk down the tree in ninety seconds: NRR → (gross retention × expansion rate) → gross retention → (logo churn × downsell) → downsell concentrated in one segment → concentrated in accounts that adopted fewer than three modules → owner: VP Customer Success, with a specific adoption programme already in flight. Without the tree, AI cannot diagnose either, because the causal structure does not exist anywhere except in people's heads.

What to deliberately *not* measure at your altitude

Activity metrics — calls, emails, tasks logged — belong at Altitude 3 and nowhere else. The moment a CRO watches activity volume, the organisation optimises for activity volume, and the metric stops carrying information. Watch *activity-to-opportunity conversion* instead: it is a quality signal and it is much harder to game. The same logic applies to "pipeline generated" without a quality filter, and to average deal size without the distribution behind it.

SECTION 03

The Ten Data Points That Decide the Year

Not the sixty you could track. The ten that carry the most decision weight, plus three that almost nobody tracks and should.

01 Net Revenue Retention (NRR)

FORMULA

```
NRR = (Starting ARR + Expansion − Downsell − Churn) ÷ Starting ARR  ·
measured on a fixed cohort, trailing 12 months
```

WHY IT WINS

It is the only metric that compounds. At 120% NRR you grow without selling anything new; at 95% you are running up a down escalator and every acquisition dollar is partly replacing what leaked. It also drives valuation more than growth rate at most scales.

FAILURE MODE

Blended NRR hides everything. A healthy enterprise cohort at 128% masks an SMB cohort at 84%. Always cut by segment, by acquisition channel, and by cohort start quarter — never look at the blended number alone.

AI JOB

Predict account-level NRR 2–3 quarters forward from usage, support, and engagement signals; identify which product-adoption paths correlate with expansion so CS can be pointed at them deliberately.

02 Qualified Pipeline Creation Rate vs. Requirement

FORMULA

```
Weekly requirement = (Target bookings ÷ Win rate) ÷ Weeks of pipeline lead
time  · compare to actual new qualified pipeline per week
```

WHY IT WINS

This is the earliest reliable leading indicator you have. Because pipeline created today closes one or two quarters out, a sustained shortfall is a future miss that is already locked in — and it is visible months before it shows up in bookings. It is the single number I would put on the wall.

FAILURE MODE

Measuring raw pipeline dollars instead of *qualified* pipeline. Reps inflate stage-1 to hit creation targets. Define qualification by observable evidence — budget confirmed, buying group identified, next meeting scheduled with a named economic buyer — not by rep assertion.

AI JOB

Score every new opportunity for qualification integrity against historical patterns; alert when creation runs below the required rate for two consecutive weeks in any segment, and translate the shortfall into dollars of future bookings at risk.

03 Forecast Confidence & Trailing Forecast Accuracy

FORMULA

```
Accuracy = 1 − |Actual − Commit| ÷ Commit  · tracked per quarter, per forecast level, per manager
```

WHY IT WINS

A forecast without a confidence interval is a wish. Reporting "$12.4M at 82% confidence, $1.1M at risk, concentrated here" changes the conversation from optimism management to risk management. And trailing accuracy per manager tells you whose judgement to weight — which is enormously valuable and almost never measured.

FAILURE MODE

Sandbagging and hockey-sticking. Both are visible if you track the *shape* of each manager's forecast movement through the quarter, not just the endpoint.

AI JOB

Produce an independent model forecast from deal attributes and behavioural signals, then explicitly show model-vs-rep delta per deal. The delta list is your pipeline review agenda.

04 Stage Conversion × Velocity Matrix

FORMULA

```
For each stage transition: Conversion % × Median days in stage × Volume × Average value
```

WHY IT WINS

Revenue leaks at handoffs. Looking at conversion alone tells you where deals die; looking at velocity alone tells you where they stall; the matrix tells you where the *money* is being lost, which is the only version that supports a resource decision. A 6-point drop in a stage handling $18M is a very different problem from the same drop in a stage handling $2M.

FAILURE MODE

Stage definitions that describe seller activity rather than buyer commitment. If a stage can be advanced without the buyer doing anything, the whole matrix is noise.

AI JOB

Continuously test every transition for statistically meaningful shifts; separate mix effects (the pipeline composition changed) from performance effects (the same deals now convert worse). This distinction is where most human analysis fails.

05 Win Rate by Segment × Structured Loss Reasons

FORMULA

```
Win rate = Won ÷ (Won + Lost + No-decision)  · never exclude no-decision, it is the largest competitor you have
```

WHY IT WINS

Win rate is the multiplier on everything else — a 4-point improvement is worth more than a 20% increase in lead volume in most models. But the win rate number alone is useless without a disciplined loss taxonomy: price, product gap, no decision, competitor, timing, wrong fit, lost champion.

FAILURE MODE

Free-text loss reasons, or a "price" bucket that absorbs every loss because it is the least uncomfortable thing to write. Price is almost never the real reason; it is the reason offered when value was not established.

AI JOB

This is where AI is genuinely transformative — it can read call transcripts, emails and proposals and classify loss reasons from evidence rather than from rep self-report, then quantify how often specific objections appear in lost versus won deals.

06 CAC Payback, by Channel and by Segment

FORMULA

```
CAC payback (months) = Fully-loaded S&M cost to acquire ÷ (New ARR × Gross margin) × 12
```

WHY IT WINS

It is the metric that decides where your next headcount and next marketing dollar goes. Growth without payback discipline is just financed decline. Fully-loaded matters: include SDR, AE, sales engineering, marketing programme spend, partner margin, and the allocated cost of onboarding.

FAILURE MODE

Blended CAC across channels that behave completely differently. Partner-sourced, outbound and inbound have different cost structures, different win rates and different retention — averaging them destroys the signal you were trying to extract.

AI JOB

Maintain live multi-touch attribution, model marginal (not average) CAC as you scale each channel, and simulate reallocation scenarios before you commit budget.

07 Slippage and No-Decision Rate

FORMULA

```
Slippage = Value of deals whose close date moved out ÷ Value of deals forecast at period start  · No-decision = deals closed-lost with no competitor selected
```

WHY IT WINS

Slippage is the quiet killer of forecasts and the earliest signal that qualification is weak. A high no-decision rate almost always means you are selling to people who cannot fund the change, not that your product lost a bake-off. These are different problems with different fixes.

