International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Conversational ERP in SAP Operations

Gururaj Veershetty

SAP Consulting Manager
gururaj.veershetty@outlook.com

Abstract:
Conversational ERP the use of natural-language assistants to query, interpret, and execute SAP business processes is advancing rapidly from pilots to production. This whitepaper presents a comprehensive governance and control framework: the Read–Advise–Act capability model, an eight-layer SAP-native control stack, four proven deployment patterns, a minimum viable operating model, and a phased implementation roadmap. The central argument is that a GenAI assistant in SAP operations is a governed control surface not a UI enhancement and must be architected accordingly.

Keywords: conversational ERP; generative AI governance; SAP operations; Read–Advise–Act model; intent gating; prompt injection; auditability; enterprise control stack; GRC; AI-enabled ERP.

1. WHAT CONVERSATIONAL ERP MEANS IN SAP OPERATIONS

Conversational ERP introduces a new operational interface through which users interact with enterprise business processes using unstructured language and through which the system may act on their behalf. Capabilities evolve across three bands with materially different governance implications [1, 2].

Figure 1: Read-Advise-Act Capability Model

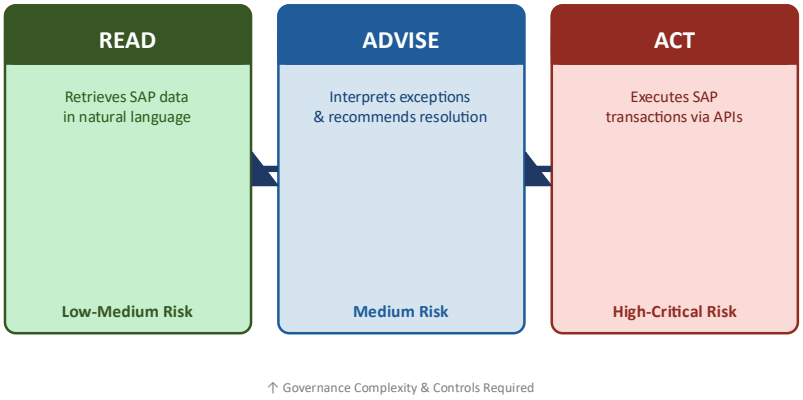

Figure 1: The Read–Advise–Act Capability Model governance complexity increases non-linearly across bands.

READ capabilities carry the lowest risk but remain vulnerable to over-broad data retrieval if field-level security is not enforced at the service layer. ADVISE capabilities require hallucination controls and evidence capture of what the assistant recommended and why. ACT capabilities introduce the full spectrum of enterprise risk unauthorized postings, SoD violations, duplicate transactions and require a categorically different governance posture [3].

A critical failure pattern in conversational ERP pilots is the assumption that 'just connecting the model to SAP' is a sufficient architecture. Each capability band must be treated as a distinct governance product with distinct controls, approvals, and evidence requirements.

2. WHY GENAI GOVERNANCE IS HARDER THAN CLASSIC SAP AUTOMATION

Classical SAP automation is deterministic a given input produces a predictable, auditable output. Generative AI is probabilistic, context-sensitive, and capable of outputs that were never explicitly programmed [4]. This introduces five failure modes with no direct precedent in traditional ERP automation:

- **Over-broad data retrieval** the assistant surfaces sensitive fields (vendor bank details, PII, financial margin) when API payloads are wider than the user's actual request.
- **Prompt injection** adversarial content embedded in SAP notes, POs, or emails attempts to override system instructions and redirect the assistant's behavior [5].
- **Hallucinated procedures and false confirmations** the assistant confidently describes incorrect resolution paths or asserts that a posting succeeded when it did not.
- **Authorization confusion** if the assistant acts via a broad service account rather than the requesting user's identity, RBAC controls are effectively bypassed [6].
- **Scale risk** a single flawed guardrail can propagate across thousands of conversations before detection, far exceeding the blast radius of any human error [7].

These failure modes translate into three non-negotiable governance imperatives: capability containment (architecturally enforced boundaries, not system prompt instructions), policy enforcement (authorization validated in SAP, not only in middleware), and evidence generation (an immutable, correlated audit trail sufficient for internal audit and regulators) [8].

