

From 61 Dashboards to 20 Numbers That Each Decide Something

A North Star metric, a driver tree, and a governed scorecard for a \$31M ARR B2B SaaS company — built so that every metric on it names the decision it exists to serve, and everything that does not is demoted rather than displayed.

CLIENT	“Arcadia Workflow” — the same Series C B2B SaaS company assessed in Work Sample 01. \$31.4M ARR, 1,180 customers, 6-person analytics team.
MANDATE	Design the metric system the audit recommended: one North Star, a driver hierarchy leadership can operate weekly, and definitions robust enough that the board number and the team number are the same number.
FORMAT	Project-based, 3 weeks. Two working sessions with the exec team, one with Analytics and Product, then this framework plus a definition repository handed to the Head of Analytics.
ROLE	I designed the hierarchy and wrote every definition; Arcadia’s analytics team built the models. I did not implement, and deliberately did not need to.

Four rules the framework is built on

1	Every metric names a decision.	If no one can say what they would do differently at a different value, the metric is interesting, not operational. It goes to Tier 3 or nowhere.
2	The North Star must be moveable and leading.	ARR is the outcome, not the North Star. A North Star that cannot be moved this quarter by a team that reports to someone in the room is decoration.
3	Definitions travel with the metric.	Owner, source, exclusions and known limitations are part of the metric. A number without them is a rumour with a decimal point.
4	Fewer tiers than teams.	Six board metrics, fourteen operating metrics, everything else diagnostic. Escalation from Tier 3 requires a stated threshold and a named decision.

The North Star: Weekly Value-Delivering Accounts

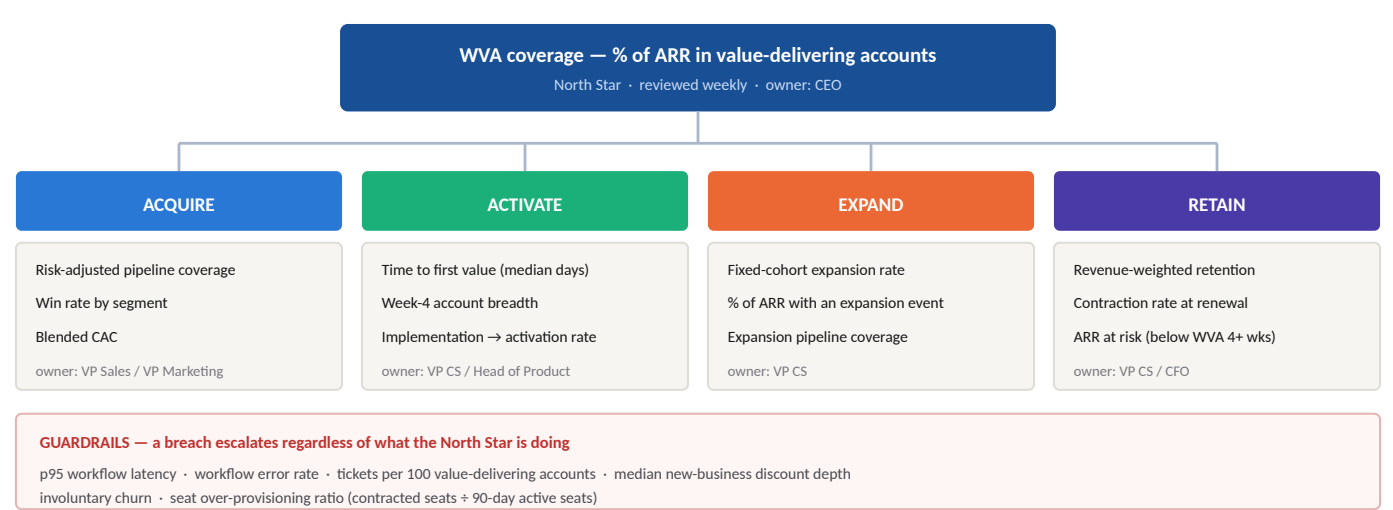
DEFINITION
An account is <i>value-delivering</i> in a given week if, within the rolling seven days, at least two distinct users each completed at least one qualifying core workflow, and the account completed at least three qualifying workflows in total. A workflow qualifies when it reaches a terminal success state <i>and</i> produces a downstream action — an export, an approval, or an integration write. Excluded: admin-only sessions, template edits, and re-runs of a failed workflow within 24 hours. The board-level form of the metric is WVA coverage: the share of ARR held in accounts that were value-delivering in at least six of the last eight weeks.

