Ping Identity®

Ultimate Guide

# The Ultimate Guide to Identity for AI

# Table of Contents


Building Trust in the New Agentic Economy 4
- Why Identity Is the New Trust Layer for AI 4
- How This Guide Helps: Strategic Insights for Executives and Practitioners 4

Understanding Agentic Identity 5
- What Is Identity for AI? 5
- Agentic Identity vs. Machine Identity and Human IAM 5
- From Access Control to Action Accountability 6
- The Four Pillars of Identity for AI 6
- Agentic Identity Unlocks Strategic Business Outcomes 7

Why Identity for AI Is Essential 8
- AI Agents Are a New Type of User 8
- Traditional IAM Isn't Enough 8
- Quantifying the Risk 8

How Agentic AI Alters the IAM Landscape 9
- Model Context Protocol (MCP) and Agent-Tool Interactions 9
- How Agents Access Systems, Data, and Services 9
- IAM's Role in Mediating Agentic Actions 9

Key Risks and Challenges in Agentic IAM 10

The Four Types of AI Agents 11

What You Need to Know 12

Equipping Your Identity and Security Teams for Agentic AI 13
- Visibility: Know Every Agent, Every Session 13
- Onboarding and Management: Define and Govern Agent Identities 13
- Authentication and Authorization: Secure Interactions at Machine Speed 13
- Human Oversight: Govern High-Risk Agentic Actions 14
- Threat Protection: Monitor and Contain Agent-Based Risks 14

# Table of Contents


Best Practices for Agentic AI Adoption 15
- Visibility: Know Every Agent, Every Session 15
- Onboarding and Management: Define and Govern Agent Identities 15
- Authentication and Authorization: Secure Interactions at Machine Speed 15
- Human Oversight: Govern High-Risk Agentic Actions 15
- Threat Protection: Monitor and Contain Agent-Based Risks 15
- Human Oversight: Govern High-Risk Agentic Actions 15
- Threat Protection: Monitor and Contain Agent-Based Risks 15
- Threat Protection: Monitor and Contain Agent-Based Risks 15

What to Watch for in the New Identity Landscape 16
- Visibility: Know Every Agent, Every Session 16
- Onboarding and Management: Define and Govern Agent Identities 16
- Authentication and Authorization: Secure Interactions at Machine Speed 16
- Human Oversight: Govern High-Risk Agentic Actions 16
- Threat Protection: Monitor and Contain Agent-Based Risks 16

Identity for AI Is the Foundation of Digital Trust 17

Appendix:
Key Terminology—Agentic Identity and IAM for AI Agents 18

# Building Trust in the New Agentic Economy

Artificial intelligence is no longer an innovation trend. It's an operational reality. AI agents are now embedded in everything from customer service to finance operations, acting on behalf of humans and enterprises alike. These agents aren't simply executing scripts. They're navigating systems, initiating transactions, and making decisions—autonomously or under delegated authority.

This evolution marks a radical shift. Agents aren't human users, but they do reason and make decisions with a growing degree of autonomy. This means they need identity. Not just credentials, but governance, oversight, and accountability. That's where Identity for AI comes in.

This guide is designed to help identity, security, and digital leaders secure and scale their AI initiatives responsibly. Whether you're enabling digital assistants for your workforce or integrating customer-facing agents, AI agents are becoming actors in your digital ecosystem. Identity is your source of trust. The control plane that ensures every agentic action is authorized, traceable, verifiably secure, and trusted.

## Why Identity Is the New Trust Layer for AI

The rise of agentic AI changes everything we know about digital interactions.

Traditional IAM systems were built for humans. But now, intelligent agents interact with sensitive systems, make decisions, and even other agents. This demands a new kind of trust framework that answers:

- **Who is the agent?**
- **Who approved its actions?**
- **What is it allowed to do?**
- **Can we trace what it did?**

Without this visibility and control, organizations risk breaches, regulatory failures, and reputational damage. They may even be blindsided by rogue agents or impersonating bots.

Identity for AI redefines the trust model for the AI era. It provides the tools to recognize agents, link them to users, enforce policies, and apply human oversight where it's needed.