FAILURE MODE

Allowing close dates to be edited freely without an audit trail. If you cannot see the history of date changes, slippage is invisible by construction.

AI JOB

Predict slippage probability per deal from engagement patterns — days since last buyer-initiated contact, buying group breadth, whether a mutual action plan exists — and surface the top-N intervention list weekly.

08 Attainment Distribution and Ramp Curve

FORMULA

```
% of reps at ≥100% quota · median attainment · P25/P75 spread · months to 80% of full productivity
```

WHY IT WINS

Average attainment is one of the most misleading numbers in the business. 100% average with two reps at 300% and eleven below 60% is a broken engine wearing a good result. The distribution tells you whether you have a repeatable motion or a few heroes — and only a repeatable motion can be scaled with hiring.

FAILURE MODE

Hiring against a plan that assumes new reps perform like tenured ones. Ramp time is a capacity constraint; if your ramp is 7 months and your average tenure is 19 months, roughly a third of your capacity is permanently ramping.

AI JOB

Detect performance anomalies at the stage level per rep (not just totals), identify which behaviours separate top and bottom quartiles from conversation data, and shorten ramp with targeted coaching rather than generic enablement.

09 Composite Customer Health / Leading Churn Signal

FORMULA

```
Weighted composite of: usage trend · breadth of adoption · support burden & sentiment · champion presence · executive engagement recency · invoice/payment behaviour
```

WHY IT WINS

Churn is a lagging metric you cannot act on. The composite is a leading one you can. The critical design choice: weight the components by their historically observed correlation with actual churn in *your* base, not by intuition. Most health scores fail because someone assigned the weights in a workshop.

FAILURE MODE

Scores that are never validated. Every quarter, check calibration: of accounts scored high-risk 90 days ago, what fraction actually churned? If it is near your base rate, the score is theatre.

AI JOB

Learn the weights from outcomes, add unstructured signals (support tone, ticket themes, review sites, champion job changes), and give you 60–120 days of notice instead of 30.

10 Partner-Sourced Share and Partner Productivity

FORMULA

```
Sourced (partner originated) vs. Influenced (partner involved) — reported separately, never summed · Productivity = qualified pipeline per active partner per quarter
```

WHY IT WINS

Channel changes your unit economics more than almost any other lever: lower CAC, higher win rate when the partner owns the relationship, and frequently better retention because the partner does the implementation. It is also the most systematically mis-measured area in revenue, which means there is usually real value sitting in a correct measurement.

FAILURE MODE

Double counting. When sourced and influenced both roll up into "partner revenue," partner contribution is routinely overstated by 40–60%, and the resulting investment decisions are wrong. Enforce a strict attribution hierarchy and audit it.

AI JOB

De-duplicate attribution automatically, segment partners into produce / activate / retire tiers from behaviour rather than tier badges, and predict which newly signed partners will actually transact based on early activity patterns.

The three nobody tracks — and you should

METRIC	DEFINITION	WHY IT MATTERS TO YOU SPECIFICALLY
11 · Cost of Revenue Motion	Fully-loaded GTM cost to acquire *and serve* a customer in a given segment, expressed as % of that customer's first-year revenue	CAC stops at acquisition. Cost of motion includes onboarding, support load, success coverage and renewal effort. Segments that look attractive on CAC frequently become unattractive here — this is how you avoid scaling a segment that loses money after the sale.
12 · Revenue Decision Latency	Median days from signal availability in a source system to a documented decision by an accountable owner	The meta-metric of this entire document. Sample ten decisions a quarter and time them. If AI investment does not move this number, it is not working, whatever the demo looked like.
13 · Data Completeness Score	% of open opportunities with all decision-critical fields populated, current, and consistent with observed activity	Your AI layer's ceiling. A model trained on 40%-complete CRM data will produce confident nonsense. Publish this score weekly by team — visibility alone typically moves it 20+ points in a quarter, and it is the cheapest AI investment you will ever make.

SECTION 04

The Data Model Underneath

Every AI failure in revenue that I have seen was a data-model failure wearing an AI costume. This is the unglamorous section that decides whether the rest works.

AI does not need *more* data. It needs data that is joined, timestamped, and honest. Six object families carry almost all of the decision weight — if these are clean and connected, a competent model can do remarkable things.

OBJECT	FIELDS THAT MUST BE RELIABLE	WHAT IT UNLOCKS
Account	Firmographics (size, industry, geography) · segment assignment with history · acquisition channel and source partner · contract dates · ARR history with change reasons · parent/child hierarchy	Cohort analysis, NRR by segment, CAC by channel, expansion targeting. Segment *history* matters: without it, accounts silently reclassify and your cohorts drift.
Buying group	Contacts mapped to roles (economic buyer, champion, technical evaluator, blocker) · engagement recency per role · seniority · departure flag	The strongest single predictor of both win rate and churn is buying-group breadth. Single-threaded deals lose. Single-threaded accounts churn when one person leaves.
Opportunity	Stage with entry/exit timestamps · close date change history · amount change history · qualification evidence · competitor · structured loss reason · next step with a date · mutual action plan present y/n · discount and approval path	Everything in Section 3. The history fields are non-negotiable — a current-state-only CRM cannot support slippage, velocity or forecast-accuracy analysis at all.
Activity & conversation	Meetings, emails, calls with direction (inbound/outbound) and participants · call transcripts · buyer-initiated flag · response latency	Deal health, coaching, objection analysis, loss classification from evidence. *Buyer-initiated* is the highest-signal field on this list and is almost never captured.
Product usage	Active users vs. licensed · feature/module adoption breadth · usage trend at 30/60/90 days · admin activity · integration health · error rates	Leading churn signals, expansion triggers, and product-qualified pipeline. For any product-led motion this is the highest-value dataset you own and typically the least joined to CRM.
Cost & capacity	Fully-loaded cost per role · marketing programme spend by channel and segment · partner margin and MDF · quota capacity, ramp status, territory coverage	CAC payback, cost of revenue motion, capacity planning, scenario simulation. This lives in Finance and is the join most companies never make — which is precisely why efficiency decisions get made on instinct.

