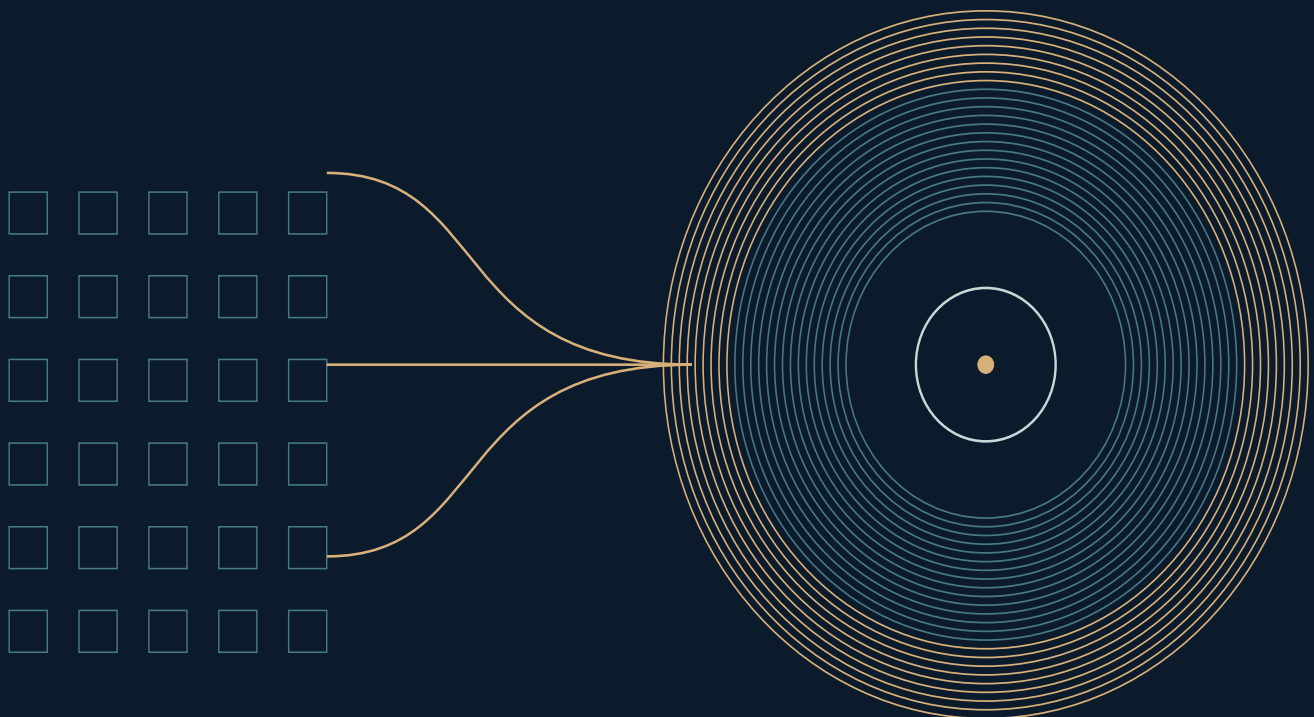


STRATEGY / OCTOBER 2026

From pilots to an operating model

Allocate decision rights before delegating work to AI



THE DECISION IN BRIEF

AI becomes an operating capability when the business assigns four rights: establish the evidence, authorize action, resolve exceptions and change the system.

- 01

A correct recommendation can still lack permission to proceed. Before expanding AI delegation, assign the authority over evidence, action, exceptions and system changes.
- 02

Reuse central controls for identity, monitoring and suppliers. Keep acceptance of business outcomes with domain owners, and give independent reviewers the evidence and standing to narrow or stop a release.
- 03

Test the operating model with a disputed case and a shift change. If the next team cannot identify the permitted action, responsible decision-maker and fallback, the deployment is still dependent on its project team.

