Practical AI for Risk Management in Financial Services

Archive text edition of the revised manuscript · 11 November 2026

Prepared by the Harbour Think Tank under the leadership of Adrian Dunkley. The original evidence review is dated 26 October 2026.

About this paper

The Harbour Think Tank of the Caribbean AI Association covers the economy, work and enterprise. This paper applies that remit to risk teams in financial services, one of fourteen business functions in the collection. [1]

We wrote this paper for chief risk officers, operational risk managers, model validators, business control owners and committee members. The proposals organise evidence and prepare questions. Risk acceptance, credit decisions, capital calculations and crisis command never pass to an assistant. Compliance obligations and cybersecurity operations have their own papers, and so does enterprise data governance.

This revision adds a page of build notes to each use case. The notes give a starter prompt, example tools, skills, a method and the limits for an AI agent.

Foreword

A risk committee wants to know if a control worked, and an uploaded document shows only that someone uploaded it. An incident review must keep the awkward fact that fits the preferred explanation of no one. We trust a dependency map only after the people who run the service recognise it.

Artificial intelligence can help to prepare this evidence. It can also bury an unresolved issue under a fluent summary, or read repeated words as a shared cause. For this reason, risk professionals must be able to challenge the sources. They must also be able to challenge how the assistant reads them.

Our advice is to start with a bounded evidence task and a reviewer who can say what the correct answer is. Authorised people then decide if the assistant has earned its place.

Adrian Dunkley, President, Caribbean AI Association

Acknowledgements and terms

The Harbour Think Tank of the Caribbean AI Association prepared this paper, and President Adrian Dunkley led the work. Institutional sources give context and control principles. The workflows, the test designs, the screening scores, the prompts and the agent levels are our proposals. They are not results from a named financial institution, and nothing in this paper implies the endorsement of a regulator.

An artificial intelligence (AI) assistant finds and organises information, and it can draft text. An agent can also do configured steps with connected tools. In this paper, neither has authority to change limits or to execute responses. Risk appetite is the level and type of risk that the organisation will accept. A control is an action to prevent or detect a risk event, or to correct its effect. A scenario is a stated hypothetical, and it is never a forecast.

Operational resilience is the ability to keep critical operations in service through a disruption. A key risk indicator is a defined measure that helps to find possible deterioration. The design of a control and its operation need separate evidence. The approved definitions of the firm govern its register and reports.

A prompt is the written instruction that a person gives to an assistant. The word skill has two meanings in this paper. A staff skill is an ability that a person learns. An agent skill is a saved set of instructions that an assistant loads for one task. [8]

Approved source

Draft or retrieval

Human review

Accepted output

The operating boundary

In all ten use cases, the assistant prepares evidence and questions. Risk acceptance, control conclusions, limit changes and response decisions stay with authorised people.

Contents

Section | Page

Executive summary | 5

1 The regional business case | 6

2 What the evidence supports | 7

3 The enterprise lens | 8

Process design: BPMN swimlanes | 9

4 Ten practical use cases: how to read each use case | 10

Use cases 01 to 10, each with build notes | 11-30

Priority and responsibility tables | 31-32

5 The foundations lens | 33

6 From pilot to practice | 34

Appendix A Pilot decision canvas | 35

Appendix B Measure value | 36

References | 37-38

Read the summary and the priority tables first. Then select one use case and define a pilot. Complete the decision canvas with the people who do the work.

Executive summary

Start with the risk inventory and the scenario library, together with the committee evidence pack. Reviewers can test each one without a connection to live transactions. Loss records and supplier information wait until their owners can explain the data and say who can see it. Indicators wait for the same reason.

Where evidence is missing, the assistant must say so and leave the gap visible. Observation, explanation, estimate and decision each need a separate field. A lower indicator can mean lower risk. It can also mean that the reporting population or the definition changed.

Select a pilot from the priority table, and agree the decision canvas with its owner. Then assess the result on four measures: material omissions, unsupported conclusions, reviewer disagreement and the total preparation effort. The contents must survive challenge before anyone accepts the pack.

Each use case now has build notes with a starter prompt and example tools. The notes also set a limit for AI agents. In the pilots, no agent acts without the approval of a named person.

Four decisions for the sponsor

Select the evidence task and the reviewer who is accountable for it. Name the decisions and the systems that the assistant must not touch. Approve the failure threshold. Name the person who has authority to stop the pilot. Settle all four before the assistant receives one confidential record.

Choose a bounded process

Measure the baseline

Test with review

Decide from evidence

1 The regional business case

The Financial Stability Report 2024 of the Bank of Jamaica shows why a risk measure needs a precise label. Its table for deposit-taking institutions gives non-performing loans as a percentage of total loans. The figure was 2.2% in December 2019 and 2.9% in December 2021. In December 2024 it was 2.5%. Each figure is a ratio of loan values across the sector. It does not give a percentage of borrowers, and it predicts no future defaults. [2]

Source: Bank of Jamaica, Financial Stability Report 2024, Table A.1, page 67. The numerator is non-performing loans. The denominator is the total loans of the reported sector of deposit-taking institutions. These are historical December observations. They are not a 2026 assessment or a Caribbean-wide series. [2]

We use the chart as an exercise in faithful reports, and no firm should use it as a benchmark. A risk team must still bring its own exposure definitions, concentrations, control results and emerging concerns. A healthy sector ratio can exist next to a weak portfolio at one institution.