3. GOVERNANCE MODEL: READ / ADVISE / ACT WITH RISK TIERS

The governance model combines two orthogonal dimensions: the capability band (READ, ADVISE, or ACT) and the risk tier of the specific intent. Together these define a governance matrix with proportionate controls for each combination [9].

Table 1: Risk Tier Governance Matrix capability band versus risk tier with required controls per combination.

Tier	Risk	Example Intents / Objects	Required Controls
0	Low	PO status, delivery tracking, open item counts, reference master data	Standard auth, field minimization, audit logging
1	Medium	Invoice block explanation, update comments, attach documents, resend confirmations	Intent gating, SAP auth, interaction logging
2	High	Invoice release, GR posting, vendor master changes, payment approvals	Maker-checker, SoD check, propose-then-execute, enhanced audit trail
3	Restricted	HR/payroll, vendor bank accounts, pricing/margin, system config, security roles	Default deny; step-up auth + GRC/legal approval required

Policy sequence: Tier 0–1 READ and ADVISE should be the starting point for all deployments. Tier 2–3 READ requires proven data minimization architecture. Tier 0–1 ACT can be enabled once propose-then-execute and SAP authorization are validated in production. Tier 2 ACT requires a full maker-checker via SAP workflow. Tier 3 actions default-deny across all capability bands and require formal GRC and legal sign-off before any enablement.

> *Most organizations should spend the first six to twelve months operating exclusively in Tier 0–1 READ and ADVISE. This generates business value, validates the control stack, and builds the evidence base required to responsibly expand into ACT.*

4. THE EIGHT-LAYER SAP-NATIVE CONTROL STACK

The following control stack provides the architectural backbone of a secure, auditable conversational ERP deployment [10]. The layers are cumulative each depends on those below it, and omitting any layer creates exploitable gaps.

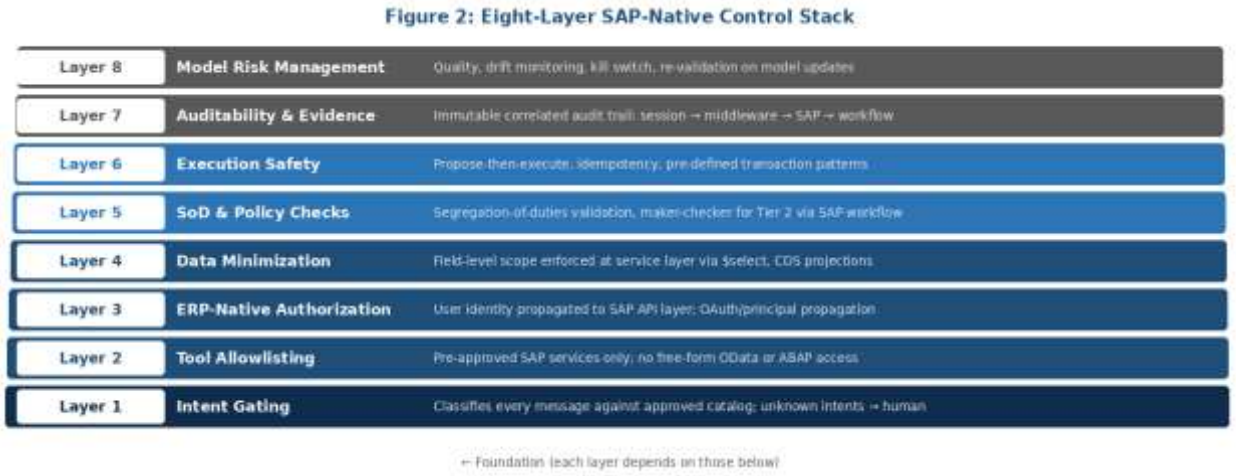

Figure 2: The eight-layer control stack each layer addresses a distinct failure mode and contributes to the evidence chain.

Table 2: Eight-Layer SAP-Native Control Stack each layer addresses a distinct failure mode in conversational ERP deployments.

