



AI for finance and accounting

Improving controlled financial workflows without
weakening review and auditability.

WORKS WITH THE ASSISTANTS TEAMS ALREADY USE



EDITION 10 / EMBER & CORAL

EXECUTIVE WORKING EDITION - 08 OCT 2026

10

AT A GLANCE

Improving controlled financial workflows without weakening review and auditability.



Finance runs on accuracy, evidence, and segregation of duties, which makes it both a strong candidate for AI and an unforgiving one. AI can speed up close support, variance commentary, invoice handling, reconciliations, and management reporting, but a probabilistic tool must never weaken the controls that auditors and boards rely on. This playbook explains where AI fits in finance, how to keep preparer and reviewer roles intact, how to evidence what the tool did, and how to measure benefit against cost of control.

WHO IT IS FOR

CFOs, controllers, heads of financial planning and analysis, accounts payable and receivable managers, internal audit, and finance transformation leads deciding how to use AI inside controlled processes.

TIME TO USE

60 minutes to read; one working session to apply



What you will have at the end

- A map of finance tasks by suitability for AI, with calculation and judgment tasks clearly separated
- A control design that preserves segregation of duties, review, and an audit trail
- A test and measurement approach that uses closed periods as ground truth
- A 90-day plan to pilot one or two workflows with finance and audit agreement

Contents

01	AI in finance: assist the preparer, never replace the control	03
02	Use cases, fit, and tasks to keep away from AI	04
03	How AI reads and drafts inside finance workflows	06
04	Roles, segregation of duties, and ownership	08
05	Accuracy, fraud, confidentiality, and evidence	09
06	Baseline, metrics, and tests against closed periods	11
07	A controlled path from pilot to production	12
08	Common failures, an illustrative scenario, and vendor questions	14
09	A first quarter that produces evidence	15
10	Terms used in this publication	16



EXECUTIVE BRIEF

AI in finance: assist the preparer, never replace the control

01



The value of AI in finance is faster preparation and better first drafts. The control framework that makes numbers trustworthy must stay exactly as strong as before.

Finance teams spend large amounts of time collecting data, matching documents, explaining movements, and writing commentary. Much of this is reading and drafting, which suits language models. But finance also demands exact arithmetic, consistent treatment, and a traceable path from a reported number back to source evidence. Language models are not calculators and do not guarantee consistency. The design principle is therefore simple: use AI to read, classify, draft, and flag, and use deterministic systems and people to calculate, post, approve, and certify.

A second principle follows from the first. Whoever prepares a number should not be the only person who approves it, and the same applies when the preparer is a tool. An AI-drafted variance explanation or a suggested journal entry is a preparer output. It needs an independent human reviewer who understands the underlying transaction and can reject the suggestion. Reviewers who rubber-stamp fluent text are the main new risk, and the controls in this playbook are designed to prevent it.

Where finance teams start

- Variance analysis: drafting first-pass explanations of budget and prior period movements from ledger detail
- Close support: checklists, status summaries, and anomaly flags for reconciliations and accruals
- Accounts payable: extracting invoice data, suggesting coding, and flagging duplicates and mismatches
- Management reporting: drafting narrative around approved numbers for monthly packs and board papers
- Policy and technical research: summarizing accounting guidance for a human accountant to interpret

WHAT STAYS DETERMINISTIC

Calculations, ledger postings, consolidations, tax computations, and payments should be performed by rules-based systems with tested logic. AI may prepare inputs or explain outputs. It should not be the system that produces the figure.

WORKING QUESTION

Which of your finance outputs would an external auditor expect to trace line by line to evidence, and could you do that if AI had touched it?

WHERE IT FITS

Use cases, fit, and tasks to keep away from AI

02

Fit depends on how checkable the output is and how much judgment or authority the task carries.

