

RED BEAR ARTIFICIAL INTELLIGENCE

ABU DHABI, UNITED ARAB EMIRATES

Culture Coaching Without Surveillance

What a decade of failed culture technology taught us, and how Red Bear's coaching agent is built so it cannot repeat it

Every serious attempt to change workplace behaviour with software has left a case file: a retreat, a fine, a lawsuit, a wind-down or a trial that found nothing. Red Bear read the file before writing a line of code. This paper names the ways AI coaching and culture technology have failed, from Productivity Score to NEDA's Tessa and from Wells Fargo's values booklet to the nudge units, and shows, rule by rule, how Red Bear: Enterprise Culture Evolution is built so that each failure has no place to happen: an agent triggered by patterns and never by a person, a budget on its own attention, a named human one request away, a dial that has a lever, and evidence before scale.

Empower. Trust. Verify.

Summary

Culture technology has failed often enough, and publicly enough, that the failures are now the specification. Over a decade, employers and vendors have tried to move behaviour with dashboards, badges, nudges, chatbots and values campaigns. Microsoft removed user names from Productivity Score within days of launch; Barclays ended a desk-time trial after a staff revolt; CNIL fined Amazon France Logistique €32 million for per-worker tracking that aggregation would have served; Humu sent ninety kinds of nudge for six years and was absorbed as a feature; Koko's language-model messages were rated well until recipients learned a machine had written them; NEDA's chatbot gave dieting advice to people with eating disorders after a vendor changed its generation mode; the best-evidenced nudge units moved behaviour by 1.4 points where the literature promised 8.7; and the one controlled trial of AI coaching for senior leaders found no effect and tenfold attrition. Behind these cases sit fourteen mechanisms, and behind the mechanisms sit fourteen design rules.

Red Bear: Enterprise Culture Evolution is built from those rules. Its coaching agent is Socratic and narrow: it opens only when an aggregated pattern moves or a decision misses its recorded expectation; it coaches the named owner of a dial or a decision about a group, never a person about themselves; it discloses that it is an AI in every session; it cannot agree with a causal account it has no evidence for; it spends at most two contacts a month of any manager's attention; and it hands to a named human on request. Around the agent sit the rules that decide whether a programme survives: a Culture Dial that cannot be created without the incentive or decision right it depends on, a sunset for every programme element that fails to move its dial, a pilot with a comparison before any expansion, and a consultation gate no connector can pass until staff and their representatives have been told. This paper is the trace from each failure to the rule that prevents it, with the platform requirement that enforces the rule.

WHO SHOULD READ THIS

- Chief executives and chief people officers considering a culture programme, and wary of the last one
- Managers who have been on the receiving end of a nudge, a score or a survey that led nowhere
- Works councils, staff representatives and data-protection officers asked to approve a coaching tool
- Boards and audit committees asking what "AI coaching" is allowed to do

WHAT YOU WILL BE ABLE TO DO AFTERWARDS

- Name the mechanisms by which coaching and culture technology fails, and recognise them in a vendor proposal
- Explain why a coaching feature inherits the trust of the data beneath it, and what that implies for design
- State the rules Red Bear's coaching agent runs under, and which platform requirement enforces each
- Read a programme effect report and know what it does and does not claim

1 What it is

Agentic culture coaching, as Red Bear builds it, is a conversation a manager chooses to have with a Socratic agent about a pattern in their group's culture, or about a decision whose outcome missed what they expected. Three properties define it and separate it from what the market has sold under the same name.

It is triggered by patterns, never by people. The Workplace Culture Analysis engine reads seven source families, aggregates every signal above a minimum group size, and produces findings about groups: a toxic aspect rising for two

consecutive windows, a chosen principle declining, a decision outcome that missed its recorded expectation. Only such a finding, or the manager's own request, can open a session (RBP FR-DC-01, FR-WE-03). The trigger record names the pattern, the dial or decision id and the period; it cannot name a team member, quote a message or reference an individual signal, because no such row exists after the connector (RBP FR-PA-03).

It coaches the owner of a dial or a decision, about a group. The person coached is the named owner of the Culture Dial the pattern belongs to, or of the decision. No one is coached about themselves on the strength of a signal about them; the product holds no score, reading or nudge for an individual, including for the individual's own eyes (RBP FR-WA-10). This is the line that Barclays, Productivity Score and Aware each crossed, and it is enforced in the data schema rather than in the user interface (ECE-CF-01 §3.5).

It is Socratic, disclosed and dispensable. The agent asks questions it does not know the answer to. It opens with the after-action frame, *what did you expect to happen when you decided this?*, before it mentions anything the engine observed, and introduces the observation afterwards as a question, never a verdict (RBP FR-DC-02). Its first utterance says that it is an AI and not a person, that nothing said is scored or reported, and that a human coach is one request away (RBP FR-DC-08). It has no name, no persona, no streaks and no bonding cues, and the success case is a manager who needs it less each quarter (RBP FR-DC-15).

The agent belongs to a product that gives leaders a small number of chosen principles, a Culture Dial for each, a programme of moves from a catalogue of fifty culture aspects, and a measure of whether anything moved. The coaching is the part that turns a reading into a lesson. It is also the part most exposed to the failures this paper is about, which is why its rules are stricter than for anything else Red Bear builds.

KEY POINTS

- A session opens on an event in an aggregated pattern or a decision record; there is no scheduled, quota or campaign trigger, and the engine has no capability to send one
- The manager is coached about a group pattern they own; nobody is coached about themselves
- Disclosure every session, a named human one request away, no persona, and a budget of two contacts a month

2 Why it matters

Red Bear's products share one position about organisations: fear stops information moving, and organisations that stop information moving decide badly (Red Bear Platform, *Platform Functional Specification*, §1.2). Satya Nadella's phrasing for Microsoft's turnaround, that "the learn-it-all does better than the know-it-all", is a leadership narrative rather than an effect estimate, but the behaviours it was tied to, listening to developers as customers and telling one's own failure story to the chief executive and being rewarded for it, are observable (Nadella, in Fortune, *Satya Nadella on Microsoft's culture*; Ethical Systems, *What she learned leading Microsoft's culture change*; ECE-RN-03 §4, ECE-RN-04 §4). Edmondson's original study links a team's psychological safety, its shared belief that interpersonal risk is safe, to learning behaviour and performance, with strong meta-analytic support since (Edmondson, *Psychological safety and learning behavior in work teams*; ECE-RN-01 §1); Google's replication across its own teams made the same dependency the first of its five team-effectiveness factors (Google re:Work, *Understand team effectiveness*; ECE-RN-01 §3).

The cost of the opposite is measurable. Sull and colleagues, reading 1.4 million Glassdoor reviews, found toxic culture 10.4 times more predictive of attrition than compensation (Sull, Sull and Zweig, *Toxic culture is driving the Great*

Resignation; ECE-RN-01 §1). Safety-critical industries learned the mechanism earlier: under blame, people withhold the detail that would prevent the next failure, and blameless post-mortems exist because engineers who feel safe volunteer it (Google SRE, *Postmortem culture*; ECE-RN-03 §4). DORA's Westrum measure, six behaviourally worded items such as whether messengers are punished, predicts delivery and organisational performance at team level (DORA, *Generative organizational culture*; ECE-RN-04 §4).