INSIDE THIS PERSPECTIVE

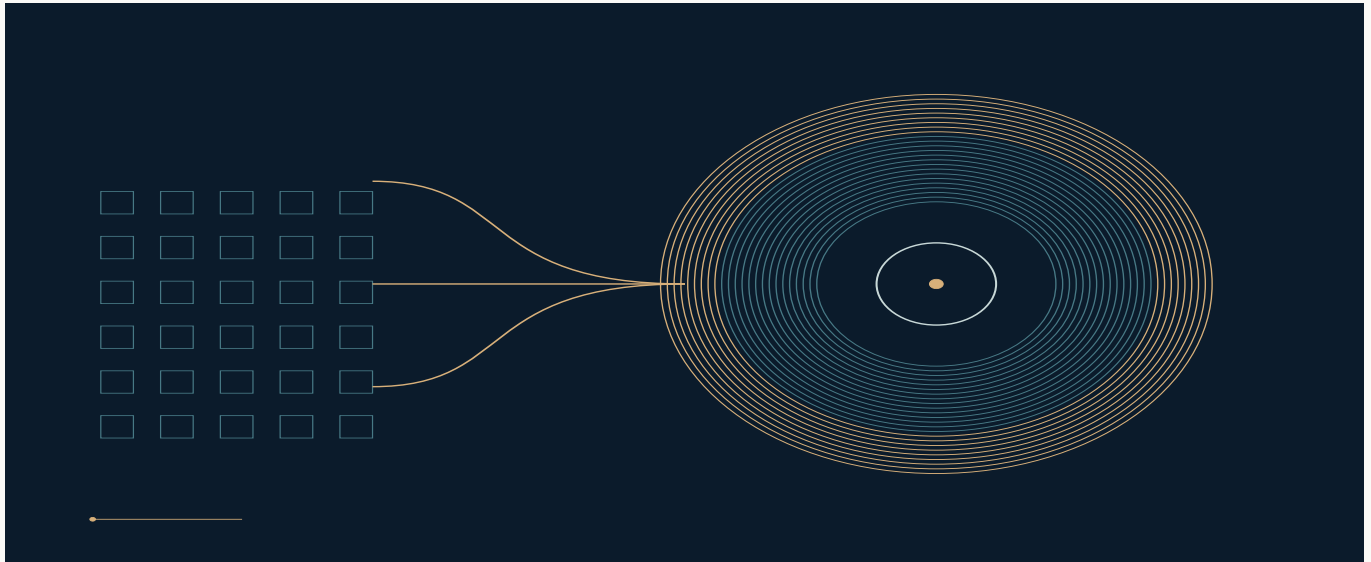
The argument at a glance	3
01 / Allocate evidence and action rights	4
02 / Make challenge effective	5
Exhibit 01 / The refund dispute needs a decision record	6
03 / Maintain the operating record	7-8
Exhibit 02 / Match oversight to the character of the decision	9
04 / Resolve ownership before scaling	10
The leadership agenda	11
Method and references	12

THE MAYDEN VIEW

Capability makes an action possible. Authority makes it a business decision.

ONE DECISION, FOUR CONNECTED QUESTIONS

From isolated pilots to one operation



OPERATING LOOP / CONCEPTUAL ARTWORK

01 Allocate evidence and action rights

A correct answer still needs authority

02 Make challenge effective

Share common controls and retain domain judgement

03 Maintain the operating record

Human oversight is a designed capability

04 Resolve ownership before scaling

Avoid turning governance into an internal monopoly

Capability makes an action possible. Authority makes it a business decision.

CHAPTER 01

01

Allocate evidence and action rights

A correct answer still needs authority

In an illustrative dispute, a customer record promises a refund while the current policy excludes it. An AI assistant can retrieve both accurately and still have no authority to settle the dispute. That is the operating-model problem in miniature. As systems move from drafting answers to changing records and issuing instructions, leaders must decide who may authorize the consequence. The interface can stay the same while the delegation changes fundamentally.

Experienced employees bridge policy gaps through judgement: they know which record to trust, whom to consult and when a familiar exception needs escalation. AI deployment exposes those informal arrangements. Use the redesign to make them explicit. Encoding a workaround without examining its legitimacy turns personal discretion into a repeatable business rule.

Begin with decision rights. Name who accepts the outcome, establishes the source record, resolves uncertainty and approves changes. A centre of excellence can support those people; it cannot assume every domain decision. The objective is a workflow that the business can operate after the project team moves on.

Assign four rights before expanding delegation

mayden separates four rights. Evidence owners establish which records are authoritative and resolve contradictions. Action owners authorize what the service may change or communicate. Exception owners decide cases outside the ordinary path. Change owners approve material updates to models, prompts, rules, data and thresholds. This is an analytical design framework; a small company may assign several rights to one person while recording them separately.

