AI Readiness Compass Methodology Brief

An organizational and workforce AI readiness diagnostic: how the assessment was built, where the items come from, and our approach to establishing validity.

Version 1.3 · 17 November 2026

1. Purpose of this brief

Consultants and clients ask three questions about any psychometric instrument: *Where did the items come from? How do you know they measure what you say they measure? How will accuracy improve over time?* This brief answers those three questions for the AI Readiness Compass.

2. Construct definitions

Two scoring planes

The Compass measures AI readiness on two distinct planes:

Organizational (8 pillars). *Strategy & Vision · Data Foundations · Technology & Infrastructure · Talent & Skills · Culture & Change Readiness · Governance, Ethics & Compliance · Operations & Use Case Portfolio · Model Management & Monitoring.* Used for Department, Division, and Enterprise stages; a standard Department engagement covers four of the eight pillars, a Division engagement six, and an Enterprise engagement all eight. A respondent answering on this plane reports on the organization's posture, not their own.

Individual (4 factors). *AI Sense-Check · AI Working Practice · AI Collaboration · AI Adaptive Mindset.* Used for the Personal Snapshot, the consultant-led Personal Deep-Dive, and the optional individual layer in any organizational engagement (departmental-plus-personal). A respondent answering on this plane reports on their own behavior.

An optional Agentic AI Readiness layer (six dimensions: Agent Governance & Accountability, Human-in-the-Loop & Oversight, Failure-Mode & Risk Awareness, Tool & Data Access Control, Autonomy Calibration, Auditability & Traceability) can be added to an organizational engagement. It is scored separately from the eight pillars.

Why two planes

The organizational instrument is designed with reference to regulatory and industry frameworks in use in the UAE and Saudi Arabia, for example the UAE Personal Data Protection Law (Federal Decree-Law No. 45 of 2021; regulator: UAE Data Office), Saudi Arabia's Essential Cybersecurity Controls (ECC-2:2024; National Cybersecurity Authority) and Saudi Arabia's national data governance regulations and

standards (SDAIA, through its National Data Management Office). The pillars are aligned with these frameworks at the level of their concepts; the Compass does not audit, certify or determine legal compliance with any of them (see §6). Readiness here is largely a property of policy, capability, and governance. The individual instrument captures the behavioral skills literature on workplace AI adoption - readiness here is a property of how a person engages with the tools in front of them. The two planes inform each other but answer different questions.

---

3. Item development

Sourcing

The current production bank (v1.1; 213 items in service as of 16 November 2026) was developed through three rounds:

1. **Initial drafting.** Items were first drafted by VIFM consultants drawing on (a) the VIFM behavioral competency framework as it stood when the items were drafted (38 competencies and 249 behavioral indicators at that time; the framework has since been consolidated into version 2, with 21 competencies in 8 clusters and 4 domains), so that the items would be recognizable to consultants who already used that framework with clients, (b) reference frameworks listed in §6, and (c) field questions consultants frequently ask in pre-assessment workshops.
2. **AI-assisted expansion.** A subset of items was drafted with the assistance of large language models, working from the construct definition and a fixed set of answer formats. Items drafted with AI assistance are reviewed by VIFM consultants. Each item in service is also reviewed individually by a subject-matter expert: as of 16 November 2026, 208 of the 213 items in service are approved in that review and 5 are pending. Results that draw on an item still under review are marked provisional.
3. **Bilingual authoring.** Every item is provided in English and Arabic. The Arabic wording was written for a Gulf audience rather than produced by back-translation, and is being reviewed by Arabic-speaking subject-matter experts; corrections from that review are applied in the next bank version.

Coverage

Every item in the question bank is tagged to a pillar (organizational) or a factor (individual), and the platform validates the tag against the list of recognized pillars and factors, so every active item has a known target construct.

For the individual instrument, the four factors map onto the four domains of the VIFM competency framework (Thinking, Results, People, Self), and each factor lists the two to four behavioral competencies it draws from. This explicit competency lineage lets a consultant trace any factor score back to the development tips that exist for those competencies (the same tips used in the Assessment Center learning plan).

Format

Individual items are mostly five-point agreement ratings, from "1 - Strongly Disagree" to "5 - Strongly Agree". The individual bank also includes a small set of scenario-based and knowledge-check items, which are scored against an expert key on the same 1 to 5 scale. Organizational items are mostly rated

on a five-point scale from "1 - Not at all" to "5 - Comprehensive"; a smaller number are yes/no or multiple-choice items whose options each carry a score on the same scale, and open-text items are not scored. Pillar results are reported on five maturity levels (Unaware, Exploring, Developing, Advancing, Leading) and individual results on three stages (Emerging, Practicing, Embedded). The score assigned to each response option is stored with the item, and every assessment records the question bank version it was created against, so scoring can be reproduced.