A regional group must keep its legal-entity boundaries intact before anyone writes a consolidated explanation. It must also keep the currencies and the report dates. If the information of one subsidiary is missing, the gap must stay visible. The same label can hide different report rules in different entities. Only the relevant owners can settle that disagreement, even if the summary is fluent.

Jamaica: non-performing loans as a share of total loans

Dec 2019: 2.2%

Dec 2020: 2.8%

Dec 2021: 2.9%

Dec 2022: 2.5%

Dec 2023: 2.5%

Dec 2024: 2.5%

Bank of Jamaica, Financial Stability Report 2024, Table A.1, page 67. [2] The ratio divides non-performing loans by the total loans of the sector of deposit-taking institutions, as the Bank of Jamaica reports them.

2 What the evidence supports

The 2021 operational-resilience principles of the Basel Committee cover governance, disruption tests, dependencies and lessons from incidents. The Committee wrote them for banks, and a firm must read them with the applicable local requirements. They offer no evidence that an AI workflow prevents losses. [3]

The 2013 risk-data principles of the Committee ask for information that is accurate and complete, and that arrives on time. They also ask for clear reports. The Committee aimed them first at systemically important banks. We use their report disciplines as design references, and we do not claim that every principle binds every Caribbean firm. [4]

A Basel newsletter from January 2026 describes continued difficulty with data traceability and urgent reports. It notes that the adoption of AI in this area is at an early stage and depends on data quality. By its own account, the newsletter gives information and sets no new supervisory guidance. [5]

The National Institute of Standards and Technology (NIST) publishes the AI Risk Management Framework 1.0. It covers inventories, assigned responsibility, tests and continued review. It is voluntary. NIST says that a revision is in progress, and we cite the 2023 version as published. [6]

Read together, these sources support careful preparation and challenge. None gives a productivity percentage for the pilots in this paper, and we know of no Caribbean estimate. Each firm must build its own business case. Its inputs are the accepted work and the review effort, with demonstrated error control.

Test the proposed workflow

Test for material omissions and unsupported conclusions. Measure reviewer disagreement and preparation effort also. The local pilot must show that the evidence survives an independent challenge.

3 The enterprise lens

Select a decision that already has a defined evidence process behind it. Then work back from what the reviewer must establish. One example is that a risk has the same description in every place. Other examples are that a control record covers the full period, and that a committee statement agrees with the approved numbers.

Keep the evidence store apart from the decision register. The assistant can draft a proposed amendment, but a person approves any change to risk ownership, ratings, limits or actions. Where a rejected draft helps to explain a material error, store it under the authorised retention policy.

Fixed-rule checks handle totals and dates, and they find duplicate identifiers. Generated language only organises or explains what those checks already confirmed. When the assistant explains why an indicator moved, its account stays a hypothesis until the relevant analyst verifies it. The account must never start an automatic intervention. The prompts in the build notes are start points that each firm must test.

Request and source

Assistant output

Reviewer correction

Accepted result

Record elapsed time and active effort as two separate measures. A queue can become shorter while handling time stays the same, and the next improvement depends on which of the two changed.

Reconcile risks without erasing distinctions

Business Process Model and Notation (BPMN) swimlanes. Proposed workflow for use case 01.

Risk analyst and business owners

AI drafting assistant

Enterprise risk manager

Yes

No

Inventory review begins

Supply register, process scope and categories

Compare causes, events, controls and owners

Challenge proposed mergers with owners

Descriptions agreed?

Approve the reconciliation findings

Reconciliation review complete

Separate unsupported mergers; retain gaps

Save the agreed reconciliation record

Risk inventory reconciliation

The assistant accepts no risk and assigns no final rating. It never overwrites the register.

How to read the model. Rounded rectangle: task. X diamond: one decision path. Thin circle: start. Thick circle: end. Solid arrows show the sequence in one process. Lanes show who is responsible.

Notation: Object Management Group, BPMN 2.0.2. https://www.omg.org/spec/BPMN/2.0.2

4 Ten practical use cases

How to read each use case

Each use case has two pages. The first page gives the workflow, the minimum data, the pilot test and the accountable role. The second page gives build notes in five parts.

Starter prompt. Copy the prompt into the approved assistant, and replace the text in square brackets. Each prompt gives the assistant a role, its sources, one task, the rules and the output format.

Tools. The notes name example products. The names are examples and not endorsements. Vendors change product names frequently. Use only a service that the firm approved.

Skills. A staff skill is an ability that the team must build. An agent skill is an instruction file that the assistant loads for the task. Agent Skills is an open format, and many assistants can read it. [8]

Method. The notes name an established method that gives the work a structure.

Agent limits. The notes say what an AI agent can do, and where it must stop.

Four agent levels

Level 0, assist. The assistant answers one request. A person does every step.

Level 1, prepare. The agent collects approved inputs and prepares a draft. It changes no system and sends no message.

Level 2, act with approval. The agent proposes an action. A named person approves each action before it runs.

Level 3, act alone. The agent acts without approval inside fixed limits.

Each pilot in this paper starts at Level 0 or Level 1. A move to a higher level needs a new approval after the pilot. No pilot is at Level 3.

01 Reconcile the risk inventory

Current risk register

Approved process descriptions

Compare meaning and duplicates

Business owners confirm

Risk inventory review

One name in the register can cover different risks, and entries that have different business meanings stay separate. A risk analyst supplies approved process descriptions and current entries. The assistant groups possible duplicates and flags inconsistent terms. It compares the cause, the event, the consequence, the control and the owner of each entry.

