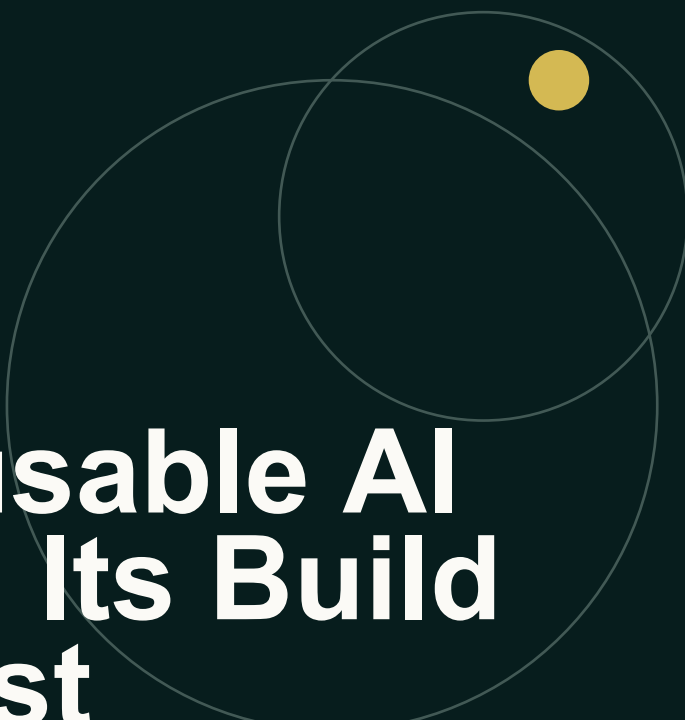


THE AI INSTITUTE / OPERATING INTELLIGENCE

EXECUTIVE DECISION BRIEF · 06 OCT / 2026



Make Reusable AI Work Earn Its Build Cost

A cheaper repeat run is not enough. Test whether useful reuse will repay development, checking and maintenance.

Boards, C-suite and senior management

Research period: September 2026 – October 2026

First published: 6 October 2026

Current to 6 October 2026

PUBLICATION RECORD

An Institute thesis and decision framework

Fund a reusable AI routine when credible savings across valid future uses can repay its full build and upkeep cost before the workflow changes.

WHAT LEADERS SHOULD TAKE FROM THIS PAPER

- | | |
|----|---|
| 01 | Count repeatable work, not total agent activity. |
| 02 | Include failed builds, checks and exception handling in the cost. |
| 03 | Test payback against a realistic useful life. |
| 04 | Keep a fallback and a named maintenance owner. |

Research method

Original Institute management analysis informed by evidence from two new research papers and an established cross-market survey. PACE supplies the research direction; the four-number approval test, enterprise cost categories and currency illustration are Institute analysis. No quoted source wording or private anecdotes are used.

First published 6 October 2026 · Updated 6 October 2026 · Research period September 2026 – October 2026 · Version 1.0 · Suggested citation: The AI Institute, *Make Reusable AI Work Earn Its Build Cost* (2026).

CONTENTS

Inside this paper

01	Do not pay for the same discovery twice Markets & Value
02	A lower running cost is only half the case The new signal
03	Count what must be recovered The economic test
04	Reuse the procedure, not an outdated assumption Where the case changes
05	Fund a bounded test, then make the choice The next approval
M	Method and limitations
R	References

How to use this brief

Start with the decision and use the framework to challenge the current plan. Figures retain their original populations and definitions. Institute analysis is labelled, and limitations are recorded before the references.

Do not pay for the same discovery twice

An AI agent can work out how to complete a task each time it receives it. That flexibility is useful when the work varies. For a stable, frequent task, it also raises a commercial question: should the organisation keep paying for that flexibility, or invest once in a repeatable procedure? The cheaper-looking option can still cost more if it is expensive to build or rarely reused.

Before funding the next automation, ask for a payback case tied to a particular routine. Identify the work it will handle, how often qualifying cases arrive and how long the procedure is likely to remain valid. Total agent activity is the wrong denominator. Only work the proposed routine can complete to the required standard contributes to its return.

The Institute's recommendation is to make this a small operating investment decision, owned jointly by the process sponsor and finance. The aim is not to replace every agent with a script. It is to stop treating either unlimited improvisation or automatic reuse as the default.

A lower running cost is only half the case

A preprint from City University of Hong Kong, submitted on 2 October and included in arXiv's 5 October release, studies when repeated computer-use tasks should become programs. Its PACE system accounts for unsuccessful build attempts as well as successful ones, and considers whether future use can repay that cost. ¹

The evidence is narrow: seven task families, three model profiles and simulations using measured costs. Its live study does not evaluate the full decision algorithm. Token savings are not a measure of enterprise profit, and the cost constraint does not guarantee task success. The useful signal is the question being measured, not a universal payback number. ¹

A separate JPMorgan Chase research paper illustrates another option: route suitable financial-chat cases through rules and reserve a language model for harder cases. The authors describe institution-specific data and adaptation requirements; the confidential source conversations prevent direct external replication. This supports testing a mixed approach, not copying its reported savings into another business case. ²

The management gap is broader than either experiment. KPMG's Q3 survey covered 2,131 senior leaders in 20 countries, territories and jurisdictions, with fieldwork from 23 July to 26 August. Its release says only 12% consistently assess AI's value against cost. This is reported practice in that sample, not audited returns or proof of causation. ³

Count what must be recovered

Start with the current cost of completing an eligible case acceptably. Include model use, staff checking, retries and correction. Compare it with the proposed routine using the same case mix and acceptance standard. A faster output that transfers more checking to another team has not necessarily lowered the organisation’s cost.

