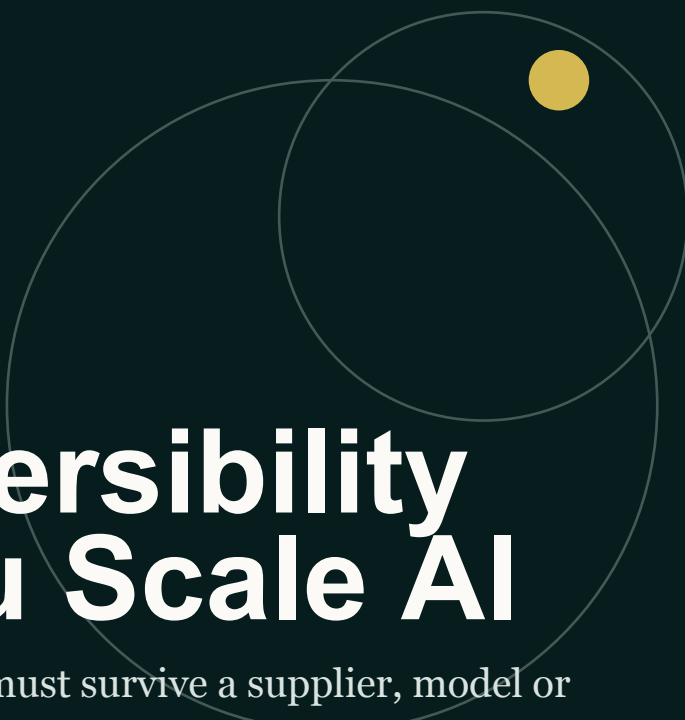


THE AI INSTITUTE / OPERATING INTELLIGENCE

FLAGSHIP EVIDENCE PAPER · 05 / 2026



Price Reversibility Before You Scale AI

A board framework for deciding what must survive a supplier, model or contract change.

Boards, C-suite and senior management

Research period: January 2013 – October 2026

First published: 6 October 2026

Current to 6 October 2026

PUBLICATION RECORD

An Institute thesis and decision framework

The board should define the maximum service, quality, data and knowledge loss it will accept when an AI dependency changes, then fund and rehearse the assets required to stay within that boundary. Reversibility is not the absence of commitment. It is a priced option to change course without losing control of the business outcome.

WHAT LEADERS SHOULD TAKE FROM THIS PAPER

- 01

Set an acceptable reversal loss for each consequential AI-enabled service before approving scale.
- 02

Fund five recoverable assets: service, state, acceptance evidence, operating knowledge and decision authority.
- 03

Treat a second supplier, an export button or a contract clause as inputs to continuity—not proof of it.
- 04

Accept provider-specific capability when its value exceeds the option cost and the dependency is visible, bounded and exercised.
- 05

Match the recovery design to local law, language, infrastructure and supplier availability while keeping the board's loss boundary explicit.

Research method

This flagship evidence paper synthesises Institute daily research briefs published from 2 September to 6 October 2026 and refreshes the recurring dependency, supplier-loss, context, partnership-learning and switching signals against primary legislation, official regulatory and government reports, multilateral evidence, one peer-reviewed systems review and a current technical preprint. The topic was scored against executive decision relevance, evidence strength, non-duplication, cross-market applicability, Institute fit and multi-channel potential. Claims preserve source population, sector, jurisdiction, date and legal status. The acceptable-reversal-loss account, five recoverable assets and three capital gates are Institute analysis proposed for testing; they are not findings attributed to the cited institutions.

First published 6 October 2026 · Updated 6 October 2026 · Research period January 2013 – October 2026 · Version 1.0 · Suggested citation: The AI Institute, *Price Reversibility Before You Scale AI* (2026).