A useful sorting question is whether the task is about language or about numbers and authority. Drafting a paragraph explaining why freight costs rose, using a table the controller supplies, is a language task. Calculating accrued freight, deciding whether it meets recognition criteria, and posting it are number and authority tasks. AI can help with the first. The second belongs to systems and accountable people, even if AI proposes inputs.

USE CASE	FIT	MAIN RISK AND CONTROL
Variance commentary drafts from supplied data	Strong	Plausible but wrong causes. Reviewer confirms with budget owners; AI cites data used
Invoice data extraction and coding suggestions	Strong	Misread amounts or wrong coding. Validate against purchase orders; confidence thresholds

USE CASE	FIT	MAIN RISK AND CONTROL
Reconciliation exceptions: matching and flagging	Strong	Missed breaks. Keep existing reconciliation control; AI adds a second look
Management report narrative	Moderate	Spin or omissions. Narrative is checked against final numbers by a named reviewer
Journal entry suggestions	Moderate	Wrong accounts or periods. Suggest only; standard approval workflow applies
Accounting policy and standards research	Moderate	Misstated guidance. Qualified accountant verifies against the source text
Calculations, tax positions, payments, and approvals	Avoid	Use deterministic systems and authorized people

Poor candidates

- Any step that releases funds or changes bank details without independent human verification
- Estimates and judgments such as impairment or provisions, where AI may draft options but management owns the conclusion
- Statements to auditors, regulators, or lenders that have not been verified by a person
- Fraud conclusions about individuals, which should be investigated by trained people with due process
- Anything that would put non-public financial information into an unapproved tool

CHECK YOUR REPORTING AND AUDIT OBLIGATIONS

Listed companies, regulated entities, and those with lender covenants have specific obligations on internal control and disclosure. Brief your external auditor early on where AI is used, and check current requirements for your jurisdiction and framework with qualified advisers.

WORKING QUESTION

Which task in your close calendar consumes the most analyst hours while carrying the lowest judgment content?

HOW IT WORKS

How AI reads and drafts inside finance workflows

03

In finance the safest architecture separates reading and drafting, done by AI, from calculating and posting, done by controlled systems.

Document intelligence tools read invoices, contracts, and statements and extract fields such as supplier, date, amount, and tax. Language models take structured data and produce text such as explanations and summaries. Matching and anomaly tools compare transactions with expected patterns. In every case the output is a suggestion with a degree of confidence. The surrounding workflow decides what happens next: auto-accept above a threshold, route to a human below it, or reject.

01 Ingest

Documents and ledger data arrive from approved sources. Access is limited by role, and sensitive fields are protected.

02 Extract or analyze

The tool pulls out fields or prepares an analysis, and attaches a confidence level and a link to the evidence.

03 Validate with rules

Deterministic checks compare results with purchase orders, vendor master data, totals, tax logic, and period rules.

04 Human review

Exceptions and low-confidence items go to an accountable person. The reviewer sees the source and can edit or reject.

05 Post and log

The accounting system posts under existing approval rules. The tool's inputs, outputs, and any human changes are logged.

Why explanations need data attached

When a model is asked why revenue rose, it will give a reason whether or not it knows one. A good design gives the model the transaction-level data and asks it to describe what the data shows, such as which customers and products drove the movement, and to say clearly where the cause cannot be determined from the data. The business reason, such as a price change or a lost contract, comes from the budget owner and is added by a person. This keeps commentary tied to evidence and avoids invented narratives.

ARITHMETIC WARNING

Language models can make arithmetic and rounding mistakes. Do not ask a model to total, convert currency, or compute percentages directly. Calculate in a spreadsheet or the ledger, then give the model the results to describe.

WORKING QUESTION

At which step in your current process does a human actually look at the evidence, and would that step survive faster volumes?



OPERATING MODEL

Roles, segregation of duties, and ownership

04

Assigning the AI tool a place in the control framework is the central design task. Treat it as a preparer with a documented scope.

