WHITE PAPER

COGILITY

Cogynt.ai: Extending Insider Risk Management to Human-Agent Systems

By Frank L. Greitzer, PhD
Stuart Booth
Michael Latta
Ethan Ayer

Executive Summary

Organizations are entering a new phase of AI adoption in which autonomous and semi-autonomous agents are no longer experimental tools but operational actors inside enterprise and mission systems. Gartner defines **agentic AI** as embedding autonomous, goal-driven behavior into products so they can complete tasks on a user's behalf. AI agents can access data, invoke tools, execute workflows, interact with APIs, and act under delegated human authority. As a result, the security question is no longer limited to whether a model generated an unsafe output; it is whether an agent's end-to-end behavior was authorized, mission-appropriate, auditable, and attributable over time. This white paper argues that agentic AI is more than a cybersecurity challenge. Effective governance therefore requires understanding the full human-agent-system relationship, rather than evaluating isolated prompts, outputs, or tool calls.

The market is moving quickly, but current offerings address only parts of the problem. Existing solutions emphasize AI observability, runtime guardrails, non-human identity governance, compliance monitoring, and integration with security operations platforms. These capabilities are important, but they are generally optimized for point-in-time visibility and policy enforcement rather than longitudinal behavioral understanding. Enterprises still need answers to higher-value questions:

- *Is an agent acting outside its authorized role?*
- *Did a human delegate inappropriate, risky, or policy-violating work to an agent?*
- *Are a human and agent behaving anomalously together?*
- *Can the organization attribute agent actions to the right human sponsor?*
- *Can the organization reconstruct a complete, defensible, and auditable history of the action?*

> ***The Core Challenge:*** *True security requires a longitudinal behavioral understanding of autonomous or semi-autonomous actors across systems over time. Effective security cannot evaluate an agent in isolation; it must analyze the* ***Human–Agent–System Triad.***

Cogility's opportunity is to extend its proven Cogynt.ai decision intelligence platform into this emerging agentic AI security domain. Rather than attempting to replace existing point solutions, Cogynt.ai can serve as the analytics, behavioral intelligence, policy-context, case-history, and human-oversight layer that fuses their telemetry with UAM, cyber, identity, mission, and policy context. In this model, agents become first-class tracked entities with persistent histories, relationships to human sponsors and teams, and measurable patterns of normal and abnormal behavior. By correlating agent actions with delegated authority, resource access, tool use, mission role, and behavioral indicators over time, Cogynt.ai can support a whole-person and whole-agent analytic approach that is directly aligned with risk management practices already proven in operational environments.

For an initial offering, the paper recommends building a minimum viable capability around seven core capabilities: an agent entity model, agent event ingestion, human-agent linkage, behavioral indicators, risk-pattern analytics, investigative timelines, and policy-context integration. This design leverages Cogynt.ai's strengths in supporting large-scale streaming analytics, stateful processing, full provenance, deterministic rule-based assessment, and expert oversight through interactive visualizations. The Core Challenge: True security requires a longitudinal behavioral understanding of autonomous or semi-autonomous actors across systems over time. Effective security cannot evaluate an agent in isolation; it must analyze the Human–Agent–System Triad. These features enable a more mature approach to agent governance: not just showing what went wrong after the fact, but continuously assessing risk across multi-agent and human-agent environments, preserving auditability, and enabling real-time enforcement or escalation when behavior departs from mission need, delegated authority, or organizational policy.

The concepts and preliminary architecture of the solution are described. A key early step in operationalizing the solution is expanding the existing **SOFIT** insider-risk knowledge base to include agentic AI indicators and patterns across the human-agent-system triad. This new taxonomy would provide the conceptual structure for detecting risky delegation, privilege expansion, anomalous tool chains, inappropriate access, coordinated multi-agent behavior, and other emerging signals of agentic misuse. From there, Cogility can execute a phased engagement framework: project planning and stakeholder alignment; requirements definition and system-boundary analysis; implementation of data integrations, behavioral models, and analyst workflows; proof-of-value testing and model calibration; and finally full operationalization with ongoing governance, enhancements, and support. This disciplined approach reduces deployment risk while creating a credible path from concept to enterprise-scale capability.

By utilizing **Hierarchical Complex Event Processing (HCEP)**, Cogynt.ai delivers unique architectural benefits:

- ***Stateful Processing.*** By maintaining state over time, Cogynt.ai can detect departures from normal multi-agent workflow.
- ***Full Provenance and Auditability.*** Cogynt.ai's stateful analysis, supported by modern LLM domain understanding, produces reliable, auditable and explainable outputs.
- ***Interactive Visualizations Supporting Expert Human Oversight.*** Cogynt.ai has visualizations and analysis allowing intervention or remediation through human oversight.
- ***Deterministic AI Framework Encoding Expert Understanding.*** The patented HCEP technology enables encoding of expert understanding in a deterministic AI framework for continuous assessment of human and agentic actors in real time at scale.

