COLLECTIVE INTELLIGENCE

A BRAINDANCE COLLECTIVE WHITE PAPER SERIES

ISSUE · NO. V

BDP-6™ · Six Phases, One Gate

Proof Before Production

How the BrainDance Pilot Process™ Turns Agentic AI Experiments into Governed, Validated Deployments

By Tracy L. Williams
The BrainDance Collective, LLC

PREPARED FOR PRIVATE EQUITY SPONSORS, INDEPENDENT SPONSORS, SEARCH FUND BUYERS AND THE EXECUTIVES ACCOUNTABLE FOR AI RESULTS

October 2026

BRAINDANCE COLLECTIVE

Executive Summary

Artificial intelligence is no longer scarce. Proof is. MIT's NANDA initiative found that roughly 95 percent of enterprise generative AI pilots deliver little or no measurable impact on the P&L. S&P Global Market Intelligence reported that 42 percent of companies abandoned most of their AI initiatives in 2025, up from 17 percent a year earlier. Gartner expects more than 40 percent of agentic AI projects to be canceled by the end of 2027 because of escalating costs, unclear business value or inadequate risk controls.

The common thread is not model quality. It is the absence of a disciplined way to decide whether an AI system is ready. Agentic systems are probabilistic by construction: the same request can produce materially different answers across runs, tool calls and context windows. Most organizations respond with demos and enthusiasm, and they discover the variance in production, where it is expensive, visible and hard to instrument after the fact.

This paper introduces the approach behind the BrainDance AI Lab. We cannot make an agent deterministic. We can make the decision about it deterministic. **The BrainDance Pilot Process™ (BDP-6™)** surrounds a probabilistic system with a fixed scope, fixed data, a fixed scoring rubric, a fixed governance gate and a fixed rollout discipline, so that the same evidence always produces the same decision:

- **Six phases, one gate.** Discover and Scope, Mirror and Provision, Data Strategy, Test and Validate, Review and Align, then Roadmap and Deploy. Nothing crosses the Go/No-Go gate on momentum.
- **A signed artifact at every phase.** A scope record, a provisioning manifest, a data lineage record, a readiness report and a signed decision record make the gate an audit, not a vote.
- **One number at the gate.** The BrainDance Readiness Index™ scores Approach, Architecture, Functionality and Outcome into a single weighted composite with explicit thresholds and a floor.
- **Two topologies, one method.** The Lab runs fully off-premises or as an isolated, network-segmented segment inside the client's own environment, with a hybrid bridge between them. The phases and the gate do not change.
- **Six to eight weeks.** A typical off-premises engagement reaches the gate at the end of Week 6 and completes a staged, monitored rollout plan by Week 8.

The AI Pilot Gap

Every portfolio company now has an AI story. Few have an AI system in production that a CFO would defend line by line. The pattern repeats across the research: capability is abundant, adoption is broad, and value is thin. McKinsey's 2026 State of AI survey found that nearly nine in ten organizations use AI regularly, yet only 44 percent are scaling it and 37 percent report any impact on EBIT.

© 2026 The BrainDance Collective, LLC

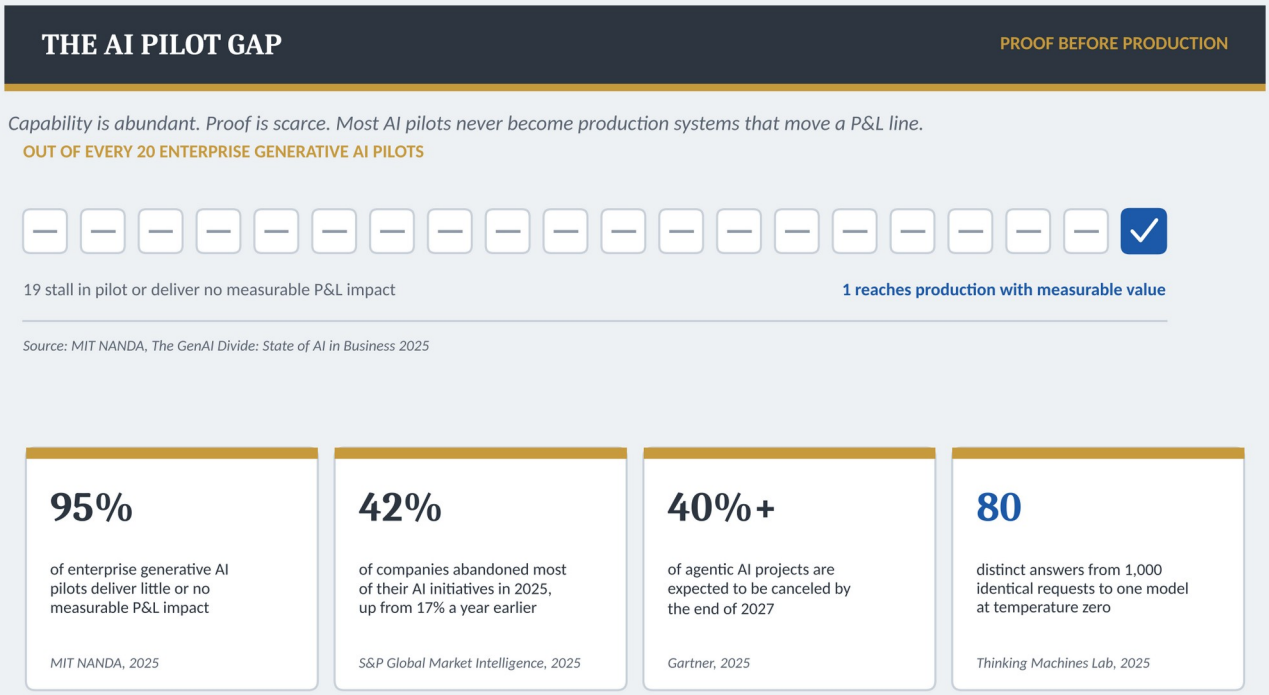


Figure 1: The AI Pilot Gap: broad experimentation, little production value, and a model that does not answer the same way twice.