Finance already has a vocabulary for this: preparer, reviewer, approver, and system owner. Apply it. The AI tool is a preparer whose work product is always reviewed by someone with the authority and competence to disagree. The tool itself has an owner, typically in finance systems, who is responsible for configuration, access, change control, and monitoring. Internal audit should see the design before go-live, not after, because retrofitting evidence is costly.

ROLE	OWNS	DECIDES
CFO	Control environment, risk appetite	Which processes use AI and the evidence standard
Controller	Close process, accounting policies	Review requirements and sign-off for AI-assisted outputs
Process owners (AP, AR, FP&A)	Day-to-day workflow quality	Exception handling, thresholds, and acceptance of results
Finance systems owner	Configuration, access, change control	What the tool can read and write, and release of changes
Internal audit	Independent assurance	Testing approach and reporting of control gaps
Data protection and security	Data handling and vendor assurance	Allowed data, retention, and approved vendors
External auditor (informed)	Audit opinion	Whether evidence supports reliance; engage early

Preserve segregation of duties

- The person who approves a payment or journal cannot be the person who configured the tool that suggested it
- The tool should not have posting rights beyond what a preparer role would have

- Changes to prompts, thresholds, or models follow the same change control as changes to financial systems
- Access to the tool is reviewed on the same cycle as access to the ledger
- Reviewers have time to review, and volume targets do not push them to approve unread

Document each AI-assisted process in the same format as other key controls: objective, scope, inputs, tool role, reviewer, frequency, evidence, and failure handling. This gives auditors something familiar to test and gives your team a clear picture of what the tool is and is not trusted to do.

WORKING QUESTION

If an auditor asked who is responsible for a specific AI-drafted journal entry, could you name the reviewer and show their evidence?

RISKS AND CONTROLS

Accuracy, fraud, confidentiality, and evidence

05

Finance risks are concentrated in wrong numbers, manipulated inputs, leaked non-public information, and missing evidence.

Financial data is among the most sensitive a company holds. Unreleased results, forecasts, deal terms, bank details, and payroll all pass through finance. Prompts can contain this data, outputs can summarize it, and logs can retain it. Decide which classes of data may go to which approved systems and block the rest. Equally, finance is a target for fraud, and AI both creates new attack routes, such as convincing fake invoices and impersonation, and offers new defenses, such as anomaly detection.

RISK	EXAMPLE	CONTROL
Misstatement	Wrong coding or amount accepted without review	Validation rules, thresholds, independent review, sampling

RISK	EXAMPLE	CONTROL
Fraud and impersonation	Fake invoice or change of bank details that looks authentic	Callback verification, vendor master controls, human approval of changes
Over-reliance	Reviewers approve fluent text without checking evidence	Evidence links, sampling of approved items, review time standards
Confidential information	Unreleased results entered into unapproved tools	Approved tools, data classification, retention limits
Audit trail gaps	No record of prompts, outputs, or changes	Logging, version control, retention aligned to record policy
Model change	Vendor update changes behavior the middle of close	Pinned versions where possible, change notices, regression test set
Manipulated documents	Hidden text in a PDF instructs the tool to approve it	Input validation, restricted tool permissions, human exception review

Control evidence to retain for each AI-assisted process

- A description of the process, the tool's role, and the reviewer
- Inputs and outputs for each run, linked to the source documents
- The reviewer's decision and any changes they made
- Test results for the current configuration, with the date
- Change records for prompts, thresholds, models, and access
- Incident and exception logs with resolution

Frameworks such as the NIST AI Risk Management Framework and ISO/IEC 42001 can help you structure governance, and established internal control frameworks apply to AI-assisted processes as they do to any other. Accounting, audit, tax, and data protection requirements vary by jurisdiction and change over time. Consult qualified advisers for your specific obligations.

WORKING QUESTION

If a vendor updated its model overnight during month-end close, how would you find out and what would you do?