#	Layer	What It Controls
Layer 1	Intent Gating	Classifies every message against an approved intent catalog; unknown intents route to human never best-effort execution.
Layer 2	Tool Allowlisting	Restricts to pre-approved SAP services (OData, RAP, BAPI wrappers); no free-form queries or direct ABAP access.
Layer 3	ERP-Native Authorization	User identity propagated to SAP API layer via OAuth 2.0 / principal propagation; assistant never acts 'as system.'
Layer 4	Data Minimization	Field scope enforced at service layer using $select, CDS projections; sensitive fields excluded structurally, not by AI redaction.
Layer 5	SoD & Policy Checks	SoD validation for high-risk action combinations; maker-checker for Tier 2 via SAP workflow; posting period and tolerance controls.

Layer 6	Execution Safety	Propose-then-execute with explicit user confirmation; idempotency keys prevent duplicate postings; pre-defined transaction patterns only.
Layer 7	Auditability & Evidence	Immutable correlated audit trail: session → middleware → SAP posting → workflow, with correlation IDs across all hops.
Layer 8	Model Risk Management	Acceptance testing per intent, continuous anomaly monitoring, kill switch for ACT (minutes to disable), model update re-validation.

Layer 1 (Intent Gating) is the most fundamental control: every message must be classified against a pre-approved catalog before any action is taken. Unknown intents must route to a human never default to best-effort execution. Layer 3 (ERP-Native Authorization) enforces that authentication at the AI or middleware layer confers no SAP authorization every object accessed and every transaction executed must produce a corresponding SAP authorization check result. This is what makes the audit trail defensible to regulators [11].

Layer 7 (Auditability) requires correlation IDs that tie together the conversational session, middleware call, SAP posting, and workflow event. Layer 8 (Model Risk Management) includes a kill switch that disables ACT across all intents within minutes this is a prerequisite for responsible ACT enablement, not an optional enhancement [12].

5. FOUR CONTROL PATTERNS THAT WORK IN SAP OPERATIONS

Pattern A Read-First Rollout with Operational Triage

Target the highest-volume, lowest-risk operational pain points: invoice block reason lookup (read-only), delivery and ATP status explanations, GR/IR mismatch investigation, and exception runbook guidance. These intents generate immediate value while proving foundational controls before any execution risk is introduced [13].

Pattern B Exception Explanation + Guided Resolution (No Direct Posting)

Extend into the ADVISE band: explain the block reason in business-readable language, identify the missing master data element, propose the resolution path, and generate a pre-populated ticket for the responsible party. Documented implementations show 40–60% reduction in exception resolution handling time with zero posting risk [14].

Pattern C Draft-and-Approve (Maker-Checker by Design)

When ACT is enabled for Tier 1 or Tier 2 actions, the assistant operates as the maker in a maker-checker model. It validates the intent, performs all pre-execution checks (authorization, SoD, posting period, tolerance), presents a structured propose-then-execute summary, and upon confirmation creates a pending posting or SAP workflow task. An authorized approver reviews and approves in SAP the assistant never bypasses an approval step.

Pattern D Document Delivery with Re-Authorization

Re-authorizes the requesting user's access to the specific document inside SAP at the moment of the request; generates the document through SAP output management (not a file cache); delivers via a time-limited authenticated link; and captures every access event as an audit record with user identity, document ID, and timestamp [15].

6. OPERATING MODEL: OWNERSHIP AND GOVERNANCE CADENCE

A technically sound control stack degrades without defined human ownership. Six functional domains must be covered and assigned before enabling production ACT capabilities [16].