The six hygiene rules that make AI possible

1. **One definition per metric, published, versioned.** If Sales and Finance compute win rate differently, every AI output will be disputed on definitional grounds and the project dies in a meeting.
2. **Timestamp everything; never overwrite.** Stage changes, amount changes, date changes and owner changes must be append-only. History is where causality lives.
3. **Stages defined by buyer evidence, not seller activity.** "Demo completed" is seller activity. "Buyer has confirmed budget and named an evaluation owner" is evidence. Only the second is predictive.
4. **Mandatory structured fields at stage gates, not at creation.** Ask for competitor and loss reason where the rep actually knows them. Forcing fields at creation produces garbage that pollutes everything downstream.
5. **Every derived metric traceable to source rows.** When AI says win rate fell 9 points, a human must be able to click through to the exact 41 opportunities. Without traceability there is no trust, and without trust nobody acts on the output.
6. **Publish the Data Completeness Score weekly by team.** Make hygiene a visible operating metric with a named owner, not a quarterly clean-up project.

THE SEQUENCING MISTAKE TO AVOID

Do not buy an AI revenue platform in month one. Buy it in month three, after the data model is defined and the hygiene score is above roughly 80%. Every deployment I have seen fail followed the same pattern: sophisticated tooling pointed at a CRM where 45% of opportunities had a stale next step, no competitor field, and a close date that had been dragged forward eleven times. The tool produced confident, wrong answers, the team stopped trusting it within a quarter, and the budget became very hard to defend the second time.

Where each dataset lives — and who has to hand it to you

Half of this work is organisational, not technical. Most CROs discover in month two that the data they need most is owned by someone who has no incentive to prioritise them. Make the asks explicit and early.

DATASET	USUALLY OWNED BY	WHAT TO ASK FOR, SPECIFICALLY
Product usage	Product / Engineering	A nightly feed of active users, module adoption and integration health, keyed to the CRM account ID. The join key is the whole negotiation — agree it in writing before anyone builds anything.
Support & sentiment	Support / CS	Ticket volume, theme tags, resolution time and sentiment by account. Ask for the free text too, not just the tags — the tags were designed for support workflow, not for churn prediction.
Fully-loaded cost	Finance	Cost per role by segment, programme spend by channel, partner margin and MDF. This is the join that makes efficiency decisions possible and it is the one most often refused on the grounds of confidentiality — escalate it early, with the CFO as an ally rather than a gatekeeper.
Conversation data	Sales / IT / Legal	Transcripts with consent flows per jurisdiction. Legal is on the critical path here, so start the conversation in week one even though you will not use the data until quarter two.
Marketing attribution	Marketing	Touch-level data, not campaign summaries. Summaries have already collapsed the detail you need to model marginal CAC.

SECTION 05

The AI Layer

A maturity ladder you can locate yourself on, and an honest account of where AI helps and where it will mislead you.

LEVEL	CAPABILITY	WHAT IT LOOKS LIKE IN PRACTICE	PREREQUISITE
L0 Report	Descriptive dashboards	Humans open dashboards, notice things, and ask questions. Where most organisations actually are, regardless of what they say.	A data warehouse
L1 Detect	Anomaly detection & narration	The system tells you what changed, across every slice, without being asked — including slices no human would have thought to check. Fastest, cheapest win available.	Clean joined metrics, defined thresholds
L2 Diagnose	Causal decomposition	Variance is broken into contributing factors with an evidence chain; mix effects separated from performance effects. Turns a meeting into a review.	Historical depth, segment history, cost joins
L3 Decide	Ranked options with expected value	"Three interventions available: A recovers ~$230K at 4% margin cost; B recovers ~$180K but pulls Q1 revenue forward; C fixes the structural cause but not this quarter."	Outcome history to calibrate uplift estimates
L4 Do	Bounded agent execution	Agents execute the mechanical steps within explicit limits: build the target list, draft the brief, update the battlecard, schedule the review, instrument the outcome.	Governance, permissions, audit log, red lines

Sequence matters and skipping is expensive. L1 → L2 → L3 → L4 in order, per use case. A common and costly error is buying an L4-marketed platform while operating at L0 data maturity.

WHERE AI GENUINELY EARNS ITS KEEP

- **Anomaly detection at scale.** Watching 400 metric-slice combinations continuously is something no human team can do, and it is where most early value lands.
- **Unstructured → structured.** Turning calls, emails, tickets and reviews into loss reasons, objection frequencies, churn signals and competitive intelligence. Historically impossible; now routine.
- **Prioritisation.** Ranking 340 open deals or 1,200 accounts by expected value of intervention. Reliable, high-leverage, low-risk.
- **Drafting.** Briefs, battlecards, QBR narratives, renewal talk tracks — the work that currently consumes your team's evenings.
- **Simulation.** "If we move two AEs to DACH, what happens to coverage?" Fast, cheap, and better than the spreadsheet that currently answers it.

WHERE IT WILL MISLEAD YOU

- **Causal claims from correlational data.** AI will confidently tell you that deals with more calls close better. Direction of causation is usually the reverse. Treat every "why" as a hypothesis to test.
- **Small-N enterprise forecasting.** With 40 deals a quarter and $500K ASPs, there is not enough statistical power. Use AI for deal-level risk flags, not for portfolio probability, in that regime.
- **Regime change.** Models trained on last year's motion break when you change pricing, segment or channel — exactly when you most need them. Retrain and widen intervals after any material GTM change.
- **Garbage-in confidence.** AI does not know your CRM is fiction. It will produce a beautifully argued answer from bad inputs, and the fluency makes it harder to catch.
- **Goodhart effects.** The moment a score drives comp or promotion, people optimise the score. Assume gaming and design against it.

SECTION 06

The Agent Catalog

Nine agents worth building, in the order I would build them. Each has an owner, a guardrail, and a metric that proves it is working.