The distinction between evidence and action is especially important. A finance database can be authoritative about an account balance without authorizing a refund. A legal policy can establish a constraint without deciding the commercial response. A model can retrieve both correctly and still lack authority to reconcile a conflict. The operating record should therefore identify the decisive source and the person or policy that grants permission to act.

The change right is frequently underestimated. A prompt adjustment can change behaviour even when no traditional application code changes. A supplier's model update can alter the balance between refusal and completion. A new source can introduce contradictions. Treat material behavioural changes as operating changes, with an owner, evidence and rollback. This does not mean every wording correction needs an executive committee; it means the change process follows consequence rather than file type.

Exhibit 01 applies the four rights to the fictional refund dispute. The evidence owner validates the promise, policy version and relevant facts. The exception authority records the permitted remedy within applicable obligations. An authorized executor carries out that disposition. The change owner separately decides whether the underlying policy should change. Assign a deadline and deputy: if the decision-maker is unavailable, hold the payment and route the case to a staffed queue. Staff may report that review is pending without promising a refund. The rehearsal tests authority, availability and handover; it does not settle the legal question.

CHAPTER 02

02

Make challenge effective

Share common controls and retain domain judgement

A central capability is valuable where consistency reduces risk or duplicated effort. Identity, access logging, supplier assessment, common evaluation infrastructure and incident reporting often qualify. Domain judgement is less portable. The acceptable evidence for a supplier approval differs from that for a service remedy. A central team should provide tools and minimum conditions while the domain owner defines what completion means and accepts the consequences of operating within those conditions.

NIST's AI RMF organizes governance alongside context, measurement and management. ^[1] This supports a continuous relationship between policy and use rather than a one-time sign-off. mayden's interpretation is to make central controls reusable but visible: a business team should understand which obligations are inherited and which decisions remain its own. A central approval must never be misread as evidence that every subsequent domain use is suitable.

The April 2026 interagency model-risk guidance explicitly excludes generative and agentic AI models from its scope. ^[3] It still provides a useful, limited comparison for traditional model governance. mayden extends the management idea of proportionality here as its own proposal: the assurance burden should reflect a workflow's consequence. This extension must not be mistaken for a regulatory requirement or an assertion that the guidance validates a generative-AI deployment.

Make independent challenge capable of changing the answer

Independent challenge must be able to change the release decision. Give reviewers representative cases, failures, acceptance criteria and the deployed configuration. Permit them to require narrower scope, further evidence or a delay. A team limited to presentation slides can challenge the story, but cannot reliably challenge the operating system.

GAO's framework places performance and monitoring beside governance and data. ^[2] That is useful for designing an assurance conversation that goes beyond policy compliance. Ask whether the measured task matches the intended work, whether the data represent the population and whether performance remains relevant over time. The reviewer should record the reasoning behind a judgement, including residual uncertainty. A green status without a defensible argument provides little help when conditions change.

Independence also needs domain competence. A reviewer who cannot recognize an incorrect commercial interpretation may approve fluent but unusable output. A domain expert who helped build the system may struggle to see its assumptions. Pairing methodological challenge with domain expertise can improve the review, but the pairing itself is not evidence of quality. Evaluate the review process through the defects it detects, the clarity of its decisions and its ability to distinguish tolerable variation from material failure.

ILLUSTRATIVE DESIGN · NO MEASURED RESULTS

The refund dispute needs a decision record

Illustrative decision record

Record field	Illustrative allocation	Evidence or boundary
Authoritative evidence	Customer-record and policy owners	Validate the previous promise, current rule and conflict
Exception decision	Named exception authority	Resolve within obligations; record reason and allowed remedy
Execution	Authorised refund executor	Act only on the recorded disposition; verify payment result
Policy change	Policy-change owner	Separate case exception from enduring rule/version change
Review and expiry	Case owner at the assigned deadline	Unresolved case stays open; permission cannot silently roll forward
Unavailable authority	Named deputy or staffed exception queue	Hold payment mutation; communicate pending review only

THE MANAGEMENT IMPLICATION

Authority to validate evidence is not authority to issue a refund.