Why this metric survived and the alternatives did not

TEST A NORTH STAR MUST PASS	HOW WVA PERFORMS
Predicts the outcome that matters	Accounts above the threshold for eight consecutive weeks renewed at 96% and contracted -2.1% at renewal. Accounts below it renewed at 71% and contracted -11.8% . That spread is the entire retention problem, visible in a weekly number.
Leads the decision	The signal is stable and readable four to six months before a renewal conversation — long enough that intervention is still cheap.
Can be moved by named teams	Onboarding moves first-week breadth, Product moves workflow completion, CS moves reactivation. Each team can point to its own input.
Resists accidental inflation	Requires two humans and a downstream action. It does not move when an administrator logs in to pull an export — which is precisely the failure that made 35% of Arcadia's "active" accounts meaningless.

Where this metric can be gamed, and what stops it. Two attacks are available. The breadth requirement can be satisfied by provisioning a second nominal user; the volume requirement by re-running trivial workflows. Mitigations are built into the definition (no duplicate credit within 24 hours; a downstream action is mandatory) and into the guardrails (a floor on the ratio of downstream actions to workflow runs). I would also audit the top twenty weekly movers by hand each quarter for the first year. A North Star with no stated attack surface has usually not been stress-tested.

The driver tree



Read it downward, operate it upward. Leadership reviews WVA coverage weekly; each function reviews the three inputs it owns. The guardrail band exists so that the North Star cannot be improved by degrading something the North Star does not measure — the standard failure mode of single-metric management.

The scorecard, by tier

TIER	COUNT	CADENCE	RULE OF ADMISSION
Tier 1 — Board	6	Quarterly	ARR & growth · NRR (fixed cohort) · GRR · WVA coverage · CAC payback · Rule of 40. Nothing enters this tier without a peer benchmark and a restated history.
Tier 2 — Operating	14	Weekly / monthly	The twelve driver-tree inputs plus gross margin and S&M magic number. Each has a single named owner — a person, not a function — and a threshold that triggers a specific action.

TIER	COUNT	CADENCE	RULE OF ADMISSION
Tier 3 — Diagnostic	unbounded	On demand	Funnel step conversion, ticket categories, feature adoption, cohort curves, campaign CPL. Escalates only when it has crossed a stated threshold <i>and</i> a named decision is attached.

Three definitions, written out

Every Tier 1 and Tier 2 metric ships as a card in this format. Three are reproduced here; the delivered repository contains twenty. The row that does the most work is the last one.

Net Revenue Retention (NRR)		TIER 1
DEFINITION	Ending ARR of a fixed cohort of accounts, divided by that cohort's ARR twelve months earlier. The cohort is frozen at the start of the period; churned accounts remain in the denominator.	
FORMULA	$(\text{start_arr} + \text{expansion} - \text{contraction} - \text{churn}) \div \text{start_arr}$	
SOURCE	Certified model <code>fct_arr_movement</code> . Contract-based ARR, not invoiced amount.	
OWNER	CFO (Head of Analytics is the technical steward)	
CADENCE	Monthly, trailing twelve months. Restated history back eight quarters.	
LIMITATIONS	Sensitive to cohort start date in a quarter with a large single renewal. Report alongside GRR — NRR alone cannot distinguish “strong expansion” from “expansion masking churn”.	
DECISION	Sizes the expansion contribution in the annual plan and sets the CS coverage investment. Below 105% on a trailing basis, the plan's expansion assumption is rebuilt rather than adjusted.	

WVA Coverage		TIER 1 — NORTH STAR
DEFINITION	Share of ARR held in accounts that were value-delivering (≥ 2 distinct users, ≥ 3 qualifying workflows, rolling 7 days) in at least six of the last eight weeks.	
FORMULA	$\sum \text{arr}[\text{account}] \text{ where } \text{wva_weeks_8} \geq 6 \div \text{total_arr}$	
SOURCE	Certified model <code>fct_account_week_value</code> , joined to contract ARR.	
OWNER	CEO at the coverage level; each driver has a functional owner (see tree)	
CADENCE	Weekly, in the Monday operating review. Reported by ARR, not by account count.	
LIMITATIONS	Seasonally soft in late December and late August; compare year-over-year, not week-over-week, across those windows. Two known gaming vectors, mitigated in the definition and guardrails.	
DECISION	Below 78%, onboarding and reactivation capacity is reallocated before any new-logo spend increase is approved. Accounts dropping out for four consecutive weeks enter the renewal-risk queue automatically.	