## How This Guide Helps:
### Strategic Insights for Executives and Practitioners

This guide is built for both executive decision-makers and technical implementers:

→ **Product and Innovation Leaders**
Learn how identity underpins scalable and secure agent deployment across customer, workforce, and partner ecosystems.

→ **Security Leaders and Practitioners**
Find actionable guidance for adapting identity systems to manage AI agents, prevent impersonation, and apply Zero Trust at machine speed.

→ **Developers**
Get technical patterns and standards (OAuth 2.0, DCR, MCP) for securely onboarding, authenticating, and authorizing agents in production systems.

Across each section, we'll explore how to secure agentic AI by assigning identity, governing access, and maintaining auditability across every interaction.

# Understanding Agentic Identity

## What Is Identity for AI?

Identity for AI refers to the frameworks, protocols, and controls that manage how AI agents (non-human actors capable of autonomous or delegated action) are authenticated, authorized, and governed across digital systems.

Agentic identity is not a repackaging of machine identity. It accounts for agency, intent, and decision-making autonomy. These agents don't just execute code. They interact, reason, and perform tasks across workflows and APIs, often acting on behalf of users or the enterprise itself.

**Identity for AI ensures that each agent is:**

- Recognized as a distinct digital entity
- Bound to permissions or delegated authority
- Governed through lifecycle, audit, and oversight controls

## Agentic Identity vs. Machine Identity and Human IAM

Traditional IAM focuses on protecting systems from humans (or vice versa), but agentic IAM must govern non-human actors that are increasingly intelligent, autonomous, and interactive. These agents demand identity constructs that can scale across real-time actions, context shifts, and multiple trust boundaries.

Agentic identity introduces a new category of digital identity, blending characteristics of both human and machine identities but introducing unique challenges due to the autonomy and decision-making power of AI agents.

| Characteristic | Human Identity | Machine Identity | Agentic Identity |
|---|---|---|---|
| **Who** | People | Systems, services | AI agents (autonomous or delegated actors) |
| **Auth Mechanisms** | Passwords, MFA, biometrics | API keys, TLS certs | OAuth 2.0, DCR, mTLS, assertion grants |
| **Behavior** | Intentional, self-aware | Deterministic, static | Adaptive, context-driven, goal-oriented |
| **Delegation** | Manual consent | None or hardcoded | Scoped, authenticated delegation |
| **Oversight** | Policy enforcement, user input | Static rules, logging | Human-in-the-loop, dynamic monitoring, audit trails |
| **Lifecycle** | HR-driven onboarding and offboarding | Automated provisioning and credential rotation | Policy-driven, ephemeral, delegation-linked, fully auditable |

## From Access Control to Action Accountability

Unlike traditional identity where a person logs in and performs a task, AI agents blur the lines between actor, approver, and executor. These agents may act autonomously, trigger workflows, or operate across systems at machine speed. That makes the core job of IAM systems fundamentally more complex. They must not only verify access, but contextualize actions and trace authority in real time.

To manage AI agents responsibly and transparently, your identity architecture must answer four core questions:

**Who is acting?**
Each agent must have its own verifiable identity, not a shared API key or a shadow integration. Agents should be explicitly recognized and uniquely authenticated.

**Who authorized the action?**
Was the agent acting on its own, or under delegated authority from a human or system? Use mechanisms like assertion grants to encode this delegation clearly.

**On whose behalf?**
If an agent is operating for a user, trace its session or action back to that human identity. This is essential for enforcing user consent and enabling human-in-the-loop verification.

**With what permissions and limits?**
Ensure agents operate within tightly scoped roles. Apply dynamic access controls like just-in-time entitlements to prevent overreach or unintended consequences.

These questions are critical checkpoints for designing trustworthy, accountable agentic systems. In the next section, we'll look at the four guiding principles that every identity program should apply to AI agents to ensure security, compliance, and control.

## The Four Pillars of Identity for AI

Once identity systems can answer who's acting and under what authority, they must enforce the right behaviors. That's where foundational IAM principles (reimagined for AI agents) come into play.

These four pillars form the baseline for securing and governing agentic actions:

**1. Delegate, Don't Impersonate**
AI agents should never log in with a human user's credentials or "borrow" their access. Instead, use authenticated delegation. Issue scoped, temporary tokens that clearly define what an agent is authorized to do, and on whose behalf. This preserves trust boundaries, maintains nonrepudiation, and eliminates the audit gaps created by credential sharing.