Figure 3: Governance Ownership Model

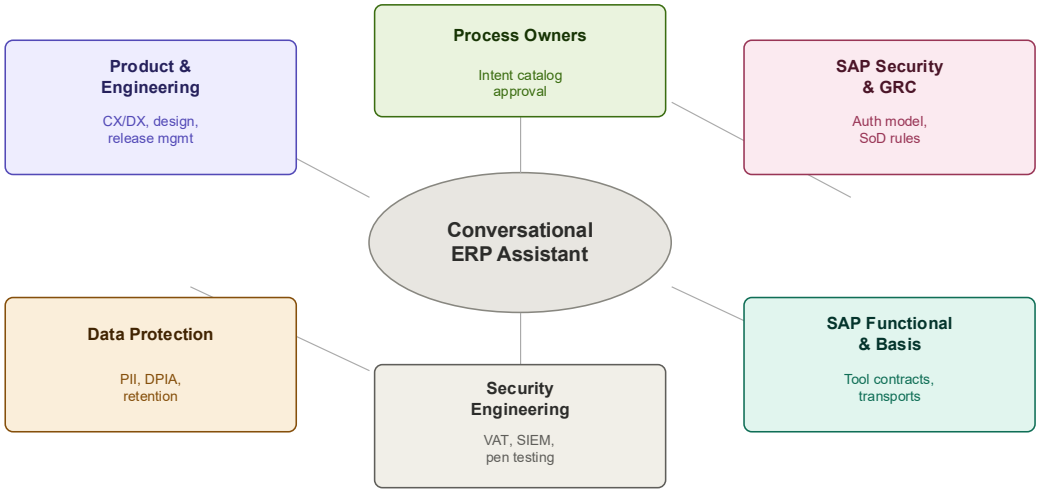

Figure 3: Six-domain governance ownership model every domain must be assigned before ACT enablement.

- **Process Owners (Finance, Supply Chain)** own and approve the intent catalog; define correct behavior and escalation thresholds; participate in weekly intent review.
- **SAP Security & GRC** own the authorization model, SoD rules for conversational intents, and integration of conversational audit logs into the enterprise GRC framework.
- **SAP Functional & Basis/Platform** design and maintain tool contracts (OData services, CDS views, BAPI wrappers); own transport and change management for tool updates.
- **Security Engineering** own channel-layer controls: WAF configuration, secrets management, SIEM integration for anomaly alerts, and penetration testing coverage.
- **Data Protection** own PII field classification, conversation log retention/deletion policies, DPIA requirements, and audit log redaction standards.
- **Product & Engineering** own conversation design, CI/CD pipeline, and the release management process that requires governance sign-off before production deployment.

Minimum viable governance cadence:

- **Weekly intent review** new intent requests, unknown-intent fallback cases, anomalies. Attendees: Product, Process Owners.
- **Monthly control review** audit log completeness, SoD exception rates, authorization failure trends, model and tool contract changes. Attendees: SAP Security/GRC, Security Engineering, Product.
- **Quarterly evidence and access review** audit trail sampling, service account access review, SoD exception approvals, control attestation for Internal Audit. Attendees: GRC, Internal Audit, Data Protection.

7. METRICS: VALUE AND CONTROL HEALTH

Value Metrics (Reported to Business Leadership)

- **Mean Time to Resolve (MTTR)** reduction in resolution time for invoice blocks, delivery exceptions, and GR/IR mismatches handled through the conversational channel versus prior baseline.

- **Deflection rate** percentage of queries resolved without escalation to a human ticket. Target: 50–70% for Tier 0–1 READ intents in mature deployments [17].
- **Intent completion rate** rate at which users successfully complete their intended task without abandonment; low rates are diagnostic of hallucination, poor disambiguation, or incorrect data scope.

Control Health Metrics (Reported to GRC, Internal Audit, CISO)

- **Authorization failure rate** elevated baseline indicates misconfigured roles; anomalous spikes indicate attempted privilege escalation or prompt injection.
- **Responses exceeding approved field set** should be zero; any non-zero value indicates a data minimization control failure requiring immediate investigation.
- **Unknown intent rate and safe fallback compliance** rate of unrecognized intents, and the rate at which they correctly trigger human routing not best-effort execution.
- **Audit log correlation ID coverage** percentage of interactions producing a complete, correlated record across the full interaction chain. Gaps are evidence chain failures.

8. PHASED IMPLEMENTATION ROADMAP

The roadmap below sequences capability expansion against control maturity, ensuring governance controls are proven at each tier before the next is enabled [18]. ACT capabilities in Phase 3 require formal GRC and Internal Audit sign-off this is not a procedural courtesy but an architectural gate.

