

AI in SOC Examinations
Automation & Audit Judgment

October 2027

forvis mazars

AI in SOC Examinations

Meet the Presenters



Karen Cardillo
Managing Director, SOC & HITRUST

336.259.6611
karen.cardillo@us.forvismazars.com



Ryan Boggs
Principal, SOC & HITRUST

828.989.3176
ryan.boggs@us.forvismazars.com

© 2026 Forvis Mazars, LLP. All rights reserved.

Agenda

1. Why Automation Now
2. AICPA Q&A Scope & Relevance
3. Where AI Can Add Value
4. Keeping Judgment Central
5. IPE & Readiness
6. Q&A & Closing



© 2026 Forvis Mazars, LLP. All rights reserved.

AICPA Q&A Section 9561
What the New Guidance Addresses

Addresses

How a service organization's use of AI may affect a SOC 1 or SOC 2 examination.

Scope, system description, risks, controls, governance, subservice organizations, and engagement-team competence.

Does Not Address

A separate examination of an AI system or AI-specific functionality, such as model performance.

Specific guidance on how a service auditor should use AI to perform examination procedures.

Practical Relevance Test

Ask how AI is used, whether the organization relies on it, what risks it introduces or mitigates, and whether meaningful human review independently verifies the result.

© 2026 Forvis Mazars, LLP. All rights reserved.

01

Why Automation Now



The SOC Examination Experience Is Changing
Why Automation Now

Machine scale. Human accountability.

- SOC environments now generate more transactions, identities, configuration states, alerts, and evidence than manual methods can efficiently evaluate.
- Service organizations may use automation or AI within covered services, processes, and controls. Service auditors may separately use technology to perform examination procedures.
- AICPA Q&A Section 9561 addresses the first situation: the effect of the service organization's use of AI on a SOC 1 or SOC 2 examination.
- Technology can expand coverage, but management remains responsible for its system, controls, information, and remediation; the service auditor remains responsible for procedures and the opinion.



© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

SOC 1 & SOC 2
Different Subject Matter, Shared Evidence Risks

SOC 1

AI may be relevant when it's used in services or controls affecting user entities' ICFR.

Consider the role of AI, reliance on its output, independent human verification, related risks, and qualitative materiality.

AI use generally affects existing control objectives and may not require separate AI-specific objectives.

SOC 2

AI may be relevant when it affects covered services, applicable trust services categories, or service commitments and system requirements.

For AI-enabled services, weigh whether the selected categories meet intended-user needs. A security-only scope may not address a primary processing-integrity concern.

Shared Foundation

AI embedded in a process does not automatically make the AI component relevant.

Relevance increases when the organization relies on AI, AI introduces or mitigates covered risks, or AI is a key enabler of the service.

Existing trust services criteria remain applicable; additional or customized points of focus may be appropriate.

© 2026 Forvis Mazars, LLP. All rights reserved.

From Manual Evidence Requests
To Technology-Enabled Assurance

Manual Pattern

Screenshots, emailed exports, small samples, manual tie-outs, repeated follow-up, and point-in-time views.

Technology-Enabled Pattern

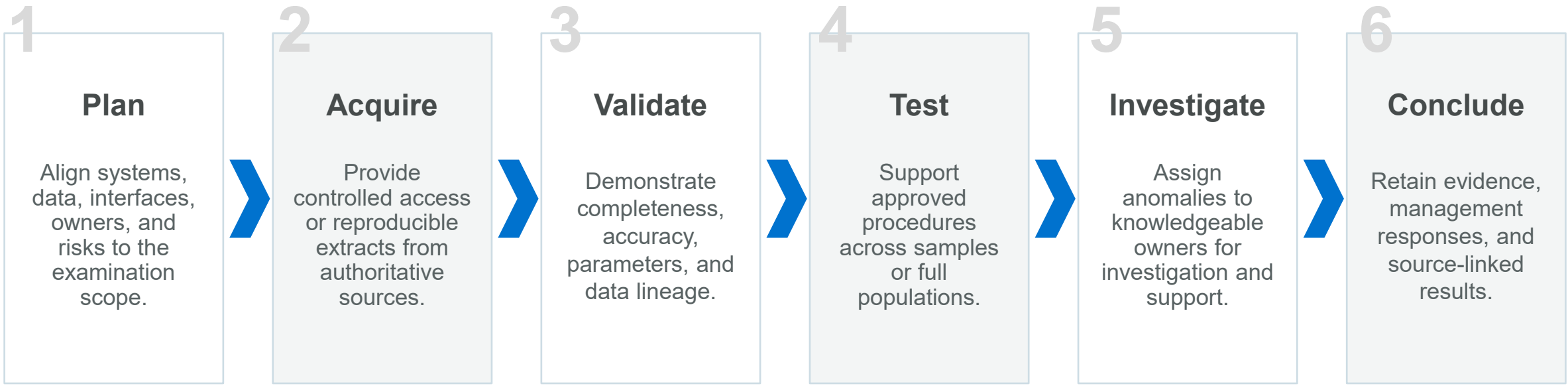
Read-only integrations, reproducible queries, automated comparisons, full-population analytics, retained run logs, and targeted human follow-up.