MIT's researchers were explicit that the failures are not mainly about models or regulation. They are about integration and learning: tools that do not fit real workflows, pilots that are never connected to the systems and data that matter, and organizations that never decide what success would look like. The RAND Corporation reached a similar conclusion, estimating that more than 80 percent of AI projects fail, roughly twice the rate of comparable technology projects, and that misunderstanding the problem to be solved is the most common root cause.

When we look inside stalled pilots, we see the same five failure modes:

Failure mode	What it looks like	What BDP-6™ fixes in place
Undefined success	The pilot is judged by a demo and a feeling	Success criteria written as instrumented KPIs in Phase 01
Wrong data, or no data	Toy data in the pilot; real data blocked by security	A sensitivity-tiered data strategy with a lineage record in Phase 03
Integration surprise	The agent works alone and breaks inside real systems	Architecture tested against mirrored systems in an isolated Lab
Variance discovered late	Inconsistent answers appear after launch	Adversarial and regression suites run before any exposure
No one decides	The pilot drifts into production or quietly dies	One named approver and a signed Go, Conditional-Go or No-Go

© 2026 The BrainDance Collective, LLC

Probabilistic Systems, Deterministic Process

Traditional software is deterministic: the same input produces the same output, so a test that passes today will pass tomorrow. Large language models and the agents built on them are not. In a widely cited 2025 experiment, Thinking Machines Lab sent the same prompt to the same model 1,000 times at temperature zero, the setting meant to remove randomness, and received 80 distinct completions. Add tool calls, retrieval and multi-step plans, and the space of possible behaviors grows quickly.

That variance is not a defect to be engineered away entirely. It is part of what makes these systems useful. The mistake is governing them as if they were ordinary software, with a demo standing in for a test and a launch date standing in for a decision. The answer is to move determinism from the model to the process around it.

The structure that does this is what we call **the Determinism Envelope™**: five fixed rings around the model. Each ring is a fixed element of the process, and each is locked by a signed artifact that an auditor, a board member or a regulator can read. The agent at the center remains probabilistic. Everything that determines whether it reaches production does not.

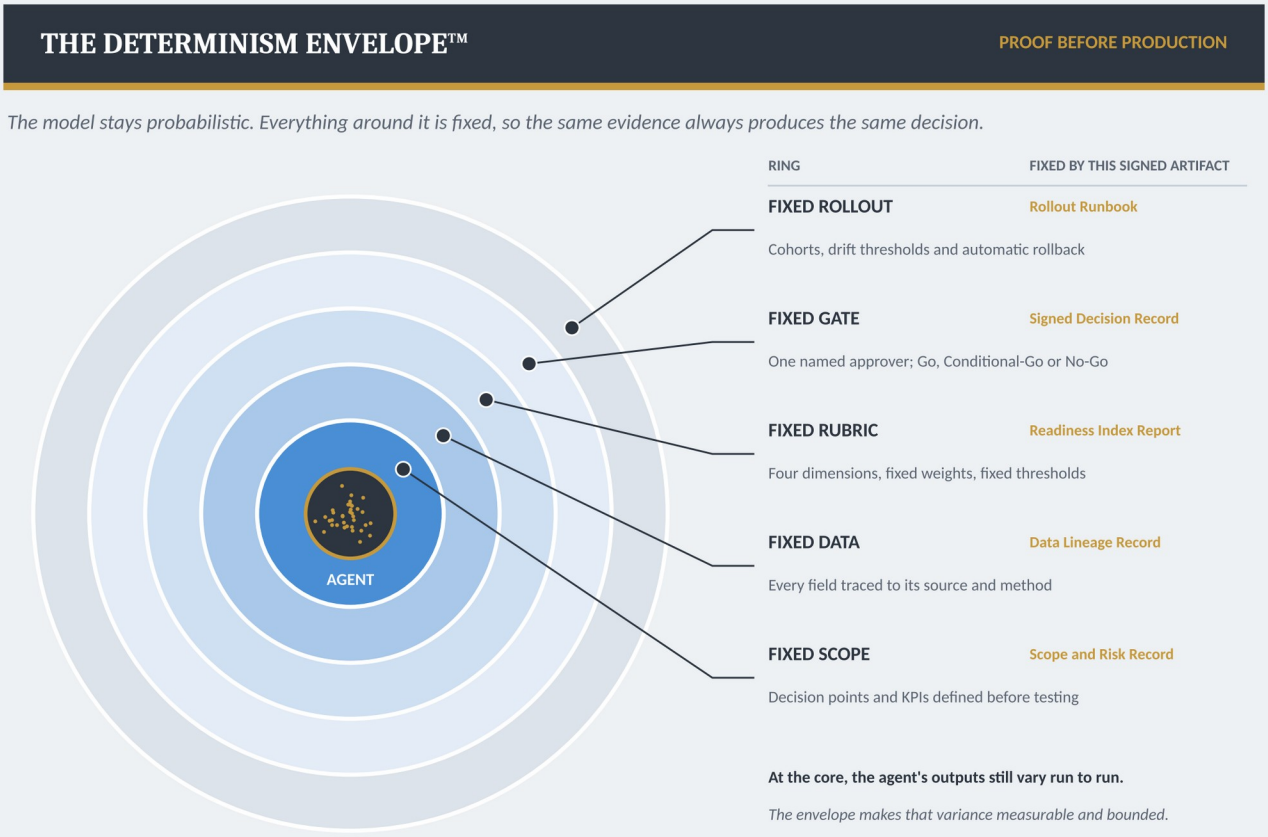


Figure 2: The Determinism Envelope™: five fixed rings, each locked by a signed artifact, surround a probabilistic core.

© 2026 The BrainDance Collective, LLC

THE DETERMINISM PRINCIPLE

You cannot make an agent deterministic. You can make the decision about it deterministic.

Fix the scope, the data, the rubric, the gate and the rollout, and the same evidence will always produce the same decision about whether an AI system is ready.