CONTENTS

Inside this paper

01	Decide how much of the operating model may become difficult to reverse The strategic choice
02	AI dependence accumulates across a chain, not one contract Why the decision has moved to the board
03	An export right is not a recovered service Do not confuse the mechanism with the outcome
04	Fund five assets that let the enterprise change course The recoverable-asset portfolio
05	Price the loss boundary beside the benefit case A board measure, not a technical score
06	Exercise one real change of course before the dependency becomes critical Evidence before assurance
07	The loss boundary travels; the substitute market does not One decision, different practical options
08	Do not pay for reversibility that destroys the value you are buying Contrary evidence and limits
09	Name the loss owner, exercise the option and gate the next capital release The 90-day executive agenda
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.

Decide how much of the operating model may become difficult to reverse

A strong AI service can deserve a deep commitment. Specialist models, integrated workflow tools and managed infrastructure may outperform a lowest-common-denominator design. The board's job is not to prevent dependence. It is to decide how much dependence the enterprise can afford before that commitment becomes hard to change.

The relevant loss is wider than an outage. A change in supplier, model, price, access, law or service quality can interrupt work, lower performance, strand data, erase institutional learning or narrow future negotiating power. Those effects can arrive together. A technically successful export may still leave the business unable to reproduce an acceptable result.

Set the boundary before scale. For each consequential service, define the longest tolerable interruption, the minimum safe level of service, the largest acceptable quality drop, the records and rights that must remain available, the knowledge that must stay with the enterprise and the commitments that can be paused. This paper calls that boundary the acceptable reversal loss.

Reversibility is not the promise that nothing will change. It is evidence that management can change course within an agreed loss boundary. The board can then compare the cost of preserving that option with the value of deeper integration.

The decision is not 'Can we leave?' It is 'What must still work, remain known and stay under our authority if we change course?'

AI dependence accumulates across a chain, not one contract

The Bank for International Settlements describes an AI supply chain with five linked layers: hardware, cloud infrastructure, training data, foundation models and applications. High fixed costs, economies of scale and network effects shape several layers, while large technology firms increasingly operate across more than one. This is a market-structure analysis, not a forecast that every buyer will be locked in. It shows why a dependency recorded only at the application contract misses important upstream exposure.¹

The Financial Stability Board's 2025 monitoring report identifies third-party dependencies and concentration among the vulnerabilities authorities need to track. Its financial-sector case study focuses on criticality, concentration and substitutability and also highlights data gaps. The finding is sectoral and supervisory; it does not measure the failure probability of an enterprise AI service.²

A 2025 US Federal Trade Commission staff study of three large cloud-provider and AI-developer partnerships found contractual and technical terms that may increase switching costs, restrict multi-cloud use or link access to compute, models and sensitive business information. The study concerns specific partnerships and competition questions, not the terms offered to every enterprise customer.³

These layers can reinforce one another. The same supplier group may influence the model, infrastructure, integration path and evaluation tooling. A buyer can therefore have several contracts yet still rely on one operating chain. Conversely, one contract can sit on components that are genuinely replaceable. The board needs a dependency map built around the business service and its upstream critical path, not a count of vendors.

EXHIBIT / FIVE LAYERS CAN CREATE ONE BUSINESS DEPENDENCY

HARDWARE Compute supply Accelerators, availability, architecture and location. BIS AI supply-chain frame	CLOUD Operating capacity Hosting, networking, identity, storage and managed services. BIS and FSB	DATA Context and state Authoritative sources, permissions, history and retrieval structures. BIS and World Bank	MODEL Behaviour and capability Weights or service, versions, policies, evaluation and change cadence. BIS and FTC
WORKFLOW Business consequence Tools, approvals, staff practice, customer promise and accepted result. Institute synthesis			

An export right is not a recovered service

The European Union's Data Act has applied since 12 September 2025. Its cloud-switching provisions require providers of data-processing services to remove specified contractual, commercial, technical and organisational obstacles, set contractual switching terms and support the portability of exportable data and digital assets. The rules include scope, technical and trade-secret limits; they do not guarantee that a buyer can reproduce every proprietary service or business outcome elsewhere.⁴

This distinction is material for AI. A portable file may not include the provider's model behaviour, orchestration, safety settings, evaluation history, hidden dependencies or operational expertise. Even the same model and prompt can vary when the numerical execution path changes. A September 2026 preprint demonstrated cross-architecture output differences and a proposed deterministic kernel approach across three NVIDIA GPU families. That bounded technical result is not enterprise switching evidence, but it is a warning against treating identical inputs as proof of identical service.¹³