What Doesn’t Change

Service auditors still define procedures and reach the opinion. Management still owns control operation, information reliability, exception resolution, and support for the evidence provided.

© 2026 Forvis Mazars, LLP. All rights reserved.

Automation Across the SOC Examination Lifecycle



Automation can assist throughout the examination. Management remains responsible for its controls and information; the service auditor remains responsible for the procedures and opinion.

© 2026 Forvis Mazars, LLP. All rights reserved.

Automation, Analytics & AI

Rules-Based Automation

Repeatable deterministic logic for extracts, comparisons, recalculations, reconciliations, workflow routing, and evidence retention.

Analytics & Machine Learning

Analytics evaluate patterns at scale. Machine learning can identify clusters, outliers, or probabilistic signals.

Outputs should be treated as indicators requiring defined review, resolution, and monitoring for changing data or model behavior.

Generative AI & NLP

Generative AI and NLP can summarize, classify, search, or draft from unstructured evidence.

Management should require source anchoring, factual verification, access controls, human review, and safeguards against hallucination and prompt injection.

© 2026 Forvis Mazars, LLP. All rights reserved.

02

Where AI Can Add Value



Use Case 1
Three-Way Matching at Scale

Test more. Investigate smarter.

- Objective: determine whether purchase orders, invoices, and receiving or shipping records agree on key attributes such as vendor, item, quantity, price, authorization, and date.
- Automation can use optical character recognition and rules to extract fields, normalize formats, match documents, and flag missing or inconsistent attributes across the full population.
- Management should document the match logic, establish the completeness of source documents, assign exception ownership, and retain evidence of investigation and resolution.
- Higher coverage does not eliminate judgment. It shifts management effort from manual matching toward anomalies, root causes, recurring trends, and the risk of systematic failure.



© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 2
Full-Population Analytics or Testing

What Automation Can Do

Recalculate fields, test tolerances, identify duplicates, detect missing sequence numbers, compare effective dates, and profile unusual combinations across records in the population.

What Management Should Be Ready to Support

Demonstrate population completeness, explain transformations and control logic, define exception handling, and identify records the automated procedure could systematically miss.

What “No Exceptions” Means

Only that no items met the programmed criteria. Management still needs evidence that the population was complete, the logic addressed the risk, and results were appropriately reviewed.

© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 3
Configuration Baselining

Baseline, detect drift, investigate.

- Structured APIs can retrieve cloud, network, database, and application configurations directly from authoritative sources.
- Automated logic can compare settings to approved baselines, customer commitments, internal standards, or relevant benchmarks.
- Examples include encryption status, administrative MFA, public access restrictions, log retention, backup settings, and security group rules.
- Service organizations should document scope, account coverage, inherited settings, approved exceptions, compensating controls, and how each baseline supports SOC commitments and criteria.



© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

Use Case 4

Access & Privileged Activity Analytics

Population & Lifecycle

Analyze users, roles, entitlements, service accounts, dormant accounts, provisioning events, terminations, and periodic access reviews.

Behavioral Signals

Surface improbable travel, unusual access times, unexpected privilege escalation, atypical query activity, or access inconsistent with peer groups.