For a sponsor, the envelope changes the question asked at the investment committee. The question is no longer “Does the demo look good?” It is “What does the evidence package say, who signed the decision, and what will make the system roll itself back?” Those questions have answers, and the answers are the same no matter who asks them.

The BrainDance Pilot Process™

BDP-6™ is a named, repeatable six-phase methodology with one formal governance gate. It is not a sandbox rental and it is not improvised per client. Every phase has an owner, a defined set of activities and a signed artifact that becomes evidence at the gate.

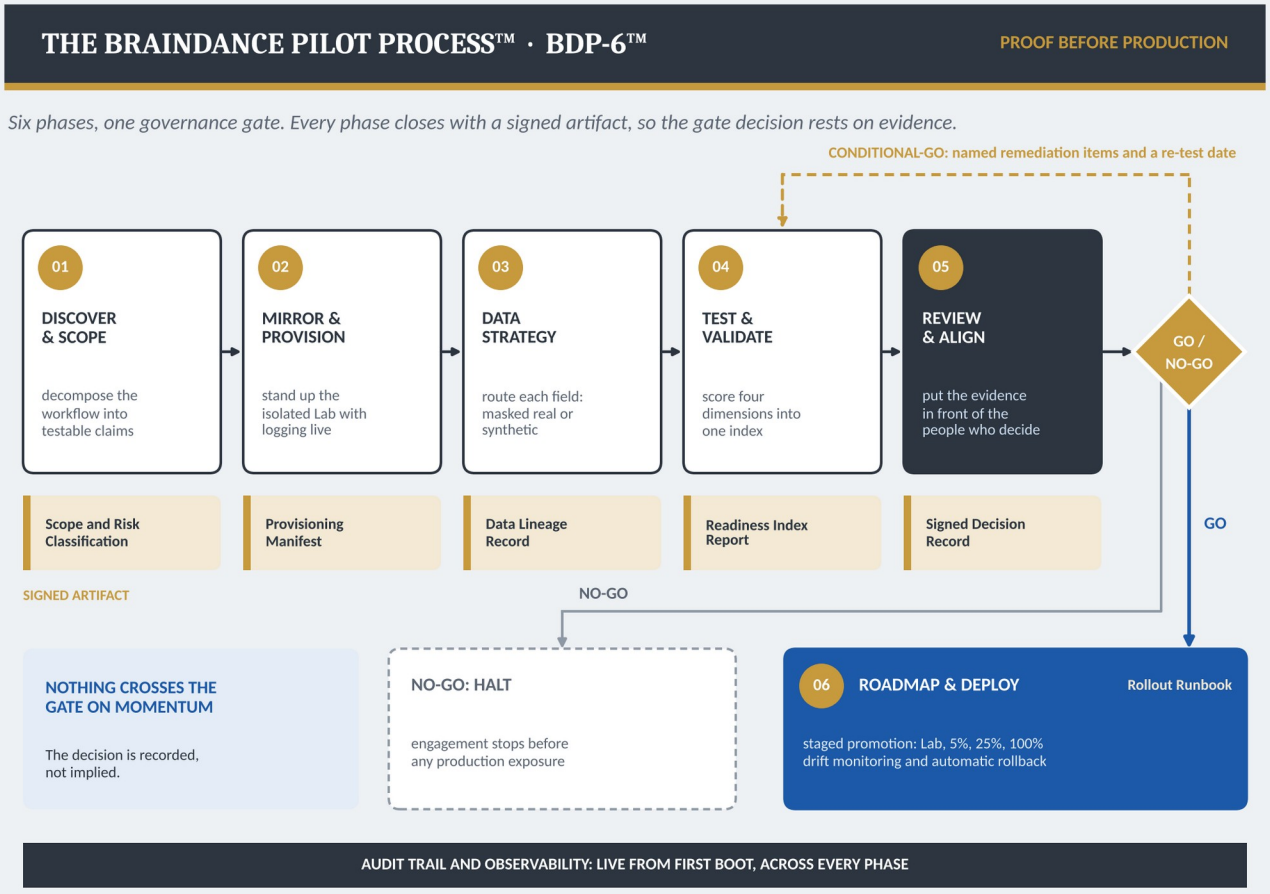


Figure 3: The BrainDance Pilot Process™ (BDP-6™): six phases, a signed artifact at each, and one gate with three possible outcomes.

© 2026 The BrainDance Collective, LLC

The six phases at a glance:

Phase	Purpose	Owner	Signed artifact
01 Discover & Scope	Decompose the target workflow into testable claims and classify data sensitivity	BrainDance PM with the client's technical sponsor	Scope and Risk Classification record
02 Mirror & Provision	Stand up the isolated Lab under the chosen topology, with logging live from first boot	BrainDance PM directs; hosting partner or client IT executes	Provisioning manifest, segmentation verified
03 Data Strategy	Route every data element to governed real data or synthetic data by tier	BrainDance PM; client approves data use	Data Lineage Record
04 Test & Validate	Score four technical dimensions into the Readiness Index	BrainDance PM	Readiness Index report and test logs
05 Review & Align	Put the evidence in front of the people who must trust it and record a decision	BrainDance facilitates; client decides	Signed Decision Record
06 Roadmap & Deploy	Promote through staged exposure with drift monitoring and automatic rollback	BrainDance leads handoff; client AI owner takes it forward	Rollout roadmap and monitoring runbook

Phase 01: Discover and Scope

The work begins by decomposing the target workflow into discrete decision points and tool-call boundaries, then writing success criteria as instrumented KPIs rather than qualitative goals. "Reduce invoice exception handling time by 40 percent at equal or better accuracy" is a testable claim. "Use AI to improve accounts payable" is not. The phase also names the Go/No-Go approver, so that the eventual decision has an owner from day one.

