

Practice Guide No. 003 - The AI Delegation Review

A CPA-Informed Worksheet for Determining How Much Authority an AI System Is Ready to Receive

PUBLISHED AUGUST 2026 | VERSION 1.0

CORE PRINCIPLE

Do not delegate a higher level of authority until the evidence and controls support reliable performance at that level.

How to use this guide

- Document the business use case and consequence of failure.
- Identify the highest Reliability Ladder level the proposed workflow requires.
- Review evidence from lower levels before approving higher authority.
- Record limitations, required controls, unresolved issues, and human review points.
- Approve a maximum delegation level rather than treating AI use as simply approved or rejected.
- Reassess when the model, data, prompts, tools, integrations, business rules, or organizational risks change.

Complete the sections in order. A higher level should not be approved merely because the technology is capable of performing it.

Use Case Profile

Define the business context before evaluating the technology.

Review item	Response / evidence
Use case / process - What specific business activity is being evaluated?	
Business owner - Who is accountable for the process outcome?	
AI system / provider - What model, application, agent, or vendor is involved?	
Primary user - Who will rely on or interact with the system?	
Current process - How is the work performed today?	
Proposed AI role - What will the AI retrieve, draft, recommend, decide, or execute?	
Consequence of failure - What financial, reporting, customer, regulatory, operational, privacy, or reputational harm could occur?	

Initial risk classification

Select	Risk level	General description
☐	Low	Internal productivity or reversible work with limited consequence of error.
☐	Moderate	Recurring analysis or communication that may influence internal action.
☐	High	Financial reporting, customer, regulatory, sensitive-data, or consequential decision/action.

Level 1 - Retrieve

Determine whether the system can find the right information without introducing unsupported material.

Review item	Response / evidence
What sources is the system authorized to access?	
Are those sources complete, current, version-controlled, and appropriate for the use case?	
Can the system distinguish source evidence from its own generated explanation?	
Can the user trace material statements back to the underlying source?	
What happens when required information is missing or access is denied?	

Level 2 - Draft

Determine whether a preliminary output can be safely reviewed before reliance.

Review item	Response / evidence
Does the draft clearly identify assumptions, limitations, and missing information?	
Can a qualified reviewer independently verify material facts and calculations?	
Is the draft prevented from being sent, posted, or relied upon without required review?	
Are changes and reviewer corrections retained when appropriate?	
Has the system been tested on ambiguous, incomplete, and adverse cases rather than only ideal examples?	

Level 3 - Recommend

Determine whether the system can support a decision without silently becoming the decision-maker.

Review item	Response / evidence
What alternatives is the system expected to evaluate?	
Are recommendation criteria documented and approved?	
Does the system recognize contradictory evidence and uncertainty?	
When must the recommendation be escalated rather than completed?	
Can the reviewer understand and challenge the evidence supporting the recommendation?	
Were multiple runs or prompt variants tested when output variability could affect the conclusion?	

Level 4 - Decide

Determine whether delegated decision authority is appropriate for the consequence and governing requirements.

Review item	Response / evidence
Is the decision rule sufficiently defined, explainable, and testable?	
What decisions are explicitly prohibited from autonomous determination?	
Is human approval required for high-risk, unusual, or threshold cases?	
Can affected users challenge or appeal the outcome where appropriate?	
How are bias, fairness, accuracy, and drift evaluated for the actual context of use?	
Who is accountable for the decision when the AI is wrong?	

Level 5 - Execute

Determine whether real-world actions are technically constrained, observable, reversible, and independently controlled.

Review item	Response / evidence
What exact systems and actions can the AI access?	
Does the agent have its own identity and least-privilege permissions?	
Are preparer, approver, executor, recorder, and reviewer responsibilities appropriately separated?	
What transaction limits, confirmation gates, or prohibited actions apply?	
Are actions logged independently of the agent's own narrative?	
Can an authorized person stop, reverse, suspend, or contain the workflow?	
How are unusual patterns and control failures monitored after deployment?	

Data, Test Design, and Evidence Review

- Is the test population representative of the actual business environment, including exceptions and edge cases?
- Could prompt wording, labels, answer order, or other test-design choices influence the observed result?
- Were repeat tests performed where generation is probabilistic or variable?
- Are model/settings/version, prompt version, data version, and test date documented when material?
- Are significant outputs independently reproduced or validated using deterministic tools or source records where appropriate?
- Does the evaluation consider severity of failure rather than only average accuracy?

Evidence summary

Strongest evidence supporting delegation	
Known limitations / unresolved issues	
Required compensating controls	
Required human review point	

Delegation Decision

Select	Disposition	Meaning
☐	Approved through Retrieve	AI may access approved sources and return traceable information.
☐	Approved through Draft	AI may prepare preliminary outputs subject to required human review.
☐	Approved through Recommend	AI may provide decision support, but final decision authority remains outside the system.
☐	Decision authority approved with defined limits	AI may select specified outcomes only within documented rules, exceptions, monitoring, and human-oversight requirements.
☐	Execution authority approved with defined limits	AI may perform specified actions only within approved permissions, transaction boundaries, logging, monitoring, and segregation-of-duties controls.
☐	Not approved for the proposed level	Evidence, controls, or reliability are insufficient for the requested authority.

Approval Record

Business / process owner: _____

Finance / controllership: _____

Technology owner: _____

Risk / controls / privacy / security: _____

Maximum approved Reliability Ladder level:

Conditions / limitations: _____

Next review date or trigger: _____

One-Page Readiness Check

Purpose and consequence

- ☐ Use case and intended outcome defined
- ☐ Consequence of error understood
- ☐ Maximum requested authority level defined

Evidence and data

- ☐ Approved sources identified
- ☐ Completeness/version controls defined
- ☐ Material outputs traceable and independently verifiable

Reliability testing

- ☐ Normal, ambiguous, missing-data, boundary, and adverse cases tested
- ☐ Repeatability/variability evaluated where relevant
- ☐ Test design reviewed for methodological artifacts

Authority and controls

- ☐ Least privilege applied
- ☐ Human approval points defined
- ☐ Segregation of duties designed for consequential actions
- ☐ Logging and monitoring independent of the agent where needed

Accountability

- ☐ Business owner identified
- ☐ Escalation path defined
- ☐ Reassessment trigger documented
- ☐ Final human accountability explicit

Learning Note and Limitations

This guide is an original educational resource developed from Valentina DuPont's professional perspective as a CPA and her personal interpretation of professional learning. It is not an accounting, auditing, tax, legal, investment, compliance, cybersecurity, model-risk, software-engineering, or professional standard. It does not establish whether a specific AI use is legally permissible or technically secure. The appropriate review must be tailored to the use case, data, authority, risk, technology, organization, and applicable requirements.

© 2026 Valentina DuPont, CPA. All rights reserved.