**2. Enforce Least Privilege**
Agents often operate at machine speed and scale. That makes overprovisioning dangerous. Assign agents only the access they need for their specific function, and no more. Implement just-in-time entitlements, ephemeral tokens, and scope-based policies to reduce exposure and limit blast radius in case of compromise.

**3. Maintain Human Oversight**
AI can move fast, but not everything should be fully automated. For high-risk, high-value, or irreversible actions, integrate human-in-the-loop (HITL) approval flows. Technologies like CIBA or push-based step-up authentication allow users to validate sensitive agent actions in real time, without stalling performance.

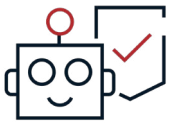
**4. Make Every Action Auditable**
To maintain accountability, every agentic action must be attributable, explainable, and reviewable. This means logging not just the action, but the actor (agent), the delegator (if any), and the policy or scope under which the action occurred. Separate agent telemetry from human logs to ensure traceability and forensics readiness.

These principles aren't just best practices—they're requirements for operating AI securely and at scale. In the next section, we'll explore why identity for AI isn't optional, and what's at stake for organizations that fail to modernize their IAM systems.

# Agentic Identity Unlocks Strategic Business Outcomes

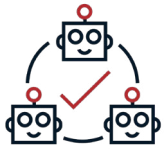
AI is fundamental to core digital initiatives that leading organizations are exploring. Identity is critical to minimizing risk and realizing the benefits promised by AI adoption.

To unlock the opportunities AI presents, organizations need to:



## Engage the Agentic Channel

AI agents are emerging as a new interface between customers and brands. With secure identity, you can confidently support personal agents, drive agentic commerce, and open new engagement channels, without compromising control.



## Secure the Autonomous Workforce

Digital workers enhance productivity by handling routine and complex tasks. Scoped, auditable identities ensure they operate within defined boundaries, which minimizes risk and maximizes operational efficiency.

## Protect Against Adversarial AI

From prompt injection to data and memory poisoning, threats are evolving. Agentic identity enforces trust boundaries, enables continuous verification, and helps detect and contain malicious behavior at machine speed.

# Why Identity for AI Is Essential

AI agents are already here. Whether they're managing IT tasks, serving customers, or making purchasing decisions, agents are quickly becoming embedded in the digital fabric of enterprise and consumer ecosystems. But most identity systems were never designed for this kind of autonomy. That gap has real consequences.

Getting it right can be a game-changing advantage for your organization. Getting it wrong can open you up to tremendous risk.

## AI Agents Are a New Type of User

Just like human users, agents interact with APIs, access sensitive data, initiate workflows, and trigger transactions. But unlike humans, they do it continuously, at scale, and they don't take breaks. AI agents reason, act, and adapt dynamically to fulfill goals, and treating them as service accounts or background processes ignores their operational complexity and risk profile.

These unique capabilities make agents powerful, but it also raises fundamental questions about how to manage their identity, access, and trust in enterprise environments. APIs, often acting on behalf of users or the enterprise itself.

## Traditional IAM Isn't Enough

Human-centric IAM relies on manual authentication, static roles, and long-lived credentials. These approaches don't translate to dynamic, autonomous agents. Most traditional IAM systems fall short in:

- Detecting when agents are active
- Differentiating good agents from malicious bots
- Enforcing context-aware access
- Maintaining traceability and accountability across delegated actions

Traditional systems authenticate users and apply static access controls, but agentic AI requires context-aware, goal-aligned identity frameworks. These agents blur the line between human and non-human user behavior, which requires new controls to ensure their actions are authorized, traceable, and policy-aligned.

## Quantifying the Risk

Without dedicated identity policies and processes in place for agentic identity, organizations face measurable exposure:

**Credential Misuse**
Shared secrets or reused user tokens undermine audit trails and open doors to impersonation and unauthorized access.

**Data Exfiltration**
Overprivileged or compromised agents can be used to extract sensitive data at scale, often without triggering traditional detection systems.

**Decision Accountability Gaps**
Without authenticated delegation, organizations can't trace which agent acted, on whose behalf, or under what authority—creating compliance and legal uncertainty.