AGENT	INPUTS	OUTPUT & TRIGGER	GUARDRAIL
1 · Pipeline Integrity BUILD FIRST	Opportunity records, stage history, activity, buyer-initiated flags	Weekly: list of deals failing qualification integrity, stale next steps, single-threaded deals, and deals aged beyond stage median. Feeds the pipeline review agenda directly.	Flags for review only; never auto-changes stage or amount
2 · Forecast BUILD FIRST	Deal attributes, historical close patterns, rep forecast history, engagement signals	Continuous: independent forecast with confidence band; alert when model-vs-rep delta exceeds threshold at deal or team level.	Model forecast is presented alongside the human one, never replaces it
3 · Churn & Expansion BUILD FIRST	Usage, support tickets and sentiment, engagement recency, contact departures, invoice behaviour	Weekly: ranked risk list with value at risk, reason codes, and days to renewal; separate ranked expansion-signal list.	Calibration published quarterly; no automated customer contact
4 · Deal Coach	Call transcripts, emails, proposals, win/loss outcomes	Per deal: risk factors, objection patterns, buying-group gaps, recommended next action. Per rep: the two behaviours most correlated with their conversion gap.	Coaching input to managers; never a direct performance-management input
5 · Win/Loss Analyst	All closed opportunities plus their conversation history	Monthly: evidence-based loss taxonomy, objection frequency trend, competitor win-rate movement, and what changed versus prior period.	Classifications sampled and human-audited monthly
6 · GTM Allocation	Revenue, CAC, payback, NRR and capacity by segment / geo / channel / product	Quarterly, or on demand: marginal-return ranking of investment options with simulated reallocation scenarios.	Recommends only; headcount and budget stay human decisions
7 · Partner Performance	Partner-sourced vs. influenced pipeline (de-duplicated), activation dates, deal registration, margin	Monthly: partner tiering by actual productivity, activation predictions for new partners, retire/invest recommendations.	Attribution hierarchy fixed and auditable; no silent re-attribution
8 · Pricing & Packaging	Discount depth, approval paths, win rate by price point, competitor pricing mentions, expansion behaviour	Quarterly: discount leakage analysis, price-elasticity signal by segment, packaging experiment proposals.	Never sets prices; floors and approvals remain human-controlled
9 · Chief of Staff COMPOSER	Output of all agents above	Daily 07:00: the morning brief — three things requiring attention, three positive signals, one recommendation with expected value. Ranked, deduplicated, capped at one page.	Hard cap on item count; every claim links to source records

> DESIGN PRINCIPLE
>
> An agent that produces a long list has failed. The scarce resource is your attention, not information. Every agent must rank, cap, and attach an expected value to what it surfaces — and it must be willing to say "nothing this week requires you." A system that always finds three urgent things is a system you will stop reading by month two.

Build sequence

Agents 1, 2 and 3 in the first quarter — pipeline integrity, forecast, churn. They pay for the programme and they need only the data you already have. Agents 4 and 5 in the second quarter, once conversation capture and the loss taxonomy are in place. Agents 6, 7 and 8 in the second half, once the cost and partner joins are complete. Agent 9 last, because it is only as good as what feeds it.

SECTION 07

Three Loops, End to End

Illustrative but arithmetically complete. The point is the shape of the reasoning, not the specific numbers.

Loop A — The quarter that was already lost in week four

Q4 TARGET	COVERAGE	MODEL FORECAST	EXPOSURE
$12.0M	**2.6×**	**$10.9M**	**$1.1M**
New bookings	vs. 3.8× required	76% confidence	Concentrated, not diffuse

DETECT · WEEK 4, AUTOMATICALLY

Enterprise qualified pipeline creation has run at $1.9M per week against a $2.6M plan for six consecutive weeks — a 27% shortfall. SMB and mid-market are at plan. No human had flagged it because total pipeline creation, blended, was only 8% below plan and looked like noise.

DIAGNOSE · EVIDENCE CHAIN ATTACHED

The shortfall is not demand. Outbound meeting-set rate in the enterprise segment fell from 4.1% to 2.6% beginning the week after the SDR territory re-cut. Inbound enterprise volume is flat. Separately, three of eleven enterprise AEs are inside 90 days of ramp and carrying full quota-equivalent territory.

QUANTIFY · TWO DISTINCT PROBLEMS, DIFFERENT TIME HORIZONS

Six weeks × $0.7M = $4.2M of missing pipeline ≈ $1.1M of future bookings at the 26% enterprise win rate. But with a 74-day median cycle, almost all of that lands in Q1 — it is a Q1 problem being detected in Q4. The *Q4* exposure is different: 14 late-stage deals worth $2.1M with no buyer-initiated activity in 14+ days, 9 of them past the median age for their stage.

DECIDE · RANKED BY EXPECTED VALUE, NOT BY URGENCY

Option A — executive intervention on the 14 stalled deals. Historic uplift on exec-sponsored late-stage enterprise deals is roughly +11 points of close rate → expected recovery ≈ $230K. Cost: your calendar.
Option B — pull forward $1.4M of Q1 renewals and expansions with a co-term offer. Recovers in-quarter revenue at a 4% discount ≈ $56K of margin, but borrows from Q1.
Option C — reverse the SDR territory re-cut and re-establish enterprise coverage. Does nothing for Q4; it is the only option that fixes Q1.
The correct answer is A and C, and explicitly *not* B unless the quarter is a board commitment — because B disguises the structural problem for one more quarter.

DO · INSTRUMENTED

Agent builds the 14-deal list with a one-page brief per deal, books the exec sessions, and sets a measurement window. Territory change reverts with a named owner and a weekly meeting-set-rate check. Outcome: quarter closes at $11.6M (97% of target) and enterprise pipeline creation returns to plan by week six — protecting Q1, which was the larger number at stake all along.

WHAT MADE THE DIFFERENCE

Not prediction. **Detection at week four instead of week ten**, and a diagnosis precise enough to separate a Q4 problem from a Q1 problem. Those are two different interventions, and organisations without this decomposition routinely apply the Q4 fix to the Q1 problem and wonder why the pattern repeats.

Loop B — The customer who was leaving quietly

```
ACCOUNT AT RISK — $420K ARR — renewal in 78 days
Composite churn probability: 74%   (base rate for this signal set: 61%, n=38)

SIGNAL                                  OBSERVED            CONTRIBUTION
Weekly active seats                     −31% over 60d       ▓▓▓▓▓▓▓▒▒  high
Admin / champion login                  absent 41 days      ▓▓▓▓▓▓▓▒▒  high
Support tickets                         +42%, 3 on one      ▓▓▓▓▓▒▒▒▒  medium
                                        integration
Competitor named in support chat        2 instances         ▓▓▓▓▓▒▒▒▒  medium
Executive engagement                    none in 94 days     ▓▓▓▒▒▒▒▒▒  medium
Expansion pipeline in same logo         $180K, stalled      ▓▓▓▒▒▒▒▒▒  context

VALUE AT RISK: $420K ARR + $180K expansion pipeline = $600K
RECOMMENDED: exec sponsor call within 5 days · SA on the integration defect ·
             24-month renewal with usage ramp · convert the 3 tickets into a
             documented roadmap commitment
```

What happened. None of these signals individually would have triggered action. The CSM saw rising tickets and read it as engagement. Product saw usage decline in a dashboard nobody reviewed weekly. Nobody knew the champion had left until the renewal call. The composite score existed because the weights were learned from 38 historically similar accounts, not assigned in a workshop.