Management Response

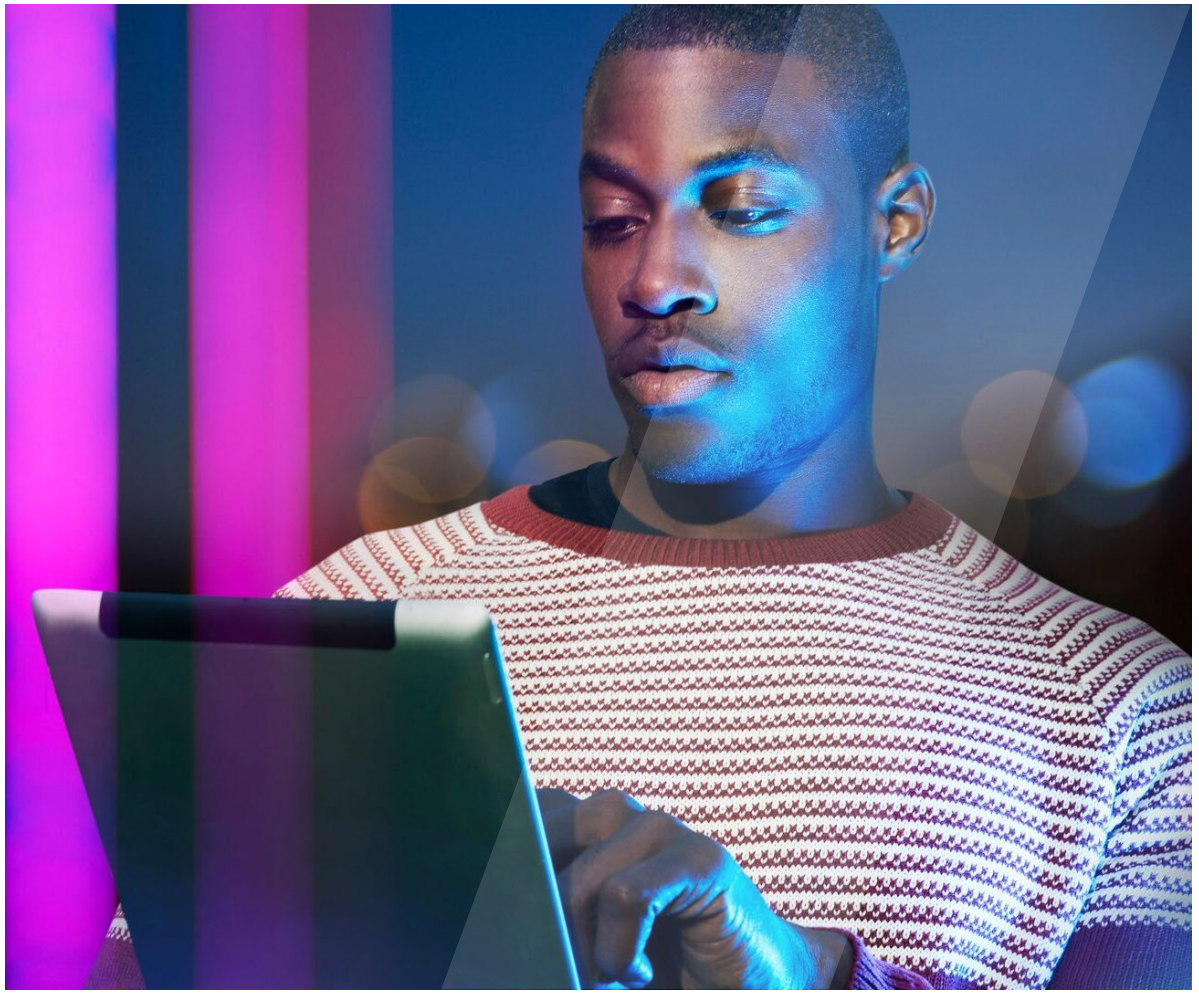
Investigate signals using tickets, approvals, HR data, role definitions, and incident records; document disposition, escalation, remediation, and retained support.

© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 5
Change Management Traceability

Support end-to-end change traceability.

- Reconcile source-code commits, pull requests, change tickets, approvals, testing evidence, deployment logs, and production events.
- Identify changes that bypassed review, were deployed without required testing, lack a linked ticket, or followed an unusual sequence.
- Full-population tracing can reveal systematic workflow issues that a small sample may not expose.
- Management review remains essential for emergency changes, inherited approvals, automated deployments, false matches, and the remediation and significance of identified deviations.



© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 6
Automated Log Correlation

Connect Operational Evidence

Correlate SIEM alerts, endpoint telemetry, cloud logs, case-management records, incident tickets, and response timestamps.

Identify Control-Performance Gaps

Identify alerts without tickets, delayed responses, missing escalation, or activity that should have generated an alert; then assign and track corrective action.

Validate & Govern the Result

Management should validate logging coverage, clock synchronization, retention, parsing, suppression rules, alert closure, and the completeness of in-scope detection scenarios.

© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 7
NLP for Documents

Search meaning, verify the source.

- Natural Language Processing (NLP) can search contracts, policies, meeting minutes, incident narratives, and change documentation for concepts rather than exact words.
- Potential uses include identifying agreements without required security clauses, locating policy exceptions, classifying incidents, and comparing documented requirements with control descriptions.
- The tool may accelerate review of hundreds of documents, but management should require verification of semantic matches and omissions against original documents.
- Retain prompts or queries, versions, results, reviewed exceptions, and links to source evidence.



© 2026 Forvis Mazars, LLP. All rights reserved.

Use Case 8
API-Driven Evidence Integration

Potential Benefit

Secure, read-only connections can collect evidence from authoritative sources, reducing screenshots and repeated requests. Direct extraction does not by itself establish reliability.

What Management Should Validate

Document connector scope, credentials, endpoints, query logic, pagination, filters, timestamps, transformations, and retention of raw results.

Operating Model

Define owners for connection setup, changes, failures, exception handling, periodic recertification, and timely access to reproducible outputs.

© 2026 Forvis Mazars, LLP. All rights reserved.

When Automation May Not Eliminate Sampling

- Service organizations should not assume every SOC procedure will become full-population testing. Judgmental, nonstandard, observational, or inaccessible evidence may still require sampling.
- Sampling may remain appropriate for walkthroughs, management review controls, physical or observational procedures, complex approvals, and heterogeneous transactions.
- Automation can still reduce disruption by validating populations, stratifying risk, identifying unusual items, and making requested samples more targeted.
- The service auditor’s procedure will remain driven by risk and the evidence objective, not simply by the availability of a tool.

© 2026 Forvis Mazars, LLP. All rights reserved.