**As Gartner puts it:**

> "By 2028, 90% of digital commerce organizations that allow humans to share credentials with AI agents will have experienced a tripling in ATO and first-party fraud."

*Gartner, How to Securely Delegate Access from Humans to AI Agents, Nathan Harris and Arif Khan, 20 August 2025*
*Gartner is a trademark of Gartner, Inc., and/or its affiliates.*

The risks carry real operational, regulatory, and reputational implications. Without clear identity and access boundaries, AI agents can become untraceable actors in systems that were never designed to govern or monitor them. Managing agentic risk starts with classification. Not all AI agents act the same, and they don't all need the same type of access. Identity strategies must adapt to how each agent operates and on whose behalf.

# How Agentic AI Alters the IAM Landscape

Agentic AI transforms the conventional IAM paradigm. Agents may act on behalf of a user, operate independently, or collaborate with other agents, all while traversing trust boundaries and invoking sensitive systems. This dynamic, autonomous behavior demands real-time evaluation of identity, authority, and context.

Where traditional IAM is centered on roles, credentials, and static policies, agentic IAM must be:

- **Delegated,** not impersonated
- **Dynamic,** not static
- **Contextual,** not binary
- **Observable,** not opaque

IAM teams must shift their mindset from controlling access to governing behavior.

## Model Context Protocol (MCP) and Agent-Tool Interactions

MCP is a foundational protocol for enabling secure, scalable interactions between agents and external tools, APIs, or data sources. It defines how agents discover, invoke, and receive responses from services in a standardized, authenticated way.

By abstracting the interface between agent and tool, MCP allows organizations to:

- Mediate access via policy-enforced gateways
- Register agent clients via Dynamic Client Registration (DCR)
- Authenticate agents via OAuth 2.0 or mTLS
- Apply fine-grained authorization using OAuth scopes and IAM roles
- Log and monitor tool use per agent session

MCP shifts agent access away from opaque automation toward controlled, auditable interactions.

## How Agents Access Systems, Data, and Services

Agents use two primary methods to interact with digital systems:

**API Interactions**
Agents use credentials or delegated OAuth tokens to call backend services such as REST or GraphQL APIs. These interactions integrate well with IAM systems and offer strong policy enforcement.

**Graphical User Interface (GUI) Interactions**
When APIs aren't available, agents may interact directly with the software's visual layer—clicking, typing, and navigating screens like a human user. These are often referred to as Computer-Using Agents (CUAs).

Each approach introduces distinct identity and security challenges. API interactions benefit from native IAM controls, but GUI interactions require additional safeguards such as session tagging, identity detection, and human-in-the-loop approvals for high-risk actions.

## IAM's Role in Mediating Agentic Actions

To become the trust broker for agentic activity, IAM must:

**Delegate**
Never impersonate. Issue scoped tokens instead of sharing credentials.

**Authenticate**
Validate agent identity using mTLS or signed assertions, with client details registered through DCR.

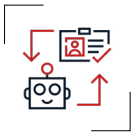
**Authorize**
Enforce policy constraints per session, task, or resource.

Just as importantly, IAM must log these interactions to maintain visibility and enable forensic review.

# Key Risks and Challenges in Agentic IAM

Agentic AI introduces a new attack surface that's often invisible to legacy IAM systems.

**Here are some of the key risks and how to solve for them:**



**Credential Sharing**

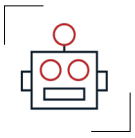
When agents use human credentials, audit trails are broken and impersonation risk rises.

***Solution:*** *Authenticated delegation using OAuth 2.0 and Dynamic Client Registration (DCR).*

**Overprivileged Agents**

Static, context-insensitive roles lead to agents accessing more than needed.

***Solution:*** *Just-in-time (JIT) entitlements and strict least privilege policies.*



**Shadow Agents**

Unauthorized or unmanaged agents act without detection.

***Solution:*** *Dynamic registration, tagging, and identity-based detection.*

**Lack of Oversight**

Automated actions can bypass compliance workflows.

***Solution:*** *Human-in-the-loop enforcement via CIBA or push-based step-up authentication.*

**Adversarial or Jailbroken AI**

Agents can be manipulated to perform unsafe or unauthorized behaviors.