The central conclusion is that AI agents are becoming a new class of digital insiders, and enterprises will need governance capabilities that go beyond access control and prompt logging. Over the next 24 months, the agentic AI security and governance market is likely to see several existing categories mature— agent observability and evaluation, runtime security and policy enforcement, agent and non-human identity governance, and security operations and response — while a distinct governance/IRM category emerges to preserve behavioral context, accountability, risk history, and investigative narratives across human, agent, and system activity. Cogility is well positioned in the last category because its differentiator is not raw telemetry collection, but the fusion of technical, behavioral, and mission context signals into explainable risk assessments and defensible investigative narratives. In that sense, this paper frames agentic AI security not simply as a technical monitoring problem, but as a strategic extension of modern insider risk management.

1. Why Agentic AI Changes Enterprise Risk

AI agents have become non-human actors inside the enterprise, and organizations need to know who/what they are, what they can access, what actions they took, whether those actions were appropriate, and who is accountable.

As mentioned previously, Gartner defines agentic AI as embedding autonomous, goal-driven behavior into products so they can complete tasks on a user's behalf. Gartner's framing is particularly useful for enterprise risk discussions because it ties agentic AI to boundaries, oversight, observability, and trust.

Solutions that address the agentic AI security problem are in formative stages, with a current focus, if any, on logging prompts or blocking bad outputs. This is at best a primitive approach to the emerging agentic AI security/Insider Risk Management (IRM) challenge. The hard problem, instead, is to achieve longitudinal behavioral understanding of autonomous or semi-autonomous actors across systems.

2. The Human-Agent-System Triad

Just as leading IRM approaches recognize the critical role of the human and the need to understand underlying psychosocial, behavioral factors in addition to more traditional technical indicators of insider exploits, effective agentic AI security programs must implement a "whole agent" behavioral analytic model within an integrated IRM approach that recognizes the important contributions of the human-agent-system triad (see Fig. 1).

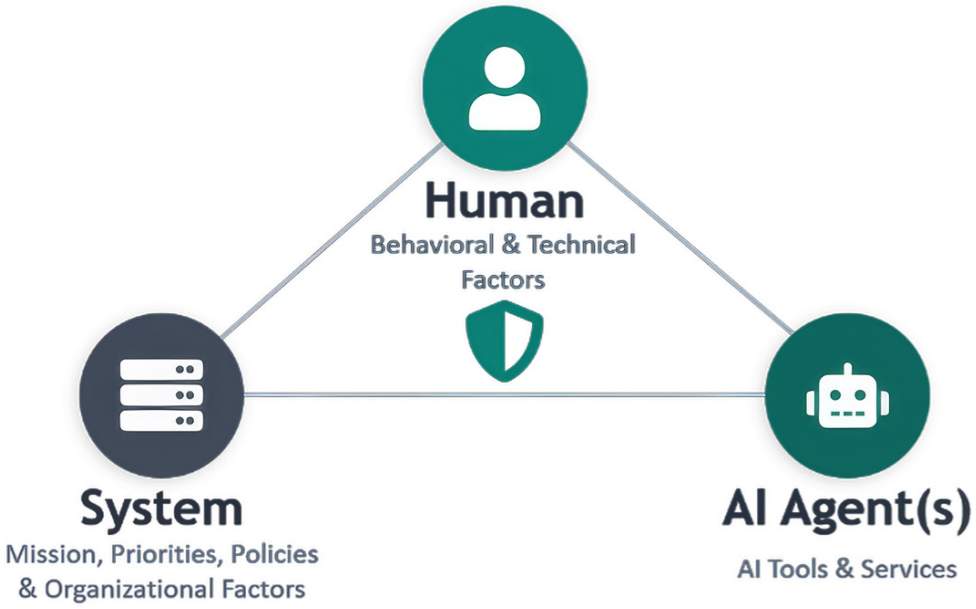


Figure 1. The Human–Agent–System Triad

This holistic approach to IRM/agentic AI security governance addresses (a) behavioral and technical indicators comprising the whole-person behavioral analytic model; (b) a "whole-agent" risk history including AI agent activities; and enterprise/system information such as the organization's mission and policies to inform the assessment of whether the combined human-agent behavior is normal, authorized, and mission-appropriate.

3. The AI Challenge: Adoption Outpacing Regulatory Control

Enterprises are deploying agents faster than governance can keep up. Microsoft's security team has described AI agents as scaling faster than companies can track, creating a visibility gap and requiring Zero Trust-style observability, governance, and security. These agents can significantly increase the security risk of enterprise because they have **tools, memory, delegated credentials, API access, file access, code execution, workflow automation, and inter-agent communication.**

U.S. AI governance is still fragmented and largely framework-driven rather than governed by a single comprehensive AI statute, with the White House setting executive policy, OMB directing Federal agency AI governance, NIST providing the AI Risk Management Framework, and CISA/DHS addressing AI-related cybersecurity and critical-infrastructure risk. The current regulatory direction is shifting from broad ethics or "responsible AI" principles toward operational controls1: agency AI governance, risk mapping, cybersecurity protection, auditability, model-security review, and enforcement against unlawful uses of AI agents to access data or systems.

