

RED BEAR ARTIFICIAL INTELLIGENCE

ABU DHABI, UNITED ARAB EMIRATES

Trust and Verify

Leadership without losing control: decision governance that gives leaders visibility instead of a signature

Almost every senior leader says they want to delegate more. Very few do, because delegation as most organisations practise it means surrendering sight of a decision as the price of not having to make it. This paper separates control from visibility, describes the Trust Dial that makes latitude explicit, personal and reversible, and answers the question every thoughtful executive asks first: is this surveillance?

Empower. Trust. Verify.

White paper EDI-WP-03 · Red Bear white paper series · Version 2.0 DRAFT
26 September 2026
Commercial in confidence

Summary

Control is deciding. Visibility is seeing what was decided, by whom, and on what basis. Leaders ask for the first and almost always want the second. An executive who insists on signing every discount is rarely doing it because their judgement on discounts is better than their sales director's. They are doing it because the signature is the only mechanism they have for knowing it happened.

The stated reason for not delegating is capability: the team is not ready, the stakes are too high. Occasionally that is true. Far more often the real reason is simpler and more forgivable: the leader cannot see what happens after they let go. Faced with a trade between visibility and throughput, an accountable executive quite rationally declines to give up the signature.

This paper argues that the trade is false. Separate control from visibility and the problem changes shape: if visibility can be had without control, then discretion can be extended without losing sight of anything, and delegation stops being an act of faith and becomes an instrument that can be set, observed and adjusted.

That instrument is the **Trust Dial**: one per decider, person or AI agent, on a scale of 0 to 10 in five bands. Guardrails on the decision type say what may ever be decided in that type. The dial says how much of it this person may decide today. Two people in the same role can hold different dials, the difference is explicit and attributed, and the dial moves one point at a time, by a person, with a reason, as evidence accumulates or as circumstances require.

The paper then answers the objection this invites, because it is the first question any thoughtful executive asks and the question their people will ask louder. A record of who decided what can support fair assessment or it can support blame. Which one it becomes is a matter of design and of leadership, and the design choices that make the difference are specific enough to be checked.

Red Bear: Enterprise Decision Intelligence (formerly GuardrAIl) implements the model: guardrails authored per decision type, one dial per decider, accountable override for the cases the rule gets wrong, a sandbox that replays history before any change takes effect, and a private view that shows each person their own record. Adjustments to a dial are recommended from evidence and never applied by the product. The platform proposes; people decide.

WHO SHOULD READ THIS

- Chief executives and chief operating officers
- Chief financial and chief risk officers
- Business unit leaders who approve too much
- Boards, and the people whose decisions will be recorded

WHAT YOU WILL BE ABLE TO DO AFTERWARDS

- Explain why a leader who stops approving sees more, not less
- Read a Trust Dial: five bands, each a review rate
- Describe how a dial is widened or narrowed, and why the conversation is not a formality
- Apply the one test that separates a development instrument from a surveillance tool

1 What it is

Delegation is the most discussed and least practised discipline in senior management. What actually happens is a partial version that satisfies nobody. Authority is delegated in principle and retrieved in practice: the manager may decide, but is expected to check first. The limits are described as "use your judgement", which is warm, unhelpful and, for anyone who intends to keep their job, an instruction to escalate.

Empower. Trust. Verify. Empower is the range you choose to give someone. Trust is extending it in the knowledge that what happens inside it is recorded. Verify is reviewing the outcomes and adjusting the range accordingly. The sequence matters: verification without prior empowerment is inspection, and empowerment without verification is abdication. Neither is leadership.

The Red Bear framework gives the three words their instruments. **Guardrails** are set per decision type, in nine fixed kinds: authority, policy, financial limits, risk appetite, compliance, segregation of duties, exceptions, human intervention and auditability. They say what may ever be decided in that type, whoever is deciding. The **Trust Dial** is set per decider, 0 to 10, and says how much of the permitted range this person or agent currently exercises. **Role scope** says which decision types the decider may touch at all. The effective range on a given decision is the product of the three (Oliver, *Empower, Trust, Verify*, Book 1, chapters 5 and 8).