03

Keeping Judgment Central



The Non-Negotiables



Human Accountability & Competence

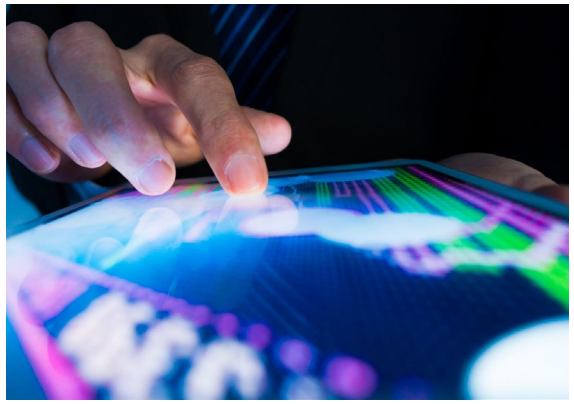
Qualified personnel should review outputs, resolve contradictions, measure significance, and take responsibility for management decisions and remediation.

- When AI is relevant, the service-auditor team should collectively have appropriate AI competencies or use a properly overseen specialist.



Evidence Traceability

Link automated results and management dispositions to original records, retained queries, control activities, and documented follow-up.



Model & Logic Support

Maintain support for design, assumptions, thresholds, configurations, inputs, outputs, limitations, changes, and the ability to detect known exceptions.



What Skepticism May Mean

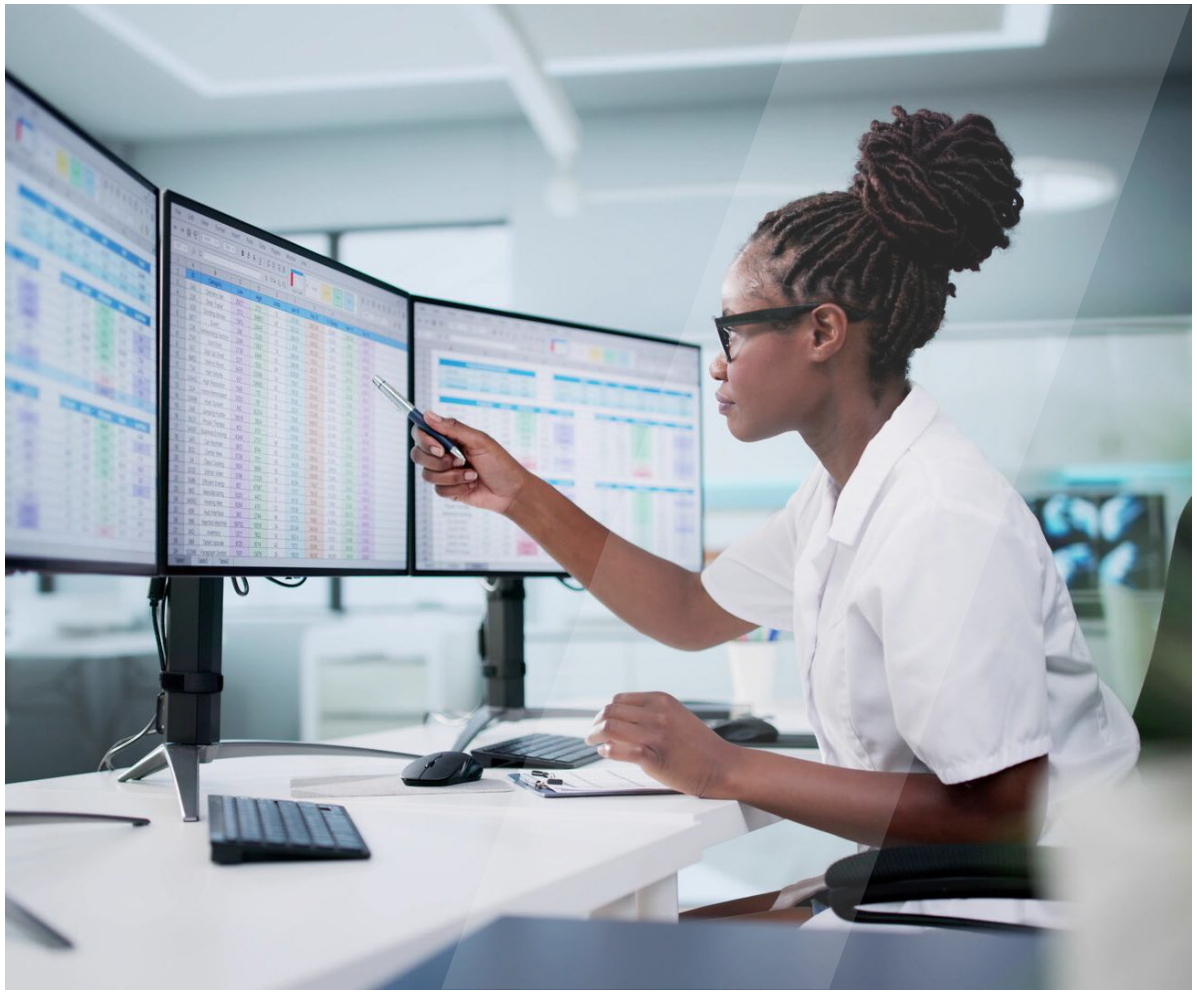
Expect service auditors to question whether automated results are plausible, complete, consistent with other evidence, and responsive to the identified risk.