In this new operational environment, the most pressing agentic AI security questions become:

- *Is an agent acting outside its authorized role, mission purpose, or historical behavioral pattern?*
- *Did a human delegate inappropriate, risky, or policy-violating work to an agent?*
- *Are a human and agent behaving anomalously together, including cases where the agent may be acting as a proxy for insider activity?*
- *Can the organization attribute agent actions to the right human sponsor, workflow, service account with context?*
- *Can the organization reconstruct a complete, defensible, and auditable history of intent, context, actions, decisions, exceptions, and consequences?*

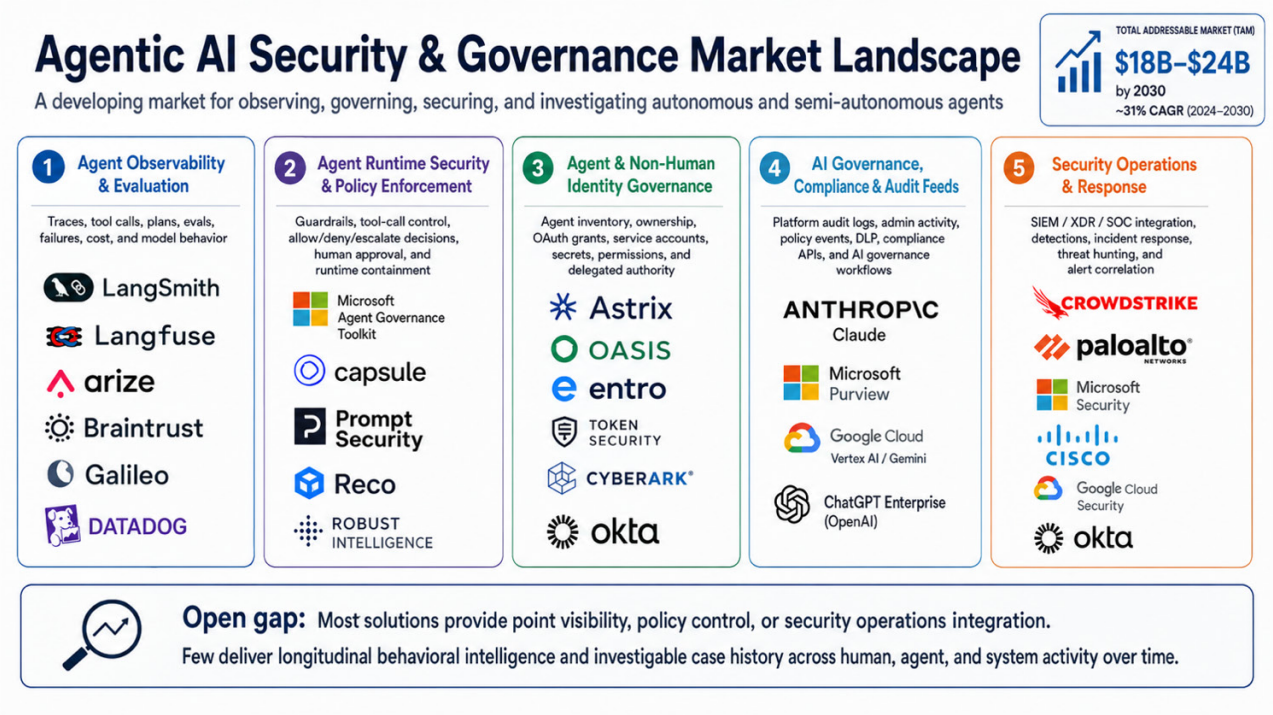
Recent Federal policy reinforces the urgency of agentic AI security. The June 2026 Executive Order, *Promoting Advanced Artificial Intelligence Innovation and Security,*[2] directs agencies to prioritize cyber defense of National Security Systems, Department of War systems, and civilian Federal Government systems; expand access to AI-enabled defensive tools; and establish mechanisms for secure frontier model deployment. The order also explicitly identifies the unlawful use of AI agents to access data or information as a federal enforcement concern. This policy direction strengthens the case for agentic AI governance capabilities that can establish attribution, preserve auditability, assess whether agent behavior is authorized and mission-appropriate, and support defensible investigation of human-agent-system activity.

---

1 NIST's AI Risk Management Framework [https://www.nist.gov/itl/ai-risk-management-framework] and GenAI Profile emphasize governance, mapping, measuring, and managing AI risks.