Every data element is classified into one of four sensitivity tiers, and the classification drives both the topology recommendation and the data strategy:

Tier	Typical examples	Default treatment in the Lab	Topology signal
Public	Published prices, product catalogs, public filings	Used as is	Off-premises
Internal	Operating metrics, internal procedures, vendor lists	Masked	Off-premises
Regulated	Personal data, payroll, health or financial records	Tokenized, with synthetic edge cases	On-premises or hybrid
Restricted	Trade secrets, classified or sovereign data	Excluded; synthetic replacement only	On-premises, air-gapped if required

Phase 02: Mirror and Provision

The Lab is provisioned under the topology selected in Phase 01. Source systems, APIs and integration points are mapped and mirrored. Tenant isolation, scoped identity and role-based access, and network

© 2026 The BrainDance Collective, LLC

segmentation are configured before any test runs, and logging, tracing and audit-trail capture are confirmed live from first boot. The phase closes only when segmentation has been verified, not assumed.

Phase 03: Data Strategy

Most pilots fail the data test in one of two ways: they run on toy data that proves nothing, or they wait months for access to real data that security will never release. Phase 03 resolves the conflict by routing each data element according to its tier. Governed real data is used where it is safe, through masking, tokenization and field-level exclusion. Synthetic data, generated by statistical modeling, rules or generative synthesis, fills every place where real data is unsafe or where rare events and edge cases are too thin to test.

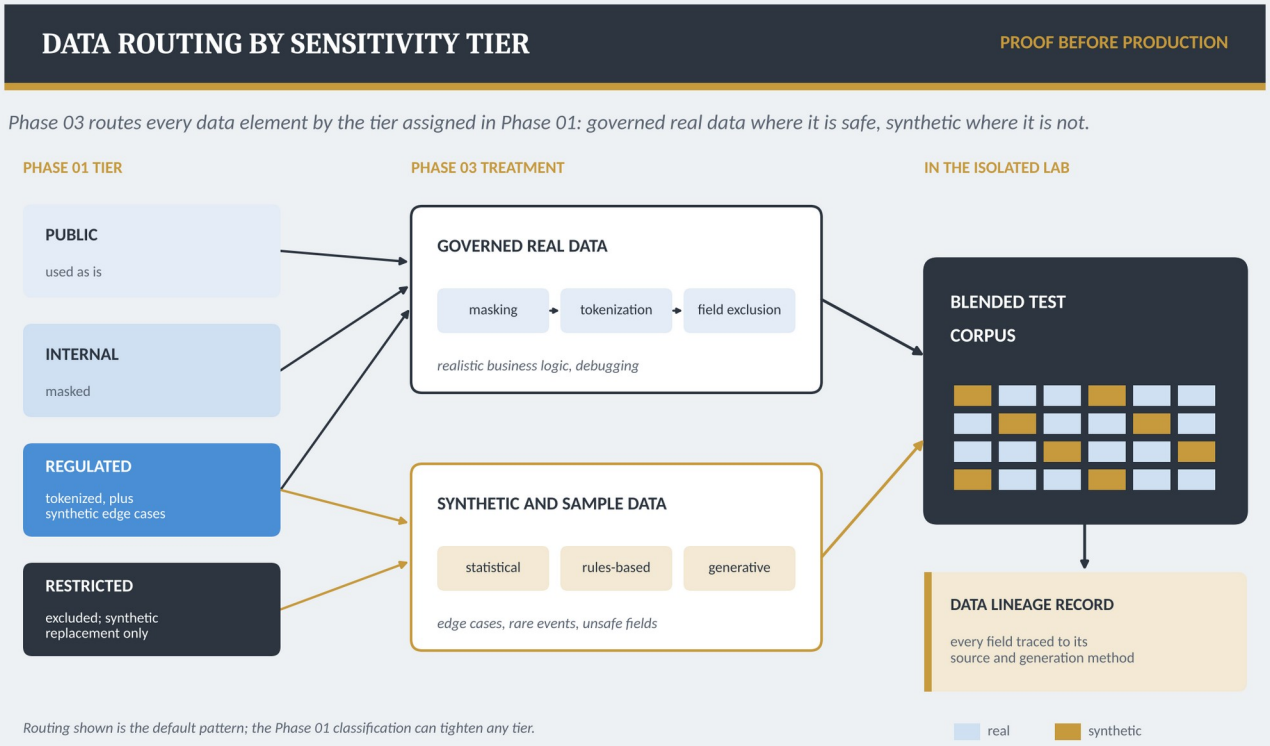


Figure 4: Data routing by sensitivity tier: governed real data and synthetic data combine into one blended test corpus with full lineage.

The output is a blended test corpus and a **Data Lineage Record** that traces every field to its source, real-masked or synthetic, and to the method that produced it. When a regulator or an acquirer asks what the system was tested on, the answer is a document, not a recollection.

Phase 04: Test and Validate

Testing is organized around four dimensions, each with a guiding question and a technical method. The results roll up into a single weighted score, the BrainDance Readiness Index™, described in the next section.

© 2026 The BrainDance Collective, LLC

Dimension	Guiding question	Technical method
Approach	Is this the right method, autonomy level and set of guardrails?	Prompt and model ablation; guardrail red-teaming
Architecture	Does it integrate cleanly with systems and security?	API contract tests; latency and throughput benchmarks
Functionality	Is it correct and consistent, including edge cases?	Adversarial suites against the synthetic corpus; regression runs
Outcome	Does it move the target business KPI?	Shadow-mode comparison against the measured baseline

Phase 05: Review and Align

The evidence package, meaning the Readiness Index plus every underlying test log, is compiled for independent review rather than summarized secondhand. Results are walked through with the department that will actually use the application, against its own workflows. The phase ends in a **Signed Decision Record**: Go, Conditional-Go or No-Go, with a named approver, a date and, if conditional, logged remediation items and a re-test date.

Phase 06: Roadmap and Deploy

A Go decision does not mean a launch. It means promotion through staged exposure with the same discipline as a modern software release pipeline: from the validated Lab to a small canary cohort, then an expanded cohort, then full production. Drift monitoring and an automatic rollback trigger sit at every step. Governance ownership transfers to the client's named AI owner along with a monitoring runbook that states every threshold.

The BrainDance Readiness Index™

A gate needs a number, and the number needs rules that do not bend for a good demo. The Readiness Index combines the four Phase 04 dimensions, each scored from 0 to 100, into one weighted composite that is carried into the Phase 05 review.