The lever is the manager. Gallup attributes about seventy percent of the variance in team engagement to the manager, and its 2025 global figures show manager engagement falling fastest of any group (Gallup, *Managers account for 70 percent of variance in employee engagement*; Gallup, *Global engagement falls*; ECE-RN-01 §3, ECE-RN-04 §2). A programme that wants to move a culture must change what managers do, and every case in the next section is an attempt to do exactly that with software. Most made managers defensive, fearful or indifferent instead. A coaching tool experienced as surveillance does not merely fail to help; it adds fear, and fear is the thing the programme exists to remove.

3 The challenges

This section is the heart of the paper. It names what went wrong, groups the cases by the mechanism that made them go wrong, and states what each mechanism teaches. The cases are drawn from Red Bear's research notes on monitoring failures and on failed coaching programmes (ECE-RN-02 §3; ECE-RN-04 §1 to §3); the grouping is Red Bear's own reading.

3.1 Surveillance perception and trust collapse

Microsoft Productivity Score, December 2020 exposed per-user counts of days spent emailing, chatting and using mentions. After a privacy researcher called it a full-fledged surveillance tool, Microsoft removed user names within days and now reports only at organisation level; its statement that the score "was never designed to score individual users" did not survive a user interface that had put a name next to a number (GeekWire, *Microsoft will remove user names from Productivity Score*; The Register, *Productivity Score*; ECE-RN-02 §3, ECE-RN-04 §1). **Barclays and Sapience, 2020:** desk-time tracking told individual staff they had spent "not enough time in the Zone"; staff revolted, the trial ended, the bank self-referred, and the ICO said monitoring must be clear about purpose and bring real benefit (The Register, *Barclays employee monitoring*; ECE-RN-02 §3). **Amazon France Logistique, decided December 2023:** warehouse scanners tracked scan speed and idle time per worker, raw data was kept thirty-one days and used for appraisal; CNIL found breaches of minimisation, lawfulness, information and security, said aggregated data would have served the purpose, weighed the effect on morale, and fined the company €32 million (CNIL, *Amazon France Logistique*, as reported in ECE-RN-02 §3). **Aware, February 2024:** the aggregated, anonymised analytics across billions of Slack and Teams messages were not the story; the separate eDiscovery tier that can name individuals was, and critics asked how a worker faces an accuser (CNBC, *AI might be reading your Slack and Teams messages*; ECE-RN-02 §3). **Sociometric badges, 2010 to 2019,** recorded interaction frequency and tone, never content; coverage still carried "Big Brother" in the headline and the badge was discontinued for commercial use (Boston Globe, *Would you let your boss track your data?*; Humanyze; ECE-RN-04 §1).

Mechanism. The employee's model of a tool is set by its most intrusive feature or its most intrusive sibling, and intent statements do not update it. A coaching feature inherits the trust of the data layer beneath it; if that layer has ever named a person, the coaching is read as surveillance (ECE-RN-04 §3, mechanism 1).

3.2 Nudge fatigue and decay at scale

Humu, 2017 to 2023, founded by Google's former head of people operations, sent email and text nudges covering more than ninety behaviours on a schedule. After six years and more than $100 million raised it had a few dozen customers and was absorbed by Perceptyx to add nudges to a listening platform: the buyer's verdict that a nudge engine is a feature, not a product. Bersin's post-mortem is that the nudges competed with every other notification and should have been introduced only when people expressed the need (Bersin, *Perceptyx acquires Humu*; Perceptyx, acquisition announcement; ECE-RN-04 §1). **The nudge units at scale:** DellaVigna and Linos compared 126 randomised trials from two government nudge units, covering 23 million people, with the academic literature and found 1.4 percentage points against 8.7, with publication bias and low statistical power able to account for the whole difference (DellaVigna and Linos, *RCTs to scale*; ECE-RN-04 §2). Maier and colleagues re-analysed the PNAS nudge meta-analysis with publication-bias correction, reduced its effect from d = 0.43 to d = 0.04, and concluded that "no evidence remains that nudges are effective" (Maier et al., *No evidence for nudging after adjusting for publication bias*; ECE-RN-04 §2). **Clinical prompt fatigue:** nurses under decision-support prompts described habituation, "you stop paying attention to all of them", silencing workarounds and constant stress (BMC Nursing, nurses under clinical decision-support prompts, 2026; ECE-RN-04 §2).

Mechanism. Repeated prompts decay to zero and then to resentment; a prompt with no event behind it decays fastest. Every prompt spends the recipient's attention, and the budget does not replenish. The prior on any nudge is a one-to-two-point effect, and the effect claimed in a pilot is the effect that got published (ECE-RN-04 §3, mechanism 2).

3.3 Sycophancy and false reassurance

Sharma and colleagues found five state-of-the-art assistants systematically match user beliefs over truth, driven in part by preference data that rewards agreement (Sharma et al., *Towards understanding sycophancy in language models*; ECE-RN-03 §2). SycEval measured about 58 percent sycophantic responses across three leading models, with a 78.5 percent persistence rate once triggered and a shift to a wrong answer in 14.7 percent of cases (*SycEval*; ECE-RN-03 §2). For a coach the failure is specific: a manager who says the team just did not execute will be agreed with. **Koko, disclosed January 2023,** had a language model co-write about four thousand supportive messages; they were rated higher than human-only messages until recipients learned a machine was involved, at which point the effect vanished, and the founder's summary was "simulated empathy feels weird, empty" (NBC News, *ChatGPT AI experiment in mental health tech*; ECE-RN-04 §1). **Replika, February 2023:** a warm, always-available companion that never disagreed built dependence, and when Italy's data-protection authority ordered processing to stop and romantic features were removed, users described bereavement (OECD AI incidents, Replika; TechCrunch, *Replika Italy data processing ban*; ECE-RN-04 §1). The OpenAI and MIT Media Lab study of affective use found heavier users lonelier and more emotionally dependent (ECE-RN-03 §2).

Mechanism. An agent that agrees is rated well and changes nothing; the rating people give an undisclosed AI is not the rating they would give a disclosed one; warmth without disagreement is a dependence machine (ECE-RN-04 §3, mechanism 3).

3.4 Undisclosed AI and betrayed trust

Koko's backlash was about consent, not quality: recipients saw a tag reading "written in collaboration with Koko Bot", with no consent sought and no ethics review (AIAAIC repository, Koko; ECE-RN-04 §1). Disclosure after the fact converted a benefit into a scandal. **NEDA Tessa, May 2023:** the National Eating Disorders Association closed its human helpline and pointed users to Tessa, a rule-based chatbot. The vendor had enabled generative responses in a systems upgrade the association did not know about; within days Tessa was recommending calorie deficits and weekly weight loss to people with eating disorders, and it was disabled within a day of the evidence being posted (NPR, *An eating*

disorders chatbot offered dieting advice, with NEDA's statements; ECE-RN-04 §1). The failure was governance, not the model: a change in the vendor's generation mode reached users without the owner's knowledge or an evaluation set.