2 https://www.whitehouse.gov/presidential-actions/2026/06/promoting-advanced-artificial-intelligence-innovation-and-security/

4. The Evolving Vendor Landscape



4.1 Why Point Controls Aren't Enough

1. **Agent Observability & Evaluation** — tracing prompts, tool calls, reasoning paths, costs, failures, and model behavior. This is the most developer-oriented capability in which vendors and open-source projects focus on traces, chains, tool calls, latency, cost, model performance, hallucination/failure analysis, and evaluation. **Drawback** *AI observability is not sufficient for IRM because it focuses on the question, "Why did the agent fail?" more than "Is this actor behaving appropriately over months across the enterprise?"*

2. **Agent Runtime Security & Policy Enforcement** — allowing, denying, or escalating agent actions before execution. The goal of the runtime security and policy enforcement capability is to evaluate tool calls, resource access, messages, egress, or inter-agent communication before or during execution. ***Product Development Opportunity:*** In describing its *open-source Agent Governance Toolkit, Microsoft observes that the infrastructure to govern autonomous behavior has not kept pace with how easy it is to build agents. Rather than seeking to replace these tools, the better product development move is to consume their logs and decisions, then add cross-source behavioral analytics, risk scoring, and investigative history.*

3. **Agent & Non-human Identity Governance** — inventorying agents, service accounts, API keys, OAuth grants, MCP servers, secrets, and permissions. This capability, which is arguably receiving the most attention, treats AI agents as identity-bearing actors. **Drawback** *While identity vendors focus on agent inventory and access-control, they do not consider long-term fusion across User Activity Monitoring ("UAM"), case history, physical/logical events, cyber indicators, mission context, and anomalous workflows.*

4. **AI Governance Compliance and Audit Feeds** — These logs will be an important source of telemetry for Cogynt.ai.

5. **Security Operations & Response** — the large security platforms are feeding agent activity into enterprise security operations to address threats such as prompt injection, memory poisoning, tool misuse, privilege escalation, data exfiltration, and shadow agents. ***Product Development Opportunity:*** *Large vendors will dominate generic security controls, but they will not necessarily solve government IRM-style questions around whole-person / whole-agent behavioral adjudication.*

Products are strong in *point visibility* or *policy control*, but weaker in *longitudinal behavioral analytics* across heterogeneous populations of agents, humans, UAM, cyber actors, systems, and mission workflows. Emerging tools are strong in *instrumentation* and *control*, but weaker in *behavioral intelligence*.

Table 1. Existing/Emerging Agentic AI Security Market Capabilities and Gaps

Most Tools Focus on…	Higher-Value Questions Not Currently Being Addressed
• "What prompt was used?" • "What LLM answered?" • "What tool did the agent call?" • "Was that tool call allowed?" • "Was sensitive data exposed?" • "Which token or service account was used?"	• "Is this agent's behavior normal for its assigned role?" • "Is this agent acting outside its historic mission pattern?" • "Did a human delegate inappropriate work to an agent?" • "Is a human-agent pair behaving anomalously?" • "Is the agent becoming a proxy for insider activity?" • "Did multiple agents produce a coordinated or cascading risk pattern?" • "Can we reconstruct a complete history of agent intent, context, access, actions, exceptions, and consequences?" • "Can we distinguish technical anomalies from missionrelevant behavioral risk?"

5. Example Use Cases

The following scenarios make the Human–Agent–System Triad tangible by showing how risk emerges not from a single prompt or tool call, but from the interaction among a sponsoring human, an operating agent, and the surrounding enterprise systems, policies, and data. In each case, the security objective is not simply to log what happened, but to determine whether the combined pattern of delegated authority, access, behavior, and outcomes remained mission-appropriate, policy-compliant, and attributable over time.

1. **Human using an agent as a proxy.** An employee directs an agent to perform bulk search, summarization, or file-transfer tasks that the employee would not normally be permitted to perform directly at scale, or that would otherwise trigger heightened DLP scrutiny if executed manually. No single action may look overtly malicious in isolation; the agent is simply searching, reading, summarizing, and moving information within its available tool set. But viewed across the human-agent relationship, the workflow may indicate proxy behavior in which a person is attempting to use automation to obscure intent or circumvent standard controls. In this scenario, Cogynt.ai's value is to link the delegated tasking to the human sponsor, compare the workflow

against normal behavioral baselines, and identify whether the agent has become a surrogate channel for risky insider activity rather than a legitimate productivity aid.

2. **Privilege expansion over time.** An agent may begin with limited OAuth grants, narrow API scopes, and tightly bounded service-account rights, yet gradually accumulate broader access through iterative configuration changes, convenience exceptions, or inherited permissions. Because each individual privilege adjustment may appear reasonable, the overall pattern can be easy to miss until the agent's effective authority no longer matches its original mission role. This is a classic longitudinal governance problem: the risk lies in the slow divergence between intended purpose and actual capability. Cogynt.ai would continuously track the agent's evolving access profile, compare it with peer agents and declared mission need, and flag when privilege growth suggests role drift, overprovisioning, or emerging opportunity for misuse.