Fictional disputed refund, not legal advice or an implemented client control. Roles may share a person but their rights remain distinct. Source: mayden analysis. Proposed framework or illustrative scenario.

CHAPTER 03

03

Maintain the operating record

Human oversight is a designed capability

Putting a person after the model does not automatically improve performance. A preregistered meta-analysis of 106 experiments found that human-AI combinations performed better than humans alone on average, but worse than the better standalone actor; results varied across tasks. ^[5] The experiments were published between January 2020 and June 2023. This supports testing the actual division of work and the relevant alternative, rather than assuming that either a review checkbox or full automation is superior.

The published Cybernetic Teammate experiment studied 791 professionals at Procter & Gamble. AI-assisted individuals matched unassisted teams on the assessed innovation tasks. Unassisted teams were better at selecting their highest-rated idea, although AI-assisted participants still produced higher-quality selected ideas overall. ^[8] Section 5.4.1 and Figure 8(b,d) distinguish these selection measures; the experiment does not identify a universal organisational mechanism. This distinction separates generation quality from selection skill. For mayden, it supports testing who should generate, evaluate and authorize an

option separately. It does not establish that teams can be removed or that the same allocation works in consequential operational decisions.

A reviewer needs sufficient time, relevant evidence and authority to disagree. If the interface presents a polished recommendation before showing the underlying record, it may make independent assessment harder. If workload targets reward quick acceptance, the review function can become ceremonial. Design the review around the decision: show the unresolved issue, the evidence that matters, known limitations and the available alternatives. Measure whether reviewers detect important errors rather than how often they click approve.

The UK AI Playbook calls for meaningful human control at an appropriate stage. ^[6] In company terms, this may involve approving a policy boundary before routine execution, reviewing a particular action or monitoring a reversible process with intervention rights. These are different control designs. Select among them according to consequence, observability and reversibility. A system whose errors are hard to detect and expensive to undo needs a different arrangement from one whose output remains a disposable draft.

CHAPTER 03

03

Maintain the operating record

Build records that support handover and recovery

Every material deployment should have a concise operating record: purpose, owner, permitted actions, authoritative evidence, evaluation results, exceptions, change history and fallback. The record lets a new manager understand what was delegated and why. It also lets an incident team distinguish a data problem, a model change and an unauthorized use without depending on the memory of the people who built the first version.

Model-card research proposes documenting intended uses and evaluation conditions. ^[4] An enterprise operating record extends that idea from a model to a complete workflow. It should state where the component evidence ends and

local validation begins. A vendor's evaluation can establish useful information about its model, but it cannot demonstrate that a customer's data, interfaces, reviewers and business rules form an effective system.

Engineering research on hidden technical debt highlights dependencies that make machine-learning maintenance harder than a narrow view of the model suggests. ^[7] The corresponding governance issue is ownership of those dependencies. If a source schema changes, who notices? If an upstream service fails, who can suspend the workflow? If a supplier retires a model, who funds substitution? These questions belong in the operating model because continuity is a business responsibility as well as a technical property.

ORIGINAL MAYDEN FRAMEWORK · ILLUSTRATIVE ARRANGEMENTS

Match oversight to the character of the decision

Illustrative design choices

Decision conditions	Candidate arrangement	What must be tested
Reversible, visible, low consequence	Bounded execution with monitoring	Detection and rollback
Material consequence; verifiable evidence	Review before action	Reviewer accuracy and capacity
Ambiguous evidence; difficult reversal	Expert-led decision with AI support	Independent judgement and source sufficiency

THE MANAGEMENT IMPLICATION

Oversight is a workload and interface design question, not simply a human-in-the-loop label.

No legal classification implied. Local laws and obligations still apply. Sources: [1], [5], [6]

CHAPTER 04

04

Resolve ownership before scaling

Avoid turning governance into an internal monopoly

The argument for explicit rights can be misused to centralize every decision. That creates a different failure: business teams wait for a specialist group that lacks enough context to judge their work. A practical model delegates within clear boundaries. Central leaders set non-negotiable conditions and provide reusable controls; domain owners can act inside them with recorded evidence. Escalation is reserved for changes in consequence, uncertainty or scope that the delegated boundary does not cover.