Outcome. Renewed at $390K on a 24-month term — a 7% concession against a $600K exposure. The expansion pipeline reactivated the following quarter once the integration defect was closed.

> **THE REAL DELIVERABLE WAS TIME**
>
> The model did not predict something a good CSM could not have inferred. It produced the inference **78 days before the renewal instead of 12**, and it did so for all 340 accounts simultaneously rather than for the handful someone happened to be worried about. Notice is the product.

Loop C — The channel that was 40% smaller than reported

MEASURE	PARTNER MOTION	DIRECT MOTION	IMPLICATION
Reported share of pipeline	38%	62%	The number the business had been planning against
After de-duplication	23%	77%	Influenced deals were being counted as sourced; contribution overstated by ~65%
Fully-loaded CAC	$18K	$31K	Partner motion remains materially cheaper even after correction
Win rate	34%	26%	Partner-originated deals arrive pre-qualified and warm
ASP	−22%	baseline	Smaller deals; the economics still work because of CAC and win rate
NRR of acquired logos	118%	106%	Partners implement, so adoption depth is higher — the strongest argument in the whole table
Concentration	Top 6 of 42 partners = 61% of sourced pipeline		28 partners produce 9% combined and consume two full-time partner managers

The decision. Two conclusions fell out that were invisible before the de-duplication. First, partner contribution was smaller than believed — an uncomfortable correction. Second, partner *economics* were better than believed, because the corrected denominator made CAC, win rate and NRR look stronger, not weaker. The reallocation: move two partner managers off the 28-partner long tail onto a structured co-sell motion with the top six, plus an activation programme for eight tier-two partners the model identified as high product-fit and low activity.

The metric that proved it. Qualified partner-sourced pipeline per partner manager per quarter, which moved from $2.1M to $3.4M over two quarters — with no additional headcount.

What the three loops have in common

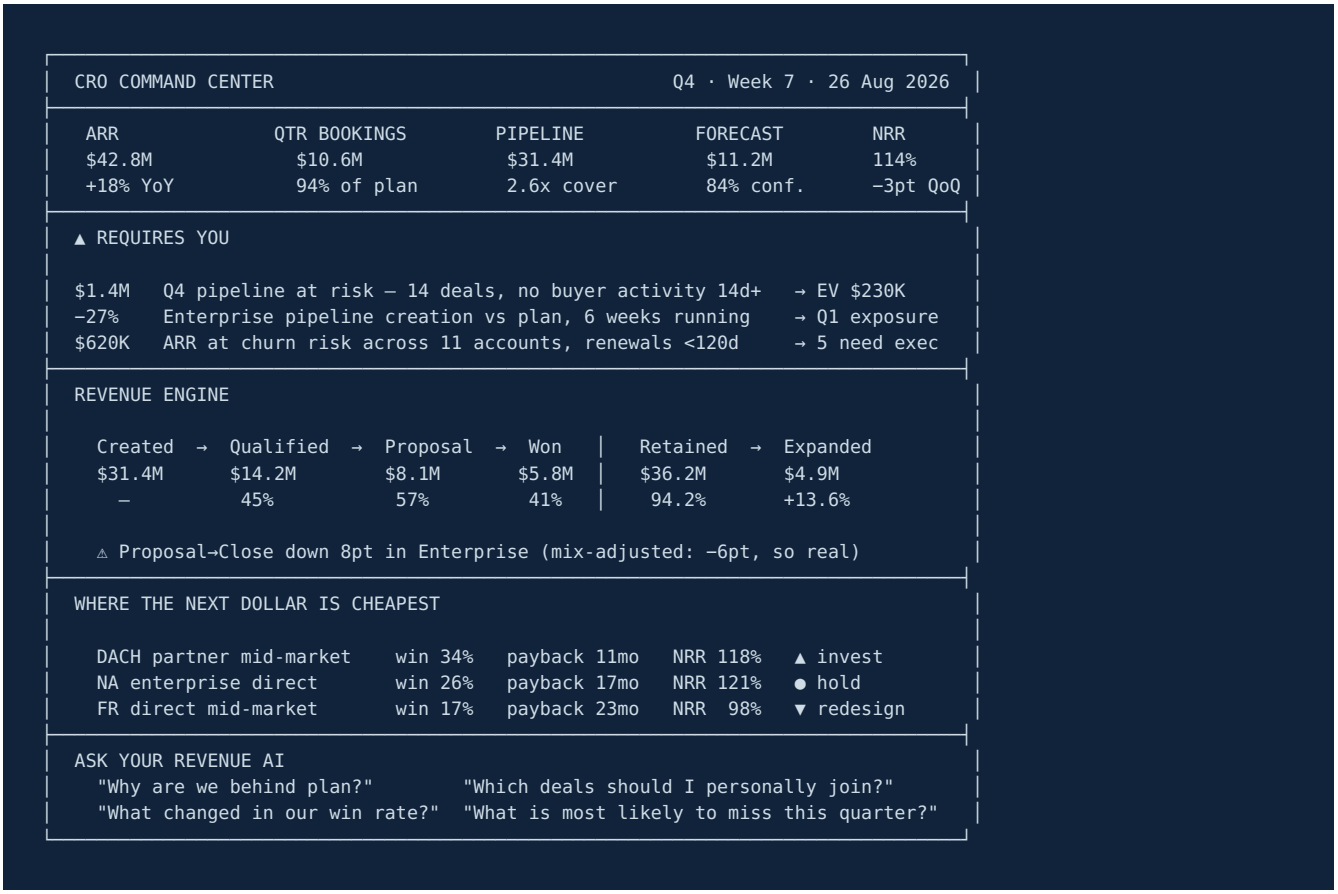
Read together, they make the same point three ways, and it is worth stating explicitly because it determines where you should spend money.