Dial	Band	What it means	Verification
0	No Discretion	Everything is referred	Every decision made by someone else
1–3	Recommend	They propose and a human decides	Every recommendation decided by a person
4–6	Supervised	They act within narrow limits	Every decision reviewed afterwards
7–9	Delegated	They act within wide limits	Decisions sampled
10	Autonomous	They act within policy	Reviewed by exception; still recorded, still inside guardrails, never ungoverned

The band is not decoration: it *is* the review rate. That is the property that makes the dial governable rather than alarming. It cannot grant authority the organisation has not sanctioned for the type; it can only determine how much of a sanctioned range a given individual exercises, and how closely what they do is looked at. Nobody, at any setting, exceeds the guardrail. An AI agent's ceiling is 9: no agent holds Autonomous on a consequential decision type.

Key points

- Control and visibility are different things; the approval step bundles them, which is why surrendering it feels like surrendering everything
- Guardrails belong to the decision type; the dial belongs to the decider; the band is the review rate
- The dial moves one point at a time, by a person, with a reason; the product proposes and never moves it

2 Why it matters

External evidence. McKinsey's research on delegated decisions found that too few decisions are explicitly delegated, meaning too few employees know what they can and cannot decide, and that even where they do

know, their managers often do not know how to support them. Just over a quarter of respondents said their organisations made good delegated decisions, the weakest result across the decision types studied. The effect of getting it right was substantial: respondents who said employees were empowered and received sufficient coaching were several times more likely to report that delegated decisions were good ones (De Smet, Lackey and Weiss, *For smarter decisions, empower your employees*, McKinsey, 2020).

Why the vague limit is the root of it. Ask a manager what they are permitted to decide alone and you will get an approximate answer, hedged, and different from the answer their director would give. That gap is the predictable result of authority that was never written down. An undocumented limit is resolved in only one direction: faced with uncertainty about whether something is theirs to settle, a sensible person escalates. It is free for them and expensive only for the organisation. Multiply that across a few hundred managers and the executive becomes a bottleneck by a process nobody chose.

What the leader is protecting. An executive who delegates a class of decision is still accountable for it. If something goes wrong they will be asked what they knew and when, and under the conventional arrangement the honest answer is *nothing, until afterwards*. So the retained signature is not really about the decision. It is about the record. The leader is using the approval step as an information channel, because it is the only one they have, and it costs the organisation a delay on every decision to give one person sight of it.

The paradox worth naming to a board. Organisations that centralise decisions to reduce risk frequently increase it. Decisions are made later, by people further from the facts, on summarised evidence, with the reasoning unrecorded because approval processes capture assent rather than argument. The organisation feels controlled and is in fact less well informed than one that delegated with a record.

Cost of under-delegation	How it presents
Executive time consumed below its level	Senior capacity spent on matters a competent manager two levels down could settle in a minute, while the decisions only the executive can make wait
Delay on every delegated-but-not-really decision	The cost is not the approval; it is the queue in front of it, and the queue is a function of one diary (EDI-WP-02)
Decisions made further from the information	Every escalation degrades the evidence available to whoever finally decides
Capable people leave	A manager who cannot decide anything is doing a smaller job than their title describes, and the ones with options notice first
Judgement never develops	People learn to decide by deciding; an organisation that reserves consequential decisions for its most senior people guarantees the next generation arrives untested
Accountability blurs in both directions	When a manager recommends and an executive approves, neither wholly owns the outcome

The signature was never the safeguard. If an executive approves forty decisions a week on a one-line summary, the approval is not scrutiny; it is a ritual that produces a defensible record. Replacing a ritual approval with a real record is not a reduction in governance. It is an increase, because the record contains what the approval never did: the reasoning, the alternatives, the expected outcome and the policy in force at the time. And under control a leader sees the decisions that came to them; under visibility they see every decision in their organisation, including the ones settled locally. **The leader who gives up approving sees more, not less.**

