akto

AI Agent Identity Security Maturity Model



The identity risks autonomous AI agents create, and a practical path from standing, ungoverned credentials to fully governed agent identities managed at machine speed.

THE PROBLEM

Every AI agent is an identity you don't control

Autonomous agents make decisions and take actions at machine speed, and every one of them acts through an identity: API keys, bearer tokens, OAuth grants, and cloud roles. Traditional IAM was built for humans and known service accounts, so these identities go unseen, over-scoped, and unowned.

50:1+

Non-human identities already outnumber humans, and AI is widening the gap fast.

No owner

Most agents run with no accountable owner and no audit trail of what they touched.

No MFA

Human MFA does not apply to agents, so most default to static secrets.

Top AI agent identity-centric risks

1. **Orphaned and unmanaged identities**
   Agents outlive their purpose & keep access no one retires, becoming quiet backdoors.
2. **Excessive permissions and privilege creep**
   Broad, inherited access grows unchecked and breaks least privilege.
3. **Static credentials and weak authentication**
   Long-lived keys and hard-coded secrets rarely rotate and cannot use MFA.
4. **Identity spoofing and impersonation**
   Weak verification between systems lets an attacker act as a trusted agent.
5. **Lack of traceability and auditability**
   Thin logging makes it impossible to tell what an agent did, or why.
6. **Inadequate behavioral monitoring**
   Tools built for humans miss anomalous machine behavior with no agent baseline.
7. **Non-human identity and secrets sprawl**
   Tokens, keys, and accounts multiply faster than anyone can manage by hand.
8. **Prompt injection and malicious instructions**
   Crafted input manipulates an agent into unauthorized actions.

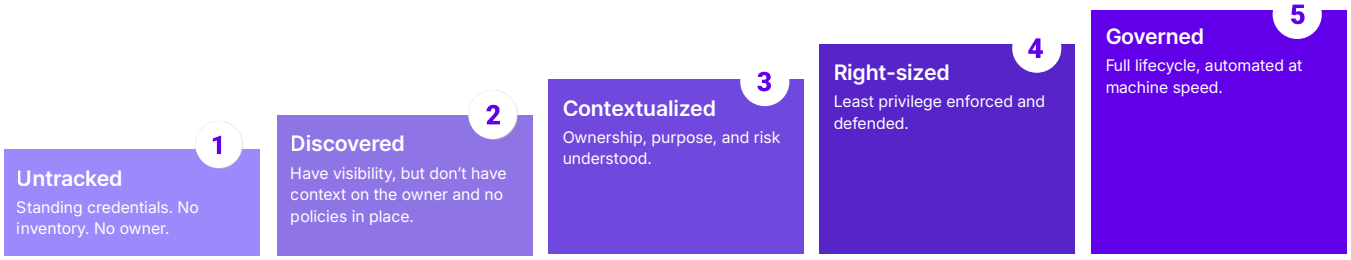
> These risks share one root cause: agent identities that are **unseen, over-scoped, and unowned**. The maturity model on the next page is the path out.

WHY TRADITIONAL IAM FALLS SHORT

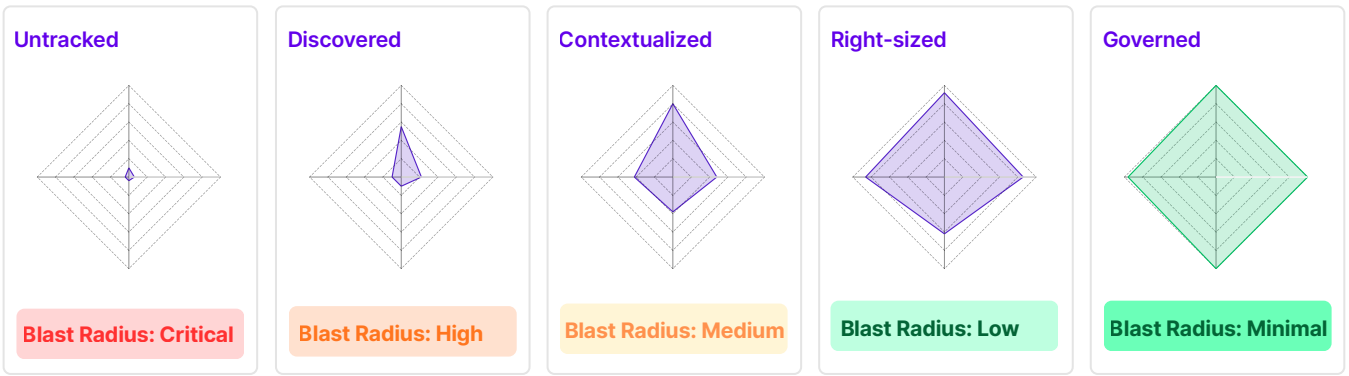
From invisible credentials to governed autonomy

Legacy IAM assumes a person logging in with a password, MFA, and a known schedule. Agents operate continuously, adapt to context, and reach across dozens of systems without pause. The result is a set of predictable failure modes.

