

amdocs STUDIOS

Agents Assuring Agents

Trust: the new currency of the agentic era



Autonomous AI needs continuous assurance

Enterprise AI is evolving from assistants into autonomous agents capable of reasoning, selecting tools, executing workflows, and making business decisions independently. As autonomy increases, organizations need continuous validation with end-to-end visibility into how agents think, act, and perform in production.

Can You Trust Your AI Agents?

Can your AI agents make business-critical decisions with consistent, explainable outcomes?

Do your agents work reliably in production – not just demos?

Can you detect reasoning failures before they impact your business?

Do you have continuous visibility into how your AI agents behave across the entire lifecycle?

How resilient are your AI agents against prompt injection, jailbreaks, and malicious inputs?

Would you know if your agents began drifting from expected behavior today?

Trust: Auditability, Observability, Explainability

As AI agents become more autonomous, Amdocs has created a continuous evaluation framework based on three pillars to validate how decisions are made, why they were made, and whether they can be trusted.

- **Auditability:** Trace every decision back to policies, guardrails and controls
- **Observability:** Continuously monitor agent behavior, execution, and drift in production
- **Explainability:** Understand why decisions were made to accelerate investigation and remediation

The Solution: Agents Assuring Agents

Amdocs' Agents Assuring Agents is an end-to-end intelligent assurance framework. It ensures agent plans and actions behave as expected by validating the quality of outputs and decisions, ensuring reliable interactions with external tools, APIs, and services, testing against adversarial attacks, security threats, and data leakage risks, and continuously validating operations in real-world production environments.

How It Works

1. Agent Onboarding

AI agents are connected to the evaluation framework through endpoint integration and define the business skills, KPIs, and risk profile that will drive the evaluation strategy.

2. Auditing Agents

Scenario Mutation and Red Teaming agents continuously test reasoning, execution, edge cases, and adversarial threats to uncover hidden risks before production.

3. AI Analytics & Governance

Evaluation results are analyzed for behavioral drift, validated against business KPIs and policies, and reviewed through governance controls with human oversight where required.

4. Continuous Evaluation

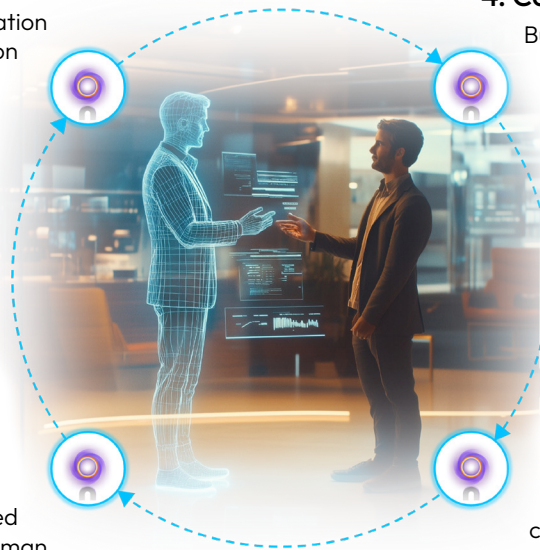
Business impact, KPI performance, drift detection, and feedback loops are continuously monitored to identify anomalies and maintain release confidence in production.

5. Business Intent

Business objectives, skills, and KPIs continuously feed the evaluation strategy, ensuring AI behavior remains aligned with evolving business priorities.

6. Alert & Adjust

Detected issues trigger alerts and corrective actions, creating a continuous assurance loop that keeps AI agents reliable as they evolve.



What We Measure

Reasoning Evaluation

Validate decision quality, reasoning accuracy, and successful task completion.

Key KPIs:

- Task Success Rate
- Goal Completion Rate
- Hallucination Rate

Integration Validation

Verify reliable tool selection, API execution, and multi-step workflow performance.

Key KPIs:

- Tool Call Accuracy
- Recovery Success Rate
- Output Validation Failure Rate

Red Teaming

Assess resilience against adversarial attacks, prompt injection, policy violations, and sensitive data exposure.

Key KPIs:

- Prompt Injection Success Rate
- Jailbreak Success Rate
- Data Leakage Rate

Resilience & Robustness

Monitor operational stability, behavioral drift, and enterprise-scale performance.

Key KPIs:

- Drift Detection Score
- Task Abandonment Rate
- Cost per Token

Measurable Impact

95%

task
success
rate

60%

reduction in
production
incidents

3x

faster agent
deployment

Build Trusted Autonomous AI

Autonomous AI should accelerate innovation, not introduce uncertainty. With Amdocs' Agents Assuring Agents, organizations can confidently scale autonomous operations while reducing risk, improving reliability, and maintaining responsible AI governance, in every stage of the agents' lifecycle.

Book a demo and discover how Amdocs helps leading enterprises deploy autonomous AI faster and increase their confidence, while reducing operating costs.

amdocs STUDIOS

Experience & Digital Engineering | Cloud | Data & GenAI | Quality Engineering

a | **make it
amazing**

www.amdocs.com

© 2026 Amdocs. All rights reserved.