3 The challenges

Challenge	Why it is hard	What goes wrong without a method
Writing the limit down	A delegation of authority matrix is accurate, rarely consulted, and silent on everything that is not a monetary amount, which is most of what needs deciding	"Use your judgement" is read as "escalate"
Varying the limit by person	The same role does not warrant the same latitude in every pair of hands, but one number per grade forces the limit to be set for the least experienced	Everyone else is constrained by the newest joiner
Making the boundary crossable	Rules are wrong at the edges; a control that cannot be exceeded gets worked around, and a workaround leaves no trace	The rule looks honoured and is not
Separating involvement from authority	Many contribute; one decides. Consultation is input, not a vote	Committees become pools of vetoes
Verifying outcomes, not activity	The verification that matters is reviewing what was decided and how it turned out, not watching the work	Verification becomes inspection, and people manage their record rather than their judgement
Sustaining it through a difficult quarter	Each of the five practices replaces a feeling with a record; sustaining all five across several hundred managers by discipline alone does not happen	The first significant error resets the behaviour for a year

4 How Red Bear: Enterprise Decision Intelligence helps

Challenge	What Enterprise Decision Intelligence does	Requirements
Writing the limit down	The policy builder composes a guardrail from the nine fixed kinds, sets thresholds and enforcement strength, and assigns it to decision types; policies are versioned, owned and tested, and every evaluation records the version it used	FR-DEL-19, FR-DEL-14
Varying the limit by person	One Trust Dial per decider, 0 to 10 in five bands, never per decision type; role scope is a separate object; a team tier seeds the dial of a joiner and never holds or moves an existing member's	FR-DEL-02, FR-DEL-06, FR-DEL-08
Moving the dial on evidence	The earned trust ledger scores trust from each actor's record and recommends dial settings with its evidence; it has no path to apply them. A move is one point at a time, by a person, with a reason, and cannot be committed without the projection having been displayed	FR-DEL-03, FR-DEL-02, RBP FR-PR-04
Making the boundary crossable	Decision interception holds a decision when a blocking policy requires and releases it the moment it clears; a blocking breach is overridden only with an explicit override and a recorded, non-blank reason, as a separate event attributed to the person	FR-DEL-05

Challenge	What Enterprise Decision Intelligence does	Requirements
Widening on evidence, not nerve	The scenario sandbox replays a named historic period under a proposed policy or dial setting through the same routing function as live routing and reports the difference in level mix, decision count by outcome and human-touch share; it applies nothing	FR-DEL-16
Loosening needs two people	A narrowing takes effect on the author's own action; any widening requires activation by a second person with authority for that decision type	FR-DEL-19, RBP FR-DI-04
Involvement and authority	The decision rights matrix holds Recommend, Decide, Own, Execute and Review per decision type; a committee may recommend and review only; Own has exactly one filled cell, a named person; no actor may decide their own decision	FR-DEL-07, RBP FR-DI-07
Verifying outcomes, not activity	The band is the review rate: every decision at Supervised, a sample at Delegated, exceptions at Autonomous; outcomes are attached from a connected system or the owner's review; quality is scored on process, apart from outcome	FR-DEL-02, FR-OUT-01, FR-OUT-05
Visibility to the leader	Every decision made under the leader's authority is in the register, with the reasoning, the policy version and the dial applied, in an append-only vault that replays as it stood	FR-REC-07, FR-REC-08, FR-REC-09
Visibility to the person	Each person reads their own history; approver oversight follows the reporting line and no further; a lesson concerning someone's own decisions is offered as a private coaching conversation, by invitation	FR-PLT-08, RBP FR-DC-01, RBP FR-DC-04

What it does not do

The product does not move anyone's dial, does not approve decisions, does not assess individuals, and does not observe how people work. A deterministic, versioned policy engine evaluates policy; models normalise, check submitted analysis and explain; people approve and people set the dial.

5 How it is used

Step	Who	What happens
1 Author the guardrails	Policy Maker	Each decision type receives its nine guardrails in the language the work uses: type, value band, risk class, conditions; a narrowing is live on save
2 Set the dials	Decision Approver	Each leader places the people under their command on the dial, one per person; a joiner is seeded by the team tier; the projection is shown before any move commits
3 Decide inside discretion	Decision Maker	The Authorise gate passes automatically; the decision is recorded with the dial and the policy version, and reviewed at the band's rate
4 Refer at the edge	Decision Maker, Decision Approver	A decision outside the effective range goes to one available authorised person, with the full record and precedent attached; three verbs; both halves recorded
5 Override at a price	Decision Approver	Proceeding over a blocking breach requires a recorded reason as a separate attributed event; override telemetry reads the pattern as a question about the policy