MEASURING VALUE

Baseline, metrics, and tests against closed periods

06



Finance has an advantage: it has ground truth. Closed periods show what the right answer was, so you can test tools before you trust them.

Run the tool on a closed period in parallel and compare its outputs with the numbers and commentary that were actually approved. For invoice extraction, compare extracted fields with the verified ledger entries. For variance commentary, have the budget owner grade whether the described drivers match what really happened. For reconciliation tools, seed known breaks into test data and see how many are caught. This gives a defensible accuracy figure before any live reliance.



METRIC	HOW TO MEASURE	WATCH FOR
Field-level extraction accuracy	Extracted values vs verified entries by field and supplier	Averages that hide a bad supplier or document type
Straight-through rate with zero errors	Share accepted automatically without later correction	Rate pushed up by loose thresholds
Review corrections	Share and type of items changed by reviewers	Very low corrections can mean rubber-stamping
Close cycle time	Working days to close and time to first draft of commentary	Savings erased by rework
Error and restatement rate	Post-close adjustments linked to AI-assisted steps	The key assurance metric
Exception ageing	Days unresolved for flagged items	Flags ignored because of volume
Cost per transaction and per close	Bully loaded labor, tool, and control cost	Leaving control cost out of the case



Set a pass mark before you look

Agree success criteria with the controller and internal audit before running tests. Illustrative criteria might be: no unreviewed item posted with a material error, field accuracy above an

agreed level for key fields such as amount, date, and supplier, and every variance explanation traceable to supplied data. Choose your own levels and document them. Retest after every change in model, prompt, supplier mix, or document format, and keep a regression set of past mistakes that must never recur.



ILLUSTRATIVE EXAMPLE, LABELED AS SUCH

A team measures that AI cut invoice keying time by half, but post-close corrections rose in one supplier group with unusual layouts. The overall saving was real, and the supplier-level view showed where to lower the confidence threshold and keep manual entry.



WORKING QUESTION

Which closed period would be your best test bed, because you already know what the correct answers were?



IMPLEMENTATION GUIDE

A controlled path from pilot to production

07

Pilot in parallel, with audit in the room, and move to reliance only when the evidence supports it.

01 Select one process

Pick a workflow with high volume, clear ground truth, and moderate risk, such as invoice extraction for a defined set of suppliers or draft variance commentary.

02 Document the control design

Write the process, the tool's scope, the reviewer, the thresholds, and the evidence to be kept. Share it with internal audit and, if relevant, the external auditor.



03 Run in parallel

Operate the tool beside the existing process for one or two cycles. Compare outputs with the approved results and record differences.

04 Go live with limits

Start with tight thresholds, human review of all exceptions, and sampling of accepted items. Keep a manual fallback ready.

05 Monitor and tune

Review accuracy, corrections, and exceptions each close. Adjust thresholds and instructions under change control.

06 Extend only on evidence

Add suppliers, entities, or processes one at a time, repeating the test and documentation steps.

Go-live readiness

- Internal audit has reviewed the design and test results
- Reviewers are trained, and review time is built into the close calendar
- A manual fallback can run if the tool is unavailable or fails a check
- Access is limited by role and reviewed
- Data handling, retention, and vendor terms are approved by security and data protection
- Month-end freeze rules cover prompts, thresholds, and vendor updates
- Avoid introducing AI in a quarter-end or year-end close for the first time
- Train reviewers to look for omitted items, wrong periods, and unsupported explanations, not just typos
- Keep a log of overrides so that patterns in mistakes turn into rule changes
- Brief the audit committee on scope, controls, and results in plain language
- Review the use cases annually against changes in tools, standards, and your own risk profile

WORKING QUESTION

Which control in your close would you be happiest to see strengthened by a second automated look, without changing who is accountable?

PITFALLS AND BUYING

Common failures, an illustrative scenario, and vendor questions

08

Failures in finance usually come from weak review, weak evidence, or unclear scope, rather than from model quality alone.

