REVYCORE

WHITE PAPER

Accountable Agents.

Certifying AI Agents as Governed Identities

REVYCORE RESEARCH

DECEMBER 2026

EXECUTIVE SUMMARY

A non-human identity remains a human responsibility.

AI agents have quietly become part of the enterprise estate. They hold or borrow identities, read document sources, call tools, act on behalf of users. They are created in collaboration platforms, cloud services, business applications and development tools, often by teams that move faster than governance.

Most organisations can say what an agent can access. Few can say who answers for it, why it exists, what it was allowed to become, and whether a human ever decided any of it.

This paper proposes to treat each agent as a governed digital entity: described in six dimensions, attested by a named owner, reassessed when it changes, and proven with the same sealed evidence as any other access.

01 Six dimensions
Responsibility, identity, access, knowledge, tools and audience, each attested separately.

02 Two readings, linked
The agent as the subject of an access, and as the object of its own attestation.

03 The fourth question
What it can access, who authorised it, whether a human decided, and whether you can prove it.

CONTENTS

01 Agents are identities that act
02 Where agents live
03 The agent in six dimensions
04 Two readings, linked
05 Access amplification
06 Tools and MCP servers
07 What to look for
08 The fourth question

01

Agents are identities that act

An AI agent is neither a robot nor a black box. From a governance point of view, it is an identity that acts: it signs in with its own credentials, a shared account or the identity of the user who calls it; it reaches data; it triggers actions.

Identity governance already knows how to handle people and service accounts. Agents add three difficulties: they multiply quickly, their capabilities change without anyone noticing, and they can give many people indirect access to information that only a few were meant to see.

Treating agents as ordinary applications misses what makes them different. Treating them as people gives them a responsibility they cannot hold. The answer is to give each agent an accountable human, and to certify what it is.

> **A non-human identity remains a human responsibility.**

02

Where agents live

Agents are scattered across the estate, in places that each have their own administration console and their own vocabulary. Governance has to see them together.

- **Collaboration platforms.** Assistants built on documents, conversations and team spaces, often by business users themselves.
- **Cloud platforms.** Agents and model endpoints deployed by engineering teams, with service identities and tool integrations.
- **Business applications.** Agents embedded in business suites, acting on records, workflows and customer data.
- **Development tools.** Coding agents and automation with access to repositories, pipelines and secrets.

The first step is an inventory that spans all of them: one view of every agent, collected by solution, reconciled with identities and owners.

03

The agent in six dimensions

An agent is described by what it answers for, what it is, where it gets in, what it knows, what it can do and who it talks to. Each dimension is attested separately, and reassessed when it changes.

DIMENSION	THE QUESTION IT ANSWERS
Responsibility	Who answers for the agent, and why does it exist? An owner and a stated purpose.
Identity	What does it act as: its own identity, a shared one, or one borrowed from the user?
Access	Which entitlements does it hold or exercise, and on which resources?
Knowledge	Which document sources and databases does it consult to answer?
Tools	Which actions can it trigger, directly or through the servers that expose them?
Audience	Who can call on it, and has that audience grown since its last attestation?

04

Two readings, linked

An agent appears in governance in two ways, and both are needed.

As the subject of an access, it reaches a resource. Its access appears in the review of that resource's owner, like a user or a service account. The owner of a contracts library certifies that the agent may read it.

As the object of an attestation, the agent itself is certified by its own owner: purpose, identity, sources, tools, audience, authority. The owner certifies that the agent does what it exists for, and nothing more.

The two readings are linked. When an access changes, the agent's attestation is reassessed; when the agent changes, the accesses it relies on are reviewed again. One view of governance, from both sides.

05

Access amplification

Depending on its execution identity and the platform's controls, an agent can create a capacity to amplify access to information. A source readable by a dozen people can become queryable by hundreds, through an agent that answers with what it can see.

This capacity can be calculated: the readers authorised on the source, the identity the agent runs under, and the audience allowed to call it. Presented to the owner, it turns an invisible exposure into a decision to take.

A capacity is not a proven leak. It is a situation that calls for a decision, and that decision belongs to a named person, on a date, with a reason.

> **Twelve people can read the source. Four hundred and eighty can query the agent.**

06

Tools and MCP servers

Agents increasingly reach their tools through MCP servers. An MCP server is a capability surface: it makes tools available to the agents that connect to it. Governing agents means governing what those surfaces expose.

- **The agent and its capabilities.** The tools an agent can invoke are part of its attestation, in the Tools dimension.
- **The shared server.** A server used by several agents is certified as a distinct component, with its own owner, within the technology scope it belongs to.
- **Advertised is not authorised.** A capability declared by a tool and an authorisation actually demonstrated are two different findings, and are reported as such.

07

What to look for

Across every platform, the same six situations deserve attention. Each one is a question for an owner, not an automatic verdict.

01

No owner

The creator has gone, the team has dissolved. Nobody can answer for the agent.

02

No purpose

There is no description of what the agent exists for, so nothing to measure its authority against.

03

Inactive

No request since the threshold set by policy. The agent keeps its access and serves no one.

04

Shared identity

Several agents behind one account: their actions can no longer be attributed.

05

Excessive authority

Tools or entitlements unrelated to the attested purpose.

06

Widened audience

The agent is open to more users than at its last attestation.

08

The fourth question

Most of the market works on the first question. Certification lives on the fourth.

FOUR QUESTIONS

01

What can it access?

The permission graph: crowded, well funded, and not where certification lives.

02

Who authorised it?

A named, accountable owner, with a reason for the identity to exist at all.

03

Did a human decide?

Not a policy, not an automation: a person, on a date, with a comment.

04

Can you prove it?

Six months later, in front of a regulator. Sealed, chained, exportable.

REVYCORE applies this discipline to its own assistant first: for every exchange, a register records what the assistant actually did, with which tools and under which session. Not what it was allowed to do, but what it did, and under whose authority.

REVYCORE

Every agent,
an accountable owner.

Human decisions. Verifiable evidence.

Request a demonstration in your context: your solutions, your scopes, your agents.

revycore.com

This paper describes an approach and the corresponding capabilities of REVYCORE for information purposes; it is not legal advice. Figures quoted are illustrative. © 2026 REVYCORE SAS.