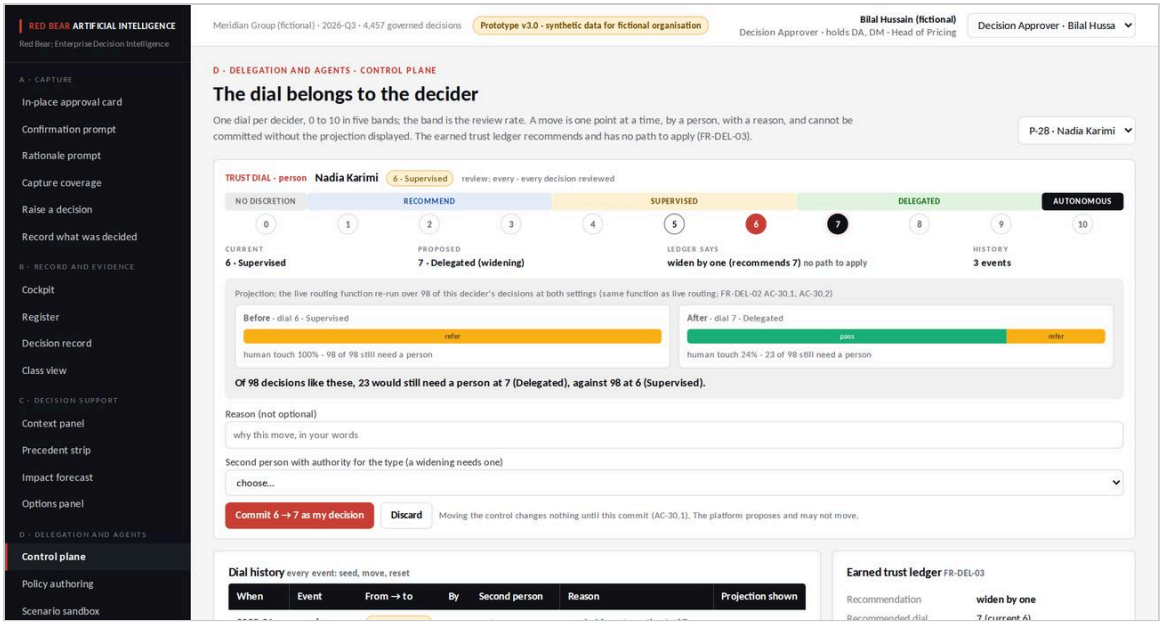
Step	Who	What happens
6 Verify	Decision Approver, director	Reviews at the band's rate; outcomes attached; the earned trust ledger puts its evidence in front of the director and the manager
7 Move the dial	Director and manager together	The recommendation is discussed; the dial moves one point, with a reason, dated and attributed; a widening is activated by a second person; a narrowing carries its reason and its review

How latitude is earned, and why the conversation is not a formality. Decisions made and outcomes recorded accumulate into evidence: volume, outcomes against expectations, override frequency, breaches, how often approvals were subsequently reversed. Where that evidence supports a change, the product raises a recommendation with the record attached, and it goes to the director and the manager together. The obvious engineering shortcut would be to apply the adjustment automatically when the evidence crosses a threshold. It would be faster, consistent, and a serious mistake. Evidence about decision quality is incomplete in ways a model cannot see: a manager inherited a difficult territory, absorbed a departing colleague's portfolio, made three sound calls in a market that moved. The record shows what happened; it does not know why. There is a second reason, and it is the more important one. An adjustment delivered by a system is a verdict. The same adjustment delivered by a leader who explains it is feedback. The mechanism is identical and the effect on the person is entirely different.