DEFINITION

RI BrainDance Readiness Index™

A weighted composite of four validation dimensions, scored from 0 to 100, with explicit decision thresholds and a floor. The default weights reflect where agentic systems most often fail in production:

- **Approach, 20 percent:** method, autonomy level and guardrails
- **Architecture, 25 percent:** integration, security, latency and throughput
- **Functionality, 30 percent:** correctness and consistency, including edge cases
- **Outcome, 25 percent:** measured movement in the target KPI against baseline

Weights are fixed in Phase 01 and recorded in the Scope and Risk record. They can be tuned to the use case, but never after testing begins.

© 2026 The BrainDance Collective, LLC

THE GATE RULES

Go: composite of 85 or above, with every dimension at 75 or above.

Conditional-Go: composite of 70 to 84, with every dimension at 60 or above. Named remediation items and a re-test date are logged.

No-Go: composite below 70, or any single dimension below 60.

The floor exists for the same reason a strong average cannot hide a failing grade in a critical subject. An agent that is brilliant at its task and cannot pass a security review is not 80 percent ready. It is not ready.

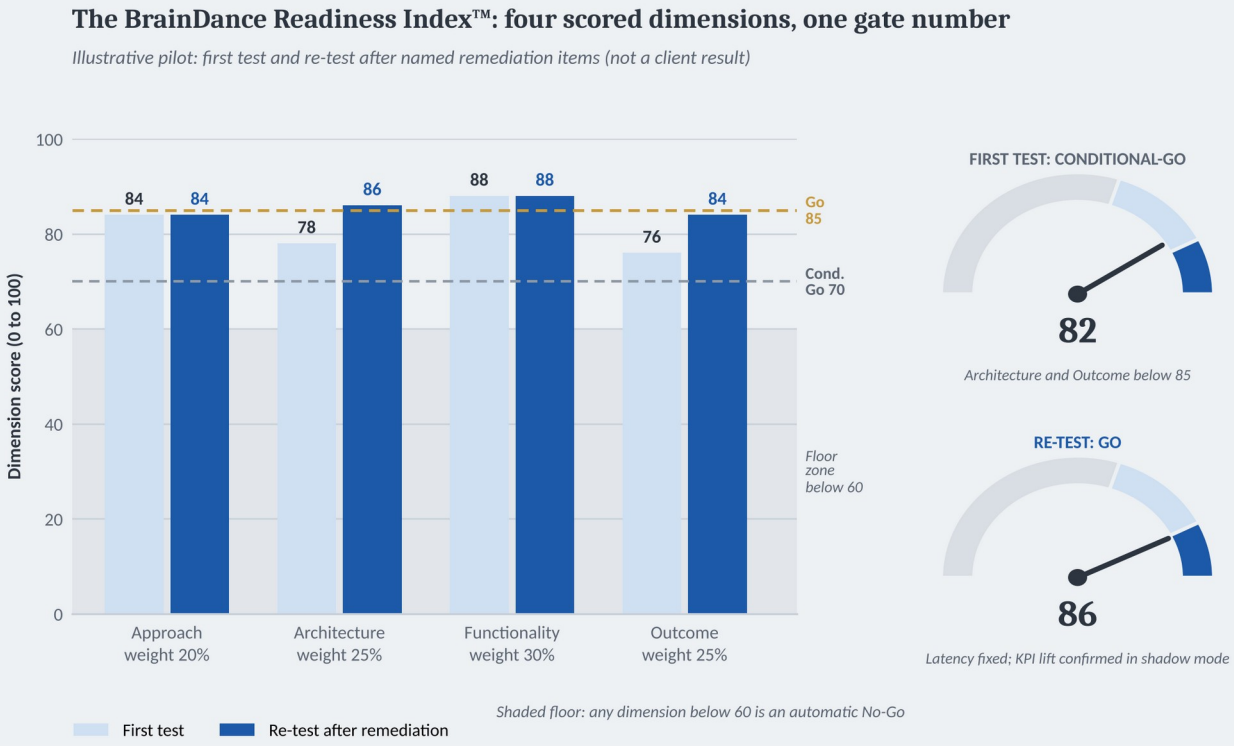


Figure 5: The BrainDance Readiness Index™ for the illustrative pilot: a Conditional-Go at 82 becomes a Go at 86 after named remediation.

© 2026 The BrainDance Collective, LLC

IN PRACTICE

Worked example: an invoice exception agent at a portfolio distributor

A lower middle market distributor wants an agent to triage accounts payable exceptions: price mismatches, missing receipts and duplicate invoices. The scores at the first Phase 04 run:

- **Approach 84, Architecture 78, Functionality 88, Outcome 76**
- **Composite:** 0.20 × 84 + 0.25 × 78 + 0.30 × 88 + 0.25 × 76 = 81.7, reported as 82
- **Decision:** every dimension clears 60, but the composite is below 85 and two dimensions are below 75: **Conditional-Go**
- **Remediation logged:** ERP lookup latency breached the service target under load; the KPI lift in shadow mode had not yet reached the 40 percent target

After a caching fix and a revised exception taxonomy, the re-test scores Architecture 86 and Outcome 84. The composite rises to **85.7, reported as 86**, every dimension is at 84 or above, and the approver signs a **Go**. The agent enters a 5 percent canary cohort the following week.

Illustrative example built to show the arithmetic; it is not a client result.

Readers of Issue No. IV will recognize the logic. The V^2 Equation™ asks whether a transformation's value is validated; the Readiness Index asks whether an AI system has earned the right to be exposed to that value. Once an agent is in production, its business results feed the same V^2 scorecard that governs the rest of the value creation plan.

From Gate to Production: Staged Promotion

Most AI incidents are not failures at launch. They are failures of drift: the inputs change, a vendor updates a model, a seasonal pattern appears that the test corpus never saw. A system that was ready in Week 6 can be wrong in Week 12. Phase 06 is built on that assumption.