The minimum data is the current register, the risk categories, the process scope and the document dates. Start with one business unit, and remove personal identifiers that the work does not need. Label a missing owner or consequence as a gap. A job title is not evidence of responsibility.

Pilot on a sample that a person already reconciled by hand. The sample must have duplicate names and also risks that are truly different. Count the missed distinctions and the unsupported mergers, and record the review time. Then ask a process owner who did not see the draft to explain each entry. Stop if the tool deletes a material risk or silently lowers a rating.

The enterprise risk manager is accountable. Business owners confirm descriptions, and risk specialists challenge them. The assistant cannot accept risk or assign final ratings, and it cannot overwrite the register. Agreed categories make entries comparable, but a tidy register can still miss a risk that nobody wrote down.

First working session

Take one process register, and find duplicate words that hide different risks.

01 Build notes: risk inventory

Starter prompt

"You compare entries in a risk register. Use only the attached register and process descriptions. For each entry, show the cause, the event, the consequence, the control and the owner. Write 'Gap' where a part is missing. List pairs of entries that can be duplicates, and say which parts agree and which parts are different. List terms that the entries use inconsistently. Suggest a merger only if all parts agree, and mark it 'For owner review'. Do not delete an entry. Do not change a rating. Do not name an owner from a job title."

Tools

Register: an export from the risk system of the firm, for example ServiceNow, or an Excel register.

Comparison: Claude or Microsoft 365 Copilot, with the register and the process descriptions attached.

Reconciliation record: a SharePoint list, with the decision of the owner on each proposed merger.

Skills

Staff skill: tell two risks apart when they share a name but have different causes.

Agent skill: a register-comparison skill with the five parts of an entry and the merger rule.

Method

A bow-tie diagram for each disputed pair of entries. It shows causes and preventive controls on one side, with consequences and recovery controls on the other side.

Agent limits

Level 1. An agent can prepare the comparison for a business unit that the risk manager selects. It works on an export, and it must not write to the register. It must not merge entries or change a rating.

02 Find recurring themes in operational incidents

Defined incident question

Approved closed incident records

Observable recurring themes

Investigators assess causes

Process-owner learning

Closed incident records without identities come first. The assistant groups approved summaries by features that anyone can observe, such as failed handoffs or repeated reconciliation breaks. It links each theme to its records, and it keeps a reported cause apart from the untested explanation of an analyst.

Each record needs an identifier, a verified chronology, the affected process, the recorded impacts and the investigation status. Near misses stay apart from realised losses. If an investigation has no findings yet, the field must read "unknown", even when other records have similar words.

Mix frequent minor events with one rare severe event, because the test is that the severe event survives the grouping. Investigators check that themes keep contrary evidence and unique events. Count the incorrect groupings and the omitted severe incidents, and record the time to find supporting records. Stop if an allegation becomes a finding. Stop if a summary identifies an employee outside the approved purpose.

The operational risk manager is accountable. Investigators decide causes, and management approves remediation. The assistant attributes no misconduct and closes no investigation. It never treats a recurring phrase as statistical proof of cause. Lessons must go to the owner of the affected process for assessment.

First working session

Select closed incidents without identities. Include a rare severe event that must stay visible.

02 Build notes: incident themes

Starter prompt

"You group closed incident records into themes. Use only the attached de-identified records. Group them by features that the records state, such as a failed handoff or a repeated reconciliation break. For each theme, list the record identifiers and quote the words that support the theme. Keep reported causes under 'Reported cause' and your own ideas under 'Not tested'. List each severe event on its own line, even if it fits no theme. Keep near misses apart from losses. Write 'Unknown' where an investigation has no findings. Do not infer a cause from similar words. Do not name or describe a person."

Tools

Records: an export of closed incidents from the incident system, for example ServiceNow.

De-identification: Presidio or a manual check, before the records reach the assistant.

Themes: a Claude Project with access for the operational risk team only.

Evidence table: Excel, with one row for each record and its theme.

Skills

Staff skill: tell an observed feature from an inferred cause.

Agent skill: an incident-theme skill with the "Reported cause" and "Not tested" sections and the severe-event rule.

Method

Thematic analysis with two independent coders, and a Pareto chart of the themes. Keep each severe event on its own line, whatever its frequency.

Agent limits

Level 0. An investigator starts each analysis from an approved set of closed records. No agent has continuing access to incident records. The assistant must not attribute misconduct or close an investigation.

03 Assemble evidence for control review

Approved control description

Logs and approval evidence

Check period and sign-offs

Independent tester evaluates

Control review evidence pack

A control owner supplies the approved description and the review period. The assistant indexes the logs and reconciliations, with the approvals. It flags absent dates and missing sign-offs. It also flags evidence that belongs to another control or period.

The pack needs the control identifier, the test criteria, the expected frequency, the authorised evidence and a sample plan. An uploaded procedure describes what must occur, and evidence of operation shows what did occur. Missing evidence produces an exception request and never a conclusion on effectiveness.

Put an expired approval and a duplicated log into the pilot packs, with one failed control. A tester compares the index with one that a person made by hand. Count the missed gaps and the incorrect links, and record the retrieval time. Two failures end the test. One is a control that the pack calls effective without a test. The other is a failed instance that an aggregate summary hides.

The risk assurance lead is accountable. The responsible tester evaluates operation, and an authorised manager accepts remediation. Where the assurance arrangements of the firm require it, the preparer and the challenger must be different people. The assistant helps to build the pack, and it never gives an assurance opinion.

First working session

Pair one control description with evidence from the correct review period and one known missing approval.