3. **Multi-agent cascade risk.** In a multi-agent environment, one agent may retrieve sensitive material, a second may summarize or transform it, and a third may transmit the result outside the originating environment. Each step may appear low-risk when viewed as an isolated tool call: retrieval, summarization, and transmission are all common operations. The real hazard emerges from the chain of activity across multiple agents and systems, where no single event fully reflects the severity of the outcome. This scenario illustrates why point-in-time observability is not enough. Cogynt.ai's stateful, cross-actor analytics can reconstruct the end-to-end workflow, attribute each stage to the relevant agent and human sponsor, and detect that the combined sequence constitutes a coordinated exfiltration or policy-violating cascade rather than a series of benign operations.

4. **Agent use after human status change.** A human sponsor who created or authorized an agent may later be placed on leave, reassigned, terminated, or lose the clearance or need-to-know that justified the delegation in the first place. If the agent continues to operate under stale credentials, retained tokens, or inherited workflows, the organization can face a significant accountability gap: the non-human actor remains active even though the human basis for its authority has changed. The triad model makes clear that agent authorization cannot be treated as static. Cogynt.ai would correlate HR or identity status changes with dependent agents, detect stale authority relationships, and surface cases where an agent's continued activity is no longer aligned with current human sponsorship, organizational policy, or mission need.

5. **Repeated blocked actions as a behavioral indicator.** Runtime guardrails may successfully stop an agent from carrying out prohibited actions such as accessing restricted data, invoking disallowed tools, or transmitting sensitive content. However, repeated blocked attempts are themselves behaviorally meaningful. An isolated denial may reflect a harmless error, but a recurring pattern can indicate that the agent's configuration, tasking, or sponsoring human is persistently pushing against organizational boundaries. In this sense, the blocked actions are not just enforcement events; they are risk indicators. Cogynt.ai would elevate this pattern into a longitudinal signal associated with the relevant agent, workflow, and human sponsor, allowing analysts to distinguish one-off mistakes from emerging misuse, misalignment, or deliberate attempts to probe for policy weaknesses.

6. How Cogynt.ai Addresses the Agentic AI Security Gap

As organizations adopt AI agents, those agents become new actors inside mission and enterprise systems. Traditional UAM tools can show technical activity, and identity tools can show permissions, but to more effectively address agentic AI security, the analysis also needs longitudinal behavioral context: who or what acted, under whose authority, against which resources, whether the action was appropriate for the mission, and whether the pattern changed over time.

Cogility is a market leader in whole-person insider risk management and offers a full-suite solution. This approach can be extended to provide risk assessment for non-human and hybrid human-agent actors by fusing agent activity, UAM, identity, cyber, application, and mission-context data into a long-lived behavioral history. Cogility will focus on building the analytics, risk scoring, policy-context mapping, case management, risk history, and human-in-the-loop oversight functionality while leaving specialized AI observability, model evaluation, and runtime-measurement functions to other vendors.

6.1 Core Capabilities

Cogility can add value by implementing the core capabilities shown in Table 2.

Table 2. Core Capabilities for Agentic AI IRM

Agent entity model	Treat agents as first-class tracked entities: agent ID, owner, sponsoring user/team, purpose, tools, permissions, model, runtime, data domains, credentials, deployment environment, and mission role.
Agent event ingestion	Ingest platform audit logs, UAM data, identity events, API calls, tool-call traces, MCP activity, DLP alerts, ticket/case data, endpoint activity, and cloud logs.
Human-agent linkage	Associate agent behavior with the human, team, service account, or workflow that created, delegated to, approved, modified, or benefited from the agent.
Behavioral indicators	Document normal patterns by agent, human-agent pair, mission function, tool class, data domain, and time window. Existing and new behavioral indicator taxonomies will inform a whole-person/whole-agent behavioral analytic model: the SOFIT[3] human and organizational insider risk indicator taxonomy can serve as a partial knowledge base that addresses human and system indicators; to complete the human-agent-system triad, an agentic AI risk indicator taxonomy must be developed.[4]
Risk patterns / HCEP-style analytics	Detect patterns such as privilege expansion, unusual tool chains, sensitive-data access outside mission need, repeated blocked actions, nighttime/off-cycle activity, anomalous delegation, agent use after employee status change, use of agents to bypass UAM controls, or coordination across multiple agents.
Investigative timeline	Produce a defensible narrative: what happened, in what order, by which actor, under whose authority, with which data, through which tools, and with what risk indicators.
Policy-context integration	Do not just say "anomaly." Map events to agency policy: acceptable use, datahandling rules, delegated authority, separation of duties, export controls, privacy rules, and mission constraints.

[3] Sociotechnical and Organizational Factors for Insider Threat (SOFIT) is a knowledge base comprising hundreds of individual human and organizational risk indicators for insider threat, originally developed in a research program funded by the Intelligence Advanced Research Projects Activity (IARPA). An updated SOFIT taxonomy is available here: https://cogility.com/blog/everything-you-always-wanted-to-know-about-sofit2/

[4] The implementation of an agentic AI risk indicator taxonomy will reflect the five key capabilities listed in the previous section and, most importantly, the security indicator gaps enumerated in Table 1.