Each stage has a defined exposure, a hold period and a reason to exist:

Stage	Exposure	What it proves
Lab validated	0 percent of production traffic	The system meets the gate rules on a blended corpus and mirrored systems
Canary cohort	About 5 percent	Behavior on live inputs matches the Lab, with a small blast radius
Expanded cohort	About 25 percent	Performance holds across more users, edge cases and load
Full production	100 percent	Steady-state monitoring, owned by the client's AI owner

THE ROLLBACK RULE

A drift-threshold breach reverts the cohort to the prior stage automatically, without waiting for a meeting or a manual call. The cohort advances again only after a full hold period inside threshold.

© 2026 The BrainDance Collective, LLC

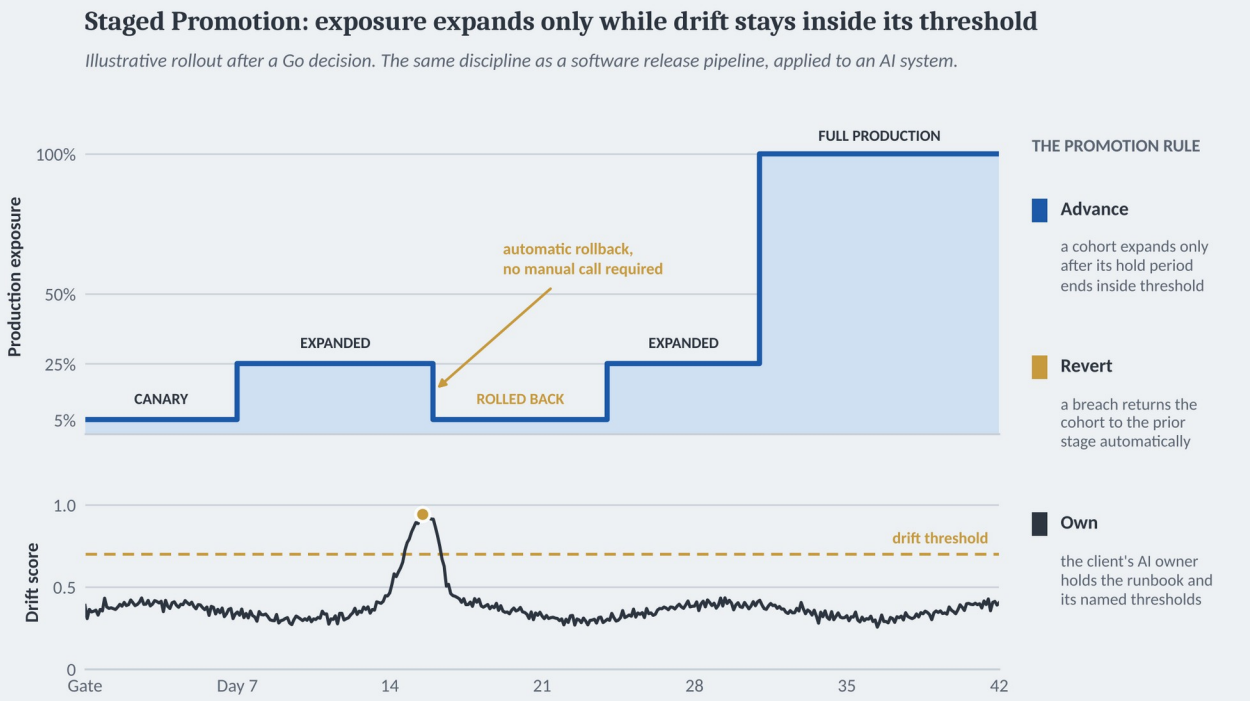


Figure 6: Staged promotion: exposure advances only while drift stays inside its threshold, and a breach reverts the cohort automatically.

One Architecture, Two Topologies

The Lab is not a fixed location. It is a provisioning pattern. Every BDP-6™ engagement runs on the same reference architecture: client systems and source APIs, a data fabric of masked real and synthetic data, isolated Lab compute, an evaluation harness built on the four-dimension rubric, the governance gate, and a promotion pipeline to production. Audit logging and observability span every layer and every phase. Only the physical location of the isolated compute layer changes.