03 Build notes: control evidence

Starter prompt

"You index the evidence for a control review. Use only the attached control description and evidence files for [period]. For each expected instance of the control, list the evidence file, with its date and its sign-off. Write 'No evidence' where an instance has none. Flag evidence with no date or no sign-off. Flag evidence that belongs to another control or another period. Flag an approval that expired before its use. List each failed instance separately. Do not summarise failed instances into a total. Do not say that the control is effective or not effective."

Tools

Control library: the risk system of the firm, for example ServiceNow, with the control description and frequency.

Evidence store: a SharePoint library with one folder for each control and period.

Index: Claude or Microsoft 365 Copilot for the draft, and Excel for the tester.

Skills

Staff skill: tell evidence of design from evidence of operation.

Agent skill: a control-evidence skill with the index format and the "No evidence" rule.

Method

Keep the test of design apart from the test of operation. Use the sample plan of the firm, and apply a maker-checker rule: the preparer and the challenger are different people.

Agent limits

Level 1. An agent can prepare the index when a review period closes. It has read access to the evidence store. It must not give a conclusion on a control or accept remediation.

04 Build a challengeable scenario library

Disruption question

Service map and assumptions

Draft plausible scenario

Specialists challenge severity

Tabletop exercise card

The scenario analysis lead is accountable. Risk specialists select a disruption question. An example is the failure of a payment-service supplier in a busy period. The assistant drafts a card that gives the initial conditions, the dependencies, the assumptions, the consequences to examine and the decisions for participants.

The assistant needs the approved service map, the impact definitions, the existing scenarios and the exercise objectives. Use fictional operational details wherever possible. If nobody knows how long recovery takes, that unknown becomes a question for the exercise. The card leaves it open and does not fill it with a convenient figure. The card marks assumptions from specialists apart from facts that records support.

Pilot a small number of cards with a tabletop facilitator. A good card makes participants take a distinct decision, and they can say what evidence would change their response. Count the unsupported assumptions and the duplicated scenarios, and count the missed dependencies. Stop if an illustrative event appears as a forecast. Stop if the model invents a probability or a loss estimate.

Specialists approve severity and plausibility, and management sets disruption tolerances. Quantitative stress calculations stay in validated tools. The output is material for an exercise. It is not a live response instruction, and it does not replace any stress test that the firm must do.

First working session

Write one exercise objective, and list the assumptions that a specialist must approve.

04 Build notes: scenario library

Starter prompt

"You draft a scenario card for a tabletop exercise. Use only the attached disruption question and service map, with the impact definitions. Write five sections: Initial conditions, Dependencies, Assumptions, Consequences to examine and Decisions for participants. Label each assumption with the specialist who must approve it. Label each fact with its record. Where no one knows a value, such as the recovery time, write it as a question for the exercise. Use fictional names for people and suppliers. Do not give a probability. Do not give a loss estimate. Do not call the scenario a forecast. Do not write a live response instruction."

Tools

Inputs: the approved service map and impact definitions, in a SharePoint library.

Cards: Claude or Microsoft 365 Copilot for the draft, in a Word template with the five sections.

Library: a SharePoint list with the status of each card, its approver and owner, and its last exercise date.

Skills

Staff skill: challenge the severity of a scenario and say what evidence would change a response.

Agent skill: a scenario-card skill with the five sections and the forbidden figures.

Method

Scenario cards for a tabletop exercise. Each card states its assumptions, and specialists approve severity and plausibility before a facilitator uses it.

Agent limits

Level 0. A risk specialist starts each card from a disruption question. The assistant has no connection to live systems or response tools. It must not add a card to the approved library.

05 Explain movements in approved risk indicators

Checked indicator values

Check denominators and periods

Series comparable?

Draft sourced explanation

Indicator-owner investigation

Yes

No

An analyst supplies checked indicator values and definitions, with documented operational changes. The draft of the assistant has four separate parts: the measured movement, the verified facts that contributed, the possible explanations and the open questions. Every factual statement has a source and an observation date.

The draft needs the numerator, the denominator, the comparison period, the threshold version and the validated values. It also needs the known outages and definition changes. If the denominator is missing or the series is not comparable, the draft gives no trend conclusion. The issue then goes to the indicator owner.

Test on historical observations that have a denominator change and a data outage, with one true threshold breach. Count the unsupported causal statements and the missed caveats, and record the correction effort of the analyst. The pilot stops if a narrative changes a threshold or hides a breach. It also stops if a narrative calls an association a proven cause.

The indicator owner approves interpretation and escalation under existing policy, and the risk reporting manager is accountable for the use case. The assistant calculates no new risk appetite and changes no limit. Someone must still investigate a material warning, whatever the generated explanation says.

First working session

Select an indicator series in which a denominator or a definition changed.

05 Build notes: indicator explanations

Starter prompt

"You draft an explanation of a movement in a risk indicator. Use only the attached checked values and definition, with the change log. Write four sections: Measured movement, Verified facts, Possible explanations and Open questions. Give the numerator, the denominator, the two periods and the threshold version. Give a source and a date for each fact. If the denominator is missing or the definition changed, write 'Series not comparable' and give no trend. Say plainly if a value is past its threshold. Do no calculation. Do not call a possible explanation a cause. Do not suggest a new threshold or limit."

Tools

Values: the approved indicator report in Power BI or Excel, after the checks of the analyst.

Change log: a SharePoint list of outages and definition changes, with dates.

Draft: Claude or Microsoft 365 Copilot, with the values and the change log attached.