Control has to live at the identity layer, where the access itself is decided. The maturity model below shows the five stages from invisible credentials to governed autonomy.



	Untracked	Discovered	Contextualized	Right-sized	Governed
Identity inventory	None	Complete	Mapped to access	Continuous	Continuous
Ownership	None	Partial	Assigned	Assigned	Automated at creation
Permissions	Broad, inherited	Broad, inherited	Assessed	Least privilege	Least privilege
Credential model	Standing	Cataloged	Flagged	Scoped, time-bound	Scoped, time-bound
Threat response	None	None	Anomaly detection	Detect and revoke	Automated
Audit trail	None	Fragmented	Reviewable	Full, Forensic-ready	Full, Forensic-ready



Each radar shows governance coverage on four axes: **visibility (top), access control (right), lifecycle (bottom), threat response (left)**. Coverage and blast radius move in opposite directions, so the fuller the shape, the lower the risk.

WHERE YOU STAND

The five stages

Every organization running AI agents sits on one of these five stages. Most are stuck in the first two, able to see some agent credentials but unable to govern them. Each stage below shows how to recognize it, and how Akto moves you up. The gap between stages is not more tooling, but a shift from seeing identities to governing them. Find where your org stands today, and take the next step up.

STAGE 1

Blast Radius: Critical

UNTRACKED

"We don't know what identities our agents carry."

SIGNALS

- No inventory of agent identities or credentials.
- Agents deployed by individuals without oversight.
- Agents inherit full permissions by default, with no owner.

HOW AKTO HELPS

- ☑ Discover every identity tied to an agent: keys, tokens, OAuth, roles.
- ☑ Detect agents operating outside security oversight.
- ☑ Attribute a clear owner to each identity from day one.

STAGE 2

Blast Radius: High

DISCOVERED

"We can see the credentials, but they are still standing and broad."

SIGNALS

- Partial discovery, static inventories that go stale.
- Credentials visible, but scope not yet assessed.
- New agents appear faster than they can be tracked.

HOW AKTO HELPS

- ☑ Continuous discovery across the entire AI stack.
- ☑ Map each credential to its agent and the scope it needs.
- ☑ Establish ownership and accountability automatically.

STAGE 3

Blast Radius: Medium

CONTEXTUALIZED

"We understand who owns each identity and what it is for."

SIGNALS

- Identities correlated to owner, permissions, and purpose.
- Excessive privilege and stale access are surfaced.
- Reviews are periodic, not continuous enforcement.

HOW AKTO HELPS

- ☑ Give every identity full context and purpose.
- ☑ Surface excessive privilege, stale access, and misconfig.
- ☑ Prioritize identity risk across the whole fleet.

STAGE 4

Blast Radius: Low

RIGHT-SIZED

"Access is scoped to intent, and compromise is caught in real time."

SIGNALS

- Least privilege enforced per identity.
- Standing credentials replaced with time-bound access.
- A leaked credential can be revoked before it is abused.

HOW AKTO HELPS

- ☑ Right-size every over-scoped identity to least privilege.
- ☑ Replace standing credentials with scoped, time-bound access.
- ☑ Detect compromised identities and revoke in real time.

STAGE 5

Blast Radius: Minimal

GOVERNED

"Identity governance operates at the speed of our AI."

SIGNALS

- Every new agent is registered and owned at creation.
- Lifecycle automated: rotation, expiration, segregation.
- A forensic-ready audit trail exists for every identity.

HOW AKTO HELPS

- ☑ Register and scope every new agent at creation.
- ☑ Automate rotation, expiration, and segregation.
- ☑ Enforce policy continuously from creation to retirement.

Identity governance, built for agents

Akto discovers every agent identity, maps its access, and governs it across the full lifecycle - at machine speed

Discovery

Discover every identity tied to an AI agent - API keys, bearer tokens, OAuth credentials, etc.

Identity Context

Understand their ownership, permissions, and purpose with complete context.

Lifecycle Management

Automate identity lifecycle and enforce policies around identity rotation, expiration and segregation.

Threat Detection

Catch real-time threats, compromised identities, and anomalous behavior across your agent fleet.

Posture Management

Surface excessive privileges, stale machine access, and misconfigurations.

Works Across your AI Stack