6.2 Reference Architecture

For Phase 1, Cogynt.ai will serve as the behavioral intelligence aggregation point for all available data sources related to humans and agents. As it does in the traditional IRM use case, Cogynt.ai will apply all the sophisticated and interrelated tracking patterns, based on SOFIT, plus the new patterns in the developing SOFIT-AI taxonomy. The platform could deploy any third-party framework or taxonomy (or multiple frameworks). The existing case management, investigative intelligence, risk histories and "drill down" audits will continue to be used so employees and agents will have their own individual PRIs, relationship links and will also roll up to aggregated dashboards. Cogility will validate available feeds for each customer and/or through partner integrations. Candidate feeds include agent runtime logs, tool-call traces, identity and entitlement data, model/platform audit logs, and cloud/API logs, along with UAM telemetry, DLP events, endpoint activity, case-management records, and policy or mission context. These feeds create persistent agent entity profiles linked to human sponsors, teams, service accounts, workflows, permissions, and organizational policies. HCEP models can then evaluate whether combined human-agent-system activity remains consistent with delegated authority, mission need, normal behavior, and policy boundaries. The resulting risk indicators, timelines, and provenance can support analyst review, case management, SIEM/XDR integration, and, where the customer architecture permits, escalation or downstream enforcement workflows. The runtime interception and policy enforcement will be introduced in Phase 2, after proving out the longitudinal tracking of human and agent behaviors.

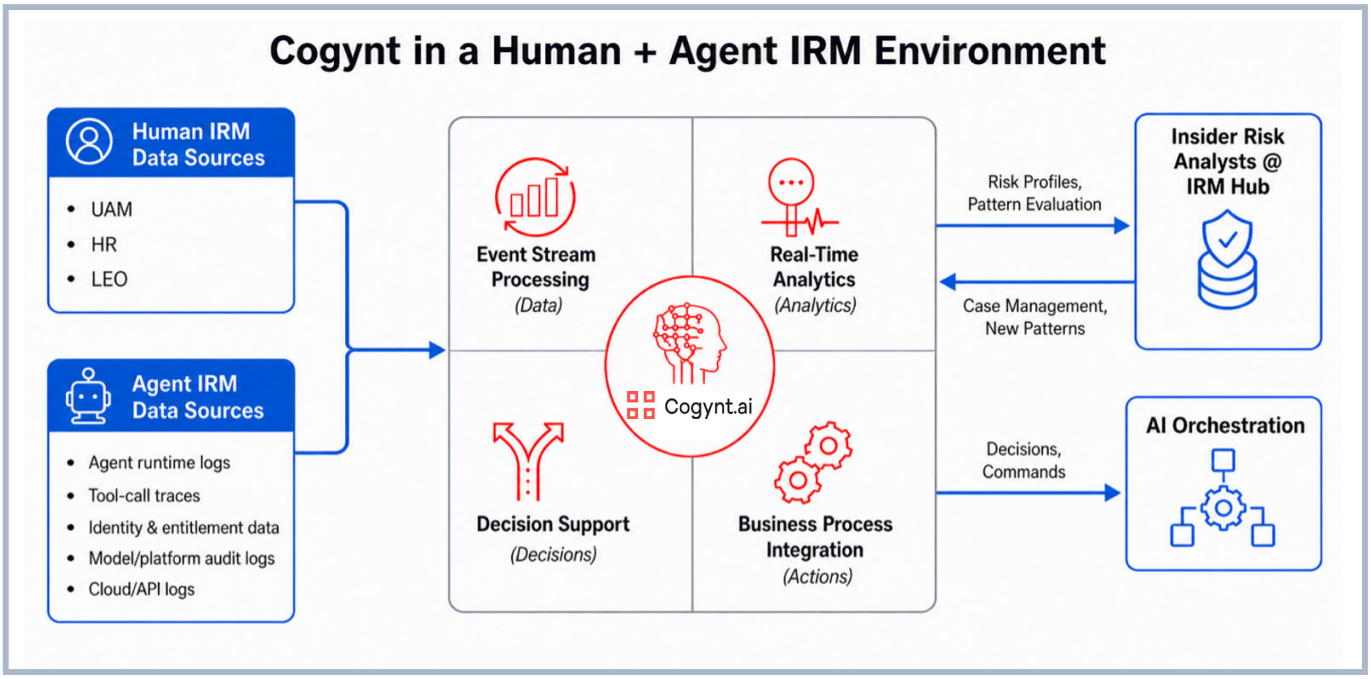


Figure 2. Architecture for Agentic IRM Solution

7. Cogility's Strategic Advantages

Cogynt.ai is uniquely suited to tackle the behavioral intelligence portion of the agentic AI security challenge with the goal of moving from siloed dashboards that merely show what went wrong after the fact into an interactive, proactive risk-management solution. This decision intelligence solution delivers proven patented technology (Table 3), policy/domain knowledge (Figure 3) and operational experience:

Table 3. Technology

Feature / Benefit	Relevance to Agentic Security Challenge
Proven Large-Scale IRM Deployments	Cogynt.ai has been successfully deployed to support whole-person IRM programs for large organizations, handling billions of events per hour. Deployed in a standard Kubernetes environment, the platform is readily managed in any environment (cloud or on premises).
Stateful Tracking - by maintaining state over time, Cogynt.ai can detect departures from normal multi-agent workflow.	Agents loop, retry and hand tasks off to other agents. Because Cogynt.ai maintains state over time, it can detect when a multi-agent workflow is drifting away from its original goal or exceeding its budget across separate sessions.
Full Provenance and Auditability - Cogynt.ai's stateful analysis produces reliable, auditable and explainable monitoring, recommendations, and case histories	If an autonomous agent makes a critical operational error, you cannot simply look at a black box LLM to understand why. Cogynt.ai provides full data lineage, letting auditors see the exact sequence of event inputs and business rules that led to a specific guardrail intervention.
Interactive Visualizations to Support Expert Human Oversight	Cogynt.ai has visualizations and analysis enabling the expert analyst to review and oversee the analysis, monitor the risk assessment process, and direct or modify intervention or proactive remediation actions as appropriate.
Deterministic Enforcement Cogynt.ai includes patented Hierarchical Complex Event Processing (HCEP) technology that allows encoding expert understanding in a deterministic AI framework to perform the continuous assessment of all actors (human and/or agentic) in real time at scale	It allows you to separate dynamic logic (creative problem-solving happening inside LLMs) from deterministic logic (the strict business boundaries). When an agent crosses a line, Cogynt.ai can trigger a Kafka sink connector to halt the host process, redact sensitive text, or force a human-in-the-loop approval.
No-Code Authoring / Faster Iteration	The no-code authoring tool allows the enterprise users to quickly and safely engage with this "expert system" to test how their policies are enforced and QA the results.

With these proven capabilities that deliver a real-time, scalable enterprise IRM solution, the Cogynt.ai decision intelligence platform functions as a first line of defense in triaging relevant human, agent, and system activity within an organization, at a very high level of confidence. The extension of this solution to address agentic AI security risk assessment fits naturally with Cogility's whole-person approach.

7.1 SOFIT - Extending the Risk Indicator Knowledge Base

Cogynt.ai provides a robust behavioral analytic framework to support the operational concept and capabilities described here. To implement the model, it will be necessary to define indicators for relevant agentic behavior and to identify risk patterns associated with security risks that come from this new multi-entity (human + agents, multiple agents, etc.) threat environment. Initial concepts for these indicators and patterns are evident in the material described in the previous sections and identified in Table 1.

The SOFIT knowledge base of Potential Risk Indicators (PRIs) can provide an initial structure with its focus on insider risk indicators associated with human users (including technical and behavioral factors) and on organizational (system) factors. To accommodate the agentic AI environment, the SOFIT taxonomy can be expanded to cover agentic AI security risks: e.g., those associated with human interactions with AI, behavior of AI agents fulfilling tasks assigned by humans, and potentially actions performed independently by AI agents.

The process aimed at defining new agentic AI PRIs would mirror the process that was used to develop SOFIT. We might designate the expanded knowledge base as "SOFIT-AI." A group of subject matter experts (SMEs) would define an initial set of agentic AI indicators and behavioral patterns. As was done with the original SOFIT knowledge base with an Individual Factors branch that was divided into nine classes and roughly thirty subclasses, the Agentic AI class will be decomposed into a 3-level hierarchy comprising a set of major classes and lower-level subclasses, with the agentic PRIs occupying the lowest level in the hierarchy. A preliminary conceptual class/subclass structure is illustrated in Fig. 3. This initial set of PRIs forming the agentic AI class of the SOFIT taxonomy is then critiqued through a SME engagement process (e.g., interviews or expert knowledge elicitation surveys) to refine the model, identify data sources for detecting PRIs, specify relevant behavioral patterns, and to gain new insights on more complex interactions among the three elements (human –agent– system) of the agentic AI security triad. The Cogynt.ai HCEP model is uniquely suited to represent this hierarchical structure and implement the SME-defined pattern-based risk assessment process.