---

4. Validity

Content validity

Every item is mapped to exactly one construct (a pillar or a factor). The construct definitions in §2 were reviewed by VIFM consultants and the competency framework owner before items were accepted. A second-round review checked that each construct had at least four items measuring it, so the score is not over-reliant on a single behavioral facet.

Face validity

Items use observable behaviors (e.g., "I check AI-generated content for factual errors before relying on it for important work") rather than abstract attributes. Respondents and consultants in pre-pilot workshops reported they could explain in their own words what each item was asking, which is the bar for face validity.

Construct validity

Construct validity asks whether the four-factor model adequately represents the proposed underlying constructs. What has been done: every item is mapped to one construct, and the construct definitions and the item-to-construct mapping were reviewed by VIFM consultants. This is content evidence from expert review. What has not been done: the four-factor structure has not been tested empirically. A confirmatory factor analysis is planned once at least 200 completed individual responses are available; responses are being accumulated through the free Personal Snapshot, and the loadings, fit indices and any item revisions that result will be published. The figure of 200 is a study-planning minimum for running the analysis; reaching it is not evidence of validity, and the analysis may support, refine or reject the four-factor structure. Until then the four factors are a working model based on content review, and factor scores should be read as descriptive summaries of groups of items. This disclosure appears in the appendix of the organizational report, including departmental-plus-personal engagements, and personal results link to this brief.

Criterion validity

The Compass is a developmental and diagnostic instrument, not a hiring screen. We do not currently make claims about predictive validity (does an Embedded score predict on-the-job AI productivity?) and we do not recommend the Compass be used for selection decisions until that work is done. A talent-acquisition view of the personal assessment exists to inform interview conversations with a candidate. It is descriptive only and must not be the sole or main basis for a selection decision.

5. Reliability

Internal consistency

Cronbach's alpha will be reported per factor and per pillar once N ≥ 200 individual responses and N ≥ 50 organizational responses. Pre-publication, we treat reliability as a property of the item-development process (multiple items per construct, expert-reviewed) rather than a property of an empirical sample.

Inter-rater reliability (organizational Phase 2)

In the Phase 2 workshop the consultant records a validated score for each pillar alongside the organization's self-assessed score, and the report shows the difference between the two (perception versus reality). This is a comparison of self-assessment with consultant judgment; it is not inter-rater reliability. Agreement between consultants is not yet measured, because the platform currently stores one validated score per pillar rather than each consultant's separate rating. We plan to add this and to report agreement once at least 30 workshops with two or more raters are complete.

Test-retest reliability

The platform supports annual reassessment by linking each organizational reassessment to the assessment it follows. Individual retakes are not currently linked in the same way. The year-on-year comparison shows how an organization's pillar scores have changed between annual assessments. This measures change over a year, during which the organization is expected to act on its results, so it is not a test-retest reliability estimate. A test-retest study, in which the same respondents repeat the assessment after a short interval, is planned.

6. Reference frameworks and published instruments

Regulatory frameworks referenced

United Arab Emirates (7): UAE Personal Data Protection Law (Federal Decree-Law No. 45 of 2021) · UAE National AI Strategy 2031 · UAE Charter for the Development and Use of Artificial Intelligence (UAE Artificial Intelligence, Digital Economy and Remote Work Applications Office) · UAE AI Ethics Guide · TDRA Digital Government Regulations · Dubai Centre for Artificial Intelligence (DCAI) Guidelines · Department of Government Enablement - Abu Dhabi (DGE) Data and AI Standards (DGE absorbed the former Abu Dhabi Digital Authority in 2023)

Saudi Arabia (9): Saudi Personal Data Protection Law (Royal Decree M/19, as amended) · SDAIA National Data Governance Framework · NCA Essential Cybersecurity Controls (ECC) · NCA Cloud Cybersecurity Controls (CCC) · SDAIA AI Ethics Principles (seven principles) · SDAIA AI Adoption Framework · Saudi Vision 2030 (AI and digital transformation targets) · SDAIA Generative AI Guidelines for Government · SAMA Cyber Security Framework

The organizational pillars are designed with reference to the frameworks above. Each framework's requirements are mapped to specific organizational pillars. The report shows, for each framework, an

indicative alignment status (aligned, partly aligned or not aligned) derived from the self-assessed score of the mapped pillar, which a consultant can review and override with evidence. This is a readiness indication to guide further review; it is not a legal opinion, an audit, or a determination of compliance with any law or standard.

Alignment is VIFM's own mapping against the published standard or law; it is not a certification, accreditation, audit or legal opinion by the body concerned.

