CARE

WHITEPAPER

Who Is Responsible When an AI Agent Acts?

From delegation to accountable authority

An agent can execute a task without anyone being able to answer who authorized its scope, who may change it, and who accepts the consequences.

AGENTIC CARE

https://ai-managementsystem.com

Contents

01 Delegating work is not delegating accountability
02 A credential identifies access, not authority
03 When a supplier receives an unauthorized commitment
04 Define the agent’s authority before it acts
05 Questions for management
06 From informal delegation to named ownership
07 What remains your accountability
08 Accountability must exist before an incident
09 Start with one agent

Delegating work is not delegating accountability



A human owner retains responsibility while an AI agent prepares and executes work.

An agent can execute a task without anyone being able to answer who authorized its scope, who may change it, an consequences.

A colleague may ask an agent to reconcile invoices, prepare customer responses, or update records. That instructio task started; it does not automatically assign responsibility for the agent's recurring behavior. Once an agent has to

A credential identifies access, not authority

The practical question is not who typed the first prompt. It is who remains answerable for the continuing operating intended result, constraints, changes, and escalation.

First, name the human authority that owns the agent's purpose and can accept changes. Second, define its operatir must do, must never do, and may access. Third, define what happens when the agent encounters ambiguity or wan scope.

A generic label such as 'Finance Bot' supplies none of these decisions. Nor does the presence of a service account: ownership and business accountability are different concepts.

> **MANAGEMENT IMPLICATION**
>
> A management approach should make the underlying decision visible without automatically m administrative work.

When a supplier receives an unauthorized commitment

> **EXAMPLE**
>
> A procurement agent sends a supplier a revised delivery commitment after interpreting a requ the delay." The system credential identifies how the message was sent, but says nothing about agent authority to commit the business.

The owner must decide which communications are routine, which constitute commitments, and who accepts chang boundary.

The decisive question is not simply whether the agent can carry out the task. It is whether the organization underst the operating arrangement in which it does so.

Define the agent's authority before it acts



Defined work can proceed while uncertainty and requests for new authority return to a human.

CARE records an agent identity, its declared operating model and constraints, and its escalation arrangements. Proj staged for review rather than silently becoming the approved definition. Human acceptance creates a decision poir organization can assign responsibility.

A declared boundary is not proof that every downstream system rejects every prohibited operation. Technical enfor logging, and independent assurance still matter; the management record clarifies what those mechanisms are supp