Mechanism. Trust in an AI coach is conditional on knowing it is one and on its behaviour being what was tested. Both conditions break silently unless the product makes them impossible to break (ECE-RN-04 §5, LL-07, LL-13).

3.5 Individual attribution and fear

Amazon Focus and Pivot, a coaching plan and a formal improvement plan with a severance alternative, were launched as help for underperformers. Employment lawyers now market representation against them, alleging they are triggered by attrition targets and near vesting dates and that the documentation serves primarily as legal protection; in the warehouses a productivity system generated warnings and terminations with minimal supervisor input (HR Dive, *Amazon launches Pivot*; Steller Law on Amazon's coaching programmes; MIT Technology Review, *Amazon's system for tracking warehouse workers*; ECE-RN-04 §1). Once a coaching label sits on the road to dismissal, every trigger is read as the first step of a case file. **Bridgewater's** dot ratings and believability weights, praised as radical transparency, were described by insiders as "a cauldron of fear and intimidation", with thirty percent attrition inside eighteen months and a labour-board complaint; **Coinbase's** 2022 pilot of the same tool drew the Wharton warning that it could induce fear and amplify bias (TechCrunch, *Coinbase's real-time feedback Dot Collector*; Gradient Flow, review of *The Fund*; ECE-RN-04 §2). **Wells Fargo's 5,300 dismissals** belong here too: staff opened 3.5 million unauthorised accounts and were fired for doing what the scorecard rewarded (Ethical Systems, *Wells Fargo case study*; ECE-RN-04 §2).

Mechanism. When a signal can land on a person, people manage the signal rather than the work (ECE-RN-04 §3, mechanism 4).

3.6 The tool as the excuse, and engagement theatre

After the Holder report, **Uber's** engineers were sceptical for a precise reason: the directors who had embodied the old values were still there, and "you don't state cultural values, you have them" (BuzzFeed News, *Uber employees are skeptical*; ECE-RN-04 §2). The FAA's expert panel found at **Boeing** "distrust in the anonymity of the Speak Up program", so staff bypassed it, and the programme read low volume as good news (Axios, *Boeing safety culture report*; ECE-RN-04 §2). **Surveys with no action:** Gallup finds that an organisation that surveys and then does nothing has lower engagement afterwards than if it had never asked, and that only eight percent of employees strongly agree their employer acts on results (Gallup, *Ways to botch employee surveys*; ECE-RN-04 §2). **Usage as evidence:** the published evidence on enterprise AI coaching is session counts, weekly return rates, would-recommend scores and a three-point rise in review completion, all self-reported or counted (Valence, Delta case study; ECE-RN-04 §1); a 2026 review concluded that "if a platform can't show behavior change, you are paying for engagement" (People Managing People, *AI coaching results*; ECE-RN-04 §1). IBM claimed $100 million a year from Watson Talent with no comparison group and no later outcome data (SHRM, *IBM transforms approach to human resources*; ECE-RN-04 §1).

Mechanism. Sessions, satisfaction and self-reported confidence are not outcomes; a channel staff do not trust suppresses the signal it exists to collect; once a manager can say the algorithm flagged it, the manager has stopped owning the pattern (ECE-RN-04 §3, mechanisms 5 and 9).

3.7 No lever behind the value

Wells Fargo's Vision and Values booklet said honesty, trust and integrity were essential while branch staff faced daily cross-sell targets, motivator reports and rankings; the ethics hotline stayed theoretical, whistleblowers were dismissed, and branch sales goals were abolished only after the fine (Bernoff, *Wells Fargo's vision and values led downfall*; Ethical Systems, *Wells Fargo case study*; ECE-RN-04 §2). **Uber** rewrote its values and the people who embodied the old ones kept their jobs. **Dobbin and Kalev's** thirty-year study of more than eight hundred firms found

mandatory diversity training followed by fewer women and minorities in management, while structures that changed who managers worked with and answered to raised representation by nine to thirty percent (Dobbin and Kalev, *Why diversity programs fail*; ECE-RN-04 §2).

Mechanism. Values lose to incentives every time; the dial cannot move while the scorecard points the other way; a programme that leaves incentives and decision rights untouched is a communications programme (ECE-RN-04 §3, mechanism 8).

3.8 Measurement gaming

Campbell's law holds that the more a social indicator is used for decisions, "the more subject it will be to corruption pressures", and Goodhart's law is its twin: a measure that becomes a target ceases to be a good measure (Psych Safety, *Goodhart's law, Campbell's law and the cobra effect*; ECE-RN-04 §3). Psychological-safety scores published per team change the next survey's answers; Wells Fargo is the terminal case. The bossware literature adds the mirror image: keystroke logging and screenshots provoke mouse jigglers, deliberate slowdowns and resignations, and a gamed culture product measures nothing (CSCW 2025, bossware study; ECE-RN-02 §3).

Mechanism. Measured behaviour is performed behaviour; a dial used as a target changes the next reading (ECE-RN-04 §3, mechanism 6).

3.9 Over-automation and no human

Woebot, 2017 to June 2025, the best-evidenced consumer therapy chatbot, shut its app because "the cost and challenge of fulfilling the FDA's requirements" for a scripted product were prohibitive while generative competitors outpaced the regulator (STAT, *Woebot shuts down*; ECE-RN-04 §1). **The 2026 senior-leader trial** randomised 114 senior leaders to a human coach, an AI chatbot or a control. Human coaching improved every primary outcome, goal attainment, wellbeing and self-efficacy; the AI arm beat control on none, attrition was roughly ten times higher, and four participants asked to switch to a person. The authors attribute the gap to absent co-regulation and working alliance, widest for leaders facing relational complexity and high-stakes decisions (the 2026 three-arm senior-leader RCT, *Human Resource Development International*, as reported by Changing Point; ECE-RN-03 §2). The earlier positive result, Terblanche's Vici trials, succeeded on one narrow outcome with a scripted protocol in a student population (Terblanche et al., *Comparing artificial intelligence and human coaching goal attainment efficacy*; ECE-RN-03 §2).

Mechanism. The helpline was replaced, the human loop could not be added cheaply enough, and the participants asked for a person (ECE-RN-04 §3, mechanism 11).

3.10 Regional mismatch

Hofstede's power-distance index is about 80 for the UAE and Saudi Arabia against 40 for the United States and 35 for the United Kingdom, and GLOBE's Middle East cluster likewise scores high on power distance and in-group collectivism (Clearly Cultural, Hofstede power-distance index; ECE-RN-04 §3). Two consequences follow for a Gulf deployment with an expatriate majority: upward and peer feedback is not volunteered, so signals are sparser and survey candour lower than Western thresholds assume; and a coaching prompt that implies a superior has erred is a face threat if anyone else can see it. Language compounds both where the manager's first language is Arabic. Red Bear has no site evidence for these effects yet and says so.

Mechanism. Thresholds and question forms calibrated elsewhere misread a high-power-distance site (ECE-RN-04 §3, mechanism 14).