***Solution:*** *Continuous monitoring, anomaly detection, and policy-based kill switches.*

# The Four Types of AI Agents

Understanding AI agent types is critical to designing the right identity strategy. There are many ways to segment agents, but there are four general overarching types:

## PERSONAL AGENT

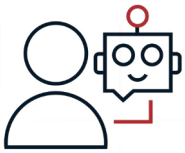
**"Your agent works for you"**

This is an agent on an individual user's device that the user deploys to external resources to complete tasks on their behalf.

**Example:** ChatGPT, Gemini, or a customer's shopping assistant

**Ownership:** Bring your own (BYO) / unmanaged

**Supervision:** Attended



## DIGITAL ASSISTANT FOR CONSUMER

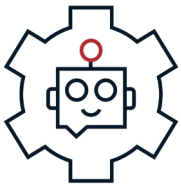
**"Our agent works for you"**

This is an agent under corporate control that has been deployed externally to serve customers.

**Example:** Brand chatbot or customer support assistant

**Ownership:** Managed

**Supervision:** Mixed (attended/unattended)



## DIGITAL ASSISTANT FOR WORKFORCE

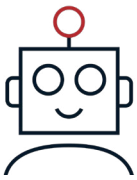
**"Our agent works for us"**

This is an agent under corporate control that has been deployed internally to serve employees.

**Example:** HR chatbot or internal IT assistant

**Ownership:** Managed

**Supervision:** Mixed



## DIGITAL WORKER

**"Our agent works autonomously for the enterprise"**

This is a semi- to fully-autonomous agent deployed by the enterprise to complete tasks internally.

**Example:** Logistics automation agent or finance assistant

**Ownership:** Managed

**Supervision:** Unattended

Each agent type requires different levels of control, delegation, oversight, and identity treatment.

Learn About Agent Types →

# What You Need to Know

## For CEOs, CIOs, Chief Product & AI Officers

**Identity for AI is foundational to scaling AI safely.**

You must know when agents are interacting with your systems and be able to trace their actions back to human users. This is critical to maintaining visibility into customer activity and scaling the agentic channel safely, without unintended fraud and security risks.

**IAM modernization is essential to monetize AI responsibly.**

Don't stretch legacy solutions. "Making do" to treat AI agents like humans or non-thinking machines will create problems that can hinder or fully halt your ability to leverage agentic AI. You need an identity solution that's fit for purpose.

## For CISOs, IAM & Security Leaders

**Delegate, don't impersonate.**

Allowing credential sharing will create massive security gaps and greatly increase risk. Agents must have their own credentials when they're autonomous, or delegated access when they're acting on behalf of a human user.

**Apply Zero Trust to all AI agent interactions.**

Agents are the new attack surface. They should never be offered implicit trust, and the principle of least privilege is key. Agents should only have the access they need to complete their tasks, and their access should be revoked when the task is complete.

**Govern agent lifecycles like human identities.**

Have a clear process to onboard, provision, monitor, and offboard agentic users.

**Use HITL for compliance, traceability, and control.**

Use out-of-band verification methods to verify explicit human consent every time agents take high-risk actions.

# Equipping Your Identity and Security Teams for Agentic AI

Now that the strategic imperatives and risks are clear, the next step is execution. Identity and security teams must evolve their tooling and operations to support agentic AI. Not just to mitigate risk, but to unlock new value. The following capabilities represent the core building blocks of an effective identity foundation for AI.

## VISIBILITY:
### Know Every Agent, Every Session

Before securing agents, organizations must first detect and classify them accurately and in context.

**Agent Discovery**
Identify and label AI agents interacting with APIs, GUIs, or external platforms, distinguishing them from traditional users or background processes.

**CUA Detection**
Use pattern recognition and device-level signals to spot computer-using agents that mimic human behavior through browser or terminal interactions.

**Platform Integration**
Connect IAM systems to agent platforms to capture real-time identity signals and behavioral data across diverse endpoints.

**Service Account Hygiene**
Ensure agents have distinct, managed identities —not shared or generic accounts—with clear attribution and auditable metadata.

## ONBOARDING AND MANAGEMENT:
### Define and Govern Agent Identities

Agents require structured provisioning, ownership assignment, and lifecycle governance tailored to their function and risk profile.