- **The intelligence was already in the building.** In all three cases, some person or system held part of the answer. What was missing was the join, the ranking, and the notice.
- **Detection timing decided the outcome, not prediction accuracy.** A 74% churn score at 78 days is worth far more than a 90% score at 12 days. Optimise your programme for earliness before you optimise it for precision.
- **The correction was uncomfortable before it was useful.** The partner numbers got worse before the partner strategy got better. Expect the first honest measurement of any area to be a bad-news conversation, and pre-frame it with the CEO so that accuracy does not get read as underperformance.

SECTION 08

The Command Center

One home screen that answers the five questions, and four drill-downs behind it.



SCREEN	ANSWERS	REFRESH	PRIMARY OWNER
Home	Are we going to make it, what needs me, where is growth cheapest	Daily 07:00	CRO
Pipeline Command	Creation vs. requirement, coverage, quality, aging, slippage, model-vs-rep forecast delta	Real time	VP Sales / RevOps
Funnel & Conversion	Every transition as volume × conversion × time × value, mix-adjusted	Daily	RevOps
Customer Revenue	NRR by cohort, health distribution, risk list, expansion signals, renewal calendar	Daily	VP Customer Success
GTM Efficiency	Revenue × growth × conversion × CAC × margin by segment, geo, channel, product	Weekly	CRO / Finance
People & Capacity	Attainment distribution, ramp curve, capacity vs. plan, coaching priorities	Weekly	VP Sales

The morning brief

The single highest-leverage artefact in the system, and the one to build first because it forces every other component to be honest.

```
Good morning. Three things changed that matter.

    REQUIRES A DECISION
1.  $840K enterprise deal moved commit → upside. Buyer's CFO joined the last call
    and pricing was reopened. Model close probability fell 71% → 44%.
    → Recommend: you join the Thursday call. Similar interventions: +11pt close rate.

2.  EMEA pipeline creation 14% below weekly target, third consecutive week.
    Cause: two of five SDRs are unramped, not demand. Inbound volume is flat.
    → Recommend: decide on contractor coverage this week. Every week costs ~$180K
      of Q1 pipeline.

3.  Customer XYZ ($420K) crossed high churn risk. Champion departed, usage −31%.
    → Recommend: exec sponsor call within 5 days.

    WORTH KNOWING
•   Partner channel generated $420K new qualified pipeline, best week of the quarter.
•   Enterprise win rate +8pt over trailing 60 days, driven by the new security
    battlecard — the pattern now holds across 3 reps, so it is likely real.
•   Expansion revenue tracking 17% above plan; DACH is the source.

    ONE RECOMMENDATION
    Temporarily move two enterprise AEs to EMEA coverage. Based on current pipeline
    and conversion, this recovers approximately $600K of Q4 pipeline at the cost of
    ~$220K of NA pipeline creation. Net positive, reversible in 30 days.
    Confidence: medium. Assumption: EMEA conversion holds at trailing-90-day rate.
```

Three design rules make this work. **Cap the items** — three red, three green, one recommendation, always, even on a quiet day. **Attach a number to every recommendation**, including the cost and the assumption it rests on. **Link every claim to source records**, so the first instinct on a surprising line is to click through rather than to distrust the system.

SECTION 09

Operating Cadence

The rhythm that converts intelligence into decisions. AI does the pre-work; humans spend the meeting deciding, not discovering.

CADENCE	FORUM	AI PRE-WORK (DELIVERED BEFORE, NOT DURING)	DECISION RIGHTS
Daily 07:00	Morning brief (async, 3 min)	Overnight anomalies ranked; deals that changed materially; new high-risk accounts; one recommendation with EV and cost	CRO: what gets personal attention today
Weekly Mon	Revenue stand-up (45 min)	Pipeline creation vs. requirement by segment; model-vs-rep forecast deltas; the intervention list with expected values; last week's committed actions and whether they worked	CRO + VPs: resource moves inside the quarter, escalations
Weekly Wed	Deal reviews (by segment)	Only deals flagged by the Pipeline Integrity and Forecast agents — typically 10–15% of the pipeline instead of a full walk-through	Managers: deal strategy, exec sponsorship requests
Weekly Thu	Customer risk review (30 min)	Risk list ranked by value at risk × probability × days to renewal; expansion signal list; calibration of last quarter's predictions	VP CS: intervention plan, exec sponsor assignment
Monthly	GTM efficiency review (90 min)	Segment/channel economics refresh; marginal CAC by channel; win/loss taxonomy shift; partner tiering update; two reallocation scenarios modelled	CRO + CFO: budget shifts, channel investment
Quarterly	GTM strategy & capacity	Cohort NRR by acquisition channel; cost of revenue motion by segment; capacity model vs. plan; scenario simulation for next-quarter allocation; model calibration report	CRO + CEO: headcount, segment strategy, pricing experiments

THE MEETING-DESIGN CHANGE THAT MATTERS MOST

Today, most revenue meetings are spent *discovering* what happened — someone presents, others react, and the decision gets deferred because the data was new to half the room. In this model the discovery happened before the meeting, in writing, with evidence attached. The meeting exists only to make the call. Expect your pipeline review to shrink from ninety minutes covering everything to forty minutes covering the 12% of deals that actually need a decision. That reclaimed time is not a side benefit — it is where the CRO's judgement gets applied to the deals that are still winnable.

SECTION 10

Governance & Failure Modes

The ways this goes wrong are predictable. Design against them before launch, not after the first incident.

FAILURE MODE	WHAT IT LOOKS LIKE	DESIGN AGAINST IT
Hallucinated causality	The system asserts a cause with fluency and confidence; the organisation acts on it; the real driver was mix, seasonality, or a comp-plan change nobody mentioned	Require an evidence chain and a stated confidence level on every "why". Frame outputs as hypotheses with a test attached. Sample and audit monthly.
Goodhart's law	A health score or engagement score starts driving comp; within a quarter the score is being farmed and no longer predicts anything	Never tie AI-derived scores directly to compensation. Recalibrate quarterly against actual outcomes and publish the calibration.
Alert fatigue	The system flags twelve urgent items every morning; by week six nobody opens the brief	Hard caps on item counts. Track open-and-act rate on the brief as a product metric. If it falls below 70%, the problem is the system, not the reader.
Silent model drift	Forecast accuracy quietly degrades after a pricing or segmentation change; nobody notices for two quarters	Publish trailing model accuracy alongside every forecast. Mandatory retrain and interval widening after any material GTM change.
Surveillance backlash	Reps discover their calls are analysed for coaching and start avoiding recording; data quality collapses at exactly the point of highest value	Transparency from day one about what is captured and how it is used. Conversation data goes to managers as coaching input, never as a direct performance-management input, and say so publicly.
Privacy and regulatory exposure	Call recording, profiling and cross-border transfer obligations under GDPR and equivalent regimes — particularly acute for EMEA-facing teams	Legal review before conversation capture goes live. Consent flows per jurisdiction, defined retention periods, and documented data-processing agreements with every vendor in the chain.
Over-automation	An agent sends customer-facing communication or changes a record without review, and the first anyone hears of it is from a customer	The red lines below. Every agent action logged, attributable, and reversible.