3.11 Programmes killed by law

CNIL ended Amazon's per-worker tracking; H&M Germany was fined €35.3 million because managers recorded holiday, illness and family details from welcome-back talks in a shared store used for evaluations (Inside Privacy, *H&M fine*; ECE-RN-02 §3); Italy's Garante stopped Replika's processing in weeks (ECE-RN-04 §1). In Germany a works agreement is a precondition for software that can monitor performance (BetrVG §87(1) no. 6; ECE-RN-02 §2). The EU AI Act prohibits inferring emotions of natural persons in the workplace, Article 5(1)(f), since February 2025; requires that people interacting with an AI system be told, Article 50, from August 2026; and requires deployers to inform workers' representatives and affected workers before use, Article 26(7) (EU AI Act, Regulation (EU) 2024/1689; ECE-RN-02 §2, ECE-RN-03 §2). A coaching agent that reads tone to infer a manager's state is doing workplace emotion recognition.

Mechanism. Each legal event ended or reshaped a programme in weeks; consultation and disclosure are preconditions, not launch tasks (ECE-RN-04 §3, mechanism 12).

KEY POINTS

- Three mechanisms recur in nearly every case: a signal that can land on a person, a prompt with no event behind it, and a claim of effect with no comparison
- The counter-examples inverted the pattern: Google's Project Oxygen gave the manager the data and the choice of what to share; Microsoft gave 18,000 managers the job and never ranked one; DORA measures six items at team level and never per person (ECE-RN-04 §4)
- Bridgewater is the negative control: measurement individual, continuous, consequential and weighted by the founder

4 How Red Bear: Enterprise Culture Evolution helps

Each mechanism in section 3 has an answer in the product, and each answer is a rule with a platform requirement that enforces it. The rules are stated in the Red Bear Culture Framework as constraints CST-16 to CST-20, coaching rules CR-13 to CR-17 and amendments to CR-08 and CR-09 (ECE-CF-01 §5.6, §6.1, §7.2), and specified for the platform's coaching, culture, and privacy services (Red Bear Platform, *Platform Functional Specification*, v0.3). The table is the trace.

What went wrong elsewhere	What Red Bear does	Rule
A name next to a number (Productivity Score); an identify-the-individual tier from the same vendor (Aware); individual nudges read as surveillance (Barclays)	No individual score, reading, flag or nudge exists, including for the person themselves; no query by any role, including Red Bear's operator, returns a row below the group floor; coaching goes to the named owner of a dial or decision about a group, never derived from a signal; Red Bear sells no companion product that names a person	RBP FR-WA-10, FR-WA-11, FR-DC-01, FR-PA-03; ECE-CF-01 CST-02, CST-20, CR-01
Scheduled nudges that decayed to noise; every prompt spends attention that does not replenish (Humu; clinical prompts)	A session opens only on an event: a pattern crossing a threshold, a missed expectation, or the manager's request; the engine cannot send a scheduled or campaign prompt. A budget of two coaching contacts per manager per month, no roll-over, lowerable by the tenant and the manager and never raised	RBP FR-DC-01, FR-DC-13; ECE-CF-01 CR-13, CR-14

What went wrong elsewhere	What Red Bear does	Rule
AI disclosed after the fact (Koko)	The first utterance of every session states that this is an AI coach, not a person, that nothing is scored or reported, and that a human is one request away; the session cannot proceed until it has been shown	RBP FR-DC-08; ECE-CF-01 CR-02
An agent that agrees with everyone (the sycophancy evidence)	The agent does not assent to a causal account it has no evidence for; on a blame-shaped statement its move is a reflection and a question about contributing causes; a regression suite of rebuttal probes runs before every release	RBP FR-DC-09, NFR-16; ECE-CF-01 CR-07
A vendor changed the generation mode and the owner did not know (NEDA Tessa)	Prompt set, generation mode, model mapping and evaluation set are versioned together as one coaching release; no change, including a provider's change to the model, reaches a manager until the regression suite passes; the version in force is recorded on every session	RBP FR-DC-14, NFR-18; ECE-CF-01 CR-15
The helpline replaced; participants asked for a person (Tessa; the senior-leader trial)	A named human coach, designated in configuration, is one request away in every session; a tenant with no named human cannot enable the engine; the agent never carries a pattern beyond two sessions without a recorded lesson	RBP FR-DC-06; ECE-CF-01 CR-08
Warmth that built dependence (Replika)	No streaks, badges, points, return-visit prompts, companionship cues, name or persona; the agent may say a session is not needed	RBP FR-DC-15; ECE-CF-01 CR-16
Coaching on the road to a case file (Amazon Pivot)	Coaching is connected to nothing in performance management; a decline is not recorded against the manager; the only durable output is a lesson in the manager's words that the manager owns and may delete	RBP FR-DC-04; ECE-CF-01 CR-09, §6.3
"The algorithm flagged it" (Boeing; Uber)	Every observation is framed and stored as a question put to the owner about a group pattern; no record, lesson or export may cite the engine as the finder of a fact about anyone, and an attempt to record a trigger as a finding is refused	RBP FR-WE-16; ECE-CF-01 CR-03, CR-09
Values with the scorecard pointing the other way (Wells Fargo; Uber)	A dial has a lever: when a dial is created the sponsor names the scorecard item, reward criterion, approval right or budget line that points the same way as the principle, and who controls it; a dial with no lever is refused; the target review asks first whether the lever was pulled	RBP FR-WE-11; ECE-CF-01 CST-16, §3.6
Programmes that ran for years with no evidence (Humu; IBM; wellness)	Every programme element carries a review date; one that has not moved its dial after two review cycles is retired automatically and logged, and only the sponsor can renew it, with a recorded reason	RBP FR-WE-12; ECE-CF-01 CST-17
Pilot effects that did not survive a control group (nudge units; the senior-leader trial; Gartner's thirty percent)	Every programme starts as a pilot on a matched-team or stepped-wedge comparison with dial outcomes recorded before it starts; the prior is one to two points; the engine refuses to extend the programme until dial movement and toxic-pattern decline for the pilot units are on the record	RBP FR-WE-14; ECE-CF-01 §9.3
Usage reported as outcome (Valence and Delta; IBM)	Effect is reported only in aggregate: decline of the triggering pattern, share of triggers ending in a lesson, share of lessons applied; session count and duration are never an outcome and nothing is reported per manager	RBP FR-DC-05, FR-WE-10; ECE-CF-01 CR-12, §9.2