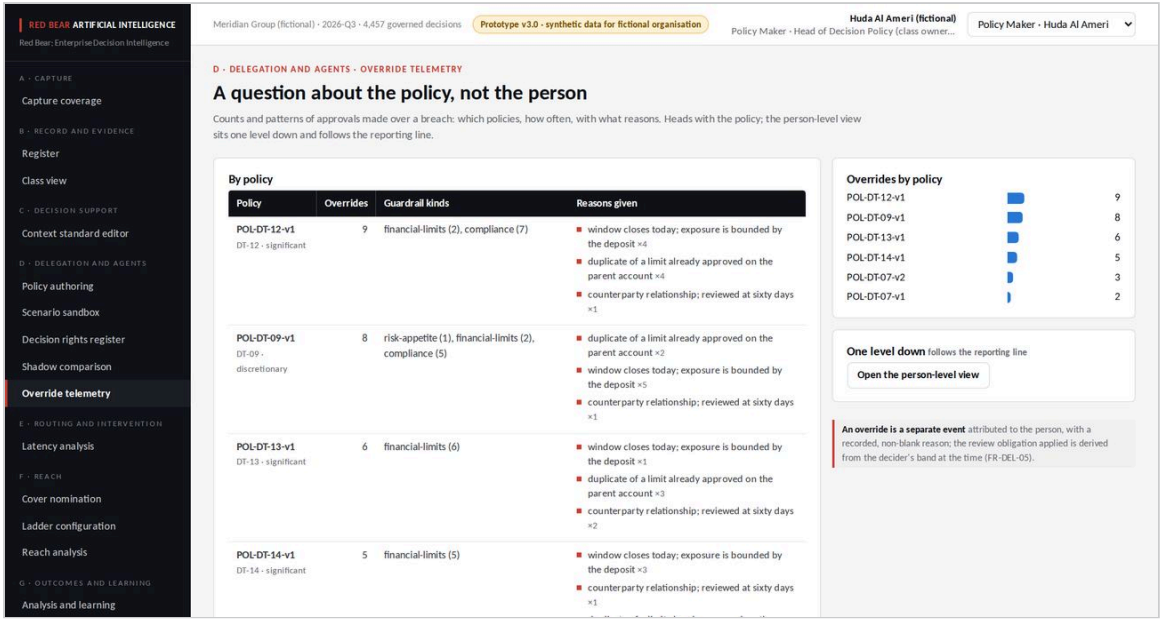
Withdrawal is the harder half. Narrowing someone's range is unpleasant and organisations avoid it, which is why it must be ordinary rather than exceptional. A reduction carries a reason, and the person can see it; a reduction someone learns about by hitting an unexpected referral is corrosive far beyond the individual. A tightening during a regulatory examination or a major incident says nothing about anyone's competence, and is written as a guardrail exception on the decision type with a sunset date, which the product refuses without one. EDI-WP-09 develops this at length.

Four illustrations, none a customer engagement. A sales director sets discount guardrails for her team. Nine of eleven account managers sit in the Delegated band and settle the standard range alone, sampled. Two sit at 9 after two years of discounts that held margin and renewed. One, four weeks into the role, sits in Supervised, seeded by the team tier, every decision reviewed. The effect is not that more discount is given; it is that nine people stop queuing for a signature on decisions the guardrail already permitted, and the director's attention moves to the few that exceed it. A contact centre manager moves eleven agents up a point after two quarters of refund decisions that held up on review; a twelfth agent, three weeks in, keeps the seeded setting, by a default everyone knows about rather than an awkward conversation. A group CFO wants to widen payment-run authority for two regional controllers, and first runs the proposed guardrail in the sandbox against eighteen months of history: it would have covered nearly all of the last hundred payment decisions and still caught the six that mattered. He widens on that evidence, a second person activates it, and he can show the audit committee precisely what changed. A business unit enters a regulatory examination; the director narrows financial limits across the unit for the duration as an exception with a sunset date, explains it in a team briefing as a response to the circumstance, and the record shows the date it expires.

6 The user experience



S-20 Control plane: S-20 Control plane. One dial per decider in five bands, the guardrails per decision type beside it, and the projection that must be displayed before any one-point move commits. Prototype render on synthetic data for fictional entities.



S-26 Override telemetry: S-26 Override telemetry. Approvals made over a breach, by policy and class, with reasons, presented as a question about the policy; the person-level view sits one level down and follows the reporting line. Prototype render on synthetic data for fictional entities.