**Control Panel**
Use centralized tools to manage agent identities, ownership, credentials, and entitlements with continuous oversight.

**Identity Classification**
Categorize agents by type (digital worker, assistant, or BYO personal agent) and link each to a responsible owner or custodian.

**Provisioning Workflows**
Apply Dynamic Client Registration (DCR) and policy-based automation to scale secure agent onboarding across environments.

**Delegated Entitlements**
Define scope and link agent permissions to human delegators, ensuring that access is purpose-specific and revocable.

## AUTHENTICATION AND AUTHORIZATION:
### Secure Interactions at Machine Speed

Agent actions must be authenticated, scoped, and traceable in real time, adapting to evolving tasks and contexts.

**OAuth 2.0 and Assertion Grants**
Use protocol-based delegation mechanisms like signed JWTs to authorize agents without user credential sharing, enabling fine-grained control.

**MCP Gateway Enforcement**
Apply policy-based controls and authorization checks when agents invoke tools via Model Context Protocol, ensuring compliance.

**Scoped Access**
Issue short-lived, narrowly scoped tokens to reduce exposure and ensure least-privilege execution aligned with current task intent.

**Inter-Agent Trust**
Support cryptographic identity verification and signed delegation chains for agent-to-agent protocols like A2A or complex workflows in MCP.

## HUMAN OVERSIGHT:
### Govern High-Risk Agentic Actions

Not all tasks should be automated. Human oversight ensures accountability and ethical safeguards.

**HITL (Human-in-the-Loop)**
Require real-time approvals for sensitive operations using mechanisms like Client-Initiated Backchannel Authentication (CIBA), integrating users via familiar devices.

**Consent Frameworks**
Define clear boundaries for what agents can do on behalf of users, with structured user consent for specific scopes of action.

**Policy Constraints**
Impose time limits, context-based rules, and step-up authentication to prevent agents from overreaching or escalating privileges.

**Audit Trails**
Maintain separate, searchable logs for agent sessions and delegated actions, tagging all agent-triggered transactions for traceability.

## THREAT PROTECTION:
### Monitor and Contain Agent-Based Risks

As agents scale, so do attack surfaces. Identity plays a central role in proactive mitigation.

**Behavioral Monitoring**
Analyze telemetry for anomalies like unusual API activity, abnormal execution times, or patterns that deviate from baseline agent behavior.

**Adversarial Detection**
Detect jailbroken, hijacked, or misconfigured agents attempting to circumvent controls or exploit system vulnerabilities.

**Automated Response**
Quarantine suspect agents, revoke access tokens, and escalate to human review within seconds to contain risks efficiently.

**Periodic Review**
Schedule regular administrative reviews of agent entitlements, ownership, and activity trends to identify risk drift and compliance gaps.

With these core capabilities in place, the focus shifts to operational discipline. The following best practices help teams apply identity principles to real-world agent interactions, balancing speed, security, and oversight.

## How to Support Your Developers

1. Provide secure token issuance and scope enforcement.
2. Use OAuth 2.x and OpenID Connect foundations.
3. Enable Dynamic Client Registration (DCR).
4. Implement assertion grant or token exchange patterns.
5. Support CIBA or device authorization flows for HITL.
6. Build MCP gateways for secure agent-tool integration.

# Best Practices for Agentic AI Adoption

Adopting agentic AI successfully requires more than just technology. It requires disciplined identity governance, aligned security protocols, and a foundation of operational trust.

These best practices provide a strategic playbook for IAM leaders:

## Know and Classify Your Agents

Build an inventory of all AI agents, categorizing them by type, ownership, and autonomy level. Distinguish personal assistants, digital assistants, and digital workers, and apply policies accordingly.

## Delegate Access, Don't Share Credentials

Replace credential-sharing with authenticated delegation using OAuth 2.0 and assertion grants. Ensure all agent actions are properly scoped, non-repudiable, and traceable to the original user or system.

## Enforce Least Privilege and Scoped Access

Assign agents only the permissions they need, for as long as they need them. Use JIT access, expiring tokens, and fine-grained scopes to minimize lateral movement and blast radius.

