

Methodology & Research Foundations

A short, sourced account of what the Arc Diagnostic measures, the reasoning behind it, and — honestly — what has and has not yet been validated. It is the organisational companion to the Judgment Premium Assessment: where the JPA measures an *individual's* judgment, the Arc Diagnostic reads an *organisation's* AI maturity.

The one-paragraph version. The Arc Diagnostic places an organisation in one of three arcs of AI maturity — adoption, optimisation, or embedded judgment at scale — using twelve questions across four dimensions (Adoption, Alignment, Autonomy, Governance), run once per track. It rests on the finding that AI maturity is *organisational, not technological*. It is a *practitioner-derived reflective instrument* — useful for planning and shared diagnosis, but *not* a validated benchmark, and not a basis for ranking organisations or people.

Why it exists

Most AI-maturity models score the technology: how many tools, how much spend, how advanced the models. That measures the wrong thing. The largest enterprise study to date found that **77% of the hardest challenges** in successful AI deployments were *invisible costs* — change management, data quality, process redesign — not model performance, and that the same use case took *weeks* in one organisation and *years* in another, the difference turning on executive sponsorship and end-user willingness.

Pereira et al. (2026), Stanford Digital Economy Lab — enterprise AI deployment analysis.

The conclusion the Arc Diagnostic is built on: **AI maturity is organisational, not technological**. What determines whether AI takes hold is culture, alignment, and governance — not the sophistication of the model. So the instrument measures the organisation, not the tooling.

What it measures — the two structures

The Three Arcs (the result)

ARC	THE WORK OF THIS ARC
Arc 1 · Adoption	Bringing AI into the organisation — getting real, trusted, sustained use.
Arc 2 · Optimisation	Optimising what adoption introduced — aligning to value, connecting across boundaries.
Arc 3 · Embedded judgment at scale	Governing autonomous systems at scale, with human accountability intact.

The Four Dimensions (the measurement)

DIMENSION	WHAT IT CAPTURES
Adoption	How widely and habitually AI is used, trusted, and sustained across the workforce.
Alignment	How well AI work is tied to business value, with stakeholders aligned on success.
Autonomy	How far the organisation has moved toward independent, connected, scalable AI.
Governance	The oversight, accountability, and data governance surrounding AI.

Three questions per dimension, twelve in total, each scored 1-5. The total (12-60) maps to Arc 1 (12-24), Arc 2 (25-42), or Arc 3 (43-60).

The two tracks — why you run it twice

AI maturity is not one number, because organisations advance on two different tracks that rarely move together:

- **Workforce Augmentation** — the organisation-wide effort to empower every professional to work with AI. Culture, leadership, and mindset do the heavy lifting; success is genuine, sustained change in how people work.
- **Enterprise Transformation at Scale** — building purpose-built agents and applications that connect across departmental boundaries. Large organisations stall here because those handoffs connect not only processes but people, whose politics and power dynamics create friction.

An organisation is often strong on one track and weak on the other — so the diagnostic is run once per track, producing an arc for each. Reporting a single blended score would hide exactly the gap that matters most.

How the questions work

The respondent selects a track, then rates twelve statements from "Not at all" (1) to "Fully" (5), interpreting each through that track's lens. Statements are written so that a higher score reflects greater maturity on that dimension. The design is deliberately simple and self-reported: it is a structured conversation-starter for a leadership team, not a covert measurement. Its value is in surfacing disagreement — where two leaders score the same dimension very differently, the diagnostic has done its job.

Best used as a team exercise. The book recommends running the Arc Diagnostic as a team exercise before a planning cycle. Individual scores reveal how differently people inside the same organisation perceive its maturity — which is itself a finding worth discussing.

What is validated — and what is not

The honest part. The Arc Diagnostic is a *practitioner-derived reflective instrument*. Its dimensions, arcs, and scoring bands are grounded in the three organisational case studies in the book (a large Australian financial-services firm, AXA China, and a large hospitality company) and corroborated by the Stanford enterprise findings — but the specific question set and score thresholds have *not* been empirically validated against organisational outcomes. It is a lens for reflection and planning. It is *not* a validated benchmark, an audit, or a basis for ranking organisations, teams, or individuals.

The path to stronger evidence:

Phase 1 — Content validity: a panel of transformation leaders reviews each question against its dimension and the arc definitions; items with weak agreement are revised.

Phase 2 — Internal consistency: with enough completed diagnostics, check that the three items in each dimension move together, and that the dimensions discriminate rather than collapse into one.

Phase 3 — Known-groups: test whether organisations independently assessed as mature (e.g. high sustained adoption, agents in production) score higher, confirming the bands separate real differences.

Phase 4 — Predictive signal: over time, examine whether a track's arc relates to deployment outcomes — the weeks-versus-years gap the Stanford study identified.

How it relates to the Judgment Premium Assessment

The two instruments share one framework at two scales. The JPA measures the **individual**: whether a professional exercises judgment when AI can do the work. The Arc Diagnostic measures the **organisation**: whether the culture, alignment, autonomy, and governance are in place for that judgment to compound. An organisation full of high-judgment individuals will still stall in Arc 1 without the organisational conditions; strong organisational maturity without individual judgment scales risk rather than value. Read together, they locate both the person and the system.

Selected references

Pereira et al. (2026). Enterprise AI deployment analysis. Stanford Digital Economy Lab.

Susan K.L. Yu (2026). *Judgment in the Age of AI: The Four Disciplines That Make You Irreplaceable*. Green Hill Publishing — Chapter 9, The Three Arcs of AI Maturity.

Industry context: McKinsey, "Superagency in the Workplace" (2025); WEF, Future of Jobs Report (2025); EU AI Act (2025).

The Arc Diagnostic, the Three Arcs, the 4C Mindset™ and the Four Disciplines are the intellectual property of Susan K.L. Yu / Altschool Australia. This is a practitioner-derived reflective instrument on the validation path described above; it is not a validated benchmark and not a basis for ranking. © 2026 Susan K.L. Yu · Altschool Australia.