7 Where the data comes from, and how it is checked

Inputs to trust and verification

Input	Source	Check
Decision rights and authority	HR and the directory, kept current; authority, accountability and ownership held as three distinct things	Accountability never transfers to an AI agent (FR-DEL-01, RBP FR-ID-02)
Guardrails	Authored per decision type in the policy builder	Conflicts between overlapping policies detected at authoring time; exceptions refused without a sunset date (FR-DEL-15)
Dial settings	Set by the person's own leader, one point at a time	Cannot commit without the projection; a widening needs a second person (FR-DEL-02, FR-DEL-19)
Decisions made under the grant	Every capture route; the agent gateway for agents	The acting user is always taken from the verified token; no interface, API or agent may assert who it is acting as (FR-PLT-02)
Overrides	An explicit override with a recorded reason	A separate event attributed to the person; the review obligation derived from the band at the time (FR-DEL-05)
Outcomes	A connected system reporting against the decision identifier, or the owner's review	An outcome the product guessed is not evidence (FR-OUT-01)
Trust evidence	The earned trust ledger, computed per actor from outcomes, overrides and breaches	A recommendation with its evidence; no path to apply it; for an agent, attached to the model version (FR-DEL-03)

How it is checked

- **Connectors are read-only and consented.** Capture routes are granted by the customer administrator under least privilege; the product records decisions, not activity: no hours, keystrokes, location, messages or attendance (RBP FR-CN-01, RBP FR-PA-02).
- **A person moves the dial.** The product proposes and shows its evidence; the move is made by a named person, one point, with a reason, dated and attributed (RBP FR-PR-04).
- **Widening needs two people.** A raised limit, a loosened guardrail or a lifted team tier requires activation by a second person with authority for the type; a narrowing does not (FR-DEL-19).
- **The record replays as it stood.** Policy and dial history are retained with the same rigour as the decision (FR-REC-09).
- **Quality is scored without outcomes.** Process quality is computed without access to any measure series, so people are not judged on luck; no interface blends the two (FR-OUT-05).
- **Visibility follows the reporting line.** An approver sees person-level detail for their own reports only; beyond that line, class-level and policy-level views (FR-PLT-08).

8 Benefits realised

Every measure is calculated from the organisation's own register against a baseline established before anything was switched on (FR-PLT-05).

Benefit	How it shows up	Measured by
Delegation actually happens	Decisions settle where the knowledge is	Share of decisions settled within the decider's own authority: the number to put in front of a board
Executives stop approving ritually	Falling volume at director level with no rise in breaches	Approval volume reaching each level of management
Rules improve instead of being flouted	A policy overridden repeatedly is read as miscalibrated	Override frequency per policy
The dial is an instrument, not a permissions table	Dials move, in both directions, on evidence	Distribution of dial settings and its movement
Verification closes the loop	Outcomes against decision quality, by band	Quality quadrant by band: slower to accumulate, and the one that justifies widening further
Judgement develops	People at Supervised move to Delegated on their record	Dial moves upward per class per year

9 Risks managed

Risk	Control in Red Bear: Enterprise Decision Intelligence
It becomes surveillance	Decisions recorded, not activity; each person reads their own record; lessons never attributed; coaching by invitation, nothing attributable kept (RBP FR-PA-02, RBP FR-DL-04, RBP FR-DC-04)
The dial is applied by an algorithm	The ledger recommends with evidence and has no path to apply; a person moves the dial one point at a time (FR-DEL-03, FR-DEL-02)
Widening costs money	Some decisions will be made that a senior person would have made differently; the sandbox shows what a widening would have permitted before it takes effect, and the review rate of the band catches what it did permit (FR-DEL-16)
The dial is set and forgotten	The ledger keeps putting evidence in front of the leader; distribution reporting shows a population stuck in one band (FR-DEL-03, FR-OUT-08)
Evidence about individuals is thin at first	No calibration figure is shown below twenty to thirty decisions; early adjustments rest on judgement, which is the correct order (FR-REC-12)
Employment and works-council exposure	Per-person latitude sits close to performance data; the individual view, the never-automatic rule and reporting-line visibility are the features that make that review straightforward (FR-PLT-08)
A league table	Slices by region and decision type, never by name; an indicator without its gaming vector may not be added to a dashboard (FR-OUT-08)

The test to apply before deployment: ask whether you would be comfortable showing any individual everything the system holds about them, and explaining how it will be used. If yes, it is a development

instrument. If no, something in the deployment has gone wrong, and the people affected will work that out considerably faster than the leadership does.