Cloud research has described portability and interoperability problems for more than a decade. A peer-reviewed ACM review of more than 120 papers found continuing challenges in security, monitoring, identity, standards and vendor lock-in across connected clouds. Because the review predates current generative AI services, it supports the persistence of the systems problem rather than any claim about today's model market.¹⁴

The practical unit of recovery is the accepted business service. The substitute does not need to reproduce every response. It needs to complete the bounded job at an agreed quality, cost and risk, with the required records and authority. That is why the enterprise needs its own acceptance pack: representative cases, thresholds, prohibited failures, human review, affected populations and evidence of which commitments can continue.

Portability moves an asset. Reversibility restores an acceptable business capability.

Fund five assets that let the enterprise change course

First, recoverable service. Define a minimum safe workflow that can operate through a supplier or model change. It may be slower, partly manual or narrower than normal. What matters is that the organisation knows which customer, safety, financial or legal commitments continue and which are paused.

Second, recoverable state. Keep authoritative source data, current instructions, identity and permission mappings, decision records and required history in forms the enterprise can retrieve and govern. Do not assume that an exported embedding store, conversation log or application database recreates the meaning, provenance or access controls of the original service.

Third, recoverable acceptance evidence. Retain the task set, baseline, thresholds, failure cases and sign-off record used to approve the service. A replacement can then be judged against the business promise rather than against the incumbent's interface. NIST's voluntary AI Risk Management Framework calls for inventory, third-party contingency, recovery, change management and safe decommissioning; it is guidance, not proof that an organisation has implemented them.⁸

Fourth, recoverable operating knowledge. Preserve workflow logic, integration documentation, prompts where they are material, evaluation decisions, incident lessons, workarounds and the expertise needed to run the fallback. Singapore's voluntary 2026 framework for agentic AI emphasises bounded authority, meaningful human accountability, lifecycle controls and user capability. It applies to organisations using in-house or third-party agents but does not certify any deployment.⁹

Fifth, recoverable authority. Name who can invoke the fallback, suspend commitments, accept reduced service, communicate the change and approve the new operating boundary. A recovery plan that requires fresh committee design during an incident is not an exercised decision right.

These assets do not all have to be owned or operated internally. Shared infrastructure, specialist evaluators, managed continuity services and sector arrangements can be stronger than a thin in-house substitute. The enterprise must be able to inspect the evidence and exercise the consequential decision.

EXHIBIT / THE FIVE RECOVERABLE ASSETS

SERVICE Minimum safe work The reduced workflow and commitments that can continue. COO	STATE Authoritative context Data, instructions, identity, permissions and required records. Data and technology owners	EVIDENCE Independent acceptance Cases, baselines, thresholds, failure tests and sign-off. Risk or assurance owner	KNOWLEDGE Operating learning Workflow logic, integrations, decisions, incidents and skills. Product and workforce owners
AUTHORITY Right to change course Named power to invoke, pause, communicate and approve. Accountable executive			

Price the loss boundary beside the benefit case

A conventional business case prices licence, implementation and expected benefit. It often leaves the cost of changing direction outside the decision. The Institute proposes a reversibility account beside the investment case. This is not an accounting treatment or a prediction of supplier failure. It makes the accepted dependency and the cost of preserving an option visible.

Record five exposures for the consequential workflow: time to a safe minimum service; expected performance or quality loss during the fallback; data, rights and records that could become unavailable; internal knowledge and staff capacity required to recover; and the financial, customer or public commitments that would be delayed or breached. State the evidence behind each estimate and the owner who accepts it.

Then record the option cost: contract and transfer rights, maintained exports, independent tests, staff training, substitute capacity and rehearsal time. Some options will not be worth buying. A low-consequence internal drafting tool may justify a simple export and a restart plan. A service that commits money, affects employment, controls critical infrastructure or determines customer access may warrant a live reduced-service path and frequent exercises.