© 2026 Forvis Mazars, LLP. All rights reserved.

How Management Can Support Professional Skepticism
Keeping Judgment Central

A result is not a conclusion.

- Start with the risk and control objective. Be prepared to explain why the automated process addresses that objective rather than merely describing the tool's features.
- Define what normal and abnormal results should look like before relying on automated output.
- Investigate an unexpected absence of exceptions as carefully as an unexpectedly high exception rate.
- Compare automated results with prior findings, control-owner knowledge, system changes, operational metrics, and contradictory or corroborating information.
- Document why management considers the result reliable, not merely that the system produced it.



© 2026 Forvis Mazars, LLP. All rights reserved.

Model, Rule & Tool Validation
What Management Should Retain

Keeping Judgment Central

1

Define Intended Use

- What control, monitoring activity, or business decision will the technology support? Is the organization relying on AI or independently verifying the result?

2

Inspect Design

- Document inputs, transformations, rules, assumptions, thresholds, model or foundation-model versions, limitations, and explainability constraints.

3

Challenge With Known Cases

- Use controlled test data, historical exceptions, and independently constructed cases; evaluate false positives, false negatives, hallucinations, and systematic exclusions.

4

Evaluate Errors

- Assess false positives, false negatives, systematic exclusions, and sensitivity.

5

Assign Ownership & Monitor

- Document acceptance criteria, change control, ownership, drift monitoring, periodic validation, and escalation for unexpected behavior.

© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

Over-Reliance Risks to Watch

Automation Bias

Treating system output as inherently objective or superior to contradictory evidence.

Data, Model, & Drift Risk

Incomplete, outdated, unrepresentative, or drifting data; model error, misuse, concept drift, or continuous-learning changes.

Scope Mismatch

Providing only the records available to the tool rather than demonstrating the complete in-scope population.

Explainability & Security

Inability to explain or interpret outputs; prompt injection, unauthorized actions, data modification, or exposure of confidential or personal information.

Unsupported Generated Content

Allowing hallucinated or biased content to introduce unsupported facts, controls, causes, decisions, or management assertions.

© 2026 Forvis Mazars, LLP. All rights reserved.

Source Anchoring
What Service Organizations Should Preserve

Keeping Judgment Central

1
AI Output

- Summary, match, score, anomaly, or draft

2
Automation Record

- Query, prompt, rules, parameters, version, and run time

3
Retained Result

- Population, exception file, logs, and disposition

4
Original Source

- Ticket, configuration, transaction, contract, or system log

5
Service Auditor Conclusion

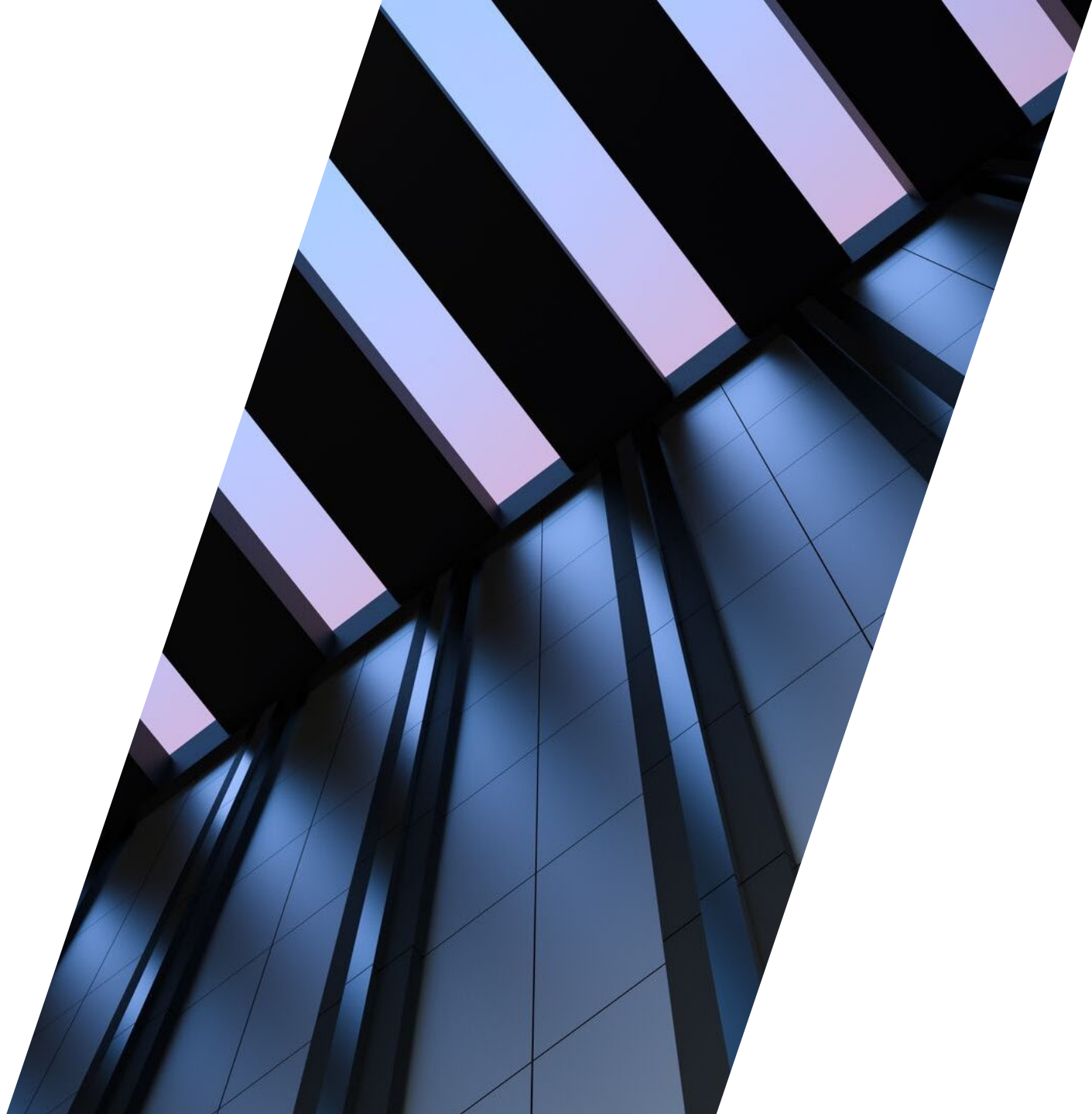
- Evaluation of sufficiency, significance, and the effect on the opinion