Assessment-design standards referenced

The Compass is designed with reference to the standards below. Alignment is VIFM's own mapping against the published standard or law; it is not a certification, accreditation, audit or legal opinion by the body concerned.

- ISO 10667 - Assessment service delivery: Procedures and methods to assess people in work and organizational settings (Parts 1 and 2)
- International Taskforce on Assessment Center Guidelines (6th Edition) - applied to the related Assessment Center service; framework consistency carries forward
- Standards for Educational and Psychological Testing (AERA / APA / NCME, 2014) - referenced for evidence-based item development

Published instruments the items content-align with

For every construct the Compass measures, each item is linked to one or more established frameworks, models or instruments from the published literature that address the same construct. The link records content alignment only; it is assigned after drafting, with AI assistance from a fixed list of sources, and confirmed by a reviewer. This is *content alignment* - items adapt the construct, not the wording - so we do not claim the same psychometric coefficients as the source instruments. Each item's validation-evidence record refers back to the works listed below. Citations are deliberately limited to seminal, well-established works so that each one can be readily checked.

Individual layer - four factors

AI Sense-Check (Thinking). Critical evaluation of AI output, detection of fabricated or false output, domain validation.

- Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations.* Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). *Conceptualizing AI literacy: An exploratory review.* Computers and Education: Artificial Intelligence, 2, 100041.

AI Working Practice (Results). Productive hands-on use of AI in real workflows.

- Davis, F. D. (1989). *Perceived usefulness, perceived ease of use, and user acceptance of information technology.* MIS Quarterly, 13(3), 319–340.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). *User acceptance of information technology: Toward a unified view.* MIS Quarterly, 27(3), 425–478. - the unified theory of acceptance and use of technology (UTAUT).
- Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). *Generative AI at work.* Quarterly Journal of Economics, 140(2), 889–942. - productivity outcomes from real-world AI use.

AI Collaboration (People). Helping the team move with AI; shaping shared norms.

- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). *Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology.* MIS Quarterly, 36(1), 157–178. - UTAUT2; social influence in technology adoption.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge University Press. - knowledge sharing within practice communities.

AI Adaptive Mindset (Self). Curiosity, openness to relearning, responsible posture.

- Parasuraman, A., & Colby, C. L. (2015). *An updated and streamlined Technology Readiness Index: TRI 2.0.* Journal of Service Research, 18(1), 59–74.
- Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House. - growth-mindset theory.

Organizational layer - eight pillars

Strategy & Vision. AI mandate, executive sponsor, budget, use case prioritization.

- Davenport, T. H., & Ronanki, R. (2018). *Artificial intelligence for the real world.* Harvard Business Review, 96(1), 108–116.
- Iansiti, M., & Lakhani, K. R. (2020). *Competing in the age of AI.* Harvard Business Review Press.

Data Foundations. Data quality, ownership, sovereignty, shadow AI risk.

- DAMA International (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd ed.). Technics Publications.
- EDM Council (2020). *Data Management Capability Assessment Model (DCAM).* Enterprise Data Management Council.

Technology & Infrastructure. Cloud sovereignty, AI tool approval, MLOps, sandbox.

- Sculley, D., Holt, G., Golovin, D., et al. (2015). *Hidden technical debt in machine learning systems.* Advances in Neural Information Processing Systems (NeurIPS) 28, 2503–2511.
- Lwakatare, L. E., Raj, A., Bosch, J., Olsson, H. H., & Crnkovic, I. (2019). *A taxonomy of software engineering challenges for machine learning systems: An empirical investigation.* International Conference on Agile Software Development (XP).

Talent & Skills. Digital literacy, AI literacy, technical skills, training plans.

- World Economic Forum (2025). *Future of Jobs Report 2025.* WEF, Geneva. - workforce skills demand projections.
- Green, A. (2024). *Artificial intelligence and the changing demand for skills in the labour market.* OECD Artificial Intelligence Papers, No. 14. OECD Publishing, Paris. https://doi.org/10.1787/88684e36-en

Culture & Change Readiness. Leadership cascade, ROI framework, employee sentiment, change plans.

- Westerman, G., Bonnet, D., & McAfee, A. (2014). *Leading Digital: Turning Technology into Business Transformation.* Harvard Business Review Press.
- Tellis, G. J., Prabhu, J. C., & Chandy, R. K. (2009). *Radical innovation across nations: The preeminence of corporate culture.* Journal of Marketing, 73(1), 3–23.

Governance, Ethics & Compliance. AI governance policy, acceptable use, compliance audits, incident response.