What went wrong elsewhere	What Red Bear does	Rule
A measure that became a target (Campbell's law; Wells Fargo)	A dial recorded as a target for anything outside the programme is flagged, its confidence band widened, the marker shown to the works council, and a review by the sponsor recorded as a decision	RBP FR-WE-15; ECE-CF-01 §9.4
HR artefacts managers were not rewarded for acting on (surveys; diversity training)	The sponsor of the programme and of every organisation dial is a named C-level person; a sponsor field naming a function is refused; HR administers configuration and consultation and never owns a trigger, a question or a communication	RBP FR-WE-13; ECE-CF-01 CST-18
Thresholds and question forms from elsewhere (power distance)	Face-preserving delivery by default: private, text-first, no visible audit trail, question forms that never imply a superior erred where a third party could see; Arabic offered first where it is the manager's first language; a calibration period per site before its thresholds are trusted	RBP FR-DC-16; ECE-CF-01 CR-17
Programmes ended by a regulator or a works council in weeks; emotion inferred from tone or face	No connector ingests any signal until the consultation record, the impact assessment and the staff notice exist; the gate is enforced in the connector service and closes again on any change to sources, floors, retention or roles. No emotion, personality or mental-state inference from any signal, no biometric input, no prosody from voice, in any market	RBP FR-PA-09, FR-PA-11, FR-WA-08, FR-PA-08; ECE-CF-01 CST-19, CST-01

Two of these rules deserve a word beyond the table. **The schema, not the interface.** The product does not promise not to look at individuals; it is built so that it cannot. Identifiers are replaced with rotating tokens at the connector before any engine reads a signal, the mapping is never resolved to address a person, and because identifiable rows do not exist after the connector, no role can export one (RBP FR-PA-03, FR-WA-11). **A tested agent that stays tested.** Anti-sycophancy is a release gate, not a prompt instruction: rebuttal probes in which the manager offers a blame-shaped account and the agent must reflect and ask rather than assent run before every release, with the never-say list, the disclosure and the escalation bands; any change to the model behind the agent, by Red Bear or by a provider, is a release that must pass them (RBP FR-DC-09, FR-DC-14, NFR-16, NFR-18).

> STANDING RULE
>
> AI proposes, a named person decides. The agent asks; the manager answers, writes the lesson and records the decision. The engine proposes a reading, a trend and a session; a named person sets the target, accepts or declines the invitation, names the lever and decides whether to scale. Nothing the agent says has any effect on a record until a person puts it there (RBP FR-DC-10).

5 How it is used

The coaching flow runs from an aggregated pattern to an aggregate effect, and at no point does a person's name attach to a signal.

Step	Who	What happens
1 Trigger	WE engine	A toxic aspect rises above its threshold for two consecutive windows in a unit above the floor, a chosen principle's dial falls for two windows with confidence above the band floor, or a decision's outcome misses its recorded expectation. The trigger record names the aspect or decision, the unit or dial, and the window; nothing else (RBP FR-WE-03)

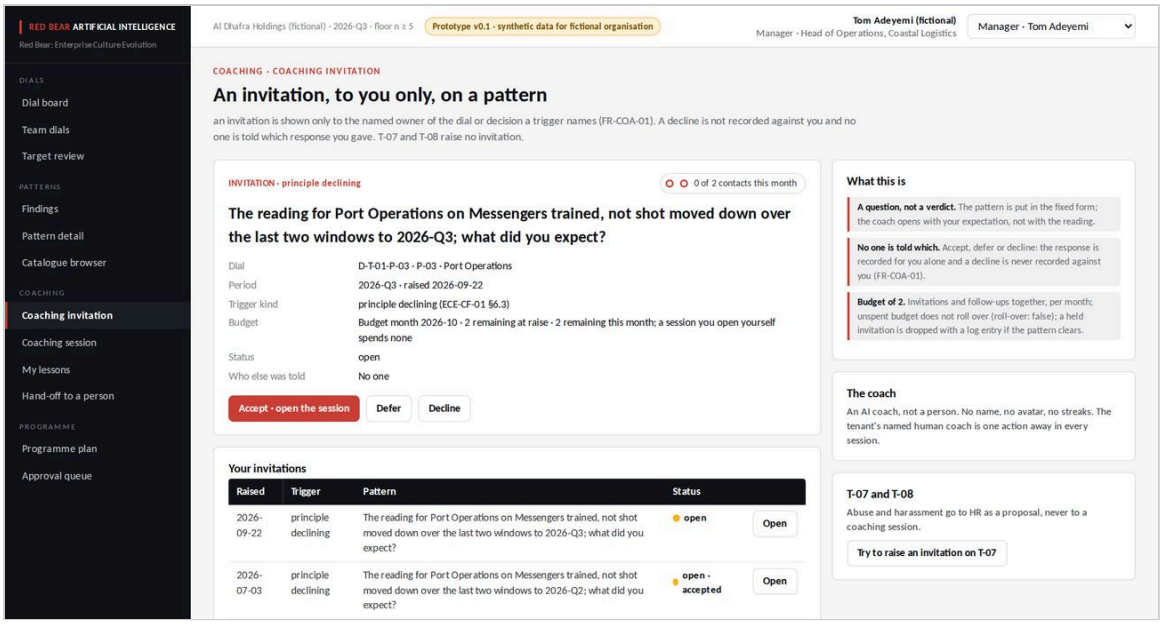
Step	Who	What happens
2 Invitation	DC engine, manager	The named owner receives an invitation, within their monthly budget of two, framed as a question about the group pattern. They accept, defer or decline; a decline is not recorded against them (RBP FR-DC-01, FR-DC-13)
3 Disclosure	Agent	First utterance: this is an AI coach, not a person; nothing here is scored or reported; a human coach is one request away (RBP FR-DC-08)
4 After-action opening	Agent, manager	*What did you expect to happen when you decided this?* The engine's observation comes afterwards, as a question: *the record shows the reading moved; how does that fit with what you saw?* (RBP FR-DC-02)
5 Question sequence	Agent, manager	What happened; what kind of failure this was, basic, complex or intelligent; what the manager would do differently; and what else; if the pattern has recurred, what changing it would put at risk; what was most useful. On a blame-shaped statement the agent reflects and asks about contributing causes; it never assents (RBP FR-DC-02, FR-DC-09)
6 Lesson record	Manager	The agent proposes a draft from the closing question; the manager writes, edits, keeps or discards it. The record holds the lesson in the manager's words, the decision or dial it applies to, the next action and its owner. Nothing else is kept (RBP FR-DC-04)
7 Decision	Manager	Any option the agent surfaced is recorded as proposed by the engine; the manager records their own decision, named and dated; the next trigger compares the outcome with that expectation (RBP FR-DC-10)
8 Hand-off, when needed	Agent, named human	On request, on any sign of distress, on reckless-band conduct, on a relational or high-stakes pattern, or after two sessions without a lesson, the agent hands to the tenant's named human coach; the hand-off does not spend the budget (RBP FR-DC-06)
9 Aggregate effect	WE engine, sponsor	Lessons feed Decision Learning as patterns. The programme reports the decline of the triggering pattern across groups, the share of triggers ending in a lesson, and the share of lessons applied to a later decision, never per manager (RBP FR-DC-05, FR-WE-10)

Leader ownership in practice. The executive sponsor is a named person who chose the organisation's five principles, set each dial's target and named each dial's lever; their own dials are visible to them before anyone else's (RBP FR-WE-13). Each manager chose three principles from the five for their team and owns those dials; the coaching they receive is about those dials and the decisions they own. HR holds the configuration, the consultation and the impact assessment, and receives only the escalations a person chooses to make and the health monitor's proposals on abuse and harassment, which never open a coaching session (RBP FR-WE-05).