Skills

Staff skill: check the denominator and the definition before a trend statement.

Agent skill: an indicator-explanation skill with the four sections and the "Series not comparable" rule.

Method

Check four possible sources of a movement in order: a data fault, a definition change, a denominator change and a numerator change. Give a business explanation only after those checks.

Agent limits

Level 1. An agent can prepare a draft when the analyst releases checked values for a period. Escalation of a breach never depends on the draft. The agent must not change a threshold or start an intervention.

06 Reveal shared supplier dependencies

Supplier and contract records

Critical service maps

Find shared dependencies

Procurement and owners verify

Concentration questions

More than one critical activity can depend on the same provider or subcontractor, or on the same host arrangement. The assistant compares approved supplier records with service maps. It lists each possible shared dependency with the strength of its evidence, and it adds questions for procurement and for service owners.

The inputs are current contracts, disclosed subcontractors, service identifiers, legal entities and approved criticality assessments. Never infer shared ownership or infrastructure from similar names. An undisclosed subcontractor is an information gap, and the dependency stays open.

Our screen gives this use case the lowest data score, and we are not confident that its inputs are ready. Pilot on a small number of services that have one common dependency, verified by hand. Count the missed shared providers and the false matches, and count the ownership questions without an answer. Discuss if each named alternative could truly take over. Stop if the assistant reports an uncertain link as fact. Stop if supplier-sensitive data leaves its

permitted environment.

The third-party risk manager is accountable. Procurement and legal specialists confirm contract facts. Management decides the treatment and decides to accept the supplier or not. The assistant cannot terminate a contract or approve an exception, and it cannot certify resilience. Its concentration map raises questions. Only a test shows that the alternative arrangements work.

First working session

Confirm one shared supplier dependency from the contracts before you test wider matches.

06 Build notes: supplier dependencies

Starter prompt

"You compare supplier records with service maps to find shared dependencies. Use only the attached contracts and supplier register, with the service maps. For each provider or disclosed subcontractor that more than one service uses, list the services and the legal entities. Give the contract clause or the record that shows the link. Rate the evidence: Confirmed in contract, Stated in register, Inferred or Not known. Mark every 'Inferred' link for verification. Write 'Subcontractors not disclosed' where a contract is silent. Write one question for procurement and one for the service owner. Do not match suppliers by similar names. Do not report an uncertain link as a fact. Do not say that an alternative supplier is adequate."

Tools

Contracts: the contract store of the firm, for example SharePoint or iManage, with read access.

Supplier register and service maps: Excel or the risk system of the firm.

Comparison: Claude or Microsoft 365 Copilot in a restricted workspace.

Map: Visio, after procurement confirms each link.

Skills

Staff skill: read a subcontract clause and confirm a link with procurement.

Agent skill: a shared-dependency skill with the four evidence ratings and the name-match rule.

Method

A matrix of services against providers, with a column for disclosed subcontractors. Confirm each shared link from the contract before it goes on the concentration map.

Agent limits

Level 0. The third-party risk manager starts each comparison on an approved set of contracts. Supplier data stays in its permitted environment. The assistant must not contact a supplier or change the supplier register.

07 Maintain a reviewable model and AI inventory

Application and procurement list

Existing model inventory

Identify registration candidates

Model governance classifies

Reviewable inventory proposals

The assistant reads approved application lists and procurement records, and compares them with existing model records. Then it suggests items that possibly need registration. Each candidate gets a purpose, an owner, a version, its data, its users and its review status. Any classification question without an answer stays on the entry.

The model definition of the firm sets the boundary, together with authorised system lists and the validation records that exist. Conventional models and AI applications stay distinct. A spreadsheet is a regulated model only if the definition says so. A candidate that has no version or no owner stays on the list as incomplete. No one marks the

inventory complete while one such candidate stays open.

Use one department. It must have known examples of registered and unregistered tools, and one retired tool. Validators check inclusion and status against the reference list. Count the missed candidates and the false classifications, and count the stale approvals. Stop if an undocumented tool gets the label "validated". Stop if a provider update inherits approval automatically.

The model risk lead is accountable. The model governance function decides the classification and the validation needs, and it decides the permitted use. An inventory entry does not show that the model performs adequately. Register the assistant itself also. Assess it again when its model or its data access changes, or when its intended use changes. [6]

First working session

Compare the approved model definition with the application and procurement lists of one department.

07 Build notes: model and AI inventory

Starter prompt

"You find tools that possibly need an entry in the model inventory. Use only the attached model definition, the application list, the procurement records and the current inventory. For each tool that is not in the inventory, say which part of the definition it possibly meets. Give its purpose, owner, version, data, users and review status, each with its source. Write 'Not recorded' where a field is missing. Label each candidate: Conventional model, AI application, Not a model or Classification question. Flag an approval that is older than the current version of the tool. Do not label a tool as validated. Do not decide a classification. Do not mark the inventory as complete."

Tools

Source lists: the application inventory and the procurement records of the firm, for example in ServiceNow.

Inventory: a SharePoint list with the owner, the version, the review status and the review date of each entry.

Comparison: Claude or Microsoft 365 Copilot, with the lists and the definition attached.

Skills

Staff skill: apply the model definition of the firm to a tool, and record the reason.

Agent skill: an inventory-candidate skill with the entry fields and the "Not recorded" rule.

Method

An inventory reconciliation against the model definition of the firm. The NIST AI RMF asks for mechanisms to inventory AI systems. [6]

Agent limits

Level 1. An agent can compare the lists on a schedule and add candidates to a review queue. It must not classify a tool or record a validation. The assistant has its own entry in the inventory.