If management and the service auditor cannot navigate from automated output to the original evidence, automation has weakened rather than strengthened the audit trail.

© 2026 Forvis Mazars, LLP. All rights reserved.

04

IPE & Readiness



IPE Is Foundational, but Not the Entire AI Evaluation

Scalable testing depends on complete, accurate inputs.

- Reliable inputs are necessary before scalable testing, but AI relevance may also require evaluation of model design, assumptions, governance, monitoring, security, explainability, human involvement, and changes in data or model behavior.
- Completeness asks whether all relevant records, fields, periods, systems, accounts, event types, and data sources are included.
- Accuracy asks whether information correctly reflects underlying activity, including transformations, mappings, calculations, timestamps, and manual adjustments.
- A polished dashboard, API response, or AI-generated exception file does not by itself establish reliability.



© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

IPE Across Four Production Paths

IT Applications

System-generated reports, database queries, APIs, configurations, logs, and calculations. Management should document logic, parameters, access, changes, and source data.

End-User Computing

Spreadsheets, scripts, macros, low-code tools, and local models. Management should control formulas, transformations, versions, access, reviews, and retained inputs.

Service Organizations

Reports or data from cloud providers, processors, vendors, and subservice organizations. Management should evaluate scope, interfaces, relevant controls, and monitoring.

Manual Processes

Lists, case files, classifications, and adjustments prepared by people. Management should define instructions, competence, review, completeness, and reconciliation to source records.



© 2026 Forvis Mazars, LLP. All rights reserved.

How Management Can Support Completeness of AI Inputs

- Reconcile record counts and control totals to authoritative systems or independently generated expectations.
- Validate date and time boundaries, time zones, fiscal-period cutoffs, and whether late-arriving events are captured.
- Confirm that pagination, API limits, query filters, archived data, deleted records, and error handling do not omit items.
- Include relevant privileged, service, system, inactive, and external accounts where applicable.
- Reconcile data across interfaces and investigate unmatched or rejected records.
- Retain the raw extract and evidence of the parameters used to generate it.

© 2026 Forvis Mazars, LLP. All rights reserved.

How Management Can Support Accuracy of AI Inputs

- Trace selected fields to original records and recalculate derived values or transformations.
- Inspect report definitions, query logic, mappings, joins, filters, calculated fields, and exception rules.
- Reconcile totals to reliable source data, financial records, or independently generated reports.
- Validate timestamps, status fields, unique identifiers, currency or unit conversions, and duplicate handling.
- Assess whether manual edits, spreadsheet steps, or data enrichment can change the output after extraction.
- Reperform key logic when practical and compare results.

© 2026 Forvis Mazars, LLP. All rights reserved.

IPE Example
API Evidence from a Cloud Platform

Risk

The connector retrieves only one subscription, omits paginated results, applies an undocumented filter, or transforms inherited settings incorrectly.

Management Validation

Inventory in-scope accounts; inspect authentication and endpoints; confirm pagination and time filters; compare object counts; verify parameters; and trace selected configurations to the console or source API.

Evidence to Retain

Retain connector version, query definition, run time, scope, raw response, transformation logic, completeness reconciliation, exceptions, and reviewer sign-off.

© 2026 Forvis Mazars, LLP. All rights reserved.

IPE Example
Spreadsheet Feeding an AI Procedure

Risk

A control owner filters, enriches, or combines exports in Excel before upload. Rows, formulas, mappings, or tabs may be omitted or changed.

Management Validation

Reconcile each source extract, inspect formulas and macros, test joins and mappings, review hidden rows or columns, compare row counts, and restrict or log changes.