Recent US public-sector evidence shows why the decision should stay proportional. A 2026 Government Accountability Office report documents multi-vendor procurement challenges. In its response, the Department of Defense argued that provider-specific capabilities can be strategically correct when dependencies are understood and mission-critical workloads have a documented exit strategy; it also warned that multi-vendor SaaS for one function can be infeasible and expensive. The report and response concern US federal cloud procurement, but the counterargument is valuable: reversibility should manage concentration risk, not erase differentiation.⁶

A separate 2025 GAO review found restrictive licensing and lock-in effects among selected agencies, including added fees, repurchase requirements, limits on architecture choice and one case in which a contractor required payment to regain data ownership. The sample is not representative of enterprise prevalence and the agencies could not always quantify the cost. It demonstrates the need to surface the actual rights and cost path before commitment.⁷

The option cost is visible spending. The reversal loss is often hidden until the organisation needs to change course. Put both in the investment case.

Exercise one real change of course before the dependency becomes critical

A second contract is not evidence of continuity. Providers can share infrastructure, identity systems, data pipelines, model families or scarce staff. A substitute can also connect successfully while failing the business acceptance test. The exercise should therefore begin with the workflow and its commitments, not with a technology failover demonstration.

Select one consequential but containable service. Freeze a representative task set and the current acceptance thresholds. Trigger a realistic change: remove the preferred model, revoke one integration, withhold a proprietary context layer, apply a new data-location constraint or assume the supplier is unavailable. Restore the minimum safe service using the documented fallback and current people.

Measure elapsed time, accepted-task completion, quality loss, manual workload, unavailable records, security and rights exceptions, customer impact and the decisions that lacked an owner. Preserve every repair. The result is not a pass/fail badge; it is a priced gap between the approved loss boundary and the present operating capability.

United Kingdom financial regulators began direct oversight of designated critical technology providers in July 2026. Their regime complements rather than replaces regulated firms' responsibility for due diligence, risk management and contingency planning. The scope is a particular financial-sector regime, but the management principle is wider: supplier supervision does not operate the buyer's fallback. ⁵

US Treasury's financial-sector cloud review likewise found that cloud can improve security and resilience while firms still face concentration, incident and migration challenges. It is older sector evidence and should not be used to claim a universal failure rate. It supports an exercise that tests both the benefits of the managed service and the practical limits of substituting it. ¹⁵

EXHIBIT / A REVERSAL REHEARSAL PRODUCES DECISION EVIDENCE

REMOVE	RESTORE	MEASURE	DECIDE
One dependency Model, provider, integration, context layer or permission. Defined scenario	Minimum safe service Current people, current records and the documented fallback. Operating exercise	Business loss Time, quality, workload, records, rights and commitments. Independent observer	Buy, redesign or accept Close the gap, narrow the use or approve the exception. Accountable executive

The loss boundary travels; the substitute market does not

The decision to set an acceptable reversal loss is portable. The recovery design is not. Law determines which records, rights and notices must survive. Language determines whether a substitute can meet the same population. Infrastructure and supplier availability determine whether parallel capacity is realistic. Contract enforceability and institutional capability determine which rights can be exercised under pressure.

The World Bank's 2025 global report describes large differences in connectivity, compute, locally relevant context and skills. High-income markets dominate data-centre capacity, while lower-income settings can face much thinner access and capability. The report concerns national foundations, not enterprise continuity. Its implication is that a company cannot copy a high-resource fallback design into every market and call the risk equivalent.¹⁰

ECLAC's 2025 index compares enabling factors, research and adoption, and governance across 19 Latin American and Caribbean countries and finds pronounced variation. The African Union's 2024 continental strategy sets shared ambitions while recognising the need for national implementation, infrastructure, data and skills. Neither source supplies firm-level exit outcomes. Together they support a simple constraint: regional labels do not reveal the substitute, expertise or rights available to one service in one country.^{11 12}

Where alternatives are scarce, the responsible answer may be a smaller service boundary, a manual minimum service, regional shared capability, a longer recovery objective or a consciously accepted dependency. A formal multi-provider design is not the global default. The board should receive the exception and its consequence, not a false promise of identical resilience everywhere.