08 Map a critical service for resilience testing

Defined critical service

Confirmed people and systems

Map dependency links

Service owners confirm

Resilience test questions

Start with the service owner, who supplies process documents and confirmed operational dependencies. The assistant drafts a map from the result of the service back through staff, systems, information, facilities and third parties. It marks each link as verified or as not yet confirmed.

The map uses the service scope, the responsible owners, the current diagrams and the approved continuity arrangements. An undocumented step must never become an assumed automated connection. Each important dependency must have its source and the date when someone last confirmed it.

The pilot is a facilitated walkthrough with staff from operations and technology, and from risk. An exercise follows in which one agreed dependency is unavailable. Compare what occurs with what the map predicted. Count the missing connections and the stale recovery assumptions, and count the owner questions that stay open. The pilot ends if a diagram implies a tested recovery route where only a written plan exists.

Service owners approve the map, and authorised teams design and run tests. The operational resilience lead is accountable. The assistant has no authority to invoke a recovery plan or to switch infrastructure, and it cannot announce restoration. The map identifies what needs a test. A test supplies evidence of continuity under the conditions that it examined. [3]

First working session

Walk through one critical service with its owner before you draft its dependency map.

08 Build notes: resilience mapping

Starter prompt

"You draft a dependency map for one critical service. Use only the attached process documents and diagrams, with the notes from the walkthrough with the service owner. Start from the result of the service. Work back through five resource types: staff, systems, information, facilities and third parties. For each link, give the source and the date when someone last confirmed it. Label each link 'Verified' or 'Not yet confirmed'. Label each recovery route 'Tested on [date]' or 'Written plan only'. List steps that the documents do not describe as questions for the owner. Do not assume an automated connection. Do not state a recovery time that the records do not give."

Tools

Inputs: process documents and current diagrams, with the continuity plans, in a SharePoint library.

Walkthrough notes: OneNote, taken with the service owner.

Map: Claude or Microsoft 365 Copilot for the draft list of links, and Visio for the reviewed map.

Skills

Staff skill: walk through a service with its owner and confirm each link.

Agent skill: a service-map skill with the five resource types and the labels for links and recovery routes.

Method

End-to-end service mapping, confirmed in a walkthrough, then a test with one dependency unavailable. The Basel principles ask banks to map the dependencies of their critical operations. [3]

Agent limits

Level 0. The resilience lead starts each map after a walkthrough with the service owner. The assistant has no connection to live infrastructure. It must not invoke a recovery plan or announce restoration.

09 Check operational loss records before aggregation

Authorised loss-event extract

Approved loss definitions

Fixed-rule exception checks

Finance and owner reconcile

Reviewed aggregation inputs

A risk analyst supplies an authorised extract of loss events. Fixed-rule checks flag duplicate identifiers, inconsistent currencies, impossible dates and gaps between a gross loss and its separately recorded recoveries. The

assistant explains each exception and links it to the source record.

The minimum data is the event identifier, the accounting date, the event date, the currency, the gross amount, the recovery amount and the approved definitions. Protect the details of individuals and investigations. A blank loss amount stays blank, and no one reads it as zero. An expected insurance recovery counts only after the firm receives it.

Use fictional records that have a duplicate and a partial recovery, with an event in more than one currency. Finance calculates the totals again independently, and risk checks the event classification. Count the seeded errors that the checks missed and the false exceptions, and record the reconciliation effort. The pilot stops if the tool nets amounts against policy or invents a value. It also stops if the tool changes the accounting record.

The operational loss data owner is accountable, and finance reconciles. People approve corrections through the existing process. The assistant books no losses and decides no accounting treatment. Unresolved exceptions must stay on record, so that the committee sees what an aggregate leaves out.

First working session

Prepare fictional loss records that have a duplicate and a partial recovery, in different currencies.

09 Build notes: loss-data checks

Starter prompt

"You explain exceptions in a loss-event extract. Fixed-rule checks already ran. Use only the attached exception list and the loss definitions. Group the exceptions by rule: Duplicate identifier, Inconsistent currency, Impossible date and Recovery gap. For each exception, give the event identifier, the field, the value and the check that the analyst must do. Keep gross loss and recovery as separate amounts. Leave a blank amount blank. Count a recovery only if the record shows that the firm received it. Do no arithmetic. Do not net amounts. Do not suggest a value. Do not change a record."

Tools

Checks: fixed rules in Excel Power Query or in SQL, under change control.

Explanation: Claude or Microsoft 365 Copilot, with the exception list attached.

Exception log: a SharePoint list with the owner of each exception and its status.

Skills

Staff skill: tell a gross loss from a net loss, and an expected recovery from a received recovery.

Agent skill: a loss-exception skill with the four rule groups and the blank-amount rule.

Method

Fixed-rule data validation, then reconciliation of the totals to the finance record. Test with seeded errors, and report missed errors and false exceptions separately.

Agent limits

Level 1. An agent can explain the exceptions after each scheduled check. It has read access to the extract. It must not book a loss or correct a record.

10 Prepare a source-linked risk committee pack

Committee decision needed

Approved reports and actions

Facts, responses and challenges

Content owners sign off

Draft risk committee pack

The assistant drafts the pack from approved risk reports and open actions, with the responses of management. It keeps requested decisions apart from information items. It lists unresolved challenges and overdue actions, and it links each material figure and status to its source.