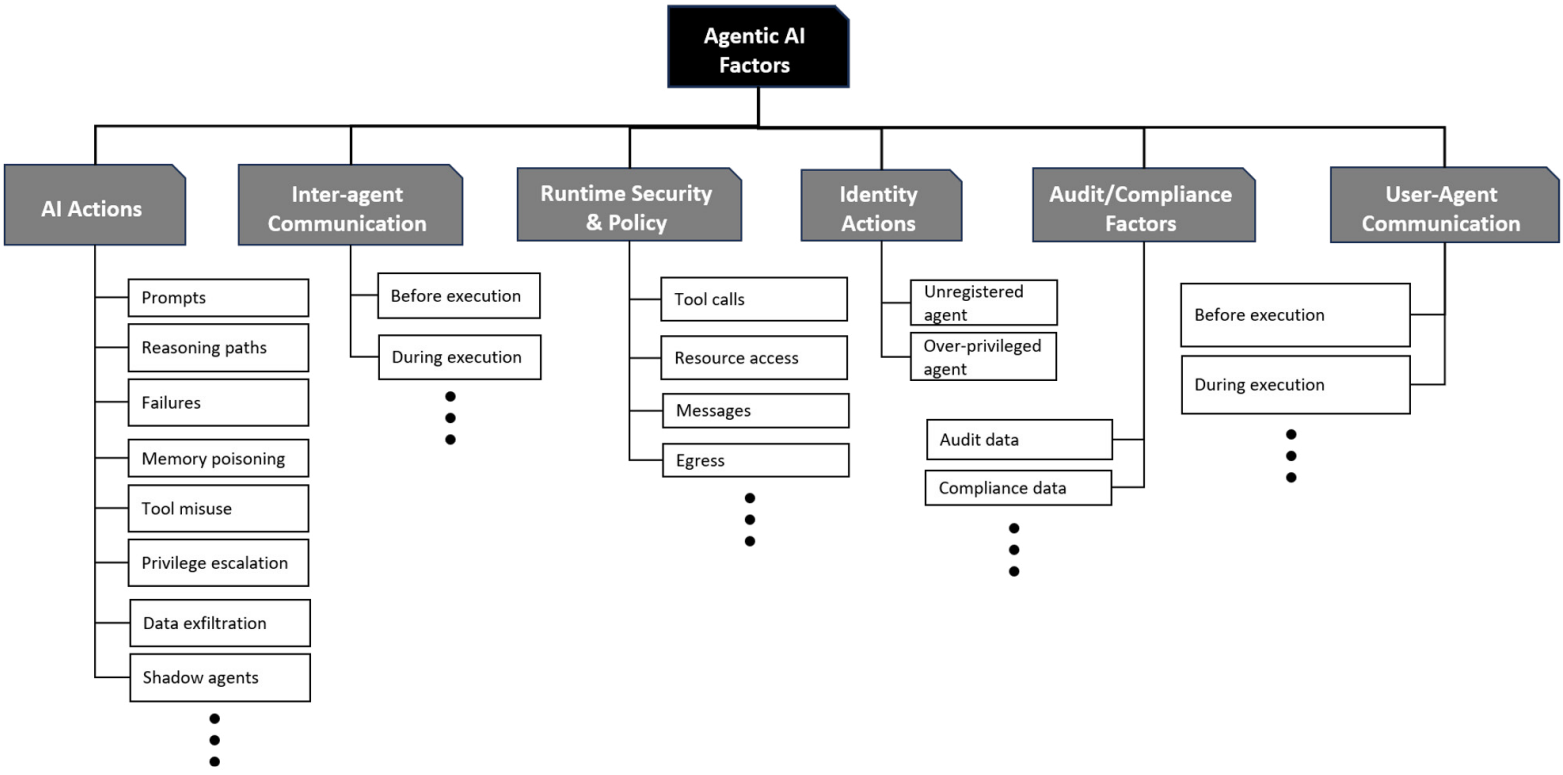


Figure 3. Notional Class/Subclass Structure for the Agentic AI Portion of an Expanded SOFIT Taxonomy

7.2 Operational Depth

- Cogility is the only company to have fully deployed whole-person IRM at large customers (1M+ people).
- The company has developed a well-tested and accepted development process for implementation.
- Whole-person IRM is still in its operational infancy, and it is the right approach to holistic AI agent risk management. All of Cogility's customers and prospects are familiar with the whole-person concept, which will also help Cogility's speed to market.

Conclusion and Path Forward

AI agents are becoming a new class of digital insiders. Organizations will need to monitor not only human employees, but also delegated, semi-autonomous actors operating under human authority. The most interesting risk object is often not the agent alone: It is the **human-agent-system triad**. That triad is where Cogility's "whole-person" strength can become a **whole-actor** or **whole-enterprise actor** strength. Cogility's advantage is not raw AI telemetry collection; it is fusing that telemetry with myriad data sources to create a long-lived risk indication and investigable history.

Given the broad industry acknowledgement of AI agents needing to be treated as "digital insiders" as reinforced by the recent White House Executive Order, there is little doubt that a large Agentic AI Governance/IRM market will quickly develop to fill this market need. The goal of this paper is for Cogility to share, with as broad an audience as possible, how the Cogynt.ai platform and our deep experience in whole-person IRM can help address this emerging challenge. We are seeking feedback and collaboration from data partners, policy experts, academic researchers, industry analysts, LLM providers, AI researchers, IAM vendors and consultants.

COGILITY

Visit **www.cogility.com/contact** to obtain more information and request a demo.

Cogility
15495 Sand Canyon Ave. #150
Irvine, CA 92618

sales@cogility.com
+1 949.398.0015

06/26