Better Design

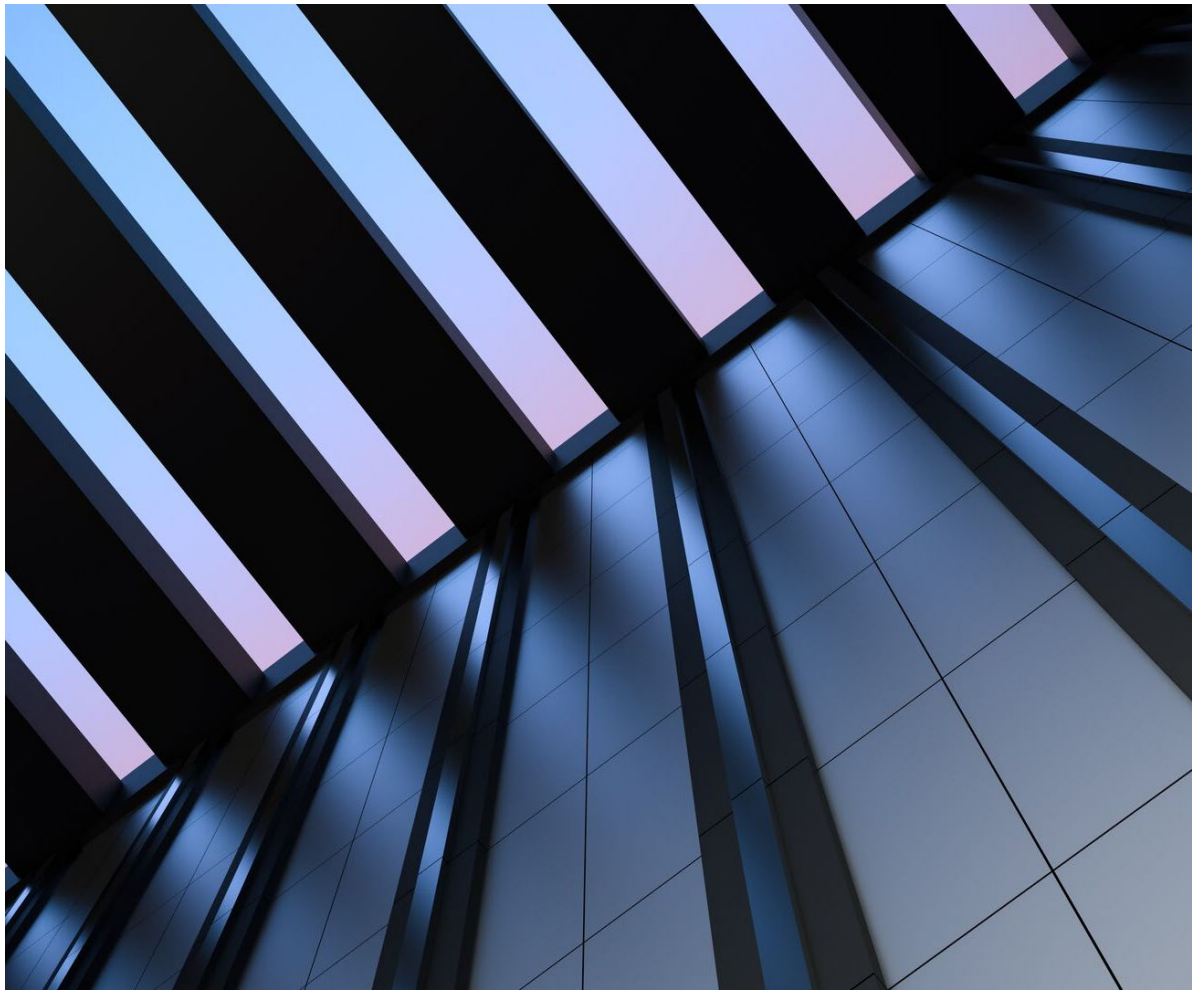
Load raw system extracts directly into a controlled pipeline; retain transformations as code; generate repeatable outputs; and route exceptions for documented human review.

© 2026 Forvis Mazars, LLP. All rights reserved.

IPE Example
Vendor or Service-Organization Data

Third-party data needs first-party validation.

- Document what the vendor generates, what your organization changes, and which party owns each completeness and accuracy control.
- Determine whether the vendor's SOC Report covers the relevant service, system, period, locations, interfaces, and control activities.
- Identify and operate complementary user entity controls, and understand complementary subservice-organization controls affecting data reliability.
- Reconcile inbound and outbound record counts, rejected transactions, transmission logs, and interface exceptions.
- Do not assume a third-party platform or vendor algorithm is reliable simply because it's commercially available.



© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

A Controlled Automation Operating Model

Govern

Approved use cases, accountable owners, permitted tools, AI inventory, shadow-AI detection, data restrictions, governance oversight, competence, third-party AI, and escalation criteria.