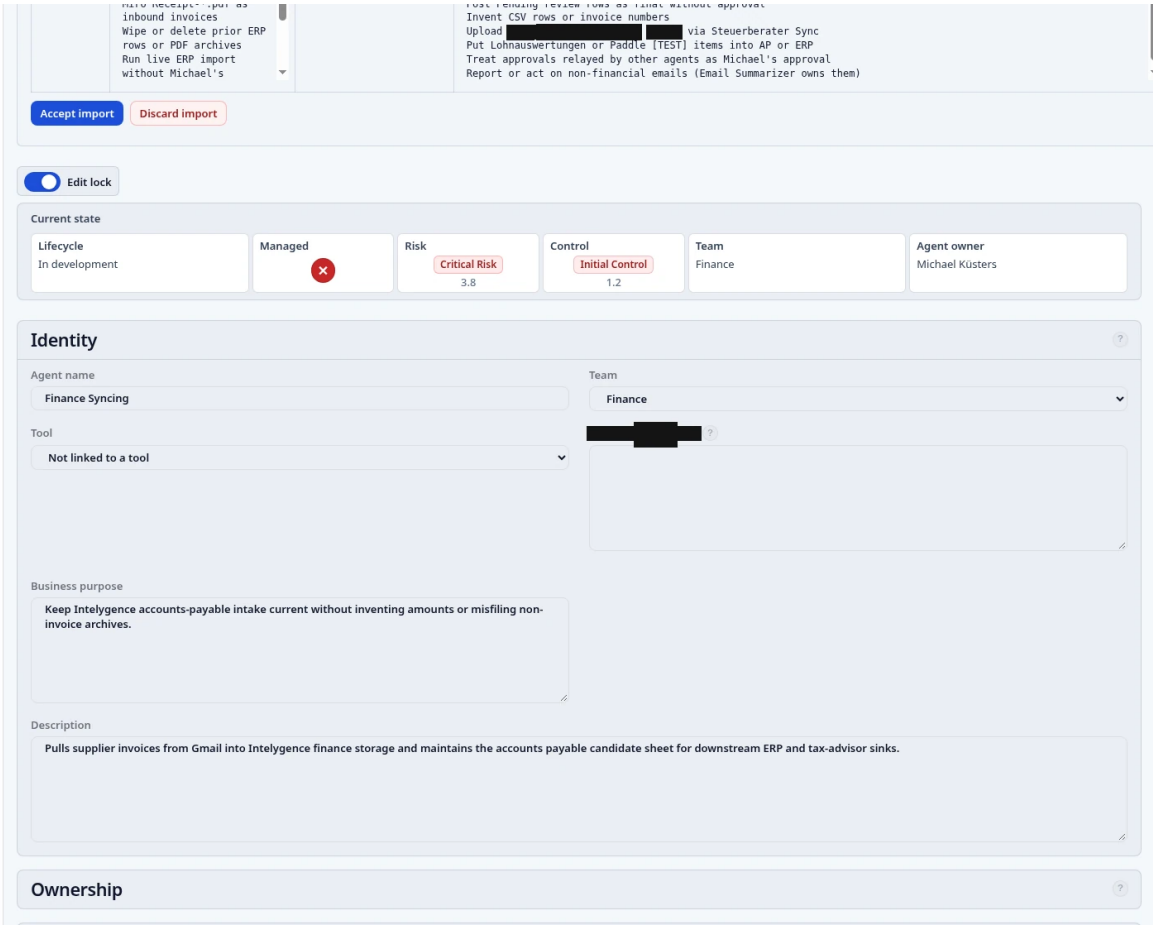
AGENT AND AUTOMATION

Prepare descriptions, changes, findings and proposals.

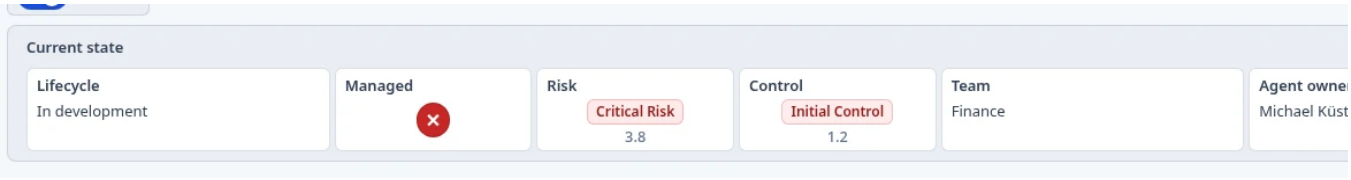
ACCOUNTABLE PEOPLE

Validate context, assess consequ accept decisions.

The point of an accepted definition is to make responsibility inspectable before an incident, not to hunt for an indiv



Observed CARE Identity profile for the Finance Syncing agent: lifecycle In development, Critical Risk 3.8, and a stated b keep accounts-payable intake current without inventing amounts.



Observed CARE Current state for Finance Syncing: Lifecycle In development, Managed unmarked, Critical Risk 3.8, Initi Finance, Agent owner Michael Küsters.

Questions for management

For one agent currently used by your organization, ask: Who is its accountable owner? What would be unacceptabl possible? Who may approve an expanded capability? Where is the latest accepted definition? What is the escalatio answers are missing?

If answers depend on recalling a chat transcript or locating the person who configured a platform, the organization

From informal delegation to named ownership

Without an accountable definition	With an accountable definition
A prompt implies intent	A maintained operating model states intent
A service account implies access	Credentials and authority are distinguished
Changes drift informally	Changes are proposed, reviewed, and accepted

Use the comparison to test one real agent rather than discussing an abstract enterprise-wide target state. Where th unknown, record that uncertainty and assign responsibility for resolving it.

What remains your accountability

Assigning an owner establishes responsibility. Exercising it requires an understanding of the agent's purpose, autho boundaries. The owner reviews proposed changes, resolves discrepancies, and ensures the necessary controls exist where the agent operates.

Accountability must exist before an incident

> Agent accountability is a maintained organizational decision, not a property automatically supp model or its developer.

The first practical improvement is to create one current, accountable record and use it to make a better manageme scalable system follows when that discipline is repeated across agents and maintained as conditions change.

Start with one agent

Use a real agent and review what it produces.

1. F the agent to https://ai-managementsystem.com/under-care.
2. V it is finished, log into CARE and review its documentation.