Next, estimate the investment required to make the routine usable. Include development attempts that fail, integration, test cases, security review and the time of the people who understand the process. These costs do not disappear because an AI system writes much of the code. Record them separately from ongoing costs so the business case can be challenged.

An illustrative calculation makes the test concrete. Suppose a routine costs 1,000 currency units to build and validate, then saves two units on each acceptable case after recurring checks and exceptions. It needs 500 such cases to recover that initial cost, before any additional maintenance. These are invented numbers for explanation, not research findings or a suggested target.

Now shorten the expected useful life. If only 300 qualifying cases are likely before the underlying process changes, lower execution cost alone cannot repay that investment. Conversely, reliable high-volume work may repay a more expensive build quickly. The decision turns on valid reuse and net savings, not the number of automations delivered.

EXHIBIT / FOUR NUMBERS TO BRING TO APPROVAL

01 Eligible use Cases the routine can complete acceptably. Institute decision framework	02 Net saving Cost avoided after checks and exceptions. Institute decision framework	03 Full investment Build, failed attempts and validation. Institute decision framework	04 Useful life Time before change requires rework. Institute decision framework
---	--	--	---

Reuse the procedure, not an outdated assumption

A working routine needs a boundary. Specify the applications, inputs, permissions and business rules under which it has been tested. A change in any of these may require revalidation. Assign an owner who can notice that change, suspend the routine and restore an approved alternative. A low-cost process that quietly produces the wrong result is not an efficient process.

Do not pool international volume until the work is genuinely comparable. An English-language routine tested in one North American business unit has not thereby passed on French documents, Arabic interfaces or a Latin American subsidiary's records. Across European, Asia-Pacific, African and Middle Eastern operations, local formats, language and systems can change how much work is eligible. Test those differences rather than assume a regional adoption rate.

Legal and institutional differences belong in the eligibility decision too. Have the relevant teams establish whether the proposed data use, permissions and retention are allowed for that workflow. This brief makes no claim that approval in one jurisdiction travels to another. Restricted data access or unavailable integration may change the available design before cost comparisons even begin.

Buying an existing integration may beat building a routine around a changing screen. A managed service may offer support that an internal script lacks, while also adding fees or dependency. Compare the actual alternatives, including keeping a human step. Where work is infrequent, highly variable or consequential, retaining flexibility can be the economically sensible choice.

Fund a bounded test, then make the choice

Choose one frequent routine with a clear definition of acceptable completion. Observe a representative set of cases before selecting only the easiest examples. Keep a record of the cases excluded, and why. That record defines the market for the proposed automation inside the organisation.

Set a capped development budget and a review point. Measure the complete cost of the current method and the candidate routine, including maintenance observed during the test. Separate measured use from forecast demand. Finance should be able to see which assumptions would have to change for the investment to stop paying back.

Then choose among building, buying, continuing with the agent or stopping the task altogether. Require a named operating owner and an approved fallback before expanding use. Revisit the decision after a material process or system change. The investment has earned its place when useful repetition repays the work required to make it dependable.

RESEARCH RECORD

Method and limitations

Method

Original Institute management analysis informed by evidence from two new research papers and an established cross-market survey. PACE supplies the research direction; the four-number approval test, enterprise cost categories and currency illustration are Institute analysis. No quoted source wording or private anecdotes are used.

Limitations

PACE is a preprint, not a validated enterprise ROI model. FinDialogLens reports EMNLP 2026 Industry Track acceptance on arXiv; proceedings were not independently checked, and its confidential corpus limits replication. KPMG findings are self-reported survey results. Regional examples identify validation questions, not measured differences in national performance. This is general management analysis, not legal or investment advice.



Sources and limitations

Source facts link to the original publication. Institute analysis reconciles definitions and turns the material into a management proposition. Frameworks are decision aids, not claims of universal causality. Read every important statistic with its population, date, definition and caveat.

Research cutoff: 6 October 2026. Review cycle: six months, or earlier after a material change in the underlying facts, standards or policy.

REFERENCES

Evidence behind the thesis

01 Ming and colleagues: When to Compile a Computer-Use Agent?

v1 submitted 2 October 2026; new 5 October batch. City University of Hong Kong preprint. HTML measurement, cost model and limitations inspected; seven task families and three model profiles. No enterprise savings claim.

02 Fu, Ni and Madahian: FinDialogLens

v1 submitted 1 October 2026; new 5 October batch. JPMorgan Chase authors; arXiv reports EMNLP 2026 Industry Track acceptance. HTML and limitations inspected; confidential financial-dialogue corpus, no universal savings claim.

03 KPMG: Global AI Pulse Q3 2026

24 September 2026 release; survey of 2,131 senior leaders in 20 countries, territories and jurisdictions, fielded 23 July–26 August. Reported practice, not an audited causal study.

THE AI INSTITUTE

Make the decision.
Then measure what changes.

Research · Readiness · Capability

theai.institute