Sources

Every claim attributed to a source names it in the text as *author, work, section*; the full references are below. Statements without a citation are Red Bear's own position, developed for Red Bear: Enterprise Decision Intelligence. Quotations are limited to fifteen words.

Cited as	Source	Authority
De Smet, Lackey and Weiss	A. De Smet, G. Lackey and L. Weiss, *For smarter decisions, empower your employees*, McKinsey (September 2020): too few decisions explicitly delegated; the effect of coaching-based empowerment	Practitioner research
De Smet, Jost and Weiss	A. De Smet, G. Jost and L. Weiss, *Three keys to faster, better decisions*, McKinsey Quarterly (May 2019): clarity about the boundaries of delegation as a defining practice	Practitioner research
Aminov, De Smet, Jost and Mendelsohn	I. Aminov, A. De Smet, G. Jost and D. Mendelsohn, *Decision making in the age of urgency*, McKinsey Global Survey (April 2019): the relative weakness of delegated decision making	Practitioner research
Blenko, Mankins and Rogers	M. Blenko, M. Mankins and P. Rogers, *Decide & Deliver*, Harvard Business Review Press (2010): decision rights and the separation of contribution from authority	Practitioner research
Duke	A. Duke, *Thinking in Bets*, Portfolio (2018): decision quality apart from outcome quality, developed in EDI-WP-01	Practitioner reference
Book 1	D. Oliver, *Empower, Trust, Verify*, Book 1 (2026), chapters 5, 8 and 17: the Trust Dial, guardrails, decision rights	Red Bear
Red Bear	EDI-FSD-01 v3.0 draft; Red Bear Platform Functional Specification RBP-FSD-01	Red Bear

Glossary

Term	Meaning in this paper
Guardrail	One of nine fixed kinds of limit set per decision type: what may ever be decided in that type, whoever is deciding
Trust Dial	One per decider, 0 to 10 in five bands, No Discretion, Recommend, Supervised, Delegated, Autonomous; the band is the review rate
Role scope	Which decision types a decider may touch at all; a separate object from the dial
Effective range	The product of the dial on the decider and the guardrail on the type, bounded by role scope
Earned trust ledger	The evidence about an actor's record, from which the product recommends a dial setting it cannot apply
Accountable override	Proceeding over a blocking breach with a recorded reason, as a separate event attributed to the person

Term	Meaning in this paper
Scenario sandbox	Replaying a historic period under a proposed policy or dial through the live routing function, applying nothing
Decision rights	Recommend, Decide, Own, Execute, Review; Own is held by exactly one named person
Verify	Reviewing outcomes of delegated decisions at the band's rate and adjusting latitude accordingly; distinct from inspecting how work is done
Widening	A raised limit, a loosened guardrail, a lifted team tier or a dial move upward; requires a second person

The series

This paper is one of ten. Each stands alone and each follows the same nine sections, so a reader who knows one knows where to look in the next.

Paper	Title	Theme
EDI-WP-01	Cartesian Decision Making	Learning what happened next
EDI-WP-02	The Cost of Organisational Indecision	The decision that never gets made
EDI-WP-03	Trust and Verify	Visibility, not control (this paper)
EDI-WP-04	Delegation Without Losing Governance	The boundary and the room
EDI-WP-05	Building a Decision-Making Culture	Judge the decision, not the result
EDI-WP-06	Decision Impact Analysis	What a decision dataset can and cannot tell you
EDI-WP-07	The Decision-Aware Enterprise	Decisions in flight
EDI-WP-08	Decision Making Anywhere	The channel is part of the governance
EDI-WP-09	When Micromanagement Is the Right Management	Oversight is a dial
EDI-WP-21	The Customer Journey	Connect, Observe, Govern, Learn, Compound

Red Bear Artificial Intelligence · Abu Dhabi, United Arab Emirates · info@redbearai.net · www.redbearai.net