## Require Human Approval for High-Risk Actions

Implement human-in-the-loop (HITL) for sensitive transactions. Use mechanisms like CIBA or push-based MFA to enable seamless human approvals without stalling agent performance.

## Monitor Agent Behavior Continuously

Collect telemetry from all agent sessions. Establish baselines and detect anomalies using behavioral analytics to catch misbehaving or compromised agents early.

## Log Agent Actions for Compliance and Auditability

Separate agent logs from human logs. Include actor identity, delegated authority, and scope in all logs. Make agent activity explainable and reviewable.

## Manage Entitlements Dynamically

Continuously evaluate agent roles and permissions. Support JIT, kill switches, and policy-based revocation to shut down access instantly in case of anomalies.

## Unify IAM Across Workforce, Customer, and Agent Channels

Treat agent identity as a core identity domain, alongside human workforce and customer identities. Ensure consistent policies and controls across all user types.

These best practices offer a stable foundation for governing agentic identity today. But implementation alone isn't enough. As AI agents become more capable and more embedded in core systems, identity leaders must stay responsive to a shifting technical landscape. The next section outlines the models and protocols shaping how agent interactions are secured, delegated, and controlled—and why staying aligned with them is essential.

# What to Watch for in the New Identity Landscape

As agentic AI systems evolve, the frameworks we use to secure and govern them must do the same. This section highlights five critical constructs that have emerged to address the challenges of secure agent behavior, identity, and collaboration.

These technologies represent the current shift in how identity systems are adapting to agent behavior, and they're actively being implemented to solve immediate needs in delegation, trust, and access control. For teams deploying or managing AI agents, these are the building blocks that are shaping secure, operational integration.

## Model Context Protocol

Model Context Protocol (MCP) is a standardized interface for how AI agents securely access external tools, APIs, or data sources. Instead of hardcoded integrations, MCP allows agents to discover and invoke capabilities through a policy-enforced gateway. It supports secure authentication (OAuth 2.0), establishes scoped access, and logs every interaction.

As organizations adopt agents at scale, MCP ensures consistency, visibility, and control over tool usage.

## Agent-to-Agent Protocols

Agent-to-agent protocols like Google's Agent2Agent (A2A) secure collaboration between AI agents. They enable agents to authenticate each other, share structured goals, and delegate tasks, all while preserving security context and policy alignment.

As multi-agent workflows grow more common (e.g., one agent researching while another summarizes), agent-to-agent protocols establish the trust and rules of engagement needed for cross-agent orchestration.

## Verifiable Credentials for Agents

Verifiable credentials (VCs) enable AI agents to present cryptographically signed claims about their identity, role, or affiliation. These credentials can be validated by relying parties without direct integration, enabling decentralized, tamper-evident verification.

VCs strengthen identity assurance, support privacy-preserving authorization, and reduce the risk of agent impersonation across systems and domains.

## Adaptive Authorization

Static roles aren't enough for AI agents that operate in dynamic, real-time environments. Adaptive authorization evaluates access decisions using contextual inputs (risk level, task type, timing, and behavioral patterns) to determine whether an agent should proceed.

This approach enables identity systems to continuously align permissions with situational needs, stepping up requirements or revoking access automatically when risk conditions change.

## Zero Trust for AI

Zero Trust principles—verify explicitly, assume breach, enforce least privilege—are even more critical when managing autonomous agents. AI agents should never have implicit trust, long-lived credentials, or unconstrained access.

Applying Zero Trust means continuously verifying agents, enforcing scoped delegation, using ephemeral credentials, and applying runtime checks to all actions, no matter how familiar the agent may seem.