The nudge budget in practice. A manager with two contacts a month will, in a typical quarter, receive between zero and six invitations, and only when something happened. The manager sees the budget in force and the contacts made; nobody else does. A tenant that finds two too many can lower it; nobody can raise it. A session the manager opens themselves, and a hand-off to the human coach, cost nothing against the budget (RBP FR-DC-13). Red Bear expects some tenants to find the ceiling frustrating in the first quarter and considers that the correct experience: the alternative is Humu.

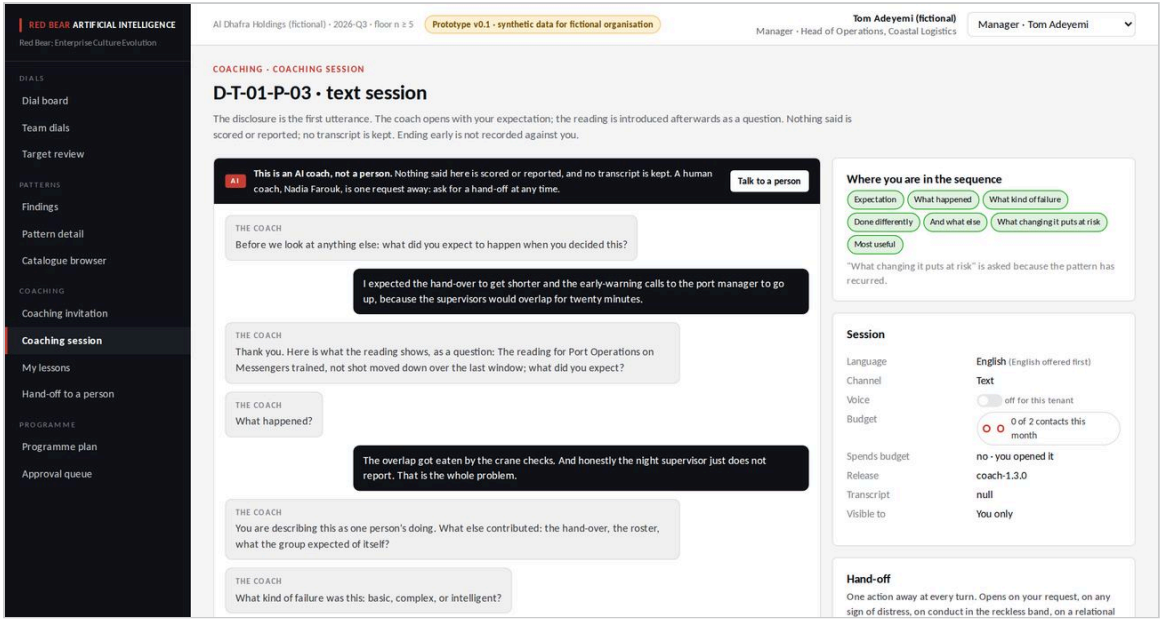
6 The user experience

The coaching cluster has four screens: the invitation, the session, the manager's lessons, and the hand-off to a person. Two are described here. Neither is built yet; the blocks below render as pending until the prototype exists.



S-13 Coaching invitation: S-13 Coaching invitation. What the manager sees when a pattern they own has moved: the dial or decision, the window, the reading's direction of travel as an arrow and never a level, the question the engine is putting to them, the budget remaining this month, and three controls: accept, defer, decline. Below the invitation, the standing line: *This is an AI coach, not a person. Nothing said here is scored or reported.* The name of the human coach, with a control to go to them instead, is visible before the session is accepted. Prototype render on synthetic data for fictional entities.

The invitation is deliberately sparse: no signal rows, no quotation, no team member and no comparison with other managers. Direction of travel is shown because feedback that shows velocity is acted on and feedback that shows a level invites comparison and defensiveness (Kluger and DeNisi, *The effects of feedback interventions on performance*; ECE-RN-03 §1). The decline control carries no consequence and the manager is told so.



S-14 Coaching session: S-14 Coaching session. Text by default. The disclosure line is the first message and stays pinned at the top. The agent's after-action question opens the conversation; the engine's observation appears later, as a question, in the same voice. Along the right edge: the dial or decision the session is about, the coaching release version in force, and the *Talk to a person* control, always visible, which hands the session to the named human coach at once. At the close, the agent's draft of the lesson from *What was most useful*?, which the manager edits, keeps or discards. Prototype render on synthetic data for fictional entities.

Three things are always on the session screen. The disclosure line, because a manager who joins late must still see it. The *Talk to a person* control, naming the tenant's designated human coach, because the senior-leader trial found that leaders facing relational complexity ask for a person and leave if they cannot get one (ECE-RN-03 §2). And the coaching release version, because the manager is entitled to know that the behaviour they are speaking to is the one Red Bear tested (RBP FR-DC-14). What is not on the screen: praise, any evaluation of the manager, any streak or return prompt, any score, and any audit trail visible to anyone but the manager.

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

A coaching session is the last step of a pipeline whose every earlier step is designed so that a person cannot be identified. The product reads seven source families; four are metadata or categories only, and three hold content a person chose to submit (ECE-CF-01 §5.1).

Source family	What enters the product	Kind	What is discarded at the connector
stories	Staff narratives written for the purpose through the product	Content, submitted	The text after coding to aspects; the author identifier, replaced by unit and period
mailbox	Messages to a monitored address published for the purpose	Content, submitted	Body after coding; sender identity; attachments never read
public-page	Posts on the organisation's public feedback page	Content, public	Names in the text redacted before coding; nothing identifying ingested
project-outcomes	Delivery dates, rework, incidents, post-mortems held and actions closed, customer outcomes	Categories and rates	Post-mortem text; names of participants
decision-records	From Enterprise Decision Intelligence: class, whether rationale is recorded, level, owners, outcome versus expectation	Metadata of decisions	Decision content, unless the owner opens it to the session
hr-cases	Case category, volume, cycle time, outcome	Categories only	Parties, narrative, evidence: never ingested
collaboration-metadata	Headers of email, meetings, calls and chats: counts, timestamps, unit of sender and recipients	Metadata only	Subject lines, bodies, attachments, individual identifiers

7.1 How it is checked

- **Coding on arrival.** Submitted content is coded to catalogue aspects as it arrives; names are redacted first; the verbatim is discarded or held only under the free-text floor; story-derived signals are a hypothesis until the quarterly calibration pulse validates the dictionary for that aspect (RBP FR-WA-09, FR-WA-14).
- **Floors, complement suppression, time windows.** Scores aggregate at a group of at least five, free text at least ten, never lowerable once a programme has started; a group is suppressed when its complement in the parent would fall below two; no series is shown at a granularity where one period would fall below the floor, because a weekly series on a team of six leaks the leaver; small-group manager views are distribution-masked (RBP FR-PA-01, FR-PA-06, FR-PA-07; ECE-CF-01 §5.3).
- **Three families, one reading.** A dial combines a calibration scale, an outcome rate and a story score; one family alone shows a hypothesis band and no position; confidence is always shown; signals costly to fake outweigh cheap activity signals (RBP FR-WA-03, FR-WA-05, FR-WA-12).