Figure 4: Phased Implementation Roadmap

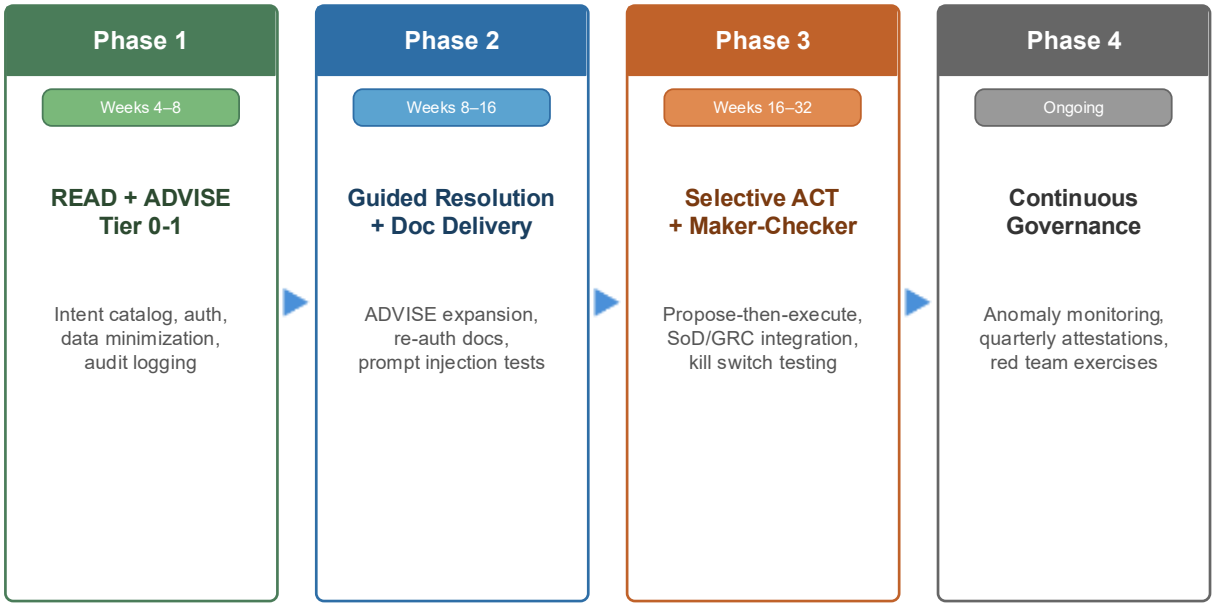

Figure 4: Four-phase implementation roadmap governance controls are proven before capability is expanded.

Table 3: Four-Phase Implementation Roadmap governance controls are validated before capability expansion at each phase.

Phase	Timeline	Key Deliverables

Phase 1: READ + ADVISE Tier 0–1	Wks 4–8	Intent catalog (5–10 intents), tool allowlist, SAP auth via principal propagation, data minimization views, immutable audit logging, negative test suite, process owner sign-off.
Phase 2: Guided Resolution + Docs	Wks 8–16	Complex ADVISE intents, re-authorized document delivery with time-limited links, prompt injection testing, ticketing/workflow integration, first monthly control attestation.
Phase 3: Selective ACT	Wks 16–32	Propose-then-execute + idempotency for Tier 1 ACT; SAP workflow maker-checker for Tier 2; SoD/GRC integration; kill switch tested in production; GRC and Internal Audit sign-off.
Phase 4: Continuous Governance	Ongoing	Anomaly monitoring with alert SLAs, quarterly evidence sampling and attestations, model update re-validation process, red team exercises, controlled intent expansion.

CONCLUSION

Conversational ERP will become a standard operational interface for SAP environments. The productivity gains are real and measurable. What remains in question is whether organizations will govern this capability responsibly [19].

Organizations that have treated conversational ERP as a UI enhancement connecting a model to SAP without an intent catalog, without ERP-native authorization enforcement, without data minimization, without an audit trail have encountered exactly the failure modes described here: data overexposure, hallucinated confirmations, SoD violations, and the absence of audit evidence. Organizations that build on ERP-native governance principles have captured the productivity benefits safely and are now expanding with confidence.