© 2026 The BrainDance Collective, LLC

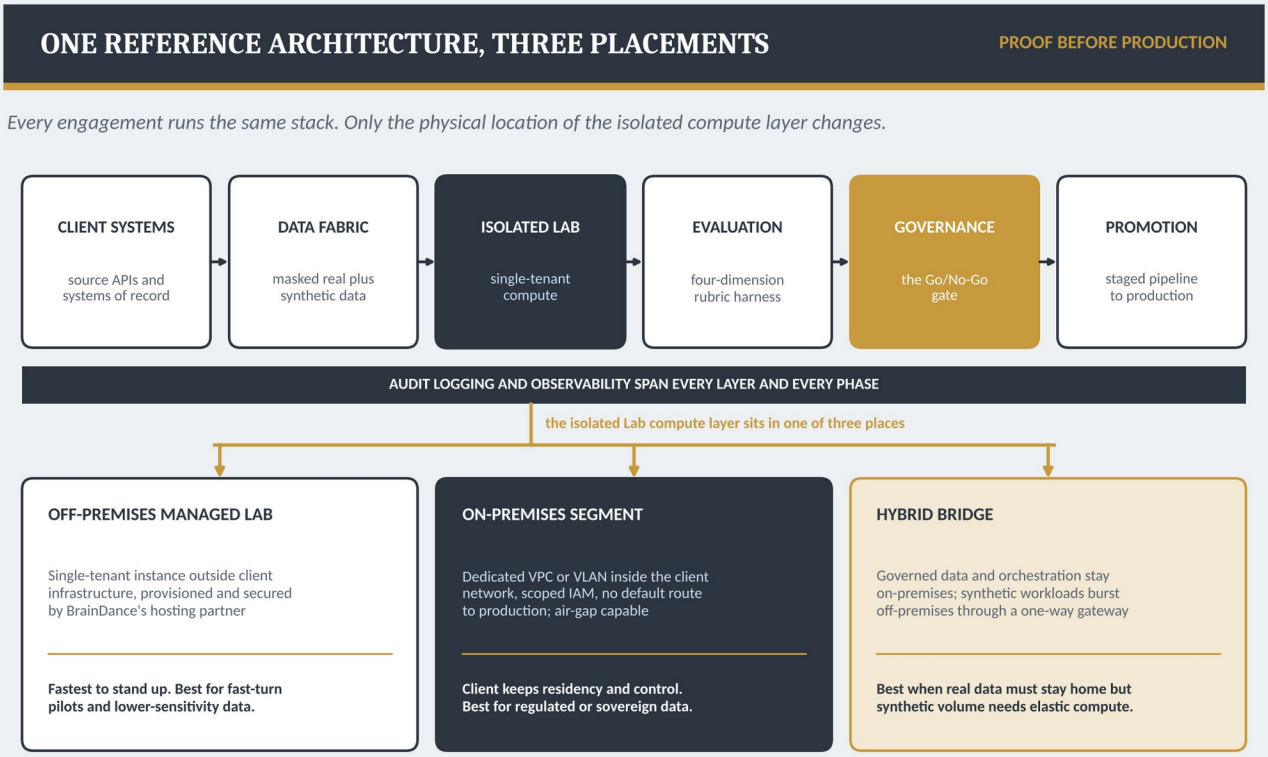


Figure 7: One reference architecture with three placements for the isolated compute layer: off-premises, on-premises or a hybrid bridge.

Choosing the topology follows from the Phase 01 classification, not from preference:

	Off-premises managed Lab	On-premises provisioned segment
Where it runs	Single-tenant instance outside client infrastructure	Dedicated VPC or VLAN inside the client's own network
Who provisions it	BrainDance's hosting partner, directed by BrainDance	Client IT, directed by BrainDance and jointly verified
Isolation	Isolated network path; no default route to client systems	Scoped identity and access; no default route to production; air-gap capable
Data control	Governed data secured in transit by the hosting partner	Client retains full data residency and infrastructure control
Best fit	Fast-turn pilots, lower-sensitivity data, teams without spare infrastructure	Regulated industries, residency mandates, sovereign or classified data
Schedule effect	Fastest to stand up	Typically adds one to two weeks to Phase 02

A **Hybrid Bridge** configuration is also available. Governed data and orchestration stay on-premises, while synthetic-data workloads burst to the off-premises Lab through a controlled, one-way gateway. It is the right answer when the real data must stay home but synthetic volume needs elastic compute.

© 2026 The BrainDance Collective, LLC

Six to Eight Weeks to a Signed Decision

A typical off-premises engagement runs six to eight weeks. The Go/No-Go gate falls at the end of Week 6, and nothing enters Roadmap and Deploy without a signed Decision Record. On-premises engagements typically add one to two weeks to Phase 02 for client IT coordination and segmentation sign-off; the remaining phases hold their duration.

A typical off-premises engagement runs six to eight weeks

Illustrative schedule. Nothing enters Roadmap & Deploy without a signed Decision Record.

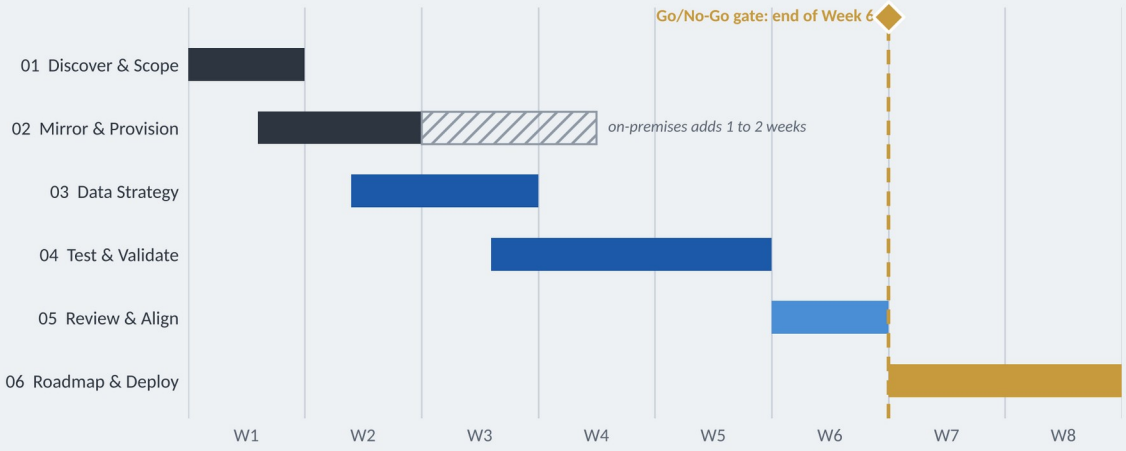


Figure 8: An illustrative BDP-6™ schedule: the gate at the end of Week 6, with on-premises provisioning shown as an extension.

Governance and Ownership

Clear ownership is the difference between a pilot and a governed program. Infrastructure responsibilities shift with the topology. Decision rights do not.

Phase	BrainDance PM	Client stakeholders	Hosting partner (off-premises)	Client IT (on-premises)
01 Discover & Scope	Leads	Consulted	Not involved	Consulted
02 Mirror & Provision	Directs	Informed	Provisions and secures	Provisions segment
03 Data Strategy	Leads	Approves data use	Secures data in transit	Enforces data residency
04 Test & Validate	Leads	Informed	Hosts environment	Hosts segment
05 Review & Align	Facilitates	Decides Go or No-Go	Not involved	Not involved
06 Roadmap & Deploy	Leads handoff	Owns going forward	Not involved	Owns monitoring infrastructure

© 2026 The BrainDance Collective, LLC

> **THE INDEPENDENCE RULE**
>
> In both topologies, BrainDance retains sole authorship of the Go/No-Go recommendation, and the client's named approver alone makes the decision. Infrastructure ownership never confers decision rights over the gate.

This separation matters most to owners of capital. The party that hosts the environment, sells the model or built the agent should never be the party that certifies it is ready. Independence at the gate is what allows a sponsor to rely on the evidence package in a board meeting, a lender discussion or a sale process.