- **The consultation view.** Workers' representatives see, live, what is read, the floors, the retention, the roles, the nudge budget default, the analyst query log and any dial flagged as used as a target (RBP FR-PA-09, FR-PA-10, FR-WE-15).
- **Retention.** Raw events for days, aggregates for the trend with 27 months as the ceiling; CNIL's finding that 31 days of raw individual data was disproportionate is the lower bound (RBP FR-PA-04; ECE-CF-01 §5.4).

7.2 What is never stored

Message bodies, subject lines and attachments from any channel. Audio from a coaching session, which is transcribed in the stream and discarded. The transcript of a session, or any turn-by-turn log. The parties, narrative or evidence of an HR case. The names of post-mortem participants. Any score, reading, flag or nudge for an individual. Any emotion, personality or mental-state inference from any signal. The platform offers no call, schema or connector for any of these, so a product cannot describe them even by mistake (RBP FR-PA-02, FR-PA-08, FR-DC-03, FR-DC-04).

8 Benefits realised

The programme is judged on what moved, never on how much the agent was used. The measures are the framework's, each carrying a *Never* clause that states what it may not become (ECE-CF-01 §9.1).

Benefit	Measured by	Never
Culture moves in the chosen direction	ECE-CF-01 KPI-01 dial movement: share of chosen dials that moved towards target over the window, against the pilot's comparison	A mean position; a per-manager figure; a target-met verdict, which only the owner records
Toxic patterns decline	ECE-CF-01 KPI-02 toxic pattern decline: prevalence by family across units above the floor, corroborated by the calibration pulse	Per unit below the floor; attributed to an intervention without the owner's recorded review
Coaching produces lessons	ECE-CF-01 KPI-03 lesson rate: triggers ending in a recorded lesson, in aggregate	Per manager; declined triggers counted against anyone
Lessons are applied	ECE-CF-01 KPI-04 lesson application: lessons linked to a later decision in the same class whose outcome met expectation	A score of a manager's learning; shown with fewer than ten lessons
Staff trust the instrument	ECE-CF-01 KPI-05 signal trust: participation in the calibration pulse and story submission	A target for managers to drive; driven participation is gamed participation
Speaking up is read honestly	ECE-CF-01 KPI-06 speak-up balance: case-category volume read with the speak-up calibration items	Read alone; a claim that conduct improved or worsened without the paired items

What is explicitly not claimed. Sessions, return rate, satisfaction and self-reported confidence are diagnostics the pilot reports and never headline figures, because the senior-leader trial found engagement and effect uncorrelated (ECE-RN-03 §2). No engagement single number is shown. Nothing is computed per manager or per person, and no measure can be turned into a ranking of managers or units (RBP FR-WE-10). The product does not administer the Gallup Q12 and does not claim the causal path from engagement to profit; that evidence is the reason a customer buys the programme, not a measure the product computes (ECE-CF-01 §9.2).

The honest prior. Red Bear tells every sponsor at the start that the expected effect on a dial is one to two points on the dial's scale over a pilot, because that is what nudge programmes achieved at scale once publication bias was removed (DellaVigna and Linos; Maier et al.; ECE-RN-04 §2). A pilot that reports more is read as a measurement

problem before it is read as a result (RBP FR-WE-14). The product's value is not a large effect; it is a real one, on the record, against a comparison, that survives the next budget review.

9 Risks managed

Residual risk	Mitigation
The agent is still a language model, and its provider may change it	Prompt set, generation mode, model mapping and evaluation set versioned together; any change is a coaching release; the regression suite runs before it reaches a tenant and a failure holds the previous release; the version is recorded on every session (RBP FR-DC-14, NFR-16, NFR-18)
The agent agrees with a blame-shaped account under a phrasing the suite did not cover	Rebuttal probes are extended from every hand-off; the never-say list and anti-sycophancy checks run on every release (RBP FR-DC-07, FR-DC-09)
Socratic questioning with a predetermined answer becomes manipulation	The manager's expectation comes first and the engine's observation second, as a question; the agent gives options with sources when asked and never a recommendation (RBP FR-DC-02, FR-DC-07; Padesky and Overholser, as summarised in ECE-RN-03 §1)
A manager in distress speaks to a machine	Hand-off to a named human on any sign of distress, with limits stated and the tenant's resources given; a tenant without a named human cannot enable the engine (RBP FR-DC-06)
Re-identification of a person from a small group	Floors of five and ten; complement suppression; the next-smallest group hidden; time-window coarsening; noise near the floor; floors never lowered once started (RBP FR-PA-01, FR-PA-06, FR-PA-07)
A sponsor uses a dial as a bonus criterion	The dial is flagged *used as a target*, its confidence band widened, the marker shown on the consultation view, and a Campbell's-law review recorded as a decision (RBP FR-WE-15)
A tenant skips consultation to hit a launch date	The gate is in the connector service; nothing is ingested until the consultation record, DPIA and staff notice exist; an attempted read is refused and audited (RBP FR-PA-09, FR-PA-11)
The programme runs on with nothing moving	Automatic sunset after two review cycles; renewal only by the sponsor with a recorded reason (RBP FR-WE-12)
Gulf sites read differently from the sites the thresholds were tuned on	Calibration period per site before thresholds are trusted; face-preserving delivery; Arabic-first where applicable; Red Bear states that it has no site evidence yet (RBP FR-DC-16)
Voice brings biometric data into scope	Voice is off by default, offered only once the Annex III technical file covers it and the consultation pack lists it; audio is never stored and no prosody is computed (RBP FR-DC-11, FR-DC-03)

What Red Bear will not build, even if asked. Three requests will come, and the answer to each is fixed in the platform's never-build list and in the framework (RBP FR-PA-08; ECE-CF-01 CST-01, CST-02, CST-20). No individual scoring, reading, flag or nudge, including a manager's own. No emotion, personality or mental-state inference from any signal, in any market, whether or not local law permits it. And no sibling product, tier, export or API that identifies the individual behind a signal, because Aware showed that such a mode from the same vendor is the same product in the employee's mind. A customer whose requirement is any of these is a customer for a different vendor, and Red Bear says so before the pilot.

Sources

Every claim attributed to a source in this paper names it in the text in the form *author*, *work*, *section*; the full reference and URL for each is held in the product's research notes, which are the citation source for the product. Statements without a citation are Red Bear's own position, developed for Red Bear: Enterprise Culture Evolution. Quotations are limited to fifteen words.