RED LINES — DECISIONS AI RECOMMENDS BUT NEVER MAKES

Compensation and quota. Hiring and termination. Pricing floors and non-standard commercial terms. Any customer-facing communication sent without human review. Territory changes affecting individual earnings. Write these down, publish them to the revenue org before the first agent goes live, and hold them. The credibility of the entire programme depends on people believing the system augments judgement rather than replacing accountability — and that belief is far easier to keep than to rebuild.

Trust is the binding constraint

The technical work is the easy half. The hard half is that a revenue organisation will only act on a system it believes, and belief is fragile in a specific way: it survives being wrong, and it does not survive being *unexplained*. A forecast that misses by 6% with a clear account of which assumption broke builds more trust than one that lands exactly and cannot say why. Optimise for traceability over accuracy in the first two quarters. Accuracy compounds later; trust, once lost, has to be rebuilt from a worse starting position than you began with.

SECTION 11

The First 90 Days

Sequenced so that each phase produces something usable, and each has a metric that proves it worked.

PHASE	DELIVERABLES	DETAIL	PROOF METRIC
Days 0–30 TRUTH	Metric dictionary · decomposition tree · data audit · baseline decision latency	Publish one definition per metric, signed off by Sales, Marketing, CS and Finance. Build the Altitude 0→1→2 tree with a named owner at every node. Audit the six data objects against the field list in Section 4 and score completeness. Time ten real decisions end to end to establish your latency baseline.	Data completeness score published; baseline latency measured
Days 31–60 DETECT	Hygiene remediation · L1 anomaly detection · first three agents in shadow mode	Fix stage definitions to buyer-evidence standard. Turn on history capture for stage, amount and close date. Launch anomaly detection on Altitude 1 metrics. Run Pipeline Integrity, Forecast and Churn agents in shadow — producing output that nobody acts on yet, so you can measure whether they would have been right.	Completeness ≥80%; shadow forecast within ±8% of actual
Days 61–90 DECIDE	Morning brief live · cadence redesigned · L2 diagnosis · calibration published	Morning brief goes daily. Redesign the weekly cadence per Section 9 so meetings decide rather than discover. Add causal decomposition to the top three recurring questions. Publish the first calibration report — of accounts flagged high-risk, how many churned; of deals the model doubted, how many closed.	Decision latency down ≥40%; brief open-and-act rate ≥70%
Quarter 2 EXTEND	Conversation intelligence · win/loss taxonomy · deal coach · GTM efficiency model	Conversation capture live with consent flows and a published policy. Evidence-based loss taxonomy replacing rep self-report. Deal Coach to managers. Complete the Finance cost join so CAC payback and cost of revenue motion become live rather than quarterly artefacts.	Loss reasons evidence-classified for ≥80% of closed-lost
Quarter 3+ COMPOUND	Allocation, partner and pricing agents · scenario simulation · L3 ranked options	Marginal-return allocation modelling. Partner attribution de-duplication and behavioural tiering. Discount leakage and packaging experiments. Recommendations begin arriving as ranked options with expected values rather than as observations.	Forecast accuracy ≥92%; CAC payback improvement by channel

THE ORDER IS THE STRATEGY

Truth before detection, detection before diagnosis, diagnosis before recommendation, recommendation before automation. Every failed revenue-AI programme I have seen inverted this — buying automation first and discovering in month five that the underlying definitions were contested, the history was overwritten, and nobody agreed on what a qualified opportunity was. The unglamorous first thirty days are what make the rest possible.

What winning looks like at the end of year one

FORECAST ACCURACY
±5%
From typical ±15–20%

DECISION LATENCY
−60%
Days, not weeks

CHURN NOTICE
90d
From ~30d

REVIEW TIME
−50%
Deciding, not discovering

None of these are AI achievements in isolation. They are the compound result of defining the metrics properly, joining the data honestly, and then applying AI to the loop rather than to the chart. The technology is the least differentiated part of the whole programme — every competitor can buy the same tools. What they cannot easily copy is a revenue organisation where the causal tree is documented, the data is trustworthy, and the operating rhythm is built around deciding quickly on good evidence.

APPENDIX A

Metric Definitions & Formulas

Reference table. Heuristics are planning starting points for B2B SaaS — validate every one against your own base rates before using it as a target.

METRIC	FORMULA	HEURISTIC	WATCH FOR
Net Revenue Retention	(Start ARR + Expansion − Downsell − Churn) ÷ Start ARR	105–125%	Blending segments; cohort drift
Gross Revenue Retention	(Start ARR − Downsell − Churn) ÷ Start ARR	85–95%	The honest twin of NRR — expansion cannot hide here
Pipeline coverage	Open qualified pipeline ÷ Remaining target	3–4×	Coverage on unqualified pipeline is meaningless
Required creation rate	(Target ÷ Win rate) ÷ Weeks of lead time	—	Must be computed per segment; blended hides the gap
Win rate	Won ÷ (Won + Lost + No-decision)	20–30%	Excluding no-decision inflates this by 5–10 points
Sales cycle	Median (not mean) days from qualified to closed-won	—	Mean is distorted by a few very long deals
Slippage rate	Value slipped out of period ÷ Value forecast at period start	<15%	Requires close-date history to compute at all
CAC payback	Fully-loaded S&M cost ÷ (New ARR × Gross margin) × 12	12–18 mo	Blending channels; excluding partner margin
LTV:CAC	(ARR × Gross margin ÷ Churn rate) ÷ CAC	3:1+	Unstable at low churn — payback is the sturdier metric
Magic number	(ΔQuarterly revenue × 4) ÷ Prior-quarter S&M spend	>0.75	Lags badly when sales cycles exceed one quarter
Quota attainment spread	P75 attainment ÷ P25 attainment	<2.5×	A wide spread means a hero motion, not a scalable one
Ramp time	Months from start to 80% of full productivity	3–9 mo	Determines how much of your capacity is permanently ramping
Forecast accuracy	1 − |Actual − Commit| ÷ Commit	≥92%	Track per manager; it tells you whose judgement to weight
Partner-sourced share	Partner-originated ARR ÷ Total new ARR	—	Never sum with influenced; audit the attribution hierarchy
Cost of revenue motion	(Acquisition + onboarding + support + success cost) ÷ Year-1 revenue	—	The metric that reveals a segment losing money after the sale
Data completeness	Opportunities with all decision-critical fields current ÷ All open	≥80%	Your AI layer's hard ceiling
Decision latency	Median days from signal availability to documented decision	—	Sample ten decisions per quarter; it is the meta-metric