What This Means for Sponsors

For private equity firms, independent sponsors and search fund buyers, AI now shows up in three places: in the value creation plan, in the diligence of targets that claim AI capability, and in the risk register. BDP-6™ gives each of them the same thing, an evidence-based decision with a named owner.

Situation	How BDP-6™ is used	What the sponsor receives
AI in the value creation plan	Validate the highest-value use case before funding a rollout across the company	A Readiness Index, a signed decision and a staged rollout plan tied to a KPI
Pre-close AI diligence	Test a target's AI claims in an isolated Lab against synthetic or masked data	Independent evidence of what the target's AI actually does
Stalled or sprawling pilots	Run each pilot through the gate and stop the ones that cannot pass	A pruned portfolio and capital redirected to proven use cases
Regulated portfolio companies	On-premises or air-gapped Lab with full lineage and audit trail	A defensible record for regulators, lenders and acquirers
Add-on integration	Validate agents against two companies' systems before consolidation	Integration risk surfaced in the Lab, not in the first close

The pilot process is also where our AI work connects to the rest of this series. Issue No. I argued for selecting AI tools strategically rather than by enthusiasm. Issue No. III described Collective Execution™, our executive-led value realization system. Issue No. IV introduced the forward deployed executive and the V^2 Equation™. BDP-6™ is how an AI use case earns its place in that system: scoped like a strategy, tested like a product, decided like an investment and governed like an operation.

Conclusion

The organizations that win with AI will not be the ones that run the most pilots. They will be the ones that can tell, quickly and defensibly, which pilots deserve to become production systems and which do not. That is a process problem before it is a technology problem, and process is something leaders can control.

The imperative for sponsors and boards is practical: define success before testing begins, test on data you can trace, score readiness against rules that do not bend, make the decision in writing, and let the system roll itself back when the world changes.

© 2026 The BrainDance Collective, LLC

Scope it. Prove it. Gate it. Then deploy it.

If you have an AI use case in your portfolio that deserves a real answer, let's connect and schedule a Lab scoping call: letstango@braindancecollective.com.

Appendix: Glossary and Sources

Term	Meaning
BrainDance Pilot Process™ (BDP-6™)	BrainDance's six-phase, single-gate methodology for validating agentic AI before production
Determinism Envelope™	The five fixed rings of scope, data, rubric, gate and rollout that surround a probabilistic AI system
BrainDance Readiness Index™	A weighted 0 to 100 composite of Approach, Architecture, Functionality and Outcome, carried to the gate
Go / Conditional-Go / No-Go	The three recorded outcomes of the governance gate, each with a named approver and date
Floor rule	Any single dimension below 60 is an automatic No-Go, regardless of the composite
Data Lineage Record	A record tracing every field in the test corpus to its source and generation method
Blended test corpus	Governed real data and synthetic data combined inside the isolated Lab
Off-premises managed Lab	A single-tenant Lab instance outside client infrastructure, run by BrainDance's hosting partner
On-premises provisioned segment	An isolated, network-segmented Lab inside the client's environment; air-gap capable
Hybrid Bridge	On-premises governed data with synthetic workloads bursting off-premises through a one-way gateway
Staged promotion	Lab, canary, expanded and full production exposure, with drift monitoring and automatic rollback
V^2 Equation™	$cD + fdE = V^2$: the series' measure of validated value (see Issue No. IV)

Sources

- Aditya Challapally et al., MIT NANDA, *The GenAI Divide: State of AI in Business 2025*, July 2025, as reported by Sheryl Estrada, *Fortune*, August 18, 2025: about 95 percent of enterprise generative AI pilots deliver little or no measurable P&L impact.
- S&P Global Market Intelligence, *Voice of the Enterprise: AI and Machine Learning, Use Cases 2025*, survey of more than 1,000 enterprises: 42 percent of companies abandoned most AI initiatives in 2025, up from 17 percent in 2024; 46 percent of proofs of concept scrapped on average.
- Gartner, press release, June 25, 2025: over 40 percent of agentic AI projects will be canceled by the end of 2027 due to escalating costs, unclear business value or inadequate risk controls.

© 2026 The BrainDance Collective, LLC

- Horace He and Thinking Machines Lab, “Defeating Nondeterminism in LLM Inference,” September 2025: 1,000 identical temperature-zero requests produced 80 unique completions under default inference settings.
- James Ryseff, Brandon De Bruhl and Sydne Newberry, RAND Corporation, *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, August 2024: more than 80 percent of AI projects fail, about twice the rate of non-AI IT projects.
- McKinsey & Company, “The state of AI,” August 2026: 44 percent of organizations scaling AI; 37 percent reporting EBIT impact.
- The BrainDance Collective, *Collective Intelligence* Issues No. I, III and IV, 2026; BrainDance AI Lab, *The BrainDance Pilot Process: BDP-6™ Methodology*, technical concept deck, 2026.

About the Author

Tracy L. Williams is Principal and Lead Strategist at The BrainDance Collective, a Wharton-educated operator whose career runs through finance, technology, operations, government contracting, entrepreneurship, consulting and private equity. She serves clients as a fractional chief financial and strategy officer and PMO lead, and leads the BrainDance AI Lab, the firm's practice for validating agentic AI before it reaches production. Her view of AI is an operator's view: a model is only as valuable as the decision it supports, and a decision is only as good as the evidence and accountability behind it. The BrainDance Pilot Process is the discipline she brings to that question, so that sponsors and executives can say yes to AI with the same confidence they bring to any other investment.

About The BrainDance Collective

The BrainDance Collective is a collection of experienced executives who design and execute solutions to complex problems, achieving the primary mission of a private company or agency shoulder to shoulder with its team, from within the organization. Guided by three commitments, Integrity Over Everything, Partners Not Vendors, and People Are the Point, BrainDance brings 150+ years of combined executive experience across Fortune 500 brands to bear on strategy, technology adoption, and operating performance.

Think Big. Solve Together.

(678) 462-9926 · letstango@braindancecollective.com

© 2026 The BrainDance Collective, LLC