The strongest counterargument is that formal rights slow experimentation. They can, if experimentation is treated like production. Separate environments and permissions allow teams to explore with synthetic or appropriately authorized material while keeping consequential action constrained. Define when a prototype becomes a business dependency and which operating responsibilities must be in place at that point. Define the transition explicitly.

Four rights do not require four departments. A compact firm may combine outcome and action ownership, use a data steward for evidence, and commission cross-functional or external evaluation challenge. Record where roles overlap and how conflicts are settled. Clarity matters more than organizational size.

Make the service independent of its project team

The CEO should settle three questions: which consequences require explicit authority, which executive owns each major outcome, and who resolves disagreements. These choices give teams a usable delegation and prevent a technical capability from quietly becoming company policy.

Apply the four-rights map to workflows already influencing business decisions, including embedded AI in purchased software. Prioritize unowned permissions, owners without sufficient evidence and material changes without acceptance tests. Fund the most consequential repairs before extending the system's authority.

Then rehearse the handover. Could another team operate, challenge, change and stop the service using its records alone? Any dependence on the original builders identifies unfinished operating work. Close that gap before scaling delegation.

FROM INSIGHT TO ACTION

Decisions to take forward

OPERATING OWNER

Separate case exception from policy change

Evidence to require: Refund conflict, authorised disposition and result

CONTROL OWNER

Choose oversight that can change the answer

Evidence to require: Reviewer competence, workload and disagreement

EXECUTIVE SPONSOR

Release only with a usable exception route

Evidence to require: Unavailable-authority rehearsal and handover

mayden connects senior judgement, research and engineering to turn strategic choices into working capabilities.

EVIDENCE YOU CAN INSPECT

Research notes

Method and boundaries

Structured synthesis of eight governance, engineering and human-AI research sources with an original decision-rights architecture.

Selected sources clarify authority, oversight, validation, documentation and maintenance. Accessed 10 October 2026, including the April 2026 banking guidance rather than presenting superseded guidance as current.

Normative frameworks do not demonstrate causal performance improvements. The human-AI meta-analysis covers heterogeneous pre-July-2023 experiments. Banking and government material is used by analogy, not as universal legal advice.

Primary sources and author research

[1] Artificial Intelligence Risk Management Framework (AI RMF 1.0)

National Institute of Standards and Technology / January 2023. Part 2; GOVERN, MAP, MEASURE and MANAGE tables. Guidance is not experimental proof that any specific governance structure works.

[Read the source](#)

[2] Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities

US Government Accountability Office / 30 June 2021. Highlights and framework principles. Normative framework, not a measured causal estimate.

[Read the source](#)

[3] Supervisory Guidance on Model Risk Management

Federal Reserve; interagency SR-26-2 / 17 April 2026. Footnote 3; governance and validation sections. Not applicable to generative or agentic AI models. Any extension in this report is maiden analysis, not regulator endorsement.

[Read the source](#)

[4] Model Cards for Model Reporting

Mitchell et al.; Google Research / 2019. Abstract. Documentation can expose limits; it does not independently validate a deployment.

[Read the source](#)

[5] When Are Combinations of Humans and AI Useful?

Vaccaro, Almaatouq and Malone; author manuscript / 9 May 2024; arXiv version 1. Sections 2.1 and 4; 106 experiments and 370 effect sizes, publications January 2020-June 2023. Studies through June 2023; task heterogeneity and publication bias; not current-model ROI.

[Read the source](#)

[6] Artificial Intelligence Playbook for the UK Government

Government Digital Service / 10 February 2025. Principles; human oversight; life cycle and commercial guidance. UK public-sector context; not a substitute for jurisdiction-specific advice.

[Read the source](#)

[7] Hidden Technical Debt in Machine Learning Systems

Sculley et al.; NeurIPS / 2015. Abstract. Predates generative AI; used for structural insight, not current failure-rate estimates.

[Read the source](#)

[8] The Cybernetic Teammate: A Field Experiment on Generative AI and Teamwork

Dell'Acqua et al.; Organization Science 37(4), 1217-1242 / 12 June 2026. Abstract; Section 3 experimental design; Section 5.4.1 and Figure 8(b,d) (idea selection versus selected-idea quality); footnote 8 (participant and complete-survey counts). One-day workshop, specified tools, one company; not a staffing or autonomy experiment. Published n differs from older working paper.

[Read the source](#)