Do not pay for reversibility that destroys the value you are buying

This paper does not argue for technology neutrality. Provider-specific models, tools and managed services can create better performance, lower operating burden, faster delivery and stronger security than a portable design. The BIS supply-chain paper and US Treasury review both describe benefits alongside concentration and resilience concerns.^{1 15}

Multi-provider designs can add identity, data-consistency, monitoring, skills and incident complexity. The 2026 GAO record includes a direct warning that multi-vendor SaaS for one business function may be infeasible and expensive. Duplicating two providers can also create a false sense of independence if both depend on the same upstream infrastructure or scarce expertise.⁶

The proportional answer is to preserve the option that matches the consequence. For low-consequence or rapidly changing work, documented reconstruction may be enough. For a high-consequence service, the enterprise may need current data exports, a maintained acceptance pack, retained skills and an exercised reduced-service path. For a uniquely valuable capability, the board may approve a larger reversal loss—provided the exposure, benefit and decision owner are explicit.

Evidence remains limited. There is no representative global dataset measuring enterprise AI exit cost, successful model substitution, quality loss or recovery time. Official frameworks define expectations; they do not prove implementation. Technical research can isolate reproducibility and portability problems without measuring business continuity. The recommendations are therefore a decision framework to test, not an empirical claim that one architecture will minimise cost.

Name the loss owner, exercise the option and gate the next capital release

Name one accountable executive for each consequential AI-enabled service—normally the COO or the business executive who owns the outcome. The board or an authorised committee sets the acceptable reversal loss and approves material exceptions. The CIO or CTO owns the dependency map and recoverable technology assets. The chief risk officer or independent assurance leader challenges the acceptance evidence. Procurement and legal establish transfer, access, assistance and termination rights. The workforce owner ensures the minimum service still has competent people.

In days 1–30, select the three most consequential AI-enabled services. Map the five dependency layers and five recoverable assets. Set the acceptable interruption, minimum safe service, quality floor, required records and rights, critical knowledge and commitments that can be paused. Record every assumption that lacks evidence.

In days 31–60, exercise one service. Remove a material dependency and restore the minimum safe workflow using current people, current exports and current decision rights. Have an independent observer record time, accepted outcomes, manual work, missing assets, security or rights exceptions and customer consequences.

In days 61–90, put the reversibility account beside the benefit case. Decide whether to buy more option capacity, redesign the service, narrow its authority, renegotiate rights or accept the dependency. Release deployment capital only when the initial loss boundary and assets are documented. Release scale capital only when the exercised result sits within the boundary or the board has approved the exception.

Review the boundary when a model, provider, contract, integration, law, language population or critical upstream service changes. The board pack should show the benefit, the option cost, the exercised reversal loss, the accountable owner and every accepted exception.

The board pack: benefit, option cost, exercised reversal loss, accountable owner and every accepted exception.

EXHIBIT / THREE CAPITAL GATES FOR A REVERSIBLE COMMITMENT

BOUND Approve the exposure Minimum service, loss boundary, dependencies, assets and owner. Release limited discovery capital	EXERCISE Prove the option Remove one dependency and measure the real recovery gap. Release deployment capital	REPRICE Accept or close the gap Fund, redesign, narrow, renegotiate or approve the exception. Release scale capital
---	--	--

RESEARCH RECORD

Method and limitations

Method

This flagship evidence paper synthesises Institute daily research briefs published from 2 September to 6 October 2026 and refreshes the recurring dependency, supplier-loss, context, partnership-learning and switching signals against primary legislation, official regulatory and government reports, multilateral evidence, one peer-reviewed systems review and a current technical preprint. The topic was scored against executive decision relevance, evidence strength, non-duplication, cross-market applicability, Institute fit and multi-channel potential. Claims preserve source population, sector, jurisdiction, date and legal status. The acceptable-reversal-loss account, five recoverable assets and three capital gates are Institute analysis proposed for testing; they are not findings attributed to the cited institutions.