- National Institute of Standards and Technology (2023). *AI Risk Management Framework 1.0 (AI RMF 1.0).* NIST AI 100-1.
- ISO/IEC 42001:2023. *Information technology - Artificial intelligence - Management system.* International Organization for Standardization.

Operations & Use Case Portfolio. Use case discovery, prioritization, ROI measurement, business outcomes.

- Software Engineering Institute, Carnegie Mellon (2010). *CMMI for Services, Version 1.3.*
- Forsgren, N., Humble, J., & Kim, G. (2018). *Accelerate: The Science of Lean Software and DevOps.* IT Revolution Press. - DORA metrics on delivery performance.

Model Management & Monitoring. Model versioning, performance monitoring, bias testing, human review.

- Breck, E., Cai, S., Nielsen, E., Salib, M., & Sculley, D. (2017). *The ML test score: A rubric for ML production readiness and technical debt reduction.* IEEE International Conference on Big Data, 1123–1132.
- Mitchell, M., Wu, S., Zaldivar, A., et al. (2019). *Model cards for model reporting.* Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAT*), 220–229.

> **A note on confidence.** Every item was authored against the construct definition, not by lifting wording from these sources. Each item's validation-evidence record labels the alignment as direct adaptation (a close adaptation of a published scale item), construct-aligned, or novel (a new item exploring the same construct). These records are reviewed by a person before they appear in any client report.

7. Limitations and disclosures

- **Self-report bias.** Both the organizational and individual instruments rely mainly on self-report; the individual instrument also includes a small number of scenario and knowledge-check items, and the personal report compares self-ratings with performance on those items when at least four were answered. We provide a perception-versus-reality analysis on Stage 2 and Stage 3 (Division and Enterprise) reports, where consultants run a validation workshop, and a response-pattern check that flags extreme responding, long runs of identical answers, and answers that diverge sharply from those of other respondents in the same assessment. Neither fully eliminates the bias inherent in self-report.
- **The norm group is GCC-skewed.** The reference group consists of everyone who has completed the personal assessment to date, most of them in the UAE and Saudi Arabia. Percentiles are shown only once that group reaches 50 people; as of 17 November 2026 it holds 178 completed personal assessments, so percentiles are shown. The group is not yet split by country or region, so a respondent from outside the Gulf is compared with a mainly Gulf group, and such a comparison may produce misleading percentile claims.
- **The AI landscape moves fast.** The bank version is dated. The active version is recorded on every assessment so reports remain reproducible. We plan to review the bank at least annually and expect to revise a substantial share of items as the technology and norms change.

8. Update cadence

- Question bank: versioned, and each assessment records the version it was created against. We plan to review the bank at least annually. Wording changes are made in a new version; while a version is in use, items may be added or withdrawn and typographical corrections made, and each such change is recorded.
- Methodology brief: re-issued with each bank version. This document is version 1.3 and describes question bank version 1.1.
- Norm groups: recomputed from the full pool of completed personal assessments each time results are produced; percentile comparisons are shown only once the pool reaches 50 respondents.

9. Contact

For methodology questions or research-collaboration inquiries, contact the VIFM consulting team: contact@viftraining.com.

Reliability and Validity Evidence

This consolidates the evidence in two parts: **(A)** established coefficients for the *methods* the Compass is built on, from the published literature; and **(B)** VIFM's own reliability procedures and their current

status. We report (A) as evidence for the method, not as a claim about VIFM's specific item set.

A. Established coefficients for the methods (published literature)

- **Technology-acceptance / readiness constructs** (which the individual factors content-align with) report acceptable to good internal consistency for their source scales in their original studies (UTAUT and UTAUT2: Venkatesh et al., 2003, 2012; Technology Readiness Index 2.0: Parasuraman & Colby, 2015).
- **Organizational maturity** is standards-based rather than coefficient-based: NIST AI RMF 1.0 (2023), ISO/IEC 42001:2023 and CMMI provide the rubric anchors, not criterion correlations.

B. VIFM's own reliability - method, threshold, and current status

- **Internal consistency:** Cronbach's **α** per factor and per pillar; target **α ≥ .70**. Status: accumulating responses via the free Personal Snapshot; to be reported once **N ≥ 200** (individual) and **N ≥ 50** (organizational). Thresholds not yet met.
- **Construct validity:** a confirmatory factor analysis of the four-factor individual model is planned at **N ≥ 200**; until then the four factors are a working model based on content review (disclosed in the appendix of the organizational report, including departmental-plus-personal engagements, and in this brief).
- **Test-retest:** not yet measured. The annual reassessment measures change over a year, not score stability; a short-interval test-retest study is planned.
- **Criterion validity:** not claimed - the Compass is diagnostic and developmental, not a selection instrument.