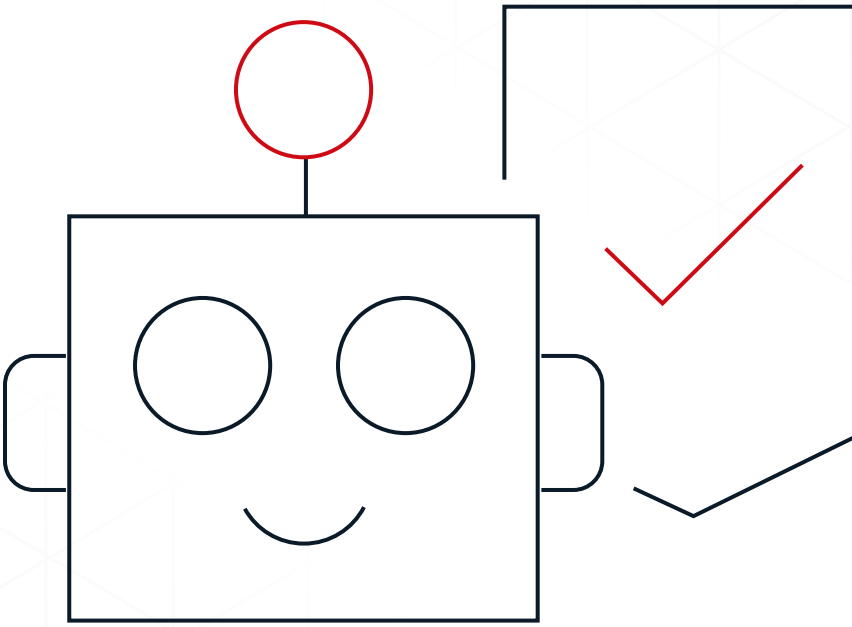
These constructs define the architecture of secure agent interactions. They're the technical foundation for enforcing trust, delegation, and control as AI agents become embedded across systems and workflows. Understanding them is essential for any team that's building or governing agentic capabilities today.

# Identity for AI Is the Foundation of Digital Trust

As you look to innovate with AI, you need identity. The constructs outlined in this guide aren't just tools. They're the mechanisms by which trust, control, and accountability are established in AI-driven systems.

If your enterprise applies these principles, you can move faster without sacrificing oversight. You can integrate agents into workflows without losing visibility. And you can deliver intelligent experiences that customers and regulators alike can trust.

In the age of agentic systems, identity isn't just infrastructure. And the organizations that get identity for AI right will quickly take the lead.



APPENDIX

# Key Terminology—Agentic Identity and IAM for AI Agents

**Agent-to-Agent Protocol**
A protocol (e.g., Google's Agent2Agent, or A2A) that enables secure communication between AI agents. They support authentication, delegation, task exchange, and collaboration across agentic workflows.

**Assertion Grant / Token Exchange**
An OAuth 2.0 flow that enables one identity token (such as a user credential) to be exchanged for a new access token that's scoped for an agent. This enables secure, traceable delegation without impersonation.

**Dynamic Client Registration (DCR)**
A mechanism that allows AI agents to register themselves as OAuth clients at runtime. DCR supports flexible and scalable onboarding of agents without manual provisioning.

**Ephemeral Credentials**
Short-lived credentials issued to agents for temporary use. These reduce the risk of exposure compared to static, long-lived secrets or API keys.

**Just-in-Time (JIT) Entitlements**
A model where agents receive only the access they need, when they need it. Entitlements are granted based on real-time context, then revoked after use.

**Model Context Protocol (MCP)**
A standard interface that allows agents to securely discover and interact with external tools and APIs through a policy-enforced gateway. MCP supports authentication, authorization, and telemetry.

**mTLS (Mutual TLS)**
A secure protocol where both the agent (client) and the service (server) verify each other's identities using digital certificates. mTLS is commonly used to ensure mutual trust in sensitive interactions.

**OAuth 2.0**
An open standard for secure access delegation. It allows AI agents to act on behalf of users by issuing access tokens with specific scopes, avoiding the need to share credentials.

**OpenID Connect (OIDC)**
An authentication protocol built on top of OAuth 2.0 that enables agents or clients to verify user identity and securely receive structured profile information.

**Policy-Based Access Control (PBAC)**
An adaptive access control model that uses contextual policies—including user role, risk signals, time, and task type—to evaluate access in real time.

**Scoped Access Tokens**
Tokens that are restricted to a defined set of actions or resources. Scoping ensures agents operate with least privilege, aligned to their delegated task.

**Secure Software Attestation (SSA)**
A method to verify the integrity and origin of agent software before granting access. SSA helps ensure only authorized, tamper-free agents can operate within a system.

**Trust Boundary**
A conceptual perimeter that separates trusted entities (like enterprise-managed agents and users) from untrusted or external entities (such as personal agents). Trust boundaries guide how identity and access policies are enforced.