Limitations

There is no representative global dataset measuring enterprise AI exit cost, model substitution, quality loss or recovery time. FSB, UK, US Treasury and some GAO evidence is concentrated in finance or public procurement. The FTC report examines three specific AI-cloud partnerships. EU switching rules have defined scope and exceptions and do not guarantee equivalent proprietary services. World Bank, ECLAC and African Union evidence describes national or regional foundations, not firm-level continuity. The current GPU study is a preprint with bounded hardware and workload conditions. The peer-reviewed cloud review predates current generative AI services. Supplier availability, law, language, infrastructure, institutional capacity and contract enforceability can materially alter the option and its cost.



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 Bank for International Settlements, The AI Supply Chain

BIS Paper 154, 18 March 2025. Economic and market-structure analysis of hardware, cloud infrastructure, training data, foundation models and applications; authors' views do not necessarily represent BIS member central banks.

02 Financial Stability Board, Monitoring Adoption of AI and Related Vulnerabilities

10 October 2025. Official financial-sector monitoring report with a case study on critical third-party dependencies, concentration and substitutability; not an enterprise failure-rate estimate.

03 US Federal Trade Commission, AI Partnerships and Investments Staff Report

January 2025 staff report based on three cloud-provider and AI-developer partnerships, with information current through January 2025; specific competition study, not a survey of enterprise contracts.

04 European Union, Data Act

Regulation (EU) 2023/2854, Chapter VI on switching between data-processing services; applies from 12 September 2025 with defined scope, transition and technical provisions.

05 Bank of England, UK Critical Third Parties Regime Begins

10 July 2026 official release. Four designated technology providers entered direct oversight from 13 July; sector-specific regime that does not replace firms' own resilience duties.

06 US GAO, Cloud Computing: Federal Government Needs to Address Procurement Challenges

GAO-26-107530, July 2026. Review of US federal cloud procurement with agency comments on provider-specific capability, exit strategy and the limits of multi-vendor SaaS; not private-sector prevalence evidence.

07 US GAO, Selected Agencies Need to Implement Updated Guidance for Managing Restrictive Licenses

GAO-25-107114, December 2024. Selected US agencies reported restrictive licensing, cost and architecture impacts; agencies did not always quantify the costs.

08 NIST, AI Risk Management Framework Core

Voluntary US framework, version 1.0 under revision at the cutoff; includes inventory, third-party contingency, response, recovery, change management and safe decommissioning outcomes.

09 Singapore IMDA, Model AI Governance Framework for Agentic AI

22 January 2026 voluntary enterprise guidance for in-house and third-party agentic systems; emphasises bounded authority, meaningful human accountability, lifecycle controls and user capability.

10 World Bank, Digital Progress and Trends Report 2025: AI Foundations

Global multilateral report on connectivity, compute, context and competency; national and income-group foundations, not enterprise continuity outcomes.

11 ECLAC, Latin American Artificial Intelligence Index 2025

Official comparison of enabling factors, research and adoption, and governance across 19 Latin American and Caribbean countries; ecosystem evidence, not firm-level resilience.

12 African Union, Continental Artificial Intelligence Strategy

Endorsed July 2024 and published 9 August 2024. Continental strategy calling for national implementation, infrastructure, data, skills and cooperation; normative rather than enterprise outcome evidence.

13 Cooper et al., Accelerating the Mitigation of LLM Inference Nondeterminism Across GPU Architectures

v1 submitted 22 September 2026. Technical preprint covering NVIDIA Ampere, Ada and Hopper with bounded workloads; not independent enterprise switching evidence.

14 Kaur, Sharma and Kahlon, Interoperability and Portability Approaches in Inter-Connected Clouds

Peer-reviewed ACM Computing Surveys article, 2017, reviewing more than 120 papers. Establishes long-running systems challenges but predates current generative AI services.

15 US Treasury, The Financial Services Sector's Adoption of Cloud Services

February 2023 sector report on cloud benefits, operational incidents, concentration, contract dynamics and migration challenges; US financial-sector evidence.

THE AI INSTITUTE

Make the decision.
Then measure what changes.

Research · Readiness · Capability

theai.institute