> *The winning pattern is ERP-native governance: intent gating, tool allowlisting, SAP-enforced authorization, data minimization, SoD enforcement, execution safety, immutable auditability, and continuous model risk management. Every one of these layers is necessary. None is optional.*

APPENDIX: GOVERNANCE READINESS CHECKLIST

Organizations should be able to answer 'yes, with evidence' to each item before enabling ACT capabilities in production.

Intent and Capability Governance

- Formally documented and approved intent catalog; each intent mapped to a capability band and risk tier.
- Defined, tested behavior for unknown intents: human routing not best-effort execution.
- Intent catalog version-controlled with an approval process for new additions.

Authorization and Access

- Requesting user identity propagated to the SAP application layer for every API call; authorization enforced at SAP layer.
- Formal access review conducted for any service accounts used in the integration pattern.
- SoD rules for conversational intents defined and integrated with SAP GRC.

Data Minimization and Security

- Each intent's approved data scope (entities, fields, time windows) formally documented and tested.

- Sensitive fields (bank details, PII, margin, payroll) excluded at the service layer not redacted by the AI model.
- Penetration test or red team exercise conducted against the conversational attack surface.

Execution Safety and Auditability

- Every ACT intent implement propose-then-execute with explicit user confirmation and idempotency controls.
- Every interaction produces a complete, correlated audit record with full correlation ID coverage.
- Kill switch for ACT capabilities has been tested under production conditions.

Governance Operations

- Ownership roles assigned across all six domains: process, SAP security, platform, security engineering, data protection, and product.
- Weekly intent review, monthly control review, and quarterly control attestation for Internal Audit all in place.
- Model update governance process triggers re-validation before any model change reaches production.

REFERENCES:

[1] SAP SE. SAP Joule: The AI Copilot for SAP Applications. SAP Product Documentation, 2024.
[2] Gartner Research. Generative AI in ERP: Emerging Capabilities and Enterprise Readiness. Gartner, Inc., 2024.
[3] SAP SE. SAP Business AI: Embedded AI Capabilities in SAP S/4HANA Cloud. SAP Help Portal, 2024.
[4] NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology, 2023.
[5] Perez, F. and Ribeiro, I. Prompt Injection Attacks Against LLM-Integrated Applications. arXiv:2302.12173, 2023.
[6] SAP SE. Identity and Access Management for SAP BTP: Principal Propagation and OAuth 2.0. SAP Help Portal, 2024.
[7] European Union Agency for Cybersecurity (ENISA). Cybersecurity of AI and Standardisation. ENISA, 2023.
[8] Institute of Internal Auditors (IIA). Global Guidance: Artificial Intelligence Considerations for the Profession of Internal Auditing. IIA, 2023.
[9] SAP SE. SAP GRC Access Control: Segregation of Duties Rule Architecture. SAP Help Portal, 2024.
[10] OWASP. OWASP Top 10 for Large Language Model Applications. OWASP Foundation, 2023.
[11] SAP SE. SAP Authorization Concept: Roles, Profiles, and Authorization Objects in S/4HANA. SAP Help Portal, 2024.
[12] Financial Stability Board (FSB). Artificial Intelligence and Machine Learning in Financial Services. FSB, 2017 (updated 2023).
[13] Forrester Research. Conversational AI in Enterprise ERP: Use Cases, ROI, and Governance Considerations. Forrester, 2024.
[14] Deloitte Insights. AI-Enabled Finance Operations: Automation, Governance, and the Path to Scale. Deloitte Development LLC, 2024.
[15] SAP SE. SAP Document Management Service: Authorization and Output Management. SAP Help Portal, 2024.
[16] ISACA. COBIT 2019 Framework: Governance and Management Objectives. ISACA, 2019.
[17] McKinsey & Company. The State of AI in 2024: Enterprise Deployments, Value Capture, and Governance Gaps. McKinsey Global Institute, 2024.
[18] SAP SE. SAP Activate Methodology: Agile Implementation Framework for SAP S/4HANA. SAP Help Portal, 2024.

[19] World Economic Forum. Responsible Use of Technology: The Microsoft INSEAD Intersection Index. WEF White Paper, 2023.