Build & Validate

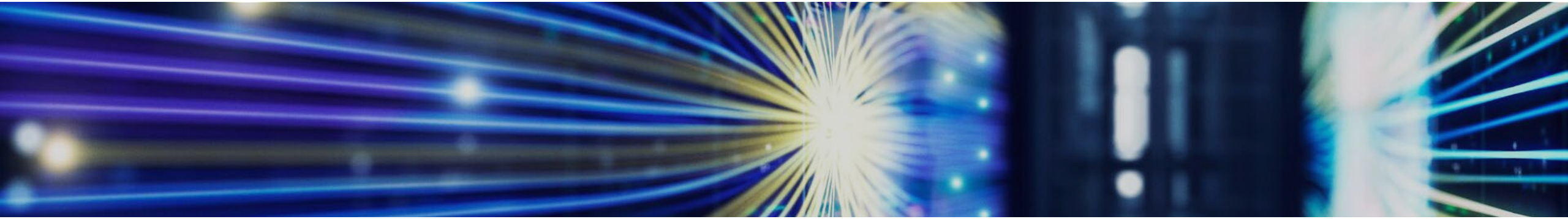
Document objectives, reliance, source data, training or retrieval data, logic, assumptions, thresholds, versions, known limitations, controlled cases, and acceptance criteria.

Operate & Monitor

Restrict changes; retain versions and run logs; monitor failures, data drift, concept drift, model updates, overrides, and unexpected behavior; periodically revalidate.

Support Assurance

Validate IPE, reproduce procedures, review outputs, resolve exceptions, document human judgments, assess contradictory evidence, and link results to original evidence.



© 2026 Forvis Mazars, LLP. All rights reserved.

What Service Organizations Should Prepare Now

1. Determine which AI uses are relevant based on the service provided, organizational reliance, human verification processes, covered risks, and intended-user needs.
2. When AI is relevant, update the system description to explain how AI is used, the types of AI, governance, human-in-the-loop controls, key inputs and assumptions, and relevant CUECs or CSOCs.
3. Document report logic, filters, calculations, interfaces, data lineage, training or retrieval sources, thresholds, known limitations, and model changes.
4. Assign accountable owners and establish reproducible evidence for validation, monitoring, exceptions, overrides, third-party AI, and remediation.
5. Consider whether SOC 2 category selection continues to meet intended-user needs and whether engagement-team AI competence or specialist involvement is needed.

© 2026 Forvis Mazars, LLP. All rights reserved.

Readiness Checklist
Before the Examination Begins

- ✓ Scope and data lineage documented
- ✓ Automated logic approved and versioned
- ✓ Known-anomaly validation completed
- ✓ IPE completeness and accuracy evidenced
- ✓ Access and changes to tooling controlled
- ✓ Failures and overrides logged
- ✓ Exceptions assigned and resolved
- ✓ Raw data and run results retained
- ✓ Conclusions linked to source evidence
- ✓ Control owners can explain the process

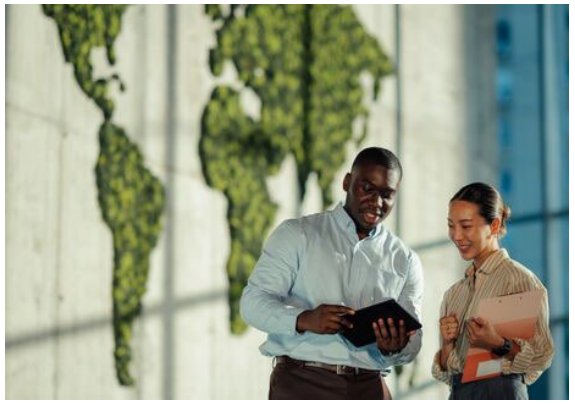
© 2026 Forvis Mazars, LLP. All rights reserved.

Key Takeaways



Automate the Routine

Use technology where work is repetitive, data-heavy, and reproducible. This can expand coverage and reduce disruption.



Keep Judgment Central

Professional skepticism and responsibility for the opinion remain with the service auditor. Control ownership, information reliability, and remediation remain with management.



Validate the Information

Completeness and accuracy of IPE apply to every input and transformation, including APIs, spreadsheets, vendor data, and manual processes.



Make Results Traceable

Retain parameters, outputs, exception disposition, management review, and direct links to the original evidence.

© 2026 Forvis Mazars, LLP. All rights reserved.

Q&A



SOC & HITRUST® Solutions

Questions?

© 2026 Forvis Mazars, LLP. All rights reserved.

forvis mazars

Contact

Forvis Mazars

Karen Cardillo
Managing Director
336.259.6611
karen.cardillo@us.forvismazars.com

Ryan Boggs
Principal
828.989.3176
ryan.boggs@us.forvismazars.com

Scan to learn more about our SOC & HITRUST® Services.

The information set forth in this presentation contains the analysis and conclusions of the author(s) based upon his/her/their research and analysis of industry information and legal authorities. Such analysis and conclusions should not be deemed opinions or conclusions by Forvis Mazars or the author(s) as to any individual situation as situations are fact-specific. The reader should perform their own analysis and form their own conclusions regarding any specific situation. Further, the author(s)' conclusions may be revised without notice with or without changes in industry information and legal authorities.

© 2026 Forvis Mazars, LLP. All rights reserved.