Time to First Value (TTFV)		TIER 2 — ACTIVATE
DEFINITION	Median calendar days from contract start to the first qualifying core workflow completed by any user in the account.	
FORMULA	median(first_qualifying_workflow_ts - contract_start_ts)	
SOURCE	Certified model fct_account_activation	
OWNER	VP Customer Success	
CADENCE	Weekly, by cohort of contract-start month. Never as an all-time average.	
LIMITATIONS	Right-censored: accounts that have not yet activated are excluded from the median and must be reported separately as a “not yet activated” count, or the metric improves every time a slow account is dropped.	
DECISION	The single leading indicator for the FY27 onboarding programme. Above 21 days for two consecutive cohorts, implementation sequencing is changed before more headcount is added.	

Guardrails

GUARDRAIL	ESCALATION THRESHOLD	WHAT IT PROTECTS AGAINST
p95 core workflow latency	> 4.0s, weekly	Growing WVA by shipping volume at the cost of the experience
Workflow error rate	> 1.5%	Counting failed-then-retried runs as value
Support tickets per 100 value-delivering accounts	> 22 / month	Activation achieved by leaning on humans rather than product
Downstream actions ÷ workflow runs	< 0.7	Inflating workflow counts with trivial or duplicated runs
Median new-business discount depth	> 18%	Buying ARR growth that will not renew at list
Involuntary churn	> 0.4% of ARR / quarter	Billing and payment failures being read as demand signal
Seat over-provisioning ratio	> 1.6×	Contracted seats drifting far ahead of used seats — the leading edge of contraction at renewal

Candidates I rejected, and why

The rejections are the part of this exercise that saves the most time later. Each of these was proposed in the working sessions by someone with a reasonable case.

CANDIDATE	REASON IT IS NOT THE NORTH STAR	WHERE IT LIVES
ARR growth	Lags by six to twelve months and cannot be moved this quarter by any single team. It is the outcome the North Star is chosen to predict — using it as the North Star closes the loop with no information in it.	Tier 1 outcome
Monthly active users / logins	Measures access, not value. Thirty-five percent of accounts Arcadia currently counts as active have nobody doing any work in them.	Retired
Total workflows run	Rewards volume and tolerates concentration: an account with one power user running 400 trivial workflows outranks one with a team of six running thirty that matter. No breadth requirement means no early warning of champion risk.	Tier 3

CANDIDATE	REASON IT IS NOT THE NORTH STAR	WHERE IT LIVES
NPS	11% response rate; the sample cannot be read at segment level; it moves for reasons no team can attribute within the quarter. Useful as listening, unusable as a target.	Tier 3
Customer count	Ignores that the top 25 accounts carry 41% of ARR. Every movement in this metric is dominated by the accounts that matter least.	Tier 3
Composite health score	A score of 62 does not tell anyone what to do. Composites hide which input moved and invite quiet reweighting when the number is unflattering. Publish the inputs; let the judgement be visible.	Rejected outright

Governance: what keeps this from decaying in six months

1	One owner, and it is a person	Every Tier 1 and Tier 2 metric names an individual accountable for its definition and its movement. “The data team” is not an owner.
2	One source of truth	If a number in a deck cannot be traced to a certified model, it does not go in the deck. This rule is what closes the \$1.4M ARR discrepancy permanently.
3	Definition changes are versioned	A change requires the metric owner plus the Head of Analytics, is recorded with a date, and restates prior periods. A silently redefined metric is worse than a wrong one.
4	Thresholds, not just targets	Each metric carries the value at which a specific action is taken. A target says how we are doing; a threshold says what happens next.
5	Quarterly use review	Each Tier 1 and Tier 2 metric must show one decision it informed during the quarter, or it is demoted to Tier 3. Scorecards grow by accretion unless something removes from them.
6	Deprecation is scheduled	Dashboards below five unique viewers in 90 days are archived on a fixed date each quarter, not debated case by case.

WHAT CHANGES ON MONDAY

The weekly operating review opens with one slide: WVA coverage, its four-week trend, the ARR that moved in or out of the value-delivering set, and any guardrail in breach. Twenty-five minutes. Function-level driver reviews happen separately and do not appear in this meeting unless a threshold has been crossed. The measure of whether the framework worked is not whether the dashboards look better in Q1 — it is whether the Q2 planning conversation reaches a decision faster than the Q4 one did.

Illustrative work sample. “Arcadia Workflow” is a fictional composite. All figures, thresholds and findings are constructed to demonstrate analytical approach and deliverable format; they are not the data of any client, past or present. · Companion piece: *Work Sample 01 — Executive Analytics & KPI Audit*, the assessment this framework answers.