- Letting the model do arithmetic, so totals and conversions drift in subtle ways
- Treating confident commentary as verified fact and publishing unexplained causes
- Giving tools broad write access to ledgers or payment systems
- Skipping early conversations with internal and external audit
- Counting time saved while ignoring the cost of review, monitoring, and rework

ILLUSTRATIVE SCENARIO, NOT A REAL COMPANY

A distribution business uses AI to draft monthly variance commentary. Reviewers, short of time, accept the text. One explanation attributes a margin drop to a supplier price rise when the real cause was a mispriced batch of orders. The company now requires the budget owner to confirm each driver and attaches the supporting data to the pack.

Questions to ask any vendor

Vendor diligence

- Where are our financial documents processed and stored, and are they used to train shared models
- Which security and assurance reports can you provide, and how recent are they
- What logs show inputs, outputs, confidence, and human changes, and can we export them for audit
- How are model and software changes announced, and can we hold a version during close
- What role-based access and segregation controls exist
- How does the product defend against manipulated or fraudulent documents
- What are the data return, deletion, and exit terms

Ask each vendor to run your closed-period test set, including awkward documents and known fraud patterns, and to show how the tool signals uncertainty. Prefer a product that clearly marks what it does not know over one that always produces a clean answer. Involve internal audit in the selection, since they will have to test the result.



WORKING QUESTION

Where in your finance function would a fluent but wrong explanation do the most damage if no one challenged it?

30 / 60 / 90

A first quarter that produces evidence

09



Pilot one controlled workflow in parallel with audit involved, then extend only as evidence and control design hold.

First 30 days: ground it

- Choose one workflow with clear ground truth, such as invoice extraction or draft variance commentary
- Document the control design and share it with internal audit
- Classify financial data and set rules for what may enter approved tools
- Capture baseline cycle time, error rates, and cost for the chosen process



Days 31 to 60: test it

- Test the tool on a closed period and grade results against approved outputs
- Run in parallel for one or two cycles and log every difference
- Train reviewers on evidence checks and set review time in the close calendar
- Agree pass marks and fallback arrangements with the controller and audit



Days 61 to 90: decide

- Go live with tight thresholds, sampling, and a manual fallback
- Review accuracy, corrections, and post-close adjustments after the first live close
- Brief the audit committee and, where relevant, the external auditor
- Decide which supplier, entity, or process to add next, if any

GLOSSARY

Terms used in this publication

10

TERM	MEANING
Accrual	An accounting entry that recognizes an expense or revenue in the period it relates to, before cash moves.
Audit trail	A traceable record showing how a figure or entry came about, who handled it, and what evidence supports it.
Confidence threshold	A set level of certainty below which a tool must send an item to a human.
Document intelligence	Tools that read documents such as invoices and extract structured fields from them.
Internal control	A policy or procedure designed to give reasonable assurance about reliable reporting, effective operations, and compliance.
Month-end close	The recurring process of completing and reviewing the books for a period so that results can be reported.
Reconciliation	Comparing two sets of records, such as the ledger and a bank statement, to confirm they agree and explain differences.
Regression test set	A fixed group of past cases, including earlier mistakes, rerun after every change to check that quality has not slipped.
Segregation of duties	A control principle that no single person or system can both initiate and approve a transaction.
Straight-through processing	Handling a transaction end to end automatically, with no manual touch.

TERM	MEANING
Variance analysis	Explaining the differences between actual results and a budget, forecast, or prior period.
Vendor master	The controlled record of supplier details, including bank accounts, used to pay invoices.



A note on use

Free assessments are available at halden.company/assessments, and you can reach the team at halden.company/contact. Figures here are illustrative, and nothing in this playbook is accounting, tax, audit, or legal advice.

TAKE THE NEXT STEP

Take the free Halden assessment at halden.company/assessments, or talk to us at halden.company/contact about turning this into a plan for your organization.