APPENDIX B

AI-Readiness Audit

Twenty-five yes/no items. Score one point each. Below 15, fix data before buying anything.

DOMAIN	ITEM
Definitions	1 · One published, versioned definition exists per Altitude 0 and 1 metric
	2 · Sales, Marketing, CS and Finance have signed off on those definitions
	3 · A decomposition tree exists from each board metric down to named drivers
	4 · Every node in the tree has a single named owner
	5 · Opportunity stages are defined by buyer evidence, not seller activity
History	6 · Stage transitions are timestamped and append-only
	7 · Close-date changes are retained with full history
	8 · Amount changes are retained with full history
	9 · Account segment assignment retains history (no silent reclassification)
	10 · At least eight quarters of comparable historical data exist
Joins	11 · Product usage data is joined to CRM accounts on a reliable key
	12 · Support tickets are joined to accounts and to renewal dates
	13 · Marketing spend is attributable to segment and channel
	14 · Fully-loaded cost data from Finance is joined to the revenue model
	15 · Partner-sourced and partner-influenced are separately and consistently recorded
Quality	16 · ≥80% of open opportunities have a next step with a future date
	17 · ≥80% of closed-lost have a structured loss reason from a fixed taxonomy
	18 · Buying-group contacts are mapped to roles on enterprise opportunities
	19 · Competitor field is populated on late-stage opportunities
	20 · A data completeness score is computed and published weekly by team
Governance	21 · Conversation capture has legal sign-off and per-jurisdiction consent flows
	22 · Data retention periods are defined and enforced
	23 · Red lines are documented and published to the revenue organisation
	24 · Model calibration is reviewed and published at least quarterly
	25 · Every AI-derived claim can be traced to the underlying source records

SCORE	READ
21–25	Ready for L3–L4. Build agents, invest in automation, expect real returns within a quarter.
15–20	Ready for L1–L2. Anomaly detection and diagnosis will work. Fix the gaps before attempting recommendation or automation.
8–14	Data work first. AI output will be plausible and wrong, which is the most dangerous state to be in. Spend 60 days on the model.
<8	Do not buy anything yet. Start with definitions, ownership and history capture — none of which requires a vendor.

APPENDIX C

CRO Prompt Library

Twelve prompts to run against your revenue data. Each is written to force evidence and confidence rather than narrative.

FORECAST & PIPELINE

1. "Give me the quarter forecast with a confidence interval. Decompose the gap to target by segment, and tell me how much of it is concentrated versus diffuse. Show the deals driving the concentration."
2. "Compare each sales manager's forecast to the model forecast. Rank the deals with the largest disagreement and explain what the model is seeing that the rep is not."
3. "Compute required weekly qualified pipeline creation by segment given current win rates and cycle lengths. Show actual versus required for the last twelve weeks and flag any segment below plan for three consecutive weeks."
4. "List open deals showing slippage risk. For each, give the specific signals — days since buyer-initiated contact, buying-group breadth, close-date change history — and rank by value at risk."

DIAGNOSIS

1. "Win rate fell in enterprise last quarter. Decompose the change into mix effects and performance effects. If it is performance, identify which stage transitions and which reps account for the majority of the decline, and show me the underlying opportunities."
2. "Analyse all closed-lost opportunities from the last two quarters using call transcripts and emails rather than the rep-entered loss reason. Classify the real reasons and show me where the rep-entered reason disagrees with the evidence."
3. "Which objections appear in lost deals at a materially higher rate than in won deals? Show frequency, trend over time, and whether the pattern differs by segment."
4. "Our NRR dropped three points. Walk down the tree: gross retention versus expansion, then segment, then cohort, then acquisition channel. Stop at the first level where the effect concentrates and show me those accounts."

ALLOCATION & GROWTH

1. "Rank every segment-geography-channel combination by marginal return on the next dollar of investment. Include win rate, CAC payback, NRR and current capacity utilisation. Tell me explicitly where the estimate is unreliable because of small sample size."
2. "Model what happens to coverage and forecast if I move two enterprise AEs from North America to EMEA for one quarter. State your assumptions and what would have to be true for this to fail."
3. "De-duplicate partner attribution and give me true partner-sourced contribution. Then tier partners by actual productivity per quarter and identify which low-activity partners have the product-fit characteristics of our top performers."

CUSTOMERS

1. "Give me accounts with renewals in the next 120 days ranked by value at risk × churn probability. For each, list the specific signals driving the score, what a comparable account did last year, and the intervention that worked."

HOW TO PHRASE EVERYTHING YOU ASK

Three habits change the quality of what comes back. **Always ask for the evidence chain** — "show me the underlying records" turns a narrative into something checkable. **Always ask what would have to be true** for the recommendation to be wrong; it surfaces the assumption you would otherwise inherit silently. **Always ask where the estimate is unreliable**, because a model that never volunteers its own weak spots is not being cautious on your behalf — it is being confident on your behalf, which is a different thing entirely.

CLOSING NOTE

There is a product thesis hiding inside this document, and it is worth naming. Most revenue tooling makes executives **ask more questions** — more dashboards, more cuts, more filters. The system described here inverts that: it continuously ingests operational signal, understands the revenue engine's causal structure, and tells the CRO what requires attention today. The interesting wedge is not another BI surface. It is the layer that decides *what is worth an executive's attention* and defends that judgement with evidence — because attention, not information, is the constrained resource in every revenue organisation.