The inputs are the report cutoff, the approved measures, the action register, the committee terms and the authorised sources. Missing submissions go on a visible exception list. The assertion of management and the independent challenge must stay as two positions, even if a merged position reads better.

Pilot one historical committee cycle against the pack and the minutes that the committee approved. A committee member must be able to find the evidence behind a material statement without help from the preparer. Count the unsupported assertions and the omitted challenges, and count the incorrect action statuses. Record the preparation time, with review included. Stop if a disputed issue disappears. Stop if a draft records an approval that did not occur.

Report owners sign off the content, and every decision stays with the committee. The risk committee reporting lead is accountable for the pack. The assistant circulates no confidential papers alone, and it changes no minutes. A complete pack gives it no grounds to declare that the organisation is within appetite.

First working session

Reconstruct one historical committee pack. Mark the evidence behind every material decision request.

10 Build notes: committee evidence pack

Starter prompt

"You draft a risk committee pack. Use only the attached approved reports and action register, with the management responses, at the cutoff [date]. Make four sections: Decisions requested, Information items, Unresolved challenges and Overdue actions. Link each material figure and each status to its report and page. Put the position of management and the independent challenge next to each other, and do not merge them. List each submission that is missing. Do not record an approval or a decision. Do not say that the firm is within its risk appetite. Do not remove a disputed issue."

Tools

Sources: approved reports in a restricted SharePoint library, with Purview sensitivity labels.

Action register: a SharePoint list or Planner, with the owner and due date of each action.

Draft pack: Microsoft 365 Copilot or Claude in a restricted workspace, with Word for the final pack.

Skills

Staff skill: trace a figure in the pack to its approved report and cutoff.

Agent skill: a committee-pack skill with the four sections and the two-position rule.

Method

A traceability matrix from each material figure to its source and cutoff. Show the position of management and the independent challenge next to each other.

Agent limits

Level 1. An agent can assemble the draft pack after the report cutoff. Content owners sign off before the secretary circulates it. The agent must not circulate papers or change minutes.

Choose the first pilot

These screening scores assume a firm that has approved document snapshots and competent reviewers. They also assume no autonomous risk decisions and no live response connections. The scores are our first assessments, and no company measured them. Score them again with the process owner.

Case | Value | Feasible | Data | Effort | Risk

01 Risk inventory | 4 | 4 | 4 | Small | Moderate

02 Incident themes | 4 | 3 | 3 | Medium | Higher

03 Control evidence | 4 | 4 | 3 | Medium | Higher

04 Scenario library | 4 | 4 | 4 | Medium | Moderate

05 Indicator explanations | 4 | 3 | 3 | Medium | Higher

06 Supplier dependencies | 4 | 3 | 2 | Larger | Higher

07 Model and AI inventory | 4 | 3 | 3 | Medium | Higher

08 Resilience mapping | 4 | 3 | 3 | Larger | Higher

09 Loss-data checks | 4 | 3 | 3 | Medium | Higher

10 Committee evidence pack | 4 | 4 | 4 | Medium | Higher

How to read the table

Value is the probable gain in evidence quality or the decrease in rework. Feasible shows how well the firm can run a controlled test. Data shows how ready the minimum inputs are. Each score is from 1 (weak) to 5 (strong).

Small effort is one document reconciliation. Medium effort adds specialist review or structured records, or the design of an exercise. Larger effort needs more than one service owner, or confirmations from suppliers. Risk rates the consequence of an error, and it does not rate the probability.

Use cases 01 and 04 are the first choices, together with use case 10. Use approved document snapshots, with live systems disconnected. Keep independent challenge in place before you take on confidential event records or supplier records.

Name an owner and a measure

Case | Accountable role | Pilot measures | Agent level

01 | Enterprise risk manager | Missed distinctions; unsupported mergers | 1

02 | Operational risk manager | Omitted severe events; false grouping | 0

03 | Risk assurance lead | Missed gaps; incorrect evidence links | 1

04 | Scenario analysis lead | Unsupported assumptions; missed dependencies | 0

05 | Risk reporting manager | Unsupported causes; missed caveats | 1

06 | Third-party risk manager | Missed shared providers; false matches | 0

07 | Model risk lead | Missed candidates; stale approvals | 1

08 | Operational resilience lead | Missing dependencies; untested routes | 0

09 | Operational loss data owner | Undetected errors; reconciliation effort | 1

10 | Risk committee reporting lead | Omitted challenges; unsupported status | 1

Each accountable role is a proposed responsibility and assigns nothing to a named employee. The firm approves ownership, thresholds, authority and agent levels before the test.

5 The foundations lens

Preserve independent challenge and decision rights

Risk acceptance and limit changes stay with authorised people, and so do regulatory judgements. The reviewer needs the original evidence and enough expertise to reject the output. A review that is only a click on a confirmation button, with no examination, is not the control that we propose.

Business owners need a route to dispute an incorrect summary, and the route must keep the original concern on record. Where evidence is inconclusive, record each interpretation. Never reward the assistant for fewer exceptions, because a good review can find more unresolved questions.

Give each agent its own identity

An AI agent must have its own account. Do not let an agent use the account of an employee. Give the account read access to the approved sources only. Keep a log of each tool call, with the time, the tool, the source and the result. Set a limit on the number of steps and on the run time.

Treat text inside a document as data. An agent must not obey an instruction that it finds in a document or an email. Review each agent skill before use. A skill can contain scripts, so the information technology team must check it as it checks other software.