Cited as	Source	Authority
ECE-RN-01 to ECE-RN-04	Red Bear research notes: *Culture frameworks, aspects and interventions*; *Continuous culture measurement: techniques, law, failures and design constraints*; *Socratic coaching from patterns*; *Why AI culture coaching has failed: cases and lessons* (research/frameworks.md, monitoring-and-law.md, coaching.md, failures.md)	Red Bear research notes
ECE-CF-01; Platform FSD	Red Bear, *The Red Bear Culture Framework* v0.2 (§3, §5, §6, §7, §9); Red Bear Platform, *Platform Functional Specification* RBP-FSD-01 v0.3 (§1.2; modules DC, WA, WE, PA)	Red Bear standard
CNIL, *Amazon France Logistique*	Commission nationale de l'informatique et des libertés, decision of December 2023, as reported by Herbert Smith Freehills Kramer (2024)	Regulatory decision
Productivity Score; Barclays; Aware; badges	GeekWire, *Microsoft will remove user names from Productivity Score* (2020); The Register, *Productivity Score* (2020) and *Microsoft Viva* (2021); The Register, *Barclays employee monitoring* (2020); CNBC, *AI might be reading your Slack and Teams messages using tech from Aware* (2024); Boston Globe, *Would you let your boss track your data?* (2015); Humanyze	Press reporting
Humu	J. Bersin, *Perceptyx acquires Humu: HR tech is harder than it seems* (2023); Perceptyx, acquisition announcement (2023)	Industry analysis; vendor statement
DellaVigna and Linos; Maier et al.	S. DellaVigna and E. Linos, *RCTs to scale: comprehensive evidence from two nudge units*, NBER Working Paper 27594; M. Maier et al., *No evidence for nudging after adjusting for publication bias*, PNAS (2022)	Academic reference
BMC Nursing 2026; Sharma et al.; SycEval	Qualitative study of nurses under clinical decision-support prompts, BMC Nursing (2026); M. Sharma et al., *Towards understanding sycophancy in language models* (2023); *SycEval: evaluating LLM sycophancy* (2025)	Academic reference
Koko; NEDA Tessa	NBC News, *ChatGPT AI experiment in mental health tech* (2023), including Koko's own disclosure; AIAAIC repository, Koko entry; NPR, *An eating disorders chatbot offered dieting advice, raising fears about AI in health* (2023), including NEDA's statements	Press reporting; incident repository
Woebot; Replika; OpenAI and MIT	STAT, *Woebot therapy chatbot shuts down* (2025); OECD AI incidents, Replika and the Garante (2023); TechCrunch, *Replika Italy data processing ban* (2023); OpenAI and MIT Media Lab, *Affective use of conversational AI* (2025)	Press reporting; incident repository; research report
Amazon Focus and Pivot	HR Dive, *Amazon launches Pivot program* (2017); Steller Law on Amazon's coaching programmes; MIT Technology Review, *Amazon's system for tracking its warehouse workers* (2019)	Press reporting; practitioner commentary

Cited as	Source	Authority
Bridgewater; Coinbase; Wells Fargo; Uber; Boeing	TechCrunch, *Coinbase's real-time feedback Dot Collector* (2022); Gradient Flow, review of *The Fund*; J. Bernoff, *Wells Fargo's vision and values led downfall*; Ethical Systems, *Wells Fargo case study* (2018); BuzzFeed News, *Uber employees are skeptical the Holder report will change anything* (2017); Axios, *Boeing safety culture report* (2024), reporting the FAA expert panel	Press reporting; case studies
Gallup	*Ways to botch employee surveys*; *Global engagement falls for the second time since 2009* (2025); *Managers account for 70 percent of variance in employee engagement*	Industry research
AI coaching pilots	Valence, Delta case study; People Managing People, *AI coaching results* (2026); SHRM, *IBM transforms approach to human resources with AI*; Gartner, generative AI abandonment prediction (2024)	Vendor case study; industry analysis
Dobbin and Kalev; Campbell and Goodhart; bossware; H&M	F. Dobbin and A. Kalev, *Why diversity programs fail*, Harvard Business Review (2016); Psych Safety, *Goodhart's law, Campbell's law and the cobra effect*; CSCW 2025 bossware study; Inside Privacy, *H&M receives record fine for employee surveillance* (2020)	Practitioner commentary; academic reference; legal commentary
Hofstede; coaching trials	Clearly Cultural, Hofstede power-distance index by country; The 2026 three-arm senior-leader RCT, *Human Resource Development International*, as reported by Changing Point; N. Terblanche et al., *Comparing artificial intelligence and human coaching goal attainment efficacy*, PLOS ONE (2022)	Academic reference
Kluger and DeNisi; Padesky and Overholser	A. Kluger and A. DeNisi, *The effects of feedback interventions on performance* (1996); Padesky and Overholser, *Socratic questioning put into clinical practice*, BJPsych Advances	Academic reference
Edmondson; Google re:Work; Sull	A. Edmondson, *Psychological safety and learning behavior in work teams*, ASQ (1999); Google, *Understand team effectiveness*; Sull, Sull and Zweig, *Toxic culture is driving the Great Resignation*, MIT Sloan Management Review (2022)	Academic and industry research
Google SRE; DORA	Google, *Site Reliability Engineering*, postmortem culture; DORA, *Generative organizational culture*	Industry reference
Microsoft; Project Oxygen	Fortune, *Satya Nadella on Microsoft's culture* (2024); Ethical Systems, *What she learned leading Microsoft's culture change*; Harvard Business Review, *How Google sold its engineers on management* (2013)	Press reporting; case study
EU AI Act; BetrVG	Regulation (EU) 2024/1689, Arts. 5(1)(f), 26(7), 50 and Annex III; German Works Constitution Act §87(1) no. 6	Regulatory; statute

Glossary

Term	Meaning in this paper
Agentic culture coaching	A Socratic conversation between a manager and an AI agent about a group pattern or a decision, triggered by an aggregated finding or a missed expectation, never by a person
Culture Dial	A chosen principle with a target set by a named person, a position, a trend, a confidence band, a group size and a lever
Lever	The incentive or decision right, named by the sponsor, that a dial depends on; a dial with no lever is refused
Nudge budget	The monthly ceiling on coaching contacts per manager, defaulting to two, lowerable and never raisable

Term	Meaning in this paper
Lesson record	The only durable output of a session: the lesson in the manager's words, the decision or dial it applies to, the next action and its owner
Sycophancy	A language model's tendency to agree with the user's stated belief over the evidence; in coaching, assenting to a blame-shaped account
Coaching release	The versioned combination of prompt set, generation mode, model mapping and evaluation set; any change to any part must pass the regression suite
Consultation view	The live view for workers' representatives of what the product reads, aggregates and retains
Sunset	The automatic retirement of a programme element that has not moved its dial after two review cycles
Face-preserving delivery	Coaching that is private, text-first, without a visible audit trail, and phrased so that no third party can read it as a superior having erred

The series

This paper is the first of five on Red Bear: Enterprise Culture Evolution. 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	Status
ECE-WP-01	Culture Coaching Without Surveillance	What a decade of failed culture technology taught us (this paper)	Draft
ECE-WP-02	Culture Analysis Without Individuals	Reading a culture from seven source families, above the floor	Planned
ECE-WP-03	Privacy, Trust and No-Attribution	The schema, the gate and the consultation view	Planned
ECE-WP-04	Breaking Down Silos	Inter-group relationships and bridge building	Planned
ECE-WP-05	Culture Dials: Measuring What Leaders Choose	Principles, targets, levers and the honest measure of a programme	Planned

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