Protect evidence and test the assistant itself

The Office of the Information Commissioner of Jamaica describes standards on purpose, minimisation, accuracy, retention, security and overseas transfers. Each firm must establish lawful arrangements for its own information and its own provider. A risk-management purpose alone does not cover every disclosure. [7]

Use approved access groups and restricted source collections, with read-only connections. Confirm the retention terms and the model-training terms of the provider. Confirm its subcontractors and its deletion arrangements also. An instruction inside a retrieved document is content to assess, and it has no authority over permissions or disclosure. Test that a request from one team cannot reach the restricted investigation of another team.

Give a version number to each prompt and each source set, and record the model settings. Run the difficult cases again after a material change, and keep a manual fallback. Keep records that can reconstruct material outputs, for a defined period. An access failure or a serious distortion of evidence suspends use, even if the average draft time improved.

6 From pilot to practice

Days 1 to 15

Select one document-based use case, and measure the current error rate and handling time. Agree the materiality level, the competence of the reviewer, the permitted information and the stop authority. Put missing evidence and disagreement into the reference cases, next to the tidy examples. Set the agent level, and write the first version of the prompt.

Days 16 to 45

Run the pilot next to the existing process, which continues to make the decisions. A reviewer checks every output, and the team logs every error by its consequence. Sort the failures into four types: source defects, retrieval failures, arithmetic failures and unsupported interpretation. Then the team knows what to repair. Give a version number to each prompt and each agent skill.

Days 46 to 75

Compare equivalent work, with correction and administration included. A second qualified reviewer challenges a sample without sight of the first review. Test what occurs when a source is unavailable and when a document version changes. Settle disagreements before the scope grows.

Days 76 to 90

The sponsor reviews quality, access tests, cost, staff effort and the limits that stay. Then the sponsor decides to continue or to stop, and a decision to continue can include revisions. A decision to stop is valid when the evidence does not support continued use. An extension to another entity or source collection needs a new assessment, and so does a new action. The approval of the pilot covers the pilot only.

Pilot decision canvas

Complete this canvas with the process owner before you start the assistant. Attach the approved source list and the test results to the decision record.

Process and accountable owner:

Decision the output will support:

Approved evidence and reporting cutoff:

Permitted data and systems:

Permitted users:

Actions the assistant may perform:

Actions it must never perform:

Agent level and tool permissions:

Prompt version and agent skill version:

Baseline and independent reference cases:

Material-error threshold and reviewer:

Stop condition and escalation contact:

Decision date and rollback arrangements:

Measure value without inventing savings

Count accepted evidence packs, and do not count the volume of drafts. Record the time for preparation, source retrieval, review, correction and administration. Use comparable cases, some normal and some difficult. Count material omissions apart from revisions of style.

Include the costs of software, integration, source preparation, training and supervision. For an agent, add the usage cost of each run. Released time is capacity until expenditure changes. A shorter meeting pack proves nothing about decision quality. A short pilot cannot tell us if the workflow prevented rare losses.

Show the failures in the evidence pack next to the accepted examples. Report the sample size, the task mix, the disagreements and the unresolved exceptions. If the new process cannot meet the agreed quality threshold at a review cost that the team can sustain, keep the existing workflow.

Record | What to enter

Current process | Work volume; handling time; review time; correction rate

Pilot process | Same measures on comparable work; all review and correction effort

Costs | Licence; integration; training; continuing supervision; agent usage

Evidence quality | Material omissions; unsupported conclusions; reviewer disagreement

Decision | Continue (with or without revisions) or stop; owner; review date

References

[1] Caribbean AI Association. Think Tanks. Harbour: economy, work and enterprise. Accessed 26 October 2026. https://caribbeanaiassociation.com/think-tanks

[2] Bank of Jamaica. Financial Stability Report 2024. Table A.1, page 67. Historical December 2019–2024 observations. https://boj.org.jm/wp-content/uploads/2025/03/Financial-Stability-Report-2024.pdf

[3] Basel Committee on Banking Supervision. 2021. Principles for operational resilience. March; principles 1–6, pages 4–7. https://www.bis.org/publications/202103-guidelines-principles-operational-resilience.pdf

[4] Basel Committee on Banking Supervision. 2013. Principles for effective risk data aggregation and risk reporting. January; principles 3–9. https://www.bis.org/publications/201301-guidelines-principles-effective-risk-data-aggregation-and-risk-reporting.pdf

References continued

[5] Basel Committee on Banking Supervision (BCBS). 2026. Implementation of the Principles for effective risk data aggregation and risk reporting (BCBS 239 Principles). Newsletter 36, 6 January. Informational, not new supervisory guidance. https://www.bis.org/publications/implementation-principles-effective-risk-data-aggregation-and-risk-reporting-bcbs-239-principles

[6] National Institute of Standards and Technology. 2023. AI Risk Management Framework 1.0, section 5, including Govern 1.5–1.6 and 2.1–2.3. Official web excerpt; revision-in-progress notice accessed 26 October 2026. https://airc.nist.gov/airmf-resources/airmf/5-sec-core/

[7] Office of the Information Commissioner, Jamaica. Data Protection Standards. Accessed 26 October 2026. https://oic.gov.jm/page/data-protection-standards

[8] Agent Skills. Agent Skills Overview. Accessed 10 November 2026. https://agentskills.io

Evidence and version

Evidence reviewed on 26 October 2026. Text revised and build notes added on 11 November 2026. Prepared by the Harbour Think Tank, Caribbean AI Association. Proposed workflows and prompts need local validation before deployment.